Evidence for primordial germinal centers in fish

by

Doaa Waly

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in
Physiology, Cell, and Developmental Biology

Department of Biological Sciences
University of Alberta

© Doaa Waly, 2022

Abstract

Affinity maturation is the process to improve antibody affinity for an antigen during an adaptive immune response; it is mediated by the immunoglobulin mutator enzyme activation-induced cytidine deaminase (AID). This process is crucial to produce high-affinity memory B-cells and plasma cells. In homeotherms, affinity maturation and the subsequent selection of B-cells expressing higher affinity antibodies occur in histologically distinct germinal centers (GCs), which are composed primarily of B-cells, T follicular helper cells, and follicular dendritic cells. Within the germinal centers, antigen-specific B-cells proliferate while acquiring AID mediated mutations in their VDJ exon; these cells then compete for a limited number of antigens trapped on the surface of follicular dendritic cells. Successful B-cells capture the antigen and present it to T follicular helper cells, which provide affinity-based help and drive B-cells proliferation and differentiation into memory B-cells and plasma cells (reviewed in Mesin et al., 2016, De Silva et al., 2015). Fish were thought to lack affinity maturation, in part because they lack histologically distinct germinal centers (reviewed in Magor, 2015). However, higher affinity antibodies (approximately 100-fold increase) were detected in the serum of trout and catfish following immunization (Ye et al., 2011, Wu et al., 2019). In addition, a functional homologue of AID has been identified in fish, where AID expressing cells co-exist with a population of pigmented cells called melano-macrophages (MMΦs) along with IgM^+ B-cells and $CD4^+$ T-cells (Saunders et al., 2010).

In my thesis, I tested the hypothesis that melano-macrophage clusters (MMΦCs) are functional analogues to GCs by doing VDJ repertoires on isolated clusters from the spleen and kidney of zebrafish. Construction of clonal lineage trees revealed: 1) Each cluster is dominated by a few B-

cell clonotypes that have generated several hundred daughter cells with acquired mutations. 2) Proliferation of B-cells within the clusters is inferred to be ongoing, with a few more Ig clones having smaller lineages. 3) There is evidence for positive selection for replacement mutations in regions encoding the antigen contact loops but not in regions encoding the framework regions. 4) Mutation patterns indicate the involvement of AID and error-prone polymerases in the mutation process. 5) Hill numbers and clonal abundance distribution revealed expansion and diversification of specific clonotypes, indicating the presence of an effective recruitment mechanism within the clusters. In addition, zebrafish vaccination with various proteins conjugated to a fluorescent tag revealed that long-term antigen retention occurs within MMΦCs. Imaging flow cytometry analysis confirmed the presence of lymphocyte-like cells, and the total number of these cells is consistent with the number of unique Ig transcripts isolated from each cluster. Identifying a distinct cellular milieu in which B-cells can undergo antibody affinity modification will provide a better understanding of how antibody affinity maturation operates in fish and may provide insights into how autoimmune-associated ectopic germinal centers or tertiary lymphoid structures in mammals can function.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor Dr. Brad Magor. I genuinely appreciate your invaluable scientific support for the past six years. Thank you for always being available to discuss and answer all my questions. I am thankful for your persistent guidance and for challenging my work to be better.

I would like to thank my committee members, Dr. Andrew Waskiewicz and Dr. Robert Ingham. Your contributions, advice, and support are greatly appreciated. Dr. Andy Mason and Dr. Troy Baldwin, thank you for being part of my candidacy exam. I would also like to thank Dr. Martin Flajnik for taking the time to be my external examiner.

I would also like to thank my lab mate, Aradana Muthupandian, for always being there for me. I am grateful for your friendship and advice.

I would also like to thank Dr. Kathy Magor for all the support and excellent suggestions. I would also like to thank everyone in Dr. Kathy Magor's lab for all the ideas and support.

I would like to give the most profound gratitude to my family; thank you for all your love and support. To my husband, thank you for your constant support and ongoing encouragement.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

CHAPTER II

MATERIALS AND METHODS

CHAPTER III

RESULTS

CHAPTER IV

DISCUSSION

CHAPTER V

CONCLUSION

CHAPTER VI

SIGNIFICANCE AND FUTURE DIRECTIONS

REFERENCES

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

- 5mC 5-methylcytosine
- Ab Antibody
- ADCC Antibody-dependent cell-mediated cytotoxicity
- Ag Antigen
- Aicda Activation-induced cytidine deaminase
- AID Activation-induced cytidine deaminase
- APCs Antigen-presenting cells
- APE Apurinic/apyrimidinic endonuclease
- APOBEC Apolipoprotein B mRNA Editing Catalytic Polypeptide-like
- BAFF B-cell-specific factor
- Batf Basic leucine zipper transcription factor
- BCL-2 B-cell lymphoma 2
- BCL-XL B-cell lymphoma-extra large
- Bcl6 B-cell lymphoma 6
- BCR B-cell receptor
- BER Base excision repair
- BIM Bcl-2-interacting mediator of cell death
- Blimp-1 B lymphocyte-induced maturation protein-1
- BRILIA B-cell repertoire inductive lineage and immunosequence annotator

- BSA Bovine serum albumin
- C-NHEJ Classical non-homologous end joining
- CCL19 C-C Motif Chemokine Ligand 19
- CCL21 C-C Motif Chemokine Ligand 21
- CCR7 CC-chemokine receptor 7
- CD40 Cluster of Differentiation 40
- CD40L Cluster of Differentiation 40 ligand
- CD83 Cluster of Differentiation 83
- CD86 Cluster of Differentiation 86
- CD95 Cluster of Differentiation 95
- CDA Cytidine deaminase
- CDRs Complementarity determining regions
- CHNV Carp haematopoietic necrosis virus
- CLPs Common lymphoid progenitors
- CLRs C-type lectin receptors
- CR1 Complement receptor type 1
- CR2 Complement receptor type 2
- CSF1-R Colony-stimulating factor 1 receptor
- CSR Class-switch recombination
- CXCL12 C-X-C Motif Chemokine Ligand 12
- CXCL13 Chemokine ligand 13
- CXCR4 C-X-C chemokine receptor type 4
- CXCR5 C-X-C chemokine receptor type 5

- Cy5 Cyanine-5 channel
- DAMPs Damage-associated molecular patterns
- dCMP Deoxycytidine monophosphate
- DCs Dendritic cells
- DIVAC Diversification Activator
- dN/dS Non-synonymous to synonymous
- DZ Dark zone
- E2A E box binding protein 2A
- EBF1 Early B cell Factor-1
- EBI2 Epstein-Barr virus-induced receptor 2
- eEF1A Elongation factor 1α
- Exo1 Exonuclease-1
- FDCs Follicular dendritic cells
- FITC Fluorescein isothiocyanate channel
- FO Follicular
- FRCs Fibroblastic reticular cells
- FSC Forward scatter
- FWRs Framework regions
- GC Germinal center
- GCV Gene conversion
- HIGM2 Hyper-IgM syndrome type 2
- HIV-1 Human Immunodeficiency Virus-1
- HR Homologous recombination

- HSCs Hematopoietic stem cells
- Hsp90 Heat shock protein 90
- i.p. Intraperitoneally
- ICAM1 Intercellular adhesion molecule 1
- ICOS Inducible T-cell co-stimulator
- ICOSL Inducible T-cell co-stimulator ligand
- IFNα Interferon-α
- IFNβ Interferon-β
- IFNγ Interferon-γ
- Ig Immunoglobulin
- IgA Immunoglobulin A
- IgD Immunoglobulin D
- IgE Immunoglobulin E
- IgG Immunoglobulin G
- IgH Immunoglobulin heavy chain
- IGHD Immunoglobulin heavy chain diversity
- IGHJ Immunoglobulin heavy chain joining
- IGHV Immunoglobulin heavy chain variable
- IgL Immunoglobulin light chain
- IgM Immunoglobulin M
- IgNAR Immunoglobulin new antigen receptor
- IgSF Immunoglobulin superfamily
- IgZ/IgT Immunoglobulin Z/T

- IL-10 Interleukin 10
- IL-1β Interleukin 1 β
- IL-21 Interleukin 21
- IL-4 Interleukin 4
- IL-6 Interleukin 6
- ILCs Innate lymphoid cells
- indels Insertions and deletions
- iNext iNterpolation/EXTrapolation
- KLH Keyhole limpet haemocyanin
- LLPCs Long-lived plasma cells
- LMPPs Lymphoid-primed multipotent progenitors
- LPS Lipopolysaccharides
- LRR Leucine-rich repeat
- LZ Light zone
- MALTs Mucosa-associated lymphoid tissues
- MFGE8 Milk fat globule epidermal growth factor 8
- MHC I Major histocompatibility complex class I
- MHC II Major histocompatibility complex class II
- mIgM Membrane Immunoglobulin M
- miRNAs Micro RNAs
- MLH1 MutL homolog 1
- MMR Mismatch repair
- MMΦCs Melano-macrophage clusters

- MMΦs Melano-macrophages
- MPPs Multipotent progenitors
- MZ Marginal zone
- ncMMR Noncanonical mismatch repair
- NES Nuclear export signal
- NK Natural killer
- NLS Nuclear localization signal
- NP Non-templated and palindromic nucleotides
- PA Bacillus anthracis
- PALS Periarteriolar lymphoid sheath
- PAMPs Pathogen-associated molecular patterns
- Pax5 Paired box protein 5
- PCNA Proliferating cell nuclear antigen
- PCR Polymerase chain reaction
- PD-1 Programmed cell death-1
- PE Phycoerythrin
- pIgR Poly-Ig receptor
- PKA Protein Kinase A
- PMS2 Postmeiotic segregation increased 2
- Pol ζ Polymerase zeta
- pol η Polymerase eta
- Pol ι Polymerase iota
- Pol κ Polymerase kappa

- PRRs Pattern recognition receptors
- R/S Replacement to silent
- RAG Recombination activating gene
- RBCs Red blood cells
- RDP4 Recombination Detection Program
- RNAPII RNA polymerase II
- RSS Recombination signal sequence
- SCS Subcapsular sinus
- SHM Somatic hypermutations
- SPR Surface plasmon resonance
- SPT5 Suppressor of Ty homolog-5
- SSC Side scatter
- ssDNA Single-stranded DNA
- TBMs Tingible body macrophages
- TCR T-cell receptor
- Tfh T follicular helper cell
- TLRs Toll-like receptors
- TLS Translesion synthesis
- TMS Tricaine methanesulfonate
- TNF Tumor necrosis factor
- TNP-KLH Trinitrophenyl keyhole limpet haemocyanin
- UNG Uracil-DNA glycosylase
- vCAM1 Vascular cell-adhesion molecule 1

- VHSV Viral Hemorrhagic Septicemia Virus
- VLR Variable lymphocyte receptors
- XP-V Xeroderma pigmentosum variant
- XRCC2 X-ray repair cross complementing 2
- XRCC3 X-ray repair cross-complementing 3

CHAPTER I

INTRODUCTION

1. Introduction

1.1. Overview

Activated B-lymphocytes mediate the humoral immune response by secreting antigen-specific antibodies in response to foreign pathogens during an adaptive immune response. The diversity of the primary antibody (Ab) repertoire is generated by a process called VDJ recombination. The basic Ab protein (also known as an immunoglobulin (Ig)) is composed of two identical heavy chains (IgH) and two identical light chains (IgL). Each chain is composed of constant and variable immunoglobulin domains (Dreyer et al., 1965). The heavy chain variable region gene is subdivided into three segments variable (V), diverse (D), and joining (J), and the light chain into variable (V) and joining (J) segments (reviewed in Di Noia & Neuberger, 2007; Schatz et al., 2011). The variety of different antibody molecules can be achieved by a random rearrangement of V, D, and J gene segments available for integration, as well as by deletion and insertion of different nucleotides at the V-D-J junctions (reviewed in Di Noia & Neuberger, 2007). As many as 10^{15} different antibody molecules can be generated by this process in humans and mice, and B-cells with low avidity to self molecules migrate to the periphery (reviewed in Ubelhart et al., 2015, Schroeder, 2006).

Mature naïve B-cells leave the bone marrow to the periphery and recirculate the secondary lymphoid organs. Following the activation of naïve B-cells by interacting with a specific antigen (Ag) via the B-cell receptor (BCR) during an immune response in the lymphoid tissue at the site of infection, B-cells interact with antigen-specific T-cells and become fully activated B-cells (plasmablasts). Plasmablasts begin to proliferate and will secrete antibodies that may have a low affinity for the foreign antigen. Plasmablasts and activated T-cells may also be recruited to specific lymphoid tissue regions within the primary follicles (reviewed in Mesin et al., 2016, De Silva et al., 2015). Typically, a subset of these activated B-cells nucleates a follicle, and each of these cells will have its own unique VDJ sequence. Once there, B-cells alter their gene expression profiles and are known as centroblasts. Centroblasts are highly proliferative, though they downregulate Ab production and upregulate expression of activation-induced cytidine deaminase (AID). The centroblasts become so plentiful, at this point, the aggregate is referred to as a secondary follicle and later as a germinal center (GC). AID expressed in the centroblasts will cause somatic hypermutations (SHM) in the VDJ exon and the surrounding introns to change (and potentially enhance) their antibodies' affinity to the exogenous antigen. Modified cells will alter their gene expression again and begin to express membrane antibodies. These cells, now referred to as centrocytes, may have altered antigen-binding abilities (reviewed in Mesin et al., 2016, De Silva et al., 2015). Centrocytes compete for a limited number of intact antigens retained on the surface of follicular dendritic cells (FDCs). Centrocytes with high affinity for the antigen endocytose it and present it via major histocompatibility complex class II (MHC II) to T follicular helper cell (Tfh), from which they receive survival signals. Thus, the GC reaction is also referred to as antibody affinity maturation. Selected cells can undergo more SHM and proliferation, or a phase of proliferation, and daughter cells will differentiate into either long-lived antibody-secreting

plasma cells or memory B-cells. Unsuccessful B-cells will undergo pre-programmed apoptosis (reviewed in Magor, 2015, Mesin et al., 2016, De Silva et al., 2015, Wing et al., 2018).

Fish and all ectothermic vertebrates lack GCs; consequently, they are thought to lack affinity maturation (reviewed in Magor, 2015). However, our lab has identified the first fish homologue of AID. It is fully functional and can induce SHM in mouse $AID^{-/-}$ B-cells (Saunders et al., 2004, Wakae et al., 2006). Moreover, *in situ* hybridization of the fish AID transcript revealed that AID is expressed within aggregates of pigmented cells called melano-macrophages (MMΦs) in the spleen and kidney of catfish. These cell aggregates are also associated with IgM^+ B-cells and $CD4^+$ T-cells (Saunders et al., 2010), and studies have shown that intact antigen is retained within melano-macrophage clusters (MMΦCs) for a long period (Lamers et al., 1985, Lamers CH, De Haas MJ 1985, Ziegenfuss et al., 1991). Melano-macrophages (MMΦs) are autofluorescent and, at least in cyprinids (goldfish and zebrafish), frequently are extensively encapsulated by reticular cells. These characteristics make it relatively easy to isolate intact MMΦCs (Diaz-Satizabal et al., 2015). Furthermore, earlier work on antibody affinity maturation in people with autoimmune diseases revealed that affinity maturation process occurs without conventional GCs in regions where T- helper cells, B-cells, and a source of antigen are present (William et al., 2002).

Taken together, these observations led to the hypothesis that MMΦCs in fish are acting as a primordial germinal center. In my thesis, I examined the three hallmarks of the GC reaction within MMΦCs, including:

1. Clonal expansion of activated B-cells.
2. Accumulation of AID-induced SHM in the VDJ exon of the Ig gene.

3. The presence of an active antigen-driven selection process.

1.2. Thesis objectives

The overall objective of my thesis was to determine if MMΦCs in fish are functional analogues to GCs of homeothermic vertebrates by doing VDJ repertoires on isolated clusters and examining the hallmarks of the GC reaction within MMΦCs.

Teleost fish are able to improve the affinity of their antibodies in response to T-cell dependent antigens (Ye et al., 2011, Wu et al., 2019), and the expression of MHC II is abundant in the spleen and kidney of zebrafish, and 70% of these MHC II expressing cells are B-cells (Wittamer et al., 2011). In addition, the key initiator of affinity maturation AID is expressed within MMΦCs in the spleen and kidney, these cell aggregates are also associated with IgM^+ B-cells, $CD4^+$ T-cells, and intact antigens are retained within MMΦCs for an extended period (Saunders et al., 2004, Saunders et al., 2010, Lamers et al., 1985, Lamers CH, De Haas MJ 1985, Ziegenfuss et al., 1991). Furthermore, affinity maturation occurs without conventional GCs in people with autoimmune diseases in regions where T- helper cells, B-cells, and a source of antigen are present (William et al., 2002).

Based on these observations, it was hypothesized that MMΦCs in fish are acting as primordial GCs. In my thesis, I investigated whether MMΦCs in fish are functionally analogous to GCs by examining the hallmarks of the GC reaction within MMΦCs, including 1) clonal expansion of activated B-cells; 2) accumulation of AID induced somatic hypermutation in the VDJ exon of the

Ig gene; 3) the presence of an active Ag-driven selection process.

Specific objectives of my thesis were: (1) to determine if B-cells clonally expand while acquiring mutations within MMΦCs; (2) to determine if an active antigen-driven selection process occurs within MMΦCs; (3) to determine if AID and error-prone polymerases are involved in the mutation process; (4) to determine if there is an effective recruitment and diversifying mechanisms within MMΦCs; (5) to determine if long-term Ag retention occurs within MMΦCs in zebrafish; (6) to confirm the presence of lymphocyte-like cells within the clusters.

1.3. Thesis outline

This thesis encompasses six chapters. Chapter I is an introduction to subjects related to my thesis work. In this chapter, I first provide a brief overview of the immune response with an emphasis on the adaptive immune response and the development and activation of B-cells. I then describe the immunoglobulin mutator enzyme, activation-induced cytidine deaminase (AID), and its role in antibody diversification. In addition, chapter I discusses the adaptive immune system and the current knowledge about melano-macrophage clusters (MMΦCs) in teleost fish. Chapter II provides a description of the methodologies and protocols used during the completion of my thesis. In chapter III I describe my results in detail. Chapter IV provides a comprehensive discussion of my results and their interpretations. Chapter V provides an integrative view of my findings and describes how these findings relate to the current understanding of antibody affinity maturation process in fish. Finally, in chapter VI, I discuss the applied significance of my results and propose further studies.

1.4. The innate and adaptive immune responses

Our immune system has the ability to eliminate pathogens and harmful substances using complex defense mechanisms. Simultaneously, to prevent damaging host cells, the immune system can differentiate between self and non-self-molecules, known as self-tolerance, through mechanisms that inhibit the immune response following the encounter with self-molecules (reviewed in Chaplin, 2010). Immune cells arise from hematopoietic stem cells in the bone marrow, and the immune system has two major lines of defense: innate immunity and adaptive immunity (reviewed in Chaplin, 2010).

Innate immunity provides a rapid response to foreign pathogens. Pattern recognition receptors (PRRs) inside and outside the cell are germline-encoded; these receptors can detect and respond to a wide range of pathogens that share common molecules known as pathogen-associated molecular patterns (PAMPs) (Lemaitre et al., 1996). In addition, PRRs can detect and eliminate damage-associated molecular patterns (DAMPs), which are endogenous molecules released from damaged or stressed cells. PAMPs and DAMPs activate PRRs signaling, which leads to the production of proinflammatory cytokines and chemokines, including type I interferon (interferon-α (IFNα) and IFNβ), interleukin 1 β (IL-1β), interleukin 6 (IL-6), and tumor necrosis factor (TNF) (Poltorak et al., 1998, reviewed in Amarante-Mendes et al., 2018). These molecules alert the immune system by activating antimicrobial activities such as immune cell recruitment. In addition, they play a key role in activating antigen-specific adaptive immune responses (reviewed in Akira et al., 2006).

Several cells derived from haematopoietic stem cells are involved in the innate immune response, including macrophages, neutrophils, dendritic cells, mast cells, basophils, eosinophils, natural killer (NK) cells, and innate lymphoid cells (ILCs) (reviewed in Chaplin., 2010). Macrophages and neutrophils are phagocytes; they phagocytose microbes and kill them using various effector molecules such as nitric oxide, superoxide, and degradative enzymes. Unlike neutrophils, macrophages are long-lived and involved in antigen presentation (reviewed in Gasteiger et al., 2017, in Chaplin., 2010). Dendritic cells (DCs) are phagocytic and antigen-presenting cells (APCs), and they play a significant role in initiating the adaptive immune response (reviewed in Chaplin et al., 2010, Marshall et al., 2018, Gasteiger et al., 2017). Mast cells, basophils, and eosinophils are involved in the initiation of the acute inflammatory responses and pathogenesis of allergic diseases (reviewed in Chirumbolo et al., 2018). NK cells play a significant role in host-rejection of tumors and the destruction of virally infected cells (reviewed in Vivier et al., 2008). ILCs lack the somatically recombined antigen receptors, and there are three types of ILCs, ILC-1, ILC-2, and ILC-3. Cytokines produced by these cells play a more regulatory role, such as regulating DCs and T-cells activities during an immune response (reviewed in Eberl et al., 2015, Gasteiger et al., 2017).

The adaptive immune system can be subdivided into T-cell mediated cellular immunity and humoral immune response, which is mediated by antibodies secreted by B-cells. T-cells express a unique T-cell receptor (TCR) that can be activated by APC (usually DCs but also macrophages and B-cells). This activation leads to rapid proliferation of the different subsets, including $CD8^+$ T-cells (cytotoxic T-cells), which kill cells infected with intracellular microbes directly, and $CD4^+$ T-cells (helper T-cells), which regulate the cellular and humoral immune responses by secreting

different cytokines such as IL4, IFNγ, and IL21 (reviewed in Luckheeram et al., 2012, Nicholson., 2016).

B-cells express a unique B-cell receptor (BCR), and unlike T-cells, they can interact with antigens directly. Based on their development and their ability to be activated by T cell-dependent or T-cell independent antigens, naïve B-cells are subdivided into follicular (FO; or B-2 type), marginal zone (MZ), and B-1 B-cells (reviewed in Allman et al., 2008).

FO B-cells ($IgM^{low}IgD^{hi}CD21^{med}CD23^{+}CD1d^{low}$) reside in the primary follicles of B-cells in the secondary lymphoid organs, and they recirculate through the blood and lymph to these follicles. Within the secondary lymphoid organs, B follicles are always adjacent to T-cell zones, which allows FO B-cells to participate in a T-cell dependent immune response. Nonetheless, FO B-cells also recirculate through the bone marrow, where they are organized around the vascular sinusoids and respond to T cell-independent blood-borne pathogens (Cariappa et al., 2005, reviewed in Allman et al., 2008). MZ B-cells are $IgM^{hi}IgD^{low}CD21^{hi}CD23^{-}CD1d^{hi}$, and they are mostly non-recirculating; they reside mainly in the vicinity of the marginal sinus of the spleen. MZ B-cells are considered to be innate-like B-cells; they express polyreactive BCRs that bind to multiple PAMPs. In addition, they participate in the early immune response due to their ability to differentiate into short-lived plasmablast in the absence of BCR ligation through various PRRs such as toll-like receptors (TLRs), scavenger receptors, and C-type lectin receptors (CLRs) (reviewed in Martin et al., 2002, Cerutti et al., 2013, Garraud et al., 2012). MZ B-cells play a significant role in T-cell independent responses to blood-borne pathogens; also, they transport antigens in immune complexes from the marginal sinus to FO B-cells in the primary follicles (reviewed in Martin et

al., 2002, Cerutti et al., 2013). Another innate-like B-cell is B-1 B-cell. These B-cells are the main B-cells population in the peritoneal and the pleural cavities; they are subdivided into B-1a ($CD5^+$) and B-1b ($CD5^-$) B cells based on the expression of CD5. B-1 B-cells spontaneously secrete polyreactive germline-encoded antibodies in the absence of exogenous antigens; these antibodies form the natural antibody repertoire (Hayakawa et al., 1985, Baumgarth et al., 2011). The mechanism leading to natural antibodies production is poorly understood; however, natural antibodies are produced in the absence of antigenic stimulation, and they are generally of low affinity. In addition to their ability to recognize self-antigens, they can bind exogenous antigens, and therefore they provide the first line of adaptive defense against pathogens (reviewed in Reyneveld et al., 2020).

A key feature of the humoral immune response is the production of BCR with an improved binding affinity to the antigen as a result of AID-induced somatic hypermutations in the antigen contact sites of the BCR and the subsequent selection of B-cells expressing higher affinity antibodies (reviewed in Nicholson., 2016).

1.5. B-cells development and activation

1.5.1. Development of B-cells and the immunoglobulin genes

B-cells express membrane-bound immunoglobulins (Igs), also known as B-cell receptors (BCRs), and they have the potential to secrete their surface receptors as antibodies (Abs) during an adaptive immune response to foreign pathogens via alternative splicing. Igs are glycoproteins, and they

belong to the immunoglobulin superfamily (IgSF) (Williams et al., 1988). The basic Ab protein is composed of two identical heavy chains (IgH) and two identical light chains (IgL); the two heavy chains are linked to each other by a number of disulfide bonds, and each heavy chain is connected to a light chain by disulfide bond. Each chain is composed of variable immunoglobulin domains at the amino-terminal and several constant immunoglobulin domains arranged from the amino- to the carboxy-terminal (e.g., CH1 and CH2: Fig.1.1) (Dreyer et al., 1965). The variable domains recognize and bind the antigen, while the constant domains are relatively constant and define the Ig isotype and effector functions. The heavy chain variable region gene is encoded by three segments variable (V), diverse (D), and joining (J), and the light chain is divided into variable (V) and joining (J) segments (Tonegawa, 1983). In mammals, the IgH locus has a translocon configuration consisting of several VH segments followed by DH, JH gene segments, and the constant (CH) regions. Similarly, their light chain has a translocon organization; however, it is only encoded by VL, JL, and CL gene segments (Fig.1.1) (Tonegawa, 1983).

In mammals, within the primary lymphoid organ (e.g., fetal liver and adult bone marrow), B-cells are generated from hematopoietic stem cells (HSCs), which give rise to multipotent progenitors (MPPs). Subsequently, MPPs differentiate into lymphoid-primed multipotent progenitors (LMPPs) and common lymphoid progenitors (CLPs). Transcription factors such as E box binding protein 2A (E2A), Early B cell Factor-1 (EBF1), Paired box protein 5 (Pax5), and Ikaros direct CLPs into B-cell developmental pathway and the initiation of VDJ recombination, the process by which the diversity of the primary Ab repertoire is generated (Schatz et al., 1989, reviewed in Ramírez et al., 2010).

During VDJ recombination, the exon of the variable regions is assembled; this process is initiated by the recombination activating gene (RAG) proteins, RAG1 and RAG2. The activity of RAG proteins is directed by the recombination signal sequence (RSS) flanking each V, D, and J gene segment; at the 3’ of V genes, 5’ of J genes and at both 5' and 3' in D genes. During the recombination process RAG proteins bind and cleave the RSS. Subsequently, the free 3' hydroxyl group in the broken single-stranded DNA (ssDNA) attacks the opposite strand and forms a covalently sealed hairpin at the end of the gene segment. Palindromic nucleotides (P nucleotides) are inserted when the hairpin ends are opened asymmetrically. Afterward, the broken ends are joined by the classical non-homologous end joining (C-NHEJ) pathway. VDJ recombination process occurs in a defined order; following the rearrangement of the immunoglobulin heavy chain (IgH), recombination begins in the light chain (IgL), and the rearrangement of D-to-J precedes V-to-DJ recombination (Schatz et al., 1989, reviewed in Schatz et al., 2011, Helmink et al., 2012, Roth et al., 2015).

Overall, the variety of different antibody molecules can be achieved by randomly selecting different germline of V, D, and J gene segments available for integration, as well as by the insertion and deletion of different nucleotides at the V-D-J junctions by terminal deoxynucleotidyl transferase (TDT) and exonucleases activity, respectively. An additional source of diversity results from the various combinations of the heavy and light chains variable regions to form the antigen contact site (Schatz et al., 1989, Nadel et al., 1997, Alt et al., 1982., reviewed in Di Noia et al., 2007, Schatz et al., 2011).

VDJ recombination allows each B-cell to produce its own unique BCR. As many as 10^{15} different antibody molecules can be generated by this process in humans and mice, and B-cells with low avidity to self molecules migrate to the periphery (reviewed in Schroeder, 2006, Ubelhart et al., 2015). B-cells with autoreactive antigen receptors undergo further rearrangement of the IgL and possibly IgH through a process called receptor editing; however, if receptor editing results in high avidity to self molecules, the cell remains in the bone marrow and dies by clonal deletion (reviewed in Nemazee, 2006). Receptor editing and clonal deletion are part of B-cells central tolerance, which occurs in the primary lymphoid organs during B-cell development (reviewed in Nemazee, 2017).

IgH and IgL V domains can be divided into three hypervariable loops, called the complementarity determining regions (CDRs), and four non-hypervariable regions of a comparatively constant sequence named the framework regions (FWRs). The IgH V gene segment contains FWR1, FWR2, FWR3, CDR1, CDR2, and the amino-terminal of CDR3, which also includes the IgH D segment and the carboxy-terminal of the CDR3 is within the J region. FWR4 is entirely within the IgH J gene segment. The three CDRs of the IgH chain are paired with the three CDRs of the IgL chain to form the antigen contact site in the Ig molecule (Fig.1.1; reviewed in Schroeder et al., 2010). On the other hand, the FWRs form the support structure for the CDRs, and they maintain the structural integrity of the Ig variable domain (Jones et al., 1986, Rada et al., 1991, Jacob et al., 1993).

The constant region of the Ab molecule is generated by a single gene; in mammals, there are five main classes of the heavy chain constant domain Igμ (IgM), Igδ (IgD), Igγ (IgG), Igα (IgA), and Igε (IgE) (reviewed in Schroeder et al., 2010). IgH constant (CH) region is responsible for a variety

of effector functions such as complement fixation, Fc receptor binding, and the isotype of the Ig (Tonegawa, 1983). The light chain constant (CL) region has two types: Kappa (κ) and Lambda (λ) (Williams et al., 1988). During B-cell development, VH domain is expressed in association with μ heavy chain to produce IgM and then IgD (Igδ) by alternative splicing, which allows co-expression of IgM and IgD. Following the encounter with an antigen, depending on the type of infection, the variable domain of the heavy chain can be associated with a different Ig constant domain (i.e., IgG, IgA, and IgE) via class switch recombination (reviewed in Yu et al., 2019).

IgM is present in all vertebrates, and it is the most conserved Ig isotype. Naïve B-cells express monomeric IgM on their cell surface upon stimulation with an antigen multimeric IgM is secreted via alternative splicing. Multimeric IgM functions by opsonizing (coating) pathogens for destruction and complement activation (reviewed in Schroeder et al., 2010). A subset of B-cells is $IgD^{+}IgM^{-}$, secreted IgD binds to bacteria and viruses and their microbial virulence factors. In addition, IgD binds to basophils and mast cells; this binding leads to the activation of these cells and the release of antimicrobial factors (reviewed in Chen et al., 2011, Schroeder et al., 2010). In mammals, the predominant Ig isotype in the body is IgG; there are four IgG subclasses (IgG1, IgG2, IgG3, and IgG4). These subclasses have different functional activities, including complement activation, opsonization of antigens, and neutralization of viruses and toxins (reviewed in Schroeder et al., 2010). While an immune response to protein antigens is generally associated with IgG1 and IgG3, IgG2 and IgG4 are induced in response to polysaccharide antigens (reviewed in Schroeder et al., 2010). Dimeric IgA is the predominant Ig isotype in mucosal surfaces and secretions; using polymeric forms of IgA and the poly-Ig receptor (pIgR) that is constitutively expressed at the basolateral side of mucosal epithelial cells, IgA is transported to the

apical side of the cells and released into the lumen (reviewed in Corthesy, 2013, Schroeder et al., 2010). IgA plays a significant role in the immune defense at these surfaces by neutralizing microbes such as viruses and bacteria and thus preventing them from penetrating the mucosal barrier (Mestecky et al., 2005). IgE is associated with allergic reactions and the response to parasitic infections and animal venoms. Furthermore, IgE binds to FcεRI, expressed on immune cells such as mast cells, basophils, and eosinophils, which leads to degranulation and the release of proinflammatory mediators, and the recruitment of immune cells to the site of infection (reviewed in Sutton et al., 2019, Schroeder et al., 2010).

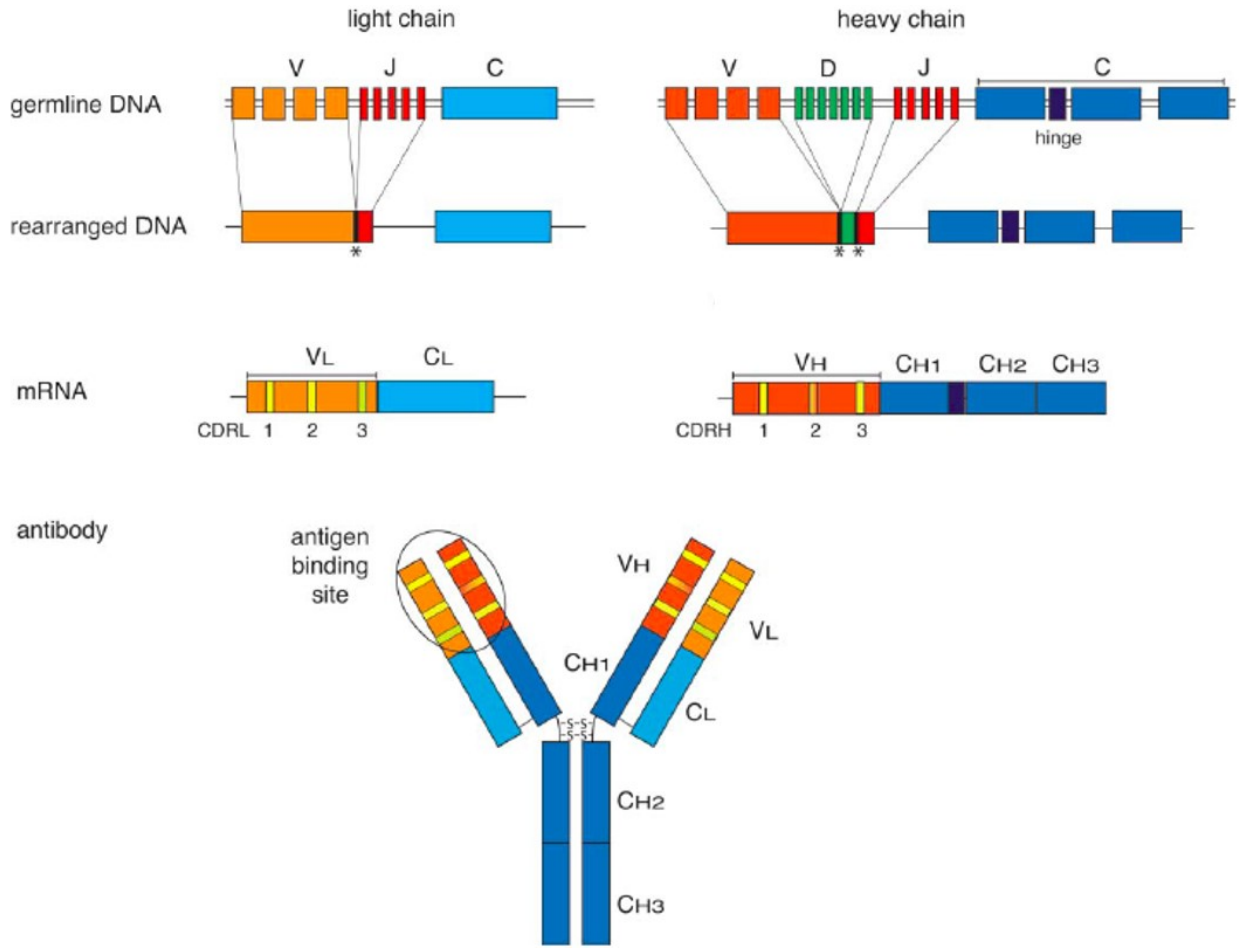


Figure 1.1. Schematic overview of the mammalian immunoglobulin gene organization and antibody molecule. Germline gene segments of the heavy and light chains are assembled by VDJ recombination; during this process, insertion and deletion of different nucleotides occur at the V-D-J junctions by terminal deoxynucleotidyl transferase and exonucleases (asterisk). Antibody molecules are composed of two identical heavy chains and two identical light chains; disulfide bonds link the two heavy chains to each other and to the light chains. Each chain is composed of variable (V) and constant (C) immunoglobulin domains. CDRH and CDRL form the antigen contact site of the antibody molecule. V, variable. D, diverse. J, joining. CDR, complementarity determining regions. (Adapted from Feederle and Schepers, 2017).

1.5.2. B-cell activation

Mature naive B-cells leave the bone marrow to the periphery and populate the secondary lymphoid organs such as the spleen and lymph node. These cells express a plasma membrane-bound Ig, termed the B-cell receptor (BCR); they co-express IgM and IgD, with identical antigen-binding domains on their surface via a splicing mechanism (reviewed in Cyster et al., 2019, Stebegg et al., 2018). Lymphocytes entering the spleen pass through the MZ, which contains MZ macrophages and MZ B-cells. Chemokines secreted by fibroblastic reticular cells within the lymphoid white pulp direct the movement of lymphocytes into the white pulp. The chemokine ligand 13 (CXCL13) is a secreted cytokine made by follicular dendritic cells (FDCs), and it is highly abundant in the B-cell follicles; it binds to cell-surface C-X-C chemokine receptor type 5 (CXCR5) and guides the migration of CXCR5 expressing B-cells into B-cell follicles. T-cell localization to the T-cell zone which is also called the periarteriolar lymphoid sheath (PALS), is mediated by the expression of CC-chemokine receptor 7 (CCR7), the receptor for CCL19 (C-C Motif Chemokine Ligand 19) and C-C Motif Chemokine Ligand 21 (CCL21) expressed by fibroblastic reticular cells (FRCs) (reviewed in Cyster et al., 2019).

B-cell antigens are divided into T-cell independent and T-cell-dependent antigens based on their ability to induce B-cells proliferation and differentiation. T-cell independent antigens are classified as either T-cell independent type 1 or T-cell independent type 2 antigens. While T-cell independent type 1 antigens trigger signal transduction in B cells via PRR such as toll-like receptors (TLRs), which recognize pathogen-associated molecular patterns (PAMPs), T-cell independent type 2 antigens can stimulate B-cells to synthesize antibodies without T-cell help by extensive

crosslinking of the BCR. On the other hand, antigens with diverse epitopes such as proteins are T-cell-dependent antigens and can only activate B-cells to synthesize specific antibodies in the presence of T-cell help (reviewed in Song et al., 2019, Allman et al., 2019, Cyster et al., 2019).

The engagement with a T-cell-dependent soluble antigen or antigen presented by professional antigen-presenting cells (APCs) primarily by FO B-cells in the follicles results in BCR-mediated signaling and initiates B-cell activation, which promotes Ag internalization. This is followed by intracellular degradation of the antigen in the endosomal compartments and the generation of peptides that are presented on the B-cells surface via MHC class II molecules. These B-cells upregulate CCR7 and Epstein-Barr virus-induced receptor 2 (EBI2) and move to the T- and B-cell zones border, where they receive cognate help from antigen-specific $CD4^+$ T-cells. In addition, activated B-cells maintain their CXCR5 expression to prevent their complete entry into T-cell zones (reviewed in Stebegg et al., 2018, Song et al., 2019, Biram et al., 2019). Antigen-specific $CD4^+$ T-cells can be activated by antigen-presenting cells such as DCs; upon stimulation through their TCR, co-stimulatory molecules and cytokines produced by DCs such as inducible T-cell co-stimulator ligand (ICOSL) and IL-6, these activated T-cells differentiate to pre- T follicular helper (Tfh) cells and rapidly upregulate the expression of CXCR5, inducible T-cell co-stimulator (ICOS), programmed cell death-1 (PD-1), and the transcription factor B-cell lymphoma 6 (Bcl6). These early pre-Tfh cells then move to the border of the T-cells and B-cells zones, where they recognize the peptide: MHC II complex via the TCR and form a stable synapse with their cognate B-cells (Garside et al., 1998, reviewed in Song et al., 2019). These synapses are essential for pre-Tfh cells to maintain their Tfh phenotype. Several molecules are involved in these synapses, including CD40, which is expressed by B-cells and binds to CD40L upregulated by pre-Tfh cells.

CD40L is a member of the tumor necrosis factor (TNF) family; it is essential for T-cell-dependent immune response and germinal center formation (Kawabe et al., 1994, reviewed in Verstegen et al., 2021, Song et al., 2019). Another molecule is the ICOS, expressed by pre-Tfh cells and binds its ligand, ICOSL, expressed by B-cells; ICOS helps form a stable immunological synapse between B- and T-cells. Ultimately, these immunological synapses stimulate B-cell proliferation and differentiation and the development of pre-Tfh cells into effector Tfh cells (reviewed in Cyster et al., 2019, Biram et al., 2019, Stebegg et al., 2018).

A subset of these fully activated B-cells retains their EBI2 expression and differentiate into extrafollicular plasmablasts, which secrete antibodies that may have a relatively low affinity for the foreign antigen. These cells are short-lived, and somatic hypermutation can be induced only at a low level during the extrafollicular response; nevertheless, they provide a rapid initial immune response (a few days after the first antigen detection) (McHeyzer-Williams et al., 1993, reviewed in Elsner et al., 2020). Activated B-cells can also differentiate to early memory B-cells; these cells are like intrafollicular memory B-cells, quiescent and long-lived (Inamine et al., 2005). Alternatively, activated B-cells upregulate Bcl6 and downregulate the expression of EBI2 to renter the follicle and commit to the germinal center (GC) formation. Studies have shown that the fate of these activated B-cells is determined by their binding affinity to the exogenous antigen and, consequently, the interactions between pre-Tfh and B-cells. B-cells with a relatively higher affinity differentiate toward plasmablasts, while cells with lower binding affinity differentiate toward memory B-cells or GC B-cells (reviewed in Hoffman et al., 2016, Batista and Harwood, 2009, Verstegen et al., 2021).

1.5.3. The germinal center reaction

The germinal center reaction is also referred to as antibody affinity maturation. A hallmark of the humoral immune response is the production of plasma cells secreting high-affinity class-switched antibodies and long-lived memory B-cells.

GCs form within B-cell follicles in the secondary lymphoid organs; these follicles also contain follicular dendritic cells (FDCs). Unlike other immune cells, FDCs are non-hematopoietic, and derive from ubiquitous perivascular precursors (Krautler et al., 2012). FDCs are non-migratory cells and express complement and Fc receptors such as CR1, CR2, and FcγRIIB in addition to high levels of the adhesion molecules vascular cell-adhesion molecule 1 (vCAM1) and intercellular adhesion molecule 1 (ICAM1). Their ability to retain antigens for a long term in the form of antibody and complement immune complexes provides the first selection step for GC B-cells (reviewed in Allen et al., 2008). These immune complexes are transported to the follicle by subcapsular sinus (SCS) macrophages in the lymph node and non-cognate B-cells in a BCR-independent way using complement receptors CR1 and CR2 (Phan et al., 2009, reviewed in Batista and Harwood, 2009). In addition, FDCs enhance the engulfment of apoptotic cells by macrophages within the GC by secreting phosphatidylserine-specific opsonin milk fat globule epidermal growth factor 8 (MFGE8), which coats dying GC B-cells (Kranich et al., 2008). FDCs within the primary follicles and GCs also produce B-cell activating factor (BAFF), which is required for maintaining B-cells homeostasis and enhancing the survival of B-cells within the GC (Lesley et al., 2004, reviewed in Allen et al., 2008).

Following the initial contact of B- and T-cells at the border of the T- and B-cells zones, a few activated B- and T-cells migrate to the primary follicle. Typically, a subset of activated B-cells nucleates a follicle, and each of these cells will have its own unique VDJ sequence. Once there, B-cells alter their gene expression profiles and are referred to as centroblasts. Centroblasts are highly proliferative, though they downregulate Ab production and upregulate AID expression (reviewed in Verstegen et al., 2021). The centroblasts become so plentiful, at this point, the aggregate is referred to as a secondary follicle and later as a germinal center. During the proliferation of centroblasts, resident naïve follicular B-cells are pushed aside to form the mantle zone around the GC. Based on their histological appearance, GCs are organized into a dark zone (DZ) and a light zone (LZ) (Fig.1.2) (Röhlich et al., 1930, reviewed in Verstegen et al., 2021). The DZ contains centroblasts (DZ B-cells) and CXCL12-expressing reticular cells, and the LZ contains centrocytes (LZ B-cells), FDCs, and Tfh cells (Fig.1.2). Centroblasts and centrocytes differ in their transcriptional program; centroblasts express high levels of AID and CXCR4 (chemokine receptor for CXCL12), while centrocytes express a low level of CXCR4 and a high level of activation markers (CD83 and CD86) and CXCR5 (reviewed in De Silva et al., 2015). GC Tfh cells express higher levels of CXCR5, PD-1, Bcl6, IL-21, and IL-4 compared to extrafollicular Tfh, and they play a significant role in the selection of B-cells expressing high-affinity antibodies through the interaction with cognate peptide antigen on MHC II molecules using their T-cell receptor (TCR) (reviewed in Biram et al., 2019). Recently, it has been shown that IL-21 secreted by Tfh plays a significant role in maintaining the GC zones distribution by promoting the proliferation of LZ B-cells, the DZ is smaller in size in the absence of IL-21, and the LZ is enlarged due to the accumulation of LZ B-cells (Zotos et al., 2021). Tfh cells secreted IL-4 and IL-21 direct the movement of B-cells within the GC; IL-4 is essential for the development of centroblasts to

centrocytes, whereas IL-21 promotes the progress of centrocytes to centroblasts (reviewed in Wishnie et al., 2021).

Within the DZ, AID expressed in the centroblasts causes somatic hypermutations (SHM) in the VDJ exon and the surrounding introns predominantly in the complementarity-determining regions (CDRs) to change (and potentially enhance) their antibodies' binding affinity to the exogenous antigen. Centroblasts that acquire deleterious mutations, mainly mutations that result in a frameshift or a stop codon, and fail to replace their surface BCR are removed by apoptosis in the DZ. Centroblasts with modified BCRs move to the LZ using CXCR5, the receptor for CXCL13 that is secreted by FDCs, where they again alter their gene expression and express membrane antibodies (reviewed in De Silva et al., 2015, Mesin et al., 2016, Biram et al., 2019). Centrocytes are non-proliferating B-cells and may have altered antigen-binding abilities due to the mutations. Within the LZ, centrocytes compete for a limited number of antigens trapped on the surface of FDCs; this provides the first step in the positive selection of GC B-cells. Centrocytes with high affinity for the antigen endocytose it and present it via class II MHC to Tfh cells, from which they receive survival signals. As a result, centrocytes with a higher affinity for the antigen present a higher density of peptide-MHC II complexes and hence receive a higher share of T-cell help and subsequently positive selection (Fig.1.2; reviewed in De Silva et al., 2015, Mesin et al., 2016). Interestingly, recent studies have shown that the GC is a more dynamic structure, where B-cells cycle between both zones. In addition, B-cell division was observed in both the dark and light zones during real-time imaging of GCs, which indicates that the DZ and LZ compartments are not as discrete as previously thought (reviewed in Allen et al., 2007, Wishnie et al., 2021).

The interaction between B-cells and Tfh cells is facilitated by co-stimulatory molecules and cytokines such as ICOS-ICOSL, CD40-CD40L, IL-4, and IL-21; interestingly, the duration of the contact between Tfh cells and GC B-cells is shorter than that between Tfh cells and extrafollicular B-cells (reviewed in Allen et al., 2007). Selected cells can return to the DZ for more SHM and proliferation; cells with intermediate affinity recycle to the DZ for further SHM and affinity maturation. It has been shown that the degree to which B-cell proliferates and acquires mutations once it re-enters the DZ depends on the amount of antigen a centrocyte presents to Tfh cell within the LZ (Li et al., 2018, reviewed in De Silva et al., 2015, Mesin et al., 2016, Verstegen et al., 2021). Alternatively, selected cells can undergo a phase of proliferation, and daughter cells will differentiate into memory cells which recirculate through the blood and the lymphoid organs and respond quickly upon the recurrence of the antigen months later. Some cells can become long-lived antibody-secreting plasma cells with up to 1000-fold higher affinity for their cognate antigen than the original B-cell that entered the GC (Fig.1.2). These plasma cells migrate to the bone marrow, where they may survive for years and provide long-term protection. Studies have shown that GC B-cells with higher affinity preferentially differentiate into long-lived plasma cells, whereas GC-derived memory B-cells are composed of cells with relatively lower affinity for the antigen. In addition, cytokines secreted by Tfh cells (IL-4 or IL-21) stimulate the fate determination of positively selected GC B-cells; intriguingly, Tfh cells with a higher density of CD40L express IL-4 which promotes plasma cell differentiation (Ise et al., 2018, reviewed in De Silva et al., 2015, Mesin et al., 2016, Magor, 2015, Verstegen et al., 2021).

Of note, SHM is a random process and can result in damaging mutations to the affinity of the Igs or even stop codons; likewise, SHM can lead to the production of B-cells secreting antibodies that

target self-antigens. These unwanted B-cells undergo pre-programmed apoptosis, a major feature of the B-cell program in the GC. GC B-cells have low expression levels of anti-apoptotic proteins such as BCL-2 and BCL-XL and high levels of the proapoptotic regulators BIM and Fas (CD95); therefore, they are prone to apoptosis, and the absence of survival signals result in their death (death by neglect; reviewed in Gatto et al., 2010). It was found that self-reactive GC B-cells can be successfully cleared when self-antigen is expressed nearby the GC microenvironment. While the deletion mechanism is not fully understood, it has been suggested that B-cells with strong BCR cross-linking alone die by apoptosis in the absence of co-stimulation signal (reviewed in Ise et al., 2019). Recently, it was demonstrated that apoptosis occurs in both the DZ and the LZ; while B-cells with non-functional Ig genes die in the DZ, B-cells in the LZ die in the absence of positive selection (Mayer et al., 2017). Apoptotic cells are rapidly cleared by tingible body macrophages (TBMs); TBMs are large macrophages that reside in the GCs within the secondary lymphoid organs and actively phagocytose B-cells labeled with MFGE8, which binds to both apoptotic cells using aminophospholipids such as phosphatidylserine and to integrins expressed by phagocytes (Gitlin et al., 2014, Kranich et al., 2008, Hanayama et al., 2002, reviewed in MacLennan et al., 1986).

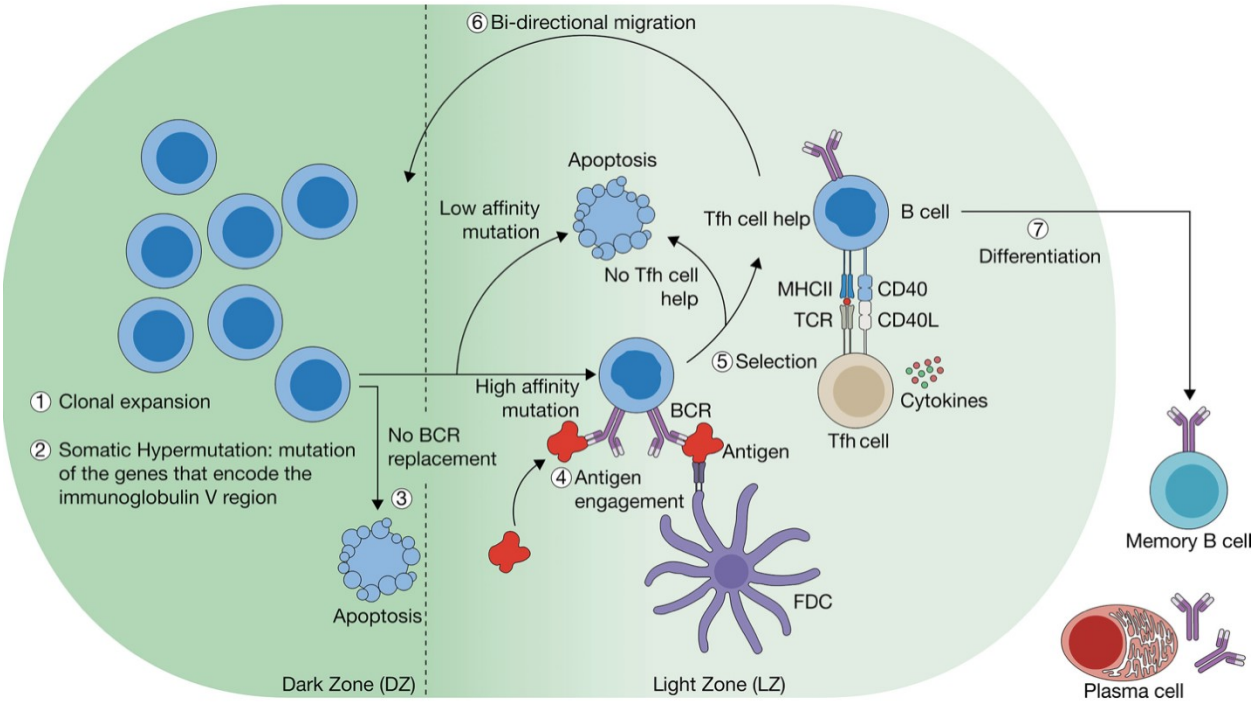


Figure 1.2. A mammalian germinal center reaction. Antigen-activated B-cells form the germinal center, in which they differentiate into highly proliferative centroblasts. During proliferation, centroblasts upregulate the expression of activation-induced cytidine deaminase (AID), which induces somatic hypermutations (SHM) in the VDJ exon to alter their antibodies affinity to the exogenous antigen. B-cells with non-functional Ig genes die in the DZ by apoptosis. Modified cells move from the dark zone to the light zone and begin again to express membrane antibodies. These cells, now referred to as centrocytes, may have altered antigen-binding abilities, and they compete for a limited number of intact antigens retained on the surface of follicular dendritic cells (FDCs). Centrocytes with high affinity for the antigen endocytose it and present it via major histocompatibility complex class II (MHC II) to T follicular helper cells (Tfh). Centrocytes with a higher affinity for the antigen present a higher density of peptide-MHC complexes and hence receive a higher share of T cell help through co-stimulatory molecules and cytokines and subsequently positive selection. Selected cells can return to the dark zone to undergo more SHM and proliferation. Selected centrocytes eventually differentiate into memory B-cells or plasma cells. Unsuccessful B-cells undergo pre-programmed apoptosis (adapted from Verstegen et al., 2021).

1.6. Activation induced cytidine deaminase (AID)

AID was discovered in 1999 by Tasuku Honjo's group using cDNA subtraction method after stimulating B-lymphocyte cell line (CH12F3) (Muramatsu et al., 1999). AID belongs to the Apolipoprotein B mRNA Editing Catalytic Polypeptide-like (APOBEC) family of RNA/DNA editing enzymes; it plays a significant role in diversifying B-cells within the GC, using AID knockout mice, studies have shown that AID is essential for both SHM and CSR (Muramatsu et al. 2000, reviewed in De Silva et al., 2015). Afterward, using chicken DT40 B lymphoma cell line, it was revealed that AID is essential for gene conversion (Arakawa et al., 2002). Initially, due to the sequence homology between AID and APOBEC1, mRNA-editing cytidine deaminases, it was thought that AID edits mRNA; subsequently, it became clear that AID deaminates cytidines (to uracils) in ssDNA, exposed during transcription (Di Noia and Neuberger, 2002, reviewed in Methot and Di Noia, 2017).

There are four conserved protein domains of AID, including nuclear localization signal (NLS) in the amino-terminal, catalytic domain, APOBEC-like domain, and nuclear export signal (NES) in the carboxy-terminal. These domains play a significant role in the function and regulation of AID by targeting specific cofactors, the carboxy-terminal is essential for CSR, and the amino-terminal is required to induce SHM (reviewed in Choudhary et al., 2018).

Interestingly, two cytidine deaminases designated CDA1 and CDA2 (cytidine deaminase 1 and 2) are present in lamprey (a jawless vertebrate; reviewed in McCurley et al., 2012). CDA1 and CDA2 generate the diversity of the variable lymphocyte receptors (VLRs) on the surface of lymphocyte-

like cells by a gene-conversion-like mechanism during lymphocyte development (reviewed in Sutoh et al., 2021). These deaminases have significant sequence homology to AID, which indicates that AID/APOBEC family members might have evolved from an ancient cytidine deaminase in the common vertebrate ancestor before the divergence of jawless and jawed vertebrates (reviewed in Trancoso et al., 2020). Additionally, several AID/APOBEC-like genes are present in the genome of the invertebrate, including echinoderm *Strongylocentrotus purpuratus* and the brachiopod *Lingula anatina* (reviewed in Trancoso et al., 2020). Phylogenetic analyses of the invertebrates AID/APOBEC-like genes revealed that these proteins cluster with the sea lamprey CDA2, indicating that CDA2 is more ancient than CDA1. Interestingly, it appears that the expression of AID/APOBEC-like genes in *S.purpuratus* is induced in response to pathogens, suggesting that these proteins are involved in the innate immune mechanisms, similar to APOBEC3 in mammals (reviewed in Knisbacher et al., 2016, Trancoso et al., 2020).

1.6.1. Regulation of AID

AID expression is tightly regulated at the transcriptional level by several proteins that can enhance or repress the expression of AID by binding to different regulatory regions within and around activation-induced cytidine deaminase (*Aicda*) gene, which encodes AID enzyme. A complex regulatory network regulates AID expression at the transcriptional level in mammals. It has been shown that AID transcription is regulated by four regions: Region 1 is located immediately upstream of the transcription start site (TSS), and it contains enhancer elements such as Sp-binding sites and a HoxC4-Oct motif, and weak suppressor element binding sites. Region 2 is within the first intron, and it also has both positive (Pax5-binding site and E-boxes) and suppressor element

binding sites. When combined, regions 1 and 2 suppress the expression of AID. Region 3 is located 17 kb downstream of *Aicda* gene, and it is required for AID expression in vivo; it has binding sites for Basic leucine zipper transcription factor (Batf), which controls CSR in B-cells (Tran et al., 2010). Lastly, region 4 is located upstream of AID promoter; it contains binding sites for NF-κB, STAT6, and Smad3/4 which are response elements for CD40L, IL-4, and TGF-β, respectively (Ise et al., 2011, reviewed in Zan et al., 2013).

AID is also regulated at the post-transcriptional level by the activity of micro RNAs (miRNAs). Five miRNAs, miRNA-93, miRNA-155, miRNA-361, miRNA-29b, and miRNA-181b, have been shown to repress the expression of human AID by targeting the mRNA encoding AID (Borchert et al., 2011, reviewed in Rios et al., 2020). Post-translational regulation of AID expression is mediated by protein kinases such as the Protein Kinase A (PKA). AID can be phosphorylated at serine-3 (Ser3), serine-38 (Ser38), threonine-140 (T140), and tyrosine-184 (Tyr184); phosphorylation can result in upregulation or downregulation of AID activity depending on which amino acid is phosphorylated (reviewed in Xu et al., 2007, Zan et al., 2013, Choudhary et al., 2018). In addition, AID activity is limited to early G1 phase of the cell cycle when the chromatin is decondensed, and thus it is regulated by the cell cycle (Wang et al., 2017).

AID is shuttled between the nucleus and cytoplasm by the nuclear localization signal (NLS) and nuclear export signal (NES) located in the amino and carboxy termini, respectively, and it is mainly localized in the cytoplasm; heat shock protein 90 (Hsp90) and elongation factor 1α (eEF1A) stabilize and keep AID in the cytoplasm. In addition, AID loses its stability within the nucleus and is targeted for proteasomal degradation (Ito et al., 2004, reviewed in Zan et al., 2013).

1.6.2. Function of AID

AID expression is required in B-cells to undergo SHM (non-templated mutation) and gene conversion (templated mutation) to diversify the secondary antibody repertoire and change the antigen-binding specificity in repones to antigenic stimulus; in addition, AID mediates CSR to alter the effector function of the antibody (Harris et al., 2002, Okazaki et al., 2002, Muramatsu et al. 2000, Arakawa et al., 2002, reviewed in Di Noia and Neuberger, 2007). AID inactivation completely blocks SHM, gene conversion, and CSR. In addition, previous studies have shown that the size of the GC and the number of GC B-cells increase in the absence of AID (Arakawa et al., 2002, reviewed in Wishnie et al., 2021).

In some species, including sheep, chickens, and rabbits, AID is involved in diversifying the primary antibody repertoire in an antigen-independent manner. While chickens use AID-mediated gene conversion solely to diversify their primary repertoire, rabbits utilize both gene conversion and SHM (reviewed in Lanning and Knight, 2015). In sheep, although at low rates, AID-induced SHM contributes to the diversity of the pre-immune antibody repertoire (reviewed in Jenne et al., 2006). Interestingly, in the nurse shark, AID is used to diversify the primary αβ T-cell repertoire by inducing SHM in the TCRα during thymic development (Ott et al., 2018). Subsequently, it has been shown that SHM contributes to the diversity of both TCRδ and TCRγ in addition to TCRα in the nurse shark (Ott et al., 2020).

AID plays a significant role in epigenetic modifications via DNA demethylation; it has the ability to deaminate 5-methylcytosine (5mC) in single-stranded DNA. This deamination process results in a thymine base opposite to guanine, the repair of the T: G mismatch leads to demethylation (reviewed in Dominguez et al., 2014). AID mRNA expression level is relatively high in oocytes and pluripotent stem cells; these cells can undergo epigenetic modifications, which indicates that AID plays a role in epigenetic reprogramming during development (Morgan et al., 2004). Furthermore, the depletion of AID /APOBEC enzymes in zebrafish embryos decreased DNA demethylation. On the other hand, overexpression of AID and MBD4 (a glycosylase that removes the thymine from a T: G mismatch during the repairing process) resulted in a widespread DNA demethylation of the embryo genome (Rai et al., 2008, Popp et al., 2010, reviewed in Dominguez et al., 2014).

Studies in mice and humans have shown that AID expression is crucial for B-cell central tolerance in both fetal liver and adult bone marrow. In the bone marrow, the expression of AID is thought to be induced in immature B-cells that express autoreactive BCRs and are undergoing receptor editing to undergo further rearrangement; if receptor editing results in autoreactive B-cells mutations induced by AID promote cell death and the deletion of autoreactive B-cells (Kuraoka et al., 2011, Cantaert et al., 2015). It has been suggested that AID-induced DNA lesions activate p53 function; P53 senses multiple types of DNA damage and leads to cell death (apoptosis; Cantaert et al., 2015). AID-deficient patients have impaired B-cell central tolerance because of their inability to remove developing autoreactive B-cells (Cantaert et al., 2015).

1.6.3. AID targeting

AID induced mutations are confined to a 1–2 kb downstream of the transcription start site in the Ig gene, and although the mechanism by which AID preferentially targets IgV or the switch (S) region for mutations is not fully understood, targeting can be attributed to different characteristics of the Ig gene (reviewed in Choudhary et al., 2018). Studies have shown that Ig super-enhancers increase AID activity; super-enhancers are 15-fold longer than typical enhancers, and they are mainly linked with highly transcribed genes (Whyte et al., 2013). These highly active enhancers result in higher chromatin accessibility of the associated genes, which provides a microenvironment suitable for AID activity (Qian et al., 2014). In addition, evolutionarily conserved elements in the Ig enhancers and enhancer-like elements known as Diversification Activator (DIVAC) efficiently target SHM to the Ig genes. DIVAC elements have multiple transcription factor binding sites; these sites are required for SHM targeting (Buerstedde et al., 2014). In addition to transcription, transcription stalling plays a key role in AID recruitment to its target. Upon transcription stalling, RNA polymerase II (RNAPII) remains associated with its processivity factor SPT5, which recruits AID and facilitates AID–RNAPII interaction. Interestingly, AID targets in the IgV and S regions are associated with stalled RNAPII and SPT5, which provides ssDNA for AID (Pavri et al., 2010, reviewed in Feng et al., 2020). Despite being nonessential, the IgV promoter enhances the efficiency of AID-mediated mutations (Shu et al., 2006). In addition, the high frequency of sequence motifs referred to as AID hotspot motifs (WRC/GYW, where W = A/T, R = A/G, and Y = C/T) in the IgV region and S region plays a critical role in recruiting AID (Zheng et al., 2005, reviewed in Choudhary et al., 2018).

AID off-target activity is also associated with highly transcribed genes and genes with super-enhancers that control the expression of several oncogenes (Lovén et al., 2013). Aberrant AID activity in the B-cell genome can target non-Ig genes, which causes reciprocal chromosomal translocation that frequently involves the Ig loci. It has been shown that abnormal AID activity targets multiple loci, including several proto-oncogenes involved in cell proliferation and apoptosis, such as Bcl2, Myc, BCL6, and ID3. AID off-targeting of these genes leads to human B-cell lymphomas such as Burkitt's lymphoma (translocation between IgH locus and c-myc) and Diffuse Large B-Cell Lymphoma (translocation between IgH and BCL2; Lohr et al., 2012, reviewed in Ramiro et al., 2004). In addition, dysregulation of AID activity can lead to non-lymphoid cancers, including pancreatic cancer, urothelial carcinoma, and breast cancer, by promoting the expression of genes associated with cancer progression through DNA demethylation (Yokoyama et al., 2016, Li et al., 2019).

1.7. Somatic hypermutation (SHM) within the dark zone

In humans and mice, the diversity of the secondary antibody repertoire is achieved primarily by SHM, this mutation process is initiated by AID, and both base excision repair (BER) and mismatch repair (MMR) pathways are involved in processing the uracils generated by AID (reviewed in Pilzecker et al., 2019). For reasons not fully understood, these repair pathways are directed to promote mutagenesis during SHM (reviewed in Feng et al., 2020). SHM is induced in the variable region of the IgH and IgL genes by up to 10^5-10^6-fold greater than the normal rate of mutation within the rest of the genome, up to 10^{-4} to 10^{-3} per base per division (reviewed in Liu et al., 2009). All four bases, including cytosine (C), guanine (G), adenine (A), and thymine (T), are targeted for

mutations, and both coding and noncoding strands are targeted for mutations (Maul et al., 2011). A previous study has shown that about 58 % of AID induced uracils were found in the non-transcribed strand, and 42 % of the uracils were observed in the transcribed strand (Maul et al., 2011). During SHM, transition (change from a purine to purine or from pyrimidine to pyrimidine) is favored over transversion (changes between purines and pyrimidines), with a ratio of transitions to transversions higher than the theoretical ratio (1:2) if mutations occurred randomly. In addition, during SHM motifs known as hotspot motifs (WRC/ GYW, WA/ TW, where W = A/T, R = A/G, and Y = C/T) accumulate more mutations compared to regular motifs (Rada et al., 2002, Zheng et al., 2005, reviewed in Di Noia et al., 2007). Interestingly, in the human Ig variable region, hotspot motifs are preferentially found in the CDRs, where replacement mutations are predominantly found to maximize the number of amino acid changes and perhaps alter the antigen-binding affinity. Also, amino acids that are encoded by only one or two codons, not four, are enriched in the CDR. Conversely, hotspot motifs are avoided in the structurally conserved FWRs, and the mutations found in FWRs are mostly silent after selection (Wagner et al., 1995, Kepler, 1997, Zheng et al., 2005).

Though centroblasts do not express surface antibodies within the DZ of the GC, they do transcribe their Ig genes. Within the transcription bubble, AID binds single-stranded DNA (ssDNA) and deaminates deoxycytidine (dC) preferentially within a WRC into deoxyuracil (dU); causing dU: dG mismatch. If not repaired, DNA replication over U, which is not a replication blocking lesion in the DNA molecule, using any DNA polymerase results in C to T transition due to the similarity between U and T, which will lead to C: G to T: A transition (Fig.1.3). During SHM, the irregular (U) base can be recognized in the DNA and processed by either the base excision repair (BER) or

noncanonical mismatch repair (ncMMR) pathways (Di Noia et al., 2002). It is believed that BER generates the majority of the mutations at C and G bases, and ncMMR pathway is responsible for most of the mutations at A and T nucleotides (reviewed in Neuberger et al., 2007).

During the BER pathway, deoxyuridine is excised from the DNA by uracil-DNA glycosylase (UNG), a component of the BER, and it is the glycosylase involved in SHM. This process results in an abasic site, translesion synthesis (TLS) polymerases such as Rev1, polymerase ζ, and polymerase η are recruited to bypass the non-instructive site; these polymerases lack proof-read activity and copy DNA with a lower fidelity which leads to transversions and transitions (Fig.1.3). However, Rev1 only inserts deoxycytidine monophosphate (dCMP) into DNA, and it is responsible for most G: C transversion mutations during SHM (Nelson et al., 1996, reviewed in Bahjat et al., 2017, Feng et al., 2020). It has been shown that monoubiquitination of the DNA sliding clamp proliferating cell nuclear antigen (PCNA) at lysine K164 (Ubi-PCNA) plays a significant role in recruiting TLS polymerases during SHM (Kannouche et al., 2004). The abasic site can be recognized and cleaved by other components of the BER, apurinic/apyrimidinic endonuclease (APE1 and APE2; Fig.1.3), which creates a nick in the phosphodiester backbone, then again Rev1 fills the nick to create mutations at C and G bases (reviewed in Krokan et al., 2013, Yu et al., 2019). Interestingly, in wild-type mice ($UNG^{+/+}$), SHM resulted in more nucleotide transitions (65 %) than transversions, and in $UNG^{-/-}$ mice, mutations significantly shifted to transitions (95 %) at dC/dG. However, $UNG^{-/-}$ mice have a normal mutation frequency at A/T pairs (Rada et al., 2002). During accurate genomic uracil repair, the removal of uracil by UNG is followed by an incision at the abasic site by APE1. Then DNA polymerase β fills the gap, and DNA ligase 3 or 1 seals the nick (reviewed in Krokan et al., 2013).

Alternatively, dU: dG mismatch can be recognized by MSH2/MSH6 mismatch recognition heterodimer, a component of the ncMMR machinery; afterward, a nick is made by PMS2 and MLH1 endonuclease complex, this nick provides an entry point for exonuclease-1 (Exo1), and a gap will be generated and extended by removing the mismatch and adjacent nucleotides (Bardwell et al., 2004). Then the gap is resynthesized by error-prone polymerases such as POLH, also known as pol η, a member of the Y-family of DNA polymerases. Pol η is responsible for most mutations at A: T and preferentially targets W<u>A</u>/<u>T</u>W for mutations on the non-transcribed and transcribed strands, respectively, and generates twofold more mutations of A than T by inserting G opposite T on the transcribed strand (Fig.1.3; Mayorov et al., 2005, reviewed in Pilzecker et al., 2019). Noticeably, patients with xeroderma pigmentosum variant (XP-V) disease who lack pol η have a significant decrease in mutations at A: T; however, the number of mutations at C: G increased, and the overall mutation frequency is normal in these patients (Zeng et al., 2001).

In addition, the ncMMR pathway can process the nick generated by the BER pathway. Exo1 is activated by components of the ncMMR and a 5′ gap, and it creates and extends patch of ssDNA from the incision made by APE2; subsequently, pol η fills the DNA gap. While mutations at A/T pairs slightly reduced in APE2$^{-/-}$ mice a dramatic reduction in mutations at A/T pairs was found in mice that lack both UNG and APE2 (Stavnezer et al., 2014, reviewed in Pilzecker et al., 2019).

Of note, a small number of mutations occur in B-cell Ig genes from patients who lack the expression of AID; however, WR<u>C</u>/<u>G</u>YW and W<u>A</u>/<u>T</u>W motifs are not targeted for mutations, and the mutations were significantly biased toward transitions which suggest that AID is involved in

the recruitment of polymerase η and UNG. Furthermore, the mutations were predominantly at G/C nucleotides, and it is believed that these mutations are caused by spontaneous cytidine deamination (Longo et al., 2008).

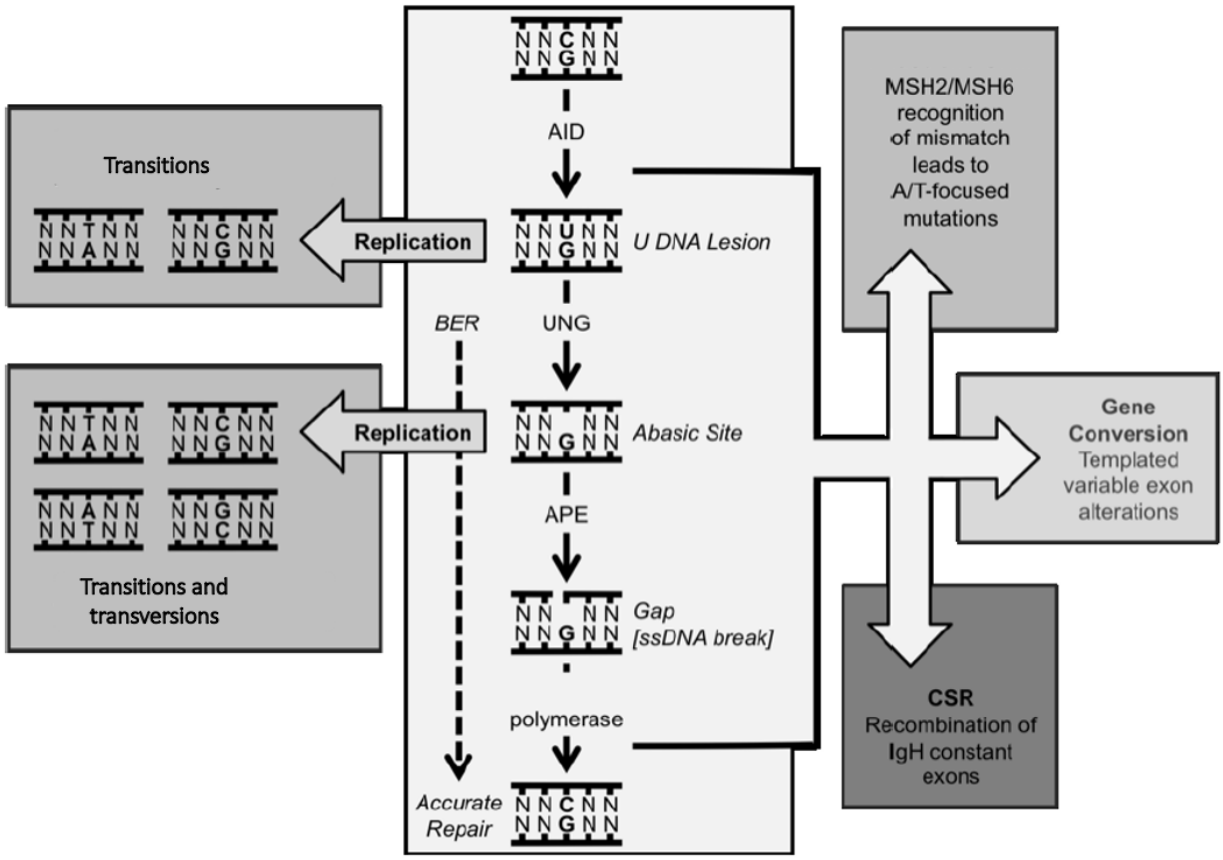


Figure 1.3. Schematic overview of antibody diversification by AID-mediated cytidine deamination and the subsequent repair mechanisms. Within the transcription bubble, AID binds single-stranded DNA (ssDNA) and deaminates C into U, causing U: G mismatch. DNA replication over U results in C to T and G to A transitions. The removal of the U during the repairing process by uracil-DNA glycosylase (UNG) will lead to the generation of an abasic site; replication opposite to the abasic site will lead to transitions or transversions. The abasic site can be recognized and cleaved by apurinic/apyrimidinic endonuclease (APE), which creates a nick in the phosphodiester backbone. The resulting single-stranded break (SSB) could directly result in gene conversion in which pseudogenes upstream of the rearranged IgV are used as donors for sequences. Incisions on opposite strands could lead to DSBs. Reinsertion of cytidine by high fidelity DNA polymerases will repair the lesion. U: G can be recognized by MSH2/MSH6, a component of the noncanonical mismatch repair (ncMMR) pathway, which will lead to A/T focused mutations. ncMMR pathway can also lead to CSR (adapted from Saunders, PhD thesis, U of A, 2020, reviewed in Pilzecker et al., 2019, Di Noia et al., 2007).

1.8. Gene conversion (GCV)

Gene conversion (GCV) is a non-reciprocal exchange of gene segments between a donor and recipient homologous DNA sequences (reviewed in Helleday et al., 2003). Some species use GCV with or instead of SHM to diversify their Ig genes. Species such as birds, rabbits, cows, and pigs have limited functional V genes to diversify their variable domain of the heavy and light chains; however, they have a large number of V pseudogenes (ΨV) (reviewed in Choudhary et al., 2018). These pseudogenes are used as a donor sequence during GCV. GCV depends almost exclusively on the BER pathway of the irregular U: G mismatch, inhibition of UNG reduces GCV events in the chicken DT40 B lymphoma cell line (Di Noia et al., 2004). In addition, GCV requires homologous recombination (HR) factors, including the RAD51 paralogues (XRCC2, XRCC3, Rad51B, Rad51C, Rad51D), and the deletion of these factors shifts mutation towards SHM in the DT40 B lymphoma cell line (Sale et al., 2001). During GCV, it is believed that the lesion generated by AID is repaired by BER and a recombinational repair process using the V pseudogenes upstream of the rearranged IgV (reviewed in Di Noia et al., 2007).

1.9. Class switch recombination (CSR)

Class switch recombination (CSR), also known as isotype switching, occurs within or outside the GCs after the first B-cell T-cell interaction at the border of the T- and B-cells zones. In addition, upon B-cells stimulation with the appropriate cytokines such as B-cell activating factor (BAFF) and IL-10, CSR occurs independently of T-cell help (Litinskiy et al., 2002). A study using a murine fibroblast cell line transfected with an artificial CSR substrate and a plasmid containing AID

complementary DNA has revealed that the only B-cell-specific factor necessary for CSR initiation is AID (Okazaki et al., 2002). Interestingly, studies have shown that CSR occurs largely outside the GCs and before SHM this could be due to the fact that CSR requires a lower expression level of AID compared to SHM (Peakman et al., 1998, Roco et al., 2019). CSR is an intrachromosomal DNA deletion and it is required to change the effector function of the Ig while maintaining the specificity of the variable region (reviewed in Stavnezer et al., 2008, Manis et al., 2002).

Mammalian Ig gene contains several constant (C) regions located downstream of the rearranged V region in the heavy chain (CH); each C_H region (except Cδ) is preceded by a G-rich repetitive region known as the switch (S) region. Each S region is associated with its own promoter and enhancer; thus, cytokines and T-cell help induce the selection of the appropriate isotype during an immune response by driving transcription and recruiting AID (Tangye et al., 2002, reviewed in Matthews et al., 2014, Yu et al., 2019). However, the Sμ promoter is constitutively transcribed and unaffected by cytokines, which allows IgM expression in unstimulated B-cells (Li et al., 1994). Mature naïve B cells express IgM and IgD on their cell surfaces; IgD expression occurs by alternative splicing of the Cμ-Cδ genes. Also, CSR to IgD can occur through the Sμ and a switch-like region 5′ to Cδ known as σδ; however, IgD CSR is infrequent and occurs only in mucosal tissues (Choi et al., 2017).

During CSR, AID-induced deoxyuridine is converted into DNA double-strand breaks (DSBs) which are essential for CSR. AID targets C in the S regions for deamination, and both BER and ncMMR pathways are involved in CSR. During BER, UNG removes AID-induced deoxyuracil, then APE1 generates a nick; incisions on opposite strands could lead to DSBs. The dU: dG

mismatch can be recognized by the ncMMR pathway, MSH2/MSH6 heterodimer recognizes the dU, then PMS2 and MLH1 endonuclease complex is recruited to generate a nick. Subsequently, Exo1 generates a gap, breaks on opposite strands could generate DSBs (reviewed in Yu et al., 2019, Matthews et al., 2014). These DSBs must be generated within the donor Sμ region and an acceptor Sx region for CSR to occur. DSBs are joined using proteins that perform nonhomologous end-joining (NHEJ) in all cell types. Mutations in components of the NHEJ pathway, including Ku70/Ku80 proteins and DNA ligase IV, led to a significant decrease in CSR (Casellas et al., 1998, Han et al., 2008, Boboila et al., 2010).

Patients with mutations in AID have an autosomal recessive form of hyper-IgM syndrome type 2 (HIGM2); these patients have normal or high levels of serum IgM; however, all the other Ig isotypes are absent because CSR is blocked (Revy et al., 2000).

1.10. Adaptive immune system in teleost fish

Teleost fish is the most diverse and largest group of extant vertebrates. In addition to having an established innate immune system, bony fish have a well-developed adaptive immunity. In teleost fish, mucosal surfaces such as the skin, gut, and gills act as the first line of defense, and their mucus contains several effector molecules such as lectins, lysozymes, complement proteins, antibacterial peptides, and IgM, which prevent pathogens invasion. However, if the barrier is breached, effector immune cells can be activated (reviewed in Peatman et al., 2015). Several studies have described and verified the function of the innate immune cells in teleost fish, including macrophages, neutrophils, dendritic cells, mast cells, basophils, eosinophils, and natural killer (NK) cells; also,

cytokines which are essential for both innate and adaptive immunity are highly conserved across vertebrates (Dezfuli et al., 2008, Lugo-Villarino et al., 2011, reviewed in Bruce et al., 2017).

A canonical recombination activation gene (RAG) -mediated adaptive immunity, including rearranged Ig, TCR, as well as the major histocompatibility complex class I and II (MHC I and MHC II), arose first in the placoderms, an extinct vertebrate class. Cartilaginous fish is the oldest living group with Ig, TCR, and MHC-based immunity. However, lymph nodes, germinal centers, and FDCs are absent in ectotherms (Bernstein et al., 1996, reviewed in Flajnik, 2018). It is believed that lymphocyte-based adaptive immunity evolved in a common ancestor of all vertebrates (reviewed in Flajnik, 2018). Nonetheless, in Atlantic cod and pipefish, MHC II genes are missing, and it was suggested that these species depend on their T-cell independent Ab response (reviewed in Magadan et al., 2015). Interestingly, loss of critical components of the adaptive immunity and thus the absence of a functional adaptive immune system was reported in the deep-sea anglerfish (*Photocorynus spiniceps*; reviewed in Isakov, 2022).

In cartilaginous fish, the primary lymphoid organs are the epigonal organ and the Leydig organ (the bone marrow equivalent), and the thymus. In addition, these species are the oldest living group that have the primordial secondary lymphoid organ, the spleen. The spleen of cartilaginous fish is divided into distinct red and white pulps; however, no defining borders are found between the red and white pulps. The white pulp of the spleen is the major secondary lymphoid tissue in cartilaginous fish, where densely packed lymphocytes that surround the arterioles are found (Rumfelt et al., 2002, reviewed in Neely et al., 2016). In amphibians, the primary lymphoid organs are the bone marrow and thymus. The secondary lymphoid organ, spleen, is separated into red and

white pulps (reviewed in Neely et al., 2016). Interestingly, while amphibians lack FDCs, it has been shown that DCs have characteristics of both FDCs and conventional DCs. These DCs (named XL cells) have a high expression level of MHC class II and trap native antigen at the cell surface (Neely et al., 2018).

The primary lymphoid tissues in teleost fish are the thymus and anterior kidney (head); the head kidney is the teleost bone marrow equivalent; it is where B-cell development and maturation occur. The spleen and various mucosa-associated lymphoid tissues (MALTs) are secondary lymphoid organs, where B-cells can be activated and differentiate into plasmablasts and plasma cells (reviewed in Bjørgen et al., 2021). Separation of the red and white pulps is found in zebrafish spleen, where macrophages surround the ellipsoids in the white pulp, with accumulation of lymphocytes (reviewed in Neely et al., 2016). In addition, the posterior kidney is a hypothesized (Zwollo et al., 2005) yet unproven secondary lymphoid organ of fish. Using transcription factors that are differentially expressed during B-cell activation (Pax-5 and Blimp-1) and the B-cell mitogen (lipopolysaccharides (LPS)), it has been shown that the posterior kidney has a high number of partially activated B-cells. In addition, it was revealed that the posterior kidney has the highest capacity to generate plasma cells following activation with LPS compared with the anterior kidney, spleen, and blood (Zwollo et al., 2005).

Similar to the mammalian long-lived plasma cells, which reside in the bone marrow, long-lived plasma cells reside mainly in the anterior kidney in teleost fish. Short-lived, dividing plasmablasts were found in the peripheral immune tissues in rainbow trout (Bromage et al., 2004, Wu et al., 2019). MHC II expression has been characterized on macrophages and B-cells, and the highest

level of MHC II expression was detected in the spleen and kidney of zebrafish, where 70% of MHC II expressing cells are B-cells (Wittamer et al., 2011, reviewed in Lewis et al., 2014). Moreover, DC-like cells in zebrafish express genes associated with DCs function and antigen presentation in mammals (Lugo-Villarino et al., 2010). Recent studies have shown that fish B-cells can be activated by CD40L, which indicates that fish B-cells are able to respond to T-cell dependent antigens like mammalian B-cells (Granja et al., 2019). In addition, it has been shown that cytokines commonly secreted by T-helper cells enhance the survival and proliferation of antibody-secreting B-cells in rainbow trout (Gorgoglione et al., 2013, Abos et al., 2020).

Like mammalian B-1 B-cells, fish B-cells have phagocytic and microbicidal activities, and they have the ability to present the internalized antigen to CD4 T-cells (reviewed in Wu et al., 2020). However, their phagocytic capacity decreases as they differentiate into antibody-secreting cells (Wu et al., 2019). Zebrafish B-cells have a strong antigen-presenting ability of both soluble antigen and bacterial particles. In addition, they can prime naïve CD4 T-cells proliferation, which indicates that fish B-cells play a significant role in initiating the adaptive immune response (like mammalian dendritic cells) (Zhu et al., 2014).

Collectively, data from rainbow trout and channel catfish led scientists to propose a model for B-cell development and activation in bony fish where B-cells mature in the anterior kidney; these mature naïve B-cells migrate to the peripheral lymphoid organs, i.e., the spleen and posterior kidney. Following antigen encounters, activated B-cell differentiates into short-lived plasmablast (Pax5+ mIgM+ (membrane IgM)), which provides the initial humoral immune response. A subset of these plasmablasts proliferates and differentiates into plasma cells (Blimp-1+ mIgM^{-}). These

plasma cells can home to the anterior kidney, where they become long-lived plasma cells and secrete high-affinity antibodies to provide antigen-specific long-term protection (Zwollo et al., 2005, Barr et al., 2011, Wu et al., 2019).

1.10.1. Ig in teleost fish

Following the divergence from the invertebrate lineage, two whole-genome duplication events (2R-WGD) took place in the genome of early vertebrates; in addition, local duplication events occurred in many teleost fish species (reviewed in Kuraku, 2013, Flajnik et al., 2010, Hikima et al., 2011). Consequently, the organization of the IgH and IgL in teleost fish is complex and varies between species (reviewed in Hikima et al., 2011).

Ig genes appeared first in jawed vertebrates, where all lineages have IgM and IgD isotypes (in cartilaginous fish, IgD is referred to as IgW; reviewed in Flajnik., 2018). In teleost fish, three Ig isotypes are found; in addition to the canonical IgM and IgD, they have a new isotype IgZ/IgT (for zebrafish/ rainbow trout or teleost) (Danilova et al., 2005, Hansen et al. 2005, reviewed in Hordvik,1998). Recently, the symbol for IgZ heavy chain in zebrafish has been officially changed to IgT (Dornburg et al., 2021). Some teleost fish, such as channel catfish and medaka, lack IgZ/IgT (Bengten et al., 2006, Magadan-Mompo et al., 2011). There are four subsets of B-cells in teleost fish, including IgM^+ IgD^+, IgM^+ IgD^-, IgD^+ IgM^-, and IgT^+ IgM^- IgD^- B cells (Perdiguero et al., 2019, reviewed in Magadan et al., 2015). Similar to higher vertebrates, bony fish diversify their primary repertoire by combinatorial, junctional diversity, and the association of different light and heavy chains (reviewed in Bilal et al., 2021). In addition, allelic exclusion, which allows each B-

cell to express one type of antibody, has been demonstrated in cartilaginous fish, bony fish, and amphibians (reviewed in Hsu, 1998, Flajnik., 2018).

IgH locus in teleost fish has a modified translocon arrangement comparable to the organization found in mammalian and other higher vertebrate IgH and IgL loci (Fig.1.4; Danilova et al., 2005, Ghaffari et al., 1997, reviewed in Flajnik, 2018). In zebrafish, IgH locus consists of 40 functional VH segments, common to all isotypes, followed by two D, two J, and C regions of IgZ/IgT (Dζ-Jζ-Cζ) located between the VH segments and five Dμδ, five Jμδ, Cμ and Cδ gene segments, which encode both IgM and IgD (Fig.1.4). This genes organization indicates that IgZ/IgT diversity (40 VH X 2 Dζ X 2 Jζ) is limited compared to IgM (40 VH X 5 Dμ X 5 Jμ). Alternative splicing is responsible for expressing IgD. However, the expression of IgZ/IgT occurs during the rearrangement process; thus, the expression of IgM isotype will result in the deletion of the Dζ, Jζ, and Cζ gene segments, and hence IgM and IgZ/IgT are expressed by two distinct populations of B-cells (Danilova et al., 2005). Interestingly, there is more than one copy of the IgH locus in most teleosts; nevertheless, zebrafish have a single copy of this locus (Danilova et al., 2005, reviewed in Hordvik,1998).

Unlike IgH locus, the light chain genes are organized in multiple clusters (multi-cluster configuration) of VL, JL, and CL genes in teleost fish, similar to the organization of both the IgH and IgL genes in cartilaginous fish (reviewed in Hsu, 1998). In teleost fish, there are four IgL isotypes, λ, κ (L1, L3), σ (L2), and σ-2. In addition to λ and κ, which are found in mammals, they have a primordial IgL isotype (σ) (reviewed in Hikima et al., 2011, Criscitiello et al., 2007). Zebrafish have three light chain types: L1, L2, and L3 (Zimmerman et al., 2008).

The primary antibody isotype in most jawed vertebrates is IgM; it is the first Ig to be expressed during development (reviewed in Buonocore et al., 2016). In zebrafish, IgM expression can be detected as early as 20 days post-fertilization (Page et al., 2013). In bony fish, the membrane form of IgM consists of three Cμ domains where the transmembrane (TM) domain is alternatively spliced to the end of Cμ3, thus splicing out the Cμ4 domain (reviewed in Hikima et al., 2011). The secreted form of IgM heavy chain contains four constant Ig domains encoding exons (μ1–μ4). Tetrameric IgM is the dominant serum Ig in teleost, as opposed to the pentameric mammalian IgM; however, there is a significant level of diversity in the degree of inter-heavy chain disulfide polymerization in teleost IgM, including monomers, dimers, trimers, and tetramers. It has been shown that antibodies with a higher affinity for the antigen have the highest degree of polymerization (Ye et al., 2010). It is thought that post-translational processes based on the affinity of BCR interaction with antigens determine the degree of inter-heavy chain disulfide polymerization. High affinity interactions promote glycosylation of the C-terminal and prevent demannosylation, thus increasing disulfide bonding activity (reviewed in Ye et al., 2011). IgM has roles in both innate and adaptive immunity, similar to mammalian IgM. Teleost fish IgM has many effector functions, including complement activation, which contributes to pathogen opsonization, agglutination for phagocytosis and removal of pathogens, and antibody-dependent cellular cytotoxicity (ADCC; reviewed in Ye et al., 2013).

IgD in teleost fish forms a hybrid of Cμ1 and different numbers of Cδ domains. In zebrafish, IgD heavy chain consists of 16 constant Ig domains and Cμ1 spliced to the third IgD exon (Cδ3.1) (Zimmerman et al., 2011). Like in mammals, two types of IgD^+ B-cells are found in catfish,

IgM^+/IgD^+ and IgM^-/IgD^+, and interestingly all secreted IgD were without a V-region (named V-less) where the leader sequence spliced to the Cδ1 domain (Edholm et al., 2010). The V-less secreted IgD is suggested to function as a pattern recognition molecule by tagging pathogens for removal (Edholm et al., 2010). Also, recently, it has been shown that IgD is involved in the humoral immune response in rainbow trout infected with *Tetracapsuloides bryosalmonae* (Abos et al., 2018).

A third Ig isotype was discovered in some teleost fish species in 2005, and like IgD, the number of Cζ/τ domains varies among teleost fish. In zebrafish, IgZ heavy chain contains four constant Ig domains encoding exons (ζ1- ζ4; Danilova et al., 2005, Hansen et al., 2005). Similar to mammalian IgA, IgZ/IgT functions in mucosal immunity against parasitic and bacterial infection in mucosal tissues, including the intestine, skin, and gill; IgZ/IgT B-cells constitute 54.3% of all trout B-cells in the gut while they only represent 16–28% of B-cells in the blood, spleen, and head kidney in rainbow trout (Savan et al., 2005, Zhang et al., 2010). However, IgZ/IgT is unrelated to IgA, and they are phylogenetically distant (reviewed in Castro et al., 2014). Also, IgZ/IgT is the main Ig isotype that coats *Tetracapsuloides bryosalmonae* (a myxozoan parasite) during the immune repones in rainbow trout with proliferative kidney disease (PKD) (Abos et al., 2018). Subsequently, a second subclass of IgZ has been identified (IgZ2) (Hu et al., 2010). A recent study showed that IgZ and IgZ2 have different tissue distribution and immune responses; IgZ was found to play a major role in both local mucosa and systemic immune responses, while IgZ2 mainly contributes to local mucosal immune responses (Ji et al., 2020).

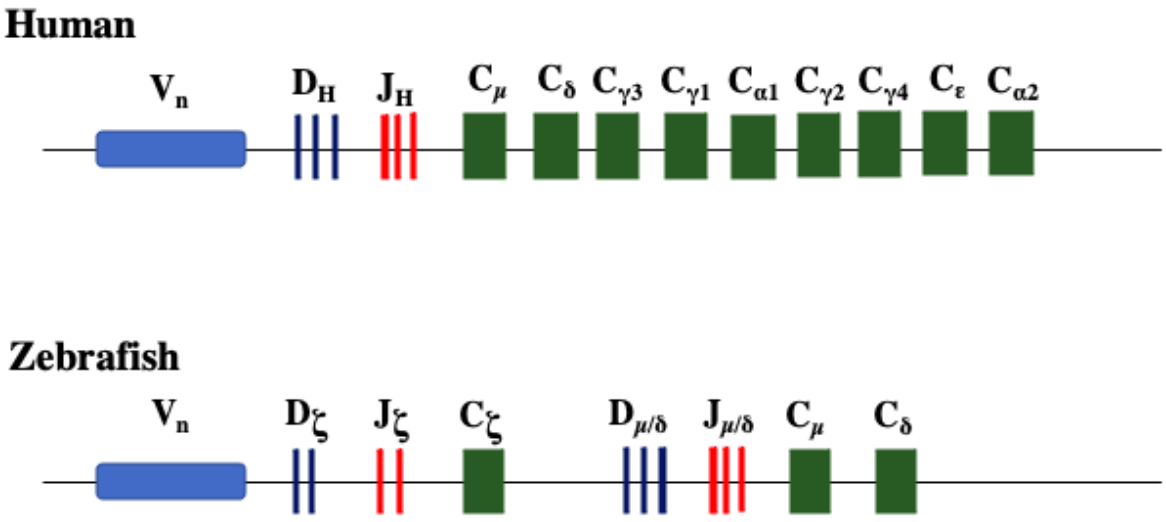


Figure 1.4. Schematic overview of the immunoglobulin heavy chain locus organization in humans and zebrafish. Human IgH locus has a translocon organization, while zebrafish IgH locus has a modified translocon configuration. V, variable; D, diversity, J, joining; C, constant (reviewed in Bilal et al., 2021).

1.10.2. T-cell receptors (TCR) in teleost fish

The expression of αβ and γδ TCRs and their associated coreceptors has been described in many species of bony fish (reviewed in Flajnik, 2018, Scapigliati., 2013, Bajoghli et al., 2019). In addition, almost all the genes associated with T helper cell subsets and cytotoxic T-cells in mammals have been identified in fish (reviewed in Scapigliati., 2013, Barraza et al., 2021, Nakanishi et al., 2011). In bony fish, T-cells play a significant role in the immune response to viral infections; the distribution and transcription factors of T-cells are regulated in response to viral infection in rainbow trout, and virus-specific cytotoxic T-cells were found in ginbuna carp infected with carp haematopoietic necrosis virus (CHNV) (Leal et al., 2016, Somamoto et al., 2009).

1.10.3. AID expression in teleost fish

In teleost fish, the expression of AID, which is the key initiator of affinity maturation, was first reported in channel catfish (Saunders et al., 2004); subsequently, several other fish, including zebrafish AID homologues, have been reported (Zhao et al., 2005). Analysis of AID sequences from birds, amphibians, and fish revealed that many conserved residues, which are functionally important for the activity of AID, are found in these species (Zhao et al., 2005). Using channel catfish, it has been shown that fish AID homologue shares 57% amino acid identity and 73% similarity with the AID proteins of mouse and human (Saunders et al., 2004). Also, AID homologue in dogfish (*Scyliorhinus caniculus*) shows 79% similarity to human AID (Conticello et al., 2005). In bony fish, but not in shark and amphibian, the region of least conservation in AID is the catalytic domain; catfish, zebrafish, and fugu each has 8 or 9 additional residues in the

catalytic domain of AID. However, this difference in the catalytic domain of bony fish AID does not seem to affect its deaminase activity in either eukaryotic or prokaryotic cells (reviewed in Barreto et al., 2011).

Aid enzyme in bony fish is temperature sensitive and works best at natural physiological temperatures; nevertheless, it is fully functional. Zebrafish and catfish AID can induce both CSR and SHM in mouse $AID^{-/-}$ B-cells, even though fish lack CSR (Dancyge et al., 2012, Wakae et al., 2006). CSR, in which a single B-cell has the ability to alter its Ig effector function, appeared first in amphibians (reviewed in Flajnik et al., 2010). Furthermore, in bony fish, AID expression can be induced by stimulation with molecules such as phorbol ester and calcium ionophore in catfish IgM^{+}/IgD^{-} cell line, and this stimulation leads to the accumulation of point mutations within the expressed VDJ exon (Saunders et al., 2010). Remarkably, zebrafish AID is more catalytically active in comparison to human and catfish AID (Dancyge et al., 2012).

It has been shown that AID NES domain in the carboxy-terminal has a pleiotropic function; in addition to being essential for transporting AID out of the nucleus, it is required for CSR. Deleting the carboxy-terminal of AID blocks its ability to perform CSR (reviewed in Choudhary et al., 2018, Barreto et al., 2011). Patients with truncated AID carboxy-terminal have Hyper IgM syndrome type 2 (Imai et al., 2005). In addition, truncated AID carboxy-terminal results in an increased mutation frequency and unregulated SHM (Ito et al., 2004). These observations suggest that the ability of fish AID to perform CSR is a by-product resulting from the functional overlap in AID carboxy-terminal (Wakae et al., 2006). In addition, certain residues in the amino-terminal

are specifically required to induce SHM; mutations in the amino-terminal have been shown to alter mutational specificity (Shinkura et al., 2004, Bransteitter et al., 2004).

AID expression in bony fish is transcriptionally regulated by five cis-regulatory regions within or flanking *Aicda* gene locus. Individually these regions are suppressive, but when some regions are coupled, they work as enhancer elements (Villota-Herdoiza et al., 2013). In addition, like mammalian AID, teleost fish AID is predominantly localized in the cytoplasm (Wakae et al., 2006). Overall, teleost fish and mammalian AID share most of their functional characteristics.

1.10.4. Affinity maturation in teleost fish

Fish were thought to lack affinity maturation in part because they lack histologically distinguishable germinal centers. In addition, early studies on the humoral immune response in fish revealed that there is no strong affinity maturation in fish, and they concluded that the slight increase in antibodies binding affinity in fish is simply due to an increase in the number of antigen-specific B-cells without clonal expansion or SHM (Arkoosh et al., 1991). In the nurse shark, all the antibodies isolated over 12 months following immunization with pneumococcal vaccines were of low average affinity (Shankey et al., 1980). Similarly, antibodies produced in the horned sharks following immunization with 2-furyloxazolone coupled to *Brucella abortus* had low affinity, and the affinity of the antibodies did not improve over a period of 25 days (Makela and Litman, 1980). Subsequently, a modest increase in antibody affinity (2-3 or 5-fold) in response to a T-cell dependent antigen was found in rainbow trout a few weeks post-immunization; in these studies, the modest increase in antibody affinity is thought to be a result of the selection and proliferation

of B-cells with natural higher antigen-binding affinity without the involvement of SHM (Cain et al., 2002, Kaattari et al., 2002). Likewise, earlier studies of affinity maturation in *Xenopus* found only a 5- to 10-fold increase in antibody affinity in response to haptenated keyhole limpet hemocyanin (DNP-KLH) four weeks after immunization (Wilson et al., 1992).

Later studies revealed that the techniques used to study affinity maturation in earlier studies assume that the affinity of the antibodies is normally distributed and might not detect alterations in specific subpopulations of B-cells following vaccination. Using affinity-based immunopartitioning assay, a shift towards the production of higher affinity antibodies (100-fold) by week 20 post-immunization was found in rainbow trout in response to TNP-keyhole limpet hemocyanin (TNP-KLH; Ye et al., 2011). A similar shift in antibodies affinity was reported recently in channel catfish in response to TNP-KLH by week two post-immunization, reaching maximum affinity at week ten after immunization (Wu et al., 2019). Subsequently, it was shown that these high-affinity antibodies are secreted by long-lived plasma cells (LLPCs). In addition, LLPCs in channel catfish migrate and reside mainly in the anterior kidney (bone marrow equivalent) later in the immune response (Wu et al., 2019, Wu et al., 2019). Despite the modest increase in antibody affinity in fish when compared to mammals (1000-fold increase), these studies confirmed that antibody affinity maturation process occurs in teleost fish.

Of note, B-cells in teleost fish have the ability to structurally modify their antibodies based on the binding affinity to the antigen. In rainbow trout, vaccination with TNP-KLH resulted in both higher binding affinity and degree of polymerization; interestingly, a higher level of disulfide polymerization is associated with an increased half-life (Ye et al., 2010).

Although the mechanism which leads to affinity maturation in teleost fish was a topic of debate in earlier studies, the development of immunological memory was established in these species (Avtalion et al., 1969). A faster and stronger secondary immune response was reported in several fish compared to the primary response, which revealed the presence of immunological memory (Tatner et al., 1987, Cossarini-Dunier et al., 1986). For example, in rainbow trout, injection with *Aeromonas salmonicida* (AS) or *Yersinia ruckeri* resulted in an enhanced secondary immune response upon re-exposure to the same antigen (Tatner et al., 1987, Cossarini-Dunier et al., 1986). However, the mechanism by which immunological memory develops in teleost fish has not been fully understood (reviewed in Stosik et al., 2021).

1.10.5. SHM in teleost fish

SHM is one of the hallmarks of affinity maturation in mammals. In ectotherms, SHM was first shown in frogs and sharks (Wilson et al., 1992, Diaz et al., 1998). Accumulation of mutations in the VDJ exon and evidence for clonal expansion were reported in *Xenopus laevis* following immunization with haptenated keyhole limpet hemocyanin (DNP-KLH), and the mutation rates were comparable to those reported in mice. Unlike mammals, there was a strong bias for mutations at G: C base pairs (41 of the 45 mutations were at G: C pairs; Wilson et al., 1992, reviewed in Hsu et al.,1998). The first evidence of SHM in sharks was reported in the horned shark using genomic and cDNA libraries; they observed mutations in the V, D, and J gene segments but not in the C_H region. C_H region sequences were identical to their corresponding germline sequences (Hinds-Frey et al., 1993). In nurse sharks, SHM was reported in the immunoglobulin new antigen receptor

(IgNAR) transcripts, and unlike the *Xenopus* Ig mutations, there was no bias for mutations at G: C base pairs and the mutation frequency was similar to that detected in human and mouse Ig (Diaz et al., 1998). Analysis of 1023 mutations in the IgNAR revealed a significant bias toward transition mutations, and only a few deletions and insertions were observed, similar to SHM in mammalian Ig (Diaz et al., 1999). Interestingly, unlike mutations found in Ig genes in other species, tandem point mutations are prevalent in the shark Ig mutations, and they occur predominantly within hotspot motifs (Diaz et al., 1999). Using the nurse shark NS3 light chain gene, which exists in the germline as joined VJ, analysis of 631 substitutions identified in 90 clones revealed that more than half the mutations (338) were found in tandem on adjacent nucleotides (stretches of 2-4 bases; Lee et al., 2002). The perinatal light chain sequences were not mutated; however, sequences from adult sharks were highly mutated (Lee et al., 2002). In addition, examination of the pattern of mutations revealed that the mutations arose from hypermutations rather than gene conversion (Lee et al., 2002). Furthermore, analysis of the mutations in the nurse shark productive and non-productive heavy and light chain genes suggested that tandem mutations result from repair by error-prone polymerases, which can generate two or more mutations before dissociating (Zhu et al., 2010). Afterward, it was suggested that these tandem mutations arise through the activities of Pol ι and Pol ζ, where Pol ι accesses the repair gap and introduces the first mutation, then Pol ζ synthesizes the second mismatch due to its unique catalytic property (Malecek et al., 2005, Saribasak et al., 2012, Maul et al., 2016). This was supported by the finding that the first mutation was predominantly A to G or A to T substitution, the two most mutations produced by Pol ι (Maul et al., 2016).

Subsequently, several studies confirmed that SHM occurs in the Ig genes of bony fish. For example, SHM was identified in unvaccinated catfish using Ig heavy μ-chain from the spleen, and the mutations were predominantly at G and C nucleotides with a bias for transition mutations (Yang et al., 2006). Their analysis of the distribution of replacement and silent mutations within the Ig gene indicated that there was no evidence of antigen-driven B-cell selection, and they concluded that perhaps SHM evolved to increase the diversity of the antibody repertoire (Yang et al., 2006). Moreover, a study using zebrafish IgL chain revealed that SHM occurs mainly at AID hotspot motifs WRCH/DGYW, similar to SHM in mammals; however, unlike in mammals, the ratio of transition to transversion mutations was much higher (5.64) than the theoretical value, and they concluded that the role of UNG is limited in zebrafish (Marianes et al., 2011). In addition, using high throughput sequencing and unvaccinated zebrafish, it was found that mutations in the VDJ exon increase with age, likely as a result of clonal expansion and accumulation of mutations in the Ig genes of activated B-cells in response to pathogens (Jiang et al., 2011).

Somatically mutated and clonally expanded IgZ/IgT and IgM B-cells were found in rainbow trout in response to parasitic infection (*T. bryosalmonae*), and most of the mutations were within AID hotspot motif (WRCY). Notably, the mutation rate increased as the immune response progressed, which was associated with an increased level of AID mRNA (Abos et al., 2018). Their analysis of the mutations at WA/TW revealed that although the mutations at WA/TW motifs were significantly lower than the mutations at WRCY motif, the mutations at WA increased throughout the immune response (Abos et al., 2018). A study in rainbow trout showed that in response to commensal bacteria, $IgD^{+}IgM^{-}$ B-cells clonally expand and mildly mutate their Ig gene in both the gills and gut, and the mutations at WRC/GYW and WA/TW increased during the clonal expansion

(Perdiguero et al., 2019). Furthermore, the distribution of replacement and silent mutations in the CDR2 and FWR3 suggested that antigen-mediated mucosal IgD selection occurs in rainbow trout (Perdiguero et al., 2019).

Even though studies of the humoral immune response in teleost fish found a modest increase in affinity following immunization (up to 100-fold; Wu et al., 2019, Ye et al., 2011); evidence of clonal lineage expansion has been reported in several fish species including, channel catfish, zebrafish, and rainbow trout (Yang et al., 2006, Marianes et al., 2011, Castro et al., 2013). For example, in rainbow trout, using CDR3 spectratyping analysis, a method used to demonstrate clonal expansion based on the shifts of the relative frequencies of B-cells with specific CDR3 size, it was shown that the immune response to the Viral Hemorrhagic Septicemia Virus (VHSV) was dominated by a few proliferating clones of IgM expressing B-cells (Castro et al., 2013). Similarly, using Ig sequencing, expansion of specific B-cell clones in response to vaccination with bacterial components was detected in Atlantic salmon (Lund et al., 2019).

Taken together, these studies demonstrated that mutations occur in a lineage-specific manner in bony fish, which indicates the presence of a germinal center-like reaction; in GC centroblasts, proliferation is associated with the accumulation of SHM (reviewed in De Silva et al., 2015).

Interestingly, a study in the teleost Ballan wrasse revealed that SHM occurs in the T-cell receptor alpha chain genes (TCRα); yet only 7% of the mutations were within AID hotspot motif (R<u>G</u>YW/WR<u>C</u>Y) (Bilal et al., 2018).

1.11. Melano-macrophage clusters (MMΦCs) in teleost fish

Melano-macrophages (MMΦs) are pigmented phagocytes; they are autofluorescent due to the accumulation of different pigments, including lipofuscin, melanin, and hemosiderin. Lipofuscin is a metabolic by-product of both phagocytosed apoptotic cells and autophagy; melanin can be derived from exogenous sources or made within the cells (Agius, 1984, Zuasti et al., 1989). Hemosiderin is a form of intracellular iron storage that forms during the breakdown of hemoglobin and serves as an intermediate step in iron recycling; this suggests that MMΦs function in iron recycling following uptake of effete RBCs (Agius et al., 2003, Kranz, 1989). MMΦs are phagocytic cells; they can phagocytose infectious materials, debris, and apoptotic cells similar to tangible body macrophages in mammalian GCs (Tsujii et al., 1990, Ferguson et al., 1976, Brattgjerd et al., 1996, reviewed in Davies et al., 2013). Similar to mammalian FDCs, MMΦs express colony-stimulating factor 1 receptor (CSF1-R), a macrophage-specific receptor (Diaz-Satizabal et al., 2015).

MMΦs in both the spleen and kidney frequently are extensively encapsulated by reticular cells (Diaz-Satizabal et al., 2015, reviewed in Wolke, 1992). In several fish species, MMΦ aggregates, known as melano-macrophage clusters (MMΦCs), are found mainly in the kidney, spleen, and sometimes the liver (Agius, 1981, Agius et al., 2003, reviewed in Steinel and Bolnick, 2017); they are observed in most vertebrates but not found in mammals, birds, and sharks (Agius, 1980, reviewed in Steinel and Bolnick, 2017). While bony fish MMΦCs form distinct structures in the spleen and kidney, in hagfish and rays MMΦs are diffusely distributed predominantly in the liver (reviewed in Steinel and Bolnick, 2017). In addition, at least in cyprinids (goldfish and zebrafish),

MMΦCs can be teased away intact from surrounding tissues (Diaz-Satizabal et al., 2015, reviewed in Muthupandian et al., 2021). Cells expressing AID were found within or near MMΦCs in the spleen and kidney of channel catfish, and almost no expression of AID was detected in the white pulp surrounding MMΦCs, and although rare, some MMΦCs in the spleen lacked AID expression (Saunders et al., 2010). In addition, AID expressing cells were found within the intestine, and they were in the vicinity of a few MMΦs; however, AID positive cells were not detected in the liver MMΦs (Saunders et al., 2010).

In bony fish, MMΦCs were seen to accumulate exogenous antigens in and around them for an extended period, and the antigen within these clusters appeared to be intact (Lamers et al., 1985, Lamers CH & De Haas MJ 1985, Ziegenfuss et al., 1991). Carp fish injected with *Aeromonas hydrophila* retained the bacterial antigen within or near MMΦCs in the spleen and kidney for at least a year (Lamers CH & De Haas MJ 1985). In addition, in 1996, Press et al. showed that after 16 weeks of intraperitoneal injection of *Aeromonas salmonicida,* lipopolysaccharide was predominantly located within MMΦCs of the spleen and kidney of Atlantic salmon (Press et al., 1996). Interestingly, studies in *Cyprinus carpio* showed that the antigen was trapped more quickly near or within MMΦCs after the injection of immune complexes, suggesting that the antigen could be trapped within these clusters using a similar mechanism to FDCs in mammals which trap antigens in the form of antibody and complement immune complexes (Secombes et al., 1980, Secombes et al., 1982, reviewed in Allen et al., 2008).

Recent work done in our lab using vaccinated goldfish revealed that MMΦs are the cells that retain the antigen within MMΦCs, and the antigen on the cell surface was in an intact form (Muthupandian MSc thesis, U of A, 2020).

Noticeably, MMΦCs in fish increase their size and numbers in response to infection or immunization, similar to mammalian GCs (Herráez et al., 1986, Agius, 1979, Secombes et al., 1982, reviewed in Steinel and Bolnick, 2017). For example, in *Clarias gariepinus*, there was a substantial increase in the frequency and size of MMΦCs in the spleen and kidney following the exposure to silver nanoparticles (Ag-NPs) (Sayed et al., 2017). In addition, the number and size of MMΦCs significantly increased in Nile tilapia in response to bacterial infections (Manrique et al., 2019). Similarly, splenic and renal MMΦCs size and frequency increased in southern bluefin tuna following treatment with *Schistosoma* (a parasitic infection), which indicates that MMΦCs participate in the immune response against pathogens (Nowak et al., 2021). These observations led scientists to use MMΦCs as a biomarker for both the immune function and health in fish (Sayed et al., 2017).

MMΦCs in teleost fish are associated with Ig expressing cells. Previous studies in Atlantic salmon showed an increase in Ig-positive cells in association with MMΦs in the spleen and kidney in response to induced infectious salmon anaemia (ISA); these cells are presumed to be B-cells (Falk et al., 1995). Also, Ig-positive cells were found adjacent to MMΦCs in turbot, and their numbers increased in response to parasitic infection (Bermudez et al., 2006). Furthermore, work on channel catfish (*Ictalurus punctatus*) showed that MMΦCs are associated with the transcripts for Ig heavy chain, TCR β, CD4, and MHC class II homologues (Saunders et al., 2010).

Taken together, these observations provide strong evidence that MMΦCs in teleost fish have the cells available to perform a similar function to mammalian GCs.

1.12. Tertiary lymphoid tissues

Tertiary lymphoid tissues (TLT) are ectopic lymphoid tissues that develop during chronic inflammation in infection, autoimmunity, graft rejection, atherosclerosis, and cancer in nonlymphoid organs. Similar to the secondary lymphoid organs, T cells, B cells, DCs, and FDCs are present in the TLT. In addition, ongoing ectopic GCs are found within these structures (reviewed in Ruddle, 2020).

1.12.1 Autoimmune diseases

In people with autoimmune diseases, affinity maturation occurs in ectopic GCs that form in the TLT at the sites of inflammation. These microenvironments contain the essential components of conventional GCs, including T- helper cells, B-cells, and they may also have FDCs; furthermore, AID is strongly expressed, and B-cells with acquired SHM and CSR are generated within ectopic GCs (Nacionales et al., 2009). For example, in patients with Sjögren's syndrome, ectopic GCs were found in the salivary gland, and like conventional GCs, signs for SHM and antigen-driven selection of high-affinity antibodies were identified (Stott et al., 1998). In addition, B-cells clonally expand, somatically mutate, and terminally differentiate into plasma cells in the synovial tissue in patients with rheumatoid arthritis; interestingly, AID expression positively correlated with the serum level

of autoantibodies, and 20% of the patients lacked FDCs (Kim et al., 1999, Xu et al.,2009, reviewed in Weyand et al., 2003). Furthermore, in patients with myasthenia gravis, which is mediated by autoantibodies directed against the postsynaptic nicotinic acetylcholine receptor (AChR), indications for clonal proliferation, somatic hypermutation, and AChR driven selection in thymic GC-like structures were found (Sims et al., 2001).

Affinity maturation occurs in areas within the secondary lymphoid organs in patients with systemic lupus erythematosus at the T-zone red pulp border rather than at the germinal centers within the B-cell follicles (William et al., 2002). These areas lacked FDC, which provides antigens for the first selection step during the GC reaction; however, DCs are widespread, and many B-cells DCs interactions occur in these sites. In addition, T-cells and proliferating B-cells with acquired mutations were found at these sites, and the mutation rate is comparable to that observed within a GC (William et al., 2002).

These studies established that affinity maturation occurs in the absence of FDCs in GC-like structures in people with autoimmune diseases; this could be the case in MMΦCs in teleost fish, which lack FDCs. Within MMΦCs, $CD4^+$ T-cells, B-cells, and a source of antigen are present (Saunders et al., 2010, Lamers CH & De Haas MJ, 1985).

1.13. High throughput sequencing and antibody repertoires

Prior to high-throughput sequencing, Sanger sequencing (first-generation DNA sequencing) allowed the analysis of the Ig variable regions at a low scale with up to a few hundred B-cells in

each experiment. In 1993, Kuppers et al. studied the development of B-cells within the GC by amplifying and sequencing B-cells isolated from stained human lymph node frozen sections (Kuppers et al., 1993). Subsequently, studies used Sanger sequencing to examine the rearrangement of the Ig genes and memory B-cells (Ehlich et al., 1994, Klein et al., 1998). However, Sanger sequencing has a low throughput and provides only a glance into the continually evolving antibody repertoire.

The development of high-throughput immune repertoire sequencing enabled quantitative insights into the diversity of lymphocyte receptors (TCR and Ig), with a wide range of applications in clinical research and biotechnology (Wu et al., 2012, Reddy et al., 2010). The first analysis of the full Ig diversity and VDJ usage was done using zebrafish, where they used validated multiplex PCR primers to amplify all the expressed transcripts for IgM and IgZ from the leader sequence to the first constant domain (Weinstein et al., 2009). Subsequently, Ig repertoires from different species, including human, mouse, rainbow trout, rabbit, chicken, and cattle, have been analyzed (Horns et al., 2019, Ippolito et al., 2012, Castro et al., 2013, Lavinder et al., 2014, Wu et al., 2012, Dong et al., 2019).

Another key focus is the ability to reconstruct and assign B-cells into clonal lineages using Ig repertoire sequencing data, where each clone is the descendant of a single B-cell responding to an antigen with mutated Ig receptors. Several software packages have been developed to assign Ig sequences into clones based on the CDR3 or the V-D-J junction of the heavy chain. CDR3 forms as the gene is recombined during development in the primary lymphoid organs where the assembly of random V, D, and J gene segments and the deletion and insertion of different nucleotides at the

V-D-J junctions take place; therefore, the CDR3 is like a fingerprint for the B-cell and its progeny (Lee et al., 2017, Gupta et al., 2017, Shlemov et al., 2017, Cortina-Ceballos et al., 2015, reviewed in Gupta et al., 2015, Schatz et al., 2011).

CHAPTER II

MATERIALS AND METHODS

2. Materials and methods

2.1. Maintenance of zebrafish

Adult zebrafish (*Danio rerio*) of the AB strain were used throughout this study. Zebrafish were maintained by the University of Alberta Department of Biological Sciences aquatic facility staff according to guidelines set by the Canadian Council of Animal Care, and protocols were preapproved by the University of Alberta's Animal Use and Care Committee. Zebrafish were kept at approximately 26-28.5 ºC.

2.2. TMS for euthanization or anesthetizing zebrafish

Zebrafish were anesthetized or euthanized using tricaine methanesulfonate (TMS) (buffered to pH 7 with NaOH). TMS concentrations of 20 mg/L and 40 mg/L were used for anesthetizing and euthanizing, respectively.

2.3. Vaccine preparation and immunization

Fish were vaccinated intraperitoneally (i.p.) with different proteins conjugated to Alexa-647, including bovine serum albumin (BSA)- Alexa fluor 647 (Invitrogen), Keyhole limpet hemocyanin (KLH)- Alexa 647 (Labelled using Alexa 647 labeling kit (Thermo Fisher Scientific)), and phycoerythrin (PE)- Alexa 647 (Thermo Fisher Scientific). In addition, KLH protein without Alexa 647 was used to vaccinate a number of fish (Table 2.1).

A total of 10 µl of vaccine containing 2 µg of protein was prepared per fish. For the first dose, one part of complete Freund's adjuvant (Sigma-Aldrich) was added to three parts of protein. The initial immunization was followed 1 month later by a second i.p. injection with either BSA or KLH conjugated to Alexa-647 emulsified in one part of incomplete Freund's adjuvant (Sigma-Aldrich). Some fish received only one dose of the vaccine (Table 2.1). BD Insulin Syringes with BD Ultra-Fine™ needle 31 G 3/10 mL 6 mm was used for the vaccination process.

2.4. Fish dissection and tissue collection

Vaccinated fish were dissected at different timepoints following immunization. Fish that received two doses of the vaccine were dissected 10 or 20 days after the second injection; on the other hand, fish that received a single dose of the vaccine were dissected one month or 20 days after the initial vaccination (Table 2.1). In addition, a number of unvaccinated zebrafish were used (Table 2.1). Using a dissecting microscope, the spleen and kidney were identified and removed. Then, I used a fluorescence microscope to identify autofluorescent melano-macrophage clusters (MMΦCs) and isolate them from the spleen and kidney. Individual clusters were photographed in brightfield and fluorescence before being processed.

For the Ig repertoires analysis, two libraries were prepared and sequenced; for the first library, I isolated 8 MMΦCs from unvaccinated zebrafish spleen and kidney. For the second run, a total of 12 MMΦCs were isolated from vaccinated fish (Table 2.1). In addition, I retained the surrounding tissues after removing MMΦCs in the kidney from vaccinated fish to compare the number and clonal relatedness of B-cells between the clusters and their surroundings (Table 2.1). A single whole kidney and intestine were used as positive controls for total V_H family expression and IgM and IgZ expression, respectively (Table 2.1; Zhang et al., 2010, reviewed in Bilal et al., 2021).

Table 2.1. Fish used to prepare Ig repertoire libraries.

Fish[1]	Vaccination[2]	Fish dissection
F1UKCa	none	–
F2UKCa	none	–
F2USCa	none	–
F3UKCa	none	–
F3USCa	none	–
F4UKCa	none	–
F4USCa	none	–
F4USCb	none	–
F5VSCa	1º PE-Alexa 2º BSA-Alexa	10 days after the booster immunization
F5VKCa	1º PE-Alexa 2º BSA-Alexa	10 days after the booster immunization
F6VKCa	1º PE-Alexa 2º BSA-Alexa	10 days after the booster immunization
F6VKCb	1º PE-Alexa 2º BSA-Alexa	10 days after the booster immunization
F7VKCa	1º PE-Alexa 2º BSA-Alexa	20 days after the booster immunization
F8VKCa	1º PE-Alexa 2º BSA-Alexa	20 days after the booster immunization
F8VKCb	1º PE-Alexa 2º BSA-Alexa	20 days after the booster immunization
F9VSCa	1º BSA-Alexa 2º KLH-Alexa	20 days after the booster immunization
F9VKCa	1º BSA-Alexa 2º KLH-Alexa	20 days after the booster immunization
F10VKCa	1º BSA-Alexa	30 days after immunization
F11VKCa	1º KLH	20 days after immunization
F12VKCa	1º KLH	20 days after immunization
F1UKW	none	–
F4UIW	none	–
F13VKS	1º BSA-Alexa	20 days after immunization
F14VKS	1º BSA-Alexa	20 days after immunization

[1]F# - Fish # in group; U/V - Unvaccinated/Vaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b; W -whole tissue; I - intestine; S - tissues surrounding the clusters.
[2]1° Primary, 2° Secondary; PE - Phycoerythrin; BSA - Bovine Serum Albumin; KLH - Keyhole Limpit Hemocyanin; Alexa - Alexa 647.

2.5. Ig repertoires Library preparation and sequencing

2.5.1. RNA isolation and cDNA preparation

Total RNA was extracted from each cluster (or tissue) using RNeasy Micro Kit (QIAGEN) following the manufacture's protocol. The concentrations of the total RNA were determined using NanoDrop (ND-1000 spectrophotometer, NanoDrop) and Qubit (RNA HS Assay Kit, Thermo Fisher Scientific).

Isolated total RNA from each sample was split into two cDNA synthesis reactions and reverse transcribed using two gene-specific primers. Previously described (Weinstein et al., 2009) primers using both IgM and IgZ constant regions were used for cDNA synthesis. Reverse transcription was carried out using SuperScript III First-Strand Synthesis kit (Invitrogen). In a 20 µl reaction cDNA synthesis reaction was carried out at 25 °C for 5 minutes, followed by 55 °C for 60 minutes, then the reaction was inactivated by heating at 70 °C for 15 minutes. Subsequently, RNase H (Invitrogen™) was added to the reactions and incubated at 37 °C for 20 minutes to remove complementary RNA.

2.5.2. Polymerase chain reaction (PCR) and size selection

I used previously described and validated primers to capture the variable domain of the zebrafish immunoglobulin heavy chain (Weinstein et al., 2009). The forward primers were designed using the consensus leader sequences of the 39 functional V gene segments of the heavy chain, and the

reverse primers using the IgM and IgZ constant regions (Weinstein et al., 2009). Initially, these primers were used to verify the expression of IgM and IgZ within zebrafish MMΦCs. Using Phusion™ High-Fidelity DNA Polymerase (Thermo Fisher Scientific), the initial denaturation was carried out at 98 °C for 30 seconds, followed by 20 cycles of denaturation at 98 °C for 10 seconds, primers annealing to DNA at 60 °C for 30 seconds, and extension at 72 °C for 30 seconds. This was followed by 10 cycles of denaturation at 98 °C for 10 seconds, primers annealing to DNA at 57.1 °C for 30 seconds, and extension at 72 °C for 30 seconds. The final extension was at 72 °C for 7 minutes.

After verifying the expression of IgM and IgZ within MMΦCs, multiplex PCR primers were modified by the addition of Illumina forward and reverse adapters to the 5' of each primer (Table 2.2). However, using the modified primers, I was not able to use the 27 forward primers in a single PCR reaction due to primer-dimers formation. Therefore, the 27 forward primers were split into 6 multiplex PCRs, and a total of 12 multiplex PCR reactions were carried out for each sample (for IgM and IgZ).

Mutations are confined to the variable region of the Ig genes in mammals (Storb, 1996); therefore, additional primers were designed to amplify the constant region of IgM as a control for sequencing error (Table 2.2). Using Phusion™ High-Fidelity DNA Polymerase (Thermo Fisher Scientific), the PCR reaction began with an initial denaturation at 98 °C for 30 seconds, followed by 30 cycles of denaturation at 98 °C for 10 seconds, annealing of primer to DNA at 58.2 °C for 30 seconds, extension at 72 ° C for 30 seconds, and a final extension at 72 ° C for 7 minutes.

Multiplex PCR reactions were followed by size selection and cleanup using a 1:1 sample to bead ratio using Ampure XP beads (Beckman Coulter) following the manufacture's protocol. The concentrations of the purified multiplex PCR products were determined using NanoDrop (ND-1000 spectrophotometer, NanoDrop). Subsequently, a second PCR was done using the purified samples and different combinations of index primers for different samples to allow multiple samples to be sequenced together. Q5® Hot Start High-Fidelity DNA Polymerase (New England Bio Labs nc) was used, and the PCR reaction began with an initial denaturation at 98 °C for 3 minutes, followed by 8 cycles of denaturation at 98 °C for 30 seconds, annealing of primer to DNA at 55 °C for 30 seconds, extension at 72 ° C for 30 seconds, and a final step at 72 ° C for 5 minutes. Then a second cleanup step was carried out using a 1:1 sample to bead ratio using Ampure XP beads (Beckman Coulter) following the manufacture's protocol. Then, the concentrations of the purified index PCR products were determined using NanoDrop (ND-1000 spectrophotometer, NanoDrop).

Table 2.2. List of primer sequences.

Primer	Sequence (5'-3')
ZF IGHV 4-1 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGGTCTCCTCTGCCTTTTGT
ZF IGHV 4-2 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAACCATGATCGCCTCATCTC
ZF IGHV 4-3 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGATGGCAACAACATCCTGTG
ZF IGHV 4-4 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGCATTTCAGTTCTGCTGCT
ZF IGHV 4-5 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACGAATGCAGGAGTCAGACA
ZF IGHV 4-6 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGTTTCAACTGTTCGTGGTCA
ZF IGHV 4-7 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGGAGTTGTGTTGATGATGATT
ZF IGHV 4-8 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTCATATGCACATGGTCAGTCA
ZF IGHV 4-9 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGTGGTGATTGTCTTTCAAGG
ZF IGHV 4-10 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGGAAAAGGAGTCAAAAAGCAT
ZF IGHV 4-11 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCTTTTGTCATGTTTGCTCTCA
ZF IGHV 4-12 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCTTACTGCTGCTCTCATTCAG
ZF IGHV 4-13 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTTCTGCTGCTGTGCTTTAC
ZF IGHV 4-14 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTGCTGTTTTCATTGGCCTTA
ZF IGHV 4-15 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGTTTATACTGTCAAGGCATGG
ZF IGHV 4-16 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAGCCTCAAGATGAAGAATGC
ZF IGHV 4-17 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTAGTGCTGTTTCTGGCAGT
ZF IGHV 4-18 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCATGATCACCTCATCTCTCTGC
ZF IGHV 4-19 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCATGATTCTGAGCATTTTATCATG
ZF IGHV 4-20 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAATAATCAACTCACTCCTGCTG
ZF IGHV 4-21 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTGCGTCCAGTGTATATTCCA
ZF IGHV 4-22 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGTATTGACTGTCAGGTTGTGC
ZF IGHV 4-23 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTCTTTCTGCAGTTGGCAG
ZF IGHV 4-24 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTCTCAAAGTTGTTGGTGTCAGA
ZF IGHV 4-25 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTCTCTAAACAAGTGCAAAGGTC
ZF IGHV 4-26 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGGACCTTAAACTTAACTGTCTG
ZF IGHV 4-27 FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCATATGTTTCTGGCATCTCCC
ZF C-μ IgM REV	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGCACTGAGACAAACCGAAG
ZF C-ζ IgZ REV	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCAGAGGCCAGACATCCAAT
ZF Ig Heavy-mu FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTGCTGGTAAAGATTTGAG
ZF Ig Heavy-mu REV	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTGATCCACCTTCTAATTCC

2.5.3. Libraries quantification, quality check, and sequencing

To quantify the intact dsDNA in each sample, I used Qubit dsDNA HS (High Sensitivity) kit in which a fluorescent dye binds to double-stranded DNA (dsDNA) (Thermo Fisher Scientific, Q32854). Then based on the different concentrations obtained using the Qubit, samples were diluted to pool the exact concentration of individual samples for each library. Afterward, in order to confirm the size range and check the quality and purity of the prepared libraries, I used the bioanalyzer high sensitivity DNA kit (Agilent Technologies, 5067–4626) following the manufacture's protocol. Then I used two MiSeq Reagent Kit v2 (500-cycles) MS-102-2003 2x250 (Illumina) to sequence the two Ig repertoire libraries (prepared from unvaccinated and vaccinated zebrafish).

PhiX is a single-stranded DNA (ssDNA) virus, and because it has a small, diverse, and well-studied genome (Sanger et al., 1977), it has been used as an effective internal control in sequencing runs. The use of PhiX spike-in is especially useful for low-diversity libraries where a significant number of the reads have similar sequences. Low diversity libraries will have unbalanced base composition and thus result in unbalanced fluorescent signals. Therefore, I used 30 % PhiX spike-in to increase the diversity of our Ig repertoire libraries, serve as an internal control, and help provide balanced fluorescent signals through the run. A total of 8 million reads were obtained from each library.

2.6. Data processing and analysis

Data processing was primarily done using tools from the Immcantation framework, which uses Python and R packages to analyze high-throughput Ig-sequencing datasets (Gupta et al., 2015, Vander Heiden et al., 2014, Stern et al., 2014, Hoehn et al., 2019). I worked with the Science & Technology IT support team at the University of Alberta to modify the Immcantation analysis framework package by adding zebrafish germline sequences to the framework. In addition, tools including, Cutadapt, Recombination Detection Program (RDP4), B-cell repertoire inductive lineage and immunosequence annotator (BRILIA), and iNext (iNterpolation/EXTrapolation), which are not part of the Immcantation framework, were used to perform different Ig repertoire data analysis (discussed below) (Martin, 2011, Martin et al., 2015, Lee et al., 2017, Hsieh et al., 2016).

2.6.1. Sequencing data quality filtering and read assembly

Remnant Illumina adapters were removed from the sequencing reads using Cutadapt (Martin, 2011), using forward and reverse adapter sequences. Quality filtering and read assembly were performed using several subcommands from pRESTO software package (Vander Heiden et al., 2014). I used align subcommand of MaskPrimers to remove the V- and C-region primers using cut mode, allowing for a maximum error rate of 0.2. Raw reads that failed primer identification by Maskprimers were removed. Then PairSeq was used to organize and match sequence records with matching coordinates in the files of the forward and reverse reads of each sample. Subsequently, I used align subcommand of AssemblePairs to merge the overlapping paired-end reads into a single sequence by aligning the ends (Vander Heiden et al., 2014).

FASTQ files encode Phred quality scores which are used to indicate the probability of a base being called correctly during sequencing (Ewing et al., 1998). A higher score indicates a higher probability that a particular base is correct; on the other hand, lower scores represent a higher probability of error. Generally, a Phred quality score of 20 or above is acceptable (Sathyanarayanan et al., 2018). To identify and remove low-quality reads, I used quality subcommand of FilterSeq (Vander Heiden et al., 2014), and reads with mean Phred quality scores less than 30 were removed. In addition, reads were filtered by length using length subcommand of FilterSeq (Vander Heiden et al., 2014). Then, to remove duplicate sequences, I used collapseSeq to collapse identical sequences without allowing any missing nucleotides while collapsing sequences (Vander Heiden et al., 2014).

Previous studies have revealed that many of the rare sequences, sequences with one copy (singletons), are artifacts and result from sequencing platform or PCR errors. They determined that around 40% to 80% of the sequences with a single copy are generated by sequencing or PCR errors (Tedersoo et al., 2010, Brown et al., 2015). Therefore, after removing duplicate sequences, I used splitseq command to keep only reads with at least two copies; singletons or reads that were observed only once were discarded (Vander Heiden et al., 2014).

2.6.2. Alignment and clonal clustering

Following quality filtering and reads assembly, VDJ annotation was performed using IMGT/High V-QUEST using zebrafish IgH locus germline sequences (Alamyar et al., 2010). In addition, IMGT/High V-QUEST provides information about the functionality of the sequence, identifies the

complementarity determining regions (CDRs) and framework regions (FWRs) and their length, analyzes the VDJ junction by providing information about the junction length, the number of nucleotides added at the V, D, and J gene ends and translated sequence of the junction. Analysis of nucleotide insertions and deletions (indels) in the V-region was also done using IMGT/High V-QUEST (Alamyar et al., 2010).

Then I used Change-O package (Gupta et al., 2015), which includes multiple tools to process the annotated Ig sequences using the output of IMGT/High V-QUEST (Alamyar et al., 2010). Initially, MakeDb was used to store sequence alignment information in a tab-delimited database file. Then I used distToNearest to determine the clustering threshold for each Ig repertoire to group sequences into clusters which will be used to assign Ig sequences into clones (see below) (Gupta et al., 2015).

To determine the clustering threshold, I used nucleotide Hamming distance (ham model) and distToNearest (Gupta et al., 2015). Hamming distance is calculated based on the number of nucleotides at which two sequences of the same length differ (Gupta et al., 2015). distToNearest command provides a bimodal distribution of the sequences by calculating the distance between every sequence and its closest neighbor (Gupta et al., 2015). The resulting histogram provides information about the clonal relatedness of the sequences within each repertoire; clonally related sequences are separated from unrelated sequences (Gupta et al., 2015). By manually inspecting the bimodal distribution histogram, I determined the clustering threshold by using the value that separates the clonally related sequences in the repertoire from the clonally unrelated.

Using the clustering threshold value determined for each Ig repertoire in the previous step, I used DefineClones to assign Ig sequences into clones, using ham model and single-linkage hierarchical clustering (Gupta et al., 2015). Initially, sequences are grouped based on the V gene, J gene, and junction length, then Hamming distance and single-linkage hierarchical clustering are used to assign Ig sequences into clones, where each clone is the offspring of a B-cell responding to an antigen with mutated Ig sequences (Gupta et al., 2015). Hamming distance was normalized by the length in which Hamming distance is divided by the length of the junction. Hierarchical clustering merges the two closest sequences into a cluster, then using single-linkage hierarchical clustering, clusters with sequences most similar to each other (a few or even single sequences) are merged even if most of the sequences in each cluster are distant from each other. It has been shown that combining length normalized Hamming distance with single linkage clustering provides the highest sensitivity and specificity in identifying the clones within Ig repertoire datasets (Gupta et al., 2017).

To reconstruct the germline sequence for each clone group, I used CreateGermlines command which reconstructs the germline sequence using the initial alignment data and IMGT-gapped zebrafish germline sequences (Gupta et al., 2015, Alamyar et al., 2010). Using CreateGermlines, Ig sequences within a clone will have the same germline sequence with identical germline V and J genes; however, because the alignment of the D-element is frequently of low confidence, the D-element was not used to generate the germline sequence; instead, I used germline with D gene masked.

2.6.3. Sample coverage

Using the total number of unique Ig sequences, the number of observed clones, and their relative abundance (frequency) within each repertoire, I estimated sample coverage, as measured by sample completeness or the proportion of the total number of Ig sequences in a repertoire that belong to the sequences represented in the sample. I used assembled sequencing reads then, the subcommand countClones in the alakazam package was used to determine the total number of sequences and clones in each clonal group and their copy number (Stern et al., 2014). Subsequently the R package iNext (iNterpolation/EXTrapolation) was used to estimate sample coverage (a measure of sample completeness) for IgM and IgZ repertoires (Hsieh et al., 2016). Using species richness (q=0) and sample abundance with a 95% confidence interval, iNext calculates sample coverage using the information provided by countClones (the number of observed sequences, clones, and their frequencies).

2.6.4. Building lineage trees and analysis of selection pressure, motif mutability, and the ratio of transition to transversion

Using the output of CreateGermlines, I used BuildTrees to remove non-functional sequences, collapse identical sequences, and convert datasets into IgPhyML-readable format (Gupta et al., 2015). Non-functional (unproductive) sequences are identified during the initial alignment using IMGT/High V-QUEST, sequences with out of frame junction or stop codon or frameshift mutation in the coding region, or changes of conserved amino acids are identified as non-functional sequences (Alamyar et al., 2010).

Then using repertoires with at least a hundred unique Ig sequencing reads, I used IgPhyML (HLP19 model) to build phylogenetic trees and to assess the evolutionary hypotheses of B-cell affinity maturation (Hoehn et al., 2019). IgPhyML is a repertoire-wide phylogenetic framework that uses maximum likelihood (ML) model to estimate tree topologies, branch lengths, and other parameters (discussed below). However, IgPhyML estimates are inaccurate for small lineages (Hoehn et al., 2019), and therefore, I only used Ig repertoires with at least a hundred unique Ig sequencing reads (the majority of the lineages in the Ig repertoires with less than a hundred unique sequences are small lineages).

I used IgPhyML to quantify the selection pressure in each Ig repertoire by determining the ratio of replacement (nonsynonymous) to silent (synonymous) mutations (dN/dS or R/S) within the CDRs and FWRs separately. In addition, using IgPhyML, I determined the fold-change in mutability for hotspot motifs which is calculated based on the frequency of mutations within the underlined base of WRC, GYW, WA, and TW (where W = A/T, R = A/G, Y = C/T) motifs compared to regular motifs minus one; therefore, a value greater than zero indicates a hotspot and less than a zero indicates a coldspot. Also, the ratio of transition (within purines and pyrimidines) to transversion (between purines/pyrimidines) mutations in the Ig repertoires was calculated using IgPhyML (Hoehn et al., 2019). Finally, to avoid biases when estimating IgPhyML model parameters, the CDR3 region was removed, as not all of it is germline coded.

To visualize IgPhyML phylogenetic trees, I used the Alakazam subcommand readIgphyml, which reads the output of IgPhyML and provides the input for igraph package to plot the lineage trees (Stern et al., 2014). To visualize the selection pressure values for CDRs and FWRs, I used

readIgphyml subcommand (Stern et al., 2014). Then to combine the selection parameters of different Ig repertoires into a single data frame, combineIgphyml, another subcommand in the Alakazam package (Stern et al., 2014), was utilized, and to create the heat maps, I used ggplot2 package (Wickham, 2016).

In addition, using the output of splitseq and sequences with at least two copies, I used B-cell repertoire inductive lineage and immunosequence annotator (BRILIA 3.5.7) software which simultaneously annotates genes, clusters sequences, identifies SHM, and assembles lineage trees based on CDR3 sequence, length, and VJ gene family (Lee et al., 2017). Using BRILIA VDJ annotation is done by maximizing the total alignment score, which is calculated based on the number of matched and mismatched nucleotides between the Ig repertoire sequences and the germline sequence (Lee et al., 2017). Furthermore, BRILIA uses the percent of Hamming distance (described in section 2.5.2) to construct lineage trees (Lee et al., 2017).

2.6.5. VDJ genes usage analysis

To determine the count and relative frequency of the IGHV, IGHD, and IGHJ genes within each Ig repertoire, I used the output from CreateGermlines and countGenes subcommand in the Alakazam package (Stern et al., 2014, Gupta et al., 2015). Then to visualize the VDJ gene usage analysis results, ggplot2 package was used (Wickham, 2016).

2.6.6. Clonal diversity indices (Hill numbers)

To estimate clonal diversity, I used the generalized diversity index (Hill numbers) and the output of CreateGermlines (Hill, 1973, Chao et al., 2014, Gupta et al., 2015). Using alphaDiversity subcommands in the Alakazam package (Stern et al., 2014), Ig repertoires diversity was calculated using a 95% confidence interval, at diversity orders (q values) from 0 to 8. The total number of clones in a sample (species richness) is calculated at q = 0; the exponential of Shannon's entropy index (q = 1) considers clones in proportion to their frequency. The inverse of Simpson's concentration index (q = 2) considers clones in proportion to their frequency and ignores rare clones; thus, it represents the dominant clones, and q = ∞ is the reciprocal of the proportional richness (abundance) of the largest clones (Chao et al., 2014). As the diversity order (q value) increases, larger clones weigh more; therefore, small clones slightly impact the diversity value at high diversity orders but have a significant effect at diversity order q = 0 (Hill, 1973, Chao et al., 2014).

To compare multiple Ig repertoires, I used repeated resampling in which samples are standardized by sample completeness (Chao et al., 2014). Therefore, the difference in total sequence count in each Ig repertoire is corrected to a common number of sequences across multiple Ig repertoires.

2.6.7. Clonal abundance

To calculate the complete clonal relative abundance, I used estimateAbundance subcommand in Alakazam and the output of CreateGermlines (Stern et al., 2014, Gupta et al., 2015). Similar to calculating clonal diversity I used a 95% confidence interval and repeated resampling to resample each Ig repertoire to the same completeness and correct for the variability in total sequence counts

between Ig repertoires (Gupta et al., 2015). To visualize the estimated complete clonal relative abundance, I used plotAbundanceCurve subcommands in Alakazam (Stern et al., 2014). plotAbundanceCurve plots the distribution of clonal abundance as a log-rank abundance distribution in which percent abundance is plotted against rank order (Stern et al., 2014). The first rank represents the most abundant clone, rank two corresponds to the clone with the second-highest abundance, and so on. In addition, the rank shows how many clones were ranked (reviewed in Matthews et al., 2015).

2.6.8. CDR3 length analysis

To calculate the average CDR3 length of the Ig repertoires, I used aminoAcidProperties subcommand in Alakazam package and the output from CreateGermlines (Stern et al., 2014, Gupta et al., 2015). Then I used ggplot2 package to visualize the distribution of the CDR3 length in each repertoire (Wickham, 2016). The CDR3 (junction) sequence is determined using IMGT/High V-QUEST during the initial alignment based on the anchor positions (Alamyar et al., 2010). Junction nucleotide sequence starts in the 3' end of the V element at the second conserved cysteine (2nd-CYS) at position 104 according to the IMGT unique numbering and ends with the conserved tryptophan (TRP) at position 118 5' of the J region (Monod et al., 2004, Alamyar et al., 2010). However, the conserved residues at the 5' and 3' are removed prior to CDR3 length analysis (Stern et al., 2014).

2.6.9. Gene conversion analysis

To determine if gene conversion event contributes to the diversity of the Ig repertoire in zebrafish, I used the output from CreateGermlines and the Recombination Detection Program (RDP4) (Martin et al., 2015, Gupta et al., 2015). Using Ig sequences from individual clones, RDP4 identifies and characterizes individual recombination events using GENECONV method, which uses polymorphic nucleotides in a pair of sequences to detect gene conversion events and determine parental sequences (Padidam et al., 1999, Martin et al., 2015). It should be noted that RDP4 is not specific for identifying gene conversion events in Ig gene sequences; however, it was used along with manual inspection for gene conversion events using zebrafish germline IGHV sequences, IGHV pseudogenes, and sequenced Ig repertoires.

2.7. Ag retention within MMΦCs in zebrafish

To determine if long-term Ag retention occurs within MMΦCs in zebrafish, I vaccinated fish with KLH or BSA conjugated to Alexa-647 (Invitrogen, Thermo Fisher Scientific), emulsified in complete Freund's adjuvant (Sigma) (described in section 2.3). Alexa-647 absorbance and emission spectra are at 650 and 665 nm, respectively, which makes it distinguishable from the autofluorescence emission spectra of MMΦ pigments in the 500–580 nm wavelength range (Saunders et al., 2010).

One month following the injection, fish were dissected, and the spleen and kidney were removed and identified. Then Zeiss Axio Imager M2 fluorescent microscope was used to isolate MMΦCs from the spleen and kidney and determine if there is long-term Ag retention in the isolated clusters.

2.8. Imaging Flow cytometry (ImageStream) analysis

To confirm the presence and the total number of lymphocyte-like cells within MMΦCs, I used clusters isolated from the spleen and kidney of unvaccinated zebrafish. A total of 16 clusters from the spleen and 15 clusters from the kidney were used in two separate imaging flow cytometry runs. Individual clusters were photographed in brightfield and fluorescence before being processed.

2.8.1. Cell preparation

To dissociate the cells from isolated MMΦCs, clusters were mashed through a 40 µm cell strainer (pluriStrainer) using a 1 ml syringe plunger. Afterward, I stained the nuclei of the dissociated cells using Hoechst 33342 fluorescent stain (Thermo Fisher Scientific). The total number of cells was assessed using manual cell count before analyzing the samples through an imaging flow cytometer (ImageStream).

2.8.2. Imaging flow cytometry (ImageStream) data analysis

After running the samples through an imaging flow cytometer (Luminex), I used IDEAS 6.1.822.0 software (Luminex) to analyze the data. Using IDEAS 6.1.822.0 software, cells that were in focus were initially selected, then, using aspect ratio, single cells were gated. Aspect ratio is calculated based on the shape of the detected object; circular cells have aspect ratio values close to 1.0 (Zuba-Surma et al., 2007). Following the removal of doublets and aggregates, a dot plot using area (size,

x-axis) vs. intensity channel 6 (internal complexity, y-axis) was created to differentiate between the subpopulations using cell size and internal complexity.

Table 2.3. Number of zebrafish used in different experiments.

Type of experiment	Number of zebrafish used
Ig repertoire libraries	14
Ag retention within MMΦCs	6
Imaging Flow cytometry and cell count analyses	10

CHAPTER III

RESULTS

3. Results

3.1. B-cell clonotypes expand while acquiring mutations within MMΦCs

3.1.1. Sample coverage and the total number of unique Ig sequences

Using iNEXT and assembled sequencing reads, VDJ recombination coverage for all the clusters was at least 93.5% for both IgM and IgZ isotypes (Table 3.1). Ig repertoires isolated from MMΦCs from unvaccinated zebrafish had a coverage between 93.5% and 99% for IgM isotype and between 98.9% and 99.6% for IgZ isotype (Table 3.1). VDJ recombination coverage for Ig repertoires isolated from clusters from vaccinated zebrafish was between 97.5% and 98.8% for IgM isotype and between 95.5% and 99.4% for IgZ (Table 3.1). The coverage for the whole kidney was 85% for IgM and 98% for IgZ, and the intestine had 96% and 99% VDJ combination coverage for IgM and IgZ isotypes, respectively. Ig repertoires isolated from the tissues surrounding kidney MMΦCs had no IgZ sequences, and VDJ combination coverage for IgM was 97% (Table 3.1).

A single whole kidney and intestine had 5539, and 3823 total unique sequencing reads, respectively. Ig repertoires isolated from both the whole kidney and intestine were mainly of IgM isotype; a total of 4389 and 3155 unique sequencing reads were of IgM isotype in the kidney and

intestine, respectively (Table 3.1). Ig repertoires isolated from unvaccinated fish MMΦCs contained between 1064 and 2725 total unique sequencing reads, and most of these clusters had IgM as the dominant isotype. Two repertoires of the Ig repertoires isolated from unvaccinated fish MMΦCs had comparable unique sequencing reads for IgM and IgZ isotypes, and a single cluster had mostly Ig sequences of IgZ isotype (Table 3.1). Ig repertoires isolated from vaccinated fish MMΦCs had between 510 and 3660 total unique Ig sequences, and all the Ig repertoires isolated from these clusters had IgM as the dominant isotype. Ig sequences of IgZ isotype were absent in a single cluster from vaccinated fish MMΦCs, and four of the repertoires from these clusters had less than a hundred total unique sequencing reads of IgZ isotype (Table 3.1). The two Ig repertoires isolated from the tissues surrounding kidney MMΦCs had 82 and 129 total unique sequences, and all the reads were of IgM isotype (Table 3.1).

3.1.2. B-cell clones

To determine the clonal relatedness of B-cells within a single repertoire, Ig sequences were assigned into clones, where each clone is the offspring of a B-cell responding to an antigen with mutated Ig sequences.

Analysis of the total number of clones within each repertoire, revealed that the number of clones varied among the different repertoires. However, Ig repertoires from both unvaccinated and vaccinated fish MMΦCs had more IgM isotype clones than the IgZ isotype (Table 3.1). Ig repertoires isolated from unvaccinated fish MMΦCs had between 154 and 437 clones of IgM isotype and between 7 and 24 clones of IgZ isotype. Most of the repertoires isolated from

vaccinated fish clusters had a higher number of clones for both IgM and IgZ isotypes than repertoires from unvaccinated fish clusters, despite having fewer total unique Ig sequences of IgZ isotype (Table 3.1). Ig repertoires isolated from vaccinated fish clusters had between 99 and 1626 clones of IgM isotype and between 2 to 84 clones for IgZ isotype (Table 3.1). The whole kidney and intestine Ig repertoires had 3511 and 1394 clones of IgM isotype and 413 and 91 clones of IgZ isotype, respectively. The tissues surrounding kidney MMΦCs had 32 and 35 clones of IgM isotype (Table 3.1).

The size of the clones varied among the different clusters; however, generally, it appears that each cluster is dominated by a few clonally expanding B-cell clonotypes, with some of the dominant clones having more than 400 mutated daughter cells; this was determined using IgPhyML (Hoehn et al., 2019) (Table 3.2, Table 3.3). Generally, clones in the repertoires from unvaccinated fish MMΦCs are larger than the clones in the repertoires isolated from vaccinated fish MMΦCs. In addition, the largest dominant clones in the repertoires from unvaccinated fish clusters were mostly of IgZ isotype, in contrast to the repertoires from vaccinated fish MMΦCs, where the largest clones were mainly of IgM isotype (Table 3.2, Table 3.3). Nonetheless, clones with fewer daughter cells compared to the dominant clones were also found within MMΦCs from unvaccinated and vaccinated fish (Table 3.2, Table 3.3).

On average, in the repertoire isolated from unvaccinated fish MMΦCs, the size of the top five largest clones ranged between 272 and 82 (Table 3.2). The size of the dominant clones varied in repertoires isolated from MMΦCs from different tissues from the same unvaccinated fish. For example, F2UKCa and F2USCa are repertoires isolated from the same unvaccinated fish kidney

and spleen, respectively (Table 3.2). The size of the largest dominant clone in the repertoire isolated from F2UKCa is 416; in contrast, the largest clone in the repertoire isolated from the spleen cluster (F2USCa) from the same fish is 50 (Table 3.2). On the contrary, in the repertoires isolated from unvaccinated fish number 3 MMΦCs (F3UKCa and F3USCa), the spleen cluster had larger clones compared to the repertoire isolated from the kidney from the same fish (F3UKCa and F3USCa, (Table 3.2)). The difference in the size of the dominant clones in the repertoires isolated from MMΦCs from the same tissue of the same unvaccinated fish (F4USCa and F4USCb (Table 3.2)), was less evident compared to the size of the dominant clones of the repertoire isolated from MMΦC from a different tissue of the same unvaccinated fish (F4USCa, F4USCb, and F4UKCa (Table 3.2)).

In the repertoires isolated from vaccinated fish MMΦCs, on average, the size of the top five largest clones ranged between 92 and 41 (Table 3.3). Similar to the repertoires isolated from MMΦCs from unvaccinated fish, the size of the dominant clones varied between repertoires isolated from MMΦCs from different tissues of vaccinated fish (F9VSCa and F9VKCa (Table 3.3)), and the size of the dominant clones was similar in the repertoires isolated from MMΦCs from the same tissue (F8VKCa and F8VKCb (Table 3.3)). However, unlike the repertoires from MMΦCs from unvaccinated fish, the size of the dominant clones varied in the two repertoires isolated from clusters from the kidney of vaccinated fish number 6 (F6VKCa and F6VKCb (Table 3.3)). In addition, clones with similar size were found in two repertoires isolated from MMΦCs from different tissues from the same vaccinated fish (F5VSCa and F5VKCa isolated from the spleen and kidney, respectively (Table 3.3)).

The top clones in the Ig repertoire isolated from the whole kidney had a smaller size compared to the dominant clones in the repertoires isolated from MMΦCs from vaccinated and unvaccinated fish (Table 3.2, Table 3.3, Table 3.4). The size of the largest clone in the whole kidney repertoire is 46; conversely, in the repertoire isolated from MMΦC from the same kidney, the size of the largest clone is 404 (F1UKW (Table 3.4), and F1UKCa (Table 3.2)). Unlike the Ig repertoire isolated from the whole kidney, the size of the clones in the repertoire isolated from the whole intestine was similar to the size of the dominant clones in the repertoires isolated from MMΦCs from unvaccinated and vaccinated fish (Table 3.2, Table 3.3, Table 3.4). The largest clone in the Ig repertoire isolated from the intestine is of IgZ isotype and had a size of 459 (F4UIW (Table 3.4)). The size of the clones in the repertoires isolated from the tissues surrounding kidney MMΦCs was significantly smaller than the size of the dominant clones in the repertoires isolated from MMΦCs from unvaccinated and vaccinated fish (Table 3.2, Table 3.3, Table 3.4). The size of the top clones, except for one clone, ranged between 2 and 15 (F13VKS and F14VKS (Table 3.4)). The size of a single clone in the repertoires isolated from the tissues surrounding kidney MMΦCs is 50 (F14VKS (Table 3.4)).

Several clones with a size equal to one were found in the Ig repertoires isolated from MMΦCs from vaccinated and unvaccinated fish. These clones had no mutations in their Ig sequences, and they were identical to their germline sequences (Fig.3.1).

Similar results were obtained using BRILIA (3.5.7) (Lee et al., 2017); Ig repertoires isolated from MMΦCs from vaccinated and unvaccinated fish had a few dominant clones, and some of the dominant clones had up to 300 daughter cells with unique mutations. The repertoire isolated from

the whole intestine had a few dominant clones, and their sizes were similar to the size of the clones in the repertoires prepared from MMΦCs. The clones in the repertoires isolated from the whole kidney and tissues surrounding kidney MMΦCs were significantly smaller than those in the Ig repertoires isolated from MMΦCs from vaccinated and unvaccinated fish. However, overall, the number of daughter cells or the size of the clones was smaller using BRILIA (3.5.7) compared to the numbers obtained using IgPhyML (Lee et al., 2017, Hoehn et al., 2019).

3.1.3 B-cell lineage trees

To provide a visualization of the evolutionary relationship between Ig sequences within a clone, I constructed phylogenetic trees. Using igraph package to visualize IgPhyML phylogenetic trees for each clone (Stern et al., 2014, Hoehn et al., 2019), I found that B-cell clonotypes proliferate and accumulate mutations in their variable region within MMΦCs.

Mean tree length or the average expected substitutions per codon site among all lineages within a single repertoire varied across the different repertoires of IgM and IgZ isotypes isolated from vaccinated and unvaccinated fish MMΦCs (Table 3.5, Table 3.6). Typically, repertoires of IgZ isotype had a higher number of expected mutations per codon than repertoires of IgM isotype (Table 3.5, Table 3.6). Ig repertoires isolated from unvaccinated fish MMΦCs had a mean tree length between 0.06 and 0.16 for IgM isotype and between 0.25 and 0.99 for IgZ isotype (Table 3.5). All the repertoires from vaccinated fish MMΦCs had similar mean tree length values for IgM isotype (between 0.1 and 0.15); however, repertoires of IgZ isotype from these clusters had values of mean tree length between 0.07 and 0.27 (Table 3.6).

Using individual clones, the tree topology revealed the ancestral evolutionary relationships among Ig sequences within a clone. Examples of lineage trees using IgPhyML and igraph (Stern et al., 2014, Hoehn et al., 2019) show the germline sequences at the root of the trees, and the numbers of expected mutations between ancestor and child nodes are shown as the branch length between them (Fig.3.2, Fig.3.3). In these lineage trees, the internal nodes are inferred using maximum likelihood (Hoehn et al., 2019), and the CDR3 sequence is used to represent each Ig sequence; however, any mutation within the variable region is used to calculate the branch length. Despite having the same number of Ig sequences in these lineage trees, sequences in the 1-Germline lineage tree are more closely related to their germline sequence compared to Ig sequences in the 2-Germline lineage tree (Fig.3.2, Fig.3.3).

Similarly, lineage trees constructed using BRILIA (3.5.7) revealed that B-cell clonotypes isolated from MMΦCs from vaccinated and unvaccinated fish proliferate while acquiring mutations in their Ig sequences. An example of a lineage tree generated using BRILIA (3.5.7) is shown in figure 3.4, where the percent of Hamming distance (HAM %) is used to show the relationship between Ig sequences within a single clone (Fig.3.4). The sequence with the smallest HAM % is assigned as the root of the tree, and all the CDR3 sequences within a lineage tree are shown using a unique dot color that corresponds to the translated sequence of the CDR3. All the sequences within a lineage tree belong to the same V, D, and J gene families and each dot is a unique sequence (Fig.3.4). The dot size reflects the copy number or the total template count for each sequence, and as shown in figure 3.4, the Ig sequence at the branch point has a high copy number, which is consistent with the accumulation of mutations in highly proliferating B-cell clonotypes (Fig.3.4).

Table 3.1. Ig repertoires isolated from MMΦCs or tissues from unvaccinated and vaccinated fish VDJ combination coverage, number of unique reads, and the total number of clones for IgM and IgZ isotypes.

Fish[1]	Vaccination[2]	IgM isotype			IgZ isotype		
		Coverage %	Number of unique reads	Total number of Clones	Coverage %	Number of unique reads	Total number of Clones
F1UKCa	none	98.8	1052	154	99.1	878	10
F2UKCa	none	97.4	1495	366	99.6	1166	24
F2USCa	none	98.4	926	256	99.1	138	7
F3UKCa	none	93.5	876	437	99.2	336	21
F3USCa	none	99	2307	332	99.5	418	9
F4UKCa	none	98.6	989	169	99.6	1335	10
F4USCa	none	98.5	1601	360	99.5	406	11
F4USCb	none	98.6	1898	358	98.9	462	14
F5VSCa	1º PE-Alexa 2º BSA-Alexa	98.5	416	137	95.5	176	84
F5VKCa	1º PE-Alexa 2º BSA-Alexa	98.7	1569	478	98.9	416	31
F6VKCa	1º PE-Alexa 2º BSA-Alexa	98.5	1112	262	-	0	0
F6VKCb	1º PE-Alexa 2º BSA-Alexa	98.3	3295	992	98.3	253	33
F7VKCa	1º PE-Alexa 2º BSA-Alexa	98.1	421	99	96.8	89	2
F8VKCa	1º PE-Alexa 2º BSA-Alexa	98.8	2826	675	96.3	48	13
F8VKCb	1º PE-Alexa 2º BSA-Alexa	98.7	2513	672	98.9	111	22
F9VSCa	1º BSA-Alexa 2º KLH-Alexa	98.5	612	125	98.8	151	8
F9VKCa	1º BSA-Alexa 2º KLH-Alexa	98.5	1862	610	99.4	189	13
F10VKCa	1º BSA-Alexa	98	2898	970	96.1	66	4
F11VKCa	1º KLH	98.4	2109	687	98.4	89	13
F12VKCa	1º KLH	97.5	3411	1626	98.7	249	77
F1UKW	none	85.9	4389	3511	98.2	1150	413
F4UIW	none	96.8	3155	1394	99.6	668	91
F13VKS	1º BSA-Alexa	97.3	82	32	-	0	0
F14VKS	1º BSA-Alexa	97.7	129	35	-	0	0

[1]F# - Fish # in group; U/V - Unvaccinated/Vaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b; W -whole tissue; I - intestine; S - tissues surrounding the clusters.
[2]1° Primary, 2° Secondary; PE - Phycoerythrin; BSA - Bovine Serum Albumin; KLH - Keyhole Limpit Hemocyanin; Alexa - Alexa 647

Table 3.2. Size of the top clones in Ig repertoires isolated using individual MMΦCs from unvaccinated zebrafish.

	F1UKCa[1]	F2UKCa[1]	F2USCa[1]	F3UKCa[1]	F3USCa[1]	F4UKCa[1]	F4USCa[1]	F4USCb[1]
Clone	Number of unique reads	Number of unique reads	Number of unique reads	Number of unique reads	Number of unique reads	Number of unique reads	Number of unique reads	Number of unique reads
1	404	416	50	77	320	449	274	189
2	301	324	48	74	162	384	170	97
3	238	186	48	74	117	316	92	87
4	215	90	45	31	97	183	90	81
5	115	79	35	29	91	174	67	65
6	60	77	30	19	74	106	49	56
7	48	69	28	13	70	63	48	55
8	42	69	27	13	65	59	44	44
9	38	67	24	12	64	34	42	38
10	37	49	24	12	46	30	28	37

[1]F# - Fish # in group; U - Unvaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.
Shaded cells are clones of IgZ isotype.

Table 3.3. Size of the top clones in Ig repertoires isolated using individual MMΦCs from vaccinated zebrafish.

	F5VSCa[1]	F5VKCa[1]	F6VKCa[1]	F6VKCb[1]	F7VKCa[1]	F8VKCa[1]
Clone	**Number of unique reads**	**Number of unique reads**	**Number of unique reads**	**Number of unique reads**	**Number of unique reads**	**Number of unique reads**
1	78	83	93	284	168	261
2	43	80	88	110	88	169
3	31	71	79	64	10	145
4	18	70	41	71	9	102
5	16	60	41	51	9	61
6	15	60	21	37	9	54
7	13	44	20	37	8	50
8	13	42	19	36	7	50
9	11	41	18	35	7	33
10	10	31	17	33	6	30
	F8VKCb[1]	**F9VSCa[1]**	**F9VKCa[1]**	**F10VKCa[1]**	**F11VKCa[1]**	**F12VKCa[1]**
Clone	**Number of unique reads**	**Number of unique reads**	**Number of unique reads**	**Number of unique reads**	**Number of unique reads**	**Number of unique reads**
1	258	105	42	109	78	205
2	165	41	41	57	77	146
3	126	38	41	45	61	113
4	75	37	40	43	53	66
5	70	30	36	37	38	48
6	52	29	36	37	33	38
7	40	27	32	36	31	24
8	31	27	28	32	28	23
9	25	26	29	31	27	23
10	22	24	26	31	27	21

[1]F# - Fish # in group; V - Vaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.
Shaded cells are clones of IgZ isotype.

Table 3.4. Size of the top clones in Ig repertoires isolated using whole kidney, intestine, and tissues surrounding kidney MMΦCs from zebrafish.

	F1UKW[1]	F4UIW[1]	F13VKS[1]	F14VKS[1]
Clone	**Number of unique reads**	**Number of unique reads**	**Number of unique reads**	**Number of unique reads**
1	46	459	15	50
2	33	121	9	14
3	23	100	7	13
4	21	87	6	7
5	18	55	6	6
6	15	46	4	4
7	15	25	4	3
8	14	24	4	2
9	14	24	3	2
10	14	21	2	2

[1]F# - Fish # in group; U/V - Unvaccinated/Vaccinated; W - whole tissue; K - kidney: I - intestine; S - tissues surrounding the clusters.
Shaded cells are clones of IgZ isotype.

```
137_Clone          1 TTGACCTCCTCTGGTTCTGAGGTCAAGAAACCCAGAGAATCAGTCACACT     50
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
137_Germline       1 TTGACCTCCTCTGGTTCTGAGGTCAAGAAACCCAGAGAATCAGTCACACT     50

137_Clone         51 GTCTTGTGTGGTTTCTGGACTCTCCCTTGCCTGGCTGCACTGGATAAGGC    100
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
137_Germline      51 GTCTTGTGTGGTTTCTGGACTCTCCCTTGCCTGGCTGCACTGGATAAGGC    100

137_Clone        101 AGAAACCGGGAAAAGGCCTGGAGTGGATTGGCCGGATTGACAGTGGCACT    150
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
137_Germline     101 AGAAACCGGGAAAAGGCCTGGAGTGGATTGGCCGGATTGACAGTGGCACT    150

137_Clone        151 GGGACTATTTTTGCTCAGTCTCTACAAGGCCAATTTACCATCACTAAAGA    200
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
137_Germline     151 GGGACTATTTTTGCTCAGTCTCTACAAGGCCAATTTACCATCACTAAAGA    200

137_Clone        201 CACCAGCAAAAACATGGTGTATTTGGAGATAAAAAGCCTGAAGGCTGAAG    250
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
137_Germline     201 CACCAGCAAAAACATGGTGTATTTGGAGATAAAAAGCCTGAAGGCTGAAG    250

137_Clone        251 ATACTGCTGTTTATTACTACTGGGGGAAAGGAACCAAAGTGACAGTTTCC    300
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
137_Germline     251 ATACTGCTGTTTATTACTACTGGGGGAAAGGAACCAAAGTGACAGTTTCC    300

137_Clone        301 TCA    303
                     |||
137_Germline     301 TCA    303
```

Figure 3.1. Sequence alignment of Ig sequence (clone number 137) to its inferred germline. Germline sequence was determined using CreateGermlines command, and the alignment was carried out using emboss needle.

Table 3.5. Mean tree length of Ig repertoires isolated from MMΦCs from unvaccinated zebrafish. The analysis was performed using IgPhyML HLP19 model.

Fish[1]	Mean tree length (IgM)	Mean tree length (IgZ)
F1UKCa	0.15	0.99
F2UKCa	0.14	0.57
F2USCa	0.12	0.26
F3USCa	0.16	0.55
F3UKCa	0.06	0.25
F4UKCa	0.15	1.41
F4USCa	0.12	0.44
F4USCb	0.14	0.43
Average	**0.13**	**0.65**
SD	**0.03**	**0.42**

[1]F# - Fish # in group; U - Unvaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.

Table 3.6. Mean tree length of Ig repertoires isolated from MMΦCs from vaccinated zebrafish. The analysis was performed using IgPhyML HLP19 model.

Fish[1]	Mean tree length (IgM)	Mean tree length (IgZ)
F5VSCa	0.11	0.07
F5VKCa	0.14	0.24
F6VKCa	0.14	_
F6VKCb	0.13	0.14
F7VKCa	0.11	_
F8VKCa	0.15	_
F8VKCb	0.14	0.11
F9VKCa	0.1	0.25
F9VSCa	0.13	0.27
F10VKCa	0.13	_
F11VKCa	0.13	_
F12VKCa	0.11	0.16
Average	**0.13**	**0.18**
SD	**0.01**	**0.07**

[1]F# - Fish # in group; V - Vaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.

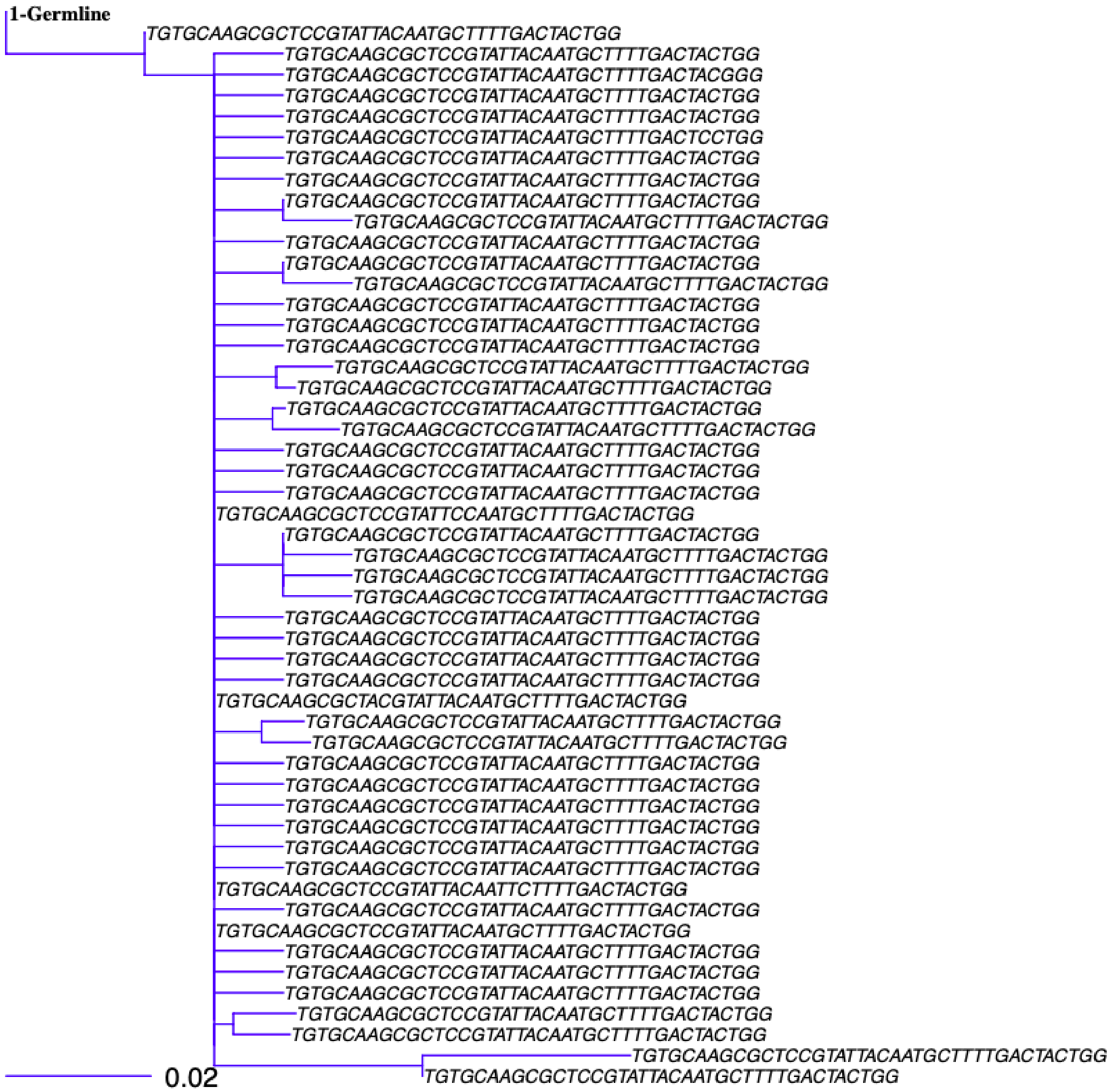


Figure 3.2. An example of a lineage tree (1-Germline) using IgPhyML (HLP19 model) and Alakazam (igraph). CDR3 nucleotide sequences are used to represent all the unique sequences in the clone (the size of this clone is 51 unique sequences). The scale bar represents the branch length (expected number of substitutions per codon site). The analysis was done using IgPhyML.

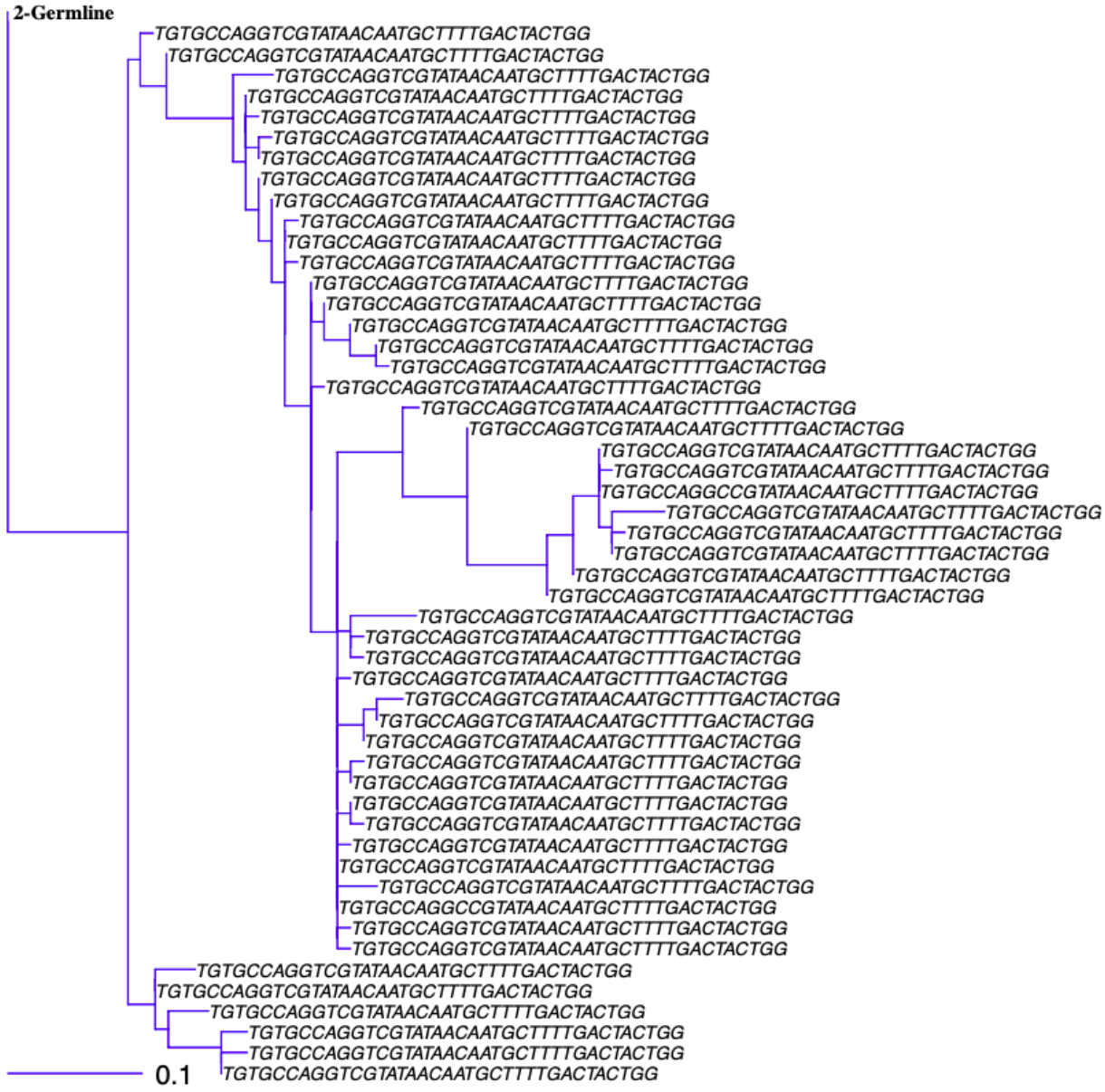


Figure 3.3. An example of a lineage tree (2-Germline) using IgPhyML (HLP19 model) and Alakazam (igraph). CDR3 nucleotide sequences are used to represent all the unique sequences in the clone (the size of this clone is 51 unique sequences). The scale bar represents the branch length (expected number of substitutions per codon site). The analysis was done using IgPhyML.

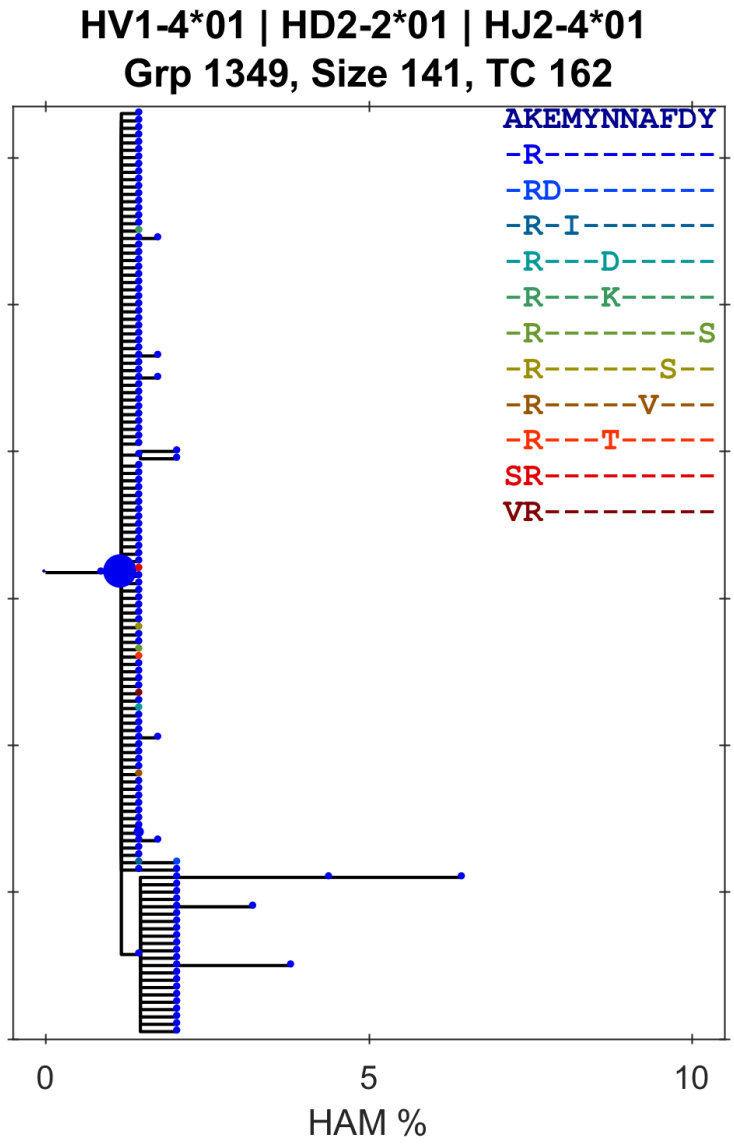


Figure 3.4. An example of a lineage tree (group 1349) using BRILIA (3.5.7). Translated sequences of the unique CDR3 sequences, only the different CDR3 sequences are shown (there are 141 unique sequences in this clone but only 12 unique CDR3 sequences). Each dot is a unique sequence, and each dot color corresponds to a unique CDR3 sequence, and the dot size reflects the total copy number of each sequence. The x-axis shows the percent of Hamming distance calculated based on the number of nucleotides at which two sequences of the same length differ.

3.2. Dominant IGHV gene usage differs between vaccinated and unvaccinated fish and IgM repertoires are more diverse than the IgZ repertoires

3.2.1. Amplification of the heavy chain V, D, and J gene segments

Ig repertoires isolated from the whole kidney and intestine were used as controls for total V_H family expression and IgM and IgZ expression, respectively (Zhang et al., 2010, reviewed in Bilal et al., 2021). All associated IgH VDJ genes were amplified from the whole kidney or intestine Ig libraries; a total of 40 IGHV, 7 IGHD, and 7 IGHJ gene segments were found in these repertoires (Table 3.7). IgM repertoires had a higher IGHV gene usage diversity than the IgZ repertoires from both the whole kidney and intestine Ig repertoires. These two repertoires are from two different fish; however, IGHV1-4 gene segment had the highest frequency in both IgM repertoires (Table 3.7). Similarly, IGHV1-4 gene segment was favored in the IgZ repertoire from the whole kidney; however, IGHV1-1 had the highest frequency among the IGHV genes in the IgZ repertoire from the intestine (Table 3.7).

The IgM-associated IGHD and IGHJ genes were amplified from the whole kidney and intestine IgM repertoires. Although the frequencies of IGHD and IGHJ genes were similar in the IgM repertoires from the whole kidney and intestine, IGHD2-4 had the highest frequency in the whole kidney IgM repertoire; however, in the IgM repertoire isolated from the intestine, the most used IGHD gene was IGHD2-1 (Table 3.7). On the other hand, IGHJ2-1 and IGHJ2-4 had the highest frequencies among all the IGHJ genes in the IgM repertoires of the whole kidney and intestine, respectively (Table 3.7). The two IgZ associated IGHD and IGHJ genes were amplified from the

IgZ repertoires from the whole kidney and intestine (Table 3.7). Unlike the IgM repertoires, the frequencies of IGHD and IGHJ genes were different in the IgZ repertoires from the whole kidney and intestine. In the IgZ repertoire from the whole kidney, IGHD1-1 and IGHD1-2 genes had similar frequencies; likewise, comparable frequencies for IGHJ1-1 and IGHJ1-1 were found in this repertoire (Table 3.7). However, IGHD1-2 had a significantly higher frequency than IGHD1-1 in the IgZ repertoire isolated from the intestine, and IGHJ1-1 was favored in this repertoire (Table 3.7).

3.2.2. Gene usage analysis

To compare the diversity of the IGHV gene segment among repertoires from unvaccinated fish MMΦCs, repertoires from vaccinated fish MMΦCs, repertoires of IgM and IgZ isotypes, and between repertoires from unvaccinated and vaccinated fish MMΦCs, I used countGenes subcommand in the Alakazam package (Stern et al., 2014, Gupta et al., 2015).

Using Ig repertoires isolated from MMΦCs from both unvaccinated and vaccinated fish spleen and kidney, I examined the IGHV gene usage for IgM and IgZ isotypes. Overall, the IGHV gene usage diversity is higher in the Ig repertoires of IgM isotype compared to the IgZ repertoires from both unvaccinated and vaccinated fish MMΦCs (Fig.3.5, Fig.3.6, Fig.3.7, Fig.3.8). Ig repertoires from MMΦCs from unvaccinated fish of IgM isotype favored some V_H-elements (Fig.3.5). In the IgM repertoires isolated from vaccinated fish MMΦCs, some of the common V_H-elements were retained, and some of the least used genes in the unvaccinated fish samples became the most used ones in specific repertoires from vaccinated fish MMΦCs (Fig.3.5, Fig.3.6). IgZ repertoires from

unvaccinated fish MMΦCs had a low IGHV gene usage diversity; a few IGHV genes were almost always favored in these repertoires (Fig.3.7). However, in the IgZ repertoires from vaccinated fish MMΦCs, the favored IGHV genes in the unvaccinated fish MMΦCs IgZ repertoires became the least used ones, and some of the rare genes became the most used ones (Fig.3.7, Fig.3.8).

Analysis of the IGHV gene usage for IgM isotype from unvaccinated fish MMΦCs revealed that a few IGHV gene segments were favored in these repertoires (e.g., IGHV5-1, IGHV1-4, IGHV4-6, IGHV13-2, IGHV9-1) (Fig.3.5). IGHV5-1 and IGHV4-6 gene segments were common to all the IgM repertoires from unvaccinated fish MMΦCs, and they were used for up to 24% in most of these repertoires (Fig.3.5). In addition, IGHV9-1 and IGHV1-4 made up to 18% of the IGHV genes in IgM repertoires from unvaccinated fish MMΦCs, and IGHV13-2 was used for up to 10% in these repertoires (Fig.3.5). IGHV4-1 gene segment was favored in most IgM repertoires from unvaccinated fish MMΦCs; 27% of the Ig sequences of IgM repertoire in F4UKCa used IGHV4-1 (Fig.3.5). IgM repertoires isolated from MMΦCs from the spleen and kidney of the same unvaccinated fish shared most of their common IGHV genes (e.g., IGHV1-4, IGHV14-1, and IGHV5-1 in F3UKCa and F3USCa); however, higher frequency of specific IGHV genes was found in the spleen or kidney of these repertoires (e.g., IGHV1-2 16% and 2% in F2UKCa and F2USCa, respectively) (Fig.3.5).

In the IgM repertoires from vaccinated fish MMΦCs, some of the common V_H-elements were retained (e.g., IGHV5-1, IGHV1-4, IGHV4-6, IGHV9-1), and some of the least used genes in the IgM repertoires from unvaccinated fish MMΦCs became one of the most used genes in specific repertoires from vaccinated fish MMΦCs (e.g., IGHV4-2, IGHV10-1, and IGHV11-1) (Fig.3.6).

In addition, IgM repertoires isolated from MMΦCs from fish that were vaccinated with the same vaccine shared most of their dominant V_H-elements; for example, IgM repertoires isolated from both F8VKCa and F8VKCb shared most of their IGHV gene usage frequency with F6VKCb, these clusters are from two different fish vaccinated with PE- Alexa 647 and boosted with BSA- Alexa 647 (Fig.3.6). Similarly, clusters from two fish vaccinated with KLH (F11VKCa and F12VKCa) had comparable IGHV genes usage frequencies (Fig.3.6). Interestingly, IgM repertoires from unvaccinated fish MMΦCs used IGHV11-1 up to 2% only (Fig.3.5); the frequency of IGHV11-1 increased in one or two clusters from vaccinated fish for up to 12%. The frequency of IGHV11-1 gene in two of the IgM repertoires from vaccinated fish MMΦCs was less than 1%; these repertoires (F11VKCa and F12VKCa) are from KLH vaccinated fish MMΦCs. All the other repertoires from vaccinated fish MMΦCs are from fish injected with KLH, PE, and BSA -Alexa-647 conjugates (Fig.3.6).

The diversity of the IGHV gene usage in the IgZ repertoires from unvaccinated fish MMΦCs was significantly lower than the diversity of the V_H-element usage in the IgM repertoires from the same fish (Fig.3.5, Fig.3.7). Two IGHV genes (IGHV1-1 and IGHV13-2) always had the highest frequency compared to the other IGHV genes in these repertoires (Fig.3.7). However, after vaccination, the favored V_H-elements in the IgZ repertoires from unvaccinated fish MMΦCs became the least used ones, and some of the rare IGHV genes became the most used ones in the IgZ repertoires from vaccinated fish MMΦCs (Fig.3.7, Fig.3.8). For example, IGHV9-1, IGHV9-3, and IGHV14-1 were among the rare IGHV genes in the IgZ repertoires from unvaccinated fish MMΦCs; however, the frequencies of these genes increased significantly in the IgZ repertoires from vaccinated fish MMΦCs (Fig.3.8).

Analysis of the IGHD and IGHJ genes usage using Ig repertoires isolated from MMΦCs from unvaccinated fish revealed that the diversity of the D and J elements is limited in the IgZ repertoires compared to repertoires of IgM isotype (Fig.3.9, Fig.3.10, Fig.3.11, Fig.3.12). For IgZ, up to 70% of IGHD gene usage in most of the repertoires were biased to IGHD1-2; similarly, between 50% and 80% of the Ig sequences in all the Ig repertoires from unvaccinated fish MMΦCs used IGHJ1-1 (Fig.3.10, Fig.3.12). However, most Ig repertoires of IgM isotype used different IGHD and IGHJ genes for less than 30% and 35%, respectively (Fig.3.9, Fig.3.11).

Table 3.7. Frequency of amplified V, D, and J genes isolated from whole kidney and intestine.

Gene	F1UKW (IgM)[1]	F1UKW (IgZ)[1]	F4UIW (IgM)[1]	F4UIW (IgZ)[1]
IGHV1-4	0.12345	0.33656	0.16420	0.00194
IGHV14-1	0.08531	0.17940	0.08919	0.00325
IGHV9-1	0.07838	0.00105	0.13633	0.00000
IGHV4-6	0.07267	0.00231	0.04776	0.00107
IGHV5-1	0.05483	0.00041	0.09607	0.00000
IGHV13-2	0.05285	0.19847	0.05654	0.07516
IGHV10-1	0.05275	0.00018	0.03912	0.00000
IGHV1-1	0.05022	0.08513	0.03846	0.85894
IGHV4-1	0.03563	0.00326	0.01914	0.00000
IGHV5-5	0.03491	0.00003	0.05531	0.00000
IGHV8-3	0.03489	0.00046	0.01233	0.00000
IGHV5-4	0.02817	0.00003	0.01442	0.00004
IGHV1-2	0.02592	0.04472	0.01516	0.00420
IGHV4-8	0.02452	0.00072	0.00511	0.00246
IGHV7-1	0.02442	0.02466	0.06393	0.00000
IGHV5-3	0.02327	0.00000	0.01838	0.00000
IGHV9-2	0.01913	0.02717	0.01071	0.01081
IGHV4-2	0.01750	0.00041	0.00738	0.00000
IGHV9-4	0.01698	0.04644	0.01287	0.00851
IGHV2-1	0.01658	0.00000	0.00175	0.00067
IGHV1-3	0.01592	0.00236	0.01475	0.00000
IGHV11-2	0.01531	0.00249	0.01539	0.00000
IGHV4-9	0.01321	0.00174	0.00887	0.00000
IGHV6-2	0.01282	0.00000	0.00636	0.00000
IGHV2-2	0.01148	0.00010	0.00379	0.00000
IGHV4-3	0.01013	0.00264	0.00311	0.03006
IGHV8-1	0.00933	0.00067	0.00525	0.00000
IGHV3-2	0.00688	0.03310	0.01162	0.00234
IGHV5-8	0.00667	0.00000	0.00515	0.00000
IGHV6-1	0.00540	0.00164	0.00233	0.00000
IGHV4-5	0.00492	0.00003	0.00441	0.00000
IGHV5-7	0.00366	0.00046	0.00343	0.00000
IGHV4-7	0.00358	0.00000	0.00188	0.00000
IGHV11-1	0.00274	0.00074	0.00525	0.00000
IGHV9-3	0.00262	0.00162	0.00285	0.00000
IGHV1-5	0.00181	0.00095	0.00081	0.00055
IGHV5-2	0.00075	0.00000	0.00014	0.00000
IGHV8-4	0.00031	0.00003	0.00000	0.00000
IGHV2-3	0.00007	0.00003	0.00000	0.00000
IGHV8-2	0.00000	0.00000	0.00046	0.00000
IGHD2-4	0.22812	--	0.19787	--
IGHD2-3	0.19493	--	0.16524	--
IGHD2-1	0.16792	--	0.20745	--
IGHD2-2	0.10336	--	0.13735	--
IGHD2-5	0.09043	--	0.07836	--
IGHD1-1	--	0.47214	--	0.09275
IGHD1-2	--	0.32221	--	0.89498
IGHJ2-1	0.31710	--	0.24211	--
IGHJ2-2	0.25033	--	0.19892	--
IGHJ2-4	0.22436	--	0.26684	--
IGHJ2-5	0.12251	--	0.13970	--
IGHJ2-3	0.08545	--	0.15220	--
IGHJ1-2	--	0.51380	--	0.12383
IGHJ1-1	--	0.48563	--	0.87597

[1]F# - Fish # in group; U/V - Unvaccinated/Vaccinated; W - whole tissue; K - kidney: I – intestine.

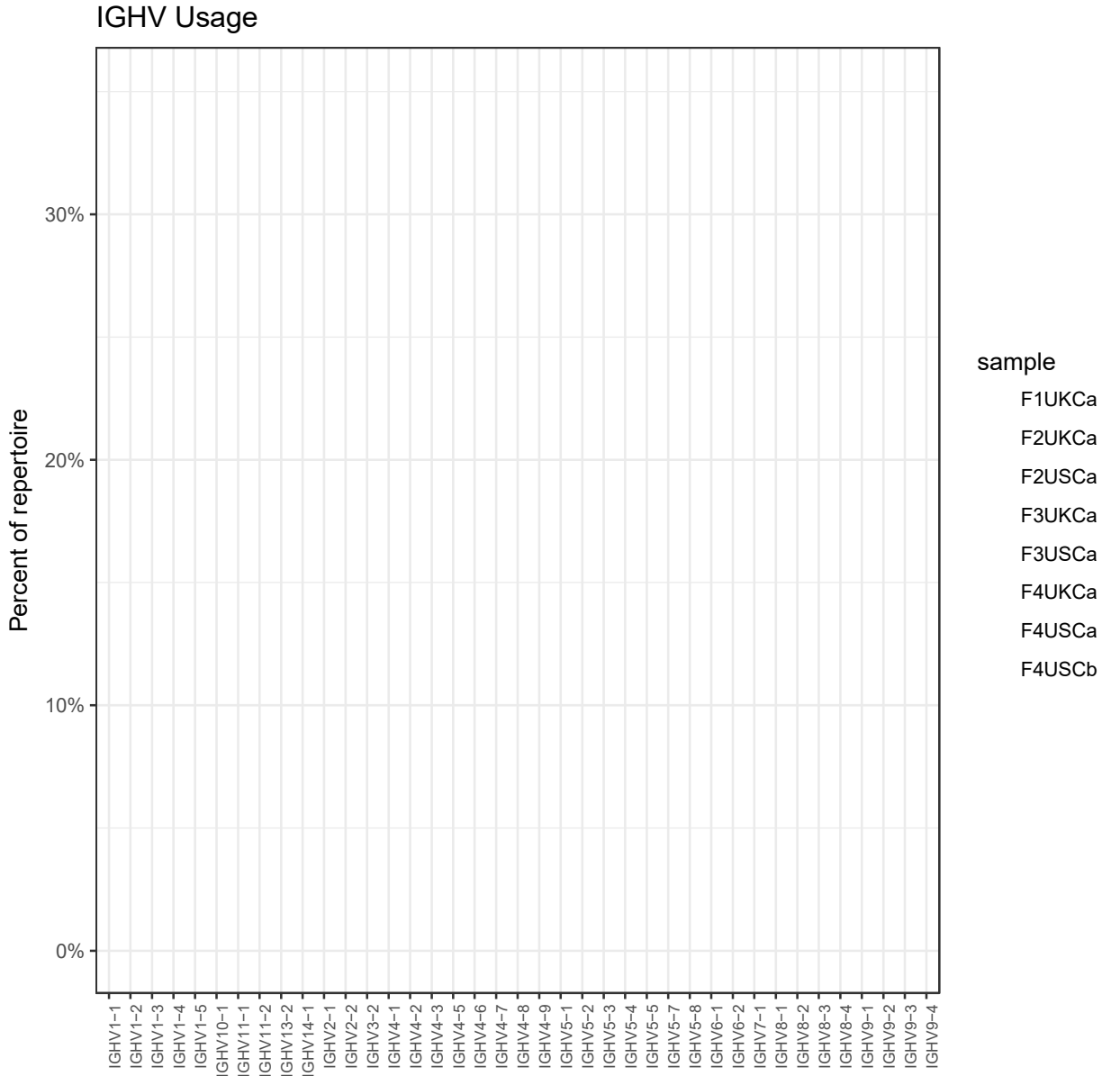


Figure 3.5. IGHV gene usage for IgM repertoires isolated from unvaccinated fish MMΦCs. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. Analysis was performed using countGenes subcommand in the Alakazam package.

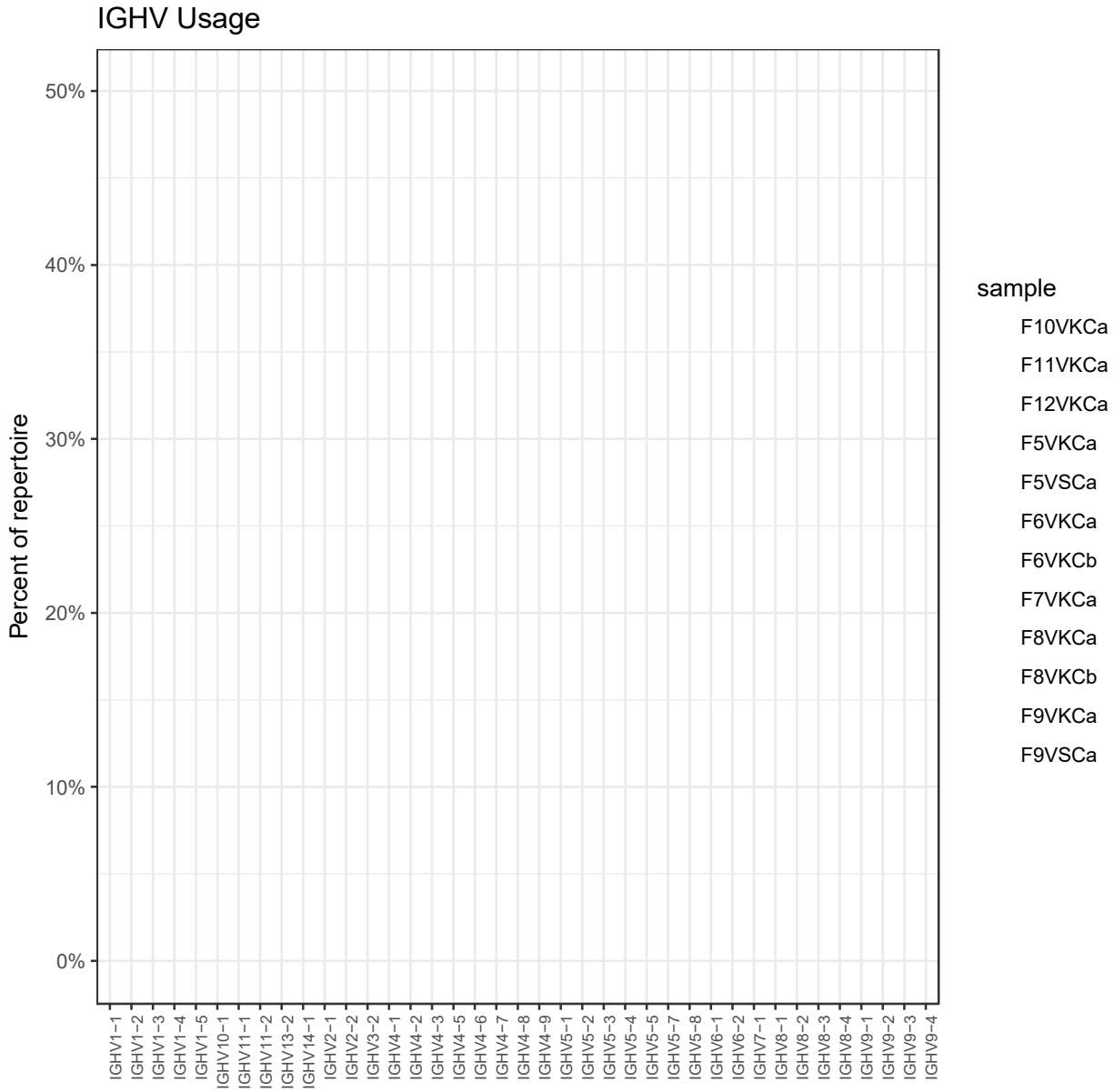


Figure 3.6. IGHV gene usage for IgM repertoires isolated from vaccinated fish MMΦCs. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. Analysis was performed using countGenes subcommand in the Alakazam package.

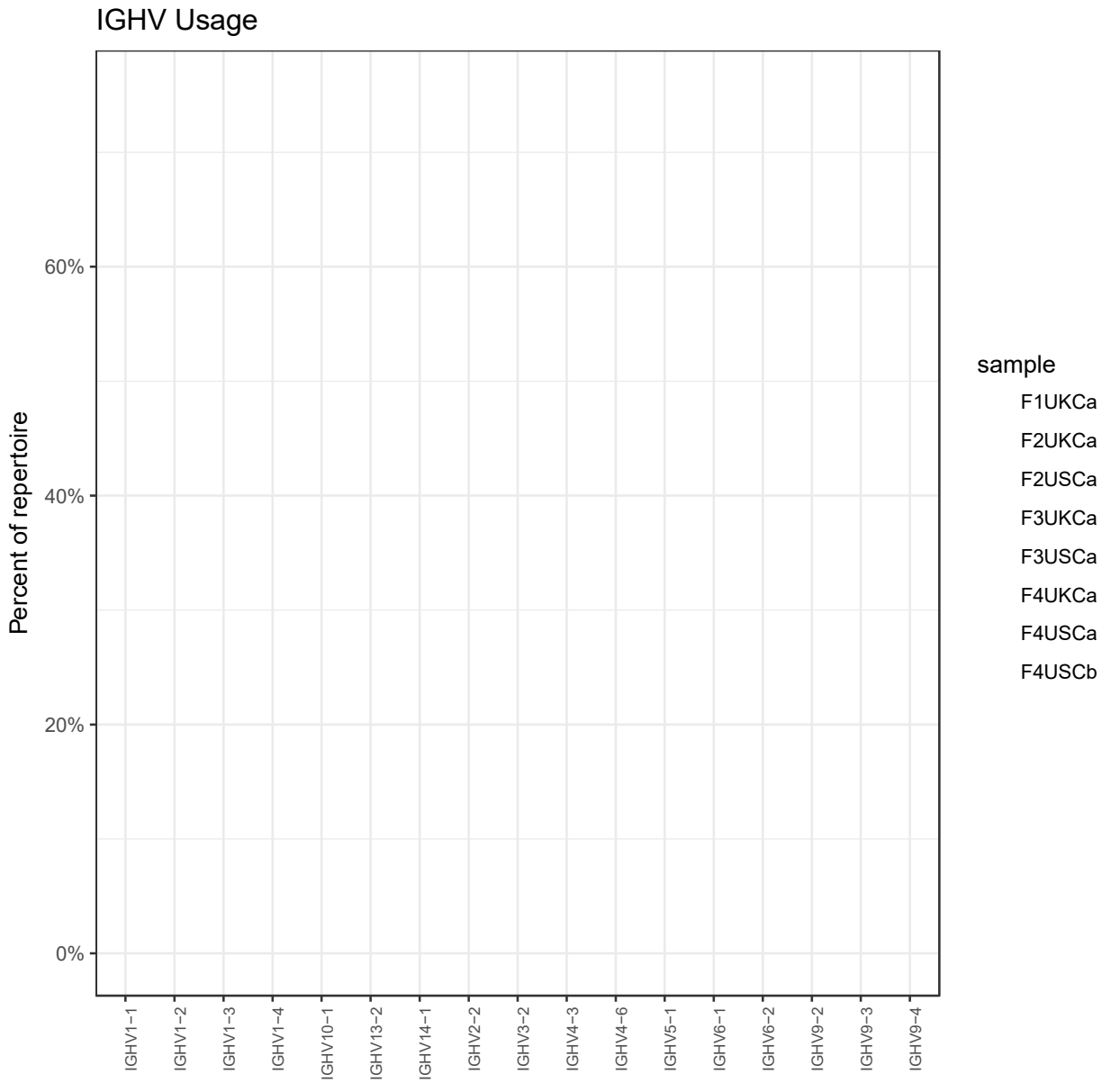


Figure 3.7. IGHV gene usage for IgZ repertoires isolated from unvaccinated fish MMΦCs. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. Analysis was performed using countGenes subcommand in the Alakazam package.

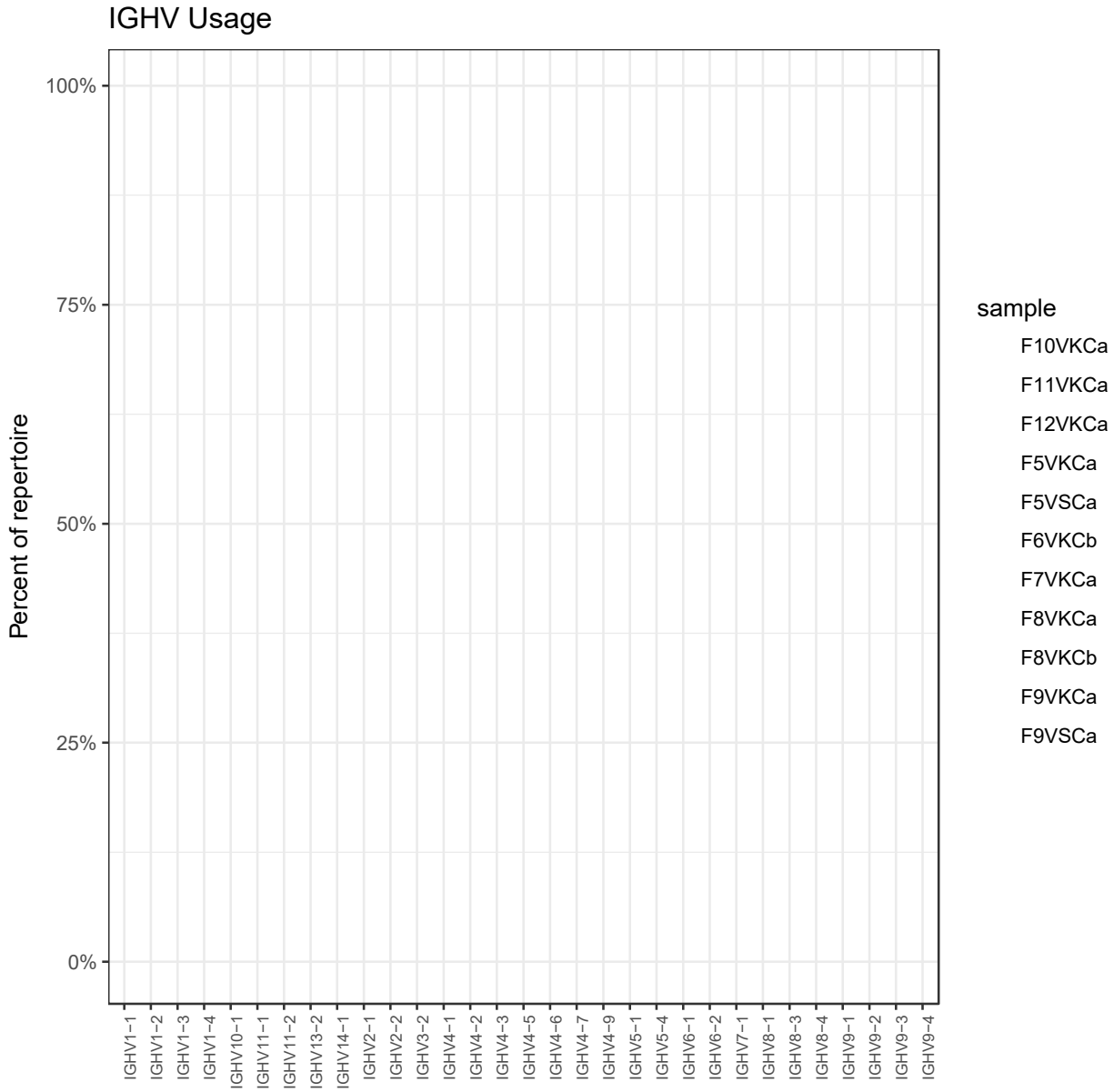


Figure 3.8. IGHV gene usage for IgZ repertoires isolated from vaccinated fish MMΦCs. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. Analysis was performed using countGenes subcommand in the Alakazam package.

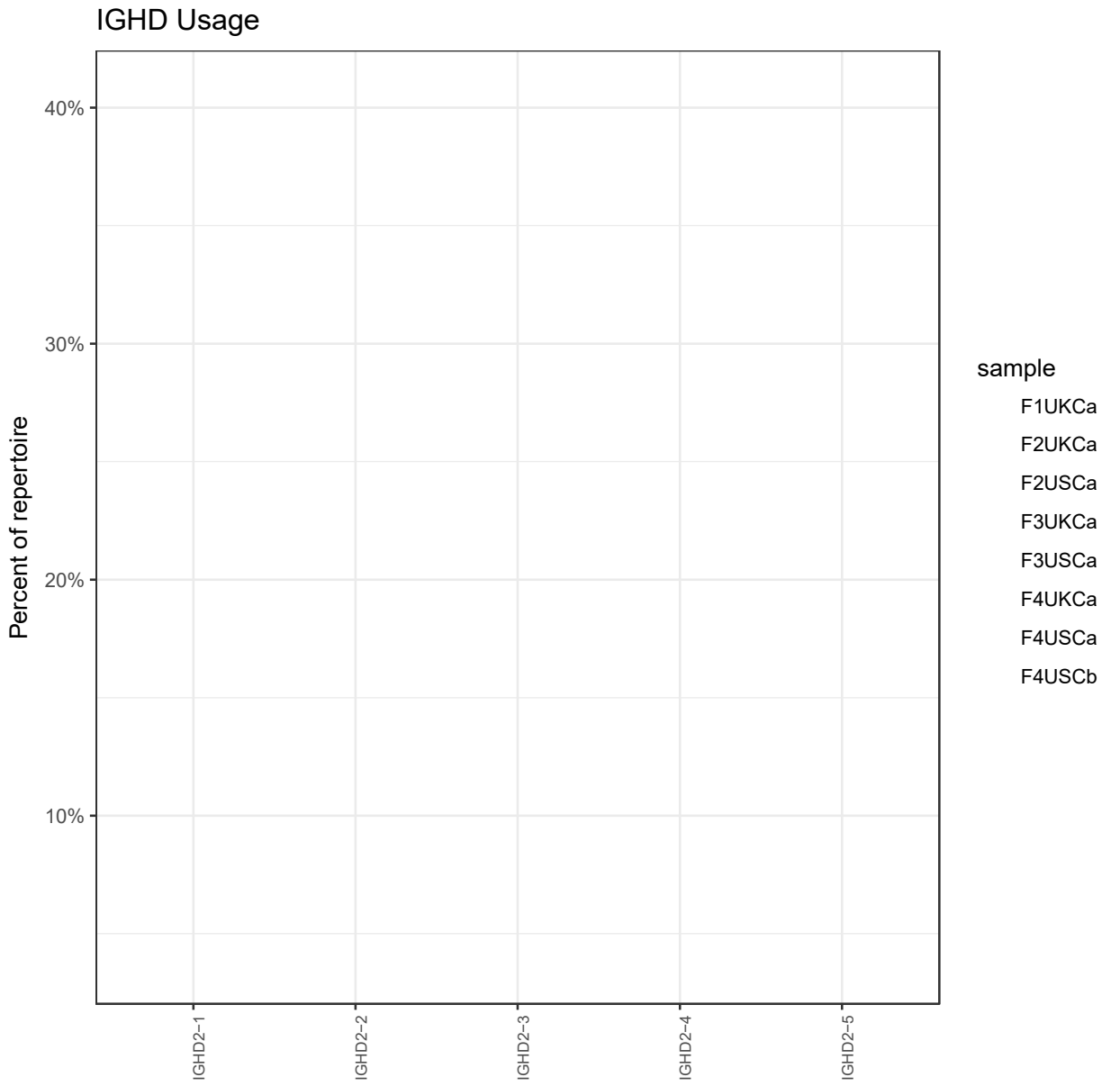


Figure 3.9. IGHD gene usage for IgM repertoires isolated from unvaccinated fish MMΦCs. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. Analysis was performed using countGenes subcommand in the Alakazam package.

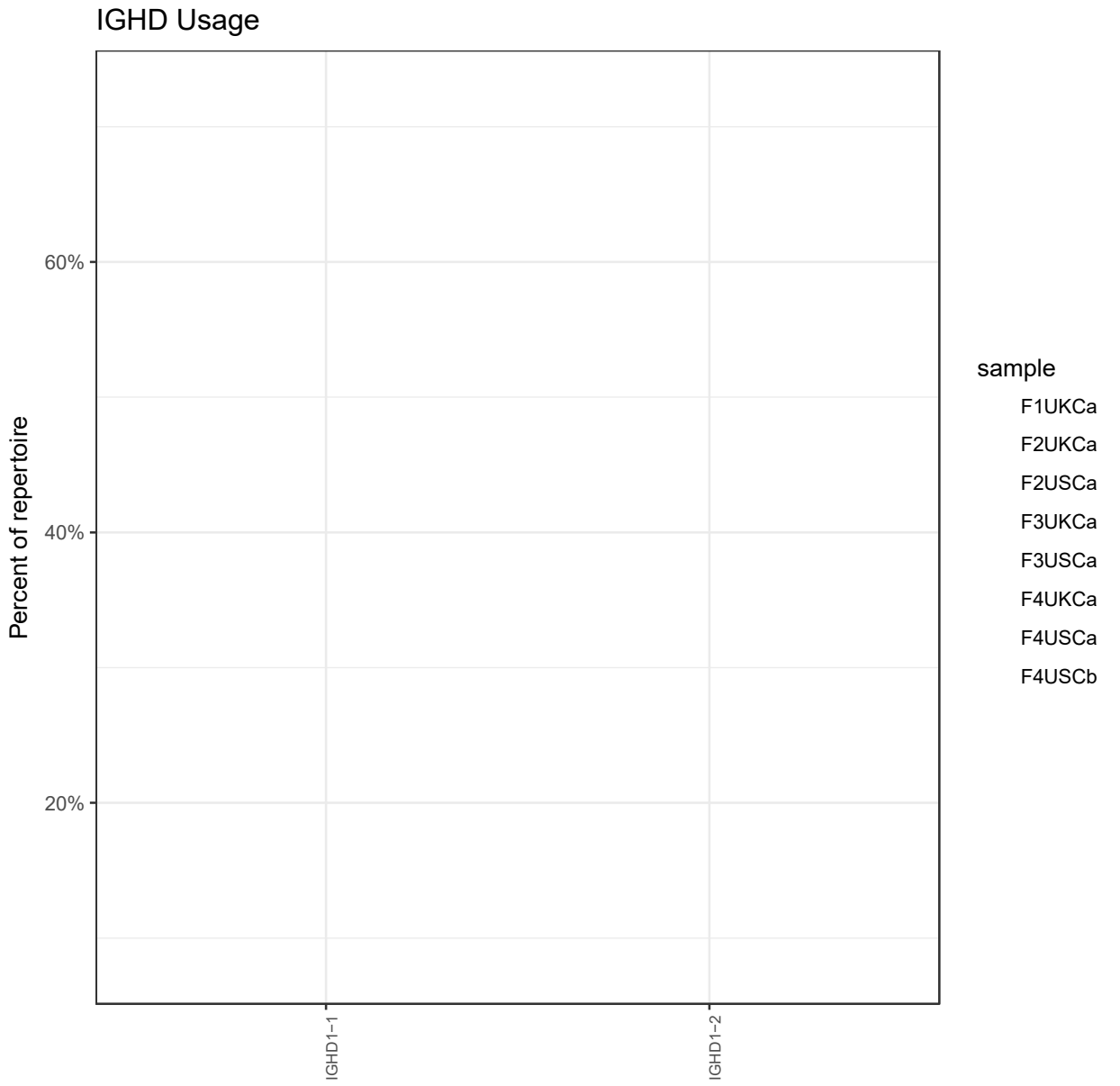


Figure 3.10. IGHD gene usage for IgZ repertoires isolated from unvaccinated fish MMΦCs. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. Analysis was performed using countGenes subcommand in the Alakazam package.

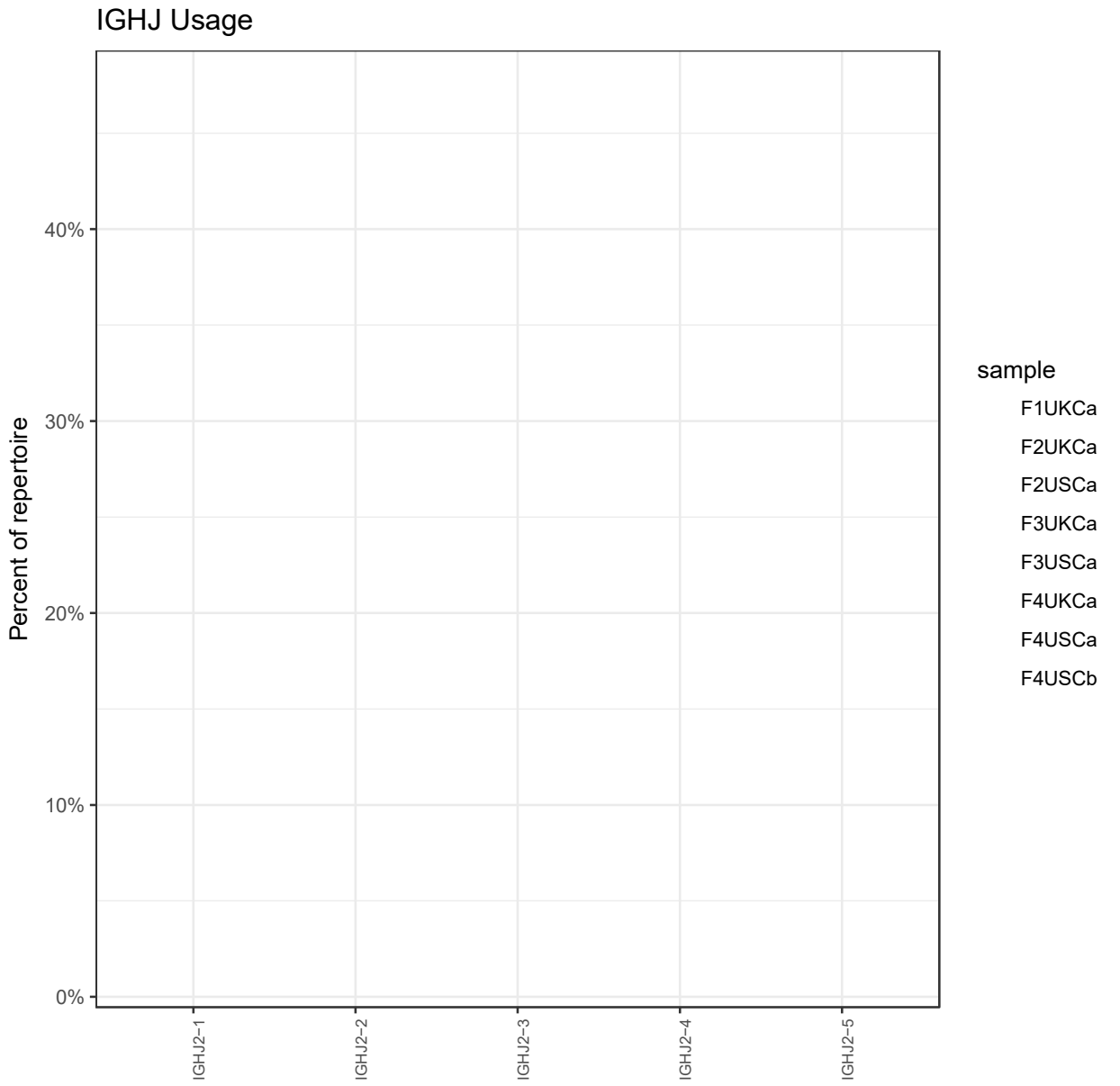


Figure 3.11. IGHJ gene usage for IgM repertoires isolated from unvaccinated fish MMΦCs. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. Analysis was performed using countGenes subcommand in the Alakazam package.

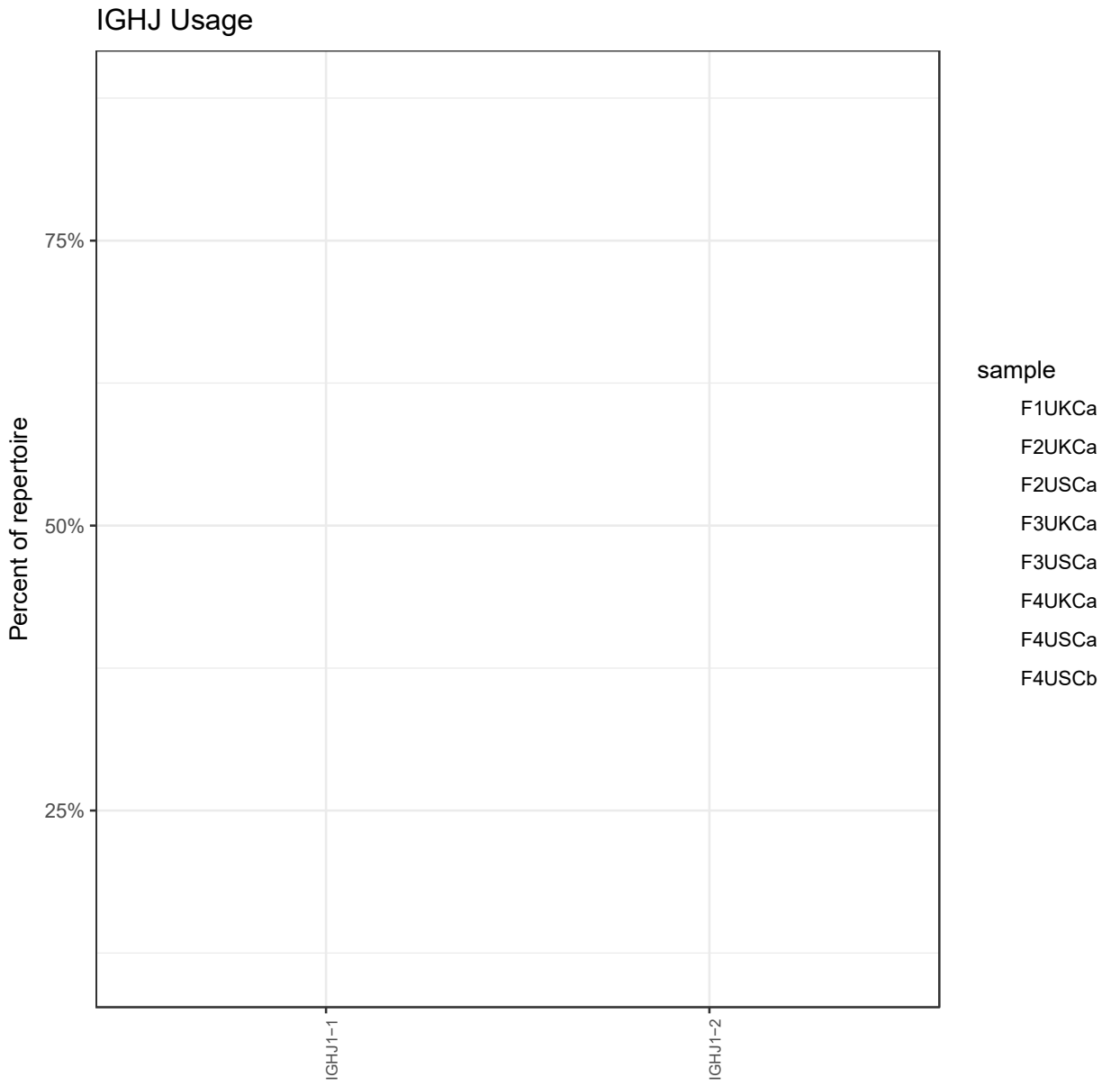


Figure 3.12. IGHJ gene usage for IgZ repertoires isolated from unvaccinated fish MMΦCs. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. Analysis was performed using countGenes subcommand in the Alakazam package.

3.3. The distribution of R/S mutations indicates that Ag-driven selection process occurs within MMΦCs

3.3.1. Estimates of the ratio of non-synonymous to synonymous substitutions (dN/dS) in the CDRs and FWRs

To determine if there is an active Ag-driven selection process within MMΦCs, I used IgPhyML HLP19 model (Hoehn et al., 2019) and the Ig repertoires generated from each cluster to examine the ratio of replacement to silent mutations (R/S) in the complementarity determining regions (CDRs) and the framework regions (FWRs). Although the R/S ratios varied among the IgM repertoires isolated from MMΦCs from unvaccinated and vaccinated fish, their CDRs always had a higher R/S estimate compared to the FWRs (Fig.3.13 A and B). Similarly, a higher R/S ratio was found in the CDRs compared to the FWRs in the Ig repertoires of IgZ isotype isolated from MMΦCs from vaccinated fish (Fig.3.13 D). However, four of the IgZ repertoires from MMΦCs from unvaccinated fish had a higher R/S ratio in their FWRs compared to their CDRs; these four IgZ repertoires are from the spleen and kidney of two different fish (F3UKCa, F3USCa, F4UKCa, and F4USCa (Fig.3.13 C)).

The R/S ratio of the Ig repertoires isolated from unvaccinated fish MMΦCs of IgM isotype was between 0.88 and 1.67 and between 0.33 and 0.53 in their CDRs and FWRs, respectively (Fig.3.13 A). The R/S ratio for IgZ repertoires from unvaccinated fish MMΦCs ranged from 0.45 to 1.77 and between 0.77 and 1.07 in their CDRs and FWRs, respectively (Fig.3.13 C). IgM repertoires isolated from MMΦCs from vaccinated fish had between 0.7 and 1.88 and between 0.28 and 0.48

R/S ratio in their CDRs and FWRs, respectively (Fig.3.13 B). The R/S ratio for the IgZ repertoires from MMΦCs from vaccinated fish was between 0.71 and 1.43 in their CDRs and between 0.32 and 0.95 in their FWRs (Fig.3.13 D).

3.3.2. Association between increased mean tree length and negative selection

Examination of the average mean tree length (i.e., the average expected substitutions per codon site) and the R/S ratio in the FWRs, revealed a positive relationship between the mean tree length and the R/S ratio in the FWRs (Fig.3.14). Ig repertoires of IgM isotype isolated from MMΦCs from unvaccinated and vaccinated fish had an average mean tree length of 0.13 (Table 3.5, Table 3.6). Similarly, IgZ repertoires from vaccinated fish MMΦCs had an average mean tree length of 0.18 (Table 3.6). On the other hand, Ig repertoires of IgZ isotype isolated from MMΦCs from unvaccinated fish had an average mean tree length equal to 0.65 (Table 3.5).

As shown in figure 3.14, Ig repertoires with longer mean tree length have a higher R/S ratio. All the IgM repertoires isolated from both unvaccinated and vaccinated fish MMΦCs had a mean tree length less than 0.25, and the R/S ratio of their FWRs was less than 0.55 (Fig.3.14). The slight increase in the mean tree length found in some Ig repertoires of IgZ isotype from MMΦCs from vaccinated fish is associated with an increased R/S ratio in their FWRs (Fig.3.14). The longest mean tree length (1.41) was found in the repertoires of IgZ isotype from unvaccinated fish MMΦCs, which also had the highest R/S ratio in the FWR (1.17) (Fig.3.14).

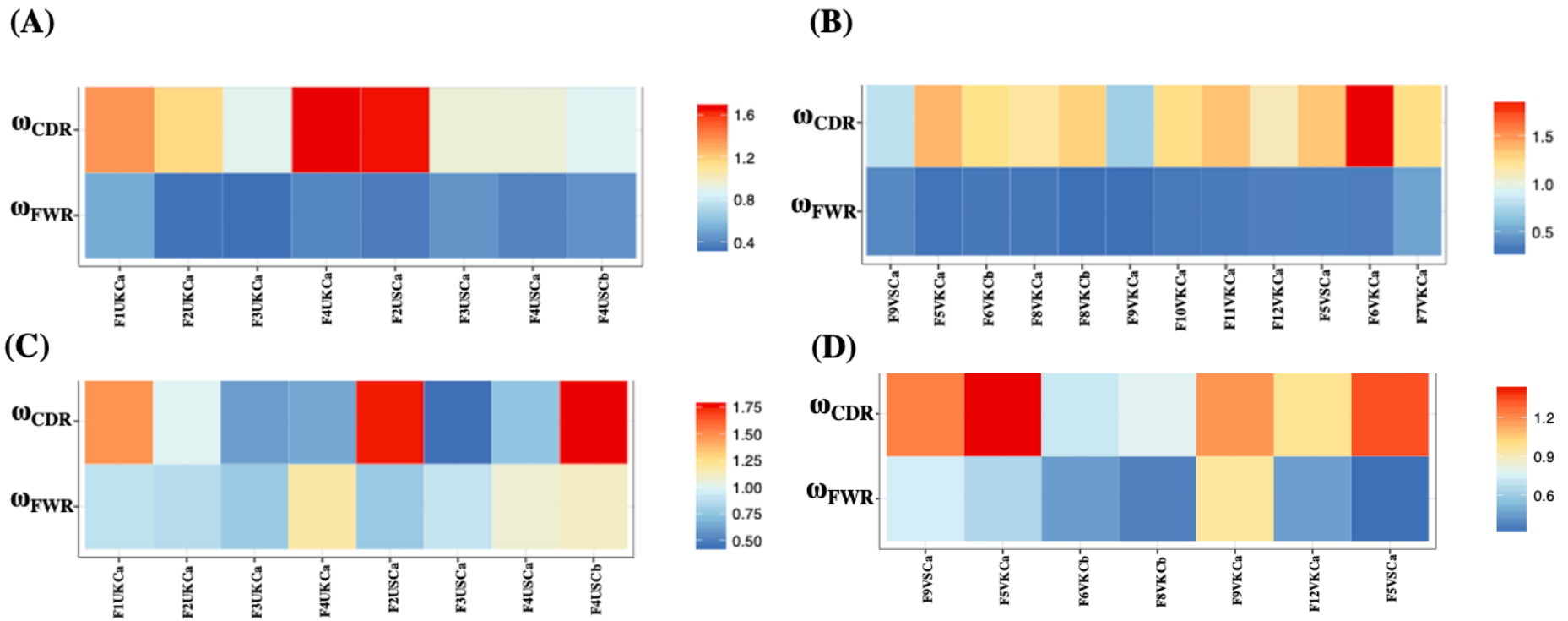

(A)
ωCDR
ωFWR
F1UKCa
F2UKCa
F3UKCa
F4UKCa
F2USCa
F3USCa
F4USCa
F4USCb
1.6
1.2
0.8
0.4
(B)
ωCDR
ωFWR
F9VSCa
F5VKCa
F6VKCb
F8VKCa
F8VKCb
F9VKCa
F10VKCa
F11VKCa
F12VKCa
F5VSCa
F6VKCa
F7VKCa
1.5
1.0
0.5
(C)
ωCDR
ωFWR
F1UKCa
F2UKCa
F3UKCa
F4UKCa
F2USCa
F3USCa
F4USCa
F4USCb
1.75
1.50
1.25
1.00
0.75
0.50
(D)
ωCDR
ωFWR
F9VSCa
F5VKCa
F6VKCb
F8VKCb
F9VKCa
F12VKCa
F5VSCa
1.2
0.9
0.6

Figure 3.13. Selection estimates (the ratio of replacement to silent mutations, R/S) for CDRs (*ω*CDR) and FWRs (*ω*FWR) using Ig repertoires isolated from MMΦCs from zebrafish. (A) Unvaccinated (IgM isotype). (B) Vaccinated (IgM isotype). (C) Unvaccinated (IgZ isotype). (D) Vaccinated (IgZ isotype). Refer to table 2.1 and table 3.1 for a detailed description of the samples used. The analysis was performed using IgPhyML HLP19 model.

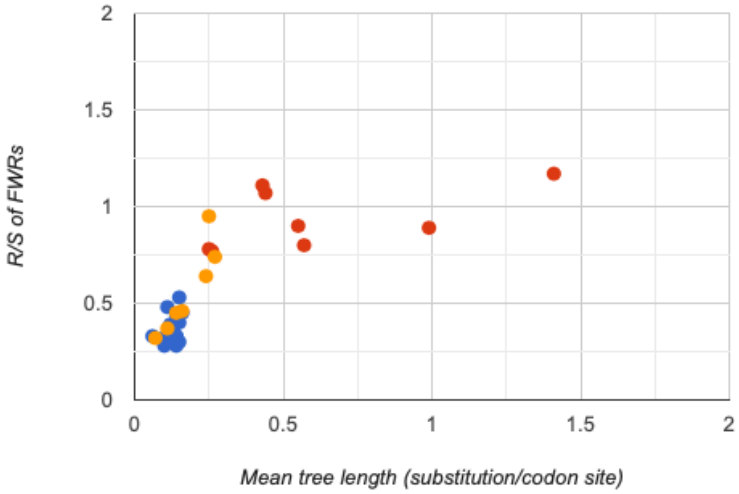


Figure 3.14. Association between the mean tree length and the ratio of R/S in the FWRs of the Ig repertoires isolated from MMΦCs from unvaccinated and vaccinated zebrafish. Blue - IgM repertoires from MMΦCs from both unvaccinated and vaccinated fish; Red - IgZ repertoires from MMΦCs from unvaccinated fish; Orange - IgZ repertoires from MMΦCs from vaccinated fish.

3.4. Nucleotide substitution patterns analysis reveals the involvement of AID and error-prone polymerases in the mutation process within MMΦCs

3.4.1. Hotspot motifs mutability

To determine if AID and error-prone polymerases are involved in the mutation process, I used IgPhyML HLP19 model (Hoehn et al., 2019) and the Ig repertoires generated from each cluster to examine the fold-change in mutability for hotspot motifs compared to regular motifs. Four hotspot motifs were examined, including WRC and its complement GYW and WA and its complement TW (where W = A/T, R = A/G, and Y = C/T); altered substitution rates occur only in the underlined bases. Altered mutation rates in WRC/ GYW and WA/ TW indicate the involvement of AID and error-prone polymerases, respectively, in the mutation process (reviewed in Pilzecker et al., 2019, Di Noia and Neuberger, 2007). Ig repertoires isolated from MMΦCs from both unvaccinated and vaccinated zebrafish had altered substitution rates in the analyzed hotspot motifs; WRC, GYW, and WA almost always had increased substitution rates with up to 4x increases in GYW substitution rate (Table 3.8, Table 3.9, Table 3.10, Table 3.11). Generally, GYW motif exhibited the highest substitution rate increases; on the other hand, TW had the lowest altered mutation rates compared to all the considered motifs in the analyzed repertoires of IgM and IgZ isotypes isolated from MMΦCs from both unvaccinated and vaccinated zebrafish (Table 3.8, Table 3.9, Table 3.10, Table 3.11).

In the IgM repertoires isolated from unvaccinated fish MMΦCs, Generally, WA motif experienced the highest altered mutation rate (1.30), and the largest substitution rate increases value was found

in WA motif (1.83) compared to the other studied motifs (Table 3.8). Altered substitution rates in GYW motif followed WA motif mutability; on average, these repertoires motif mutability of GYW was equal to 0.8 (Table 3.8). Mutation rates in the WRC motif followed the rates found in the GYW motif, motif mutability values of WRC were found to be between -0.34 and 0.8 in these repertoires (Table 3.8). The lowest altered mutation rates were found in the TW motif; substitution rate increases value was 0.03 for this motif in the IgM repertoires from unvaccinated fish MMΦCs (Table 3.8).

Ig repertoires isolated from MMΦCs from unvaccinated fish of IgZ isotype showed significantly higher substitution rates in WRC and its complement GYW compared to WA and its complement TW (Table 3.9). In these repertoires, substation rates of WRC and GYW motifs were, on average, 1.3 and 1.37, respectively (Table 3.9). Lower mutation rates were found in WA motif compared to WRC and GYW motifs; motif mutability of WA was between -0.20 and 0.34 (Table 3.9). The lowest mutation rate values were found in the TW motif, and they were between -0.49 and 0.62 (Table 3.9).

In the IgM repertoires isolated from vaccinated fish MMΦCs, GYW motif exhibited the highest substitution rate increases; the mutation rate values for this motif ranged from 1.04 to 3.06 (Table 3.10). Altered substitution rates in WA motif followed GYW motif mutability in these repertoires; on average, WA mutation rate was equal to 1.03 (Table 3.10). All the IgM repertoires from vaccinated fish MMΦCs had increased substitution rates in the WRC motif, and the substation rate values for this motif were between 0.1 and 2 (Table 3.10). Similar to the Ig repertoires from unvaccinated fish MMΦCs, the lowest altered mutation rates of all the motifs considered were

found in TW motif. substitution rate increases value for TW motif was between -0.12 and 0.53 (Table 3.10).

Ig repertoires isolated from MMΦCs from vaccinated fish of IgZ isotype overall had lower motif mutability values in all the analyzed motifs compared to the substitution rate increases found in the Ig repertoires from unvaccinated fish MMΦCs and the IgM repertoires from vaccinated fish MMΦCs (Table 3.11). In these repertoires, GYW motif experienced the largest substitution rate increases; on average, GYW motif mutability value was equal to 0.66. Mutation rates in GYW motif were followed by WA motif mutability; substitution rates for WA motif ranged from -0.27 to 0.8 (Table 3.11). Unlike the Ig repertoires from unvaccinated fish MMΦCs and IgM repertoires from vaccinated fish MMΦCs, the lowest mutation rate values were found in the WRC motif compared to the other studied motifs and the substation rate values for this motif were between -0.45 and 0.61 (Table 3.11). TW motif mutability value was between -0.27 and 1.22 in these repertoires (Table 3.11).

3.4.2. The ratio of transitions to transversions

Using Ig repertoires of both isotypes (IgM and IgZ) from MMΦCs from unvaccinated and vaccinated fish on average, the ratio of transition to transversion mutations was equal to 1.86 (Table 3.8, Table 3.9, Table 3.10, Table 3.11). Typically, the lowest transitions to transversions ratio values were found in the IgZ repertoires from unvaccinated fish MMΦCs, and the highest values of transitions to transversions ratio were detected in the IgM repertoires from vaccinated fish MMΦCs (Table 3.9, Table 3.10). In the Ig repertoires from unvaccinated fish MMΦCs, the

ratio of transitions to transversions ranged from 1.68 to 2.25 and from 1.33 to 1.74 for IgM and IgZ, respectively (Table 3.8, Table 3.9). The ratio of transitions to transversions in the repertoires isolated from MMΦCs from vaccinated fish was between 1.91 and 2.54 and between 1.56 and 2.37 for IgM and IgZ, respectively (Table 3.10, Table 3.11).

3.4.3. Nucleotide insertions and deletions (indels)

Using Ig repertoires isolated from MMΦCs from unvaccinated and vaccinated fish, I analyzed the frequency of nucleotide insertions and deletions (indels) in the productive Ig sequences. Examination of indels revealed that, in general, indels are infrequent in the Ig repertoires from both unvaccinated and vaccinated fish MMΦCs (Table 3.12, Table 3.13). On average, only 2.2% of the functional Ig sequences in these repertoires had indels (Table 3.12, Table 3.13). IgM repertoires from both unvaccinated and vaccinated fish MMΦCs had a higher average percentage of indels than IgZ repertoires (Table 3.12, Table 3.13).

In the Ig repertoires from MMΦCs isolated from unvaccinated fish, the percentage of indels was between 0.8% and 4.8% and between 0.59% and 1.94% of the productive Ig sequences for IgM and IgZ, respectively (Table 3.12). The frequency of indels in the repertoires from vaccinated fish MMΦCs ranged from 1% to 11.6% and from 0% to 4.49% for IgM and IgZ, respectively (Table 3.13). Some of the repertoires isolated from MMΦCs from different tissues of the same fish had a similar frequency of indels for IgM and IgZ isotypes, or in some cases, they had a similar frequency of indels in IgM or IgZ repertoires (Table 3.12, Table 3.13). As shown in table 3.13, repertoires isolated from MMΦCs from the spleen and kidney of the same fish (F9VSCa and F9VKCa) had

similar indels frequencies for both IgM and IgZ repertoires (Table 3.13). The highest values of indels were found in the IgM repertoires from vaccinated fish MMΦCs; the repertoires F6VKCa and F6VKCb are from two different clusters from the kidney of the same fish, 11.6% and 10.3%, respectively, of their productive Ig sequences had indels (Table 3.13).

3.4.4. Gene conversion

Using individual clones in the Ig repertoires from MMΦCs isolated from unvaccinated and vaccinated fish, I examined the contribution of gene conversion to the diversity of the Ig repertoire in zebrafish. While major gene conversion-like events or stretch of mutations were not detected in the Ig sequences from both unvaccinated and vaccinated fish MMΦCs, a few minor gene conversion-like events were found in a few Ig sequences.

Table 3.8. Fold change in mutability for WR<u>C</u>, <u>G</u>YW, W<u>A</u>, and <u>T</u>W hotspot motifs and transitions to transversions ratio (Ts: Tv), using Ig repertoires isolated from MMΦCs from unvaccinated zebrafish (IgM isotype). The analysis was performed using IgPhyML HLP19 model.

Fish[1]	WR<u>C</u>[2]	<u>G</u>YW[2]	W<u>A</u>[2]	<u>T</u>W[2]	(Ts:Tv)
F1UKCa	-0.11	1.10	0.77	0.03	2.06
F2UKCa	0.81	1.38	1.64	0.44	1.96
F2USCa	0.51	0.45	1.06	0.16	2.25
F3UKCa	-0.34	-0.06	0.76	-0.29	1.82
F3USCa	-0.04	1.23	1.02	-0.08	1.69
F4UKCa	0.30	0.04	1.83	0.07	1.68
F4USCa	0.11	1.05	0.96	0.23	1.84
F4USCb	0.00	0.44	0.94	-0.07	1.70
Average	**0.16**	**0.80**	**1.30**	**0.03**	**1.88**
SD	**0.37**	**0.49**	**0.36**	**0.25**	**0.19**

[1]F# - Fish # in group; U - Unvaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.
[2]only the underlined nucleotide experiences a change in mutability.

Table 3.9. Fold change in mutability for WR<u>C</u>, <u>G</u>YW, W<u>A</u>, and <u>T</u>W hotspot motifs and transitions to transversions ratio (Ts: Tv), using Ig repertoires isolated from MMΦCs from unvaccinated zebrafish (IgZ isotype). The analysis was performed using IgPhyML HLP19 model.

Fish[1]	WR<u>C</u>[2]	<u>G</u>YW[2]	W<u>A</u>[2]	<u>T</u>W[2]	(Ts:Tv)
F1UKCa	2.12	0.90	0.07	-0.49	1.33
F2UKCa	1.27	1.95	0.34	0.62	1.74
F2USCa	0.91	1.56	0.11	-0.48	1.42
F3UKCa	0.90	1.35	0.12	0.28	1.69
F3USCa	1.84	1.77	0.18	-0.32	1.68
F4UKCa	0.95	0.74	-0.11	0.51	1.51
F4USCa	0.35	0.52	-0.20	-0.01	1.52
F4USCb	2.00	1.52	0.21	-0.18	1.74
Average	**1.29**	**1.37**	**0.09**	**-0.08**	**1.58**
SD	**0.59**	**0.46**	**0.16**	**0.38**	**0.15**

[1]F# - Fish # in group; U - Unvaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.
[2]only the underlined nucleotide experiences a change in mutability.

Table 3.10. Fold change in mutability for WR<u>C</u>, <u>G</u>YW, W<u>A</u>, and <u>T</u>W hotspot motifs and transitions to transversions ratio (Ts: Tv), using Ig repertoires isolated from MMΦCs from vaccinated zebrafish (IgM isotype). The analysis was performed using IgPhyML HLP19 model.

Fish[1]	WR<u>C</u>[2]	<u>G</u>YW[2]	W<u>A</u>[2]	<u>T</u>W[2]	(Ts:Tv)
F5VKCa	0.60	2.40	1.35	0.39	1.94
F5VSCa	0.63	1.85	1.03	0.14	2.33
F6VKCa	0.70	2.48	1.12	-0.02	2.33
F6VKCb	1.00	3.06	1.12	0.03	2.40
F7VKCa	2.0	1.74	1.04	0.53	1.99
F8VKCa	0.83	2.52	0.78	0.09	2.44
F8VKCb	0.82	2.61	0.63	0.05	2.54
F9VKCa	0.10	1.04	0.88	-0.02	2.03
F9VSCa	0.61	1.89	1.25	0.27	1.91
F10VKCa	1.18	2.76	1.10	0.19	2.02
F11VKCa	0.57	1.34	1.12	0.44	2.23
F12VKCa	0.15	1.16	0.70	-0.12	2.29
Average	**0.60**	**1.98**	**1.03**	**0.18**	**2.20**
SD	**0.30**	**0.59**	**0.22**	**0.21**	**0.21**

[1]F# - Fish # in group; V - Vaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.
[2]only the underlined nucleotide experiences a change in mutability.

Table 3.11. Fold change in mutability for WR<u>C</u>, <u>G</u>YW, W<u>A</u>, and <u>T</u>W hotspot motifs and transitions to transversions ratio (Ts: Tv), using Ig repertoires isolated from MMΦCs from vaccinated zebrafish (IgZ isotype). The analysis was performed using IgPhyML HLP19 model.

Fish[1]	WR<u>C</u>[2]	<u>G</u>YW[2]	W<u>A</u>[2]	<u>T</u>W[2]	(Ts:Tv)
F5VKCa	-0.35	-0.20	0.35	-0.14	1.56
F5VSCa	0.61	0.54	0.81	-0.27	2.17
F6VKCb	-0.39	1.24	0.57	0.04	1.93
F8VKCb	0.14	1.33	0.41	-0.04	2.37
F9VKCa	0.44	1.32	-0.27	1.22	1.56
F9VSCa	-0.70	-0.19	-0.04	-0.01	1.83
F12VKCa	-0.45	0.59	0.11	0.43	1.57
Average	**-0.10**	**0.66**	**0.28**	**0.18**	**1.86**
SD	**0.46**	**0.62**	**0.34**	**0.47**	**0.30**

[1]F# - Fish # in group; V - Vaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.
[2]only the underlined nucleotide experiences a change in mutability.

Table 3.12. The percentage of nucleotide insertions and deletions (indels) in the Ig repertoires isolated from MMΦCs from unvaccinated zebrafish.

Fish[1]	Indels % (IgM)	Indels % (IgZ)
F1UKCa	3.20	1.60
F2UKCa	4.80	1.00
F2USCa	3.50	0.59
F3UKCa	0.80	1.40
F3USCa	1.40	1.39
F4UKCa	2.30	1.17
F4USCa	1.60	1.94
F4USCb	0.80	1.60
Average	**2.30**	**1.34**
SD	**1.34**	**0.39**

[1]F# - Fish # in group; U - Unvaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.

Table 3.13. The percentage of nucleotide insertions and deletions (indels) in the Ig repertoires isolated from MMΦCs from vaccinated zebrafish.

Fish[1]	Indels % (IgM)	Indels % (IgZ)
F5VSCa	5.10	0.45
F5VKCa	2.50	0.64
F6VKCa	11.60	–
F6VKCb	10.30	3.60
F7VKCa	1.13	0.00
F8VKCa	2.60	0.00
F8VKCb	2.00	1.40
F9VKCa	1.00	1.88
F9VSCa	1.20	1.86
F10VKCa	1.70	4.49
F11VKCa	6.00	0.89
F12VKCa	1.89	1.60
Average	**3.92**	**1.53**
SD	**3.48**	**1.36**

[1]F# - Fish # in group; V - Vaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.

3.5. Diversity analysis indicates the presence of an effective recruitment mechanism within MMΦCs

To determine if there is an effective recruitment and diversifying mechanism within MMΦCs, I used Hill numbers and clonal abundance to examine the clonal diversity of B-cell clonotypes within individual clusters.

3.5.1. Clonal diversity indices (Hill numbers)

Using Ig repertoires generated from unvaccinated and vaccinated fish MMΦCs and the whole kidney and intestine, the generalized diversity index (Hill numbers) revealed that Ig repertoires (IgM and IgZ combined) isolated from individual MMΦCs from unvaccinated and vaccinated fish have low clonal diversity and more related clones (Fig.3.15, Fig.3.16). All the repertoires isolated from unvaccinated and vaccinated fish MMΦCs had low diversity numbers at all the diversity orders (q values) compared with the Ig repertoire isolated from the whole kidney (Fig.3.15, Fig.3.16). In order to be able to compare multiple Ig repertoires, I used repeated resampling to correct for the variability in total sequence counts among the repertoires.

In the Ig repertoires isolated from unvaccinated fish MMΦCs, between 150 and 425 clones were sampled (species richness at q = 0) (Fig.3.15). Ig repertoires isolated from the whole kidney and intestine had around 900 and 600 clones, respectively (Fig.3.15). At the inverse of Simpson's concentration index (q = 2), which considers species in proportion to their frequency and ignores rare species, the diversity number for the Ig repertoires from unvaccinated fish MMΦCs and the

intestine was less than 80 (q = 2, Fig.3.15). On the other hand, the diversity number for the Ig repertoire from the whole kidney at the same order (q = 2) was around 650 (Fig.3.15). As the diversity order increased, the diversity number for the Ig repertoires decreased; however, the diversity number for the whole kidney Ig repertoire remained relatively high even at higher diversity orders (e.g., q = 8) (Fig.3.15).

Species richness of the Ig repertoires isolated from vaccinated fish MMΦCs was between 100 and 425 clones (Fig.3.16). At the diversity order, which represents the dominant clones, the diversity number for the Ig repertoires from vaccinated fish MMΦCs was around 120 (q = 2, Fig.3.16). Similar to the Ig repertoires from unvaccinated fish MMΦCs, the diversity number for the Ig repertoires from vaccinated fish MMΦCs decreased as the order increased (Fig.3.16).

3.5.2. Clonal abundance

Using complete clonal abundance distribution, I found a few highly abundant clones in the Ig repertoires isolated from unvaccinated and vaccinated fish MMΦCs and the intestine (Fig.3.17, Fig.3.18). Conversely, clones in the Ig repertoire isolated from the whole kidney had similar abundance and high evenness, as indicated by the shallow gradient or the lower slope (Fig.3.17). Using repeated resampling, all the Ig repertoires were standardized by sample completeness to correct for the variability in total sequence counts.

Between 500 and 1200 clones were ranked in the Ig repertoires from unvaccinated and vaccinated fish MMΦCs, and around 1500 and 1200 clones were ranked in the Ig repertoires from the whole

kidney and intestine, respectively (Fig.3.17, Fig.3.18). In the Ig repertoires from unvaccinated fish MMΦCs, the abundance of the most abundant clone (at 10^0) was between 7% and 25% (Fig.3.17), and in the repertoires from vaccinated fish MMΦCs the abundance of the clone with the highest-ranking (10^0) was between 4% and 35% (Fig.3.18). The abundance of the most abundant clone in the whole kidney and intestine was around 2% and 25%, respectively (Fig.3.17). The steep gradient found in all the Ig repertoires from unvaccinated and vaccinated fish MMΦCs, and the intestine indicates that the abundance of the clones in these repertoires is not evenly distributed; a few clones have much higher abundance than the low-ranking clones (i.e., clones at 10^2 and 10^3) (Fig.3.17, Fig.3.18).

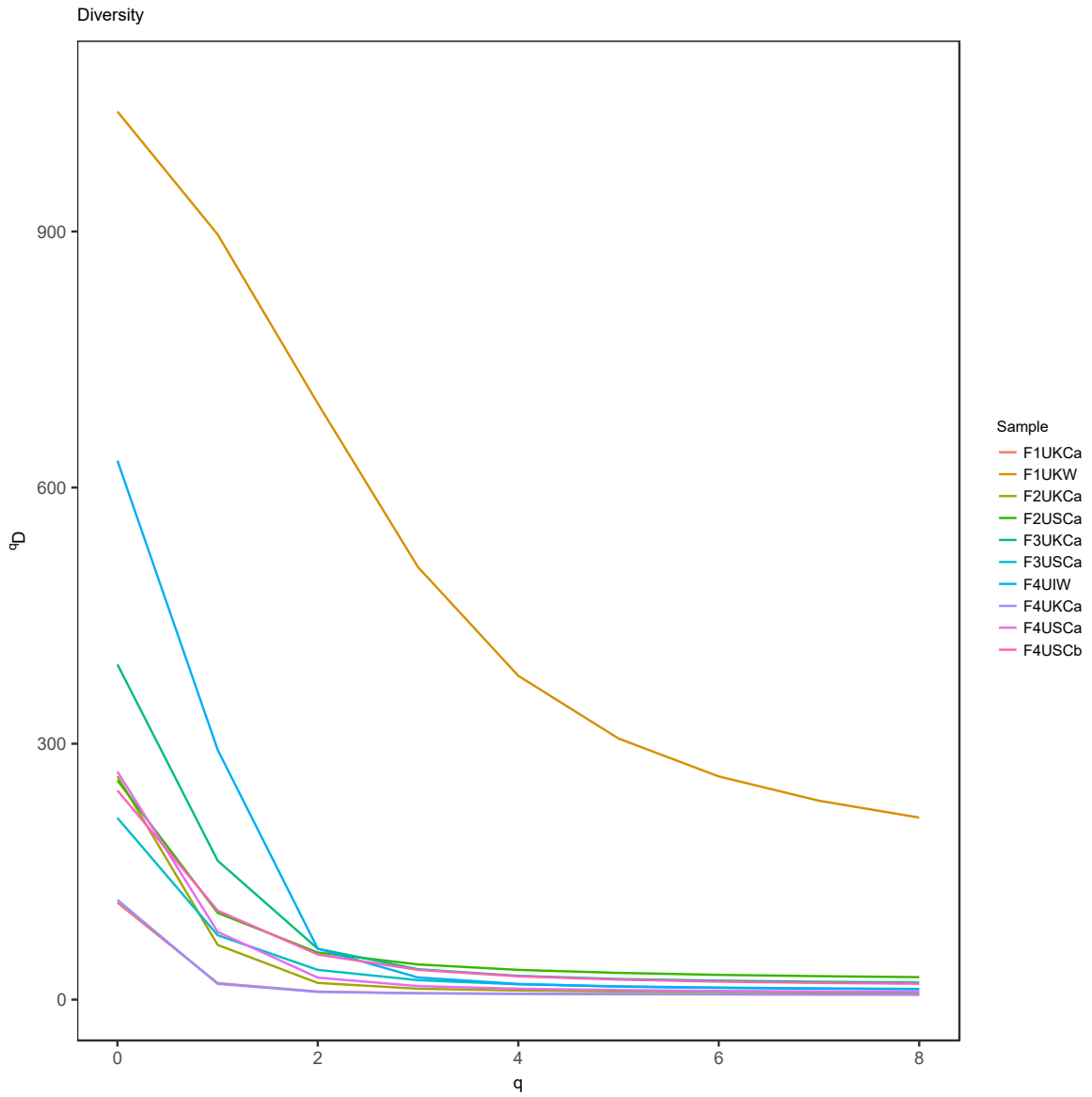


Figure 3.15. Hill numbers (diversity index) using Ig repertoires isolated from unvaccinated zebrafish MMΦCs and the whole kidney and intestine. ^{q}D is the diversity number, q values are the orders of the diversity number. q = 0 is the total number of clones in a sample, q = 1 is the exponential of Shannon's entropy index (considers clones in proportion to their frequency), and q = 2 is the inverse of Simpson's concentration index (considers clones in proportion to their frequency and ignores rare clones). q= ∞ is the reciprocal of the proportional richness of the commonest clones. Each line represents an individual repertoire; the shaded areas are 95% confidence intervals. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. The analysis was performed using alphaDiversity subcommand in the Alakazam package.

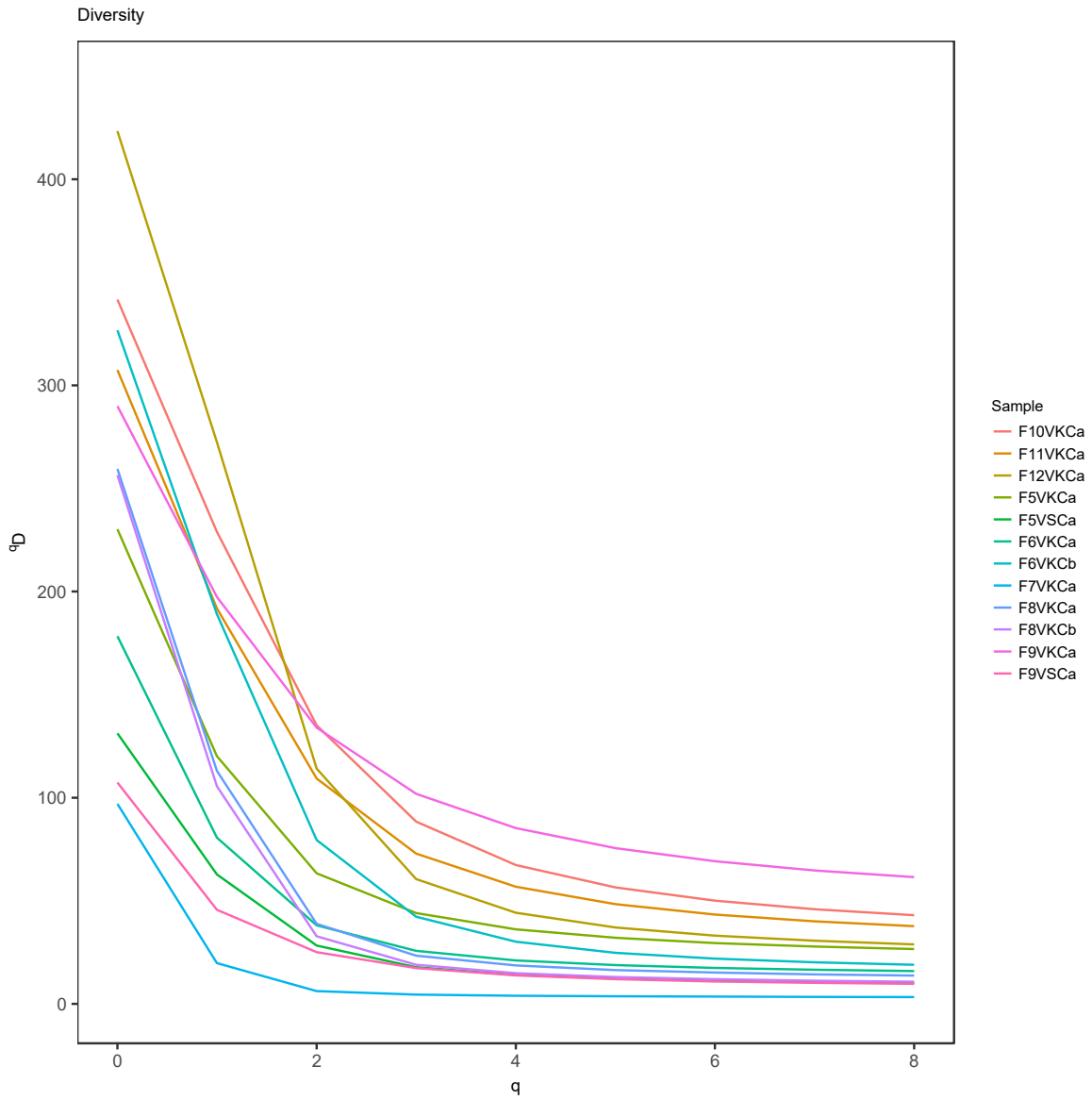


Figure 3.16. Hill numbers (diversity index) using Ig repertoires isolated from vaccinated zebrafish MMΦCs. ^{q}D is the diversity number, q values are the orders of the diversity number. q = 0 is the total number of clones in a sample, q = 1 is the exponential of Shannon's entropy index (considers clones in proportion to their frequency), and q = 2 is the inverse of Simpson's concentration index (considers clones in proportion to their frequency and ignores rare clones). q= ∞ is the reciprocal of the proportional richness of the commonest clones. Each line represents an individual repertoire; the shaded areas are 95% confidence intervals. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. The analysis was performed using alphaDiversity subcommand in the Alakazam package.

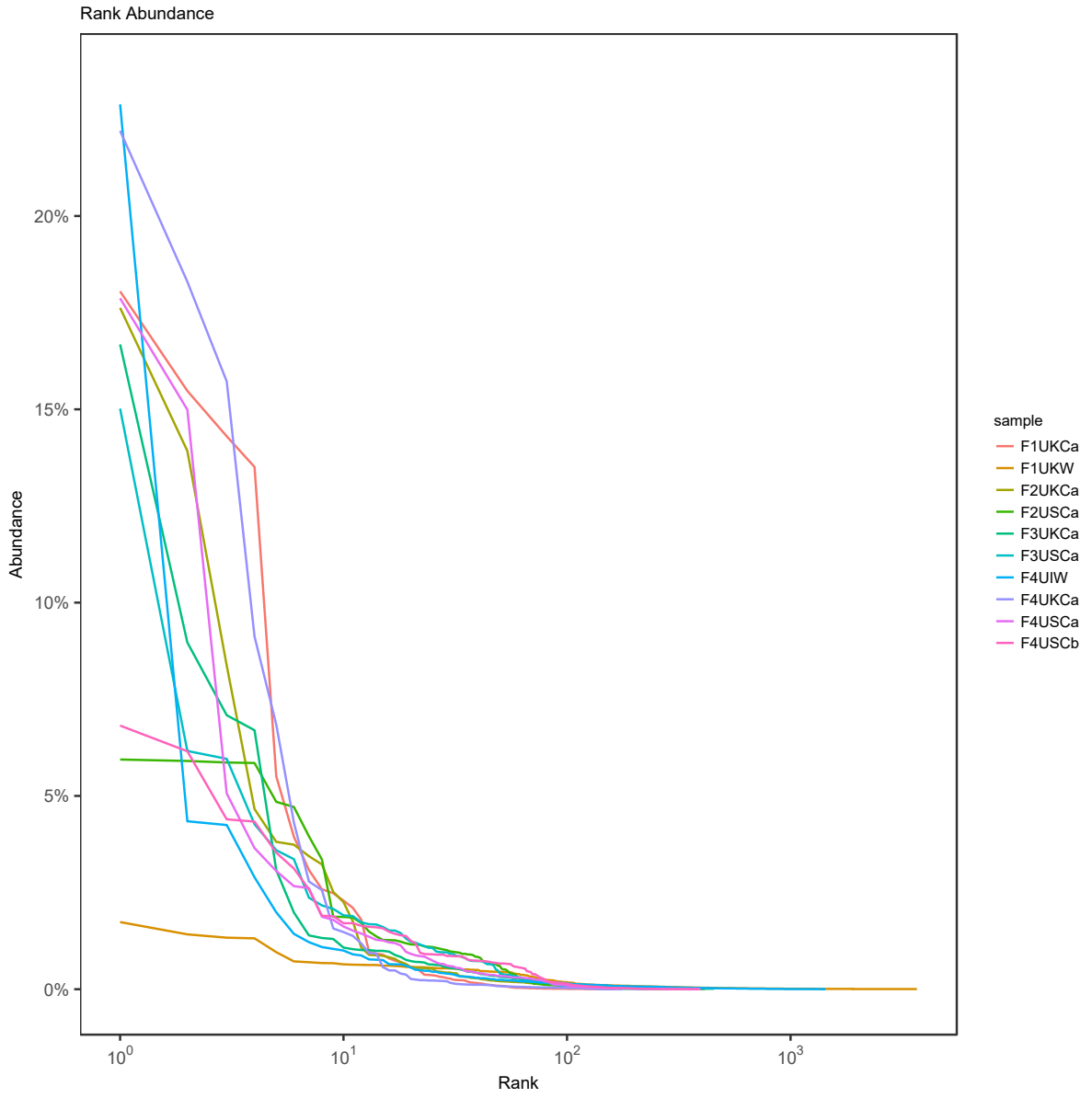


Figure 3.17. Clonal abundance curve using Ig repertoires isolated from unvaccinated zebrafish MMΦCs and the whole kidney and intestine. Abundance is the clone size as a percent of the repertoire; clones abundances are plotted in decreasing order from the most abundant on the left to the least abundant clones on the right. Each line represents an individual repertoire; the shaded areas are 95% confidence intervals. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. The analysis was performed using estimateAbundance subcommand in the Alakazam package.

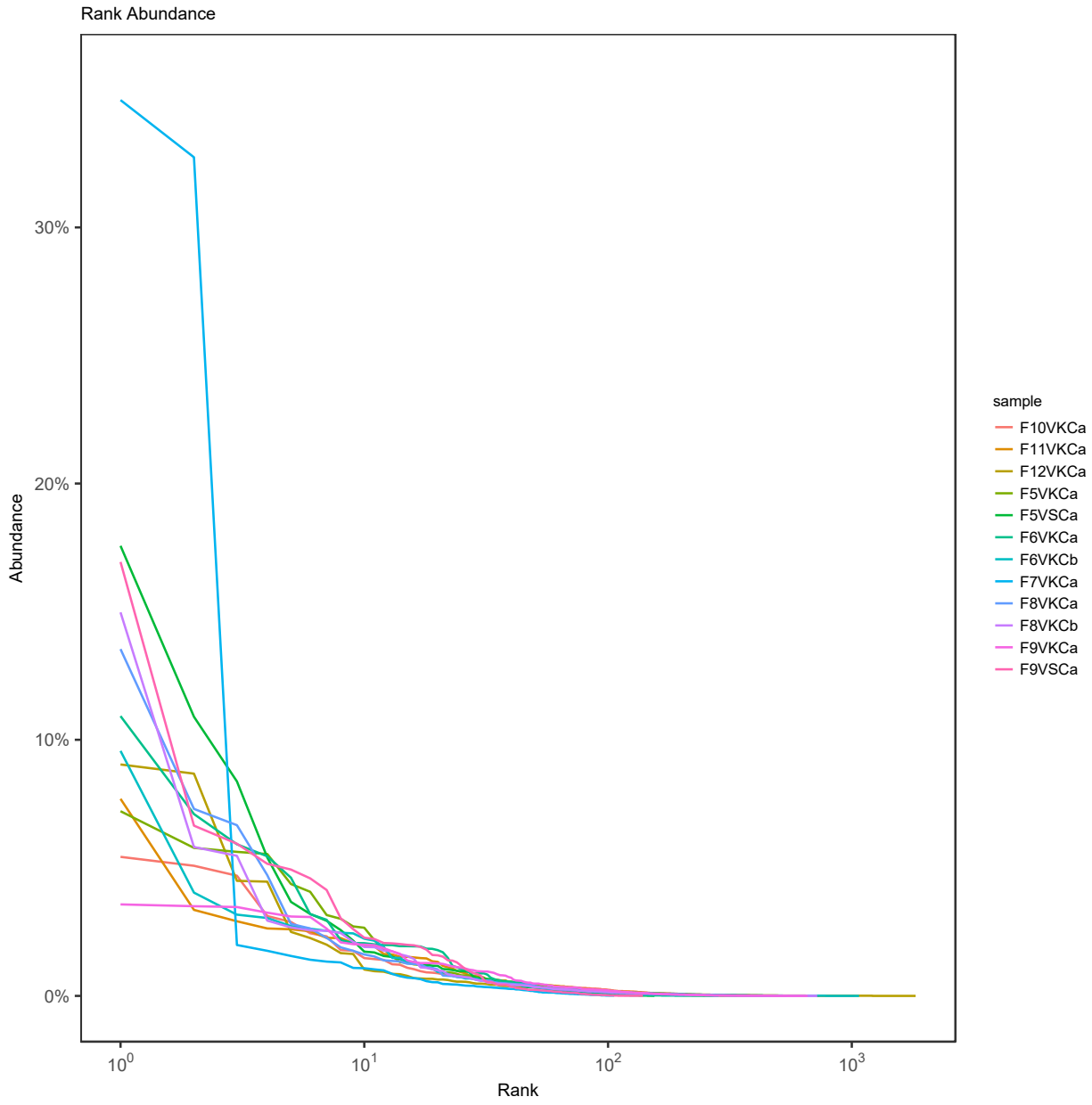


Figure 3.18. Clonal abundance curve using Ig repertoires isolated from vaccinated zebrafish MMΦCs. Abundance is the clone size as a percent of the repertoire; clones abundances are plotted in decreasing order from the most abundant on the left to the least abundant clones on the right. Each line represents an individual repertoire; the shaded areas are 95% confidence intervals. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. The analysis was performed using estimateAbundance subcommand in the Alakazam package.

3.6. IgM and IgZ junction (CDR3) region analysis

3.6.1. Non-templated nucleotides (N- nucleotides) and palindromic nucleotides (P-nucleotides)

Analysis of the N- and P-nucleotides additions using productive Ig sequences isolated from unvaccinated and vaccinated fish MMΦCs revealed that overall IgZ repertoires accumulated more N- and P-nucleotides than IgM repertoires (Table 3.14, Table 3.15, Table 3.16, Table 3.17). Furthermore, the percentage of Ig sequences in the IgM repertoires from both unvaccinated and vaccinated fish MMΦCs that had N- and P-nucleotides additions at the VD and DJ junctions was similar (Table 3.14, Table 3.15). It is necessary to notice that all the productive, unique Ig sequences within a repertoire were used to calculate the percentage of N- and P-nucleotides additions; therefore, these numbers are biased by clone size (Table 3.14, Table 3.15, Table 3.16, Table 3.17).

In the IgM repertoires from unvaccinated fish MMΦCs, on average, 57.8% and 58.4% of the productive sequences had N- and P-nucleotides additions between V and D and D and J segments, respectively (Table 3.14). In addition, 63.9% and 53.5% of the functional IgM sequences from vaccinated fish MMΦCs had N- and P-nucleotides insertions at the VD and DJ junctions, respectively (Table 3.15).

On the other hand, 70.9% and 61.6% of the Ig sequences of IgZ isotype from unvaccinated fish MMΦCs contained N- and P-nucleotides additions between VD, and DJ segments, respectively

(Table 3.16). In the IgZ repertoires from vaccinated fish MMΦCs, N- and P-nucleotides additions were found in 71% and 81.4% of the productive Ig sequences at the VD and DJ junctions, respectively (Table 3.17).

3.6.2. CDR3 length

Examination of the CDR3 length using Ig repertoires from unvaccinated and vaccinated fish MMΦCs showed that, on average, IgZ repertoires from unvaccinated fish MMΦCs had a longer CDR3 compared to IgM repertoires from unvaccinated and vaccinated fish MMΦCs (Fig.3.19 A and B, Fig.3.20 A). CDR3 length was determined based on the number of amino acids between the first and last conserved residues.

The CDR3 length in all the IgM repertoires from MMΦCs isolated from vaccinated fish was between 10 and 11 aa (Fig.3.20 A). Similarly, except for one repertoire, CDR3 length ranged from 10 to 11 aa in the IgM repertoires from unvaccinated fish MMΦCs; the CDR3 length was between 11 and 12 aa in one of these repertoires (Fig.3.19 A). The CDR3 length for the IgZ repertoires from unvaccinated fish MMΦCs was between 10 and 15 aa (Fig.3.19 B). For the IgZ repertoires from vaccinated fish MMΦCs, CDR3 length ranged from 9 to 12 aa (Fig.3.20 B).

Table 3.14. The percentage of N- and P-nucleotide additions in the IgM repertoires isolated from MMΦCs from unvaccinated zebrafish.

Fish[1]	NP1 %[2]	NP2 %[2]
F1UKCa	40.50	87.10
F2UKCa	53.60	67.30
F2USCa	70.60	57.60
F3UKCa	63.00	50.40
F3USCa	53.20	52.20
F4UKCa	41.70	41.80
F4USCa	71.80	55.79
F4USCb	68.00	55.60
Average	**57.80**	**58.47**
SD	**11.66**	**12.73**

[1]F# - Fish # in group; U - Unvaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.
[2]NP1 - Nucleotides between V and D segments; NP2 - Nucleotides between D and J segments.

Table 3.15. The percentage of N- and P-nucleotide additions in the IgM repertoires isolated from MMΦCs from vaccinated zebrafish.

Fish[1]	NP1 %[2]	NP2 %[2]
F5VKCa	65.20	60.00
F5VSCa	47.60	46.20
F6VKCa	72.40	44.00
F6VKCb	69.90	47.70
F7VKCa	81.20	78.40
F8VKCa	66.60	42.50
F8VKCb	66.20	56.10
F9VKCa	59.90	51.30
F9VSCa	52.30	52.20
F10VKCa	65.10	52.90
F11VKCa	55.60	59.90
F12VKCa	65.70	50.70
Average	**63.98**	**53.49**
SD	**8.72**	**9.23**

[1]F# - Fish # in group; V - Vaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.
[2]NP1 - Nucleotides between V and D segments; NP2 - Nucleotides between D and J segments.

Table 3.16. The percentage of N- and P-nucleotide additions in the IgZ repertoires isolated from MMΦCs from unvaccinated zebrafish.

Fish[1]	NP1 %[2]	NP2 %[2]
F1UKCa	71.10	70.50
F2UKCa	75.68	81.10
F2USCa	64.40	49.70
F3UKCa	86.20	49.50
F3USCa	62.70	60.60
F4UKCa	74.90	63.60
F4USCa	54.00	44.50
F4USCb	78.50	73.40
Average	**70.94**	**61.61**
SD	**9.53**	**12.17**

[1]F# - Fish # in group; U - Unvaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.
[2]NP1 - Nucleotides between V and D segments; NP2 - Nucleotides between D and J segments.

Table 3.17. The percentage of N- and P-nucleotide additions in the IgZ repertoires isolated from MMΦCs from vaccinated zebrafish.

Fish[1]	NP1 %[2]	NP2 %[2]
F5VKCa	82.50	63.20
F5VSCa	36.90	74.80
F6VKCb	68.10	54.70
F7VKCa	98.30	100.00
F8VKCa	96.70	91.80
F8VKCb	96.40	88.50
F9VKCa	51.69	63.70
F9VSCa	74.70	100.00
F10VKCa	98.80	98.80
F11VKCa	3.50	100.00
F12VKCa	74.50	60.00
Average	**71.10**	**81.41**
SD	**28.73**	**17.47**

[1]F# - Fish # in group; V - Vaccinated; cluster from K/S - Kidney/Spleen and cluster # - a/b.
[2]NP1 - Nucleotides between V and D segments; NP2 - Nucleotides between D and J segments.

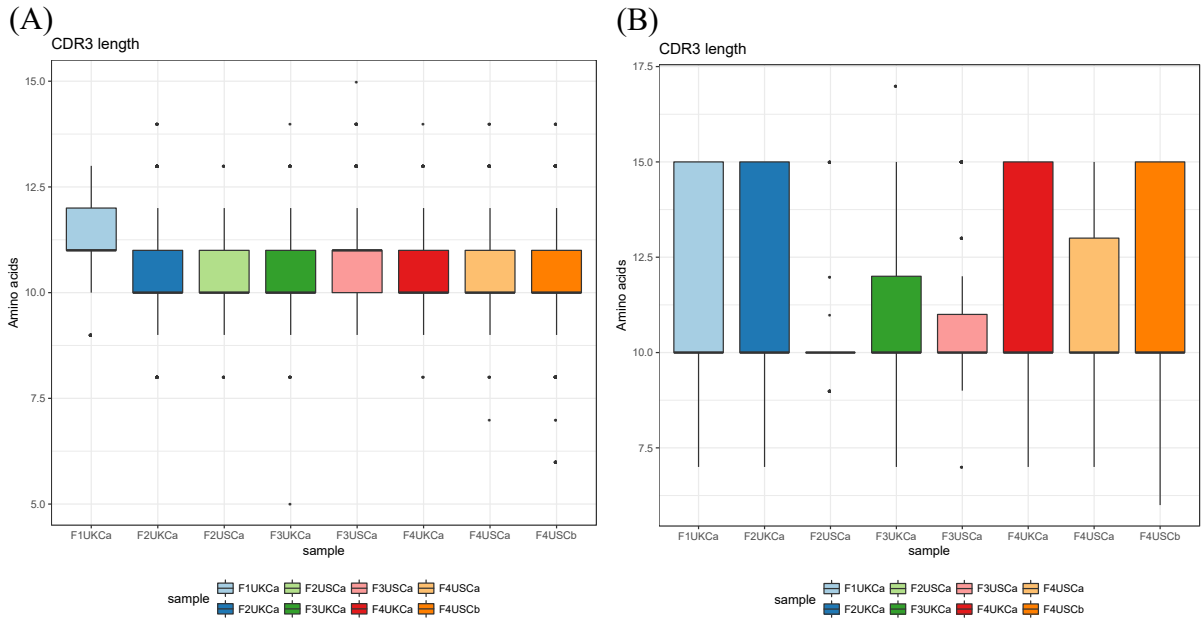


Figure 3.19. CDR3 length using Ig repertoires isolated from unvaccinated zebrafish MMΦCs. (A) IgM repertoires. (B) IgZ repertoires. Boxes show the distribution of the CDR3 length, the whiskers extend to the most extreme values, and the dots show the outliers. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. The analysis was performed using aminoAcidProperties subcommand in the Alakazam package.

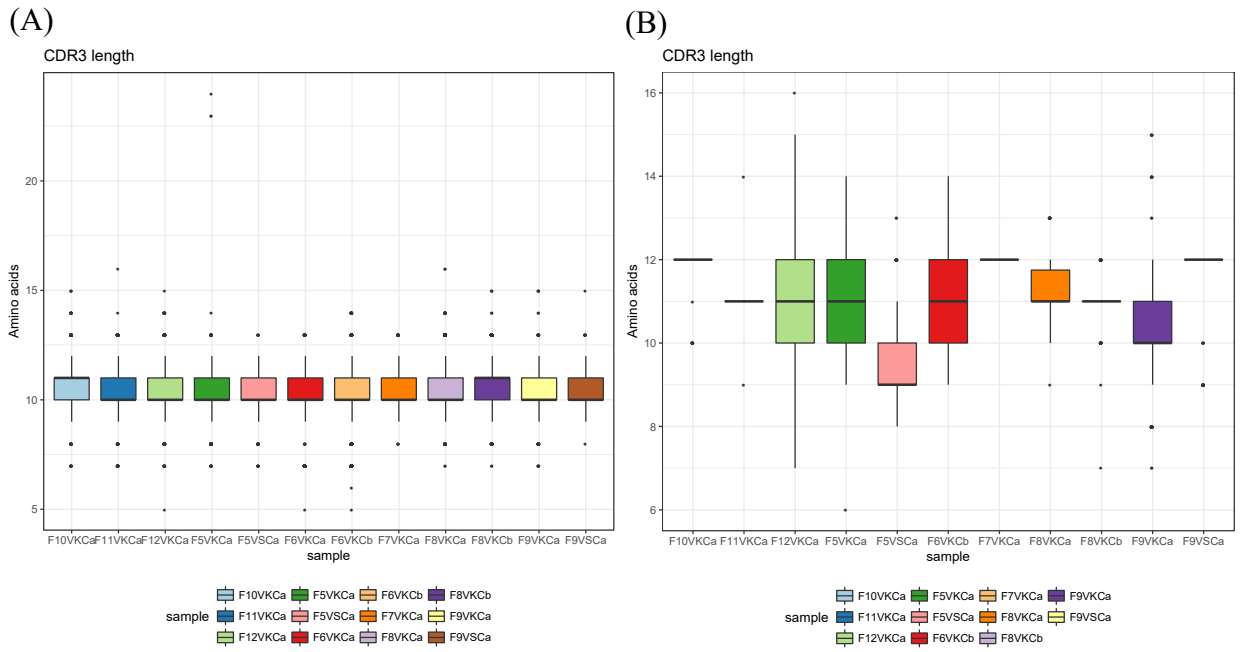


Figure 3.20. CDR3 length using Ig repertoires isolated from vaccinated zebrafish MMΦCs. (A) IgM repertoires. (B) IgZ repertoires. Boxes show the distribution of the CDR3 length, the whiskers extend to the most extreme values, and the dots show the outliers. Refer to table 2.1 and table 3.1 for a detailed description of the samples used. The analysis was performed using aminoAcidProperties subcommand in the Alakazam package.

3.7. Long-term Ag retention occurs within MMΦCs isolated from vaccinated zebrafish

To determine if long term Ag retention occurs within MMΦCs in zebrafish, I injected fish with antigen conjugated to a fluorescent label.

Using isolated MMΦCs from zebrafish vaccinated with KLH or BSA conjugated to Alexa-647, microscopy results revealed that long-term Ag retention occurs within MMΦCs in the spleen and kidney one month after the initial vaccination (Fig.3.21, Fig.3.22). The autofluorescence emission spectra of MMΦ pigments in the 500–580 nm wavelength range (Saunders et al., 2010) make it distinguishable from Alexa-647 absorbance and emission spectra at 650 and 665 nm, respectively.

Fig.3.21 shows MMΦC isolated from the kidney from zebrafish vaccinated with KLH -Alexa-647; after one month, the injected Ag can be detected at 633 nm wavelength (Cy5 laser) and the merged image. Similarly, the injected Ag (BSA -Alexa-647) is found within MMΦC isolated from the spleen from zebrafish one month after the initial vaccination (Fig.3.22).

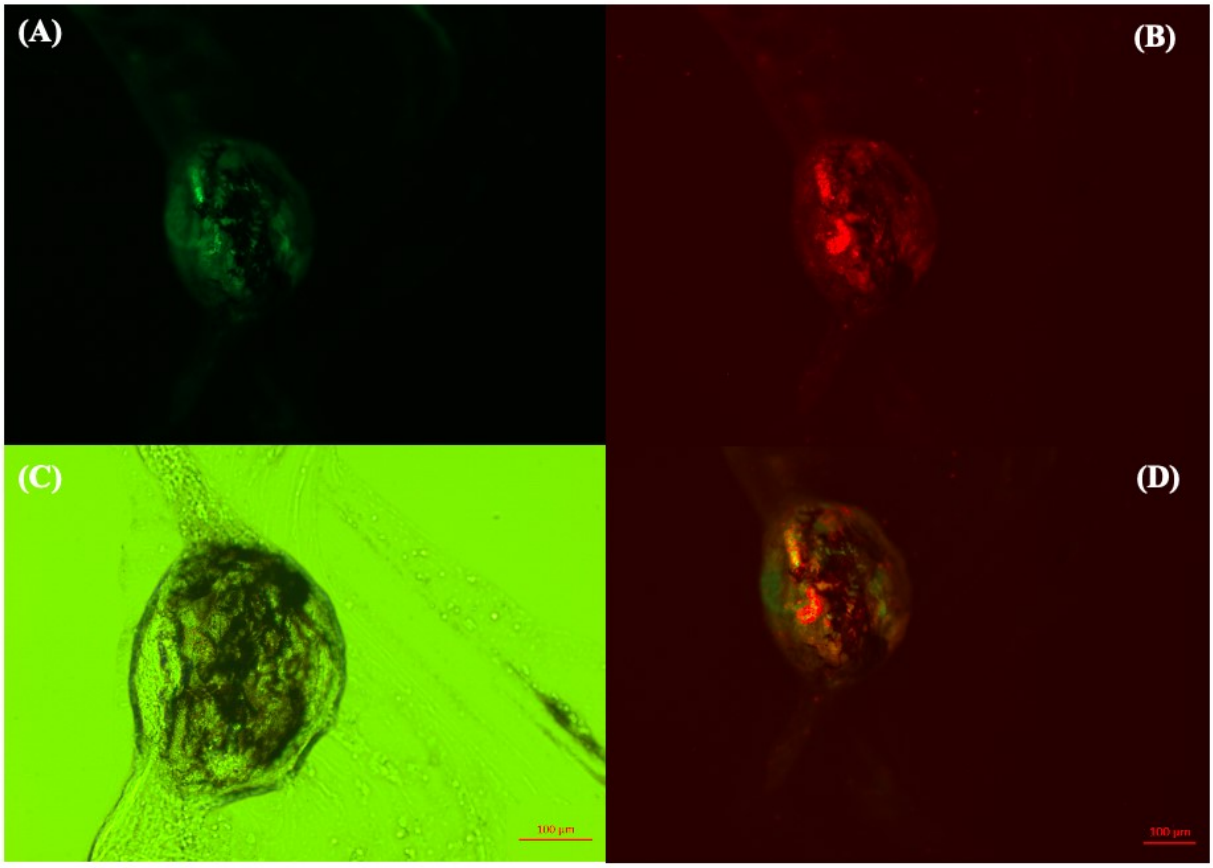


Figure 3.21. Antigen retention within MMΦC isolated from zebrafish kidney. Kidney MMΦC isolated from zebrafish immunized with keyhole limpet haemocyanin (KLH) -Alexa-647 one month following the vaccination. (A) Autofluorescent MMΦs observed in the FITC channel. (B) Alexa-647 in the Cy5 channel. (C) Bright-field image. (D) Merged image. MMΦC was imaged using Zeiss Axio Imager M2 fluorescent microscope.

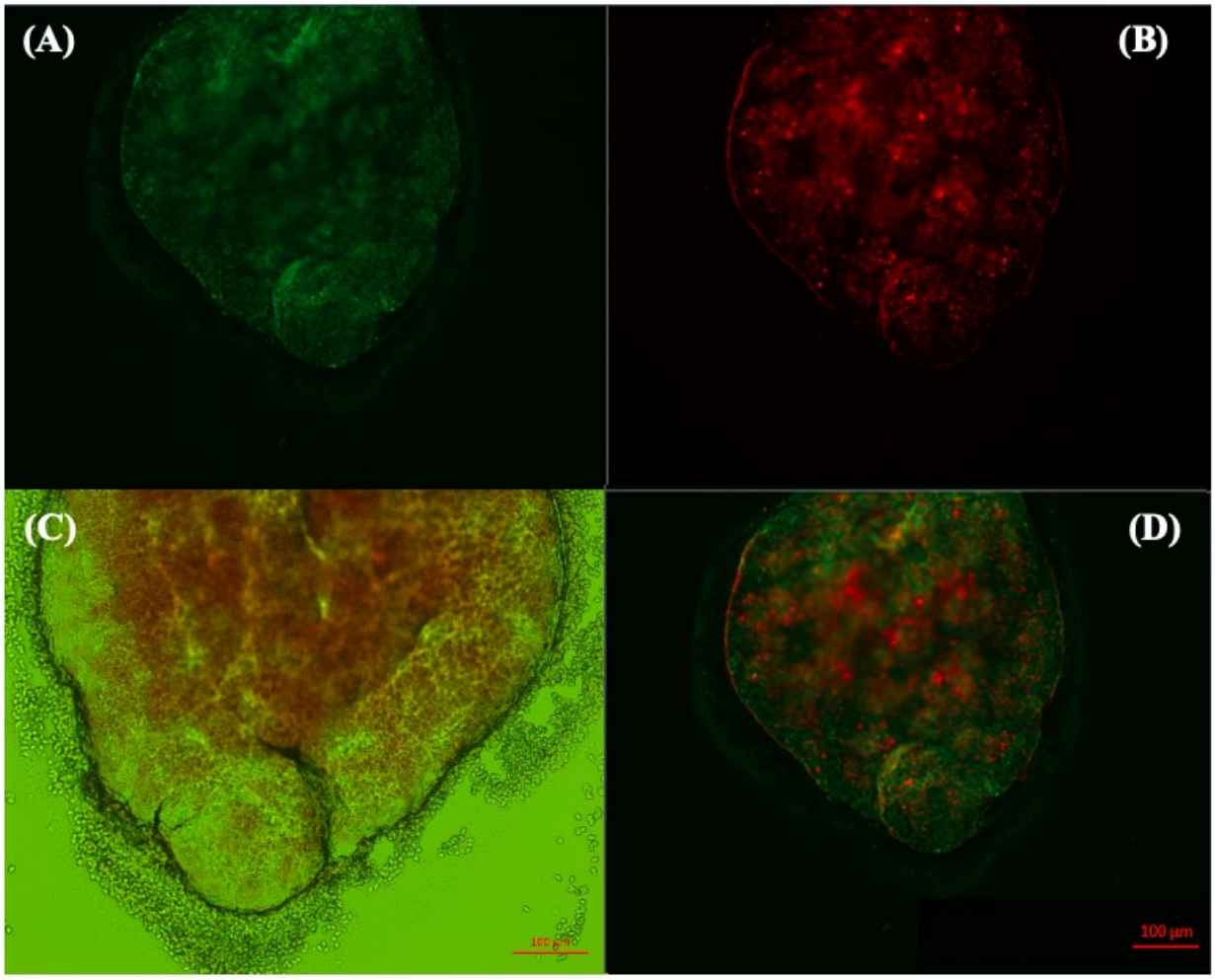


Figure 3.22. Antigen retention within MMΦC isolated from zebrafish spleen. Spleen MMΦC isolated from zebrafish immunized with bovine serum albumin (BSA) -Alexa-647, one month following the vaccination. (A) Autofluorescent MMΦs observed in the FITC channel. (B) Alexa-647 in the Cy5 channel. (C) Bright-field image. (D) Merged image. MMΦC was imaged with Zeiss Axio Imager M2 fluorescent microscope.

3.8. lymphocyte-like cells within MMΦCs

3.8.1. Total number of cells within MMΦCs using manual cell count

Using intact, isolated zebrafish MMΦCs from the spleen and kidney and manual counting on a hemocytometer, I estimated the total number of cells in a cluster to be 41,420 cells/cluster. Four clusters from the spleen and four clusters from the kidney were isolated, and after dissociating the cells, the total number of cells was found to be 33,040 cells/cluster and 49,800 cells/cluster for the spleen and kidney, respectively. Estimation of the total number of cells using manual counting was performed to ensure that imaging flow cytometry analysis will not underestimate the number of cells as a result of damaging too many cells.

3.8.2. Total number of cells and lymphocyte-like cells within MMΦCs using imaging flow cytometry (ImageStream)

Intact MMΦCs isolated from zebrafish spleen and kidney were used to verify the total number of cells and the presence of lymphocyte-like cells within these clusters. As shown in figure 3.23 and figure 3.24, the diameter of the analyzed clusters varied and ranged from ≈200 to 350 μm (Figure.3.23, Figure.3.24). Isolated clusters were first passed through a cell strainer to dissociate the cells then the nuclei of the dissociated cells were stained. Subsequently, the samples were run through an imaging flow cytometry (ImageStream), and data were analyzed using IDEAS 6.1.822.0 software (Luminex). Initially, I kept only the cells in focus; cells out of focus were discarded (R1, Figure 3.25 A). Then, single cells were selected; doublets, debris, and aggregates

were removed (R2, Figure 3.25 B). After selecting focused cells and removing doublets and aggregates, I found the total number of cells in MMΦCs isolated from the spleen and kidney to be 33,252 cells/cluster and 37,091 cells/cluster, respectively (Table 3.18). Two separate imaging flow cytometry runs were performed using eight MMΦCs from zebrafish spleen and eight clusters from the kidney in the first run and eight MMΦCs from zebrafish spleen and seven clusters from the kidney in the second run. In the first run, the total number of cells was 33,231 cells/cluster and 35,062 cells/cluster in MMΦCs isolated from the spleen and kidney, respectively. A total of 33,273 cells and 39,120 cells were found in individual MMΦCs from the spleen and kidney, respectively, during the second run of the imaging flow cytometry (Table 3.18).

Subsequently, the number of lymphocyte-like cells in individual clusters was determined based on the cell size and internal complexity. Previously, leukocyte populations from whole kidney cell suspension were described in zebrafish using flow cytometry, and they were able to identify three populations, monocytes/macrophage and granulocytes, hematopoietic precursors, and lymphocyte-like cells (Petrie-Hanson et al., 2009). Lymphocyte-like cells have low internal complexity, and they are small; however, they are larger in size than erythroid cells (Petrie-Hanson et al., 2009, Traver et al., 2003). MMΦ is the primary cell type within MMΦCs, and it has been shown that MMΦs vary widely in their size and internal complexity; nonetheless, MMΦs have higher internal complexity and are larger in size than lymphocyte-like cells (Diaz-Satizabal et al., 2015). Therefore, using imaging flow cytometry (ImageStream), I was able to identify lymphocyte-like cells which have a small size and low internal complexity (R3, Figure 3.25 C, representative images of lymphocyte-like cells are shown in Figure.3.25).

I found 5305 lymphocyte-like cells in individual MMΦCs isolated from the spleen and 6027 lymphocyte-like cells/cluster in MMΦCs isolated from the kidney (Table 3.19). A total of 4,293 and 6,473 lymphocyte-like cells were found in individual MMΦCs isolated from the spleen and kidney, respectively, during the first imaging flow cytometry run. In the second run, 6,318 lymphocyte-like cells/cluster and 5,581 lymphocyte-like cells/cluster were found in MMΦCs isolated from the spleen and kidney, respectively (Table 3.19). In addition, the diameter of the analyzed lymphocyte-like cells isolated from MMΦCs varied and ranged from 5 to 7 μm (Figure.3.25).

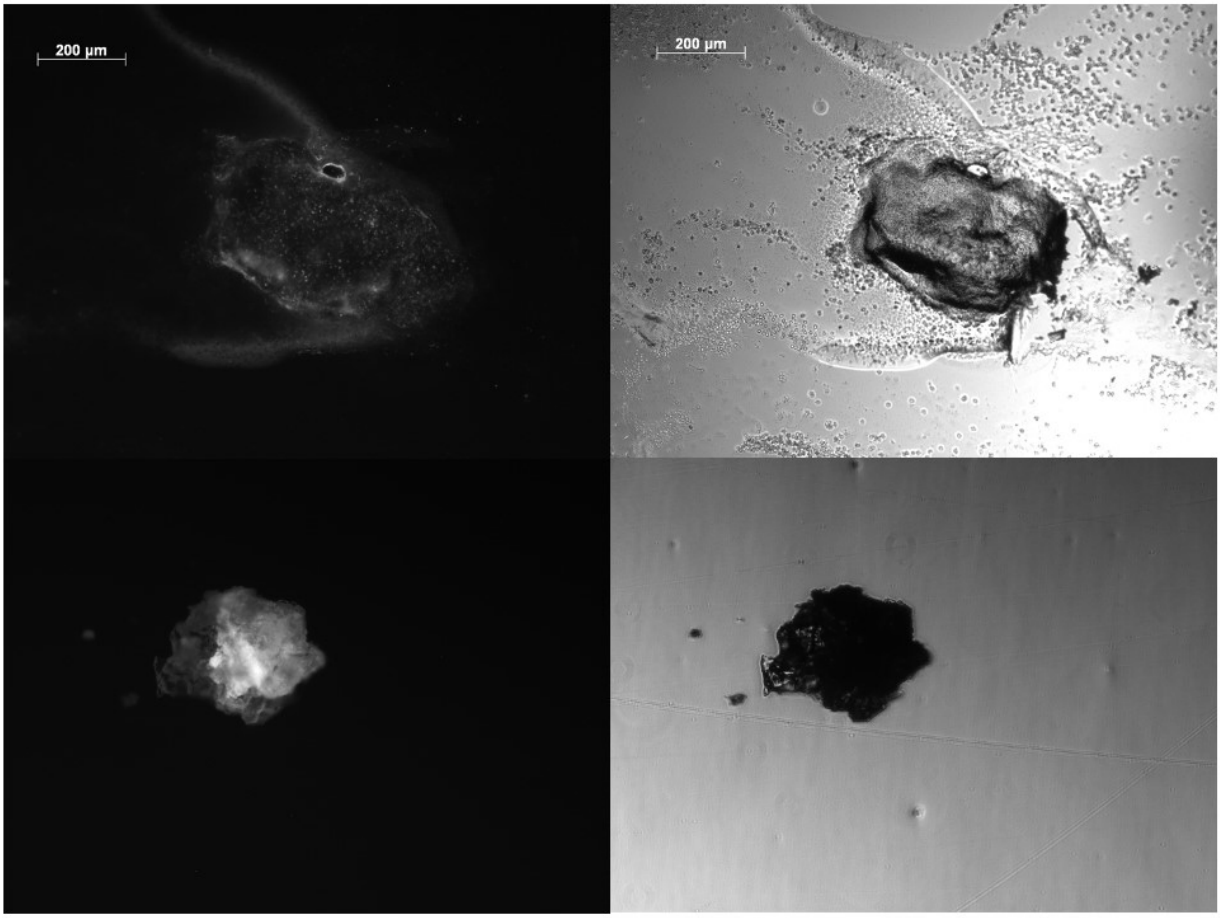


Figure 3.23. Representative images of MMΦCs isolated from zebrafish spleen and used for imaging flow cytometry analysis. Left column autofluorescent MMΦs in the FITC channel, and on the right column are the bright-field images. Scale bar = 200 µm.

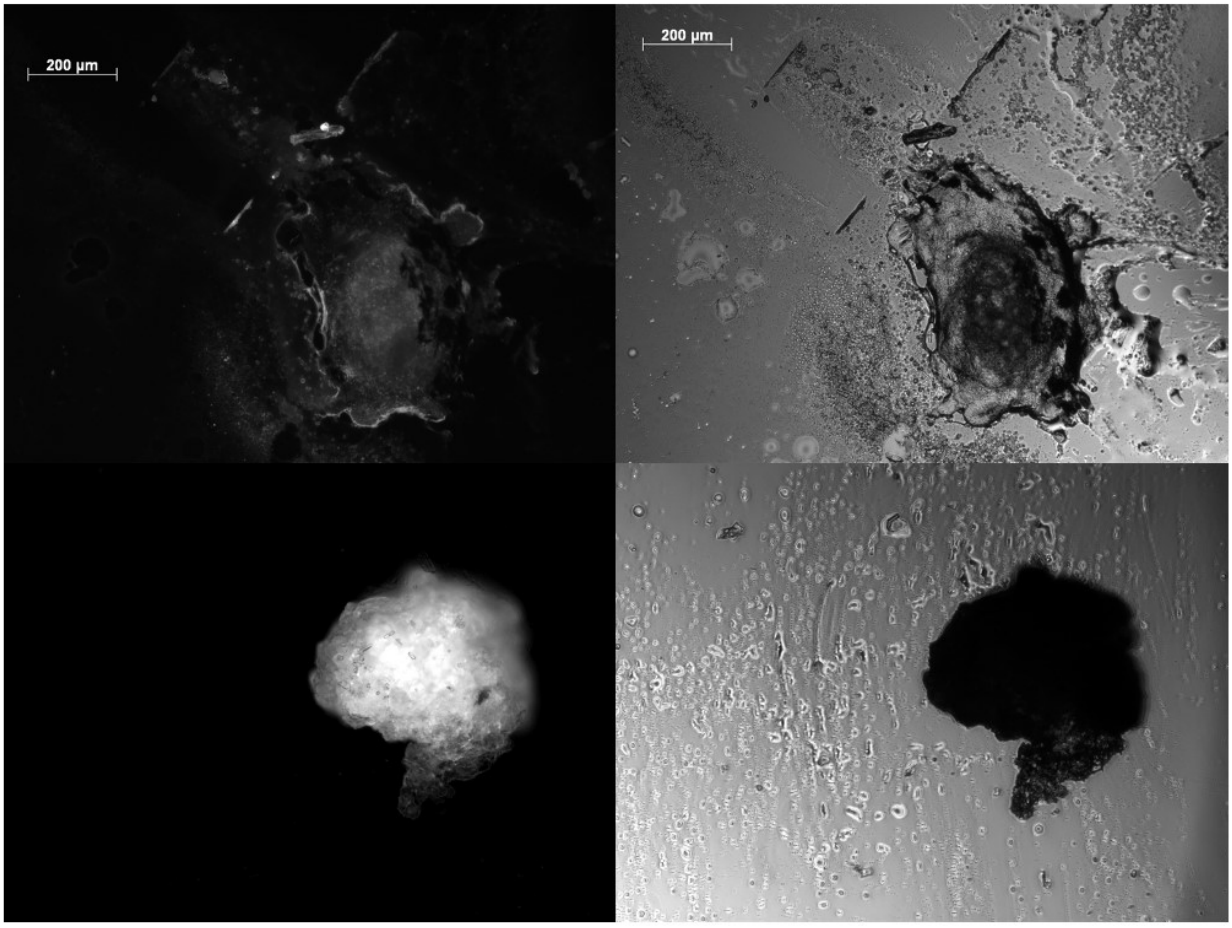


Figure 3.24. Representative images of MMΦCs isolated from zebrafish kidney and used for imaging flow cytometry analysis. Left column autofluorescent MMΦs in the FITC channel, and on the right column are the bright-field images. Scale bar = 200 µm.

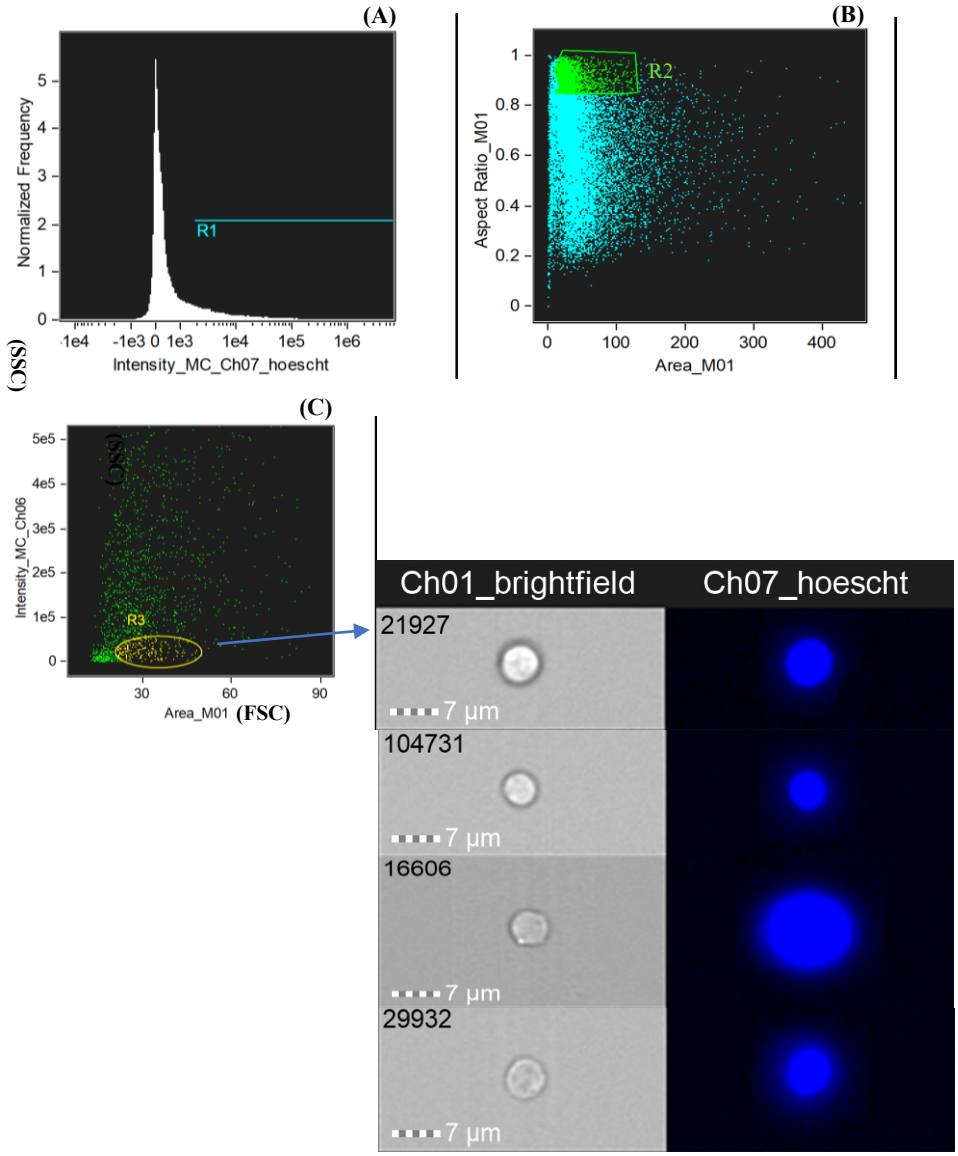


Figure 3.25. Workflow of the analysis to identify lymphocyte-like cells using imaging flow cytometry and cells isolated from MMΦCs from zebrafish kidney. (A) R1, Selected focused cells. (B) R2, Focused single cells, aspect ratio is calculated based on the shape of the detected object; circular cells have aspect ratio values close to 1.0. (C) R3, lymphocyte-like cells; area (cell size) vs. intensity (cell internal complexity) dot plot, and representative images of lymphocyte-like cells. Cell images were captured with 40 × objective, Scale bar = 7µm. FSC, forward scatter; SSC, side scatter. Analysis was done using IDEAS 6.1.822.0 software.

Table 3.18. Estimated total number of cells isolated from individual MMΦCs (R2, Figure 3.25).

	kidney MMΦCs	spleen MMΦCs
Imagestream 1st run	35062	33231
Imagestream 2nd run	39120	33273
Average	37091	33252
SD	2029	21

Table 3.19. Estimated total number of lymphocyte-like cells isolated from individual MMΦCs (R3, Figure 3.25).

	Kidney MMΦCs	Spleen MMΦCs
Imagestream 1st run	6473	4293
Imagestream 2nd run	5581	6318
Average	6027	5305.5
SD	446	1012.5

3.9. IgM constant region

Using the first Cμ exon (Cμ1), 260 nucleotides were sequenced. Analysis of the sequenced Cμ1 region revealed that point mutations, insertion, and deletion mutations accumulate within this region. Interestingly, many insertion mutations were observed in this region. However, the primers that were used to amplify this region were not verified, and thus, although we were able to sequence the Cμ1 region, the majority of the sequences were from genes other than our target gene. Therefore, I was not able to perform further analysis on the observed mutations in the Cμ1 region.

3.10. V-less IgM and IgZ transcripts

Using Ig repertoires from unvaccinated and vaccinated fish MMΦCs, several Ig (IgM and IgZ) sequences that lack the VH region were found. In these sequences, the leader sequence is spliced to the first exon of the C-μ or C-ζ (Fig.3.26 A and B).

(A)

TGCATTTCAGTTCTGCTGCTCTTGTTAAGTCTACAGCCTCAACCATCTGCGCCCCAGTCAGT
CTTCGGTTTGTCTCAGTGCA

(B)

GGTTTATACTGTCAAGGCATGGAATATACACTGGACGCAGGTGACCGTGACAAATGAAACT
CTCACAGCACCAGTTGTGTTCAAAATGTCTCAGTGTAGTTCTTCTACTGATTCCTTAATTATT
GGATGTCTGGCCTCTGA

Figure 3.26. Examples of the Ig sequences which lack the VH region. The leader sequence is spliced to the first exon of the C-μ (A) or C-ζ (B); the primer sequences for the leader sequences and the constant regions are underlined. Shaded nucleotides match the constant regions of (A) IgM (B) IgZ.

CHAPTER IV

DISCUSSION

4. Discussion

4.1. Overview

In homeotherms, during a humoral immune response to a T-cell dependent antigen, following the interaction with antigen-specific T-cell, a few activated B-cells migrate back to the B-cell follicle within the secondary lymphoid organs, where they form transient structures known as germinal centers (GCs) (reviewed in Stebegg et al., 2018). These activated B-cells rapidly divide and undergo clonal expansion within the follicle. During this process, activated B-cells displace naïve B-cells, which form the mantle zone around the GC. By day 7 post-immunization, the GC can be divided into two distinct compartments known as the dark zone and the light zone (reviewed in De Silva et al., 2015). Large, proliferating B cells, now referred to as centroblasts, downregulate Ab production and upregulate AID expression within the dark zone. AID expressed in the centroblasts causes somatic hypermutations (SHM) in the VDJ exon to alter (and potentially enhance) their antibodies' binding affinity to the exogenous antigen. Modified B-cells, now known as centrocytes, move to the light zone. Centrocytes are small, nondividing B-cells expressing membrane antibodies. Within the light zone, centrocytes compete for a limited number of antigens trapped on the surface of FDCs; centrocytes with high affinity for the antigen endocytose it and present it via class II MHC to Tfh cells, from which they receive survival signals based on their affinity to the

exogenous antigen. Subsequently, selected B-cells can return to the dark zone for more SHM and proliferation or undergo a phase of proliferation, and daughter cells will differentiate into antibody-secreting plasma cells and memory B-cells (Ise et al., 2018, reviewed in De Silva et al., 2015, Stebegg et al., 2018).

Cold-blooded vertebrates lack GCs; however, they have a fully functional Ig mutator enzyme (AID) (reviewed in Magor, 2015). In bony fish, AID is expressed within melano-macrophage clusters (MMΦCs), which are aggregates of highly pigmented melano-macrophages in the spleen and kidney (Saunders et al., 2010). Within MMΦCs, melano-macrophages (MMΦs) are extensively encapsulated by reticular cells, and at least in cyprinids (goldfish and zebrafish), intact clusters can be isolated from surrounding tissues (Diaz-Satizabal et al., 2015). These clusters are also associated with IgM^+ B-cells and $CD4^+$ T-cells (Saunders et al., 2010); in addition, studies have shown that antigen is retained within MMΦCs for a long period (Lamers et al., 1985, Lamers CH, De Haas MJ 1985, Ziegenfuss et al., 1991). These observations led to the hypothesis that MMΦCs in fish are functionally analogous to GCs in mammals and birds.

Hence, the current study sought to directly test the hypothesis that MMΦCs are acting as primordial GCs in fish by examining the hallmarks of the GC reaction within MMΦCs, including

1. Clonal expansion of activated B-cells.
2. Accumulation of AID and error-prone polymerases induced mutations in the VDJ exon of the Ig gene of activated B-cells.
3. The presence of an active antigen-driven selection process.

4.2. Proliferation and clonal expansion of B-cell clonotypes within MMΦCs

The results shown in table 3.2 and table 3.3 revealed that B-cell clonotypes clonally expand within MMΦCs, with some of the clonotypes having more than 400 daughter cells. While the size of the clones varies among the different clusters, generally, it appears that each cluster is dominated by a few clonally expanding B-cell clonotypes. Despite having similar VDJ combination coverage to the repertoires from the clusters, the clones isolated from the whole kidney were smaller than the clones from the clusters. Even though the size of the largest clone from the whole kidney is comparable to the size of some clones in a few repertoires isolated from MMΦCs, multiple clones with the size of the largest clone from the whole kidney Ig repertoire were found in the repertoires from the clusters. On the other hand, B-cell clonotypes isolated from the whole kidney generally had a smaller size than the clones found in the clusters. In addition, lineage trees showed that B-cell clonotypes clonally expand within MMΦCs while acquiring mutations in their Ig VDJ sequences (Fig.3.2, Fig.3.3, Fig.3.4). These observations indicate that MMΦCs in fish are the sites where B-cell clonotypes proliferate and accumulate mutations in their variable region, which is consistent with the hypothesis that MMΦCs in fish are functionally analogous to GCs. Within the dark zone of the GC, B-cells (centroblasts) extensively proliferate while acquiring AID and error-prone polymerases mediated SHM (reviewed in De Silva et al., 2015, Mesin et al., 2016, Verstegen et al., 2021).

In agreement with the variability in the total number of clones and the size of the dominant clones among MMΦCs, it has been shown that following immunization, the degree of clonal dominance varies among GCs (Tas et al., 2016); in addition, they demonstrated that affinity maturation occurs

within GCs that have low clonal dominance (Tas et al., 2016). Moreover, a study using complex antigens (*Bacillus anthracis* protective antigen and influenza hemagglutinin) revealed that antibody affinity maturation process was associated with increased clonal diversity (Kuraoka et al., 2016).

Interestingly, clones in the Ig repertoires isolated from MMΦCs from unvaccinated fish were, in general, larger than the clones from MMΦCs from vaccinated fish. In addition, the number of clones for both IgM and IgZ isotypes was higher in the Ig repertoires isolated from MMΦCs from vaccinated fish compared to the number of clones in the clusters from unvaccinated zebrafish. While we dissected the vaccinated fish at specific timepoints following the vaccination, B-cell clonotypes could have been proliferating for a longer period of time within MMΦCs from unvaccinated fish. Nevertheless, it has been demonstrated that B-cells rapid clonal expansion (referred to as clonal bursts), which leads to a significant loss of clonal diversity, occurs at different rates within different GCs (Tas et al., 2016). Consistent with this, the size and number of clones varied among MMΦCs isolated from vaccinated fish that were dissected at the same timepoint (Table 2.1, Table 3.1, Table 3.3); similarly, variability in the size and number of the clones was also found in MMΦCs isolated from unvaccinated zebrafish (Table 3.1, Table 3.2).

The findings that the size of the dominant clones varied among some of the clusters that were isolated from different tissues from the same fish and between MMΦCs isolated from the same tissue from the same fish is consistent with the heterogeneity in the clonal dominance between GCs, even when they were isolated from the same lymph node (Tas et al., 2016).

Recently, computational models estimated the total number of B-cells within a GC to be around 2000 cells at the peak of the GC reaction (Pélissier et al., 2020), which is similar to the total number of unique Ig sequencing reads found within MMΦCs (Table 3.1). Nonetheless, an earlier study using stained human lymph node frozen sections estimated the total number of cells in a GC to be between 12,000 and 14,000 cells, and around 5,000 B-cells are estimated to be within a GC (Kuppers et al., 1993). On average, Ig repertoires isolated from unvaccinated and vaccinated fish MMΦCs had 2058 total unique VDJ sequencing reads; however, some repertoires had as many as 3660 and as low as 510 unique VDJ sequencing reads. Of note, the size of MMΦCs in fish varies and increases in repose to infection or immunization, similar to mammalian GCs (Herráez et al., 1986, Agius, 1979, Secombes et al., 1982, reviewed in Steinel and Bolnick, 2017). For example, in Nile tilapia, the size of MMΦCs significantly increased in response to bacterial infections (Manrique et al., 2019). Similarly, MMΦCs size increased in southern bluefin tuna following treatment with *Schistosoma* (a parasitic infection; Nowak et al., 2021). These studies indicate that MMΦCs in fish participate in the immune response against pathogens, and perhaps the increase in MMΦCs size in response to infection is associated with an increase in the number of B-cells within a cluster.

In addition to the dominant clones, smaller clones were also found within MMΦCs from the spleen and kidney, suggesting that the clusters are continually evolving structures and develop toward the dominance of a few clones and their progeny (Table 3.2, Table 3.3). Similarly, it has been shown that subclones are found within GCs at different timepoints (Tas et al., 2016). Moreover, using two-photon laser-scanning microscopy, it has been reported that GCs are open and dynamic structures, and recently activated follicular B-cells have the ability to enter an ongoing GC

(Schwickert et al., 2007). Also, antigen-specific B-cells with high affinity for the exogenous antigen can be recruited to an ongoing GC (Schwickert et al., 2007). They suggested that the open structure of the GC improves the competition between high-affinity antigen-specific B-cells (Schwickert et al., 2007). More recently, it has been shown that clonal expansion of a specific B-cell clonotype in a GC is followed by the appearance of a similar B-cell clonotype in a neighboring GC, further supporting the dynamic, open structure of the GC (Firl et al., 2018). Therefore, the subclones in the Ig repertoires isolated from MMΦCs from unvaccinated and vaccinated zebrafish could indicate that these clusters are open structures and activated B-cells can enter and proliferate within the clusters at varying times.

During the GC formation, proliferating B-cells push and replace naïve follicular B-cells within the follicle, which leads to the formation of the mantle zone around the GCs (reviewed in De Silva et al., 2015). Similarly, we found several B-cell clonotypes with no mutations in their Ig sequences associated with isolated clusters (Fig.3.1); these clones could be naive B-cells within or surrounding MMΦCs.

Tissues immediately surrounding MMΦCs significantly had fewer total unique Ig sequencing reads than the clusters, and no B-cell clonotypes of IgZ isotype were found in these tissues. In addition, the clones isolated from the tissues surrounding MMΦCs were much smaller than the clones isolated from MMΦCs; only one of the clones from the surrounding tissues had a size comparable to the size of the clones isolated from MMΦCs (Table 3.2, Table 3.3, Table 3.4). These findings are consistent with MMΦCs being the sites where B-cells proliferate and acquire mutations before migrating to the surrounding or mucosal tissues. A recent study in rainbow trout

found a high number of IgT expressing cells in the spleen and kidney of immersion vaccinated fish (Hoare et al., 2017), and our results suggest that these IgT expressing B-cells are predominantly within MMΦCs. The mucosal antibody in bony fish is known as IgZ or IgT for zebrafish or rainbow trout/teleost, respectively (Danilova et al., 2005, Hansen et al., 2005).

Consistent with the previous suggestion that the posterior kidney is a secondary lymphoid organ in fish (Zwollo et al., 2005), the results from MMΦCs from the spleen and kidney are similar. Using transcription factors that are differentially expressed during B-cell activation, it has been shown that the kidney has a high number of partially activated B-cells and plasma cells (Zwollo et al., 2005). My results support the hypothesis that, similar to the spleen, the posterior kidney is a secondary lymphoid organ in fish; B-cell clonotypes proliferate while acquiring mutations within MMΦCs isolated from the spleen and kidney. B-cells activation and proliferation following antigens encounter occur predominantly within the secondary lymphoid organs (reviewed in Cyster et al., 2019). In addition, in mammals and birds, B-cells proliferate while acquiring mutations within the GC, which forms in the secondary lymphoid tissues (reviewed in Stebegg et al., 2018).

MMΦCs from vaccinated fish had mainly IgM^+ B-cell clonotypes; in contrast, most of MMΦCs from unvaccinated fish had IgZ as the predominant isotype. This finding indicates that the role of MMΦCs in the development of Ag-specific B-cells is not limited to IgM^+ B-cells but also involves IgZ^+ B-cells, which has been described as the primary antibody isotype in mucosal immunity in teleost (Zhang et al., 2010). While we induced a systemic immune response in the vaccinated fish, unvaccinated fish could have been responding to mucosal infection. A recent study demonstrated

that the production of IgZ in zebrafish is largely dependent on αβ/CD4^{+} T cells; depletion of T-cells in vivo significantly reduced IgZ production, indicating that IgZ production is T-cell-dependent (Ji et al., 2021). Previously it has been revealed that MMΦCs are associated with the transcripts for TCR β and CD4 (Saunders et al., 2010). Our results suggest that T cell-dependent IgZ proliferation occurs mainly within MMΦCs, where both IgM^{+} and IgZ^{+} B-cell clonotypes proliferate and acquire mutations.

In contrast to the size of the clones that were isolated from the whole kidney, several large clones of both IgM and IgZ isotypes, with up to 459 clone size, were found in the Ig repertoire isolated from the intestine. A single cluster was isolated from the kidney before preparing the whole kidney Ig repertoire; the size of the largest clone in the Ig repertoire from MMΦC isolated from the whole kidney was 404, which is much larger than the largest clone in the whole kidney Ig repertoire (Table 3.2, Table 3.4). These results indicate that mutations and clonal expansion of B-cell clonotypes occur somewhere in the intestine. Subsequent fluorescence microscopy analysis revealed that MMΦCs are found in the intestine (unpublished observation), in addition to the spleen and kidney in zebrafish. Consistent with our results, a study has shown that IgD^{+} IgM^{-}, IgD^{-} IgM^{+}, and IgT^{+} (IgD^{-}IgM^{-}) B-cells clonally expand within the intestine in rainbow trout (Perdiguero et al., 2019).

In summary, the data presented in this section showed that B-cell clonotypes clonally expand within MMΦCs while acquiring mutations in their Ig VDJ sequences. In addition, the presence of subclones in the Ig repertoires isolated from MMΦCs from unvaccinated and vaccinated fish indicates that these clusters are continually evolving structures and develop toward the dominance

of a few clones and their progeny. Moreover, it appears that MMΦCs are open structures, and activated B-cells can enter and proliferate within the clusters at varying times. Also, my results revealed that the role of MMΦCs in the development of Ag-specific B-cells is not limited to IgM expressing B-cells but also involves IgZ expressing B-cells. Collectively, these results support the hypothesis that MMΦCs in fish function as primordial germinal centers. Moreover, my data support the hypothesis that the posterior kidney is a secondary lymphoid organ in fish similar to the spleen. Finally, analysis of the Ig repertoire isolated from the intestine revealed that clonal expansion of B-cell clonotypes with concomitant somatic hypermutation occurs somewhere in the intestine.

4.3. Evidence for common V_H-elements usage in distinct repertoires

In zebrafish, IgH locus consists of several V_H segments followed by two D, two J, and C regions of IgZ (Dζ-Jζ-Cζ) between the V_H segments and five Dμδ, five Jμδ, Cμ, and Cδ gene segments, which encode both IgM and IgD (Danilova et al., 2005). All associated IgH VDJ genes were captured from whole kidney or intestine Ig libraries. Even though the whole kidney and intestine are from two different fish, their IgM repertoires had similar V, D, and J genes usage frequencies suggesting that the two fish Ig repertoires were under a common selection force (Table 3.7). Similarly, a previous study using zebrafish Ig repertoires isolated from the whole fish revealed evidence for convergent evolution in which different fish had comparable VDJ usage frequencies (Weinstein et al., 2009).

In zebrafish, the V_H-elements are shared and rearranged to Dζ-Jζ-Cζ or Dμδ-Jμδ-Cμ/Cδ; however, my results revealed that the diversity of the IGHV gene is significantly lower in the Ig repertoires of IgZ isotype compared to the IgM repertoires (Fig.3.5, Fig.3.6, Fig.3.7, Fig.3.8). All the IgZ repertoires isolated from unvaccinated fish MMΦCs favored a few IGHV gene segments, indicating the presence of common selection pressure on these repertoires. Similar observations were reported using IgT repertoires isolated from unvaccinated rainbow trout spleen, where they found a few shared B-cell clonotypes in the IgT repertoires from different individuals; on the other hand, the most abundant B-cell clonotypes differ from fish to fish (Magadan et al., 2018). They suggested that perhaps the IgT repertoires in the unvaccinated fish are under common selection pressure, and because IgT/IgZ is involved in mucosal immunity, this common selection pressure could be shared components of the microbiota (Magadan et al., 2018). Consistent with this, my data revealed that the usage of the IGHD and IGHJ genes in the IgZ repertoires was biased in most of the repertoires from unvaccinated fish MMΦCs compared to the IgM repertoires. Recent examination of the IgM and IgT repertoires diversity in rainbow trout demonstrated that biases in the V, D, and J genes usage are prevalent in the IgT repertoires; however, IgM repertoires are more diverse and less restricted in their VDJ genes usage (Costa et al., 2021). IgM in fish is the primary Ig class in the serum and functions mainly as a systemic Ig. On the other hand, IgT is focused mainly on mucosal immunity; therefore, it was hypothesized that the abundant IgM would have less pressure to use specific VDJ genes compared to the IgZ isotype, which has a more specialized function (Costa et al., 2021, Zhang et al., 2010, reviewed in Hikima et al., 2011).

My data showed that IgM and IgZ repertoires from unvaccinated fish MMΦCs favored some of the V_H-elements; a few IGHV genes were common to all the IgM or IgZ repertoires from

unvaccinated fish MMΦCs (Fig.3.5, Fig.3.7). Previously, using unvaccinated zebrafish, it was demonstrated that different individuals share a common frequency of VDJ usage, perhaps in response to common selection pressures in their environment (Weinstein et al., 2009). Moreover, the common IGHV genes found in different Ig repertoires from unvaccinated fish MMΦCs could result from combinatorial biases during the recombination process. Consistent with this, biases in V, D, and J genes usage during the recombination process have been established in mammals; for example, in the human naïve B-cell repertoire, different IGHV genes are used as little as 0.1% to over 10% during the rearrangement process (reviewed in Jackson et al., 2013). The bias in gene usage during the recombination process is also reported in the IGHD and IGHJ genes; distinct IGHD genes are used at frequencies between 1% and 15% of the rearrangements, and different IGHJ genes are used from less than 1% to more than 50% of the total rearrangements in human B-cell repertoire (reviewed in Jackson et al., 2013). In addition, combinatorial biases were also reported in unvaccinated rainbow trout (Magadan et al., 2018). Similarly, a previous study using unvaccinated zebrafish demonstrated that a few V, D, and J genes are preferentially used in the primary repertoires in young and adult fish, which indicates that VDJ combinatorial biases are present in zebrafish. However, the biases in VDJ usage are different between young and adult fish (Jiang et al., 2011). Biases in the V, D, and J genes usage during the recombination process are genetically determined; however, the mechanism responsible for the combinatorial biases is not fully understood (reviewed in Jackson et al., 2013). Nevertheless, several factors, including variation in the enhancers and RSS, have been associated with biased V, D, and J genes usage (reviewed in Jackson et al., 2013). My results could support the hypothesis that combinatorial biases during VDJ recombination evolved early in the vertebrate. However, it should be noted that the size of the clones influences VDJ usage analysis, and thus, it is likely that the observed biases

in VDJ usage reflect the expansion of B-cell clonotypes with the appropriate VDJ combination within unvaccinated fish MMΦCs (Fig.3.5, Fig.3.7).

Vaccination induced a shift in the IGHV gene usage, as shown in figure 3.5 and figure 3.6; following vaccination, some of the common V_H-elements were retained, and some of the least used genes in the IgM repertoires from unvaccinated fish MMΦCs became one of the most used genes in specific repertoires from vaccinated fish MMΦCs. Likewise, after vaccination, the favored IGHV genes in the IgZ repertoires from unvaccinated fish MMΦCs became the least used ones, and some of the rare IGHV genes became the most used ones in the IgZ repertoires from vaccinated fish MMΦCs. These results suggest that the injected vaccine induced the expansion of certain IGHV genes in the repertoires from vaccinated fish MMΦCs. However, it is necessary to notice that the number of Ig sequences of IgZ isotype was much lower in the repertoires from vaccinated fish MMΦCs compared to the number of IgZ sequences isolated from unvaccinated fish MMΦCs. A study using rainbow trout vaccinated with an attenuated strain of the rhabdovirus Viral Hemorrhagic Septicemia Virus (VHSV) showed that less frequent IGHV genes in the unvaccinated fish expanded following vaccination (Magadan et al., 2018). Subsequent analysis demonstrated that the antibodies associated with these common IGHV genes (named public antibodies) are able to neutralize the virus and thus protect against VHSV infection (Castro et al., 2021).

IgM repertoires isolated from MMΦCs from fish vaccinated with the same vaccine shared most of their dominant V_H-elements with a relative increase in specific IGHV genes compared to the Ig repertoires from unvaccinated fish MMΦCs (Fig.3.5, Fig.3.6). The finding that two fish using the

same IGHV elements for their activated B-cells following immunization is unlikely to occur by chance alone; therefore, this could be indicative of common selection pressures on the repertoires of different individuals. Also, these results demonstrate that the injected vaccine induced common modifications on the Ig repertoires of the individuals that were vaccinated with the same vaccine. A study using rainbow trout vaccinated intraperitoneally with the enteric red mouth (ERM) vaccine, a *Yersinia ruckeri* bacterin, revealed that the frequency of the dominant clones in the unvaccinated fish was altered after vaccination (Magadan et al., 2019). Vaccination resulted in increased frequencies of certain IGHV genes, and some of the common IGHV genes in the unvaccinated fish Ig repertoires significantly decreased after vaccination. In addition, vaccination-induced IGHV gene usage modifications were shared among different individuals (Magadan et al., 2019). However, the shift in the IGHV usage was significantly lower in the IgT repertoires in the spleen compared with the IgM repertoires, indicating that the systemic route of immunization has a much lower effect on IgT than IgM (Magadan et al., 2019). Consistent with this, IgZ repertoires from fish that were vaccinated with the same vaccine shared only a few of their dominant IGHV genes; however, the total number of Ig sequences of IgZ isotype isolated from vaccinated fish is much lower than the Ig sequences in the IgZ repertoires from unvaccinated fish MMΦCs.

Interestingly, among all the V_H-elements identified in a repertoire, Ig repertoires from unvaccinated fish MMΦCs used IGHV11-1 up to 2% only; the use of IGHV11-1 increased in one or two clusters from vaccinated fish for up to 12% after Alexa-647 conjugated protein vaccination, regardless of whether the conjugating protein was KLH, PE, or BSA. Conversely, the frequency of IGHV11-1 in two repertoires isolated from MMΦCs from unconjugated keyhole limpet

haemocyanin (KLH) vaccinated fish was less than 1% (F11VKCa and F12VKCa (Fig.3.6)). One interpretation for these results is that IGHV11-1 is binding to Alexa-647 as an epitope.

Taken together, my data revealed that the usage of the IGHV, IGHD, and IGHJ genes is more restricted and less diverse in the IgZ repertoires compared to the IgM repertoires; this could be due to the specialized function of IgZ in mucosal immunity compared to the abundant IgM (reviewed in Hikima et al., 2011). In addition, the finding that some of the V_H-elements are favored in the Ig repertoires from unvaccinated fish indicates that the repertoires from these individuals are under common selection pressures in their environment; however, the bias in VDJ usage could result from combinatorial biases during the recombination process. Lastly, my VDJ usage analysis using Ig repertoires isolated from vaccinated fish MMΦCs revealed that the injected vaccine induced the expansion of particular IGHV genes in these repertoires, and the results from Ig repertoires from different fish vaccinated with the same vaccine demonstrate that the repertoires of the different individuals were under common selection pressure.

4.4. MMΦCs have low clonal diversity and more related clones

Within the GCs, FDCs secrete the chemo-attractant CXCL13 to recruit activated B-cells. Nucleating B-cells proliferate rapidly while mutating their Ig genes; as a result, these structures are dominated by a few clonally expanding B-cells (reviewed in De Silva et al., 2015, Mesin et al., 2016). To study the dynamics of B-cell clonotypes within MMΦCs, I used Hill numbers (clonal diversity) and clonal abundance curve (Hill et al., 1973, Stern et al., 2014) using repeated resampling to correct for the variability in total sequence counts among the repertoires.

Using Hill numbers where higher orders (q values) give more weight to larger clones, all the Ig repertoires isolated from unvaccinated and vaccinated fish MMΦCs had low diversity numbers at all the diversity orders (q values) compared with the Ig repertoire isolated from the whole kidney. Between 100 and 425 clones were sampled (species richness at q = 0) from the repertoires isolated from unvaccinated and vaccinated fish MMΦCs (q = 0, Fig.3.15, Fig.3.16). At the inverse of Simpson's concentration index (q = 2), which considers species in proportion to their frequency and ignores rare species, the diversity number for the Ig repertoires from unvaccinated and vaccinated fish MMΦCs was around 120, which is much lower than the diversity number for the whole kidney (≈ 650) at the same order (q = 2; Fig.3.15, Fig.3.16). As the diversity order increased, the diversity number for the Ig repertoires decreased; however, the diversity number for the whole kidney Ig repertoire remained relatively high even at higher diversity orders (e.g., q = 8). These results indicate that MMΦCs have low clonal diversity and more related clones, suggesting that B-cell clonotypes are recruited to the clusters where they clonally expand. In agreement with the observed slight differences in clonal diversity found in different Ig repertoires from MMΦCs, a previous study revealed that B-cells rapid clonal expansion, which leads to a significant loss of clonal diversity, occurs at different rates within different GCs (Tas et al., 2016). We also observed clonal expansion of B-cell clonotypes and loss of clonal diversity in the Ig repertoire isolated from the whole intestine comparable to those detected in the repertoires from MMΦCs, which indicates that B-cell recruitment and clonal expansion occur somewhere in the intestine.

Similarly, using complete clonal abundance distribution, we found a few highly abundant clones in each cluster; low ranking (10^3) clones have much lower abundance than high-ranking (10^0)

clones (Fig.3.17, Fig.3.18). In addition, the steep gradient observed in the Ig repertoires isolated from MMΦCs from unvaccinated and vaccinated fish compared to the gradual gradient of the Ig repertoire isolated from the whole kidney demonstrates that Ig repertoires from unvaccinated and vaccinated fish MMΦCs are primarily dominated by a few highly abundant clones. Overall, the majority of the clones in each MMΦCs had low abundance; nevertheless, a few clones were found at high frequencies in these clusters.

Previously, it has been shown that B-cells clonally expand in rainbow trout following vaccination, which results in reduced clonal diversity in the Ig repertoire (Magadan et al., 2018). In addition, in mammals, FDCs secrete the chemo-attractant CXCL13 to recruit activated B-cells to the GCs, where B-cells clonally expand while mutating their Ig genes which results in low clonal diversity and more related clones within the GCs. Thus, the successful recruitment of B-cells leads to the clonal expansion and accumulation of mutations in the Ig genes of B-cells with the appropriate antigen specificity (reviewed in De Silva et al., 2015, Mesin et al., 2016). Interestingly, work done in our lab using goldfish has shown that CXCL13 is expressed by MMΦs (Waly et al., in preparation).

In summary, the data presented in this section revealed that MMΦCs have low clonal diversity and more related clones and indicate the presence of an effective recruitment mechanism within these clusters, where few B-cells (perhaps with the appropriate antigen specificity) are recruited and diversified. Further supporting the hypothesis that MMΦCs in fish are functionally analogous to GCs.

4.5. Ag-driven selection process occurs within MMΦCs

To determine if there is an active Ag-driven selection process within MMΦCs, I used the Ig repertoires generated from each cluster to examine the ratio of replacement to silent mutations (R/S) in the complementarity determining regions (CDRs) and the framework regions (FWRs). The residues within the CDRs form the sites that contact the antigen and determine the antigen-binding affinity and specificity; on the other hand, FWRs provide the structural backbone of the Ig (Jones et al., 1986, Rada et al., 1991, Jacob et al., 1993). Therefore, if there is an active antigen-driven selection process, we expect to see a higher ratio of R to S mutations in the CDRs compared to the FRWs that are less tolerant to mutations.

Preferential targeting of mutations to the CDRs occurs in part due to biased codon usage in the CDRs and FWRs. Immunoglobulin variable region sequences have evolved to enhance the number of mutations in the antigen contact loops. Accordingly, hotspot motifs are preferentially found in the CDR loops, where replacement mutations are predominantly found to maximize the number of amino acid changes and perhaps alter the antigen-binding affinity (Wagner et al., 1995, Jolly et al., 1996). Conversely, hotspot motifs are avoided in the structurally conserved regions, and the codons within the FWRs have a higher level of degeneracy which leads to mainly silent mutations in the FWRs (Wagner et al., 1995, Jolly et al., 1996, Kepler, 1997, Zheng et al., 2005). Moreover, mutations within the FWRs could disturb the structural integrity of the Ig variable domain which subsequently leads to B-cells death; B-cells that fail to express surface Ig die by apoptosis (reviewed in Rathmell et al., 2002). Recent work has shown that B-cells with non-functional Ig genes die by apoptosis in the dark zone of the germinal center (Mayer et al., 2017).

In mammals and birds, antigen-driven selection process occurs within the GCs where centrocytes compete for a limited number of antigens trapped on the surface of FDCs. Centrocytes with a high affinity for the antigen endocytose it and present it via MHC II to Tfh cells, from which they receive survival signals and consequently positive selection (reviewed in De Silva et al., 2015, Mesin et al., 2016). Therefore, if MMΦCs in fish are functionally analogous to GCs, we would expect Ag-driven selection processes to occur within these clusters.

Although the R/S ratio varied among the different IgM repertoires from unvaccinated and vaccinated fish MMΦCs, their CDRs always had higher R/S estimates compared to the FWRs. In these repertoires, the R/S ratios ranged from 0.7 to 1.88 and from 0.28 to 0.53 in their CDRs and FWRs, respectively (Fig.3.13 A and B). Likewise, a higher R/S ratio was found in the CDRs compared to the FWRs in the Ig repertoires of IgZ isotype isolated from MMΦCs from vaccinated fish. R/S estimates were between 0.71 and 1.43 in the CDRs and between 0.32 and 0.95 in the FWRs of the IgZ repertoires from vaccinated fish MMΦCs (Fig.3.13 D). CDRs are the sites that determine the Ag-binding affinity and specificity, and they are structurally more plastic while FWRs maintain the structural integrity of the Ig variable domain (Jones et al., 1986, Rada et al., 1991). Therefore, my analysis of the R/S ratio demonstrates that the mutations in the Ig sequences isolated from MMΦCs are not randomly distributed and are consistent with the hypothesis that an active antigen-driven selection process of B-cell clonotypes with mutated Ig sequences occurs within MMΦCs.

Even though fish lack FDCs which retain antigens for a long term in the form of antibody and complement immune complexes and provide the first selection step for B-cells within the GCs (reviewed in Flajnik et al., 2018, Allen et al., 2008), it has been demonstrated in several studies that exogenous antigens accumulate in and around MMΦCs in fish for an extended period (Lamers et al., 1985, Lamers CH & De Haas MJ 1985, Ziegenfuss et al., 1991). In addition, in *Cyprinus carpio*, the antigen was trapped more quickly near or within MMΦCs after the injection of immune complexes (Secombes et al., 1980, Secombes et al., 1982). Recent work done in our lab using goldfish showed that MMΦs are the cells that retain the antigen within MMΦCs, and the antigen on the cell surface appeared to be intact (Muthupandian MSc thesis, U of A, 2020). Collectively, these studies indicated that MMΦs in fish perform a similar function to mammalian FDCs and perhaps provide the intact antigens for the emerging mutated B-cells within MMΦCs.

A previous study analyzed the distribution of replacement and silent mutations in the CDRs and FWRs using catfish IgH; they found that the ratio of R/S mutations was similar in the CDRs and FWRs and, as a result, they concluded that bony fish lack antigen-driven selection mechanisms (Yang et al., 2006). However, they only analyzed 187 unique IgH sequences, and the size of the clones in their dataset was only between 2 and 9 (Yang et al., 2006). Similarly, analysis of 80 mutations within zebrafish IgL chain revealed that the distribution of replacements and silent mutations is not significantly different between the CDRs and FWRs, and they suggested that antigen-driven selection process is restricted in zebrafish (Marianes et al., 2011). Subsequent attempts to explain the absence of Ag-driven selection processes in fish suggested that this could be due to the lack of the appropriate microenvironment for selecting mutated B-cells in fish (reviewed in Fillatreau et al., 2013).

Examination of the distribution of the mutations in the CDRs and FWRs using IgNAR gene in the nurse shark revealed that CDR1 had a significant bias toward replacement mutations, indicating the presence of an active antigen-driven selection process (Diaz et al., 1998). These observations suggest that antigen-driven selection mechanisms appeared early in the gnathostome lineage.

Consistent with my data, a recent study using $IgD^{+}IgM^{-}$ B-cells from rainbow trout showed that the distribution of replacement and silent mutations in the CDR2 and FWR3 suggests that antigen-mediated IgD selection occurs in rainbow trout (Perdiguero et al., 2019).

Four of the IgZ repertoires from unvaccinated fish MMΦCs had a higher R/S ratio in their FWRs compared to their CDRs; these four IgZ repertoires are from the spleen and kidney of two different fish (Fig.3.13 C). Intriguingly, these repertoires had mean tree lengths (i.e., the average expected substitutions per codon site) equal to 0.55, 0.25, 1.41, and 0.44 (Table 3.5). These mean tree lengths are long compared to IgM repertoires from MMΦCs from unvaccinated and vaccinated fish; average mean tree lengths of 0.13 (Table 3.5, Table 3.6). Similarly, IgZ repertoires from vaccinated fish MMΦCs had an average mean tree length of 0.18 (Table 3.6). In addition, the repertoire with the longest mean tree length also had the highest R/S ratio in the FWRs compared to the CDRs. Examination of the relationship between the mean tree length and the ratio of replacement to silent mutations in the FWRs revealed a positive correlation between the mean tree length and the R/S ratio in the FWRs; all the IgZ repertoires with longer mean tree length had a high R/S ratio in their FWRs. IgM repertoires from unvaccinated and vaccinated fish MMΦCs had R/S ratio in their FWRs between 0.28-0.5; however, this ratio reached 1.17 in the IgZ repertoires

from unvaccinated fish MMΦCs. The slight increase in the mean tree length found in some IgZ repertoires isolated from vaccinated fish MMΦCs is associated with an increased R/S ratio in their FWRs (Fig.3.14).

Together, these results indicate that the negative selection found in some of the IgZ repertoires from unvaccinated fish MMΦCs (higher R/S ratio in the FWRs compared to the CDRs) suggests that perhaps these B-cell clonotypes had been acquiring mutations for a long time and the binding affinity of their immunoglobulins reached an optimal affinity. Thus, the addition of mutations is not beneficial and possibly selected against. Similar observations were reported recently in response to the influenza vaccine; Ig repertoires isolated at the peak of the vaccine response had an increased mean tree length (i.e., the average expected substitutions per codon site) which was associated with significantly decreased R/S ratios in the CDRs compared to the FWRs (a signature of negative selection; Hoehn et al., 2019). However, their analysis also showed that the ratio of replacement to silent mutations in the FWRs slightly decreases as the mean tree length increases (Hoehn et al., 2019). Nevertheless, their results overall indicate that increased mean tree length is associated with negative selection, and they suggested that eventually shifting to negative selection is a general feature of affinity maturation (Hoehn et al., 2019). In addition, using HIV-1 broadly neutralizing antibodies (bnAbs), it has been shown that selection pressure changes from positive selection to negative selection over time (Sheng et al., 2016). The authors concluded that as antigen-binding affinity becomes high during the immune response, additional replacement mutations are less likely to increase the binding affinity (Sheng et al., 2016). Therefore, shifting to negative selection instead of positive selection could be more beneficial; they called this hypothesis the affinity maturation selection (AMS) model (Sheng et al., 2016). Recent analysis

using Ig sequences from both patients with asthma and non-asthmatic controls showed that IgD^+ IgM^- B-cells from the bronchial mucosa had a significant negative selection pressure in their CDRs (Ohm-Laursen et al., 2021). The frequency of mutations in these sequences was considerably high, and they revealed that the high number of mutations within a clone is associated with negative selection pressure (Ohm-Laursen et al., 2021).

In summary, in support of the hypothesis that MMΦCs are acting as primordial germinal centers in fish, the data presented in this section indicate that, similar to germinal centers, antigen-driven selection processes occur within MMΦCs. A higher ratio of replacement to silent mutations in the CDRs compared to the FWRs in the Ig sequences isolated from unvaccinated and vaccinated fish MMΦCs is consistent with the presence of active antigen-driven selection mechanisms in these clusters. MMΦs, which retain intact antigens on their surface (Muthupandian MSc thesis, U of A, 2020), perhaps provide a selection platform within these clusters. In addition, I found a positive correlation between the number of mutations within a clone and the ratio of replacement to silent mutations in the FWRs. Mutations within the FWRs affect the structural integrity of the Ig and consequently can be deleterious and lead to B-cells death; my results indicate that as the number of mutations in the Ig sequence increases, a shift from positive selection to negative selection occurs within MMΦCs in fish.

4.6. AID and error-prone polymerases are involved in the mutation process within MMΦCs

In mammals, SHM occurs within the germinal center and is initiated by AID; both base excision repair (BER) and mismatch repair (MMR) pathways are involved in processing the mutation generated by AID (reviewed in Pilzecker et al., 2019, Di Noia and Neuberger, 2007). Analysis of

the mutation patterns using Ig repertoires isolated from unvaccinated and vaccinated fish MMΦCs revealed that AID and error-prone polymerases are involved in the mutation process.

Within the dark zone of the germinal center, AID binds single-stranded DNA (ssDNA) and deaminates deoxycytidine (dC) preferentially within a WRC into deoxyuracil (dU); causing dU: dG mismatch. DNA replication over U results in C to T transition due to the similarity between U and T, which will lead to C: G to T: A transition. However, the removal of the deoxyuridine by uracil-DNA glycosylase (UNG) results in an abasic site, translesion synthesis (TLS) polymerases such as Rev1, polymerase ζ, and polymerase η are recruited to bypass the non-instructive site; these polymerases lack proof-read activity and copy DNA with a lower fidelity which leads to transversions and transitions (reviewed in Krokan et al., 2013, Yu et al., 2019, Pilzecker et al., 2019, Feng et al., 2020). In addition, the abasic site can be recognized and cleaved by apurinic/apyrimidinic endonuclease 1 (APE1), which creates a nick in the phosphodiester backbone. The resulting single-stranded break (SSB) could directly result in gene conversion in which a set of homologous pseudogenes upstream of the rearranged IgV are used as donors for sequences (reviewed in Di Noia et al., 2007). Alternatively, dU: dG mismatch can be recognized by MSH2/MSH6, a component of the ncMMR machinery; afterward, a nick is made by PMS2 and MLH1 endonuclease complex, this nick provides an entry point for exonuclease-1 (Exo1), and a gap will be generated and extended by removing the mismatch and adjacent nucleotides (Bardwell et al., 2004). Subsequently, the gap is resynthesized by error-prone polymerases such as pol η, which is responsible for the majority of mutations at A: T and preferentially targets WA/TW for mutations on the non-transcribed and transcribed strand, respectively (reviewed in Pilzecker et al., 2019).

The analyzed hotspot motifs are WRC and its complement GYW and WA, and its complement TW (where W = A/T, R = A/G, and Y = C/T); altered substitution rates occur only in the underlined bases. The results presented in table 3.8, table 3.9, table 3.10, and table 3.11 show that all the analyzed hotspot motifs had altered substitution rates; when all the Ig repertoires from unvaccinated and vaccinated fish MMΦCs considered, GYW motif exhibited the highest substitution rate increases (1.3) followed by WA (0.6) and closely by WRC (0.5) then TW (0.08). These mutation frequencies are similar to the values reported in a previous study using 27 healthy individuals of different age and sex; they revealed that GYW motif had the largest substitution rate increases (2.46), followed by WRC (1.87), then WA (1.71) and TW motif exhibited the lowest substitution rate increases (1.19) (Hoehn et al., 2019).

Generally, WRC/GYW motifs had increased substitution rates in the analyzed Ig repertoires, with up to 4x increases in GYW substitution rate. Altered substitution rates in WRC/GYW motifs compared to regular motifs indicate the involvement of AID in the mutation process, which has the ability to mutate both the coding and non-coding strands (reviewed in Feng et al., 2020, Di Noia et al., 2007, Pilzecker et al., 2019). Since AID requires a single-stranded DNA, it was suggested that AID preferentially targets the non-transcribed strand, which will be single-stranded within the transcription bubble; however, only 58 % of AID induced uracils were in the non-transcribed strand, and 42 % of the uracils were found in the transcribed strand (Maul et al., 2011). During the repairing process, Rev1 is the translesion synthesis (TLS) polymerase responsible for most C: G to G: C transversion mutations; the active site of Rev1 allows it only to insert deoxycytidine monophosphate (dCMP) into DNA (reviewed in Bahjat et al., 2017, Pilzecker et al.,

2019, Feng et al., 2020). However, it is still not clear which polymerase is responsible for G: C to T: A transversion (reviewed in Pilzecker et al., 2019).

Examination of the fold-change in mutability for WA and its complement TW revealed that Ig repertoires isolated from unvaccinated and vaccinated fish MMΦCs had altered substitution rates at WA/TW motifs, which indicates the involvement of error-prone polymerases in the repairing process of the lesion generated by AID. Mutations at A: T pairs are introduced by recruiting error-prone polymerases to resynthesize the patch of DNA that has been degraded by Exo1 activity. Pol η is the polymerase responsible for the majority of mutations at A: T; it is a member of the Y-family of DNA polymerases and a translesional polymerase (reviewed in Yang et al., 2007, Di Noia et al., 2007, Pilzecker et al., 2019). In addition, it has been demonstrated that Pol η primarily resynthesis the non-transcribed strand using the transcribed strand as a template and introduces twofold more mutations at WA compared to TW by inserting G opposite to T on the transcribed strand (Mayorov et al., 2005, reviewed in Pilzecker et al., 2019). Consistent with this, my data revealed a strong bias to mutations at WA compared to TW motif; on average, the fold change in mutability values were 0.6 and 0.08 for WA and TW, respectively (Table 3.8, Table 3.9, Table 3.10, Table 3.11). In patients with xeroderma pigmentosum variant (XP-V) disease who lack Pol η, the overall frequency of mutation was normal; however, mutations at A: T pairs significantly decreased, and this was associated with an increase in the number of mutations at C: G (Zeng et al., 2001). Nonetheless, in the absence of Pol η, mutations at A: T pairs can be generated by other low fidelity polymerases such as polymerase kappa (Pol κ), polymerase ζ (Pol ζ), and DNA polymerase ι (Pol ι; Faili et al., 2009, Saribasak et al., 2012, Maul et al., 2016).

In agreement with my hotspot motifs mutability data, earlier studies using immunoglobulin sequences isolated from zebrafish and rainbow trout revealed that SHM occurs preferentially at AID and error-prone polymerases (Pol η) hotspot motifs in bony fish (Marianes et al., 2011, Abos et al., 2018, Perdiguero et al., 2019). For example, a previous study using zebrafish IgL chain revealed that SHM occurs mainly at AID hotspot motifs (WRC/GYW) (Marianes et al., 2011). In addition, it has been revealed that during an immune response to *T. bryosalmonae* in rainbow trout, IgT and IgM expressing B-cells clonally expand while acquiring mutations, mainly at WRC motif; also, mutations at WA/TW motifs increased throughout the immune response (Abos et al., 2018). Furthermore, a recent study using rainbow trout showed that in response to commensal bacteria, IgD expressing B-cells clonally expand and mildly mutate their Ig gene in both the gills and gut, and the mutations at WRC/GYW and WA/TW increased during the clonal expansion (Perdiguero et al., 2019).

Using Ig repertoires isolated from MMΦCs from unvaccinated and vaccinated fish, the ratio of transition to transversion mutations ranged from 1.33 to 2.37. During SHM, DNA replication over AID induced uracil by any of the DNA polymerases results in C: G to T: A transition due to the similarity between U and T (reviewed in Feng et al., 2020, Di Noia et al., 2007, Pilzecker et al., 2019). Therefore, transition (change from a purine to purine or from pyrimidine to pyrimidine) is favored over transversion (changes between purines and pyrimidines) during SHM, with a ratio of transitions to transversions higher than the theoretical ratio (1:2) if mutations occurred randomly (reviewed in Di Noia et al., 2007). In wild-type mice ($UNG^{+/+}$), 65% of the mutations induced by AID during SHM were transition mutations (instead of 33% of the total mutations if mutations were random), and in $UNG^{-/-}$ mice, mutations significantly shifted to transitions (95%) at dC/dG

(Rada et al., 2002). The ratio of transition to transversion mutations found in our Ig repertoires suggests that the mutations observed in these Ig repertoires are induced by AID rather than being randomly generated.

Consistent with my results, using Ig heavy μ-chain from catfish spleen, it was revealed that transitions are more frequent than transversions; 60.3% of the total mutations were transitions (Yang et al., 2006). Similarly, a previous study using zebrafish IgL chain revealed that transition mutations are substantially more abundant (85%) than transversions. Furthermore, they found the ratio of transitions to transversions to be much higher (5.64) than the theoretical value (1:2), and they concluded that the role of UNG is limited in zebrafish (Marianes et al., 2011). However, in their analysis, they only examined 93 VL mutations (Marianes et al., 2011). The values of transition to transversion mutations ratio presented in table 3.8, table 3.9, table 3.10, and table 3.11 suggest that UNG is involved in the mutation process in zebrafish, given the mechanism of action of AID and the subsequent repair mechanisms. AID deaminates cytidine into uracil, causing U: G mismatch. DNA replication over uracil results in C to T transition; nonetheless, the removal of uracil during the repairing process by UNG will lead to transitions or transversions at the abasic site (reviewed in Zanotti et al., 2016, Pilzecker et al., 2019, Feng et al., 2020, Di Noia et al., 2007). Therefore, if UNG activities were limited in zebrafish, we would expect our Ig repertoires to have a much higher ratio of transition to transversion mutations.

Examination of nucleotide insertions and deletions (indels) in the productive Ig sequences isolated from unvaccinated and vaccinated fish MMΦCs revealed that indels are rare and are found in only 2.2 % of the functional Ig sequences in each repertoire (Table 3.12, Table 3.13). Previously, using

IgH sequences isolated from human tonsils, it has been demonstrated that indels are associated with somatic hypermutation process; similar to SHM indued point mutations, indels are predominantly localized to the CDRs of the variable region, and no indels are found in the constant region of the Ig sequences (Wilson et al., 1998). In addition, indels are observed in association with somatically mutated Ig sequences; however, they are not detected in the unmutated (germline-encoded) Ig sequences (Wilson et al., 1998, Goossens et l., 1998). Moreover, it has been shown that indel mutations tend to occur at SHM associated hotspot motifs (RGYW/WRCY) (Wilson et al., 1998). Although the mechanisms that lead to the generation of indels during SHM are poorly understood, it has been suggested that indels occur due to misalignment of short-repeated DNA sequences, which creates an unpaired loop between the repeats. Subsequently, deletions are formed when the repeated sequences are removed; on the other hand, insertions are generated as a result of sequence duplication (Bowers et al., 2014, de Wildt et al., 1999, reviewed in Garcia-Diaz et al., 2006).

Consistent with my indels frequency data (Table 3.12, Table 3.13), several studies reported that SHM associated indels in humans are infrequent; it is estimated that only 1.3 to 6.5 % of circulating B-cells contain indels (reviewed in Briney et al., 2013). Moreover, using human Ig sequences generated in response to the capsular polysaccharides of *Streptococcus pneumoniae* or *Bacillus anthracis* vaccine, it was shown that indels are present in 9.7 % of the rearranged Ig sequences (Reason et al., 2006). Furthermore, using zebrafish VJ-C cDNA clones, it was revealed that indels are rare in zebrafish IgL chain; analysis of 93 cDNA clones found only one deletion and a few insertion mutations (Marianes et al., 2011). Taken together, these observations indicate that the frequency of indel mutations found in our Ig repertoires from unvaccinated and vaccinated fish

MMΦCs suggests that the observed indels are not random and perhaps are generated during SHM. Despite being rare, indels play a substantial role in diversifying the Ig repertoires; in addition, it has been shown that indel mutations are critical during the immune repones against viruses and bacteria (reviewed in Briney et al., 2013). A higher estimate of indels was reported among HIV-1 broad neutralizing antibodies (HIV-1 bnAbs); about 40 % of the examined bnAbs had indels (Kepler et al., 2014).

Our preliminary analysis of the contribution of gene conversion to the diversity of the Ig sequences in zebrafish revealed that gene conversion does not seem to be a major contributor to Ig diversity in zebrafish. Using Ig repertoires isolated from unvaccinated and vaccinated fish MMΦCs, major gene conversion-like events were not found in the Ig sequences, and only a few minor gene conversion-like events were detected in a few Ig sequences. In some species, such as birds, rabbits, cows, and pigs, the number of functional V genes is very limited; however, they have a large number of V pseudogenes (ΨV) (reviewed in Choudhary et al., 2018). These pseudogenes are used as a donor sequence during GCV to diversify their Ig genes (Di Noia et al., 2004, reviewed in Di Noia et al., 2007). The finding that zebrafish IgH locus has only eight V pseudogenes, which is considerably lower than the number of the functional V genes in this locus (40 potentially functional VH segments; Danilova et al., 2005), further indicates that gene conversion is not a major contributor to Ig diversity in zebrafish.

Unlike mutations found in Ig genes in other species, mutations exist in tandem nucleotide stretches of 2–4 bases in sharks Ig genes (Diaz et al., 1999). Manual examination of the mutations in a few clones suggested that tandem mutations are not dominant in bony fish Ig genes as in sharks; only

a few mutations were found on adjacent nucleotides. The availability of the tools to perform large-scale analysis on tandem mutations will verify these observations.

In summary, using Ig repertoires isolated from unvaccinated and vaccinated fish MMΦCs, analysis of the mutation patterns strongly suggest that the mutations are generated by AID and error-prone polymerases during the subsequent repair of the lesion generated by AID. Examination of the fold change in mutability within hotspot motifs that are preferentially targeted for mutations by AID and error-prone polymerases, WRC/ GYW and WA/ TW, respectively, compared to regular motifs showed that all the analyzed hotspot motifs had altered substitution rates. In addition, the frequency of mutations within the hotspot motifs is consistent with a previously reported hotspot motifs mutability (Hoehn et al., 2019). The ratio of transitions to transversions further suggests that the mutations found in our Ig repertoires are not randomly generated; instead, they are induced by AID and indicate that UNG is involved in the mutation process in zebrafish. Moreover, similar to SHM associated indels in humans, examination of indels in the productive Ig sequences isolated from MMΦCs showed that indels are infrequent. Collectively, these observations indicate that AID and error-prone polymerase induced somatic hypermutations occur within MMΦCs, which further support the hypothesis that these clusters in fish are functionally analogous to the germinal centers of homeotherms. In mammals and birds, SHM occurs predominantly within the dark zone of the germinal center (reviewed in De Silva et al., 2015, Mesin et al., 2016).

4.7. IgM and IgZ CDR3 length distribution

Analysis of the non-templated and palindromic nucleotides (N- and P-nucleotides) additions using productive Ig sequences isolated from unvaccinated and vaccinated fish MMΦCs revealed that overall IgZ repertoires accumulated more N- and P-nucleotide additions than IgM repertoires (Table 3.14, Table 3.15, Table 3.16, Table 3.17). This was associated with a longer CDR3 length of the IgZ repertoires from unvaccinated fish MMΦCs compared to IgM repertoires from unvaccinated and vaccinated fish MMΦCs (Fig.3.19 A and B, Fig.3.20 A). These observations indicate that the variability in the CDR3 length could be related to the compartmentalization and the functional differences of the two Ig isotypes.

N- and P-nucleotides are added during VDJ recombination process within the primary lymphoid organ. VDJ recombination begins when RAG proteins bind and cleave the recombination signal sequence (RSS) flanking each V, D, and J gene segment; subsequently, the free 3' hydroxyl group in the broken ssDNA attacks the opposite strand and forms a covalently sealed hairpin at the end of the gene segment. P-nucleotides are added when the Artemis endonuclease opens the hairpin ends asymmetrically. Another form of junctional insertion, N-nucleotides, is provided by the action of the enzyme terminal deoxynucleotidyl transferase (TdT), which adds nucleotides randomly to the 3′-ends (Schatz et al., 1989, reviewed in Schatz et al., 2011, Helmink et al., 2012, Roth et al., 2015). Analysis of N- and P-nucleotides addition using zebrafish IgL VJ-C cDNA clones showed that N- and P-nucleotides were absent at the CDR3 (junction), and they suggested that junctional diversity is limited in zebrafish IgL chain (Marianes et al., 2011). Nevertheless, consistent with my findings, using high throughput sequencing, analysis of zebrafish IgH chain revealed that nucleotides insertion at the VDJ junctions is widespread in adult zebrafish compared to young, immature fish, which indicates that, similar to mammals, the activity of terminal

deoxynucleotidyl transferase is developmentally regulated in zebrafish (Jiang et al., 2011, Feeney et al., 1990). Earlier studies in mammals, using IgH sequences from fetal and newborn mice, have established that N-nucleotides are absent in neonatal B-cells (Feeney et al., 1990).

In fish, IgM is the primary Ig isotype in the serum and the most widespread Ig class during systemic immune responses; on the other hand, IgZ, which is produced by a distinct subset of B-cells, has been described as the primary antibody isotype in mucosal immunity in teleost (Zhang et al., 2010, reviewed in Hikima et al., 2011, Salinas et al., 2021). Using rainbow trout vaccinated with *Flavobacterium columnare*, a mucosal pathogen, it has been demonstrated that pathogen-specific IgT expressing B-cells proliferate and accumulate at the mucosal barrier (buccal mucosa) in response to the bacterial pathogen; on the other hand, pathogen-specific IgM titers were detected in the serum of the vaccinated fish (Xu et al., 2020). Similarly, it has been shown that the number of IgT^+ B cells significantly increases at the mucosal barrier of fish infected by immersion with a parasitic pathogen (*Ichthyophthirius multifiliis*); however, the abundance of IgM^+ B cells did not change significantly in the buccal surface (Yu et al., 2019). Nevertheless, in channel catfish, intraperitoneal injection with TNP-KLH leads to the production of detectable levels of antigen-specific IgM in the serum by week two post-immunization (Wu et al., 2019). Collectively, these studies further support the key role of IgT/IgZ and IgM in mucosal and systemic immunity, respectively.

The results presented in figure 3.19 show the difference in CDR3 length distribution between IgM and IgZ repertoires from unvaccinated fish MMΦCs and indicate that IgZ sequences with a longer CDR3 were selected. Similar observations were reported recently using unvaccinated rainbow

trout, where they found that IgT repertoires are skewed towards longer or shorter CDR3 compared with IgM repertoires (costa et al., 2021). Interestingly, the CDR3 length distribution was similar for IgM repertoires isolated from both unvaccinated and vaccinated fish MMΦCs (Fig.3.19 A, Fig.3.20 A). Nevertheless, the CDR3 length distribution varied among the IgZ repertoires isolated from vaccinated fish MMΦCs; however, these repertoires had only a few unique Ig sequences compared to IgM and IgZ repertoires from unvaccinated fish MMΦCs and IgM repertoires from vaccinated fish MMΦCs. Likewise, in general, IgZ repertoires from unvaccinated fish MMΦCs with fewer unique Ig sequences had a diverse CDR3 length distribution compared with IgZ repertoires with a higher number of unique Ig sequences from unvaccinated fish MMΦCs (Fig.3.19 B, Table 3.1). It should be noted that the CDR3 length and N- and P-nucleotides were calculated using individual Ig sequences, and thus they are biased by clone size.

In summary, examinations of the CDR3 length distribution and N- and P-nucleotide additions suggest that the differences in CDR3 length distribution are caused by N- and P-nucleotides addition. Moreover, IgZ repertoires from unvaccinated fish MMΦCs had a longer CDR3 length compared to IgM repertoires from the same fish, which could indicate the selection of IgZ sequences with a longer CDR3. The dissimilarity in the CDR3 length between IgM and IgZ repertoires could be related to the compartmentalization and functional differences of teleost IgM and IgZ.

4.8. Antigen retention occurs within MMΦCs isolated from zebrafish

Within the light zone of the germinal center, FDCs retain antigens for long periods of time in the form of antibody and complement immune complexes and provide the first selection step for GC B-cells (reviewed in Allen et al., 2008). During the germinal center reaction, high-affinity centrocytes endocytose the antigen, internalize and present it via class II MHC to Tfh cells, from which they receive further BCR stimulation in addition to other survival signals (reviewed in Allen et al., 2008, Kranich et al., 2016). Ectotherms lack FDCs; however, several studies have established that long-term antigen retention occurs in fish within or near MMΦCs in the spleen and kidney (Lamers et al., 1985, Lamers CH & De Haas MJ 1985, Ziegenfuss et al., 1991, Press et al., 1996, reviewed in Flajnik et al., 2018). Moreover, it was revealed that the antigen was trapped more rapidly in *Cyprinus carpio* after the injection of immune complexes, suggesting that the antigen is trapped within these clusters using a similar mechanism to FDCs in mammals (Secombes et al., 1980, Secombes et al., 1982).

To determine if long-term antigen retention occurs within MMΦCs in zebrafish, I vaccinated zebrafish with different proteins conjugated to Alexa-647. The results shown in figure 3.21 and figure 3.22 confirmed that long-term antigen retention occurs within MMΦCs in the spleen and kidney in zebrafish. Additionally, work done in our lab using goldfish injected with various proteins conjugated to a fluorescent tag and confocal microscopy revealed that autofluorescent MMΦs are involved in antigen retention. Further experiments have shown that the retained antigen on the cell surface is in an intact form (Muthupandian MSc thesis, U of A, 2020). These observations verified the presence of a selection platform within MMΦCs in zebrafish, which further supports the hypothesis that these clusters in fish are the sites where B-cells affinity mature their antibodies and thus are germinal centers analogous.

4.9. Lymphocyte-like cells within MMΦCs isolated from zebrafish

Using intact, isolated zebrafish MMΦCs and high throughput sequencing, the total number of B-cell clonotypes within a cluster ranged from 510 to 3660. To confirm the presence and verify the numbers of lymphocyte-like cells within zebrafish MMΦCs, I used imaging flow cytometry (ImageStream). Based on the cell size and internal complexity, a total of 5,305 and 6,027 lymphocyte-like cells were found in individual MMΦCs isolated from the spleen and kidney, respectively (Table 3.19).

Previous studies have established that MMΦCs in fish vary in pigmentation intensity, size, and number, and their size and number increase in response to infection or immunization (Herráez et al., 1986, Agius, 1979, Secombes et al., 1982, reviewed in Steinel and Bolnick, 2017). For example, in Nile tilapia, the number and size of MMΦCs significantly increased in response to *Aeromonas hydrophila* and *Mycobacterium marinum* (Manrique et al., 2019). Likewise, MMΦCs size and frequency increased in southern bluefin tuna in response to parasitic infection (Nowak et al., 2021). Consistent with these observations, the diameter, and the pigmentation intensity, of the analyzed clusters varied, and their diameter ranged from ≈200 to 350 μm; thus, the total number of cells within these clusters was estimated to be between 33,231 and 39,120 cells/cluster (Figure.3.23, Figure.3.24, Table 3.18).

Lymphocyte-like cells were identified based on their size and internal complexity; lymphocyte-like cells have low internal complexity and are small in size (Petrie-Hanson et al., 2009).

Furthermore, using a nuclear stain, we were able to verify that these lymphocyte-like cells have a high nuclear to cytoplasmic ratio. In figure 3.25, cells that are larger than lymphocyte-like cells are perhaps monocytes/macrophages-like cells (Figure 3.25 C, to the right of R3). In addition, it has been established that MMΦs vary widely in their size and internal complexity; nonetheless, MMΦs have higher internal complexity and are larger than lymphocyte-like cells (Diaz-Satizabal et al., 2015).

Lymphocyte-like cells were estimated to represent between 13% to 19% of the total number of cells within each cluster; interestingly, the diameter of lymphocyte-like cells isolated from MMΦCs varied and ranged from 5 to 7 μm. In agreement with this, B-lymphocytes within the germinal center vary in size; large, rapidly diving centroblasts within the dark zone differentiate into small, non-proliferating centrocytes as they enter the light zone (reviewed in Gatto et al., 2010, De Silva et al., 2015). The primary cell type within the germinal centers is B-cell, and using spleen sections isolated from mice immunized with 2-phenyl-5-oxazolone (phOx)-chicken serum albumin (CSA), it has been shown that while the numbers of T-cell within the germinal center change over time, in a typical germinal center the ratio of T-cells to B-cells appears to be 12.5:100 (Wittenbrink et al., 2011). Interestingly, this ratio is not influenced by the size of the germinal center (Wittenbrink et al., 2011). Therefore, if MMΦCs in fish are germinal centers analogue, then most of the observed lymphocyte-like cells within these clusters are B-cells which is consistent with the total number of unique Ig transcripts isolated from individual MMΦCs (Table 3.1, Table 3.19).

Overall, imaging flow cytometry data verified the presence of lymphocyte-like cells within zebrafish MMΦCs. In addition, the numbers of lymphocyte-like cells within zebrafish MMΦCs found using imaging flow cytometry are consistent with the numbers of unique B-cell clonotypes within a cluster obtained using high throughput sequencing.

4.10. V-less IgM and IgZ

Several Ig transcripts without a VH region where the leader sequence is spliced to the first exon of the C-μ or C-ζ were found using Ig repertoires isolated from unvaccinated and vaccinated zebrafish MMΦCs. Similarly, previous work done in our lab using 5' RACE of catfish MMΦCs revealed that about 20% of the Ig sequences isolated from these clusters lacked the V-element (C-W Fan and Magor, unpublished). A previous study revealed that all secreted IgD in catfish is without a V-region (named V-less) where the leader sequence is spliced to the Cδ1 domain (Edholm et al., 2010). They suggested that perhaps V-less IgD functions as a pattern recognition molecule through its Fc region and targets pathogens for destruction by binding to an IgD-binding receptor on granulocytes (Edholm et al., 2010). However, we did not perform any functional analysis on the V-less transcripts of IgM and IgZ.

4.11. Mutations within the Cμ1

Using the first Cμ exon (Cμ1), 260 nucleotides were sequenced. Analysis of the sequenced Cμ1 region revealed that point mutations, insertion, and deletion mutations accumulate within this region. These mutations could be a result of the IgH enhancer, Eμ3′enhancer, in fish where it is

found in the intron between the Cμ and Cδ genes in contrast to mammals where it is located in the JH and Cμ intron, and it has been shown that SHM is directed to the Ig loci by Ig enhancers and enhancer-like elements (Buerstedde et al., 2014, Qian et al., 2014, reviewed in Hikima et al., 2011). However, as mentioned in section 3.9, we could not perform further analysis on the observed mutations, and thus we were not able to use the Cμ1 region as a control for sequencing error. Nonetheless, it has been established that removing low-quality reads (Q score <30) significantly reduces sequencing errors. Moreover, discarding singletons is very effective in reducing errors (discussed in section 2.6.1; Scott et al., 2018, Rosenfeld et al., 2018, Tedersoo et al., 2010, Brown et al., 2015, reviewed in Chaudhary et al., 2018).

CHAPTER V

CONCLUSION

5. conclusion

The divergence of jawless vertebrates (agnathans) and jawed vertebrates (gnathostomes) occurred around 550 million years ago (reviewed in Sutoh et al., 2021). While jawless vertebrates lack BCR, TCR, and MHC, early studies reported evidence for the presence of an adaptive immune-like response in these species (Hildemann et al., 1969, Fujii et al., 1979). For example, studies using various antigens, including killed *Brucella* cells and sheep red blood cells (SRBC), demonstrated that antigenic challenge induced the formation of antigen-specific immune response and immunological memory (Hildemann et al., 1969, Fujii et al., 1979). Subsequent studies revealed that lymphocyte-like cells exist in jawless fish, and they have a distinct lymphocyte-based adaptive immune system (reviewed in Cooper et al., 2006, Flajnik et al., 2010). BCRs and TCRs in jawed vertebrates belong to the immunoglobulin superfamily (IgSF); on the other hand, jawless fish antigenic receptors, named variable lymphocyte receptors (VLRs), belong to the leucine-rich repeat (LRR) receptor family (reviewed in Flajnik et al., 2018). Despite being structurally unrelated, the diversity of the VLRs is generated through somatic recombination similar to BCRs and TCRs in jawed vertebrates (reviewed in Sutoh et al., 2021). In addition, functional analyses have established that distinct types of VLRs in jawless vertebrates perform functions equivalent to gnathostomes BCRs and TCRs (Alder et al., 2008, Guo et al., 2009). Although the origin and evolution of the two systems are not fully understood, scientists believe that perhaps a VLR-based

adaptive immune system existed in a common ancestor of all vertebrates. Subsequently, this system was kept in jawless vertebrates but replaced in jawed vertebrates by BCR- TCR-based adaptive immune system (reviewed in Sutoh et al., 2021).

Similar to the existence of two distinct forms of adaptive immunity in jawless and jawed vertebrates, the findings in this thesis provide strong evidence that antibody affinity modification in fish occurs within MMΦCs, which are structurally distinct from the microenvironment in which affinity maturation process takes place in mammals and birds. MMΦCs are found in both jawless and jawed vertebrates, which indicates that these clusters evolved in a common ancestor of all vertebrates. However, MMΦCs are absent in birds and mammals. While the origin and evolution of MMΦCs are not fully understood, some scientists believe that these clusters evolved into germinal centers in early mammals (reviewed in Steinel et al., 2017).

Analysis of VDJ repertoires using intact isolated MMΦCs from the spleen and kidney of zebrafish revealed that the hallmarks of the germinal center reaction are found within these clusters and strongly supports the hypothesis that MMΦCs in fish are analogous to germinal centers. Specifically, the construction of clonal lineage trees revealed that B-cell clonotypes clonally expand within MMΦCs while acquiring mutations in their Ig VDJ sequences, with some of the dominant clones having more than 400 mutated daughter cells. In addition, the presence of smaller lineages indicates that these clusters are continually evolving structures and develop toward the dominance of a few clones and their progeny. The presence of subclones could also indicate that MMΦCs are open structures, and activated B-cells can enter and proliferate within these clusters at varying times. Moreover, my results indicate that, similar to germinal centers, antigen-driven

selection processes occur within MMΦCs. Also, mutation patterns analysis strongly suggests that the observed mutations in the Ig sequences are generated by AID and error-prone polymerases within these clusters. Furthermore, diversity analysis indicates the presence of an effective recruitment mechanism within these clusters, where few B-cells (perhaps with the appropriate antigen specificity) are recruited and diversified.

My results revealed that the role of MMΦCs in the development of Ag-specific B-cells is not limited to IgM expressing B-cells but also involves IgZ expressing B-cells. Moreover, VDJ usage analysis showed that some of the V_H-elements are favored in the Ig repertoires from unvaccinated fish, which indicates that these repertoires were under common selection pressures in their environment; though, the bias in VDJ usage could result from combinatorial biases during the recombination process. Also, VDJ usage analysis using Ig repertoires isolated from vaccinated fish MMΦCs revealed that the injected vaccine induced the expansion of certain IGHV genes in these repertoires, and the results from the repertoires isolated from MMΦCs from different fish vaccinated with the same vaccine demonstrate that the repertoires of the different individuals were under common selection pressure.

Interestingly, my data support the hypothesis that the posterior kidney is a secondary lymphoid organ similar to the spleen in fish. In addition, analysis of the Ig repertoire isolated from the intestine revealed that mutations and clonal expansion of B cell clonotypes occur in the intestine.

Imaging flow cytometry data confirmed the presence of lymphocyte-like cells within zebrafish MMΦCs. In addition, the numbers of lymphocyte-like cells within zebrafish MMΦCs found using

imaging flow cytometry are consistent with the numbers of unique B-cell clonotypes within a cluster obtained using high throughput sequencing. Furthermore, I was able to verify that long-term antigen retention occurs within MMΦCs in the spleen and kidney in zebrafish using labeled antigens.

Overall, the findings of my thesis strongly suggest that melano-macrophage clusters in fish are functional analogues of homeotherm germinal centers. They also indicate that the posterior kidney of bony fish is a *bona fide* secondary lymphoid tissue.

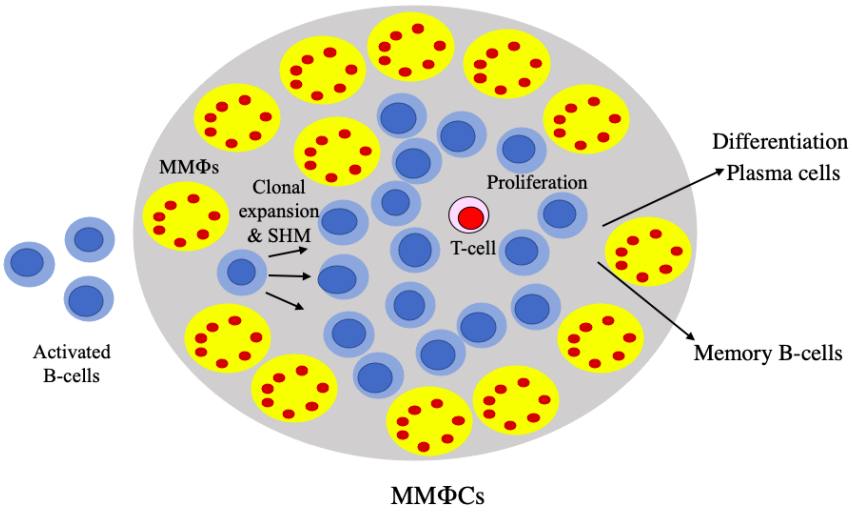
Recent work done in our lab using vaccinated goldfish has established that MMΦs are the cells that retain the antigen within MMΦCs, and the antigen on the cell surface appeared to be in an intact form (Muthupandian MSc thesis, U of A, 2020). In addition, several genes which could potentially aid in FDC-like functions, such as BAFF, FcR-like isoform, and CR-like isoform, were found to be expressed by MMΦs using RNA-sequencing of goldfish kidney MMΦs (Muthupandian MSc thesis, U of A, 2020). BAFF is produced by FDCs, and it is crucial for maintaining B-cells homeostasis and enhancing their survival within the GCs (Lesley et al., 2004, reviewed in Allen et al., 2008). In addition, FDCs express complement and Fc receptors such as CR1, CR2, and FcγRIIB; these receptors are required for trapping the unprocessed antigen as an immune complex on the surface of FDCs (reviewed in Allen et al., 2008).

Previous studies revealed that antibody affinity maturation process occurs in bony fish (Ye et al., 2011, Wu et al., 2019). Using rainbow trout and channel catfish, it was shown that following immunization with a T-cell dependent antigen (TNP-KLH), a shift towards the production of

higher affinity antibodies occurs in these fish a few weeks post-immunization (Wu et al., 2019; Ye et al., 2011). However, the increase in antibody affinity found in fish (100-fold increase) is considerably modest compared to mammals (1000-fold increase). Interestingly, a unique feature of bony fish IgM is that the number of interchain disulfide is determined by the affinity of the BCR for a specific antigen, where antibodies with a higher affinity for the antigen have the highest degree of polymerization (reviewed in Ye et al., 2011). My analysis of the Ig repertoires prepared using MMΦCs from unvaccinated and vaccinated fish revealed that clonal lineages of B-cells accumulate new mutations within the VDJ exon, and the analysis of the distribution of mutations indicated the presence of an active antigen-driven selection process. While we have not verified if the observed mutations lead to an increase in the antibody binding affinity to the foreign antigens, my results suggest that AID-induced SHM is essential to increase the binding affinity. The higher degree of polymerization found in the antibodies with higher binding affinity could also contribute to the overall strength of the antibody-antigen interaction.

Collectively the data presented in this thesis and the finding that MMΦs are the cells that retain the antigen within MMΦCs suggest a model of affinity maturation in fish where activated B and Tfh -cells are recruited to MMΦCs, perhaps using CXCL13 expressed by MMΦs (Waly et al., in preparation). Within the clusters, B-cells rapidly proliferate while acquiring AID and error-prone polymerases mediated mutations in their VDJ exon. In this model, B-cells with mutated Ig receptors compete for intact antigens on the surface of MMΦs. However, unlike the mammalian germinal center, where B-cells represent around 35% to 41% of the total number of cells (Kuppers et al., 1993), my results revealed that lymphocyte-like cells represent between 13% to 19% of the total number of cells within MMΦCs. Moreover, earlier studies revealed that there is a relatively

high number of intact antigens available on the surface of MMΦs within these clusters compared to the number of trapped antigens on the surface of FDCs in the germinal centers (reviewed in Muthupandian et al., 2021). Therefore, it appears that within MMΦCs, any B-cells that can bind the antigen will internalize and present it to Tfh-cells from which they receive a survival signal (positive selection); subsequently, selected cells will undergo further proliferation and eventually differentiate into plasma cells and memory B-cells (Figure 5.1; reviewed in Muthupandian et al., 2021). Furthermore, MMΦs are phagocytic cells, and thus they could phagocytose apoptotic cells that fail to bind the intact antigen within these clusters similar to tangible body macrophages (TBM), which are found within the germinal centers (reviewed in Steinel et al., 2017). This model explains the modest increase in antibody affinity in fish (100-fold increase) when compared to mammals (1000-fold increase) following immunization with T-cell dependent antigens (Ye et al., 2011, Wu et al., 2019).



	GCs	MMΦCs
B-cells	More	Less
Antigen	Less	More
Affinity	1000-fold	100-fold

Figure 5.1. Model for antibody affinity maturation in fish. Activated B and T-cells are recruited to MMΦCs. Within the clusters, B-cells rapidly proliferate while acquiring AID and error-prone polymerases mediated mutations in their VDJ exon. B-cells with mutated Ig receptors compete for intact antigens on the surface of MMΦs. Within MMΦCs, any B-cells that can bind the antigen will internalize and present it to Tfh-cells from which they receive a survival signal (positive selection); subsequently, selected cells will undergo further proliferation and eventually differentiate into plasma cells and memory B-cells. MMΦs are phagocytic cells, and thus they could phagocytose apoptotic cells that fail to bind the intact antigen within these clusters.

CHAPTER VI

SIGNIFICANCE AND FUTURE DIRECTIONS

6. Significance and future direction

6.1. Significance

The work presented in this thesis is the first to directly test the hypothesis that MMΦCs are functional analogous of homeotherm germinal centers. The findings in my thesis support that MMΦCs in fish are acting as primordial germinal centers and verified their role in the adaptive immune system, which indicate that MMΦCs can provide a valuable tool for comparative studies of the adaptive immunity in lower vertebrates. Furthermore, the finding that MMΦCs are the sites where affinity modification process occurs in fish will allow further studies into the different cell types involved in this process. In addition, this work will help develop comprehensive models of how antibody affinity maturation evolved and functions in fish, which can aid in the development of more effective vaccines for aquaculture and fish husbandry industries. Overall, this work will provide a better understanding of the evolutionary history of the adaptive immune system. Moreover, the findings of my thesis may provide an insight into processes occurring in ectopic GCs that form in the absence of FDCs in the lymphoid tissues of patients with certain auto-immune or pro-inflammatory diseases (William et al., 2002).

6.2. Future directions

Using Ig repertoires isolated from unvaccinated and vaccinated fish MMΦCs, we revealed that B-cell clonotypes proliferate within the clusters and accumulate AID and error-prone polymerases induced mutations. In addition, I found evidence for Ag-driven selection process within MMΦCs. However, we have not verified if the observed mutations lead to affinity maturation and increase the antibodies' binding affinity to the foreign antigens. To determine if the observed mutations lead to antibody affinity maturation, the binding affinity of antibodies isolated from MMΦCs should be examined. Ig repertoires from vaccinated fish MMΦCs were isolated 30 or 40 days after the first injection; therefore, MMΦCs from vaccinated fish should be isolated 14 days following vaccination and placed in a small volume of culture media for 14-20 days, then the supernatant from these cultures can be collected to measure the binding affinity. Biacore system, which uses surface plasmon resonance (SPR), can be used to measure the antibodies binding affinity, where the target antigen can be attached to a thin layer of gold (Murphy et al., 2006).

REFERENCES

References

Abos, B., Estensoro, I., Perdiguero, P., Faber, M., Hu, Y., Díaz Rosales, P., Granja, A.G., Secombes, C.J., Holland, J.W. & Tafalla, C. (2018). Dysregulation of B cell activity during proliferative kidney disease in rainbow trout. *Frontiers in immunology*, 9, 1203.

Abos, B., Wang, T., Secombes, C. J., & Tafalla, C. (2020). Distinct modes of action of CD40L and adaptive cytokines IL-2, IL-4/13, IL-10 and IL-21 on rainbow trout IgM+ B cells. *Developmental & Comparative Immunology*, 111, 103752.

Agius, C. (1979). The role of melano-macrophage centres in iron storage in normal and diseased fish. *Journal of Fish Diseases*, 2(4), 337-343.

Agius, C. (1980). Phylogenetic development of melano–macrophage centres in fish. *Journal of zoology*, 191(1), 11-31.

Agius, C. (1981). Preliminary studies on the ontogeny of the melano-macrophages of teleost haemopoietic tissues and age-related changes. *Developmental & comparative immunology*, 5(4), 597-606.

Agius, C., & Agbede, S. A. (1984). An electron microscopical study on the genesis of lipofuscin, melanin and haemosiderin in the haemopoietic tissues of fish. *Journal of fish biology*, 24(4), 471-488.

Agius, C., & Roberts, R. J. (2003). Melano-macrophage centres and their role in fish pathology. *Journal of fish diseases*, 26(9), 499-509.

Akira, S., Uematsu, S., & Takeuchi, O. (2006). Pathogen recognition and innate immunity. *Cell*, 124(4), 783-801.

Alamyar, E., Giudicelli, V., Duroux, P., & Lefranc, M. P. (2010). IMGT/HighV-QUEST: A high-throughput system and Web portal for the analysis of rearranged nucleotide sequences of antigen receptors-High-throughput version of IMGT/V-QUEST. *V-QUEST 11èmes Journées Ouvertes en Biologie, Informatique et Mathématiques (JOBIM)*, 7-9.

Alder, M. N., Herrin, B. R., Sadlonova, A., Stockard, C. R., Grizzle, W. E., Gartland, L. A., ... & Cooper, M. D. (2008). Antibody responses of variable lymphocyte receptors in the lamprey. *Nature immunology*, 9(3), 319-327.

Allen, C. D. C., & Cyster, J. G. (2008). Follicular dendritic cell networks of primary follicles and germinal centers*:* Phenotype and function. *Seminars in immunology* (Vol. 20, No. 1, pp. 14-25). Academic Press.

Allen, C. D., Okada, T., & Cyster, J. G. (2007). Germinal-center organization and cellular dynamics. *Immunity*, 27(2), 190-202.

Allman, D., & Pillai, S. (2008). Peripheral B cell subsets. *Current opinion in immunology*, 20(2), 149-157.

Allman, D., Wilmore, J. R., & Gaudette, B. T. (2019). The continuing story of T-cell independent antibodies. *Immunological reviews*, 288(1), 128-135.

Alt, F. W., & Baltimore, D. (1982). Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-JH fusions. *Proceedings of the National Academy of Sciences*, 79(13), 4118-4122.

Amarante-Mendes, G. P., Adjemian, S., Branco, L. M., Zanetti, L. C., Weinlich, R., & Bortoluci, K. R. (2018). Pattern recognition receptors and the host cell death molecular machinery. *Frontiers in immunology*, 2379.

Arakawa, H., Hauschild, J., & Buerstedde, J. M. (2002). Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. *Science*, 295(5558), 1301-1306.

Arkoosh, M. R., & Kaattari, S. L. (1991). Development of immunological memory in rainbow trout (Oncorhynchus mykiss). I. An immunochemical and cellular analysis of the B cell response. *Developmental & comparative immunology*, 15(4), 279-293.

Avtalion, R. R. (1969). Temperature effect on antibody production and immunological memory, in carp (cyprinus carpio) immunized against bovine serum albumin (BSA). *Immunology*, 17(6), 927-931.

Bahjat, M., & Guikema, J. E. (2017). The complex interplay between DNA injury and repair in enzymatically induced mutagenesis and DNA damage in B lymphocytes. *International Journal of Molecular Sciences*, 18(9), 1876.

Bardwell, P.D., Woo, C.J., Wei, K., Li, Z., Martin, A., Sack, S.Z., Parris, T., Edelmann, W. & Scharff, M.D. (2004). Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. *Nature Immunology, 5*(2), 224-229.

Barr, M., Mott, K., & Zwollo, P. (2011). Defining terminally differentiating B cell populations in rainbow trout immune tissues using the transcription factor XbpI. *Fish and Shellfish Immunology,* 31(6), 727-735.

Barraza, F., Montero, R., Wong-Benito, V., Valenzuela, H., Godoy-Guzmán, C., Guzmán, F., Köllner, B., Wang, T., Secombes, C.J., Maisey, K. & Imarai, M. (2021). Revisiting the teleost thymus: current knowledge and future perspectives. *Biology*, 10(1), 8.

Barreto, V. M., & Magor, B. G. (2011). Activation-induced cytidine deaminase structure and functions: A species comparative view. *Developmental and Comparative Immunology*, 35(9), 991-1007.

Batista, F. D., & Harwood, N. E. (2009). The who, how and where of antigen presentation to B cells. *Nature* Reviews *Immunology*, 9(1), 15-27.

Baumgarth, N. (2011). The double life of a B-1 cell: Self-reactivity selects for protective effector functions. *Nature Reviews Immunology,* 11(1), 34-46.

Bengten, E., Quiniou, S., Hikima, J., Waldbieser, G., Warr, G. W., Miller, N. W., & Wilson, M. (2006). Structure of the catfish IGH locus: Analysis of the region including the single functional IGHM gene. *Immunogenetics*, 58(10), 831-844.

Bermudez, R., Vigliano, F., Marcaccini, A., Sitja-Bobadilla, A., Quiroga, M. I., & Nieto, J. M. (2006). Response of ig-positive cells to enteromyxum scophthalmi (myxozoa) experimental infection in turbot, scophthalmus maximus (L.): A histopathological and immunohistochemical study. *Fish & Shellfish Immunology*, 21(5), 501-512.

Bernstein, R. M., Schluter, S. F., Bernstein, H., & Marchalonis, J. J. (1996). Primordial emergence of the recombination activating gene 1 (RAG1): sequence of the complete shark gene indicates homology to microbial integrases. *Proceedings of the National Academy of Sciences*, 93(18), 9454-9459.

Bilal, S., Etayo, A., & Hordvik, I. (2021). Immunoglobulins in teleosts. *Immunogenetics*, 73(1), 65-77.

Bilal, S., Lie, K. K., Sæle, Ø., & Hordvik, I. (2018). T cell receptor alpha chain genes in the teleost ballan wrasse (Labrus bergylta) are subjected to somatic hypermutation. *Frontiers in immunology*, 9, 1101.

Biram, A., Davidzohn, N., & Shulman, Z. (2019). T cell interactions with B cells during germinal center formation, a three-step model. *Immunological reviews*, 288(1), 37-48.

Boboila, C., Jankovic, M., Yan, C.T., Wang, J.H., Wesemann, D.R., Zhang, T., Fazeli, A., Feldman, L., Nussenzweig, A., Nussenzweig, M. & Alt, F.W. (2010). Alternative end-joining catalyzes robust IgH locus deletions and translocations in the combined absence of ligase 4 and Ku70. *Proceedings of the National Academy of Sciences of the United States of America,* 107(7), 3034-3039.

Borchert, G. M., Holton, N. W., & Larson, E. D. (2011). Repression of human activation induced cytidine deaminase by miR-93 and miR-155. *BMC Cancer*, 11, 347-347.

Bowers, P. M., Verdino, P., Wang, Z., da Silva Correia, J., Chhoa, M., Macondray, G., ... & King, D. J. (2014). Nucleotide insertions and deletions complement point mutations to massively expand the diversity created by somatic hypermutation of antibodies. *Journal of Biological Chemistry*, 289(48), 33557-33567.

Bransteitter, R., Pham, P., Calabrese, P., & Goodman, M. F. (2004). Biochemical analysis of hypermutational targeting by wild type and mutant activation-induced cytidine deaminase. *The Journal of Biological Chemistry*, 279(49), 51612-51621.

Brattgjerd, S., & Evensen, Ø. (1996). A sequential light microscopic and ultrastructural study on the uptake and handling of vibrio salmonicida in phagocytes of the head kidney in experimentally infected atlantic salmon (salmo salar L.). *Veterinary pathology*, 33(1), 55-65.

Briney, B. S., & Crowe Jr, J. E. (2013). Secondary mechanisms of diversification in the human antibody repertoire. *Frontiers in immunology*, 4, 42.

Bromage, E. S., Kaattari, I. M., Zwollo, P., & Kaattari, S. L. (2004). Plasmablast and plasma cell production and distribution in trout immune tissues. *The Journal of Immunology,* 173(12), 7317-7323.

Brown, S.P., Veach, A.M., Rigdon-Huss, A.R., Grond, K., Lickteig, S.K., Lothamer, K., Oliver, A.K. and Jumpponen, A. (2015). Scraping the bottom of the barrel: are rare high throughput sequences artifacts?. *fungal ecology*, 13, 221-225.

Bruce, T. J., & Brown, M. L. (2017). A review of immune system components, cytokines, and immunostimulants in cultured finfish species. *Open Journal of Animal Sciences,* 07(03), 267-288.

Buerstedde, J. M., Alinikula, J., Arakawa, H., McDonald, J. J., & Schatz, D. G. (2014). Targeting of somatic hypermutation by immunoglobulin enhancer and enhancer-like sequences. *PLoS biology*, 12(4), e1001831.

Buonocore, F., & Gerdol, M. (2016). Alternative adaptive immunity strategies: Coelacanth, cod and shark immunity. *Molecular Immunology,* 69, 157-169.

Cain, K. D., Jones, D. R., & Raison, R. L. (2002). Antibody-antigen kinetics following immunization of rainbow trout (oncorhynchus mykiss) with a T-cell dependent antigen. *Developmental and Comparative Immunology,* 26(2), 181-190.

Cantaert, T., Schickel, J. N., Bannock, J. M., Ng, Y. S., Massad, C., Oe, T., . . . Meffre, E. (2015). Activation-induced cytidine deaminase expression in human B cell precursors is essential for central B cell tolerance. *Immunity, 43*(5), 884-895.

Cariappa, A., Mazo, I. B., Chase, C., Shi, H. N., Liu, H., Li, Q., . . . Pillai, S. (2005). Perisinusoidal B cells in the bone marrow participate in T-independent responses to blood-borne microbes. *Immunity, 23*(4), 397-407.

Casellas, R., Nussenzweig, A., Wuerffel, R., Pelanda, R., Reichlin, A., Suh, H., . . . Nussenzweig, M. C. (1998). Ku80 is required for immunoglobulin isotype switching. *The EMBO journal*, 17(8), 2404-2411.

Castro, C. D., & Flajnik, M. F. (2014). Putting J chain back on the map: how might its expression define plasma cell development?. *The Journal of Immunology*, 193(7), 3248-3255.

Castro, R., Jouneau, L., Pham, H. P., Bouchez, O., Giudicelli, V., Lefranc, M. P., ... & Boudinot, P. (2013). Teleost fish mount complex clonal IgM and IgT responses in spleen upon systemic viral infection. *PLoS pathogens*, 9(1), e1003098.

Cerutti, A., Cols, M., & Puga, I. (2013). Marginal zone B cells: virtues of innate-like antibody-producing lymphocytes. *Nature Reviews Immunology*, 13(2), 118-132.

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological monographs*, 84(1), 45-67.

Chaplin, D. D. (2010). Overview of the immune response. *Journal of allergy and clinical immunology*, 125(2), S3-S23.

Chaudhary, N., & Wesemann, D. R. (2018). Analyzing immunoglobulin repertoires. *Frontiers in immunology*, 9, 462.

Chen, K., & Cerutti, A. (2011). The function and regulation of immunoglobulin D. *Current opinion in immunology*, 23(3), 345-352.

Chirumbolo, S., Bjørklund, G., Sboarina, A., & Vella, A. (2018). The role of basophils as innate immune regulatory cells in allergy and immunotherapy. *Human vaccines & immunotherapeutics*, 14(4), 815-831.

Choi, J. H., Wang, K. W., Zhang, D., Zhan, X., Wang, T., Bu, C. H., . . . & Beutler, B. (2017). IgD class switching is initiated by microbiota and limited to mucosa-associated lymphoid tissue in mice. *Proceedings of the National Academy of Sciences of the United States of America, 114*(7), E1196-E1204.

Choudhary, M., Tamrakar, A., Singh, A. K., Jain, M., Jaiswal, A., & Kodgire, P. (2018). AID biology: A pathological and clinical perspective. *International reviews of immunology*, 37(1), 37-56.

Conticello, S. G., Thomas, C. J., Petersen-Mahrt, S. K., & Neuberger, M. S. (2005). Evolution of the AID/APOBEC family of polynucleotide (deoxy) cytidine deaminases. Molecular biology and evolution, 22(2), 367-377.

Corthesy, B. (2013). Multi-faceted functions of secretory IgA at mucosal surfaces. *Frontiers in Immunology*, 4, 185.

Cortina-Ceballos, B., Godoy-Lozano, E. E., Sámano-Sánchez, H., Aguilar-Salgado, A., Velasco-Herrera, M. D. C., Vargas-Chávez, C., ... & Martínez-Barnetche, J. (2015, May). Reconstructing and mining the B cell repertoire with ImmunediveRsity. *In MAbs* (Vol. 7, No. 3, pp. 516-524). Taylor & Francis.

Cossarini-Dunier, M. (1986). Secondary response of rainbow trout (Salmo gairdneri Richardson) to DNP-haemocyanin and Yersinia ruckeri. *Aquaculture*, 52(2), 81-86.

Costa, G. R., Poirier, A., & Bromage, E. (2021). A comparison of the IgM and IgT repertoires reveals the structure of B cell communities in rainbow trout is a consequence of mechanistic and evolutionary forces. *bioRxiv*.

Criscitiello, M. F., & Flajnik, M. F. (2007). Four primordial immunoglobulin light chain isotypes, including λ and κ, identified in the most primitive living jawed vertebrates. *European journal of immunology*, 37(10), 2683-2694.

Cyster, J. G., & Allen, C. D. C. (2019). B cell responses: Cell interaction dynamics and decisions. *Cell,* 177(3), 524-540.

Dancyger, A. M., King, J. J., Quinlan, M. J., Fifield, H., Tucker, S., Saunders, H. L., . . . & Larijani, M. (2012). Differences in the enzymatic efficiency of human and bony fish AID are mediated by a single residue in the C terminus modulating single-stranded DNA binding. *The FASEB Journal,* 26(4), 1517-1525.

Danilova, N., Bussmann, J., Jekosch, K., & Steiner, L. A. (2005). The immunoglobulin heavy-chain locus in zebrafish: Identification and expression of a previously unknown isotype, immunoglobulin Z. *Nature Immunology,* 6(3), 295-302.

Davies, L. C., Jenkins, S. J., Allen, J. E., & Taylor, P. R. (2013). Tissue-resident macrophages. *Nature Immunology*, 14(10), 986-995.

De Silva, N. S., & Klein, U. (2015). Dynamics of B cells in germinal centres. *Nature Reviews Immunology, 15*(3), 137-148.

de Wildt, R. M., van Venrooij, W. J., Winter, G., Hoet, R. M., & Tomlinson, I. M. (1999). Somatic insertions and deletions shape the human antibody repertoire. *Journal of molecular biology*, 294(3), 701-710.

Dezfuli, B. S., Giovinazzo, G., Lui, A., & Giari, L. (2008). Inflammatory response to dentitruncus truttae (acanthocephala) in the intestine of brown trout. *Fish and Shellfish Immunology, 24*(6), 726-733.

Di Noia, J. M. D., & Neuberger, M. S. (2004). Immunoglobulin gene conversion in chicken DT40 cells largely proceeds through an abasic site intermediate generated by excision of the uracil produced by AID-mediated deoxycytidine deamination. *European Journal of Immunology*, 34(2), 504–508.

Di Noia, J. M., & Neuberger, M. S. (2007). Molecular mechanisms of antibody somatic hypermutation *Annu. Rev. Biochem*., 76, 1-22.

Di Noia, J. M., Williams, G. T., Chan, D. T., Buerstedde, J. M., Baldwin, G. S., & Neuberger, M. S. (2007). Dependence of antibody gene diversification on uracil excision. *The Journal of Experimental Medicine, 204*(13), 3209-3219.

Di Noia, J., & Neuberger, M. S. (2002). Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature*, 419(6902), 43-48.

Diaz-Satizabal, L., & Magor, B. G. (2015). Isolation and cytochemical characterization of melanomacrophages and melanomacrophage clusters from goldfish (carassius auratus, L.). *Developmental and Comparative Immunology, 48*(1), 221-228.

Diaz, M., Greenberg, A. S., & Flajnik, M. F. (1998). Somatic hypermutation of the new antigen receptor gene (NAR) in the nurse shark does not generate the repertoire: possible role in antigen-driven reactions in the absence of germinal centers. *Proceedings of the National Academy of Sciences*, 95(24), 14343-14348.

Diaz, M., Velez, J., Singh, M., Cerny, J., & Flajnik, M. F. (1999). Mutational pattern of the nurse shark antigen receptor gene (NAR) is similar to that of mammalian Ig genes and to spontaneous mutations in evolution: the translesion synthesis model of somatic hypermutation. International Immunology, 11(5), 825-833.

Dominguez, P. M., & Shaknovich, R. (2014). Epigenetic function of activation-induced cytidine deaminase and its link to lymphomagenesis. *Frontiers in Immunology*, 5, 642.

Dong, J., Finn, J. A., Larsen, P. A., Smith, T. P., & Crowe Jr, J. E. (2019). Structural diversity of ultralong CDRH3s in seven bovine antibody heavy chains. *Frontiers in immunology*, 10, 558.

Dornburg, A., Ota, T., Criscitiello, M. F., Salinas, I., Sunyer, J. O., Magadán, S., ... & Yoder, J. A. (2021). From IgZ to IgT: A Call for a Common Nomenclature for Immunoglobulin Heavy Chain Genes of Ray-Finned Fish. *Zebrafish*, 18(6), 343-345.

Dreyer, W. J., & Bennett, J. C. (1965). The molecular basis of antibody formation: A paradox. *Proceedings of the National Academy of Sciences of the United States of America*, 54(3), 864-869.

Eberl, G., Colonna, M., Santo, J. P. D., & McKenzie, A. N. J. (2015). Innate lymphoid cells: A new paradigm in immunology. *Science*, 348(6237).

Edholm, E., Bengtén, E., Stafford, J. L., Sahoo, M., Taylor, E. B., Miller, N. W., & Wilson, M. (2010). Identification of two IgD + B cell populations in channel catfish, ictalurus punctatus. *The Journal of Immunology,* 185(7), 4082-4094.

Ehlich, A., Martin, V., Müller, W., & Rajewsky, K. (1994). Analysis of the B-cell progenitor compartment at the level of single cells. *Current Biology*, 4(7), 573-583.

Elsner, R. A., & Shlomchik, M. J. (2020). Germinal center and extrafollicular B cell responses in vaccination, immunity, and autoimmunity. *Immunity*, 53(6), 1136-1150.

Ewing, B., & Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome research*, 8(3), 186-194.

Faili, A., Stary, A., Delbos, F., Weller, S., Aoufouchi, S., Sarasin, A., ... & Reynaud, C. A. (2009). A backup role of DNA polymerase κ in Ig gene hypermutation only takes place in the complete absence of DNA polymerase η. *The Journal of Immunology*, 182(10), 6353-6359.

Falk, K., Press, C. M., Landsverk, T., & Dannevig, B. H. (1995). Spleen and kidney of atlantic salmon (salmo salar L.) show histochemical changes early in the course of experimentally induced infectious salmon anaemia (ISA). *Veterinary immunology and immunopathology*, 49(1-2), 115-126.

Feederle, R., & Schepers, A. (2017). Antibodies specific for nucleic acid modifications. *RNA biology*, 14(9), 1089-1098.

Feeney, A. J. (1990). Lack of N regions in fetal and neonatal mouse immunoglobulin VDJ junctional sequences. *The Journal of experimental medicine*, 172(5), 1377-1390.

Feng, Y., Seija, N., D:i Noia, J. M., & Martin, A. (2020). AID in antibody diversification: There and back again. *Trends in immunology*, 41(7), 586-600.

Ferguson, H. W. (1976). The relationship between ellipsoids and melano-macrophage centres in the spleen of turbot. *Journal of Comparative Pathology*, 86(3), 377-380.

Fillatreau, S., Six, A., Magadan, S., Castro, R., Sunyer, J. O., & Boudinot, P. (2013). The astonishing diversity of Ig classes and B cell repertoires in teleost fish. *Frontiers in immunology*, 4, 28.

Firl, D. J., Degn, S. E., Padera, T., & Carroll, M. C. (2018). Capturing change in clonal composition amongst single mouse germinal centers. *Elife*, 7, e33051.

Flajnik, M. F. (2018). A cold-blooded view of adaptive immunity. *Nature Reviews.Immunology, 18*(7), 438-453.

Fujii, T., Nakagawa, H., & Murakawa, S. (1979). Immunity in lamprey I. Production of haemolytic and haemagglutinating antibody to sheep red blood cells in Japanese lampreys. *Developmental & Comparative Immunology*, 3, 441-451.

Garcia-Diaz, M., & Kunkel, T. A. (2006). Mechanism of a genetic glissando*: structural biology of indel mutations. *Trends in biochemical sciences*, 31(4), 206-214.

Garraud, O., Borhis, G., Badr, G., Degrelle, S., Pozzetto, B., Cognasse, F., & Richard, Y. (2012). Revisiting the B-cell compartment in mouse and humans: more than one B-cell subset exists in the marginal zone and beyond. *BMC immunology*, 13(1), 1-17.

Garside, P., Ingulli, E., Merica, R. R., Johnson, J. G., Noelle, R. J., & Jenkins, M. K. (1998). Visualization of specific B and T lymphocyte interactions in the lymph node. *Science,* 281(5373), 96-99.

Gasteiger, G., D'osualdo, A., Schubert, D. A., Weber, A., Bruscia, E. M., & Hartl, D. (2017). Cellular innate immunity: an old game with new players. *Journal of innate immunity*, 9(2), 111-125.

Gatto, D., & Brink, R. (2010). The germinal center reaction. *Journal of Allergy and Clinical Immunology*, 126(5), 898-907.

Ghaffari, S. H., & Lobb, C. J. (1997). Structure and genomic organization of a second class of immunoglobulin light chain genes in the channel catfish. *The Journal of Immunology*, 159(1), 250-258.

Gitlin, A. D., Shulman, Z., & Nussenzweig, M. C. (2014). Clonal selection in the germinal centre by regulated proliferation and hypermutation. *Nature,* 509(7502), 637-640.

Goossens, T., Klein, U., & Küppers, R. (1998). Frequent occurrence of deletions and duplications during somatic hypermutation: implications for oncogene translocations and heavy chain disease. *Proceedings of the National Academy of Sciences*, 95(5), 2463-2468.

Gorgoglione, B., Wang, T., Secombes, C. J., & Holland, J. W. (2013). Immune gene expression profiling of proliferative kidney disease in rainbow trout Oncorhynchus mykiss reveals a dominance of anti-inflammatory, antibody and T helper cell-like activities. *Veterinary research*, 44(1), 1-16.

Granja, A. G., Perdiguero, P., Martín-Martín, A., Díaz-Rosales, P., Soleto, I., & Tafalla, C. (2019). Rainbow trout IgM+ B cells preferentially respond to thymus-independent antigens but are activated by CD40L. *Frontiers in immunology*, 10, 2902.

Guo, P., Hirano, M., Herrin, B. R., Li, J., Yu, C., Sadlonova, A., & Cooper, M. D. (2009). Dual nature of the adaptive immune system in lampreys. *Nature*, 459(7248), 796-801.

Gupta, N. T., Adams, K. D., Briggs, A. W., Timberlake, S. C., Vigneault, F., & Kleinstein, S. H. (2017). Hierarchical clustering can identify B cell clones with high confidence in ig repertoire sequencing data. *The Journal of Immunology,* 198(6), 2489-2499.

Gupta, N. T., Vander Heiden, J. A., Uduman, M., Gadala-Maria, D., Yaari, G., & Kleinstein, S. H. (2015). Change-O: A toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics,* 31(20), 3356-3358.

Han, L., & Yu, K. (2008). Altered kinetics of nonhomologous end joining and class switch recombination in ligase IV-deficient B cells. *Journal of Experimental Medicine,* 205(12), 2745-2753.

Hanayama, R., Tanaka, M., Miwa, K., Shinohara, A., Iwamatsu, A., & Nagata, S. (2002). Identification of a factor that links apoptotic cells to phagocytes. *Nature*, 417(6885), 182-187.

Hansen, J. D., Landis, E. D., & Phillips, R. B. (2005). Discovery of a unique ig heavy-chain isotype (IgT) in rainbow trout: Implications for a distinctive B cell developmental pathway in teleost fish. *Proceedings of the National Academy of Sciences of the United States of America*, 102(19), 6919-6924.

Harris, R. S., Sale, J. E., Petersen-Mahrt, S. K., & Neuberger, M. S. (2002). AID is essential for immunoglobulin V gene conversion in a cultured B cell line. *Current Biology*: CB, 12(5), 435-438.

Hayakawa, K. Y. O. K. O., Hardy, R. R., Herzenberg, L. A., & Herzenberg, L. A. (1985). Progenitors for Ly-1 B cells are distinct from progenitors for other B cells. *The Journal of experimental medicine*, 161(6), 1554-1568.

Helleday, T. (2003). Pathways for mitotic homologous recombination in mammalian cells. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 532(1-2), 103-115.

Helmink, B. A., & Sleckman, B. P. (2012). The response to and repair of RAG-mediated DNA double-strand breaks. *Annual review of immunology*, 30, 175-202.

Herraez, M. P., & Zapata, A. G. (1986). Structure and function of the melano-macrophage centres of the goldfishCarassius auratus. *Veterinary Immunology and Immunopathology*, 12(1-4), 117-126.

Hikima, J. I., Jung, T. S., & Aoki, T. (2011). Immunoglobulin genes and their transcriptional control in teleosts. *Developmental & Comparative Immunology*, 35(9), 924-936.

Hildemann, W. H., & Thoenes, G. H. (1969). IMMUNOLOGICAL RESPONSES OF PACIFIC HAGFISH: I. Skin Transplantation Immunity: 1. *Transplantation*, 7(6), 506-521.

Hill M. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54(2):427-32.

Hinds-Frey, K. R., Nishikata, H., Litman, R. T., & Litman, G. W. (1993). Somatic variation precedes extensive diversification of germline sequences and combinatorial joining in the evolution of immunoglobulin heavy chain diversity. *The Journal of experimental medicine*, 178(3), 815-824.

Hoare, R., Ngo, T. P., Bartie, K. L., & Adams, A. (2017). Efficacy of a polyvalent immersion vaccine against Flavobacterium psychrophilum and evaluation of immune response to vaccination in rainbow trout fry (Onchorynchus mykiss L.). *Veterinary research*, 48(1), 1-13.

Hoehn, K. B., Vander Heiden, J. A., Zhou, J. Q., Lunter, G., Pybus, O. G., & Kleinstein, S. H. (2019). Repertoire-wide phylogenetic models of B cell molecular evolution reveal evolutionary signatures of aging and vaccination. *Proceedings of the National Academy of Sciences*, 116(45), 22664-22672.

Hoffman, W., Lakkis, F. G., & Chalasani, G. (2016). B cells, antibodies, and more. *Clinical Journal of the American Society of Nephrology, 11*(1), 137-154.

Hordvik, I. (1998). The impact of ancestral tetraploidy on antibody heterogeneity in salmonid fishes. *Immunological Reviews*, 166, 153-157.

Horns, F., Vollmers, C., Dekker, C. L., & Quake, S. R. (2019). Signatures of selection in the human antibody repertoire: Selective sweeps, competing subclones, and neutral drift. *Proceedings of the National Academy of Sciences of the United States of America, 116*(4), 1261-1266.

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods in Ecology and Evolution*, 7(12), 1451-1456.

Hsu, E. (1998). Mutation, selection, and memory in B lymphocytes of exothermic vertebrates. *Immunological reviews*, 162(1), 25-36.

Imai, K., Zhu, Y., Revy, P., Morio, T., Mizutani, S., Fischer, A., . . . Durandy, A. (2005). Analysis of class switch recombination and somatic hypermutation in patients affected with autosomal dominant hyper-IgM syndrome type 2. *Clinical Immunology* (Orlando, Fla.), 115(3), 277-285.

Inamine, A., Takahashi, Y., Baba, N., Miyake, K., Tokuhisa, T., Takemori, T., & Abe, R. (2005). Two waves of memory B-cell generation in the primary immune response. *International Immunology,* 17(5), 581-589.

Ippolito, G. C., Hoi, K. H., Reddy, S. T., Carroll, S. M., Ge, X., Rogosch, T., ... & Georgiou, G. (2012). Antibody repertoires in humanized NOD-scid-IL2Rγnull mice and human B cells reveals human-like diversification and tolerance checkpoints in the mouse. *PloS one*, 7(4), e35497.

Isakov, N. (2022). Histocompatibility and Reproduction: Lessons from the Anglerfish. Life, 12(1), 113.

Ise, W., & Kurosaki, T. (2019). Plasma cell differentiation during the germinal center reaction. Immunological Reviews, 288(1), 64-74.

Ise, W., Fujii, K., Shiroguchi, K., Ito, A., Kometani, K., Takeda, K., . . . Kurosaki, T. (2018). T follicular helper cell-germinal center B cell interaction strength regulates entry into plasma cell or recycling germinal center cell fate. *Immunity, 48*(4), 702-715.e4.

Ise, W., Kohyama, M., Schraml, B. U., Zhang, T., Schwer, B., Basu, U., . . . Murphy, K. M. (2011). The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. *Nature Immunology, 12*(6), 536-543.

Ito, S., Nagaoka, H., Shinkura, R., Begum, N., Muramatsu, M., Nakata, M., & Honjo, T. (2004). Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proceedings of the National Academy of Sciences*, 101(7), 1975-1980.

Jackson, K. J., Kidd, M. J., Wang, Y., & Collins, A. M. (2013). The shape of the lymphocyte receptor repertoire: lessons from the B cell receptor. *Frontiers in immunology*, 4, 263.

Jacob, J., Przylepa, J., Miller, C., & Kelsoe, G. (1993). In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl) acetyl. III. The kinetics of V region mutation and selection in germinal center B cells. *The Journal of experimental medicine*, 178(4), 1293-1307.

Jenne, C. N., Kennedy, L. J., & Reynolds, J. D. (2006). Antibody repertoire development in the sheep. *Developmental & Comparative Immunology*, 30(1-2), 165-174.

Ji, J. F., Hu, C. B., Shao, T., Fan, D. D., Zhang, N., Lin, A. F., ... & Shao, J. Z. (2020). Differential immune responses of immunoglobulin Z subclass members in antibacterial immunity in a zebrafish model. *Immunology*, 162(1), 105-120.

Jiang, N., Weinstein, J. A., Penland, L., White, R. A., Fisher, D. S., & Quake, S. R. (2011). Determinism and stochasticity during maturation of the zebrafish antibody repertoire. *Proceedings of the National Academy of Sciences of the United States of America,* 108(13), 5348-5353.

Jolly, C. J., Wagner, S. D., Rada, C., Klix, N., Milstein, C., & Neuberger, M. S. (1996). The targeting of somatic hypermutation. *In Seminars in immunology* (Vol. 8, No. 3, pp. 159-168).

Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S., & Winter, G. (1986). Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature*, 321(6069), 522-525.

Kaattari, S. L., Zhang, H. L., Khor, I. W., Kaattari, I. M., & Shapiro, D. A. (2002). Affinity maturation in trout: Clonal dominance of high affinity antibodies late in the immune response. *Developmental and Comparative Immunology, 26*(2), 191-200.

Kannouche, P. L., Wing, J., & Lehmann, A. R. (2004). Interaction of human DNA polymerase η with monoubiquitinated PCNA: a possible mechanism for the polymerase switch in response to DNA damage. *Molecular cell*, 14(4), 491-500.

Kawabe, T., Naka, T., Yoshida, K., Tanaka, T., Fujiwara, H., Suematsu, S., . . . Kikutani, H. (1994). The immune responses in CD40-deficient mice: Impaired immunoglobulin class switching and germinal center formation. *Immunity*, 1(3), 167-178.

Kepler, T. B. (1997). Codon bias and plasticity in immunoglobulins. *Molecular Biology and Evolution*, 14(6), 637-643.

Kepler, T. B., Liao, H. X., Alam, S. M., Bhaskarabhatla, R., Zhang, R., Yandava, C., ... & Haynes, B. F. (2014). Immunoglobulin gene insertions and deletions in the affinity maturation of HIV-1 broadly reactive neutralizing antibodies. *Cell host & microbe*, 16(3), 304-313.

Kim, H. J., Krenn, V., Steinhauser, G., & Berek, C. (1999). Plasma cell development in synovial germinal centers in patients with rheumatoid and reactive arthritis. *The Journal of Immunology*, 162(5), 3053-3062.

Klein, U., Rajewsky, K., & Küppers, R. (1998). Human immunoglobulin (Ig) M+ IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. *The Journal of experimental medicine*, 188(9), 1679-1689.

Knisbacher, B. A., Gerber, D., & Levanon, E. Y. (2016). DNA editing by APOBECs: a genomic preserver and transformer. *Trends in Genetics*, 32(1), 16-28.

Kranich, J., & Krautler, N. J. (2016). How follicular dendritic cells shape the B-cell antigenome. *Frontiers in immunology*, 7, 225.

Kranich, J., Krautler, N. J., Heinen, E., Polymenidou, M., Bridel, C., Schildknecht, A., . . . Aguzzi, A. (2008). Follicular dendritic cells control engulfment of apoptotic bodies by secreting Mfge8. *Journal of Experimental Medicine,* 205(6), 1293-1302.

Kranz, H. (1989). Changes in splenic melano-macrophage centres of dab Limanda limanda during and after infection with ulcer disease. *Diseases of aquatic organisms*, 6(3), 167-173.

Krautler, N. J., Kana, V., Kranich, J., Tian, Y., Perera, D., Lemm, D., . . . Aguzzi, A. (2012). Follicular dendritic cells emerge from ubiquitous perivascular precursors. *Cell,* 150(1), 194-206.

Krokan, H. E., & Bjørås, M. (2013). Base excision repair. *Cold Spring Harbor Perspectives in Biology, 5*(4), 1-22.

Kuppers', R., Zhao2, M., Hansmann2, M., & Rajewsky, K. (1993). Tracing B cell development in human germinal centres by molecular analysis of single cells picked from histological sections. *The EMBO journal*, 12(13), 4955-4967.

Kuraku, S. (2013). Impact of asymmetric gene repertoire between cyclostomes and gnathostomes. *Seminars in cell & developmental biology* (Vol. 24, No. 2, pp. 119-127). Academic Press.

Kuraoka, M., Holl, T. M., Liao, D., Womble, M., Cain, D. W., Reynolds, A. E., & Kelsoe, G. (2011). Activation-induced cytidine deaminase mediates central tolerance in B cells. *Proceedings of the National Academy of Sciences of the United States of America, 108*(28), 11560-11565.

Kuraoka, M., Schmidt, A. G., Nojima, T., Feng, F., Watanabe, A., Kitamura, D., ... & Kelsoe, G. (2016). Complex antigens drive permissive clonal selection in germinal centers. *Immunity*, 44(3), 542-552.

Lamers, C. H. J., & Parmentier, H. K. (1985). The fate of intraperitoneally injected carbon particles in cyprinid fish. *Cell and tissue research*, 242(3), 499-503.

Lamers, C. H., & De Haas, M. J. (1985). Antigen localization in the lymphoid organs of carp (cyprinus carpio). *Cell and Tissue Research, 242*(3), 491-498.

Lanning, D. K., & Knight, K. L. (2015). Diversification of the primary antibody repertoire by AID-mediated gene conversion. *results Probl Cell Differ*, 279-293.

Lavinder, J. J., Hoi, K. H., Reddy, S. T., Wine, Y., & Georgiou, G. (2014). Systematic characterization and comparative analysis of the rabbit immunoglobulin repertoire. *PloS one*, 9(6), e101322.

Leal, E., Granja, A. G., Zarza, C., & Tafalla, C. (2016). Distribution of T cells in rainbow trout (Oncorhynchus mykiss) skin and responsiveness to viral infection. *PLoS One*, 11(1), e0147477.

Lee, D. W., Khavrutskii, I. V., Wallqvist, A., Bavari, S., Cooper, C. L., & Chaudhury, S. (2017). BRILIA: Integrated tool for high-throughput annotation and lineage tree assembly of B-cell repertoires. *Frontiers in Immunology,* 7, 681.

Lee, S. S., Tranchina, D., Ohta, Y., Flajnik, M. F., & Hsu, E. (2002). Hypermutation in shark immunoglobulin light chain genes results in contiguous substitutions. Immunity, 16(4), 571-582.

Lemaitre, B., Nicolas, E., & Michaut, L. (1996). The dorsoventral regulatory gene cassette spä tzle/toll/cactus controls the potent antifungal response in drosophila adults. *Cell*, 86(6), 973-983.

Lesley, R., Xu, Y., Kalled, S. L., Hess, D. M., Schwab, S. R., Shu, H. B., & Cyster, J. G. (2004). Reduced competitiveness of autoantigen-engaged B cells due to increased dependence on BAFF. *Immunity*, 20(4), 441-453.

Li, H., Li, Q., Ma, Z., Zhou, Z., Fan, J., Jin, Y., ... & Liang, P. (2019). AID modulates carcinogenesis network via DNA demethylation in bladder urothelial cell carcinoma. *Cell death & disease*, 10(4), 1-15.

Li, S. C., Rothman, P. B., Zhang, J., Chan, C., Hirsh, D., & Alt, F. W. (1994). Expression of iμ-cγ hybrid germline transcripts subsequent to immunoglobulin heavy chain class switching. *International Immunology,* 6(4), 491-497.

Li, X., Gadzinsky, A., Gong, L., Tong, H., Calderon, V., Li, Y., . . . Gu, H. (2018). Cbl ubiquitin ligases control B cell exit from the germinal-center reaction. *Immunity,* 48(3), 530-541.e6.

Litinskiy, M. B., Nardelli, B., Hilbert, D. M., He, B., Schaffer, A., Casali, P., & Cerutti, A. (2002). DCs induce CD40-independent immunoglobulin class switching through BLyS and APRIL. *Nature Immunology, 3*(9), 822-829.

Liu, M., & Schatz, D. G. (2009). Balancing AID and DNA repair during somatic hypermutation. *Trends in Immunology*, 30(4), 173-181.

Liu, M., Duke, J. L., Richter, D. J., Vinuesa, C. G., Goodnow, C. C., Kleinstein, S. H., & Schatz, D. G. (2008). Two levels of protection for the B cell genome during somatic hypermutation. *Nature,* 451(7180), 841-845.

Lohr, J. G., Stojanov, P., Lawrence, M. S., Auclair, D., Chapuy, B., Sougnez, C., . . . Golub, T. R. (2012). Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proceedings of the National Academy of Sciences of the United States of America, 109*(10), 3879-3884.

Longo, N. S., Satorius, C. L., Plebani, A., Durandy, A., & Lipsky, P. E. (2008). Characterization of ig gene somatic hypermutation in the absence of activation-induced cytidine deaminase. *The Journal of Immunology,* 181(2), 1299-1306.

Lovén, J., Hoke, H. A., Lin, C. Y., Lau, A., Orlando, D. A., Vakoc, C. R., . . . Young, R. A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell,* 153(2), 320-334.

Luckheeram, R. V., Zhou, R., Verma, A. D., & Xia, B. (2012). CD4+ T cells: differentiation and functions. *Clinical and developmental immunology*, 2012.

Lugo-Villarino, G., Balla, K. M., Stachura, D. L., Bañuelos, K., Werneck, M. B. F., & Traver, D. (2010). Identification of dendritic antigen-presenting cells in the zebrafish. *Proceedings of the National Academy of Sciences of the United States of America,* 107(36), 15850-15855.

Lund, H., Bakke, A. F., Sommerset, I., Afanasyev, S., Schriwer, G., Thorisdottir, A., . . . Krasnov, A. (2019). A time-course study of gene expression and antibody repertoire at early time post vaccination of atlantic salmon. *Molecular Immunology,* 106, 99-107.

MacLennan, I. C., & Gray, D. (1986). Antigen-driven selection of virgin and memory B cells. *Immunological reviews*, 91, 61-85.

Magadan-Mompo, S., Sanchez-Espinel, C., & Gambon-Deza, F. (2011). Immunoglobulin heavy chains in medaka (oryzias latipes). *BMC Evolutionary Biology*, 11, 165-165.

Magadan, S., Jouneau, L., Boudinot, P., & Salinas, I. (2019). Nasal vaccination drives modifications of nasal and systemic antibody repertoires in rainbow trout. *The Journal of Immunology*, 203(6), 1480-1492.

Magadan, S., Jouneau, L., Puelma Touzel, M., Marillet, S., Chara, W., Six, A., ... & Boudinot, P. (2018). Origin of public memory B cell clones in fish after antiviral vaccination. *Frontiers in immunology*, 2115.

Magadan, S., Sunyer, O. J., & Boudinot, P. (2015). Unique features of fish immune repertoires: Particularities of adaptive immunity within the largest group of vertebrates. *Results and Problems in Cell Differentiation*, 57, 235-264.

Magor, B. G. (2015). Antibody affinity maturation in fishes-our current understanding. *Biology, 4*(3), 512-524.

Maizels, N. (2005). Immunoglobulin gene diversification. *Annu. Rev. Genet*., 39, 23-46.

Makela, O., & Litman, G. W. (1980). Lack of heterogeneity in anti-hapten antibodies of a phylogenetically primitive shark. *Nature*, 287(5783), 639-640.

Malecek, K., Brandman, J., Brodsky, J. E., Ohta, Y., Flajnik, M. F., & Hsu, E. (2005). Somatic hypermutation and junctional diversification at Ig heavy chain loci in the nurse shark. *The Journal of Immunology*, 175(12), 8105-8115.

Manis, J. P., Tian, M., & Alt, F. W. (2002). Mechanism and control of class-switch recombination. *Trends in Immunology*, 23(1), 31-39.

Manrique, W. G., Pereira Figueiredo, M. A., Charlie-Silva, I., Antonio de Andrade Belo, Marco, & Dib, C. C. (2019). Spleen melanomacrophage centers response of nile tilapia during aeromanas hydrophila and mycobacterium marinum infections. *Fish and Shellfish Immunology,* 95, 514-518.

Marianes, A. E., & Zimmerman, A. M. (2011). Targets of somatic hypermutation within immunoglobulin light chain genes in zebrafish. *Immunology,* 132(2), 240-255.

Marshall, J. S., Warrington, R., Watson, W., & Kim, H. L. (2018). An introduction to immunology and immunopathology, *Allergy, Asthma & Clinical Immunology*, 14(2), 1-10.

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus evolution*, 1(1).

Martin, F., & Kearney, J. F. (2002). Marginal-zone B cells. *Nature Reviews Immunology*, 2(5), 323-335.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet* J 17, 10–12.

Matthews, A. J., Zheng, S., DiMenna, L. J., & Chaudhuri, J. (2014). Regulation of immunoglobulin class-switch recombination: choreography of noncoding transcription, targeted DNA deamination, and long-range DNA repair. *Advances in immunology*, 122, 1-57.

Matthews, T. J., & Whittaker, R. J. (2015). On the species abundance distribution in applied ecology and biodiversity management. *Journal of Applied Ecology*, 52(2), 443–454.

Maul, R. W., MacCarthy, T., Frank, E. G., Donigan, K. A., McLenigan, M. P., Yang, W., . . . Gearhart, P. J. (2016). DNA polymerase ι functions in the generation of tandem mutations during somatic hypermutation of antibody genes. *Journal of Experimental Medicine,* 213(9), 1675-1683.

Maul, R. W., MacCarthy, T., Frank, E. G., Donigan, K. A., McLenigan, M. P., Yang, W., ... & Gearhart, P. J. (2016). DNA polymerase ι functions in the generation of tandem mutations during somatic hypermutation of antibody genes. *Journal of Experimental Medicine*, 213(9), 1675-1683.

Maul, R. W., Saribasak, H., Martomo, S. A., McClure, R. L., Yang, W., Vaisman, A., . . . Gearhart, P. J. (2011). Uracil residues dependent on the deaminase AID in immunoglobulin gene variable and switch regions. *Nature Immunology,* 12(1), 70-76.

Mayer, C. T., Gazumyan, A., Kara, E. E., Gitlin, A. D., Golijanin, J., Viant, C., ... & Nussenzweig, M. C. (2017). The microanatomic segregation of selection by apoptosis in the germinal center. *Science*, 358(6360), eaao2602.

Mayorov, V. I., Rogozin, I. B., Adkison, L. R., & Gearhart, P. J. (2005). DNA polymerase η contributes to strand bias of mutations of A versus T in immunoglobulin genes. *The Journal of Immunology*, 174(12), 7781-7786.

Mcheyzer-Williams, M., Mclean, M. J., Lalor, P. A., & Nossal, G. J. V. (1993). Antigen-driven B cell differentiation in vivo. *The Journal of experimental medicine*, 178(1), 295-307.

Mesin, L., Ersching, J., & Victora, G. D. (2016). Germinal center B cell dynamics. *Immunity,* 45(3), 471-482.

Mestecky, J., Moro, I., Kerr, M. A., & Woof, J. M. (2005). Mucosal immunoglobulins. *Mucosal immunology* (pp. 153-181). Academic Press.

Methot, S. P., & Di Noia, J. M. (2017). Molecular mechanisms of somatic hypermutation and class switch recombination. *Advances in immunology*, 133, 37-87.

Monod, M. Y., Giudicelli, V., Chaume, D., & Lefranc, M. P. (2004). IMGT/JunctionAnalysis: the first tool for the analysis of the immunoglobulin and T cell receptor complex V–J and V–D–J JUNCTIONs. *Bioinformatics*, 20(suppl_1), i379-i385.

Morgan, H. D., Dean, W., Coker, H. A., Reik, W., & Petersen-Mahrt, S. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: Implications for epigenetic reprogramming. *Journal of Biological Chemistry, 279*(50), 52353-52360.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., & Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, 102(5), 553-563.

Muramatsu, M., Sankaranand, V. S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N. O., & Honjo, T. (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *Journal of Biological Chemistry*, 274(26), 18470-18476.

Murphy, M., Jason-Moller, L., & Bruno, J. (2006). Using Biacore to Measure the Binding Kinetics of an Antibody-Antigen Interaction. *Current protocols in protein science*, 45(1), 19-14.

Muthupandian, A. (2020). Identification and characterization of antigen retaining cells in the putative germinal centres of goldfish (Carassius auratus)

Muthupandian, A., Waly, D., & Magor, B. G. (2021). Do ectothermic vertebrates have a home in which to affinity mature their antibody responses?. *Developmental & Comparative Immunology*, 119, 104021.

Nacionales, D. C., Weinstein, J. S., Yan, X. J., Albesiano, E., Lee, P. Y., Kelly-Scumpia, K. M., . . . Reeves, W. H. (2009). B cell proliferation, somatic hypermutation, class switch recombination, and autoantibody production in ectopic lymphoid tissue in murine lupus. *Journal of Immunology*, 182(7), 4226-4236.

Nadel, B., & Feeney, A. J. (1997). Nucleotide deletion and P addition in V(D)J recombination: A determinant role of the coding-end sequence. *Molecular and Cellular Biology*, 17(7), 3768-3778.

Nakanishi, T., Toda, H., Shibasaki, Y., & Somamoto, T. (2011). Cytotoxic T cells in teleost fish. *Developmental and Comparative Immunology,* 35(12), 1317-1323.

Nathanael McCurley, Masayuki Hirano, Sabyasachi Das, & Max D. Cooper. (2012). Immune Related Genes Underpin the Evolution of Adaptive Immunity in Jawless Vertebrates. *Current Genomics*, 13(2), 86–94.

Neely, H. R., & Flajnik, M. F. (2016). Emergence and evolution of secondary lymphoid organs. *Annual review of cell and developmental biology*, 32, 693-711.

Neely, H. R., Guo, J., Flowers, E. M., Criscitiello, M. F., & Flajnik, M. F. (2018). "Double-duty" conventional dendritic cells in the amphibian Xenopus as the prototype for antigen presentation to B cells. *European journal of immunology*, 48(3), 430-440.

Nelson, J. R., Lawrence, C. W., & Hinkle, D. C. (1996). Deoxycytidyl transferase activity of yeast REV1 protein. *Nature*, 382(6593), 729-731.

Nemazee, D. (2006). Receptor editing in lymphocyte development and central tolerance. *Nature Reviews Immunology*, 6(10), 728-740.

Nemazee, D. (2017). Mechanisms of central tolerance for B cells. *Nature Reviews Immunology*, 17(5), 281-294.

Neuberger, M. S., & Rada, C. (2007). Somatic hypermutation: Activation-induced deaminase for C/G followed by polymerase eta for A/T. *The Journal of Experimental Medicine, 204*(1), 7-10.

Nicholson, L. B. (2016). The immune system. *Essays in Biochemistry,* 60(3), 275-301.

Nowak, B. F., Dang, M., Webber, C., Neumann, L., Bridle, A., Bermudez, R., & Evans, D. (2021). Changes in the splenic melanomacrophage centre surface area in southern bluefin tuna (Thunnus maccoyii) are associated with blood fluke infections. *Pathogens*, 10(1), 79.

Ohm-Laursen, L., Meng, H., Hoehn, K., Nouri, N., Jiang, Y., Clouser, C., ... & Gould, H. (2021). B cell Mobilisation, Dissemination, Fine Tuning of Local Antigen Specificity and Isotype Selection in Asthma. *Frontiers in immunology*, 3891.

Okazaki, I. M., Kinoshita, K., Muramatsu, M., Yoshikawa, K., & Honjo, T. (2002). The AID enzyme induces class switch recombination in fibroblasts. *Nature*, 416(6878), 340-345.

Ott, J. A., Castro, C. D., Deiss, T. C., Ohta, Y., Flajnik, M. F., & Criscitiello, M. F. (2018). Somatic hypermutation of T cell receptor α chain contributes to selection in nurse shark thymus. *Elife*, 7, e28477.

Ott, J. A., Harrison, J., Flajnik, M. F., & Criscitiello, M. F. (2020). Nurse shark T-cell receptors employ somatic hypermutation preferentially to alter alpha/delta variable segments associated with alpha constant region. *European journal of immunology*, 50(9), 1307-1320.

Padidam, M., Sawyer, S. & Fauquet, C. M. (1999). Possible emergence of new geminiviruses by frequent recombination. *Virology,* 265, 218-225.

Page, D. M., Wittamer, V., Bertrand, J. Y., Lewis, K. L., Pratt, D. N., Delgado, N., . . . Traver, D. (2013). An evolutionarily conserved program of B-cell development and activation in zebrafish. *Blood*, 122(8), 1.

Pavri, R., Gazumyan, A., Jankovic, M., Di Virgilio, M., Klein, I., Ansarah-Sobrinho, C., . . . Nussenzweig, M. C. (2010). Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell,* 143(1), 122-133.

Peakman, M. C., & Maizels, N. (1998). Localization of splenic B cells activated for switch recombination by in situ hybridization with Iγ1 switch transcript and Rad51 probes. *The Journal of Immunology*, 161(8), 4008-4015.

Peatman, E., Lange, M., Zhao, H., & Beck, B. H. (2015). Physiology and immunology of mucosal barriers in catfish (ictalurus spp.). *Tissue Barriers, 3*(4), 1-14.

Pélissier, A., Akrout, Y., Jahn, K., Kuipers, J., Klein, U., Beerenwinkel, N., & Rodríguez Martínez, M. (2020). Computational model reveals a stochastic mechanism behind germinal center clonal bursts. *Cells*, 9(6), 1448.

Perdiguero, P., Martín-Martín, A., Benedicenti, O., Díaz-Rosales, P., Morel, E., Muñoz-Atienza, E., . . . Tafalla, C. (2019). Teleost IgD+IgM− B cells mount clonally expanded and mildly mutated intestinal IgD responses in the absence of lymphoid follicles. *Cell Reports,* 29(13), 4223-4235.e5.

Petrie-Hanson, L., Hohn, C., & Hanson, L. (2009). Characterization of rag 1 mutant zebrafish leukocytes. *BMC immunology*, 10(1), 1-8.

Phan, T. G., Green, J. A., Xu, Y., & Cyster, J. G. (2009). Immune complex relay by subcapsular sinus macrophages and noncognate B cells drives antibody affinity maturation. *Nature Immunology, 10*(7), 786-796.

Pilzecker, B., & Jacobs, H. (2019). Mutating for good: DNA damage responses during somatic hypermutation. *Frontiers in immunology*, 10, 438.

Poltorak, A., He, X., Smirnova, I., Liu, M. Y., Van Huffel, C., Du, X., . . . Beutler, B. (1998). Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: Mutations in Tlr4 gene. *Science, 282*(5396), 2085-2088.

Popp, C., Dean, W., Feng, S., Cokus, S. J., Andrews, S., Pellegrini, M., . . . Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature,* 463(7284), 1101-1105.

Press, C. M., Evensen, Ø., Reitan, L. J., & Landsverk, T. (1996). Retention of furunculosis vaccine components in Atlantic salmon, Salmo salar L., following different routes of vaccine administration. *Journal of Fish Diseases*, 19(3), 215-224.

Qian, J., Wang, Q., Dose, M., Pruett, N., Kieffer-Kwon, K. R., Resch, W., . . . Casellas, R. (2014). B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. *Cell,* 159(7), 1524-1537.

Rada, C., Gupta, S. K., Gherardi, E., & Milstein, C. U. S. A. R. (1991). Mutation and selection during the secondary response to 2-phenyloxazolone. *Proceedings of the National Academy of Sciences*, 88(13), 5508-5512.

Rada, C., Williams, G. T., Nilsen, H., Barnes, D. E., Lindahl, T., & Neuberger, M. S. (2002). Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Current Biology*, 12(20), 1748-1755.

Rai, K., Huggins, I. J., James, S. R., Karpf, A. R., Jones, D. A., & Cairns, B. R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell*, 135(7), 1201-1212.

Ramírez, J., Lukin, K., & Hagman, J. (2010). From hematopoietic progenitors to B cells: mechanisms of lineage restriction and commitment. *Current opinion in immunology*, 22(2), 177-184.

Ramiro, A. R., Jankovic, M., Eisenreich, T., Difilippantonio, S., Chen-Kiang, S., Muramatsu, M., . . . Nussenzweig, M. C. (2004). AID is required for c-myc/IgH chromosome translocations in vivo. *Cell,* 118(4), 431-438.

Rathmell, J. C., & Thompson, C. B. (2002). Pathways of apoptosis in lymphocyte development, homeostasis, and disease. *Cell*, 109(2), S97-S107.

Reason, D. C., & Zhou, J. (2006). Codon insertion and deletion functions as a somatic diversification mechanism in human antibody repertoires. *Biology Direct*, 1(1), 1-14.

Reddy, S. T., Ge, X., Miklos, A. E., Hughes, R. A., Kang, S. H., Hoi, K. H., . . . Georgiou, G. (2010). Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. *Nature Biotechnology,* 28(9), 965-969.

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., ... & Durandy, A. (2000). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell*, 102(5), 565-575.

Reyneveld, G. I., Savelkoul, H. F. J., & Parmentier, H. K. (2020). Current understanding of natural antibodies and exploring the possibilities of modulation using veterinary models. A review. *Frontiers in Immunology*, 11, 2139.

Rios, L. A. S., Cloete, B., & Mowla, S. (2020). Activation-induced cytidine deaminase: In sickness and in health. *Journal of Cancer Research and Clinical Oncology*, 146(11), 2721-2730.

Roco, J. A., Mesin, L., Binder, S. C., Nefzger, C., Gonzalez-Figueroa, P., Canete, P. F., . . . Vinuesa, C. G. (2019). Class-switch recombination occurs infrequently in germinal centers. *Immunity,* 51(2), 337-350.e7.

Röhlich, K. (1930). Beitrag zur Cytologie der Keimzentren der Lymphknoten. *Z. Mikrosk. Anat. Forsch*, 20, 287-297.

Rosenfeld, A. M., Meng, W., Chen, D. Y., Zhang, B., Granot, T., Farber, D. L., ... & Luning Prak, E. T. (2018). Computational evaluation of B-cell clone sizes in bulk populations. *Frontiers in immunology*, 9, 1472.

Roth, D. B. (2015). V (D) J recombination: mechanism, errors, and fidelity. *Mobile DNA III*, 311-324.

Ruddle, N. H. (2020). Basics of inducible lymphoid organs. *Inducible Lymphoid Organs*, 1-19.

Rumfelt, L. L., McKinney, E. C., Taylor, E., & Flajnik, M. F. (2002). The development of primary and secondary lymphoid tissues in the nurse shark Ginglymostoma cirratum: B-cell zones precede

dendritic cell immigration and T-cell zone formation during ontogeny of the spleen. *Scandinavian journal of immunology*, 56(2), 130-148.

Sale, J. E., Calandrini, D. M., Takata, M., Takeda, S., & Neuberger, M. S. (2001). Ablation of XRCC2/3 transforms immunoglobulin V gene conversion into somatic hypermutation. *Nature*, 412(6850), 921-926.

Salinas, I., Fernández-Montero, Á., Ding, Y., & Sunyer, J. O. (2021). Mucosal immunoglobulins of teleost fish: A decade of advances. *Developmental & Comparative Immunology*, 121, 104079.

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., ... & Smith, M. (1977). Nucleotide sequence of bacteriophage φX174 DNA. *nature*, 265(5596), 687-695.

Saribasak, H., Maul, R. W., Cao, Z., Yang, W. W., Schenten, D., Kracker, S., & Gearhart, P. J. (2012). DNA polymerase ζ generates tandem mutations in immunoglobulin variable regions. *Journal of Experimental Medicine*, 209(6), 1075-1081.

Sathyanarayanan, A., Manda, S., Poojary, M., & Nagaraj, S. H. (2018). Exome sequencing data analysis.

Saunders, H. L. (2020). Ictalurus Punctatus Activation-Induced Cytidine Deaminase Expression.

Saunders, H. L., & Magor, B. G. (2004). Cloning and expression of the AID gene in the channel catfish. *Developmental and Comparative Immunology,* 28(7-8), 657-663.

Saunders, H. L., Oko, A. L., Scott, A. N., Fan, C. W., & Magor, B. G. (2010). The cellular context of AID expressing cells in fish lymphoid tissues. *Developmental and Comparative Immunology,* 34(6), 669-676.

Savan, R., Aman, A., Sato, K., Yamaguchi, R., & Sakai, M. (2005). Discovery of a new class of immunoglobulin heavy chain from fugu. *European Journal of Immunology,* 35(11), 3320-3331.

Sayed, A. H., & Younes, H. M. (2017). Melanomacrophage centers in clarias gariepinus as an immunological biomarker for toxicity of silver nanoparticles. *Journal of Microscopy and Ultrastructure, 5*(2), 97.

Scapigliati, G. (2013). Functional aspects of fish lymphocytes. *Developmental and Comparative Immunology,* 41(2), 200-208.

Schatz, D. G., & Ji, Y. (2011). Recombination centres and the orchestration of V(D)J recombination. *Nature Reviews.Immunology*, 11(4), 251-263.

Schatz, D. G., Oettinger, M. A., & Baltimore, D. (1989). The V (D) J recombination activating gene, Rag-1. *Cell*, 59(6), 1035-1048.

Schroeder Jr, H. W., & Cavacini, L. (2010). Structure and function of immunoglobulins. *Journal of Allergy and Clinical Immunology*, 125(2), S41-S52.

Schroeder, H. W. (2006). Similarity and divergence in the development and expression of the mouse and human antibody repertoires. *Developmental and Comparative Immunology*, 30(1-2), 119-135.

Schwickert, T. A., Lindquist, R. L., Shakhar, G., Livshits, G., Skokos, D., Kosco-Vilbois, M. H., ... & Nussenzweig, M. C. (2007). In vivo imaging of germinal centres reveals a dynamic open structure. *Nature*, 446(7131), 83-87.

Scott, R., Zhan, A., Brown, E. A., Chain, F. J., Cristescu, M. E., Gras, R., & MacIsaac, H. J. (2018). Optimization and performance testing of a sequence processing pipeline applied to detection of nonindigenous species. *Evolutionary Applications*, 11(6), 891-905.

Secombes, C. J., & MANNING, M. J. (1980). Comparative studies on the immune system of fishes and amphibians: antigen localization in the carp Cyprinus carpio L. *Journal of Fish Diseases*, 3(5), 399-412.

Secombes, C. J., Manning, M. J., & Ellis, A. E. (1982). Localization of immune complexes and heat-aggregated immunoglobulin in the carp Cyprinus carpio L. *Immunology*, 47(1), 101.

Shankey, T. V., & Clem, L. W. (1980). Phylogeny of immunoglobulin structure and function—VIII: Intermolecular heterogeneity of shark 19S IgM antibodies to pneumococcal polysaccharide. *Molecular Immunology*, 17(3), 365-375.

Shinkura, R., Ito, S., Begum, N. A., Nagaoka, H., Muramatsu, M., Kinoshita, K., . . . Honjo, T. (2004). Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nature Immunology*, 5(7), 707-712.

Shlemov, A., Bankevich, S., Bzikadze, A., Turchaninova, M. A., Safonova, Y., & Pevzner, P. A. (2017). Reconstructing antibody repertoires from error-prone immunosequencing reads. *The Journal of Immunology*, 199(9), 3369-3380.

Shu, Y. Y., Fugmann, S. D., & Schatz, D. G. (2006). Control of gene conversion and somatic hypermutation by immunoglobulin promoter and enhancer sequences. *Journal of Experimental Medicine,* 203(13), 2919-2928.

Sims, G. P., Shiono, H., Willcox, N., & Stott, D. I. (2001). Somatic hypermutation and selection of B cells in thymic germinal centers responding to acetylcholine receptor in myasthenia gravis. *The Journal of Immunology,* 167(4), 1935-1944.

Somamoto, T., Okamoto, N., Nakanishi, T., Ototake, M., & Nakao, M. (2009). In vitro generation of viral-antigen dependent cytotoxic T-cells from ginbuna crucian carp, carassius auratus langsdorfii. *Virology,* 389(1-2), 26-33.

Song, W., & Craft, J. (2019). T follicular helper cell heterogeneity: time, space, and function. *Immunological reviews*, 288(1), 85-96.

Stavnezer, J., Guikema, J. E., & Schrader, C. E. (2008). Mechanism and regulation of class switch recombination. *Annual Review of Immunology*, 26, 261-292.

Stavnezer, J., Linehan, E. K., Thompson, M. R., Habboub, G., Ucher, A. J., Kadungure, T., . . . Schrader, C. E. (2014). Differential expression of APE1 and APE2 in germinal centers promotes error-prone repair and A:T mutations during somatic hypermutation. *Proceedings of the National Academy of Sciences of the United States of America,* 111(25), 9217-9222.

Stebegg, M., Kumar, S. D., Silva-Cayetano, A., Fonseca, V. R., Linterman, M. A., & Graca, L. (2018). Regulation of the germinal center response. *Frontiers in immunology*, 2469.

Steinel, N. C., & Bolnick, D. I. (2017). Melanomacrophage centers as a histological indicator of immune function in fish and other poikilotherms. *Frontiers in Immunology,* 8, 827.

Stern, J. N., Yaari, G., Vander Heiden, J. A., Church, G., Donahue, W. F., Hintzen, R. Q., ... & O'Connor, K. C. (2014). B cells populating the multiple sclerosis brain mature in the draining cervical lymph nodes. *Science translational medicine*, 6(248), 248ra107-248ra107.

Storb, U. (1996). The molecular basis of somatic hypermutation of immunoglobulin genes. *Current opinion in immunology*, 8(2), 206-214.

Stosik, M., Tokarz-Deptuła, B., & Deptuła, W. (2021). Immunological memory in teleost fish. *Fish & Shellfish Immunology*, 115, 95-103.

Stott, D. I., Hiepe, F., Hummel, M., Steinhauser, G., & Berek, C. (1998). Antigen-driven clonal proliferation of B cells within the target tissue of an autoimmune disease. The salivary glands of patients with Sjögren's syndrome. *The Journal of clinical investigation*, 102(5), 938-946.

Sutton, B., Davies, A., Bax, H., & Karagiannis, S. (2019). IgE antibodies: From structure to function and clinical translation. *Antibodies, 8*(1), 19.

Tangye, S. G., Ferguson, A., Avery, D. T., Ma, C. S., & Hodgkin, P. D. (2002). Isotype switching by human B cells is division-associated and regulated by cytokines. *The Journal of Immunology,* 169(8), 4298-4306.

Tas, J. M., Mesin, L., Pasqual, G., Targ, S., Jacobsen, J. T., Mano, Y. M., ... & Victora, G. D. (2016). Visualizing antibody affinity maturation in germinal centers. *Science*, 351(6277), 1048-1054.

Tatner, M. F., Adams, A., & Leschen, W. (1987). An analysis of the primary and secondary antibody response in intact and thymectomized rainbow trout, Salmo gairdneri richardson, to human gamma globulin and Aeromonas salmonicida. *Journal of fish biology*, 31(2), 177-195.

Tedersoo, L., Nilsson, R.H., Abarenkov, K., Jairus, T., Sadam, A., Saar, I., Bahram, M., Bechem, E., Chuyong, G. and Kõljalg, U. (2010). 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New phytologist*, 188(1), 291-301.

Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature*, 302(5909), 575-581.

Tran, T. H., Nakata, M., Suzuki, K., Begum, N. A., Shinkura, R., Fagarasan, S., . . . Nagaoka, H. (2010). B cell-specific and stimulation-responsive enhancers derepress aicda by overcoming the effects of silencers. *Nature Immunology,* 11(2), 148-154.

Trancoso, I., Morimoto, R., & Boehm, T. (2020). Co-evolution of mutagenic genome editors and vertebrate adaptive immunity. *Current opinion in immunology*, 65, 32-41.

Traver, D., Paw, B. H., Poss, K. D., Penberthy, W. T., Lin, S., & Zon, L. I. (2003). Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nature immunology*, 4(12), 1238-1246.

Tsujii, T., & Seno, S. (1990). Melano-macrophage centers in the aglomerular kidney of the sea horse (teleosts): Morphologic studies on its formation and possible function. *The Anatomical Record*, 226(4), 460-470.

Ubelhart, R., & Jumaa, H. (2015). Autoreactivity and the positive selection of B cells. *European Journal of Immunology,* 45(11), 2971-2977.

Vander Heiden, J. A., Yaari, G., Uduman, M., Stern, J. N., O'Connor, K. C., Hafler, D. A., ... & Kleinstein, S. H. (2014). pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. *Bioinformatics*, 30(13), 1930-1932.

Verstegen, N. J., Ubels, V., Westerhoff, H. V., Van Ham, S. M., & Barberis, M. (2021). System-Level Scenarios for the Elucidation of T Cell-Mediated Germinal Center B Cell Differentiation. *Frontiers in Immunology*, 3829.

Villota-Herdoiza, D., Pila, E. A., Quiniou, S., Waldbieser, G. C., & Magor, B. G. (2013). Transcriptional regulation of teleost aicda genes. part 1 - suppressors of promiscuous promoters. *Fish & Shellfish Immunology,* 35(6), 1981-1987.

Vivier, E., Tomasello, E., Baratin, M., Walzer, T., & Ugolini, S. (2008). Functions of natural killer cells. *Nature immunology*, 9(5), 503-510.

Wagner, S. D., Milstein, C., & Neuberger, M. S. (1995). Codon bias targets mutation. *Nature*, 376(6543), 732.

Wakae, K., Magor, B. G., Saunders, H., Nagaoka, H., Kawamura, A., Kinoshita, K., . . . Muramatsu, M. (2006). Evolution of class switch recombination function in fish activation-induced cytidine deaminase, AID. *International Immunology,* 18(1), 41-47.

Wang, Q., Kieffer-Kwon, K., Oliveira, T. Y., Mayer, C. T., Yao, K., Pai, J., . . . Robbiani, D. F. (2017). The cell cycle restricts activation-induced cytidine deaminase activity to early G1. *Journal of Experimental Medicine,* 214(1), 49-58.

Weinstein, J. A., Jiang, N., White, R. A., Fisher, D. S., & Quake, S. R. (2009). High-throughput sequencing of the zebrafish antibody repertoire. *Science,* 324(5928), 807-810.

Weyand, C. M., & Goronzy, J. J. (2003). Ectopic germinal center formation in rheumatoid synovitis. *Annals of the New York Academy of Sciences*, 987, 140-149.

Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., . . . Young, R. A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell,* 153(2), 307-319.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*, ISBN 978-3-319-24277-4

William, J., Euler, C., Christensen, S., & Shlomchik, M. J. (2002). Evolution of autoantibody responses via somatic hypermutation outside of germinal centers. *Science, 297*(5589), 2066-2070.

Williams, A. F., & Barclay, A. N. (1988). The immunoglobulin superfamily—domains for cell surface recognition. *Annual review of immunology*, 6(1), 381-405.

Wilson, M., Hsu, E., Marcuz, A., Courtet, M., Du Pasquier, L., & Steinberg, C. (1992). What limits affinity maturation of antibodies in Xenopus--the rate of somatic mutation or the ability to select mutants?. *The EMBO Journal*, 11(12), 4337-4347.

Wilson, P. C., Bouteiller, O. D., Liu, Y. J., Potter, K., Banchereau, J., Capra, J. D., & Pascual, V. (1998). Somatic hypermutation introduces insertions and deletions into immunoglobulin V genes. *The Journal of experimental medicine*, 187(1), 59-70.

Wing, J. B., Tekguc, M., & Sakaguchi, S. (2018). Control of germinal center responses by T-follicular regulatory cells. *Frontiers in Immunology,* 9, 1910.

Wishnie, A. J., Chwat-Edelstein, T., Attaway, M., & Vuong, B. Q. (2021). BCR Affinity Influences TB Interactions and B Cell Development in Secondary Lymphoid Organs. *Frontiers in Immunology*, 12.

Wittamer, V., Bertrand, J. Y., Gutschow, P. W., & Traver, D. (2011). Characterization of the mononuclear phagocyte system in zebrafish. Blood, *The Journal of the American Society of Hematology*, 117(26), 7126-7135.

Wittenbrink, N., Klein, A., Weiser, A. A., Schuchhardt, J., & Or-Guil, M. (2011). Is there a typical germinal center? A large-scale immunohistological study on the cellular composition of germinal centers during the hapten-carrier–driven primary immune response in mice. *The Journal of Immunology*, 187(12), 6185-6196.

Wolke, R. E. (1992). Piscine macrophage aggregates: a review. *Annual Review of Fish Diseases*, 2, 91-108.

Wu, L., Fu, S., Yin, X., Guo, Z., Wang, A., & Ye, J. (2019). Long-lived plasma cells secrete high-affinity antibodies responding to a T-dependent immunization in a teleost fish. *Frontiers in Immunology*, 2324.

Wu, L., Fu, S., Yin, X., Leng, W., Guo, Z., Wang, A., & Ye, J. (2019). Affinity maturation occurs in channel catfish (ictalurus punctaus) following immunization with a T-cell dependent antigen. *Fish and Shellfish Immunology,* 84, 781-786.

Wu, L., Kong, L., Yang, Y., Bian, X., Wu, S., Li, B., ... & Ye, J. (2019). Effects of cell differentiation on the phagocytic activities of IgM+ B cells in a teleost fish. *Frontiers in Immunology*, 10, 2225.

Wu, L., Oficjalska, K., Lambert, M., Fennell, B. J., Darmanin-Sheehan, A., Ní Shúilleabháin, D., . . . Finlay, W. J. J. (2012). Fundamental characteristics of the immunoglobulin V H repertoire of chickens in comparison with those of humans, mice, and camelids. *The Journal of Immunology,* 188(1), 322-333.

Wu, L., Qin, Z., Liu, H., Lin, L., Ye, J., & Li, J. (2020). Recent advances on phagocytic B cells in teleost fish. *Frontiers in Immunology*, 11, 824.

Xu, H. Y., Dong, F., Zhai, X., Meng, K. F., Han, G. K., Cheng, G. F., ... & Xu, Z. (2020). Mediation of mucosal immunoglobulins in buccal cavity of teleost in antibacterial immunity. *Frontiers in immunology*, 2406.

Xu, X., Hsu, H. C., Chen, J., Grizzle, W. E., Chatham, W. W., Stockard, C. R., . . . Mountz, J. D. (2009). Increased expression of activation-induced cytidine deaminase is associated with anti-CCP and rheumatoid factor in rheumatoid arthritis. *Scandinavian Journal of Immunology*, 70(3), 309-316.

Xu, Z., Pone, E. J., Al-Qahtani, A., Park, S. R., Zan, H., & Casali, P. (2007). Regulation of aicda expression and AID activity: relevance to somatic hypermutation and class switch DNA recombination. *Critical Reviews™ in Immunology*, 27(4).

Yang, F., Waldbieser, G. C., & Lobb, C. J. (2006). The nucleotide targets of somatic mutation and the role of selection in immunoglobulin heavy chains of a teleost fish. *The Journal of Immunology,* 176(3), 1655-1667.

Yang, W., & Woodgate, R. (2007). What a difference a decade makes: insights into translesion DNA synthesis. *Proceedings of the National Academy of Sciences*, 104(40), 15591-15598.

Ye, J., Bromage, E. S., & Kaattari, S. L. (2010). The strength of B cell interaction with antigen determines the degree of IgM polymerization. *The Journal of Immunology,* 184(2), 844-850.

Ye, J., Bromage, E., Kaattari, I., & Kaattari, S. (2011). Transduction of binding affinity by B lymphocytes: a new dimension in immunological regulation. *Developmental & Comparative Immunology*, 35(9), 982-990.

Ye, J., Kaattari, I. M., & Kaattari, S. L. (2011). The differential dynamics of antibody subpopulation expression during affinity maturation in a teleost. *Fish and Shellfish Immunology,* 30(1), 372-377.

Ye, J., Kaattari, I. M., Ma, C., & Kaattari, S. (2013). The teleost humoral immune response. *Fish and Shellfish Immunology,* 35(6), 1719-1728.

Yokoyama, S., Higashi, M., Kitamoto, S., Oeldorf, M., Knippschild, U., Kornmann, M., ... & Hollingsworth, M. A. (2016). Aberrant methylation of MUC1 and MUC4 promoters are potential prognostic biomarkers for pancreatic ductal adenocarcinomas. *Oncotarget*, 7(27), 42553.

Yu, K., & Lieber, M. R. (2019). Current insights into the mechanism of mammalian immunoglobulin class switch recombination. *Critical reviews in biochemistry and molecular biology*, 54(4), 333-351.

Yu, Y. Y., Kong, W. G., Xu, H. Y., Huang, Z. Y., Zhang, X. T., Ding, L. G., ... & Xu, Z. (2019). Convergent evolution of mucosal immune responses at the buccal cavity of teleost fish. *Iscience*, 19, 821-835.

Zahn, A., Eranki, A. K., Patenaude, A. M., Methot, S. P., Fifield, H., Cortizas, E. M., ... & Di Noia, J. M. (2014). Activation induced deaminase C-terminal domain links DNA breaks to end protection and repair during class switch recombination. *Proceedings of the National Academy of Sciences*, 111(11), E988-E997.

Zan, H., & Casali, P. (2013). Regulation of aicda expression and AID activity. *Autoimmunity,* 46(2), 83-101.

Zanotti, K. J., & Gearhart, P. J. (2016). Antibody diversification caused by disrupted mismatch repair and promiscuous DNA polymerases. *DNA repair*, 38, 110-116.

Zeng, X., Winter, D. B., Kasmer, C., Kraemer, K. H., Lehmann, A. R., & Gearhart, P. J. (2001). DNA polymerase η is an AT mutator in somatic hypermutation of immunoglobulin variable genes. *Nature immunology*, 2(6), 537-541.

Zhang, Y. A., Salinas, I., Li, J., Parra, D., Bjork, S., Xu, Z., . . . Sunyer, J. O. (2010). IgT, a primitive immunoglobulin class specialized in mucosal immunity. *Nature Immunology,* 11(9), 827-835.

Zhao, Y., Pan-Hammarström, Q., Zhao, Z., & Hammarström, L. (2005). Identification of the activation-induced cytidine deaminase gene from zebrafish: An evolutionary analysis. *Developmental and Comparative Immunology,* 29(1), 61-71.

Zheng, N. Y., Wilson, K., Jared, M., & Wilson, P. C. (2005). Intricate targeting of immunoglobulin somatic hypermutation maximizes the efficiency of affinity maturation. *The Journal of Experimental Medicine,* 201(9), 1467-1478.

Zhu, C., & Hsu, E. (2010). Error-prone DNA repair activity during somatic hypermutation in shark B lymphocytes. *The Journal of Immunology*, 185(9), 5336-5347.

Zhu, L., Lin, A., Shao, T., Nie, L., Dong, W., Xiang, L., & Shao, J. (2014). B cells in teleost fish act as pivotal initiating APCs in priming adaptive immunity: An evolutionary perspective on the origin of the B-1 cell subset and B7 molecules. *The Journal of Immunology,* 192(6), 2699-2714.

Ziegenfuss, M. C., & Wolke, R. E. (1991). The use of fluorescent microspheres in the study of piscine macrophage aggregate kinetics. *Developmental & Comparative Immunology*, 15(3), 165-171.

Zimmerman, A. M., Moustafa, F. M., Romanowski, K. E., & Steiner, L. A. (2011). Zebrafish immunoglobulin IgD: Unusual exon usage and quantitative expression profiles with IgM and IgZ/T heavy chain isotypes. *Molecular Immunology, 48*(15-16), 2220-2223.

Zimmerman, A. M., Yeo, G., Howe, K., Maddox, B. J., & Steiner, L. A. (2008). Immunoglobulin light chain (IgL) genes in zebrafish: Genomic configurations and inversional rearrangements between (VL-JL-CL) gene clusters. *Developmental and Comparative Immunology, 32*(4), 421-434.

Zotos, D., Quast, I., Li-Wai-Suen, C. S. N., McKenzie, C. I., Robinson, M. J., Kan, A., . . . Tarlinton, D. M. (2021). The concerted change in the distribution of cell cycle phases and zone composition in germinal centers is regulated by IL-21. *Nature Communications,* 12(1), 7160.

Zuasti, A., Jara, Jose,R., Ferrer, C., & Solano, F. (1989). Occurrence of melanin granules and melanosynthesis in the kidney of sparus aumtus. *Pigment Cell Research,* 2(2), 93-99.

Zuba-Surma, E. K., Kucia, M., Abdel-Latif, A., Lillard, J. W., & Ratajczak, M. Z. (2007). The ImageStream System: a key step to a new era in imaging. *Folia histochemica et cytobiologica*, 45(4), 279-290.

Zwollo, P., Cole, S., Bromage, E., & Kaattari, S. (2005). B cell heterogeneity in the teleost kidney: evidence for a maturation gradient from anterior to posterior kidney. *The Journal of Immunology*, 174(11), 6608-6616.