Role of activation-induced cytidine deaminase (AID) in follicular lymphoma biology

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Omar Ahmed Saadoon Alishlash

May, 2017

Abstract

Follicular lymphoma is the second most common non-Hodgkin's lymphoma (NHL). The clinical course of disease is heterogeneous, typically with multiple relapses. Most patients live 10 years or more. However, another group of patients deteriorate rapidly and may progress to death within two years. Activation induced cytidine deaminase (AID) is an enzyme that plays an important role in somatic hypermutation (SHM) and class switch recombination (CSR) of immunoglobulin genes (IG). It induces mutations in IG and non-IG genes leading to genomic instability and chromosomal breaks that are important in the pathogenesis of B-cell malignancies.

In this study, we wanted to first measure AID mRNA and protein levels and its biological function in follicular lymphoma (FL) and then correlate each of these variables with clinical features. Our cohort consisted of 87 patients recruited into the Purine-Alkylator Combination In Follicular lymphoma Immuno-Chemotherapy for Older patients (PACIFICO) trial which is comparing alternative frontline chemoimmunotherapy regimens in older patients with FL. The patient samples were in the form of formalin fixed, paraffin embedded (FFPE) biopsies, which are notorious for nucleic acid degradation. We first chose the best kits for extracting RNA and DNA from FFPE biopsies then optimized the procedure to obtain higher quantity of RNA and DNA from the minimum amount of tissue. We then degraded RNA from an AID positive cell line by heating and compared the degraded material with intact material obtained from the same cells to identify a cut-off point for RNA degradation to be applied in a quantitative polymerase chain reaction (qPCR) experiments. This was followed by a qPCR experiment to identify *AID* mRNA expression in 59 patients. AID protein was then quantified by Immunohistochemistry (IHC) in all samples. We also aimed to measure the functional readout of AID, first by exploring the nuclear/cytoplasmic (N/C) ratio of AID

(AID is stored in the cytoplasm and translocates to the nucleus to function), which was calculated for 20 patients using confocal microscopy. A second AID functional measurement was applied using cloning and PCR to detect ongoing mutation and AID-induced mutation in the immunoglobulin heavy variable gene (*IGHV*) in 18 cases. Finally, we correlated AID expression and functional readouts with available baseline and longitudinal clinical data obtained from the Clinical Trials Unit.

In summary, a significant positive correlation was found between AID mRNA and protein expression (P= 0.001). We also found a significantly higher AID N/C ratio in the patient group with higher total AID mRNA and protein expression (P= 0.025 and 0.023 respectively). No correlation was identified between AID mRNA or protein levels and baseline or longitudinal clinical data. However, AID functionality measured as N/C ratio of AID and AID-related or ongoing *IGHV* mutation was positively correlated with disease status, treatment response and patient survival times.

In conclusion, we found that functional readouts of AID are more strongly associated with adverse clinical features in FL compared to AID mRNA or protein expression.

Acknowledgments

First, I would like to acknowledge my three supervisors: Professor Andrew Pettitt, Professor Sarah E. Coupland and Dr. Ke Lin for their help and support throughout my Ph.D. journey. They built my knowledge step by step to where I am now. Without their precious help and guidance, I would not be able to perform this work.

In addition, I would like to thank all the members of the haematology group in the Department of Molecular and Clinical Cancer Medicine for their support in my laboratory work. Everyone in the group helped me at some point in my Ph.D. I am particularly thankful to Dr. Kathy Till and Dr. Gina Eagle for their help with confocal microscopy and to members of the Liverpool Ocular Oncology Research Group (LOORG), especially Dr. Helen Kalirai, for helping me with immunohistochemistry and immunofluorescent staining of my slides. Also, I am grateful to Dr. Geetha Menon for her help throughout the project.

I want to thank Fotis Polydoros from the Liverpool Cancer Trials Unit (LCTU) for his aid with statistical analyses of chapter 7 of the thesis, as well as Dr. Melanie Oates and the Good Clinical Laboratory Practice (GCLP) technicians for providing me with samples from the PACIFICO trial.

I am grateful to Prof. Robert Sutton and Dr. Muhammad Awais for allowing me to use the confocal microscope, performing the associated training and alleviating the cost.

I thoroughly enjoyed the friendship of many who have supported me during my Ph.D. study. I thank all my colleagues, Dr. Sozan Kareem, Mr. Faris Tayeb, Mr. Moses Lugos, Mr. Jehad Alhmoud, Mr. Venkateswarlu Perikala, Dr. Ishaque Mohammad, Mr. Ahmed Alshatti, Ms. Sofia Karatasaki, Ms. Faten Yasen and Dr. Ola Alsanabra.

I would also like to thank the Higher Committee for Education Development in Iraq for funding this project and Dr. Awaad Farhan, Dr. Miqdam Aljumaily and Mr. Khalid Al-Angood from Al-Sherqat General Hospital for their kind support.

Special thanks should go to my parents Mr. Ahmed Alishlash and Mrs. Alia Hamid, my brothers and sister who have inspired me to be a hardworking person.

Finally, I would like to dedicate my thesis to my wife Dr. Reem Saleh and my sons Mohammad and Zakariya for their limitless patience and their encouragement since the start of this study.

Declaration

With the exception of the statistical analysis of the clinical trial data presented in Chapter 7, which was performed by Mr. Fotis Polydoros (LCTU statistician), all of the work presented in this thesis is entirely my own.

Publications arising from this work

1- Optimizing Technical Protocols to Detect AID Expression in Archived FFPE Samples of Patients with Follicular Lymphoma. Journal of Pathology. March 2017, Volume 241, Issue Supplement S1:S11.

2- Optimization of RNA extraction and RT-qPCR for the quantification of AID in archived FFPE biopsy samples: only two sections of 4µm tissue are required; British journal of haematology. April 2015, Vol. 169: 74.

Presentations arising from this work at University of Liverpool, national and international conferences

1- Omar Alishlash, Ke Lin, Sarah Coupland, Melanie Oates, Helen Kalirai and Andrew Pettitt. AID expression from gene to protein to function and its correlation with clinical data of Follicular Lymphoma. (Oral presentation) 2017, Belfast pathology 2017 meeting, Belfast, UK.

2- Omar Alishlash, Ke Lin, Sarah Coupland, Melanie Oates, Helen Kalirai and Andrew Pettitt. Nuclear AID expression in archived FFPE tissue samples may associate with ongoing *IGHV* mutation and advanced stage in follicular lymphoma. (Poster presentation) 2017, NWCR-University of Liverpool Cancer Research Centre's Annual Scientific Symposium, 28th April 2017, Liverpool, UK.

3- Omar Alishlash, Ke Lin, Sarah Coupland, Melanie Oates, Helen Kalirai and Andrew Pettitt. Correlation of FL clinical features with AID gene and protein expression. (Poster presentation) 2017, Genomics and Oncology Study Day, Liverpool, UK.

4- Omar Alishlash, Ke Lin, Sarah Coupland, Melanie Oates, Helen Kalirai and Andrew Pettitt. Optimizing Technical Protocols to Detect AID Expression in Archived FFPE Samples of Patients with Follicular Lymphoma. (Poster presentation) 2017, Pathological Society of Great England and Ireland winter meeting, London, UK. (This poster received the Cambridge University press poster prize).

5- Omar Alishlash, Ke Lin, Sarah Coupland and Andrew Pettitt. Optimisation of RNA extraction and RT-qPCR for the quantification of AID in archived FFPE biopsy samples: only two sections of 4µm tissue are required. (Poster presentation) 2015, BSH Conference, Edinburgh, UK.

6- Omar Alishlash, Ke Lin, Sarah Coupland and Andrew Pettitt. Validation of qPCR for AID expression in FFPE samples of follicular lymphoma. University of Liverpool poster day (Poster presentation), 11th of April, 2014.

Grants to support attending conferences and courses related to Ph.D. project

1- Grant from Pathological Society of Great Britain to present a poster in Pathological Society winter meeting 19-20 January 2017, the poster won Cambridge University press poster prize.

2- Grant from Pathological Society of Great Britain to participate in the "National Undergraduate Histopathology and Molecular Pathology Introduction" in Leeds, 10-11/April/2017.

3- Full grant from European Association for Cancer Research to attend "Accelerating Cancer Immunotherapy" conference in London at 29-30 March/2017.

4- Grant from European Association for Cancer Research (EACR) to attend Accelerating Cancer Immunotherapy conference in London (19-20/ March/2017)

5- The Royal College of Pathologists travel grant to attend Pathology Summer School in London at 19 –20 August 2016.

6- Grant from European Society for Medical Oncology to participate in ESMO Preceptorship on Immuno-oncology in Sweden, Lund, 4-5 December 2015.

7- Grant from The Organisation of European Cancer Institutes (OECI) to participate in Eurocan Platforms Summer course in Translation Cancer Research, Portugal, October 12th - 16th 2015.

8- British Society for Haematology bursary to present a poster at 55th BSH conference in Edinburgh, Scotland, 20th-22nd April, 2015.

9- European Society for Pathology grant to attend 5th EACR-OECI Joint Training Course (Molecular Pathology Approach to Cancer) in Amsterdam 11 - 13 May 2015.

10- Two free of charge places from “Hartley Taylor Medical Communications Ltd” to attend “Current Treatment Options in Haematological Malignancies” workshop in 2013 and “Manchester Post-ASH Meeting” on 8th February 2017.

Contents

List of Figures

List of Tables

List of abbreviations

ABC	Activated B-Cell germinal centre
AID	Activation Induced cytidine Deaminase
AID	Activation Induced cytidine Deaminase gene
ANC	Absolute neutrophil count
APE	Apurinic-apyrimidinic Endonuclease
ASHM	A Somatic Hypermutation
bp	Base pair
BCR	B Cell Receptor
BSA	Bovine Serum Albumin
CDR	Complementary Determining
CI	Confidence Interval
CLL	Chronic Lymphocytic Leukemia
CML	Chronic Myeloid Leukemia
CPCs	Common Precursor Clones
CR	Complete Response
CRM1	Chromosomal Maintenance 1
CSR	Class Switch Recombination
C_t	Cycle Threshold
CTNNBL1	Catenin Beta Like 1
DAB	Diaminobenzidine
DLBCL	Diffuse Large B Cell Lymphoma
DSB	Double Strand Break

EC	Endogenous Control
EPP	Error Prone Polymerases
ESR	Erythrocyte Sedimentation Rate
FDCs	Follicular Dendritic Cells
FFS	Failure-Free Survival
FISH	Fluorescence In Situ Hybridisation
FL	Follicular Lymphoma
FLIPI	Follicular Lymphoma International Prognostic Index
FLLC	Follicular Lymphoma Like Cell
FLIS	Follicular Lymphoma Insitu
FR	Framework
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase gene
GC	Germinal Centre
GCB	Germinal Centre B-cell-like of diffuse large B-cell lymphoma
gDNA	genomic DNA
GELF	Groupe d'Etude des Lymphomes Folliculaires
GIT	Gastrointestinal Tract
HCT	Hematopoietic Cell Transplantation
HIV	Human Immunodeficiency Virus
HPF	High Power Field
HR	Hazard Ratio
HRP	Horseradish Peroxidase
IF	Immunofluorescence
IG	Immunoglobulin
IGH	Immunoglobulin heavy gene

IGHV	Immunoglobulin heavy variable gene
IGV	Immunoglobulin Variable gene
IMGT	International ImMunoGeneTics information system
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IHC	Immunohistochemistry
IMS	Industrial Methylated Spirits
kb	Kilobase
LB	Luria-Bertani agar
LCTU	Liverpool Cancer Trials Unit
LDH	Lactate Dehydrogenase
LN	Lymph Node
LBIH	Liverpool BioInnovation Hub
mAb	Mono Clonal Antibody
MHC	Major histocompatibility complex
MMR	Mismatch Repair
MRD	Minimal Residual Disease
mTOR	Mammalian Target Of Rapamycin
N/C	Nuclear/cytoplasmic
NGS	Next Generation Sequencing
NHL	Non Hodgkins Lymphoma
OR	Odds ratio
OS	Overall Survival
PACIFICO	Purine-Alkylator Combination In Follicular lymphoma Immuno-Chemotherapy for Older patients.
PBS	Phosphate Buffer Saline

PCR	Polymerase Chain Reaction
PD	Progressive Disease
PET	Positron emission tomography
PFS	Progression Free Survival
PI3K	Phosphatidylinositol 3-kinase
qPCR	quantitative Polymerase Chain Reaction
PR	Partial Response
PTM	Post-translational Modification
R	Rituximab
RAG	Recombination Activating Gene
RIN	RNA Integrity Number
RPM	Revolutions Per Minute
RPMI	Roswell Park Memorial Institute
RR	Response Rate
SD	Standard Deviation
SDS	Stable Disease
SEM	Slandered Error of Mean
SHM	Somatic Hypermutation
SOC	Super Optimal broth with Catabolite repression
TBS	Tris buffered saline
TLR	Toll-like Receptors
TME	Tissue Microenvironment
TP53	Tumour Protein p53
UK	United Kingdom
U/L	Unit per Liter

UNG	Uracil DNA Glycosylase
VDJ	Variable, Diversity and Joining gene segments
WBC	White Blood Cells
WHO	World Health Organization

Chapter 1

General introduction

1.1 Normal B-cell biology:

My thesis addresses the role of activation-induced cytidine deaminase (AID) in follicular lymphoma (FL), which is a B-cell non-Hodgkin's lymphoma (NHL) and the most common indolent B-cell lymphoma (1).

To understand the role of B-cell development and AID in this disease, it is essential to initially discuss B-cell biology (Section 1). The details of FL will be discussed in Section 2, and AID in Section 3 of this introduction.

1.1.1 B-cell differentiation, maturation and immunoglobulin gene recombination:

B-cells are essential in the immune system. They carry the B-cell receptor (BCR) that identifies antigens, allowing these cells to produce specific antibodies against them (2).

There are two major stages of B-cell development; one is antigen independent, occurring in bone marrow, and the other is antigen dependent, located in the lymph node (LN) (3).

B-cells originate from haematopoietic stem cells in the bone marrow, which gives rise to progenitor B-cells. BCR formation begins at this stage, by combining diversity (D) and joining (J) gene segments with immunoglobulin variable region (*IGV)* gene segments, followed by joining one of the constant genes (4). BCRs consist of heavy and light chains, and are initially expressed with light chain-like molecules (surrogate light chains) to form a pre-B-cell receptor (5). The stages of B-cell development are illustrated in Figure 1.1.

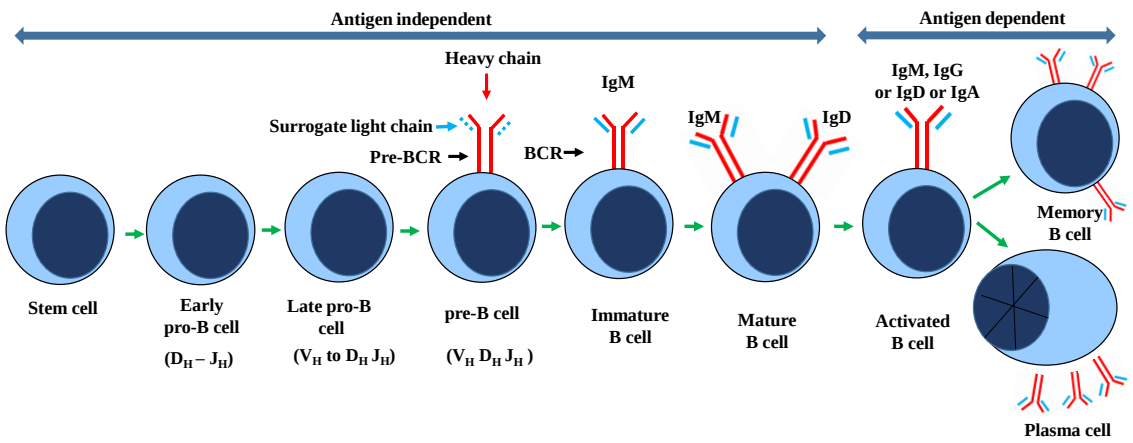

Figure 1.1: Stages of B-cell development from stem cells to memory or plasma cells. Figure based on reviews by Elgert and Parham (6, 7).

1.1.2 Somatic hypermutation in immunoglobulin variable genes:

Somatic hypermutation (SHM) is the process of generating highly specific antibodies against foreign antigens by inducing mutations (mainly as single nucleotide replacements) in *IGV* genes. SHM occurs at a frequency of 10^{-2} to 10^{-3} in immunoglobulin heavy variable (*IGHV)* and immunoglobulin light variable **(***IGLV*) regions (8). AID is a key enzyme in this process (9).

The targeted *IGHV* gene is approximately 1.25 mega base in length. It is located in the *IGHV* gene cluster on chromosome 14q, and consists of four framework regions (FR) and three complementarity-determining regions (CDR) (10, 11).

In addition to SHM, the process of VDJ recombination occurs in early stages of B-cell development and further diversifies BCR to identify antigens. This recombination occurs between any of the 50 V (variable), 30 D and 6 J gene segments, creating the possibility of up to 10^4 different combinations with any of the constant genes (12). *IGHV* gene rearrangement and SHM are shown in Figure 1.2.

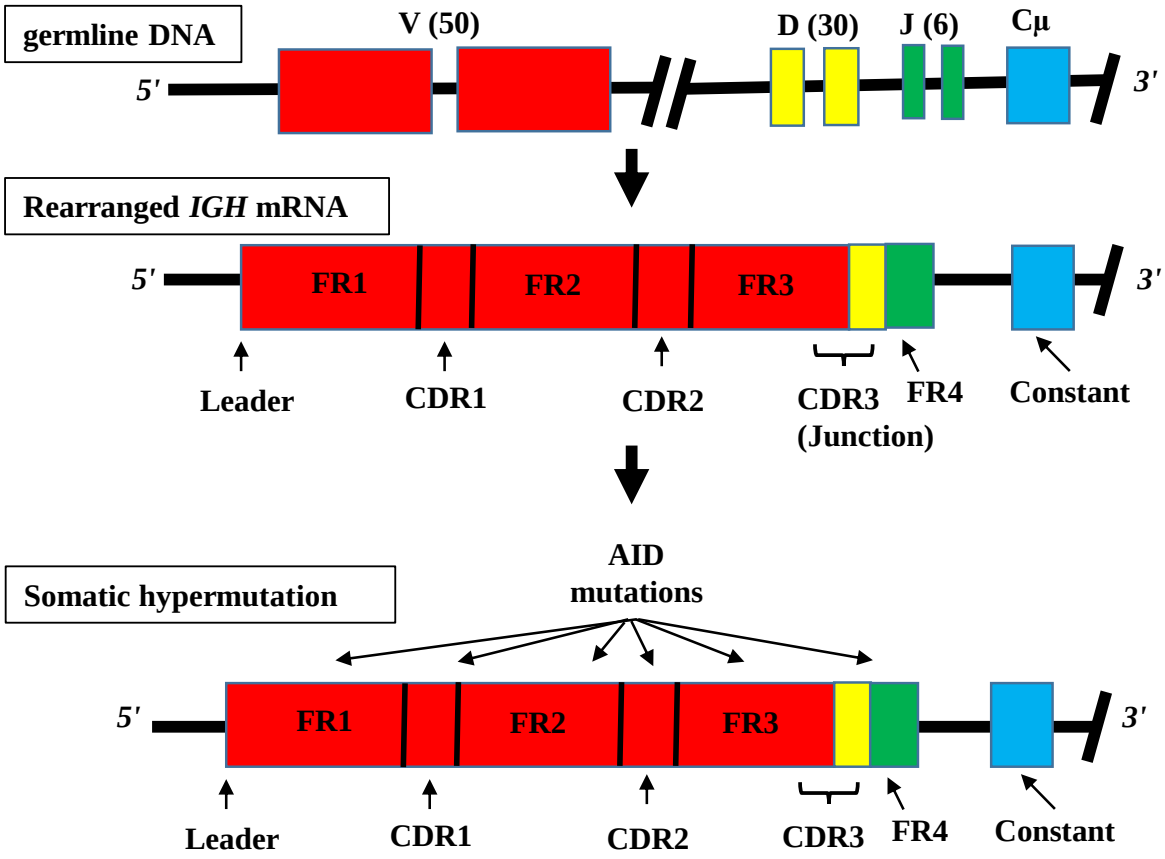

Figure 1.2: *IGHV* gene rearrangement, VDJ recombination occurs in bone marrow, rearranged mRNA contains 4 FR regions and 3 CDR regions. Figure based on Figure 1 of McClure and colleagues' work (2) and Figure 7 of Gagyi Eva Ph.D. thesis (13).

BCR created by VDJ recombination plays a role in the innate immune system. This receptor is expressed on the B-cell surface by adding either kappa (*IGK*) or lambda (*IGL*) light chains to the heavy chain (14). On the other hand, in comparison with the innate immune system, the germinal centre (GC) B-cell *IGHV* mutations produce antibodies with higher specificity (2).

Germinal centre-associated nuclear protein is responsible for directing AID to the *IGV* gene during the process of SHM (15), while the SPT5 protein is important for AID recruitment toward both immunoglobulin (IG) and non-immunoglobulin genes (16). Then, AID cofactor

replication protein A helps AID to induce conversion of deoxy Cytidine to deoxy Uridine within the *IGHV* gene during transcription (17).

In B-cell malignancies, a B-cell clone with unique *IGHV* rearrangement will be produced. This *IGHV* can be amplified by PCR reaction, and adds to microscopic morphology for confirming diagnosis in difficult cases of B-cell malignancy (18).

1.1.3 Class-Switch Recombination:

Class-Switch Recombination (CSR) is the process of exchanging the constant region (Cμ) in the immunoglobulin heavy gene (*IGH)* with either Cγ or Cε or Cα, allowing B-cells to produce antibodies of IgG, IgE, or IgA, respectively, rather than IgM (19). The AID enzyme is the keystone for CSR. It deaminates cytosine in two switch regions of the *IGH* gene, then double-stranded DNA breaks occur by a DNA repair mechanism (20) (Figure 1.3). AID can be directed to switch regions in the *IGH* gene by 14-3-3 adaptor proteins (21).

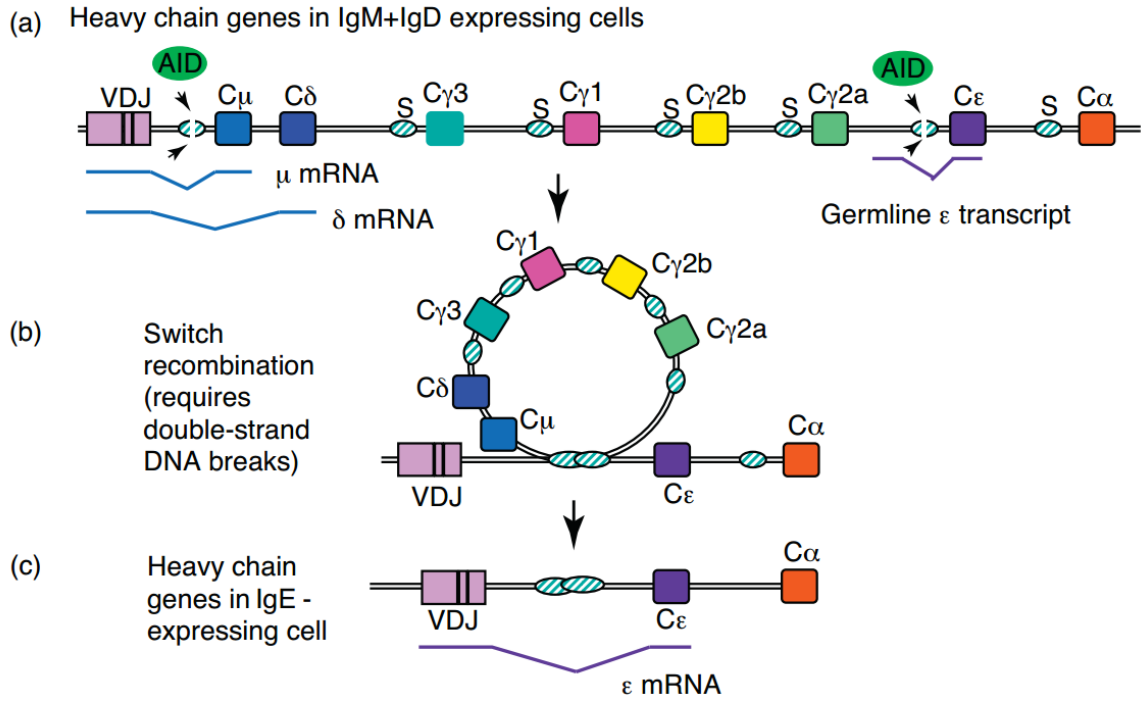

Figure 1.3: CSR from IgM to IgE. AID enzyme deaminates cytidine residues in the S region of the *IGH* gene to create a DNA double-strand break. This process is followed by deletion of the chromosomal part between VDJ and required constant segment (in this case, Cε to produce IgE Antibody). Figure taken with permission from Stavnezer and colleagues (22).

1.1.4 The germinal centre reaction:

This is the reaction in the lymphoid follicle initiated by antigens, and results in a highly specific antibody response against these antigens (23).

The reaction starts when the cells of the innate immune system catch an antigen, and present some of its components on their cell surface. Then, these cells migrate to lymphoid organs where they are recognized by T-helper cells. At the same time, B-cells, upon antigenic stimulation, migrate to and bind T cells to initiate the GC reaction, which produces specific antibodies against the offending antigen (24, 25).

B-cells proliferate extensively, forming the dark zone of secondary lymphoid follicle, which consists of centroblasts surrounding a less condensed zone of cells called the light zone (26). After that, B-cells enter the light zone to become centrocytes and the proliferation rate decreases. It is the place where only the higher affinity antibody-producing B-cells are selected to survive, while cells with low affinity are directed for apoptosis (27). The end result of this GC reaction will be specific antibody-producing plasma cells and memory B-cells (28) (Figure 1.4).

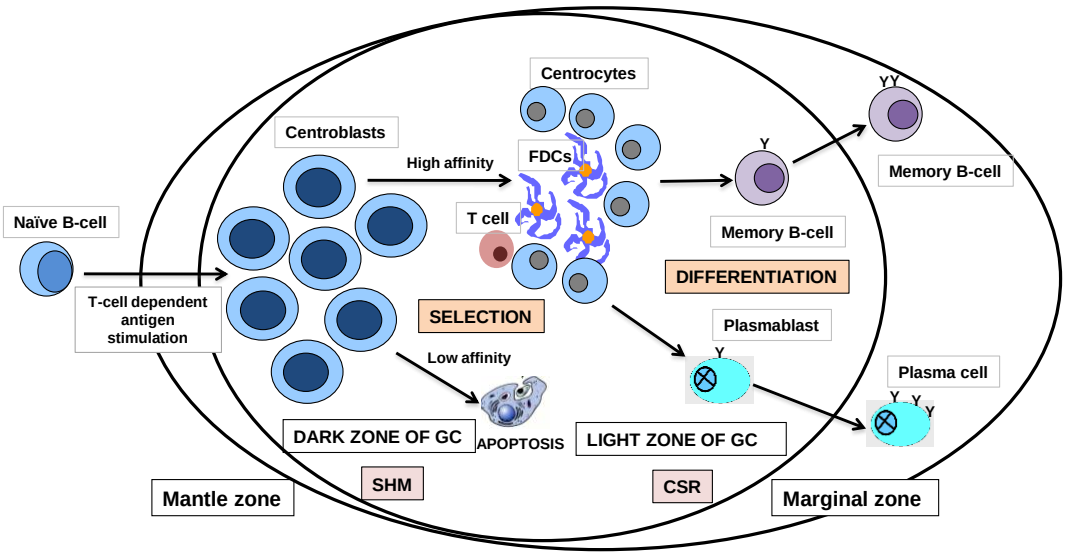

Figure 1.4: Germinal Centre reaction. After encountering an antigen, T cell and B-cell interaction stimulates B-cells and directs them to the GC. B-cells proliferate at the dark zone of the GC and undergo SHM with the help of AID. If the B-cells have high affinity against antigen, then they will enter the light zone and proliferate. Otherwise they will undergo apoptosis. B-cells will then become either memory B-cells or plasma cells and produce specific antibodies (29). Figure courtesy of Professor Sarah E. Coupland.

Other essential cells in the GC, in addition to B-cells and stromal cells, include:

1- Follicular dendritic cells (FDCs), which play at least two roles in GC reactions. First, they prolong retention of antigen to help the SHM process in selecting B-cells with highest affinity against antigen (30). Second, FDCs maintain B-cell survival and proliferation (31).

2- Macrophages, which contain apoptotic bodies of dead B-cells (32).

3- T cells, which play a pivotal role in formation and maintenance of the GC. They help B-cells to initiate the GC reaction, trigger migration cycles of B-cells between dark and light zones to achieve high specificity against antigen, and increase B-cell proliferation (33, 34).

In addition to these cells, AID and BCL6 play an important role in GC reactions. AID induces the SHM process to increase antibody affinity (35), and is highly expressed in the GC (36). In addition, the AID enzyme may induce genomic instability leading to B-cell lymphomas in long-lived GC associated with chronic infections (37).

On the other hand**,** *BCL6* is a proto-oncogene and a transcription repressor that plays a major role in the creation of GC reactions (26, 38). BCL6 increases cell tolerance to GC-associated DNA damage and prevents cells from apoptosis by suppressing DNA damage checkpoints (38, 39). Mutations and translocations of the BCL6 gene are associated with lymphomas (40).

1.2 Follicular lymphoma:

1.2.1 Definition and epidemiology of lymphoma:

Lymphomas involve malignant proliferation of lymphocytes, and can occur at any stage of lymphocyte maturation. About 80% arise from B lymphocytes, and the remaining lymphomas arise from T lymphocytes and natural killer cells (41, 42).

The World Health Organization (WHO) classified lymphomas into two major groups: Non-Hodgkin's lymphomas (NHL) forming about 90% of all lymphomas, and Hodgkin's lymphomas covering the remainder. Both types consist of more than 50 lymphoma subclasses (41).

In NHL, eleven major subtypes were identified with different geographical distribution over the world. The two most common types are diffuse large B-cell lymphoma (DLBCL) and FL, both comprising more than two thirds of all Non-Hodgkin's lymphomas (41, 42).

In the UK, NHL is the fifth most common cancer. The annual incidence reached 23 per 100,000 males and 19 per 100,000 females. In 2013 alone, about 10,000 new cases of NHL had been registered, of which 1,884 were FL, according to the UK Office of National Statistics (43).

More than 20 subtypes of lymphoma have been registered in the UK, and in almost all of these, males have a higher incidence rate and are affected at younger ages (44). To distinguish from the disease in children, the updated WHO classification of lymphoma includes paediatric-type FL as a definitive entity (45).

FL predominantly affects patients aged between 60 and 70, with overall survival (OS) shortening in older-age groups. Unlike most other NHLs, FL is more common in women than men (46, 47). However, the progression free survival (PFS) is longer in women, and both OS and lymphoma-related mortality are better in female patients younger than 60 years of age (48).

Regarding geographical distribution, the highest incidences of FL are in the developed world and the lowest in Asia (49).

1.2.2 Aetiology of follicular lymphoma:

Despite intensive efforts into determining the causative factors of FL, the aetiology of this disease remains unclear, and needs to be investigated in larger multicentre studies (50).

To date, FL has been correlated with different factors, examples include:

Farrell and colleagues suggested that using immunosuppressive drugs is correlated with increased risk of NHL (50); Morton and colleagues showed that only the FL type of NHL is higher in smokers than non-smokers (51).

Viral infections – e.g. Epstein-Barr virus (EBV), human herpesvirus-8 (HHV-8) and Human T-cell leukemia virus-1 (HTLV-1) - were all suggested as causing genomic changes which facilitate lymphoma occurrence (52).

Furthermore, lack of daily exercise, obesity, increased ingestion of milk and meat were suggested to be associated with an increased incidence of FL (53). This was supported by data suggesting that the intake of sufficient antioxidants, fruit, vegetables and vitamin D were correlated with lower risk of FL (53).

Hair dyes used prior to 1980 were also proposed to lead to an increased risk of FL, with only black colour hair dyes after 1980 being associated with an increased risk (54).

Exposure to pesticides and chemical solvents such as benzene, toluene and xylene were also found to be associated with an increased incidence of FL (55, 56).

1.2.3 Pathology of FL:

The development of FL is hypothesized to require multiple genomic aberrations. The early precursors of FL may include variants, such as “follicular lymphoma in situ”, duodenal follicular lymphoma and t(14;18)-positive circulating B-cells.

It was proposed that t(14;18), which is the hallmark of FL, is the initial and underlying genetic alteration (see below). This is followed by ‘second hits’, or additional genetic aberrations (discussed below). It was hypothesized that the risk for subsequent mutation that resulted in FL development occurs when cells re-enter the GC and are exposed to the effect of AID and somatic hypermutations (57).

1.2.3.1 t(14;18) in FL:

The most common genetic aberration in FL is the translocation of genes on chromosomes 14 and 18, i.e. t(14;18), which occurs in about 85-90% of cases. This results in the translocation of the *BCL2* gene from chromosome 18 to the enhancer loci of the *IGHV* gene on chromosome 14 (58). This leads to increased expression of the anti-apoptotic BCL2 protein, and therefore facilitates the survival and accumulation of lymphoma cells initially within the germinal centre (Figure 1.5) (58, 59). Other rare translocations in FL, such as t(2;18) and t(18;22), also lead to a dysregulation of the *BCL2* gene and the abnormal overexpression of its protein (60).

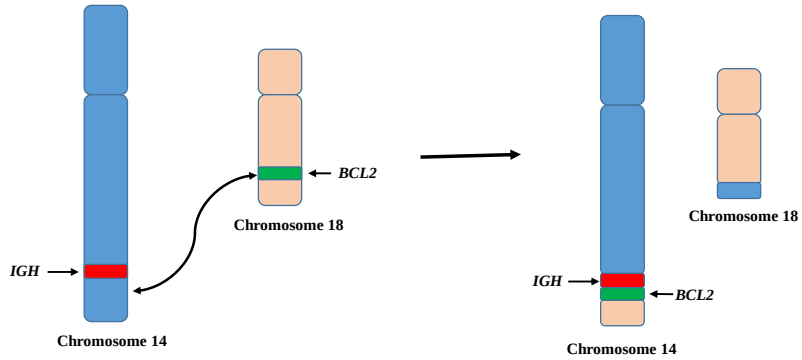

Figure 1.5: t(14;18) present in 85-90% of FL patients. The *BCL2* gene is translocated from chromosome 18 to chromosome 14 to be enhanced by *IGH* gene. This leads to elevated expression of anti-apoptotic BCL2 protein. Figure based on translocation review of Blacker and Leich papers (58, 59).

This t(14;18) can occur as a result of a defect in repair of DNA double strand breaks, mediated by recombination activating gene (RAG) in either both *BCL2* and *IGH* loci, or the latter gene alone. It can also result from CpG gene segment fragility in *BCL2* (59).

In addition to RAG, AID was also found to contribute to t(14;18) in B-cells (61). The expression of AID in B-cells is associated with double-strand breaks (DSB) in points across the whole genome (62).

Furthermore, AID affects other oncogenes: about 25% of expressed genes in B-cells were not fully protected against AID mutagenic effects (63).

Breaks in the *BCL2* gene are located in a number of locations across the gene. The most important break points are major breakpoint regions, minor cluster regions and intermediate cluster regions (64-66). Other uncommon breakpoints are the 3' *BCL2* (67) and 5' ends of *BCL2* (68). These points can be detected by PCR to provide a rapid, sensitive and reproducible method for minimal residual disease (MRD) detection.

Fluorescence In Situ Hybridisation (FISH) is the preferred method for routine detection of t(14;18) (69). In addition to t(14;18), more genetic hits are required for FL development (70).

1.2.3.2 Other genomic aberrations may also contribute to FL pathogenesis:

It is hypothesised that t(14;18) occurs during the early stages of FL. Subsequently, the t(14;18)+ B-cells acquire other mutations in genes that ultimately lead to FL. Later on, other genomic aberrations – e.g. loss (or mutation) of tumour protein p53 (*TP53)* and *CDKN2A* may occur that cause the transformation of FL from a low-grade to a high-grade or aggressive disease (71).

Besides t(14;18), genomic aberrations in FL include losses of 1p, 6q, 10p, 17p, genetic gains in chromosomes: 1,3q, 6p, 7, 8, 12q, X and 18q, t(2:18) and t(18;22) (72).

In addition to structural aberration of chromosomes, mutations in multiple genes contribute to FL. The most important genes affected are histone modifiers, including *EZH2*, *MLL2*, *CREBBP*, *MEF2B*, and *EP300* and *TNFRSF*14. Deletion of tumour suppressors *TNFAIP3/A20* and *EPHA7* are associated with pathogenesis of this disease (73). A recent study revealed that recurrent somatic mutations in the *RRAGC* gene, which encodes an important protein in the mTOR (Mammalian target of rapamycin) complex, occurred in 17% of patients with FL, suggesting a possible role of these mutations in survival of the neoplastic B-cell clones, and also the possibility of targeting them in FL treatment (74).

1.2.3.3 Early lesions and clonal expansion in FL:

It was thought that the malignant FL cells are derived from transformed GC lymphocytes. This hypothesis was based on clinical, microscopical and molecular similarities between normal GC and FL cells. However, some studies suggested that FL occurs after years of slow growth of common precursor clones (CPCs). These clones might also be responsible for disease relapse after treatment (75).

It was proposed that t(14:18) that occurs in bone marrow in the ‘pre-B’ cells is the initiating step in the pathogenesis of FL, rather than in the GC (76-79). Nevertheless, the observations that t(14:18) is present in normal individuals, and that at least one additional mutation occurs

in almost all FL patients, would suggest that t(14:18) is a crucial step for the pathogenesis of the disease, but that it is not necessarily sufficient – i.e. additional hits are required for a malignant transformation to commence (78, 80, 81).

It was suggested that in early stages of FL development, AID targets genes other than *IGH,* especially proto-oncogenes involved in FL pathogenesis in a process named aberrant somatic hypermutation (ASHM). These genes include *MYC, BCL6, PAX5, ARHH* and *PIM1*(59, 82).

It was also hypothesised that the origin of FL could be a group of cells named "follicular lymphoma in situ (FLIS)", which express BCL2 protein found accidently in the GC, with no evidence of dissemination and risk of about 5% of evident FL (83).

Another cell suggested to be the origin of FL are circulating t(14;18) positive cells, which are termed FL-like cells (FLLC).

FLLCs are reported to accumulate within the GC in 14% of reactive LN specimens as non-proliferative centrocytes (84). Molecular testing confirmed a clonal relationship between FL and FLLCs in some patients (69), and it was suggested that these cells may re-enter the GC to undergo multiple SHM reactions, increasing the chance for additional genomic hits required for conversion to overt FL (85).

1.2.3.4 Ongoing mutations in FL:

FL is a counterpart of normal GC B-cells that maintains its reactions including functional BCR and ongoing somatic hypermutation. Ongoing SHM leads to intraclonal variability in FL, and both SHM and BCR stimulation are necessary for FL viability and probably arise from lectins that react with IG variable regions on the surface of FL cells (86-88).

Loeffler and colleagues studied the *IGHV* gene in different clones in FL by next generation sequencing (NGS) and confirmed the ongoing SHM process. They also suggested that a functional BCR and continuous AID stimulation induced ongoing *IGHV* mutation in

addition to ASHM to other important genes involved in FL pathology. This is important for genomic and epigenetic evolution and clonal diversity in FL (88).

IGHV ongoing mutation was studied mainly by NGS in diagnostic, relapsed or transformed FL sequential samples, and it was found that relapsed or transformed FL mainly occur by divergent evolution from CPCs (70, 88-90).

The first event in FL pathogenesis is suggested to be t(14;18). This is followed by acquiring additional mutations to make CPC population of FL precursor cells. Studies recently suggested that transformed FL arises from CPC by divergent evolution rather than by acquisition of new mutations in the same pathogenic clone (linear evolution) (Figure 1.6) (70).

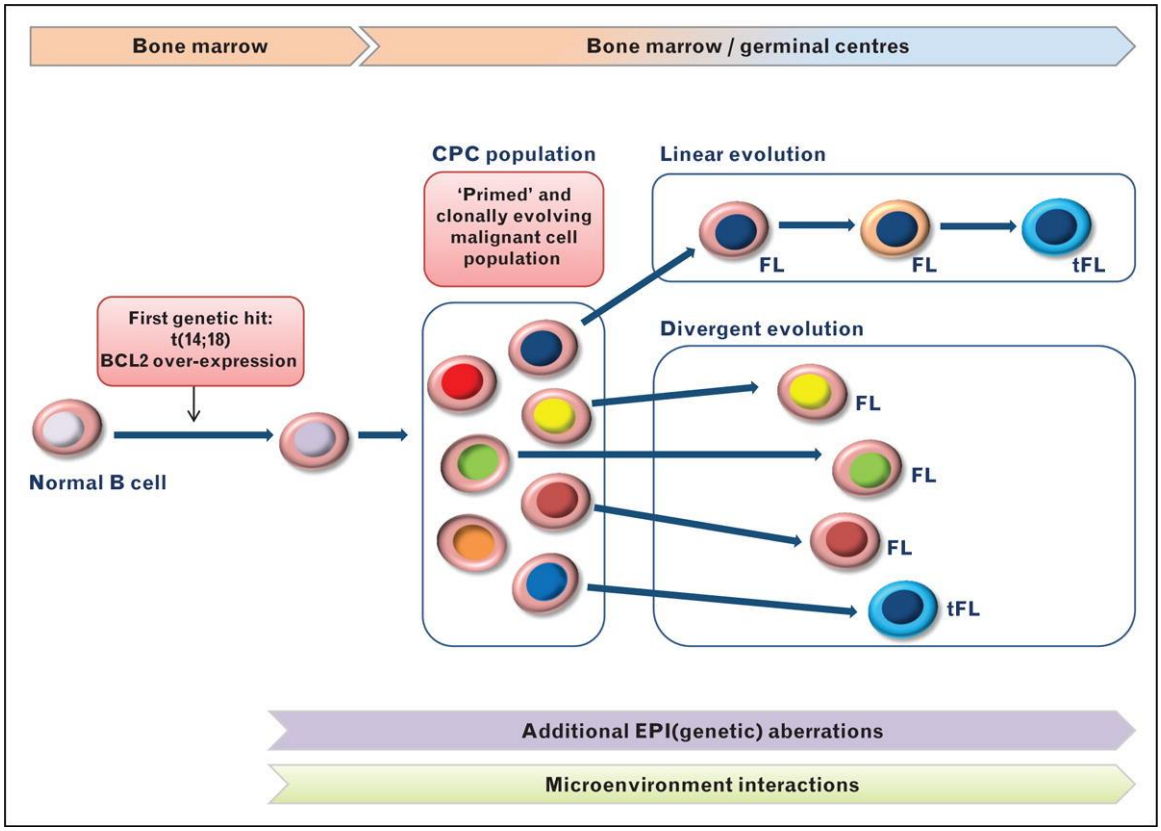

Figure 1.6: Ongoing mutation and divergent evolution of CPC clones in FL transformation and relapse. Ongoing SHM was studied in FL and transformed biopsies of the same patients. It was found that the branched divergent evolution rather than accumulation of mutations in the same clone was the principal route of relapse in many cases. Figure taken with permission and legend adapted from Okuson's review (70).

1.2.3.5 The tumour microenvironment in FL:

In B-cell lymphomas, the tumour microenvironment (TME) is a key player in their pathogenesis: this can be done by playing a role in controlling tumour cell proliferation and survival, helping malignant cells to escape immune surveillance and develop drug resistance (91).

TME is composed of vasculature, cancer-associated fibroblasts, extracellular matrix and immune cells such as B and T lymphocytes, macrophages, FDCs, mast cells, neutrophils, and natural killer cells (92, 93).

Chemicals released by TME cells can induce drug resistance, in addition drug therapy can be affected by adhesion between malignant cells to extracellular matrix or fibroblast, and effective malignant eradication probably can be achieved by targeting both malignant cells and specific TME components (92).

The relationship between malignant cells and TME in B-cell lymphomas can involve any of the following three models:

1- Re-education, as in the case of FL, where malignant cells are dependent on the microenvironment for growth and survival; the microenvironment here is similar both in arrangement and composition to normal lymphoid tissues.
2- Recruitment, as in classical Hodgkin's lymphoma, the malignant Reed-Sternberg cells recruit and get benefit from non-malignant supporting cells that are not similar in composition to normal lymphoid tissues.
3- Effacement, as in Burkitt's lymphoma and DLBCL, whereby genetic divergence, such as *MYC* translocation in malignant cells leads to self-determining, microenvironment-independent growth and survival (94).

In FL, prognosis can be predicted from the composition of TME at diagnosis. Gene expression profiling divides TME into two immune signatures that predict survival (95).

In addition, a correlation between different components of TME and malignant cells is important for predicting malignant transformation of FL to DLBCL or length of survival.

A high percentage of tumour-infiltrating lymphocytes in FL that are PD1+ are less likely to undergo high-grade malignant transformation (96). Exhausted T cells in TME are also associated with poor survival (97).

Time of transformation to DLBCL at diagnosis can be associated with high intrafollicular CD4+ T cells, CD68+ and PD-L1+ macrophage infiltration (98).

Survival is decreased if the malignant follicles contain fewer dendritic cells and more T regulatory cells compared to hyperplastic lymphoid tissues (99).

Malignant B and TME cells express programmed cell death ligand (PD-L1). This ligand binds to programmed cell death 1 (PD-1) of T cells. The PD-L1/PD-1 binding rescues cancer cells from T‑cell-mediated cellular cytotoxicity. Immune checkpoint inhibitors prevent this binding (100).

Other therapeutic agents targeting the TME include chimeric antigen receptor (CAR) T-cell therapy, B-cell receptor signalling inhibitors and immunomodulatory drugs (91).

1.2.4 Clinical features of FL:

Painless lymphadenopathy, especially of cervical and supraclavicular nodes, is the most common presentation of FL, with bone marrow involvement usually occurring in about half of the patients at diagnosis (101).

Involvement of retroperitoneal or mediastinal LNs can cause pressure effects on adjacent organs with large-sized nodes. Furthermore, dissemination of malignant FL cells to the circulation, bone marrow, spleen, liver and gastrointestinal tract (GIT) lymphoid tissues can occur (102, 103).

Exranodal involvement of the GIT, skin, lung, thyroid gland, ocular adnexa and other organs has been reported (104).

The disease course differs among patients. Most commonly it runs an indolent course with multiple relapses and fluctuating lymphadenopathy over years; and about 80% of patients survive for about 8 years or more even without treatment (105, 106), while in others the disease becomes aggressive with rapid deterioration and sometimes death occurs rapidly (95).

1.2.5 Laboratory features:

1.2.5.1 Blood findings:

In about 10% of FL patients, a high number of lymphomatous cells are found in the blood – i.e. a FL 'leukaemia' is seen (107). In the peripheral blood smear, these cells are small with a high degree of pleomorphism, with nuclear notches or clefts and a small rim of cytoplasm (108). Leukopenia, anaemia, a high erythrocyte sedimentation rate (ESR) and lactate dehydrogenase (LDH), are other possible serological findings associated with FL (109).

1.2.5.2 Radiological investigations and imaging techniques:

The simple, rapid and low cost diagnostic imaging techniques, such as conventional radiography and ultrasound, are yet to be frequently used in lymphoma clinical diagnostics. Radiological techniques mentioned below are used to check for deep nodal masses, whilst ultrasound is helpful for examining superficial and deep LNs in addition to abdominal visceral organs (110).

The more advanced radiological techniques applied for the deeply-located masses include computed tomography (CT), used for both diagnostics and response to therapy assessment, are more commonly used compared to a decade ago (111). These also include positron emission tomography (PET) together with 18-F-fluorodeoxyglucose (FDG-PET) or with computed tomography (PET-CT). The latter techniques are regarded as the most sensitive

and specific imaging techniques for staging, detecting therapy response and disease surveillance in NHL (112). More recently, whole body magnetic resonance imaging (MRI) is suggested as an attractive technique for FL early diagnosis, staging and detecting bone marrow involvement (113-115).

1.2.5.3 Histopathological diagnosis of FL:

The diagnosis of FL only can be confirmed by histopathological examination of biopsied or excised LN; fine needle aspiration is usually not sufficient for reaching an unequivocal diagnosis (116).

The biopsy should be adequate and core biopsies are only to be considered if excisional biopsies cannot be performed. Compared to excisional biopsies, core biopsies are less reliable in achieving a definitive diagnosis, and particularly for identifying the grade of the FL (117).

Histopathological findings of a typical LN in FL includes: complete or partial effacement of the LN and replacement of its usual architecture by back-to-back atypical follicles, composed mainly of expanded neoplastic GC composed of large centroblasts with prominent nuclei and smaller centrocytes (Giemsa stain). The GCs are surrounded by an attenuated mantle zone (118, 119) (Figure 1.7).

1.2.5.4 Immunohistochemical staining:

A full or mostly partial panel of Immunohistochemical (IHC) staining in FL must be done to confirm the diagnosis before starting therapy.

FL biopsies upon IHC staining are usually positive for B-cell markers (CD20 or CD79a) with the neoplastic cells also staining for CD10, BCL6 and BCL2. Only few CD3 positive T cells are seen in the GC (Figure 1.7) (93). They are negative for CD5 and CD23, differentiating them from chronic lymphocytic leukaemia (CLL), and for Cyclin D1,

differentiating them from mantle cell lymphoma (120-122). BCL2 protein is abnormally expressed by the neoplastic B-cells in the GCs in 65%-85% of FL cases (121-123).

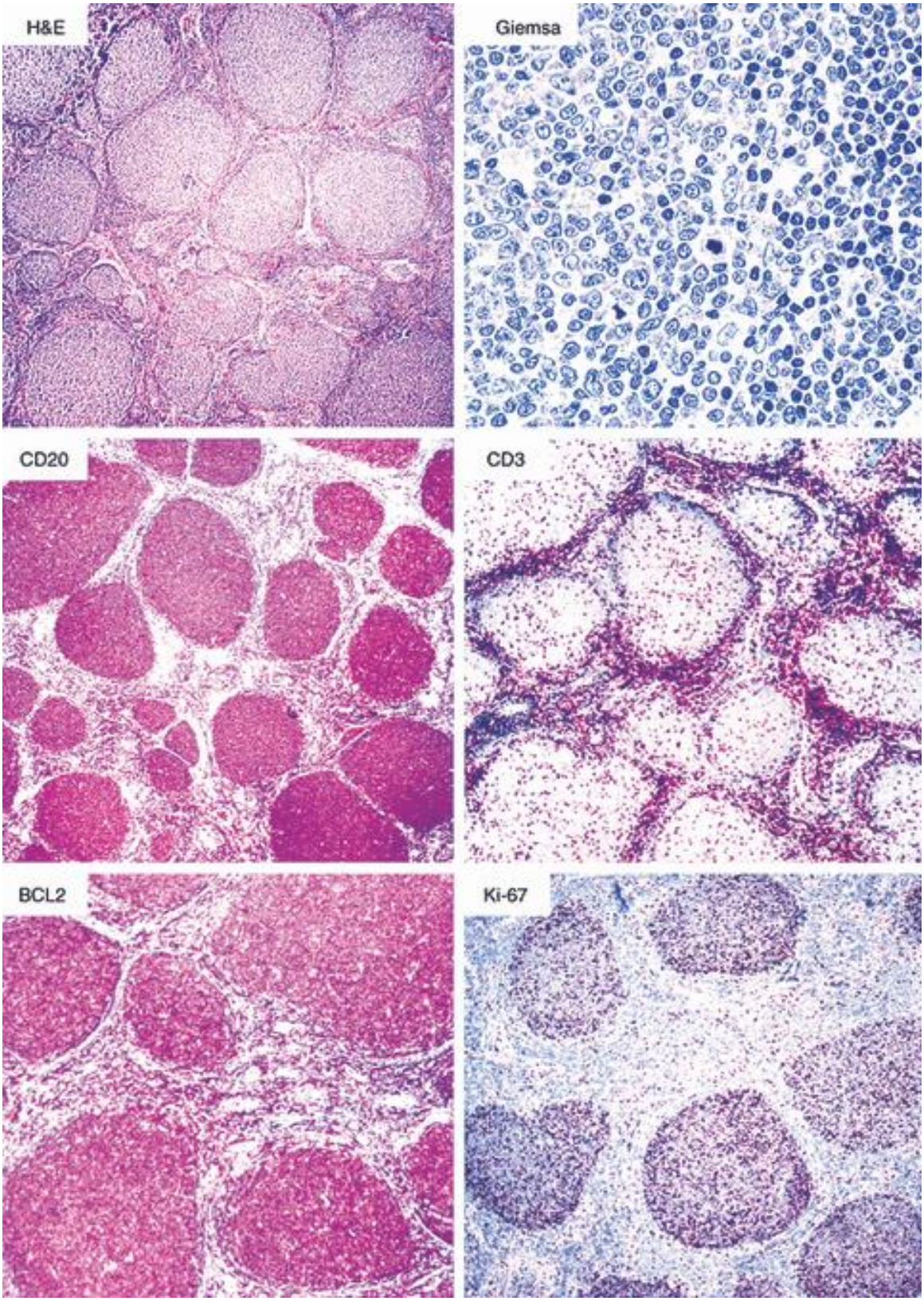

Figure 1.7: Histopathology of FL, haematoxylin and eosin shows nodular follicles of malignant FL cells. These cells are positive for CD20 and BCL2 antibodies. CD3-stained T cells in the interfollicular area and a few T cells in the GC. Ki67 is usually used to illustrate proliferation in FL. Figure taken from and legend based on Coupland's paper (93).

1.2.5.5 Grading of FL:

FL are categorised according to the WHO classification into three grades, FL grade 3 is further divided into 3A and 3B:

Grade	Histopathological microscopic criteria
Grade 1	0-5 centroblasts per high power field (HPF)
Grade 2	5-15 centroblasts per HPF
Grade 3	>15 centroblasts per HPF
Grade 3A	Centrocytes are present
Grade 3B	Solid sheets of centroblasts

Table 1.1: Grading of FL (41).

It is recommended that at least 10 high power fields (HPF) are viewed for the grading of FL. Hence, it is difficult to ascertain the grade of an FL in small fine needle biopsies.

1.2.5.6 Cytogenetics and molecular genetics:

Investigating genetic abnormalities is performed as routine practice in diagnosing B-cell lymphomas: FISH can be performed on fresh or formalin-fixed paraffin-embedded (FFPE) tissue to confirm the presence of t(14;18). *IGHV*-polymerase chain reaction (PCR) may be undertaken to demonstrate clonality; however, due to the large number of somatic mutations, the primers may not bind, leading to false negative results. Further tests that may be undertaken, include analysis of further translocation breakpoints, microRNA profiles, gene

copy number variations, as well as NGS; however, these are undertaken more in the research area rather than for diagnostic practice (72).

1.2.5.7 Next generation sequencing:

NGS is a technique for DNA sequencing that revolutionized biology in the last decade; it increased speed, accuracy, diversity of genome sequencing and enabled sequencing the whole human genome at a cost of around $1,000 (124).

Some advantages of NGS over older sequencing methods are, preparing NGS libraries instead of laborious bacterial cloning, multimillion sequencing reactions could be run at the same time instead of hundreds only by older sequencing techniques, and results are seen directly without the need for electrophoresis (125).

However, an enormous amount of NGS data needs careful examination and analysis. For this reason, bioinformatics is considered an important part of the team generating NGS data, especially in laboratories applying NGS for clinical genetic testing (126).

In FL, NGS research has allowed major discoveries in understanding disease pathology, and has enabled a better expectation of prognosis. It also facilitates understanding clonal evolution, disease transformation, enabling MRD detection and researching future targeted therapies in FL (72, 88, 127-129).

1.2.6 Ann Arbor staging of FL:

For the staging of FL, the Ann Arbor staging system is used (130). As shown in Table 1.2, this system involves the following: the number and locations of involved lymph node site as detected by physical examination and CT scanning, particularly of the abdomen and pelvis; a bone marrow aspirate and/or trephine; a complete blood picture and serological assessment of LDH, uric acid; and virological testing for human immunodeficiency virus (HIV),

hepatitis B and C virology screen. PET-CT increases the precision of staging of FL and needs to be used routinely (116).

Stage	Definition
Stage I	single LN group
Stage II	Two or more LN groups either above or below the diaphragm
Stage III	Involvement of lymph regions or structures on both sides of the diaphragm
Stage IV	Extranodal involvement
	In any stage of disease
A	No symptoms
B	Fever (>38ºC), drenching sweats, weight loss of more than 10% of body weight within 6 months

Table 1.2: Staging of FL (116, 130).

1.2.7 Prognosis of FL:

About 80% of FL patients can live with their disease for 10 years; however, some patients respond poorly to therapy and die early within two years (105, 106).

Those patients with an indolent course have periods of disease progression with LN enlargement and systemic symptoms, and respond less well to therapy with each relapse (131).

Studies have identified several prognostic factors for FL: these include patient age and gender, size of tumour, extent of systemic manifestation, LDH levels, tumour stage, and β2-microglobulin levels (132).

Different prognostic algorithms have been proposed for FL. The most well accepted and widely used algorithms is the "Follicular Lymphoma International Prognostic Index" (FLIPI) score. It was established following a study on 4,167 patients, which took into account several clinical factors, including patient age, the Ann Arbor stage, LDH levels, the number of involved LNs, and haemoglobin level (133, 134). A revised version of the test called (FLIPI 2) was subsequently introduced (132), whereby different LN measurements were used, β2-microglobulin replaced LDH and bone marrow involvement replaced Ann Arbor stage III or IV (See Table 1.3).

Parameter	**FLIPI 1**	**FLIPI 2**
LNs	>4 LN sites	Longest diameter of biggest LN > 6 cm
Age	>60 years	>60 years
Serum level	LDH > upper normal limit	β2-microglobulin > upper normal limit
Haemoglobin concentration	<12.0 g/d	<12.0 g/d
Clinically	Ann Arbor Stage III or IV	Bone marrow involvement

Table 1.3: Comparison of FLIPI1 and FLIPI2 prognostic parameters for FL.

Another system, called the Groupe d'Etude des Lymphomes Folliculaires (GELF) Criteria, takes into account the number and size of LNs, the presence/absence of organomegaly

particularly splenomegaly, the presence/absence of B symptoms, the presence/absence of plural or peritoneal effusions, and the LDH or β2-microglobulin levels (116, 135).

Recently, a scoring system based on the presence/absence of mutation in 7 genes in FL has been introduced, and termed (m7-FLIPI). These genes are *EZH2, CREBBP, MEF2B, EP300, FOXO1, ARID1A and CARD11,* in addition to FLIPI, The Eastern Cooperative Oncology Group (ECOG) demonstrated that it offers a better prediction of OS and failure-free survival (FFS) in FL patients receiving the first chemoimmunotherapy (129).

1.2.8 Transformation of FL to high-grade aggressive disease:

Transformation of FL is the histological conversion of low-grade FL (grade 1 or 2) to an aggressive high-grade NHL (i.e. grade 3 FL) or to DLBCL (71). DLBCL have 2 major subgroups, these are activated B-cell germinal centre subtype (ABC) and germinal centre B-cell-like (GCB). The two groups are totally different in disease pathogenesis, molecular signature and survival rate (45, 136).

FL transforms to DLBCL, with 80% of the cases to the GCB subtype compared to 16% of patients who transform to the ABC subtype (137).

This high-grade transformation is diagnosed by histopathological examination, centrocytes and centroblasts are converted to irregular sheets of malignant cells with loss of follicular pattern. However, if no tissue is available, the transformation can be detected clinically by enlarging nodes or the appearance of extranodal growth, elevated LDH, B symptoms, hypercalcaemia or the detection of active sites by PET (59).

Some recent genetic studies on serial or paired indolent and transformed FL found that the transformation clone is related to a common progenitor clone. Furthermore, a divergent transformation rather than a linear transformation of FL is probably a common rather than rare model; this was discussed earlier in Section 1.2.3.4 of this chapter (88, 138-140).

Pasqualucci and colleagues found that many FL transformation-associated genomic changes were seen in genes that are targeted by AID. She found that the mutations, deletions and duplications in these genes are typical for AID-induced aberrations, and these changes were absent before transformation, suggesting AID aberration played a role in transformation of FL (139).

Loeffler and colleagues studied clonal evolution in FL, and suggested that *AID* and *CREBBP* genes are the key genes in co-evolution in FL (88). Correia and colleagues also observed that AID expression and *BCL-2* mutation are correlated with higher transformation risk and death in FL (141).

Further studies are needed to explore more the role of *AID* in FL transformation, particularly through inducing ASHM (71, 139).

1.2.9 Treatment of FL:

For the minority of patients with Ann Arbor Stages I and II disease at diagnosis, localised radiotherapy is used. For most patients with Stage 3 or above disease, the treatment in the last 3 decades has typically been anthracycline-based, with the most common protocols being either cyclophosphamide, vincristine and prednisone (CVP) or cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) (142).

The introduction of rituximab (R), which is an anti-CD20 antibody, in the last decade and adding it to chemotherapy has improved OS and response rate in FL patients in many clinical trials (143-145).

1.2.9.1 Treatment options:

There is no standard first-line nor curative treatment for FL, therapy options often rely on and are not limited to stage of disease, tumour burden, age and associated illness; also, treatment choices differ between different physicians and oncology centres (146, 147).

However, treatment can be divided into first-line therapy, maintenance and relapse treatment of patients, and in each stage of treatment, the therapy is largely dependent on stage of FL and tumour burden (116).

1.2.9.1.1 First-line therapy:

Induction treatment of FL differs according to disease stage, as follows:

1-Stage I and II: In the case of localised disease, radiotherapy using 24 Gray is the favoured approach currently (148). Sometimes a 'watch-and-wait' approach is chosen if the patient declines to undergo radiotherapy, or if there is associated comorbidity (149).

2- Stage III and IV disease:

These represent the majority of patients, curative treatment is not yet available, and only symptomatic nodal and extranodal patients are treated, those with B symptoms, affected organ function or cytopaenia (150).

No survival advantage of therapy over the 'watch and wait' approach was noted whether before rituximab treatment (135, 151) or during the rituximab era, despite the improved PFS endpoint on rituximab treatment (152). However, one study showed that for a symptomatic patient, rituximab induction with no maintenance decreased the life-time cost of treatment (153).

So, commonly used induction treatment includes rituximab plus chemotherapy including either CHOP (154).

R with CVP, is an alternative to R-CHOP; however, a study found that R-CHOP has better PFS and longer time to treatment failure but still no difference in overall survival (155).

Bendamustine with rituximab showed longer PFS and less toxic effects than R+CHOP (156) with quality of life being better for R and bendamustine compared to R-CHOP or R-CVP (157).

1.2.9.1.2 Maintenance therapy for FL:

Rituximab maintenance after chemoimmunotherapy has been shown to improve PFS with less drug toxicities and a better tolerance (158, 159). However, in elderly patients, the PFS is inferior and the risk of infection is higher compared to younger people on rituximab maintenance (160).

Radioimmunotherapy maintenance in FL was demonstrated to prolong PFS in one trial (161); however, in another trial, which compared using radioimmunotherapy after either CHOP or R-CHOP, it was shown that there was no difference in PFS (162).

1.2.9.1.3 Treatment of relapsed FL:

When relapse of disease occurs, a repeat biopsy is required to exclude transformation of FL to DLBCL. If the tumour burden is low on recurrence, a 'watch and wait' strategy is an option for asymptomatic patients, whilst rituximab monotherapy is advised in symptomatic patients.

Using another treatment regimen is recommended to avoid the development of tumour resistance, and the relapsed treatment regimen administered will depend on the response to first-line therapy (116).

Rituximab with bendamustine is considered to be more effective than rituximab with fludarabine in relapsed patients (163).

Another recently introduced anti-CD20 antibody, Obinutuzumab, was shown to be effective when administered with bendamustine. This combination is followed by obinutuzumab maintenance showed better efficacy compared to bendumastine alone in relapsed patients not responding to rituximab (164). It was also shown that this antibody is effective and tolerable if combined with bendamustine or CHOP in induction treatment of FL followed by obinutuzumab maintenance (165).

Allogeneic or autologous hematopoietic cell transplantation (HCT) in relapsed patients is controversial in FL (166).

The PI3Kδ inhibitor, Idelalisib monotherapy, was shown to have a good effect with an acceptable safety profile in patients with multiple FL relapses (167).

1.2.9.2 Updated research and new approaches for treating FL:

Drugs with different mechanisms of action against FL are under clinical trials including:

1-Monoclonal antibodies:

Several monoclonal antibodies are under trials, but their effects are yet to be confirmed by phase III trials. One of the most promising is 'next generation' anti-CD20 as obinutuzumab (168).

Other than anti-CD20 antibodies, there are trials on many other antibodies that are given with rituximab; for example, Pidilizumab (anti-PD1) and rituximab (169), Epratuzumab (anti-CD22) +rituximab (170).

2-Immunomodulators:

An example of an immunomodulatory drug is lenalidomide, which acts as a tumoricidal and immunomodulatory agent (171).

Many clinical trials have been performed to investigate lenalidomide performance against FL, either alone (172) or in a synergistic combination with rituximab, with promising results (173).

3. Phosphatidylinositol 3-kinase (PI3K) inhibitors:

The PI3K enzymes are vital for many functions, such as cellular metabolism, apoptosis, cell cycle regulation and DNA repair. By activating these enzymes, various intracellular pathways are activated, resulting in cell growth and proliferation (174).

Idelalisib, which targets delta isoform of PI3K, showed promising results in FL patients that are refractory to rituximab and alkylating agents, and it was approved by the US Food and Drug Administration (FDA) recently as a single agent for treating relapsed FL (167, 175).

4. Bruton tyrosin-kinase (BTK) inhibitors:

BTK is a vital component of B-cell receptor signalling in B-cell lymphoma (176). Ibrutinib irreversibly inhibits BTK, and thereby block signals for cell growth and survival. It is also under scrutiny in clinical trials as a monotherapy or part of combination therapies in FL (174).

5- Other drugs:

Several other drugs are under study: for example, drugs activating the immune response against cancer using anti-PD-1 antibodies or inhibiting EZH2 or BCL2, using CAR T cells etc. (177). Finally, it is important to know that not only new drugs are under investigation, but also new methods of administration, a study showed that subcutaneous rather than intravenous R administration in FL and other NHLs reduced the cost of treatment in addition to decreasing time required for patient hospitalization and health care professionals' engagement with patients (178).

1.2.9.3 The criteria for treatment responses:

To assess the prognosis of patients in FL, imaging techniques using PET–CT predict poor prognosis with positivity in PET after chemotherapy induction (179). As mentioned above, FDG-PET is also found to be an important prognostic and PFS predictor (180).

1.2.10 PACIFICO clinical trial:

The PACIFICO (Purine-Alkylator Combination In Follicular lymphoma Immuno-Chemotherapy for Older patients) trial is a phase III clinical trial that compares first-line R with cyclophosphamide, vincristine, and prednisone (R-CVP) versus R, fludarabine and cyclophosphamide (R-FC). Patients who achieve a complete or partial response to induction therapy receive maintenance therapy with rituximab for 2 years and are followed up for at least a minimum of 3 years from trial entry. The primary endpoint is PFS. Recruitment ceased on 30th April 2016 with a total of 369 patients randomised from across the UK. A sufficient number of primary events is expected to have accumulated to enable reporting of

results in October 2018. The trial is led by my Ph.D. primary supervisor (ARP). The FFPE tissue samples for most of the patients were stored in the trial tissue bank and were available for this project. Clinical and laboratory data are collected and managed by the LCTU and were available for analysis in this project.

Criteria for including a patient in the trial were:

1- Histological confirmation of FL.

2-Ann Arbor stage II-IV.

3-Aged 60 years or over, or aged less than 60 but anthracycline-based therapy is contra-indicated.

4-No systemic cancer therapy before.

5-Acceptable haematological parameters (Haemoglobin ≥ 8.0 g/dl, Absolute neutrophil count (ANC) ≥ 1.5 x 10^9/L, Platelet count ≥ 100 x 10^9/L.

6- Written informed consent.

The objective of this trial is to explore the best treatment approach for elderly patients with FL, who represent the majority of patients.

The PACIFICO trial is testing the hypothesis that R-FC may be more effective than R-CVP in older patients with FL without necessarily being more toxic, R-CVP being the current gold-standard for such patients. Only few data are currently available for tolerance and side effects of R-FC treatment in elderly patients with FL and other NHLs (181).

1.3 AID:

1.3.1 Structure and chromosomal localization:

AID is a member of the AID/APOBEC (Apolipoprotein B mRNA Editing enzyme Catalytic polypeptide-like) family, which is able to deaminate cytidine, converting it to uracil in vertebrates; it was discovered by the Muramatsu group in 1999 (182).

It is encoded by its gene *AID* which is located on chromosome 12p13. This gene consists of five exons and encodes AID protein. AID protein is 24kD, and consists of 198 amino acids that contain the active deaminase site, and its structure is closely related to the APOBEC1 mRNA-editing enzyme (182-184). AID consists of an N-terminal domain which is essential for SHM, and a C-terminal that is essential for CSR (185-187).

AID moves between the cytoplasm and nucleus by a nuclear-localization sequence (NLS) at the N-terminal end and a nuclear export signal (NES) at the C-terminal end. AID also contains a 39-amino-acid cytidine deaminase domain and a larger APOBEC-like domain (about 60 amino acids) (185, 187) (Figure 1.8).

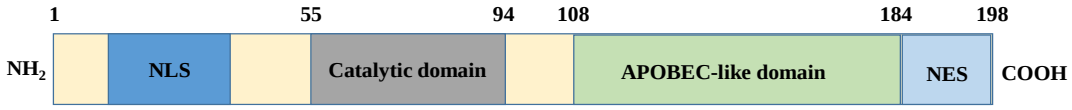

Figure 1.8: Structure of AID protein, NLS; nuclear-localization sequence, NES; nuclear export signal. Figure based upon a review and box 1 of Zhenming Xu and colleagues (188).

For efficient CSR, DNA breaks require a functioning C-terminal end of AID (183).

Due to the structural similarity of AID to APOBEC 1, it had been thought that AID was an RNA editing enzyme (182); however, it was later determined that it deaminates single-stranded DNA (189).

AID acts mainly on *IGV* genes. Recently, several elements were found to be responsible for making AID target this gene, including non-coding RNA and genomic architecture (190). AID-deficient mice fail to develop SHM or CSR and have large GC in LNs (191). In addition, mutated AID has been found in patients with hyper IgM syndrome who failed to develop CSR, and had high IgM level and immune deficiency (35, 192).

1.3.2 Function of AID in normal B-cell maturation:

AID targets single-stranded DNA and deaminates deoxycytidine, converting it to deoxyuridine (193). This creates a defect in the DNA sequence, which induces subsequent changes through four mechanisms (Figure 1.9):

1-A change of uridine to adenine and DNA replication with mutation (194).

2-Uracil DNA glycosylase (UNG) removes the deoxyuridine, which is further cleared by apurinic-apyrimidinic endonuclease (APE) nicking the ssDNA, then it will be acted on by error-free replication or polymerase mediated repair to create any mutation (193).

3-DNA mismatch repair enzymes (MMR) repair with excision of nucleotides surrounding the mutated one, followed by error-prone polymerases (EPP) to produce any mutation, and deficiency of either of those sets of enzymes affects the process of SHM (195, 196).

4-Repair by homologous recombination if the twin chromatid template is available (197).

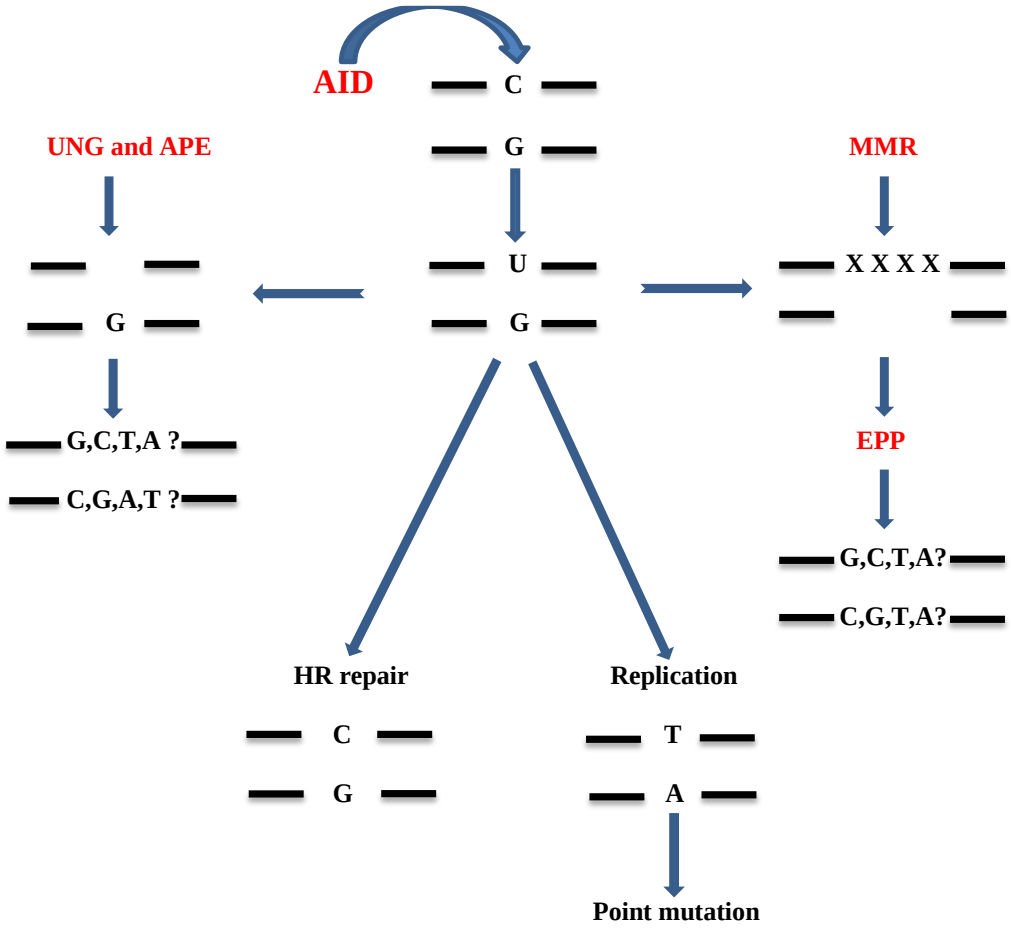

Figure 1.9: Consequences of AID deamination of cytidine to uridine. Either a replication of uridine occurs leading to point mutation, or UNG and APE remove the uridine leading to any mutation. Alternatively MMR

enzymes remove the mutated uridine and surrounding nucleotides causing any mutation. Finally a homologous recombination repair may occur. UNG; Uracil DNA glycosylase, APE: Apurinic-apyrimidinic endonuclease, MMR; Mismatch repair enzymes, EPP; Error-prone polymerase, HR; homologous recombination. Figure based on reviews by Casellas and colleagues, and Kumar and colleagues (8, 198).

1.3.3 Potential role of AID in FL:

1.3.3.1 Initiating DNA damage in B-cells and its potential roles in genomic instability:

Transition mutation of C to T is the commonest mutation found in cancers, and cytosine deamination was suggested to be the possible mechanism. Thus, studies in B-cell lymphoma cell lines found that uracil was positively correlated with AID mRNA and protein expressions. In addition, it was discovered that stimulating AID expression in B-cells increases the genomic uracil many fold. On the other hand, knocking out AID reduces uracil (199).

Additionally, AID is able to induce DNA mutations at non-IG gene loci including genes involved in B-cell transformation, as *BCL6*, *MYC*, *BCL7A* and *MSH6* (200). Furthermore, AID induces genomic instability, gene structural lesions and translocations that are important in the pathogenesis of B-cell lymphoma (201).

Single-stranded DNA in the R-loop (which is a three-strand gene structure composed of RNA-DNA hybrid and single stranded DNA) is particularly targeted by AID in *IG* and non *IG* genes, increasing mutations and is associated with DNA damage (202).

AID is hypothesised to induce DNA double-strand breaks in both *IG* and *MYC* genes, allowing translocation to occur and inducing pathogenesis of Burkett's lymphoma (203).

To our knowledge, there are no detailed studies that have examined the role of AID in FL pathogenesis.

1.3.3.2 Role of AID in ongoing *IGV* SHM:

The study of genomic clonal evolution of FL by NGS revealed a correlation between ongoing mutation in the *IGHV* gene and tumour evolution, suggesting a significant role of AID with disease progression (88).

The mechanism behind ongoing mutation is suggested to be either as a result of reaction with self-antigens (204), or by binding of oligomannose in the tumour microenvironment to the variable region of BCR (205, 206), where BCR is essential for survival of B-cells (36).

1.3.3.3 Hot motifs of AID-induced mutation:

During CSR, it was noted that AID targets WRC/GYW motifs in the switch region of the *IG* gene (R = purine nucleotide, Y = pyrimidine nucleotide, and W = A or T base). This is called canonical AID induced mutation (190, 207, 208).Another hotspot for AID is the non-canonical AID induced mutation targeting WA/TW (208-210) .

Such motifs are more common in *IGV* genes compared to the whole genome (211). In SHMs, AID is directed to act upon these motifs in productive and non-productive alleles of the *IGHV* gene (212). While in CSR, transcription targets AID to these spots (213).

1.3.4 Regulation of AID:

The body can control the expression and activity of AID by the following adopted mechanisms.

1.3.4.1 AID Transcription:

Several factors play a role in controlling AID transcription. Four regions of *AID* DNA control AID expression. In region I, the gene promotor contains binding sites for NF-κB,

HOXC4 and Sp1 that induce AID expression by activating the promoter (214, 215). Region II contains binding sites for E2f and c-Myb, and these silence AID (216).

While AID expression is induced by IL-4 and CD40 ligation (217), E proteins (218) PAX5 activate AID transcription (217, 219).

Region III is a BATF-binding gene segment downstream of the *AID* gene, and is responsible for maintaining physiological AID expression by an unclear mechanism (220).

Region IV is upstream of the *AID* gene and has enhancers binding NF-κB, STAT6, and SMAD3/4 (221).

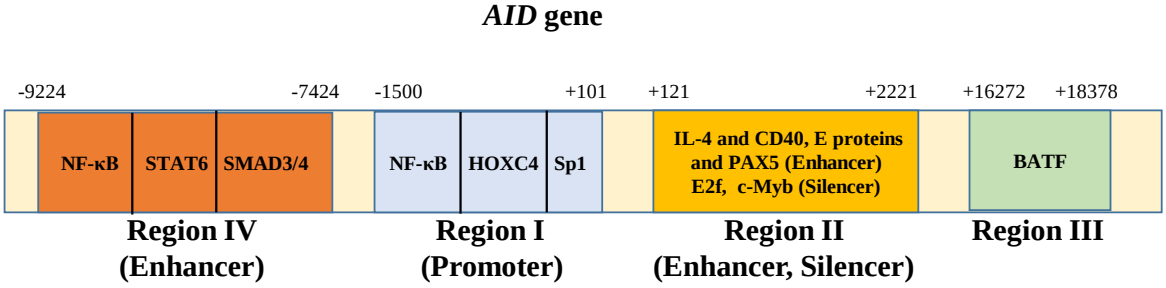

Figure 1.10: Binding regions for factors controlling *AID* transcription; distances of these regions were in base pair (bp) from transcription start of the gene. Figure based on review by Casali and colleagues, and other papers (215, 217, 219, 222).

Further, there are miRNAs that downregulate AID, such as miR-155 and miR-181b, by binding to the 3′ untranslated region of *AID* mRNA (223, 224). It is believed that miR-181b downregulates AID in earlier stages, while miR-155 acts to control excessive AID expression and accumulation (225).

The presence of AID splice variants inactivates SHM and CSR, suggesting it as a protective mechanism against mutagenic effects of AID (226).

1.3.4.2 Post-translational modification:

Another way of regulating AID activity is by post-translational modifications (PTM).

Phosphorylation of many amino acids have been identified as a PTM of AID. S38 phosphorylation in a mouse model is essential for SHM and CSR, with more than 70%

decrease in SHM and reduction of CSR to nearly background level in S38-mutated mice (227).

Other sites of phosphorylation involved in AID regulation include S3, T27, S41, S43, T140 and Y184 (214).Ubiquitination is another post-translational control mechanism of AID reducing its half-life (228).

1.3.5 Cellular localization of AID:

Although AID functions in the nucleus, it is mostly located in the cytoplasm, and it can shuttle in and out of the nucleus due to the NLS at its N terminal and NES at the C terminal (185).

Because AID induces mutations in many genes, the cellular localization of this protein is strictly controlled. This includes storing of the protein in cytoplasm rather than nucleus, in addition to active and efficient nuclear import and export mechanisms (190).

The NES binds to Chromosomal Maintenance 1 (CRM1) protein, directing AID out of the nucleus. NES is also essential for AID stabilization and CSR (229, 230).

Sub-nuclear AID location is found in conjugation with essential cellular spliceosome and splicing organelle suggesting the regulation of AID function by transcription associated with splicing (231).

NLS directs AID to the nucleus to be associated with CTNNBL1, nucleolin and nucleophosmin, and it is directed out of the nucleoli by the C-terminal end. It has been found that CSR is directly associated with the extent of nuclear import (232).

1.3.6 AID targets:

At least 236 genes have been found to be targeted by AID, using sensitive techniques (200).

AID targets the *IG* gene primarily; however, it can target and induce mutations in a large number of non *IG* genes in B-cells, including *CD79A* and *B, BCL6, PAX5,* and *Myc*, causing genomic instability (233, 234).

Researchers have found that AID targets a larger range of tumour-associated genes than originally thought. It induces genomic instability and chromosomal translocations, and only about 25% of these mutations are protected by DNA repair mechanisms (63).

Further ectopic AID expression in non-B-cells was found to be associated with tumour development (235).

1.3.7 Correlation between AID and cancers:

Aberrant AID activity is closely correlated to oncogenic translocations, which is the pathology behind many B-cell malignancies (198). Also, B-cell lymphoproliferative disorders of the head and neck were associated with increased AID expression and lymphogenesis (236).

In DLBCL, AID expression was found to be linked to poor prognosis in patients treated with CHOP chemotherapy (237).

A high level of AID is suggested to be involved in the pathogenesis of a subset of human lung cancers, and a positive correlation between *AID* mRNA level and *TP53* mutation had been found in these tumours (238).

Also, high AID had been found in malignant cells of oral squamous cell carcinoma (239).

Regarding skin cancer, inflammation-associated aberrant AID expression was found to enhance skin cancer development, in a mechanism that is unrelated to ultraviolet light damage to the skin (240).

Malaria infection has also been found to be associated with inflammation-related increase of AID level, leading to B-cell malignancies (37).

Chronic active gastritis by Helicobacter pylori infection causing chronic active gastritis was found to be associated with high AID expression that mutates *TP53* , *CDKN2B* and *CDKN2A*, and promotes the development of gastric cancers (241).

Finally, ovarian serous carcinoma was hypothesised to be linked with inflammation-induced aberrant AID expression (242).

Because of the association of AID with a large diversity of cancer types, efforts are ongoing to investigate the possibility of targeting AID expression in a mouse model and cell lines (243, 244).

1.3.8 Methods used for measuring levels of AID expression:

1.3.8.1 *AID* mRNA level:

Small RNA and miRNA extracted from FFPE specimens are the best material to be used for retrospective molecular analysis (245).

In addition, FFPE tissue is the largest source for tissues that can preserve tissue architecture and allows protein-based analysis as IHC (246). Many factors affect quantity and quality of nucleic acids extracted from FFPE samples. These include the time period between the surgical extraction and fixation, the type and duration of the fixative, storage conditions and sample thickness (247).

DNA and RNA are never intact in FFPE tissues compared to those in fresh samples. However, selected nucleic acid extraction kits may improve the quantity and quality.In addition, placing tissues in storage reagents gives a better yield than in FFPE blocks alone (248).

1.3.8.2 AID at protein level:

Different methods have been used to detect AID protein expression. Pettersen and colleagues measured AID protein in B-cell lymphoma cell lines using western blotting and mass spectrometer, and found that expression is similar using both methods, and that AID protein was significantly positively correlated to *AID* mRNA expression (199).

Cantaert and colleagues measured AID protein levels from healthy human B-cell lines by western blotting, and from human tissues by IHC and immunofluorescence (IF) (249).

Antibodies specific for AID protein have been used for IHC and IF in many studies with FFPE tissues from B-cell lymphomas (237, 250), for western blotting using cell lysates (37, 237, 251-253), or for immunoprecipitation of AID protein (254).

1.4. Hypotheses of this thesis and approaches applied:

1.4.1 Hypothesis:

The level of AID activity varies between individual patients with follicular lymphoma, high levels of AID activity being associated with more gene mutations and adverse clinical features.

1.4.2 Experimental approach:

1. Measure AID expression at mRNA (by qPCR) and protein (by IHC) levels.

2. Measure AID function as the nuclear/cytoplasmic (N/C) ratio of AID, and as total, AID-related and ongoing *IGHV* mutation.

3. Correlate AID expression and function with clinical features and the expression of genes in the FL microenvironment that are associated with prognosis.

1.4.3 Outline of thesis:

1- Chapter 1: Introduction to thesis.

2- Chapter 2: Materials and methods used.

3- Chapter 3: Optimization of RNA and DNA extraction from FFPE tissue samples of FL.

4- Chapter 4: AID expression in FL.

5- Chapter 5: Correlation between AID expression and ongoing *IGHV* mutation in FL.

6- Chapter 6: Investigation of relationships between transcriptional levels of genes associated with T cells and macrophages and AID expression in lymph node tissue, and their correlations with clinical status of FL.

7- Chapter 7: Correlations between AID expression and function against clinical outcome of FL.

8- Chapter 8: General discussion.

Chapter 2

Materials and methods

This chapter describes the patient samples, materials and procedures used throughout this study.

2.1 Tissue samples used in experiments:

2.1.1. Patients:

All patients included in this study had typical FL as diagnosed using the criteria of the WHO classification (45, 255). They were from a national clinical trial termed the "PACIFICO" (n=87), with the Principle Investigator being one of my supervisors (ARP). Informed written consents had been obtained from all patients to donate their samples for research.

2.1.2 FFPE tissue samples from local Biobank:

Three FL and 2 NHL, in addition to two hyperactive tonsil samples, were taken from the Liverpool BioInnovation Hub (LBIH) Biobank to undertake the optimization experiments for RNA and DNA extraction from FFPE specimens. Clinical data of the 5 samples are presented in Appendix Table 2.

2.1.2.1 Storage conditions and duration:

FFPE samples from LBIH had been prepared according to standard procedure guidelines (256), stored in the dark at room temperature for a period between 132 and 252 (132, 180, 192, 240 and 252) months before being analysed.

2.1.3 Tissue samples from the PACIFICO trial:

2.1.3.1 Patient consent and diagnosis criteria:

FFPE lymph node samples from 87 patients of PACIFICO trial were used in this project. The clinical data of these patients at time of sampling are included in Appendix Table 1.

2.1.3.2 Storage conditions and length:

Samples were stored at room temperature in the dark for 1-108 months (the median being 24 months).

2.2 Control cell lines and tissue culturing:

2.2.1 Raji and K562 cell lines:

Both Raji (a B-cell line, derived from a Burkitt′s lymphoma patient) (257) and K562 (a myeloid cell line, derived from a chronic myeloid leukaemia clone) (258) cells were used as positive and negative controls, respectively, in a Real Time-quantitative Polymerase Chain Reaction (qPCR) study on *AID* mRNA expression. They were kindly provided by Dr. Chloe Clapham and Dr. Claire Lucas (University of Liverpool), and were cultured under conditions described in Section 2.2.2 (below) before being harvested for RNA extraction. Both cell lines were free of mycoplasma infection, as tested before and during culture by departmental technicians using the e-Myco TM Mycoplasma PCR Detection Kit (Intro Biotechnologies, Seongnam-si Gyeonggi-do, Korea). DNA of Raji cells was also used as a positive control in analysis of ongoing mutations in *IGHV* genes (Chapter 2, Section 2.4.5).

2.2.2 Culture, passaging and quantitation of Raji and K562 cells:

The cell lines were cultured in Roswell Park Memorial Institute medium-1640 (RPMI-1640) + 10% Foetal Calf Serum, 100 U/ml penicillin, 0.1 mg/ml streptomycin and 2mM L-Glutamine) (all obtained from Sigma Aldrich, Gillingham, UK) in vented Nunc T75 flasks (Thermo Fisher Scientific, Warrington, UK) in a tissue-culture incubator (Panasonic MCO-170AICUV, Gunma, Japan) set at 37^{o}C and supplied with 5% CO_2.

For growth maintenance, these cells were split every three days at a ratio of 1:3 cells to pre-warmed media in the same culture conditions.

Cell viability and density were measured using a Cellometer TM AutoT4 cell counter (Nexce1om, Massachusetts, USA). Briefly, the diluted cells were mixed with 0.1% Trypan blue (Thermo Fisher Scientific, UK), this allowed the trypan blue to be taken into non-viable cells and visualised under microscope (259). Then, dead blue-stained cells and the shiny bright viable cells were quantified using a Cellometer counting chamber (Nexce1om, Massachusetts, USA), according to the manufacturer's instructions. $1x10^7$ cells were harvested after a few passages of culture, in order to produce sufficient amounts of RNA and DNA.

2.2.3 RNA and DNA extraction from cell lines:

DNA and RNA extraction from Raji and K562 cell lines was performed using an AllPrep DNA/RNA Mini Kit (Qiagen, Manchester, UK), following the manufacturer's instructions. $1x10^7$ cells were used for each DNA/RNA column. They were transferred into 15ml centrifuge tubes (Greiner Bio-One, Stonehouse, UK), then centrifuged at 300 x g for 5 mins. With the culture medium being removed, the cells were lysed in 600µl Buffer RLT and then centrifuged through a Qiashredder column (Qiagen, Manchester, UK) for homogenization and removal of cell debris. The homogenized lysate was subsequently transferred to an AllPrep DNA spin column and centrifuged for 30 sec at 8000 x g. The flow-through was collected for RNA preparation.

For RNA isolation, the flow-through was mixed with 350µl of 70% ethanol. Following this, 700µl of the mixture was centrifuged through an RNeasy spin column. The column was washed with 700µl Buffer RW1 and 500µl Buffer RPE sequentially by centrifuge at 8000 x g for 15 sec. The RNeasy spin column was then placed in a new collection tube and the total

cellular RNA was eluted from the column with 50µl RNase-free water in a centrifuge at 8000 x g for 1 min.

For continuing genomic DNA preparation, 500µl Buffer AW1 was added to the AllPrep DNA spin column. After centrifugation, the column was washed with 500µl Buffer AW2. The spin column was next put into a new collection tube, and the DNA in this column was finally eluted using 50µl Buffer EB twice by centrifugation after incubation at room temperature for 1 min.

2.3 RT-PCR experiments:

2.3.1 RNA extraction from FFPE tissues:

RNA was needed to synthesize cDNA to perform qPCR for quantifying expression of *AID* and endogenous control gene (EC gene). It was also required for measuring expression of other genes in cells of the FL microenvironment which might correlate with expression of *AID*.

Total cellular RNA was extracted using the RNeasy FFPE kit (Qiagen, Manchester, UK). The method was optimized for a starting amount of FFPE tissue as low as 2 sections of 4µm, instead of up to 8 sections of 10 µm in the recommended protocol. Each section was deparaffinized in 160µl of deparaffinization solution (Qiagen, Manchester, UK) in a separate microcentrifuge tube in a 56°C heating block for 3mins, 150µl Buffer PKD was added and centrifuged. This was followed by digestion of proteins in the RNA-containing lower phase with 10µl proteinase K and incubation sequentially at 56°C and at 80°C each for 15mins. After pooling the lower phase from the two sections into a new microcentrifuge tube, the RNA-containing fraction was incubated on ice for 3 mins and then centrifuged for 15 mins. In the same tube, the DNA was digested with 20µl DNase I and 10µl DNase I booster buffer at room temperature for 15 mins, before it was mixed with 320µl Buffer RBC and 700µl ethanol (100%).

700µl of the RNA-containing solution were next centrifuged through an RNeasy MinElute spin column at ≥10,000 rpm in a benchtop microcentrifuge. The flow-through was discarded and the step was repeated once for the remaining solution. In the wash step, the RNeasy MinElute spin column was washed twice, each with 500µl of Buffer RPE and centrifugation at the same speed.

After centrifugation at 14,000 rpm in the same centrifuge for 5 mins with the tube lid open, RNA in the RNeasy MinElute spin column was finally eluted into a fresh tube with 30µl of RNase-free water by centrifugation at 14,000 rpm for 1 min.

RNA isolation with deparaffinization was performed also by ReliaPrep™ FFPE Total RNA Miniprep System (Promega, Southampton, UK) during comparison to choose the best extraction kit, the steps are almost similar RNeasy FFPE kit but the deparaffinization was done using mineral oil rather than Qiagen depparafinization solution.

2.3.2 Detection of RNA quantity by Qubit® 2.0 fluorometer:

RNA and DNA quantity was measured using a Qubit® 2.0 Fluorometer instrument and a Qubit™ RNA BR Assay Kit for RNA detection in the range of 2-1000 ng/µl and Qubit™ dsDNA BR Assay Kits for DNA with a scope of detection between 100 pg/µl to 1000 ng/µl. The instrument and both kits were obtained from (Thermo Fisher Scientific, UK). The concept involves using fluorescent dyes that bind specifically to RNA and DNA molecules, producing a fluorescent signal. It is more specific and accurate than the Nanodrop spectrophotometry method, which tends to overestimate RNA and DNA (260, 261).

Working solutions were prepared by mixing Qubit® dsDNA BR Reagent with buffer in a ratio of 1:200. RNA BR Low and high RNA standards were supplied to obtain a standard curve; 10µl standards and 1µl of RNA totalled 190µl and 199µl of working solution, respectively; mixtures were vortexed and incubated for 2 mins, fluorescence measured and concentrations given in ng/µl.

2.3.3 Detection of RNA quality:

RNA quality can be detected accurately using an Agilent Bioanalyzer 2100 (Agilent technologies, Shrewsbury, UK). We reported RNA quality by using a value termed 'RNA integrity number' (RIN). RIN ranges from 1 for totally degraded RNA to 10 for intact RNA (262).

The method is based on gel electrophoresis separation of molecules according to molecular weight, and associated fluorescence is detected by laser; 18S and 28S ribosomal RNA peaks can then be visualised (262).

The RNA 6000 Nano Kit was used, which detects RNA at a range of 25–500 ng/µl, and the experiment was performed according to the manufacturer's instructions. We prepared the gel mix by adding 1µl of dye to 65µl of the gel, vortexing and centrifuging at 13,000 g for 10 mins.

Machine electrodes were cleaned using RNaseZap (Thermo Fisher Scientific) and deionised water. 9µl of the gel were dispensed into three gel wells in the chip, then 5µl of the marker was loaded into the ladder well and 12 sample wells. Finally, 1µl of denatured ladder and RNA samples were loaded into the chip, and this chip was placed into the machine for bioanalysis, and then run.

2.3.4 Reverse transcription of mRNA to cDNA:

In order to measure the transcriptional level of *AID* (as compared to that of EC gene) by qPCR, cDNA was initially synthesised from mRNA. This was achieved by performing reverse transcription in a 0.5-ml RNase-free tube (Greiner Bio-One, Gloucester, UK), containing 1µg of total cellular RNA and 1µl of random primers (0.5 µg/µl) (Promega, UK) and sterile water to make up a final volume of 13.5µl. The tube was then heated at 70°C for 5 mins to denature the secondary structure of the RNA. After cooling it down on ice for 5 mins, 0.5µl recombinant RNasin® Ribonuclease Inhibitor (20-40 u/µl), 4µl of 5X reverse

transcriptase buffer, 1µl (10mM) dNTPs and 1µl of M-MLV reverse transcriptase (all from Promega, UK) were added to initiate the reaction in a 37°C water bath for 1 hr, followed by heating at 65°C for 10 mins to inactivate the enzyme.

2.3.5 Polymerase Chain Reaction (PCR):

PCR was performed to detect the mRNA levels of *AID*, *CD19* (B lymphocyte marker) and Glyceraldehyde-3-Phosphate Dehydrogenase *(GAPDH)* genes in FL samples using the resulting cDNA prepared as described in Section 2.3.4 in this chapter. All PCRs were performed in 50µl reaction in a 0.5-ml tube in an Eppendorf Mastercycler (Hamburg, Germany). A GoTaq® Flexi DNA Polymerase kit (Promega, UK) was used. The components in the PCR are listed in Table 2.1. The forward and reverse primers for each target gene are presented in Table 2.2, and the PCR amplification protocol is shown in Table 2.3.

Reagent	Amount per sample in µl
5X PCR Buffer	10
$MgCl_2$ (25mM)	3
Nuclease-free water	33.75
Forward primers (20µm)	1
Reverse primers (20µm)	1
DNA polymerase (5u/µl)	0.25
cDNA	1

Table 2.1: Reagents for PCR amplification of *AID, GAPDH and CD19.*

Primer Name	Primer Sequence (5'- 3')	Product size (bp)	Source of mRNA sequence templates
AID For (F)	5'-GAGGCAAGAAGACACTCTGGAC-3'	406	NM_020661.2
AID Rev (R)	5'-TGCGGTCCT CACAGAAGTAGAG-3'		NM_020661.2
GAPDH For	5'-ATCACCATCTTCCAGGAGCGAG-3'	315	NM_002046.3
GAPDH Rev	5'-GACTGTGGTCATGAGTCCTTCC-3'		NM_002046.3
CD19 For	5'-TTC AAA GAG CCC TGG TCC TG-3'	441	NM_001178098.1
CD19 Rev	5'-ACT CAG CGT TGG AGA AGG AG-3'		NM_001178098.1

Table 2.2: Primers used for PCR amplification of *AID, GAPDH and CD19.*

The PCR conditions had been optimised for amplification of all of the three target genes in the laboratory before this study, and are recorded in Table 2.3.

Steps	Temperature	Time	Number of cycles
Initial denaturation	94ºC	3mins	1
Denaturation	94ºC	30s	29
Annealing	58 ºC	30s	
Extension	72^{0}C	30s	
Final extension	72^{0}C	5mins	1

Table 2.3: PCR protocol for PCR amplification of *AID, GAPDH and CD19.*

2.3.6 Agarose gel electrophoresis:

This is a technique that enables separation of DNA fragments. The principle is that charged DNA fragments with smaller size move more quickly in the agarose gel's electrical field. Thus, DNA fragments with different sizes can be visualised and purified after staining (263). In this study we used 1-2% agarose gels for DNAs with different sizes. We prepared the gel by melting the desired amount of agarose (Sigma-Aldrich, Dorset, UK) in 1X TBE buffer, which contained 89 mM Tris-borate and 2 mM EDTA, pH 8.3 and diluted from the 10X stock prepared by dissolving Tris-Borate-EDTA buffer power (Sigma-Aldrich, UK) in distilled water.

5μL of Midori Green Advance DNA stain (Geneflow, Lichfield, UK) were added to the melted agarose, which was then left to cool to room temperature before being cast.

After solidifying, the gel was covered in the 1X TBE buffer in an electrophoresis tank. DNA samples with 1:5 (v/v) 6X gel-loading dye (New England BioLabs, Hitchin, UK) were mixed, and then loaded along with DNA size markers into slots of the gel. After 45 to 90 mins of electrophoresis in an electrical field of 100-150 constant voltage, separated DNA

bands in the gel were visualised and recorded using the INGenius Imaging System (Syngene, Cambridge, UK).

2.3.7 qPCR measurement of *AID* transcriptional level:

Real-time PCR is a method developed to overcome the inaccuracy of conventional PCR in quantifying relative or total copy number of gene transcripts in cells. It records product signals after each cycle of PCR, and therefore enables analysis before plateau of the amplification. Compared with a pre-quantified standard curve or simultaneously amplified EC gene(s), either absolute copy number or relative level of gene transcripts can be measured (264).

A LightCycler ® 480 (Roche, Hertfordshire, UK) and LightCycler ® 480 Multiwell Plate 96, white (Roche, UK) were used to perform qPCR for quantifying *AID* expression.

A Qiagen QuantiTect SYBR Green PCR master mix (Qiagen, UK) was used in the reactions. *GAPDH* was selected as an EC gene after an experimental comparison with two other EC (*TBP and HPRT1*). PCR primers are listed in Table 2.4.

Primer Name	Primer Sequence (5'- 3')	Product size (bp)	Source of mRNA sequence templates
AID For (F)	5'-GAGGCAAGAAGACACTCTGGAC-3'	90	NM_020661.2
AID Rev (R)	5'-GCCCAGCGGACATTTTTGAATTG-3'		NM_020661.2
GAPDH For	5'-ATCACCATCTTCCAGGAGCGAG-3'	101	NM_002046.3
GAPDH Rev	5'-TTCTCCATGGTGGTGAAGACGC-3'		NM_002046.3
TBP For	5'-GCCCGAAACGCCGAATAT-3'	73	NM_001172085.1
TBP Rev	5'-CCGTGGTTCGTGGCTCTCT-3'		NM_001172085.1
HPRT1 For	5'-TGACCTTGATTTATTTTGCATACC-3'	102	NM_000194.2
HPRT1 Rev	5'-CGAGCAAGACGTTCAGTCCT-3'		NM_000194.2

Table 2.4: Primers used for qPCR amplification of *AID, GAPDH, TBP and HPRT1*.

The designed primers amplified *AID* product spanning exons one and two and exons four and five of *GAPDH*.

qPCR components and protocols are listed in Tables 2.5 and 2.6. Briefly, both forward and reverse primers for *AID* or *GAPDH* were added to 12.5µl of PCR master mix, and 2µl of ¼

diluted (v/v, in sterile water) cDNA of patients (Table 2.5.). Tests for each sample were performed three times, under conditions presented in Table 2.6.

The ratio of *AID*/*GAPDH* gene transcription levels was calculated using the following equation: (265)

$\Delta Ct = Ct(AID) - Ct(GAPDH)$

$\Delta Ct \text{ expression} = 2^{-\Delta Ct}$

Reagent	Amount per sample in µl
Nuclease-free water	9.5
SYBR Green PCR master mix	12.5
Forward Primer (20µm)	0.5
Reverse Primer (20µm)	0.5
Diluted cDNA	2
Total	25

Table 2.5: Components of qPCR reaction.

Steps	Temperature	Time	Number of cycles
Initial denaturation	95ºC	15mins	1
Denaturation	95ºC	15s	45
Annealing	59 ºC	20s	
Extension	72^{0}C	20s	
Final extension	72^{0}C	5mins	1

Table 2.6: Cycle details of qPCR reaction.

2.4 DNA-based experiments:

2.4.1 DNA extraction from FFPE tissues:

DNA extraction from FFPE tissues was performed using a QIAamp DNA FFPE Tissue Kit (Qiagen, UK). The protocol was optimized to produce more DNA from FFPE tissue sections; in brief, two sections of 4µm paraffinized tissue were dissolved in 1ml xylene at room

temperature by vortexing for 10 sec, and centrifuged at full speed for 2 mins, followed by removing the liquid phase and mixing the spin-down by a brief vortex with 1ml of ethanol. The tube was centrifuged again for 2 mins at full speed and the ethanol was removed. The tube was left opened at room temperature for 10 mins to allow evaporation of remaining ethanol.

Subsequently, 180µl of Buffer ATL and 20µl proteinase K were added to the tube, which was then vortexed and kept at 56ºC in the heating block of an Eppendorf ® Thermomixer (Sigma-Aldrich, UK) with shaking at 500 rpm overnight.

The next day, 5µl proteinase K was added to the tube. After incubation under the same conditions for 1 hour, the tube was incubated at 90º C for 1 hour to reverse DNA formaldehyde modifications. Once cooled down, RNAs were digested with 2µl of RNAse-A at room temperature for 2 mins.

After mixing with 200µl Buffer AL and 200µl ethanol by vortex, the lysate was transferred onto a QIAamp MinElute column followed by centrifugation at 6000 x g for 1 min. After the flow through was discarded, the column was washed with 500µl of Buffer AW1 and then 500µl of Buffer AW2 sequentially by centrifuging for 1 min at 6000 x g. The column was dried by centrifugation at 20000 x g for 3 mins.

Finally, 50µl of ATE buffer was loaded on the column, which was then left at room temperature for 5 mins, before being centrifuged at the same speed for 1 min to elute the DNA.

2.4.2 Measuring concentration and purity of DNA extracted from PACIFICO FFPE tissues:

The DNA quantity was measured using the Qubit™ dsDNA BR Assay Kits (life technologies, UK), details of which are in Section 2.3.2.

Purity of DNA was detected using a Nanodrop 2000 spectrophotometer (Thermo scientific, UK). The purity was acceptable if the ratio of spectrometric absorbance of 260 vs 280 was between 1.8 and 2.0.

2.4.3 Detecting length of DNA extracted from FFPE specimens:

In addition to purity, length is another important parameter of DNA quality. For measuring the quality and level of degradation of DNA, we used PCR to amplify *GAPDH* (DNA template code NC_000012.12) fragments with different sizes to measure the size of DNA extracted from those FFPE tissue samples.

This was achieved by applying a consensus forward primer and 6 reverse primers designed to amplify 101, 206, 299, 396, 505 and 598-bp fragments of this EC gene. The details of these primers are listed in Table 2.7. The PCR cycle conditions are presented in Table 2.8.

Primer name	Sequences (5'- 3')
GAPDH For	5'-AGGTGGAGCGAGGCTAGC-3'
GAPDH Rev 101bp	5'-CAA GAG GAC CTC CAT AAA CC-3'
GAPDH Rev 206 bp	5'-AGG GGC CAT TTT GCG GTG-3'
GAPDH Rev 299 bp	5'-AGC CTG GCC TTT GGG GTC-3'
GAPDH Rev 396 bp	5'-AGG GTG CGG TGG GAG ATC-3'
GAPDH Rev 505 bp	5'-GGAGTGGAGCACAGGTAAG-3'
GAPDH Rev 598 bp	5'-GGTGGGAAGAGGGGAAGC-3'

Table 2.7: Primers used to amplify different sizes of *GAPDH* PCR bands.

Steps	Temperature	Time	Number of cycles
Initial denaturation	94ºC	1min	1
Denaturation	95ºC	15sec	35
Annealing	58 ºC	30 sec	
Extension	72^0C	30 sec	
Final extension	72^0C	1min	1

Table 2.8: PCR protocol of *GAPDH* PCR bands.

After a pilot experiment to explore the range of DNA sizes in a number of FFPE samples, it was decided to amplify 206, 299, and 396 bp products, as these represented the DNA size of most of the samples tested. The PCR products were then semi-quantified based on their relative band intensity as visualized on the agarose gel (Chapter 3, Section 3.2.2.2).

2.4.4 Full length *IGHV* amplification in the FFPE samples:

Because SHM can occur anywhere across the *IGHV* gene, we initially attempted to amplify the full length of the gene by PCR, using primers annealing *IGHV* leader and heavy chain constant regions, with the expected product length being about 500 bp (Chapter 1, Figure 1.2).

Primers were designed in our laboratory for amplifying the *IGHV* gene from intact DNA, and conditions were optimized and published previously (266).

Primers and all components in PCR reactions are shown in Tables 2.9 and 2.10, respectively.

The PCR program was as follows:

1-Initial denaturation was one cycle at 94ºC for 3 mins.

2- (Denaturation at 94ºC for 45 sec, followed by annealing at 63ºC for 45 sec, and then extending at 72ºC for 45 sec) repeated twice.

3- (Denaturation at 94ºC for 45 sec, followed by annealing at 62ºC for 45 sec, then 72ºC for 45 sec) repeated twice.

4- (Denaturation at 94ºC for 45 sec, followed by annealing at 61ºC for 45 sec, then 72ºC for 45 sec) repeated twice.

5- (Denaturation at 94ºC for 45 sec, followed by annealing at 60ºC for 45 sec, then 72ºC for 45 sec) repeated 24 times.

6- Final extension at 72ºC for 5 mins for 1 cycle.

Primer name	**Sequences (5'- 3')**
Forward primers	
VH1 Leader	5'-CCA TGG ACT GGA CCT GGA -3'
VH2 Leader	5'- ATG GAC ATA CTT TGT TCC AC -3'
VH3 Leader	5'-CCA TGG AGT TTG GC TGA GC -3'
VH4 Leader	5'-ATG AAA CAC CTG TGG TTC TT -3'
VH5 Leader	5'- ATG GGG TCA ACC GCC ATC CT -3'
VH6 Leader	5'-ATG TCT GTC TCC TTC CTC AT -3'
Reverse primers	
IgA	5'- GAG GCT CAG CGG GAA GAC CTT -3'
IgG	5'-GGG GAA GTA GTC CTT GAC CAG -3'
IgM	5'- CAG GAG AAA GTG ATG GAG TCG -3'

Table 2.9: Primers to amplify full length *IGHV* gene from leader to constant region.

Reagent	**Amount per reaction in µl**
Nuclease free water	33.25
dNTP(10mM)	0.5
5X Taq Buffer	10
$MgCl_2$ (25mM)	3
GoTaq® DNA polymerase (Promega)	0.25
Forward primers (20µm)	1
Reverse primers (20µm)	1
cDNA	1

Table 2.10: Contents of PCR master mix to amplify full length *IGHV* gene from leader to constant region.

However, as expected from degraded DNA that is extracted from FFPE, we could not amplify the full length of the gene. So, we tried the BIOMED-2 protocol to amplify a shorter length of the *IGHV* gene.

2.4.5 BIOMED-2 protocol for study of *IGHV* clonality by multiplex PCR:

BIOMED-2 is a PCR-based technique to test clonality of the immunoglobulin gene receptor that aids in diagnosis of B-cell malignancies, including FL (11). This technique was used to amplify the *IGHV* gene for cloning study on ongoing mutations of the same sample by comparing this gene nucleotides between different clones.

2.4.5.1 PCR amplification of *IGHV* genes in families:

In total, there are more than 52 functional VH segments; these have been divided into 6 families that have similar structure, and can be amplified by PCR using the same primers (11).

The protocol allowed us to use primers to start the amplification from FR1, FR2 or FR3 to the junction (Chapter 1, Figure 1.2). Owing to limited size of the DNA in FFPE specimens, we could not amplify the *IGHV* genes with a combination of primers targeting leader and constant regions. But with modification of the PCR conditions, we successfully amplified the rearranged *IGHV* genes between their FR1 and junction regions, which were long enough to provide information about the clonality and ongoing mutations.

So, the first round of PCR was used to apply primers of all families to detect B-cell clonality of the lymphoma. In the second round, the family with productive results was re-amplified in order to purify templates for Sanger sequencing.

A Thermo-Start DNA polymerase kit (Life technologies, UK) was used for PCR, which was performed in a TC512 Techne PCR thermal cycler (Bibby Scientific™, Staffordshire, UK).

As mentioned above, we modified the BIOMED-2 protocol in order to amplify the large region of this gene. Thus, 5µl $Mgcl_2$ (2.5 mM) instead of 3µl (1.5 mM), as suggested by BIOMED-2, was used in the 50µl reactions; and the number of PCR cycles was increased from 35 to 38. Tables 2.11, 2.12 and 2.13 show details of primers used, components in the reaction and PCR cycle conditions.

Primer name	Sequences (5'- 3')
VH1-FR1	5'-GGCCTCAGTGAAGGTCTCCTGCAAG-3'
VH2-FR1	5'-GTCTGGTCCTACGCTGGTGAAACCC-3'
VH3-FR1	5'-CTGGGGGGTCCCTGAGACTCTCCTG-3'
VH4-FR1	5'-CTTCGGAGACCCTGTCCCTCACCTG-3'
VH5-FR1	5'-CGGGGAGTCTCTGAAGATCTCCTGT-3'
VH6-FR1	5'-TCGCAGACCCTCTCACTCACCTGTG-3'
VH1-FR2	5'-CTGGGTGCGACAGGCCCCTGGACAA-3'
VH2-FR2	5'-TGGATCCGTCAGCCCCCAGGGAAGG-3'
VH3-FR2	5'-GGTCCGCCAGGCTCCAGGGAA-3'
VH4-FR2	5'-TGGATCCGCCAGCCCCCAGGGAAGG-3'
VH5-FR2	5'-GGGYGCGCCAGATGCCCGGGAAAGG-3'
VH6-FR2	5'-TGGATCAGGCAGTCCCCATCGAGAG-3'
VH7-FR2	5'-TTGGGTGCGACAGGCCCCTGGACAA-3'
VH1-FR3	5'-TGGAGCTGAGCAGCCTGAGATCTGA-3'
VH2-FR3	5'-CAATGACCAACATGGACCCTGTGGA-3'
VH3-FR3	5'-TCTGCAAATGAACAGCCTGAGAGCC-3'
VH4-FR3	5'-GAGCTCTGTGACCGCCGCGGACACG-3'
VH5-FR3	5'-CAGCACCGCCTACCTGCAGTGGAGC-3'
VH6-FR3	5'-GTTCTCCCTGCAGCTGAACTCTGTG-3'
VH7-FR3	5'-CAGCACGGCATATCTGCAGATCAG-3'
JH consensus	5'-CCAGTGGCAGAGGAGTCCATTC-3'

Table 2.11: BIOMED-2 primers (11).

Reagents	Amount
PCR buffer (10X)	5µl
$Mgcl_2$ (25mM)	5µl
dNTP (10mM)	2µl
Thermo-Start DNA Polymerase (5U/µl)	0.2µl
DNA	50ng
Nuclease-free water	Add to complete volume to 50µl

Table 2.12: BIOMED-2 PCR components.

Steps	Temperature	Time	Number of cycles
Initial denaturation	95ºC	7mins	1
Denaturation	95ºC	30 sec	38
Annealing	60 ºC	30 sec	
Extension	70^{0}C	50 sec	
Final extension	72^{0}C	10mins	1

Table 2.13: BIOMED-2 PCR protocol.

2.4.5.2 Agarose gel electrophoresis and PCR purification of *IGHV* gene:

In order to study ongoing mutations of *IGHV* in FL clones by Sanger sequencing, we needed to subclone the PCR product of the clonal *IGHV* into a plasmid vector. Thus, an agarose electrophoresis gel was run, in order to identify and purify PCR products from such clones. To do so, 400µl of *IGHV* PCR product (details in Section 2.4.5.1) were loaded into slots of a 2% agarose gel (details of agarose gel preparation and running is described in Section 2.3.6). DNA bands were visualised and cut out under UV light using a blade. Next, DNA from the agarose gel was purified using a Wizard® SV Gel and PCR Clean-Up System (Promega, UK) according to the manufacturer's instructions.

In brief, the gel slice was covered with membrane binding solution provided (each 10mg of gel in 10µl of solution) and then incubated for 10 mins at 65°C until all gel was thoroughly dissolved. The dissolved gel mixture was then transferred onto the SV minicolumn connecting to a collection tube and incubated at room temperature for 1 min. Then, the SV minicolumn was centrifuged at 16,000 × g for 1 min and the 'flow through' was discarded; the column was washed twice, first with 700µl and then 500µl using membrane wash solution and centrifugation at 16,000 × g for 1 min in the first wash and 5 mins in the second one.

Another centrifugation of the minicolumn was done for 1 min with the lid off to allow ethanol to evaporate, then we added 50µl of nuclease-free water to the Minicolumn and left

it at room temperature for 1 min. Finally, the purified DNA was eluted into a fresh tube by centrifugation at 16,000 × g for 1 min. The concentration of DNA was measured by using a Qubit™ dsDNA BR Assay Kit (Life Technologies, UK) and method as in Section 2.3.2. of this chapter.

2.4.5.3 Cloning of *IGHV* DNA:

To study the ongoing mutations of *IGHV* in FL, we subcloned *IGHV* PCR products into a pGEM-T-easy Vector system (Promega, UK), according to the manufacturer's recommendations. Following this, the recombinant vector was transformed into JM109 competent cells. According to the manufacturer, we made 10µl of ligation in a micro tube containing the following components:

2x ligation buffer	5.0µL
T-easy vector (50ng/µl)	1µL
T4 ligase	1.0µL

Then, 15ng PCR products were added at a 3:1 ratio to the amount of vector as determined in the pilot experiment. The required 15ng of DNA in PCR products was calculated as follows:

$$\frac{\text{ng of vector x kb size of insert}}{\text{kb size of vector}} \text{ x insert: vector molar ratio = ng of insert}$$

(50ng vector × 0.3kb insert)/3 × 3/1 = 15ng of DNA. This is followed by adding nuclease free water to a final volume of 10µL. Following the ligation reaction in 4°C overnight, the transformation reaction was performed in a 1.5ml microcentrifuge tube, which contained 2µl of the ligated product and 50µl of JM109 competent cells (Promega, UK). Contents were flicked to mix, put on ice for 20 mins and then into a 42°C water bath for 45-50 sec;

the tube was then transferred onto ice for 2 mins. Then, 950µl of Super Optimal broth with Catabolite repression (SOC) medium (Thermo scientific, UK) were added followed by incubation at 37°C with shaking at 150 rpm for 1.5 hours.

200µl of the transformation culture was plated onto duplicate Luria-Bertani agar (LB)/ampicillin/ isopropylethio-ß-D galactoside (IPTG)/X-Gal plates and then incubated at 37°C for 16-24 hours. LB/ampicillin/IPTG (Sigma Aldrich, UK) and X-gal (Promega, UK) plates had been prepared as follows: after melting LB agar tablets in distilled water, we added Ampicillin to the LB agar when it was cooled to 55°C or less, to be at a concentration of 100µg/ml. When the LB agar had set in the petri dishes (Thermo Fisher Scientific), 5µl of stock IPTG (200mg/ml) and 40µl of X-gal (20mg/ml) were spread on the agar an hour before the transformed bacteria were inoculated onto the plates.

2.4.5.4 Clonal selection:

Following incubation, white clones that contained the DNA insert were distinguished from blue clones that had no DNA insert. At least 15 white clones were selected for each sample. Each clone was cultured overnight in a falcon tube that contained 1ml of LB agar at 37°C. Then, the clones were sent for DNA to be prepared from these clones and sequenced with Sanger sequencing as commercial services provided by Genewiz company, UK.

2.4.5.5 Analysis of *IGHV* gene sequences:

Using IMGT/V-QUEST software (the international ImMunoGeneTics information system), we were able to align our *IGHV* gene sequences against their germline counterparts, and identify somatic mutations in them (267). After applying the sequence results to IMGT/V-QUEST software, the obtained sequence was aligned to the closest Ig germline. Further, our junction segment was translated into amino acids: first, sequences of junction segments of all *IGHV* subclones were compared, and only those with identical junction segment sequences as originated from the same clone were considered.

Then, an excel file was used to record mutations in each codon manually (for example, GCA to GAA), and then these mutations were translated into amino acids (Alanine to Glutamine) to determine whether the mutation changed the amino acid or not. Next, it was identified whether each mutation was related to AID activities, by analysis of AID hotspots, as suggested in previous publications (198, 208).

2.5 Protein-based experiments:

2.5.1 Immunohistochemical detection of AID protein:

2.5.1.1 IHC staining of AID:

To detect the protein level of AID in FL, 4µm FFPE tissue sections were cut and placed on X-tra® slides (Leica Biosystems); these slides were dried at 40°C overnight and subsequently stained using mouse anti-human monoclonal antibody against AID protein clone: ZA001, (Thermo scientific, UK).

Heat-induced deparaffinization, rehydration and antigen retrieval was performed using a DAKO PT-Link machine. IHC was undertaken using a Dako autostainer and Dako Flex System kit for FFPE tissues, both from Dako UK Ltd (Cambridgeshire, UK).

With each batch of slides stained, a positive control (hyperplastic tonsil) and a negative control (hyperplastic tonsil with added mouse IgG1 isotype-specific negative control) (Dako UK Ltd, UK) were included. The steps of AID staining by IHC are shown in Figure 2.1.

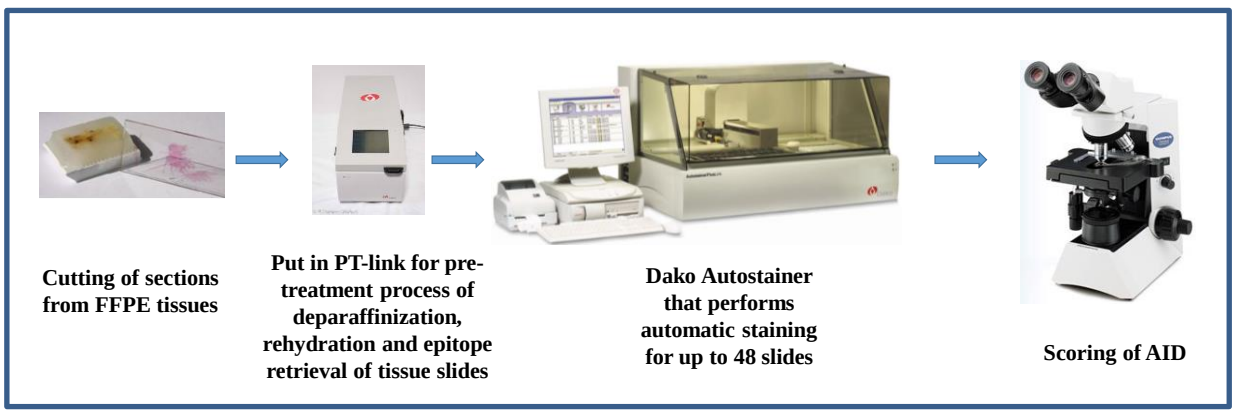

Figure 2.1: Workflow experiment for AID staining by IHC.

2.5.1.1.1 PT link epitope retrieval:

Antigen retrieval was undertaken using the PT link solution by adding 30ml of Envision Flex target retrieval solution high pH (50x) (Dako UK Ltd, UK) to 1.5 L of distilled water. The PT-link was heated to 65°C, and the slides were placed in PT link; the temperature was raised to 96ºC and then reduced to 65ºC, when the slides were removed, washed using EnVision™ FLEX Wash (Dako UK Ltd, UK) 3 times, and then moved to the DAKO autostainer (Dako UK Ltd, UK).

2.5.1.1.2 Staining slides with AID antibody using Dako autostainer:

The DAKO Autostainer is capable of undertaking automatic IHC staining of up to 48 tissue slides in each run.

It requires FLEX Wash Buffer, which is composed of Tris buffered saline (TBS) with Tween pH 7.6 (50mM Tris, 150mM sodium chloride, 0.05% Tween). Hence, 60.6 gm of Tris (Hydroxymethyle-methylamine) (VWR 271195Y) and 80.9 gm of sodium chloride (VWR 27800.291), 34ml of hydrochloric acid sp.gr. 1.18 (VWR 28507AE), and Tween 20 (polyoxyethelyne (20) sorbitan monolaurate) (VWR 663684B) 5ml was added to 10L distilled water and the pH adjusted to 7.6 using sodium hydroxide and hydrochloric acid.

The slide number was entered and the appropriate reagents were used as per software instructions. Envision-Flex block, which blocks nonspecific antibody binding sites was added for 5 mins, then AID antibodies (1/500 dilution in Phosphate buffered saline (PBS)) applied for 30 mins, followed by the mouse linker that amplifies the signal of primary antibody, for 15 mins, Horseradish Peroxidase (HRP) enzyme was added for 15 mins: this converts diaminobenzidine (DAB) in the presence of hydrogen peroxide (H_20_2) into a brown pigment; DAB was added for 20 mins.

Slides were washed by Envision-Flex buffer and rinsed in distilled water; subsequently slides were transferred to filtered haematoxylin solution for 30 sec to stain the nucleus, and

placed into a rack under a running tap water bath until the water ran clear.

The slides were dipped in acid alcohol to remove excess stain, and then put in Scott's tap water/ammonia water for 30 sec to change the pH to alkaline, and change the nucleus stain to blue. Then slides were put into a rack and placed into a running tap water bath for 1 min. Subsequently, the slides were dehydrated in a series of 5 times IMS (Industrial Methylated Spirits) and 2 times xylene (both from Sigma Aldrich, UK), before one drop of mountant was added to a coverslip, and covered the stained section.

2.5.1.2 Scoring system for AID quantification:

The protein expression of AID was scored separately by two individuals (myself and Professor Sarah E. Coupland (a haematopathologist)); if there was any discordance of scores, at a combined meeting the slides were re-reviewed and consensus obtained.

The scoring was undertaken according to the number and intensity of stained cells, which is the general scoring method used in IHC as follows: percentage of lymphocytes stained: 0% Negative, 1-24% scored 1, 25-49% scored 2, 50-74% scored 3, and more than 75% scored 4 Intensity: Negative scored zero, weak scored one, moderate scored 2 and strong scored 3.

The final score was decided by multiplying percentage score X intensity score, with a minimum score being 0 and a maximum being 12 (268-270).

2.5.2 Immunofluorescent double staining of AID with CD3, CD20 or CD68:

The protocol for double staining of each of the antibody combinations was optimised.

Two protocols were followed: either the 2 antibodies were mixed together followed by AID secondary antibody, then a mixture of fluorescent detection dyes of 2 colours; or a complete protocol for staining of one antibody was followed by the staining of the other antibody.

The concentration of AID in all experiments was 1/50 dilution in PBS.

For the fluorescent dyes, Alexa Fluor 555 (Red) (Thermo Fisher Scientific, UK) was applied in a concentration of 1/200 to visualize CD3, CD20 and CD68. For AID, the secondary biotinylated Goat Anti-Mouse IgG Antibody was used, and the green colour DyLight 488 Streptavidin, both of Vector Laboratories Ltd (Peterborough, UK), were added to accelerate the signal diluted in PBS in a 1/200 concentration.

For the double staining of AID with CD3 (Rabbit polyclonal anti-human CD3, Agilent Technologies, UK), the antibodies were applied 'one after the other' whilst for the combination of CD20 and CD68, the antibodies were mixed together at a concentration of 1/50 for AID with 1/100 PBS dilution of Rabbit Monoclonal antibody to human CD20 (AMS Biotechnology (Europe) Ltd, Abingdon, U.K) and 1/50 PBS diluted Rabbit polyclonal IgG to human CD68 Antibody (Santa Cruz Biotechnology, Heidelberg, Germany).

Prior to the immunofluorescence staining, the slides were put into a DAKO PT Link system for antigen retrieval (see details in Section 2.6.1.1.1)**,** and the blocking reagent was added to prevent nonspecific binding of AID antibody to other proteins; the blocking reagent used was 10% Normal Goat Serum in PBS (Vector Laboratories Ltd, UK) with 1% bovine serum albumin (BSA) for 1 hour at room temperature. The slides were then incubated with ready to use CD3 antibody for 1 hour (in the case of AID and CD3 double staining), then goat anti-rabbit Alexa fluor 555 (Red). Monoclonal mouse AID antibody was then added followed by Biotinylated Goat Anti-Mouse IgG Antibody for 1 hour, then DyLight 488 Streptavidin (green) in 1/200 PBS dilution for 30 min.

In the case of double staining of AID with CD20 and CD68, AID was mixed with CD20 or CD68, then incubated for 1 hour at room temperature; then washed three times each for 3 mins in FLEX wash buffer (preparation details in Section 2.6.1.1.2 of this chapter).

Then, biotinylated Goat Anti-Mouse IgG Antibody was placed on the slides for 1 hour, and then goat anti-rabbit Alexa fluor 555 (Red) and DyLight 488 (Green) were added, both at

1/200 PBS dilution in the dark at room temperature for 30 mins. Subsequent to this, they were washed and mounted in Cygel mountant (Biostatus, Leicestershire, UK). Positive- (hyperplastic tonsil) and negative controls (hyperplastic tonsil with added monoclonal mouse IgG1 antibody 1/50 dilution) were included in all experiments.

The slides were visualised using an Olympus BX61 microscope (Olympus, Tokyo, Japan).

Steps involved for AID immunofluorescent co-expression with CD3, CD20 or CD68 are shown in Figure 2.2.

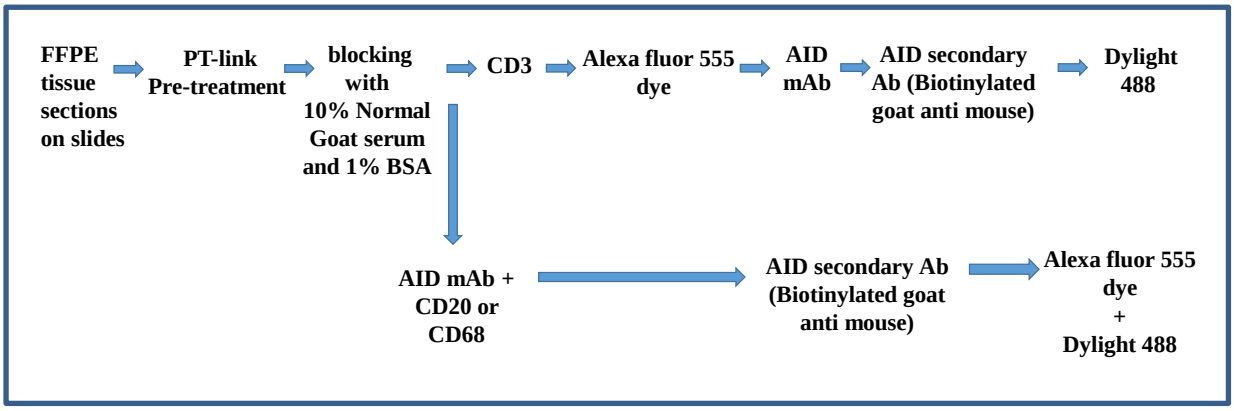

Figure 2.2: Workflow for Immunofluorescent double staining of AID with CD3, CD20 or CD68 experiments.

2.5.3 Subcellular localization of AID by immunofluorescence:

2.5.3.1 Protocol for immunofluorescent staining and confocal microscopy visualization:

To identify the subcellular localization of AID protein by immunofluorescent expression, 4µm sections from FFPE tissues were cut and placed on super frost plus slides (Thermo Fisher Scientific). The slides were stained with nuclear stain and AID stain to identify nuclear co-expression of AID; the same protocol as in Section 2.6.2 (above) was applied.

Initially, antigen retrieval using the DAKO PT Link system was performed and then a blocking reagent was added, followed by the AID Monoclonal Antibody at 1/50 dilution. Subsequently, a secondary biotinylated goat anti-mouse IgG antibody was added with a 1/100 dilution (Vector Laboratories Ltd, UK) to amplify the signal. This was followed by

1/200 PBS diluted DyLight 488 Streptavidin (Vector Laboratories Ltd, UK). Finally, Hoechst 33258 Fluorescent Stain (Thermo Fisher Scientific, UK) in 1/200 dilution for 15 mins was added. The slides were coverslipped using 10µl of Cygel mountant (Biostatus, Leicestershire, UK).

The slides were visualised using an LSM 710 confocal microscope (Carl Zeiss Vision UK Ltd, Cambridge, UK), using a 10x lens first to identify the GCs that are closest to 12, 3, 6, and 9 o'clock directions and the centre itself. A higher power (25x objective) was used to identify upper and lower parts of the GC, and then the 40x objective to ultimately take the images.

2.5.3.2 Analysis of immunofluorescent images:

To analyse the subcellular localization of AID, we modified a previously-published method (271, 272). Image j1x software was used to measure the nuclear intensity:

The colours were changed into the grey scale, because of the better appearance and contrast in Image j1x by (Image> Color> channel tools> grey scale), and then the image was split into 2 images (Image> color> split channels): one having the nuclear stain and the other with the AID stain. The command (Image > Adjust > Threshold) was used to focus on the nuclear stain image to identify nuclei, then (Edit > Selection > Create Selection) was used to select nuclei only. Following this, the AID positive image was focused on and selected [Edit > Selection > Restore Selection] and then 'm' was pressed to measure the region. This is for the nuclear measurement.

Using [Edit > Selection > Select None] the area was 'de-selected', and the command 'm' given to measure the whole image. A total measurement was obtained. Finally, 5 different areas that were free of fluorescence were used to calibrate for (and remove) background staining.

The corrected total cell fluorescence (CTCF) was then calculated according to the following equation:

CTCF = Integrated Density – (Area of selected cell X Mean fluorescence of background readings) (271, 272).

The ratio of nuclear to total AID was calculated using the following equation:

= (Nuclear CTCF / (nuclear + cytoplasmic CTCF)) X 100 (Figure 2.3)

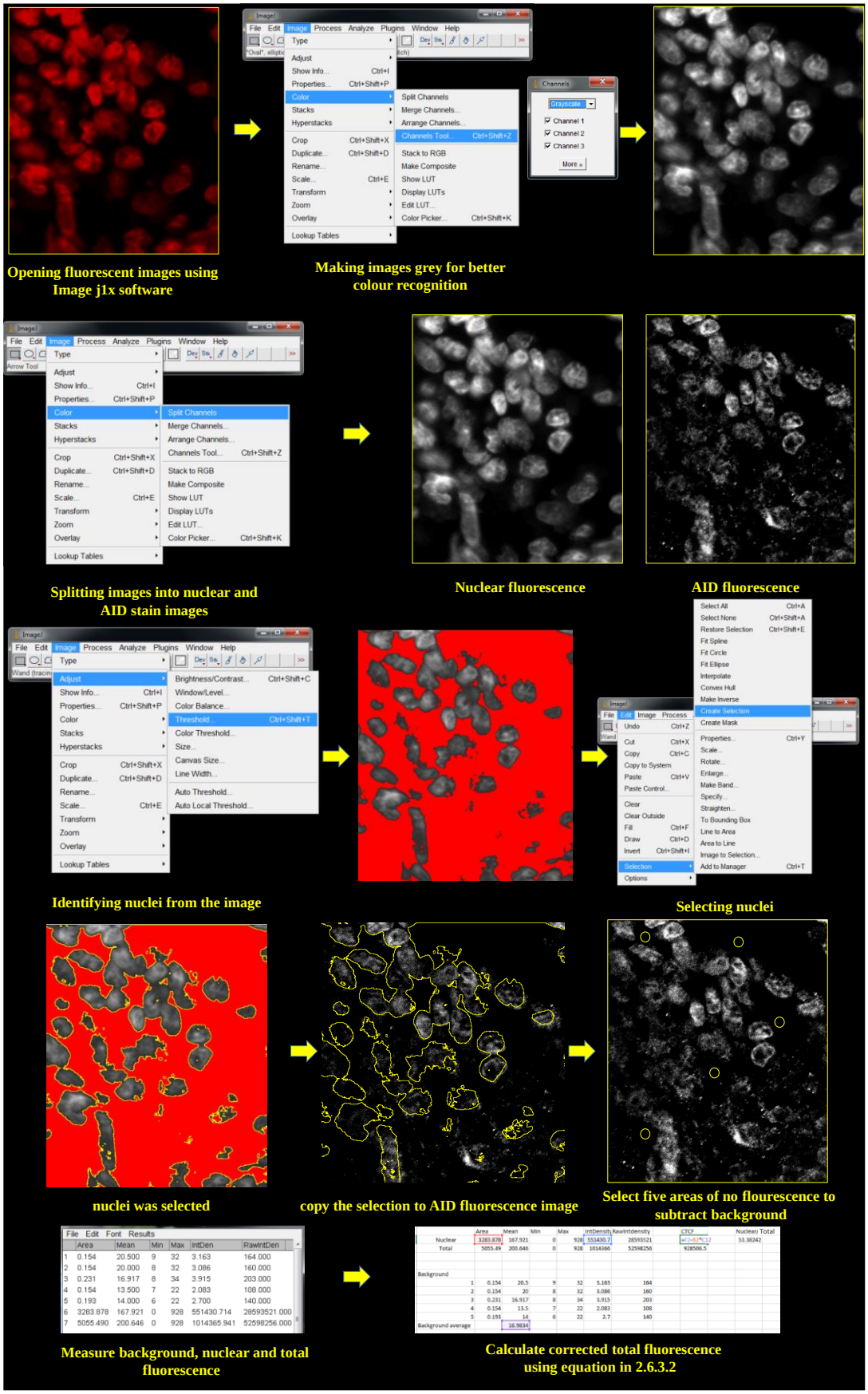

Figure 2.3: Calculating nuclear cytoplasmic/AID ratio using Image j1x software. The flow of arrows between the screenshots explains the technique used to quantify and calculate the intensity of staining.

2.6 qPCR quantification of genes associated with microenvironment in FL:

2.6.1 Optimizing qPCR protocol:

This experiment was done to explore the correlation between *AID* and expression signature of genes involved in tumour microenvironment in FL that had been proven to correlate with disease outcome (95).

To quantify expression of these genes as shown in Table 2.13, we optimized a method of qPCR using Taqman primers (Thermo Fisher Scientific, UK), published by Byers and colleagues (273).

Because of limited time we had for this last chapter, also, due to the limited availability of remaining cDNA of the PACIFICO samples, we only included six out of 66 candidate genes in this study. Three of the six genes (*STAT4*, *LEF1* and *CD2*) were selected as they were associated with good prognosis and related to T cells, while the other three (*C3AR1*, *LGM* and *MAPK1)*, were associated with bad prognosis. The details of these six genes are listed in Table 2.14.

Taqman Primer Probe Set	**Cell marker**	**Pathway***	**Prognosis**	**Taqman primers catalogue number**
MAPK1(Mitogen-activated protein kinase 1)	Many mammalian cells	proliferation, differentiation, transcription regulation and development	Bad	(Hs01046830_M1)
C3AR1	Macrophage and dendritic cells	Signalling by GPCR, complement and coagulation cascades.	Bad	(Hs00269693_s1)
LGMN	Macrophage and dendritic cells	processing of bacterial peptides and endogenous proteins for MHC (Major histocompatibility complex class) II	Bad	(Hs00271599_m1)
CD2	T cell	Adhesion and costimulatory molecule	Good	(Hs00233515_m1)
LEF1	T cell	transcription factor	Good	(Hs01547250_m1)
STAT4	T cell	transcription factor	Good	(Hs01028017_m1)
GAPDH	All mammalian cells	EC gene		(HS02758991_G1)
ACTB	All mammalian cells	EC gene		(HS01060665_G1)

Table 2.14: FL microenvironment-related genes studied in this thesis.

* Information about pathways these genes are involved in was taken from GeneCards (http://www.genecards.org/) and Byers and colleagues' paper (273).

20µl qPCR reaction was composed of 10µl of Taqman ® Universal PCR Master Mix (Thermo Fisher Scientific, UK), 8µl water, 1µl cDNA and 1µl of Taqman assay which includes primers and a specific probe. The qPCR was conducted in a LightCycler ® 480 and LightCycler ® 480 Multiwell Plate 96 (Roche, UK) with a programme detailed in Table 2.15.

Step	Temperature	Time	Cycle
Initial denaturation	95ºC	10 min	1
Denaturation	94ºC	40 sec	40
Annealing	56 ºC	30 sec	
Extension	72^0C	1 min	
Final extension	72^0C	1 mins	1

Table 2.15: qPCR protocol used for Taqman reaction.

2.6.2 Selection of cell lines to be used as a control for expression of these genes:

To amplify genes representing TME in PACIFICO samples, we needed to choose positive control cell lines for the qPCR reaction; to do that, we ran the qPCR reaction for six malignant cell lines available at the department: those are Raji cell line (Burkett's lymphoma) (257), Karpas 422 cell line (DLBCL) (274), Daudy cell line (Burkett's lymphoma) (275) and Jurkat cell line (Acute T cell Leukaemia) (276). We found that genes *MAPK1*, *C3AR1* and *LGMN* were expressed in the Daudy cell line, while *LEF1*, *CD2* and *STAT4* were best expressed in Jurkat cell lines, so these cell lines were used as a positive control for those genes. Each qPCR measurement was repeated three times (technical repeats) (Table 2.16).

Gene	Cell line	Average qPCR cycle threshold
MAPK1	Daudy	18.29
C3AR1	Daudy	21.89
LGMN	Daudy	23.85
LEF1	Jurkat	22.94
CD2	Jurkat	26.74
STAT4	Jurkat	30.34

Table 2.16: Shows cell lines amplifying TME genes that are used as a positive control for these genes' expression in PACIFICO samples.

RNA was extracted and cDNA were prepared from these cell lines in the same way as described in Sections 2.2.3 and 2.3.4, respectively, and the cDNA was then used in the optimized qPCR reaction.

2.7 Statistical analysis:

Statistical analyses were performed using IBM SPSS 22 software. A *P* value of 0.05 or less was regarded as statistically significant. Details of tests used and explanations can be found in relevant results sections.

Chapter 3

Optimization of RNA and DNA extraction from FFPE tissue samples of FL

3.1 Introduction:

The golden age of molecular cancer research began with the discovery of DNA's structure (277). This led us to understand many aspects of cancer development and progression. We began to understand the relationship between gene mutations and cancer, and this knowledge, in turn, increased our chances of developing cancer therapies. For example, discovery of defects in DNA repair mechanisms in cancer cells enabled us to understand how cancers can be initiated, and how we can target tumour cells (277).

With respect to RNA, it was considered for a long time to be merely involved in the translation from DNA to protein. However, scientists have recently established that it has another functional importance, with the discovery of non-coding RNAs. Scientists have shown that expression of both forms of RNA (coding and non-coding) are closely related to the formation, progression and metastasis of cancer (278).

One of the most significant advances in molecular biology during the last decade has been the application of NGS in sequencing DNA to unprecedented depths, to reveal further astonishing knowledge of the human genome. Initially NGS was performed on fresh tissue material or fluids. For decades, however, surgical biopsies and post-mortem tissues had been preserved using formaldehyde followed by tissue dehydration and paraffin embedding.

These FFPE tissues with the associated clinical data represent a major resource for DNA and RNA that can be used for molecular testing. Such preservation prevents tissue from degradation to keep tissue structures as close as possible to a living state (4). However, nucleic acids are often partially degraded in those FFPE tissues, compared to fresh or frozen tissues (279).

Many factors affect the integrity of nucleic acid in FFPE samples, these include the time interval between surgical procedure/sampling and fixation, size of the tissue and storage conditions (247).

Additionally, formaldehyde fixation of FFPE tissues is associated with cross-linking of RNA and DNA with proteins and with each other. Moreover, the current methods of nucleic acid extraction from FFPE samples affects the integrity of DNA and RNA. Therefore, use of better, less aggressive extraction methods is required to improve the quality of nucleic acid extracted from FFPE samples (280, 281).

One of the aims of this thesis was to determine whether the AID level at diagnosis of FL is important in FL diagnostics and/or predicting its progression. For that purpose, it was important to examine AID at mRNA and protein levels, as well as its sub-cellular localization. It was also important to determine the correlation of AID with ongoing *IGHV* mutation in FL, and its associations with mutations in other genes detected by NGS in a parallel study that may play a role in FL pathogenesis.

With the exception of AID protein expression that was examined by IHC and IF in tissue sections, all other experiments for this thesis were undertaken using DNA and RNA.

As stated in the General Introduction, we used tissue samples from the PACIFICO trial in this study. Since materials for this study were limited, we show our efforts in this chapter to optimize conditions of RNA and DNA extraction from FFPE tissue samples, with the aim

of obtaining higher quantity and better quality of nucleic acids for molecular testing in this study.

3.2. Procedures and Results:

3.2.1 Optimization of RNA extraction from FFPE samples:

3.2.1.1 Choosing the best kit for extraction of RNA from FFPE tissues:

As described in the literature, multiple kits are commercially available for RNA extraction from FFPE samples, including the RNeasy FFPE kit (Qiagen, Valencia, CA, U.S.A.), RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE tissues (Ambion, Austin TX), ReliaPrep™ FFPE Total RNA Miniprep System (Promega, UK), PureLink FFPE Total RNA Isolation Kit (Invitrogen, Carlsbad, CA) Absolutely RNA® FFPE Kit (Stratagene, La Jolla, CA), high Pure RNA paraffin kit (Roche Diagnostics, Mannheim, Germany), and Agencourt FormaPure (Beckman Coulter, Beverly, MA).

The most commonly used kits, RNeasy FFPE kit of Qiagen and ReliaPrep™ FFPE Total RNA Miniprep System of Promega were chosen in the following experiments, to compare the quantity and quality of the resulting RNA.

These two kits were used to extract RNA from FFPE tissue samples of three cases with FL and two with NHL stored in the LBIH according to the manufactures' instructions, and three individual experiments were performed for each sample.

It was found that the Qiagen RNeasy FFPE kit produced higher concentrations of RNA in the same elution volume (20μl) from the same amount of tissue (two sections of 4μm thickness) compared to the Promega ReliaPrep™ FFPE kit in most experiments (13/15) using samples from five cases (Figure 3.1). Thus, the averages in all of the five cases were higher with the former kit. Using the paired t test, this difference was statistically significant (P= 0.003). No noticeable change was found between the two kits in the remaining single experiments in two cases as presented in red and blue, and in both patients the average was

better when applying the Qiagen kit compared to the Promega kit (average: 128 ng/µl vs 83 ng/µl for the brown and 130 ng/µl vs 124 ng/µl for the red case for the Qiagen and Promega kits, respectively). The Qiagen kit was therefore used for RNA extraction in all subsequent experiments, as detailed in Chapter 2, Section 2.3.1.

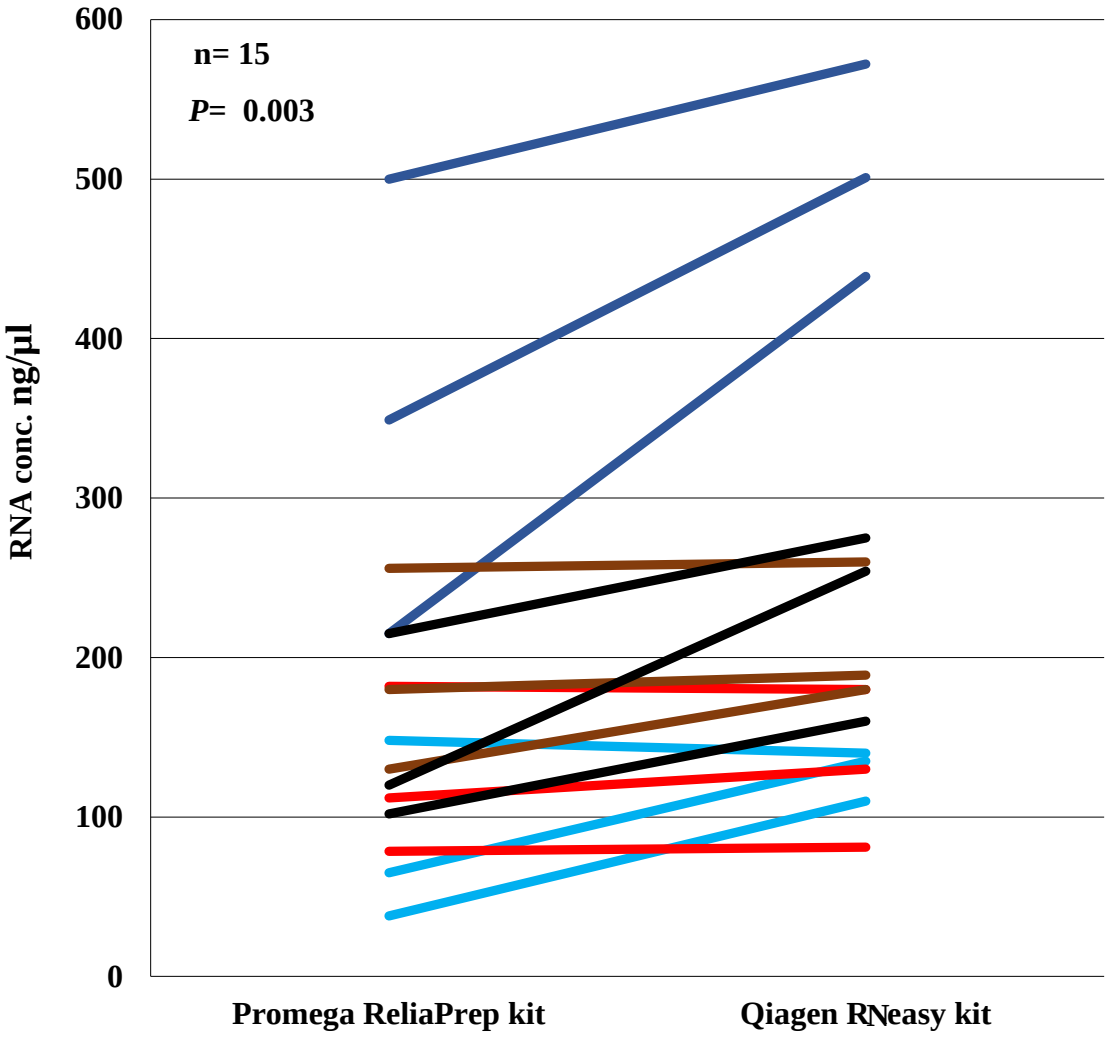

Figure 3.1: Comparison of yields of total cellular RNA extracted from FFPE samples with two different kits. RNA was extracted from FFPE tissue samples of five patients with NHL (two sections of 4µm thickness) in three independent experiments with ReliaPrep™ FFPE Total RNA Miniprep System of Promega and RNeasy FFPE kit of Qiagen. Concentration of the RNA in the same volume was compared using a paired t test. Experiments on the sample of each of the five patients are colourised.

3.2.1.2 Thickness of FFPE tissue sections used for RNA extraction:

Because the yield of RNA amount depends on many factors, including the size of tissue material in the FFPE block, a large variation of RNA yield was found between the examined cases (0.3-24 µg).

Moreover, the kit manufacturer stated that too much starting material reduces RNA yield, due to overloading the spin column. For each extraction, the RNeasy FFPE kit handbook suggested to use a maximum of four sections per FFPE sample, each with a thickness of 10µm. However, there was no suggestion of a minimum amount to be used.

Given that the use of thick sections not only reduces nucleic acid yield as suggested in the manufacturer's protocol, but also wastes precious FFPE tissue blocks, RNA from only two sections of 4µm per sample was extracted. In an experiment using samples from three FL cases, it was established that the RNA yield was 2.25-3 µg, which is sufficient for the planned experiments. Consequently, this amount of starting material was used in all subsequent RNA extractions.

3.2.1.3 RNA degradation in FFPE samples during storage:

To examine the effects of FFPE block storage time on tissue RNA quality, three FFPE samples stored for six months or less, two samples stored for about one year, and four samples stored for 10 years or more, were compared. All samples had been stored in the dark at room temperature.

The RIN was determined using a 2100 bioanalyzer instrument, with the RIN ranges being between one and 10 representing highly degraded and intact RNA, respectively. Examples of degraded RNA measured by 2100 bioanalyzer are shown in Figure 3.2.

We found that there was no difference in RNA integrity between samples stored for about one year and those stored for 10 years or longer (P= 0.438) (Mann-Whitney U test). However, FFPE samples stored for six months had much better RNA quality compared to both groups of tissues stored for one year or more (P < 0.024) (Mann-Whitney U test) (Figure 3.3).

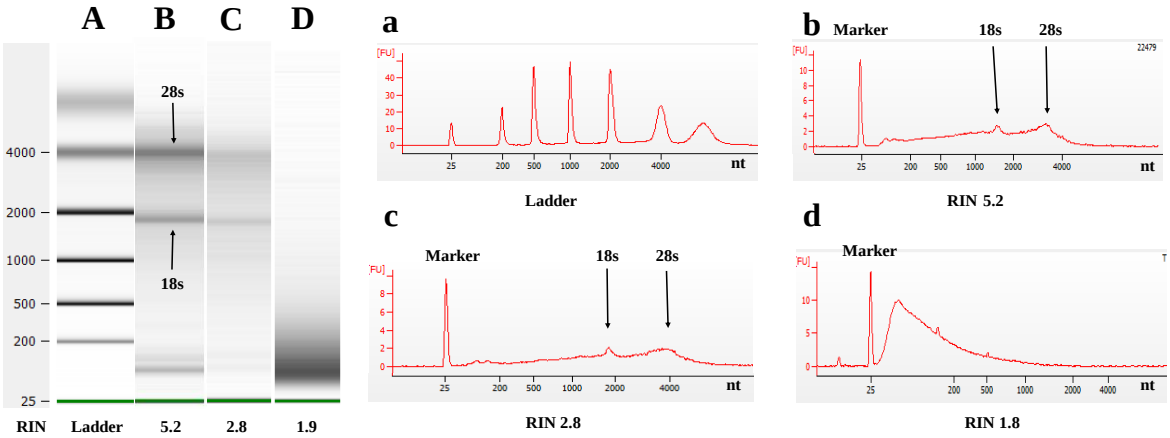

Figure 3.2: Examples of RNA degradation measured by 2100 bioanalyzer. The RNA ladder as shown in gel image (A) and electropherogram (a); RNA extracted from a FFPE sample with RIN of 5.2 as shown (B) and (b); RNA from a FFPE tissue sample with RIN of 2.8 as shown in (C) and (c); and RNA from a FFPE sample with RIN of 1.9 as shown in (D) and (d).

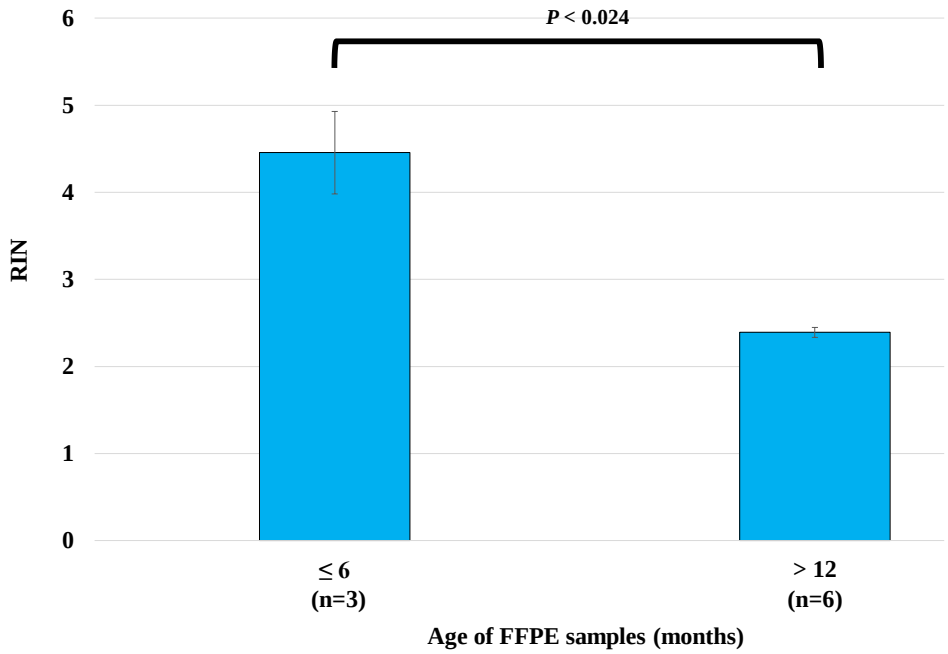

Figure 3.3: Comparison of RIN of RNA extracted from FFPE samples stored for different periods of time. Results are presented as mean ± SE measured in triplicate, Mann-Whitney U test was used.

3.2.2. Optimization of DNA extraction from FFPE tissues:

It is well known that DNA is more stable than RNA (282, 283). This is because compared to de-oxy ribose in DNA, the RNA contains an additional OH group in 2', which reduces its stability making RNA more prone to hydrolysis (284). In order to examine our study hypothesis presented in Chapter 1, we needed DNA to:

1. Explore the role of AID in ongoing mutations in FL.

2. Identify the correlation between AID levels and the mutation of other important genes in the pathogenesis of FL using NGS.

Efforts were made initially to isolate better quality and higher quantity of DNA from the least amount of tissues of precious PACIFICO clinical trial FFPE specimens. To select the best kit for DNA extraction from FFPE samples, a literature review established that the QIAamp DNA FFPE Tissue Kit (Qiagen) had proven to be one of the best kits that yielded the highest amount and the best quality of DNA from FFPE tissues among all commercially

available kits (285-287). Hence, we decided to use this kit for DNA extraction in all of the planned experiments requiring DNA.

3.2.2.1 Optimizing DNA extraction kit to obtain higher quality and quantity of DNA:

In order to optimize the protocol for obtaining more DNA from the FFPE samples, DNA yields achieved with different duration of protein digestion were compared. In experiments using 3 FL and 2 NHL samples stored in the LBIH in addition to one of the PACIFICO trial FL samples, proteins in one section of 4µm thickness of FFPE tissues was digested for one hour and overnight before extraction of gDNA, as described in Chapter 2, Section 2.4.1. Each experiment was repeated twice. The DNA yield (presented as concentration in the same volume) was greater with the overnight digestion in 11 out of the 12 experiments. In only one case, DNA showed a lower concentration as compared to the shorter digestion. However, the average level of the two experiments for this case was almost the same (63 versus 61.5) (brown in Figure 3.4.). Taken together, the overnight treatment was associated with an increase in the DNA extracted (Figure 3.4) ($P = 0.017$, Paired t-test). This was also seen when six sections from each of the six cases were tested (data not shown).

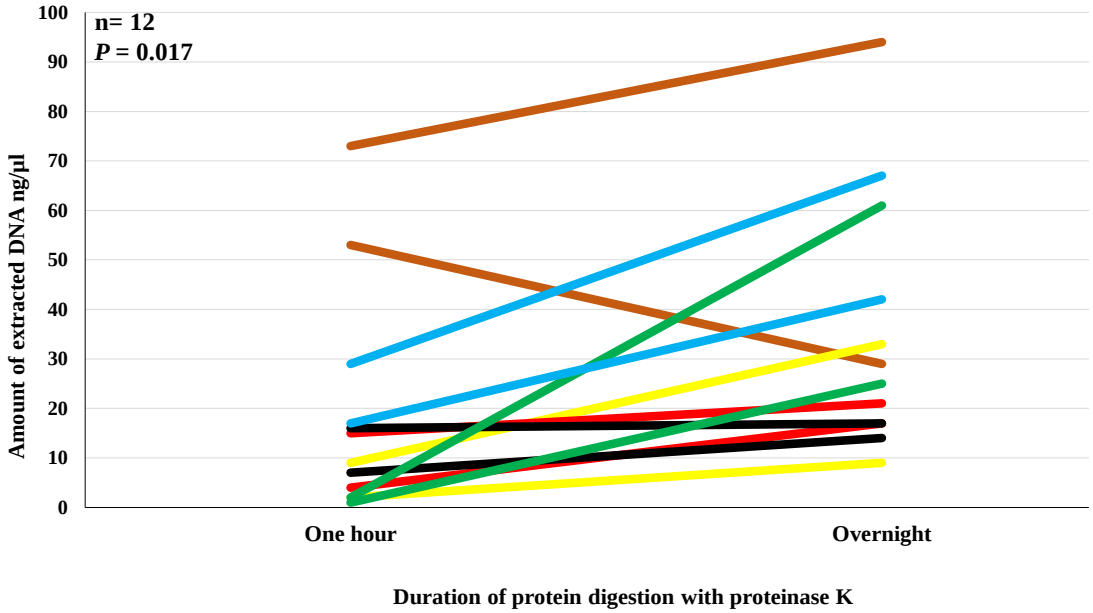

Figure 3.4: Comparison of DNA yield (presented as the concentration of DNA in the same volume of eluent) extracted from FFPE samples after protein digestion for different durations. The FFPE tissue sample sections from each of the six patients with FL were incubated with proteinase K for one hour and overnight. This procedure was repeated twice in different experiments as colourised (biological repeats). The paired t-test was employed to compare the difference between the two digestion durations.

The quality of extracted DNA can be determined by examining the length of gene sequences amplifiable in PCR. In order to do this, about 100-bp, 200-bp, 300-bp, 400-bp and 500-bp PCR products within the EC gene *GAPDH* were amplified using DNA extracted from the FFPE samples treated with proteinase K, for one hour and overnight.

It was found that ≥100bp PCR fragments were produced in all of the six samples. In 4/6 patient samples, sharper bands were obtained from DNA prepared by overnight incubation with proteinase K (Table 3.1). This suggested the prolonged treatment allowed more time for reversal of protein-DNA crosslinking in FFPE tissues, and consequently released an increased amount of DNA with good amplifiable quality. This can be seen in Figure 3.5 below: compared to one-hour treatment, the stronger and/or larger PCR bands are presented in samples of three cases incubated overnight with proteinase K (Figure 3.5).

Hence, it was decided to use overnight incubation with proteinase K at 56°C for all of the samples in the planned experiments.

LTB Samples	One hour proteinase k (PCR product size (bp))					Overnight proteinase k (PCR product size (bp))				
	100	200	300	400	500	100	200	300	400	500
Patient (217)	++	+++	++	++	+	+++	++++	++++	++++	++
Patient (T247/99)	++	++	+	+	-	++	++	+	+	-
Patient (T15/98)	++	+++	+++	+++	+	++	++++	+++	+++	++
Patient (T106/93)	++	+++	+++	++	-	++	+++	++	++	-
Patient (T305/03)	+	++	+	+	+	+	+++	++	+	-
Patient (T54/94)	++	+++	+++	+++	+	+++	+++	+++	+++	+++

Table 3.1: Comparison of PCR products of *GAPDH* between FFPE samples with protein digestion for one hour and overnight. -: no product; +, ++, +++ and ++++: PCR product band intensity change from weak to strong.

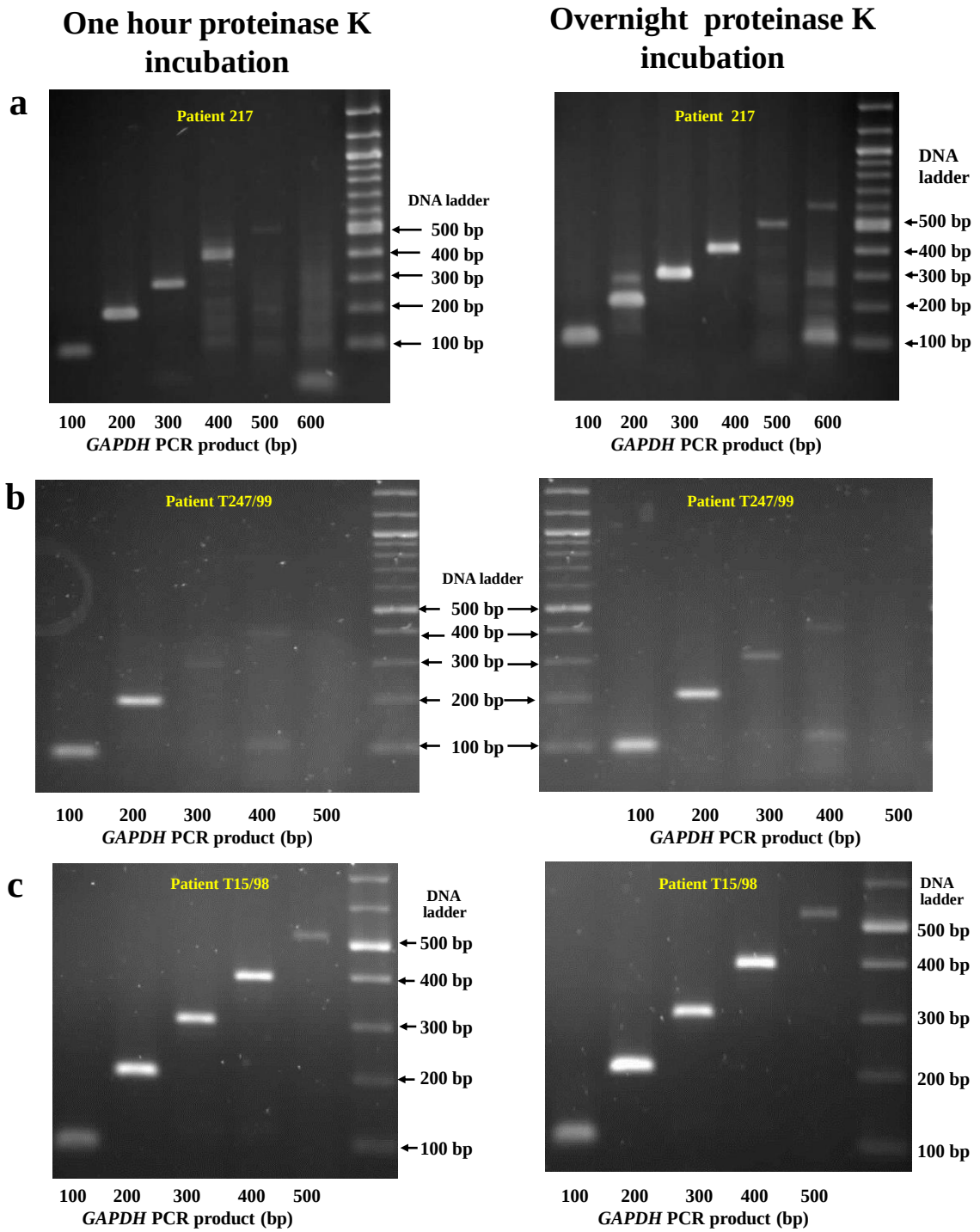

Figure 3.5: Comparison of length of *GAPDH* gene fragments amplified using gDNA extracted from FFPE tissue samples pre-treated with proteinase K for different periods of time. FFPE tissues from three FL patients, (a, b and c) were incubated with proteinase K for one hour (left) and overnight (right) before DNA extraction. The PCR products amplified with primer sets designed for 100- to 500-bp gDNA fragments of *GAPDH* were visualised on agarose gels following electrophoresis.

3.2.2.2 Adaption of a system to measure DNA integrity:

The quality of DNA is important for selecting the best samples for ongoing mutation analysis of *IGHV* and NGS experiments. Since we needed DNA templates between 200-400 bp in subsequent experiments, we designed PCR primers that amplify about 200-bp, 300-bp and 400-bp DNA fragments of *GAPDH*, and used them to detect DNA integrity in all samples. Intact DNA prepared from the Karpas 422 cell line (DLBCL) was used as a positive control. Its PCR band intensity was scored as ++++, whereas the bands representing partially degraded DNA were scored as +++, ++ or +. A negative score was given to any PCR that failed to amplify the gene fragment (Figure 3.6).

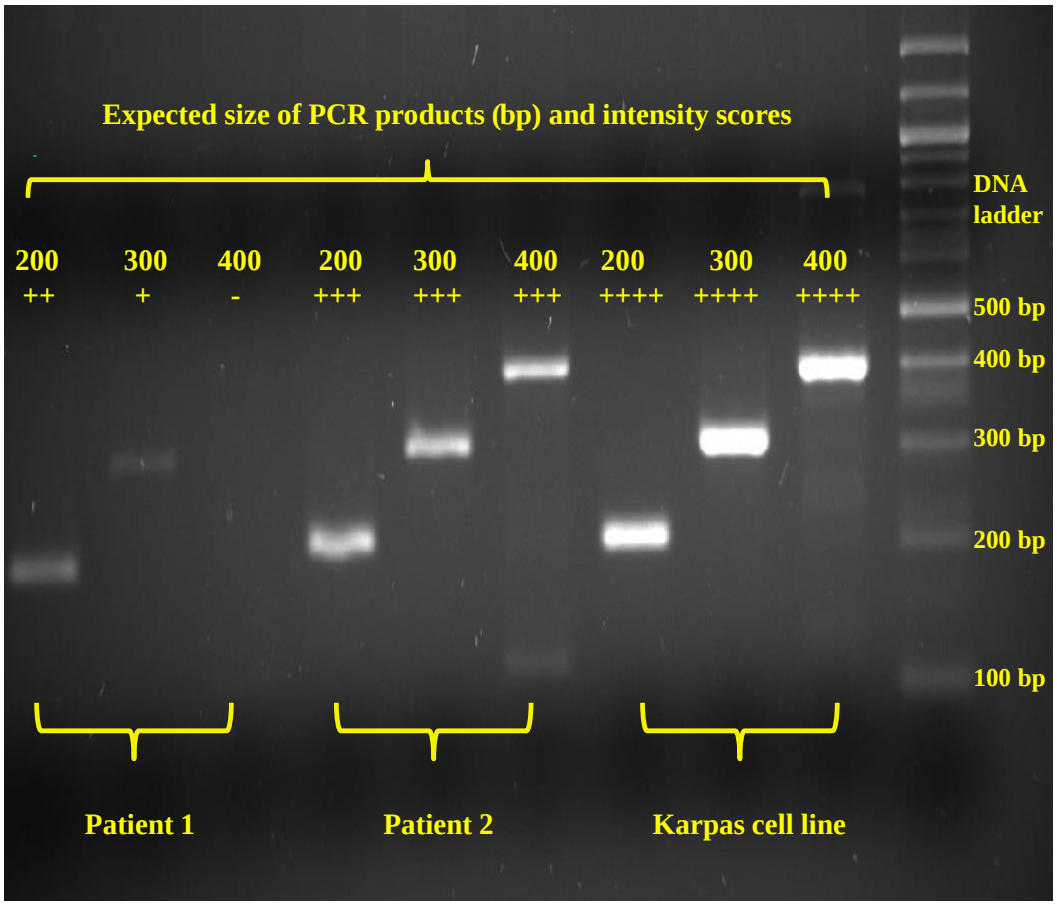

Figure 3.6: PCR primers to amplify *GAPDH* gene at 200, 300 and 400 bp were designed to detect DNA quality from FFPE tissues of PACIFICO samples, intact DNA of Karpas 422 cell line was scored (++++), partially degraded DNA were scored as +++, ++ or +.

In total, DNA from all 87 samples originating from the PACIFICO trial was extracted, with an average DNA amount of 18.8 µg (range, 0.15 and 101µg), and the DNA yield was high enough for NGS (≥ 0.5µg) in 79 (91%) of the 87 cases. A sufficient quality of DNA was able to be obtained for most of the samples to analyse ongoing mutations (Table 3.2).

Band intensity / **Product size**	-	+	++	+++	++++	**Total**	**% of PCR positive samples**
200bp	0	14	27	36	10	87	100.00%
300bp	7	24	25	25	6	87	91.95%
400bp	23	29	20	13	2	87	73.56%

Table 3.2: The number (%) of DNA samples with different length and amount as estimated with PCR amplified *GAPDH* gene fragments.

3.3 Discussion:

In this Chapter, the extraction procedures for both RNA and DNA from initial test samples were able to be optimised. The DNA and RNA required for the main study are from FFPE tissues with all of the associated limitations of formaldehyde and paraffin embedding

With respect to RNA degradation during the storage in FFPE, it was established that samples that were stored for up to six months were associated with a significantly better RNA quality than those stored for up to one year. However, there was no difference in the RIN for those samples stored for one year and >1 year (up to 19 years) (Figure 3.2). These results are in agreement with Ahlfen and colleagues, who found a similar RNA integrity between a series of FFPE samples stored for > 1-10 years under the same storage conditions (288).

Regarding DNA extraction from FFPE samples, we found that overnight incubation with proteinase K at 56°C improved both the quantity and quality of DNA compared to one hour

incubation at the same temperature (Figures 3.4 and 3.5). Proteinase K digestion is a critical step for proper isolation of the largest possible amount of DNA from FFPE tissues with commonly used commercial kits (289). In this step, both concentration of proteinase K and duration of treatment are important factors to optimize (290).

When comparing proteinase K incubation durations of 39, 68 and 116 hours, Hassani and colleagues found that 68 hours yielded the largest DNA fragments (290). While Sengüven and co-authors found that 72 hours' incubation yielded better quality and quantity of DNA compared to overnight incubation (286). These protein digestions were longer than those used in our study. However, the temperature was one degree lower than ours (55°C versus 56°C). It is difficult to conclude which procedure was better for DNA extraction from the FFPE tissue samples.

Nevertheless, it became a general practice to prolong incubation time with proteinase K to improve DNA yield from FFPE samples (291). In fact, other people also prolonged the protein digestion to overnight or longer either using the same (292) or different extraction kits (293).

We used PCR amplification of an EC gene to estimate the integrity of DNA extracted from FFPE samples (Figure 3.6). This is a commonly used quality control technique by many research groups to detect DNA quality (294-297).

We achieved a 100% amplification for DNA product sized 200 bp. This was greater than that reported by Lin and colleagues, who achieved successful DNA amplification in 97% (range: 88–100%) of samples only for a PCR product (using the beta actin gene rather than the *GAPDH* that we used). This was probably because the optimizations performed in our work improved the quality and quantity of DNA extracted from FFPE tissues (298). Thus, our results of DNA extraction from the FFPE samples were comparable to achievements by others. For example, Hassani and colleagues, with conditions of DNA extraction optimised,

managed to have all DNA samples amplified for a 104-bp PCR product, 50% of samples a 351-bp (20% faint bands) and 20% of samples a 560-bp products (290).

In conclusion, suitable kits were selected and conditions optimized for extracting higher amount and better quality of RNA and DNA with the less starting material of FFPE specimen. These optimised methods were applied in subsequent studies in this project, outlined in the following chapters.

Chapter 4

AID expression in FL

4.1 Introduction:

The AID enzyme is well known for inducing *IGHV* mutations in SHM and CSR (198). These are physiological processes required to produce highly specific antibodies against antigens, as illustrated previously in Chapter 1 (Figures 1.2 and 1.3).

Strict control of AID is important for preventing mutations affecting genes other than *IGHV*. However, molecular biological techniques have revealed that AID targets other genes involved in pathogenesis of lymphomas (198).

Measuring AID expression is an attractive approach for studying its roles in B-cell malignancies. Unfortunately, this is hampered by technical difficulties in extracting good quantity and quality of nucleic acids and proteins from FFPE tissue samples.

Different methods for AID expression in cancers have been used. For *AID* mRNA expression, qPCR was measured either using SYBR Green (299) or Taqman primers and reagents (300).

To detect AID protein levels, different techniques have been employed, including FACS analysis, western blotting (37, 299), IHC (141) and immunofluorescence (301).

However, to our knowledge, no publication to date has provided data concerning AID expression at the levels of mRNA, IHC, IF and subcellular localization in FL for one cohort of patients.

To identify AID expression at all of these levels, we undertook all of the optimizations required to make a reliable measurement of AID level from partially degraded FFPE as the starting material.

4.2 Patients and Methods:

87 patients diagnosed with FL and recruited in the PACIFICO clinical trial were included in this study. Most patients had advanced disease at trial entry: Stage I, n=0 patients; stage II, n=11; stage III, n=40; stage IV, n=36. Clinical data of patients were blinded until completion of experimental work and the commencement of thesis writing.

4.2.1 *AID* mRNA expression using qPCR:

We optimized RNA extraction methods from FFPE to obtain more RNA; details are described in Chapter 2, Section 3.2.1.

4.2.1.1 Quality and quantity of RNA extracted from FFPE:

Before starting qPCR experiments for AID expression of FL, we wanted to establish a cut-off level of RNA quality that can be reliably used in our experiments. To detect the quantity of RNA extracted from FFPE, we used the Qubit fluorometer 2. To analyse the integrity and quality of RNA, we used the 2100 Bioanalyzer instrument of Agilent.

In 83 out of 87 (95%) PACIFICO FFPE samples, RIN values ranged between one and 5.2 with an average of 2.2. Four samples had a RIN value of ‘NA’ (not available), which is usually an indication of highly degraded RNA. The total RNA extracted from FFPE tissue ranged between 0.225-36 µg, with the average being 7.36 µg.

4.2.1.2 Storage conditions and duration of PACIFICO FFPE tissues:

PACIFICO FFPE tissues used in our work were stored in the dark at room temperature.

The age of the 87 FFPE samples ranged between 1-108 months in paraffin, with a median of 24 months. We explored the correlation of sample age and the RIN of all samples. We established a weak but significant negative correlation between sample age and RIN, with poor RNA quality being associated with increasing duration of FFPE samples storage, (r= - 0.265, *P*= 0.016) (Spearman's correlation), (Figure 4.1 (a))

In addition, there was a statistically significant difference between RIN of samples aged 12 months or less compared to those stored for more than one year (Figure 4.1(b)) (*P*= 0.02), (Mann-Whitney U test).

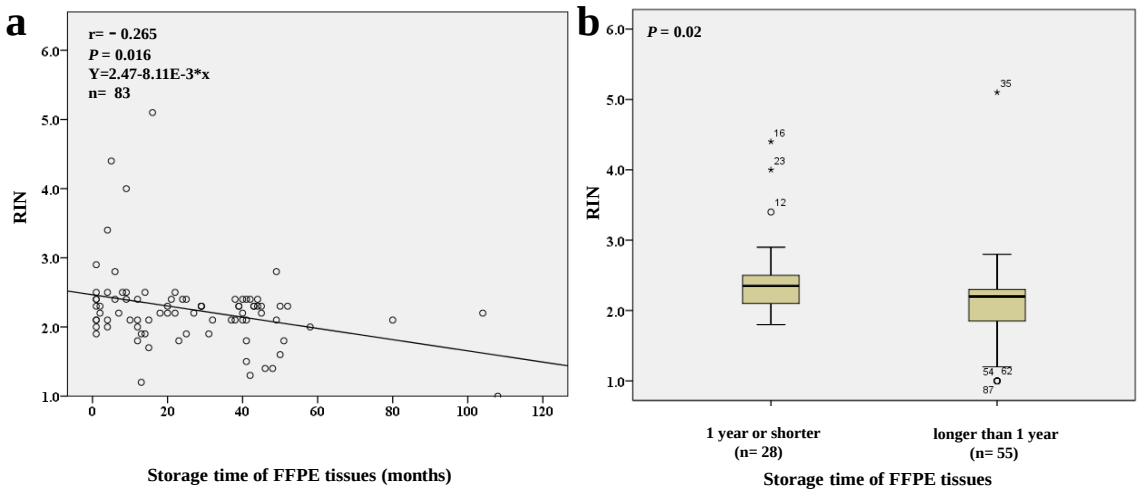

Figure 4.1: RNA integrity was reduced during storage of FFPE samples. (a) A negative correlation between RNA integrity as presented as RIN quantified with Bioanalyser measurement and duration of sample storage. This was identified amongts the 83 samples with the Spearman's correlation analysis (r = - 0.265, *P*= 0.016); (b) A significantly lower level of RIN in FFPE samples stored more than one year compared to those stored for one year or less (*P*= 0.02), (Mann-Whitney U test).

To identify whether we could use the RNA extracted from FFPE to amplify *AID*, an EC gene (*GAPDH*) and one of germinal B-cell markers (*CD19*) were evaluated.

As a positive control, we used cDNA from a CLL patient blood sample, which was known to express AID protein.

We reversed transcribed RNA extracted from two FFPE tissues of the LBIH that were stored for a period >10 years, and then performed PCR using primers to amplify 300-400 base pairs for the three genes. We repeated this experiment twice, with method details being described in Chapter 2, Section 2.3.5 (Figure 4.2).

We did not obtain PCR products of *GAPDH*, *AID* or *CD19* on agarose gel analysis in both LBIH samples. This was expected to be due to degradation of RNA extracted from the FFPE tissue, so we subsequently optimized RNA extraction and designed primers to amplify smaller PCR products.

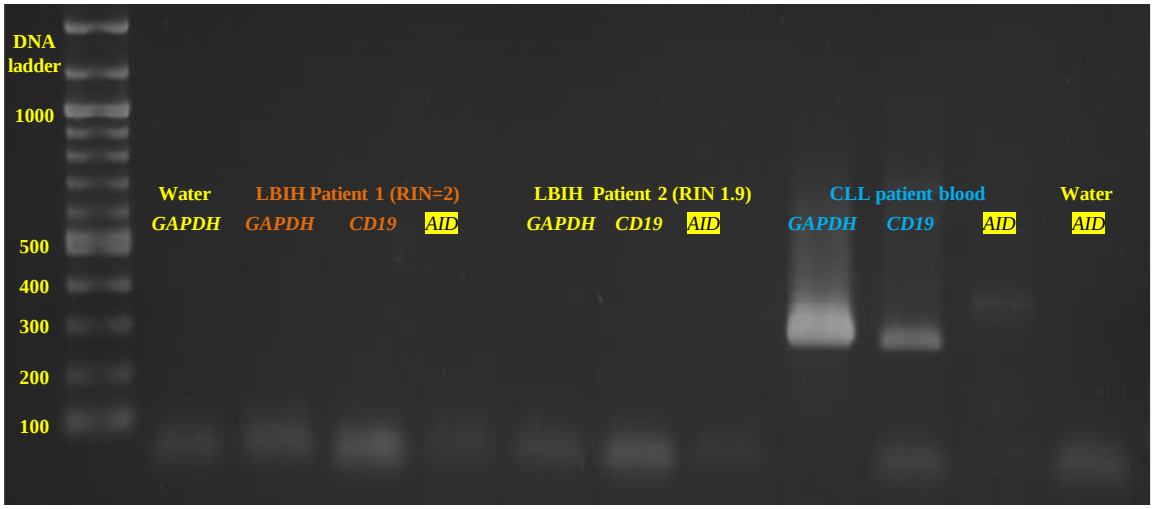

Figure 4.2: No amplification of genes from RNA of FFPE when primers amplifying large PCR products were used. *AID*, *GAPDH* and *CD19* could not be amplified in two old FFPE samples when PCR primers of 300-400 bp were used (orange and yellow), The fresh blood CLL sample amplified all three genes (blue); In each of the three different samples used, the experiment was technically repeated twice.

4.2.1.3 Checking the quality of RNA required for reliable qPCR results:

Fixation and paraffin embedding degrades RNA in FFPE tissues. To determine the degree of RNA degradation that would still allow us to use the sample for *AID* expression, we needed to compare expression of *AID* to the EC gene in fresh and degraded counterparts of the same sample. Unfortunately, this could not be done on our FFPE tissues, as we did not have the respective fresh frozen biopsy of the sample.

Hence, we decided to grow a B-cell lymphoma cell line that is known to express AID. We then extracted RNA and degraded it by heating for differing time periods. This allowed us to measure the qPCR expression ratio of *AID* to EC gene in fresh and degraded series of RNAs from the same cell line. Subsequently, we compared qPCR expression in intact and degraded series, with determine to what degree of degradation it was still possible to use the RNA.

For the above, we chose the Raji cell line (Burkitt's lymphoma cell line), which is known for high expression of AID (302, 303) . We degraded the cell line by heating at 95°C in batches from 3-30 minutes; this led to a degraded RNA series with RIN ranging from 9.1-NA, with NA being an indication of severely degraded RNA (Figure 4.3 (a)).

To understand the effect of RNA degradation on *AID* expression by qPCR, we used the qPCR reaction that was optimized by our research group (Chapter 2, Section 2.3.7).

For further optimization for detecting suitable quality of RNA for qPCR, we compared oligo-dt primers and random primers to amplify *AID* and *GAPDH* in the series of degraded Raji RNA samples (RIN from 9.1 - NA) to produce small (90 bp versus 101 bp)

and large (386 bp versus 315 bp) *AID* and *GAPDH* PCR products, respectively, random primers amplified mRNA products in the above samples.

We also saw that cDNA derived from RNA samples with RIN of 2.6 produced the same qPCR ratio of *AID/GAPDH* to the intact RNA; however, this ratio was not the same for samples of RIN of 1.9.

So, we further degraded two batches of samples to obtain RNA samples with RIN of 2.4 and 2.1. We found that the ratio of *AID/GAPDH* genes was achieved in degraded samples down to RIN of 2.1 was close to the intact RNA.

Each experiment was performed at least twice and each time three qPCR readings were taken (n=6).

The *AID/GAPDH* mRNA ratio in an intact sample is statistically significantly different from samples with RIN of 1.9 (P= 0.028) (Wilcoxon signed-rank test). None of the other ratios from 8.1 – 2.1 were statistically different from the intact RNA (Figure 4.3(b)).

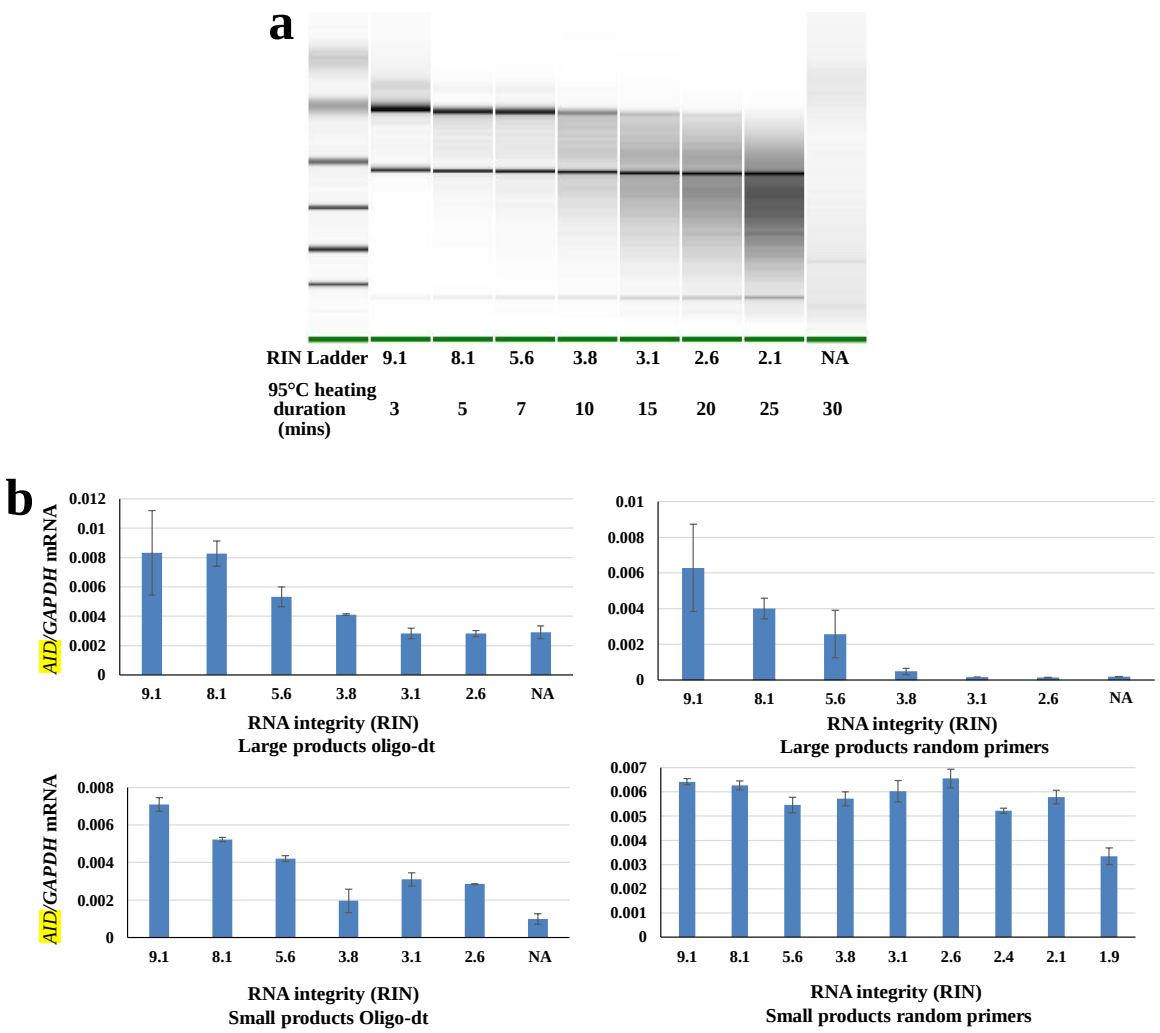

Figure 4.3: RNA RIN ≥2.1 produced reliable *AID/GAPDH* qPCR results. (a) We degraded RNA extracted from a Raji cell line by heating at 95°C for 3-30 mins. RIN measured by a bioanalyzer 2100 ranged from 9.1 – NA, NA usually results from severely degraded RNA; (c) Only small products of random primers can produce a ratio of *AID/GAPDH* in degraded RNA down to 2.1 that does not statistically differ from the ratio produced by intact RNA. A significantly lower ratio of *AID/GAPDH* was found for RNA RIN of 1.9 compared to intact RNA (P= 0.028) (Wilcoxon signed-rank test) (mean ± SD, n=6).

Hence, as a result of these experiments, we used random primers in RNA reverse transcription reactions and PCR primers that amplify products less than 101 bp. The suitable samples for mRNA expression of *AID* were only those with RNA of RIN ≥ 2.1. In our cohort of patients, out of 87 patients, 59 samples had a RIN of 2.1 or more (≈68%).

4.2.1.4 Selecting the best EC gene for qPCR reaction of *AID* gene:

The *AID* transcriptional level was measured as its ratio to an EC gene in qPCR. To do so, we needed firstly to determine what EC gene is most homogeneously expressing among different patients with FL, and therefore whether it was suitable to be used to as an internal control.

A review of the literature showed that other authors used the following genes as endogenous controls in FL: *GAPDH*, *HPRT1* and *TBP* (273, 304-306).

We used the same primers as in the above papers and performed our qPCR expression using our optimized protocol (Chapter 2, Section 2.3.7) of all three genes in nine cases of FL from the PACIFICO trial. Each experiment was repeated technically three times using our optimized qPCR protocol.

In comparison to transcriptional levels, as presented by the threshold of PCR cycles (C_t), a significant positive correlation was observed between any two of the three EC genes (Figure 4.4).

However, the slope in the linear regression and the correlation efficient (r) between *GAPDH* and *HPRT1*or *TBP* was closer to one (Figure 4.4. A and B) than those between *HPRT1*and *TBP* (Figure 4.4. C), suggesting that the variation of *GAPDH* expression among FL patients is smallest among the three EC genes. In addition, the C_t of *GAPDH* was smaller than 30 in eight out of the nine patient samples, with the C_t for the remaining one being closer to 32.5, whereas it was bigger than 30 in four, or all, patients for *HPRT1*and *TBP*, respectively (Figure 4.4). This indicated that the measurement of *GAPDH* mRNA in this qPCR was more reliable as compared to the other two EC genes tested. So, we subsequently used *GAPDH* as an EC gene in our qPCR to quantify *AID* transcriptional levels.

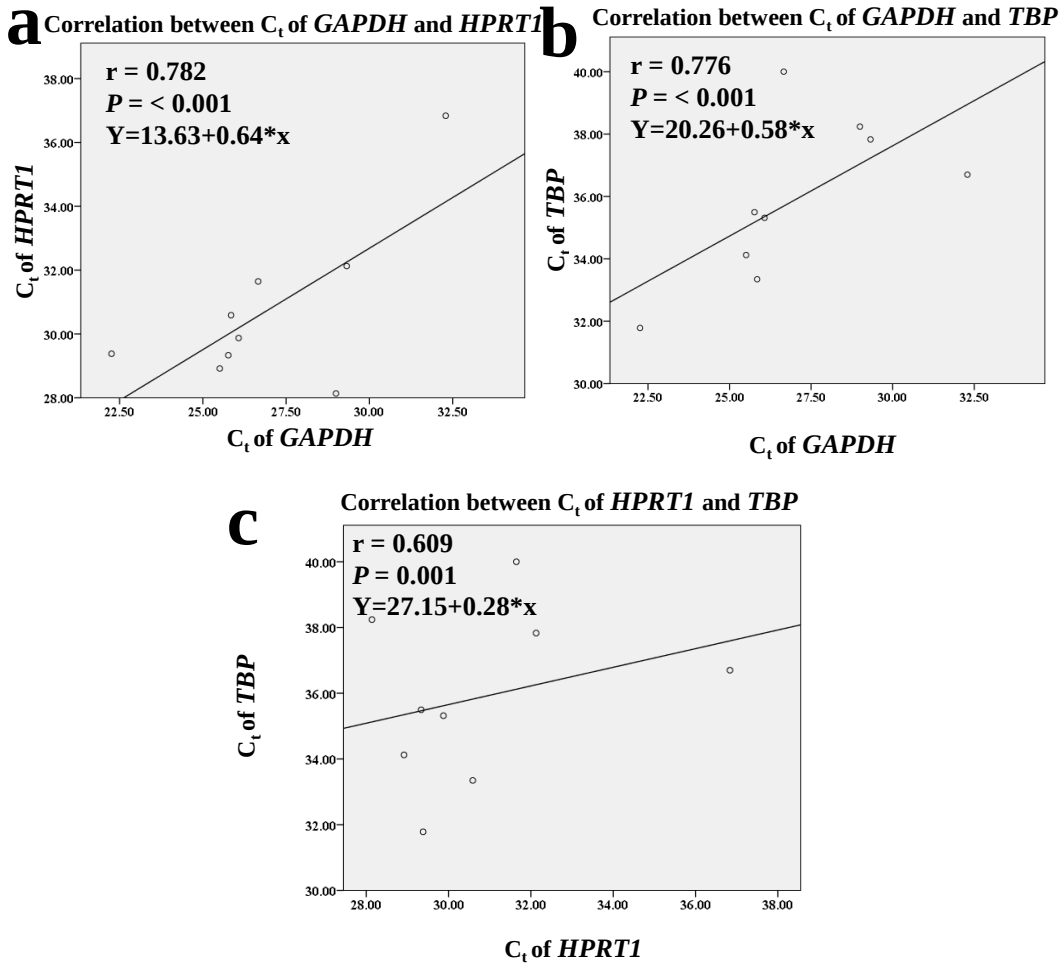

Figure 4.4: Comparison between different EC genes to choose the most suitable gene to produce the *AID*/EC gene ratio. Nine FL cases were used to explore this issue: we found a significantly positive correlation between all three genes. However, the linear regression slope and r value were closer to one between *GAPDH* against *HPRT1* and *TBP* (a and b) than between the latter two genes (c) (n= 9).

To recap, at this stage, the selection criteria for RNA quality to give reliable qPCR results was completed, the best EC genes were selected, and the qPCR reaction for estimation of *AID* mRNA by qPCR was optimised.

Our total cohort was 87 cases, 28 samples (≈ 32%) (≈ is approximately equal) had a RIN less than 2.1, and so we performed *AID* mRNA expression by qPCR in 59 out of 87 samples (≈ 68%).

As mentioned above, the Raji cell line (Burkitt's lymphoma) was used as a positive control (302, 303), and the K562 Chronic Myeloid Leukemia (CML) cell line as a negative control

(307, 308) for *AID* qPCR expression (details of methods provided in Chapter 2, Section 2.3.7).

The *AID* qPCR products were sent for sequencing in forward and reverse directions, and we established that the qPCR sequence of *AID* was consistent with the primer design. The qPCR ratio of *AID/GAPDH* varied widely in the 59 cohort samples, ranging between 0.0001- 0.72; the median was 0.02. Each experiment was repeated technically twice (Figure 4.7).

4.2.2 AID protein expression by IHC:

To elucidate the role of AID in FL pathogenesis, we wanted to identify total protein expression of AID by IHC. Details of this protocol are provided in Chapter 2, Section 2.6.1.1. To determine the best concentration of AID antibody to be used for IHC, we tried different dilutions of the AID monoclonal antibody (mAb). We found that a 1/500 dilution of AID in antibody diluent was the most suitable.

We used hyperplastic tonsil tissue that was stained by AID as a positive control, and IgG1 as an isotype control (Figure 4.5).

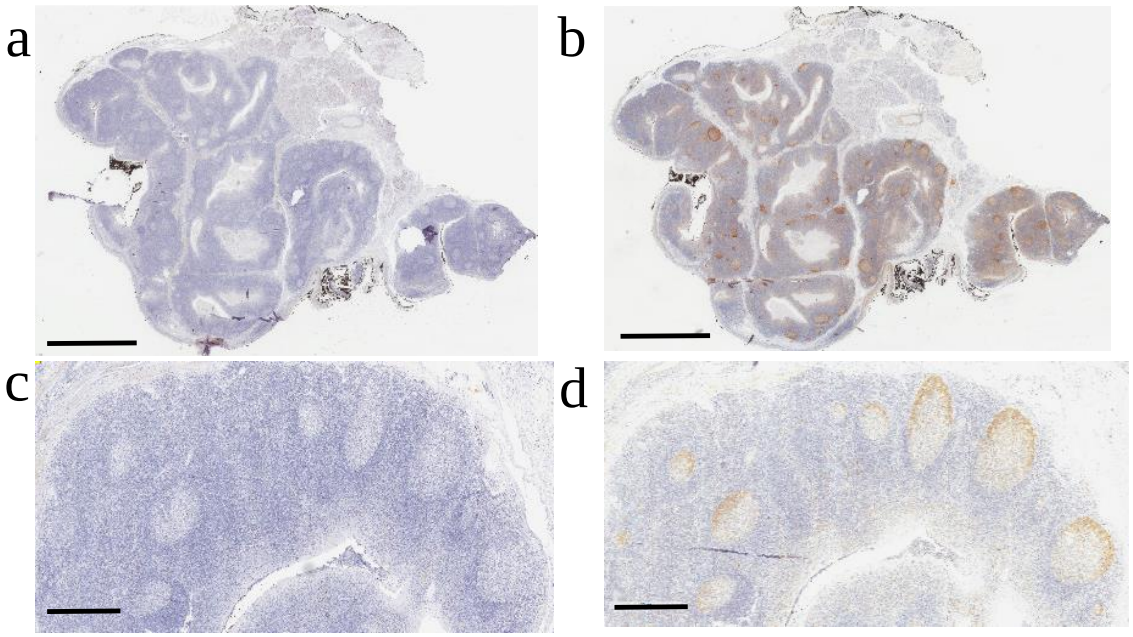

Figure 4.5: Hyperplastic tonsil was used as a positive and negative controls for AID protein expression by IHC. (a) Hyperactive tonsil stained with IgG1 isotype control (c) is a magnified image of the same stained tissue in (a); (b) The same tonsil was stained by AID mAb (d) is a magnified part of the image in (b). In (a) and (b) the scale bar is 3mm, in (c) and (d) it is 100µm.

We then stained all of our 87 patient cohort samples with the AID mAb; examples for different levels of AID protein expression by IHC can be seen in Figure 4.6.

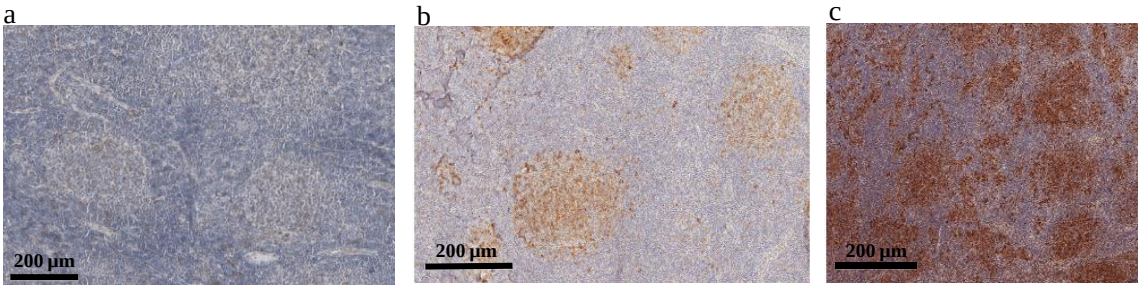

Figure 4.6: Different FL LNs stained with AID mAb using IHC technique, (a) LN scored 1, (b) LN scored 4, (c) LN scored 9.

We scored the slide sections according to the number of stained cells and intensity of staining using the following formula: % of lymphocytes stained X Intensity (Table 4.1) This scoring system is widely used to identify IHC score of a protein in lymphoma and other cancers (268-270, 309-312).

Lymphocytes		Intensity of stain		Final score
Percentage	Score	Degree	Score	= % of lymphocytes stained X Intensity
0	0	Negative	0	
1-24	1	Weak	1	
25-49	2	Moderate	2	
50-74	3	Strong	3	
>75	4			

Table 4.1: Scoring system used to score IHC slides stained with AID mAb.

The scoring of the slides was performed by two observers (myself and a specialist pathologist, Prof. Sarah E. Coupland). All of the 87 slides were scored. The AID score ranged from 1-9 with a median of 2.

4.2.3. AID mRNA and protein expression profiles in the cohort of patients with FL studied:

AID/GAPDH mRNA expression and AID IHC expression varied widely among FL PACIFICO patients (Figure 4.7).

AID IHC expression was calculated in all cases. *AID/GAPDH* mRNA expression was measured in about 2/3 of patients (59 out of 87) and 18/20 cases of nuclear to cytoplasmic (N/C) ratio of AID have *AID/GAPDH* mRNA data

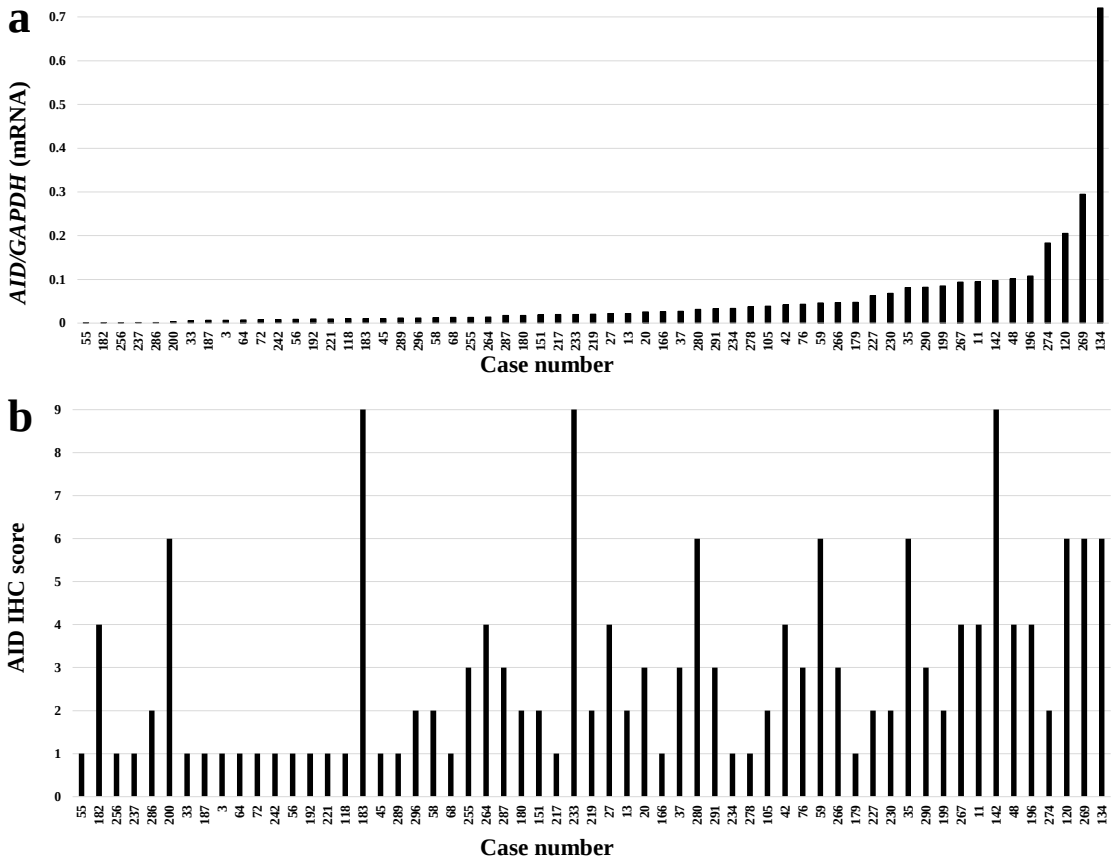

Figure 4.7: Expression of AID is variable among FL cases (a) Variable expression of *AID/GAPDH* mRNA, and (b) AID protein in PACIFICO FL cases.

4.3 Correlation between AID protein and *AID* mRNA expression:

We were able to obtain sufficient quality of RNA to perform qPCR in about 68% of our samples (59 out of 87).

When we divided IHC into 2 groups based on the median, we found a significantly higher ratio of *AID/GAPDH* mRNA in patients with high AID protein ($P = 0.001$) (Mann-Whitney U-test) (Figure 4.8).

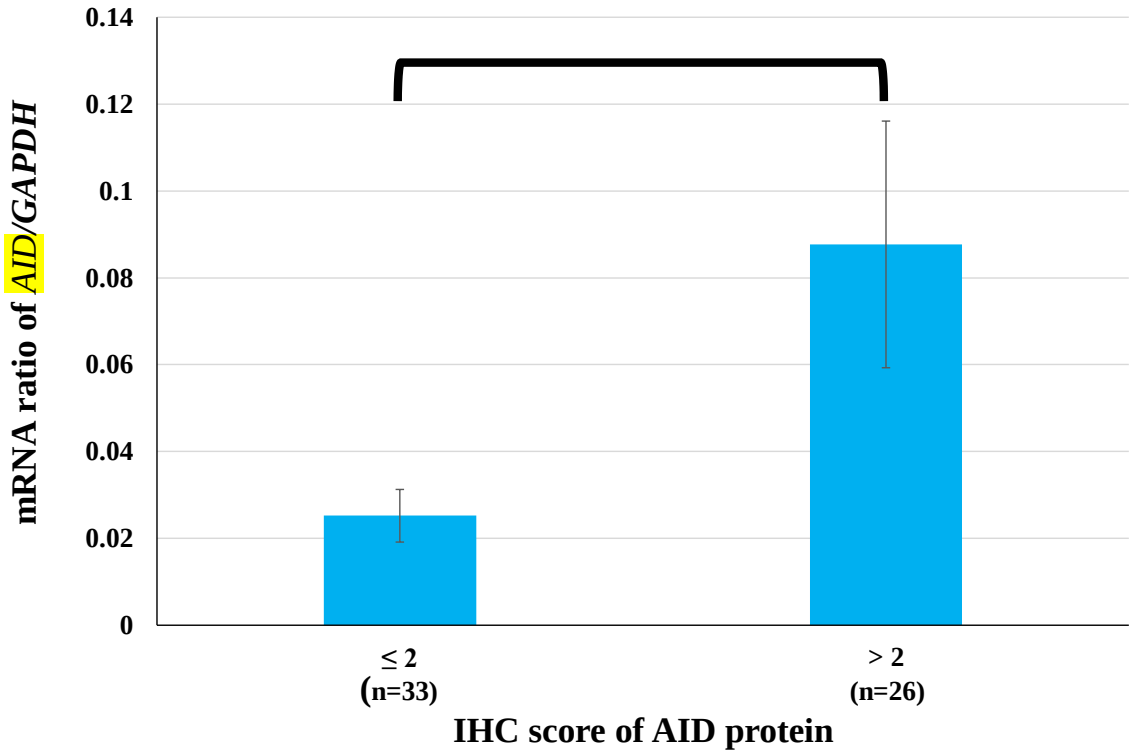

Figure 4.8: A positive significant correlation was obtained by comparing mRNA and protein level of AID. We devided AID into high and low groups, the ratio of *AID/GAPDH* in mRNA was significantly higher in pateints with high AID protein by IHC (*P*= 0.001, Mann-Whiteney U test).

4.4 Immunofluorescent expression of AID:

Total AID protein expression was determined using IHC. We wanted next to explore AID expression by immunofluorescence (IF). This was undertaken to identify exactly which cells of the GC were expressing AID, and to quantify the nuclear/cytoplasmic (N/C) ratio of AID.

4.4.1 Immunofluorescent co-expression of AID with CD3, CD20 and CD68:

In order to identify cells expressing AID in FL LNs, we undertook an immunofluorescent co-expression of AID protein with markers of common cells in GC.

We performed dual-immunostaining of AID with CD3 (T cells), CD20 (B-cells) and CD68 (macrophages). The method details are outlined in Chapter 2, Section 2.6.2.

When FL GC were stained with AID and CD3 immunofluorescence, we established that there was no co-expression of AID and CD3 on the same cells, i.e. on T-lymphocytes. This

was the same with CD68, i.e on macrophages. However, when we performed the dual staining of AID with CD20, we found that the colour became yellow in most GC cells, due to co-expression of AID (green) and CD20 (red), i.e. on the B-lymphocytes. We did this experiment in two cases expressing high AID and in hyperplastic tonsil tissue, and the results were the same (Figure 4.9).

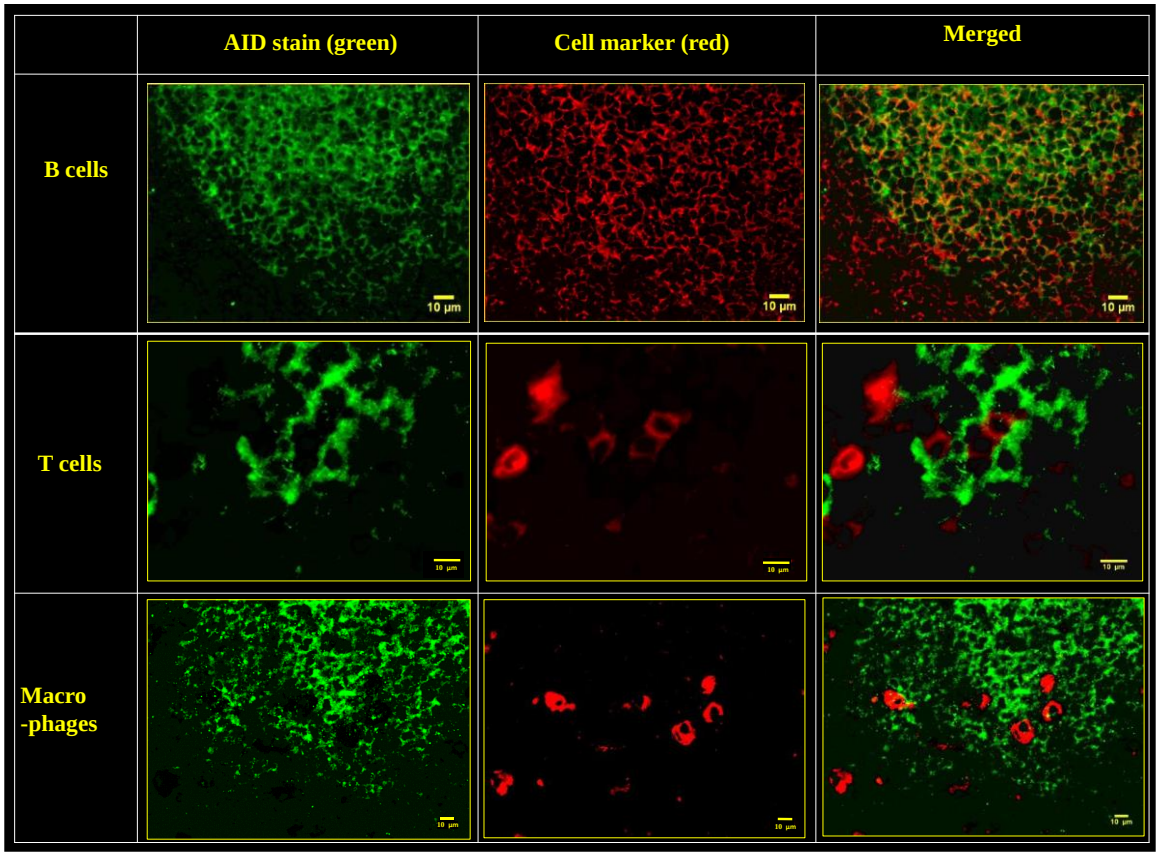

Figure 4.9: AID is expressed in B-cells of the GC, and not T cells or macrophages. CD20 is the B-cell marker used, CD3 is (T cell marker) and CD68 (Macrophage marker). Only CD20 is co-expressed with AID in the same cells giving an orange colour (upper right).

4.4.2 Subcellular localization of AID:

AID protein is synthesized and stored in a stable form in the cytoplasm, but shuttles to and from the nucleus where it functions (198). Hence, we wanted to determine the N/C ratio of AID in patients with high and low total AID expression. Our aim was to better explore functional AID. This is because total AID represents the stored and functional AID, while nuclear AID probably represents more the active AID.

To do this, three nuclear immunofluorescent stains (TO-PRO-3, 7-AAD and Hoechst 33258) were tested. In addition, different antibodies dilutions and protocols for double immunostaining with AID were used. The Hoechst 33258 antibody gave the best images, so it was employed in all IF slide staining. The method details are given in Chapter 2, Section 2.6.3.1 (Figure 4.10).

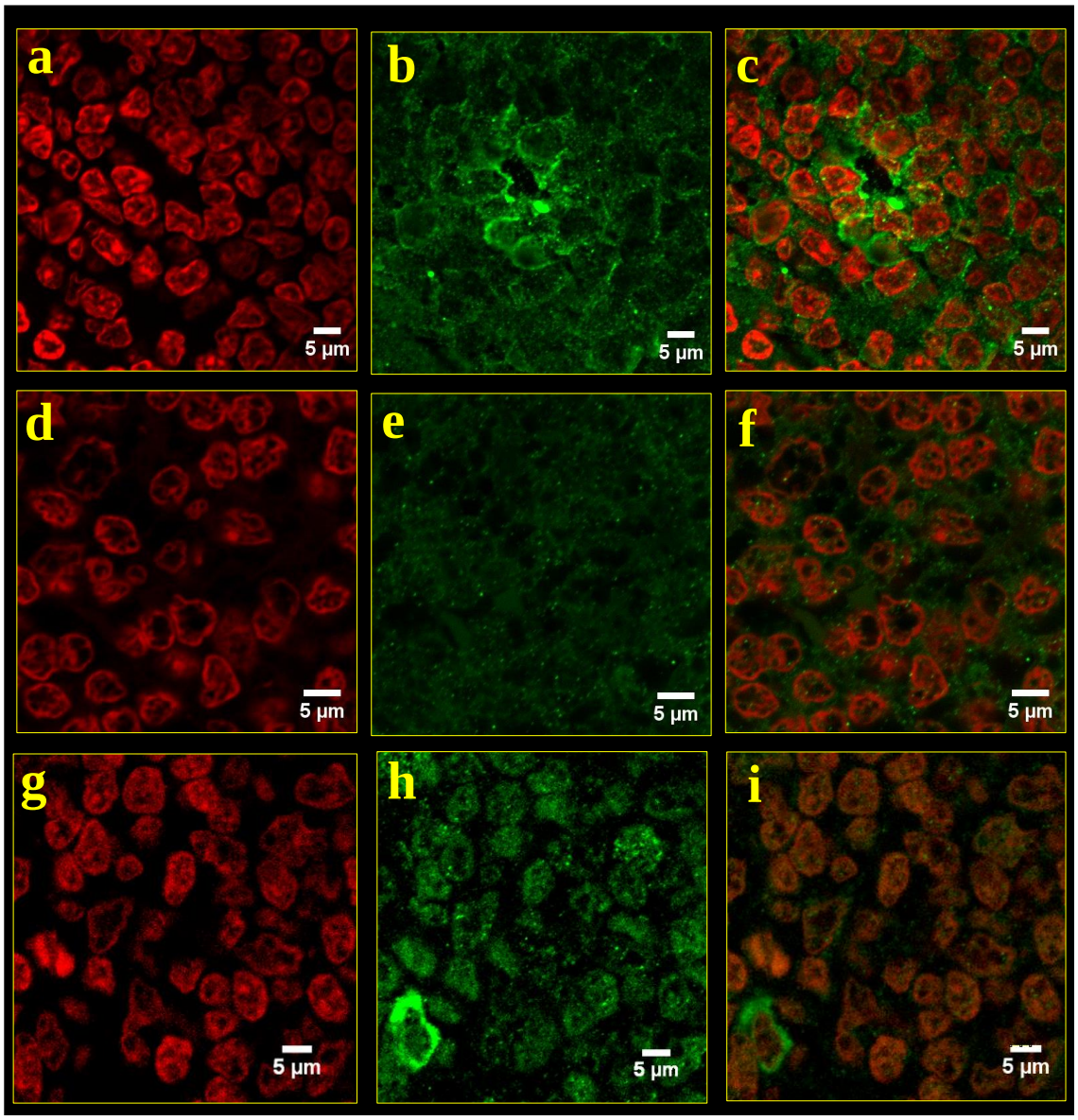

Figure 4.10: Three different nuclear stains were tried (red) in co-expression with AID (green) (a) Hoechst nuclear stain (red) (b) AID (green) (c) Hoechst and AID (d) TO-PRO-3 (red) (e) AID (green) (f) TO-PRO-3 and AID (g) 7-AAD (red) (h) AID (i) 7-AAD and AID.

Ten cases of FL with high AID protein by IHC and another ten cases of FL with low AID were selected. The slides were stained and visualised by confocal microscopy. After obtaining AID and nuclear co-expression IF images, a review of the literature was undertaken to identify a way of measuring N/C ratio of AID.

Published methods using Image j1x software were modified to analyse nuclear and cytoplasmic portions of AID (271, 313, 314).

Image j1x software was used to measure the ratio of N/C AID expression. This was done by removing the background IF stain. Then, an automatic selection of the nuclei was performed, followed by measurements of the nuclear AID expression. Cytoplasmic AID in the image was finally calculated by excluding the nuclear AID. The method details are supplied in Chapter 2, Section 2.6.3.2.

The positive control was a hyperactive tonsil, and the negative control was the IgG1 isotype staining of the hyperactive tonsil (Figure 4.11).

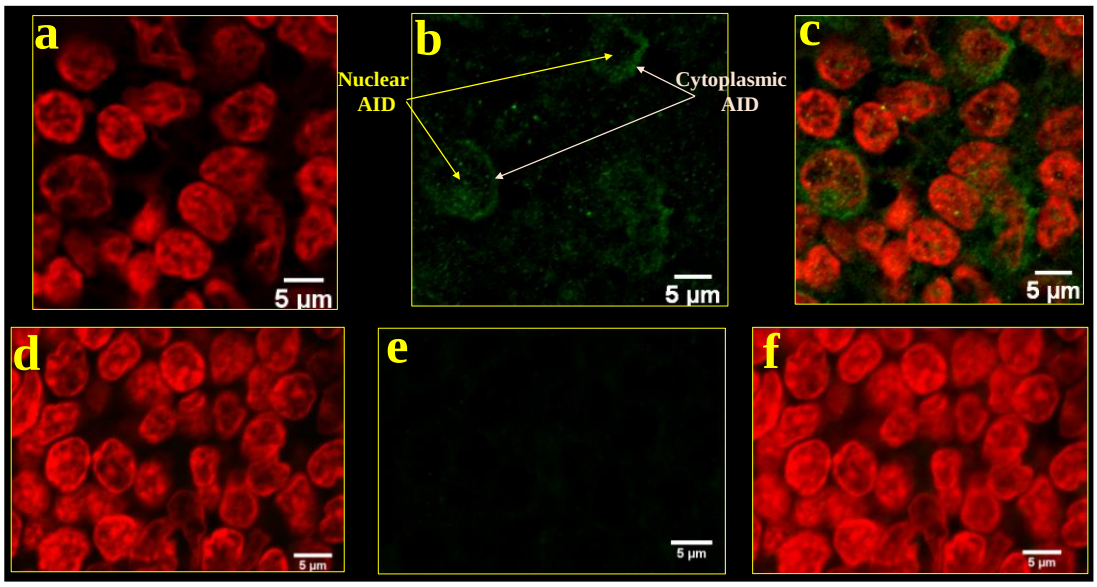

Figure 4.11: Positive and negative controls for subcellular AID IF. (a, b and c) are hyperactive tonsil positive controls (a) Hoechst nuclear stain (red) (b) is AID antibody stain (green) showing nuclear and cytoplasmic AID (c) is co-expression of both stains; (d, e and f) are IgG1 isotype control staining of hyperactive tonsil; (d) is the nuclear stain (red); (e) IgG1 stain (no AID) and (f) is a combination of both.

To identify the GCs under the IF microscope, a low magnification was initially used (10x). Five GCs were then selected: these were closest to 12, 3, 6 and 9 o'clock positions, in addition to one central GC within the LN. This is to take a representative GC from each region of LN because it was very difficult, time consuming and costly to take all or the majority of GCs (Figure 4.12).

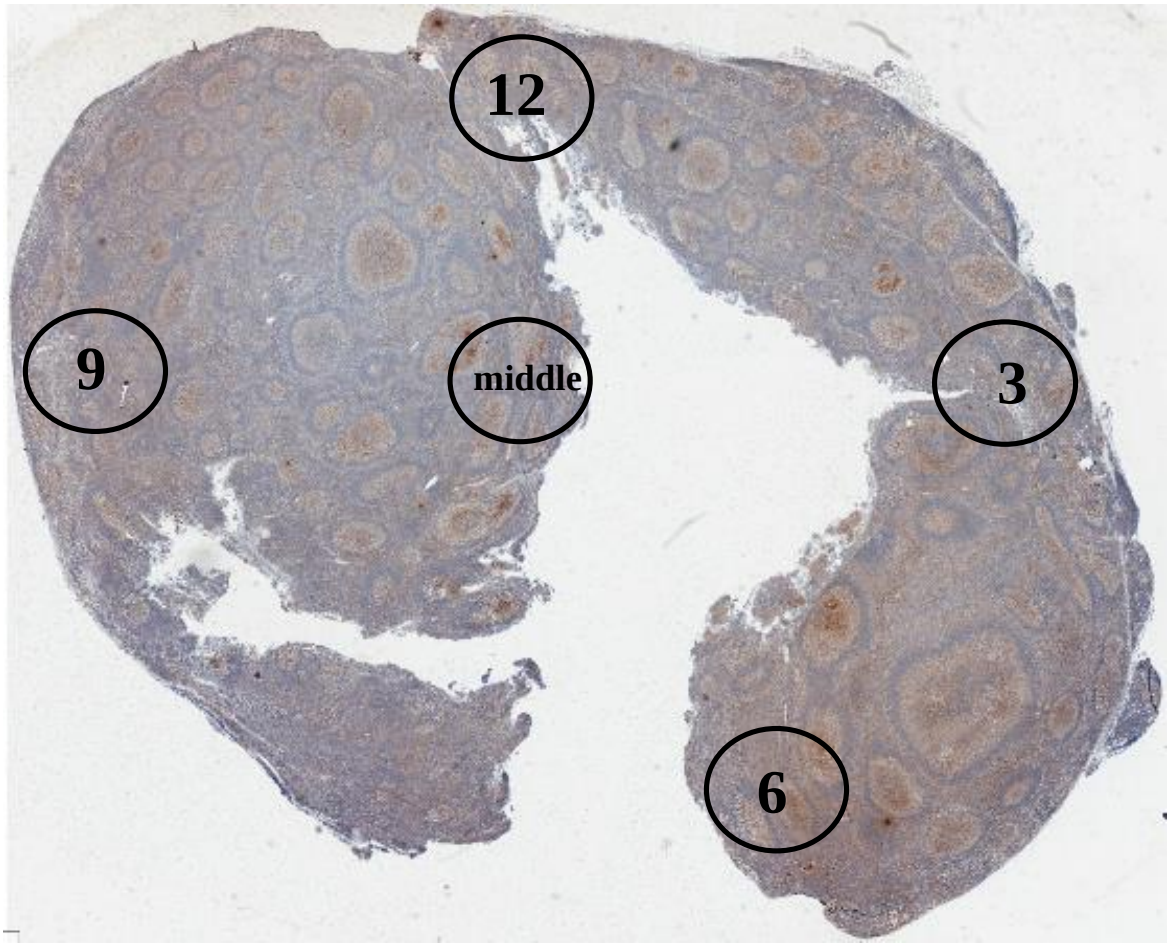

Figure 4.12: To study subcellular localization of AID in FL we needed to take a good representative areas of LN, so we took from five different GCs by choosing closest GC to areas 12, 3, 6 and 9 o'clock positions in addition to the middle of LN.

Within each GC selected, it was attempted to examine two immunofluorescent images if possible in the upper and lower parts of the GC, so that the number of images for each sample ranged between "6-15", with an average of nine.

To carry out the subcellular localization of AID in FL, ten cases were selected with high AID protein by IHC (≥3), the range was 3 - 9 and ten cases with low AID. The N/C ratio of AID in the 20 cases ranged between 46.91 and 70.69; the median of N/C AID ratio was 60.13%. The expression of N/C ratio of AID is shown in Figure 4.13.

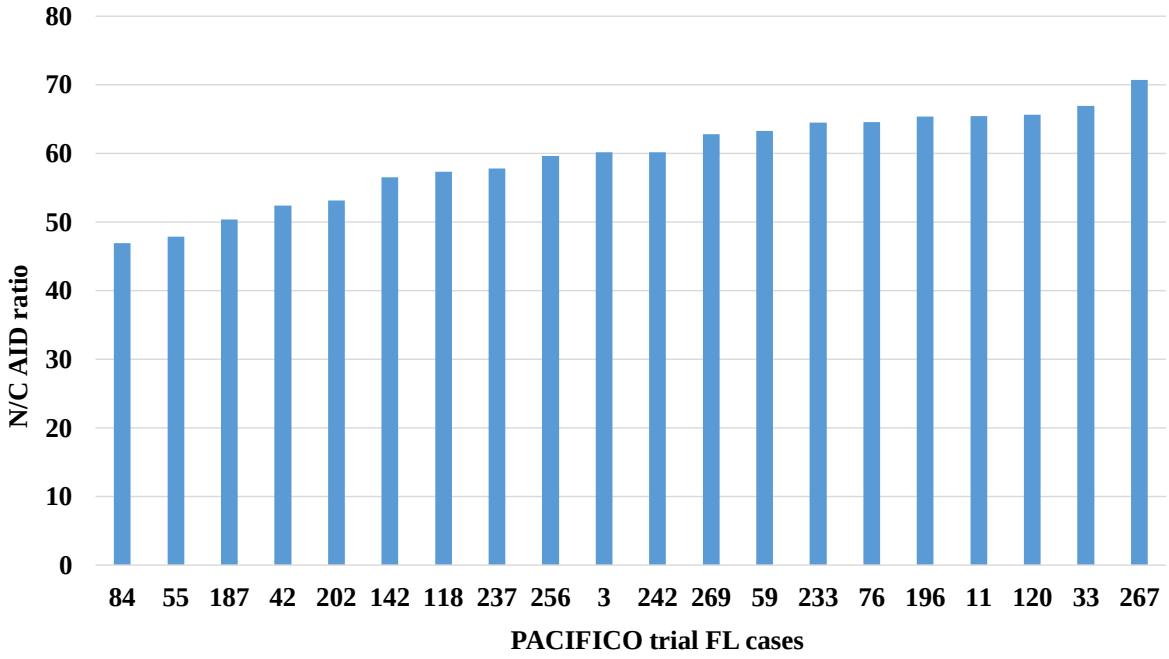

Figure 4.13: Expression of N/C ratio of AID protein in 20 cases of the PACIFICO trial FL. We divided patients into two groups, these were low and high N/C ratio of AID using the median as a cut-off point.

The *AID/GAPDH* mRNA expression data were available for 18/20 cases with subcellular AID staining. A significant positive correlation was found between the N/C ratio of AID and *AID/GAPDH* mRNA (r=0.513, *P*= 0.025) (Spearman's correlation).

The same observation was made when we examined the correlation between N/C ratio of AID compared to total AID protein. We found a significantly higher N/C ratio of AID in the group of high AID protein score measured by IHC (*P*= 0.023) (Mann-Whitney U test).

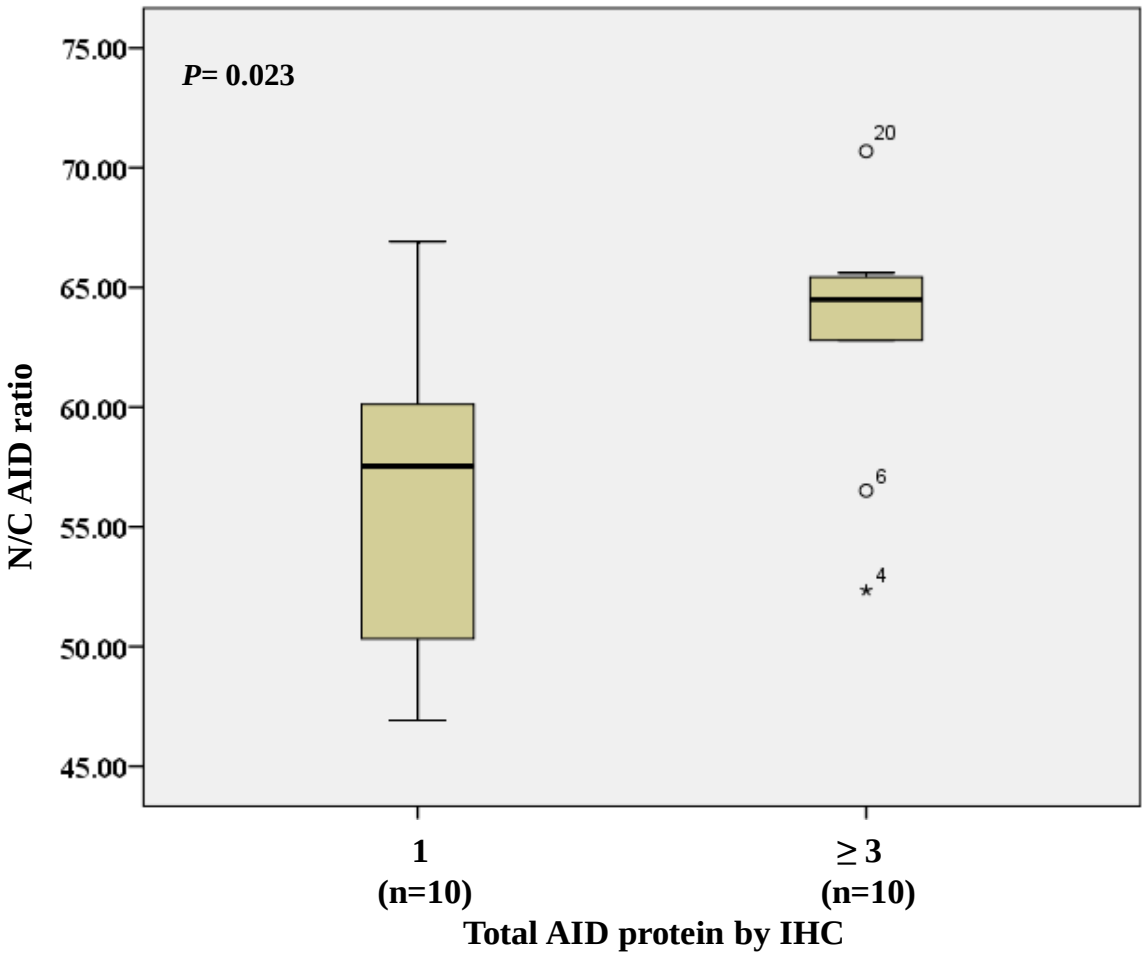

Figure 4.14: A significantly higher N/C AID ratio was found in cases with higher AID total protein by IHC (*P*= 0.023) (Mann-Whitney U test).

4.5: Correlation between clinical data at diagnosis and AID level:

We were able to obtain some clinical and diagnostic data from the LCTU. These data included the age, sex, number of enlarged LNs (patients were divided into two groups, five or less nodes in one group and six or more LNs is the second group), LDH levels, histological grade, anatomical stage and FLIPI prognostic score.

Details of correlations between *AID/GAPDH* mRNA expression, total AID protein and N/C ratio of AID and clinical data are found in Tables 4.2, 4.3 and 4.4, respectively.

Parameter	Range	*AID/GAPDH* mRNA Mean ± SD	*AID/GAPDH* mRNA	
			***P* value**	**Statistical test**
Age	57- 89	0.053 ± 0.103	0.786 (r= 0.036)	Spearman's correlation
Sex	Male 34 (57.62%)	0.07 ± 0.132	0.407	Mann-Whitney U test
	Female 25 (42.37%)	0.028 ± 0.027		
Number of LN	1 – 5 22 (37.29%)	0.041 ± 0.048	0.402	Mann-Whitney U test
	≥ 6 37 (62.71%)	0.059 ± 0.126		
LDH (U/L)	156 – 862	0.053 ± 0.103	0.108 (r= - 0.213)	Spearman's correlation
Histological grade	1 14 (23.73%)	0.038 ± 0.048	0.966	Kruskal - Wallis test
	2 29 (49.15%)	0.07 ± 0.1418		
	3 16 (27.12)	0.035 ± 0.032		
Ann Arber stage	2 8 (13.56%)	0.018 ± 0.011	0.429	Kruskal - Wallis test
	3 26 (44.06%)	0.065 ± 0.141		
	4 25 (42.37%)	0.051 ± 0.068		
FLIPI score	0 5 (8.47%)	0.03 ± 0.032	0.73	Kruskal - Wallis test
	1 13 (22.03%)	0.046 ± 0.059		
	2 41 (69.49)	0.057 ± 0.12		

Table 4.2: Correlation between *AID/GAPDH* mRNA expression and clinical data of patients.

Parameter	Range	AID total protein Mean ± SD	AID total protein by IHC 2 groups: ≤2 (55 patients (63.21%)) >2 (32 patients (36.78%))	
			***P* value**	**Statistical test**
Age	54 - 89	3.25 ± 2.73	0.888	Mann-Whitney U test
Sex	Male 46 (52.87%)	3 ± 2.76	0.35	Chi-square test
	Female 41 (47.12%)	3.63 ± 2.83		
Number of LN	1 – 5 32 (36.78%)	3.67 ± 2.92	0.49	Chi-square test
	≥ 6 55 (63.22%)	2.91 ± 2.66		
LDH (U/L)	109 – 862	3.25 ± 2.73	0.398	Mann-Whitney U test
Histological grade	1 23 (26.44 %)	3 ± 4	0.17	Chi-square test
	2 42 (48.28%)	2.91 ± 2.02		
	3 22 (25.29%)	4.2 ± 3.42		
FLIPI score	0 9 (10.34%)	1.33 ± 1.32	0.077	Fisher-Freeman-Halton exact test
	1 20 (22.99 %)	2.3 ± 2.36		
	2 58 (66.66)	2.64 ± 2.15		

Table 4.3: Relation between total AID protein expression and clinical data of the patients.

Parameter	Range	N/C ratio of AID Mean ± SD	N/C ratio of AID 2 groups' ≤ 60.13 (11 patients (55%)) > 60.13 (9 patients (45 %))	
			P value	Statistical test
Age	57 - 84	59.56 ± 6.69	0.28	Mann-Whitney U test
Sex	Male 12 (60%)	58.12 ± 6.48	0.169	Fisher Exact test
	Female 8 (40%)	61.71 ± 6.83		
Number of LN	1 – 5 9 (45%)	58.26 ± 7.09	1	Fisher Exact test
	≥ 6 11 (55%)	60.62 ± 6.48		
LDH (U/L)	109 – 830	59.56 ± 6.69	0.971	Mann-Whitney U test
Histological grade	1 4 (20%)	56.62 ± 7.25	0.6	Fisher Exact test
	2 11 (55%)	60.81 ± 7.05		
	3 5 (25%)	59.14 ± 5.98		
FLIPI score	0 3 (15%)	58.83 ± 6.31	0.533	Kruskal -Wallis test
	1 3 (15%)	56.29 ± 9.26		
	2 14 (70%)	60.41 ± 6.53		

Table 4.4: Correlation between N/C AID ratio expression and clinical data of PACIFICO patients.

No statistically significant correlations were found between AID expression at all levels and any of the clinical parameters. However, when we compared patients with N/C AID protein ratio against stage of disease, a tendency towards higher N/C ratio of AID in patients with advanced stages of disease, (*P*= 0.089) (Fisher Exact test). Note that stage IV was compared to stages II and III: this was because, in agreement with the fact that most patients with FL present advanced disease at time of diagnosis (315), and particularly at entrance to the

clinical trials, 18 of the 20 (90%) cases studied were at stages III (n = 9) or IV (n = 9), with the remaining two being at stage II (Table 4.5).

Also, there is a non-significant positive association between AID total protein expression and FLIPI score (*P*= 0.077)

Stage	**N/C ratio of AID**	
	≤ median*	**> median**
II or III	8	3
IV	2	7

Table 4.5: Comparison of frequency of stage IV FL between patients with high and low N/C ratio of AID protein by IF.* The median of N/C ratio of AID protein was 60.13.

4.6 Discussion:

Using gene mRNA expression by qPCR is one of the key technologies for biomarker discovery and clinical research to date (316). However, it is well known that intact RNA in fresh frozen tissues are preferred for qPCR mRNA levels. Unfortunately, these samples are usually of limited availability in research centres, in contrast to FFPE tissues that are stored for decades in most clinical centres and pathology departments (317).

The aim of this chapter was to measure AID expression in FL at the levels of mRNA, total protein by IHC and subcellular localization of AID protein. This can be used to correlate AID expression to clinical data, and to my knowledge, there is no publication which includes these three levels of AID expression in FL.

As discussed in Chapter 3 previously, RNA is degraded in FFPE tissues that are available for our research project. We wanted first to explore the effect of time storage of FFPE on RNA quality. It was found that the longer the time storage, the worse was the quality of

RNA. In addition, a significant difference was found in RNA quality between samples stored for less than or more than one year (Figure 4.1).

In contrast to our results, Bradley and colleagues found that the storage time of FFPE samples was not correlated with RNA quality for samples stored up to 3 years (318). This contradiction to our results was probably due to their small cohort of samples (18 compared to our 87 samples). Furthermore, they did not publish the storage time of each sample; probably the majority of their samples were in one of the categories of less than or more than one year and yielded no statistically significant correlation between RNA quality and sample age. In addition, this difference may also have resulted from other factors known to affect RNA quality such as fixation process or storage conditions (319, 320).

Subsequent to this, we tried to amplify large PCR products of *AID* and *GAPDH* genes (between 300-400 bp) from cDNA that was reversed transcribed from RNA of FFPE tissues of FL patients. Unfortunately, we did not obtain PCR products, mostly due to RNA degradation (Figure 4.2).

These findings were similar to many other studies: for example, Mathieson and colleagues used FFPE tissue samples for PCR amplification (321). They were able to amplify PCR in all samples when they used 65bp. However, by using primers to amplify 265bp products, only about half of samples were amplified (321). Belder and colleagues failed to amplify fragments longer than 150 bp from cDNA derived from FFPE tissues RNA (322). Similarly, Kong and co-workers found a poor PCR amplification efficiency and reliability for products around 200bp size (317).

For this reason, we tried next to identify a cut-off point of RNA degradation that produce a reliable results for *AID/GAPDH* ratio in qPCR reaction. To do this, we degraded a series of Raji cell line RNA by heating at 95°C for 3-30 mins (Figure 4.3 a), then synthesised cDNA and performed qPCR; after that, we compared qPCR ratios across all samples.

In comparison to our method of artificial RNA degradation, different procedures were used in these publications. Examples of these techniques include adding RNAse enzymes (323, 324) or RNA hydrolysis by incubating RNA with high pH alkaline buffer (325). Other mechanisms for RNA degradation are exposing nucleic acid to ultraviolet light for extended time periods (326) or heating RNA in a thermal block using different temperatures and times, and this method was used in our experiment (327).

However for FFPE tissues, the common method for estimating degradation of RNA was by comparing the samples stored for different time periods (288). So, we used the heating method in RNA extracted from cell lines because we did not have enough time to wait for our FFPE tissue samples to degrade.

To determine the most suitable primers for reverse transcription reactions, a comparison was carried out between using random primers and oligo-dt primers in RT qPCR reactions. It was found that random primers produced better results (Figure 4.3 b). This could be due to RNA degradation in FFPE tissue samples. As a result, the poly A tail of mRNA is degraded in many templates and oligo-dt primers cannot start reverse transcription. This is in contrast to random primers, which can start reverse transcription from any segment of the gene. So, although Oligo-dt primers are more specific and preferred in two-step PCR reactions, random primers are the best option for degraded RNA in FFPE tissues and produce a large amount of cDNA (328, 329). Some researchers prefer using a combination of oligo-dt primers and random primers in the same reverse transcription reaction to obtain the benefits of both types. Examples include Olsen et al. (330), Yao et al. (331), Vuong et al. (332) and Medimegh et al. (333).

In our study, it was shown that short PCR products were more suitable and produce reliable and consistent results for qPCR performed on nucleic acid of FFPE (Figure 4.3 b). Recent

studies also used short amplicons around 100bp for amplifying PCR from FFPE tissues (334, 335). This is in agreement with Zeka et al. (336), and Kocjan et al. (337).

In order to determine which EC gene should be targeted for our qPCR expression of *AID*, we undertook qPCR expression of three most commonly used genes in FL. It was found that *GAPDH* is the most suitable (Figure 4.4). In agreement with us, Walter and colleagues found that *GAPDH* and not *HPRT1* was the best reference gene in FFPE tissue gene expression experiments (338). In addition, *GAPDH* is still being used by the most well-known research groups involved in FL as a reference gene for qPCR reactions in FL (339-343). However, in contrast, B de Kok and colleagues found that *HPRT* was a better EC gene than *GAPDH* (344). Further, lso Lossos and co-workers noted that *PRKG1* and/or *TBP* were also better than *GAPDH* as an EC gene in lymphoma (345). This is probably because of the use of intact RNA in those two studies (frozen tissues) compared to our degraded tissues using *TBP* and *HPRT1*, which appeared to be more degraded and found in later cycles of qPCR.

After we performed AID expression at both mRNA and protein level, we found that there was a significant positive correlation between mRNA and protein levels of AID in FL (Figure 4.8). These results were in agreement with Rush et al. (346), who also noted a positive correlation between AID mRNA and protein level in GC B-cells.

In this study, it was observed that only B-cells but neither T cells nor macrophages of GC expressed AID (Figure 4.9). This was in agreement with Cattoretti and colleagues. These authors discovered that AID was expressed in B-cells and not T cells or FDCs. In addition, they found that AID positive B-cells were in contact with T cells or FDCs (347). However, in the presence of chronic inflammation, FDCs start to form follicular structures that are similar to GC. It has been noted that these cells may express AID, and this was observed in FDCs of rheumatoid arthritis synovium (348) and Sjogren's Syndrome (349). In T cells, although less than 1% of T cells express AID at any given time point, Qin and co-workers

found that in mouse LNs, a small subset of T cells may express AID, and these cells increase with age; the authors suggest that these cells are connected to long-term immunity (350).

It was observed in this study that AID was present in the nucleus and those cells with high AID mRNA and high AID total protein have high N/C AID ratio (Figure 4.14).

Excess AID in the nucleus is mutagenic and different molecular mechanisms to prevent nuclear damage by the additional AID were elucidated. For instance, Methot and colleagues found that AID was kept out of the nucleus by two mechanisms: (1) AID cytoplasmic sequestration by eukaryotic elongation factor 1 α (eEF1A), and (2) CRM1-mediated AID nuclear export (351). Hu and colleagues discovered the C-terminal region of AID includes a nuclear-export signal which is responsible for AID's nuclear cytoplasmic (232).

In addition, (ubiquitin-dependent) proteosomal degradation pathways also prevent AID nuclear damage (228, 352, 353).

We used Image j1x software to measure N/C ratio of AID in FL FFPE tissue sections; however, we faced some difficulties with these samples. This is because of the nature of cells in FL that are tightly packed close to each other, with the nucleus appearing to extend into the cytoplasm of the same cells and even into that of nearby cells. Consequently, this made automatic nuclear selection problematic.

When AID expression was compared to all of the clinical data at entry to clinical trial, it was found that there was no direct correlation between AID expression and clinical outcome of patients. This was with the exception of a tendency towards higher N/C ratio of AID in patients with more advanced stage of disease. This could be explained by the following factors:

1- Most patients reached advanced disease. In stage II disease, we had 11 FL patients only compared to 40 patients in stage III, and 36 patients in stage IV. So, the patient

cohort was skewed with most patients having advanced disease. This may have led to no clear changes between AID expression and clinical data.

2- Although the cohort of patients was low in N/C AID ratio (20 patients), we still observed tendency toward higher AID N/C protein ratio with advanced stage of disease. This may indicate that the function of AID correlates better with disease progression than AID mRNA or protein expression. However, a larger cohort of patients is required to validate our results and provide a robust conclusion.

In agreement with our findings, Willenbrock and colleagues found that total AID protein expression by IHC was not related to prognosis and stage in FL (354).

Further support for our results comes from very recent data from Kawamura and co-workers. They found that in contrast to diffuse large B-cell lymphomas, there was no correlation between AID expression at RNA and protein levels with disease stage and prognosis in FL (237).

In summary, in this chapter an optimization of methods for AID expression at the levels of mRNA, protein and N/C AID expression was carried out. Further, it was established that AID expression at the mRNA, total and N/C protein levels were correlated to each other. A tendency towards higher N/C ratio of AID and stage of disease was observed. Consequently, this study was further extended to explore the correlation between AID expression and the number of ongoing mutations in FL: this work is to be found in the next chapter. This process of ongoing mutation of the *IGHV* gene is of relevance as it is associated with formation of subclones, relapse and progression of disease in addition to resistance to chemotherapy.

Chapter 5

Correlation between AID expression and ongoing *IGHV* mutation in FL

5.1 Introduction:

This thesis is focused on studying the role of AID in pathology and pathogenesis of FL. FL FFPE tissue samples were used to identify AID expression, which was then correlated with various biological and clinical parameters of patients.

In this chapter, the relationship between AID expression and ongoing *IGHV* mutation in FL is explored, with the relationship being connected to disease progression and transformation (78).

5.1.1 Background:

FL is predominantly a monoclonal proliferation of neoplastic B-cells. However, the *IGHV* gene which encodes an important component of the BCR, has been found to carry significant varying SHM in the tumour cells within a single tumour. In addition, the SHM process is ongoing, so that each malignant subclone may have a slightly different *IGHV* gene sequence (355).

Continuing AID activation in the SHM throughout tumour progression was suggested to be a possible mechanism for ongoing mutations occurring in the GC. Continuous evolutionary changes at genomic and epigenomic levels in FL were suggested to be associated with AID activities producing continuous GC reaction and mutations in *IGHV* and other genes. (88, 356).

Ongoing AID activity is also associated with the acquisition and loss of genetic material, intraclonal heterogeneity and drug resistance in FL (357, 358).

The mutational status of the *IGHV* gene could be used as an indication for disease prognosis in FL (359). Because of the importance of ongoing mutations, researchers have started to investigate ongoing mutation in *IGHV* and other genes in FL using recently developed techniques, including next generation sequencing (360).

5.1.2: Aims:

To explore the relationship between ongoing mutations of *IGHV* that represent AID function and clinical parameters of FL.

5.1.3: Hypothesis:

We hypothesised that high ongoing *IGHV* mutations represents AID function and this ongoing mutational process is associated with advanced disease and clinical outcome in FL.

5.2: Approaches to study ongoing *IGHV* mutation in FL:

To test our hypothesis, the ongoing mutation status was studied in two cohorts of patients: the first group were those with high levels of AID expression (at both AID protein assessed by IHC and mRNA by qPCR), whilst the second group were cases with low levels of AID expression.

5.3 Materials and methods:

5.3.1 Patients and samples:

In this Chapter, 20 of 87 FL samples from patients enrolled in the PACIFICO clinical trial were selected, in which the AID protein expression in the FFPE tissue samples had been successfully assessed with IHC (see Chapter 4, Section 4.2.2).

Based on material availability and to exclude those with intermediate levels, ten samples at the low-level end and ten at the high-level end of the profile of AID protein expression were

selected for study in this chapter. The group with low levels of AID protein had IHC scores of one (with the exception of one case which had an AID IHC score of 2. This case was then removed from analysis due to failure of cloning). The group with high level of AID protein scored between 4 and 9.

Selection of samples in each cohort was undertaken according to the following:

1-Possibility of having high or low both AID total protein and mRNA in the same sample.

2-The presence of sufficient DNA materials.

3-Better quality of DNA.

Details of samples, AID expression levels both at mRNA by qPCR and protein levels by IHC, amount and quality of DNA of samples used can be found in Table 5.1.

Patient ID	Total cellular AID protein (IHC score)	*AID* mRNA (by qPCR)	DNA available (µg)	Length of DNA (by PCR)		
				200 bp	300 bp	400 bp
142	9	0.0978	78.4	+++	+++	+++
233	9	0.02	9.75	+++	+	-
120	6	0.2052	3	+	+	-
269	6	0.2946	2.668	+++	+	+
200	6	0.0037	25.2	+++	++	+
182	4	0.0001	7.8	++++	++++	++
11	4	0.0948	5.5	++	+	+
267	4	0.094	46.8	+++	++	-
48	4	0.1015	15.25	+++	++	+
42	4	0.0421	2.2	++	+	-
187	1	0.0065	66.8	++++	++++	++++
202	1	0.004	20	++++	++++	+
33	1	0.0057	11.25	++	++	+
237	1	0.0012	10.2	+++	+++	+
221	1	0.0098	24.15	+++	+++	++
118	1	0.0102	14.25	+++	+++	++
56	1	0.0091	10.25	++	+	+
46	1		4.8	+++	+++	++
234	1	0.0336	35.42	++	++	+
105	2	0.0387	0.5	+++	++	+

Table 5.1: Samples used for analysis of correlation between AID expression and ongoing mutation, there are ten cases with high AID total protein by IHC and ten cases with low AID; cases 234 and 105 failed to produce *IGHV* and *IGK* clones and were removed from the analysis. DNA ranges from negative for no gel bands to 4+ for strong bands.

5.3.1.1 Clinical data of cases:

Clinical data of the PACIFICO patients at time of diagnosis were stored at the LCTU and became available by request.

5.3.2 Technical approaches to study SHM in *IGH* gene:

DNA was first extracted from FFPE tissues (as discussed in Chapter 2, Section 2.4.1), and then *IGHV* PCR amplification was performed. Next, PCR products were run on agarose gel and PCR products were extracted from gel bands.

The PCR products were cloned in *E. coli* bacteria and clones were grown on agar. Fifteen bacterial clones were selected and sent for sequencing to Genewiz Company using T7P primer (Figure 5.1). Cloning method details are outlined in Chapter 2, Section 2.4.5.3.

In the case of *IGHV* of non-clonal B-cells being sequenced, more colonies were screened until 15 clones were successfully sequenced for each sample.

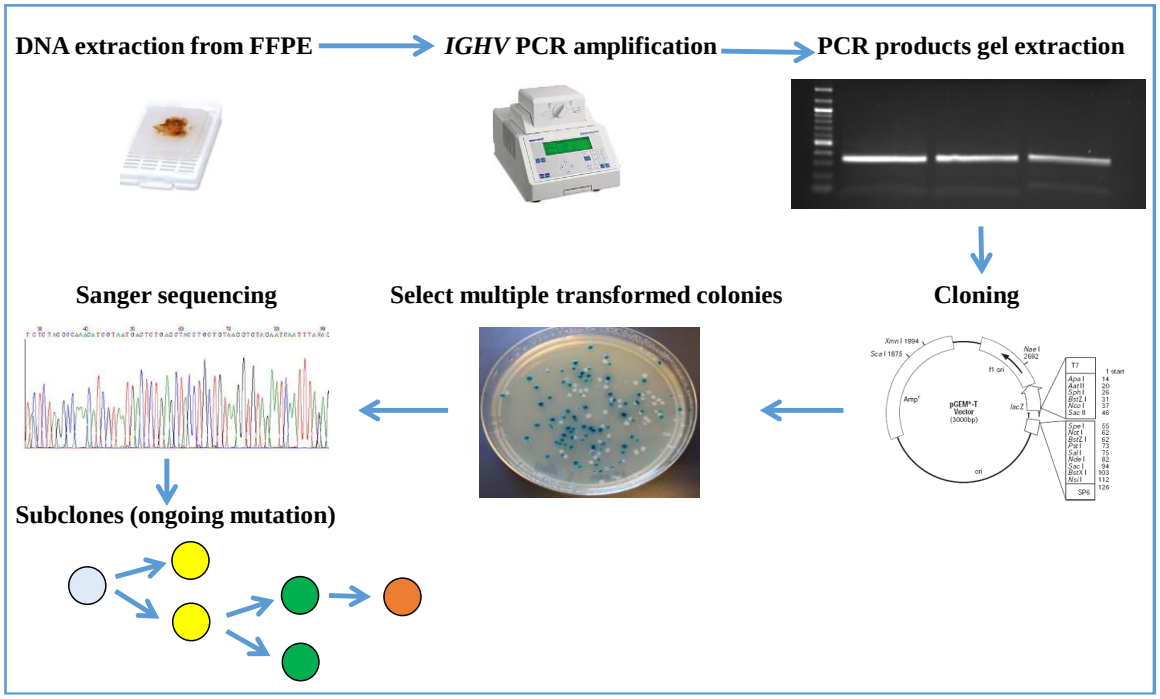

Figure 5.1: Platform of work to identify ongoing *IGHV* mutation in FL, starting with DNA extraction from FFPE to pedigree clone screening.

5.3.3: PCR amplification of clonal *IGH* in the FFPE samples using BIOMED-2 protocol:

Van Dongen and colleagues designed primers and optimized PCR conditions for detecting clonality in immunoglobulin and T-cell receptor genes in lymphoproliferative diseases. In combination with a junction primer, primers designed to anneal FR1, FR2 and FR3 produce a 310 – 350bp, 250 – 295bp and 100 – 170bp PCR band, respectively. This strategy is particularly useful for amplification of *IGHV* fragments from degraded DNA, such as those prepared from FFPE samples. Details are given in Chapter 2, Section 2.4.5.1. In the case of no amplification by FR1 primers, FR2 or FR3 primers can be used. *IGK* primers were also used to amplify *IGK* genes in two-tube PCR in our study using a Raji cell line (positive control). *IGK* amplification was performed to be used in samples failing to produce *IGHV* amplification. The expected sizes of *IGK* in tube A is 120-160bp and in tube B *IGK* is 270-300bp (Figure 5.2) (11).

The first step was to put one set of primers for each of the 6 FR1 families into an individual PCR tube (six PCR tubes in total) to determine which FR1 family was amplified.

Then, the positive *IGHV* gene family was reamplified in a 400µl PCR reaction performed in eight tubes to obtain enough PCR product for subsequent experiments. The PCR products were then extracted from agarose gel before being used for subcloning to detect ongoing mutation in *IGHV* (Method provided in Section 2.4.5).

5.3.3.1: Optimizations for BIOMED-2 protocol for *IGHV* amplification:

With the BIOMED-2 protocol, we obtained *IGHV* bands. However, the bands were faint, so we tried the following optimizations to improve PCR bands from BIOMED-2 reactions:

1-$MgCl_2$ was tried in different concentrations between 1.5mM to 2.5 mM (3-5 µl).

2-Annealing temperature was increased gradually from 55-60°C.

3-PCR cycle number was increased gradually from 35 to 40.

4. dNTP (10mM) was increased from 2 to 3 µl.

5-DNA input was increased from 50ng to 100 ng.

6- Thermo-Start DNA Polymerase (5U/µL) was increased from 0.2 to 0.5 µl.

5.3.3.2: Investigating the faint bands resulted from BIOMED-2 PCR amplification:

In half of the samples (ten out of 20 cases), we obtained a prominent PCR band in one *IGHV* family in addition to a faint band in another family.

So, we decided to investigate the origin of the faint band. To do this, we chose one of the samples (Patient 221) with an additional faint band, then performed six tubes of PCR for each family to identify the positive *IGHV* gene family. Then, we cloned the PCR product from that family in *E. coli*, and then compared the *IGHV* sequences from the predominant and faint PCR bands.

5.3.4 SHM analysis:

5.3.4.1 Database used:

The international ImMunoGeneTics information systemR (IMGTR) has been the most commonly used reference software for analysis of the *IGHV* gene for over 28 years now (267).

The IMGTR consists of seven databases and 17 online tools to study the complicated diversity of immunoglobulin and T cell receptor major histocompatibility genes (267).

IMGT/V-QUEST program version: 3.4.2 was used for studying the *IGHV* gene; this software is able to align and analyse our *IGHV* gene in comparison to its closest germline counterpart.

PCR sequences between the two ends of the vector were used to run in IMGT/V-QUEST using the (http://www.imgt.org/IMGT_vquest/share/textes/) website. After inserting our cloned PCR sequence, we chose human species and *IG* gene type, then ran the software. In this software, we also obtained the percentage of *IGHV* mutation.

5.3.5: Strategy to identify ongoing mutations:

To identify the degree of clonal evolution as reflected by ongoing *IGHV* mutation, the first step was to identify clonality of these multiple *IGHV* sequences. We considered all cloned sequences with the same VDJ junction segment being derived from a B-cell malignant clone, as discussed in various publications (361-363).

Details of *IGHV* mutational analysis were discussed in Section 2.4.5.5.

To study ongoing mutation, we used the following parameters:

1-Number of subclones with ongoing mutation activity identified among the 15 *IGHV* sequences for each case using phylogenetic tree analysis software (364).

2-Ongoing mutation was analysed using a commonly used calculation in publications. It was determined by dividing the cumulative number of mutations occurring in a single clone and in more than one clone (but not in all of the 15 clones) with the number of artificial mutations resulting from DNA polymerase-induced error in PCR (358, 365).

Error rate mutation was calculated with the expected DNA Polymerase-induced rate of error, which is 2.28×10^{-5} per base per PCR cycle for the Thermo-Start DNA Polymerase used in this study (366-368).

5.3.6 Identifying clone branching:

Performing a phylogenetic tree and its branching can be used as a read out of ongoing mutation (364).

To identify the phylogenetic tree of the resultant clones, the NCBI nucleotide blast website was used for aligning nucleotide sequences of the 15 clones (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=MegaBlast&PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch&BLAST_SPEC=blast2seq&DATABASE=n/a&QUERY=&SUBJECTS=). We then stored the aligned sequences in FASTA form.

Phylogeny.fr software was then used for creating the phylogenetic tree.

This was done by using software on the appropriate website (http://www.phylogeny.fr/); then we clicked on Phylogeny analysis one click mode which is the simplest mode, loaded our aligned nucleotides sequences in FASTA form and clicked submit to obtain our phylogenetic tree (369).

5.3.7 AID hotspot mutation in *IGHV* gene analysis:

AID is known to target specific sequences of nucleotides in the *IGHV* gene; these sequences are called AID hotspots. AID targets WA/TW sequences, in addition to WRC/GYW sequences, W is either (A or T), R is (A or G) and Y is (T or C) (209, 210).

WRC/GYW sequences produced by the action of AID while WA/TW mutation is produced by error prone DNA polymerase as a result of DNA repair following AID action (207, 370). WRC/GYW mutations are called canonical AID induced mutations, while WA/TW were named non-canonical AID mutations (208).

5.3.8 Statistical analysis:

Statistical analysis was undertaken by using IBM SPSS statistics v22. Spearman's correlation, Mann-Whitney U test, Kruskal-Wallis test and Fisher exact test were used to examine the correlation between AID hotspots, percentage (%) of SHM and ongoing mutation with patient clinical parameters. Statistical significance was defined as $P< 0.05$.

5.4 Results:

5.4.1 Optimizations of SHM identification in *IGHV* gene:

5.4.1.1 Successful *IGHV* gene amplification using BIOMED-2 PCR reaction:

Full *IGHV* gene amplification failed using primers from leader to constant regions in four FL FFPE samples. The positive control (Raji cell line) successfully amplified the gene, and each experiment of the four FFPE samples was repeated three times.

For shorter segments of the *IGHV* gene from FR1 to VDJ junction, BIOMED-2 PCR amplification was successful; however, PCR bands were not strong enough in some samples. The weak PCR bands appeared on using $Mgcl_2$ concentration of 1.5mM and 35 cycles of the PCR reaction.

5.4.1.2: Optimizations for BIOMED-2 PCR reaction:

Optimizations to improve PCR efficiency were carried out using the Raji cell line that was also used as a positive control for *IGHV* amplification.

We found that increasing $Mgcl_2$ concentration from 1.5mM to 2.5 mM and increasing the PCR cycles from 35 to 38 produced the strongest specific bands. We therefore used these optimized BIOMED-2 PCR conditions to amplify the *IGHV* gene in our study cohort of 20 cases with FL in this chapter (Figure 5.2).

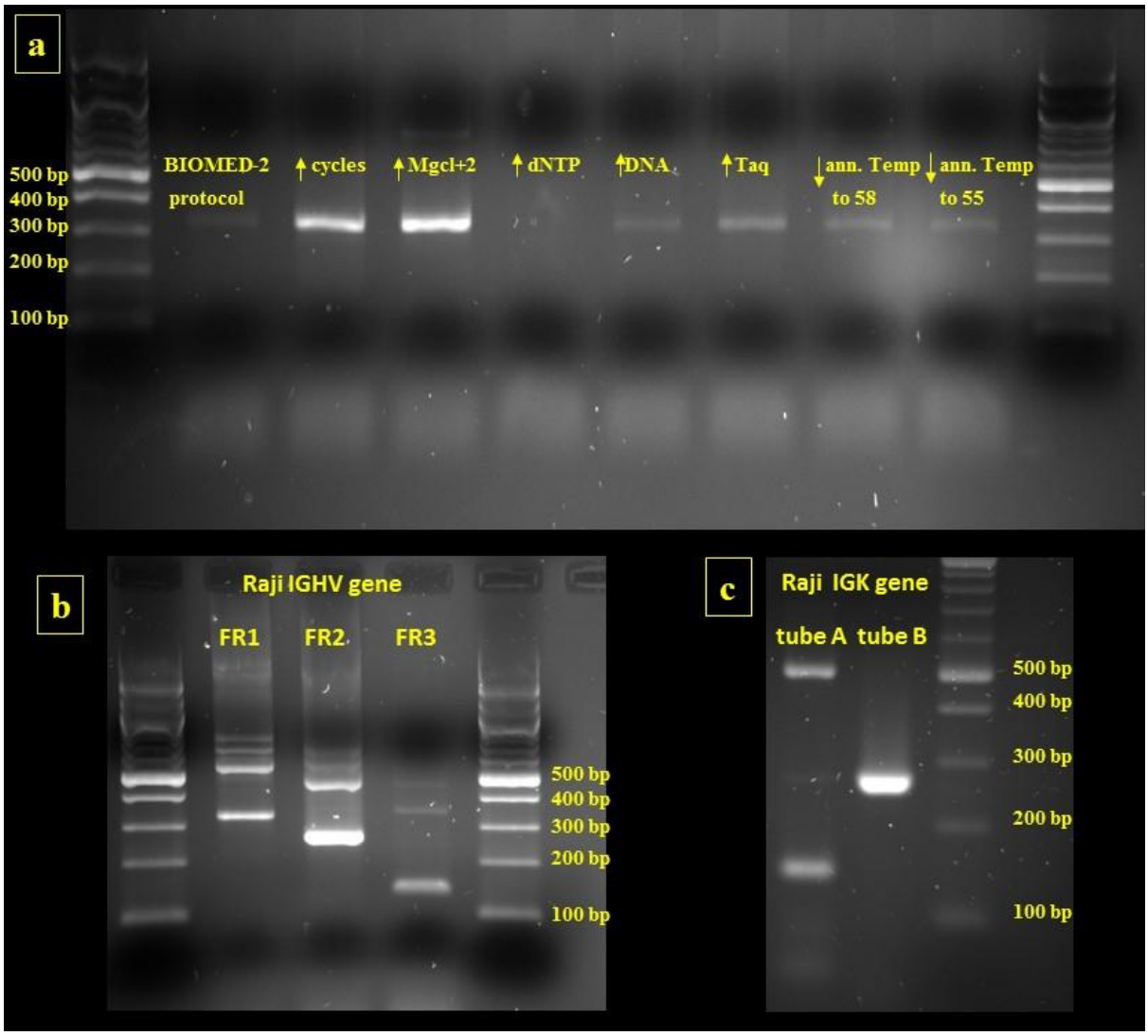

Figure 5.2: Optimization of PCR conditions harmonising BIOMED-2 protocol to amplify *IGHV* in Raji cell line (329 bp in size). (a) increasing the PCR cycles from 35 to 38 or $Mgcl_2$ concentration from 1.5mM to 2.5 mM, improved intensity of PCR products; (b) Successful PCR amplification of Raji *IGHV* gene using modified BIOMED-2 protocol from FR1 (product = 329 bp), FR2 (≈250 bp) or FR3 (≈150 bp) to VDJ junction segment; (c) Raji IGK gene amplification in two-tube PCR (products in tubes A and B were ≈150 and ≈280 bp, respectively). All PCR gel bands sizes were similar to that expected from primer design and to other publications (371-373).

5.4.2 Investigating the extra faint bands generated in BIOMED-2 PCR reactions for amplifying *IGHV* gene of PACIFICO trial samples:

B-cell monoclonality is a feature of FL (374); however, in half of the cases we found a faint band in addition to the main PCR band of monoclonal B-cells. The prominent and faint bands in one of the FL patients (Sample 221) had been cloned and sequenced. It was found that the prominent PCR band produced clones with the same VDJ junction segment, although some

of them had different mutations within the *IGHV* gene fragment compared to others. On the other hand, on cloning the faint band, 11 clones was successfully sequenced. Nine out of the 11 clones derived from the faint PCR band had different *IGHV* gene and different VDJ junction segments from each other. Another two clones had the same *IGHV* gene and the same VDJ segment, but one of them was 100% similar to the wild type gene and the other one had an additional single mutation. We concluded that the faint bands represented the background *IGHV* expression in non-clonal normal B-cells, and that they had wild type *IGHV* genes with different junction segments for individual clones (Figure 5.3).

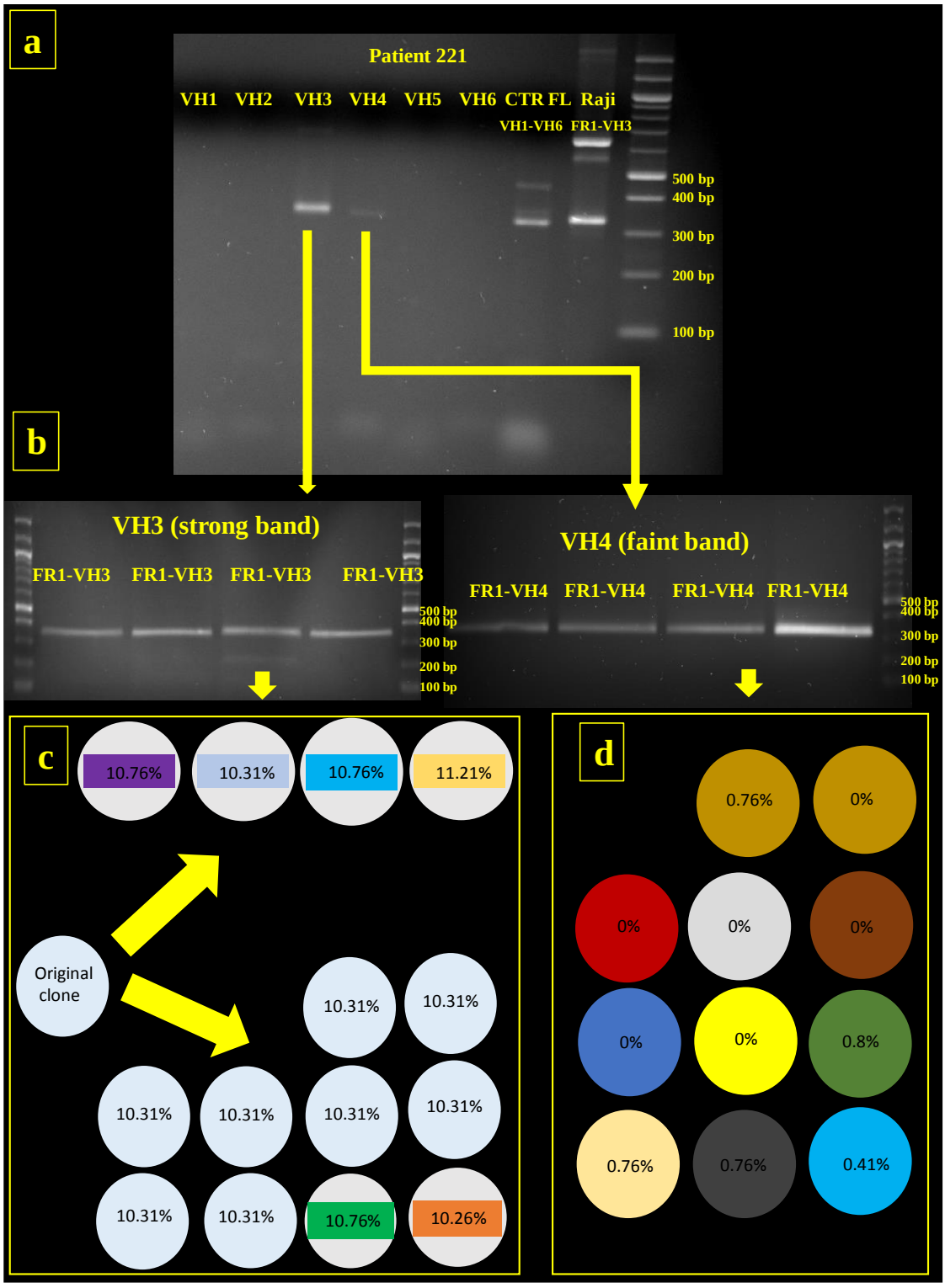

Figure 5.3: PCR faint bands in *IGHV* gene amplification are background normal B-cells. (a) Main and faint PCR bands of BIOMED-2 protocol to amplify DNA of *IGHV* gene in sample (221) PCR products on agarose gel (b) reamplifiable main PCR products in family 3 and then weak PCR in family 4 for cloning (c) ongoing mutation detected in different subclones of the same malignant clone as defined using criteria described in Section 5.3.5 above (d) background B-cells identified by different VDJ junction segment between different clones and low degree of mutation compared to FL samples. Circles of the same colour have similar VDJ segment of the *IGHV* gene.

5.4.3 Performing BIOMED-2 reaction, cloning and analysing FL cohort:

Biomed-2 PCR reaction and cloning was successful in 10 patients with high and eight out of 10 patients with low AID protein. Unfortunately, in the low AID cohort, two cases (patients 234 and 105) failed to produce clones, so they were excluded from further analysis.

After cloning, we sent clones for sequencing and ensured that we had 15 clone sequences for each patient. The cloned PCR sequences in all clones of all patients were clean, and sequences were analysed using IMGT/V-QUEST software as discussed in Section 5.3.4.1**.** An example of the clean PCR sequences and IMGT software use are shown in Figure 5.4.

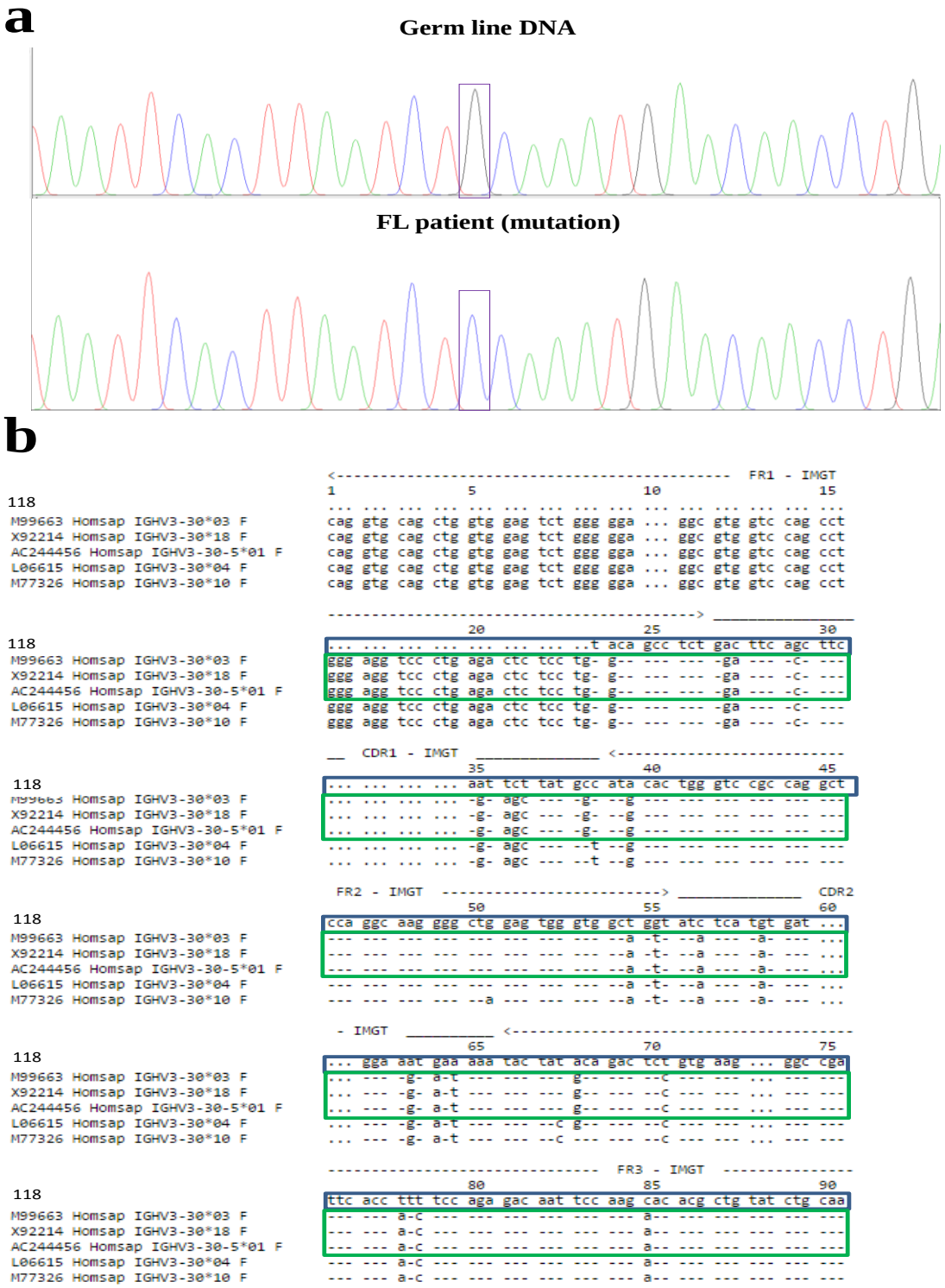

Figure 5.4: (a) A representative Sanger sequence of a part of an *IGHV* gene cloned shows a clean sequence and a clear point mutation; (b) An example of *IGHV* sequence analysis in one of the clones of sample 48, the nucleotide sequence in the purple box is the cloned sequence compared to the closest germline sequence in the green box (Homsap *IGHV*3-30*03 F, or Homsap *IGHV*3-30*18 F or Homsap *IGHV*3-30-5*01 F). The sequence identity to the germline counterpart is 89.24% (199/223 nt). Identical nucleotides are represented by hyphens.

5.4.4 Percentage of SHM:

In all 18 samples, the percentage of mutation ranged between 2.89 and 19.22, with an average of 14.86 (Figure 5.5).

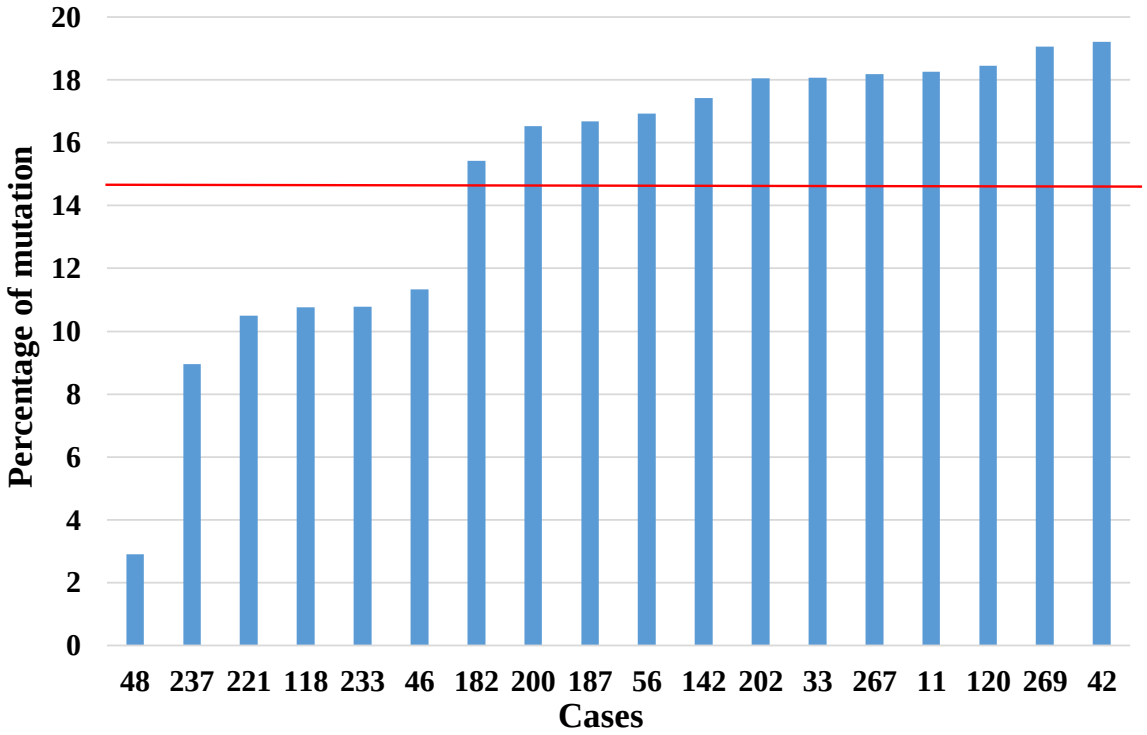

Figure 5.5: Percentage of *IGHV* mutations in the 18 patients studied. Red line represents the mean of mutations (14.86%).

5.4.5 Divergence of ongoing *IGHV* mutation:

In addition to percentage mutations in *IGHV*, there is a large difference in the intraclonal variation among the cohort of patients. For example, all of the 15 cloned *IGHV* in patient (118) shared the same sequence, while patient (269) had 12 different subclone sequences due to ongoing mutation. The number of different subclones derived from the original clone ranged from (0 to 12) with a median of eight. An example of ongoing mutation in one of these patients is shown as intraclonal variation in Figure 5.6.

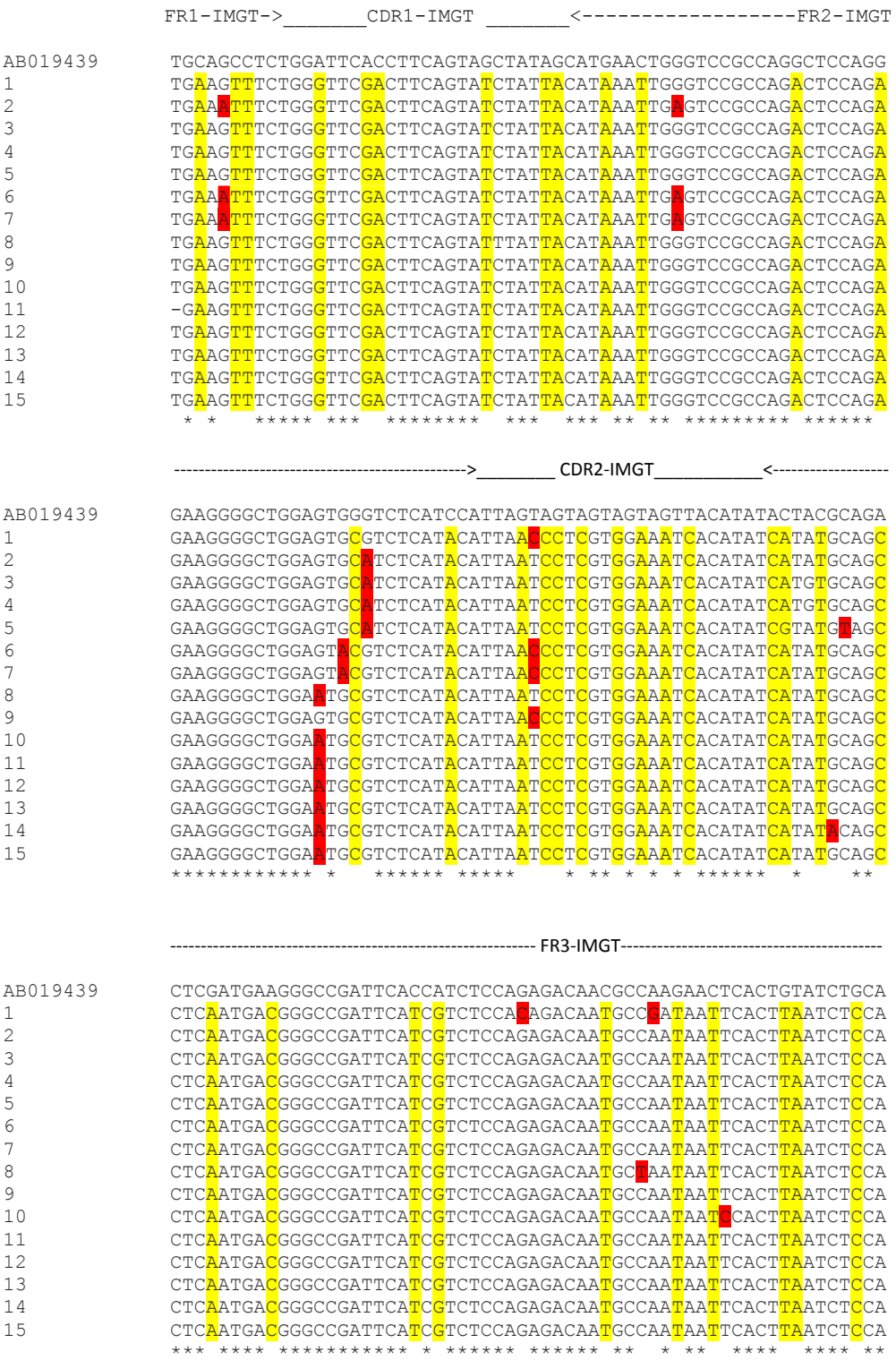

Figure 5.6: An example of ongoing mutations shown in *IGHV* sequences of different subclones originating from one original clone. The yellow lines represent mutations shared by all clones that are different from the closest *IGHV* germline in the first raw data, while the red coloured nucleotides represent single nucleotide or partial mutation across different clones. The sequence on top is that of germline counterpart.

5.4.6 Number of subclonal generations:

A phylogenetic tree of *IGHV* clones for each sample clone was produced using the phylogene.fr software. The range of generations derived from parental clones was between zero and four with a median of two in the 18 patients.

An example of phylogenetic tree generation is illustrated in Figure 5.7.

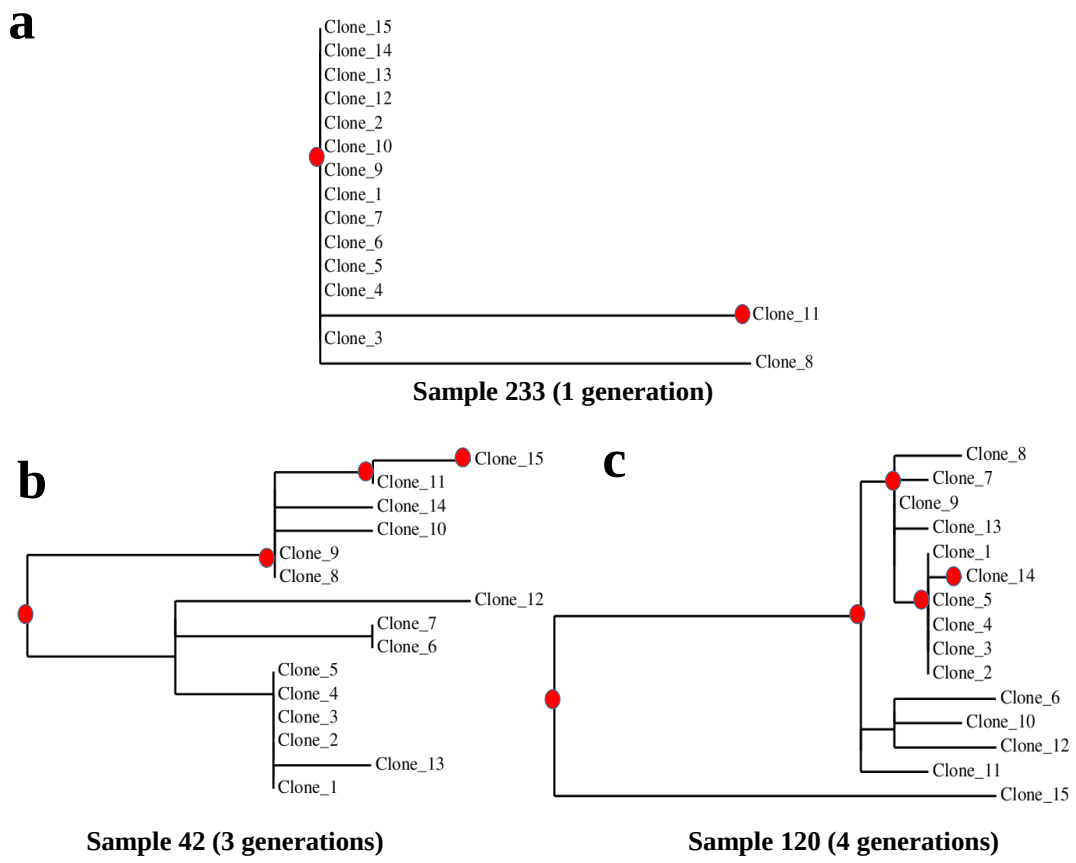

Figure 5.7: Phylogenetic tree of three different patients produced by phylogeny.fr software. This was used to estimate generation of FL subclones based on ongoing mutation in *IGHV* for all of the 18 patients studied.

5.4.7: Types of mutations in *IGHV*:

We performed an analysis of different types of mutations in ten cases of high AID and eight cases with low AID.

Replacement mutations were more common than silent mutations in both CDR and FR regions; however, replacement to silent mutations were more in CDR than FR regions. (R/S mutations =7/1.5 median in CDR compared to 13/7 median in FR region) (Table 5.2).

All mutations in all of 15 clones of 18 patients were point mutations with the exception of the following:

1- One nucleotide insertion in one of the clones in sample 142.
2- Three nucleotides inserted and one deleted in one of sample 200 clones.
3- Three nucleotides inserted in one of the clones of sample 120.
4- Deleted one nucleotide in one of the clones of sample 182.
5- All clones with insertion or deletion mutations were from samples with high AID protein by IHC.

Ongoing mutations were identified in 17 out of 18 cases in our cohort. The ongoing mutations were measured using techniques described earlier in Section 5.3.5 of this chapter. The percentage of *IGHV* mutation range was (2.89 - 19.22). The number of subclones derived from 15 clones analysed in each patient was between (0 and 12). The ongoing analysis measured by excluding PCR error rate from unique and partially shared mutations was between (0 and 16.37) (Table 5.3).

Pt ID	Close *IGHV* Germline	No. of mut.	CDR1 and CDR2					FR1,FR2 and FR3				
			R	S	TR	TV	C to T	R	S	TR	TV	C to T
142	*IGHV*3-7*02 F, or -7*03 F	39	4	1	2	6	1	13	10	17	14	7
233	*IGHV*3-30*03 F, or -30*18 F or 30-5*01 F	25	8	2	6	7	0	8	2	5	6	1
120	*IGHV*3-21*01 F or 21*02 F	42	9	1	6	8	0	17	7	16	12	4
269	*IGHV*3-21*01 F, or *02 F	42	9	1	7	7	0	20	6	15	12	7
200	*IGHV*3-23*01 F, or23*04 F or 23D*01 F or -23D*02 F	38	8	1	8	5	0	12	8	20	4	9
182	Homsap *IGHV*3-48*02 F	37	5	2	10	2	3	14	7	13	11	4
11	Homsap *IGHV*3-69-1*02 P	38	7	2	7	5	0	17	7	14	12	8
267	*IGHV*3-48*03 F	43	8	2	5	10	0	17	7	17	11	8
48	Homsap *IGHV*3-23*01 F, or*04 F or *IGHV*3-23D*01 F or *02 F	7	2	0	0	2	0	3	0	0	4	0
42	*IGHV*3-21*01 F, or 21*02 F or 48*03 F	42	9	1	7	7	0	17	7.5	15.5	12	8
	Average	35.3	6.9	1.3	5.8	5.9	0.4	13.8	6.15	13.25	9.8	5.6
187	*IGHV*3-23*01 F, or23*04 F or 23D*01 F or23D*02 F	38	7	2	9	6	2	12	7	19	4	10
202	*IGHV*3-69-1*02 P	41	8	2	7	8	1	17	6	14	11	8
33	Homsap *IGHV*3-21*03 F	42	9	1	7	7	0	18	7	16	12	8
237	Homsap *IGHV*3-23*01 F, or23*04 F or 23D*01 F or 23D*02 F	21	5	2	5	5	1	6.5	3	5	5	3
221	*IGHV*3-21*01 F, or *02 F	23	7	1	6	5	1	6	3	7.5	5	2
118	*IGHV*3-30*03 F, or *18 F or5*01 F	26	7	2	6	7	1	8	2	6	7	2
56	*IGHV*3-23*01 F, or-23*04 F or 23D*01 F	38	6	3	9	5	1	13	7	19	5	9
46	*IGHV*3-23*05 F	26	6	1	3	7	1	9	3	9	7	0
	Average	31.88	6.88	1.75	6.5	6.25	1	11.19	4.75	11.94	7	5.25

Table 5.2: Data derived from ongoing mutation analysis in 10 samples with high total AID protein by IHC (red colour) and 8 samples with low AID protein (blue). Pt, patient; mut, mutation; R, replacement mutation; S, silent mutation; TR, Transition mutation; TV, Transversion mutation.

Sample ID	AID protein (IHC score)	*AID/ GAPDH* mRNA (qPCR)	Average num. of nucleotide	* PCR error	Unique mut.s	Partially Shared mut.s	**Total PCR error	*** Ongoing mutation	No of subclones	Average % of mut.
142	9	0.0978	226.6	0.013	9	9	2.95	6.11	10	17.42
233	9	0.02	233	0.013	2	1	3.03	0.99	3	10.79
120	6	0.2051	225.36	0.013	15	34	2.93	16.73	10	18.45
269	6	0.2946	223	0.013	9	39	2.90	16.56	12	19.06
200	6	0.0037	235.86	0.013	9	7	3.07	5.22	12	16.529
182	4	0.0001	230.86	0.013	3	2	3.00	1.67	4	15.42
11	4	0.0948	221.64	0.013	2	6	2.88	2.78	4	18.266
267	4	0.094	230.4	0.013	7	10	3.00	5.68	12	18.19
48	4	0.1015	232.86	0.013	11	0	3.03	3.63	8	2.89
42	4	0.0421	223	0.013	7	14	2.90	7.24	10	19.22
187	1	0.0065	225.43	0.013	8	15	2.93	7.85	9	16.69
202	1	0.004	218.67	0.013	4	3	2.84	2.46	8	18.06
33	1	0.0057	229.67	0.013	6	5	2.99	3.68	5	18.08
237	1	0.0012	225.07	0.013	4	1	2.93	1.71	5	8.96
221	1	0.0097	223.79	0.013	7	4	2.91	3.78	8	10.49
118	1	0.0102	223	0.013	0	0	2.90	0.00	0	10.76
56	1	0.0091	226.36	0.013	1	0	2.94	0.34	2	16.93
46	1		226.14	0.013	7	4	2.94	3.74	8	11.33

Table 5.3: AID expression and ongoing mutation in FL, *Total PCR error= PCR enzyme error (2.28×10^{-5}) X 38 PCR cycles X 15 clones per sample; mut, mutation; nuc, nucleotides.** Total PCR error= PCR error X Average number of total nucleotides; Ongoing mutation= (Unique mutations +Partially shared mutations)/Total PCR error.

5.4.8: AID-induced *IGHV* mutations:

The types of mutations induced by AID are either canonical or non-canonical, as discussed in Section 5.3.7. In order to accurately identify the percentage of AID induced mutations among the total nucleotides in the 15 clones of each patient, we used the following formula:

= (Number of AID induced mutations/Total number of nucleotides in the 15 clones) X 100

The percentage of AID induced mutation among the total nucleotides ranged between (1.38 % and 6.67 %). The mean for all cases was 4.22.

Canonical AID mutation ranged between (0.92% - 4.51%), while non canonical mutations ranged between (0 - 3.08%).

5.4.9 Correlation of AID expression with different parameters tested in this chapter:

To explore the relationship between AID expression and ongoing mutations in *IGHV,* statistical analyses were conducted in following subsections.

5.4.9.1 Relation between *AID/GAPDH* mRNA expression and ongoing mutation:

The correlation between AID expression levels and ongoing mutation in FL was explored. A suggestive of a trend if a larger sample size could be analysed in future between higher *AID/GAPDH* mRNA and bigger % of *IGHV* mutation (r= 0.419, *P*= 0.094) (Spearman's correlation). A very similar positive trend (very close to statistical significance) was determined between *AID/GAPDH* mRNA and ongoing mutation (r= 0.483, *P*= 0.05) (Spearman's correlation).

No correlation was identified between *AID/GAPDH* mRNA and % of AID induced mutations (canonical and non-canonical) among total nucleotides (Table 5.4).

Parameter	*AID/GAPDH* mRNA Range	*AID/GAPDH* mRNA Mean ± SD	*AID/GAPDH* mRNA	
			***P* value**	**Statistical test**
Percentage of *IGHV* mutation	0.00012 - 0.295	0.0589 ± 0.0834	0.094 (r= 0.419)	Spearman's correlation
Number of subclones	(≤8) 0.0001- 0.102	0.0257 ± 0.0387	0.133	Mann-Whitney U test
	(>8) 0.0037 - 0.295	0.1063 ± 0.1081		
Ongoing mutation	0.00012 - 0.295	0.0589 ± 0.0834	0.05 (r= 0.483)	Spearman's correlation
% of AID induced mutations among total nucleotides	0.00012 - 0.295	0.0589 ± 0.0834	0.225 (r= 0.386)	Spearman's correlation

Table 5.4: Correlations between *AID/GAPDH* mRNA and % of *IGHV* mutation, AID induced mutations and measures of ongoing mutation**.**

Furthermore, out of 15 clones in each patient, the cases were divided by the median number of subclones (8) into two groups (≤8 and >8). We also found a positive but non-significant trend of higher number of subclones in patients with higher *AID/GAPDH* mRNA (*P*= 0.133) (Mann-Whitney U test) (Figure 5.8).

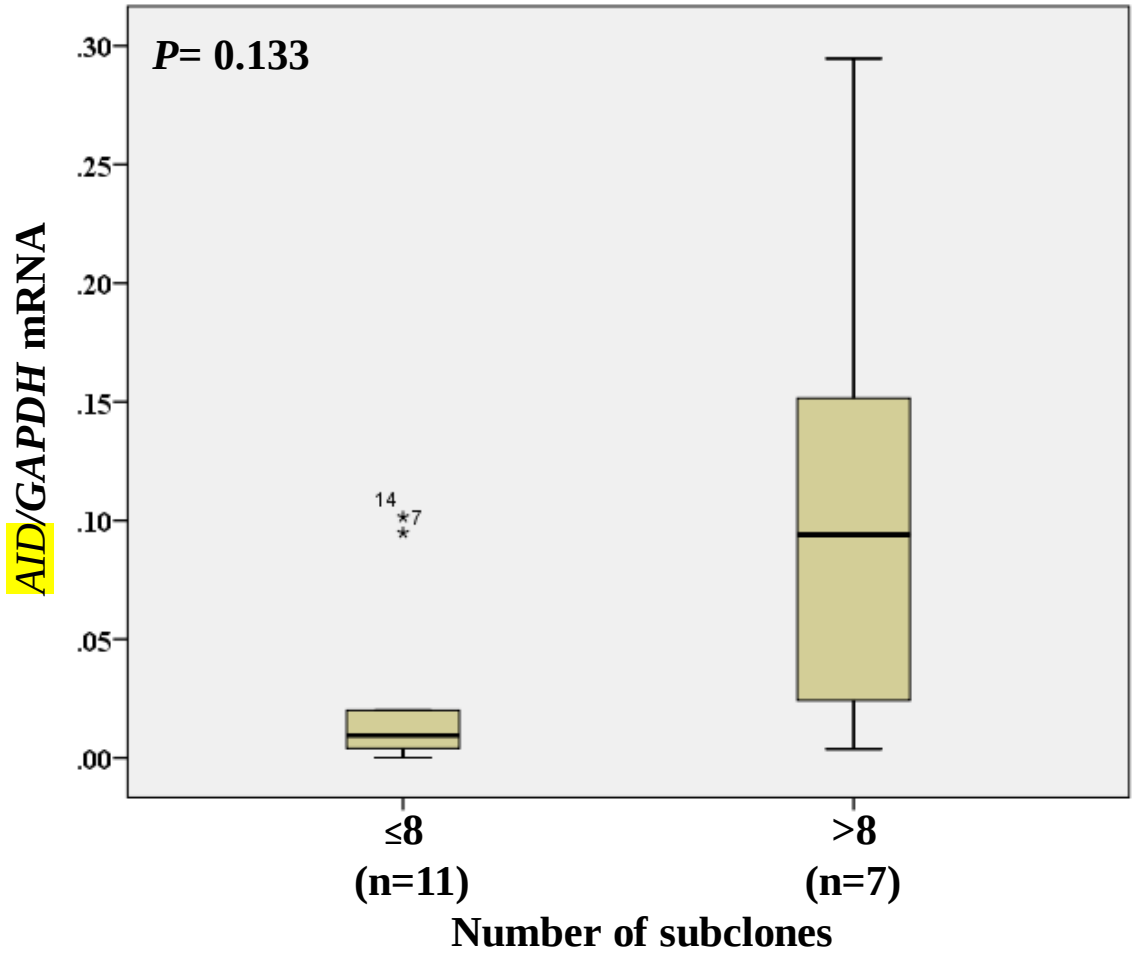

Figure 5.8: Examining the relationship between *AID/GAPDH* mRNA and number of subclones (*P*= 0.133) (Mann-Whitney U test). Subclones were divided into high and low by the median of subclones (8) in all cases.

5.4.9.2 Correlation between total AID protein expression and ongoing mutation:

Ongoing mutations were compared to AID total protein. Two groups of AID protein expression were used. Eight samples with low AID protein (IHC score =1) and ten cases with high AID (≥4).

When the % of *IGHV* mutations between high and low AID IHC groups was compared, the high level group showed a larger % of mutations, although the difference was not statistically significant (*P*= 0.146) (Mann-Whitney U test) (Figure 5.9).

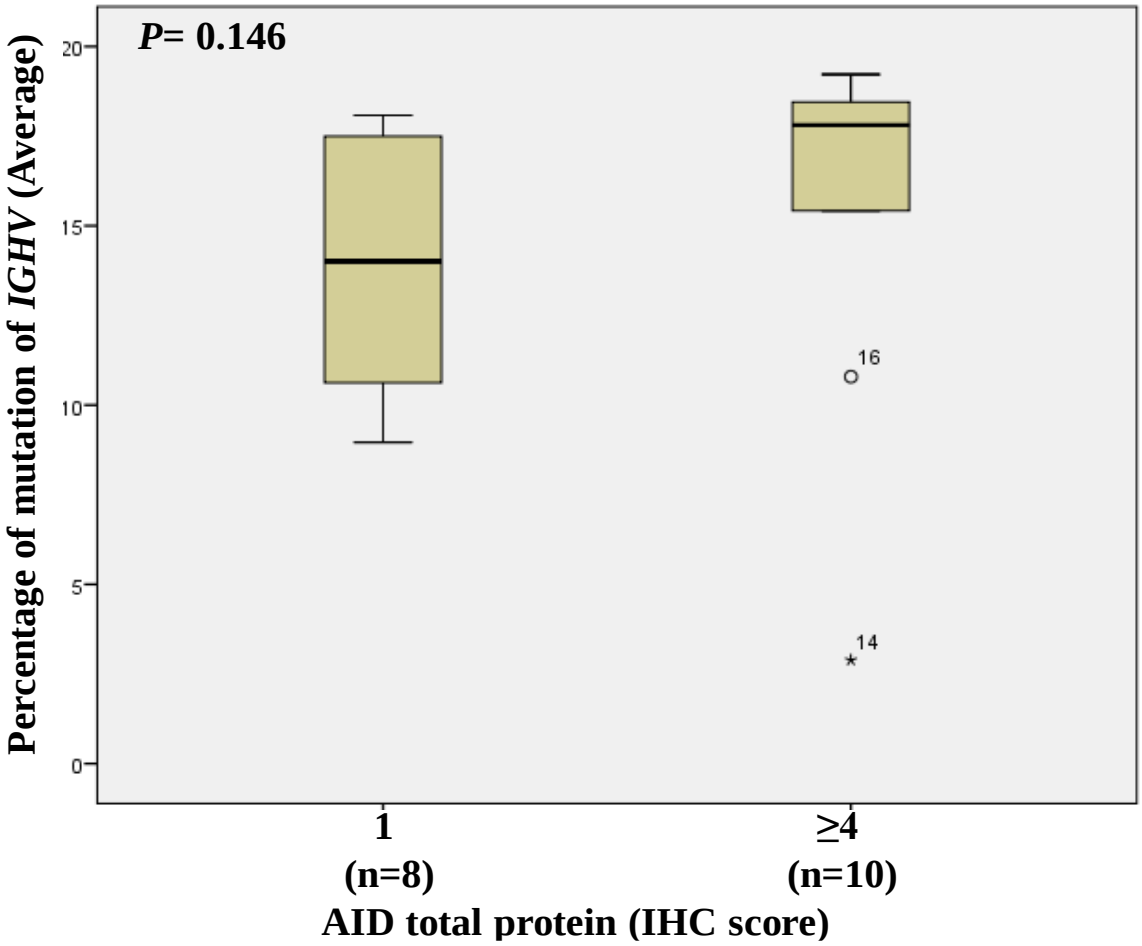

Figure 5.9: Examining the relationship between AID total protein IHC level and percentage of mutation (*P*= 0.146) (Mann-Whitney U test).

Also, when the subclones were divided into two groups by the median (8), a tendency towards higher AID IHC score and more number of subclones was observed (*P*= 0.066) (Fisher exact test) (Table 5.5). Thus, six of the ten (60%) patients with a high level of AID protein had > 8 mutated *IGHV* subclones, while only one of the eight (13%) cases with a low level had a similar number of the subclones.

Number of subclones (Median is 8)	Total AID IHC score	
	1	**≥ 4**
≤ 8	7	4
>8	1	6

Table 5.5: Comparison of number of subclones between patients with high and low total AID protein by IHC (*P*= 0.066) (Fisher exact test).

The correlation between AID IHC and ongoing mutations was explored, (*P*= 0.173) (Mann-Whitney U test) (Figure 5.10).

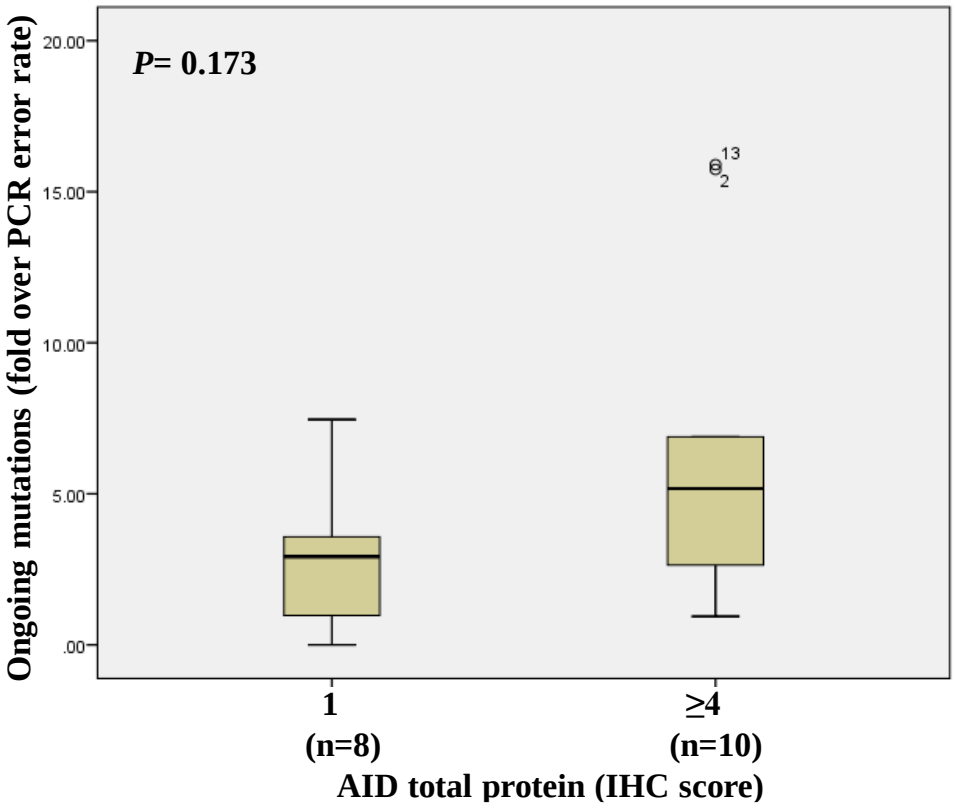

Figure 5.10: Examining the relationship of ongoing mutation in the group of patients with low and high total AID protein by IHC score (*P*= 0.173) (Mann-Whitney U test).

When the percentages of AID induced mutations among total nucleotides was explored in the two groups of low and high AID, no difference in the % of AID-induced mutations was noted between the two groups (*P*= 0.897) (Mann-Whitney U test).

Ratio of N/C AID was available for 12/18 cases cloned. No association was identified between ongoing mutation and N/C ratio of AID (r= 0.056, *P*= 0.863) (Spearman's correlation).

5.4.10: Correlation between clinical parameters of FL at time of entry to clinical trial and parameters tested in this chapter:

We compared AID induced mutations, analysing the % of *IGHV* mutation and ongoing mutation against clinical data at time of enrolling the patients to the clinical trial.

5.4.10.1 Relation between percent of AID induced mutations and FL clinical data:

The correlation between percentage of AID induced mutations and FL clinical data at time of entry to the clinical trial was explored.

No correlation was found between % of AID induced mutations and age (*P*= 0.408, r= 0.208) (Spearman's correlation) and sex (*P*= 0.503) (Mann-Whitney U test) of patients (Table 5.6).

Furthermore, it was found that there was no relation between % of AID induced mutations and number of LNs (*P*= 0.211) (Mann-Whitney U test).

Additionally, no correlation was identified between % of AID induced mutation and LDH level (*P*= 0.462, r= - 0.191) (Spearman's correlation), histological grade of patients (*P*= 0.244) (Kruskal -Wallis test) and Ann Arbor stage (*P*= 0.914) (Table 5.6).

Also no relation was identified between % of AID induced mutation and FLIPI score (*P*= 0.939) (Kruskal -Wallis test).

Parameter	Range or number and % of total	% of AID induced mutations (Mean ± SD)	% of AID induced mutations	
			P value	Statistical test
Age	57- 78	4.22 ± 1.63	0.408 (r = 0.208)	Spearman's correlation
Sex	Male 13 (72%)	4.21 ± 1.61	0.503	Mann-Whitney U test
	Female 5 (28%)	4.25 ± 1.89		
Number of LN	1 – 5 7 (39%)	3.81 ± 1.92	0.211	Mann-Whitney U test
	≥ 6 11 (61 %)	4.49 ± 1.46		
LDH (U/L)	109 – 862	4.22 ± 1.63	0.462 (r = -0.191)	Spearman's correlation
Histological grade	1 3 (17%)	3.1 ± 1.34	0.244	Kruskal -Wallis test
	2 10 (56%)	4.66 ± 1.79		
	3 5 (28%)	4.04 ± 1.344		
Ann Arber stage	II+III 11 (61%)	4.42 ± 1.54	0.914	Mann-Whitney U test
	IV 7 (39%)	3.92 ± 1.85		

Table 5.6: Correlation between percent of AID induced mutations and clinical features of FL at time of entry to clinical trial.

5.4.10.1 Correlation of percentage of *IGHV* mutation with FL clinical data:

The range of percentage of *IGHV* mutation in all 18 patients in our cohort was 2.89% – 19.22%, and the average was 14.86%. Although we did not find a significant difference in any of the clinical variables (including age, sex, number of lymph nodes, LDH level and Ann Arbor stage) recorded at time of trial entry (Table 5.7), a non-significant positive trend probably is present between percentage of *IGHV* mutation and histological grade of the disease in the 18 patient cohort. But no relation was noted between % of *IGHV* mutation and FLIPI score (P= 0.945) (Kruskal -Wallis test).

Parameter	Range or number and % of total	Mean ± SD of *IGHV* percentage mutation	Percentage of *IGHV* mutation	
			***P* value**	**Statistical test**
Age	57- 78	14.86 ± 4.56	0.672 (r = 0.111)	Spearman's correlation
Sex	Male 13 (72%)	15.31 ± 4.79	0.566	Mann-Whitney U test
	Female 5 (28%)	13.7 ± 4.14		
Number of LN	1 – 5 7 (39%)	15.48 ± 4	0.93	Mann-Whitney U test
	≥ 6 11 (61%)	14.47 ± 5.034		
LDH (U/L)	109 – 862	14.86 ± 4.56	0.462 (r = - 0.191)	Spearman's correlation
Histological grade	1 3 (17%)	10.17 ± 1.05	0.116	Kruskal -Wallis test
	2 10 (56%)	15.23 ± 5.3544		
	3 5 (28%)	16.95 ± 1.04		
Ann Arber stage	II+III 11 (61%)	14.99 ± 3.81	0.775	Kruskal -Wallis test
	IV 7 (39%)	14.66 ± 5.88		

Table 5.7: Results of statistical analyses for relations between percentage of *IGHV* gene mutation and clinical parameters of FL patients at time of entry to the clinical trial.

5.4.10.2 Relationship between number of subclones and FL clinical data:

The number of subclones in all patients ranged between (0 and 12) with an average of 7.22 and a median of 8. Patients were divided by the median into two groups. The first group consisted of 11 patients and the other group had seven patients. Differences in clinical variables were compared between the two groups of patients.

The correlation between number of subclones and age of the patients was first examined. No difference in age was identified between the two groups (*P*= 0.285) (Mann-Whitney U test) Similarly, no difference in LDH level was seen between the groups of low and high subclones (*P*= 0.364) (Mann-Whitney U test) (Table 5.8).

Parameter	**Range**	**Mean ± SD for number of subclones**	**Number of subclones 2 groups' ≤ 8 (11 patients (61.11%)) > 8 (7 patients (38.88%))**	
			***P* value**	**Statistical test**
Age	57 - 84	7.22 ± 3.64	0.285	Mann-Whitney U test
LDH (U/L)	109 – 830	7.22 ± 3.64	0.364	Mann-Whitney U test

Table 5.8: Correlation between age and LDH level against the two groups of low and high number of subclone.

The correlation between sex of patients and the 2 groups of subclones was tested. No relationship was identified (*P*= 0.59) (Fisher exact test) (Table 5.9).

Number of subclones (Median is 8)	**Sex of patient**	
	Male	**Female**
≤ 8 (n=11)	7	4
>8 (n=7)	6	1

Table 5.9: No correlation between sex of patients and number of subclones (*P*= 0.59) (Fisher exact test).

At entry to clinical trial, the actual number of enlarged LNs was noted if there were five or less. If six or more LNs were enlarged, the patients was given the number six. So, we divided patients into two groups, those with five or less enlarged LNs (seven patients) vs those with six or more LNs (11 patients). No correlation was found between number of enlarged lymph nodes and the groups of low and high subclones (*P*= 1) (Fisher exact test) (Table 5.10).

Number of subclones (Median is 8)	**Grade of disease**		
	2	3	4
≤ 8 (n=11)	2	7	2
>8 (n=7)	0	2	5

Table 5.10: No relationship between number of LNs and subclones of mutated *IGHV* (*P*= 1) (Fisher exact test)

No difference was found in the histological grade of disease between the groups of low and high subclones (*P*= 0.358) (Fisher-Freeman-Halton exact test) (Table 5.11).

Number of subclones (Median is 8)	**Number of LNs**	
	LNs < 6	**LNs ≥ 6**
≤ 8 (n=11)	4	7
>8 (n=7)	3	4

Table 5.11: No Correlation was found between grade of disease and subclones of mutated *IGHV* (*P*= 0.358) (Fisher-Freeman-Halton exact test).

Then, we compared the difference in the stage of disease among the groups of patients with low and high number of subclones. As mentioned in Chapter 4, Section 4.5, 18 of the 20 cases studied were at stage III (n = 9) or IV (n = 9), with only two cases being at stage II. So, we combined stages II and III into one group against stage IV disease in the other group (Table 4.6).

A significant positive correlation was identified between patients advanced stage of disease (stage IV) and number of subclones (*P*= 0.049) (Fisher exact test) (Table 5.12).
Thus, nine of the 11 (82%) patients with less advanced disease had ≤ 8 mutated *IGHV* subclones, while only two of the seven (29%) cases with stage 4 disease had a similar number of the subclones.

Number of subclones (Median is 8)	**Stage of disease (Ann Arber staging)**	
	II+III	**IV**
≤ 8 (n=11)	9	2
>8 (n=7)	2	5

Table 5.12: A positive significant correlation was identified between number of subclones in mutated *IGHV* and Ann Arbor stage (*P*= 0.049) (Fisher exact test)**.**

Finally, when we compared the number of subclones with FLIPI score, we found that there was no correlation between the two (*P*= 0.81) (Fisher-Freeman-Halton exact test).

5.4.10.3 Correlation of ongoing mutation with FL clinical data:

Following the above analyses, the ongoing *IGHV* mutation was measured by dividing the sum of unique mutations and partially shared mutations with PCR error rate (Section 5.3.5 of this chapter). The range of ongoing mutations was between (0-16.37) with a mean of 5.

The relationship between age of patients and ongoing mutation was first explored. It was found that there was no correlation between the two, (*P*= 0.371) (r= 0.224) (Spearman's correlation) (Table 5.13).

Next, we wanted to determine whether there was a difference in ongoing mutation between males and females. No difference of ongoing mutation was identified between the two sexes (*P*= 0.246) (Mann-Whitney U test).

Also no difference in the ongoing mutation was found between patients with high or low number of enlarged LNs (*P*= 0.375) (Mann-Whitney U test).

No correlation was found between LDH level and the degree of ongoing mutation (*P*= 0.127) (r = - 0.385) (Spearman's correlation). Also, no association was identified between ongoing mutation and FLIPI score (P=0.9) (Kruskal -Wallis test) (Table 5.13)..

Parameter	Range or number and percentage of total	Ongoing mutation Mean ± SD	Ongoing mutation	
			***P* value**	**Statistical test**
Age	57- 78	4.76 ± 4.54	0.371 (r = 0.224)	Spearman's correlation
Sex	Male 13 (72%)	5.62 ± 4.98	0.246	Mann-Whitney U test
	Female 5 (28%)	2.51 ± 2.14		
Number of LN	1 – 5 7 (39%)	5.62 ± 4.98	0.375	Mann-Whitney U test
	≥ 6 11 (61 %)	4.21 ± 4.40		
LDH (U/L)	109 – 862	4.76 ± 4.54	0.127 (r = - 0.385)	Spearman's correlation
Histological grade	1 3 (17%)	0.86 ± 0.82	0.037	Kruskal -Wallis test
	2 10 (56%)	6.53 ± 5.05		
	3 5 (28%)	3.56 ± 2.98		
Ann Arber stage	II+III 11 (61%)	2.9 ± 2.11	0.069	Mann-Whitney U test
	IV 7 (39%)	7.69 ± 5.89		
FLIPI score	0 3 (15%)	6.62 ± 8.04	0.9	Kruskal -Wallis test
	1 4 (20%)	3.9 ± 1.35		
	2 14 (65%)	4.57 ± 4.5		

Table 5.13: Correlation between ongoing mutations analysed by method from published papers and clinical parameters of FL at entry to clinical trial.

Ongoing mutation was compared among the histological grades of FL in our cohort of 18 patients. It was found that the ongoing mutation in patients with histological grade 2 FL significantly increased by more than two fold as compared to patients with grade 1.

In cases with grade 3, although slightly lower than that in grade 2, it still nearly doubled compared to that in the group with grade 1 (*P*= 0.037) (Kruskal -Wallis test) (Figure 5.11).

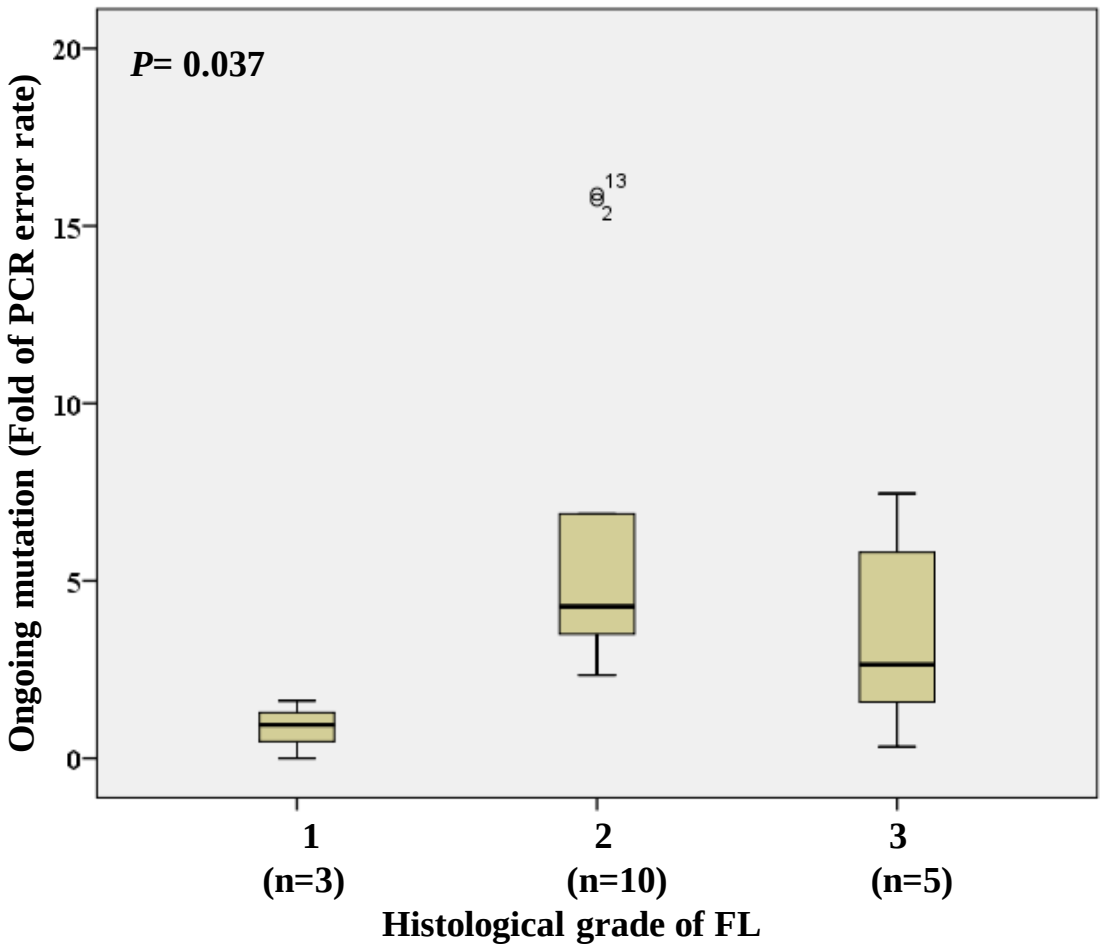

Figure 5.11: A positive significant correlation between ongoing mutations with advanced disease grade (*P*= 0.037) (Kruskal -Wallis test).

Furthermore, the average ongoing mutation level in patients at stage IV was nearly twice as high in cases at stage II and III. This showed a tendency towards increase in ongoing *IGHV* mutation in advanced FL (*P*= 0.069) (Mann-Whitney U test) (Figure 5.12).

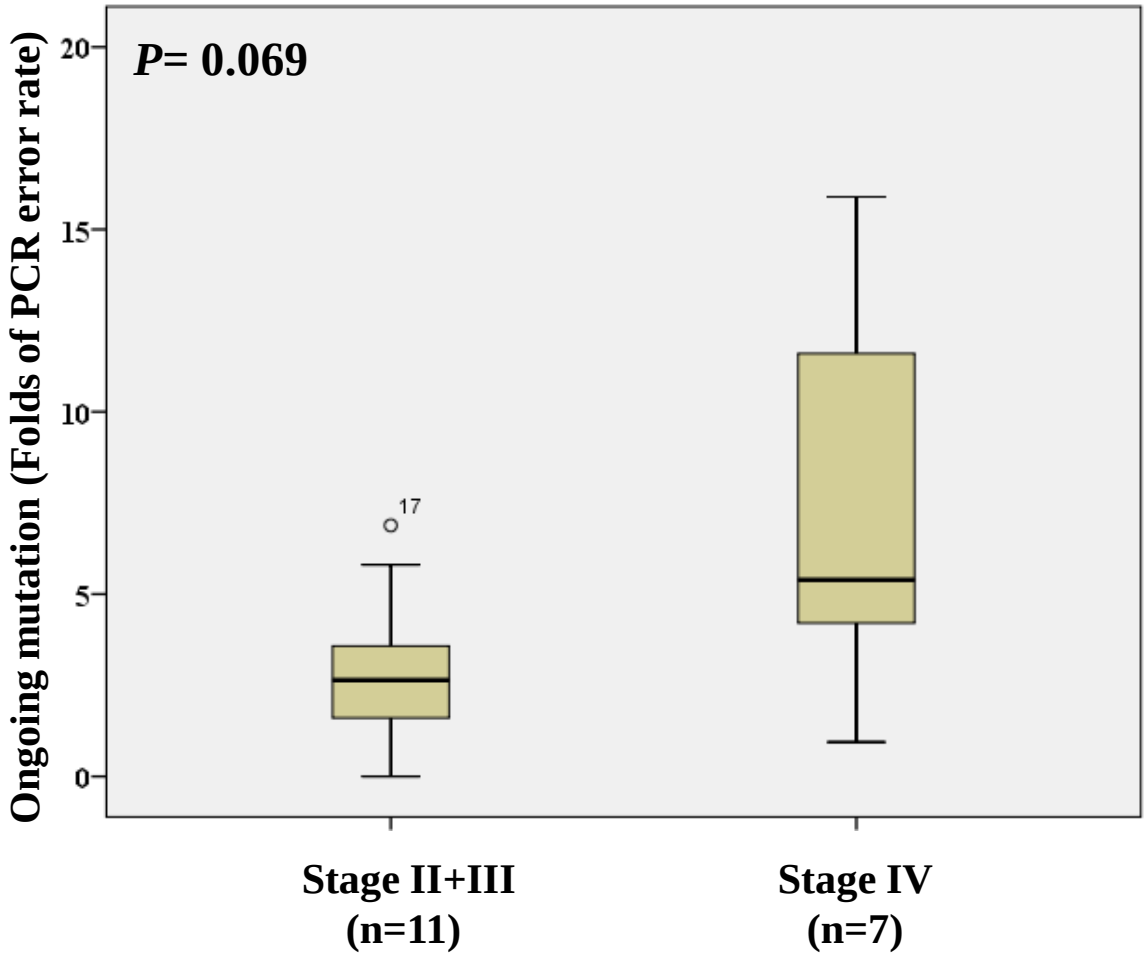

Figure 5.12: Difference in ongoing *IGHV* mutation between FL patients at different disease stages as analysed with a Mann-Whitney U test.

Discussion:

The aims of our work in this chapter were: (1) to explore the correlation between AID expression and ongoing *IGHV* mutations, and (2) to determine whether there was a relationship between the ongoing mutation and clinical variables in FL at time of patient enrolment in the clinical trial.

Although lacking a statistical significance, a positive trend was observed between both *AID* mRNA and protein expression with ongoing mutation in *IGHV* of the patients studied.

With the optimised protocols for DNA preparation and PCR in combining with BIOMED-2 primers, the *IGHV* gene was successfully amplified from 90% (18/20) of patient FFPE samples, and all of these PCR products were successfully subcloned and sequenced. The success rate was markedly higher than those previously reported (between 35% (375) and 79.7% (376)) in studies using similar FFPE samples from patients with FL.

The improvement made in this thesis was probably due to the following reasons:

1-Our optimizations of DNA extraction from FFPE tissue samples produced a very good quality of DNA. All of our samples produced at least a PCR band with a size of 200 bp or more (details given in Chapter 3, Section 3.2.2.2)

2-Optimizations of BIOMED-2 PCR conditions by increasing the $Mgcl_2$ concentration from 1.5-2.5 mM or increasing the number of PCR reaction cycles to 38 rather than 35. The Berget group found also that increasing PCR cycles to 38 produced better results (376), and this may explain better results of *IGHV* clonality detection of this group compared to the HALLDO´ RSDO´ TTIR paper (79.7% compared to 35%) (375).

So, it can be concluded that optimizing the BIOMED-2 PCR reaction is important for avoiding false negative cases associated with clonality detection in clinical services.

We followed the same analysis and also found the positive trend between *AID* mRNA expression and ongoing mutation reported by Hardianti et al. (358).

In agreement with Bose and colleagues, Loeffler and co-workers, and Jackson and colleagues, the number of mutations are more in FR regions because it covers a larger area of the *IGHV* gene (377, 378).

Although we found a trend of higher AID expression with ongoing mutation, However, Takata and colleagues and Smit and co-workers found that there was no correlation between AID expression, particularly at the mRNA level, with percentage of *IGHV* mutation and ongoing mutation in FL (365, 379). This is probably due to different tissues used (duodenal

rather than nodal FL) in the first paper or different method for ongoing mutation detection in the second publication. Willenbrock and colleagues and Kawamura and co-workers found that there was no correlation between AID expression and stage of disease in FL (237, 354). In conclusion, we found a positive correlation between ongoing *IGHV* mutation and both histological grade and Ann Arbor stage of FL.

Again, and similar to findings in the previous chapter, function of AID rather than expression level was correlated better with clinical parameters. However, increasing the size of the cohort of patients is necessary to prove the relationship between the functional readout of AID (N/C ratio of AID and ongoing mutation) with FL clinical features.

Finally, in agreement with Rebhandl and colleagues, our results suggest the possible role of AID not only in malignant transformation as thought earlier, but also in clonal evolution of B-cell malignancies (380).

Chapter 6

Investigation of relationships between transcriptional levels of genes associated with T cells and macrophages and AID expression in lymph node tissue, and their correlations with clinical status of FL

6.1 Introduction:

The tissue microenvironment plays a vital role in lymphoma pathogenesis. It helps malignant cells to grow, survive and escape the immune system (94). We can predict the prognosis of FL more accurately from knowledge of TME composition than from features of neoplastic B-cells (95).

In addition to its role in pathogenesis, the TME protects some lymphoma cells from destruction by therapeutic drugs. Malignant cell clones that are resistant to particular therapies are able to proliferate, resulting in disease relapse. Furthermore, with the help of the TME, resistant malignant lymphocytes can acquire further mutations and genomic instability that may lead in turn to the development of aggressive, complex and drug-resistant malignant clones (381).

The TME of FL contains different types of cells: most of them have been reported to be associated with disease outcome, and some of them are known to be associated with malignant transformation (94). TME cells in FL include reactive T lymphocytes of differing subtypes (helper T cells, tumour-infiltrating CD8+ T cells and follicular regulatory T cells). Other TME cells include macrophages, mast cells, FDCs, and fibroblastic reticular cells (382).

The different types of TME cells have varying correlations with FL prognosis: for example,

macrophages are associated with a bad prognosis in FL patients not treated with rituximab, and a good prognosis if patients were treated with rituximab (382).

Neoplastic follicles with lower dendritic cell densities and higher T regulatory cells are found to be associated with a decreased overall survival. This is probably caused by the inhibition of other beneficial T-cell types from proliferating within the FL TME; these other T-cell types would otherwise destroy the malignant lymphocytes (99).

Although gene expression profiling studies have provided clues relating TME gene expression to FL prognosis, IHC studies examining T-cell subsets and macrophages on FFPE tissues and correlating these IHC data to clinical outcome have been inconsistent, and hence have yet to be translated into clinical practice (73).

The process of SHM in FL is important for producing N-glycosylation sites for BCR in malignant FL cells (383), and the process of ongoing SHM favours maintaining these mannose molecules (384). The BCR mannose is capable of binding to mannose-binding lectins on bacteria, dendritic cells and macrophages, and commonly supporting survival of malignant lymphoma cells (87, 385). The correlation between AID expression and TME was studied in some other B-cell malignancies such as CLL. In CLL, it was found that in the group of patients with bad prognosis, AID levels were high. This high AID expression was associated with ongoing CSR and the expression was affected by TME signals (386). However, we did not find studies exploring the relationship between AID expression and TME in FL in the literature.

So, we wanted to explore the correlation between several genes expressed in TME and both AID expression (as a readout of SHM) and the process of ongoing SHM. The selected list of genes and number of samples were affected by the amount and quality of RNA available for our study.

6.1.1 Aim and hypothesis:

The aim of Chapter 6 was to examine the correlation of AID expression and function at diagnosis of FL with TME gene expression.

This is because it was found that in other B-cell malignancies, active TME induced AID expression (386). Also, cytokines secreted by inflammatory cells such as TNF, IL-1β, IL-4, IL-13 and TGFβ were found to be associated with ectopic expression of AID, inducing genomic instability and cancer development (387). We wanted to explore whether there is a relationship between AID and genes predicting prognosis and representing TME cells. Thus, AID expression or function can probably be added to TME to predict prognosis of disease.

6.2 Materials and methods:

6.2.1 Samples used for qPCR measurement of TME gene expression:

To carry out cDNA synthesis for qPCR gene expression, only cases with good RNA amount and quality (RIN ≥ 2.1) were used. Therefore, of the 87 cases from the PACIFICO trial included in this thesis, only 31 samples met such criteria, and were therefore used for research in this chapter. Seventeen of these cases had low AID protein expression (IHC score=1) and 14 high AID (IHC score of ≥4).

6.2.2 Approaches to identify genes expression of TME in FL:

mRNA expression of 6 genes in the TME that predicts prognosis in FL were explored (see below for details). The transcription level of these genes to EC gene (*GAPDH*) by qPCR was measured in the same way, and using the same equation, as presented in Chapter 2, Section 2.3.7.

6.2.3 Selecting genes for Taqman qPCR reaction for TME genes in FL:

The most important publication that initiated investigation of the impact of TME in FL with disease prognosis is Dave and colleagues' paper published in the New England Journal of

Medicine (NEJM) in 2004. These authors identified two sets of genes representing TME cells. They named the first group "immune-response 1 signature genes"; these were 42 genes predicting good prognosis and represent genes expressed mainly in T cells and sometimes monocytes. The other group were "immune-response 2 signature genes" (24 genes). This second group of genes are expressed in monocytes and/or FDCs but not T cells. The immune-response 2 signature genes were associated with poor disease outcome (95).

In 2008, Byers and colleagues (273) designed a Taqman qPCR reaction to measure expression of 35 genes studied by Dave et al. and Glas et al. (95, 388). From lists presented in these papers, six indicator genes were selected to explore their correlation with AID expression. The selection of these genes was based on the following factors: First, these genes were studied in at least two of these publications. Second, they are proposed to be expressed in specific types of TME cells (T cells vs macrophages or FDCs). Third, the qPCR measurement of gene transcription had been optimized for these genes. Three genes were supposed to be associated with good prognosis and mainly expressed in T cells or monocytes (*CD2*, *LEF1* and *STAT4*), and 3 genes were associated with bad prognosis and are expressed mainly in macrophages or FDCs (*LGMN*, *C3AR1* and *MAPK1*). However, *MAPK1* is not specific to macrophages and can be expressed in other human cells (273). Some of these cells are NK cells, eosinophils and epithelial cells (389-391).

Details of types of cells expressing these genes and cellular pathways associated with these genes can be seen in Chapter 2, Table 2.13. These genes represent specific types of cells in TME and the Taqman qPCR expression for these genes has been optimized by Byers and colleagues. In addition, these genes were found to be clearly associated with disease prognosis in FL in Dave and colleagues' work.

6.2.4 Selecting cell lines to be used as a control for selected genes:

To identify controls for these six genes representing TME expression signatures, cDNA from different haematological cell lines in our department was utilised. An optimized and published qPCR protocol to identify these genes' expression in cell lines was used (273).

TME genes were tested in these cell lines: Raji cell line (Burkett's lymphoma) (257), Karpas 422 cell line (Diffuse Large B Cell Lymphoma) (274), Daudy cell line (Burkett's lymphoma) (275) and Jurkat cell line (Acute T cell Leukaemia) (276).

It was confirmed that genes that are associated mainly with macrophages and FDCs (*MAPK1*, *C3AR1*and *LGMN*) have lower Ct values of qPCR in the Daudi cell line.

On the other hand, genes associated with T cell expression (*CD2*, *LEF1* and *STAT4*) appeared in earlier qPCR cycles in a human T lymphocyte cell line (Jurkat cell line).

6.3 Results of TME gene expression profile:

qPCR reaction details and optimizations were discussed in Chapter 2, Section 2.7. For every patient in the cohort, qPCR expression (Ct value) for each of the six genes and the EC gene (*GAPDH*) was repeated three times (technical replicates). Positive control cell lines qPCR expression was repeated three times also, and a water control was added to identify possible qPCR contamination.

Regarding genes associated with good prognosis of FL, the ratio of *STAT/GAPDH* among all cases ranged between 0.0001-0.07, with a mean of 0.011. *LEF1/GAPDH* ranged between 0.0001- 0.046, the average was 0.008. *CD2/GAPDH* expression was 0.0002-0.211 and the mean was 0.041.

With respect to those genes associated with bad prognosis of FL, firstly we had *C3AR1/GAPDH* average of 0.0073 and range of 0.0001 - 0.025. *LGMN/GAPDH* mean in all cases was 0.0077 with a range of 0.0009 - 0.033 and finally *MAPK1/GAPDH* expression

range was between 0.0048 - 0.31 and the mean was 0.089. Expression of genes was variable among cases. A heat map representing gene expression among all cases is seen in Figure 6.1.

AID protein	Good prognosis genes			Bad prognosis genes		
	STAT4	*LEF1*	*CD2*	*C3AR1*	*LGMN*	*MAPK1*
Low AID	0.0216	0.0159	0.0938	0.0003	0.0202	0.0450
	0.0470	0.0186	0.1813	0.0240	0.0086	0.2062
	0.0003	0.0002	0.0018	0.0033	0.0029	0.0344
	0.0130	0.0319	0.1081	0.0179	0.0128	0.3059
	0.0073	0.0103	0.0157	0.0101	0.0091	0.0927
	0.0087	0.0093	0.0629	0.0083	0.0131	0.0987
	0.0677	0.0355	0.1955	0.0159	0.0112	0.1632
	0.0194	0.0024	0.0158	0.0002	0.0015	0.0379
	0.0198	0.0038	0.0668	0.0146	0.0325	0.1241
	0.0090	0.0091	0.0756	0.0070	0.0149	0.0619
	0.0007	0.0003	0.0037	0.0248	0.0060	0.0504
	0.0003	0.0003	0.0022	0.0006	0.0028	0.0048
	0.0004	0.0004	0.0024	0.0025	0.0133	0.0659
	0.0268	0.0459	0.2110	0.0072	0.0161	0.2849
	0.0014	0.0007	0.0041	0.0095	0.0029	0.0897
	0.0022	0.0010	0.0041	0.0034	0.0023	0.0886
	0.0004	0.0002	0.0015	0.0037	0.0020	0.0704
High AID	0.0003	0.0003	0.0025	0.0048	0.0027	0.0736
	0.0001	0.0001	0.0014	0.0053	0.0009	0.0624
	0.0497	0.0249	0.0912	0.0069	0.0138	0.1411
	0.0009	0.0003	0.0022	0.0026	0.0024	0.0661
	0.0007	0.0018	0.0023	0.0069	0.0063	0.0776
	0.0207	0.0128	0.0604	0.0001	0.0027	0.0522
	0.0001	0.0001	0.0002	0.0040	0.0029	0.0423
	0.0092	0.0114	0.0483	0.0048	0.0083	0.0589
	0.0004	0.0002	0.0024	0.0062	0.0019	0.0902
	0.0007	0.0003	0.0032	0.0050	0.0069	0.0832
	0.0001	0.0001	0.0008	0.0028	0.0020	0.0230
	0.0003	0.0003	0.0018	0.0095	0.0025	0.0261
	0.0029	0.0025	0.0075	0.0037	0.0042	0.0902
	0.0020	0.0014	0.0103	0.0091	0.0095	0.0561

0.0001 **Colour scale** **0.3059**

Figure 6.1: Heat map of genes representing TME and predicting prognosis in FL among cases of low and high AID. The ratio of gene expression to *GAPDH* increased gradually from low (green colour) to yellow then red (highest expression).

6.3.1 Correlations between different gene expressions:

The correlation of expression level among the three genes reported to be associated with good prognosis in FL patients was explored. A strong significant positive correlation was observed between all of these genes as analysed using Spearman's correlation test. As shown in Figure 6.2, *LEF1/GAPDH* and *STAT4/GAPDH* were significantly correlated (r= 0.933, *P*< 0.001). *LEF1/GAPDH* and *CD2/GAPDH* were also positively and significantly correlated (r= 0.943, *P*< 0.001). *STAT4/GAPDH* and *CD2/GAPDH* had the strongest positive correlation (r= 0.95, *P*< 0.001).

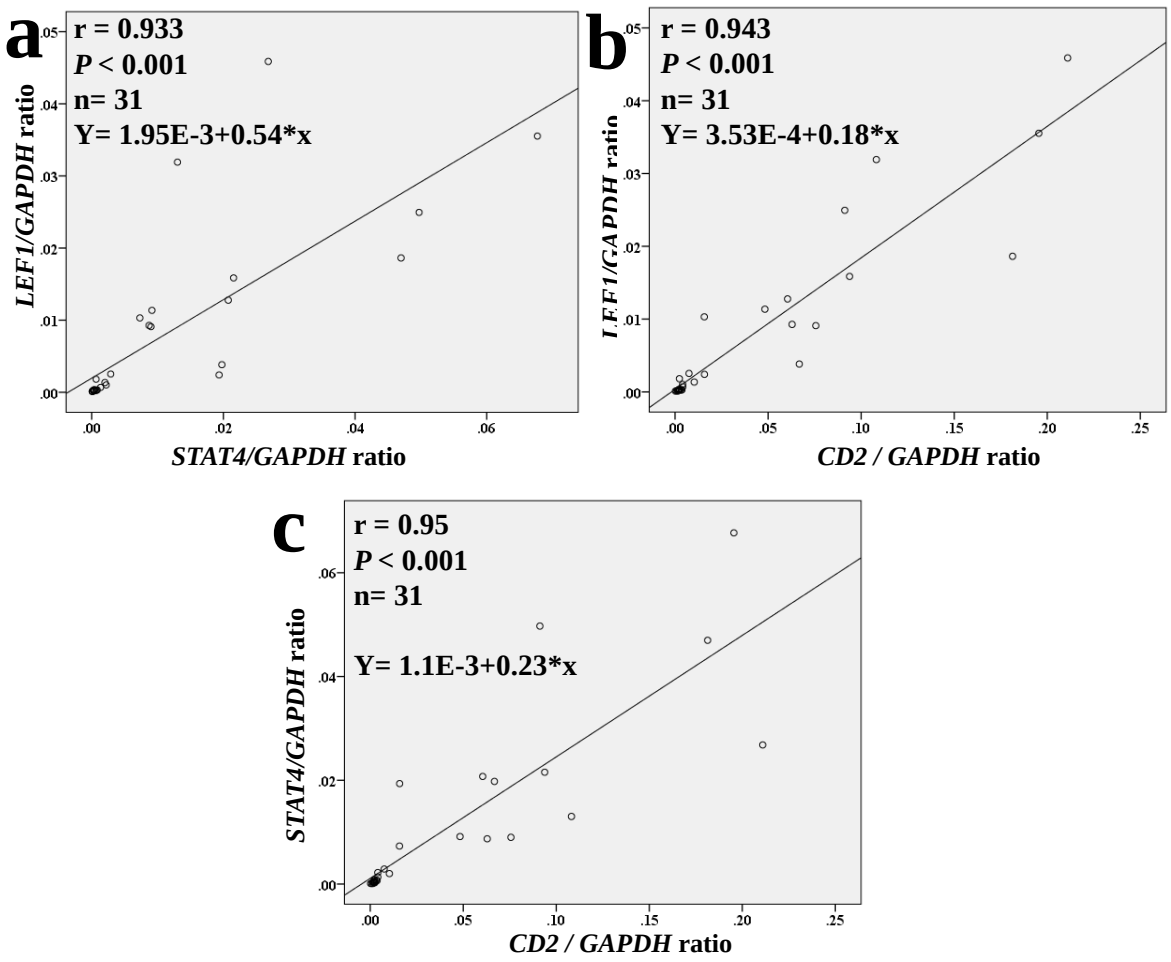

Figure 6.2: A significant positive correlation was observed between all of the three genes expressed in T cells in TME and reportedly associated with good prognosis in FL. Spearman's correlation analysis was used in all figures.

In addition, weak correlations were also observed between genes associated with bad prognosis. *C3AR1/GAPDH* and *LGMN/GAPDH* were significantly correlated (r= 0.419, *P*= 0.019). *C3AR1/GAPDH* and *MAPK1/GAPDH* were also positively and significantly correlated (r= 0.569, *P*= 0.001), as were *LGMN/GAPDH* and *MAPK1/GAPDH* (r= 0.441, *P*= 0.013) (Figure 6.3).

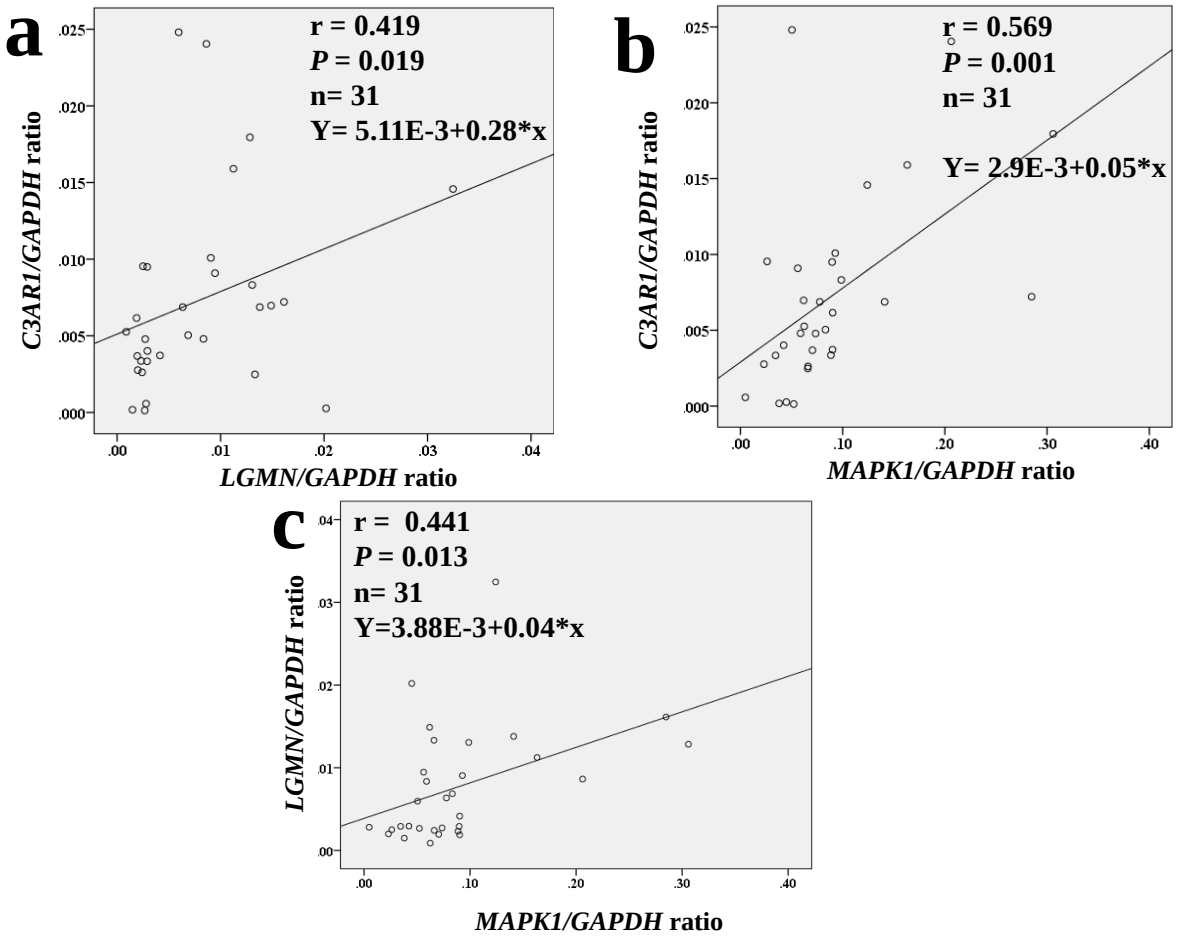

Figure 6.3: Results of Spearman's correlation analyses between genes expressed in TME and associated with poor prognosis in FL.

6.3.1 Correlation between AID expression and gene signatures of FL microenvironment:

To explore the relationship between AID expression and TME genes, the expression of these genes was correlated with mRNA qPCR and total IHC protein expression of AID.

6.3.1.1 Relation between *AID/GAPDH* mRNA expression and TME genes:

No correlation was found between *AID/GAPDH* mRNA ratio and genes representing TME and predicting prognosis in FL. Correlations between *AID/GAPDH* mRNA and *CD2/GAPDH* (r= 0.031, *P*= 0.868), *STAT4/GAPDH* (r= 0.102, *P*= 0.587) and *LEF1/GAPDH* (r= 0.102, *P*= 0.578), *C3AR1/GAPDH* (r= - 0.002, *P*= 0.991), *LGMN/GAPDH* (r= - 0.151, *P*= 0.417) and *MAPK1/GAPDH* were also not correlated (r =0.096, *P*= 0.608) (Spearman's correlation).

6.3.2 Relationship between AID protein and TME gene signatures of FL:

Based on AID protein expression by IHC, patients were divided into two groups. Seventeen patients with low AID protein (IHC=1) and 14 cases with high AID protein (IHC ≥ 4) were selected, then TME genes were compared between the two groups.

A significant negative correlation was identified between two of the genes associated with good prognosis and AID protein. Higher *STAT4/GAPDH* (*P*= 0.048) and *CD2/GAPDH* (*P*= 0.026) gene expression was found in cases with lower total AID protein. A non-significant negative trend was also identified between AID protein expression and the 3rd gene of good prognosis in FL (*LEF1/GAPDH*) (*P*= 0.071). A Mann-Whitney U test was used in all calculations (Figure 6.4).

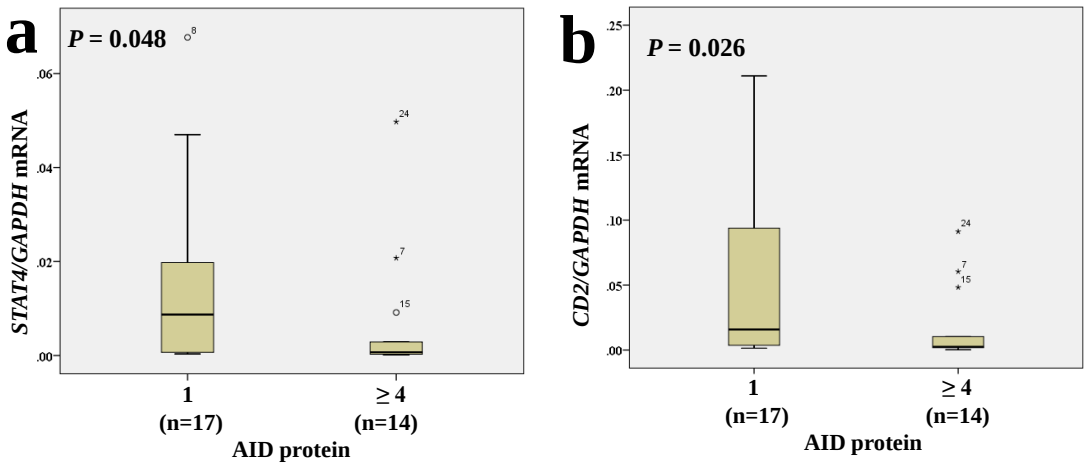

Figure 6.4: A significant negative correlation was found between AID protein expression and two genes predicting good prognosis in FL. A Mann-Whitney U test was performed.

No correlation was identified between any of the bad prognosis genes and AID protein. However, a negative trend was identified between AID expression and *LGMN/GAPDH* gene (*P*= 0.053). The correlations between AID protein level and *C3AR1/GAPDH* was (*P*= 0.297) and *MAPK1/GAPDH* association with AID was (*P*= 0.262) (Mann-Whitney U test).

It is important to note that, for the good prognosis genes, which are negatively correlated with AID protein, the degree of increase in the median of gene expression with decreasing AID protein is 6.32 – 12.7 fold. However, in the *LGMN* gene, which had no statistically significant negative correlation with AID protein, the change in expression was only 3.3 fold (Table 6.1).

	Gene	**Median of gene expression in low AID protein group**	**Median of gene expression in high AID protein group**	**Folds of change**	***P* value**
Good prognosis genes	*STAT4*	0.0087	0.0007	12.4	0.048
	CD2	0.0158	0.0025	6.32	0.026
	LEF1	0.0038	0.0003	12.7	0.071
Bad prognosis genes	*LGMN*	0.0091	0.0028	3.3	0.053
	C3AR1	0.0072	0.0049	1.5	0.297
	MAPK1	0.0885	0.0643	1.4	0.262

Table 6.1: Folds of increase in gene expression of the examined genes with decreased AID protein.

6.3.3: Association between TME genes and nuclear/cytoplasmic ratio of AID:

Only 14 out of 31 patients had data for N/C AID ratio. The relationship between this ratio and TME was explored. No significant relationship was identified between any of these genes and N/C ratio of AID.

Correlations between N/C ratio of AID and *CD2/GAPDH* (r= - 0.152, *P*= 0.605), *STAT4/GAPDH* (r= - 0.332, *P*= 0.246) and *LEF1/GAPDH* (r= - 0.068, *P*= 0.817), *C3AR1/GAPDH* (r= 0.231, *P*= 0.427), *LGMN/GAPDH* (r= - 0.09, *P*= 0.759) and *MAPK1/GAPDH* (r =0.363, *P*= 0.203) were found. Spearman's correlation was used in all tests.

6.3.4 Correlation between EC gene in qPCR and Taqman qPCR reactions:

In order to confirm the accuracy of gene transcription level using Taqman qPCR, Ct value of *GAPDH* gene using SYBR Green dye in qPCR that was discussed in Chapter 2, Section 2.3.7 was compared to *GAPDH* Ct values using Taqman reagents of qPCR. A positive significant correlation was noted between the *GAPDH* gene measured by qPCR for mRNA expression of *AID/GAPDH* and the Taqman method (r= 0.865, p< 0.001, n= 31) (Spearman's correlation) (Figure 6.5).

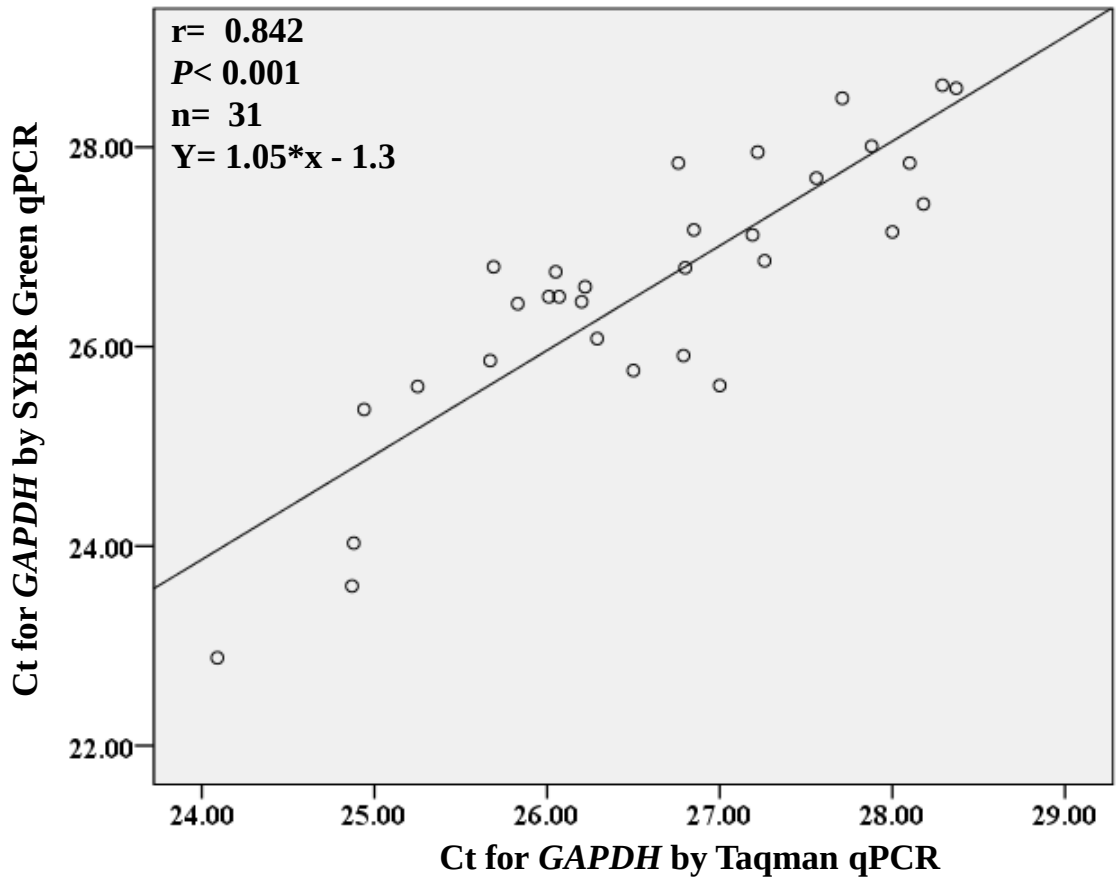

Figure 6.5: A positive significant correlation between Ct values of *GAPDH* measured by qPCR and Taqman qPCR (Spearman's correlation).

6.3.5 Association betweenTME genes' expression and ongoing mutation:

Out of 31 samples used in this chapter, we had ongoing mutation results for 11 FL samples only. The association between TME genes and *IGHV* ongoing mutation was explored. Ongoing mutation represents folds of PCR error rate mutations; the method of quantification was discussed in Chapter 5, Section 5.3.5. No correlation was found between ongoing mutation and any of the expressed genes (Table 6.2).

	TME Gene	Correlation between ongoing mutation and TME genes	
		r value	***P* value**
Good prognosis genes	*STAT4/GAPDH*	- 0.418	0.201
	LEF1/GAPDH	- 0.355	0.285
	CD2/GAPDH	- 0.291	0.385
Bad prognosis genes	*C3AR1/GAPDH*	- 0.136	0.689
	LGMN/GAPDH	- 0.055	0.873
	MAPK1/GAPDH	0.036	0.915

Table 6.2: Correlations between ongoing mutation in *IGHV* and TME gene expression. Spearman's correlation was used in all tests.

6.3.6 Relationship between TME gene expressions with clinical data at entry to clinical trial:

The correlations between TME gene expression and patient clinical data was explored. This is because there is controversy in the literature regarding the correlation of these genes with clinical outcome. Some of those genes were correlated with good and poor prognosis of FL in some studies, while not in others (details provided in the discussion section). We wanted to explore this association in more detail, as outlined below.

6.3.6.1 Relationship between sex of patients and TME genes:

Out of 31 patients in our cohort, 16 were males (52%) and 15 (48%) were females. No correlation was identified between sex of patients and any of the TME genes. A Mann-Whitney U test was used to identify all relationships (Table 6.3).

	TME Gene	Correlation between sex of patients and TME genes (*P* value)
Good prognosis genes	*STAT4/GAPDH*	0.8
	LEF1/GAPDH	0.47
	CD2/GAPDH	0.711
Bad prognosis genes	*C3AR1/GAPDH*	0.232
	LGMN/GAPDH	0.922
	MAPK1/GAPDH	0.545

Table 6.3: Correlations between ongoing mutation and TME gene expression. Mann-Whitney U test used in all correlations.

6.3.6.2 Correlation between age of patients and TME genes:

The age of patients ranged between (64 to 84) with an average of 70.6 years. Age of patients was compared to TME genes. A weak negative correlation was identified between age of patients and *LGMN/GAPDH* expression (r= - 0.376) (*P*= 0.037) (Spearman's correlation). No correlation was identified between the other five genes and age of patients. Relationships between age and *STAT4/GAPDH* (r= - 0.136) (*P*= 0.465), *LEF1/GAPDH* (r= - 0.209) (*P*= 0.259), *CD2/GAPDH* (r= - 0.13) (*P*= 0.487), *C3AR1/GAPDH* (r= - 0.118 (*P*= 0.4529), *MAPK1/GAPDH* (r= - 0.017) (*P*= 0.929) are as reported here. Spearman's correlation was used in all tests.

6.3.6.3 Association between number of LNs and TME genes:

The number of LNs of patients were divided into two groups; those with five or less enlarged LNs were group one, forming ten out of 31 patients (32%). Patients with six or more LNs were group 2; those were 21 cases (68%). Patients were divided into these two groups

because on trial data sheets, for any patient with six or more LNs, the number 6 was written. No correlation was observed between any of the six genes expressed in TME and the two groups of enlarged LNs. A Mann-Whitney U test was used to explore correlations between the number of enlarged LNs and genes (Table 6.4).

	TME Gene	**Mean ± SD in patients with ≤ 5 enlarged LNs**	**Mean ± SD in patients with ≥ 6 enlarged LNs**	***P* value**
Good prognosis genes	*STAT4/GAPDH*	0.014 ± 0.024	0.009 ± 0.012	0.852
	LEF1/GAPDH	0.007 ± 0.012	0.008 ± 0.012	0.884
	CD2/GAPDH	0.049 ± 0.079	0.038 ± 0.053	0.917
Bad prognosis genes	*C3AR1/GAPDH*	0.009 ± 0.009	0.006 ± 0.004	0.724
	LGMN/GAPDH	0.008 ± 0.006	0.008 ± 0.008	0.466
	MAPK1/GAPDH	0.084± 0.058	0.092 ± 0.08	0.724

Table 6.4: Association between genes predicting good and bad prognosis in FL and groups of patients with low and high number of enlarged LNs. Mann-Whitney U test used.

6.3.6.4 Relationship between TME genes and LDH level:

The LDH level ranged between (109 and 862 U/L), with an average of 407 U/L in our 31 patient cohort. When the correlation between LDH level and expression of TME genes was explored, it was found that there was no correlation between LDH level and any of the TME genes. Spearman's correlation was used in all tests (Table 6.5).

	TME Gene	Correlation between TME genes and LDH level	
		r value	***P* value**
Good prognosis genes	*STAT4/GAPDH*	0.018	0.922
	LEF1/GAPDH	0.11	0.556
	CD2/GAPDH	0.013	0.944
Bad prognosis genes	*C3AR1/GAPDH*	- 0.335	0.066
	LGMN/GAPDH	- 0.125	0.501
	MAPK1/GAPDH	- 0.16	0.389

Table 6.5: Correlations between LDH level and TME gene expression. Spearman's correlation was used in all tests.

6.3.6.5 Association between grade of FL and TME genes:

Patients with FL were divided into three grades. Our cohort in this chapter was 31 samples. Six patients were in grade 1 disease (19%), 17 samples had grade 2 by IHC (55%) and eight patients were grade 3 disease (26%). No correlation was observed between any of the TME genes and grade of FL. A Kruskal-Wallis test was used in all correlations (Table 6.6).

	TME Gene	Mean ± SD for grade 1 disease (n= 6)	Mean ± SD for grade 2 disease (n= 17)	Mean ± SD for grade 3 disease (n= 8)	TME genes and grade of disease correlation (*P* value)
Good prognosis genes	*STAT4/GAPDH*	0.016 ± 0.06	0.01 ± 0.02	0.01 ± 0.01	0.324
	LEF1/GAPDH	0.006 ± 0.01	0.007 ± 0.012	0.012 ± 0.016	0.772
	CD2/GAPDH	0.032 ± 0.038	0.04 ± 0.07	0.053 ± 0.073	0.477
Bad prognosis genes	*C3AR1/GAPDH*	0.011 ± 0.008	0.008 ± 0.0068	0.004 ± 0.003	0.149
	LGMN/GAPDH	0.011 ± 0.011	0.007 ± 0.005	0.008 ± 0.007	0.629
	MAPK1/GAPDH	0.083 ± 0.042	0.1 ± 0.07	0.074 ± 0.088	0.12

Table 6.6: Association between grade of disease and TME gene expression. Kruskal-Wallis test used in all correlations.

6.3.6.6 Association between stage and FLIPI score of FL with TME genes:

Patients were divided into two groups. The low stage disease group (stage Π and III) were 18 patients and the high stage group (stage IV) were 13 patients. Only four out of 31 patients (13 %) were in stage Π disease. So, these patients were combined with those at stage III. The association between TME and stage of disease was explored. No correlation between the two groups was identified (Table 6.7).

	TME Gene	Mean ± SD for stage II and III disease	Mean ± SD for stage IV disease	TME genes and stage of disease correlation (*P* value)
Good prognosis genes	*STAT4/GAPDH*	0.012 ± 0.019	0.009 ± 0.014	0.737
	LEF1/GAPDH	0.01 ± 0.015	0.005 ± 0.006	0.54
	CD2/GAPDH	0.043 ± 0.067	0.039 ± 0.055	0.953
Bad prognosis genes	*C3AR1/GAPDH*	0.008 ± 0.007	0.007 ± 0.0067	0.89
	LGMN/GAPDH	0.007 ± 0.005	0.009 ± 0.009	0.737
	MAPK1/GAPDH	0.097 ± 0.08	0.079 ± 0.05	0.921

Table 6.7: Correlations between stage of disease and TME gene expression. Mann-Whitney U test used in all correlations.

Again, no correlation was identified between the FLIPI score and TME genes (Table 6.8)

	TME Gene	Mean ± SD for FLIPI 0	Mean ± SD for FLIPI 1	Mean ± SD for FLIPI 2	correlation (*P* value)
Good prognosis genes	*STAT4/GAPDH*	0.023±0.039	0.007±0.001	0.01±0.014	0.605
	LEF1/GAPDH	0.012 ± 0.02	0.004±0.001	0.008±0.012	0.806
	CD2/GAPDH	0.067±0.111	0.02 ± 0.003	0.043±0.058	0.565
Bad prognosis genes	*C3AR1/GAPDH*	0.015±0.011	0.008±0.004	0.007±0.006	0.423
	LGMN/GAPDH	0.007±0.0047	0.007±0.006	0.008±0.008	0.976
	MAPK1/GAPDH	0.083 ± 0.07	0.066±0.032	0.094±0.073	0.696

Table 6.8: No association between TME genes and FLIPI score, Kruskal-Wallis test was used in all correlations.

6.4 Discussion:

In this chapter, an exploration into the correlation between AID expression and function with immune signature of FL was initiated by measuring qPCR expression of six genes. Three of

these genes (*C3AR1*, *LGMN* and *MAPK1*) were found to be mainly expressed in macrophages and are associated with poor prognosis, while *CD2*, *STAT4* and *LEF1* are expressed mainly by T cells and are associated with good prognosis (273).

Only 17 cases with low and 14 samples of high AID protein had sufficient concentration of good quality mRNA to be used in the experiments undertaken. In this study, it was found that genes linked to good prognosis were strongly associated with each other; also, genes involved with poor prognosis in FL were also correlated with one another, but to a weaker extent (Figures 6.2 and 6.3). This is in agreement with Dave and colleagues, and is expected because genes that are expressed in the same cell correlated positively with each other (95). In this study, no significant correlation was found between AID mRNA expression and any of the TME genes. On the other hand, a negative significant correlation was identified between *STAT4* and *LEF1* with AID protein expression. However, the current results may support our hypothesis of poor disease outcome in cases with higher AID expression or function. Nevertheless, increasing the size of the cohort is required to further support these results.

In this study, we did not identify a correlation between the functional 'read out' of AID (ongoing mutation and N/C ratio of AID) with the six genes' expression. Also, no correlation was identified between these genes and clinical parameters at entry to clinical trial. The only exception is that we found decreased expression of *LGMN* with age. No studies were found in the literature that investigated the transcription level of this gene and lymphoma. However, it was found that *LGMN* gene expression decreases with aging of brain tissues (392).

In agreement with most of our findings, Byers and colleagues used the same qPCR Taqman technique and the same genes that we used. They found that none of the six examined genes were associated with prognosis of FL (273).

Glas and co-workers, as well as Chevalier, found that not only the amount of T cells in the TME is important in prognosis, but also the location with early FL transformation if T helper cells are located within the neoplastic follicle (99, 393). So, it is likely that studying expression of TME genes alone, without determining the exact location of TME cells, is insufficient for predicting the prognosis in FL.

In summary, we identified a correlation of lower AID protein expression and good prognosis genes in FL. However, the TME is a complex structure of stromal and blood cells. The number, composition and location of TME cells play a role in prognosis of disease (394).

The addition of rituximab to chemotherapy reversed the association of poor prognosis with high macrophage content in the pre-rituximab era to a good prognosis after adding rituximab to chemotherapy in FL patients (395).

Hence, for this reason, data from a larger number of genes, involvement of other methods of gene and protein expression, and determining the number, type and location of TME cells is required before drawing a definite conclusion about any associations of TME with AID expression, and with FL prognoses.

Chapter 7

Correlations between AID expression and function against clinical outcome of FL

7.1 Introduction:

AID activity is associated with ongoing SHM, ASHM and chromosomal translocations. All of these processes are associated with DNA damage that is linked with the pathogenesis and/or transformation of FL (20, 78, 139, 237, 396).

It was of interest to correlate AID expression and activity with clinical outcome in patients with FL for the following reasons:

1-AID expression and activity was heterogeneous across our cohort of patients (Chapter 4, Section 4.2.3), raising the possibility that it could account for the known variability in clinical outcome in FL.

2- Correlations and trends were identified between AID activity and baseline clinical features at the time of trial entry. In particular, a significant positive correlation was identified between ongoing *IGHV* mutation and histological grade. In addition, a tendency towards higher ongoing mutation and more advanced anatomical stage (Chapter 5, Section 5.4.10.3). Furthermore, a suggestive of a trend if a larger sample size could be analysed in future between the N/C ratio of AID and stage of disease (Chapter 4, Section 4.5). These correlations and trends suggest that AID activity influences the course of the disease prior to the point of trial entry. It was therefore logical to examine the possible role of AID in influencing the course of the disease after the point of trial entry.

3-The cohort of samples used to characterise AID expression and function were from the PACIFICO clinical trial, and was therefore well-defined with linked longitudinal data (although since the trial was still in follow-up mode, not all of this data was available).

4- Studies analysing the correlation between AID activity and clinical outcome in FL are not found within the literature, although one paper did report a failure to identify any correlation between FL prognosis and AID protein expression by IHC (354).

5- A significant positive correlation has previously been identified between AID expression/ activity and outcome in other B-cell malignancies such as DLBCL and CLL (237, 397).

Based on the considerations outlined above, it was hypothesised that high AID activity (rather than expression) is associated with genomic instability and poor prognosis in FL. To test this hypothesis, laboratory data presented in Chapters 4, 5 and 6 were related to available longitudinal clinical data from the same patients. In these previous chapters, an optimization of nucleic acid extraction from FFPE tissues was carried out, followed by measurement of overall *AID* mRNA expression, overall AID protein expression, and the N/C AID ratio in these tissues. AID activity, in the form of ongoing mutation, and AID-specific mutations in the *IGHV* gene was also measured, and the correlation of AID expression with genes associated with the TME, were also determined. Correlations between these laboratory variables and longitudinal clinical data are presented in the current chapter.

7.2 Materials and methods:

7.2.1 Patients and data:

Longitudinal clinical data was obtained for the 87 patients in the PACIFICO trial whose samples were used to generate the laboratory data presented in Chapters 4, 5 and 6. The data were provided by and analysed in collaboration with the trial statistician.

7.3 Statistical analysis:

Progression-free survival (PFS), overall survival (OS) and the response rate (RR) at the end of the planned eight cycles of induction treatment were compared to laboratory data. Cox proportional hazards model was used to explore the effect of laboratory variables on PFS and OS. Boxplots and logistic regression were used to explore the effect of those data on RR after induction treatment (just after or before cycle 8).

PFS is defined as disease progression or death from any cause. Response to treatment was based on principal investigators' assessment, and takes four possible values: Complete Response (CR), Partial Response (PR), Stable Disease (SDS) and Progressive Disease (PD). Age was adjusted for the analysis of this chapter to eliminate the confounding effect of aging on clinical endpoints.

7.4 Results:

7.4.1 Relationship between AID expression and trial endpoints:

In our cohort of 87 patients, only 17 have experienced a PFS event (9 progressions and 8 deaths from any cause). The number of PFS events is smaller for the qPCR analysis (11 only). This is because qPCR results were missing for 28 samples where the quantity or quality of RNA was low. The median follow-up time was 28 months while the maximum was 70 months.

No association was noted between AID expression at both mRNA and protein level with PFS in our cohort of patients (Table 7.1).

Variable	HR	95% CI for HR	*P* value
IHC score	0.94	0.74, 1.20	0.624
log(qPCR)	0.88	0.65, 1.20	0.433

Table 7.1: Presents the results from the cox proportional hazards model fitted separately for each variable; the age-adjusted cox proportional hazards model yields a non-significant *P* value in both cases.HR; Hazard ratio, CI; Confidence Interval

Out of the 87 patients, only eleven have experienced an OS event so far. The number of OS events is even smaller for the qPCR analysis (7). No correlation was identified between AID expression and OS (Table 7.2).

Variable	HR	95% CI for HR	*P* value
IHC score	0.86	0.61, 1.22	0.404
log(qPCR)	1.16	0.70, 1.92	0.576

Table 7.2: No association was identified between AID expression at both mRNA and protein level with OS; the age-adjusted cox proportional hazards model was used.

Out of the 87 patients, the response status is missing for ten patients. Among the remainder, 31 achieved a CR, 41 achieved a PR, 3 had SDS, and 2 had PD. Hence, our analysis compared the patients who achieved CR against those who did not. No correlation was seen between AID expression and RR (Table 7.3).

Variable	OR	95% CI for HR	*P* value
IHC score	1.07	0.86, 1.33	0.523
log(qPCR)	0.89	0.62, 1.25	0.494

Table 7.3: Presents the results from the logistic model fitted separately for each variable; the age-adjusted logistic regression yields a non-significant *P* value in both cases.

7.4.2. Relationship between AID N/C ratio and trial endpoints:

AID N/C ratio, AID-specific mutation, and ongoing mutation of *IGHV* gene were used as readouts of AID function.

The N/C ratio of AID was measured in 20 out of the 87 patients. Among these, only four have experienced a PFS event at the time of statistical analysis (one progression and three deaths). The median follow-up time is 34 months while the maximum is 70 months. No

correlation was identified between N/C ratio of AID and PFS in our cohort. HR= 0.92, 95% CI for HR= 0.78, 1.08 and *P*= 0.311 (Age-adjusted cox proportional hazards model).

Out of the 20 patients, only three have experienced an OS event so far. No correlation was found between N/C ratio and OS in our cohort. HR= 0.93, 95% CI for HR= 0.78, 1.1 and *P*= 0.385 (Age-adjusted cox proportional hazards model).

Box-plots and logistic regression were used to explore the effect of the N/C ratio of AID on RR. Out of the 20 patients, the response status was missing for one patient. Among the rest, five achieved a CR and 14 achieved a PR. Hence, our analysis compared patients who achieved a CR with those who did not.

Although not statistically significant, these was a trend for patients who achieved a CR to a lower AID N/C ratio compared to those who did not achieve a CR. HR= 0.86, 95% CI for HR= 0.7, 1.02 and *P*= 0.11 (Age-adjusted logistic regression model).

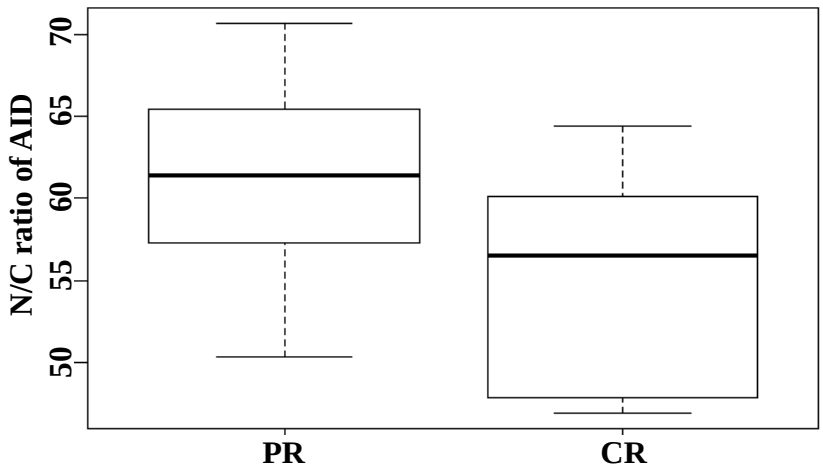

Figure 7.1: Lower N/C ratio of AID is associated with better response rate.

7.4.3 Relationship between overall cumulative *IGHV* mutation and trial endpoints:

In this study, the total *IGHV* mutation, AID induced mutation and *IGHV* ongoing mutation were examined in 18 cases of FL. Logistic regression was used to explore the effect of ongoing *IGHV* mutation and AID-induced mutations on RR after induction treatment.

It was found that there was a trend that was close to statistical significance between mutation load and superior PFS (HR= 0.9, 95% CI for HR= 0.81, 1.01 and *P*= 0.063).

As was the case with PFS, here there was a statistically significant association between mutation load and superior OS (HR= 0.87, 95% CI for HR= 0.76, 0.99 and *P*= 0.044).

However, the age-adjusted logistic regression did not yield any significant association at 5% level between RR with the total *IGHV* mutations: HR= 0.99, 95% CI for HR= 0.89, 1.14 and *P*= 0.939.

We also explored the correlation between % of *IGHV* mutation and the trial endpoints. The % of *IGHV* mutation represents the difference from the germline. In our 18 patients, the range of % of *IGHV* mutation was between 2.89 and 19.22 with an average of 14.86 (details provided in Chapter 5, Section 5.4.4). No correlation was identified between the % of *IGHV* mutation and PFS in our cohort (HR= 0.86, 95% CI for HR= 0.69, 1.08 and *P*= 0.188).

Similar to PFS, there was no association between % of *IGHV* mutation and OS; HR= 0.83, 95% CI for HR= 0.65, 1.07 and *P*= 0.145.

Also, there was no correlation between the percentage of *IGHV* mutation and RR (HR= 0.94, 95% CI for HR= 0.73, 1.23 and *P*= 0.593).

7.4.4 Relationship between AID-related cumulative *IGHV* mutation and trial endpoints:

Interestingly, when the percentage of AID-related *IGHV* mutations was related to PFS, a trend for worse PFS was found in patients with higher percentage of AID-related mutations (HR= 1.22, 95% CI for HR= 0.97, 1.53 and *P*= 0.086).

However, there was no association between OS and the percentage of *IGHV* mutations due to AID (HR= 1.64, 95% CI for HR= 0.8, 3.36 and *P*= 0.173).

Also, there was no correlation between the percentage of AID induced mutations among the total mutations and the RR: HR= 0.93, 95% CI for HR= 0.72, 1.13 and *P*= 0.484.

7.4.5 Relationship between AID-unrelated cumulative *IGHV* mutation and trial endpoints:

In contrast to results of correlation between percentage of AID related mutations and PFS, a higher percentage of non-AID-induced *IGHV* mutations was associated with a better PFS, although this association did not quite reach statistical significance (HR= 0.081, 95% CI for HR= 0.63, 1.04 and *P*= 0.086). Cox proportional hazards were used, and age was adjusted in the analysis.

Similarly, there was no correlation between the percentage of AID unrelated mutations among the total mutations and the RR: HR= 1.08, 95% CI for HR= 0.88, 1.39 and *P*= 0.484.

7.4.6. Relationship between ongoing *IGHV* mutation and trial endpoints:

Ongoing mutation of the *IGHV* gene is the process of continuing mutation as indicated by intraclonal variation of the Ig sequence (398). The strategy to identify ongoing mutations was discussed in Chapter 5, Section 5.3.5. This *IGHV* ongoing mutation has previously been used by others as a read out of AID function (88, 356). In this thesis, the ongoing mutation and the accumulation of AID-induced mutations was examined in 18 cases of FL. Among

these patients, only five have experienced a PFS event at the time of analysis (two progressions and three deaths). The median follow-up time is 29 months, while the maximum is 67 months. After age adjustment, no association was identified between ongoing *IGHV* mutation and PFS, HR= 1.02, 95% CI for HR= 0.83, 1.27 and *P*= 0.841.

Among the 18 patients studied for ongoing *IGHV* mutation, only four have experienced an OS event so far. The age-adjusted analysis revealed no correlation between OS and ongoing mutation in *IGHV* (HR= 1.08, 95% CI for HR= 0.86, 1.35 and *P*= 0.51).

Out of the 18 patients, response data is missing for one patient, four achieved a CR, 11 achieved a PR, one had SDS, and one had PD. Hence, our analysis compared patients who achieved CR with those who did not.

The age-adjusted logistic regression did not yield any significant association between RR and ongoing mutation: HR= 0.93, 95% CI for HR= 0.61, 1.21 and *P*= 0.641.

7.4.7 Relationship between genes expressed in the TME and trial endpoints:

As discussed in Chapter 6, the TME plays an important role in pathogenesis and progression of FL (94). Three genes expressed in TME were found to be associated with clinical outcome in a previous study; these were (*CD2*, *LEF1* and *STAT4*), and another three genes were associated with bad prognosis (*LGMN*, *C3AR1* and *MAPK1*) (95).

Higher expression of *STAT4* and *LEF1* in patients with lower expression of AID protein was previously reported in this thesis (Chapter 6, Section 6.3.2). However, no association between these genes and baseline clinical variables was found.

The age-adjusted cox proportional hazards analysis yields a significantly higher hazard for progression among cases with higher *LGMN* gene expression. HR= 1.15, 95% CI for HR= 1.02, 1.29 and *P*= 0.019. No correlation was observed between any of the rest of the five genes and PFS (Table 7.4).

	Variable	HR	95% CI for HR	*P* value
Good prognosis genes	***STAT4***	0.98	0.92, 1.04	0.549
	LEF1	0.95	0.85, 1.06	0.385
	CD2	0.99	0.98, 1.01	0.721
Bad prognosis genes	***C3AR1***	0.99	0.86, 1.14	0.911
	LGMN	1.15	1.02, 1.29	0.019
	MAPK1	0.96	0.81, 1.13	0.600

Table 7.4: A significantly higher hazard for disease progression in patients with higher expression of *LGMN* gene. Cox proportional hazards test was used.

Using the cox proportional hazards model to explore the effect of gene expression on OS, the age-adjusted model did not yield any significant correlation. However, a trend of higher hazard of death was observed in patients with higher *LGMN* expression. However, the *P* value is not statistically significant (Table 7.5).

	Variable	HR	95% CI for HR	*P* value
Good prognosis genes	***STAT4***	1.01	0.92, 1.09	0.987
	LEF1	1.02	0.89, 1.17	0.741
	CD2	1.01	0.98, 1.03	0.743
Bad prognosis genes	***C3AR1***	0.94	0.72, 1.22	0.628
	LGMN	1.16	0.98, 1.38	0.093
	MAPK1	0.96	0.73, 1.28	0.799

Table 7.5: No correlation was observed between any of the genes expressed in TME of FL and OS. Cox proportional hazards were used in all tests.

No correlation between any of the genes expressed in TME of FL and RR after induction treatment (just after or before cycle 8) was observed (Table 7.6). Logistic regression was used to explore the effect of those variables on RR.

	Variable	OR	95% CI for OR	*P* value
Good prognosis genes	***STAT4***	1.01	0.97, 1.07	0.527
	LEF1	0.99	0.91, 1.08	0.980
	CD2	1.01	0.99, 1.02	0.563
Bad prognosis genes	***C3AR1***	1.02	0.90, 1.16	0.693
	LGMN	0.92	0.75, 1.06	0.317
	MAPK1	0.99	0.98, 1.01	0.655

Table 7.6: No association was observed between any of the genes expressed in TME of FL and RR. Logistic regression was used in all tests.

A summary of all associations between clinical features at diagnosis and follow up with our laboratory data can be found in Figure 7.2 below.

Clinical feature	High level of AID expression		AID functions			
	AID protein (n=87)	*AID* mRNA (n=59)	Big AID N/C ratio (n=20)	More total *IGHV* mutation (n=18)	High % of AID-related *IGHV* mutation (n=18)	More ongoing *IGHV* mutation (n=18)
High Grade						
Advanced Stage						
Low RR						
Short PFS				*		
Short OS				*		

	Statistically significant association	$P < 0.05$
	Non-significant trend	P 0.05 – 0.11
	No association or trend	$P > 0.11$

Figure 7.2: Summarises all correlations between AID expression and function with the clinical features and outcome in the cohort of FL patients studied in this thesis. The significant associations and non-significant trends correlates high AID function and AID related *IGHV* mutations with poor disease outcome; Patients with higher total *IGHV* mutations were associated with a better PFS and OS. * Negative correlation observed only between total *IGHV* mutation and shorter PFS or OS.

7.5 Discussion:

The aim of this chapter was to explore correlations of our laboratory data with clinical outcomes of our cohort of patients. Our main hypothesis was to identify the role of AID in FL. We hypothesised that AID expression at the genomic or protein level, or the functional readout of AID (N/C ratio of AID protein, ongoing *IGHV* mutation and AID-induced *IGHV* mutations) are associated with genomic instability, aberrant SHM and disease progression in FL. With the exception of a two reports that did not find a correlation between overall AID protein expression by IHC and prognosis, no previous studies could be found in the literature that had tested this hypothesis in such an FL setting (237, 354). In one of these papers, it was found that AID expression (both mRNA and protein) is correlated with poor prognosis in patients with DLBCL but not FL treated with CHOP-based chemotherapy (237).

In previous chapters, we did not identify any correlations between AID expression and clinical variables at the point of trial entry. However, we did find correlations between such clinical variables and functional readouts of AID activity. For example, a tendency towards a higher N/C AID ratio in patients with advanced anatomical stage was noted, and a significant association between higher histological grade and a suggestion of a trend of more advanced stage with more ongoing *IGHV* mutation will probably become evident if more samples are used in the future. In this chapter, we also did not identify a correlation between AID expression both at mRNA and protein levels and any of the longitudinal trial endpoints examined. This was in agreement with the two studies of AID correlation with prognosis mentioned earlier in the discussion (237, 354)

Regarding the function of AID, we did not identify a correlation between ongoing *IGHV* mutation and clinical outcome. However, there was a trend for a lower N/C AID ratio in patients who achieved a CR compared to those who did not. This is in keeping with the idea

that AID functions predominantly in the nucleus, and supports our hypothesis that high AID activity has a deleterious effect on outcome in FL.

In support of this conclusion, we found a trend for worse PFS in patients with higher levels of AID-related *IGHV* mutation. In contrast, we found a trend for better PFS in patients with higher levels of non-AID-related *IGHV* mutation.

In conclusion, the findings of this chapter are consistent with the other chapters, i.e. that the functional readout of AID, rather than its expression level, is more important in predicting the clinical features and outcomes of FL patients.

Chapter 8

General discussion

AID expression has been found to be responsible for chromosomal translocations that are involved in the pathogenesis of B-cell lymphomas (198). In addition, AID has also been correlated with the pathogenesis, progression and relapse in many B-cell and non B-cell tumours (37, 236-240).

However, to our knowledge, no comprehensive study has so far examined the expression of AID at different levels in FL patients and correlated this to ongoing mutations, clinical features at diagnosis and clinical outcome, particularly when the only sample type available in most oncology centres is FFPE LN tissues that are taken at the time of diagnosis.

So, in my thesis, the first aim was to identify and optimize methods of assessing AID expression in terms of mRNA, total protein and cellular localization of AID in B-cells in FFPE tissues from FL patients. To do this, I used pre-treatment FFPE biopsy samples obtained from the phase III PACIFICO trial which is comparing alternative frontline chemoimmunotherapy regimens in older patients with follicular lymphoma.

After obtaining reliable measurements of AID expression from FFPE biopsies, my next aim was to explore AID function in the form of ongoing and AID-specific mutations in the *IGHV* gene. I also examined the relationship between AID expression and the tissue microenvironment. Finally, I wanted to see whether there was an association between any of the AID-related laboratory data I obtained and the clinical data obtained from the PACIFICO trial.

I hypothesised that AID expression and/or function is associated with genomic instability and progression or transformation of FL. I also predicted that AID function in the form of ongoing mutation would be associated with continuing subclonal evolution and disease relapse in FL. If my hypothesis was proved to be correct, AID could potentially serve as a predictive or prognostic biomarker for disease stratification in FL, as well as a therapeutic target.

I spent a considerable amount of time optimizing RNA and DNA extractions from FFPE tissue sections. Because the test samples were obtained from a clinical trial, these tissues were precious and only a limited amount of material was available.

Initially, I compared two commonly used RNA extraction kits, and found that the RNeasy FFPE kit from Qiagen was more suitable. I then optimized RNA extraction methods to obtain higher quantity from less tissue material (Chapter 3, Section 3.2.1).

With respect to DNA extraction from FFPE tissues, I optimized the commercially available kits to produce a larger amount of DNA (Chapter 3, Section 3.2.2.1).

I measured RNA quantity by using a Qubit fluorometer and quality with a bioanalyzer 2100 in all of our cohort of 87 patients. In addition, the DNA quantity (also by Qubit fluorometer) and quality was measured by designing primers for PCR amplification of 200-400 bp fragments of the *GAPDH* gene in all samples.

Although my endeavours to optimise of nucleic acid extraction from FFPE biopsies were primarily aimed at investigating AID expression in FL, their relevance extends to all conditions where characterisation of the expression of specific genes is required using FFPE biopsy material.

qPCR was then chosen to estimate *AID* mRNA expression because of its simplicity, reliability, specificity and cost effectiveness for estimation of gene expression (399). It was found that degraded RNA with a RIN of < 2.1 was not suitable for reliable qPCR results,

and this RIN number was therefore used as a cut-off for excluding RNA samples from this study (Chapter 4, Section 4.2.1.3). I then compared different EC genes and used the best of them as a control in qPCR experiments.

After optimization and validation of all steps before and during qPCR, *AID* mRNA expression was quantified in 59 (with RIN of 2.1 or more) out of 87 samples (68%) from the PACIFICO trial.

Subsequently, protein expression of AID in FL was determined using immunohistochemistry in FFPE tissue samples. This was performed by staining and scoring all of the 87 FL samples. A good correlation was observed between *AID* mRNA and protein levels (Chapter 4, Section 4.3). This provided cross-validation of the 2 techniques and increased the veracity of the findings.

I then sought to establish which cells in the GC were expressing AID. To do this, I performed IF to examine co-expression of AID with CD3 (T cells), CD68 (macrophages) and CD20 (B-cells). Essentially, AID expression was confined to B-cells.

Because AID is stored in the cytoplasm and performs its function in the nucleus (400), I postulated that the function of AID could be better assessed if a method of exploring the N/C ratio of AID could be established. To do this, confocal microscopy was used to quantify the co-expression of AID with a nuclear stain (Chapter 4, Section 4.4.2). A positive correlation was identified between this N/C ratio with both mRNA and protein expression levels of AID.

I then related AID expression to clinical features at trial entry. No statistically significant correlation was found between AID mRNA or protein levels and any of the clinical features recorded. However, a trend for more advanced-stage disease was found in patients with a higher AID N/C ratio.

Because ongoing *IGHV* mutational activity has been associated with the acquisition and loss of genetic material, intraclonal heterogeneity and drug resistance in FL (357, 358), I studied

ongoing mutations and AID-induced mutations in this gene in my cohort of patient samples. I found a statistically significant association between higher rates of ongoing *IGHV* mutation and unfavourable histological grade; there was also a non-significant trend for cases with higher rates of ongoing *IGHV* mutation to have more advanced disease (Chapter 5, Section 5.4.10.3).

In addition, I examined whether there was an association between AID expression and/or function and the expression of genes associated with the FL tissue microenvironment (TME) that are known to correlate with patient survival. It was found that higher AID protein expression was associated with reduced expression of 2 out of 3 genes that are linked with better survival in FL.

Finally, when I compared my laboratory data to longitudinal clinical data from the trial, it became apparent that a higher % of AID-related mutations in the *IGHV* gene was associated with worse survival (Chapter 7, Section 7.4.4).

Probably one of the most important aspects achieved in this work is the optimization of protocols to extract larger amounts of nucleic acids from FFPE tissues. Another key achievement is the identification of AID expression at the mRNA and protein level including its subcellular localisation.

I also established a positive correlation between AID mRNA and protein expression. However, one of the most important findings is that functional readouts of AID (N/C ratio, ongoing *IGHV* mutation and AID-related *IGHV* mutation) correlate much more closely than AID expression with adverse baseline and longitudinal clinical features.

On the other hand, I faced several difficulties during the Ph.D. For example, despite optimising the method for extracting nucleic acids from FFPE tissues, I could not reverse the degradation of those substances that took place prior to nucleic acid extraction. Although

it was not available for my project, fresh frozen patient material is likely to have produced a much better yield of intact DNA and RNA.

The other difficulty I encountered was using the Image j1x software to measure intra-nuclear AID. Specifically, I found that cells from FFPE tissue sections were very close to each other, and it was therefore difficult to identify individual cell nuclei.

Additionally, regarding the correlation between TME and AID expression, I found that many updated published works noted that it was not simply the expression level of genes that played a role in disease outcome but rather alterations in the number, type or location of TME cells (91).

Due to time limits for this study, further future work may be required to confirm these results and build solid conclusions.

Finding that the functional readout of AID, rather than its expression, is correlated to adverse clinical features in FL opens the door for further work. An increase in the size of my patient cohort to confirm these results is required, because *IGHV* mutation was studied in only 18 patients and AID N/C ratio measured in only 20 patients. Furthermore, the discrepancy between AID expression and function calls for studies to investigate functionally important posttranslational modifications of AID.

Also, in our clonal analysis of *IGHV* families, I found that in about half of the patients there was a faint band of IG in addition to the main band. We examined this faint band in one patient only and found that these were background B-cells. These faint bands require further examination, as they appeared not in one or 2 patients but in approximately half of our cohort.

It would also be of interest to compare the impact of AID expression and function on the outcome of the 2 treatment arms of the PACIFICO trial. This was not possible within the framework of my Ph.D. project as patients are still being followed-up for disease progression

and the number of cases examined would be too small to draw any meaningful conclusions. However, if a larger number of patient samples could be examined for AID expression and function and the results related to therapy response, PFS and OS when the final trial data are available, it should be possible to address the question of whether the effect of AID expression and function on FL outcome is differentially modulated by specific drug regimens. If we are able to perform this future work and confirm the correlation of higher AID function with worse disease outcome, it may be possible to explore AID function as a potential biomarker for determining disease outcome in FL. Completion of this work may also provide us with a better understanding of the role of AID in the biology and progression of FL.

Another line of investigation that would be interesting to pursue would be to correlate AID expression and function with mutations in genes other than *IGHV* that are important in FL biology. This work was planned in collaboration with Professor Jude Fitzgibbon at Bart's cancer institute but was delayed owing to a strategy to expand the NGS panel as much as possible before applying it to the trial samples.

The methodology for this work will be done as follows, DNA samples from FFPE tissues of the PACIFICO trial were prepared for performing NGS experiments, in collaboration with Bart's Cancer Institute's lymphoma research group led by Professor Jude Fitzgibbon; however, the time limit for this Ph.D. did not allow me to perform this experiment, and it will be accomplished as a future work.

To select our panal of genes, a total of 161 genes (Appendix Table 3) that are frequently mutated in FL and DLBCL had been selected for NGS mutation analysis (127, 401-415). Chosen genes were based upon importance of pathophysiology of FL formation and transformation in addition to mutation frequency. Agilent SureDesign software was used to design 28,261 probes that cover a genome size of 1.067 Mb.

To perform target sequencing, a capture-based approach for NGS will be performed using a SureSelect XT Target Enrichment System (Agilent Technologies, UK). A total of 0.5-1 µg of DNA is needed, physical shearing of DNA by sonication to a size range of 150-200bp will be done using a Covaris E210 instrument (Covaris Inc); this is followed by blunting gene ends, addition of adenine overhang, ligation to adapters and PCR. Biotinylated DNA probes designed to cover exons of our genes of interest will be included in the hybridization solution. After post-capture enrichment, libraries will be sequenced using the MiSeq platform at a coverage of 500-1000X.

I wanted to compare the two groups of patients with high and low AID function against the clinical outcome, but unfortunately the number of patients with PFS or OS events were very low. This will be the future goal of a further study.

Finally, it will be important to confirm an association between ongoing *IGHV* mutation and adverse clinical outcome in the PACIFICO trial as I have so far only been able to show a non-significant trend.

Conclusion

In this Ph.D., I have carried out comprehensive work to explore AID expression by mRNA, IHC and IF that will be useful for researchers in future studies exploring AID in any FFPE tissues. Moreover, I improved the protocols of commercially-available kits to be able extract the higher yield and better quality of RNA and DNA from FFPE tissues. I demonstrated that AID mRNA, total protein and AID nuclear expression levels were correlated with each other. More importantly, I found that functional readouts of AID (rather than AID expression per se) correlate with clinical features. Specifically, the N/C ratio of AID and ongoing *IGHV* mutation were associated with adverse histological grades and more advanced anatomical stages of disease. Additionally, I found a trend for worse longitudinal outcome in patients with a higher N/C ratio of AID or higher % of AID-related *IGHV* mutations, although a larger cohort of patients is required to validate these results. These findings could be particularly important for clinicians since AID is mutagenic and may be associated with clonal evolution and FL transformation.

References

1. Stathis A, Owens C. Risk factors, etiology, and pathogenesis. In: Younes A, editor. Handbook of Lymphoma. Cham: Springer International Publishing; 2016. p. 11-20.

2. Bonifer C, Cockerill PN. Transcriptional and Epigenetic Mechanisms Regulating Normal and Aberrant Blood Cell Development Preface. Transcriptional and Epigenetic Mechanisms Regulating Normal and Aberrant Blood Cell Development. 2014:V-Vii.

3. Shapiro-Shelef M, Calame K. Regulation of plasma-cell development. Nat Rev Immunol. 2005;5(3):230-42.

4. Laiosa CV, Stadtfeld M, Graf T. Determinants of lymphoid-myeloid lineage diversification. Annu Rev Immunol. 2006;24:705-38.

5. Pillai S, Baltimore D. Formation of Disulfide-Linked Mu-2-Omega-2 Tetramers in Pre-B Cells by the 18k Omega-Immunoglobulin Light Chain. Nature. 1987;329(6135):172-4.

6. Elgert KD. Immunology: Understanding The Immune System: Wiley; 2009.

7. Parham P. The Immune System, Fourth Edition: Taylor & Francis Group; 2014.

8. Kumar R, DiMenna LJ, Chaudhuri J, Evans T. Biological function of activation-induced cytidine deaminase (AID). Biomed J. 2014;37(5):269-83.

9. Teng G, Papavasiliou EN. Immunoglobulin somatic hypermutation. Annual Review of Genetics. 2007;41:107-20.

10. Magdelaine-Beuzelin C, Kaas Q, Wehbi V, Ohresser M, Jefferis R, Lefranc MP, et al. Structure-function relationships of the variable domains of monoclonal antibodies approved for cancer treatment. Crit Rev Oncol Hematol. 2007;64(3):210-25.

11. van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL, et al. Design and standardization of PCR primers and protocols for detection of clonal

immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia. 2003;17(12):2257-317.

12. Krebs JE, Lewin B, Goldstein ES, Kilpatrick ST, Lewin B. Lewin's essential genes. 3rd ed. Burlington, MA: Jones and Bartlett Publishers; 2013. xxi, 847 p. p.

13. Gagyi É. Somatic hypermutation of IgVH genes and aberrant somatic hypermutation in follicular lymphoma without bcl2 gene rearrangement and expression [PhD thesis]. Budapest: Semmelweis University; 2012

14. Gorman JR, Alt FW. Regulation of immunoglobulin light chain isotype expression. Advances in Immunology, Vol 69. 1998;69:113-81.

15. Maeda K, Singh SK, Eda K, Kitabatake M, Pham P, Goodman MF, et al. GANP-mediated recruitment of activation-induced cytidine deaminase to cell nuclei and to immunoglobulin variable region DNA. J Biol Chem. 2010;285(31):23945-53.

16. Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, et al. Activation-Induced Cytidine Deaminase Targets DNA at Sites of RNA Polymerase II Stalling by Interaction with Spt5. Cell. 2010;143(1):122-33.

17. Chaudhuri J, Alt FW. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. Nat Rev Immunol. 2004;4(7):541-52.

18. Hartmann S, Helling A, Doring C, Renne C, Hansmann ML. Clonality testing of malignant lymphomas with the BIOMED-2 primers in a large cohort of 1969 primary and consultant biopsies. Pathol Res Pract. 2013;209(8):495-502.

19. Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. Annu Rev Immunol. 2008;26:261-92.

20. Kenter AL, Kumar S, Wuerffel R, Grigera F. AID hits the jackpot when missing the target. Current Opinion in Immunology. 2016;39:96-102.

21. Xu Z, Fulop Z, Wu G, Pone EJ, Zhang J, Mai T, et al. 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. Nat Struct Mol Biol. 2010;17(9):1124-35.

22. Stavnezer J. Complex regulation and function of activation-induced cytidine deaminase. Trends Immunol. 2011;32(5):194-201.

23. McHeyzer-Williams LJ, Driver DJ, McHeyzer-Williams MG. Germinal center reaction. Curr Opin Hematol. 2001;8(1):52-9.

24. Tomura M, Itoh K, Kanagawa O. Naive CD4(+) T Lymphocytes Circulate through Lymphoid Organs To Interact with Endogenous Antigens and Upregulate Their Function. J Immunol. 2010;184(9):4646-53.

25. Mesin L, Ersching J, Victora GD. Germinal Center B Cell Dynamics. Immunity. 2016;45(3):471-82.

26. Coupland SE. Molecular pathology of lymphoma. Eye (Lond). 2013;27(2):180-9.

27. Natkunam Y. The biology of the germinal center. Hematology Am Soc Hematol Educ Program. 2007:210-5.

28. Allen CD, Okada T, Cyster JG. Germinal-center organization and cellular dynamics. Immunity. 2007;27(2):190-202.

29. Heesters BA, Myers RC, Carroll MC. Follicular dendritic cells: dynamic antigen libraries. Nat Rev Immunol. 2014;14(7):495-504.

30. Heesters BA, Chatterjee P, Kim YA, Gonzalez SF, Kuligowski MP, Kirchhausen T, et al. Endocytosis and recycling of immune complexes by follicular dendritic cells enhances B cell antigen binding and activation. Immunity. 2013;38(6):1164-75.

31. Garin A, Meyer-Hermann M, Contie M, Figge MT, Buatois V, Gunzer M, et al. Toll-like Receptor 4 Signaling by Follicular Dendritic Cells Is Pivotal for Germinal Center Onset and Affinity Maturation. Immunity. 2010;33(1):84-95.

32. Allen CD, Okada T, Tang HL, Cyster JG. Imaging of germinal center selection events during affinity maturation. Science. 2007;315(5811):528-31.

33. Gitlin AD, Mayer CT, Oliveira TY, Shulman Z, Jones MJ, Koren A, et al. HUMORAL IMMUNITY. T cell help controls the speed of the cell cycle in germinal center B cells. Science. 2015;349(6248):643-6.

34. Gitlin AD, Shulman Z, Nussenzweig MC. Clonal selection in the germinal centre by regulated proliferation and hypermutation. Nature. 2014;509(7502):637-40.

35. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell. 2000;102(5):553-63.

36. Victora GD, Dominguez-Sola D, Holmes AB, Deroubaix S, Dalla-Favera R, Nussenzweig MC. Identification of human germinal center light and dark zone cells and their relationship to human B-cell lymphomas. Blood. 2012;120(11):2240-8.

37. Robbiani DF, Deroubaix S, Feldhahn N, Oliveira TY, Callen E, Wang Q, et al. Plasmodium Infection Promotes Genomic Instability and AID-Dependent B Cell Lymphoma. Cell. 2015;162(4):727-37.

38. Basso K, Dalla-Favera R. Roles of BCL6 in normal and transformed germinal center B cells. Immunol Rev. 2012;247(1):172-83.

39. Ranuncolo SM, Polo JM, Dierov J, Singer M, Kuo T, Greally J, et al. Bcl-6 mediates the germinal center B cell phenotype and lymphomagenesis through transcriptional repression of the DNA-damage sensor ATR. Nat Immunol. 2007;8(7):705-14.

40. Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM. The generation of antibody-secreting plasma cells. Nat Rev Immunol. 2015;15(3):160-71.

41. Swerdlow SH, Cancer IAfRo, Organization WH. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues: International Agency for Research on Cancer; 2008.

42. Shankland KR, Armitage JO, Hancock BW. Non-Hodgkin lymphoma. Lancet. 2012;380(9844):848-57.

43. UK Cr. Non-Hodgkin lymphoma incidence statistics. Cancer research UK website. 2014; http://www.cancerresearchuk.org/cancer-info/cancerstats/types/nhl/incidence/uk-nonhodgkin-lymphoma-incidence-statistics#world.

44. Smith A, Crouch S, Lax S, Li J, Painter D, Howell D, et al. Lymphoma incidence, survival and prevalence 2004-2014: sub-type analyses from the UK's Haematological Malignancy Research Network. British Journal of Cancer. 2015;112(9):1575-84.

45. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood. 2016;127(20):2375-90.

46. Maartense E, Le Cessie S, Kluin-Nelemans HC, Kluin PM, Snijder S, Wijermans PW, et al. Age-related differences among patients with follicular lymphoma and the importance of prognostic scoring systems: analysis from a population-based non-Hodgkin's lymphoma registry. Ann Oncol. 2002;13(8):1275-84.

47. Kane EV, Roman E, Becker N, Bernstein L, Boffetta P, Bracci PM, et al. Menstrual and reproductive factors, and hormonal contraception use: associations with non-Hodgkin lymphoma in a pooled analysis of InterLymph case-control studies. Ann Oncol. 2012;23(9):2362-74.

48. Nabhan C, Zhou XL, Day BM, Dawson K, Zelenetz AD, Friedberg JW, et al. Disease, treatment, and outcome differences between men and women with follicular lymphoma in the United States. American Journal of Hematology. 2016;91(8):770-5.

49. Muller AM, Ihorst G, Mertelsmann R, Engelhardt M. Epidemiology of non-Hodgkin's lymphoma (NHL): trends, geographic distribution, and etiology. Ann Hematol. 2005;84(1):1-12.

50. Farrell RJ, Ang Y, Kileen P, O'Briain DS, Kelleher D, Keeling PW, et al. Increased incidence of non-Hodgkin's lymphoma in inflammatory bowel disease patients on immunosuppressive therapy but overall risk is low. Gut. 2000;47(4):514-9.

51. Morton LM, Hartge P, Holford TR, Holly EA, Chiu BC, Vineis P, et al. Cigarette smoking and risk of non-Hodgkin lymphoma: a pooled analysis from the International Lymphoma Epidemiology Consortium (interlymph). Cancer Epidemiol Biomarkers Prev. 2005;14(4):925-33.

52. Lyons SF, Liebowitz DN. The roles of human viruses in the pathogenesis of lymphoma. Semin Oncol. 1998;25(4):461-75.

53. Ambinder AJ, Shenoy PJ, Malik N, Maggioncalda A, Nastoupil LJ, Flowers CR. Exploring risk factors for follicular lymphoma. Adv Hematol. 2012;2012:626035.

54. Zhang Y, Sanjose SD, Bracci PM, Morton LM, Wang R, Brennan P, et al. Personal use of hair dye and the risk of certain subtypes of non-Hodgkin lymphoma. Am J Epidemiol. 2008;167(11):1321-31.

55. Chiu BC, Blair A. Pesticides, chromosomal aberrations, and non-Hodgkin's lymphoma. J Agromedicine. 2009;14(2):250-5.

56. Cocco P, t'Mannetje A, Fadda D, Melis M, Becker N, de Sanjose S, et al. Occupational exposure to solvents and risk of lymphoma subtypes: results from the Epilymph case-control study. Occup Environ Med. 2010;67(5):341-7.

57. Mamessier E, Broussais-Guillaumot F, Chetaille B, Bouabdallah R, Xerri L, Jaffe ES, et al. Nature and importance of follicular lymphoma precursors. Haematologica. 2014;99(5):802-10.

58. Leich E, Hoster E, Wartenberg M, Unterhalt M, Siebert R, Koch K, et al. Similar clinical features in follicular lymphomas with and without breaks in the BCL2 locus. Leukemia. 2016;30(4):854-60.

59. Klapper W, Klein U, De Silva NS, Zha S, Crowe JL, Dai B, et al. Malignant Lymphomas: Biology and Molecular Pathogenesis: De Gruyter; 2016.

60. Horsman DE, Connors JM, Pantzar T, Gascoyne RD. Analysis of secondary chromosomal alterations in 165 cases of follicular lymphoma with t(14;18). Genes Chromosomes Cancer. 2001;30(4):375-82.

61. Tsai AG, Lu HH, Raghavan SC, Muschen M, Hsieh CL, Lieber MR. Human Chromosomal Translocations at CpG Sites and a Theoretical Basis for Their Lineage and Stage Specificity. Cell. 2008;135(6):1130-42.

62. Robbiani DF, Bunting S, Feldhahn N, Bothmer A, Camps J, Deroubaix S, et al. AID Produces DNA Double-Strand Breaks in Non-Ig Genes and Mature B Cell Lymphomas with Reciprocal Chromosome Translocations. Molecular Cell. 2009;36(4):631-41.

63. Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH, et al. Two levels of protection for the B cell genome during somatic hypermutation. Nature. 2008;451(7180):841-U11.

64. Weinberg OK, Ai WZ, Mariappan MR, Shum C, Levy R, Arber DA. "Minor" BCL2 breakpoints in follicular lymphoma: frequency and correlation with grade and disease presentation in 236 cases. J Mol Diagn. 2007;9(4):530-7.

65. Cleary ML, Galili N, Sklar J. Detection of a second t(14;18) breakpoint cluster region in human follicular lymphomas. J Exp Med. 1986;164(1):315-20.

66. Albinger-Hegyi A, Hochreutener B, Abdou MT, Hegyi I, Dours-Zimmermann MT, Kurrer MO, et al. High frequency of t(14;18)-translocation breakpoints outside of major breakpoint and minor cluster regions in follicular lymphomas: improved polymerase chain reaction protocols for their detection. Am J Pathol. 2002;160(3):823-32.

67. Buchonnet G, Lenain P, Ruminy P, Lepretre S, Stamatoullas A, Parmentier F, et al. Characterisation of BCL2-JH rearrangements in follicular lymphoma: PCR detection of 3 ' BCL2 breakpoints and evidence of a new cluster. Leukemia. 2000;14(9):1563-9.

68. Hillion J, Mecucci C, Aventin A, Leroux D, Wlodarska I, Van Den Berghe H, et al. A variant translocation t(2;18) in follicular lymphoma involves the 5' end of bcl-2 and Ig kappa light chain gene. Oncogene. 1991;6(1):169-72.

69. Roulland S, Kelly RS, Morgado E, Sungalee S, Solal-Celigny P, Colombat P, et al. t(14;18) Translocation: A predictive blood biomarker for follicular lymphoma. J Clin Oncol. 2014;32(13):1347-55.

70. Okosun J, Montoto S, Fitzgibbon J. The routes for transformation of follicular lymphoma. Curr Opin Hematol. 2016;23(4):385-91.

71. Casulo C, Burack WR, Friedberg JW. Transformed follicular non-Hodgkin lymphoma. Blood. 2015;125(1):40-7.

72. Iqbal J, Naushad H, Bi C, Yu J, Bouska A, Rohr J, et al. Genomic signatures in B-cell lymphoma: How can these improve precision in diagnosis and inform prognosis? Blood Rev. 2016;30(2):73-88.

73. Kahl BS, Yang DT. Follicular lymphoma: evolving therapeutic strategies. Blood. 2016;127(17):2055-63.

74. Okosun J, Wolfson RL, Wang J, Araf S, Wilkins L, Castellano BM, et al. Recurrent mTORC1-activating RRAGC mutations in follicular lymphoma (vol 48, pg 183, 2016). Nature Genetics. 2016;48(6):700-.

75. Nadel B. Early Steps of Follicular Lymphoma Pathogenesis. Blood. 2016;128(22):SCI-5-SCI-.

76. Adam P, Schoof J, Hartmann M, Schwarz S, Puppe B, Ott M, et al. Cell migration patterns and ongoing somatic mutations in the progression of follicular lymphoma. Cytogenet Genome Res. 2007;118(2-4):328-36.

77. Dogan A, Bagdi E, Munson P, Isaacson PG. CD10 and BCL-6 expression in paraffin sections of normal lymphoid tissue and B-cell lymphomas. Am J Surg Pathol. 2000;24(6):846-52.

78. Kridel R, Sehn LH, Gascoyne RD. Pathogenesis of follicular lymphoma. J Clin Invest. 2012/10/02 ed2012. p. 3424-31.

79. Roulland S, Faroudi M, Mamessier E, Sungalee S, Salles G, Nadel B. Early steps of follicular lymphoma pathogenesis. Adv Immunol. 2011;111:1-46.

80. de Jong D. Molecular pathogenesis of follicular lymphoma: a cross talk of genetic and immunologic factors. J Clin Oncol. 2005;23(26):6358-63.

81. Gaulard P, d'Agay MF, Peuchmaur M, Brousse N, Gisselbrecht C, Solal-Celigny P, et al. Expression of the bcl-2 gene product in follicular lymphoma. Am J Pathol. 1992;140(5):1089-95.

82. Robbiani DF, Nussenzweig MC. Chromosome Translocation, B Cell Lymphoma, and Activation-Induced Cytidine Deaminase. Annual Review of Pathology: Mechanisms of Disease, Vol 8. 2013;8:79-103.

83. Kishimoto W, Nishikori M. Molecular pathogenesis of follicular lymphoma. J Clin Exp Hematop. 2014;54(1):23-30.

84. Tellier J, Menard C, Roulland S, Martin N, Monvoisin C, Chasson L, et al. Human t(14;18)positive germinal center B cells: a new step in follicular lymphoma pathogenesis? Blood. 2014;123(22):3462-5.

85. Kluin PM. Origin and migration of follicular lymphoma cells. Haematologica. 2013;98(9):1331-3.

86. Linley A, Krysov S, Ponzoni M, Johnson PW, Packham G, Stevenson FK. Lectin binding to surface Ig variable regions provides a universal persistent activating signal for follicular lymphoma cells. Blood. 2015;126(16):1902-10.

87. Schneider D, Duhren-von Minden M, Alkhatib A, Setz C, van Bergen CA, Benkisser-Petersen M, et al. Lectins from opportunistic bacteria interact with acquired variable-region glycans of surface immunoglobulin in follicular lymphoma. Blood. 2015;125(21):3287-96.

88. Loeffler M, Kreuz M, Haake A, Hasenclever D, Trautmann H, Arnold C, et al. Genomic and epigenomic co-evolution in follicular lymphomas. Leukemia. 2015;29(2):456-63.

89. Carlotti E, Wrench D, Matthews J, Iqbal S, Davies A, Norton A, et al. Transformation of follicular lymphoma to diffuse large B-cell lymphoma may occur by divergent evolution from a common progenitor cell or by direct evolution from the follicular lymphoma clone. Blood. 2009;113(15):3553-7.

90. Carlotti E, Wrench D, Rosignoli G, Marzec J, Sangaralingam A, Hazanov L, et al. High Throughput Sequencing Analysis of the Immunoglobulin Heavy Chain Gene from Flow-Sorted B Cell Sub-Populations Define the Dynamics of Follicular Lymphoma Clonal Evolution. Plos One. 2015;10(9):e0134833.

91. Fowler NH, Cheah CY, Gascoyne RD, Gribben J, Neelapu SS, Ghia P, et al. Role of the tumor microenvironment in mature B-cell lymphoid malignancies. Haematologica. 2016;101(5):531-40.

92. Wu T, Dai Y. Tumor microenvironment and therapeutic response. Cancer Lett. 2016.

93. Coupland SE. The challenge of the microenvironment in B-cell lymphomas. Histopathology. 2011;58(1):69-80.

94. Scott DW, Gascoyne RD. The tumour microenvironment in B cell lymphomas. Nature Reviews Cancer. 2014;14(8):517-34.

95. Dave SS, Wright G, Tan B, Rosenwald A, Gascoyne RD, Chan WC, et al. Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. N Engl J Med. 2004;351(21):2159-69.

96. Carreras J, Lopez-Guillermo A, Roncador G, Villamor N, Colomo L, Martinez A, et al. High Numbers of Tumor-Infiltrating Programmed Cell Death 1-Positive Regulatory Lymphocytes Are Associated With Improved Overall Survival in Follicular Lymphoma. J Clin Oncol. 2009;27(9):1470-6.

97. Kiaii S, Clear AJ, Ramsay AG, Davies D, Sangaralingam A, Lee A, et al. Follicular Lymphoma Cells Induce Changes in T-Cell Gene Expression and Function: Potential Impact on Survival and Risk of Transformation. J Clin Oncol. 2013;31(21):2654-+.

98. Blaker YN, Spetalen S, Brodtkorb M, Lingjaerde OC, Beiske K, Ostenstad B, et al. The tumour microenvironment influences survival and time to transformation in follicular lymphoma in the rituximab era. Br J Haematol. 2016;175(1):102-14.

99. Chevalier N, Mueller M, Mougiakakos D, Ihorst G, Marks R, Schmitt-Graeff A, et al. Analysis of dendritic cell subpopulations in follicular lymphoma with respect to the tumor immune microenvironment. Leuk Lymphoma. 2016;57(9):2150-60.

100. Ansell SM. Targeting immune checkpoints in lymphoma. Curr Opin Hematol. 2015;22(4):337-42.

101. Salles GA. Clinical features, prognosis and treatment of follicular lymphoma. Hematology Am Soc Hematol Educ Program. 2007:216-25.

102. Provan D. Oxford handbook of clinical haematology. 3rd ed. Oxford ; New York: Oxford University Press; 2009. xli, 820 p., 16 p. of plates p.

103. Fernandez de Larrea C, Martinez-Pozo A, Mercadal S, Garcia A, Gutierrez-Garcia G, Valera A, et al. Initial features and outcome of cutaneous and non-cutaneous primary extranodal follicular lymphoma. Br J Haematol. 2011;153(3):334-40.

104. Lee JC, Hoehn D, Schecter J, Murty VV, Mansukhani MM, Alobeid B, et al. Lymphoid follicle colonization by Bcl-2(bright+)CD10(+) B-cells ("follicular lymphoma in situ") at nodal and extranodal sites can be a manifestation of follicular homing of lymphoma. Hum Pathol. 2013.

105. Nooka AK, Nabhan C, Zhou X, Taylor MD, Byrtek M, Miller TP, et al. Examination of the follicular lymphoma international prognostic index (FLIPI) in the National LymphoCare study (NLCS): a prospective US patient cohort treated predominantly in community practices. Ann Oncol. 2013;24(2):441-8.

106. Maurer MJ, Bachy E, Ghesquieres H, Ansell SM, Nowakowski GS, Thompson CA, et al. Early event status informs subsequent outcome in newly diagnosed follicular lymphoma. Am J Hematol. 2016;91(11):1096-101.

107. O'Malley DP, Fedoriw Y. Bone Marrow. In: Cheng L, Bostwick GD, editors. Essentials of Anatomic Pathology. Cham: Springer International Publishing; 2016. p. 821-68.

108. Bain BJ. A beginner's guide to blood cells. 2nd ed. ed. Malden, Mass.: Blackwell Pub.; 2004.

109. Dehghani M, Haddadi S, Vojdani R. Signs, Symptoms and Complications of Non-Hodgkin's Lymphoma According to Grade and Stage in South Iran. Asian Pac J Cancer Prev. 2015;16(8):3551-7.

110. Leventaki V, Khoury JD, Voss SD. Tumors of Lymphoid and Hematopoietic Tissues. In: Parham MD, Khoury DJ, McCarville BM, editors. Pediatric Malignancies: Pathology and Imaging. New York, NY: Springer New York; 2015. p. 103-49.

111. Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised response criteria for malignant lymphoma. J Clin Oncol. 2007;25(5):579-86.

112. Thanarajasingam G, Bennani-Baiti N, Thompson CA. PET-CT in Staging, Response Evaluation, and Surveillance of Lymphoma. Curr Treat Options Oncol. 2016;17(5):24.

113. Albano D, Patti C, La Grutta L, Agnello F, Grassedonio E, Mule A, et al. Comparison between whole-body MRI with diffusion-weighted imaging and PET/CT in staging newly diagnosed FDG-avid lymphomas. European Journal of Radiology. 2016;85(2):313-8.

114. Galia M, Albano D, Narese D, Patti C, Chianca V, Di Pietto F, et al. Whole-body MRI in patients with lymphoma: collateral findings. Radiol Med. 2016;121(10):793-800.

115. Perry C, Lerman H, Joffe E, Sarid N, Amit O, Avivi I, et al. The Value of PET/CT in Detecting Bone Marrow Involvement in Patients With Follicular Lymphoma. Medicine (Baltimore). 2016;95(9):e2910.

116. Dreyling M, Ghielmini M, Rule S, Salles G, Vitolo U, Ladetto M, et al. Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2016;27(suppl 5):v83-v90.

117. Johl A, Lengfelder E, Hiddemann W, Klapper W, German Low-grade Lymphoma Study G. Core needle biopsies and surgical excision biopsies in the diagnosis of

lymphoma-experience at the Lymph Node Registry Kiel. Ann Hematol. 2016;95(8):1281-6.

118. McNamara C, Davies J, Dyer M, Hoskin P, Illidge T, Lyttelton M, et al. Guidelines on the investigation and management of follicular lymphoma. Br J Haematol. 2012;156(4):446-67.

119. Jaffe ES HN, Vardiman JW et al. Hematopathology. Harris NL LL, and Ferry JA, editor: Elsevier.Saunders; 2011. 1058 p.

120. Xerri L, Dirnhofer S, Quintanilla-Martinez L, Sander B, Chan JK, Campo E, et al. The heterogeneity of follicular lymphomas: from early development to transformation. Virchows Arch. 2016;468(2):127-39.

121. West RB, Warnke RA, Natkunam Y. The usefulness of immunohistochemistry in the diagnosis of follicular lymphoma in bone marrow biopsy specimens (vol 117, pg 636, 2002). Am J Clin Pathol. 2002;118(1):145-.

122. Olsen B, Srkalovic G, Hall M, McPhail E. Cyclin D1 positive follicular lymphoma. Journal of Clinical Pathology. 2009;62(9):855-7.

123. Masir N, Jones M, Abdul-Rahman F, Florence CS, Mason DY. Variation in BCL2 protein expression in follicular lymphomas without t(14;18) chromosomal translocations. Pathology. 2012;44(3):228-33.

124. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics. 2016;17(6):333-51.

125. van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. Trends in Genetics. 2014;30(9):418-26.

126. Oliver GR, Hart SN, Klee EW. Bioinformatics for Clinical Next Generation Sequencing. Clinical Chemistry. 2015;61(1):124-35.

127. Fehniger TA, Krysiak K, White BS, Matlock M, Miller C, Fulton R, et al. Recurrent Somatic Genomic Alterations in Follicular NHL (FL) Revealed By Exome and Custom-Capture Next Generation Sequencing. Blood. 2015;126(23).

128. Friedman J, Guttapalli A, Ma C, Thodima V, Padmanabhan R, Kamalakaran S, et al. A 220-Gene Targeted Next-Generation Sequencing Panel for the Detection of Variants in Diffuse Large B-Cell Lymphoma, Follicular Lymphoma, and Mantle Cell Lymphoma: Application to a Cohort of 85 Formalin-Fixed Paraffin-Embedded Diffuse Large B-Cell Lymphoma Biopsies. Blood. 2015;126(23).

129. Pastore A, Jurinovic V, Kridel R, Hoster E, Staiger AM, Szczepanowski M, et al. Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. Lancet Oncology. 2015;16(9):1111-22.

130. Carbone PP, Kaplan HS, Musshoff K, Smithers DW, Tubiana M. Report of the Committee on Hodgkin's Disease Staging Classification. Cancer Res. 1971;31(11):1860-1.

131. Federico M, Vitolo U, Zinzani PL, Chisesi T, Clo V, Bellesi G, et al. Prognosis of follicular lymphoma: a predictive model based on a retrospective analysis of 987 cases. Intergruppo Italiano Linfomi. Blood. 2000;95(3):783-9.

132. Federico M, Bellei M, Marcheselli L, Luminari S, Lopez-Guillermo A, Vitolo U, et al. Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project. J Clin Oncol. 2009;27(27):4555-62.

133. Solal-Celigny P, Roy P, Colombat P, White J, Armitage JO, Arranz-Saez R, et al. Follicular lymphoma international prognostic index. Blood. 2004;104(5):1258-65.

134. van de Schans SA, Steyerberg EW, Nijziel MR, Creemers GJ, Janssen-Heijnen ML, van Spronsen DJ. Validation, revision and extension of the Follicular Lymphoma International Prognostic Index (FLIPI) in a population-based setting. Ann Oncol. 2009;20(10):1697-702.

135. Brice P, Bastion Y, Lepage E, Brousse N, Haioun C, Moreau P, et al. Comparison in low-tumor-burden follicular lymphomas between an initial no-treatment policy, prednimustine, or interferon alfa: A randomized study from the Groupe D'Etude des Lymphomes Folliculaires. J Clin Oncol. 1997;15(3):1110-7.

136. Armitage JO, Gascoyne RD, Lunning MA, Cavalli F. Non-Hodgkin lymphoma. Lancet. 2017.

137. Kridel R, Mottok A, Farinha P, Ben-Neriah S, Ennishi D, Zheng Y, et al. Cell of origin of transformed follicular lymphoma. Blood. 2015;126(18):2118-27.

138. Okosun J, Bodor C, Wang J, Araf S, Yang CY, Pan CY, et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. Nature Genetics. 2014;46(2):176-+.

139. Pasqualucci L, Khiabanian H, Fangazio M, Vasishtha M, Messina M, Holmes AB, et al. Genetics of follicular lymphoma transformation. Cell Rep. 2014;6(1):130-40.

140. Bouska A, McKeithan TW, Deffenbacher KE, Lachel C, Wright GW, Iqbal J, et al. Genome-wide copy-number analyses reveal genomic abnormalities involved in transformation of follicular lymphoma. Blood. 2014;123(11):1681-90.

141. Correia C, Schneider PA, Dai H, Dogan A, Maurer MJ, Church AK, et al. BCL2 mutations are associated with increased risk of transformation and shortened survival in follicular lymphoma. Blood. 2015;125(4):658-67.

142. Bello C, Zhang L, Naghashpour M. Follicular lymphoma: current management and future directions. Cancer Control. 2012;19(3):187-95.

143. Herold M, Haas A, Srock S, Neser S, Al-Ali KH, Neubauer A, et al. Rituximab added to first-line mitoxantrone, chlorambucil, and prednisolone chemotherapy followed by interferon maintenance prolongs survival in patients with advanced follicular lymphoma: An East German Study Group hematology and oncology study. J Clin Oncol. 2007;25(15):1986-92.

144. Salles G, Mounier N, de Guibert S, Morschhauser F, Doyen C, Rossi JF, et al. Rituximab combined with chemotherapy and interferon in follicular lymphoma patients: results of the GELA-GOELAMS FL2000 study. Blood. 2008;112(13):4824-31.

145. Marcus R, Imrie K, Solal-Celigny P, Catalano JV, Dmoszynska A, Raposo JC, et al. Phase III study of R-CVP compared with cyclophosphamide, vincristine, and prednisone alone in patients with previously untreated advanced follicular lymphoma. J Clin Oncol. 2008;26(28):4579-86.

146. Reagan PM, Friedberg JW. Follicular Lymphoma: First-Line Treatment Without Chemotherapy for Follicular Lymphoma. Current Treatment Options in Oncology. 2015;16(7).

147. Nabhan C, Byrtek M, Rai A, Dawson K, Zhou XL, Link BK, et al. Disease characteristics, treatment patterns, prognosis, outcomes and lymphoma-related mortality in elderly follicular lymphoma in the United States. British Journal of Haematology. 2015;170(1):85-95.

148. Hoskin PJ, Kirkwood AA, Popova B, Smith P, Robinson M, Gallop-Evans E, et al. 4 Gy versus 24 Gy radiotherapy for patients with indolent lymphoma (FORT): a randomised phase 3 non- inferiority trial. Lancet Oncology. 2014;15(4):457-63.

149. Wirth A, Seymour J, MacManus M, Ng A, Hicks R. Stage I-II follicular non-Hodgkin's lymphoma (FL): Impact of 18F FDG PET (PET) on disease staging and patient management. Ann Oncol. 2005;16:145-.

150. Freedman A. Follicular lymphoma: 2015 update on diagnosis and management. American Journal of Hematology. 2015;90(12):1172-8.

151. Ardeshna KM, Smith P, Norton A, Hancock BW, Hoskin PJ, MacLennan KA, et al. Long-term effect of a watch and wait policy versus immediate systemic treatment for asymptomatic advanced-stage non-Hodgkin lymphoma: a randomised controlled trial. Lancet. 2003;362(9383):516-22.

152. Ardeshna KM, Qian WD, Smith P, Braganca N, Lowry L, Patrick P, et al. Rituximab versus a watch-and-wait approach in patients with advanced- stage, asymptomatic, non-bulky follicular lymphoma: an open-label randomised phase 3 trial. Lancet Oncology. 2014;15(4):424-35.

153. Prica A, Chan K, Cheung M. Frontline rituximab monotherapy induction versus a watch and wait approach for asymptomatic advanced-stage follicular lymphoma: A cost-effectiveness analysis. Cancer. 2015;121(15):2637-45.

154. Hiddemann W, Kneba M, Dreyling M, Schmitz N, Lengfelder E, Schmits R, et al. Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. Blood. 2005;106(12):3725-32.

155. Federico M, Luminari S, Dondi A, Tucci A, Vitolo U, Rigacci L, et al. R-CVP Versus R-CHOP Versus R-FM for the Initial Treatment of Patients With Advanced-Stage Follicular Lymphoma: Results of the FOLL05 Trial Conducted by the Fondazione Italiana Linfomi. J Clin Oncol. 2013;31(12):1506-13.

156. Mondello P, Steiner N, Willenbacher W, Wasle I, Zaja F, Zambello R, et al. Bendamustine plus rituximab versus R-CHOP as first-line treatment for patients with

indolent non-Hodgkin's lymphoma: evidence from a multicenter, retrospective study. Annals of Hematology. 2016;95(7):1107-14.

157. Burke JM, van der Jagt RHC, Kahl BS, Wood P, Hawkins TE, MacDonald D, et al. Differences in Quality of Life Between Bendamustine-Rituximab and R-CHOP/R-CVP in Patients With Previously Untreated Advanced Indolent Non-Hodgkin Lymphoma or Mantle Cell Lymphoma. Clinical Lymphoma Myeloma & Leukemia. 2016;16(4):182-+.

158. Salles GA, Seymour JF, Feugier P, Offner F, Lopez-Guillermo A, Belada D, et al. Updated 6 Year Follow-Up Of The PRIMA Study Confirms The Benefit Of 2-Year Rituximab Maintenance In Follicular Lymphoma Patients Responding To Frontline Immunochemotherapy. Blood. 2013;122(21).

159. Salles G, Seymour JF, Offner F. Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial (vol 377, pg 42, 2011). Lancet. 2011;377(9772):1154-.

160. Savard M-F, Johnson NA. Risks and Benefits of Rituximab Maintenance in Elderly Patients with Advanced Follicular Lymphoma. Blood. 2016;128(22):5329-.

161. Morschhauser F, Radford J, Van Hoof A, Botto B, Rohatiner AZS, Salles G, et al. (90)Yttrium-Ibritumomab Tiuxetan Consolidation of First Remission in Advanced-Stage Follicular Non-Hodgkin Lymphoma: Updated Results After a Median Follow-Up of 7.3 Years From the International, Randomized, Phase III First-Line Indolent Trial. J Clin Oncol. 2013;31(16):1977-+.

162. Press OW, Unger JM, Rimsza LM, Friedberg JW, LeBlanc M, Czuczman MS, et al. Phase III Randomized Intergroup Trial of CHOP Plus Rituximab Compared With CHOP Chemotherapy Plus (131)Iodine-Tositumomab for Previously Untreated Follicular Non-Hodgkin Lymphoma: SWOG S0016. J Clin Oncol. 2013;31(3):314-20.

163. Rummel M, Kaiser U, Balser C, Stauch M, Brugger W, Welslau M, et al. Bendamustine plus rituximab versus fludarabine plus rituximab for patients with relapsed indolent and mantle-cell lymphomas: a multicentre, randomised, open-label, non-inferiority phase 3 trial. Lancet Oncology. 2016;17(1):57-66.

164. Sehn LH, Chua N, Mayer J, Dueck G, Trneny M, Bouabdallah K, et al. Obinutuzumab plus bendamustine versus bendamustine monotherapy in patients with rituximab-refractory indolent non-Hodgkin lymphoma (GADOLIN): a randomised, controlled, open-label, multicentre, phase 3 trial. Lancet Oncology. 2016;17(8):1081-93.

165. Grigg A, Dyer MJS, González Díaz M, Dreyling M, Rule S, Lei G, et al. Safety and efficacy of Obinutuzumab with CHOP or bendamustine in previously untreated follicular lymphoma. Haematologica. 2016.

166. Montoto S, Corradini P, Dreyling M, Ghielmini M, Kimby E, Lopez-Guillermo A, et al. Indications for hematopoietic stem cell transplantation in patients with follicular lymphoma: a consensus project of the EBMT-Lymphoma Working Party. Haematologica. 2013;98(7):1014-21.

167. Gopal AK, Kahl BS, de Vos S, Wagner-Johnston ND, Schuster SJ, Jurczak WJ, et al. PI3K delta Inhibition by Idelalisib in Patients with Relapsed Indolent Lymphoma. New England Journal of Medicine. 2014;370(11):1008-18.

168. Sehn LH, Goy A, Offner FC, Martinelli G, Caballero MD, Gadeberg O, et al. Randomized Phase II Trial Comparing Obinutuzumab (GA101) With Rituximab in Patients With Relapsed CD20(+) Indolent B-Cell Non-Hodgkin Lymphoma: Final Analysis of the GAUSS Study. J Clin Oncol. 2015;33(30):3467-+.

169. Westin JR, Chu FL, Zhang M, Fayad LE, Kwak LW, Fowler N, et al. Safety and activity of PD1 blockade by pidilizumab in combination with rituximab in patients with

relapsed follicular lymphoma: a single group, open-label, phase 2 trial. Lancet Oncology. 2014;15(1):69-77.

170. Czuczman MS, Leonard JP, Jung S, Johnson JL, Hsi ED, Byrd JC, et al. Phase II trial of galiximab (anti-CD80 monoclonal antibody) plus rituximab (CALGB 50402): Follicular Lymphoma International Prognostic Index (FLIPI) score is predictive of upfront immunotherapy responsiveness. Ann Oncol. 2012;23(9):2356-62.

171. Gribben JG, Fowler N, Morschhauser F. Mechanisms of Action of Lenalidomide in B-Cell Non-Hodgkin Lymphoma. J Clin Oncol. 2015;33(25):2803-U127.

172. Witzig TE, Wiernik PH, Moore T, Reeder C, Cole C, Justice G, et al. Lenalidomide Oral Monotherapy Produces Durable Responses in Relapsed or Refractory Indolent Non-Hodgkin's Lymphoma. J Clin Oncol. 2009;27(32):5404-9.

173. Ahmadi T, Chong EA, Gordon A, Aqui NA, Nasta SD, Svoboda J, et al. Combined Lenalidomide, Low-Dose Dexamethasone, and Rituximab Achieves Durable Responses in Rituximab-Resistant Indolent and Mantle Cell Lymphomas. Cancer. 2014;120(2):222-8.

174. Sorigue M, Ribera JM, Motllo C, Sancho JM. New drugs for follicular lymphoma. Leuk Res. 2016;49:38-46.

175. Miller BW, Przepiorka D, de Claro RA, Lee K, Nie L, Simpson N, et al. FDA Approval: Idelalisib Monotherapy for the Treatment of Patients with Follicular Lymphoma and Small Lymphocytic Lymphoma. Clinical Cancer Research. 2015;21(7):1525-9.

176. Ponader S, Burger JA. Bruton's Tyrosine Kinase: From X-Linked Agammaglobulinemia Toward Targeted Therapy for B-Cell Malignancies. J Clin Oncol. 2014;32(17):1830-U115.

177. Ng SY, Abramson JS. Chemotherapy-sparing treatment strategies for follicular lymphoma: current options and future directions. Curr Opin Hematol. 2016;23(4):371-6.

178. Rule S, Collins GP, Samanta K. Subcutaneous vs intravenous rituximab in patients with non-Hodgkin lymphoma: a time and motion study in the United Kingdom. Journal of Medical Economics. 2014;17(7):459-68.

179. Trotman J, Luminari S, Boussetta S. Prognostic value of PET-CT after first-line therapy in patients with follicular lymphoma: a pooled ana lysis of central scan review in three multicentre studies (vol 1, pg e17, 2014). Lancet Haematology. 2014;1(1):E16-E.

180. Kostakoglu L, Goy A, Martinelli G, Caballero D, Crump M, Gaidano G, et al. FDG-PET is prognostic and predictive for progression-free survival in relapsed follicular lymphoma: exploratory analysis of the GAUSS study. Leuk Lymphoma. 2016:1-10.

181. Shah N, Tam C, Seymour JF, Rule S. How applicable is fludarabine, cyclophosphamide and rituximab to the elderly? Leukemia Lymphoma. 2015;56(6):1599-610.

182. Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, et al. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J Biol Chem. 1999;274(26):18470-6.

183. Sabouri S, Kobayashi M, Begum NA, Xu J, Hirota K, Honjo T. C-terminal region of activation-induced cytidine deaminase (AID) is required for efficient class switch recombination and gene conversion. Proc Natl Acad Sci U S A. 2014;111(6):2253-8.

184. Salter JD, Bennett RP, Smith HC. The APOBEC Protein Family: United by Structure, Divergent in Function. Trends Biochem Sci. 2016;41(7):578-94.

185. Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, Nakata M, et al. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(7):1975-80.

186. Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM, Nussenzweig MC. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. Mol Cell. 2003;12(2):501-8.

187. Durandy A, Peron S, Taubenheim N, Fischer A. Activation-induced cytidine deaminase: structure-function relationship as based on the study of mutants. Hum Mutat. 2006;27(12):1185-91.

188. Xu Z, Zan H, Pone EJ, Mai T, Casali P. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. Nat Rev Immunol. 2012;12(7):517-31.

189. Delker RK, Fugmann SD, Papavasiliou FN. A coming-of-age story: activation-induced cytidine deaminase turns 10. Nat Immunol. 2009;10(11):1147-53.

190. Chandra V, Bortnick A, Murre C. AID targeting: old mysteries and new challenges. Trends Immunol. 2015;36(9):527-35.

191. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell. 2000;102(5):565-75.

192. Quartier P, Bustamante J, Sanal O, Plebani A, Debre M, Deville A, et al. Clinical, immunologic and genetic analysis of 29 patients with autosomal recessive hyper-IgM syndrome due to Activation-Induced Cytidine Deaminase deficiency (vol 110, pg 22, 2004). Clinical Immunology. 2004;113(2):220-.

193. Petersen-Mahrt SK, Harris RS, Neuberger MS. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. Nature. 2002;418(6893):99-103.

194. Kavli B, Otterlei M, Slupphaug G, Krokan HE. Uracil in DNA--general mutagen, but normal intermediate in acquired immunity. DNA Repair (Amst). 2007;6(4):505-16.

195. Bardwell PD, Woo CJ, Wei K, Li Z, Martin A, Sack SZ, et al. Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. Nat Immunol. 2004;5(2):224-9.

196. Diaz M, Verkoczy LK, Flajnik MF, Klinman NR. Decreased frequency of somatic hypermutation and impaired affinity maturation but intact germinal center formation in mice expressing antisense RNA to DNA polymerase zeta. J Immunol. 2001;167(1):327-35.

197. Jungnickel B. False moves for survival: error-prone DNA repair in adaptive immunity. Cell Cycle. 2006;5(24):2856-61.

198. Casellas R, Basu U, Yewdell WT, Chaudhuri J, Robbiani DF, Di Noia JM. Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. Nat Rev Immunol. 2016;16(3):164-76.

199. Pettersen HS, Galashevskaya A, Doseth B, Sousa MML, Sarno A, Visnes T, et al. AID expression in B-cell lymphomas causes accumulation of genomic uracil and a distinct AID mutational signature. DNA Repair. 2015;25:60-71.

200. Qian J, Wang Q, Dose M, Pruett N, Kieffer-Kwon KR, Resch W, et al. B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. Cell. 2014;159(7):1524-37.

201. Compagno M, Wang Q, Pighi C, Meng FG, Cheong TC, Yeap LS, et al. PI3K delta Inhibitors Increase Genomic Instability by Upregulating AID Expression. Cytogenetic and Genome Research. 2016;148(2-3):132-.

202. Sollier J, Cimprich KA. Breaking bad: R-loops and genome integrity. Trends Cell Biol. 2015;25(9):514-22.

203. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. Nat Rev Genet. 2015;16(10):583-97.

204. Sachen KL, Strohman MJ, Singletary J, Alizadeh AA, Kattah NH, Lossos C, et al. Self-antigen recognition by follicular lymphoma B-cell receptors. Blood. 2012;120(20):4182-90.

205. Radcliffe CM, Arnold JN, Suter DM, Wormald MR, Harvey DJ, Royle L, et al. Human follicular lymphoma cells contain oligomannose glycans in the antigen-binding site of the B-cell receptor. Journal of Biological Chemistry. 2007;282(10):7405-15.

206. Mamessier E, Drevet C, Broussais-Guillaumot F, Mollichella ML, Garciaz S, Roulland S, et al. Contiguous follicular lymphoma and follicular lymphoma in situ harboring N-Glycosylated sites. Haematologica. 2014.

207. Lindsley AW, Saal HM, Burrow TA, Hopkin RJ, Shchelochkov O, Khandelwal P, et al. Defects of B-cell terminal differentiation in patients with type-1 Kabuki syndrome. Journal of Allergy and Clinical Immunology. 2016;137(1):179-+.

208. Kasar S, Kim J, Improgo R, Tiao G, Polak P, Haradhvala N, et al. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. Nature Communications. 2015;6.

209. Gu XW, Booth CJ, Liu ZZ, Strout MP. AID-associated DNA repair pathways regulate malignant transformation in a murine model of BCL6-driven diffuse large B-cell lymphoma. Blood. 2016;127(1):102-12.

210. Bohannon C, Powers R, Satyabhama L, Cui A, Tipton C, Michaeli M, et al. Long-lived antigen-induced IgM plasma cells demonstrate somatic mutations and contribute to long-term protection. Nature Communications. 2016;7.

211. Alt FW. Advances in Immunology: Elsevier Science; 2009.

212. Wei L, Chahwan R, Wang S, Wang X, Pham PT, Goodman MF, et al. Overlapping hotspots in CDRs are critical sites for V region diversification. Proc Natl Acad Sci U S A. 2015;112(7):E728-37.

213. Alt FW, Zhang Y, Meng FL, Guo C, Schwer B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. Cell. 2013;152(3):417-29.

214. Vuong BQ, Chaudhuri J. Combinatorial mechanisms regulating AID-dependent DNA deamination: interacting proteins and post-translational modifications. Semin Immunol. 2012;24(4):264-72.

215. Casali P. Epigenetics of B Cells and Antibody Responses2016.

216. Tran TH, Nakata M, Suzuki K, Begum NA, Shinkura R, Fagarasan S, et al. B cell-specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. Nat Immunol. 2010;11(2):148-54.

217. Dedeoglu F, Horwitz B, Chaudhuri J, Alt FW, Geha RS. Induction of activation-induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NF kappa B. International Immunology. 2004;16(3):395-404.

218. Sayegh CE, Quong MW, Agata Y, Murre C. E-proteins directly regulate expression of activation-induced deaminase in mature B cells. Nat Immunol. 2003;4(6):586-93.

219. Gonda H, Sugai M, Nambu Y, Katakai T, Agata Y, Mori KJ, et al. The balance between Pax5 and Id2 activities is the key to AID gene expression. Journal of Experimental Medicine. 2003;198(9):1427-37.

220. Ise W, Kohyama M, Schraml BU, Zhang TT, Schwer B, Basu U, et al. The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. Nat Immunol. 2011;12(6):536-U245.

221. Vaidyanathan B, Yen WF, Pucella JN, Chaudhuri J. AIDing Chromatin and Transcription-Coupled Orchestration of Immunoglobulin Class-Switch Recombination. Front Immunol. 2014;5:120.

222. Tran TH, Nakata M, Suzuki K, Begum NA, Shinkura R, Fagarasan S, et al. B cell-specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. Nat Immunol. 2010;11(2):148-U67.

223. Teng G, Hakimpour P, Landgraf P, Rice A, Tuschl T, Casellas R, et al. MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. Immunity. 2008;28(5):621-9.

224. de Yebenes VG, Belver L, Pisano DG, Gonzalez S, Villasante A, Croce C, et al. miR-181b negatively regulates activation-induced cytidine deaminase in B cells. J Exp Med. 2008;205(10):2199-206.

225. Musilova K, Mraz M. MicroRNAs in B-cell lymphomas: how a complex biology gets more complex. Leukemia. 2015;29(5):1004-17.

226. Wu X, Darce JR, Chang SK, Nowakowski GS, Jelinek DF. Alternative splicing regulates activation-induced cytidine deaminase (AID): implications for suppression of AID mutagenic activity in normal and malignant B cells. Blood. 2008;112(12):4675-82.

227. Cheng HL, Vuong BQ, Basu U, Franklin A, Schwer B, Astarita J, et al. Integrity of the AID serine-38 phosphorylation site is critical for class switch recombination and somatic hypermutation in mice. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(8):2717-22.

228. Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill JC, et al. Proteasomal degradation restricts the nuclear lifespan of AID. Journal of Experimental Medicine. 2008;205(6):1357-68.

229. McBride KM, Barreto V, Ramiro AR, Stavropoulos P, Nussenzweig MC. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. Journal of Experimental Medicine. 2004;199(9):1235-44.

230. Geisberger R, Rada C, Neuberger MS. The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(16):6736-41.

231. Hui Y, Ericsson I, Doseth B, Liabakk NB, Krokan HE, Kavli B. Activation-induced cytidine deaminase (AID) is localized to subnuclear domains enriched in splicing factors. Experimental Cell Research. 2014;322(1):178-92.

232. Hu Y, Ericsson I, Torseth K, Methot SP, Sundheim O, Liabakk NB, et al. A combined nuclear and nucleolar localization motif in activation-induced cytidine deaminase (AID) controls immunoglobulin class switching. J Mol Biol. 2013;425(2):424-43.

233. Peng HZ, Du MQ, Koulis A, Aiello A, Dogan A, Pan LX, et al. Nonimmunoglobulin gene hypermutation in germinal center B cells. Blood. 1999;93(7):2167-72.

234. Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature. 2011;471(7337):189-95.

235. Chiba T, Marusawa H. A novel mechanism for inflammation-associated carcinogenesis; an important role of activation-induced cytidine deaminase (AID) in mutation induction. J Mol Med (Berl). 2009;87(10):1023-7.

236. Kikuchi K, Ishige T, Ide F, Ito Y, Saito I, Hoshino M, et al. Overexpression of Activation-Induced Cytidine Deaminase in MTX- and Age-Related Epstein-Barr Virus-Associated B-Cell Lymphoproliferative Disorders of the Head and Neck. Journal of Oncology. 2015.

237. Kawamura K, Wada A, Wang JY, Li QH, Ishii A, Tsujimura H, et al. Expression of activation-induced cytidine deaminase is associated with a poor prognosis of diffuse large B cell lymphoma patients treated with CHOP-based chemotherapy. Journal of Cancer Research and Clinical Oncology. 2016;142(1):27-36.

238. Shinmura K, Igarashi H, Goto M, Tao H, Yamada H, Matsuura S, et al. Aberrant expression and mutation-inducing activity of AID in human lung cancer. Ann Surg Oncol. 2011;18(7):2084-92.

239. Nakanishi Y, Kondo S, Wakisaka N, Tsuji A, Endo K, Murono S, et al. Role of Activation-Induced Cytidine Deaminase in the Development of Oral Squamous Cell Carcinoma. Plos One. 2013;8(4).

240. Nonaka T, Toda Y, Hiai H, Uemura M, Nakamura M, Yamamoto N, et al. Involvement of activation-induced cytidine deaminase in skin cancer development. Journal of Clinical Investigation. 2016;126(4):1367-82.

241. Shin CM. Gastric Cancer: Genetic Alternations Induced by H. pylori Infection: The Role of Activation-Induced Cytidine Deaminase. In: Kim N, editor. Helicobacter pylori. Singapore: Springer Singapore; 2016. p. 251-6.

242. Sapoznik S, Bahar-Shany K, Brand H, Pinto Y, Gabay O, Glick-Saar E, et al. Activation-Induced Cytidine Deaminase Links Ovulation-Induced Inflammation and Serous Carcinogenesis. Neoplasia. 2016;18(2):90-9.

243. Rebhandl S, Geisberger R. AIDing cancer treatment: Reducing AID activity via HSP90 inhibition. European Journal of Immunology. 2015;45(8):2208-11.

244. Montamat-Sicotte D, Litzler LC, Abreu C, Safavi S, Zahn A, Orthwein A, et al. HSP90 inhibitors decrease AID levels and activity in mice and in human cells. European Journal of Immunology. 2015;45(8):2365-76.

245. Muciaccia B, Vico C, Aromatario M, Fazi F, Cecchi R. Molecular analysis of different classes of RNA molecules from formalin-fixed paraffin-embedded autoptic tissues: a pilot study. Int J Legal Med. 2015;129(1):11-21.

246. Frankel A. Formalin fixation in the '-omics' era: a primer for the surgeon-scientist. ANZ J Surg. 2012;82(6):395-402.

247. Bass BP, Engel KB, Greytak SR, Moore HM. A review of preanalytical factors affecting molecular, protein, and morphological analysis of formalin-fixed, paraffin-embedded (FFPE) tissue: how well do you know your FFPE specimen? Arch Pathol Lab Med. 2014;138(11):1520-30.

248. Granato A, Giantin M, Ariani P, Carminato A, Baratto C, Zorzan E, et al. DNA and RNA isolation from canine oncologic formalin-fixed, paraffin-embedded tissues for downstream "-omic" analyses: possible or not? J Vet Diagn Invest. 2014;26(1):117-24.

249. Cantaert T, Schickel JN, Bannock JM, Ng YS, Massad C, Oe T, et al. Activation-Induced Cytidine Deaminase Expression in Human B Cell Precursors Is Essential for Central B Cell Tolerance. Immunity. 2015;43(5):884-95.

250. Leuenberger M, Frigerio S, Wild PJ, Noetzli F, Korol D, Zimmermann DR, et al. AID protein expression in chronic lymphocytic leukemia/small lymphocytic lymphoma is associated with poor prognosis and complex genetic alterations. Modern Pathol. 2010;23(2):177-86.

251. Sala C, Mattiuz G, Pietrobono S, Chicca A, Conticello SG. Splice Variants of Activation Induced Deaminase (AID) Do Not Affect the Efficiency of Class Switch Recombination in Murine CH12F3 Cells. Plos One. 2015;10(3).

252. Kim JH, Kim WS, Park C. Epstein-Barr virus latent membrane protein 1 increases genomic instability through Egr-1-mediated up-regulation of activation-induced cytidine deaminase in B-cell lymphoma. Leukemia Lymphoma. 2013;54(9):2035-40.

253. Kalchschmidt JS, Bashford-Rogers R, Paschos K, Gillman ACT, Styles CT, Kellam P, et al. Epstein-Barr virus nuclear protein EBNA3C directly induces expression of AID and somatic mutations in B cells. Journal of Experimental Medicine. 2016;213(6):921-8.

254. Hu WJ, Begum NA, Mondal S, Stanlie A, Honjo T. Identification of DNA cleavage- and recombination-specific hnRNP cofactors for activation-induced cytidine deaminase. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(18):5791-6.

255. Dreyling M, Ghielmini M, Marcus R, Salles G, Vitolo U, Ladetto M, et al. Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2014;25:76-82.

256. Cree IA, Deans Z, Ligtenberg MJL, Normanno N, Edsjo A, Rouleau E, et al. Guidance for laboratories performing molecular pathology for cancer patients. Journal of Clinical Pathology. 2014;67(11):923-31.

257. Karpova MB, Schoumans J, Ernberg I, Henter JI, Nordenskjold M, Fadeel B. Raji revisited: cytogenetics of the original Burkitt's lymphoma cell line. Leukemia. 2005;19(1):159-61.

258. Klein E, Ben-Bassat H, Neumann H, Ralph P, Zeuthen J, Polliack A, et al. Properties of the K562 cell line, derived from a patient with chronic myeloid leukemia. Int J Cancer. 1976;18(4):421-31.

259. Altman SA, Randers L, Rao G. Comparison of trypan blue dye exclusion and fluorometric assays for mammalian cell viability determinations. Biotechnol Prog. 1993;9(6):671-4.

260. Simbolo M, Gottardi M, Corbo V, Fassan M, Mafficini A, Malpeli G, et al. DNA qualification workflow for next generation sequencing of histopathological samples. Plos One. 2013;8(6):e62692.

261. Robin JD, Ludlow AT, LaRanger R, Wright WE, Shay JW. Comparison of DNA Quantification Methods for Next Generation Sequencing. Sci Rep. 2016;6:24067.

262. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol. 2006;7:3.

263. Smith DR. Agarose Gel Electrophoresis. In: Harwood AJ, editor. Basic DNA and RNA Protocols. Totowa, NJ: Humana Press; 1996. p. 17-21.

264. Overturf K. Quantitative PCR. Molecular Research in Aquaculture: Wiley-Blackwell; 2009. p. 39-61.

265. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402-8.

266. Lin K, Sherrington PD, Dennis M, Matrai Z, Cawley JC, Pettitt AR. Relationship between p53 dysfunction, CD38 expression, and IgV(H) mutation in chronic lymphocytic leukemia. Blood. 2002;100(4):1404-9.

267. Lefranc MP, Giudicelli V, Duroux P, Jabado-Michaloud J, Folch G, Aouinti S, et al. IMGT(R), the international ImMunoGeneTics information system(R) 25 years on. Nucleic Acids Res. 2015;43(Database issue):D413-22.

268. Zhao XD, Lu YY, Guo H, Xie HH, He LJ, Shen GF, et al. MicroRNA-7/NF-kappaB signaling regulatory feedback circuit regulates gastric carcinogenesis. J Cell Biol. 2015;210(4):613-27.

269. Liu JY, Li F, Wang LP, Chen XF, Wang D, Cao L, et al. CTL- vs Treg lymphocyte-attracting chemokines, CCL4 and CCL20, are strong reciprocal predictive

markers for survival of patients with oesophageal squamous cell carcinoma. Br J Cancer. 2015;113(5):747-55.

270. Tang Z, Ma J, Zhang W, Gong C, He J, Wang Y, et al. The Role of Prion Protein Expression in Predicting Gastric Cancer Prognosis. J Cancer. 2016;7(8):984-90.

271. McCloy RA, Rogers S, Caldon CE, Lorca T, Castro A, Burgess A. Partial inhibition of Cdk1 in G 2 phase overrides the SAC and decouples mitotic events. Cell Cycle. 2014;13(9):1400-12.

272. Gavet O, Pines J. Progressive activation of CyclinB1-Cdk1 coordinates entry to mitosis. Dev Cell. 2010;18(4):533-43.

273. Byers RJ, Sakhinia E, Joseph P, Glennie C, Hoyland JA, Menasce LP, et al. Clinical quantitation of immune signature in follicular lymphoma by RT-PCR-based gene expression profiling. Blood. 2008;111(9):4764-70.

274. Dyer MJS, Fischer P, Nacheva E, Labastide W, Karpas A. A New Human B-Cell Non-Hodgkins Lymphoma Cell-Line (Karpas 422) Exhibiting Both T(14-18) and T(4-11) Chromosomal Translocations. Blood. 1990;75(3):709-14.

275. Seong RH, Clayberger CA, Krensky AM, Parnes JR. Rescue of Daudi cell HLA expression by transfection of the mouse beta 2-microglobulin gene. J Exp Med. 1988;167(2):288-99.

276. Schneider U, Schwenk HU, Bornkamm G. Characterization of EBV-genome negative "null" and "T" cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma. Int J Cancer. 1977;19(5):621-6.

277. Jeggo PA, Pearl LH, Carr AM. DNA repair, genome stability and cancer: a historical perspective. Nat Rev Cancer. 2016;16(1):35-42.

278. Fatima R, Akhade VS, Pal D, Rao SM. Long noncoding RNAs in development and cancer: potential biomarkers and therapeutic targets. Mol Cell Ther. 2015;3:5.

279. Greytak SR, Engel KB, Bass BP, Moore HM. Accuracy of Molecular Data Generated with FFPE Biospecimens: Lessons from the Literature. Cancer Res. 2015;75(8):1541-7.

280. Karmakar S, Harcourt EM, Hewings DS, Scherer F, Lovejoy AF, Kurtz DM, et al. Organocatalytic removal of formaldehyde adducts from RNA and DNA bases. Nat Chem. 2015;7(9):752-8.

281. Wieczorek D, Hook B, Vincent E, Schagat T. Abstract 4738: Maximize recovery of quality nucleic acid from formalin-fixed paraffin-embedded tissue samples using a novel, flexible purification technology. Cancer Res. 2015;75(15 Supplement):4738-.

282. Blazewicz SJ, Barnard RL, Daly RA, Firestone MK. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. Isme J. 2013;7(11):2061-8.

283. Scholz C, Wagner E. Therapeutic plasmid DNA versus siRNA delivery: Common and different tasks for synthetic carriers. J Control Release. 2012;161(2):554-65.

284. van de Rijn M, Guo XQ, Sweeney RT, Beck AH, West RB. Molecular pathological analysis of sarcomas using paraffin-embedded tissue: current limitations and future possibilities. Histopathology. 2014;64(1):163-70.

285. Janecka A, Adamczyk A, Gasinska A. Comparison of eight commercially available kits for DNA extraction from formalin-fixed paraffin-embedded tissues. Anal Biochem. 2015;476:8-10.

286. Senguven B, Baris E, Oygur T, Berktas M. Comparison of methods for the extraction of DNA from formalin-fixed, paraffin-embedded archival tissues. Int J Med Sci. 2014;11(5):494-9.

287. Potluri K, Mahas A, Kent MN, Naik S, Markey M. Genomic DNA extraction methods using formalin-fixed paraffin-embedded tissue. Analytical Biochemistry. 2015;486:17-23.

288. von Ahlfen S, Missel A, Bendrat K, Schlumpberger M. Determinants of RNA quality from FFPE samples. Plos One. 2007;2(12):e1261.

289. Huijsmans CJ, Damen J, van der Linden JC, Savelkoul PH, Hermans MH. Comparative analysis of four methods to extract DNA from paraffin-embedded tissues: effect on downstream molecular applications. BMC Res Notes. 2010;3:239.

290. Hassani A, Khan G. A simple procedure for the extraction of DNA from long-term formalin-preserved brain tissues for the detection of EBV by PCR. Exp Mol Pathol. 2015;99(3):558-63.

291. Harada S. DNA and RNA Extractions from Mammalian Samples. In: Micic M, editor. Sample Preparation Techniques for Soil, Plant, and Animal Samples. New York, NY: Springer New York; 2016. p. 125-38.

292. Atanesyan L, Steenkamer MJ, Horstman A, Moelans CB, Schouten JP, Savola SP. Optimal Fixation Conditions and DNA Extraction Methods for MLPA Analysis on FFPE Tissue-Derived DNA. Am J Clin Pathol. 2017;147(1):60-8.

293. Arbeithuber B, Makova KD, Tiemann-Boege I. Artifactual mutations resulting from DNA lesions limit detection levels in ultrasensitive sequencing applications. DNA Res. 2016;23(6):547-59.

294. Wang M, Beck CR, English AC, Meng Q, Buhay C, Han Y, et al. PacBio-LITS: a large-insert targeted sequencing method for characterization of human disease-associated chromosomal structural variations. BMC Genomics. 2015;16:214.

295. Sohrabi M, Nair RG, Samaranayake LP, Zhang L, Zulfiker AH, Ahmetagic A, et al. The yield and quality of cellular and bacterial DNA extracts from human oral rinse

samples are variably affected by the cell lysis methodology. J Microbiol Methods. 2016;122:64-72.

296. Didelot A, Kotsopoulos SK, Lupo A, Pekin D, Li X, Atochin I, et al. Multiplex picoliter-droplet digital PCR for quantitative assessment of DNA integrity in clinical samples. Clin Chem. 2013;59(5):815-23.

297. Jiang L, Hu X, Xu T, Zhang H, Sheng D, Yin D. Prevalence of antibiotic resistance genes and their relationship with antibiotics in the Huangpu River and the drinking water sources, Shanghai, China. Sci Total Environ. 2013;458-460:267-72.

298. Lin J, Kennedy SH, Svarovsky T, Rogers J, Kemnitz JW, Xu A, et al. High-quality genomic DNA extraction from formalin-fixed and paraffin-embedded samples deparaffinized using mineral oil. Anal Biochem. 2009;395(2):265-7.

299. Swaminathan S, Klemm L, Park E, Papaemmanuil E, Ford A, Kweon SM, et al. Mechanisms of clonal evolution in childhood acute lymphoblastic leukemia. Nat Immunol. 2015;16(7):766-74.

300. Wilmore JR, Asito AS, Wei C, Piriou E, Sumba PO, Sanz I, et al. AID expression in peripheral blood of children living in a malaria holoendemic region is associated with changes in B cell subsets and Epstein-Barr virus. Int J Cancer. 2015;136(6):1371-80.

301. Dominguez-Sola D, Kung J, Holmes AB, Wells VA, Mo T, Basso K, et al. The FOXO1 Transcription Factor Instructs the Germinal Center Dark Zone Program. Immunity. 2015;43(6):1064-74.

302. Duan Z, Zheng H, Liu H, Li M, Tang M, Weng X, et al. AID expression increased by TNF-alpha is associated with class switch recombination of Igalpha gene in cancers. Cell Mol Immunol. 2015.

303. Leeman-Neill RJ, Johnson DE, Swerdlow SH. The Herbicide Isoproturon Induces Activation-Induced Cytidine Deaminase Expression in Germinal Center B Cells. Blood. 2015;126(23).

304. Shikata H, Yakushijin Y, Matsushita N, Sakai A, Sugita A, Nakamura N, et al. Role of activation-induced cytidine deaminase in the progression of follicular lymphoma. Cancer Sci. 2012;103(3):415-21.

305. Levy R. Optimization of quantitative real-time RT-PCR parameters for the study of lymphoid malignancies. Leukemia. 2003;17(4):796-7.

306. Deutsch AJ, Rinner B, Wenzl K, Pichler M, Troppan K, Steinbauer E, et al. NR4A1-mediated apoptosis suppresses lymphomagenesis and is associated with a favorable cancer-specific survival in patients with aggressive B-cell lymphomas. Blood. 2014;123(15):2367-77.

307. Tsai CT, Yang PM, Chern TR, Chuang SH, Lin JH, Klemm L, et al. AID downregulation is a novel function of the DNMT inhibitor 5-aza-deoxycytidine. Oncotarget. 2014;5(1):211-23.

308. Klemm L, Duy C, Iacobucci I, Kuchen S, von Levetzow G, Feldhahn N, et al. The B Cell Mutator AID Promotes B Lymphoid Blast Crisis and Drug Resistance in Chronic Myeloid Leukemia. Cancer Cell. 2009;16(3):232-45.

309. Xu YL, Jiang L, Fang JC, Fang R, Morse HC, Ouyang GF, et al. Loss of IRF8 Inhibits the Growth of Diffuse Large B-cell Lymphoma. Journal of Cancer. 2015;6(10):953-61.

310. Shi MY, He XD, Wei W, Wang J, Zhang T, Shen XH. Tenascin-C induces resistance to apoptosis in pancreatic cancer cell through activation of ERK/NF-kappa B pathway. Apoptosis. 2015;20(6):843-57.

311. Liu JJ, Liu JY, Chen J, Wu YX, Yan P, Ji CD, et al. Scinderin promotes the invasion and metastasis of gastric cancer cells and predicts the outcome of patients. Cancer Lett. 2016;376(1):110-7.

312. Ho Y, Li X, Jamison S, Harding HP, McKinnon PJ, Ron D, et al. PERK Activation Promotes Medulloblastoma Tumorigenesis by Attenuating Premalignant Granule Cell Precursor Apoptosis. Am J Pathol. 2016;186(7):1939-51.

313. Hough KP, Rogers AM, Zelic M, Paris M, Heilman DW. Transformed cell-specific induction of apoptosis by porcine circovirus type 1 viral protein 3. J Gen Virol. 2015;96(Pt 2):351-9.

314. Jensen EC. Quantitative analysis of histological staining and fluorescence using ImageJ. Anat Rec (Hoboken). 2013;296(3):378-81.

315. Armitage JO, Longo DL. Is watch and wait still acceptable for patients with low-grade follicular lymphoma? Blood. 2016;127(23):2804-8.

316. Tournoud M, Larue A, Cazalis MA, Venet F, Pachot A, Monneret G, et al. A strategy to build and validate a prognostic biomarker model based on RT-qPCR gene expression and clinical covariates. BMC Bioinformatics. 2015;16:106.

317. Kong H, Zhu M, Cui F, Wang S, Gao X, Lu S, et al. Quantitative assessment of short amplicons in FFPE-derived long-chain RNA. Sci Rep. 2014;4:7246.

318. Bradley WH, Eng K, Le M, Mackinnon AC, Kendziorski C, Rader JS. Comparing gene expression data from formalin-fixed, paraffin embedded tissues and qPCR with that from snap-frozen tissue and microarrays for modeling outcomes of patients with ovarian carcinoma. BMC Clin Pathol. 2015;15:17.

319. van Maldegem F, de Wit M, Morsink F, Musler A, Weegenaar J, van Noesel CJ. Effects of processing delay, formalin fixation, and immunohistochemistry on RNA

Recovery From Formalin-fixed Paraffin-embedded Tissue Sections. Diagn Mol Pathol. 2008;17(1):51-8.

320. Ludyga N, Grunwald B, Azimzadeh O, Englert S, Hofler H, Tapio S, et al. Nucleic acids from long-term preserved FFPE tissues are suitable for downstream analyses. Virchows Arch. 2012;460(2):131-40.

321. Mathieson W, Betsou F, Myshunina T, Pushkarev V, Pushkarev V, Shinkarkina A, et al. The effect of long-term -80 degrees C storage of thyroid biospecimens on RNA quality and ensuring fitness for purpose. J Clin Pathol. 2016.

322. Belder N, Coskun O, Doganay Erdogan B, Ilk O, Savas B, Ensari A, et al. From RNA isolation to microarray analysis: Comparison of methods in FFPE tissues. Pathol Res Pract. 2016.

323. Bjorkman J, Svec D, Lott E, Kubista M, Sjoback R. Differential amplicons (DeltaAmp)-a new molecular method to assess RNA integrity. Biomol Detect Quantif. 2016;6:4-12.

324. Brisco MJ, Morley AA. Quantification of RNA integrity and its use for measurement of transcript number. Nucleic Acids Res. 2012;40(18):e144.

325. Antonov J, Goldstein DR, Oberli A, Baltzer A, Pirotta M, Fleischmann A, et al. Reliable gene expression measurements from degraded RNA by quantitative real-time PCR depend on short amplicons and a proper normalization. Lab Invest. 2005;85(8):1040-50.

326. Simonet J, Gantzer C. Inactivation of poliovirus 1 and F-specific RNA phages and degradation of their genomes by UV irradiation at 254 nanometers. Appl Environ Microbiol. 2006;72(12):7671-7.

327. Opitz L, Salinas-Riester G, Grade M, Jung K, Jo P, Emons G, et al. Impact of RNA degradation on gene expression profiling. BMC Med Genomics. 2010;3:36.

328. Li Y, Wang K, Chen L, Zhu X, Zhou J. Quantification of mRNA Levels Using Real-Time Polymerase Chain Reaction (PCR). In: Cao J, editor. Breast Cancer: Methods and Protocols. New York, NY: Springer New York; 2016. p. 73-9.

329. Mustafa DA, Sieuwerts AM, Smid M, de Weerd V, van der Weiden M, Meijer-van Gelder ME, et al. A Method to Correlate mRNA Expression Datasets Obtained from Fresh Frozen and Formalin-Fixed, Paraffin-Embedded Tissue Samples: A Matter of Thresholds. Plos One. 2015;10(12):e0144097.

330. Olsen IH, Langer SW, Federspiel BH, Oxbol J, Loft A, Berthelsen AK, et al. (68)Ga-DOTATOC PET and gene expression profile in patients with neuroendocrine carcinomas: strong correlation between PET tracer uptake and gene expression of somatostatin receptor subtype 2. Am J Nucl Med Mol Imaging. 2016;6(1):59-72.

331. Yao J, Xu C, Fang Z, Li Y, Liu H, Wang Y, et al. Androgen receptor regulated microRNA miR-182-5p promotes prostate cancer progression by targeting the ARRDC3/ITGB4 pathway. Biochem Biophys Res Commun. 2016;474(1):213-9.

332. Vuong HG, Kondo T, Oishi N, Nakazawa T, Mochizuki K, Inoue T, et al. Genetic alterations of differentiated thyroid carcinoma in iodine-rich and iodine-deficient countries. Cancer Med. 2016.

333. Medimegh I, Omrane I, Privat M, Uhrhummer N, Ayari H, Belaiba F, et al. MicroRNAs expression in triple negative vs non triple negative breast cancer in Tunisia: interaction with clinical outcome. Plos One. 2014;9(11):e111877.

334. Lung J, Lin YC, Hung MS, Jiang YY, Lee KD, Lin PY, et al. A sensitive and high throughput TaqMan-based reverse transcription quantitative polymerase chain reaction assay efficiently discriminates ALK rearrangement from overexpression for lung cancer FFPE specimens. Lung Cancer. 2016;94:114-20.

335. Chamizo C, Rojo F, Madoz-Gurpide J. Determination of True ERBB2 Gene Amplification in Breast Cancer by Quantitative PCR Using a Reference and a Novel Control Gene. Appl Immunohistochem Mol Morphol. 2016;24(3):179-87.

336. Zeka F, Vanderheyden K, De Smet E, Cuvelier CA, Mestdagh P, Vandesompele J. Straightforward and sensitive RT-qPCR based gene expression analysis of FFPE samples. Sci Rep. 2016;6:21418.

337. Kocjan BJ, Hosnjak L, Poljak M. Detection of alpha human papillomaviruses in archival formalin-fixed, paraffin-embedded (FFPE) tissue specimens. J Clin Virol. 2016;76 Suppl 1:S88-97.

338. Walter RFH, Mairinger FD, Wohlschlaeger J, Worm K, Ting S, Vollbrecht C, et al. FFPE tissue as a feasible source for gene expression analysis - A comparison of three reference genes and one tumor marker. Pathology Research and Practice. 2013;209(12):784-9.

339. Pan Y, Li HX, Guo Y, Luo Y, Li H, Xu Y, et al. A pilot study of long noncoding RNA expression profiling by microarray in follicular lymphoma. Gene. 2016;577(2):132-9.

340. Le KS, Thibult ML, Just-Landi S, Pastor S, Gondois-Rey F, Granjeaud S, et al. Follicular B lymphomas generate regulatory T cells via the ICOS/ICOSL pathway and are susceptible to treatment by anti-ICOS/ICOSL therapy. Cancer Res. 2016.

341. Bosl MW, Osterode E, Bararia D, Pastore A, Staiger AM, Ott G, et al. STAT6 Is Recurrently and Significantly Mutated in Follicular Lymphoma and Enhances the IL-4 Induced Expression of Membrane-Bound and Soluble CD23. Blood. 2015;126(23).

342. Yildiz M, Li H, Bernard D, Amin NA, Ouillette P, Jones S, et al. Activating STAT6 mutations in follicular lymphoma. Blood. 2015;125(4):668-79.

343. Lykken JM, Horikawa M, Minard-Colin V, Kamata M, Miyagaki T, Poe JC, et al. Galectin-1 drives lymphoma CD20 immunotherapy resistance: validation of a preclinical system to identify resistance mechanisms. Blood. 2016;127(15):1886-95.

344. de Kok JB, Roelofs RW, Giesendorf BA, Pennings JL, Waas ET, Feuth T, et al. Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. Laboratory Investigation. 2005;85(1):154-9.

345. Lossos IS, Czerwinski DK, Wechser MA, Levy R. Optimization of quantitative real-time RT-PCR parameters for the study of lymphoid malignancies. Leukemia. 2003;17(4):789-95.

346. Rush JS, Liu M, Odegard VH, Unniraman S, Schatz DG. Expression of activation-induced cytidine deaminase is regulated by cell division, providing a mechanistic basis for division-linked class switch recombination. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(37):13242-7.

347. Cattoretti G, Buttner M, Shaknovich R, Kremmer E, Alobeid B, Niedobitek G. Nuclear and cytoplasmic AID in extrafollicular and germinal center B cells. Blood. 2006;107(10):3967-75.

348. Humby F, Bombardieri M, Manzo A, Kelly S, Blades MC, Kirkham B, et al. Ectopic Lymphoid Structures Support Ongoing Production of Class-Switched Autoantibodies in Rheumatoid Synovium. Plos Medicine. 2009;6(1):59-75.

349. Bombardieri M, Barone F, Humby F, Kelly S, McGurk M, Morgan P, et al. Activation-induced cytidine deaminase expression in follicular dendritic cell networks and interfollicular large B cells supports functionality of ectopic lymphoid neogenesis in autoimmune sialoadenitis and MALT lymphoma in Sjogren's syndrome. J Immunol. 2007;179(7):4929-38.

350. Qin HY, Suzuki K, Nakata M, Chikuma S, Izumi N, Huong LT, et al. Activation-Induced Cytidine Deaminase Expression in CD4(+) T Cells is Associated with a Unique IL-10-Producing Subset that Increases with Age. Plos One. 2011;6(12).

351. Methot SP, Litzler LC, Trajtenberg F, Zahn A, Robert F, Pelletier J, et al. Consecutive interactions with HSP90 and eEF1A underlie a functional maturation and storage pathway of AID in the cytoplasm. Journal of Experimental Medicine. 2015;212(4):581-96.

352. Uchimura Y, Barton LF, Rada C, Neuberger MS. REG-gamma associates with and modulates the abundance of nuclear activation-induced deaminase. Journal of Experimental Medicine. 2011;208(12):2385-91.

353. Delker RK, Zhou YJ, Strikoudis A, Stebbins CE, Papavasiliou FN. Solubility-based genetic screen identifies RING finger protein 126 as an E3 ligase for activation-induced cytidine deaminase. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(3):1029-34.

354. Willenbrock K, Renne C, Rottenkolber M, Klapper W, Dreyling M, Engelhard M, et al. The expression of activation induced cytidine deaminase in follicular lymphoma is independent of prognosis and stage. Histopathology. 2009;54(4):509-12.

355. Bognar A, Csernus B, Bodor C, Reiniger L, Szepesi A, Toth E, et al. Clonal selection in the bone marrow involvement of follicular lymphoma. Leukemia. 2005;19(9):1656-62.

356. Loeffler M, Kreuz M, Haake A, Hasenclever D, Trautmann H, Arnold C, et al. Integrated Somatic Mutation and DNA Methylation Analysis Reveal Genomic and Epigenomic Co-Evolution In Follicular Lymphomas. Blood. 2013;122(21).

357. Staszewski O, Baker RE, Ucher AJ, Martier R, Stavnezer J, Guikema JEJ. Activation-Induced Cytidine Deaminase Induces Reproducible DNA Breaks at Many Non-Ig Loci in Activated B Cells. Molecular Cell. 2011;41(2):232-42.

358. Hardianti MS, Tatsumi E, Syampurnawati M, Furuta K, Saigo K, Nakamachi Y, et al. Activation-induced cytidine deaminase expression in follicular lymphoma: association between AID expression and ongoing mutation in FL. Leukemia. 2004;18(4):826-31.

359. Berget E, Molven A, Lokeland T, Helgeland L, Vintermyr OK. IGHV gene usage and mutational status in follicular lymphoma: Correlations with prognosis and patient age. Leukemia Research. 2015;39(7):702-8.

360. Spence JM, Spence JP, Abumoussa A, Burack WR. Ultradeep analysis of tumor heterogeneity in regions of somatic hypermutation. Genome Medicine. 2015;7.

361. Jiang YW, Redmond D, Nie K, Eng KW, Clozel T, Martin P, et al. Deep sequencing reveals clonal evolution patterns and mutation events associated with relapse in B-cell lymphomas. Genome Biol. 2014;15(8).

362. Lee SE, Kang SY, Yoo HY, Kim SJ, Kim WS, Ko YH. Clonal relationships in recurrent B-cell lymphomas. Oncotarget. 2016;7(11):12359-71.

363. Boice M, Salloum D, Mourcin F, Sanghvi V, Amin R, Oricchio E, et al. Loss of the HVEM Tumor Suppressor in Lymphoma and Restoration by Modified CAR-T Cells. Cell. 2016;167(2):405-+.

364. Brauninger A, Spieker T, Willenbrock K, Gaulard P, Wacker HH, Rajewsky K, et al. Survival and clonal expansion of mutating "forbidden" (immunoglobulin receptor-deficient) Epstein-Barr virus-infected B cells in angioimmunoblastic T cell lymphoma. Journal of Experimental Medicine. 2001;194(7):927-40.

365. Takata K, Sato Y, Nakamura N, Kikuti YY, Ichimura K, Tanaka T, et al. Duodenal and nodal follicular lymphomas are distinct: the former lacks activation-induced cytidine

deaminase and follicular dendritic cells despite ongoing somatic hypermutations. Modern Pathol. 2009;22(7):940-9.

366. Hestand MS, Van Houdt J, Cristofoli F, Vermeesch JR. Polymerase specific error rates and profiles identified by single molecule sequencing. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis. 2016;784:39-45.

367. Lee DF, Lu J, Chang S, Loparo JJ, Xie XS. Mapping DNA polymerase errors by single-molecule sequencing. Nucleic Acids Res. 2016;44(13):e118.

368. Rawson JMO, Clouser CL, Mansky LM. Rapid Determination of HIV-1 Mutant Frequencies and Mutation Spectra Using an mCherry/EGFP Dual-Reporter Viral Vector. In: Prasad RV, Kalpana VG, editors. HIV Protocols. New York, NY: Springer New York; 2016. p. 71-88.

369. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Research. 2008;36:W465-W9.

370. Zhu DX, Lossos C, Chapman-Fredricks JR, Matthews JM, Ikpatt OF, Ruiz P, et al. Biased Use of the IGHV4 Family and Evidence for Antigen Selection in Chlamydophila psittaci-Negative Ocular Adnexal Extranodal Marginal Zone Lymphomas. Plos One. 2011;6(12).

371. Nihal M, Mikkola D, Wood GS. Detection of clonally restricted immunoglobulin heavy chain gene rearrangements in normal and lesional skin: analysis of the B cell component of the skin-associated lymphoid tissue and implications for the molecular diagnosis of cutaneous B cell lymphomas. J Mol Diagn. 2000;2(1):5-10.

372. Pai RK, Chakerian AE, Binder JM, Amin M, Viswanatha DS. B-cell clonality determination using an immunoglobulin kappa light chain polymerase chain reaction method. J Mol Diagn. 2005;7(2):300-7.

373. Pambuccian SE, Bardales RH. Lymph Node Cytopathology: Springer US; 2010.

374. Stein H, Lennert K. Histopathology of Non-Hodgkin's Lymphomas: Based on the Kiel Classification: Springer Berlin Heidelberg; 2013.

375. Halldorsdottir AM, Zehnbauer BA, Burack WR. Application of BIOMED-2 clonality assays to formalin-fixed paraffin embedded follicular lymphoma specimens: Superior performance of the IGK assays compared to IGH for suboptimal specimens. Leukemia Lymphoma. 2007;48(7):1338-43.

376. Berget E, Helgeland L, Molven A, Vintermyr OK. Detection of clonality in follicular lymphoma using formalin-fixed, paraffin-embedded tissue samples and BIOMED-2 immunoglobulin primers. Journal of Clinical Pathology. 2011;64(1):37-41.

377. Jackson KJL, Wang Y, Collins AM. Human immunoglobulin classes and subclasses show variability in VDJ gene mutation levels. Immunology and Cell Biology. 2014;92(8):729-33.

378. Bose B, Sinha S. Problems in using statistical analysis of replacement and silent mutations in antibody genes for determining antigen-driven affinity selection. Immunology. 2005;116(2):172-83.

379. Smit LA, Bende RJ, Aten J, Guikema JEJ, Aarts WM, van Noesel CJM. Expression of activation-induced cytidine deaminase is confined to B-cell non-Hodgkin's lymphomas of germinal-center phenotype. Cancer Res. 2003;63(14):3894-8.

380. Rebhandl S, Huemer M, Greil R, Geisberger R. AID/APOBEC deaminases and cancer. Oncoscience. 2015;2(4):320-33.

381. Shain KH, Dalton WS, Tao J. The tumor microenvironment shapes hallmarks of mature B-cell malignancies. Oncogene. 2015;34(36):4673-82.

382. Sugimoto T, Watanabe T. Follicular Lymphoma: The Role of the Tumor Microenvironment in Prognosis. J Clin Exp Hematop. 2016;56(1):1-19.

383. Zhu D, McCarthy H, Ottensmeier CH, Johnson P, Hamblin TJ, Stevenson FK. Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma. Blood. 2002;99(7):2562-8.

384. McCann KJ, Ottensmeier CH, Callard A, Radcliffe CM, Harvey DJ, Dwek RA, et al. Remarkable selective glycosylation of the immunoglobulin variable region in follicular lymphoma. Mol Immunol. 2008;45(6):1567-72.

385. Coelho V, Krysov S, Ghaemmaghami AM, Emara M, Potter KN, Johnson P, et al. Glycosylation of surface Ig creates a functional bridge between human follicular lymphoma and microenvironmental lectins. Proc Natl Acad Sci U S A. 2010;107(43):18587-92.

386. Palacios F, Moreno P, Morande P, Abreu C, Correa A, Porro V, et al. High expression of AID and active class switch recombination might account for a more aggressive disease in unmutated CLL patients: link with an activated microenvironment in CLL disease. Blood. 2010;115(22):4488-96.

387. Elinav E, Nowarski R, Thaiss CA, Hu B, Jin CC, Flavell RA. Inflammation-induced cancer: crosstalk between tumours, immune cells and microorganisms. Nature Reviews Cancer. 2013;13(11):759-71.

388. Glas AM, Kersten MJ, Delahaye LJ, Witteveen AT, Kibbelaar RE, Velds A, et al. Gene expression profiling in follicular lymphoma to assess clinical aggressiveness and to guide the choice of treatment. Blood. 2005;105(1):301-7.

389. Mainiero F, Soriani A, Strippoli R, Jacobelli J, Gismondi A, Piccoli M, et al. RAC1/P38 MAPK signaling pathway controls beta1 integrin-induced interleukin-8 production in human natural killer cells. Immunity. 2000;12(1):7-16.

390. Cuvelier SL, Paul S, Shariat N, Colarusso P, Patel KD. Eosinophil adhesion under flow conditions activates mechanosensitive signaling pathways in human endothelial cells. J Exp Med. 2005;202(6):865-76.

391. Brand S, Sakaguchi T, Gu X, Colgan SP, Reinecker HC. Fractalkine-mediated signals regulate cell-survival and immune-modulatory responses in intestinal epithelial cells. Gastroenterology. 2002;122(1):166-77.

392. Lu T, Pan Y, Kao SY, Li C, Kohane I, Chan J, et al. Gene regulation and DNA damage in the ageing human brain. Nature. 2004;429(6994):883-91.

393. Glas AM, Knoops L, Delahaye L, Kersten MJ, Kibbelaar RE, Wessels LA, et al. Gene-expression and immunohistochemical study of specific T-cell subsets and accessory cell types in the transformation and prognosis of follicular lymphoma. J Clin Oncol. 2007;25(4):390-8.

394. Lee AM, Clear AJ, Calaminici M, Davies AJ, Jordan S, MacDougall F, et al. Number of CD4(+) cells and location of forkhead box protein P3-positive cells in diagnostic follicular lymphoma tissue microarrays correlates with outcome. J Clin Oncol. 2006;24(31):5052-9.

395. Taskinen M, Karjalainen-Lindsberg ML, Nyman H, Eerola LM, Leppa S. A high tumor-associated macrophage content predicts favorable outcome in follicular lymphoma patients treated with rituximab and cyclophosphamide-doxorubicin-vincristine-prednisone. Clin Cancer Res. 2007;13(19):5784-9.

396. Scherer F, Navarrete MA, Bertinetti-Lapatki C, Boehm J, Schmitt-Graeff A, Veelken H. Isotype-switched follicular lymphoma displays dissociation between activation-induced cytidine deaminase expression and somatic hypermutation. Leukemia Lymphoma. 2016;57(1):151-60.

397. McCarthy H, Wierda WG, Barron LL, Cromwell CC, Wang J, Coombes KR, et al. High expression of activation-induced cytidine deaminase (AID) and splice variants is a distinctive feature of poor-prognosis chronic lymphocytic leukemia. Blood. 2003;101(12):4903-8.

398. Klinger M, Zheng J, Elenitoba-Johnson KS, Perkins SL, Faham M, Bahler DW. Next-generation IgVH sequencing CLL-like monoclonal B-cell lymphocytosis reveals frequent oligoclonality and ongoing hypermutation. Leukemia. 2016;30(5):1055-61.

399. Saunders NA, Lee MA. Real-time PCR: Advanced Technologies and Applications: Caister Academic Press; 2013.

400. Hasler J, Rada C, Neuberger MS. Cytoplasmic activation-induced cytidine deaminase (AID) exists in stoichiometric complex with translation elongation factor 1alpha (eEF1A). Proc Natl Acad Sci U S A. 2011;108(45):18366-71.

401. Intlekofer AM, Younes A. Precision therapy for lymphoma-current state and future directions. Nat Rev Clin Oncol. 2014;11(10):585-96.

402. Stover J, Yellore V, Miller A, Sharman JP. Validation of a Next-Generation Sequencing Cancer Panel for Clinical Mutation Profiling in Patients with Diffuse Large B-Cell Lymphoma (DLBCL). Blood. 2014;124(21).

403. Ennishi D, Hoffer C, Shulha H, Mottok A, Farinha P, Chan FC, et al. Clinical Significance of Genetic Aberrations in Diffuse Large B Cell Lymphoma. Blood. 2014;124(21).

404. Asmann YW, Maurer MJ, Wang C, Sarangi V, Ansell SM, Feldman AL, et al. Genetic diversity of newly diagnosed follicular lymphoma. Blood Cancer J. 2014;4.

405. Taskinen M, Meriranta L, Pasanen A, Louhimo R, Cervera A, Karjalainen-Lindsberg ML, et al. Somatic Mutations in E3 Ubiquitin Ligase Deltex 1 Are Associated with Survival in Diffuse Large B-Cell Lymphoma. Blood. 2014;124(21).

406. Bouamar H, Abbas S, Lin AP, Wang L, Jiang DF, Holder KN, et al. A capture-sequencing strategy identifies IRF8, EBF1, and APRIL as novel IGH fusion partners in B-cell lymphoma. Blood. 2013;122(5):726-33.

407. Talaulikar D, Jeelall Y, Ziolkowski A, Cook M, Jain S, Andrews D, et al. Genetics of Disease Progression in Diffuse Large B-Cell Lymphoma: Clonal Selection and Acquisition of Newly Acquired Somatic Mutations at Relapse. Blood. 2014;124(21).

408. Melchardt T, Hufnagl C, Weigert O, Weinstock DM, Kopp N, Weiss L, et al. Clonal Evolution in Relapsed or Refractory Diffuse Large B Cell Lymphoma. Blood. 2014;124(21).

409. Healy JA, Nugent A, Rempel RE, Moffitt AB, Davis NS, Jiang XY, et al. GNA13 loss in germinal center B cells leads to impaired apoptosis and promotes lymphoma in vivo. Blood. 2016;127(22):2723-31.

410. van Krieken JH. New developments in the pathology of malignant lymphoma. A review of the literature published from January-April 2016. J Hematop. 2016;9(2):73-83.

411. Mareschal S, Dubois S, Viailly PJ, Bertrand P, Bohers E, Maingonnat C, et al. Whole Exome Sequencing of Relapsed/Refractory Patients Expands the Repertoire of Somatic Mutations in Diffuse Large B-Cell Lymphoma. Gene Chromosome Canc. 2016;55(3):251-67.

412. Oricchio E, Nanjangud G, Wolfe AL, Schatz JH, Mavrakis KJ, Jiang M, et al. The Eph-Receptor A7 Is a Soluble Tumor Suppressor for Follicular Lymphoma. Cell. 2011;147(3):554-64.

413. Demosthenous C, Han JJ, Hu GZ, Stenson M, Gupta M. Loss of function mutations in PTPN6 promote STAT3 deregulation via JAK3 kinase in diffuse large B-cell lymphoma. Oncotarget. 2015;6(42):44703-13.

414. Novak AJ, Asmann YW, Maurer MJ, Wang C, Slager SL, Hodge LS, et al. Whole-exome analysis reveals novel somatic genomic alterations associated with outcome in immunochemotherapy-treated diffuse large B-cell lymphoma. Blood Cancer J. 2015;5.

415. Ying ZX, Jin M, Peterson LF, Bernard D, Saiya-Cork K, Yildiz M, et al. Recurrent Mutations in the MTOR Regulator RRAGC in Follicular Lymphoma. Clin Cancer Res. 2016;22(21):5383-93.

Appendices

Appendix table 1: Clinical data of PACIFICO trial patients used in our study.

Patient no	Age	Sex	No. Lymph Nodes	Ann Arbor Staging	Histological Grade	Date of Diagnosis	LDH (IU/L)
159	79	F	4	II	2	24/11/2009	394
182	70	M	6	III	3	26/07/2012	NULL
187	75	M	1	IV	3	03/01/2013	365
261	67	F	6	IV	2	24/02/2009	258
290	75	M	6	IV	2	10/07/2014	313
202	68	M	5	II	2	30/04/2013	353
310	76	M	6	IV	1	03/07/2013	258
278	73	M	3	IV	2	07/05/2014	296
166	64	F	5	II	3	03/09/2012	531
003	68	M	6	III	2	20/01/2010	315
013	76	M	6	III	1	19/04/2010	380
286	73	M	6	IV	2	16/06/2014	297
051	64	F	4	IV	2	20/12/2010	289
120	57	M	3	IV	2	26/01/2012	158
33	69	M	6	III	2	18/09/2010	491
291	69	M	1	IV	3	17/07/2014	1.25
53	70	F	6	IV	2	16/12/2010	349
211	77	F	6	III	3	16/03/2010	427
199	66	F	6	IV	1	22/03/2013	476
309	73	M	5	III	1	20/06/2013	546
256	64	M	6	IV	2	17/01/2014	658
266	78	F	6	IV	2	12/03/2014	490
008	78	M	4	III	3	23/02/2010	300
016	66	F	6	III	2	27/05/2010	382
035	66	M	6	III	3	13/09/2010	481
045	70	M	2	II	2	28/10/2010	273
063	54	M	6	III	1	18/02/2011	434
142	67	M	5	III	3	23/11/2010	193
118	72	F	6	III	1	16/01/2012	586
156	81	F	6	III	3	02/08/2012	398
192	64	F	6	IV	1	17/11/2012	148
151	76	M	4	II	1	01/06/2012	167
180	72	M	5	II	1	29/11/2012	181
160	64	F	6	III	3	11/05/2012	151
255	66	M	6	IV	1	19/02/2014	166
274	89	M	6	III	1	02/05/2014	127
289	69	F	6	IV	2	09/07/2014	213
292	63	F	6	IV	1	02/07/2014	147
20	82	M	6	IV	2	22/03/2010	656

Patient no	Age	Sex	No. Lymph Nodes	Ann Arbor Staging	Histological Grade	Date of Diagnosis	LDH (IU/L)
237	70	M	5	II	1	30/10/2013	367
134	67	M	6	III	2	04/01/2012	203
280	71	F	6	III	3	11/06/2014	543
242	76	F	6	IV	3	04/12/2013	830
84	65	M	4	III	1	19/05/2011	335
72	77	F	6	III	3	01/03/2011	354
221	66	F	5	III	2	02/09/2013	326
71	63	F	6	III	2	11/06/2004	136
11	70	F	5	III	3	29/04/2010	382
56	64	M	6	III	3	17/02/2011	456
157	64	M	6	IV	2	26/07/2012	494
196	79	M	5	III	2	12/10/2012	375
217	64	F	6	III	3	22/07/2013	411
227	60	M	1	III	3	09/05/2013	324
259	66	M	6	IV	2	03/03/2014	560
179	67	M	6	IV	1	27/09/2012	251
307	78	M	6	IV	2	25/09/2014	219
287	61	F	4	IV	2	06/06/2014	156
267	78	F	6	IV	2	30/01/2014	109
296	78	F	4	III	3	07/07/2014	373
234	76	M	2	II	2	01/11/2013	168
269	75	M	6	IV	2	20/03/2014	358
183	72	F	5	II	2	15/03/2012	626
64	79	F	6	III	1	04/04/2011	552
230	75	M	6	III	1	21/10/2013	303
324	77	M	1	III	1	02/02/2011	419
58	68	M	5	III	2	28/11/2005	317
205	72	F	3	II	1	04/06/2013	213
28	59	F	6	III	2	15/07/2010	180
80	72	F	4	III	1	26/05/2011	178
27	70	M	6	III	1	20/11/2007	422
59	84	F	4	IV	3	19/08/2010	367
113	78	F	6	IV	1	09/11/2011	333
264	67	M	6	III	2	06/03/2014	327
48	69	M	6	IV	2	29/12/2010	487
200	68	M	6	IV	2	23/04/2013	862
68	71	F	6	IV	2	08/03/2011	424
201	54	F	6	III	3	01/05/2013	303
244	74	F	5	IV	3	16/07/2013	443
219	77	F	3	IV	3	22/10/2013	479
233	64	F	6	IV	1	01/07/2013	333
105	67	F	6	III	2	04/10/2011	168
37	64	F	2	II	2	25/08/2010	447

Patient no	Age	Sex	No. Lymph Nodes	Ann Arbor Staging	Histological Grade	Date of Diagnosis	LDH (IU/L)
76	73	F	6	IV	2	14/04/2011	219
55	78	F	6	III	2	08/02/2011	156
42	75	M	6	III	2	13/10/2010	476
46	72	M	6	III	2	07/12/2010	263
38	74	M	6	IV	2	09/11/2010	563

Appendix Table 2: Clinical data of FL FFPE samples borrowed from Liverpool BioInnovation Hub;* NK is not known.

Block	Sex	Age at diagnosis	Site	Diagnosis	Grade	IHC stains
T54/94	Female	60	Femoral Node	NHL	NK*	NK
T305/03	Female	60	Left axillary	FL	1	bcl-6+, bc2-2+, CD10-/+, CD20+ CD79A+. Negative for bcl-1, CD5, CD23 & light chains
T247/99	Female	72	Spleen	FL	1	B Cell type CD10+ bcl-2 negative
T106/93	Female	66	Left axillary	NHL	NK	NK
T15/98	Female	51	Left axillary	FL	2	IgM+, lambda+, Kappa-, IgD-) and unequivocal stainings for CD79A,CD2 + CD45. BCL2+ except in residual GCs which contain CD21 positive dendritic cells. There is no CD30 positive cells. CD3 in interfollicular regions. Scattered CD68 positive cells are present.

Appendix Table 3: 161 important genes in pathogenesis of FL that will be studied by NGS.

	Gene	Ensembl	Transcipt ID	Reference Sequence	Function
1	*ACTB*	ENSG00000075624	ENST00000331789	NM_001101	Actin protein/cell adhesion/chromatin remodeling
2	*APC*	ENSG00000134982	ENST00000257430	NM_000038	Tumour suppressor/WNT signalling/ B catenin
3	*ARHGEF1*	ENSG00000076928	ENST00000354532	NM_004706	GTPase activating protein for GNA12 and GNA13
4	*ARID1A*	ENSG00000117713	ENST00000324856	NM_006015	Chromatin remodelling/transcrition coactivator activity
5	*ARID1B*	ENSG00000049618	ENST00000346085	NM_020732	Chromatin remodelling
6	*ATM*	ENSG00000149311	ENST00000278616	NM_000051	DNA damage response/signal transduction
7	*ATP10A*	ENSG00000206190	ENST00000356865	NM_024490	ATPase/transport of phospholipids
8	*ATP13A4*	ENSG00000127249	ENST00000342695	NM_032279	ATPase activity
9	*ATP6AP1*	ENSG00000071553	ENST00000369762	NM_001183	V-ATPase
10	*ATP6AP2*	ENSG00000182220	ENST00000378438	NM_005765	Renin and prorenin cellular receptor
11	*ATP6V0A1*	ENSG00000033627	ENST00000343619	NM_001130021	V-ATPase
12	*ATP6V0C*	ENSG00000185883	ENST00000330398	NM_001694	V-ATPase
13	*ATP6V0D1*	ENSG00000159720	ENST00000290949	NM_004691	V-ATPase
14	*ATP6V1A*	ENSG00000114573	ENST00000273398	NM_001690	V-ATPase
15	*ATP6V1B2*	ENSG00000147416	ENST00000276390	NM_001693	V-ATPase
16	*ATP6V1F*	ENSG00000128524	ENST00000492758	NM_001198909	V-ATPase
17	*ATRX*	ENSG00000085224	ENST00000373344	NM_000489	Transcriptional regulation/chromatin remodelling
18	*B2M*	ENSG00000166710	ENST00000558401	NM_004048	Antigen presentation
19	*BCL10*	ENSG00000142867	ENST00000370580	NM_003921	NF-κB/pro apoptotic
20	*BCL2*	ENSG00000171791	ENST00000398117	NM_000633	Anti-apoptosis
21	*BCL6*	ENSG00000113916	ENST00000406870	NM_001706	Transcriptional repressor
22	*BCL7A*	ENSG00000110987	ENST00000538010	NM_020993	Translocated in Burkett's Lymphoma
23	*BCOR*	ENSG00000183337	ENST00000342274	NM_001123383	Transcriptional repressor/inhibits H3K7 methylation
24	*BCR*	ENSG00000186716	ENST00000305877	NM_004327	Serine threonine kinase activity/GTPase
25	*BRD4*	ENSG00000141867	ENST00000263377	NM_058243	Chromatin reader/NF-κB
26	*BTG1*	ENSG00000133639	ENST00000256015	NM_001731	Anti-proliferative
27	*BTG2*	ENSG00000159388	ENST00000290551	NM_006763	Anti-proliferative/cell cycle regulation
28	*CARD11*	ENSG00000198286	ENST00000396946	NM_032415	BCR/NF-κB
29	*CCND3*	ENSG00000112576	ENST00000372991	NM_001760	Cell cycle/interacts with rb1
30	*CD22*	ENSG00000012124	ENST00000085219	NM_001771	B-cell signalling

	Gene	Ensembl	Transcipt ID	Reference Sequence	Function
31	*CD36*	ENSG00000135218	ENST00000447544	NM_001001548	Cell adhesion/TLR (Toll like receptor) signalling/myd88 signalling
32	*CD58*	ENSG00000116815	ENST00000457047	NM_001144822	T cell activation/immune evasion/cell adhesion
33	*CD70*	ENSG00000125726	ENST00000245903	NM_001252	T cell activation/immune response
34	*CD74*	ENSG00000019582	ENST00000353334	NM_001025159	Immune evasion/MHC class II presentation
35	*CD79A*	ENSG00000105369	ENST00000221972	NM_001783	B-cell receptor signalling
36	*CD79B*	ENSG00000007312	ENST00000006750	NM_000626	B-cell receptor signalling
37	*CD83*	ENSG00000112149	ENST00000379153	NM_001040280	Antigen presentation/immune response
38	*CDKN2A*	ENSG00000147889	ENST00000304494	NM_000077	Cell cycle progression/negative regulator of proliferation
39	*CDKN2B*	ENSG00000147883	ENST00000276925	NM_004936	Cell cycle progression
40	*CIITA*	ENSG00000179583	ENST00000324288	NM_000246	Immune response/HLA class II
41	*CREBBP*	ENSG00000005339	ENST00000262367	NM_004380	H3k27/18AC
42	*CRIPAK*	ENSG00000179979	ENST00000324803	NM_175918	PAK 1 inhibitor
43	*CXCR4*	ENSG00000121966	ENST00000409817	NM_001008540	Chemokine receptor CXCL12
44	*DTX1*	ENSG00000135144	ENST00000257600	NM_004416	Ubiquitan ligase/notch pathway
45	*DUSP2*	ENSG00000158050	ENST00000288943	NM_004418	Regulate map kinases
46	*EBF1*	ENSG00000164330	ENST00000313708	NM_024007	B-cell development/transcriptional coactivator
47	*EEF1A1*	ENSG00000156508	ENST00000316292	NM_001402	Translation elongation factor/protein synthesis
48	*EP300*	ENSG00000100393	ENST00000263253	NM_001429	*H3K27/18AC*
49	*EPHA7*	ENSG00000135333	ENST00000369303	NM_001288629	Tyrosine kinase/ERK signalling/neuro development
50	*ETS1*	ENSG00000134954	ENST00000392668	NM_001143820	Transcription factor/cell differentiation and survival
51	*ETV1*	ENSG00000006468	ENST00000430479	NM_004956	Transcriptional activator
52	*ETV6*	ENSG00000139083	ENST00000396373	NM_001987	Transcriptional repressor
53	*EZH2*	ENSG00000106462	ENST00000320356	NM_001203247	H3K27meth3
54	*FAS*	ENSG00000026103	ENST00000355740	NM_000043	Apoptosis
55	*FAT2*	ENSG00000086570	ENST00000261800	NM_001447	Cell adhesion/proliferation/ cerebellar devlopment
56	*FAT4*	ENSG00000196159	ENST00000394329	NM_024582	Regulates cell polarity
57	*FBXO11*	ENSG00000138081	ENST00000403359	NM_001190274	The SCF(FBXO11) complex mediates ubiquitination
58	*FBXW7*	ENSG00000109670	ENST00000281708	NM_033632	Ubiquitination/degradation of MYC and notch
59	*FOXO1*	ENSG00000150907	ENST00000379561	NM_002015	Transcription factor
60	*FOXP1*	ENSG00000114861	ENST00000318789	NM_001244810	Transcriptional repressor/b cell development
61	*GNA12*	ENSG00000146535	ENST00000275364	NM_007353	Cell migration/guanine nucleotide binding protein
62	*GNA13*	ENSG00000120063	ENST00000439174	NM_006572	Cell migration/guanine nucleotide binding protein
63	*GNAI2*	ENSG00000114353	ENST00000313601	NM_001282618	GtPase

	Gene	Ensembl	Transcipt ID	Reference Sequence	Function
64	*HDAC7*	ENSG00000061273	ENST00000080059	NM_015401	Histone deacetylase
65	*HIST1H1B*	ENSG00000184357	ENST00000331442	NM_005322	Chromatin remodelling
66	*HIST1H1C*	ENSG00000187837	ENST00000343677	NM_005319	Chromatin remodelling
67	*HIST1H1D*	ENSG00000124575	ENST00000244534	NM_005320	Chromatin remodelling
68	*HIST1H1E*	ENSG00000168298	ENST00000304218	NM_005321	Chromatin remodelling
69	*HIST1H2AC*	ENSG00000180573	ENST00000377791	NM_003512	Chromatin remodelling
70	*HIST1H2AE*	ENSG00000277075	ENST00000303910	NM_021052	Chromatin remodelling
71	*HIST1H2AG*	ENSG00000196787	ENST00000359193	NM_021064	Chromatin remodelling
72	*HIST1H2AL*	ENSG00000276903	ENST00000613174	NM_003511	Chromatin remodelling
73	*HIST1H2AM*	ENSG00000278677	ENST00000359611	NM_003514	Chromatin remodelling
74	*HIST1H2BC*	ENSG00000180596	ENST00000396984	NM_003526	Chromatin remodelling
75	*HIST1H2BD*	ENSG00000158373	ENST00000289316	NM_138720	Chromatin remodelling
76	*HIST1H2BG*	ENSG00000273802	ENST00000541790	NM_003518	Chromatin remodelling
77	*HIST1H2BJ*	ENSG00000124635	ENST00000541790	NP_066402	Chromatin remodelling
78	*HIST1H2BK*	ENSG00000197903	ENST00000396891	NM_080593	Chromatin remodelling
79	*HIST1H2BO*	ENSG00000274641	ENST00000616182	NM_003527	Chromatin remodelling
80	*HIST1H3B*	ENSG00000274267	ENST00000621411	NM_003537	Chromatin remodelling
81	*HIST1H3C*	ENSG00000278272	ENST00000612966	NM_003531	Chromatin remodelling
82	*HIST1H3G*	ENSG00000273983	ENST00000614378	NM_003534	Chromatin remodelling
83	*HIST1H3H*	ENSG00000278828	ENST00000369163	NM_003536	Chromatin remodelling
84	*ID3*	ENSG00000117318	ENST00000374561	NM_002167	Transcriptional regulator/apoptosis/ differntiation
85	*IKZF3*	ENSG00000161405	ENST00000346872	NM_012481	Transcription factor/role in B-cell differentiation
86	*IRF4*	ENSG00000137265	ENST00000380956	NM_001195286	Transcriptional activator/binds to MHC class I promoter
87	*IRF8*	ENSG00000140968	ENST00000268638	NM_002163	Regulates interferon induced MHC class I genes
88	*ITPKB*	ENSG00000143772	ENST00000429204	NM_002221	Inositol phosphate metabolism
89	*KDM2B*	ENSG00000089094	ENST00000377069	NM_001005366	Histone demethylase
90	*KLHL14*	ENSG00000197705	ENST00000359358	NM_020805	Unknown
91	*KLHL6*	ENSG00000172578	ENST00000341319	NM_130446	Germinal centre formation
92	*LRRN3*	ENSG00000173114	ENST00000451085	NM_001099660	Unknown/highly expressed in brain/protein phosphorylation
93	*LYN*	ENSG00000254087	ENST00000520220	NM_001111097	B-cell receptor
94	*LYST*	ENSG00000143669	ENST00000389793	NM_000081	Intracellular protein trafficking
95	*MALT1*	ENSG00000172175	ENST00000348428	NM_006785	NF-κB
96	*MAP2K1*	ENSG00000169032	ENST00000307102	NM_002755	Map kinase signalling pathway/proliferation
97	*MCL1*	ENSG00000143384	ENST00000369026	NM_001197320	Apoptosis regulation
98	*MEF2B*	ENSG00000213999	ENST00000424583	NM_005919	Histone acetylation
99	*MEF2C*	ENSG00000081189	ENST00000340208	NM_001193347	Transcriptiomal activator/B-cell development
100	*MLL/KMT2A*	ENSG00000118058	ENST00000534358	NM_001197104	Chromatin structure
101	*MLL2/KMT2D*	ENSG00000167548	ENST00000301067	NM_003482	Histone methylation
102	*MLL3/KMT2C*	ENSG00000055609	ENST00000262189	NM_170606	Histone methylation

	Gene	Ensembl	Transcipt ID	Reference Sequence	Function
103	*MPEG1*	ENSG00000197629	ENST00000361050	NM_001039396	Membrane protein/expressed in macrophages
104	*MTOR*	ENSG00000198793	ENST00000361445	NM_004958	mTOR signalling
105	*MYC*	ENSG00000136997	ENST00000377970	NM_002467	Transcription factor
107	*NF-KBIA*	ENSG00000100906	ENST00000216797	NM_020529	Inhibits NF-κB
108	*NF-KBIE*	ENSG00000146232	ENST00000275015	NM_004556	NF-κB
109	*NOTCH 1*	ENSG00000148400	ENST00000277541	NM_017617	Notch signalling/cellular differentiation and proliferation
110	*NOTCH 2*	ENSG00000134250	ENST00000256646	NM_024408	Notch signalling/cellular differentiation and proliferation
111	*NOTCH 3*	ENSG00000074181	ENST00000263388	NM_000435	Notch signalling/cellular differentiation and proliferation
112	*NOTCH 4*	ENSG00000204301	ENST00000375023	NM_004557	Notch signalling/cellular differentiation and proliferation
113	*P2RY8*	ENSG00000182162	ENST00000381297	NM_178129	Purinoceptor
114	*PAX5*	ENSG00000196092	ENST00000358127	NM_016734	B-cell differentiation/target of ASHM
115	*PCLO*	ENSG00000186472	ENST00000333891	NM_033026	Synaptic vesicles
116	*PDL1/CD274*	ENSG00000120217	ENST00000381577	NM_014143	T cell proliferation
117	*PIK3C2G*	ENSG00000139144	ENST00000433979	NM_004570	PI3K
118	*PIK3CD*	ENSG00000171608	ENST00000377346	NM_005026	PI3K pathway/immune response
119	*PIK3R1*	ENSG00000145675	ENST00000521381	NM_181523	PI3K
120	*PIM1*	ENSG00000137193	ENST00000373509	NM_001243186	Serine threonine protein kinase/cell proliferation/MYC
121	*PLCB1*	ENSG00000182621	ENST00000378641	NM_182734	PI3K
122	*POU2F2*	ENSG00000028277	ENST00000389341	NM_001207025	Transcription factor/IG gene expression
123	*PRDM1*	ENSG00000057657	ENST00000369096	NM_001198	Plasmacytic differentiation
124	*PRKCB*	ENSG00000166501	ENST00000303531	NM_002738	Serine threonine protein kinase/b cell activation
125	*PTEN*	ENSG00000171862	ENST00000371953	NM_000314	PI3K/AKT (Protein kinase B)
126	*PTPN6*	ENSG00000111679	ENST00000456013	NM_080549	Haematopoietic cell signalling
127	*RB1*	ENSG00000139687	ENST00000267163	NM_000321	Tumour suppressor/cell cycle/histone methylation
128	*REL*	ENSG00000162924	ENST00000295025	NM_002908	NF-κB
129	*ARHH*	ENSG00000168421	ENST00000610353	NM_001278359	Negative regulator of haematopoietic progenitor cell
130	*RRAGC*	ENSG00000116954	ENST00000373001	NM_001271851	mTOR
131	*S1PR2*	ENSG00000267534	ENST00000590320	NM_004230	Growth regulation
132	*SETD2*	ENSG00000181555	ENST00000409792	NM_014159	Histone methylation
133	*SGK1*	ENSG00000118515	ENST00000367858	NM_001143676	Serine threonine kinase/notch/NF-κB
134	*SLC22A16*	ENSG00000004809	ENST00000368919	NM_033125	Doxorubicin transporter
135	*SMARCA4*	ENSG00000127616	ENST00000429416	NM_001128844	Chromatin remodelling/NURF
136	*SMARCB1*	ENSG00000099956	ENST00000263121	NM_003073	Chromatin remodelling
137	*SOCS1*	ENSG00000185338	ENST00000332029	NM_003745	Jak-Stat
138	*SPEN*	ENSG00000065526	ENST00000375759	NM_015001	Transcriptional repressor/notch

	Gene	Ensembl	Transcipt ID	Reference Sequence	Function
139	*STAT3*	ENSG00000168610	ENST00000264657	NM_003150	Transcription activator
140	*STAT6*	ENSG00000166888	ENST00000300134	NM_001178078	Transcription activator
141	*SYK*	ENSG00000165025	ENST00000375754	NM_003177	B-cell receptor
142	*TAF1*	ENSG00000147133	ENST00000423759	NM_024665	Transcription factor complex/histone acetylation
143	*TBL1XR1*	ENSG00000177565	ENST00000457928	NM_024665	Transcription activator
144	*TCF3*	ENSG00000071564	ENST00000262965	NM_003200	Transcription regulation/b cell differentiation
145	*TET2*	ENSG00000168769	ENST00000380013	NM_001127208	DNA methylation
146	*TLR2*	ENSG00000137462	ENST00000260010	NM_003264	Toll like receptor/myd88
147	*TMEM30A*	ENSG00000112697	ENST00000230461	NM_018247	ATPase
148	*TMSB4X*	ENSG00000205542	ENST00000451311	NM_021109	Actin cytoskeleton/cell migration and proliferation
149	*TNFAIP3*	ENSG00000118503	ENST00000237289	NM_001270507	NF-κB
150	*TNFRSF14*	ENSG00000157873	ENST00000355716	NM_003820	Immune response
151	*TP53*	ENSG00000141510	ENST00000269305	NM_000546	Tumour suppressor/apoptosis
152	*TRAF3*	ENSG00000131323	ENST00000560371	NM_003300	NF-κB
153	*UBE2A*	ENSG00000077721	ENST00000371558	NM_001282161	Ubiquitination/transcription regulation
154	*XPO1*	ENSG00000082898	ENST00000401558	NM_003400	Nuclear export of proteins
155	*ZFHX3*	ENSG00000140836	ENST00000268489	NM_006885	Transcriptional repressor
156	*ZMYM3*	ENSG00000147130	ENST00000373998	NM_001171162	Chromatin structure
157	*NF-KBIZ*	ENSG00000144802	ENST00000326172	NM_031419	NF-κB
158	*HVCN1*	ENSG00000122986	ENST00000242607	NM_001040107	Voltage gated proton channel
159	*POU2AF1*	ENSG00000110777	ENST00000393067	NM_006235	Transcriptional coactivator
160	*VMA21*	ENSG00000160131	ENST00000330374	NM_001017980	Required for the assembly of the V0 complex of the vacuolar ATPase
161	*BCL2*	chr18:60985900-60987443		NM_030766	Regulate cell death

Appendix table 4: Permissions to use published figures in my Ph.D. thesis.

Publisher: Elsevier
Date: March 2015

Order Completed

Thank you for your order.

This Agreement between Omar A lishlash ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

Printable details.

License Number	4047531177143
License date	Feb 14, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	The Journal of Molecular Diagnostics
Licensed Content Title	High-Throughput Sequencing Using the Ion Torrent Personal Genome Machine for Clinical Evaluation of Somatic Hypermutation Status in Chronic Lymphocytic Leukemia
Licensed Content Author	Rebecca McClure,Ming Mai,Scott McClure
Licensed Content Date	March 2015
Licensed Content Volume	17
Licensed Content Issue	2
Licensed Content Pages	10
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	
Original figure numbers	Figure 1
Title of your thesis/dissertation	Role of AID in FL
Expected completion date	Jul 2017
Estimated size (number of pages)	400
Elsevier VAT number	GB 494 6272 12
Requestor Location	Omar A lishlash 42 St Andrew street Liverpool, L3 5XZ United Kingdom Attn: Omar A lishlash
Publisher Tax ID	GB 494 6272 12
Total	0.00 GBP

ORDER MORE CLOSE WINDOW

Title: Complex regulation and function of activation-induced cytidine deaminase
Author: Janet Stavnezer
Publication: Trends in Immunology
Publisher: Elsevier
Date: May 2011

Logged in as:
Omar lishlash
LOGOUT

Order Completed

Thank you for your order.

This Agreement between Omar A lishlash ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

Printable details.

License Number	4047291171012
License date	Feb 13, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	Trends in Immunology
Licensed Content Title	Complex regulation and function of activation-induced cytidine deaminase
Licensed Content Author	Janet Stavnezer
Licensed Content Date	May 2011
Licensed Content Volume	32
Licensed Content Issue	5
Licensed Content Pages	8
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	
Original figure numbers	Figure 1
Title of your thesis/dissertation	Role of AID in FL
Expected completion date	Jul 2017
Estimated size (number of pages)	400
Elsevier VAT number	GB 494 6272 12
Requestor Location	Omar A lishlash 42 St Andrew street Liverpool, L3 5XZ United Kingdom Attn: Omar A lishlash
Publisher Tax ID	GB 494 6272 12
Total	0.00 GBP

ORDER MORE CLOSE WINDOW

microenvironment in B-cell lymphomas

Author: Sarah E Coupland

Publication: Histopathology

Publisher: John Wiley and Sons

Date: Jan 24, 2011

Omar lishlash

Account #: 3001112598

LOGOUT

Order Completed

Thank you for your order.

This Agreement between Omar A lishlash ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

Printable details.

License Number	4047690593904
License date	Feb 14, 2017
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Histopathology
Licensed Content Title	The challenge of the microenvironment in B-cell lymphomas
Licensed Content Author	Sarah E Coupland
Licensed Content Date	Jan 24, 2011
Licensed Content Pages	12
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 3
Will you be translating?	No
Title of your thesis / dissertation	Role of AID in FL
Expected completion date	Jul 2017
Expected size (number of pages)	400
Requestor Location	Omar A lishlash 42 St Andrew street Liverpool, L3 5XZ United Kingdom Attn: Omar A lishlash
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing address	Omar A lishlash 42 St Andrew street Liverpool, United Kingdom L3 5XZ Attn: Omar A lishlash
Total	0.00 GBP

License Details

This Agreement between Omar A lishlash ("You") and Wolters Kluwer Health, Inc. ("Wolters Kluwer Health, Inc.") consists of your license details and the terms and conditions provided by Wolters Kluwer Health, Inc. and Copyright Clearance Center.

printable details

License Number	4047810184873
License date	Feb 14, 2017
Licensed Content Publisher	Wolters Kluwer Health, Inc.
Licensed Content Publication	Current Opinion in Hematology
Licensed Content Title	The routes for transformation of follicular lymphoma.
Licensed Content Author	Okosun, Jessica; Montoto, Silvia; Fitzgibbon, Jude
Licensed Content Date	Jul 1, 2016
Licensed Content Volume	23
Licensed Content Issue	4
Type of Use	Dissertation/Thesis
Requestor type	Individual Account
Portion	Figures/table/illustration
Number of figures/tables/illustrations	1
Figures/tables/illustrations used	Figure 1
Author of this Wolters Kluwer article	No
Title of your thesis / dissertation	Role of AID in FL
Expected completion date	Jul 2017
Estimated size(pages)	400
Requestor Location	Omar A lishlash 42 St Andrew street Liverpool, L3 5XZ United Kingdom Attn: Omar A lishlash
Publisher Tax ID	EU826013006
Billing Type	Invoice
Billing address	Omar A lishlash 42 St Andrew street Liverpool, United Kingdom L3 5XZ Attn: Omar A lishlash
Total	**0.00 GBP**

License Details

This Agreement between Omar A lishlash ("You") and Nature Publishing Group ("Nature Publishing Group") consists of your license details and the terms and conditions provided by Nature Publishing Group and Copyright Clearance Center.

printable details

License Number	4047821073504
License date	Feb 14, 2017
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Reviews Immunology
Licensed Content Title	Immunoglobulin class-switch DNA recombination: induction, targeting and beyond
Licensed Content Author	Zhenming Xu, Hong Zan, Egest J. Pone, Thach Mai and Paolo Casali
Licensed Content Date	Jul 1, 2012
Licensed Content Volume	12
Licensed Content Issue	7
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Box 1 | Activation-induced cytidine deaminase
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	Role of AID in FL
Expected completion date	Jul 2017
Estimated size (number of pages)	400
Requestor Location	Omar A lishlash 42 St Andrew street Liverpool, L3 5XZ United Kingdom Attn: Omar A lishlash
Billing Type	Invoice
Billing address	Omar A lishlash 42 St Andrew street Liverpool, United Kingdom L3 5XZ Attn: Omar A lishlash
Total	**0.00 GBP**