ANALYSIS OF THE GENOME AND VIRAL TRANSCRIPTOME OF EPSTEIN-BARR VIRUS ASSOCIATED PAEDIATRIC B CELL LYMPHOMAS

By **Pradeep Ramagiri**

A thesis submitted to the school of Cancer Sciences

of University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Institute of Cancer and Genomic Sciences

College of Medical and Dental Sciences

University of Birmingham

May 2017

ABSTRACT

Paediatric Hodgkin lymphoma (HL) and endemic Burkitt lymphoma (eBL) are Epstein-Barr virus (EBV) associated childhood malignancies.

In HL, malignant Hodgkin/Reed-Sternberg (HRS) cells, are characterised by aneuploidy, and comprise <1% of tumour which is challenging for global genetic studies. In chapter 3, I describe the mutational landscape of paediatric HL using in vitro and in silico methods I established, including tumour DNA amplification from low numbers of (~150) microdissected cells. I found that protein-altering mutations converge on mitotic spindle functions and that EBV contributes to transcriptionally downregulate mitotic spindle genes.

The contribution of the EBV to eBL pathogenesis remains poorly understood. In chapter 4, I show that the virus G protein-coupled receptor, BILF1 is apparently a latent gene expressed by tumour cells in a subset of BL. BILF1expression induced a transcriptional profile in primary human germinal centre (GC) B cells that partially recapitulates the transcriptional programme of EBV-positive eBL, and genes up-regulated by BILF1 have functions in oxidative phosphorylation and are also targets for MYC.

In summary, my data identify novel pathogenic pathways contributing to the pathogenesis of two common paediatric B cell malignancies that are likely to have clinical, including therapeutic, relevance.

DEDICATION

I dedicate my thesis to my parents, Bhagyalaxmi Ramagiri and Rajanna Ramagiri, for their enduring love and support.

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my supervisor Professor Paul Murray for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. A very special gratitude goes to my ex-supervisor Dr. Wenbin Wei for his enduring support, motivation and knowledge input. Without their precious support it would not be possible to conduct this research.

My sincere thanks go to Dr. Andrew Bell and Dr. Andrew Beggs, who provided me support, guidance, and access to their laboratory and research facilities. I also would like to thank Dr. Ghada Abdelsalam Ahmed and Dr. Katerina Vrzalikova for her contribution to the microdissection of HRS cells work and viral expression in germinal centre B cells work, respectively. I would also like to thank Paul Murray group members, especially my colleague and friend Robert Hollows, and staff at the Institute of Cancer and Genomic sciences for their support and encouragement.

Very special thanks goes out to all down at MRC (Medical Research Council) for providing Ph.D. fellowship, and at BCHRF (Birmingham Children's Hospital Research Foundation) for providing the necessary funds for this work. I would also like to thank CCLG (Children's Cancer and Leukaemia Group) for providing paediatric Hodgkin lymphoma samples. I also thank the University of Birmingham for providing me this opportunity, guidance and resources for this endeavour.

Finally, I would like to convey my gratitude to my friends and family members for their love and support throughout various stages of my life.

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

APPENDIX

Abbreviations

aa	amino acid
ABF1	activated B-cell factor 1
ABVD	Adriamycin, Bleomycin, Vinblastine, and Dacarbazine
aCGH	array based comparative genomic hybridization
α-KG	α-ketoglutarate
ALL	acute lymphoblastic leukaemia
amiGO	gene ontology data base
ASCT	autologous stem cell transplantation
AYAs	adolescents/young adults
BAM	Binary Alignment/Map format
BEACOPP	Bleomycin, Etoposide, Adriamycin, Cyclophosphamide, Oncovin, Procarbazine, Prednisone
BFM	Berlin- Frankfurt-Muenster group
BL	Burkitt lymphoma
Btk	Bruton tyrosine kinase
CALGB	Cancer and Leukemia Group B
CCND2	cyclin D2
CDK4	cyclin-dependent kinase 4
cDNA	complementary DNA
c-FLIP	cellular FLICE-inhibitory protein
c-FLIP	FLICE-inhibitory protein
cHL	classical HL
chr	chromosome

CIN	indicating chromosomal instability
CT	computerised tomography
CTAR1-3	C-terminal activating region 1-3
CtBP	C-terminal-binding protein
CTLs	CD8+ cytotoxic T lymphocytes
DA	dose-adjusted
DAB	3,3’-diaminobenzidine
DDR2	discoidin receptor family, member 2
DHAP	dexamethasone, high-dose Ara-C and cisplatin
DISC	death-inducing signalling complex
E	early
EBER1-2	EBV-encoded RNAs
EBF	B-cell factor
eBL	endemic Burkitt lymphoma
EBNA1-3	Epstein-Barr nuclear antigens 1-3
EBNA3A-C	Epstein–Barr nuclear antigens 3A-3C
EBV	Epstein-Barr virus
ENO1	enolase 1
EPHB1	ephrin receptor EphB1
EPOCH-R	etoposide, prednisone, vincristine, cyclophosphamide, doxorubicin, rituximab
ESR	erythrocyte sedimentation rate

EuroNet-PHL-C2	European Network-Paediatric Hodgkin Lymphoma Study Group Second International Inter-Group Study for Classical Hodgkin Lymphoma in Children and Adolescents
FFPE	formalin-fixed paraffin-embedded
Fp	F promoter
GC	germinal centre
GEO	Gene Expression Omnibus
GLUT1	glucose transporter
GLS	glutaminase
GMALL	German Multicenter Study Group for Adult ALL
GO	gene ontology
Gp	glycoprotein
GPCR	G-protein-coupled receptor
H&E	Haematoxylin and Eosin
HDACs	histone deacetylases
HIV	human immunodeficiency virus
HK2	hexokinase 2
HL	Hodgkin lymphoma
HLA	human leukocyte antigen
HRS	Hodgkin/Reed-Sternberg
HSC	haematopoietic stem cells
IAP	inhibitors of apoptosis
ICE	ifosfamide, carboplatin and etoposide
ICTV	International Committee on Taxonomy of Viruses

ID2	inhibitor of differentiation and DNA binding 2
IE	immediate early
IFRT	involved field radiotherapy
IKK	IkB kinase complex
IL	interleukins
IPS	international Prognostic Score
ISH	in situ hybridisation
ITAM	immuno-receptor tyrosine-based activation motif
IκBβ	NF-κB inhibitors
Kb	kilo base pair
K-Da	kilo Dalton
L	late lytic cycle
LCLs	lymphoblastoid cell lines
LCM	laser capture microdissection
LCV	lymphocryptovirus
LDCHL	lymphocyte-depleted classical HL
LDHA	lactate dehydrogenase A
LMP	latent membrane proteins
LP	lymphocyte-predominant cells
LRCHL	lymphocyte-rich classical HL
MCCHL	mixed cellularity classical HL
MDFNE	median distance from the nearest end
miRNA	microRNA
MTX	methotrexate

MYB	myeloblastosis gene
NCI	National Cancer Institute
NIH	National Institutes of Health
NILG	Northern Italy Leukemia Group
NLPHL	nodular lymphocyte predominant HL
NPC	nasopharyngeal carcinoma
NSCHL	nodular sclerosis classical HL
ORF	open reading frame
OriLyt	lytic origin of replication
OriP	The origin of plasmid replication
PDGFRA	platelet-derived growth factor receptor A
PDH	pyruvate dehydrogenase.
PDK1	pyruvate dehydrogenase kinase 1
PEN	polyethylene naphthalate
PET	positron emission tomography
PETHEMA	Programa Español de Tratamientos en Hematologia.
PFKM	phosphofructokinase
PKB	protein kinase B
PKM2	pyruvate kinase M2
PKR	protein kinase R
PTLD	post-transplant lymphoproliferative disease
Qp	Q promoter
REAL	revised European-American lymphoma classification
RIP	receptor-interacting protein

RON	macrophage-stimulating protein receptor
SAM	Sequence Alignment/Map format
SCAs	segmental chromosomal aberrations
SFKs	Src-family protein tyrosine kinases
SKIP	Ski-interacting protein
SRA	sequence read archive
SYK	spleen tyrosine kinase
TCA	tricarboxylic acid cycle
TES	transformation effector sites
TMCs	tonsillar mononuclear cells
TMM	trimmed mean of M values
TNFR	tumour necrosis factor receptor
TR	terminal repeats
TRADD	TNFR-associated death domain
TRKA/B	tyrosine kinase receptor A/B
USF	upstream stimulatory factor
UTR	untranslated regions
vBCL2	viral homolog of human BCL2 gene
VCA	viral capsid antigen
vIL-10	viral IL-10
WGA	whole genome amplification
WHO	World Health Organization
Wp	W-promoter
XIAP	X-linked inhibitor of apoptosis

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

1.1 The Epstein-Barr virus

The Epstein-Barr virus (EBV) infects approximately 95% of humans. Infection usually occurs in early childhood, and then persists for the rest of an individual's life with no symptoms (Baer et al., 1984, Fleisher et al., 1979). More than 200,000 new cases of EBV-associated cancer are registered each year (2% of all cancers), and these malignancies account for about 1.8% of all cancer deaths world-wide (zur Hausen and de Villiers, 2015, de Martel et al., 2012, Khan and Hashim, 2014). EBV is causally associated with several types of cancer, including B cell tumours such as Burkitt lymphoma (BL) and Hodgkin lymphoma (HL); and epithelial cancers, including nasopharyngeal carcinoma (NPC) and gastric cancer (Rickinson and Kieff, 2007). EBV-associated tumours are more prevalent in the developing world compared with developed regions and contribute to cancer development in immunosuppressed patients (de Martel et al., 2012).

1.2 The discovery of Epstein-Barr virus

The events that led to the discovery of Epstein–Barr virus (EBV) as the first human oncovirus began on 22 March 1961 when Anthony Epstein, a young pathologist working at the Middlesex Hospital in London, attended a lecture entitled 'The commonest children's cancer in tropical Africa: a hitherto unrecognized syndrome', given by Dennis Burkitt, a surgeon working in Uganda. The disease in question was an unusual extranodal tumour of the jaws of African children, which later known as Burkitt lymphoma (BL) (Burkitt, 1958). Following discussions with Epstein, Burkitt agreed to send fresh tumour biopsies to Epstein's laboratory for analysis by electron microscopy to identify the putative virus resident within BL cells

(Crawford et al., 2014). After a series of unsuccessful attempts to culture BL cells, Epstein and Yvonne Barr accidentally found viable tumour cells floating free of the main lymphoma mass in a delayed shipment of a BL biopsy sent by Burkitt in December 1963. Epstein and Barr grew these tumour cells in fresh culture medium, and thus the first cell line (now named the EB1 cell line after Epstein and Barr) from a human lymphoma was established in 1964 (Epstein and Barr, 1964). Analysis of the EB1 line by electron microscopy revealed herpes-like virus particles (Epstein, 1999). Unlike the three other human herpesviruses known at that time, this virus was biologically inert for all activities associated with known herpesviruses. It was formally identified as a new herpes virus in 1965 (Epstein, 1979).

In the following years, a serological assay was developed against the viral capsid antigen (VCA) to screen for EBV infection. This screening assay revealed the presence of EBV infection in 90% of the adult human population (Henle et al., 1969).

1.3 Classification and structure of EBV

EBV (also known as human herpes virus 4) is a lymphotropic virus and belongs to the family *Herpesviridae.* The International Committee on Taxonomy of Viruses (ICTV) revised the family *Herpesviridae* in 2016, which was incorporated into the new order *Herpesvirales.* Family *Herpesviridae*, which retains the mammal, bird and reptile viruses, is sub-divided into *Alphaherpesvirinae*, *Betaherpesvirinae* and *Gammaherpesvirinae* based on genome sequence homology. EBV is the only human *Lymphocryptovirus* (LCV) of the *Gammaherpesvirinae* and like other LCVs EBV can establish a latent infection (Bilello et al., 2006, Davison et al., 2009, Marr-Belvin et al., 2008).

EBV has a toroid-shaped protein-core wrapped with viral genome which is encapsulated by an icosahedral nucleocapsid composed of 162 capsomers containing tubular 160, 47, and 28 k-Da major viral proteins as well as a number of minor virion proteins. A 152 k-Da protein

tegument, containing a mixture of cellular and virally encoded proteins, is found between the nucleocapsid and an inner envelope as well as external glycoproteins on the surface of the outer envelope. The protein content of the host derived envelope is dominated by the virally encoded 350 or 220 k-Da glycoprotein (gp350/220) (Dolyniuk et al., 1976a, Dolyniuk et al., 1976b, Johannsen et al., 2004).

The complete sequence of an EBV strain (the B95-8 strain) was published in 1984 (Baer et al., 1984). Sequencing of the EBV genome was carried out on a *Bam* HI fragment cloned library which revealed that it is composed of 172.76 kilo base (kb) of linear double stranded DNA and contains 60% GC residues (Figure 1). *Bam* HI fragments containing open reading frames, genes and sites for transcription or RNA processing are depicted in descending order of fragment size as shown in the Figure 2. Subsequent studies identified 82 EBV genes encoding 80 proteins, and 2 non-coding RNAs; however the number of genes and transcripts are regularly updated and are likely to increase due the advances in sequencing technology (Dolan et al., 2006, O'Grady et al., 2016).

Coordinated expression of a subset of these viral genes, including the so-called latent genes, is responsible for in vitro B cell immortalization. Sequencing of various virus strains revealed that the prototypic B95-8 sequence was unusual in that it had a 12 kb deletion in the BamHI A region. In 2003, de Jesus et al., replaced the 12-kb deletion in the B95-8 EBV sequence using the sequences obtained from the Raji strain (de Jesus et al., 2003). The latest reference EBV sequence is available at NCBI (GenBank accession number: AJ507799.2; last update on 26-JUL-2016). Although it is widely believed that slow evolutionary drift is responsible for genetic polymorphisms in the viral genome, the origin of the two EBV types, type 1 and type 2, remains enigmatic as both types are found in the oropharynx of most human populations. These subtypes differ mainly in the sequences of the Epstein-Barr nuclear antigen 2 (EBNA2)

and EBNA3 genes (Baer et al., 1984, Dambaugh et al., 1984). Type 1 EBV is more prevalent worldwide except in areas endemic for BL (equatorial Africa and New Guinea) and has greater transforming potential (Barton et al., 2007, Zimber et al., 1986) than Type 2 EBV (Young et al., 1987, Sixbey et al., 1989).

Despite a high degree of sequence similarity between EBV strains, additional variations exist in some viral genes, mainly in the latent genes, which give rise to functional differences (Table 1).

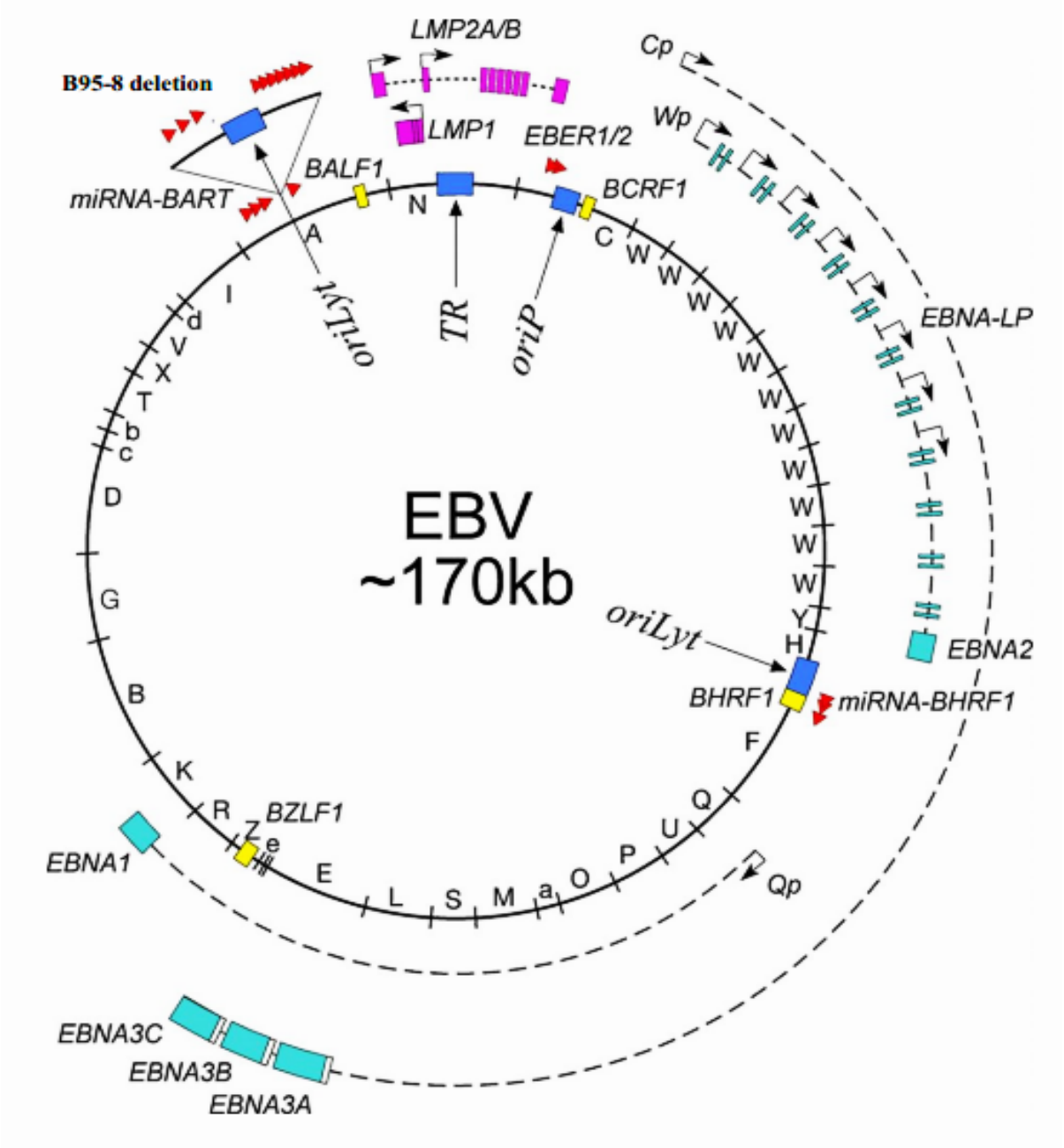

Figure 1. Schematic representation of the EBV genome. The origin of plasmid replication (OriP), Terminal repeats (TR) and origin of lytic replication (OriLyt) are shown in blue boxes. Exons encoding latent proteins: Epstein–Barr nuclear antigens (EBNA1, 2, 3A, 3B and 3C, and EBNA-LP) are shown in light blue boxes, latent membrane proteins (LMP1, 2A and 2B) are shown in pink boxes. Lytic cycle genes BALF1, BCRF1, BHRF1, and BHRF1 are shown in yellow boxes. Highly transcribed EBER1 and 2, and the EBV encoded microRNAs are depicted as red arrows (taken from Kalla and Hammerschmidt, 2012).

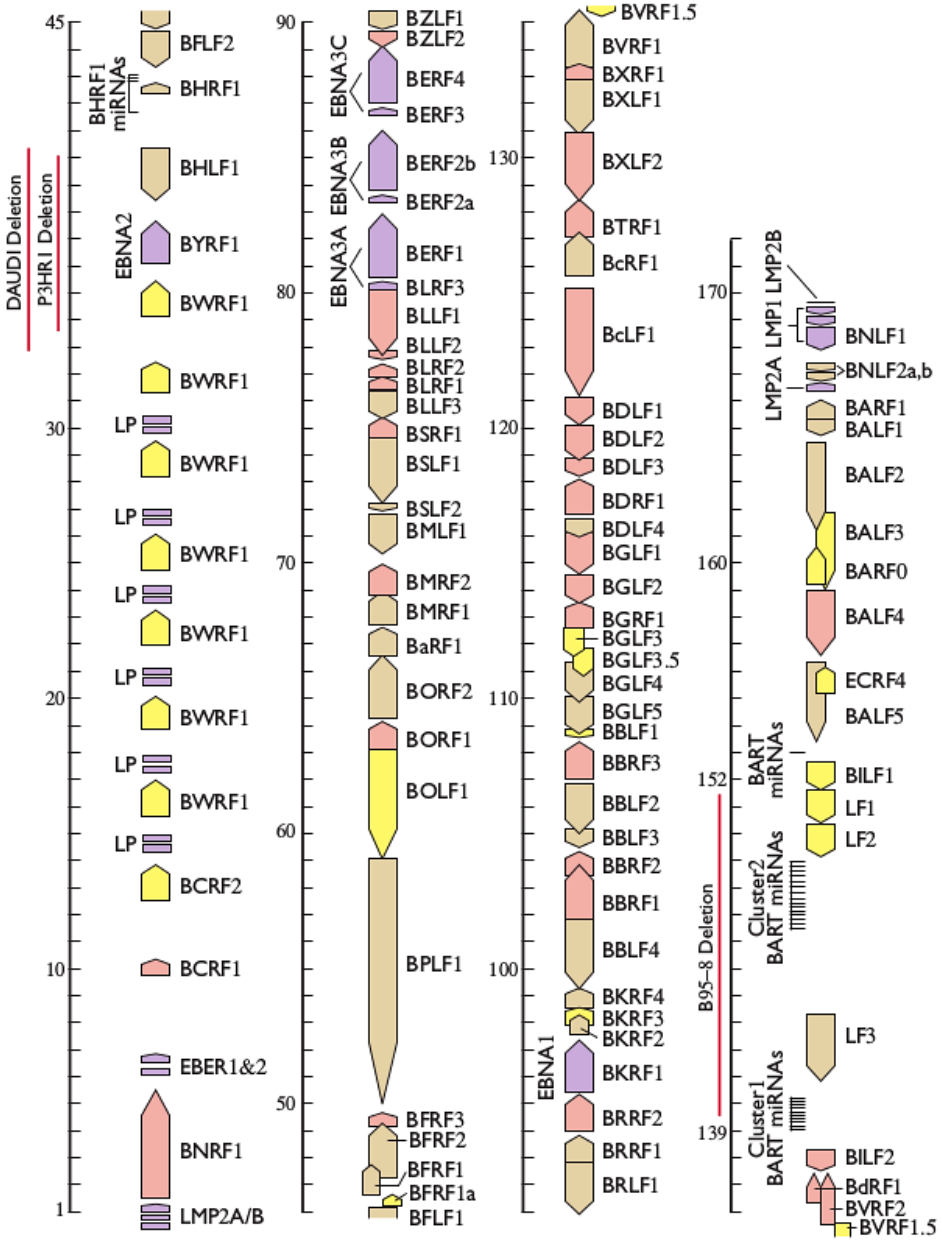

Figure 2. Schematic depiction of the linear EBV genome with open reading frames (ORFs). EBV ORFs are named based on their location (ORF1, ORF2, etc.) and direction (rightward or leftward) within the Bam H1fragment (A to Z, in decreasing size). The EBV genes are colour coded as: latent gene ORFs or exons (purple, the early lytic ORFs (beige), late lytic ORFs (red). Yellow arrows depict ORFs of unknown gene class or are hypothetical. LMP2A and LMP2B are encoded across the circularized genome. The location of microRNA (miRNAs) and the noncoding RNAs EBER1and EBER2 encoded by the genome is also shown. Numbers refer to kilo-base pairs. Correction: BILF1, depicted as uncharacterized gene here in yellow is now known to be express during lytic cycle (taken from Longnecker and Kieff, 2013).

Table 1. Reported EBV sequences. BL, Burkitt lymphoma; EBV-GC, EBV-associated gastric carcinoma; HL, Hodgkin lymphoma; IM, infectious mononucleosis; LCLs, lymphoblastoid cell lines; NPC, nasopharyngeal carcinoma; PTLD, post-transplant lymphoproliferative disease (taken from Young et al., 2016).

EBV strain	**Type**	**Origin**	**Disease**	**Year**	**References**
B95-8/Raji	1	USA and Nigeria	IM and BL	2003	Sample et al., 1990
GD1	1	China	NPC	2005	de Jesuset al., 2003
AG876	2	Ghana	BL	2006	Zeng et al., 2005
GD2	1	China	NPC	2011	Dolan, 2006
HKNPC1	1	Hong Kong	NPC	2012	Liu et al., 2011
Akata	1	Japan	BL	2013	Kwok et al., 2012
Mutu	1	Kenya	BL	2013	Kwok et al., 2012
M81	1	Hong Kong	NPC	2013	Lin et al., 2013
K4123-Mi and K4413-Mi	1	USA	LCLs	2013	Tsai et al., 2013
HKNPC2–HKNPC9	1	Hong Kong	NPC	2014	Lei et al., 2013
NA19114, NA19315 and NA19384	1	USA	LCLs	2014	Kwok et al., 2014
71 different EBV strains	60 type 1 and 11 type 2	UK, Germany, Hong Kong, China, Japan, S.Korea, Australia, Kenya and Nigeria	Healthy saliva, LCLs, PTLD, HL, BL, NPC and EBV-GC	2015	Santpere et al., 2014

1.4 EBV induced B cell growth transformation *in vitro*

The discovery of the ability of EBV to infect and transform resting B cells in vitro was critical to understanding its role as an oncovirus (Henle et al., 1967, Pope et al., 1968). EBV glycoproteins gp350 and gp42 bind to the CD21 receptor and human leukocyte antigen (HLA) class II molecules on the surface of B cells, resulting in infection followed by conversion of primary B cells to continuously proliferating lymphoblastoid cell lines (LCLs). LCLs usually express latency III or the 'growth programme'; where multiple copies of extrachromosomal circular genetic material known as episomes express all latent genes (Figure 1), including six Epstein–Barr nuclear antigens (EBNAs 1, 2, 3A, 3B and 3C and EBNA leader protein (EBNA-LP)), latent membrane proteins LMP1 and LMP2 (which encodes two isoforms, LMP2A and LMP2B) and the non-coding EBV-encoded RNAs (EBER1 and EBER2), and viral microRNA (miRNA) (Rickinson and Kieff, 2007) .

In LCLs, differential splicing of the same primary transcript expressed from either the Cp or Wp promoters, located in the *Bam* HI C and W fragment of the EBV genome, respectively, generates all EBNA mRNAs (Longnecker et al., 2013, Speck and Strominger, 1985). The LMP1 promoters, the proximal ED-L1 and the distal TR-L1, are regulated by different viral and cellular factors and can be transcribed into LMP1 mRNA (Kis et al., 2010, Noda et al., 2011, Goormachtigh et al., 2006). For example, EBNA1 can stimulate LMP1 transcription from both promoters through a distal enhancer (Gahn and Sugden, 1995), whereas EBNA2 can only activate ED-L1 (Johannsen et al., 1995). Cellular proteins such as NOTCH1, IRF7, STATs, upstream stimulatory factor (USF), and ATF-1/cAMP responsive element binding protein (CREB)-1 can also activate the ED-L1 promoter (Chang et al., 2000). Cellular transcription factors, including SP1/SP3 and STATs can activate the TR-L1 promoter (Tsai et al., 1999).

Two separate promoters are involved in the transcription of LMP2A and LMP2B genes; the LMP2A promoter is present directly upstream of the first coding exon while the LMP2B promoter shares a bidirectional promoter with LMP1. LMP2A and LMP2B promoters are separated by 3 kb. These two isoforms differ only in their first exons. The exon 1 of LMP2A encodes a 119-amino-acid-N-terminal tail. The exon 1 of LMP2B is non-coding (Laux et al., 1988, Laux et al., 1989, Longnecker and Kieff, 1990, Sample et al., 1989). A novel LMP2 transcript originating in the TRs has been described recently in NK/T cell lymphomas (Fox et al., 2010).

High levels of B cell activation markers such as CD23, CD30, CD39, and CD70: and cellular adhesion molecules such as FA1, LFA3, and ICAM1, similar to that of activated B cells, have been observed in LCLs, suggesting that EBV-induced immortalization in LCLs can be achieved through the constitutive activation of cellular pathways that normally drive B cell proliferation (Rowe et al., 1987a, Gregory et al., 1988, Young et al., 2000).

1.5 EBV infection and persistence *in vivo*

Although it is now well established that EBV persists as a latent infection within the memory B cells, it has been suggested that primary infection occurs in mucosal epithelial cells of the oropharynx and upon viral replication EBV spreads to B cells (Morgan et al., 1979). However, it has also been suggested that EBV infection of epithelial cells could also involve mechanism distinct from that required for B cell infection (Shannon-Lowe et al., 2009).

NK cells, latent-antigen-specific CTLs and CD4 T helper cells recognize the B cells expressing a latency III programme (Williams and Greene, 2005, Hislop et al., 2007a). However, a fraction of these B cells can down-regulate the expression of most viral antigens, thereby evading immune recognition and establishing a pool of resting EBV-positive memory B cells, as shown in Figure 3 (Young and Rickinson, 2004). Two mechanisms have been

proposed to explain how this occurs. In the first model, EBV infects naïve B cells and drives naive cells into memory by mimicking the physiological process of antigen-driven memory cell development in lymphoid tissues, a process involving somatic immunoglobulin gene hyper-mutation during transit through a germinal centre. In the second model, EBV directly infects pre-existing memory cells as a direct route into memory. EBV-infected memory B cells (which express a so called latency 0 pattern of EBV gene expression) act as a reservoir and might be recruited into germinal centres of secondary follicles and switch into the viral latency II programme. EBV-positive GC B cells receive the necessary signals for antigen-independent proliferation with the help of EBV-encoded surrogate T cell signals (provided by LMP1) and B-cell receptor (BCR) engagement stimulus (provided by LMP2A) (Caldwell et al., 1998, Caldwell et al., 2000, Roughan and Thorley-Lawson, 2009). Subsequently, the EBV-infected B cells might either re-enter the reservoir as memory cells (latency 0) or commit to plasma-cell differentiation and move to mucosal sites in the oropharynx and, in the process, activate the viral lytic cycle.

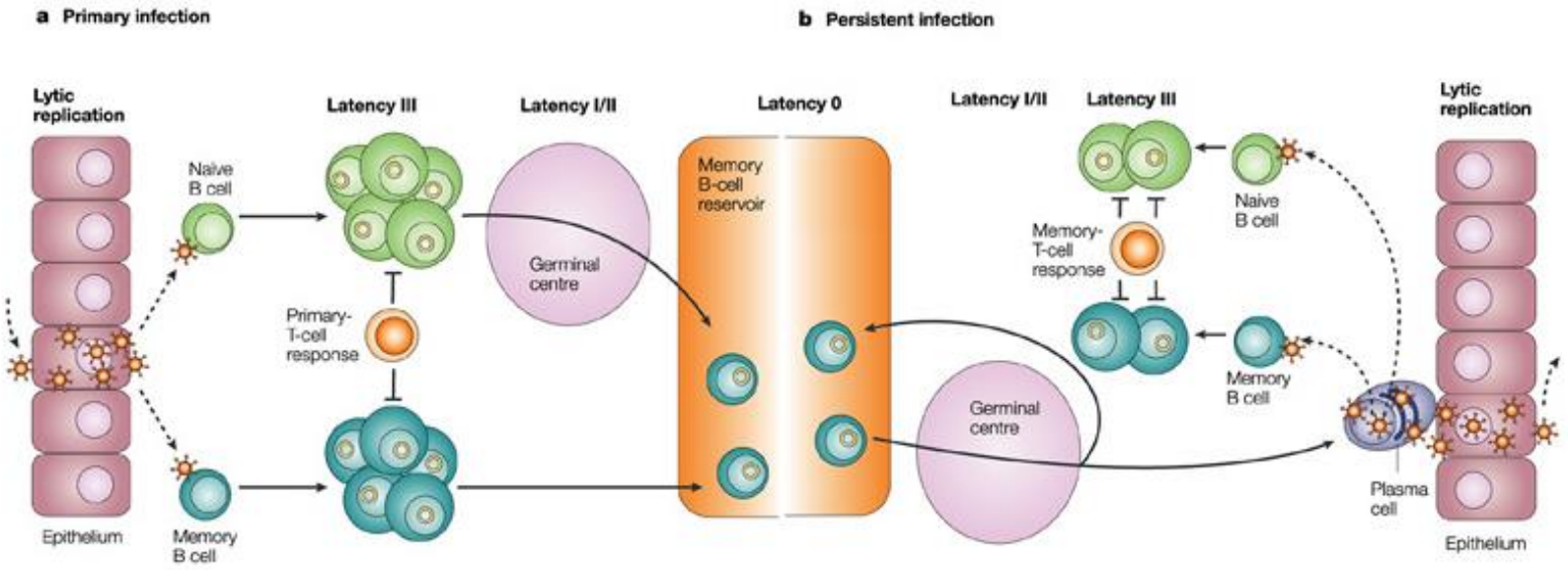

Figure 3. Primary infection and persistent infection of EBV *in vivo*. a) Primary infection. Upon infection EBV establishes initial lytic replication in the oropharynx, after which the virus spreads throughout the lymphoid tissues as a growth-transforming (latency III) infection of B cells. Latent-antigen-specific primary-T-cell response removes many of these proliferating B cells, however some B cells escape by downregulating antigen expression. These B cells then establish a stable reservoir of resting EBV-positive memory B cells, in which viral antigen expression is mostly suppressed (latency 0). Two possible mechanisms of B cell infection are shown: EBV infects naïve B cells and drives naive cells into memory by mimicking the physiological process of antigen-driven memory-cell development in lymphoid tissues, a process involving somatic immunoglobulin-gene hyper-mutation during transit through a germinal centre or alternatively EBV might directly infect pre-existing memory cells as a direct route into memory. b) Persistent infection. EBV-infected memory B cells acts as a reservoir and might be recruited into germinal-centre reactions, entailing the activation of different latency programmes. Subsequently they might either re-enter the reservoir as memory cells or commit to plasma-cell differentiation and move to mucosal sites in the oropharynx and, in the process, activating the viral lytic cycle (taken fromYoung and Rickinson, 2004).

1.6 Alternative forms of EBV latent infection

Figure 4 taken from Rowe et al., 2009) shows the four different forms of EBV latent infection.

Latency 0

In this form of latency, EBV persists in memory B cells of EBV-infected healthy humans where all latent genes with the exception of EBERs and the viral miRNA are suppressed (Babcock et al., 1999). Upon reactivation of EBV latent transcription, memory B cells enter germinal centres and express low levels of LMP1, LMP2A, and Qp-initiated EBNA1 (Babcock and Thorley-Lawson, 2000) (Figure 4).

Latency I

Latency I viral expression is observed in EBV-positive BL and BL derived EBV-positive cell lines (Rowe et al., 1987b) Latency I is also a restricted viral latent gene expression pattern, characterized by the activity of the Q promoter (Qp) and suppression of Cp, Wp, and possibly LMP promoter activity (Schaefer et al., 1995, Nonkwelo et al., 1996). In this latency form, non-coding EBERs, and EBV-encoded micro RNAs (BARTs) (Cai et al., 2006) are expressed along with the EBNA1 protein from the Qp initiated Q-U-K transcripts.

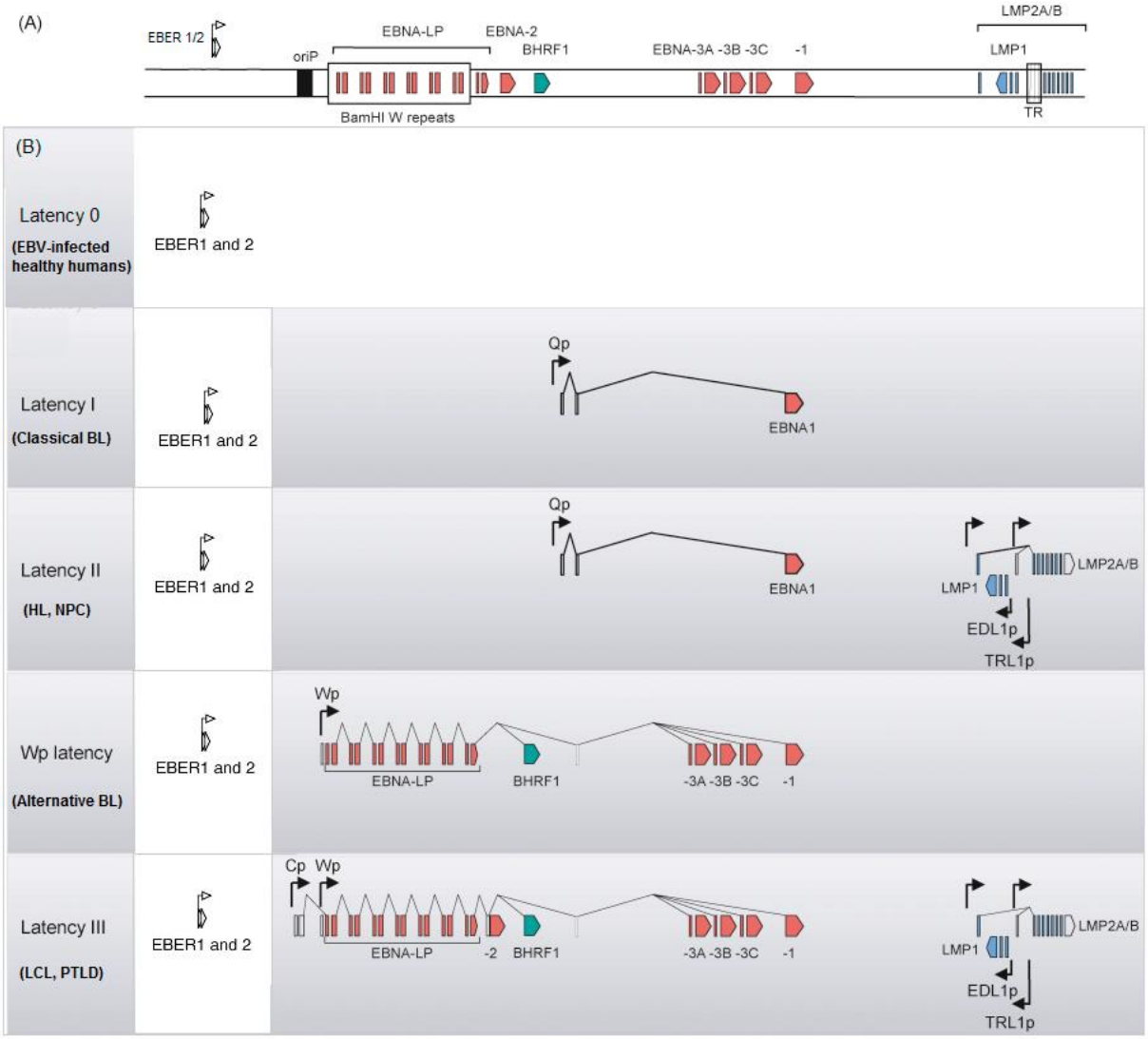

Figure 4. Schematic diagram of EBV latent viral transcription in the different forms of latency. (A) Schematic illustration of the EBV genome showing the location of coding exons for the nuclear antigens: EBNA1, 2, 3A, 3B, 3C, and EBNA-LP; ORF for BHRF1; and the three latent membrane proteins LMP1, 2A and 2B. Also shown are the latent origin of replication (oriP), and the terminal repeats (TR). Shown below (B) are the structures of viral mRNA transcripts expressed in different forms of latency. Promoters are identified by arrowheads, coding exons are coloured and non-coding exons are unshaded. The BART microRNAs are expressed in all latency forms (not shown) (adapted from Rowe et al., 2009).

Latency II

Latency II is another restricted form of viral latency and is observed in 'normal' EBV-positive GC B cells and in EBV-positive cancers such as HL, NPC, and gastric cancer (Fahraeus et al., 1988, Young et al., 1988). Latency II infected cells express variable levels of LMP1 and LMP2 along with the non-coding EBERs, viral miRNA (BARTs), and EBNA1 (Brooks et al., 1992, Brooks et al., 1993, Kerr et al., 1992, Deacon et al., 1993).

Wp latency

An alternative pattern of viral gene expression was identified in a minority of BL cell lines, which have deletions of around 6–8 kb that span the EBNA2 gene, and in their corresponding tumour biopsies (Kelly et al., 2002a, Kelly et al., 2005, Rowe et al., 2014a). In this form of latency, the W-promoter (Wp) is active and drives expression of all EBNAs except EBNA2; instead of Q-promoter (Qp)-driven EBNA1 expression. Importantly, latent viral gene expression in these Wp-restricted BLs, several latent viral genes are expressed including high levels of the BHRF1 gene encoding a viral homolog of human BCL2 gene (vBCL2) that was previously thought to be exclusively a lytic cycle protein (Kelly et al., 2009b, Rowe, et al. 2014). Interestingly, gene expression in Wp-restricted BL cell lines seems to be derived entirely from the deleted genomes while the wild-type genomes remain silenced (Kelly et al., 2005).

Latency III

In this form of latency, in addition to the genes expressed in the latency II, the remaining 5 EBV nuclear antigens: EBNA2, 3A, 3B, 3C, and LP are expressed from large highly spliced transcripts initiated from Cp or Wp promoters (Alfieri et al., 1991, Brooks et al., 1993). This

type of latency is observed in LCLs and post-transplant lymphoproliferative disease (PTLD) (Young et al., 1989, Gratama et al., 1991).

1.7 Epstein-Barr virus latent gene products

Epstein-Barr nuclear antigen 1 (EBNA1)

EBNA1 is a nuclear antigen encoded by the EBV BKRF1 ORF. EBNA1 binding to the plasmid origin of viral replication (OriP) on the EBV genome is essential for EBV replication during latency (Lupton and Levine, 1985, Yates et al., 1985). EBNA1 is the most ubiquitously expressed EBV-encoded protein that is found in all forms of viral infection, with the exception of latency 0 (Reedman and Klein, 1973, Babcock et al., 1999). During cell division, EBNA1 also ensures the partitioning of viral genomes by tethering viral episomes to host chromosomes (Marechal et al., 1999, Sears et al., 2004). In latency I and II infections, EBNA1 transcripts are initiated from the Qp promoter, while in latency III infection, EBNA1 transcripts are initiated by splicing of the large Cp or Wp promoter initiated transcripts (Bodescot et al., 1986, Sample et al., 1991, Speck and Strominger, 1985). EBNA1 is also expressed during the lytic cycle, where lytic EBNA1 transcripts are directly initiated from the lytic F promoter (Fp) which is located upstream of Qp. Recent studies have suggested that EBNA1 can regulate EBER expression by increasing pol III transcription following its binding to and activation of ATF2, c-MYC and TFIIIC (Owen et al., 2010).

Evidence provided by Humme et al., contradicts the indispensable role of EBNA1 in the EBV-induced transformation of B cells. This group managed to grow LCLs, albeit less efficiently, from an EBNA1-deficient EBV mutant (Humme et al., 2003). Although EBNA1 has been shown to enhance tumour growth in NOD-SCID mice following its expression in

HL cells (Kube et al., 1999), conflicting reports of the role of EBNA1 in oncogenesis exist where other researchers have not able to reproduce this oncogenic effect (Wilson et al., 1996, Kang et al., 2005, Tsimbouri et al., 2002). Furthermore, inhibition of EBNA1 by dnEBNA1 or RNAi reduced the proliferation and survival of EBV-positive BL cells (Kennedy et al., 2003, Hong et al., 2006).

Epstein-Barr nuclear antigen 2 (EBNA2)

EBNA2, which is encoded by the BYRF1 ORF, is one of the first viral proteins to be expressed during the transformation of primary B cells by EBV. EBNA2 and EBNA-LP activate the adjacent Cp promoter following their expression from Wp (Bodescot et al., 1987, Evans et al., 1996, Jin and Speck, 1992).

The Cp promoter acts as the dominant promoter for expression of all six EBNAs; EBNAs 1, 2, 3A, 3B and 3C and EBNA-LP (Speck and Strominger, 1985, Woisetschlaeger et al., 1989). The growth transforming ability of the EBNA2 protein was demonstrated by the inability of the P3HR-1 strain, in which the EBNA2 gene and the last two exons of EBNA-LP are deleted, to transform B cells *in vitro* (King et al., 1982, Rickinson and Kieff, 2001, Skare et al., 1985). P3HR-1 showed full EBV growth transforming ability after the restoration of EBNA2 gene into its genome (Hammerschmidt and Sugden, 1989, Cohen et al., 1989).

EBNA2 acts as a potent and specific transcriptional activator of both viral and cellular genes mainly through three domains, the RBP-Jk binding domain, the acidic activation domain, and the homotypic association domains (Rickinson and Kieff, 2001)). RBP-Jk, either alone or with Ski-interacting protein (SKIP), binds EBNA2 to its responsive elements upstream of cellular and viral promoters, as shown in LCLs where it has been involved in the regulation of Cp (Bodescot et al., 1987, Evans et al., 1996), as well as the up-regulation of LMP1 (Abbot et

al., 1990, Wang et al., 1990a), LMP2 (Zimber-Strobl et al., 1993), and the cellular genes CD21 (Cordier et al., 1990), CD23 (Wang et al., 1990a), and c-Fgr (Knutson, 1990). The RBP-Jk homolog in Drosophila is involved in signal transduction from the NOTCH receptor, which has also been linked to the development of T cell tumours in humans (Artavanis-Tsakonas et al., 1995). EBNA2 can mimic a constitutively activated NOTCH receptor (Rabson et al., 1982) leading to the transactivation of viral and cellular downstream targets. EBNA2 also targets the transcription of cellular genes that are likely to be important for EBV-induced B cell proliferation; these include c-MYC, cyclin D2 (CCND2), and cyclin-dependent kinase-4 (CDK4) (Jayachandra et al., 1999, Kaiser et al., 1999). EBNA2 also shown to induce the formation of new binding for both RBP-jκ and EBF1, two master regulators of B-cell fate. A strong correlation between these newly induced binding sites co-occupied by EBNA2-EBF1-RBP-jκ and transcriptional activation of genes involved in the B-lymphoblast growth and survival was observed (Lu et al., 2016).

The Epstein-Barr nuclear antigen 3 (EBNA3) family members

The EBNA3 family members are the least abundant EBNA mRNAs in latently infected cells and are transcribed from the Cp or Wp promoters. The EBNA3A, 3B, and 3C proteins can interact with several cellular factors involved in transcriptional repression or silencing, including histone deacetylases (HDACs), and C-terminal binding protein (CtBP) (Hickabottom et al., 2002, Radkov et al., 1999). All three EBNA3 proteins contain a conserved region which is able to stably bind RBP-Jk and compete with EBNA2 and Notch for RBP-Jk, thus repressing the EBNA2-mediated transactivation by disrupting RBP-Jk binding to Notch and to EBNA2 (Johannsen et al., 1996, Robertson and Ambinder, 1997). EBNA3A and EBNA3C, but not EBNA3B, are reported to be essential for B cell

transformation in vitro (Maruo et al., 2009, Lee et al., 2009). However, Hertle et al showed transformation of B cells by EBNA3A-deleted EBV, albeit with reduced efficiency (Hertle et al., 2009). Both EBNA3A and EBNA3C can transform embryonic rat fibroblasts by co-operating with the *ras* oncogene (Parker et al., 1996). Recent studies have shown that EBNA3C can inhibit the accumulation of the CDK inhibitors p27KIP1 and p16INK4A, and mediates Rb degradation via the SCF ubiquitin ligase. Furthermore, EBNA3C also shown to stabilizes c-myc, and also involved in the inactivation of cyclinA-CDK complexes (Bajaj et al., 2008, Saha et al., 2009, Saha et al., 2011).

A recent study showed that EBNA3B might act as a tumour suppressor gene, since infection with an EBNA3B-knockout virus (EBNA3BKO) led to aggressive, immune-evading monomorphic DLBCL-like tumours in NOD/SCID/γc-/- mice. Furthermore, EBNA3BKO-infected B cells showed faster expansion and reduction in the secretion of T cell-chemoattractant CXCL10, reducing T cell recruitment in vitro and T cell-mediated killing in vivo (White et al., 2012).

Epstein-Barr nuclear antigen leader protein (EBNA-LP)

EBNA-LP is the first latent protein to be expressed during B cell transformation by EBV (Alfieri et al., 1991), and its mRNA is transcribed from C1 C2 or W0 exons initiated by Cp or Wp promoters followed by transcription from a variable number of W2W1 repeats and unique Y1 and Y2 exons. Although EBNA-LP is not absolutely essential for the *in vitro* transformation of B cells, it is required for the efficient outgrowth of LCLs (Allan et al., 1992). EBNA-LP can potentiate the effects of EBNA2-mediated transcriptional regulation, as transient transfection of EBNA-LP and EBNA2 into primary B cells induced transition of cells from G0 to G1 via up-regulation of CCND2 expression (Sinclair et al., 1994). EBNA-LP

also enhances the EBNA2 transactivation of its downstream targets, including LMP1 (Harada and Kieff, 1997, Nitsche et al., 1997).

Latent membrane protein 1 (LMP1)

LMP1 is a 66 K-Da integral membrane protein encoded by the BNLF1 ORF and is essential for the EBV-induced transformation of B cells. LMP1 has been shown to transform rat fibroblasts and induce tumours in transgenic mice (Kaye et al., 1993, Moorthy and Thorley-Lawson, 1993, Wang et al., 1985). LMP1 is a plasma membrane protein that acts as a constitutively active tumour necrosis factor receptor (TNFR), similar to CD40 or a combination of TNFRI and TNFRII (Izumi and Kieff, 1997, Mosialos et al., 1995). LMP1 expression is critical for the transformation of B cells into LCLs and also for the growth and survival of LCLs, as loss of LMP1 expression in LCLs has been shown to halt cell cycle and result in cell death (Dirmeier et al., 2003, Kaye et al., 1993, Kilger et al., 1998).

In primary B cells, LMP1 induces many of the changes associated with EBV infection or antigen presentation, including cell clumping, increased villous projections, and increased expression of cell-surface adhesion molecules (ICAM-1, LFA-1, and LFA-3), B cell activation genes (CD23, CD30, CD40, CD44, and MHC Class II), interleukin 6 (IL-6), IL-8, and IL-10 (Eliopoulos et al., 1997, Eliopoulos et al., 1999, Nakagomi et al., 1994, Wang et al., 1990b). LMP1 has also been shown to up-regulate several anti-apoptotic proteins, including MCL-1, TNFAIP3, and BCL2 (Henderson et al., 1991, Laherty et al., 1992, Wang et al., 1990a).

The long cytoplasmic C-terminal domain of LMP1 contains two critical transformation sites; C-terminus transformation effector sites (TES) 1 and 2 (Izumi et al., 1999). TES1 and TES2 also correlate closely to the LMP1 carboxy-terminal NFκB activation domains: the

membrane-proximal C-terminal activating region 1 (CTAR1) and the distal C-terminal activating region 2 (CTAR2) (Kieser, 2008, Li and Chang, 2003, Huen et al., 1995, Mitchell and Sugden, 1995). The membrane-proximal CTAR1 domain contains an essential motif that recruits several members of the TNFR associated factor (TRAF) family, including TRAF1, TRAF2, TRAF3, and TRAF5 (Yi et al., 2015, Li and Chang, 2003, Mosialos et al., 1995, Brodeur et al., 1997). The TES2/CTAR2 interacts with TNFR-associated death domain (TRADD) protein and receptor-interacting protein (RIP), which are involved in TNFRI signalling (Wu et al., 2006a, Li and Chang, 2003, Izumi et al., 1999). CTAR1 has been shown to be critical for EBV-mediated transformation, while CTAR2 is required for long-term growth of EBV-infected cells (Wu et al., 2006b, Izumi and Kieff, 1997). It has been shown that the LMP1 activates NFκB via two distinct pathways. A major (canonical) pathway activates NFκB through the CTAR2 domain by engaging TRAF6/TAK1/IKKβ and a minor (non-canonical) pathway activates NFκB through the CTAR1 domain by engaging TRAF3/NIK/IKKα (Wu et al., 2006a). A recent study has found that the LMP1 mediates sumoylation of cellular proteins through its C-terminal activating region 3 (CTAR3), a distinct functional regulatory region of LMP1, to repress of EBV lytic promoters and maintain the viral genome in its latent state (Bentz et al., 2011). In LCLs, inhibition of CTAR3-mediated protein sumoylation abrogates the binding of KAP1 (KRAB-associated protein-1) to EBV OriLyt and immediate early promoters; and increases sensitivity to ganciclovir treatment (Bentz et al., 2015).

Latent membrane protein 2 (LMP2)

In latency III, the LMP2 gene encodes two LMP2 proteins, LMP2A and LMP2B from unique EBNA2 responsive promoters containing multiple RBP-Jk binding sites (Laux et al., 1988,

Laux et al., 1994, Meitinger et al., 1994). Both of these proteins have been shown to be essential for the EBV-induced transformation of primary B cells (Longnecker et al., 1992, Longnecker et al., 1993b, Longnecker et al., 1993a, Speck et al., 1999). Although it may not be critical for transformation, LMP2A has been shown to provide important survival and anti-differentiation signals in B cells (Wasil et al., 2013). LMP2A expression in BL and gastric cancers cell lines was able to protect cells from BCR-cross linking and TGFβ associated apoptosis (Fukuda and Longnecker, 2004, Fukuda and Longnecker, 2005). LMP2 expression during B cell development in mice also decreased the expression of B cell development factors, E2A (a.k.a. TCF3), EBF, and Pax5 (Portis et al., 2003).

The N-terminal domain of LMP2A is able to interact with the Src-family protein tyrosine kinases (SFKs), especially Lyn and Syk (Pang et al., 2009, Johnson et al., 1995, Pleiman et al., 1994). LMP2A binding to Lyn through its immuno-receptor tyrosine-based activation motif (ITAM) leads to phosphorylation of all the LMP2A tyrosine residues and to the recruitment of Syk and the activation of PI3K, Bruton tyrosine kinase (Btk), BLNK, and Protein kinase B (PKB) (Akt) (Merchant et al., 2000, Merchant and Longnecker, 2001, Swart et al., 2000, Pang et al., 2009).

The function of the other LMP2 protein, LMP2B, is still largely unknown. However, Lynch et al., showed that LMP2B may act as an antagonist to LMP2A by co-localizing with it on the plasma membrane (Lynch et al., 2002). Co-expression of both these proteins has been shown to attenuate the LMP2A-mediated block on the activation of the EBV-lytic cycle and restore normal BCR signal transduction (Rechsteiner et al., 2008, Rovedo and Longnecker, 2007). Infection of B cells with an LMP2B deletion mutant EBV showed similar activation to wild-type viruses, but fewer numbers of proliferating B cells (Wasil et al., 2013).

EBV encoded microRNAs (miRNAs)

There are currently 69 unique miRNAs encoded by EBV (Kozomara and Griffiths-Jones, 2014), which show a complex pattern of expression depending on the cell type and overall pattern of EBV gene expression. Furthermore, most of the EBV miRNAs have also been shown to be expressed during the EBV lytic cycle. The BART miRNAs are abundantly expressed in all forms of latency, but particularly high levels in NPC cell lines. By contrast, BHRF1 miRNAs are exclusively expressed in B cells with latent III form of infection as these are derived from an intron generated by splicing of the Cp and Wp initiated transcripts (Cai et al., 2006, Xing and Kieff, 2007), although miR-BHRF1-2 and BHRF1-3 are also expressed in lytic cycle from the lytic BHRF1 promoter.

Although EBV miRNAs are not essential for B cell transformation, as shown in the experiments with a B95-8 virus which has lost the majority of the EBV miRNAs, these EBV miRNAs are highly conserved in various lymphocryptoviruses suggesting a possible role in some stage of the EBV life cycle (Cai et al., 2006, Xing and Kieff, 2007).

EBV miRNAs have been shown to modulate expression of at least two EBV genes, BALF5 and LMP1. BALF5, which encodes the EBV DNA polymerase, is transcribed from the strand opposite to the strand encoding miR-BART2. Overexpression of miR-BART2 genes resulted in the reduction of BALF5 expression and also in the release of infectious EBV particles (Barth et al., 2008). The 3' UTR region of LMP1 and miR-BART2 share partial sequence complementarity (Lo et al., 2007), and over-expression of miR-BART2 suppresses LMP1 protein production, protects against apoptotic stimuli and prevents NFκB inhibition. A potential role for many of these EBV-encoded miRNAs is becoming increasingly apparent as we begin to understand how they modulate expression of broad range of viral and cellular genes.

Furthermore, recent studies have shown that BHRF1 miRNA promote cell cycle progression and prevent apoptosis of primary human B cells, and that BHRF1, PTEN and p27 may be the direct or indirect targets of the BHRF1 miRNA and thus regulate the expression of PTEN, p27 and of a bcl-2 Homolog (Seto et al., 2010, Bernhardt et al., 2016). Increasing number of novel mRNA targets for the anti-apoptotic miR-BARTs have been identified through systematic screening (Kang et al., 2015).

Epstein-Barr virus encoded small RNAs (EBERs)

EBV expresses two non-polyadenylated/non-coding EBER transcripts (EBER1 and EBER2) in all forms of latency (Alfieri et al., 1991). These are the most abundantly expressed EBV transcripts in latently infected cells (Arrand and Rymo, 1982, Howe and Steitz, 1986). There are conflicting reports of the role of EBERs during B cell transformation *in vitro* (Ruf et al., 2000). Initial studies using recombinant virus lacking EBERs showed that the EBERs are not critical for B cell transformation (Swaminathan et al., 1991). However, more recent studies suggest that EBER2, but not EBER1, could be required for B cell transformation (Wu et al., 2007, Yajima et al., 2005).

EBERs share similar primary sequence and secondary structures to the adenovirus VA1 (Arrand et al., 1983, Rosa et al., 1981), and have been shown to inhibit protein kinase R (PKR) with similar efficiencies to VA1 (Bhat and Thimmappaya, 1985, Sharp et al., 1993)As PKR mediates the anti-viral effect of interferons, inhibition of PKR by EBERs could be important for viral persistence (Clemens et al., 1994, Spanggord et al., 2002). However, Ruf et al., suggested EBER inhibition of alpha-interferon-induced apoptosis, and potentially other PKR-mediated events, is unlikely to be mediated through direct inhibition of PKR. Using multiple phosphorylation state-specific antibodies to monitor PKR activation within cells in

response to interferon, it was demonstrated that the EBERs are unable to inhibit the phosphorylation of either cytoplasmic or nuclear PKR. (Ruf et al., 2005). Despite the non-coding nature of these EBERs, they are likely to play an important role in normal EBV infection and EBV-mediated pathogenesis. Recent study on proteome and transcriptome of BJAB cells (an EBV-negative B lymphoma cell line) transfected with EBERs has identified upregulation of several genes and proteins that promote cell proliferation and are associated with EBV-mediated lymphomas, including PIK3AP1 protein, a B-cell specific protein adapter known to activate the PI3K-AKT signalling pathway, and oncogenes such as CCL3, CCR7, IL10 and VEGFA (Pimienta et al., 2015).

1.8 Epstein-Barr virus lytic cycle

Schematic depiction of the location for lytic cycle genes on the B95-8 viral genome is shown in Figure 5.

The cascade of events in the EBV lytic cycle can be subdivided according to the expression of distinct lytic genes, into the immediate early (IE), early (E), and late (L) lytic cycle (Knipe et al., 2007; Tierney et al., 2015). The IE genes are expressed immediately following lytic cycle induction and are characterized by the activation of transactivator proteins which subsequently stimulate expression of early lytic genes including enzymes required for viral DNA replication. Finally, the late lytic genes are expressed, including viral structural proteins which are then assembled into virus particles enclosing newly synthesized viral DNA prior to their release as infectious virions.

Immediate-early lytic cycle genes

Activation of the EBV lytic cycle *in vivo* is thought to be initiated by the differentiation of infected squamous epithelial cells (Tovey et al., 1978, Young et al., 1991) or by the differentiation of infected B cells into plasma cells following antigen stimulation of BCR (Laichalk and Thorley-Lawson, 2005). Furthermore, lytic replication can be stimulated by certain cytokines (e.g. TGFβ) (Fahmi et al., 2000), or by toxic stimuli including chemotherapy and radiation therapy, or by interaction between EBV-infected B cells and CD4 T lymphocytes (Feng et al., 2002). *In vitro,* lytic reactivation can be induced by Ig cross linking, treatment with phorbol ester, sodium butyrate, or a calcium ionophore (Angel et al., 1987, Flemington and Speck, 1990, Takada and Ono, 1989).

During the IE lytic cycle two IE genes, BZLF1and BRLF1, are activated triggering the switch from latency to lytic cycle (Biggin et al., 1987, Countryman and Miller, 1985, Hardwick et al., 1988, Rooney et al., 1989). BZLF1and BRLF1 proteins are induced simultaneously from their respective promoters, Z-promoter (Zp) and R-promoter (Rp), of bicistronic R-Z RNA (Sinclair et al., 1991). In some cell lines the EBV lytic cycle can be initiated by direct expression of BZLF1, a homolog of the AP-1 transcription factor, whereby BZLF1 binds into the lytic origin of replication (OriLyt) and triggers the entire lytic cascade (Flemington and Speck, 1990; Rickinson and Kieff, 2007; Hammerschmidt and Sugden, 1988, Rooney et al., 1989). The BZLF1 and BRLF1 proteins function as transcription factors which can activate their own and each other's promoters amplifying the inducing effect of the initial lytic stimulus (Adamson et al., 2000, Flemington et al., 1991, Liu and Speck, 2003, Ragoczy and Miller, 2001, Sinclair et al., 1991, Speck et al., 1997, Zhang et al., 1996, Zalani et al., 1996). The BZLF1 and BRLF1 proteins transactivate many viral promoters of early, and subsequently late viral genes; while some of these genes are either activated either by BZLF1

or by BRLF1 alone, others are induced by the synergistic effects of both genes (Feederle et al., 2000, Gradoville et al., 1990, Kudoh et al., 2003, Ragoczy and Miller, 1999, Tsurumi et al., 2005).

Recent study described transcriptional reprogramming of the expression of cellular genes by viral bZIP transcription factor, Zta (BZLF1, ZEBRA, EB1) using genome-wide ChIP-seq (chromatin immunoprecipitation coupled to next-generation DNA sequencing). In this study they identified significant change in the expression of 2263 cellular genes during the EBV lytic replication cycle and that Zta binds mostly to sites that are distal to transcription start sites of 278 of the regulated genes they analysed (Ramasubramanyan et al., 2015).

Early lytic cycle genes

In Akata cells, cross linking of surface Ig results in the expression of at least 38 early lytic genes (Lu et al., 2006, Yuan et al., 2006), including the EBV DNA polymerase gene, BALF5 (Holley-Guthrie et al., 2005) and the major DNA binding protein, BALF2, which is essential for EBV DNA replication in Raji cells (Tsurumi et al., 1996, Decaussin et al., 1995, Polack et al., 1984). Other EBV early lytic genes include BMFR1, a transcriptional activator which is a viral DNA polymerase co-factor; and BMLF1 which shuttles un-spliced EBV lytic mRNA and is also required for the assembly of infectious virions (Hiriart et al., 2003).

During the early lytic cycle, EBV also encodes two structural homologs of the anti-apoptotic Bcl2-protein, BHRF1 and BALF1 (Bellows et al., 2002, Henderson et al., 1993, Marshall et al., 1999, Pearson et al., 1983). BHRF1 is expressed in the first few days following B cell infection and has been shown to suppress apoptosis in lymphoid cells in response to various external stimuli, including serum deprivation (Henderson et al., 1993), DNA damaging

agents, chemotherapeutic drugs (McCarthy et al., 1996) and cytokines (Foghsgaard and Jaattela, 1997). Although BHRF1 is not essential for EBV replication (Lee and Yates, 1992, Marchini et al., 1991), it is highly conserved in several EBV strains (Khanim et al., 1997) and is important for optimal transformation efficiency (Altmann and Hammerschmidt, 2005). Although the role of the BALF1 protein is largely unknown, it seems to counteract the anti-apoptotic function of BHRF1 (Bellows et al., 2002). BALF1 has been shown to exhibit anti-apoptotic functions in HeLa cells (Marshall et al., 1999), while in BL cells it enabled growth in low serum (Cabras et al., 2005).

Some of the EBV lytic genes counteract host immune responses. For example, BGLF5 acts as a competitive antagonist for the production of HLA class I and II molecules (Rowe et al., 2007). The BNLF2a protein blocks both the ATP and peptide binding capacity of the transporter associated with antigen processing (TAP), and thus reduces the expression of surface HLA class I molecules (Hislop et al., 2007b).

Late lytic cycle genes

Following EBV DNA replication, lytically active cells begin to express late lytic antigens. 40 EBV late lytic mRNAs have been identified in Akata cells following lytic cycle induction by engagement of the B cell receptor (Lu et al., 2006; Yuan et al., 2006). Some of these late lytic genes are spliced while others are not. 28 of the mRNAs are translated into virion proteins and assembled into viral particles ready for the encapsulation of viral DNA. For example, the BVRF2 gene encodes a viral protease that provides a scaffold for the assembly of virions (Donaghy and Jupp, 1995).

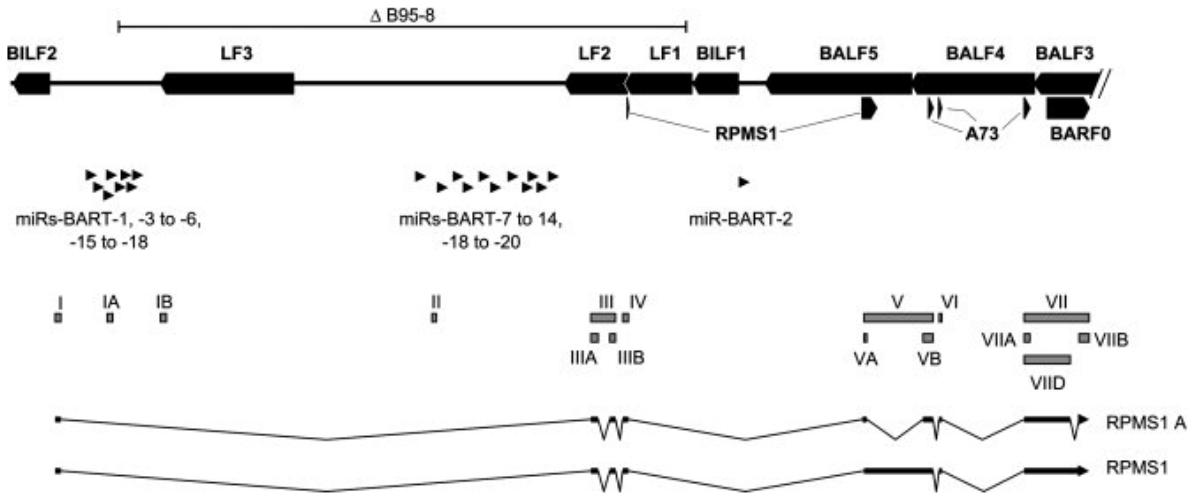

Figure 5. Schematic depiction of the location for lytic cycle genes on the B95-8 viral genome. The dark boxes represent ORFs; small arrows represent microRNAs.

Late lytic cycle genes encode two major classes of protein, glycoproteins and non-glycoproteins (Figure 1.8.1). The glycoproteins include BLLF1, BILF1, BILF2, BDLF2, BDLF3, BMRF2, and BXLF2, some of these are present in the classical membrane antigen (MA) complex (Edson and Thorley-Lawson, 1981). The BLLF1 gene encodes gp350/220, which is the most abundantly expressed viral glycoprotein and is a major target of host neutralizing antibodies (Thorley-Lawson and Poodry, 1982). The gp350/220protein mediates the binding of EBV to its B cell receptor, CD21 (Fingeroth et al., 1984, Nemerow et al., 1989, Torrisi et al., 1989). This interaction induces the activation of the Wp promoter (Sinclair and Farrell, 1995, Sugano et al., 1997) as well as the penetration of B cells by EBV (Tanner et al., 1987).

Conflicting reports on the early/late lytic cycle status of BILF1 have been published (Paulsen et al., 2005, Beisser et al., 2005, Yuan et al., 2006). The BILF1 gene encodes a heavily glycosylated G-protein-coupled receptor (GPCR) which is localized on the plasma membrane similar to KSHV-encoded GPCR, ORF74, indicating that the BILF-GPCR and ORF74 may have similar functions during γ-herpesvirus replication progression (Paulsen et al., 2005). Furthermore, BILF1 sequence and genomic positions are conserved among all of the currently known γ-1-herpesviruses and in a subset of γ-2-herpesvirus species suggesting a possible important role in some stage of EBV lytic replication (Beisser et al., 2005). BILF1 can hetero-dimerize with chemokine receptors CXCR4, CXCR3, CCR9, and CCR10 thereby alter the responsiveness of B lymphocytes to chemokines thereby altering homing and homeostasis of infected B lymphocytes. This process might be essential for EBV dissemination and/or be involved in EBV-induced pathogenesis (Vischer et al., 2008). Several studies have shown that BILF1 can down-regulate surface HLA class I expression thereby obstructing the presentation of viral antigens and cytotoxic T cell recognition of the infected cell (Zuo et al., 2009, Griffin

et al., 2013). BILF1 can also modulate various intra-cellular signalling pathways (Beisser et al., 2005), as the BILF1 gene has been shown to inhibit forskolin-triggered CREB activation via stimulation of endogenous G proteins, thereby affecting disease progression (Paulsen et al., 2005). Furthermore, BILF1-encoded protein increases CRE-mediated transcription in EBV-positive BL cells and LCL transfected with BILF1, through Gi/Go coupling (Beisser et al., 2005).

The heterodimeric complex containing the BXLF2 (gH), BKRF2 (gL) and BZLF2 (gp42) glycoproteins participates in the endocytosis of EBV into B cells using HLA class II as co-receptor. BALF4 is also a highly conserved gene which is thought to be involved in the release of virions from EBV-infected cells (Herrold et al., 1996). The non-glycoproteins include structural proteins, such as BcLF1, BNRF1 and BXRF1 that make up the viral capsid (Kieff, 1996); some of these proteins are part of the immunologically defined viral capsid antigen (VCA) complex. Non-structural late lytic proteins include BCRF1 or viral IL-10 (vIL-10), which shares 90% collinear amino acid sequence with human IL-10 (Hsu et al., 1990, Vieira et al., 1991), and which is an autocrine growth factor capable of paracrine signalling to surrounding B cells. BCRF1 signalling modulates the expression of NF-κB and the JAK-STAT signalling pathways, increases the expression of the Th1 cytokines, and enhances cell survival and proliferation (Stuart and Gruss, 1995, Ding et al., 2000, Ding et al., 2001).

1.9 Epstein-Barr virus associated B cell tumours

In this section, I describe two EBV-associated cancers, Burkitt lymphoma and Hodgkin lymphoma, as these are the major focus of this thesis.

1.9.1 Burkitt lymphoma

General introduction

Endemic Burkitt lymphoma (eBL) was first described by Denis Burkitt in 1958 (Burkitt, 1958). It has a highly unusual pathogenesis which is linked to EBV in almost every case and to the activation of the cellular oncogene, c-MYC. As described in the alternative forms of EBV latent infection section, EBV-positive BL tumours express only the EBNA1 protein along with the non-coding EBERs, and a collection of EBV-encoded microRNAs (Cai et al., 2006, Rowe et al., 1987a). Two more forms of BL are also described which have a lower association with EBV; sporadic BL (sBL) and AIDS-associated BL. All three subtypes of BL display a similar histological appearance (a characteristic ‘starry sky pattern) as shown in Figure 6, in which the malignant population of medium to large round monomorphic B cells are interspersed with macrophages (Kelly and Rickinson, 2007). All three forms of BL also have a cell differentiation phenotype closely resembling that of the cells populating the proliferative compartment of germinal centres (normal centroblasts). The detection of ongoing Ig mutations and centroblast cell surface markers (CD10 and CD77) suggests that BL cells are derived from germinal centre (GC) B cells (Rickinson and Kieff, 2007).

Epidemiology of Burkitt lymphoma

The clinicopathological features of Burkitt lymphoma are detailed in Table 2.

The incidence of BL in Africa ranges from 3-6 new cases per year per 100,000 children aged 0–14 years, accounting for 30-50% of all childhood cancers in equatorial Africa (Martha et al., 2012). The high-incidence eBL form is found exclusively in equatorial Africa and Papua New Guinea where *Plasmodium falciparum* malaria is holoendemic (Burkitt, 1962a, Burkitt, 1962b, Rainey et al., 2007a, Rainey et al., 2007b). The eBL cases almost always have EBV infection, and typically present in young children. Endemic BL affects extranodal sites including the jaw, abdomen and endocrine organs. Suppression of malaria infection correlates with reduced incidence of eBL (Geser et al., 1989) suggesting that co-infection with EBV and *Plasmodium falciparum* malaria is important in the development of eBL (Rochford et al., 2005, Rowe et al., 2014b). Malaria probably increases the pool of EBV-infected BL progenitors by inducing an intense polyclonal B cell activation and loss of T cell control of EBV (de-Thé et al., 1978, Whittle et al., 1984, Moormann et al., 2007, Moormann et al., 2011). There is an association between the incidence of eBL in Africa and high antibody titres to both EBV and malaria (Carpenter et al., 2008, Mutalima et al., 2008), yet only a relatively small number of co-infected children develop eBL suggesting the role of other factors in eBL pathogenesis.

Sporadic BLs also occurs in children but is a less prevalent form and only 15% of sBLs are EBV-positive. AIDS-associated BL often presents as the first AIDS-associated illness in relatively immunocompetent patients, and accounts for 30-40% of all AIDS-associated lymphomas; 30-40% of these AIDS-BL are EBV-positive (Rickinson and Kieff, 2001).

Table 2. Clinicopathological features of Burkitt lymphoma (taken from Hecht and Aster, 2000).

	Endemic BL	**Sporadic BL**	**HIV-associated BL**
Incidence	5-15/100000-year	2-3/106-year	24%-35% of HIV non-Hodgkin lymphoma CD4 usually > 200/μL
Age	More prevalent in children	More prevalent in children and males	More prevalent in male adults
Site	More extranodal Jaw/orbit than abdominal	More extranodal Abdominal (e.g. terminal ileum) than in jaw	Often extranodal (e.g. GI tract, bone marrow, liver) and nodal at presentation
Bone marrow involvement	~ 10% at presentation	~ 30% at presentation	
Meningeal involvement	20%-30% at presentation		
EBV association	> 90% EBV	10%-85% EBV	Variable (30%-80%)
c-MYC translocation	~ 80% t(8;14), ~ 15% t(2;8), ~ 5% t(8;22)	9, 16, 18, 19	

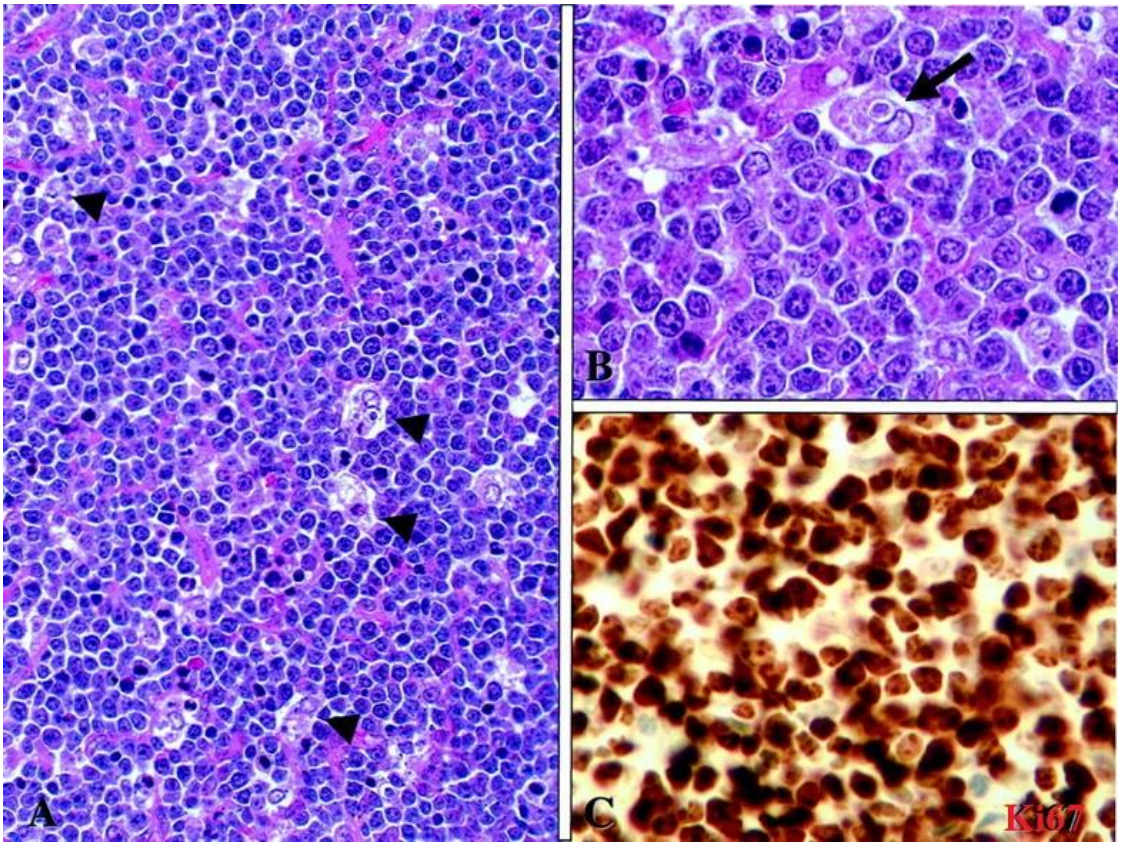

Figure 6. Histopathology of BL. (A) Arrowheads denote some of the phagocytic macrophages (also showing the dark coloured nuclear remnant of phagocytosed tumour cells) responsible for the characteristic "starry sky" appearance that is displayed by all three subtypes of BL, in which the malignant population of medium to large round monomorphic B cells are interspersed with macrophages (×200). (B) An enlarged photograph showing a monomorphous population of intermediate-sized BL cells and a macrophage containing nuclear debris (×400). (C) Strong nuclear reactivity for Ki-67, a marker of proliferating cells, shown in virtually all BL cells (×400) (taken from Hecht and Aster, 2000).

Role of EBV in the pathogenesis of Burkitt lymphoma

The precise role of EBV in BL pathogenesis is still debatable, however BL patients (in particular the endemic form) with high antibody titres against the EBV capsid antigen (VCA) have increased susceptibility to BL (Henle et al., 1969, de-Thé et al., 1978, Geser et al., 1982). BL tumours show a restricted EBV latency programme (latency I) which lacks expression of the majority of latent antigens, including the major EBV growth transforming proteins, EBNA2 and LMP1 (Neri et al., 1991). As described in section 1.6, an alternative pattern of viral gene expression was identified in a minority of Wp-restricted BL cell lines where, instead of Q-promoter (Qp)-driven EBNA1 expression, the W-promoter (Wp) is highly active and drives expression of all EBNAs, except EBNA2 (Kelly et al., 2002b, Kelly et al., 2005). The Wp-restricted BLs also express high levels of the BHRF1 gene encoding a viral homolog of BCL2 gene (vBCL2) that was previously thought to be exclusively a lytic cycle protein (Kelly et al., 2009a, Rowe et al., 2014a). In both forms of latency, non-coding EBERs, and some virally encoded microRNAs are expressed.

Although 95% of humans are infected with EBV, only a very small minority of infected individuals develop BL. Additionally, not all cases of BL are EBV-associated. Furthermore, the growth-transforming genes necessary for establishing and maintaining LCLs are repressed in BL tumours, indicating that the role of EBV exceed further than its ability to growth-transform B lymphocytes. These data suggest that besides EBV infection several genetic changes are required for the development of lymphomas.

Genetic changes in the pathogenesis of Burkitt lymphoma

Deregulated c-MYC expression is a characteristic feature of all three subtypes of BL, and results from a reciprocal translocation of the long arm of chromosome 8 (8q24) in the c-MYC gene with either part of chromosome 14 including the IgH gene (accounting for 80% of all translocations) or chromosome 2 or 22 in the IGL or IGK genes (termed 'variant' translocations) (Magrath, 1990; Kroenlein et al., 2012). Deregulated c-MYC expression in BL cells affects cell proliferation, growth and apoptosis. Studies of the 'common' t(8;14) translocations show differences between the BL subtypes, where breakpoints in eBLs usually occur more than 100 kb upstream of the first exon of the c-MYC gene and in the VDJ region of the IgH locus, whereas in sBLs and AIDS-associated BL, break points commonly occur within the first exon of the c-MYC gene and in the class switch region of the IgH locus, indicating that the translocations may occur at a different stage of B cell differentiation in these tumours (Jacobs et al., 1999, Magrath, 1990, Pelicci et al., 1986, Shiramizu et al., 1991).

Although the mechanisms behind translocations in the c-MYC gene are still largely unknown, these are thought to arise from mistakes during the maturation of antibody responses in GCs (Kuppers, 2003; Kuppers et al., 2005); this could explain the role of malaria and HIV infections as high risk factors for BL development. As discussed in the previous section, malaria may act as a chronic stimulus to the B-cell system and increase the GC activity and also suppress EBV-specific T cell responses resulting in very high levels of EBV DNA in children living in holo-endemic malarial areas (Geser et al., 1989, Moormann et al., 2007, Moormann et al., 2011, Rasti et al., 2005). Similarly, HIV infection has also been shown to act as a chronic stimulus to the B cell system and increase the EBV load in the HIV infected patients long before any obvious T-cell impairment (Piriou et al., 2004). HIV infection can

induce persistent generalized lymphadenopathy in these patients, upon which histological examination has revealed lymphoid hyperplasia with greatly expanded GC regions which will clearly increase the number of cells at risk of a c-MYC translocation (Hecht et al., 2000).

Several genes have been reported to be frequently mutated in BL (Figure 7), including MYC, TP53, FBXO11, ARID1A, RHOA, TCF3 (a.k.a. E2A), and MYD88 (Forbes et al., 2015). Next generation sequencing (NGS) studies have confirmed the presence of mutations in these genes along with novel mutated genes (Abate et al., 2015, Love et al., 2012, Richter et al., 2012). For example, Love et al., identified 70 recurrently mutated genes (Love et al., 2012) including MYC (40%) and ID3 (34%) which were the most frequently mutated genes. Other frequently mutated genes included known suppressor genes ARID1A, SMARCA4 and TP53, as well as the oncogene PIK3R1 and NOTCH1. Gene silencing events, such as nonsense and frameshift mutations constituted a substantial proportion (~30% or more) of mutations in ID3, GNA13, ARID1A, CREBBP, and CCT6B, suggesting that these genes are inactivated. Abate et al., studied 20 BL samples, and also observed recurrent mutations in the genes reported by previous NGS studies; including MYC (50% of cases), DDX3X (35%), ID3 (30%), ARID1A (25%), RHOA (20%), TCF3 (a.k.a. E2A) and TP53 (15%), and CCND3 (5%) (Abate et al., 2015, Forbes et al., 2015, Love et al., 2012, Richter et al., 2012).

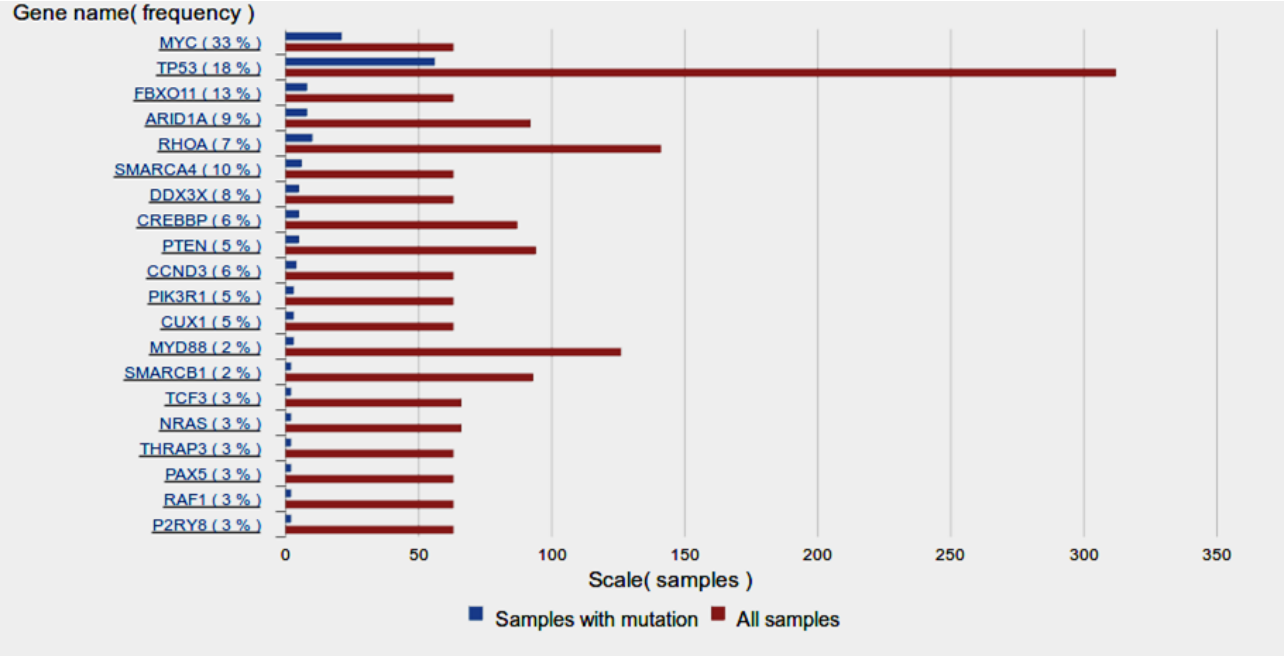

Figure 7. The top 20 most frequently mutated genes in BL. (data from COSMIC data base; cancer.sanger.ac.uk obtained on 18 April 2017).

Role of c-MYC in the pathogenesis of Burkitt lymphoma

Multiple cellular processes that can contribute positively or negatively to cellular transformation have been shown to be influenced by c-MYC through binding of MYC to its heterodimeric partner, MAX. Myc/Max heterodimers bound to DNA recruit protein complexes (TRRAP) associated with histone acetylase activity that modify chromatin and activate transcription; binding of c-MYC/Max heterodimers to the promoter elements of a discrete set of downstream genes either induce or repress their expression (Grandori et al., 1996, Hecht and Aster, 2000). Coller et al., screened genes regulated by c-MYC in cultured human fibroblasts and detected 27 genes that were induced and nine genes that were repressed by c-MYC; some of these c-MYC target genes encode proteins that regulate cell growth, division, death, metabolism, adhesion, and motility, all of which are potentially important in cellular transformation (Coller et al., 2000).

A summary of the biological effects of c-MYC over-expression is shown in Figure 8, which is taken from Hecht and Aster, 2000 (Hecht and Aster, 2000). Some of the most relevant (to my project) biological effects of c-MYC are discussed here.

Cell cycle regulation

The expression of c-MYC is strongly correlated with proliferation as c-MYC is normally expressed in all dividing cells, where it enhances cell cycle progression (Yokoyama et al., 1987, Prochownik et al., 1988, Karn et al., 1989, Goruppi et al., 1994, Mateyak et al., 1999, Schrader et al., 2012). The expression of c-MYC is downregulated in cells undergoing cell cycle arrest and/or terminal differentiation (Marcu et al., 1992).

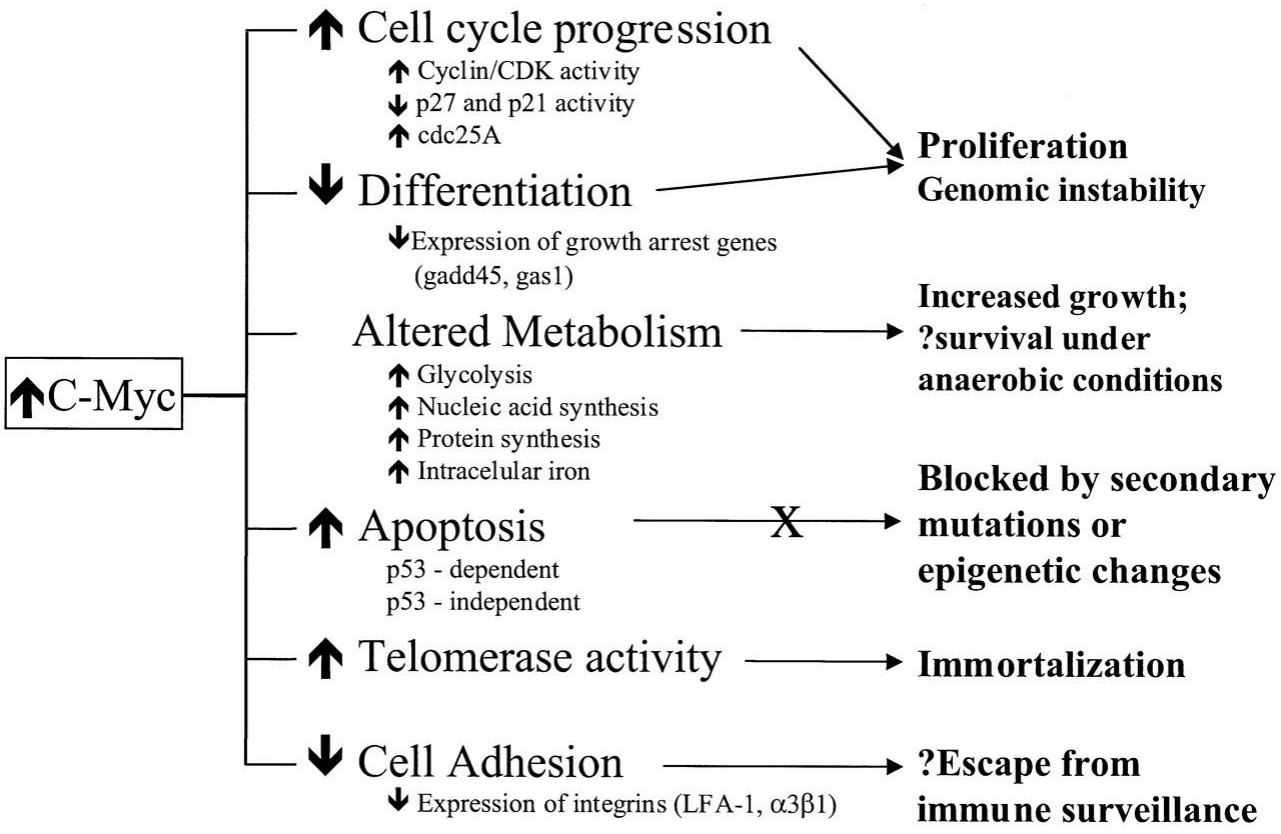

Figure 8. Summary of the biological effects of c-MYC over-expression. Upregulation of c-MYC disrupts many aspects of cell function, which leads to tumorigenesis. (taken from Hecht and Aster, 2000).

In c-MYC knockout fibroblasts, a marked lengthening of both G1 and G2 and a 12-fold reduction in cyclin D1/CDK4 and cyclin D1/CDK6 activity, a pair of protein kinases that promote cell cycle progression, was observed which demonstrates the global nature of the effects of c-MYC on cell cycle progression (Mateyak et al., 1999).

In the human EBV positive B-cell line P493-6, in which c-MYC is expressed under the control of a tetracycline regulated promoter, P493-6 cells arrest in G0/G1 in the presence of serum when the expression of c-MYC is switched off. Re-expression of c-MYC activates the cell cycle, without inducing apoptosis, in these cells (Pajic et al., 2000). Furthermore, c-MYC overexpression can induce genomic instability, which may contribute to subsequent transformation. Transient up-regulation of c-MYC increased the transformation of Rat-1 fibroblasts at least 50-fold, which is correlated with the appearance of chromosomal abnormalities, gene amplification, and hypersensitivity to DNA-damaging agents (Felsher and Bishop, 1999).

Differentiation

As terminal differentiation requires exit from the cell cycle, it is not surprising that the overexpression of c-MYC inhibits differentiation in a number of cellular assays. It can also inhibit cellular differentiation through the repression of genes that direct differentiation, such as C/EBPα (Li et al., 1994, Mink et al., 1996, Kovalchuk et al., 2000).

Metabolism

BL is usually associated with very high levels of serum lactate dehydrogenase A (LDHA), which correlate with the Warburg effect (the tendency for tumour cells to produce increased lactic acid under aerobic conditions) and seem to help tumour cells to survive under hypoxic

conditions. MYC regulates diverse metabolic pathways; of most obvious pertinence to BL, is the observation that c-MYC activates LDH-A transcription (Shim et al., 1997, Vettraino et al., 2013). These effects of c-MYC and LDH-A on metabolism and survival have been demonstrated in Rat fibroblasts, in which stable overexpression of either LDH-A or c-MYC produced increased amounts of lactate under aerobic conditions, whereas antisense LDH-A RNA antagonizes the growth of c-MYC–transformed fibroblasts under hypoxic conditions and reduces the growth of BL cell lines in soft agar (Shim et al., 1997). Furthermore, c-MYC has also been shown to regulate genes involved in nucleotide synthesis (Bello-Fernandez et al., 1993, Pusch et al., 1997), protein synthesis (Rosenwald et al., 1993, Jones et al., 1996), and iron metabolism (Wu et al., 1998).

Apoptosis

Several studies have reported that c-MYC-induces p53-dependent and p53-independent apoptotic pathways. The p53-dependent pathway is triggered by withdrawal of growth factors and is accompanied by the upregulation of the regulatory protein phosphatase cdc25A (Galaktionov et al., 1996, Hermeking et al., 1994). In the alternative p53-dependent pathway, c-MYC over-expression inactivates ARF and thus suppresses p53 expression (Zindy et al., 1998). The mechanism of the p53-independent pathway is unclear, but is thought to be induced upon glucose deprivation and increased levels of LDH-A (Shim et al., 1997). However, overexpression of LDHA alone in certain cell types is sufficient to cause apoptosis upon withdrawal of glucose but not serum.

Clinical features of BL

Patients with BL are clinically staged and classified as low or high risk based on various criteria, such as the number of involved sites, the presence of bulky disease, and LDH concentration. Various staging systems were developed over the last few decades, and the most widely used systems include the Murphy staging system developed for childhood B-NHL (Murphy, 1980), and more recent staging systems developed in Europe by the French-American-British/Lymphomes Malins B group and the Berlin-Frankfurt-Muenster (BFM) group (Waxman et al., 2011). These systems, along with the classic Ann Arbor classification which is still employed in adult patients, are commonly used in children and adolescents/young adults (AYAs).

In the Murphy staging system the clinical stages are defined as follows:

a) Stage 1: presence of a single nodal/extranodal site excluding mediastinum or abdomen
b) Stage 2: presence of a single extranodal site with regional involvement, or presence of two or more nodal sites with or without regional node involvement on the same side of the diaphragm, or presence of a completely resected primary gastrointestinal tumour with or without regional involvement
c) Stage 3: all primary intrathoracic tumours (mediastinum, pleura), by nodal/extranodal tumours on both sides of the diaphragm, by un-resectable extensive intra-abdominal disease, and by paraspinal/epidural tumours regardless of other involved areas
d) Stage 4: CNS or bone marrow involvement.

Treatment of BL

Effective treatment strategies for BL were initially pioneered in paediatric cases and in HIV-negative patients, who still represent the majority of BL cases in the AYA group. Most of the

current treatment regimens for BL include the anti-CD20 antibody rituximab (Table 3), which resulted in improved outcomes, bringing the 3-year survival close to 90% in younger adults (<55–60 years). Favourable outcome with rituximab usage in younger patients was reported by Meinhardt et al. In this study, anti-CD20 rituximab was administered, prior to chemotherapy, to 136 patients (22 aged ≥15 years), which yielded a response rate of 40% in BL/Burkitt leukaemia (Meinhardt et al., 2010). For adults, rituximab is generally added to the Hyper-CVAD regimen which includes cyclophosphamide, vincristine, doxorubicin, and dexamethasone alternating with HD-MTX and cytarabine. Studies have shown that this regimen improved outcome with no additional toxicity compared to Hyper-CVAD alone (Jain et al., 2013).

Dunleavy et al., reported better response and survival rates in a small patient cohort with an infusional MTX-free regimen, dose-adjusted (DA) EPOCH-R (etoposide, prednisone, vincristine, cyclophosphamide, doxorubicin, rituximab), retaining a high therapeutic potential even without dose adjustment (Dunleavy et al., 2013). Recent improvements in the understanding of the pathogenesis of BL led to therapies targeting newly identified pathways that are essential for BL pathogenesis, such as ID3/E2A (a.k.a. TCF3)/cyclinD3 pathway (Schmitz et al., 2012, Richter et al., 2012, Love et al., 2012).

Despite the better cure rates with the current multi-agent chemotherapy and immunotherapy, relapse following treatment occur in small proportion of BL cases and current results in refractory/relapsed BL are extremely poor. Furthermore, the outcome of patients with chemotherapy-resistant disease and sporadic adult BLs are poor. Children with BL suffer long-term toxic effects of chemotherapy in later stages of their life (Jacobson and LaCasce, 2014, Ferry, 2006, Oriol et al., 2008, Ngoma et al., 2012, Magrath, 2012, Appelbaum et al., 1978, Thomas et al., 2006, Molyneux et al., 2012).

Table 3. Comparative features of highly effective, modern treatment regimens for BL (taken from Dozzo et al., 2016).

Regimens	Main characteristics	Number of cycles	Surplus	Main limits
R-CODOX-M/IVAC	First intensive regimen developed for BL, high-dose MTX-based	4 (2 alternating cycles of each regimen)	Reference regimen; including etoposide and ifosfamide	Myelosuppression, mucositis, infections; 20% unable to complete therapy with high-dose MTX
R-Hyper-CVAD	Hyperfractionated cyclophosphamide, high-dose cytarabine/MTX	8, alternating schedule A (cycles 1, 3, 5, 7) and schedule B (cycles 2, 4, 6, 8)	Smaller MTX dose to minimize side effects (1 g/m2)	Myelosuppression, mucositis, infections related to high-dose cytarabine (3 g/m2)
R-NHL 2002	Multidrug regimen, BFM-based	4–6 (A1–B1–C1–A2; B2–C2)	Introduced by BFM and GMALL, used by CALGB, NILG, and PETHEMA	Derived from BFM paediatric protocol; toxicity >55 years (drug reduction)
R-LMB	Dose dense, with high-dose MTX	3–8	Escalating treatment intensity (MTX) according to early response	Myelosuppression, mucositis, infections
DA-EPOCH-R/SC-EPOCH-RR	Dose-adjusted chemotherapy, continuous 24-hour infusion (vincristine, doxorubicin, etoposide), no MTX	4–6	More tolerable by older/HIV+/frail patients; no MTX-related toxicity; outpatient administration feasible (after CR)	Few patients treated, mostly used at NCI/NIH

MTX, methotrexate; BFM, Berlin-Frankfurt-Muenster; GMALL, German Multicenter Study Group for Adult ALL; NCI/NIH, National Cancer Institute/National Institutes of Health; CALGB, Cancer and Leukemia Group B; NILG, Northern Italy Leukemia Group; PETHEMA, Programa Español de Tratamientos en Hematologia.

1.9.2 Hodgkin lymphoma

Hodgkin lymphoma (HL) is a malignant lymphoid tumour that was first described in 1932 by Thomas Hodgkin in his paper entitled "morbid appearances of the absorbent glands and spleen" and which reported the post-mortem findings of patients with enlarged lymph nodes and spleen (T, 1832). Carl Sternberg (1898) and Dorothy Reed (1902) had independently described the "presence of bi-nucleated and multi-nucleated giant cells within an inflammatory milieu" as a characteristic feature of HL. These cells are now known as Reed-Sternberg (RS) cells. Along with RS cells, variable numbers of mononuclear cells known as Hodgkin (H) cells are present in the tumour microenvironment. Neoplastic cells in classical HL (cHL) are collectively referred to as Hodgkin and Reed-Sternberg (HRS) cells (Figure 9).

Epidemiology of HL

HL is one of the most common cancers diagnosed in children and adolescents in the UK and accounts for 4.5% of all cancer types in children. In the UK, 2,106 new cases of paediatric HL were registered in the year 2014 and 355 children died due to HL in2014 (Cancer Research, UK). HL accounts for approximately 0.9% of all cancers diagnosed in the world with a 0.4% mortality rate among all cancers (GLOBOCAN, 2012). Incidence rate for HL vary across different parts of the world, with higher incidence rates are seen in southern Europe and Northern America and lower rates in Asia and Africa. Children treated for HL are prone to develop secondary malignancies in later stages of life. These children also suffer from low quality of life, infertility and hormonal imbalances (Lukes and Butler, 1966, Piccaluga et al., 2011).

The age distribution of HL for both men and women in the UK shows two distinct peaks. The first peak in men occurs in age group 20-34 and the second peak occurs in older men aged 75-

79 years. For women these two peaks occur in the 20-24 and 70-74 year age groups, respectively. The lifetime risk of developing HL is approximately 1 in 370 for men and 1 in 490 for women in the UK (Cancer Research, UK).

Histological classification

Various histopathological classification schemes have been devised since the first comprehensive classification by Jackson and Parker in 1944, in which they described three disease entities: paragranuloma, granuloma, and sarcoma. The recently revised World Health Organization (WHO) scheme (2016) (Swerdlow et al., 2016), was initially adopted from the revised European-American lymphoma (REAL) classification, subdivides HL into two main entities: nodular lymphocyte predominant HL (NLPHL) and classical HL (cHL) and further divides cHL into nodular sclerosis classical HL (NSCHL), diffuse form of lymphocyte-rich (LRCHL), lymphocyte-depleted classical HL (LDCHL), and mixed cellularity classical HL (MCCHL) subtypes based on morphological, phenotypic, genotypic, and clinical evidence (Figure 9). Classical HL accounts for about 95% of all HL of which 20-90% harbour EBV infection (Piccaluga et al., 2011).

NSCHL is more common among young adults in the Western populations. Contrary to the other types of cHL, it is more prevalent in women and less frequently associated with EBV. NSCHL differs from other cHL and NLPHL types in its etiology and histology. The presence of broad bands of collagen is a hallmark feature of NSCHL, the fibrosis being due to expression of pro-fibrotic cytokines. The NSCHL tumour microenvironment consists of abundant lymphocytes and variable numbers of neutrophils and eosinophils (Ikeda et al., 2010).

MCCHL is predominant in men, occurring in both young and old patients, whereas LDCHL occurs mainly in older individuals and is often associated with immune-suppression. Both MCCHL and LDCHL are frequently associated with human immunodeficiency virus (HIV) infection and are more common among lower socio-economic groups. Both subtypes share common clinical, epidemiological, and biological features while differing mainly in the extent of depletion of background lymphocytes. The neoplastic cells in both these subtypes are usually infected with EBV (Eberle et al., 2009).

As the name indicates, MCCHL consists of a mixture of different inflammatory cell types, such as lymphocytes, plasma cells, histiocytes, eosinophils, T-cells, and neutrophils. These infiltrating cells usually form rosettes around neoplastic elements. MCCHL usually shows disruption of lymph node architecture and sometimes “mummified” cells can be seen due to apoptotic changes (Piccaluga et al., 2011). LDCHL appearance varies from patient to patient. However, the most striking feature of LDCHL is the presence of abundant HRS cells among a relatively lower number of lymphocytes (Eberle et al., 2009).

The LRCHL subtype accounts for approximately 5% of all cHL cases. LRCHL was first described as a follicular Hodgkin disease because of the localization of neoplastic cells in the far mantle and marginal zones of the follicles. In Haematoxylin and Eosin staining, LRCHL resembles NLPHL due to the predominant lymphocyte background but differ immunophenotypically, as it expresses CD30 and CD15 (Piccaluga et al., 2011). Neoplastic cells in LRCHL show the typical HRS cell features with small germinal centres at the periphery of nodular structures. LRCHL occurs predominantly in men and the median age is higher than for NLPHL.

NLPHL represents 4-5% of all HL cases and shows distinct morphological, phenotypic, genotypic, and clinical features. It also differs from cHL with its unimodal age distribution

with a peak incidence around 40 years of age. The neoplastic cells in NLPHL are known as ‘Lymphocyte-predominant (LP) cells (Mani and Jaffe, 2009). In NLPHL, LP cells are often surrounded by activated helper-inducer memory T-cells forming a T-cell rosette also observed in cHL. NLPHL usually shows nodular growth affecting single cervical or axillary or inguinal nodes rather than groups of nodes. The neoplastic cells are polylobular with finely dispersed chromatin and small nucleoli, and are often found adjacent to the nuclear membrane. These cells are also known as “popcorn” cells (Mani and Jaffe, 2009).

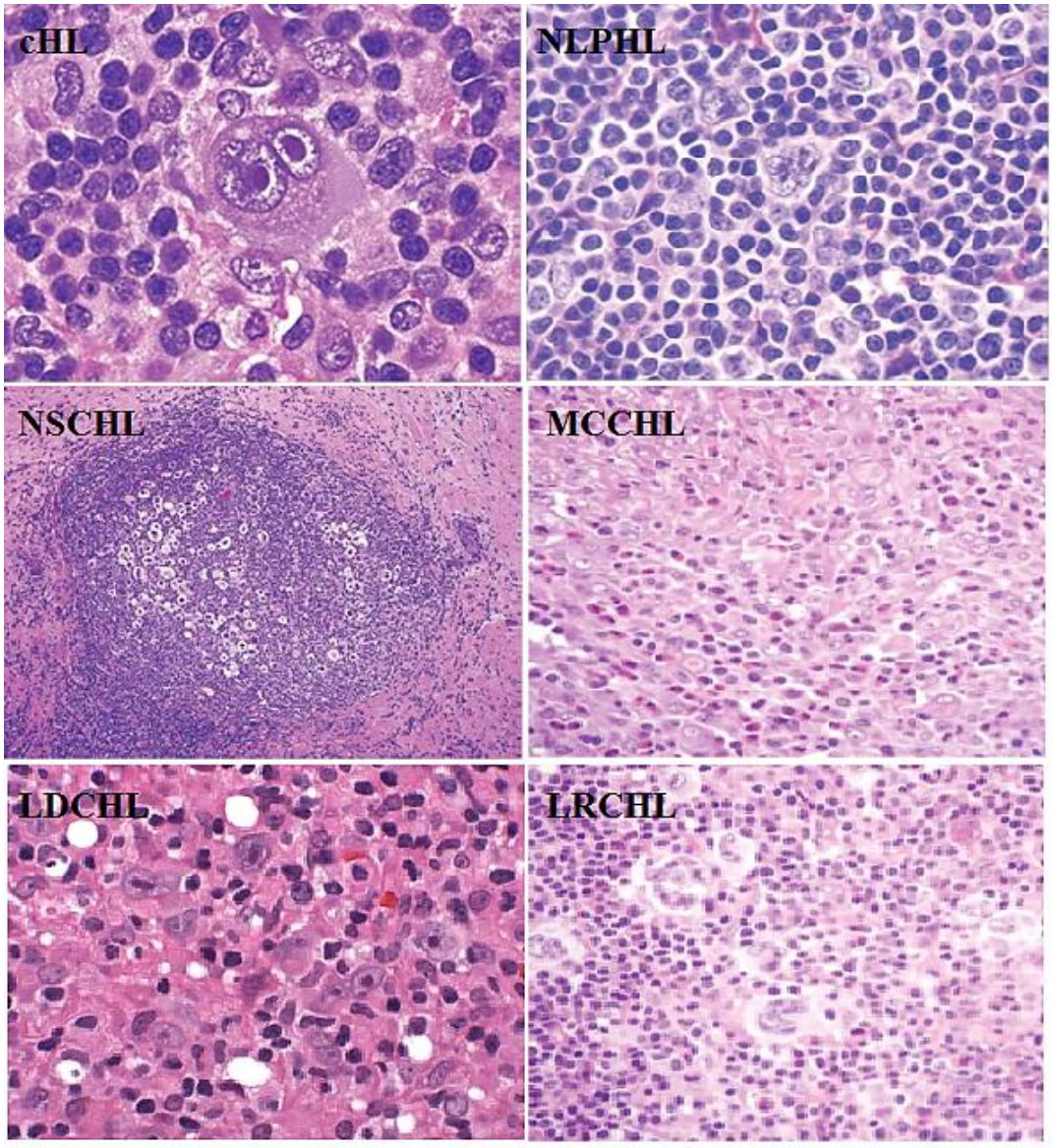

Figure 9. HL subtypes. cHL: Classic Reed-Sternberg bi-nucleated cell with owl-eye inclusion-like nucleoli; each nucleus has a prominent eosinophilic nucleolus with perinucleolar halos (x1000). NLPHL: characteristic multi-lobate nucleus and smaller basophilic nucleoli (resembling popcorn) are present within the lymphocyte rich background (x400). NSCHL: characteristic broad bands of birefringent collagen and with clustering of lacunar variants of HRS cells within the nodules. (x100). MCCHL: with a mixture of different inflammatory cell types in the background (x400). LDCHL: abundant HRS cells and histiocytes along with a few lymphocytes in the background (x200). LRCHL: classic HRS cells present in lymphocyte rich background (x400). (taken from Eberle et al., 2009 (LDCHL); Mani et al., 2009 (cHL, NSCHL, and NLPHL); Piccaluga et al., 2011 (MCCHL and LRCHL).

Origin and clonality of HRS cells

HRS cells in cHL are mainly derived from the clonal proliferation of B lymphocytes that have undergone somatic hyper-mutation in the immunoglobulin V genes. In rare cases (~2%), HRS cells originate from T lymphocytes (Muschen et al., 2000a). On the other hand, LP cells are derived from GC B cells that show evidence of ongoing somatic hypermutation (Brauninger et al., 2006).

Most B-cell lymphomas retain their B-cell phenotype during malignant transformation. However, the HRS cells of cHL have lost their B-cell phenotype due to extensive 'reprogramming' of B- cell gene expression. HRS cells very rarely, express B-cell typical genes, such as immunoglobulins, CD20, spleen tyrosine kinase (SYK), myeloblastosis gene (MYB), and CD19. Although recent studies have shed light on several contributing factors for HL pathogenesis, underlying genetic mechanisms behind the downregulation of numerous B cell transcription factors are still poorly understood (Kuppers, 2009, Kuppers et al., 2012).

Chromosome instability in HL

Hyper-ploidy is a characteristic feature of HL where an abnormal karyotype is usually observed in HRS cells (Figure 10). Structural rearrangements involving chromosome 1 are frequently observed in HRS cells and a recurring numerical abnormality in chromosomes 1, 2, 5, 12, and 21 is seen (Thangavelu and Le Beau, 1989). HRS cells also show multiple sub-clonal aberrations, indicating chromosomal instability (CIN) of the tumour (Weber-Matthiesen et al., 1995). Mutation and dysregulation of kinetochore components are linked to chromosomal missegregation in HL, where deletions in the kinetochore protein BUB1 have been demonstrated in HL patients (Ru et al., 2002). PTTG1 gene, a homolog of yeast securins, has been reported to be over-expressed in HL cases (73%) and in other B cell

lymphomas (Sáez et al., 2002). PTTG1 regulates synchronous sister chromatid separation at the metaphase-to-anaphase transition. PTTG1 over-expression is usually seen in the highly proliferative cells and is linked to aneuploidy (Yu et al., 2003).

Furthermore, centrosome amplification is present in nearly all solid and haematological malignancies (Zyss and Gergely, 2009), and dysregulation of the centrosome cycle is strongly linked to lymphoid malignancy. In Myc-driven B-cell lymphomas, c-MYC has been shown to up-regulate Aurora A and B kinases (mitotic serine/threonine kinase that contributes to the regulation of cell cycle progression) causing centrosomal amplification and CIN (den Hollander et al., 2010). It has been shown that the inability to complete cytokinesis is a cause of CIN in B-cell neoplasms. Long-term time-lapse microscopy and single-cell tracking of HL cell lines confirmed that multinucleated Reed-Sternberg cells arise from incomplete cytokinesis of mitotic daughter cells that remain joined via the mid-body bridge (Rengstl et al., 2013). Haploinsuffiency in KLHDC8B, a kelch domain-containing 8B protein which localizes to the midbody, is linked to the delay in cytokinesis resulting in mitotic errors and aneuploidy, as seen in familial HL (Krem et al., 2012).

Chromosomal abnormalities in HL

Multiple segmental chromosomal aberrations (SCAs), which represent a mechanism of gene amplification by translocating or inserting multiple copies of a certain chromosomal segment into different parts of the genome, were detected in the cHL cell lines HDLM-2, KM-H2, L428 and L1236; involving mostly telomeric sequences like 4p16, 8q24, 9p24 and 12p13, but also intra-chromosomal sequences like 2p13–16 (Joos et al., 2003). Furthermore, chromosomal translocations involving the Ig loci are seen in about 20% of classical HLs (Martin-Subero et al., 2006), some of which involve the known oncogenes such as BCL1, BCL2, BCL3, BCL6, MYC and REL (Martin-Subero et al., 2006, Szymanowska et al., 2008). As Ig loci in HRS cells are usually silenced (Kuppers, 2009), the translocations may result in the deregulated expression of these oncogenes.

Array based comparative genomic hybridisation (aCGH) experiments on cHL cases rich in HRS cells have revealed genomic imbalances in several components of constitutively activated signalling pathways, such as REL/ NF-κB, JAK/STAT and AP1/JUNB pathways. These studies have reported recurrent multiple chromosomal break points in chromosomal arms 1p, 6q, 7q, 11q, 12q, and 14q. Frequent recurrent gains have been observed on chromosomes 2, 5, 9, and 12, and frequent losses are reported on chromosomes 13, 21 and Y. The chromosomal regions that have frequent recurrent gains/losses harbour genes that play vital roles in the pathways that are important for the survival and proliferation of HRS cells, such as NF-κB and JAK/STAT pathways.

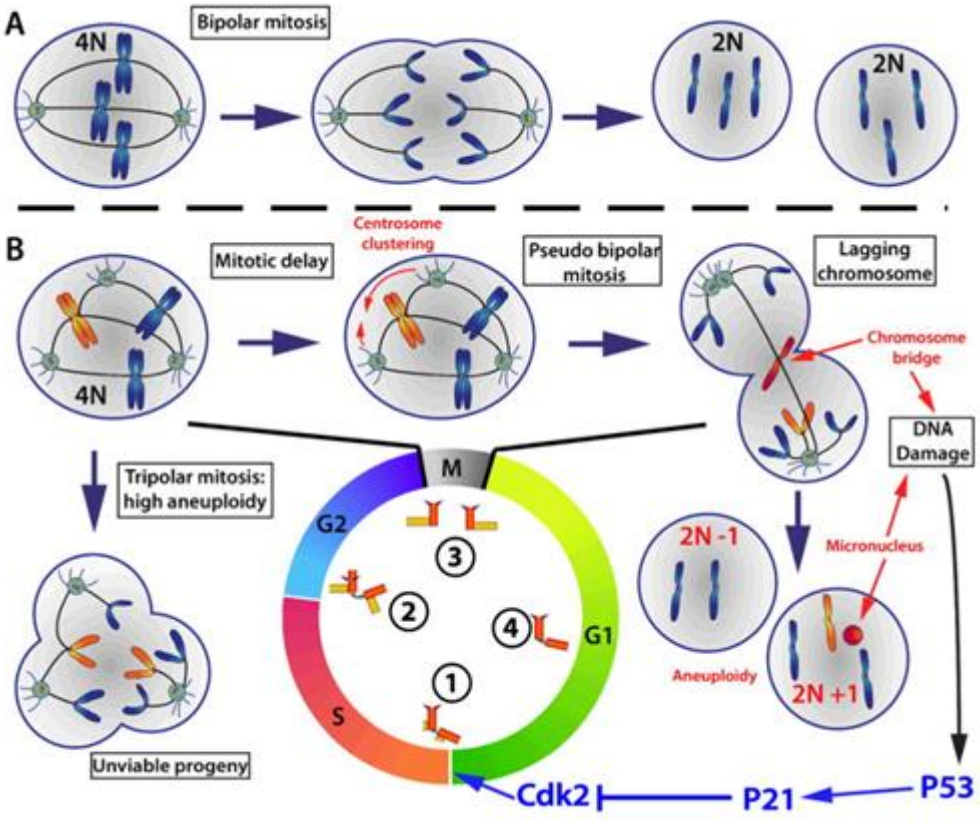

Figure 10. Chromosomes segregation in presence of extra centrosomes. A. Chromosome segregation in normal cells where two centrosomes form a bipolar spindle, accurately segregated, and cellular ploidy is maintained after proliferation. B. Cells entering mitosis with more than two centrosomes can either go through a multipolar mitosis or cluster their centrosomes to form a bipolar spindle. Progeny of multipolar divisions are highly aneuploid and are usually inviable. Centrosome clustering favours the formation of merotelic kinetochore/microtubule attachments resulting in lagging chromosomes. Lagging chromosomes can be included in DNA bridges or micronuclei during cell cleavage to produce daughter cells. These phenomena induce p53- dependent activation of p21, which inhibits Cdk2 and produces cell cycle arrest at the G1/S transition. (taken from Vitre and Cleveland, 2012).

Deregulation of transcription factors

The peculiar phenotype of HRS cells could be a result of the constitutive expression of multiple key transcription factors, such as polycomb group family 2 complexes, Notch1, NF-κB; and JAK/STAT (Table 4). Downregulation of B-cell genes by the polycomb group of proteins and irregular expression of distinct haematopoietic cell markers of haematopoietic stem cells (HSC) and lymphoid progenitor cells might play vital roles in the distinct phenotype of HRS and LP cells (Kuppers, 2009, Kuppers et al., 2012).

NF-κB pathway activation

The NF-κB transcription factor family play a key role in various cellular responses (Figure 11). This family consists of five members, Rel A (p65), Rel B, c-Rel, NF-κB1 (p50 and p100), and NF-κB2 (p52 and p100), which can act as homo-and/or heterodimers. In the unstimulated state, NF-κB is inactive and bound to IκBα, IκBβ, IκBε and p100 and p105 in the cytoplasm (Bonizzi and Karin, 2004). In the classical NF-κB pathway the p50/p65 dimer is kept in the cytoplasm by binding to IκBα, an NF-κB inhibitor. Cellular receptors such as TNF receptor super-family member CD40 activate IKK (IkB kinase complex) kinases. These activated IKK kinases phosphorylate IκBα and thereby induce its proteasomal degradation. This results in the translocation of NF-κB dimers into the nucleus where they subsequently activate multiple target genes.

Constitutive NFκB activation, which is essential for tumour-cell survival, has been observed in HL cell lines and HL primary tumours (Bargou et al., 1997). Furthermore, HRS cells express several receptors such as CD30, CD40 and RANK associating with TRAFs (TNF receptor-associated factors), which activate the classical NFκB pathway. Ligands for these

receptors are frequently expressed by surrounding cells in the tumour microenvironment (Carbone et al., 1995b, Carbone et al., 1995a, Fiumara et al., 2001, Horie et al., 2002).

The activated IKK complex targets IκBα and IκBβ (NF-κB inhibitors) for ubiquitination and proteasomal degradation resulting in the translocation of NF-κB transcription factors into the nucleus, which subsequently activate multiple target genes. TNFAIP3 and CYLD also act as the negative regulators of the NF-κB signalling pathway. In the alternative NF-κB pathway, CD40, BCMA and TACI receptors, upon activation, stimulates the kinase NIK (MAP3K14), which then activates IKKα complex. Here, the NIK activity is negatively regulated by TRAF3. IKKα targets RELB and p100 for proteasomal degradation, which translocate into nucleus as p52/RLB heterodimers. Both pathways are constitutively active in HRS cells. In EBV-positive HL cases, EBV-encoded latent membrane protein 1 (LMP1) and other genetic lesions play an important role in the deregulation of NF-κB signalling pathway.

Several genetic aberrations affecting the NF-κB pathway have been identified in HRS cells. The REL gene has copy number gains or amplification in 40-50% of HL cases. IκBα and IκBε genes, inhibitors of NF-κB, are mutated in 10-20% of HL cases. Recent studies have also demonstrated a role for tumour suppressor genes, such as TNFAIP3 and CYLD, in the negative regulation of the NF-κB pathway. The TNFAIP3 gene encodes the A20 protein that has ubiquitinase and deubiquitinase functions and is a negative regulator of NF-κB signalling acting upstream of the IKK kinases (Schmitz et al., 2009). Recent studies have revealed mutations and deletions in TNFAIP3 in 30-40% of HL (Kuppers, 2009, Kuppers et al., 2012, Bräuninger et al., 2005).

Table 4. Signalling pathways and transcription factors constitutively active in HRS cells (taken fromBrauninger et al., 2006).

Pathway	**Proposed effect of activation**	**Mechanisms of activation**
JAK-STAT	Supports proliferation; promotes survival	Autocrine stimulation via IL-13/IL-13R
AP-1	Supports proliferation	Autoregulation of c-Jun; NFκB activation of JunB
Notch1	Induces proliferation; inhibits apoptosis; downregulates B-cell phenotype?	Expression of ligand Jagged1 in HRS cells and surrounding cells
MAPK/ERK	Supports proliferation; inhibits apoptosis	CD30-, CD40-, RANK-signalling; RTK signalling?
PI3K/AKT	Promotes survival	CD40 signalling?; LMP1 and LMP2a in EBV-positive cases; RTK signalling?
NFκB	Inhibits apoptosis; supports proliferation; recruits T cells to lymphoma; promotes dissemination	Mutations in IκBα; mutations in IκBε genomic c-Rel amplifications; CD40 stimulation by CD40 ligand expressing T cells; CD30 stimulation; ligand independent or via CD30 ligand expressing T cells; LMP1 mimicry of active CD40 in EBV-positive cases
Receptor tyrosine kinases*	Supports proliferation; others?	Autocrine or paracrine stimulation through ligands; no indication for mutations

*DDR2 (discoidin receptor family, member 2); EPHB1 (ephrin receptor EphB1); MET; PDGFRA (platelet-derived growth factor receptor A); RON (macrophage-stimulating protein receptor), TRKA/B (tyrosine kinase receptor A/B).

JAK/STAT pathway in HL

Chromosomal gains (in ~ 20% of HL cases) and translocations JAK2 gene, an activator of STAT signalling, are seen in HRS cells (Joos et al., 2000, Van Roosbroeck et al., 2011). In HRS cells, several members of the STAT family of transcription factors, including STAT3, STAT5 and STAT6, are frequently activated (Figure 1.9.6) (Kube, 2001, Skinnider, 2002). STATs can be activated by cytokine receptors such as IL-13 and IL-21 via the JAK kinases, by RTKs and by 7 transmembrane receptors (Ilhe, 2001). These activated STATs undergo phosphorylation to form dimers, which translocate to the nucleus and activate several target genes (Levy and Darnell, 2002).

The constitutive activation of STAT6 in HRS cells, by simultaneous expression of IL-13 and its receptor was shown to increase the proliferation of HRS cell lines (Kapp et al., 1999, Skinnider, 2001). STAT3 downregulation in an HRS cell line induces apoptosis, indicating that STAT3 activity may be essential for HRS cell survival (Holtick et al., 2005). Genomic gains or translocations of JAK2 and frequent silencing mutations in SOCS1 (Weniger et al., 2006b) have been shown to be able to mediate activation of the JAK/STAT pathway in HRS cells.

In HRS cells, NFκB and the AP-1 transcription factors act synergistically and mediate cell cycle promoting and cell-death inhibiting pathways. AP-1 dimer consisting of c-Jun and JunB is strongly expressed in HRS cells (Mathas, 2002), and NFκB seems to be essential for the strong JunB expression in HL cell lines (Watanabe et al., 2003). Furthermore, AP-1 activates CD30 transcription in HL cell lines (Watanabe et al., 2003), creating a positive feedback loop to enhance JunB expression.

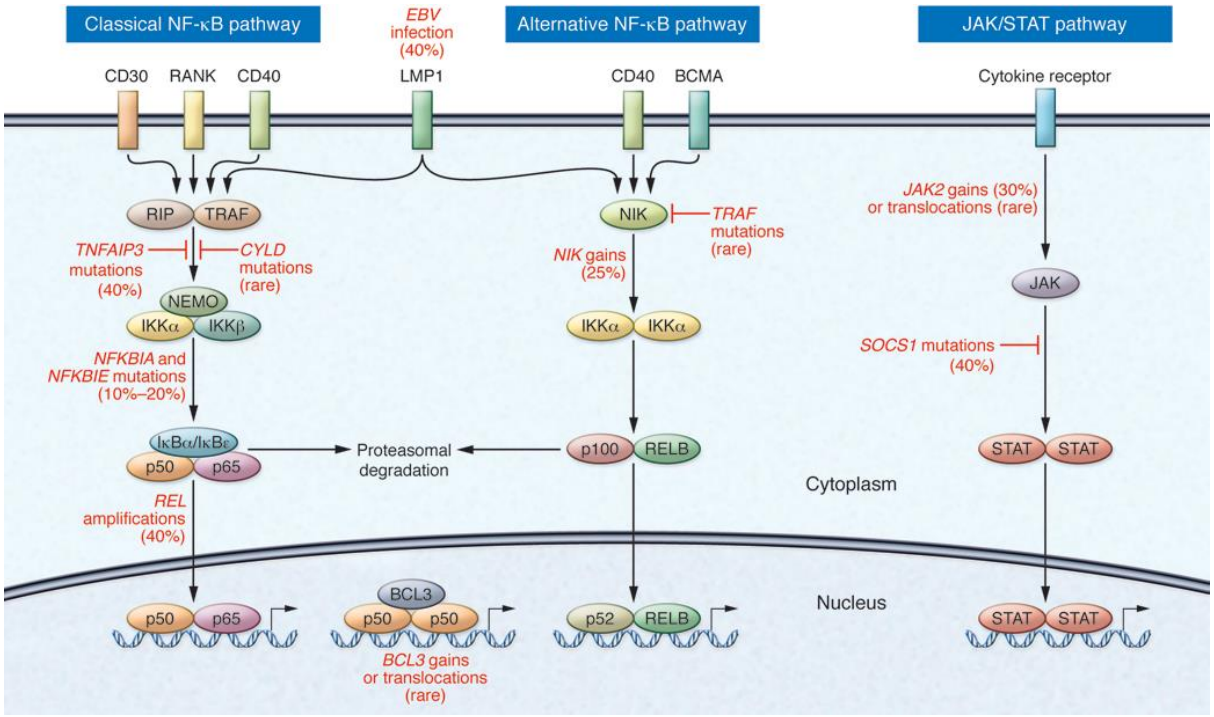

Figure 11. The role of NF-κB and JAK/STAT pathways in pathogenesis of HL. The extent of viral infections and genetic lesions affecting NF-κB and JAK/STAT pathways are shown in percentages (taken fromKuppers et al., 2012).

Notch1 promotes the degradation of E2A (a.k.a. TCF3), and blocks DNA binding of early B-cell factor (EBF). This results in the suppression of the B cell differentiation programme, while inducing the transcription of T-cell genes. In addition, activation of Notch1 by Jagged1 inhibits PAX5 at transcriptional and post-translational stages. Overexpression of activated B-cell factor 1 (ABF1) and inhibitor of differentiation and DNA binding 2 (ID2) also results in the degradation of E2A (a.k.a. TCF3) (Hartmann et al., 2008).

Epigenetic modification is another important factor in the transcriptional downregulation of B-cell lineage-specific genes. Simultaneous silencing of master transcriptional regulator(s), such as Oct2, BOB.1/OBF.1, and PU.1 suggests the importance of epigenetic modification in the reprogramming of HRS cells (Ushmorov et al., 2006).

Anti-apoptotic mechanisms in HRS cells

Evading apoptosis is a key event in the transformation process of pre-apoptotic GC B-cells, leading to the formation of HRS cells. Deregulation of several signalling pathways, including NFκB, contribute to the survival of HRS cells.

Two major apoptotic signalling pathways are reported in B cells (Mizuno et al., 2003), the extrinsic pathway of apoptosis involving TNF receptor super-family members carrying a death domain (TRADDs), and the intrinsic pathway involving the apoptosome. In the extrinsic pathway of apoptosis, TNF receptor super-family members carrying a death domain (TRADDs), upon activation, form a multi-molecular death-inducing signalling complex (DISC), in which caspase-8 is activated by auto-proteolytic cleavage, leading to the sequential activation of caspases-3 and -7 followed by destruction of substrates essential for cell survival

(Mizuno et al., 2003). In the intrinsic pathway, pro-apoptotic proteins form a multi-protein complex, the apoptosome, where caspase-9 is activated.

Although CD95 (Fas), CD95-ligand, and TRADDs (such as TNF-R1, TRAIL-R1 and TRAIL-R2) are expressed in HRS cells (Maggio et al., 2003, Xerri et al., 1995), resistance to CD95 mediated death was observed in these cells *in vitro*, indicating CD95 death signalling pathway might be impaired in HRS cells (Metkar et al., 1999, Verbeke et al., 2001, Re et al., 2000)Mutations in CD95, resulting in the impairment of CD95-mediated cell death, was also observed in some rare cases of HL (Maggio et al., 2003, Muschen et al., 2000b). In HRS cells, high levels of the cellular FLICE-inhibitory protein (c-FLIP), a proximal negative regulator of the CD95-induced death program that interferes with formation of the DISC, was observed (Maggio et al., 2003, Dutton et al., 2004) Inhibition of c-FLIP expression in HL cell lines restored CD95-mediated apoptosis, suggesting that c-FLIP plays a central role in preventing CD95-mediated apoptosis of HRS cells (Dutton et al., 2004, Mathas et al., 2004).

HRS cells have been reported to evade the intrinsic pathway of apoptosis by inhibiting the proteolytic activity of caspases-3, 7, and 9 through the over-expression of several inhibitors of apoptosis (IAP), including XIAP (X-linked inhibitor of apoptosis) (Kashkar et al., 2003, Kreuz et al., 2001).

Interaction of HRS cells with the inflammatory milieu

Figure 12 illustrates the tumour microenvironment and various cellular interactions in HL.

Historically, it was believed that the tumour-induced immune responses simply inhibit tumour growth and progression. However, recent studies have demonstrated that the tumour-induced inflammatory response has a paradoxical effect which can also contribute to tumour growth and survival by providing growth factors to sustain proliferative signalling and anti-apoptotic

signals. Reactive cells in the tumour microenvironment may also facilitate tumour angiogenesis, and invasion and metastasis by secreting extracellular matrix-modifying enzymes (Hanahan and Weinberg, 2011).

About 99% of the tumour microenvironment in HL is composed of lymphocytes, eosinophils, fibroblasts, plasma cells, mast cells, and macrophages. Recent studies have reported that the tumour cells actively recruit these infiltrating cells by secreting cytokines and chemokines (Kuppers, 2009). The recruitment of these cells and their interaction with HRS cells seems to be very important for the survival and proliferation of HRS cells, as HRS cells are usually not found in the peripheral blood and HRS cells do not grow well *in vitro* or *in vivo* (Kapp et al., 1993, Vockerodt et al., 1998). The majority of these infiltrating cells are $CD4^+$ T-cells that often demonstrate a T-helper type 2 (Th2)-phenotype. These Th2 cells are recruited via the secretion of chemokines such as TARC, CCL5 and CCL22, by the HRS cells (Kuppers, 2009). These Th2 cells express CD40 ligand which is essential in the expression of CD40 receptors in the NF-κB pathway (Carbone et al., 1995a, Nozawa et al., 1998). Furthermore, the infiltrating Treg cells suppress the tumour-antigen specific NK cells and $CD8^+$ cytotoxic T lymphocytes (CTLs) and contribute to the immune escape of HRS cells (Marshall, 2004). Cytotoxic T cells are also inhibited through expression of the PDL-1, PDL-2 and CD95 ligands and secretion of IL-10, TGF-β, and galectin1 by the HRS cells (Yamamoto et al., 2008, Juszczynski et al., 2007, Gandhi et al., 2007, Newcom and Gu, 1995).

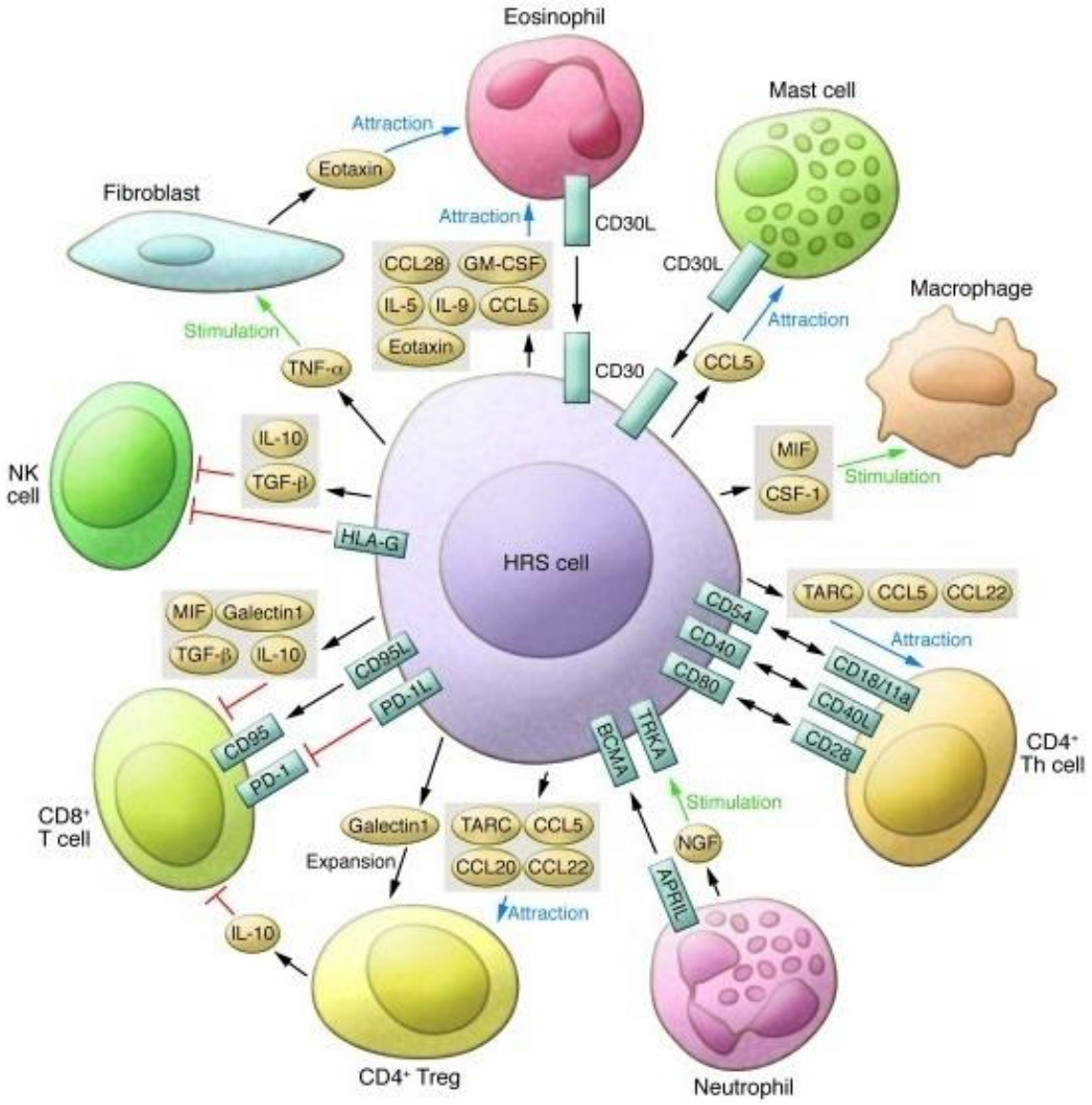

Figure 12. Tumour microenvironment and cellular interactions in HL. Infiltration and activation of various reactive elements into the tumour microenvironment is facilitated by the HRS cells via secretion of chemokines and cytokines. The majority of the infiltrate comprises of Th2 and Treg cells, which are attracted to HRS cells by TARC, CCL5, CCL20, CCL22, and Galectin secreted by the HRS cells. These Th2 and Treg cells interact with HRS cells and assist in the activation of some key pathways and protect HRS cells from immune shock (taken fromKuppers et al., 2012).

Recurrent mutations in Hodgkin lymphoma

Liu et al., 2014, described the mutational spectrum of 7 HL cell lines; SUPHD1, L540, KMH2, L428, L1236, DEV, and L591 using exome sequencing technology. This study identified 463 recurrently mutated genes of which 373 were unique to cHLs. Of the 463 genes that were mutated in at least two HL cell lines, 346 had a high-impact mutation or a SIFT score of 0.05 (Figure 13). The authors identified several genes, including PTPN1, TNFAIP3, NFKBIE, CYLD and NFKBIA, which were previously reported to be mutated in primary cHL. Furthermore, this study identified recurrent mutations affecting the ATG start codon in B2M, and mutations in CIITA. Consistent with the B2M mutations affecting the ATG start codon, the authors observed no or very low membrane B2M and HLA class I expression in L428 and DEV cell lines. This finding is of importance because several studies have shown that in some cases primary HRS cells lack of expression of MHC-I and B2M (Oudejans et al., 1996, Murray et al., 1998, Diepstra et al., 2005, Lee et al., 1998). Furthermore, lack of MHC-I expression in HRS cells was also reported as an independent adverse prognostic factor in cHL (Diepstra et al., 2005).

Reichel et al., described mutation spectrum of purified HRS cells from ten primary HL tumours using exome sequencing (Figures 14-15) (Reichel et al., 2015). This study identified recurrent gains of a region in chromosome (chr) 2 containing REL (5/10), BCL11A, XPO1, and variably MYCN (4/10); focal amplifications involving only NSD1 (chr 5, 4/10); gains involving CD274 (chr 9, 4/10) and variably JAK2 and MLLT3 (3/10); gains involving UBE2A (chr X, 3/10); gains involving CDK4 (chr 12, 2/10); losses of gene segments involving TNFAIP3 (chr 6, 5/10) and variably MLL, MLLT4, PRDM1 (3/10), and MLL; losses of ATM and BIRC3 (chr 11, 5/10); and losses of RB1 (chr 13, 4/10), and BRCA2 (chr 13, 3/10).

This study identified 99 recurrently mutated genes with protein-altering mutations, including B2M, TNFAIP3, HIST1H1E, EEF1A1, ITPKB, SOCS1, BCL7A, CSF2RB and RANBP2. In line with Kato et al., study which reported frequent mutual exclusivity of EBV-infection and TNFAIP3 mutation (NF-κB-activating events) (Kato et al., 2009), this study did not observe any TNFAIP3 mutation in the only EBV-positive case (the only case with NS morphology that lacked any TNFAIP3 mutation), indicating that EBV provides alternative molecular mechanisms for avoiding tumour immunity. Additionally, they also identified some novel genes that have not been previously linked to lymphoid malignancies but have been seen in solid cancers and myeloid stem cell disorders, including CSF2RB, NEK1, HECW2, SENP7, TBC1D15, TICRR, and ZPF36L1. Mutated genes identified in this study had enrichment in the genes responsible for interactions with the immune system, preservation of genomic stability, and transcriptional regulation.

Furthermore, they found statistically significant association with older age among the B2M-positive cases (median age of 47 vs 30 years; $\boldsymbol{p} \leq 0.0001$) and stage III/IV disease ($\boldsymbol{p} = 0.001$), and with male predominance ($\boldsymbol{p} = 0.037$) and bulk disease ($\boldsymbol{p} = 0.029$) and that the cases that lacked B2M expression belonged to a better clinical outcome categor0y (10-year progression-free survival of 74% vs 49%; $\boldsymbol{p} = 0.026$; and overall survival of 87% vs 66%; $\boldsymbol{p} = 0.013$).

The Catalogue of Somatic Mutations in Cancer (COSMIC)

COSMIC is a comprehensive resource for exploring the impact of somatic mutations in human cancer. The COSMIC database provides list of mutations based on tumour type and the frequency of these mutations in each cancer. SOCS1, TNFAIP3, NRAS, EZH2, PIK3R1, NFKBIE, TP53, CDKN2A, CYLD, and HRAS genes were reported to be mutated frequently in HL (Forbes et al., 2015) (Table 5).

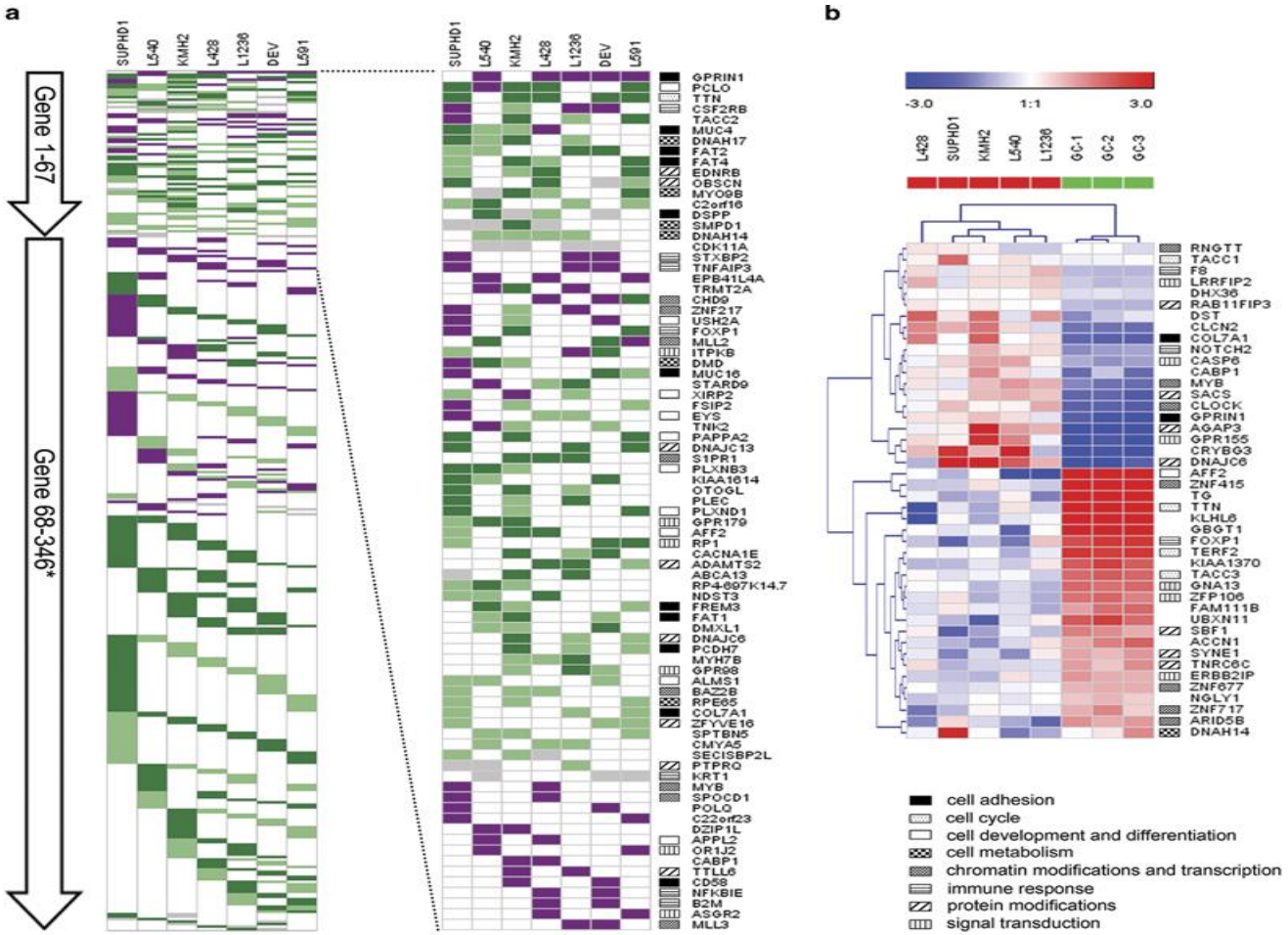

Figure 13. Recurrently mutated genes and differentially expressed genes in HL cell lines**.** (a) In all, 463 genes were mutated in at least two out of seven HL cell lines. 81 genes (shown in the middle panel) included 67 genes that were mutated in at least three HL cell lines and 14 genes showing high-impact (stop gain, frame shift and splice site) mutations in two HL cell lines (b) Heat map of the gene expression levels of the 44 genes that were mutated specifically in cHL and differentially expressed in cHL vs GC B cells. (taken from Liu et al., 2014)

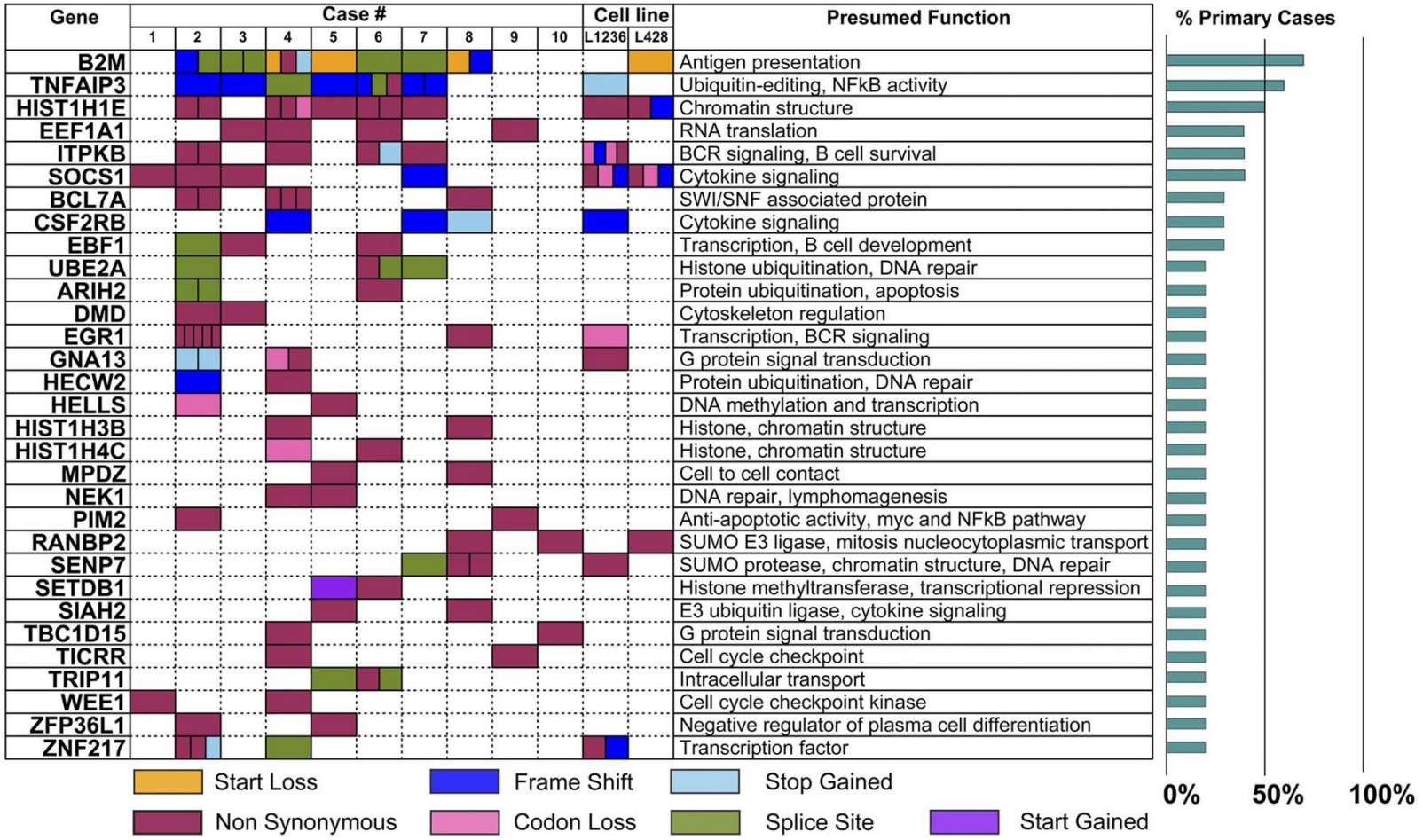

Figure 14. Recurrently mutated genes with potential pathogenic functions in Reichel et al., 2015 (taken from Reichel et al., 2015).

Figure 15. Recurrently mutated genes in Reichel et al., 2015 (taken fromReichel et al., 2015).

Table 5. Recurrent mutations in HL (obtained from COSMIC data base; 20 Jan 2017).

Gene	Number of samples tested	Number of samples mutated	Mutation frequency
SOCS1	33	11	33.3
TNFAIP3	83	24	28.9
NRAS	49	6	12.2
EZH2	10	1	10
PIK3R1	10	1	10
NFKBIE	21	2	9.5
TP53	35	3	8.6
CDKN2A	14	1	7.1
CYLD	27	1	3.7
HRAS	48	1	2.1

.

Role of EBV in the pathogenesis of Hodgkin lymphoma

The frequency of EBV association in HL patients depends on certain factors such as geographical location; the prevalence rate varying across the world, with lower rates in the West (approximately 40%) and higher rates in the developing world e.g. Central Africa (90% of paediatric cases), age; occurring more frequently in children and elderly patients, sex; more prevalent in men, and histological subtype; most often associated with MCCHL and LDCHL (Eberle et al., 2009).

Early reports documented an increased risk of cHL in people with a prior history of IM (Connelly and Christine, 1974, Rosdahl et al., 1974). Subsequently, those with a prior history of either self-reported or laboratory-confirmed IM were shown to have an increased risk of developing EBV-positive cHL; but not EBV-negative cHL (Hjalgrim et al., 2007, Hjalgrim, 2012). In HRS cells the EBV genome is monoclonal and is maintained throughout the course of disease suggesting it is involved in both the initiation and maintenance of disease. EBV infected HRS cells express EBNA1, LMP1 and LMP2A/B, together with the non-coding EBERs and miRNAs (Deacon, 1993, Young et al., 1991, Niedobitek et al., 1997, Murray et al., 1992a, Grasser et al., 1994).

HRS cells lack pro-survival signals mediated by a functional BCR, instead, EBV latent gene products contribute to the survival of HRS progenitors through multiple routes, most of which converge on a limited number of cell signalling pathways. CD40 and BCR are required for the survival and proliferation of germinal centre B cells. LMP1 and LMP2 have been shown to mimic the physiological functions of CD40 and BCR, respectively, thereby provide essential signals for survival. LMP1, functions as a constitutively active homologue of the cellular CD40 receptor (Lam and Sugden, 2003), and LMP1 expression in HRS cells has been shown to activate multiple cell signalling pathways, including the NF-κB, JAK/STAT and

phospatidylinositol-3-kinase (PI3K)/AKT pathways (Bargou et al., 1997, Kube, 2001, Dutton et al., 2005), which could potentially rescue HRS progenitors from apoptosis (Gires et al., 1999, Huen et al., 1995, Kieser et al., 1997).

LMP1 also induces the overexpression of FLICE-inhibitory protein (c-FLIP) thus providing further protection from death induced by the extrinsic pathway to apoptosis (Dutton et al., 2004). In HRS cells, LMP1 was also shown to up-regulate pro-survival genes, such as BCL2 and BFL-1, and downregulate B-cell transcription factors and BCR signalling components that are required to maintain B-cell identity; thereby protecting B cells from apoptosis (Vockerodt et al., 2008, Henderson et al., 1991, D'Souza et al., 2004). These transcriptional changes may be mediated through the induction of ID2 (Vockerodt et al., 2008), or via epigenetic mechanisms involving polycomb proteins, DNA methyltransferases and protein arginine methyltransferases (Anderton et al., 2011, Leonard et al., 2011, Dutton et al., 2007, Leonard et al., 2012).

The unique N-terminus of LMP2A includes an ITAM motif, which resembles the signalling domain of the BCR (Longnecker and Kieff, 1990), and thus LMP2A can mimic a functional BCR and can allow B cells to develop in the absence of normal BCR signalling, as shown in transgenic mice experiments (Caldwell et al., 1998, Merchant et al., 2000). Importantly, like a functional BCR, LMP2A can also induce lytic cycle entry, and can do so even in the absence of a functional BCR. However, it cannot do so when certain components of the BCR signalling machinery are missing which is the case in HRS cells (Vockerodt et al., 2013). Thus, the loss of BCR, as well as of BCR signalling components could prevent both BCR- and LMP2A-induced virus replications in HRS cell progenitors. LMP2A was also shown to induce several signalling pathways important for B-cell survival, including the activation of RAS/PI3K/AKT (Dutton et al., 2005, Swart et al., 2000, Fukuda and Longnecker, 2004). In

latently infected HRS cells, LMP2A can also induce the upregulation of genes involved in proliferation, such as MKI67 and PCNA, and anti-apoptotic genes, such as BCLXL.

Clinical presentation and staging of HL

HL patients usually present with painless lymphadenopathy or non-specific features such as weight loss, fluctuating fever, night sweats or fatigue. Accurate staging of the disease in patients using pathological, clinical, and radiologic data is the most important aspect in deciding prognosis. The Ann Arbor system (Carbone et al., 1971) and Costwold modification (Lister et al., 1989) are widely used for staging HL patients.

Currently radiological examination with computerised tomography (CT) and positron emission tomography (PET) is widely used for the initial staging assessment. PET is often used to assess response to treatment and to evaluate residual tumour tissue upon completion of therapy (Rathore and Kadin, 2010).

In addition to the above prognostic factors, the presence of bulky disease (>10 cm in diameter), multiple sites of disease, elevated erythrocyte sedimentation rate (ESR), extranodal disease, advanced age, are important factors in stratifying patients (Rathore and Kadin, 2010). They include:

1. Early favourable (Stages I-II, no B symptoms)
2. Early unfavourable (Stages I-II with at least one risk factor), and
3. Advanced (Stages III-IV or any stage with bulky disease or intra-abdominal disease)

Long-term survival rates and morbidity due to treatment-related toxicity are significant among HL patients. So, there is a need for better prognostic models which would help to identify patients with poor prognosis. The international Prognostic Score (IPS) model proposed by

Hasenclever and Diehl is widely used for assessing HL prognosis (Hasenclever and Diehl, 1998).

Current treatment strategies for HL

For early favourable HL, patients are generally treated with two to four chemotherapy courses and involved-node radiotherapy. A combined modality therapy with ABVD (Adriamycin, Bleomycin, Vinblastine, and Dacarbazine) followed by involved field radiotherapy (IFRT) or extended field radio therapy, and with regular PET assessment is a standard treatment approach. The above regimen with additional cycles or with BEACOPP (Bleomycin, Etoposide, Adriamycin, Cyclophosphamide, oncovin, Procarbazine, Prednisone) treatment regime is often used in early unfavourable HL. For advanced HL, the above regimen is applied, albeit with more cycles (Rathore and Kadin, 2010).

These treatments allow for high cure rates but relapse can occur in around 10% of early HL cases and 20% of advanced HL cases. The current treatment regimen for patients with relapsed/refractory HL includes salvage chemotherapy such as ICE (ifosfamide, carboplatin and etoposide) or DHAP (dexamethasone, high-dose Ara-C and cisplatin) followed by high-dose chemotherapy and autologous stem cell transplantation (ASCT). However, this treatment is not tolerable in some patients, and in others is not sufficient for cure.

Identifying the risk-groups who will relapse after initial treatment and finding the right treatment regime is necessary as these patients (especially young patients) are prone to develop secondary malignancies, hormonal imbalances, and sterility (Delwail, et al. 2002; Ibrahim, et al. 2012). Recent advances in the understanding of pathogenic mechanisms have led to the approval of several novel agents (Figure 16), such as brentuximab vedotin, the antibody–drug conjugate which targets CD30 (Adcetris®, Seattle Genetics, Inc., USA); and nivolumab, an immune checkpoint inhibitor (Opdivo®, Bristol-Myers Squibb, NY, USA).

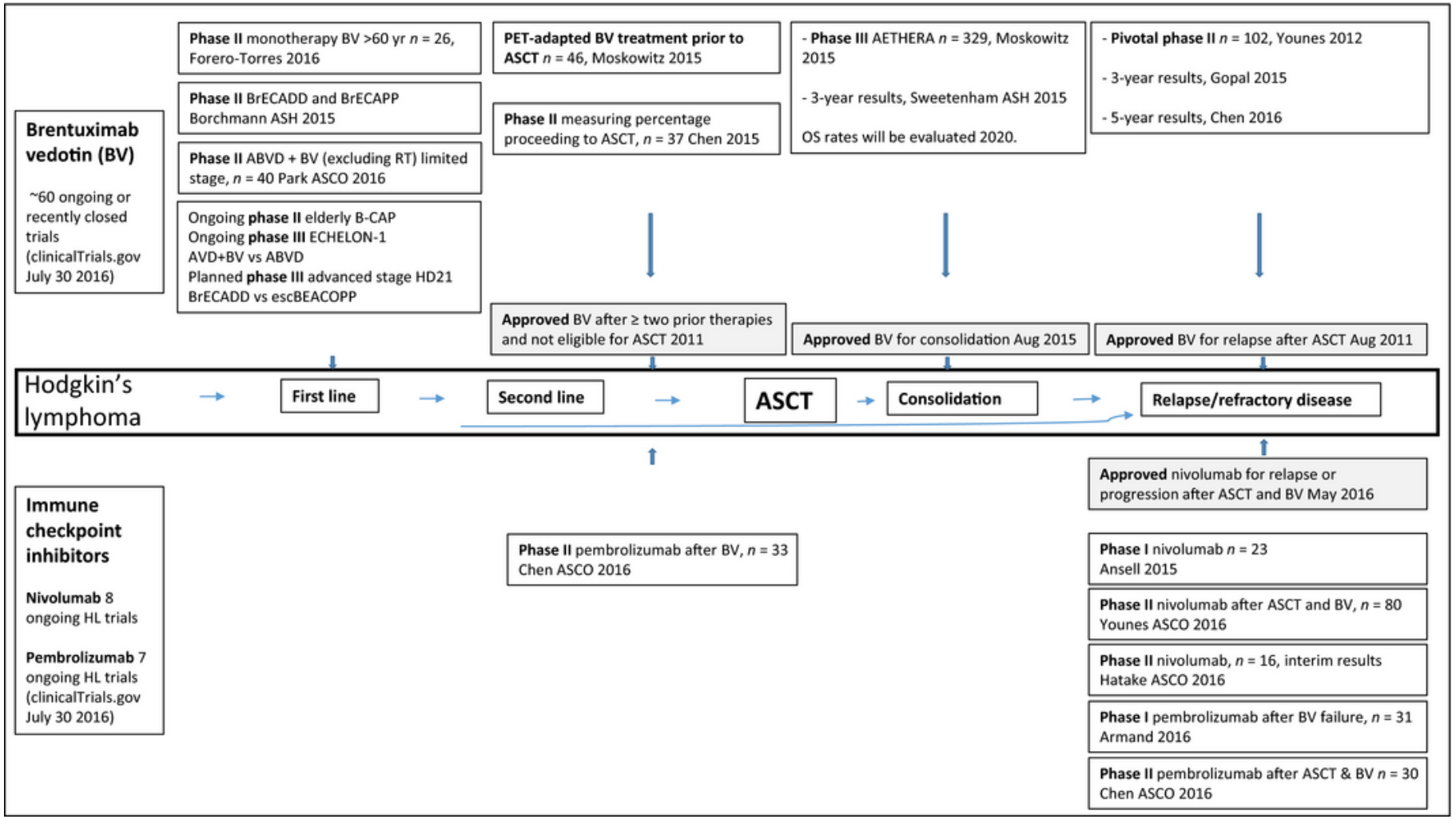

Figure 16. An overview of published Phase II and Phase III studies (taken from Glimelius and Diepstra, 2017).

1.10 Project aims

In the foregoing sections I have outlined the biology and clinical characteristics of two B cell lymphomas, Hodgkin lymphoma and Burkitt lymphoma, both of which are associated with infection with the Epstein-Barr virus.

In the case of HL, treatment with optimised combination chemotherapy regimens offers cure as a realistic probability for the majority. However, around 15-20% of HL patients will either fail to attain a complete response or will relapse early; these patients have significantly poorer long-term survival. Furthermore, multi-agent chemotherapy regimens and radiotherapy are associated with high rates of acute and long term toxicity. The appearance of secondary malignancies is a particular concern among paediatric HL sufferers as these patients are at the highest risk among all childhood cancer survivors. Identifying biomarkers of response to conventional chemotherapy in paediatric HL would enable patients to be stratified for more or less intensive regimens in turn maximising survival rates while at the same time reducing late effects. The development of targeted therapies for HL sufferers, alongside better stratification of patients for conventional therapies, should not only improve cure rates but could also help intensify conventional treatments.

For many cancer types the development of new targeted therapies and biomarkers of response has been underpinned by improved knowledge of disease biology, especially the contribution of genomic changes. This has not been the case for HL, since to date there have only been limited studies of the genomic landscape of primary tumours, mainly due to the low frequency of HRS cells in tumour biopsies. In Chapter 3, I set out to establish a new methodology for the analysis of the whole exomes of HRS cells micro-dissected from tumour biopsies of patients with HL, as the only exome study on HL primary tumours was done using

viable suspensions, however, majority of the patient samples are archived as flash frozen or paraffin-embedded.

Endemic BL is the most common malignancy of children in equatorial Africa comprising around 50% of all paediatric malignancies. Sporadic BL is also the most common type (30-40%) of childhood non-Hodgkin lymphoma. Although better performance was achieved with the current multi-agent chemotherapy and immunotherapy, relapse following treatment occur in small proportion of BL cases and current results in refractory/relapsed BL are extremely poor. Despite the fact that almost all cases of eBL and a sizeable fraction of sporadic BL are EBV-positive, there are currently no new EBV targeted therapies available for these patients. This is in part due to the previously reported limited expression of viral genes that can be targeted by pharmacological or by immunological based therapies. In Chapter 4, I set out to address the issue of viral gene expression in eBL, initialling describing global virus gene expression then focussing on the potential contribution of a novel latent gene, BILF1.

Specific objectives

In Part I, I will-

1. Establish the methodology for the whole exome sequencing of HRS cells micro-dissected from HL biopsies of paediatric HL.
2. Use this approach to describe the genomic landscape of paediatric HL
3. Identify key genes, and biological mechanisms relevant to the pathogenesis of HL using established bioinformatics approaches.

In Part II, I will-

1. Analyse the viral transcriptome of EBV-associated cancers, with a focus on eBL
2. Study the expression of a novel latent EBV gene, BILF1, in eBL.

3. Identify the cellular transcriptional targets of BILF1 in B cells and use these data to explore the biological impact of this EBV gene in BL.

CHAPTER 2

MATERIALS AND METHODS

2. MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 Hodgkin and Burkitt lymphoma cell lines

Hodgkin lymphoma (HL) derived cell lines KM-H2 and L-591, and Burkitt lymphoma (BL) cell lines DG-75 and AKATA were maintained in 90% RPMI 1640 (Sigma-Aldrich, UK) supplemented with 10% foetal calf serum (FCS), 2mM L-glutamine and 1% penicillin-streptomycin solution (Sigma-Aldrich, UK), which was exchanged every 72 hours.

2.1.2 Paediatric Hodgkin lymphoma samples

35 paediatric HL biopsies (flash frozen) and their matched blood DNA, collected from untreated patients, were provided by National Children's Cancer and Leukaemia Group (CCLG) Biological Studies Steering Group (CCLG reference number 2012BS07). This study fulfils the criteria mandated by MREC as approved by CCLG Biological Studies Steering Group. This study also obtained approval from National Research Ethics Committee (Reference number: ERN_13-0570**).** Samples were stored at -70^{o}C. Results shown in my PhD thesis were from 13 paediatric patient samples (Table 6).

2.1.3 Tonsillar samples

Paediatric tonsils were used to obtain germinal centre B (GC B) cells. These samples were obtained from the Children's Hospital Birmingham following informed consent (reference number for ethical approval: 06/Q2702/50).

Table 6. Paediatric Hodgkin lymphoma patient samples. NS= Nodular Sclerosis; MC= Mixed Cellular; NS-MC= Nodular Sclerosis-Mixed Cellular.

Sample	**HL subtype**	**Gender**	**EBV status**
S1	NS-MC	Male	EBV-positive
S2	NS	Male	EBV-positive
S3	NS	Male	EBV-negative
S4	NS	Female	EBV-negative
S5	NS	Female	EBV-negative
S6	NS	Female	EBV-positive
S7	NS	Female	EBV-negative
S8	NS	Female	EBV-negative
S9	NS	Male	EBV-positive
S10	NS	Male	EBV-positive
S11	NS	Male	EBV-negative
S12	MC	Male	EBV-positive
S13	NS	Male	EBV-negative

2.1.4 Samples used in the EBV gene expression analysis

I performed EBV gene expression analysis, using high-throughput q-PCR (Fluidigm) method, on 4 EBV-associated cancers; HL, eBL, NPC, and gastric cancers. All samples used in this study were initially diagnosed as EBV-infected.

HL samples included 24 primary HL tumours and L-591 cell line. Primary HL tumour samples were provided by National Children's Cancer and Leukaemia Group (CCLG) Biological Studies Steering Group (CCLG reference number 2012BS07). Work on these HL samples is approved by the NRES (National Research Ethics Committee) of the UK (Reference number: ERN_13-0570 which falls under an existing ethical approval ERN_12-0985).

A total of 41 primary eBL tumour samples were used in this study. These samples were collected from the Department of Human Pathology of the Lacor Hospital (Uganda, Africa), in endemic areas, with a written permission and informed consent in accordance with the Declaration of Helsinki. Ethics approval for this study was obtained from the Ethics and Research Committee at the Lacor Hospital (Uganda), from the Ethics and Research Committee at Moi University, Eldoret (Kenya), and from the Institutional Review Board at the University of Siena (Italy).

50 primary NPC tumour samples, which also included 9 RNA-amplified samples, and 9 NPC cell lines were used in this study. Complementary DNA (cDNA) of 41 primary NPC tumour samples and 9 NPC cell lines (which include C666.1 which is a naturally infected cell line, and *in vitro* Akata-infected cell lines CNE1, CNE2, HK1, HONE1, NP460, SUNE1, TWO1, and TWO4) were received from Dr. Yap Lee Fah, Faculty of Dentistry, University of Malaya, 50603, Kuala Lumpur, Malaysia. These NPC samples were obtained with an approval (code: KKM/NIHSEC/P13-494) from Medical Research and Ethics Committee, Ministry of Health

Malaysia. 9 NPC primary tumour samples from amplified RNA were provided by a group at the Institute of Cancer and Genomic Sciences, University of Birmingham. Ethical information for these samples can be obtained from a previously published paper by this group (Hu et al., 2012).

Four gastric cell lines used in this study were either naturally infected with EBV (SNU-719 and YCCEL-1) or infected *in vitro* with the Akata (AGS and OE19). CDNA for these samples was provided by Dr. Helena Carley at the Institute of Cancer and Genomic Sciences, University of Birmingham.

2.1.5 Samples used in the RNA *in situ* hybridization experiment

I performed RNA *in situ* hybridization experiment (RNAScope) (Wang et al., 2012) on 15 eBL and two paediatric HL samples that were initially diagnosed as EBV-infected. Source and ethical approval for these samples are described in the previous section.

2.1.6 RNA sequencing data obtained from SRA and GEO

RNA sequencing data for 20 eBL samples was obtained from NCBI sequence Read Archive (SRA) service (accession number PRJNA292327). RNA sequencing data for four germinal centre (GC) B cell samples was obtained from NCBI Gene Expression Omnibus (GEO accession number GSE45982).

2.2 METHODS

2.2.1 Sample preparation for immunohistochemistry and LCM

HL cell line KM-H2 was used to optimize methods to obtain good quality tumour DNA for exome sequencing, including CD30 staining and laser capture microdissection (LCM) of CD30-positive HRS cells. Culture media with KM-H2 cells was centrifuged at 3000 rpm for 5 minutes and the pellet was washed and resuspended in 500 µL PBS (pH 7.6).

Approximately 200,000 cells were used to generate monolayer cell preparations onto RNAse and DNAse-free polyethylene naphthalate (PEN) membrane slides (Sample MembraneSlide 1.0 PEN (D), CarlZeiss Ltd.) by cyto-centrifuging at 1000 rpm for 5 minutes in the Cytospin2 (Shandon, UK). PEN membrane slides were air dried for 1 minute and fixed in cold ethanol (75%) for 2 minutes.

Frozen sections of paediatric HL primary tumour samples were prepared at 8 µm using a decontaminated, RNAse and DNAse-free standard cryostat and placed onto RNAse and DNAse-free PEN membrane slides. PEN membrane slides were air dried for 10-30 minutes and fixed in cold ethanol (75%) for 2 minutes.

2.2.2 Immunohistochemistry

Tissue sections were stained with Haematoxylin and Eosin (H&E) to investigate morphology and abundance of the HRS cells. These sections were later stained for EBER to determine EBV status. Tissue sections were stained for CD30 positivity using CD30 antibody and Haematoxylin on Leica BOND-MAXTM autostainer (Leica Biosystems) wherein sections were treated for antigen retrieval at 65 C for 30 minutes. Endogenous blocking was performed by placing 0.3% hydrogen peroxide H_2O_2 solution on sections for 15 minutes. Primary antibody (monoclonal mouse anti-human CD30 (Clone Ber-H2), Dako, UK Ltd) at 1:10 dilution was applied on sections for 15 minutes. Secondary antibody (Dako REAL EnVision HRP Rabbit/Mouse (ENV)) was applied on sections for 5 minutes. AB substrate (980ul Dako REAL substrate buffer + 20ul Dako REAL DAB+ chromogen) was placed on sections for 1 minute. Slides were washed twice between each of the above step.

For staining Formalin-fixed and paraffin-embedded (FFPE) samples, tissue sections (4µ) were deparaffinized in xylene and rehydrated by washing in serial dilutions of ethanol (96%, 80% and 70%) and rinsed in deionized water. For blocking endogenous peroxidase activity,

sections were treated with 0.3% H_2O_2 solution for 15 min. Sections were then treated for antigen retrieval by immersing sections in Citrate buffer (pH6.0; 1.26g sodium citrate+0.25g citric acid+1L distilled water) and boiled using microwave for 30 minutes. Sections were rinsed in tap and deionized water. Sections were washed in deionized water for 5 min, then twice in 0.05M Tris-buffer (pH 7.4-7.6), once in Tris-buffer with 0.5% Tween solution and incubated with primary antibody (monoclonal mouse anti-human CD30 (Clone Ber-H2), Dako, UK Ltd) at 1:10 dilution was applied on sections for 60 minutes at room temperature in a humid chamber. Slides were washed twice in Tris-buffer, once in Tris-buffer with 0.5% Tween solution, and incubated with secondary antibody (Dako REAL EnVision HRP Rabbit/Mouse (ENV)) was applied on sections for 5 minutes. Slides were washed twice in Tris-buffer and AB substrate (980ul Dako REAL substrate buffer + 20ul Dako REAL DAB+ chromogen) was placed on sections for 1 minute and then counterstained in haematoxylin, dehydrated through alcohols and xylene.

2.2.3 Laser capture microdissection of HRS cells

I used ultraviolet (UV) cutting system (PALM MicroBeam, CarlZeiss Ltd) to micro-dissect six sets of approximately 50 HRS cells from each patient sample mounted on PEN membrane slides. HRS cells were identified with the help of a trained pathologist (Dr. Ghada Mohammed). The microdissected cells were captured on PALM Adhesive cap tubes (AdhesiveCap 500 clear (D), CarlZeiss, Ltd). A schematic diagram of the UV-LCM procedure is depicted below (Figure 1). For each sample, the six sets were split into two groups, each containing 3x50 HRS cells, and the two groups were treated as biological replicates from here onwards.

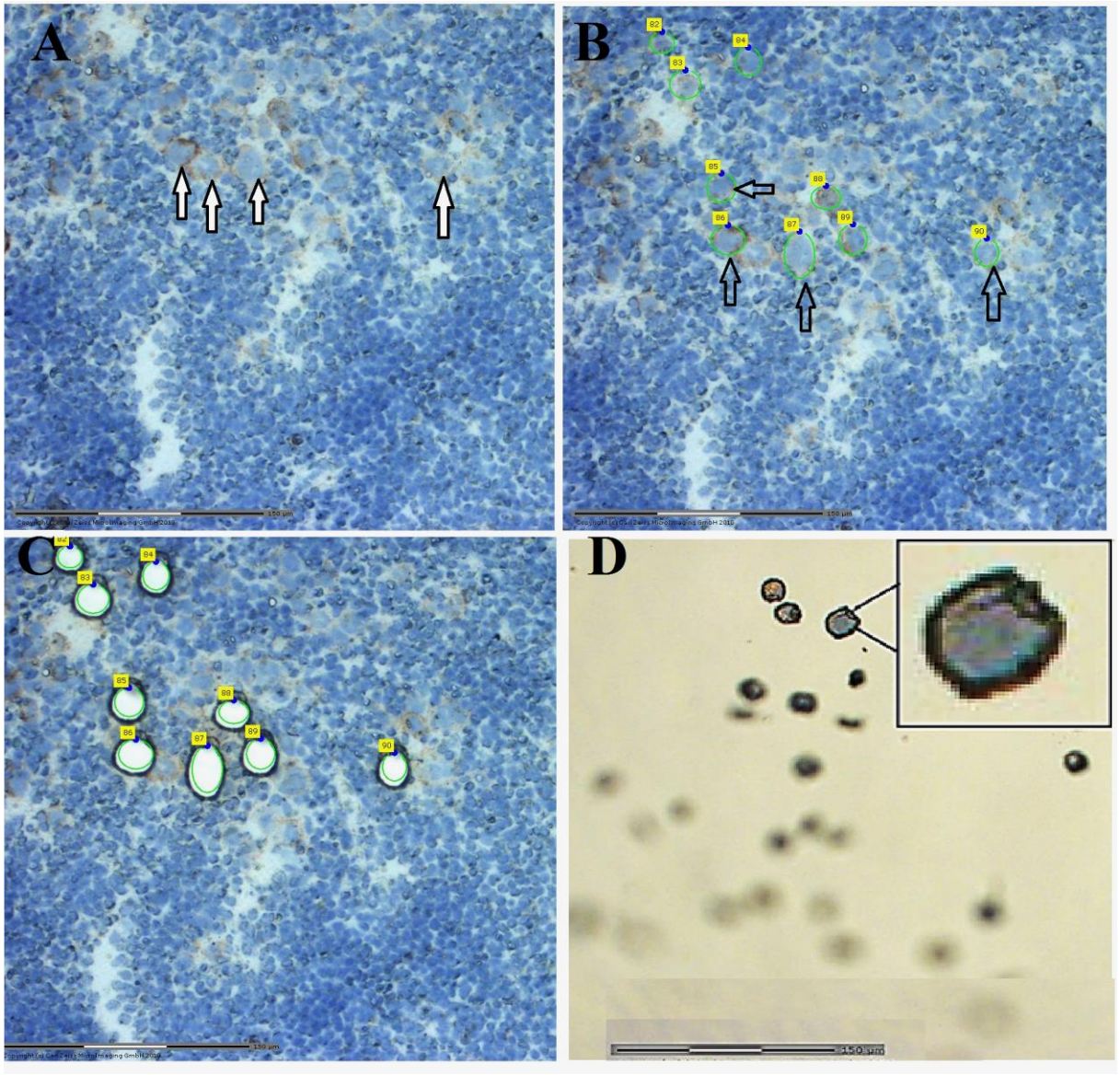

Figure 17. Laser capture microdissection of HRS cells. The above figure depicts the schematic of LCM method. (Cells were pointed with arrows) A. Visualisation of the cells of interest, in this case CD30 stained HRS cells (large multi-nucleated cells with membrane stained in brown) in a frozen tissue section of a patient biopsy. B. Selection of the cells of interest C. Cutting of the selected cells using laser energy D. Visualisation of intact HRS cells that were captured on the adhesive cap (images scale to 150 µm and HRS cell shown in a box in D is enlarged further to x4).

2.2.4 Whole genome amplification

HL cell line KM-H2 was used to optimize methods including whole genome amplification (WGA) of DNA from microdissected HRS cells, purification of WGA DNA samples, and quality and quantity assessment of whole genome amplified DNA samples. Genomic DNA from 50 HRS cells in each tube were amplified in a total of 50 µl reaction volume using GenomePlex® Single Cell Whole Genome Amplification Kit (WGA4) according to manufacturer's protocol (Sigma-Aldrich Ltd.), along with a positive control and negative control reactions. WGA products from three 50 µl reactions were pooled to obtain a technical replicate for that paediatric HL sample. Thus I have two technical replicates from the six sets containing 50 HRS cells. These two technical replicates for each sample were sequenced separately.

I performed quality and quantity assessment of WGA products prior to sample submission for exome sequencing. WGA products were cleaned according to manufacturer's protocol using Genomic DNA Clean & ConcentratorTM (Zymos Res. Corp.) and eluted with nuclease-free distilled water. Quality and quantity of WGA products and matched blood DNA samples was assessed using Nanodrop spectrophotometer (Thermo Scientific) and Qubit (ThermoFisher Scientific), respectively. Fragment sizes of the whole genome amplified products ranged between 200 base pairs (bp) and 1 kbp when separated on 2% agarose gel confirmed by Bioanalyzer (Agilent Technologies). WGA products were also randomly checked for coverage by PCR and Sanger sequencing using randomly selected 18 housekeeping genes present on various chromosomes (Appendix 2).

2.2.5 Whole exome sequencing

Initial sequencing on some samples was performed at Oxford Genomics Centre (Oxford, UK) and the remaining samples were sequenced at Edinburgh Genomics Centre (Edinburgh, UK).

Exome sequencing required at least 1 μg of DNA with quality 1.8-2.0 for 260/280 ratio and 2.0-2.2 for 260/230 ratio. 3/26 tumour exomes and 4/13 normal exomes were captured using NimbleGen SeqCap EZ Exome Kit v3.0, Roche (Hiatt et al., 2010) and the remaining exomes in this study were captured using Nextera Rapid Capture Exome enrichment kit, Illumina Inc (Caruccio, 2011). The captured exome library was sequenced using Illumina HiSeq 2000 to obtain 100-fold average coverage of each base in the targeted region and greater than 95% of the exome sequence was expected to be covered at sufficient depth for variant calling (≥20X coverage).

2.2.6 Whole exome sequencing data analysis

Exome sequence reads for each sample were received in FastQ format from sequencing service providers. Exome sequencing data for technical replicates for each sample were analysed separately and common variations between these two technical replicates for each sample were considered for further investigation.

Preparation of input files

Human reference sequence hg19 was downloaded from UCSC (Speir et al., 2016) in order to align raw exome sequence reads in FastQ files. Nextera Rapid Capture Exome target BED files and Nimblegen exome capture target BED files were used to assess the coverage in the exome sequences. Nextera Rapid Capture Exome targeted regions manifest file (.bed) lists the 212,158 targeted exome regions with start and stop chromosome locations in hg19. Nimblegen exome capture target BED file contains 368,146 target regions covering more than 20,000 genes in the human reference genome (hg19).

Quality assessment of the reads by FastQC software

FastQC software was used to examine read base quality, sequence quality, length, over-represented reads and extent of sequence duplication in FastQ files. Sequences with a mean Phred quality score greater than 33 were considered as good quality sequence reads and were used in further analysis.

Trimming of adapters and overrepresented sequences

I used Trimmomatic tool (Bolger et al., 2014) to remove bad quality sequences (Phred quality score less than 33), adapters that were used in library preparation, and overrepresented sequences from the raw FastQ files. This process is designed to improve mapping of sequence reads to the reference sequence.

Alignment of exome sequencing reads to human reference genome

I aligned exome sequences to the human reference genome (hg19) using Burrows Wheeler Aligner BWA (Li and Durbin, 2009). Reads uniquely mapped to hg19 were considered for further statistical analysis. Mate pair reads were used to create SAM (Sequence Alignment/Map format) files and these SAM files were then converted to BAM (Binary Alignment/Map format) files using Picard tools (McKenna et al., 2010). Reads were sorted according to the genomic coordinates and an index file (bai) was created for each BAM output file. Multiple BAM files (of same replicate sample) originating from different sequencing lanes were merged using SAMtools (McKenna et al., 2010).

Marking PCR duplicates in exome sequencing reads

I marked PCR duplicates using Picard tools. These duplicate reads share the same sequence and alignment position and were generated due to artefacts introduced during PCR amplification during WGA and sequencing process. Marking and removal of PCR replicates will remove most of these artefacts and reduce the number of false positives in variant calling.

Karyotypic sorting of reads in exome sequencing reads

I sorted reads in the marked BAM file as per Genome Analysis Tool Kit (GATK) guidelines on the ordering of the chromosome contigs in BAM file. This was done to make contigs identical to the canonical ordering in the human reference file (hg19) I used to create these BAM files.

Local alignment of exome sequences around Indels using tumour and normal samples

To reduce false positives during variant calling, I performed local alignment of reads around indels using GATK software. Mate pair information between the mates in the BAM files was checked and fixed with Picard tools.

Quality score recalibration and coverage statistics

I calibrated quality scores using GATK software before calling for variants in these BAM files. I calculated the coverage (breadth and depth) of the aligned, de-duplicated, karyotypically sorted, realigned, and calibrated sequence reads using Nextera Rapid Capture Exome target BED files or Nimblegen exome capture target BED files with BEDTools. Coverage statistics was performed were calculated using R program.

Variant calling in exome sequencing reads

I used Strelka (Saunders et al., 2012) (Bayesian), MuTect1.0 (Cibulskis et al., 2013) (Bayesian with 7 additional filters), VarScan2.0 (Koboldt et al., 2013) (Heuristic/Statistic), and Mutect2.0 to identify somatic mutations in exome sequencing reads. I used default settings when calling somatic mutations using Strelka v1.0.14, Mutect v1.0 and Mutect2.0. VarScan2.0 needs input files in pileup file format and can be generated from BAM files using SAMTools software. After initial analysis on variants generated using criteria at different stringency levels, I adopted the following variant calling criteria that is more sensitive in

obtaining point mutations: normal read depth of ≥ 8, tumour read depth of ≥ 1, and tumour allele frequency of ≥ 0.05, and tumour allele frequency of ≥ 0.005.

Filtering somatic mutations identified by variant callers

Following several iterations of analysis on various stringency levels for filtering mutations, I adopted the following filtering criteria. Normal allele frequency ≤ 0.005 was applied to remove most of the germline mutations. Somatic indels from VarScan2.0 were further filtered using median distance from the nearest end (MDFNE) ≥ 10 criteria using VariantTools to remove PCR and sequencing artefacts usually found at the end of the reads (San Lucas et al., 2012).

Annotation of somatic mutations

I used Ensemble GRCh37.75 annotations to annotate somatic mutations using snpEff software (version 4.1g (build 2015-05-17)).

Validation of somatic mutations by Sanger sequencing

I validated a set of selected somatic mutations called by Strelka, MuTect1.0, and VarScan2.0 using Sanger sequencing method. PCR and sequencing primers were designed using Primer-BLAST in NCBI (Ye et al., 2012). Primers were designed with a maximum PCR product length of 200 bp and sequencing products less than 160 bp. PCR was performed on WGA products of replicates for each sample. A list of PCR primers can be found in Appendix 2.

2.2.7 Gene Ontology analysis of mutated genes

I performed gene ontology (GO) analysis of mutated genes using DAVID version 6.8 (Huang da et al., 2009) Functional annotations for biological processes enriched in the mutated genes were obtained using GOTERM_ BP-FAT, which filters very broad GO terms based on a measured specificity of each term.

2.2.8 Gene set enrichment analysis of mutated genes

Gene set enrichment analysis of mutated genes was performed using R program. Odds ratios and ***p***-values were calculated using Chi square test and t-Test.

2.2.9 RNA extraction

RNA was isolated from cultured cells using the QIAGEN RNeasy™ MICRO ($<5x10^5$ cells) kit according to manufacturer's instructions (Qiagen Ltd, Crawley, UK). Briefly, cells were disrupted in RLT lysis buffer containing a denaturant guanidinium isothiocyanate. β-mercaptoethanol was added to inactivate RNAse enzymes by reducing disulfide bonds in their native conformation. Cell lysates were then homogenized by vigorous vortexing for 1 minute. Samples were thoroughly mixed with 1 volume of 70% ethanol to provide ideal binding conditions and loaded onto the RNeasy silica membrane of MICRO spin columns. This was followed by centrifugation at 10000 rpm and washing of the spin column membrane with RWI buffer. Samples were then incubated for 15 minutes with DNAse solution (RNase-Free DNase Set; QIAGEN Ltd., West Sussex, UK) at room temperature to remove contaminating bound DNA. Salts, metabolites and cellular components were washed off in several washing steps with RWI, RPE and 80% ethanol buffers. Finally, the spin columns were centrifuged with their lids open in order to completely dry column membranes and RNA was eluted with DNase/RNase-free.

RNA from formalin-fixed paraffin-embedded (FFPE) samples RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE (ThermoFisher SCIENTIFIC Ltd.). Several FFPE sections at 80 µm thickness were incubated in 1ml 100% xylene for 3 minutes at 50 C for 3 minutes in RNAse-free Eppendorf tube. This tube was centrifuged at maximum speed for 2 minutes and xylene was discarded. Then, the pellet was washed twice with 1 mL 100% ethanol. Pellet was air dried to remove any residual ethanol. To this pellet protease-digestion buffer was added

and incubated at 50 C for 30 minutes. Then, isolation buffer was added to the tube and mixed before passing through a filter cartridge. Filter cartridge was then washed with 700 µl of Wash1 buffer and centrifuge for 1 minute; I repeated this step with 500 µl Wash2 buffer followed by Wash3 buffers. DNase buffer was added to the filter cartridge and incubated for 30 minutes. I then repeated the earlier wash up steps with Wash 1-3 buffers. Cartridge was then centrifuged for 1 minute to remove any residual fluid. Then, I added 60 µl of nuclease-free water at room temperature. This eluate was stored at -20 C until further use.

2.2.10 Specific target amplification and 48:48 dynamic (Fluidigm) array analysis

A subset of EBV transcripts (n=45) that include examples of immediate early (IE), early (E) and late (L) lytic cycle genes in addition to latent cycle genes were quantified by Real-Time PCR using a 48:48 dynamic array as described (Spurgeon et al., 2008, Tierney et al., 2015). These transcripts include 23 latent, 2 immediate early (IE), 11 early (E), 6 late (L) lytic cycle transcripts; 3 poorly characterized RNAs (LF1, LF2 and LF3) and 4 non-coding RNAs (BARTs and EBERs)). 3 cellular housekeeping genes (B2M, PGK1 and GAPDH) were also included in this analysis. Briefly, 50 ng RNA was used for cDNA synthesis according to the manufacturer's protocol (qScript, VWR). An aliquot of cDNA was then subjected to specific target amplification (STA) to pre-amplify 45 EBV amplicons and 3 cellular amplicons in the cDNA, RNA (non-reverse transcription control) and AQ-plasmid standards (ranging from 1 to 10^5 copies/µl) prior to Fluidigm assay (12- 14 cycles of PCR). Pre-amplified samples were loaded onto 48.48 Dynamic Array™ IFC and Real-Time PCR was performed using BioMark™ HD System according to manufacturer's guidelines (Fluidigm Corporation). Real-Time PCR data was analysed using Biomark & EP1 software (Fluidigm Corporation). After subtraction of non-RT values (where present), individual gene expression values of each sample were normalized for PGK1 expression. Samples with low or no cellular gene

expression (B2M, GAPDH, and PGK1) and/or with low EBER expression were excluded from further analysis.

2.2.11 Plasmid vectors

BILF1 gene expression-vector with a 5′-HA tag and the alternative vector, pLZRS-IRES-GFP, were obtained from Dr. Jianmin Zuo, Institute of Genomics and Cancer Sciences, University of Birmingham (Zuo et al., 2009).

2.2.12 BILF1-transfection in DG75 and germinal centre B cells

In order to identify cellular target genes for EBV gene BILF1 in BL cell line DG75 and germinal centre B (GC B) cells, cultured DG-75 cells and GC B cells obtained from tonsillar tissue were transfected with BILF1 expression-plasmid. GFP-positive cells were sorted and RNA sequencing was performed on the either RNA (DG-75) or amplified RNA (GC B cells) from these cells. These methods are described in detail in the following sections.

All the experimental work on GC B cells in the next sections was performed by Dr. Katerina Vrzalikova at the Institute of Cancer and Genomic Sciences, University of Birmingham, UK.

2.2.13 Purification of tonsillar mononuclear cells (TMCs)

Tonsils were minced with a scalpel in cold RPMI 1640 medium (Sigma-Aldrich Ltd., Gillingham, UK) without foetal calf serum, supplemented with 2mM L-glutamine (Invitrogen Ltd., Paisley, UK), 0.5% ciprofloxacin (Bayer, Newbury, UK) and 1% penicillin-streptomycin solution (Sigma-Aldrich Ltd., Gillingham, UK). Tonsillar mononuclear cells (TMCs) were then isolated by Ficoll-Isopaque centrifugation. 15 ml of Lymphoprep® solution (Axis-Shield Diagnostics Ltd. UK, Dundee, UK) was pipetted into universal tubes and carefully overlaid with media containing a mixture of tonsillar cells. Tubes were then centrifuged at 2200 rpm at room temperature for 30 minutes creating the following visible layers: plasma and other

constituents (top), mononuclear cells (middle), Ficoll-Paque, and erythrocytes and granulocytes debris (pellet). The layer of mononuclear cells was transferred to a fresh tube, washed twice with cold media and once with cold autoMACS™ Rinsing Solution (autoMACS; Miltenyi Biotec Ltd, Surrey, UK) supplemented with 1% penicillin-streptomycin solution (Sigma-Aldrich Company Ltd., Gillingham, UK), 0.5% ciprofloxacin (Bayer, Newbury, UK) and 5% MACS® BSA Stock Solution (Miltenyi Biotec Ltd, Surrey, UK); each time centrifuged at 900 rpm at 4°C for 10 minutes.

2.2.14 Purification of GC B cells

GC B cells (CD10+) were isolated from TMCs by positive enrichment of CD10+ cells using magnetic separation with anti-CD10-Phycoerythrin (eBioH35-17.2 (H35-17.2) PE) (eBioscience, San Diego, CA, USA), anti-PE microbeads and LS columns (both Miltenyi Biotec Ltd, Surrey, UK). Magnet and autoMACS buffer were pre-cooled at 4°C and all purification steps performed on ice to prevent rapid apoptosis of GC B cells and unspecific antibody binding. First, 10^7 TMCs were resuspended in 100 µl of autoMACS and CD10-PE antibody was added at 1:50 dilution. Following incubation at 4°C for 15 minutes, cells were washed with 10 volumes of autoMACS and centrifuged at 900 rpm at 4°C for 10 minutes. 10^7 CD10-PE-labelled TMCs were then resuspended in 80 µl of autoMACS and 20 µl of anti-PE beads were added. Following incubation at 4°C for 15 minutes, cells were washed with 10 volumes of autoMACS and centrifuged at 900 rpm at 4°C for 10 minutes. Finally, 10^6 CD10-PE-anti-PE-beads-labelled TMCs were resuspended in 500 µl of autoMACS and transferred on pre-washed filters (Partec UK Limited, Canterbury, UK) placed on LS columns. When the cells passed through; columns and filters were washed three times with 3 ml of autoMACS. Finally, columns were removed from the magnet and retained population of CD10+ positive GC B cells was eluted with 5 ml of autoMACS.

2.2.15 Transfection of GC B cells and DG-75 cell line by nucleofection

10 µg of BILF1 or control plasmid with the marker gene GFP were transiently transfected into GC B cells using non-viral Nucleofector technology with the Nucleofector II device (Amaxa GmbH, Cologne, Germany). For each reaction, 10^7 cells were pelleted by centrifugation in universal tubes at 636 rpm at room temperature for 10 minutes. Supernatant was completely removed and the cell pellets were resuspended in 100 µL of Nucleofector reaction mix consisting of Cell Line Nucleofector Solution B mixed with Supplement I. The plasmids were combined with the cell suspensions and transferred to a 0.1cm cuvette. The cuvette was placed in to the Nucleofector II device and U-15 program was selected. Following the pulse, the entire contents of the cuvette were immediately transferred into 24-well plates (Nalge Europe Ltd., Hereford, UK) using plastic pipettes and a small amount of pre-warmed RPMI. The 24-well plates were prepared in advance and filled with 2 ml of RPMI containing supplemented with 20% of foetal calf serum. Cells were then incubated at 37°C in 5% CO2 for 16 hours (overnight) prior to MoFlo enrichment of transfected cells.

2 µg of BILF1 or control plasmid with the marker gene GFP were transiently transfected into DG-75 cells using the Nucleofector II device. For each reaction, $2x10^6$ cells were centrifuged at 720 rpm at room temperature for 10 minutes. After removing the supernatant, the cell pellets were resuspended in 100 µL of Nucleofector reaction mix consisting of Cell Line Nucleofector Solution V mixed with Supplement I. The plasmids were combined with the cell suspensions and transferred to a 0.1cm cuvette. The cuvette was placed in to the Nucleofector II device and R-013 program was selected. The latter steps were same as for the GC B cell transfection, mentioned in the paragraph above.

2.2.16 MoFlo enrichment of transfected GC B cells and DG-75 cell line

Following incubation, transfected GC B cells were pooled together, washed twice with cold autoMACS and pelleted at 636 rpm at 4°C for 10 minutes The cell pellets were resuspended in 250 µl of autoMACS and re-stained with anti-CD10-Phycoerythrin (PE) (eBioscience, San Diego, CA, USA) antibody. Following 10 minutes incubation at 4°C in the dark, cells were washed with 6 ml of autoMACS and centrifuged again at 636 rpm at 4°C for 10 minutes. Cells were passed through pre-washed filters into sorting tubes and kept on ice in the dark. The PE-labelled GFP-positive cells were collected by FACS on a MoFlo sorter (Dako Cytomation, Colorado, USA) using propidium iodide (Sigma-Aldrich Ltd., Gillingham, UK) to distinguish between viable and dead cells. The transfection efficiency of the living cells was generally between 10%-20% and purity of the collected cells was >95%.

Transfected DG-75 cells were enriched in a similar manner to those of GC B cells as described above. Transfection efficiency of the living cells was generally between 15%-20% and purity of the collected cells was >95%. Real-time PCR were performed to validate successful BILF1 gene expression in transfected cells.

2.2.17 RNA extraction

RNA was isolated from BILF1-transfected and empty vector-transfected GC B-cells (triplicates) and DG-75 cells (triplicates) as described in methods section 2.7. Quality of isolated RNA was determined with a Bioanalyzer 2100 (Agilent, Waldbronn, Germany). Only those samples with a 260/280 nm ratio between 1.8 and 2.1 and a 28S/18S ratio within 1.5-2.0 were processed further.

2.2.18 Amplification of RNA extracted from BILF1-transfected GC B cells

As the amount of RNA obtained from BILF1-transfected GC B-cells was not sufficient for RNA sequencing, Dr. Katerina Vrzalikova amplified the extracted RNA using Ovation RNA-

Seq System V2 (NuGEN Technologies, San Carlos, California), according to the manufacturer's protocol.

2.2.19 RNA sequencing

CDNA obtained from amplification of BILF1-transfected GC B-cells was sent for library preparation (TruSeq DNA Nano gel free library (350 bp insert)) and RNA sequencing to Edinburgh Genomics (HiSeq2500 50 base paired end sequencing). RNA isolated from BILF1 and empty vector-transfected DG-75 cells (triplicates) were sent to Macrogen Inc, Seoul, Republic of South Korea (http://www.macrogen.com/eng/business/ngs_overview.html) for Illumina 100 bp paired-end RNA sequencing.

2.2.20 RNA sequencing data analysis

I aligned RNA sequence reads to hg19 and EBV (NC_007605.1) reference sequence using Rsubread aligner (Liao et al., 2013). Mapped sequencing reads were assigned to hg19 refGene and EBV genes using featureCounts function. RefGene exon coordinates were obtained from UCSC table browser. Gene symbol and description was obtained from NCBI gene database. EBV exon coordinates were obtained from Eric Flemington's lab (http://www.flemingtonlab.com/Resources/chrEBV(B_95_8_Raji).ann).

Refinement of EBV-annotations

I made the following changes to the EBV gene annotations obtained from NCBI (AJ507799.2), so that it can be used for calculating featureCounts from RNA sequencing data. A). I limited the annotation to the known EBV ORFs (coding sequence or CDS) as these coordinates are known exactly. I removed the ORFs of BWRF1 (in the W repeats) which has recently been shown to give rise to a non- coding RNA of unknown function and does not produce protein to the best of my knowledge. B). For RPMS1 and A73, I included the coordinates of the exons as the ORFs of RPMS1 and A73 transcripts do not produce

detectable protein. C). I excluded the annotations that refer to just exons as, in many cases, the true starts or ends of these transcripts are not precisely mapped. The exceptions are the alternatively spliced latent transcripts which are well characterised and here I have included notes on the untranslated regions (UTR). Thus for the EBNA transcripts I have labelled the C1, C2 and W0 exons. W1 and W2 exons encode EBNA-LP but are also present in all EBNA transcripts as a common 5' leader sequence. Thus the presence of W1 and W2 reads does not imply expression of EBNA-LP. D). I used coordinates of the miRNA precursors not the mature miRNAs, obtained from miRBase (Griffiths-Jones et al., 2006). E). I excluded all annotation of regulatory regions and TR. F). The UTR of LMP1 overlapped with CDS of BNLF2A and BNLF2B and was therefore removed. G). Redundant overlapping entries for the same genes were removed, including LMP2A (genomic coordinate position: 1574-1682), LMP2B (59-272), LMP2B (1574-1682).

Data was normalized using TMM (trimmed mean of M values) method using edgeR (Robinson et al., 2010). EBV genes were not used in the normalization of human genes. The experiment was paired with each BILF1 treated replicate paired with one control-treated replicate. Differentially expressed genes were identified using edgeR with pair number and BILF1 status as explanatory variables. 26438 genes (from hg19 refGene) were used to generate these differentially expressed genes. For transfected samples, the following criteria were applied for obtaining details of differentially expressed genes:

1. First, differentially expressed gene lists were generated without any CPM criterion,

2. For upregulated genes, at least two BILF1-transfected samples (in triplicate) must have >=1 CPM, and

3. For downregulated genes: at least two Vector-transfected samples (in triplicate) must have >=1 CPM.

2.2.21 Pathway analysis

I used Molecular Signatures Database (MSigDB) (Liberzon et al., 2015) to identify molecular functions, biological processes, and pathways that were significantly enriched in differentially expressed genes. I calculated P value and FDR q-value for the above analyses to determine statistical significance of enrichment.

2.2.22 Lytic cycle induction in Akata cells

I induced Akata cells into lytic cycle using anti-IgG Capell Ab as follows: Cultured Akata cells were split 24 hours before required. Cells were then transferred to a 50ml centrifuge tube and spun down at 1600 RPM for 5 minutes. Supernatant was discarded and cell pellet was washed once in PBS, and cells were spun-down one more time and supernatant was discarded. Cells were resuspended in RPMI + 1% FCS at a concentration of 4 x 106 per ml and an aliquot was removed as a negative control. To the remaining cells, Capell anti-IgG (ThermoFisher Scientific Inc.)was added to a final concentration of 0.1% v/v. Cells were transferred to a 12-well plate (4 x 10^6 cells per well) and incubated at 37°C for 3 h. 3 ml of RPMI + 10% FCS to each well and incubated at 37°C. Cells were harvested after 48 hours using FACS which was performed by Dr. Rosemary Tierney at the Institute of Cancer and Genomic Sciences, University of Birmingham, UK.

2.2.23 RNA in situ hybridization experiment by RNAScope

This experiment was performed by Dr. Max Robinson at the Centre for Oral Health Research, Newcastle University, UK, using the manufacturer's protocol.

BILF-1 RNA in situ hybridisation (ISH) was carried out on 4μm formalin-fixed paraffin-embedded (FFPE) sections using proprietary reagents (RNAscope, Advanced Cell Diagnostics Inc.). Sections were deparaffinised and pre-treated with heat and protease before hybridisation with target-specific probes: BILF-1, PPIB (constitutively expressed endogenous

gene; positive control probe) and dapB (bacterial mRNA; negative control probe) in a dedicated hybridization oven (HybEZ oven, Advanced Cell Diagnostics Inc.). Probe hybridization was detected using the chromogen 3,3'-diaminobenzidine (DAB). BILF-1 RNA ISH was optimized using FFPE cell pellets, namely Akata cells (negative control for EBV gene expression), EBV-positive Akata cells (positive control for EBV expression), lytic cycle-induced Akata cells (as lytic cycle positive control). PPIB RNA ISH was used to demonstrate the presence of hybridisable RNA and dapB RNA ISH was used assess non-specific hybridisation. Test samples were classified as BILF-1 positive or negative, a positive result was defined as any brown reaction product that co-localised with the malignant cells, a negative result was defined as no staining with BILF-1, but evidence of staining with PPIB RNA ISH (positive control probe).

2.2.24 BILF1 antibody peptide design and Immunohistochemistry

BILF1 antibody was commercially not available, so I designed a custom antibody against Epstein-Barr virus BILF1 (YP_401711.1) using the following BILF1 peptide sequence (312 amino acids):

MLSTMAPGSTVGTLVANMTSVNATEDACTKSYSAFLSGMTSLLLVLLILLTLAGILFII
FVRKLVHRMDVWLIALLIELLLWVLGKMIQEFSSTGLCLLTQNMMFLGLMCSVWTH
LGMALEKTLALFSRTPKRTSHRNVCLYLMGVFCLVLLLIIILLITMGPDANLNRGPNM
CREGPTKGMHTAVQGLKAGCYLLAAVLIVLLTVIIIWKLLRTKFGRKPRLICNVTFTG
LICAFSWFMLSLPLLFLGEAGSLGFDCTESLVARYYPGPAACLALLLIILYAWSFSHFM
DSLKNQVTVTARYFRRVPSQST

BILF1 is a membrane protein and the predicted topology is shown in Figure 18.

A homology search (BLAST) was performed for BILF1 peptide sequence in order to optimise any wanted or minimise any unwanted homology to other proteins. The following two peptides were selected for immunization programme in rabbits:

1) aa 298-312: h- C+TVTARYFRRVPSQST –oh (cytoplasmic tail),

2) 159-174: ac- ITMGPDANLNRGPNMC –nh2 (part of an extracellular loop)

These two peptides are 100% conserved in B95, Akata, GD1, C666 and Ag876 (type 2).

BILF1 protein expression in FFPE sections was performed using the method described in section 1.12.2 using BILF1 antibody at 1:20 dilutions.

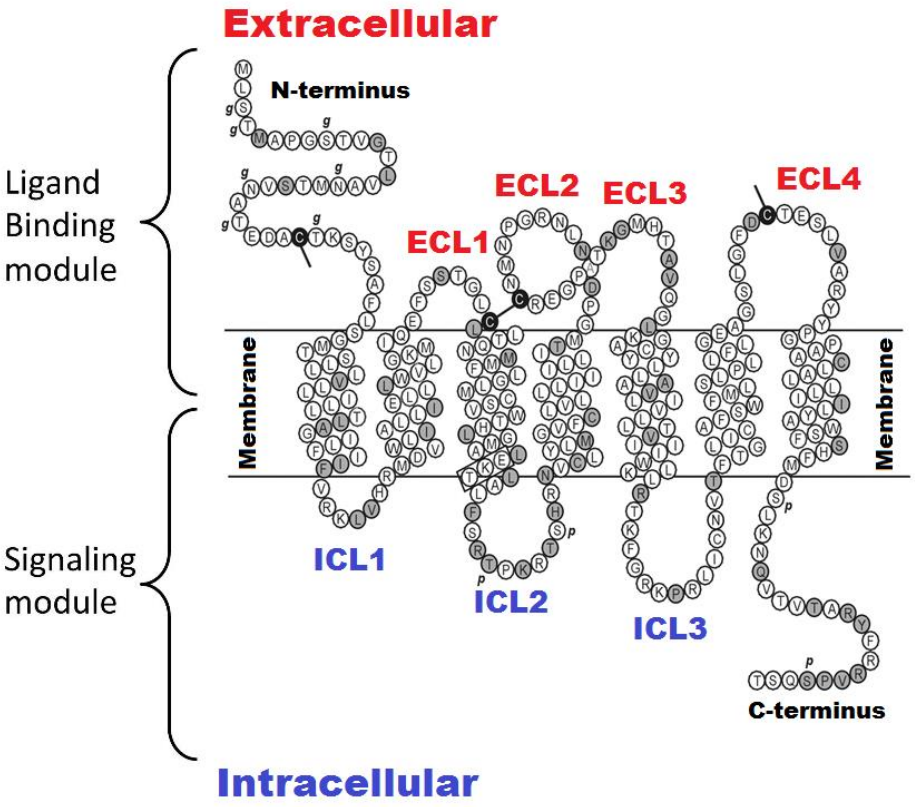

Figure 18. Serpentine diagram of the human EBV BILF1 receptor (taken from Paulsen et al., 2005). In this figure, seven transmembrane helices of BILF1 receptor can be seen. Differences between human EBV and rhesus EBV are indicated in black on grey, and identical amino acids are indicated in black on white. Computer-predicted glycosylation sites are indicated with g, and predicted phosphorylation sites are indicated with p. The alternative DRY box, EKT, is marked with a rectangle.

CHAPTER 3

IDENTIFICATION AND ANALYSIS OF SOMATIC MUTATIONS IN PAEDIATRIC HODGKIN LYMPHOMA

3. IDENTIFICATION AND ANALYSIS OF SOMATIC MUTATIONS IN PAEDIATRIC HODGKIN LYMPHOMA

3.1 Introduction

Hodgkin lymphoma (HL) is a malignancy derived from the germinal centre or post-germinal centre stages of B cell differentiation that is characterised by the presence of large, often multinucleated, malignant cells known as Hodgkin/Reed-Sternberg (HRS) cells. HRS cells often show chromosomal instability with frequent aneuploidy, including tetraploidy or near-tetraploidy (Falzetti et al., 1999, Watanabe et al., 2004). Although chromosomal instability is strongly implicated in the pathogenesis of HL, the mechanisms responsible have yet to be elucidated.

Unusually for a cancer, the malignant HRS cells of HL comprise only a small fraction (< 1%) of the total tumour mass and are surrounded by a prominent tumour microenvironment that promotes HRS cell survival and evasion of tumour-specific immune responses (Kuppers et al., 2005, Kuppers et al., 2002). The relative paucity of HRS cells has so far restricted global genetic analyses of this tumour to only two exome sequencing studies comprising a total of ten primary cases and seven HRS-derived cell lines (Reichel et al., 2015, Liu et al., 2014). For this reason our understanding of the contribution of recurrent somatic genetic alterations to the pathogenesis of HL has been gained mainly from the sequencing of individual genes (Cabannes et al., 1999, Jungnickel et al., 2000, Emmerich et al., 2003, Lake et al., 2009, Gunawardana et al., 2014, Joos et al., 2003, Weniger et al., 2006a, Schmitz et al., 2009). These studies have revealed the existence of recurrent somatic mutations that deregulate multiple cell signalling pathways contributing to HRS cell survival and proliferation. For example, regulatory and signal transducing proteins in the nuclear factor kappa B (NF-κB)

pathway are frequently mutated, resulting in constitutive NF-κB signalling (Jungnickel et al., 2000, Emmerich et al., 2003, Lake et al., 2009, Schmitz et al., 2009). Altered signalling through JAK/STAT proteins in HRS cells also frequently results from either gain-of-function mutations in JAK2 or inactivation of SOCS1 or PTPN1/PTPB1 (Gunawardana et al., 2014, Joos et al., 2003, Weniger et al., 2006a).

The Epstein-Barr virus (EBV) is present in HRS cells in around one-third of cases (Murray et al., 1992b, Pallesen et al., 1991a). EBV latent genes expressed in HRS cells include the EBV oncogene, latent membrane protein-1 (LMP1), which alone is capable of driving the constitutive activation of many of the signalling pathways known to be aberrantly expressed in HRS cells, including both NF-κB and JAK/STAT (Huen et al., 1995, Eliopoulos et al., 2003, Gires et al., 1999, Bolger et al., 2014). The observation that mutations in these cell signalling pathways are more frequent in EBV-negative, compared with EBV-positive, cases of HL suggests that EBV and cellular mutations provide mutually exclusive means to the same pathogenic end-point (Schmitz et al., 2009).

In this study I have used whole-exome sequencing following whole genome amplification (WGA) of HRS cells micro-dissected from 13 paediatric HL to reveal the genomic landscape of childhood HL. The data point to the mitotic spindle as a key pathogenic target in HL, the disruption of which can occur either by somatic mutation or EBV infection. The results not only provide a potential mechanistic explanation for the chromosome instability and aneuploidy frequently observed in HRS cells, but also offer potential new insights into the mechanisms underpinning resistance to spindle-damaging and topoisomerase poisons routinely used for the treatment of patients with HL.

3.2 Identification and analysis of somatic mutations

3.2.1 Microgram amounts of good quality WGA products were obtained from 150 microdissected tumour cells

Using the methods set out in section 2 of Materials and Methods, I amplified the genomes of 150 (2 replicates) HRS cells microdissected from tissue sections of each sample. The whole genome amplified products ranged between 200 base pairs (bp) and 1000 bp (1kbp) (Figure 19A) and the majority of these fragments ranged between 200 and 500 bp (Figure 19B). I obtained sufficient quantity (between 2.4 and 4.7 μg of DNA when measured on Qubit) and quality (between 1.8 and 2 for 260/280 ratio when measured on Nanodrop) of DNA for exome sequencing. Quality measurements for the whole genome amplified samples along with their matched blood DNA are detailed in the exome sequencing submission table (Appendix 1). WGA products were checked for coverage by PCR and Sanger sequencing using 8 housekeeping genes (PTEN, BRAF, PCM1, JAK2, ATM, PIK3CA, TNFAIP3, and BCR) (Appendix 2). Similar experiments on the microdissected HRS cells from FFPE samples yielded more fragmented (WGA product size lower than 200 bp) (Figure 19C). Size distribution of WGA products obtained from the DNA extracted FFPE samples has similar WGA product size distribution to those from microdissected cells from flash frozen samples, however these DNA extracted FFPE samples needed greater amount of input DNA (at least 1 ng) (Figure 19D).

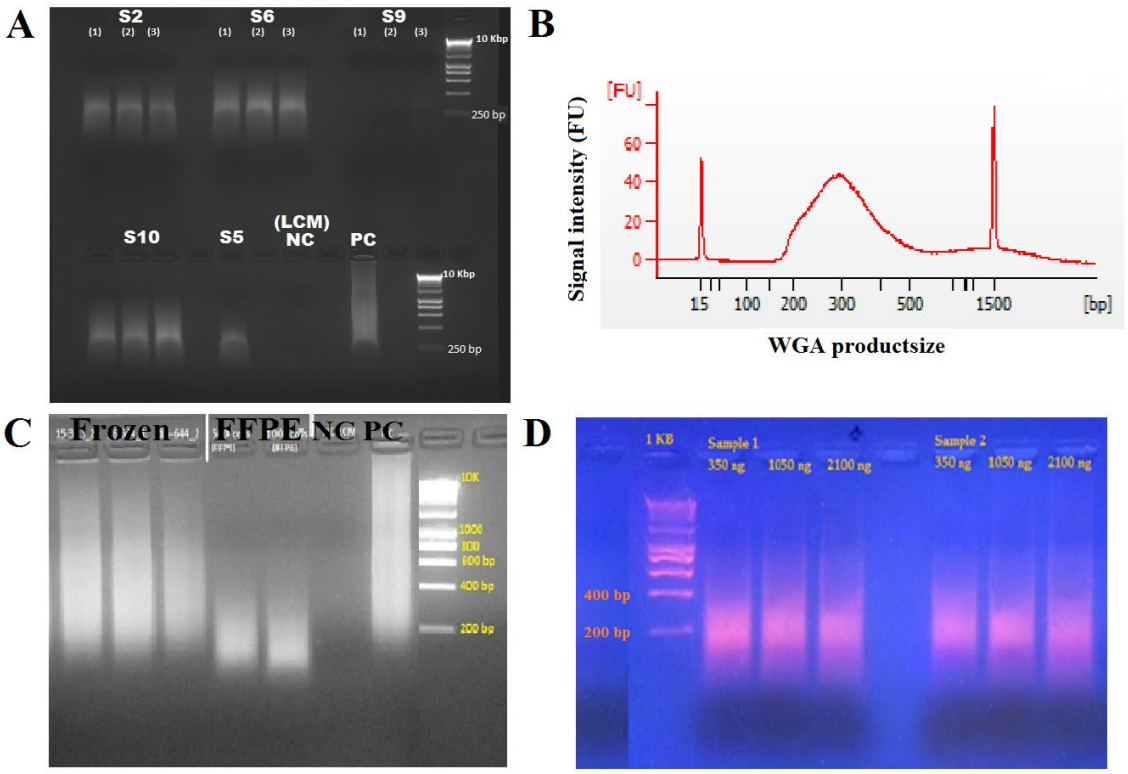

Figure 19. Size distribution for whole genome amplified products. A. WGA products from 150 HRS cells for patients S2, S6 and S9 in the top lane and S10, S5, NC-LCM (negative control from LCM procedure onwards, i.e. no DNA), PC (positive control-genomic DNA of L-591 cell line) in the bottom lane were separated on 2% agarose gel. 1Kbp ladder used to mark the size of WGA fragments. B. Bioanalyzer image for a typical WGA product showing the majority of the WGA fragments were between 200 bp and 500 bp in length. C. Size distribution of WGA products obtained from microdissected FFPE samples, showing smaller fragment size compared to the frozen samples used in the same experiment set up (first 3 wells). D. Size distribution of WGA products obtained from the DNA extracted FFPE samples. WGA product sizes are much lower in the microdissected FFPE samples (C) compared to those amplified from the DNA extracted FFPE samples (D).

3.2.2 Good quality exome sequence reads were obtained for all samples

Mean base quality (Phred) scores for reads across all positions (1-100) were higher than 33 in all samples, as shown in a typical FastQ file (Figure 20A). Base quality (Phred) scores for positions on either ends (~5 bases) typically ranged between 30 and 33 in all tumour exome sequences, whereas in matched normal exome sequences these scores were higher than 33 at these positions. All the reads had Phred quality score higher than 26 and the majority (> 90%) of the reads had a Phred quality score of 37 or higher (Figure 20B).

3.2.3 Coverage of sequencing reads of tumour DNA

Exome sequencing data were obtained from one blood DNA sample and two separate whole genome amplification products of tumour DNA for each of the 13 patients. The number of raw exome sequencing reads in tumour samples ranged approximately from 13.68 million to 78.02 million, whereas in normal blood samples this range was from 26.54 million to 173.37 million. The expected number of reads from NGS providers for non-microdissected and un-amplified samples is ~37 million; however, I received less than 30 million reads for more than half of the tumour samples. In tumour samples the number of de-duplicated reads mapped to hg19 ranged from 10.16 million to 39.09 million, whereas in matched normal samples this range was from 23.07 million to 144.25 million (Appendix 3). As expected, the number of mapped reads in the majority (12/13) of normal samples (mean= 60.54 million reads) was higher than those of the tumour samples (mean= 20.10 million reads). The percentage of reads mapped to the human genome after removing PCR duplicates ranged between 71.83 and 86.73 (mean = 79.41%) in tumour exomes, whereas a higher percentage of reads were mapped in matched normal (ranged between 81.56 and 90.74; mean = 85.29%).

3.2.4 Effective coverage of sequencing reads for the identification of somatic mutations in the exomes

In the approach undertaken here, the identification of a somatic mutation at a particular base position requires that the base is covered by at least one read from both WGA products and at least 8 reads from the matched blood DNA (defined variant calling criteria for tumour sequences see Materials and Methods section 2). The percentage of target regions in the exome covered by tumour sequences at 1x read depth in 13 samples ranged from 44.80 to 91.24 (mean = 70.65%) as shown in Table 7. The percentage of target regions in exome covered by both tumour replicate sequences at 1x read depth in 13 samples ranged from 44.18 to 71.16 (mean=56.42 %) as shown in Figure 21 . As expected, higher coverage was observed at x8 read depth in normal blood sequences (ranged from 82.25% to 96.50%) with a mean coverage of 91.54. Furthermore, the percentage of the target regions covered by tumour replicates and matched normal for each sample at the defined variant calling criteria ranged from 40.59 to 69.16 with a mean coverage of 54.49%.

This low coverage is due to the very high stringency criteria that I have applied to filter false positive mutations arising from the PCR amplification. I understand that I may miss some valuable information but my priority was to obtain good quality mutation data.

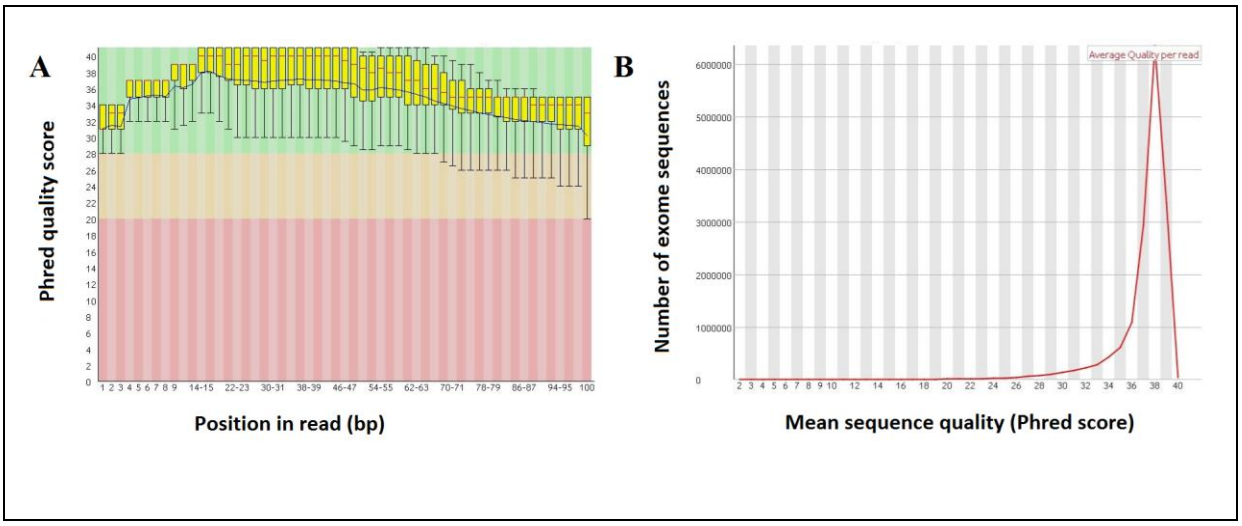

Figure 20. Base and sequence quality scores for a typical FastQ file in my samples. A. Inter-quartile range (25-75%) was shown in yellow box. Upper and lower whiskers represent the 10% and 90% points. Central red line was the median value and the blue line represents mean quality. Bases on either end of the 100 bp reads show a mean base quality Phred score below 33. B. Sequence quality score distribution for the same sample as shown in Figure 6 (see later) showing majority of reads with Phred quality score of 37 or higher.

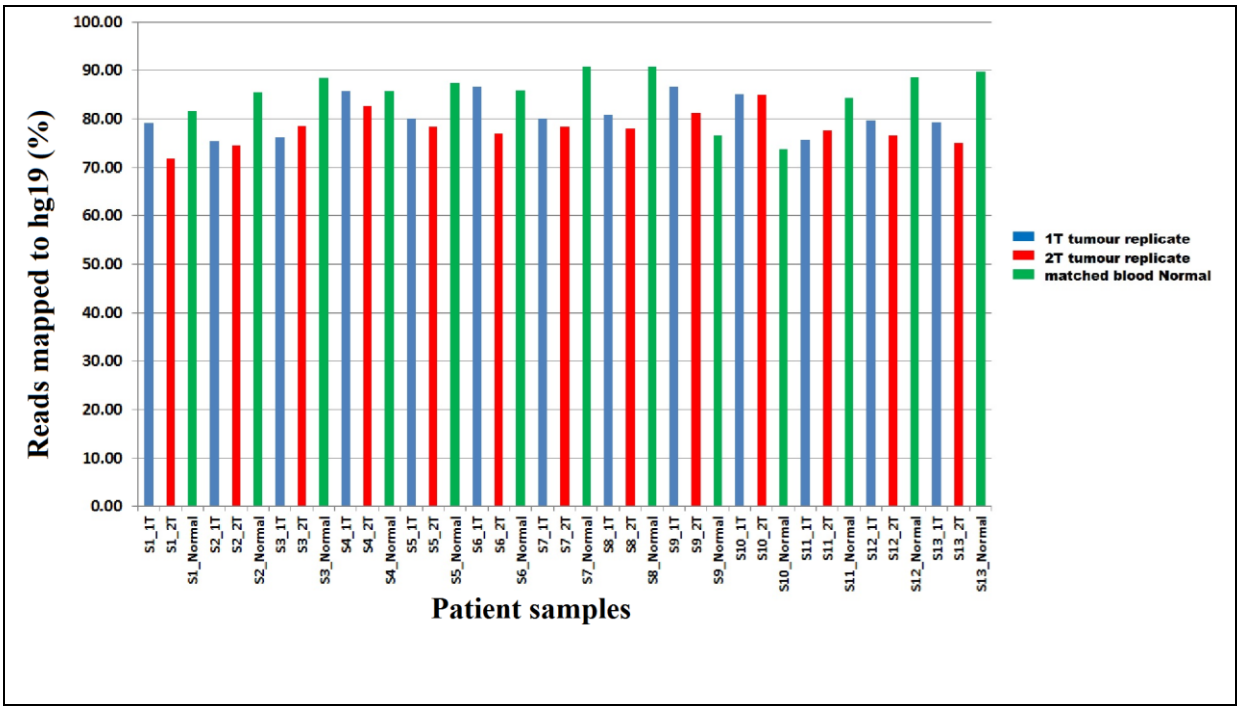

Figure 21. Coverage graph showing de-duplicated sequences mapped to hg19 in 13 HL samples. Over 70% of tumour exome reads mapped to hg19 (at read depth of x1) were retained after removing duplicate sequencing reads, whereas a higher percentage (>81%) of de-duplicated reads were mapped to hg19 in matched normal exomes in 13 HL samples (at read depth of x8).

Table 7. Coverage of 13 samples at variant calling criteria.

Sample	Coverage in N* at x8 read depth (%)	Coverage in 1T* at x 1read depth (%)	Coverage in 2T* at x1 read depth (%)	Coverage in 1T+2T* at x1 read depth (%)	Coverage in N+1T+2T* (%)
S1	90.40	73.11	58.75	50.58	49.02
S2	91.69	91.24	44.80	44.18	42.88
S3	82.25	58.94	63.90	44.69	40.59
S4	91.26	51.42	89.69	51.11	50.07
S5	90.21	81.85	71.57	66.69	63.87
S6	90.30	72.94	58.40	47.16	46.77
S7	96.50	64.94	59.99	47.35	47.08
S8	91.90	75.57	82.26	71.16	69.16
S9	90.07	74.14	84.09	67.86	62.62
S10	91.87	77.44	77.89	65.53	61.37
S11	95.31	65.98	67.03	53.15	52.63
S12	93.52	71.16	71.64	60.21	59.28
S13	94.71	76.33	71.83	63.71	63.03

*N = matched normal samples; 1T = first tumour replicate; 2T = second tumour replicate; 1T+2T = covered in both tumour replicates; N+1T+2T = covered in both tumour replicates (1T and 2T) at read depth of x1 and matched normal (N) at read depth of x8.

3.3 Identification of somatic variants in 13 Hodgkin lymphoma exomes

As described in the Materials and Methods (section 2.6.9), four variant callers (Strelka v1.0.14, Mutect v1.0, VarScan2.0, and MuTect v2.0) were used for calling mutations in my cohort. I briefly summarise below the analyses I performed with these variant callers.

3.3.1 Variant calling by Strelka v1.0.14 and validation of selected mutations

Strelka v1.0.14 identified a total of 2356 somatic point mutations (sSNVs) in 1647 genes and 305 somatic indels (insertions/deletions) in 236 genes. Of these, 1331 were protein-altering SNVs and 182 were somatic protein-altering indels. Seven mutations in two genes called by Strelka (Table 8) with a tumour allele frequency (TAF) of > 30% were selected for validation by Sanger sequencing. 2/6 mutations (33.33%) with good quality Sanger sequences were validated. An example of a mutation detected by exome sequencing but not validated by Sanger sequencing is shown in Figure 22.

3.3.2 Variant calling by Mutect v1.0 and validation of selected mutations

Mutect v1.0 called a total of 654 sSNVs in 858 genes including 454 protein-altering mutations in 433 genes. A total of 18 sSNVs identified by Mutect 1.0 and with a tumour allele frequency (TAF) of > 30% were selected for validation by Sanger sequencing. Good quality Sanger sequences were obtained for 14/18 sSNVs. 13/14 of the sequence variants were detected by Sanger sequencing (Table 9). Figure 23 shows an example of a mutation that was detected by exome sequencing and validated by Sanger sequencing in at least one of two WGA replicates.

3.3.3 Variant calling by VarScan 2.0 and validation of selected mutations

Mutect v1.0 software is designed to call only somatic point mutations. For this reason VarScan2.0[73] was used to identify indels. 681 somatic indels in 856 genes, including 521

protein-altering indels in 470 genes were called by VarScan2.0. VarScan2 also called 3294 sSNVs in 2075 genes. Eight deletions called by VarScan2.0 and with a TAF >30% were selected for validation by Sanger sequencing. I obtained good quality Sanger sequences for only 4/8 deletions and these four indels showed the same base changes in Sanger sequences to those from exome data (Table 10).

3.3.4 Variant calling by Mutect v2.0

Mutect v2.0 (which can call both SNVs and indels) was available at the later stages of analysis, and it identified only 76 sSNVs and 11 somatic indels in 13 HL exomes. This was attributed to the more stringent criteria that Mutect v2.0 employs for variant calling and filtering. Unfortunately, further documentation was not available on this software at that time and so I was unable to deduce the factors responsible for the low mutation call rate in Mutect v2.0. For this reason I did not use Mutect v2.0 further.

Figure 24A shows the overlap of somatic point mutations detected by these variant callers and Figure 24B the overlap of insertions/deletions. 167 somatic point mutations were shared among Mutect v1.0, Strelka v1.0.14, and VarScan2. Only 2 indels were shared between Mutect v2.0, Strelka v1.0.14, and VarScan2.0. The large variation in the number of calls and highly variable agreement from the four variant callers used in this analysis in my sample cohort could be due to differences in the sensitivity of the callers to sequencing depth as reported in previous studies (Krøigård et al., 2016).

Mutect v1.0 and VarScan2.0 were used for all subsequent analyses in this study as these two callers had high validation rate by Sanger sequencing method.

Table 8. Sanger sequencing validation of sSNVs called by Strelka (*chromosome_base position_refrence allele_variant allele; **AA= amino acid).

Primer	Mutations*	AA** change	Samples	Sanger sequencing
ACVR1_PCR1	chr2_158594977_A_G	I457T	S3	NOT VALIDATED
	chr2_158594977_A_G	I457T	S4	NOT VALIDATED
	chr2_158594968_C_G	R460T	S4	NOT VALIDATED
	chr2_158594968_C_G	R460T	S8	NOT VALIDATED
	chr2_158594968_C_T	R460K	S11	VALIDATED
ACVR1_PCR2	chr2_158595062_G_T	P429T	S5	NOT VALIDATED
MSH2_PCR1	chr2_47656991_G_A	R396K	S7	NOT VALIDATED
MSH2_PCR1	chr2_47656991_G_A	R396K	S12	NOT VALIDATED
MSH2_PCR2	chr2_47707858_A_G	I828V	S8	NOT VALIDATED
MSH2_PCR3	chr2_47705572_C_T	A791V	S11	VALIDATED

Table 9. Sanger sequencing data for 14 sSNVs detected by Mutect v1.0, which confirmed 13 of the point mutations identified by exome sequencing. (*Stop-gained mutation)

Gene	Mutation	AA change	Sample	Sanger validation
AKT2	chr19_40743884_C_T	D275N	S3	VALIDATED
FERMT1	chr20_6057912_C_T	G648S	S3	VALIDATED
HLA-DQA1	chr6_32609192_A_G	E63G	S3	VALIDATED
NR1H4	chr12_100926334_T_C	C192R	S3	VALIDATED
PIWIL4	chr11_94335040_G_A	W487*	S3	VALIDATED
S100B	chr21_48022204_G_T	S42Y	S3	VALIDATED
ABCA1	chr9_107576477_A_T	F1275I	S4	VALIDATED
BCR	chr22_23596126_C_T	R474W	S5	VALIDATED
DPYD	chr1_97547923_C_T	G957D	S5	VALIDATED
F5	chr1_169511277_C_A	P1274L	S5	VALIDATED
DPYD	chr1_97915760_G_T	S587Y	S7	VALIDATED
LIFR	chr5_38496642_G_T	S576*	S11	VALIDATED
PTPN14	chr1_214551424_C_T	A856T	S11	VALIDATED
MLLT4	chr6_168348563_C_A	A1187E	S12	NOT VALIDATED

Table 10. Sanger sequencing data was obtained for four somatic deletions identified by Varscan 2.0 and all of which confirmed the deletions identified by exome sequencing (*Stop-gained mutation).

Gene	Mutation	AA change	Samples	Sanger validation
WBP11	chr12_14949846_TC_T	R94*	S3	VALIDATED
CTSA	chr20_44521061_TCC_T	P147*	S7	VALIDATED
ST8SIA6	chr10_17363320_CT_C	K251*	S11	VALIDATED
SMG1	chr16_18847426_TG_T	G2629*	S12	VALIDATED

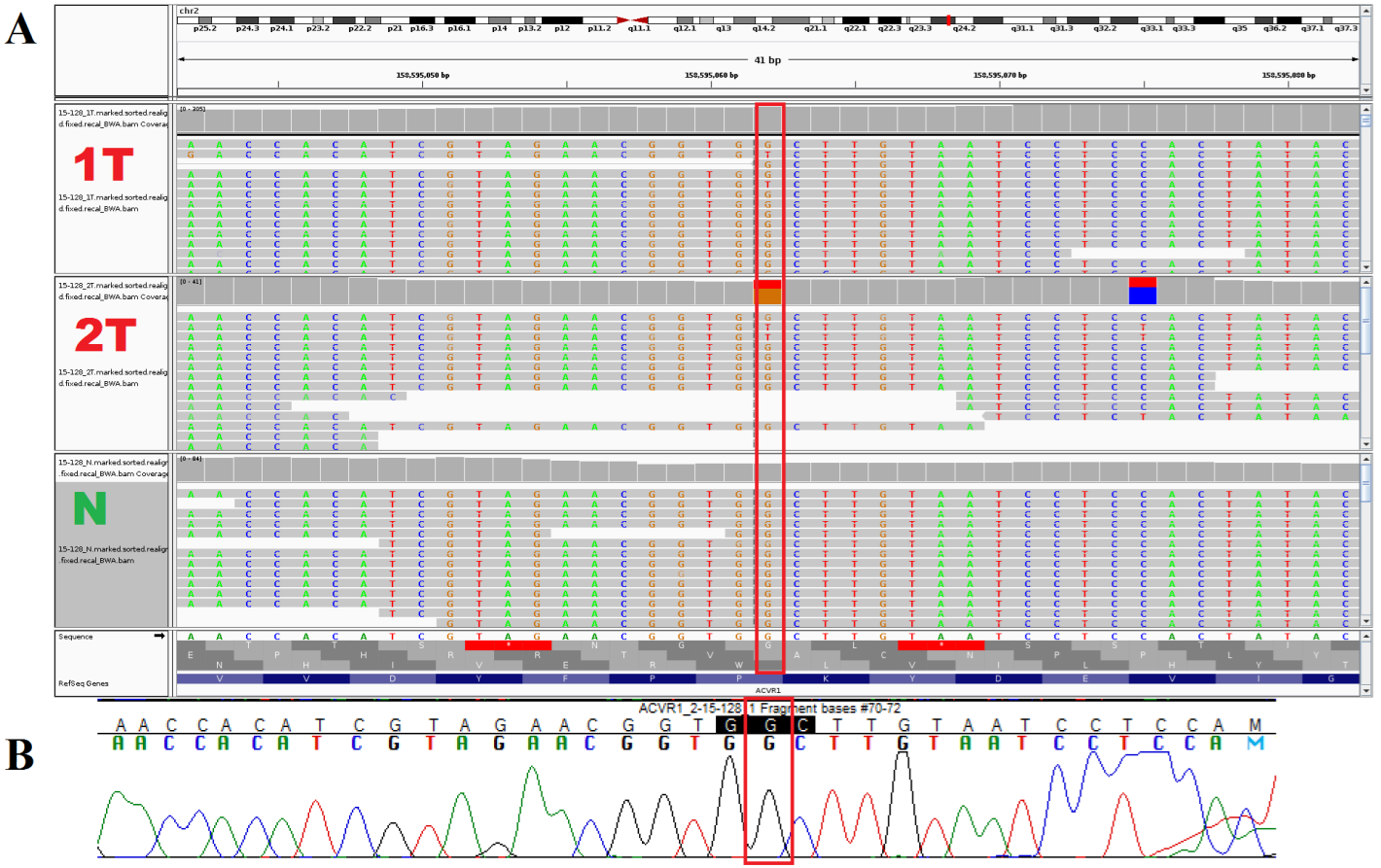

Figure 22. Example of a variant called by Strelka which was not detected by Sanger sequencing. A. G->T transversion (highlighted in rectangle) in ACVR1 gene in sample S5 at 158595062 in Chr2 was absent in (B) sequences obtained by Sanger sequencing method for ACVR1 gene.

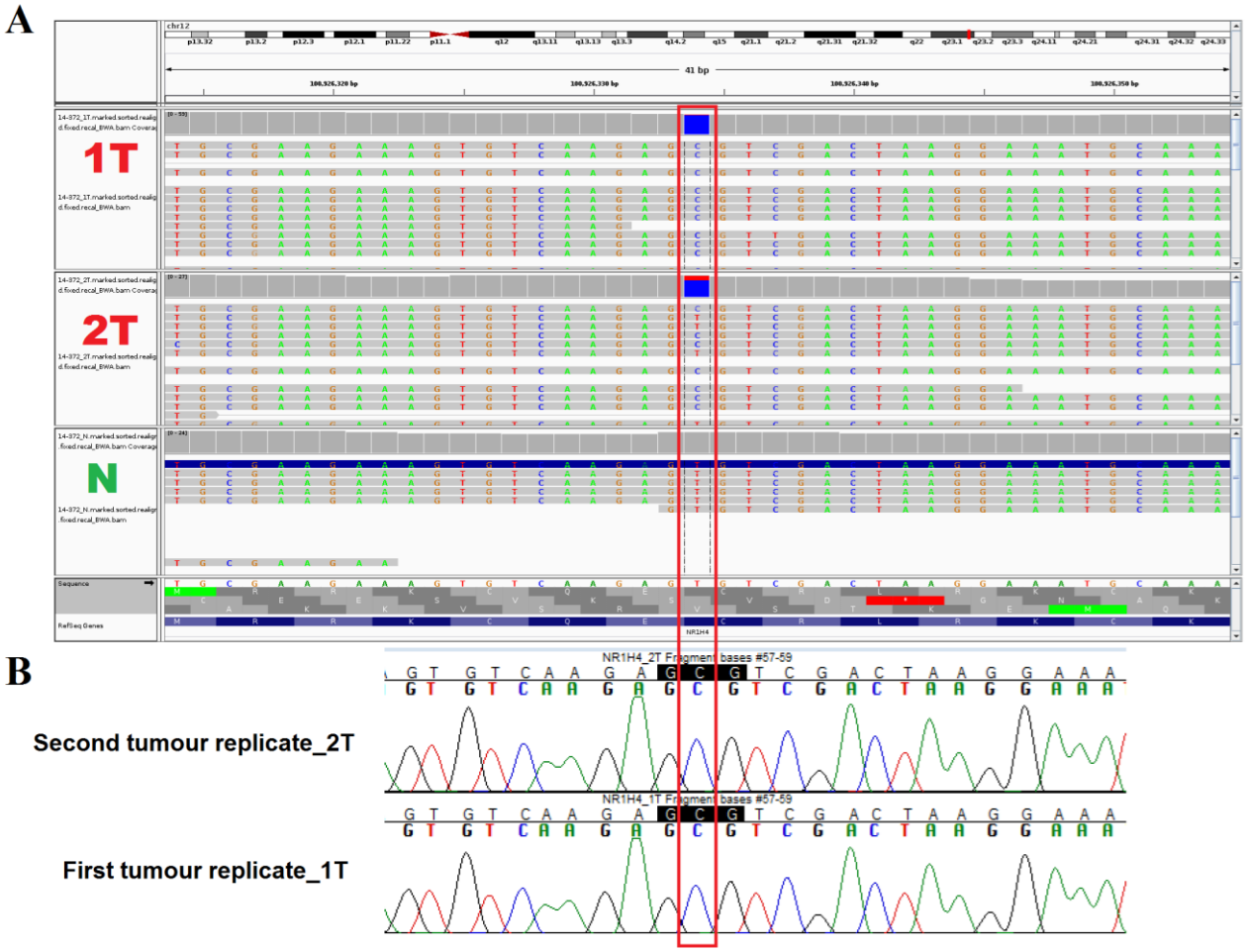

Figure 23. Detection of a mutation in NR1H4 detected by Mutect 1.0 by Sanger sequencing. A. Integrated Genomic Viewer image showing T->C transition in NR1H4 at position 100926334 in Chromosome 12 and also detected in both replicates of sample S3 bySanger sequencing (B) .

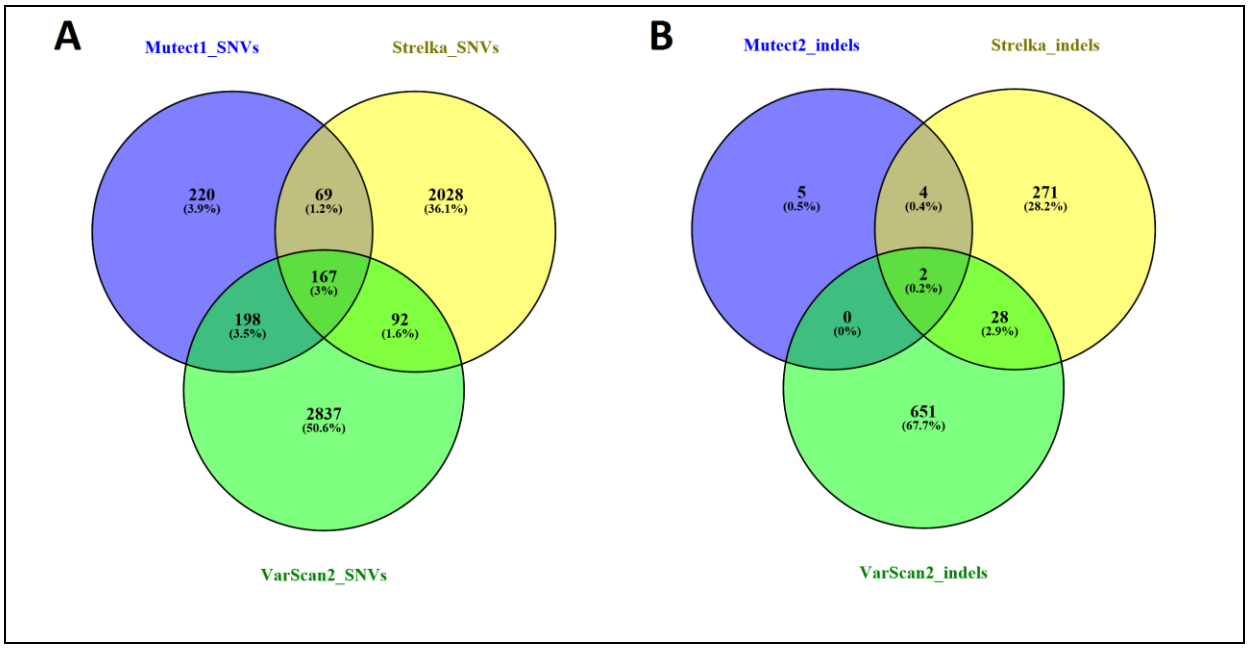

Figure 24. Variant caller comparison. A. Overlap of somatic point mutations among Mutect v1.0, Strelka v1.0.14, and VarScan2. B. Overlap of somatic insertions/deletions among Mutect v2.0, Strelka v1.0.14, and VarScan2.

3.4 Whole exome sequencing reveals recurrent somatic mutations in my cohort of 13 paediatric Hodgkin lymphoma samples

3.4.1 Types of mutations observed

A total number of 1335 somatic mutations were observed in 13 paediatric HL samples of which 654 were point mutations and 681 were indels. 16 major categories of mutations that were observed in my sample cohort were: 3' Flank, 3' UTR, 5' Flank, 5' UTR, Intron, Intergenic, RNA, Targeted Region, Silent, In Frame Deletions, In Frame Insertions, Missense, Splice site, Frame Shift Deletions, Frame Shift Insertions, and Nonsense mutations (Table 11). Most of the mutations were either missense or frameshift indels (Table 11). Annotation of these mutations was performed using snpEff and GRCh37.75 database which revealed that the 1335 mutations are within or near 1622 genes. A total of 975 protein-altering mutations were identified in 857 genes. The number of protein-altering mutations was significantly higher than the non-protein-altering mutations (t-Test; $\boldsymbol{p}$= 0.0016). The number of mutations per sample ranged between 36 and 213 (range 30-156 for protein-altering mutations). An average of 102 mutations (median = 104) were observed per sample, whereas the mean and median number of protein-altering mutations were 75 and 75, respectively. The mean mutation rate per million bases (Mb) was 5.18/Mb (range 1.55-9.11). The mean mutation rate per Mb for protein-altering mutations was 3.77 (1.29-6.67).

Table 11. Types of mutations observed in my cohort of 13 paediatric Hodgkin lymphoma samples.

Type of Mutation	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	
	EBV (+)	EBV (+)	EBV (-)	EBV (-)	EBV (-)	EBV (+)	EBV (-)	EBV (-)	EBV (+)	EBV (+)	EBV (-)	EBV (+)	EBV (-)	**Total**
Frame Shift Del	14	31	28	33	55	25	34	25	23	21	36	63	85	**473**
Frame Shift Ins		2	1	2	20		1	4		1	2	1	5	**39**
Missense Mutation	23	39	49	10	48	14	35	8	6	14	39	49	55	**389**
Nonsense Mutation	1	4	10	4	10	1	5	2	1		5	8	6	**57**
Splice Site	2	1	2		2						4	1	5	**17**
3'Flank	2	1			3		3	1			2	2	1	**15**
3'UTR		4	6	4	5		1	4			3	1	4	**32**
5'Flank	1	2	2	2	8	1	6	1	1		2	4	1	**31**
5'UTR			1	1										**2**
IGR			1										1	**2**
In Frame Del	6	4	8	4	9	1	2	3	1	2	4	6	6	**56**
In Frame Ins			1						1				2	**4**
Intron	3	5	8	9	6		2	2	1	1	4	7	6	**54**
RNA		1		1	1		1		1	1	2	1	2	**11**
Silent	10	16	14	6	17	9	14	2		5	8	15	34	**150**
Targeted Region	1				1				1					**3**
Total	**63**	**110**	**131**	**76**	**185**	**51**	**104**	**52**	**36**	**45**	**111**	**158**	**213**	**1335**
Mutation rate per Mb	3.46	6.91	8.70	4.09	7.81	2.94	5.95	2.03	1.55	1.98	5.68	7.18	9.11	

3.4.2 Frequently mutated genes in paediatric Hodgkin lymphoma samples

Two hundred genes with somatic mutations in two or more paediatric Hodgkin lymphoma samples are shown in Appendix 4, of which KLF6 and SMG1 were mutated in six and five samples, respectively. Analysis of genes with protein-altering mutations revealed considerable heterogeneity between samples. Fifty genes with protein-altering mutations in two or more paediatric Hodgkin lymphoma samples are shown in Figure 25 (complete list in Appendix 5). SMG1 was the gene most frequently harbouring protein-altering mutations (5/13 samples).

Here onwards I focussed only on protein-altering mutations.

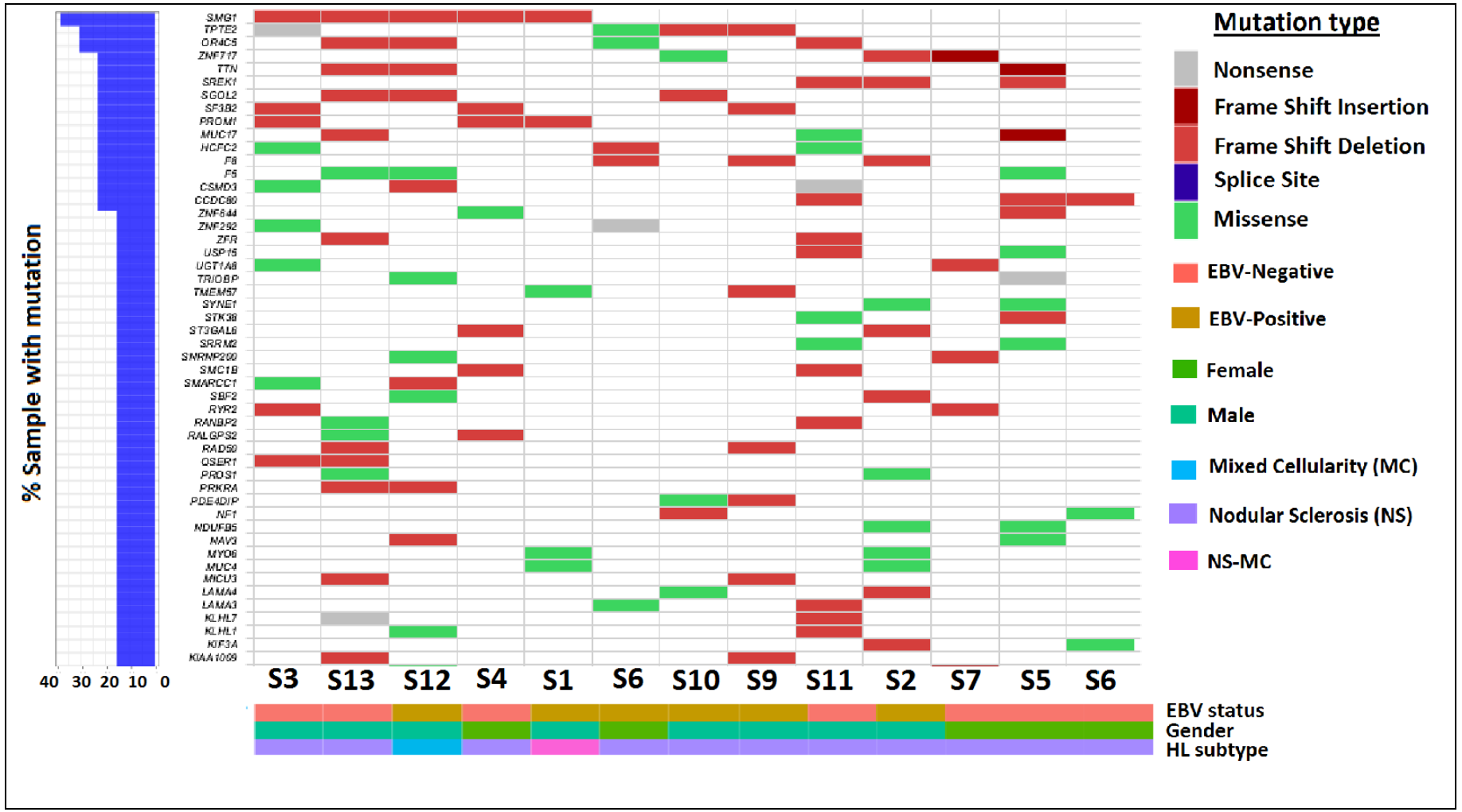

Figure 25. 50 genes with the most frequent protein-altering mutations in 13 paediatric HL samples are shown here.

3.4.3 Identification of genes previously reported in other Hodgkin lymphoma studies

Having described the genomic landscape of paediatric HL, I next studied the extent to which genes mutated in paediatric HL (from here onwards these are referred to as 'my series') had also been reported to be mutated in other HL studies. I compiled a list of genes reported to be mutated in primary HRS cells and HL cell lines from Reichel et al., and Liu et al., studies (Reichel et al., 2015, Liu et al., 2014) (here onwards these are referred to as 'Reichel series' and 'Liu series'). I found that genes mutated in my series were significantly enriched in genes previously reported to be mutated in HRS cells flow sorted from ten cases of primary HL in Reichel series (odds ratio=2.23; $\boldsymbol{p}< 0.0001$) and among genes reported to be mutated in HL cell lines in Liu series (odds ratio=1.92; $\boldsymbol{p}< 0.0001$).

42 genes with protein-altering mutations ere common in all three HL studies as shown in Figure 26 (gene list in Table 12). To understand the biological importance of these 42 genes, gene ontology analysis was performed using DAVID (version 6.8) functional annotation tool (Huang da et al., 2009) with GOTERM BP-FAT as described in Materials and Methods (Section 2.2.7). This analysis revealed a significant enrichment of biological processes which included "negative regulation of NF-kappaB transcription factor activity" (GO:0032088; $\boldsymbol{p}$= 0.00037) , "cellular component assembly" (GO:0022607; $\boldsymbol{p}$= 0.00089), "regulation of tumour necrosis factor-mediated signalling pathway" (GO:0010803; $\boldsymbol{p}$= 0.0045) "sister chromatid segregation" (GO:0000819; $\boldsymbol{p}$= 0.009), "nuclear chromosome segregation" (GO:0098813; $\boldsymbol{p}$= 0.018), "mitotic cell cycle process" (GO:1903047; $\boldsymbol{p}$= 0.029), and "centrosome localization" (GO:0051642; $\boldsymbol{p}$= 0.046) (Appendix 6).

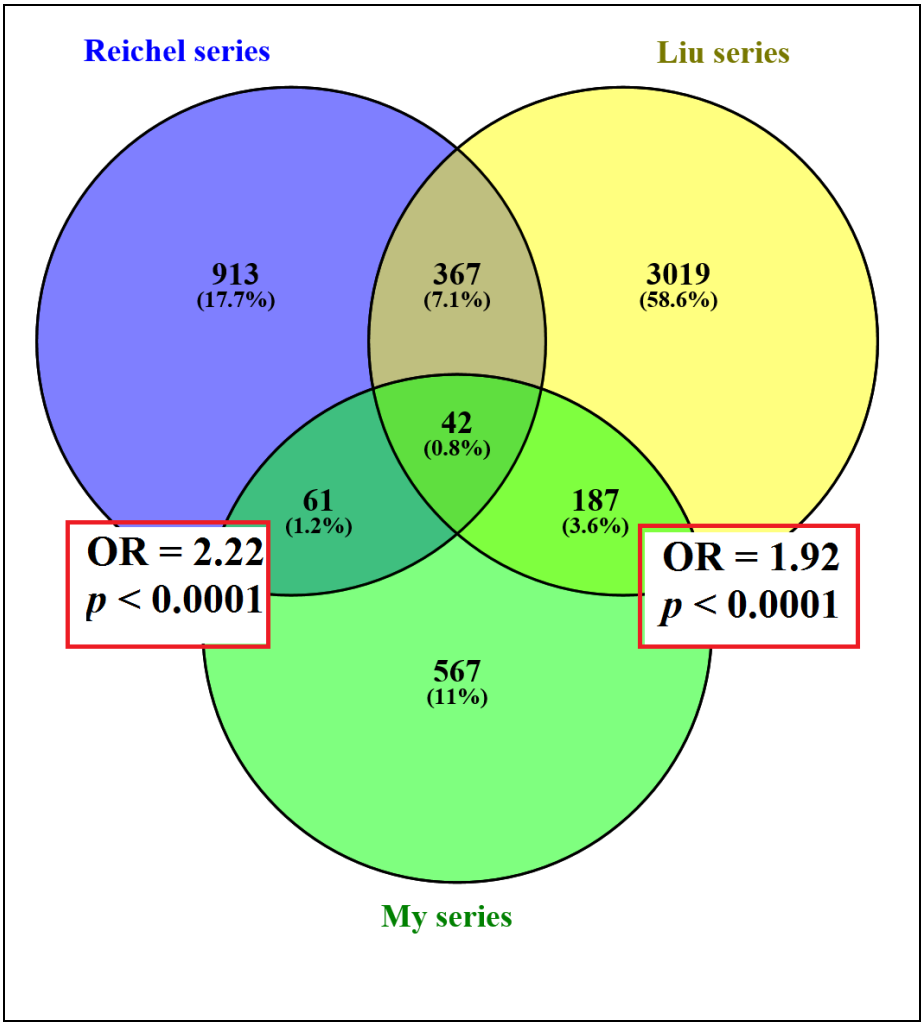

Figure 26. Overlap between all 3 HL studies for protein-altering mutations. Significant overlap was seen between the three HL studies.

103 genes with protein-altering mutations were common between my series and 'Reichel series' (Table 13). Significant enrichment of "cellular component assembly" (GO:0022607 ; ***p***= 0.0004), "negative regulation of NF-kappaB transcription factor activity" (GO:0032088; ***p***= 0.0057) , "microtubule cytoskeleton organization" (GO:0000226; ***p***= 0.006), and "chromosome segregation" (GO:0007059; ***p***= 0.028) processes was observed for these 103 genes (Appendix 7).

The 25 most frequently mutated genes in Reichel and Liu series are listed in Table 14, together with their frequency in my series. The most frequently mutated genes that are common in all 3 HL studies are shown in Table 15, together with their frequency in my series. It can be seen that several genes frequently mutated in Reichel and Liu series were either not reported as mutated or mutated at a lower frequency in my series. I measured the sequence coverage in my series for TNFAIP3, B2M, and SOCS1 genes that were reported to be most frequently mutated genes in COSMIC, Reichel and Liu series. Mean coverage for these genes in my series was 54.85%, 79.68%, and 0%, respectively. Besides TNFAIP3 gene, which was deleted in one sample in my series, the other two genes were not mutated in my series. This could either be due to low coverage (as in the case of SOCS1gene) or tumour heterogeneity (for B2M gene).

Gene ontology analysis of genes with protein-altering mutations in all three HL studies revealed an overlap in 12 biological processes significantly enriched in the genes with protein-altering mutations in these studies (Table 16). These include 'cellular component assembly' and 'circulatory system development' processes.

Table 12. 42 genes that were mutated in all 3 HL studies.

ACACA	CSMD3	HNRNPC	MACF1	NR1H4	TNFAIP3
ALMS1	CUBN	ITPR1	MUC16	PKHD1	TTN
ANKRD36	CYLD	ITPR2	MUC17	RANBP2	UNC13C
ANXA2	FAM228A	LAMA3	MUC4	SON	USH2A
APOB	FAT3	LMLN	MYCBP2	SPAG17	VAV2
CCDC175	FNBP1	LRRIQ1	MYO18B	SRRM2	WDR87
CDH23	GPR98	LST3	NIN	STAG1	XPO1

Table 13. 103 genes mutated in my series and Reichel series.

ACACA	CTC-432M15.3	IGHV1-69	NF1	SON
ADAMTS12	CTD-2370N5.3	IGHV2-70	NHS	SPAG17
ALK	CTTNBP2	INVS	NIN	SRRM2
ALMS1	CUBN	ITPR1	NOL4	STAG1
ANKHD1	CYLD	ITPR2	NPIPB5	SYCP2
ANKHD1-EIF4EBP3	DDX20	KIAA1462	NR1H4	TNFAIP3
ANKRD36	ECT2	KIAA1715	NUP50	TRAPPC8
ANXA2	EGFLAM	LAMA3	OR52E2	TRDN
APOB	EPG5	LMLN	PDE4DIP	TTN
ARHGAP25	FAM228A	LRRIQ1	PKHD1	UBR5
ATL1	FAT3	LST3	PLS3	UNC13C
C1orf87	FNBP1	MACF1	PRRC2C	USH2A
CCDC175	FNDC7	MKS1	PTK2	VAV2
CCDC73	FRG1B	MRRF	PTPN21	WDR87
CD109	GLUD2	MUC16	PTPRD	XPO1
CDH23	GPR98	MUC17	RANBP2	XRCC4
CEACAM8	HBS1L	MUC19	RAPGEF6	ZBTB21
CHD1	HECTD1	MUC4	RNF145	ZNF292
CHD1L	HEPH	MYCBP2	SLC14A2	ZNF430
CHD4	HNRNPC	MYO18B	SLC25A13	
CSMD3	HNRNPDL	MYO5B	SMARCAL1	

Table 14. Top 25 most frequently mutated genes previously reported in HL exomes.

Gene	Number of samples mutated in combined Reichel and Liu series (n=17)	Number of samples mutated in my series (n=13)
TTN	10	3
MUC4	9	2
TNFAIP3	9	1
HIST1H1E	8	0
B2M	7	0
CSF2RB	7	0
DNAH14	7	0
IGHV2-70	7	2
IGLL5	7	0
IGLV3-1	7	0
ITPKB	7	0
PCLO	7	0
SOCS1	7	0
FAT4	6	0
GPR112	6	0
GPR98	6	1
IGHE	6	0
LRRN3	6	0
MUC16	6	1
ABCA13	5	0
ACTB	5	0
C2orf16	5	0
DMD	5	0
DNAH11	5	0
FAT1	5	0

Table 15. Most frequently mutated genes that were common in all 3 HL studies.

Most frequently mutated genes common in all 3 HL studies	Number of samples mutated in my series (n=13)	Number of samples mutated in combined Reichel and Liu series (n=17)
CSMD3	3	3
F8	3	2
MUC17	3	4
SREK1	3	2
TTN	3	10
ZNF717	3	3
C20orf194	2	2
ITPR1	2	2
ITPR2	2	4
LAMA3	2	2
LAMA4	2	2
MUC4	2	9
NAV3	2	2
RANBP2	2	4
SRRM2	2	3
SYNE1	2	2

Table 16. Biological processes common in all 3 HL studies.

Gene ontology term
GO:0072359~circulatory system development
GO:0072358~cardiovascular system development
GO:0060249~anatomical structure homeostasis
GO:0000902~cell morphogenesis
GO:0048646~anatomical structure formation involved in morphogenesis
GO:0032990~cell part morphogenesis
GO:0048858~cell projection morphogenesis
GO:0022607~cellular component assembly
GO:0048871~multicellular organismal homeostasis
GO:0042592~homeostatic process
GO:0007507~heart development
GO:0051128~regulation of cellular component organization

3.4.4 Identification of genes previously reported in other B cell lymphomas

I next compared genes with protein-altering mutations in my series with those previously reported to be mutated in other B-cell lymphoma data sets. For GC-derived non-Hodgkin lymphomas, I had available a number of datasets for each tumour type. For this analysis I compiled a list of genes mutated in one or more samples of a specified tumour type. Using these gene lists I observed a significant enrichment of genes mutated in my series among genes reported to be mutated in DLBCL (odds ratio=2.39; $\boldsymbol{p}< 0.0001$) (Park et al., 2016, Morin et al., 2016, Zhang et al., 2013, Lohr et al., 2012), BL (odds ratio=1.87; $\boldsymbol{p}< 0.0001$) (Love et al., 2012, Abate et al., 2015, Richter et al., 2012), and FL (odds ratio=2.40; $\boldsymbol{p}< 0.0001$) (Pasqualucci et al., 2014, Krysiak et al., 2017, Ozawa et al., 2016).

Furthermore, I compiled a list of protein-altering mutations in HL (from now on called as HL series), which included genes mutated in all 3 HL studies and also two genes (HRAS and NRAS) reported to be frequently mutated in HL as seen in COSMIC data base but not present in all the three HL series (Forbes et al., 2015) to define a set of genes with protein-altering mutations that were present only in HL (including mutations reported in Reichel et al and Liu et al) and not in any other B-cell lymphomas (including those tumour types described above but also including mantle cell lymphoma (MCL), marginal zone lymphoma (MZL), and chronic lymphocyte leukaemia (CLL). 1310 HL-specific genes were mutated in at least 3/30 samples from the three 3 HL studies (Table 17). 141 HL-specific genes present only in my series are shown in Table 18. This group represent a set of putative paediatric specific mutations[1]. The 11 most frequently mutated HL-specific genes unique to my series are shown in Table 19. Gene ontology analysis of 1310 HL-specific genes is presented in Table 20.

[1] 9/10 patients reported in the Reichel series above 21 years of age.

Table 17. HL-specific genes mutated in at least 10% of samples in all three HL studies.

Gene	Number of samples mutated in 3 HL studies (n=30)
IGHV2-70	9
IGHE	6
OR4C5	5
TRMT2A	5
PLXND1	4
ALOX12	3
APPL1	3
ASIC2	3
BIVM-ERCC5	3
COL26A1	3
CYB5R2	3
CYLD	3
DYNC1LI1	3
EPDR1	3
FAM169B	3
IGHJ5	3
IGHV1-69	3
IGKV4-1	3
KMT2B	3
LONRF2	3
LST3	3
SMCR8	3
TDRD15	3
TRIM66	3
ZNF681	3
ZNF835	3

Table 18. 141 HL-specific genes unique to my series.

AC012215.1	CWC27	IGKV1-6	OR1N1	RRH	TULP3
AC112715.2	DLEU2L	IGLC3	OR4F4	S100B	UGT1A5
ADAT2	DMTF1	IKZF5	OR4Q2	SC5D	UGT1A8
ALDH3A2	DUS2	IL17RB	OR52R1	SCOC	USH1C
ARL14EP	DYX1C1	IMPACT	PANK1	SEC23B	USP46
ATPAF1	EDRF1	INSIG2	PAQR5	SLC25A14	UTS2B
ATXN10	EXD1	IPP	PBX4	SLC25A3	VAMP4
B3GALNT1	FAM210A	KIAA0125	PDCD2L	SLC25A33	VCAM1
BSPH1	FAM98A	KLHL7	PDCL2	SLC5A6	WBP11
C14orf23	FANCL	KLRD1	PGAP1	SNAP29	WNT2B
C20orf24	FBXO45	LAMTOR2	PGBD1	SPATS1	ZBED6
C2orf66	FNDC3B	LRRC9	PGM1	SRXN1	ZCRB1
C7orf25	FSIP1	MAN1A2	PIGK	STK17B	ZFR
C8orf37	FUNDC1	MESDC2	PITPNC1	SUN3	ZNF12
CARF	GAPT	MFAP5	PLSCR1	TAB2	ZNF143
CASP2	GCNT2	MIER3	PPP3CB	TARDBP	ZNF567
CATSPERD	GK5	MLEC	PRKRA	TCTN2	ZNF625-ZNF20
CDKL4	GLUL	MLLT4	PROSER1	TFDP1	ZNF682
CKLF-CMTM1	GSS	MRPS18B	PUS7	TIGD7	ZNF69
CLCN3	HIATL2	MSANTD2	RGS6	TMEM136	ZNF816
CMPK1	HLF	MUC7	RNF138	TMEM50A	ZSCAN32
CMTR2	IFIT3	NPIPA5	RNF7	TMEM97	
CSE1L	IGHV1-18	NUGGC	RP11-53I6.2	TTC12	
CTCFL	IGHV2OR16-5	OCIAD2	RPL9	TUBA1C	

Table 19. Eleven most frequently mutated HL-specific genes unique to my series.

Most frequently mutated HL-specific genes unique to my series	**Number of samples mutated**
BSPH1	2
CMPK1	2
DLEU2L	2
DMTF1	2
EDRF1	2
IPP	2
KLHL7	2
PRKRA	2
UGT1A5	2
UGT1A8	2
ZFR	2

Table 20. Biological process significantly enriched in 1310 genes unique to Hodgkin lymphoma.

Gene ontology term	***p*-value**
Complement activation, classical pathway	3.55E-05
Fc-gamma receptor signalling pathway involved in phagocytosis	8.99E-05
Immune response-regulating cell surface receptor signalling pathway involved in phagocytosis	8.99E-05
Fc receptor mediated stimulatory signalling pathway	1.24E-04
Fc-gamma receptor signalling pathway	1.37E-04
Humoral immune response mediated by circulating immunoglobulin	1.73E-04
Complement activation	2.96E-04
Protein activation cascade	5.62E-04
Immunoglobulin mediated immune response	0.001576
B cell mediated immunity	0.001813
Cellular response to fatty acid	0.004092
Receptor-mediated endocytosis	0.004728
Fc receptor signalling pathway	0.005479
Fc-epsilon receptor signalling pathway	0.005982
Lymphocyte mediated immunity	0.01089
Leukocyte mediated immunity	0.011056
Adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	0.011355
Response to fatty acid	0.012326
Humoral immune response	0.024109
Phagocytosis	0.024586
Cellular response to acid chemical	0.02544
Immune response-activating cell surface receptor signalling pathway	0.026786
Sensory perception of chemical stimulus	0.027145
Positive regulation of protein secretion	0.033817
Glycolipid biosynthetic process	0.034139
Positive regulation of insulin secretion	0.039717
Membrane lipid metabolic process	0.040114
Cell wall macromolecule metabolic process	0.040817
Immune response-regulating cell surface receptor signalling pathway	0.04429
Phagocytosis, recognition	0.044989
Positive regulation of cytokine secretion	0.046605
Oligosaccharide biosynthetic process	0.048942
Cell wall organization or biogenesis	0.048942
Positive regulation of cellular response to insulin stimulus	0.04942

3.4.5 Expression of mutated genes in primary HRS cells

I next investigated whether the expression of mutated genes was more or less likely to be up - or down-regulated in Hodgkin lymphoma using gene expression datasets GSE12453 (Brune et al., 2008) and GSE39133 (Steidl et al., 2012). 377 genes mutated in my series were down-regulated (Appendix 8) and 141 mutated in my series were up-regulated (Appendices 9) in Hodgkin lymphoma compared with GC B cells in at least one dataset. The genes with somatic mutations in paediatric HL were significantly enriched (odds ratio = 2.34; $\boldsymbol{p}< 0.0001$) in genes down-regulated in primary HL and significantly depleted (odds ratio = 0.76; $\boldsymbol{p}= 0.019$) in genes up-regulated in HL.

To understand the biological importance of protein-altering genes that were also downregulated in either Brune or Steidl arrays, gene ontology analysis was performed. This analysis revealed a significant enrichment of "mitotic cell cycle process" (GO:1903047; $\boldsymbol{p}< 0.0001$), "chromosome organization" (GO:0051276; $\boldsymbol{p}<0.0001$), "chromosome segregation" (GO:0007059; $\boldsymbol{p}< 0.0001$), and "mitotic cell cycle checkpoint" (GO:0007093; $\boldsymbol{p}< 0.0001$) (Appendix 10 showing 125 most significantly enriched biological processes). Significant enrichment of "blood vessel development" (GO:0001568; $\boldsymbol{p}= 0.0045$), "cell adhesion" (GO:0007155; $\boldsymbol{p}= 0.013$), and "angiogenesis" (GO:0001525; $\boldsymbol{p}= 0.018$) was observed in protein-altering genes that were upregulated in either Brune or Steidl arrays (Appendix 11).

3.4.6 Differences in the mutational landscape of EBV-positive and EBV-negative HL

322 genes with protein-altering mutations were present in EBV-positive (n=6) cases (Appendix 13) and 587 genes with protein-altering changes were mutated in EBV-negative (n=7) cases (Appendix 14). The mutation rate was higher, but not statistically significant (t-Test; ***p***= 0.167), in EBV-negative cases (Mean=4.46/Mb) compared to EBV-positive cases (Mean=2.96/Mb). However, a significantly higher mutation rate in 16 EBV-negative cases (Mean=8.92/Mb) compared to 7 EBV-positive cases (Mean=4.82/Mb) was observed in my series and Reichel series (from here onwards called as 'HL primary series') (t-Test; ***p***= 0.022). I also wanted to know if those cellular functions disrupted by mutation were different between EBV-positive and EBV-negative cases. To do this I performed a gene ontology analysis on the genes mutated in either EBV-positive or EBV-negative HLs, but not in both, in my series. Gene ontology analysis showed the enrichment of 'cell cycle' and 'DNA repair' related genes in both EBV positive and negative cases. Biological processes enriched in the genes mutated in both EBV-negative and EBV-positive cases in my series are shown in Table 21.

I observed enrichment of similar GO terms when I repeated the gene ontology analysis comparing genes mutated in EBV-negative cases with EBV-positive cases in the HL primary series. Table 22 shows the top 20 genes most frequently mutated in either EBV-positive or EBV-negativetumours, but not in both, in the HL primary series (see Appendix 15 for a complete list).

Table 21. Biological processes that are common in both EBV-positive and EBV-negative paediatric HL tumours.

Gene ontology terms
GO:0060249~anatomical structure homeostasis
GO:0001894~tissue homeostasis
GO:0045494~photoreceptor cell maintenance
GO:0042592~homeostatic process
GO:1903047~mitotic cell cycle process
GO:0051276~chromosome organization
GO:0000723~telomere maintenance
GO:0044770~cell cycle phase transition
GO:0006259~DNA metabolic process
GO:0022402~cell cycle process
GO:0006974~cellular response to DNA damage stimulus
GO:0006281~DNA repair
GO:0007049~cell cycle
GO:0032508~DNA duplex unwinding
GO:0032392~DNA geometric change
GO:0051592~response to calcium ion
GO:0042384~cilium assembly
GO:0006302~double-strand break repair

Table 22. 20 genes most frequently mutated exclusively in either EBV-positive or EBV-negative primery HL series (n=23).

Gene	Number of samples mutated	EBV status	Gene	Number of samples mutated	EBV status
F8	3	Positive	IGHV2-70	8	Negative
TPTE2	3	Positive	TNFAIP3	7	Negative
AGPAT9	2	Positive	IGHE	5	Negative
CCDC175	2	Positive	IGLV3-1	5	Negative
CHD4	2	Positive	MUC17	5	Negative
CMPK1	2	Positive	HIST1H1E	5	Negative
COL9A1	2	Positive	B2M	4	Negative
DNAH12	2	Positive	IGLL5	4	Negative
EML4	2	Positive	ITPKB	4	Negative
EPG5	2	Positive	EEF1A1	4	Negative
ITPR2	2	Positive	SOCS1	4	Negative
LAMA4	2	Positive	ITPR1	3	Negative
MYO6	2	Positive	LRRN3	3	Negative
OR4C5	2	Positive	CSMD3	3	Negative
PDE4DIP	2	Positive	HECTD1	3	Negative
SBF2	2	Positive	UBE2A	3	Negative
SGOL2	2	Positive	DNAH14	3	Negative
TMEM57	2	Positive	SRRM2	3	Negative
TRDN	2	Positive	C1orf87	3	Negative
ZNF717	2	Positive	CCDC80	3	Negative

Given that it has been shown that TNFAIP3, a negative regulator of NF-κB signalling, is more frequently mutated in EBV-negative HL than in EBV-positive HL, I wanted to determine if mutations in negative regulators of NF-κB signalling were also more frequent among the EBV-negative cases compared with the EBV-positive cases from the HL primary series. To do this, I compiled a list of genes (Table 23) reported to negatively regulate NF-κB signalling by Warner et al., Gewurz et al., and also the gene ontology terms available in amiGO (Gene ontology data base (Warner et al., 2013, Gewurz et al., 2012, Gene Ontology, 2015). Gene enrichment analysis revealed a significant enrichment of negative regulators of NF-κB in the EBV-negative HL but not in EBV-positive HL.

Table 23. 174 genes reported to negatively regulate NF-κB signalling.

ACOD1	CNOT6	FBXO31	MINA	PIAS4	STAT2
ACOX1	COMMD1	FLJ12443	MPZL1	PKHD1	STAT3
AIM1L	COMMD4	FLJ35776	MRPL55	PKNOX2	TAX1BP1
AIM2	COMMD6	FOXJ1	MUC1	PLEKHG2	TCEAL7
AKAP14	COMMD7	FOXP3	MYLK	POGZ	TLR9
ANXA4	CRLF1	FUS	MYST4	PPM1B	TNFAIP2
ARRB1	CSDE1	G3BP2	NAP1L2	PPP1R14C	TNFAIP3
ARRB2	CYLD	GADD45G	NDFIP2	PPP1R3B	TNFSF18
ASB9	CYP1B1	GCKR	NFKBIA	PPP4R1L	TNIP1
BAG2	DAB2IP	GFI1	NFKBID	PRMT2	TRAF3
BEX1	DAP	GUCA2B	NFKBIL1	PSMD10	TRIB3
BMP7	DET1	H2AFY	NFRKB	PTGIS	TRIM21
BRMS1	DKFZP762E1312	HAVCR2	NLRC3	PTPN23	TRIM37
BZRPL1	DRD2	HEXIM2	NLRC5	PTPRK	TRIM40
C10ORF76	DRD3	HIST1H2BN	NLRP12	PYCARD	TRY1
C22ORF4	DRD4	IFIT1	NLRP2B	PYDC1	TSG101
CA2	DUSP16	IL16	NLRP3	PYDC2	UBE2D1
CABP7	DUSP4	IRAK1	NR1H4	RASA4	UBE2L3
CACTIN	DYRK2	IRAK2	OATL1	RBCK1	UBXN2A
CAD	EDG2	IRAK3	OTUD6B	RNF13	UFL1
CAT	ENDOG	ITCH	OTULIN	ROBO4	UGDH
CCR7	ERBIN	KIAA0522	PARP10	RWDD3	USP7
CDK5RAP3	ERK8	KL	pcdh19	SERPINE1	USP8
CDKN2A	EVI5	KLF4	PDLIM2	SETD6	YWHAG
CHCHD3	EZH1	LOC644168	PDXDC1	SIRT1	ZC3H12A
CHP1	FASTKD3	MAGEA3	PELI1	SLA	ZC3H12B
CKAP2	FBX039	MAP2K5	PELO	SMARCC1	ZFC3H1
CMKLR1	FBXL2	MAP3K1	PFKP	SMARCC2	ZNF675
CNOT10	FBXL7	MAPK15	PHF7	SSR2	ZNF830

3.4.7 Identification of biological functions enriched in genes with protein-altering mutations in paediatric HL cases

Gene ontology analysis of 857 genes with protein-altering mutations in my series revealed the significant enrichment of a number of key biological processes (Appendix 12) of potential relevance to HL, including "chromosome organization" (GO: 0051276; $p < 0.00001$), "negative regulation of DNA metabolic process" (GO: 0051053; $p = 0.001$), "DNA repair" (GO: 0006281; $p = 0.001$), "negative regulation of telomere maintenance" (GO:0032205; $p = 0.001$), and "mitotic cell cycle process" (GO: 1903047; $p = 0.001$).

Given that gene ontology analysis performed here had shown a consistant enrichment of GO terms related to the mitotic spindle associated genes, I decided to pursue this further. I compiled a list of 513 mitotic spindle associated genes (spindle (GO:0005819); mitotic spindle check point (GO:0031577); nuclear chromosome segregation (GO:0098813); and sister chromatid segregation (GO:0000819) from amiGO. (Gene Ontology, 2015). I found significant enrichment of these 513 mitotic spindle associated genes in the mutated genes in my series (odds ratio=1.50; $p < 0.0001$). Furthermore, I saw a significant enrichment of mitotic spindle associated genes in mutated genes that were also downregulated in primary HL (odds ratio = 4.12 and $p < 0.0001$) but not in the mutated genes that were upregulated in HL (odds ratio = 0.17 and $p = 0.674$) in my series.

Having shown that the mitotic spindle-associated genes were significantly enriched in genes with protein-altering mutations and also in a subset of these genes that were also downregulated in HRS cells, I next focussed on the subset of mitotic spindle-associated genes comprising the mitotic spindle checkpoint, a multi-protein cascade that prevents separation of the duplicated chromosomes until all chromosomes are properly attached to the mitotic spindle. I derived a comprehensive mitotic spindle checkpoint signature, comprising 103

genes (Table 24) including those classified under the GO term ‘GO:0031577’, those reported by Bieche et al., (Bieche et al., 2011) and those identified in an unbiased Pubmed search using the search term ‘mitotic spindle checkpoint’. I found that genes with protein-altering mutations in 13 paediatric HL exomes were significantly enriched in mitotic spindle checkpoint genes (odds ratio=7.87; $\boldsymbol{p}$<0.0001). Figure 27 shows a heatmap of mitotic spindle checkpoint genes across the 13 paediatric HL exomes. I found that at least one mitotic spindle checkpoint gene was mutated in 9/13 paediatric HL samples.

Table 24. 103 mitotic spindle check point genes.

ANAPC1	BUB1	CENPE	KNTC1	PTTG1	TPR
ANAPC10	BUB1B	CENPF	LCMT1	RAD21	TPX2
ANAPC11	BUB3	CEP250	MAD1L1	RAE1	TTK
ANAPC15	CASC5	CLASP1	MAD2L1	RAN	UBD
ANAPC2	CCNA1	CLASP2	MAD2L2	RASSF1A	UBE1C
ANAPC4	CCNA2	CSE1L	MRE11A	RCC1	UBE2B
ANAPC5	CCNB1	CSNK2A1	NDC80	RNF2	UBE2N
ANAPC7	CCNB2	CSNK2A2	NDRG1	SGOL2	USP44
APC	CCNB3	DUSP1	NEDD8	SMC1A	XRCC3
ATM	CDC16	DYNC1LI1	NEK2	SMC1B	ZNF207
AURKA	CDC2	ESPL1	NUP133	SMC3	ZW10
AURKAIP1	CDC20	FBXO5	PCID2	SPDL1	ZWILCH
AURKB	CDC23	FBXW7	PLK1	STAG1	ZWINT
AURKC	CDC26	FZR1	PPP1CA	STAG2	
BCL2L11	CDC27	GEN1	PPP1R2	TACC1	
BIRC5	CDC34	GSG2	PRCC	TACC2	
BRCA1	CDK5RAP2	HSPB1	PRPF4B	TACC3	
BRCA2	CDKN1A	KLHL22	PSMG2	TEX14	

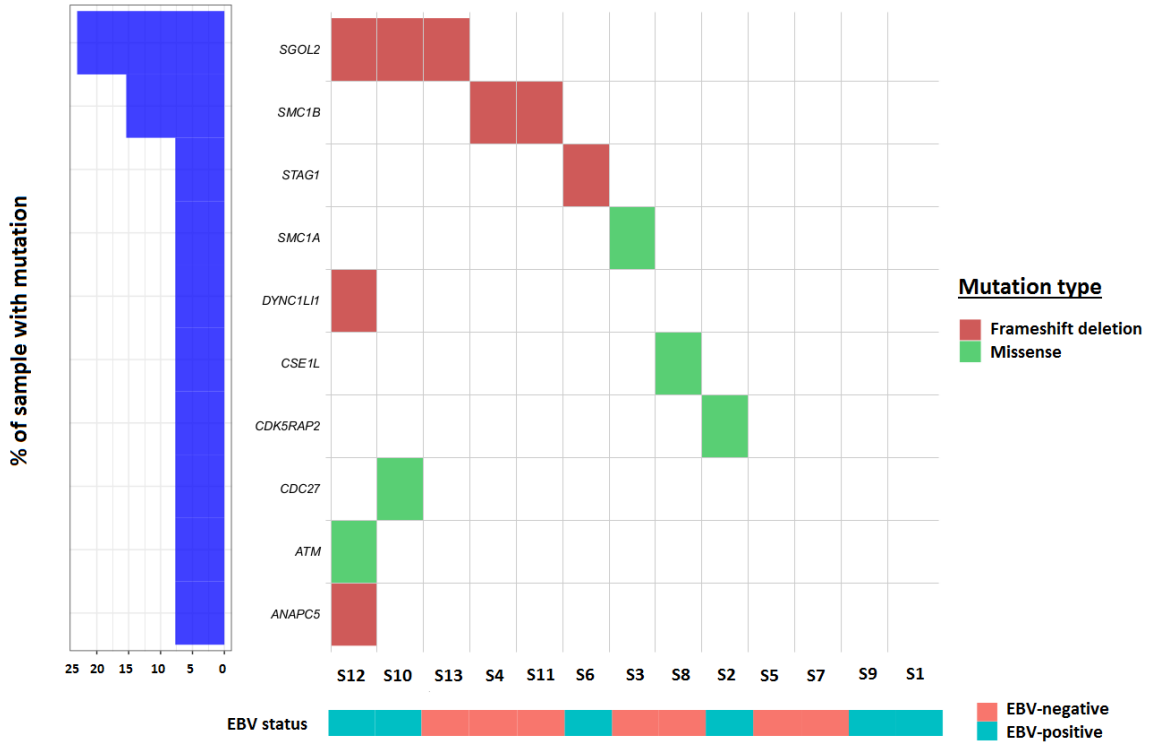

Figure 27. Nine of the thirteen paediatric Hodgkin lymphoma samples contain protein-altering mutations in at least one mitotic spindle checkpoint gene.

3.4.8 Transcriptional regulation of mitotic spindle-associated genes by Epstein-Barr virus infection

Mutations in mitotic spindle-associated genes were more frequent in EBV-negative (mean=3.85) compared with EBV-positive tumours (mean=3.16). I next explored the possibility that EBV might contribute to the inactivation of mitotic spindle functions in HRS cells. To do this I re-analysed a previously published dataset which had reported the global gene expression changes that follow the infection of primary germinal centre B cells with EBV (Leonard et al., 2011). This analysis revealed that genes down-regulated (odds ratio=2.23; $\boldsymbol{p}< 0.0001$), but not those up-regulated (odds ratio=1.14; $\boldsymbol{p}= 0.293$), by EBV in primary GC B cells, were significantly enriched for mitotic spindle-associated genes. I repeated this analysis using a previously published dataset that had reported the gene expression changes that follow the transfection of GC B cells with the EBV oncogene, latent membrane protein-1 (LMP1) (Vockerodt et al., 2008). I found that genes down-regulated (odds ratio=1.80; $\boldsymbol{p}< 0.0001$), but not those up-regulated (odds ratio=1.48; $\boldsymbol{p}= 0.778$), by LMP1 in primary GC B cells, were significantly enriched for mitotic spindle-associated genes.

Furthermore, I showed that the subset of mitotic spindle-associated genes, mitotic spindle checkpoint genes, that were transcriptionally down-regulated by either EBV or LMP1, or both, and mutated in at least one HL tumour were also significantly more likely to be down-regulated (for EBV, odds ratio=8.14; $\boldsymbol{p}= 0.0006$; and for LMP1, odds ratio=17.28; $\boldsymbol{p}= 0.0001$) and not upregulated (for EBV, odds ratio=2.92; $\boldsymbol{p}= 0.267$ and for LMP1, odds ratio=0.0; $\boldsymbol{p}= 0.806$) at the transcriptional level in primary HRS cells (Figure 28).

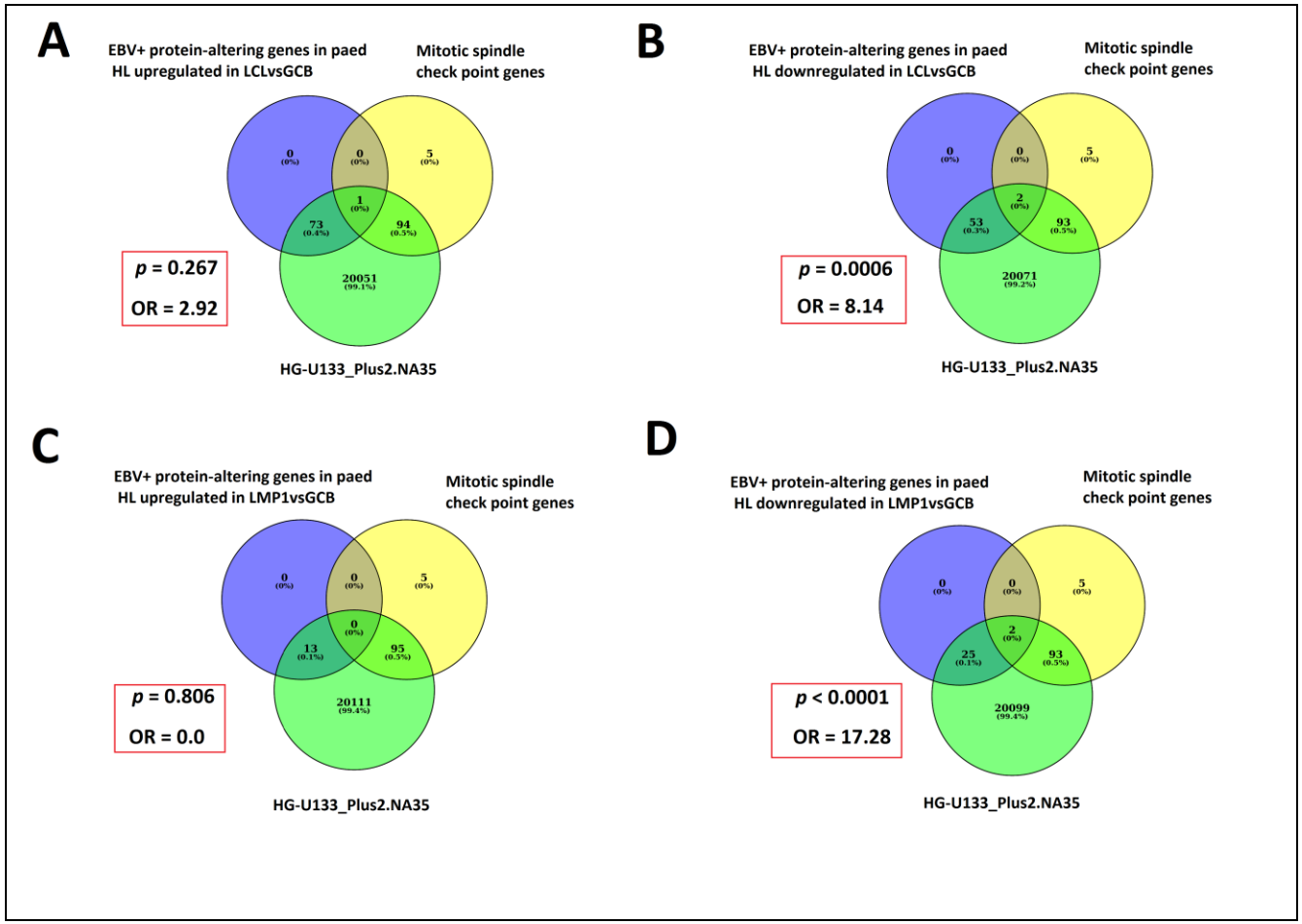

Figure 28. Mitotic spindle checkpoint genes enriched in EBV infected HRS cells. Significant enrichment (B and D) of mitotic spindle checkpoint genes enriched in EBV infected HRS cells that also results in downregulated transcription in EBV infected GCB cells. Mitotic spindle checkpoint genes are either depleted or not enriched in EBV infected HRS cells (A and C) that also results in upregulated transcription in EBV infected GCB cells.

3.5 Discussion

In this chapter I have described the genomic landscape of paediatric Hodgkin lymphoma.

I have shown that paediatric Hodgkin lymphoma has considerable heterogeneity at the gene level. For this reason I focussed on a pathway analysis which would take into account the possibility that mutations in different genes might converge on the same biological function (s). This analysis revealed the striking enrichment of genes with function in the mitotic spindle. This observation is of importance as dysregulation of the mitotic spindle associated genes are reported to be associated with several malignancies, including HL. Additionally, several compounds, including microtubule targeting agents such as Vinblastine and Vincristine, have been designed to target several mitotic spindle associated genes such as AURKA, AURKB, BUB1, CDK1, and shown in the Figure 29 (Dominguez-Brauer et al., 2015).

I was interested in the possibility that mitotic spindle functions were mutational targets in HL for two main reasons. The first is because HL is characterized by frequent aneuploidy and multinucleated condition in HRS cells. As suspected, several mitotic spindle genes were mutated in my study cohort as highlighted in purple in Figure 29. The defects of mitotic spindle function are well described in chromosome missegregation and in turn aneuploidy (Ru et al., 2002, Yu et al., 2003, Zyss and Gergely, 2009, den Hollander et al., 2010, Rengstl et al., 2013). The second reason is that the mitotic spindle checkpoint is a target for chemotherapy in children with HL such as Doxorubicin, a topoisomerase II inhibitor (Hande, 2008).

Although I was unable to deduce the biological effect of mutations affecting mitotic spindle functions in HL, analysis of genes differentially expressed in HL revealed that genes downregulated in HL but not upregulated in HL were significantly enriched in mitotic spindle

associated genes. These data suggest that many of the observed mutations in mitotic spindle associated genes are likely to be inactivating. In support of this, I observed a high frequency of frameshift mutations (23/43) in the mitotic spindle associated genes.

I also observed that the mutation rate in EBV-negative tumours is significantly higher when compared to EBV-positive tumours. This is in keeping with data from other tumours in which the frequency of mutations in viral negative forms and viral positive forms were described. For example, in gastric cancer the EBV-negative tumours have more frequent mutations than in the EBV-positive tumours (Network, 2014), similarly HPV-negative head and neck tumours have more frequent mutations than in HPV-positive tumours (Riaz et al., 2014). These observations can be explained by the requirement for cellular mutations to replace viral functions in virus negative tumours. There are already good examples in HL, where it has been found that mutations in TNFAIP3 gene are almost exclusively found in EBV-negative tumours (Schmitz et al., 2009).

I observed mutation in TNFAIP3 gene in only one sample in my cohort. This could be explained by the relatively low coverage across this gene in my analysis. However, I found that genes mutated in EBV-negative tumours but not in EBV-positive tumours were enriched in functions involved in NF-κB signalling, including many genes reported to be negative regulators of NF-KB activation. Therefore, an alternative possibility is that mutations in one or more of different negative regulators of NF-κB activation are required to substitute TNFAIP3 mutation. Similar to observation made by Schmitz et al., all of the mutations in negative regulators of NF-κB activation in my analysis were only seen in EBV-negative tumours.

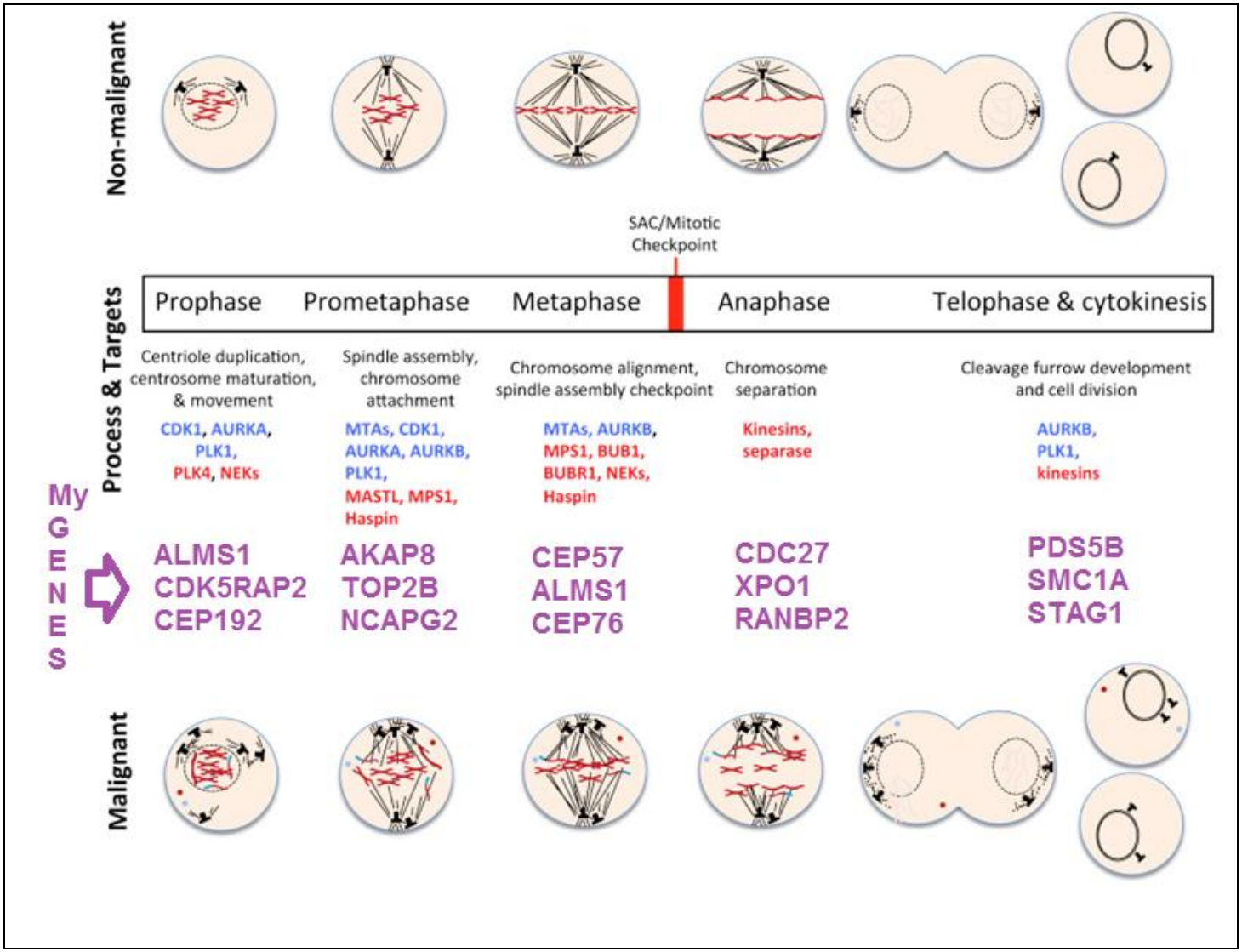

Figure 29. Mitotic Components Targeted for Cancer Therapy (modified from Dominguez-Brauer et al., 2015).

I considered the possibility that in the absence of a cellular mutation, EBV infection could disrupt mitotic spindle function through transcriptional downregulation of target genes. I found that mutated mitotic spindle associated genes were significantly enriched in transcriptionally downregulated genes but not in upregulated genes by LMP1 or EBV infection. These data suggest that EBV infection and cellular mutations might provide mutually-exclusive means to the pathogenic end points.

CHAPTER 4

INVESTIGATING EPSTEIN-BARR VIRUS GENE EXPRESSION AND THE ROLE OF THE EPSTEIN-BARR VIRUS-ENCODED BILF1 IN ENDEMIC BURKITT LYMPHOMA

4. INVESTIGATING EPSTEIN-BARR VIRUS GENE EXPRESSION AND THE ROLE OF THE EPSTEIN-BARR VIRUS-ENCODED BILF1 IN ENDEMIC BURKITT LYMPHOMA

4.1 INTRODUCTION

The association between human cancers and viruses was first described in Epstein-Barr virus (EBV) associated tumours (Epstein et al., 1964). EBV was shown to cause cancers originating from lymphocytes (e.g. Hodgkin lymphoma and Burkitt lymphoma), epithelial cells (nasopharyngeal carcinoma and gastric cancers), and mesenchymal cells (e.g. EBV-associated smooth muscle tumour) (Gutensohn and Cole, 1980, Desgranges et al., 1975, Shibata and Weiss, 1992).

Most of the studies that have explored the contribution of individual EBV genes to oncogenesis have focussed on the latent genes, making the assumption that induction of the lytic cycle is incompatible with transformation. However, recent evidence points to a role for the lytic cycle genes in transformation and cancer progression (Woellmer et al., 2012, Wille et al., 2012). Furthermore, genes previously classified as EBV lytic genes have also recently been shown to be expressed during latency. For example, BHRF1 the EBV-BCL2 homolog has been shown to be expressed as a latent gene in a subset of Burkitt lymphoma where it provides a potent anti-apoptotic signal (Kelly et al., 2009a). Recent studies using either high-throughput qPCR or RNA sequencing suggest that other 'lytic' cycle genes might also be expressed during latency in tumours (Tierney et al., 2015) to measure the expression of 45 EBV genes, not only in eBL, but also in NPC, HL, and GC. Having shown that all these tumours are characterized by the expression of subsets of 'lytic' cycle genes, I then focussed on the potential contribution of one of these, BILF1, to the pathogenesis of BL.

4.2 EBV gene expression in endemic Burkitt lymphoma, Hodgkin lymphoma, nasopharyngeal carcinoma, and gastric cancers

To measure EBV gene expression in different EBV-associated cancers, I utilized the same assays and qPCR platform described by Tierney et. al., to quantify absolute numbers of a subset (n=45) of EBV transcripts that include examples of immediate early (IE), early (E) and late (L) lytic cycle genes in addition to latent cycle genes. All these 45 assays have been assessed for expression in lytic cycle-induced Akata cells and all the lytic genes are detectable at high levels during viral replication (Tierney et al., 2015). This approach has the advantage over other approaches that can measure global EBV gene expression, such as RNA sequencing, insofar as it is cheaper and uses only nanogram amounts of RNA. Full details of this method are described in Section 2.2.10 of the Materials and Methods.

4.2.1 EBV gene expression in primary endemic Burkitt lymphoma samples

High-throughput qPCR was performed on 41 EBV-positive endemic BL samples (for details of samples, see section 2.1.4 of the Materials and Methods). Three of 41 endemic BL samples were excluded from the final analysis due to poor signal for housekeeping genes (GAPDH, PGK1, and B2M); while a further 10 samples that contained only very low levels of EBV transcripts were also excluded. In the remaining 28 eBL samples, expression levels of 45 EBV amplicons (23 latent, 2 immediate early (IE), 11 early (E), 6 late (L) lytic cycle transcripts; 3 poorly characterized RNAs (LF1, LF2 and LF3) and 4 non-coding RNAs (BARTs and EBERs)) and 3 cellular housekeeping genes (B2M, PGK1 and GAPDH) is shown in Figure 30. These samples exhibited a Latency I pattern similar to those previously reported (Rowe et al., 1987a), in which Qp-driven QUK spliced EBNA1 transcripts were detected in the majority of BL samples, albeit at variable levels. In agreement with previous reports, LMP1 transcripts were not detected in the majority (24/28) of the samples

(Niedobitek et al., 1995). LMP2A transcripts were detected in 19/28 cases, albeit at low levels. The remaining EBNAs were virtually absent, as were Cp/Wp transcripts. The two small non-coding RNAs, EBER 1 and 2, were also expressed. However, in addition to these latent transcripts previously reported, many lytic transcripts were also expressed, albeit at variable levels. In particular, there were high levels of BILF1, BALF4 and the LF transcripts. The lack of other lytic signals in particular BZLF1, indicating that this was not a low level of lytic reactivation but selective expression of certain genes.

For a subset of tumours (n = 9) I had the opportunity to compare the results obtained from the high-throughput q-PCR assays of primary eBL with those previously reported by RNA sequencing for these samples. I edited annotations obtained from NCBI in order to make this comparison. I compared gene expression values normalized to PGK1 gene expression from the high-throughput q-PCR assays to those obtained by RNA sequencing that were normalized to counts per million human reads (See section 2.2.10 of the Materials and Methods). As it can be seen from Figure 31, there was a high similarity between the expression values obtained by the two approaches, especially for the latent genes. However, the expression for some lytic genes shows different expression pattern in some cases. This may be due to the sensitivity of each method, particularly the expression of non-coding RNAs (EBERs) which are usually absent from the RNA sequencing data, or the relative scales used in each heatmap.

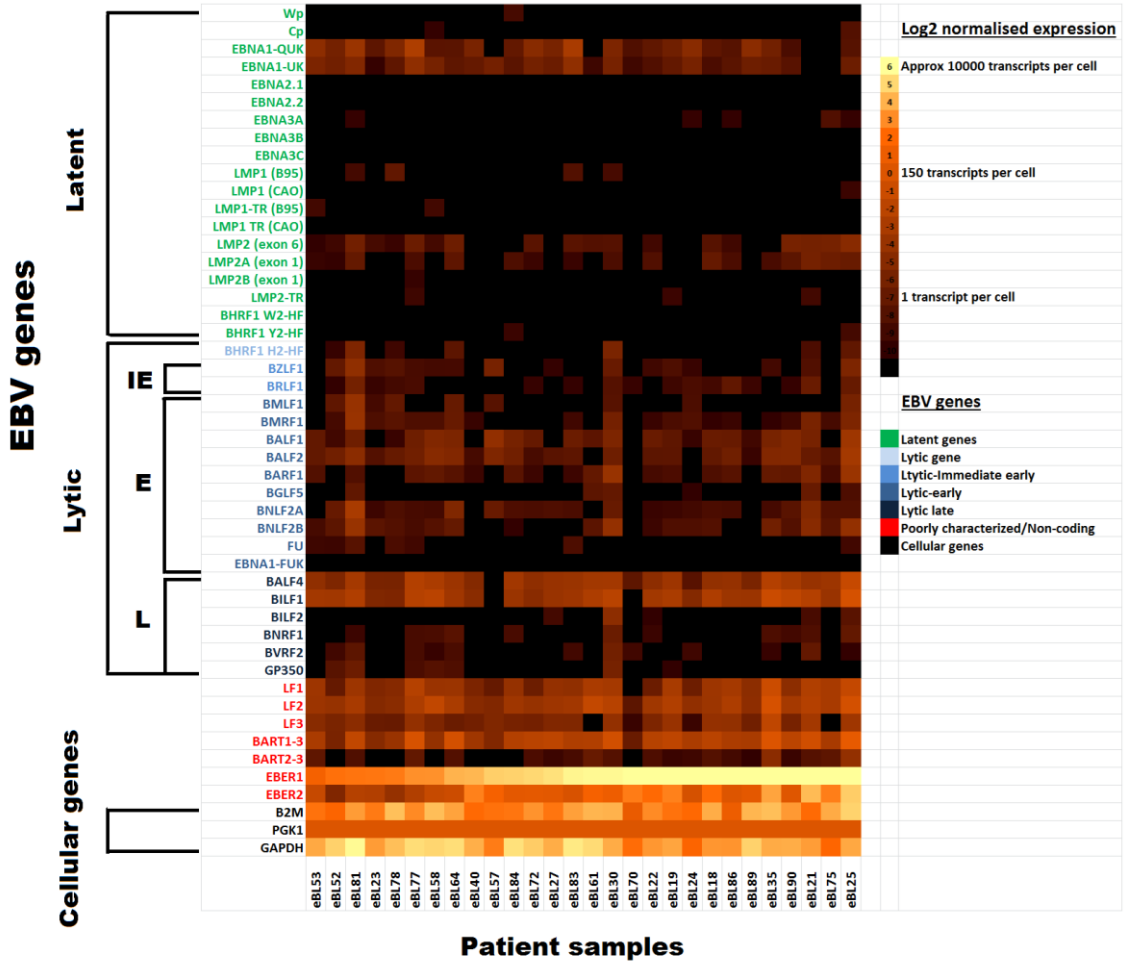

Figure 30. Heatmap summarizing EBV gene expression in endemic BL samples. Expression levels of 45 EBV amplicons (23 latent, 2 immediate early (IE), 11 early (E), 6 late (L) transcripts; 3 poorly characterized RNAs (LF1, LF2 and LF3) and 4 non-coding RNAs (BARTs and EBERs)) and 3 cellular genes (B2M, PGK1 and GAPDH) in 28 endemic Burkitt lymphoma samples was shown here. Absolute expression of the 48 amplicons was normalized using the PGK1 expression levels. The scale on the right hand side of the figure shows how these numbers correlate to transcripts per cell, based on our earlier finding that BL and LCLs cells contain an average of 150 PGK1 transcripts.

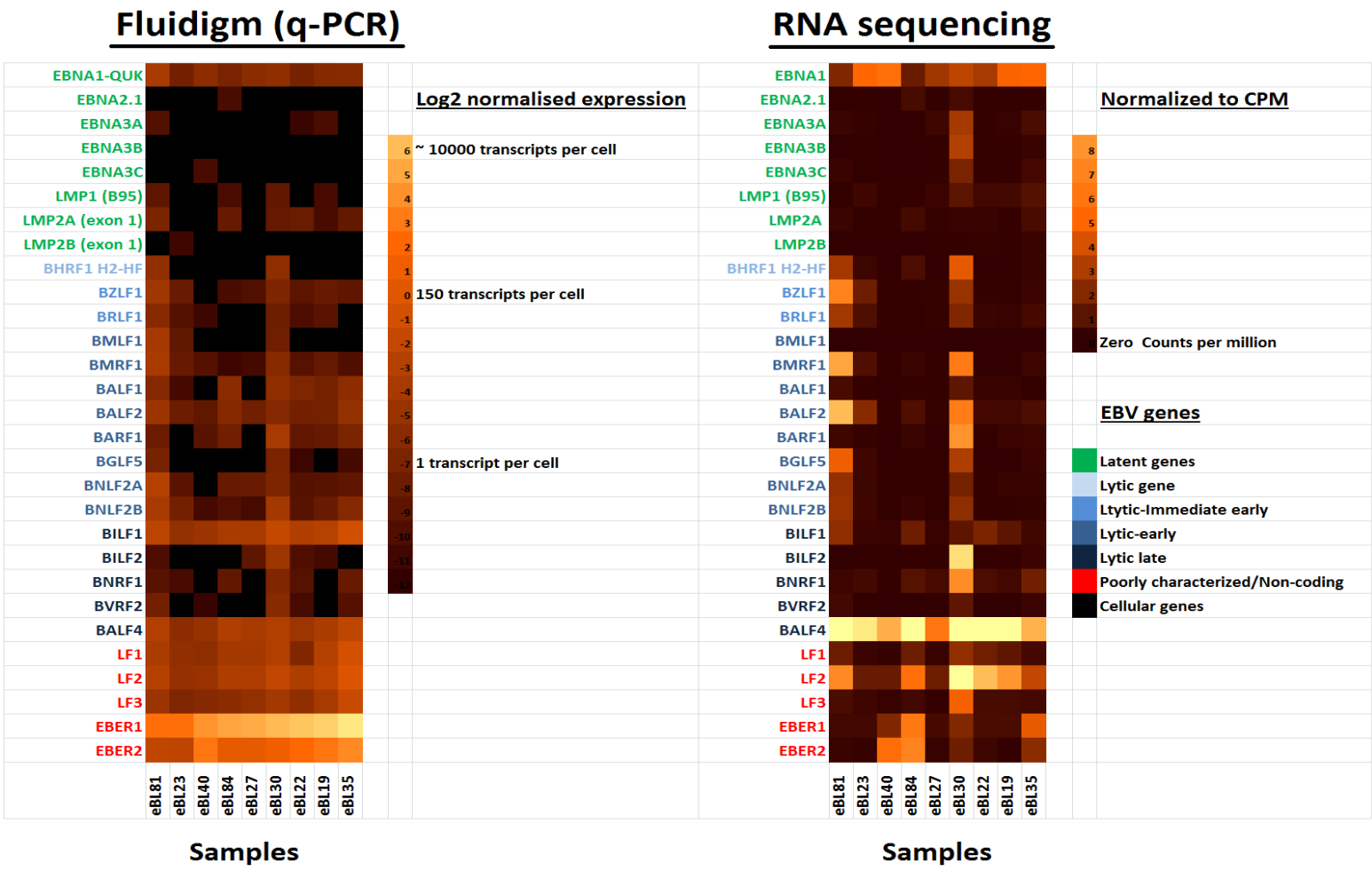

Figure 31. High concordance between the Fluidigm and RNA sequencing methods. Similar expression patterns for most of the EBV genes was observed, with the exception of small non-coding RNAs (EBERs) that were mostly absent in the RNA sequencing data.

4.2.2 EBV gene expression in primary nasopharyngeal carcinoma

A total of 50 NPC primary tumour samples (including 9 samples with amplified RNA) and 9 cell lines were available. 15/50 NPC primary samples were excluded from the analysis due to low EBER1 expression. Expression levels of 45 EBV amplicons and 3 cellular housekeeping genes in 33 Nasopharyngeal carcinoma (NPC) primary tumours (in green colour), in 3 NPC primary tumours from amplified RNA (in blue colour), and 9 NPC cell lines (in purple colour) were shown in the Figure 32. Absolute expression of these 48 amplicons was normalized using the PGK1 expression levels. EBV expression in these 44 NPC samples was consistent, but not identical, with the latency II form of infection (Brooks et al., 1992, Young et al., 1988). I detected EBER1 expression as well as variable levels of Qp-driven QUK spliced EBNA1 transcripts in the majority of NPC samples. High levels of LMP2A transcripts were also detected in most NPC and LMP1 CAO was expressed in 9/35 primary tumours. The remaining EBNAs were either absent or expressed at very low levels, similar to the Cp/Wp transcripts. Similar to eBL I observed high level expression of EBV lytic transcripts, including BILF1, BALF4 and the LF transcripts.

I was also interested to determine if the NPC cell lines recapitulate the pattern of EBV gene expression observed in primary tumours. The nine cell lines included C666.1 which is a naturally infected cell line, and *in vitro* Akata-infected cell lines CNE1, CNE2, HK1, HONE1, NP460, SUNE1, TWO1, and TWO4. Although similar EBV gene expression was observed between the primary NPC primary tumours and cell lines, expression of early lytic gene BHRF11 transcripts with the H2-HF splice (7/9), early lytic gene BGLF5 (6/9), late lytic genes BILF2 (7/9) and BVRF2 (8/9), and Fp-initiated transcripts and FUK (5/9) spliced EBNA1 transcripts (5/9) were detected more frequently in the NPC cell lines than in primary tumours. EBER2 transcripts were also abundantly expressed in the Akata-infected NPC cell lines, indicating that EBER2 expression is associated with Akata-strain.

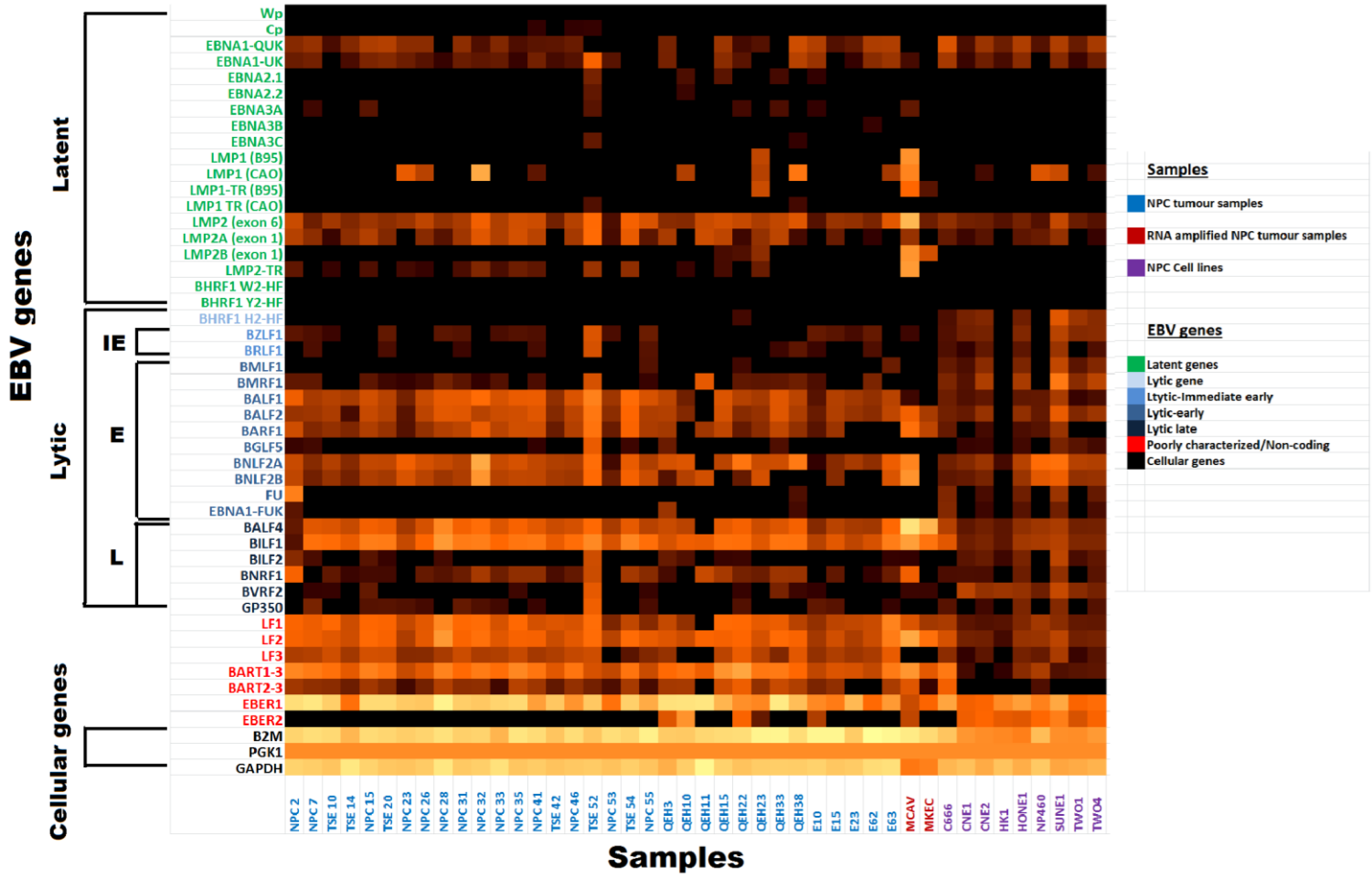

Figure 32. Heatmap of EBV gene expression in NPC samples. Expression levels of 45 EBV amplicons and 3 cellular genes in 33 NPC primary NPC tumours (in green colour), in 3 RNA amplified primary NPC tumours (in blue colour), and 9 NPC cell lines (in purple colour), were shown here. Absolute expression of the 48 amplicons was normalized using the PGK1 expression levels.

4.2.3 EBV gene expression in Hodgkin lymphoma samples

Fifteen primary HL samples (5 FFPE and 10 FF) and one EBV-positive HL cell line L-591 were available. However, two FFPE HL samples and five FF HL samples were excluded from the final analysis as they contained very low levels of EBV transcripts. As shown in Figure 33, EBV expression pattern in the remaining eight primary HL samples is consistent with latency type II pattern previously reported in HL (Deacon et al., 1993, Pallesen et al., 1991b), and latency III reported in L-591 (Baumforth et al., 2008). Qp-driven QUK spliced EBNA1 transcripts were detected in 3/8 primary HL tumours and in L-591 cells, however at very low levels. The remaining EBNAs were virtually absent in primary tumours, as were Cp/Wp transcripts. However, L-591 cells had high levels of EBNA2 expression. High levels of LMP1 (B95-8) transcripts were detected in four primary HL cases, but surprisingly were absent in the L-591 cells. Very low levels of LMP2 expression was observed in half of primary tumours and in the L-591 cell line (~10-20 transcripts per cell). Varying levels of BNLF2A and BNLF2B transcripts were present in primary HL cases; however this expression may also have come from the LMP1 RNA since LMP1 and BNLF2A/B share the same 3' end. High level of BNLF2A expression was observed in the L-591, whereas very low levels of BNLF2B expression was observed in the L-591 cell line. Lower levels of BILF1, BALF4 and the LF transcripts were observed in two primary HL samples and the L-591 cells. The level of expression for these transcripts in the remaining samples was almost negligible. There was no significant expression of other lytic genes in these samples.

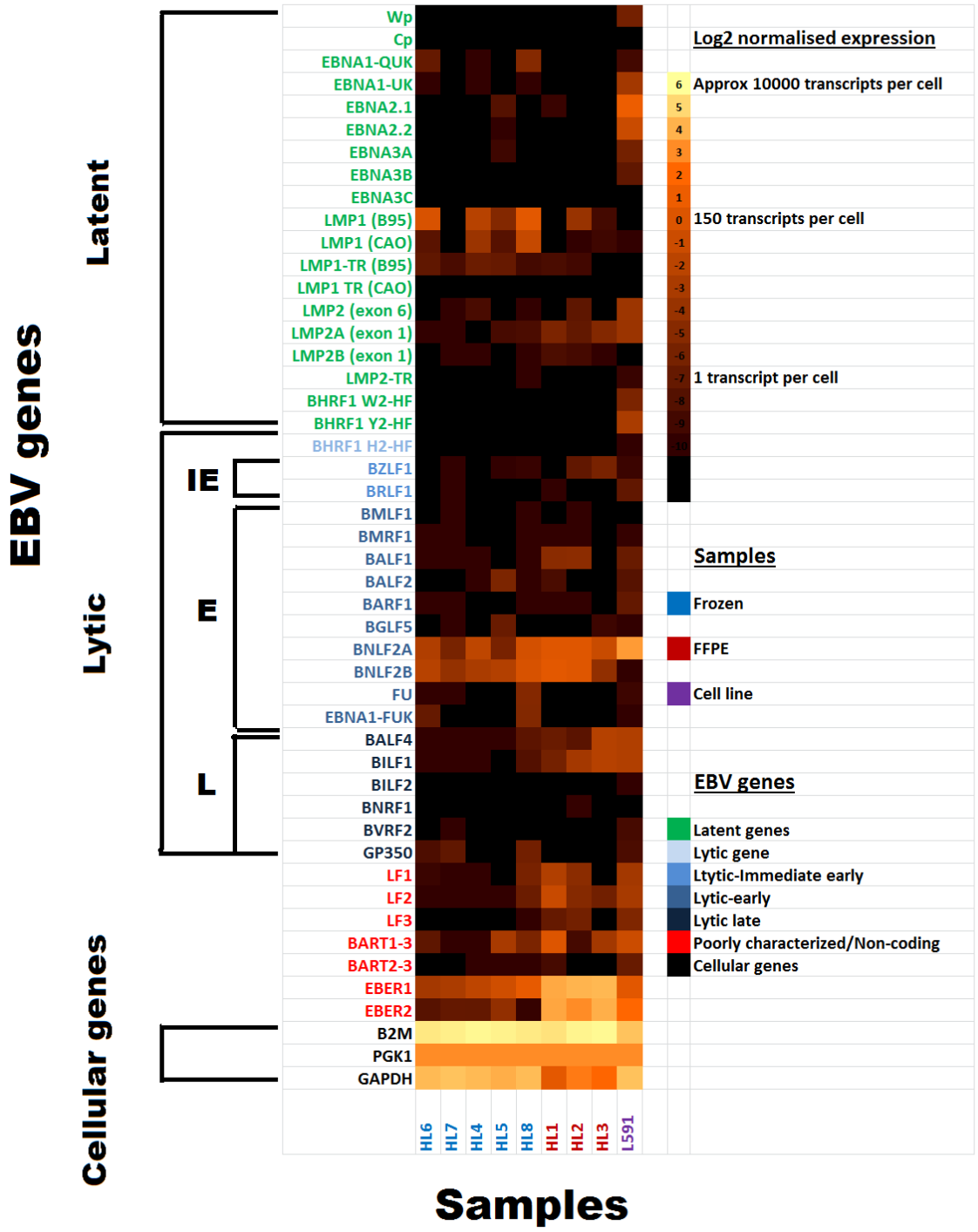

Figure 33. EBV gene expression in HL samples. Expression levels of 45 EBV amplicons (23 latent, 2 immediate early (IE), 11 early (E), 6 late (L) transcripts; 3 poorly characterized RNAs (LF1, LF2 and LF3) and 4 non-coding RNAs (BARTs and EBERs)) and 3 cellular genes (B2M, PGK1 and GAPDH) in 8 Hodgkin lymphoma samples and a HL cell line L591 were shown here. FF: Flash Frozen; FFPE: Formalin-fixed, paraffin-embedded.

4.2.4 EBV gene expression in gastric cancer cell lines

EBV gene expression data was obtained from four gastric cancer cell lines, which include two naturally infected cell lines (SNU-719 and YCCEL-1) and two *in vitro* EBV-infected cell lines (AGS and OE19). EBV gene expression showed latency type I pattern in all four cancer cell lines (Figure 34), in agreement with the previously published data (Decaussin et al., 2000, zur Hausen et al., 2000) including low levels of Qp-driven QUK spliced EBNA1 and LMP2A, but no LMP1 transcripts were detected. High expression of early lytic transcripts, BNLF2A and BNLF2B, were observed. High levels of BILF1, BALF4 and the LF transcripts were also observed.

In summary, Fluidigm analysis of EBV expression indicates that the pattern of EBV gene expression is broader than previously recognised, with lytic genes being present in addition to the well described latent genes. However lytic expression is selective, i.e. not all genes are expressed suggesting this is not full virus replication. This is consistent with the proposed hypothesis that virus reactivation must be blocked to allow tumour formation, while still allowing certain genes to be expressed that may provide a growth advantage.

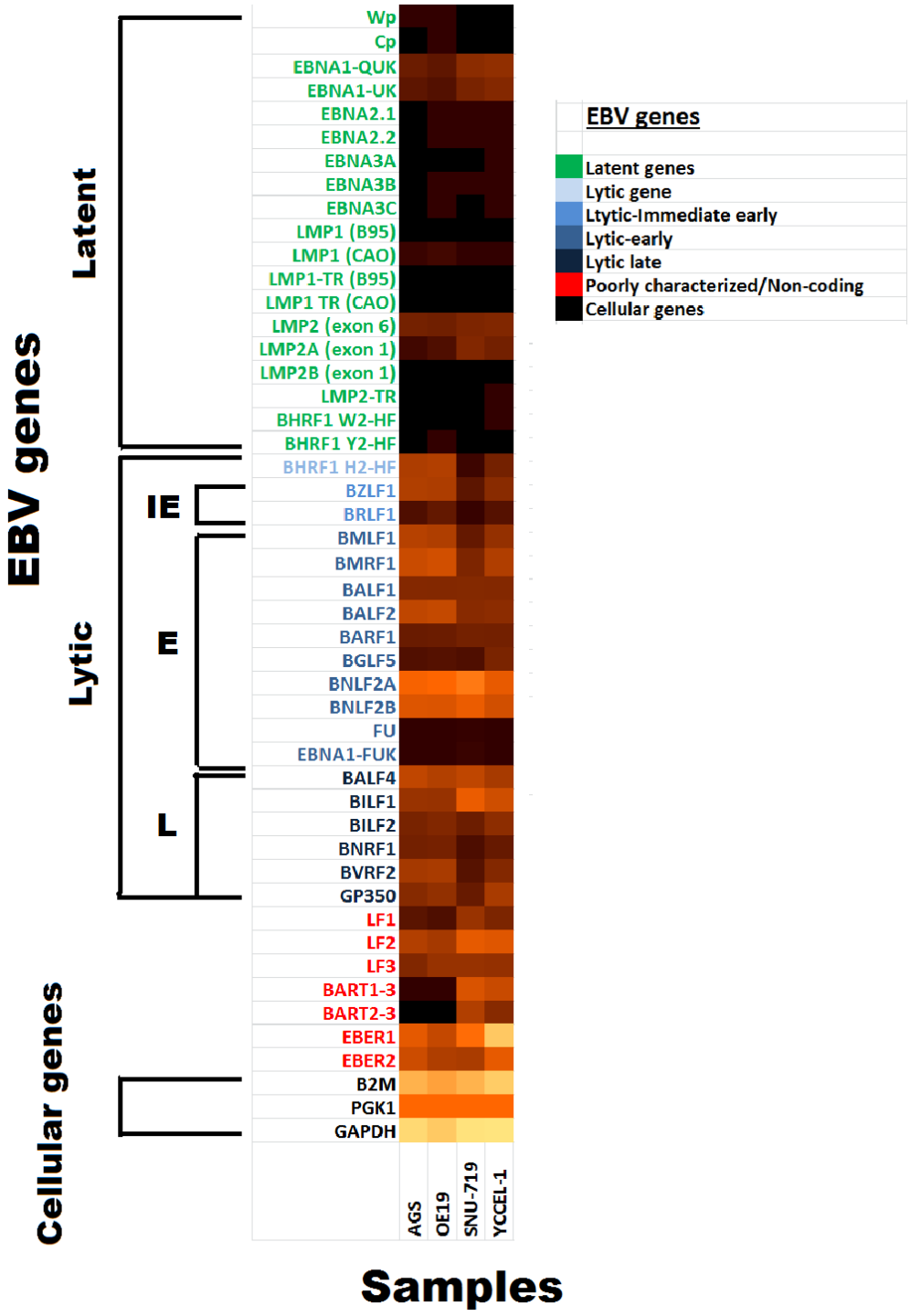

Figure 34. Heatmap summarizing EBV gene expression in gastric cancer cell lines.
Expression levels of 45 EBV amplicons (23 latent, 2 immediate early (IE), 11 early (E), 6 late (L) transcripts; 3 poorly characterized RNAs (LF1, LF2 and LF3) and 4 non-coding RNAs (BARTs and EBERs)) and 3 cellular genes (B2M, PGK1 and GAPDH) in four gastric cancer cell lines were shown here. Absolute expression of the 48 amplicons was normalized using the PGK1 expression levels.

4.3 BILF1 detection using RNAScope and immunohistochemistry

The results presented in the foregoing section provide further convincing evidence that some genes formerly reported as lytic cycle genes are expressed from moderate to high levels in primary EBV-associated malignancies. I decided to focus on BILF1, the only G protein coupled receptor (GPCR) encoded by EBV. BILF1 has the highest sequence similarity to the chemokine receptor CXCR4 and has been suggested to be important in escaping immune recognition by downregulating the surface expression of MHC class I proteins (Zuo et al., 2009, Griffin et al., 2013). Other studies suggest pathogenic roles for BILF1 through its ability to hetero-dimerize with various chemokine receptors in turn altering the responsiveness of B lymphocytes to chemokines (Nijmeijer et al., 2010a, Vischer et al., 2008). BILF1 also inhibits phosphorylated RNA-dependent protein kinase R (PKR), which plays a role in anti-viral immune responses (Beisser et al., 2005). BILF1 may also directly contribute to tumour development as it has been shown to induce secretion of VEGF and tumour formation in nude mice (Lyngaa et al., 2010).

I utilized RNA *in situ* hybridization, RNAScope, (Wang et al., 2012) and immunohistochemistry methods to detect BILF1 as these methods can localize BILF1 expression. Furthermore, a major limitation with both RNA sequencing and PCR method is that the data are an average of the cell population and does not provide information on what proportion of cells are BILF1 positive. This is important as BILF1 expression in the majority of cells may be biologically significant while high expression in just 1-2% of tumour cells may not. Details of these methods were provided in sections 2.2.22 and 2.2.23 of the Materials and Methods. I first made a FFPE block comprising pelleted cells of EBV loss Akata cells line, EBV-positive Akata cells, and EBV-positive Akata cells induced into the lytic cycle by method described in section 2.2.22 of the Materials and Methods. PPIB

(Cyclophilin B) and DapB were used as positive and negative controls for *in situ* hybridization method, respectively. I detected BILF1 expression in both Akata un-induced and induced Akata cells. In un-induced Akata cells BILF1 transcripts were present in all cells at low levels and in 1-5% of cells at higher levels (Figure 35) representing the small proportion of cells spontaneously undergoing lytic cycle. In induced Akata around 20% of cells showed higher levels of expression (Figure 35D) consistent with the expression of BILF1 in cells induced into lytic cells although I did not detect the expression of the positive control gene, PPIB (Figure 35B).

I had available to me FFPE sections of 15 BL samples which had already been tested for BILF1 expression by qPCR, of which 13 were BILF1-positive and 2 were BILF1-negative. These 15 samples were then subjected to RNA *in situ* hybridization. 7 samples that were BILF1-positive by qPCR were also positive by RNA *in situ* hybridization. 2 samples were negative for BILF1 expression in both methods. Four samples that were BILF1-positive by qPCR did not show BILF1 expression in the RNA *in situ* hybridization method, suggesting RNA *in situ* hybridization method may not be as sensitive as qPCR based methods. One sample failed in the RNA *in situ* hybridization method. Example for RNA *in situ* hybridization in BL samples was shown in the Figure 36, where EBV-positive samples (Figures 36A-B) show positive staining for BILF1, where majority of tumour cells show low levels of BILF1 expression while few cells show higher level BILF1 expression indicating BILF1 is expressed in latency as well.

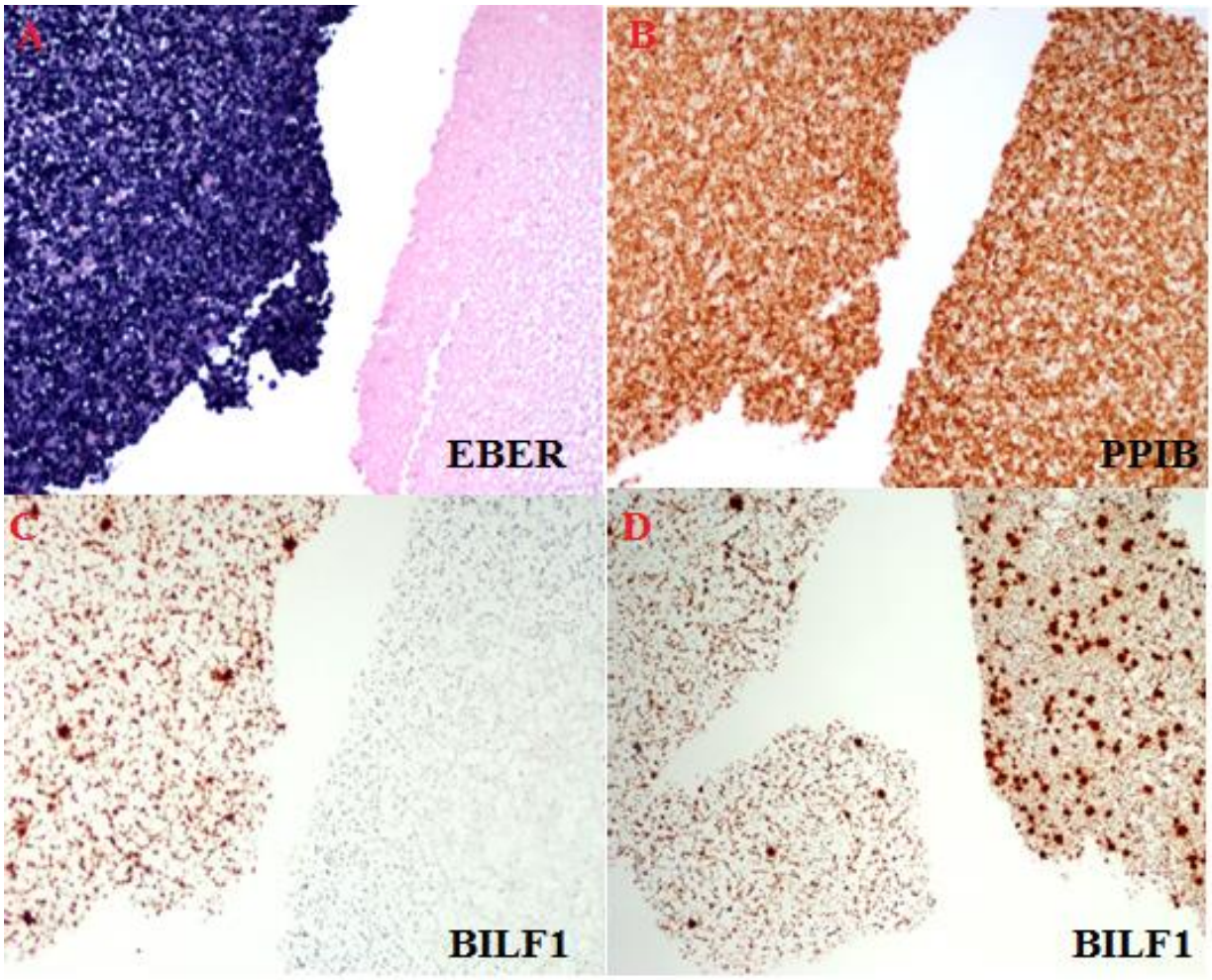

Figure 35. Validation of the specificity of BILF1 RNAScope. A- EBER in situ hybridization on EBV-positive Akata cells. B. Image Showing the results of RNAScope with the positive control probe which detects expression of the housekeeping gene, PPIB. C. No expression of BILF1 by RNAScope in Akata-loss cells. D. A small number of cells spontaneously entering the lytic cycle, after treatment with a lytic cycle inducing agent, in EBV-positive Akata cells show very high levels of BILF1 mRNA.

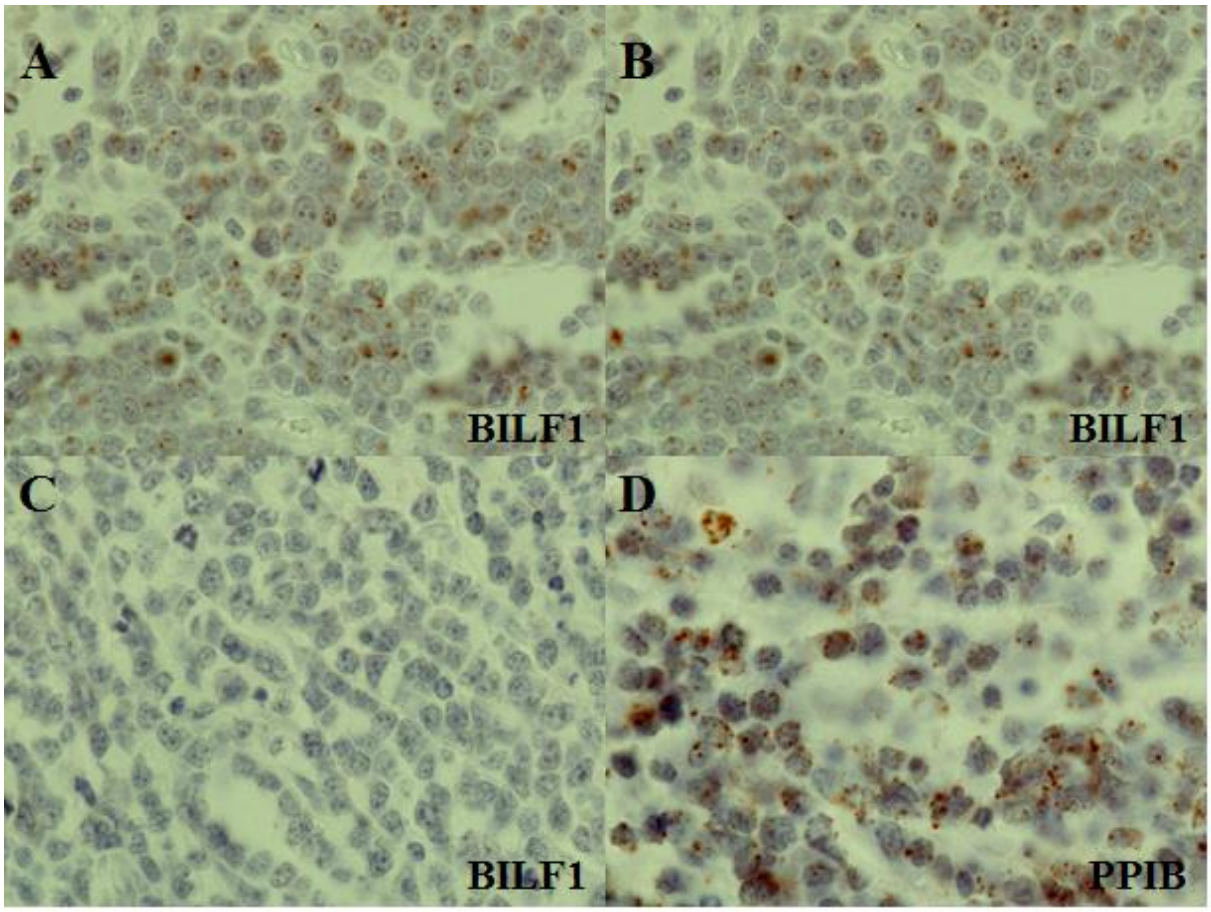

Figure 36. BILF1 RNAScope in BL. A-B shows two cases of BL positive for BILF1 mRNA showing typical punctate staining. C- shows a case lacking BILF1 mRNA. D. Positive staining of this case with the PPIB probe confirms that this was not due to poor RNA quality.

Next, I wanted to study the expression of BILF1 protein in tumour sample. As there was no commercially available antibody to BILF1, I designed two peptides for BILF1 antibody against BILF1 (YP 401711.1) and obtained rabbit polyclonal ab for two BILF1 peptides and BILF1 antibody staining was optimized in EBV-negative, EBV-positive, and lytic cycle induced Akata cells as described in section 2.2.24 of the Materials and Methods.

Figure 37A shows no expression of BILF1 in the EBV-negative Akata cells, negative control, and the lytic cycle induced Akata cells, positive control, show a strong expression of BILF1 on membrane and cytoplasm (Figure 37B). An example for the absence of BILF1expression was shown in the BL sample BL-782 (Figure 37C) and BL sample BL-668 (Figure 37D) show membranous expression of BILF1.

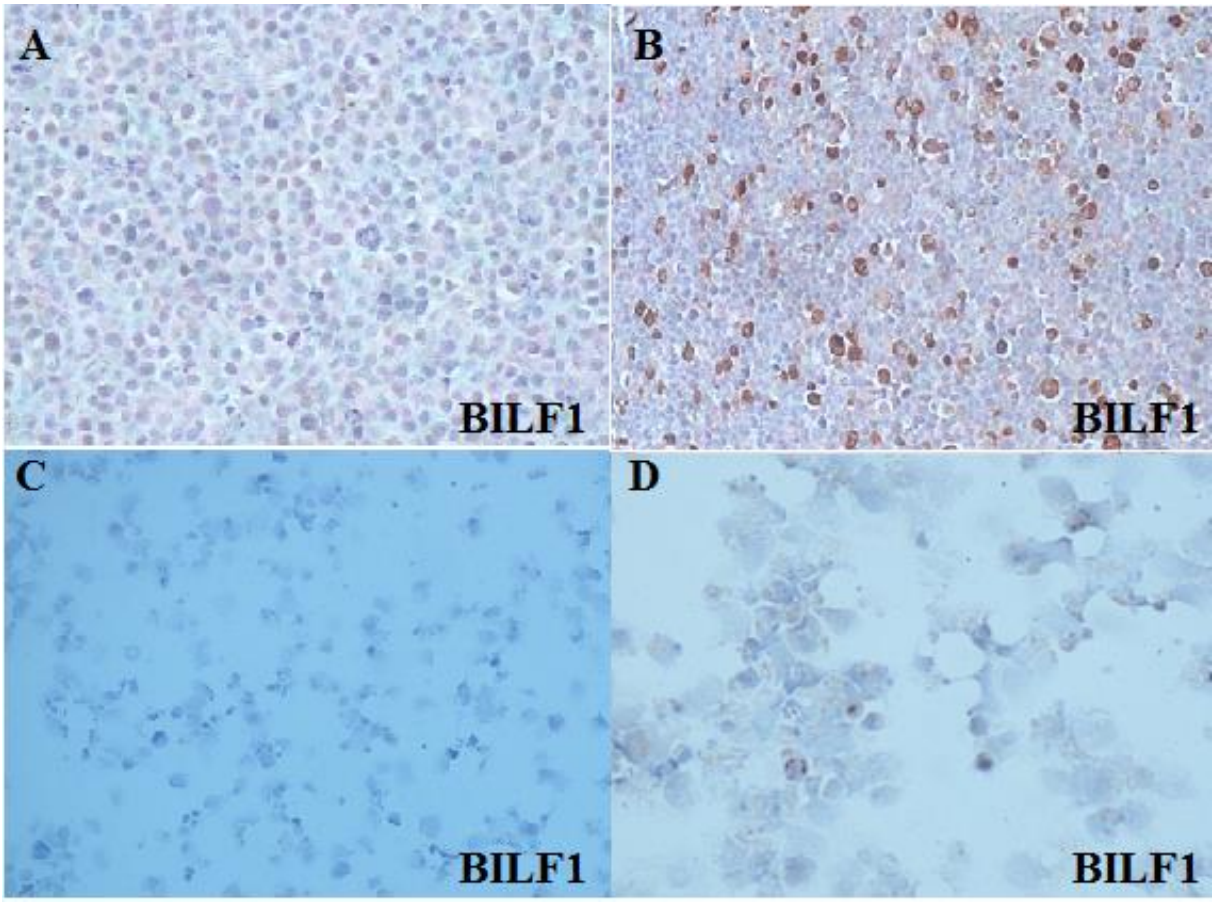

Figure 37. BILF1 expression shown by immunohistochemistry method. BILF1 expression was absent in the (A) Akata loss cells (negative control); (B) and present in in the lytic cycle-induced Akata cells, (positive control); (C) No expression for BILF1 was seen in sample BL-782; whereas (D) BL sample BL-668 show nuclear expression of BILF1.

4.4 Genes regulated by BILF1 in DG-75 cells and in primary human germinal centre B cells

4.4.1 BILF1 expression in DG-75 cells

Since BILF1 is a constitutively active GPCR I reasoned that it is likely to be a potent transcriptional regulator. Therefore, initially I elected to investigate the transcriptional consequences of BILF1 over-expression in the EBV-negative BL cell line, DG-75.

To do this, DG-75 cells were transfected with BILF1-plasmid and empty vector-plasmid. After 48 hours following transfection, GFP-positive cells were flow-sorted. GFP-positive cells were about 22% of total population in vector-transfected DG-75 and 18% in BILF1-plasmid transfected cells. BILF1 protein was extracted from these flow sorted DG-75 cells. Western blotting was performed on these HA-tagged BILF1-GFP and vector-GFP proteins. A 50 K-Da BILF1 protein is seen in western blotting for DG-75 cells transfected with BILF1-GFP, but not in the protein extracted from empty vector transfected DG-75 cells (Figure 38). RNA sequencing data for three biological replicates of DG-75 cells following transfection with BILF1 or vector, and this data was analysed using the methods described in section 2.2.12 of the Materials and Methods.

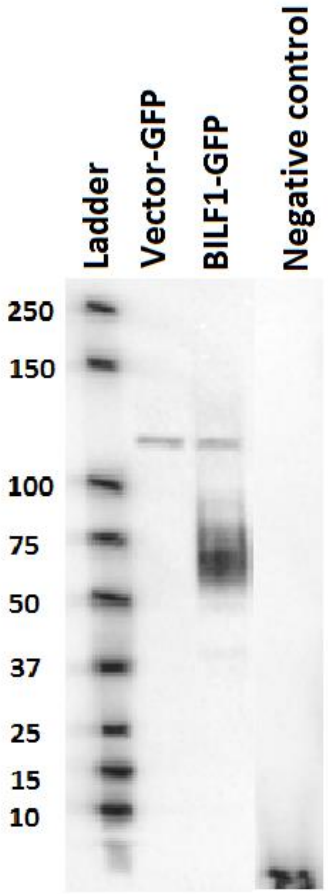

Figure 38. Detection of BILF1 protein in DG-75 cells transfected with BILF1-GFP plasmid. BILF1 protein extracted from flow sorted DG-75 cells transfected with BILF1 or empty vector. Western blotting was performed on these HA-tagged BILF1-GFP and vector-GFP proteins. A 50 K-Da BILF1 protein is seen in western blotting for DG-75 cells transfected with BILF1-GFP, but not in the protein extracted from empty vector transfected DG-75 cells. BILF1 protein is also absent in the negative control (NC).

4.4.2 Gene ontology analysis of genes regulated by BILF1 in DG-75 cells

I observed 38, 83 and 96 BILF1 CPM in each of the triplicates of BILF1-transfected DG-75. In the case of the vector-transfected DG-75 cells I observed 1.24, 1.35, and 1.22 BILF1 CPM, which could have been due to the contamination during the preparation of BILF1 and empty-vector plasmids.

132 genes were up-regulated and 72 genes were down-regulated in BILF1-transfected DG-75 cells compared to vector-transfected cells (absolute fold change ≥ 1.5 and $\boldsymbol{p}<0.05$) (Table 25 and Table 26).

As a first step to understand the biological relevance of the genes I had shown to be regulated by BILF1 in DG-75 cells, I performed a gene ontology analysis using DAVID6.8. This analysis revealed enrichment of several biological processes relevant to lymphomagenesis, including 'regulation of viral life cycle' (GO: 1903900; $\boldsymbol{p}= 0.0021$), 'programmed cell death' (GO: 0012501; $\boldsymbol{p}= 0.0051$), 'response to cAMP' (GO: 0051591; $\boldsymbol{p}= 0.013$), 'regulation of G-protein coupled receptor protein signalling pathway' (GO: 0008277; $\boldsymbol{p}= 0.027$), 'regulation of ERK1 and ERK2 cascade' (GO: 0070372; $\boldsymbol{p}= 0.038$), and 'I-kappaB kinase/NF-kappaB signalling' (GO:0007249; $\boldsymbol{p}= 0.042$) (Table 27).

Biological processes enriched in down-regulated genes included 'positive regulation of macromolecule biosynthetic process' (GO: 0010557; $\boldsymbol{p}= 0.01$), 'positive regulation of nitrogen compound metabolic process' (GO: 0051173; $\boldsymbol{p}= 0.016$), and 'regulation of cell differentiation' (GO: 0045595; $\boldsymbol{p}= 0.043$) (Table 28).

Table 25. 132 genes up-regulated by BILF1 in DG-75 cells.

Gene	Fold Change	Gene	Fold Change	Gene	Fold Change
POTEI	3.61	CDRT1	1.83	GPR89B	1.59
EGR1	3.32	CRB3	1.82	LINC01465	1.59
OVCA2	2.62	BTBD19	1.81	DFNA5	1.59
TNF	2.60	EDRF1-AS1	1.80	PLEKHG5	1.59
EGR2	2.55	POLR2J2	1.79	MAFF	1.59
RUNDC3B	2.54	FBXL2	1.78	PPP1R15A	1.59
TIAF1	2.51	LOC441454	1.78	BAIAP3	1.58
PTPN6	2.45	DRICH1	1.78	ZSWIM8-AS1	1.58
CD69	2.42	ZP1	1.78	SH3TC1	1.58
C1QTNF6	2.31	NHLRC4	1.77	C10orf95	1.58
SSUH2	2.27	FAM53B-AS1	1.76	RPL13AP20	1.58
F2	2.19	FTH1P3	1.75	THBS4	1.57
ITGB2-AS1	2.17	MIR22HG	1.75	MTRNR2L5	1.57
MUC6	2.14	NUTM2B	1.74	TCF7	1.56
PCSK4	2.13	SLIT1	1.74	TCTN1	1.56
LBX2	2.10	CMTM4	1.74	CLTCL1	1.56
FCMR	2.08	POPDC2	1.73	TSC22D3	1.56
TMCC2	2.08	TPH1	1.73	LOC101928100	1.56
GLYATL2	2.08	JUN	1.71	RPS14P3	1.55
CXCL8	2.07	FAM20A	1.70	MAP3K12	1.55
PKD2L2	2.06	POMZP3	1.70	ANGPTL1	1.55
TRPV3	2.06	CPLX3	1.68	C16orf86	1.55
LPAR6	2.05	JOSD2	1.68	PTPRE	1.55
PIP5K1P1	2.03	ARID5A	1.68	ANKRD30BL	1.55
MIR663A	2.02	LOC439994	1.67	CKMT2	1.55
BRD7P3	1.99	DUSP1	1.67	MIR8072	1.54
BCL2L15	1.99	KIAA0513	1.67	ZFP36	1.54
LOC644936	1.98	CILP2	1.66	ZNF192P1	1.54
ADHFE1	1.96	FERMT2	1.65	IL10RB-AS1	1.53
FOS	1.96	CDC42EP2	1.65	TAGAP	1.53
SERPINE3	1.96	GJB7	1.64	LOXL4	1.53
MIR186	1.92	FFAR1	1.64	SH2D2A	1.53
FOSB	1.89	RGS1	1.64	PLCH2	1.53
ZNF382	1.89	ROCK1P1	1.64	STAT5A	1.52
POTEE	1.88	LINC00116	1.63	ZSWIM4	1.52
OSCAR	1.87	EFCAB2	1.62	VPS37B	1.52
CYP3A5	1.86	DOCK10	1.61	SPAG5-AS1	1.52
FLRT1	1.86	LOC100506990	1.61	PRKRIP1	1.52
COLQ	1.85	PRKXP1	1.61	CHMP3	1.52

BEST1	1.85	GREB1L	1.60	MAGEH1	1.51
ST3GAL1	1.84	TUBA3D	1.60	RSPH4A	1.51
NPB	1.83	NCF1	1.60	KCNJ13	1.51
KLHL22	1.83	PRDM11	1.60	RFX3-AS1	1.50
KLF6	1.83	HTR2B	1.60	B3GNT9	1.50

Table 26. 72 genes down-regulated by BILF1 in DG-75 cells.

Gene	Fold Change	Gene	Fold Change	Gene	Fold Change
SNORA80A	-2.79	SNORA52	-1.75	HMGN5	-1.63
ID4	-2.65	TSPAN15	-1.74	TUBB4A	-1.62
SNORA54	-2.58	SNORA73B	-1.72	HIST2H2BA	-1.61
MIR4539	-2.52	LINC01560	-1.72	KATNBL1P6	-1.61
FEZF2	-2.15	LINC01226	-1.72	SCARNA5	-1.60
SNORA23	-2.10	SNORD30	-1.71	SNORD116-14	-1.59
SCARNA22	-2.06	SNORD75	-1.71	SNORA12	-1.58
SNORD50A	-2.04	SNORA76A	-1.70	PLEKHB1	-1.57
SNORA57	-2.04	MORF4L2-AS1	-1.70	LINC01225	-1.57
SNORD54	-2.03	PFKFB4	-1.70	SH3BP5	-1.56
CATIP-AS2	-2.03	SERPINA9	-1.69	SNORA21	-1.56
NUDT6	-2.02	SNORA71D	-1.69	SNORD60	-1.55
LOC101927415	-2.00	CTAGE7P	-1.68	SLC16A13	-1.55
SNORA63	-1.99	SNORD38A	-1.67	SNORD10	-1.54
SNORD94	-1.99	CXorf21	-1.67	NRCAM	-1.54
SNORA48	-1.95	SAMD13	-1.67	MYLK-AS1	-1.53
SNORA71A	-1.93	SNORA70	-1.66	SNORD97	-1.53
SNORD1B	-1.93	CCDC67	-1.66	KIAA1107	-1.53
SCARNA28	-1.86	EGFR	-1.66	ANKRD19P	-1.53
SNORA37	-1.84	LOC100506801	-1.65	HNRNPA1	-1.52
SNORA49	-1.83	KRT17	-1.65	VSIG10	-1.51
SNORA53	-1.77	SNORA7B	-1.64	P2RX2	-1.51
LOC100129697	-1.76	TGFBI	-1.64	SNORD78	-1.50
SNORD100	-1.75	ANP32AP1	-1.63	SNORD133	-1.50

Table 27. Selected biological processes enriched in the 132 genes up-regulated by BILF1 in DG-75 cells.

Gene ontology term	*p*-value
GO:1903900~regulation of viral life cycle	0.002
GO:0012501~programmed cell death	0.005
GO:0006915~apoptotic process	0.007
GO:1903901~negative regulation of viral life cycle	0.011
GO:0051591~response to Camp	0.014
GO:0008277~regulation of G-protein coupled receptor protein signalling pathway	0.027
GO:0070372~regulation of ERK1 and ERK2 cascade	0.039
GO:0007249~I-kappaB kinase/NF-kappaB signalling	0.043

Table 28. All biological processes enriched in the 72 genes down-regulated by BILF1 in DG-75 cells.

Gene ontology term	*p*-value
GO:0021543~pallium development	0.009
GO:0010557~positive regulation of macromolecule biosynthetic process	0.011
GO:0031328~positive regulation of cellular biosynthetic process	0.015
GO:0032846~positive regulation of homeostatic process	0.015
GO:0009891~positive regulation of biosynthetic process	0.016
GO:0051173~positive regulation of nitrogen compound metabolic process	0.016
GO:0021537~telencephalon development	0.018
GO:0007413~axonal fasciculation	0.019
GO:0021895~cerebral cortex neuron differentiation	0.021
GO:0008038~neuron recognition	0.029
GO:0051240~positive regulation of multicellular organismal process	0.034
GO:0030900~forebrain development	0.041
GO:0045595~regulation of cell differentiation	0.044
GO:0007417~central nervous system development	0.045
GO:0010605~negative regulation of macromolecule metabolic process	0.047
GO:0031324~negative regulation of cellular metabolic process	0.049
GO:0001558~regulation of cell growth	0.050

4.4.3 Pathway analysis of the genes regulated by BILF1 in DG-75 cells

I performed a gene set enrichment analysis using canonical pathways (BIOCARTA CP) and Hallmark pathways gene sets from MSigDB database. As described in section 2.2.21 of the Materials and Methods, this approach enables the identification of genes enriched in signalling pathways or other biological processes.

I found significant enrichment in lymphomagenesis-related pathways in the genes up-regulated by BILF1 in DG-75 cells, including 'EPO signalling pathway' (EPO pathway; ***p***= $9E^{-08}$), 'oxidative stress induced gene expression via Nrf2' (ARENRF2 pathway; *p*= $3E^{-06}$), 'PDGF signalling pathway' (***p***= $5E^{-05}$), 'pertussis toxin-insensitive CCR5 signalling in macrophage' (CCR5pathway; ***p***= $7E^{-04}$), and 'signalling pathway from G-Protein families' (GPCR pathway; ***p***= $3E^{-03}$) (Table 29). A complete list of pathways enriched in the genes up-regulated by BILF1 in DG-75 cells is given in Appendix 18. Hallmark pathway analysis for these genes revealed significant enrichment of 'TNFA signalling via NF-κB' (***p***= $1.35E^{-12}$), 'hypoxia' (***p***= $6.16E^{-05}$), 'inflammatory response' (***p***= 0.0031), 'apoptosis' (***p***= 0.0086), and 'p53 pathway' (***p***= 0.017) (Table 29).

In the case of genes down-regulated by BILF1 in DG-75 cells, canonical pathway analysis revealed enrichment in several pathways, including 'Map Kinase Inactivation of SMRT Corepressor' (***p***= 0.013), 'CBL mediated ligand-induced downregulation of EGF receptors' (***p***= 0.016), 'sprouty regulation of tyrosine kinase signals' (***p***= 0.022), 'role of ERBB2 in signal transduction and oncology' (HER2 pathway; (***p***= 0.026), and 'Erk1/Erk2 MAPK signalling pathway' (***p***= 0.033) (Table 30). Hall mark pathway analysis for down-regulated by BILF1 in DG-75 cells revealed a significant enrichment of pathways related to 'hypoxia' (***p***= 0.019), 'glycolysis' (***p***= 0.019), 'apical junction' (***p***= 0.02), and 'Hedgehog signalling' (***p***= 0.037) (Table 30).

Table 29. Selected canonical and all hallmark pathways enriched in the 132 genes up-regulated by BILF1 in DG-75 cells.

Canonical pathway	***p*-value**
EGFR SMRTE pathway	0.013
CBL pathway	0.016
Cardiac EGF pathway	0.022
SPRY pathway	0.022
TEL pathway	0.022
TFF pathway	0.025
HER2 pathway	0.026
MCALPAIN pathway	0.030
ERK pathway	0.033
EGF pathway	0.037
AT1R pathway	0.040
Hallmark pathway	***p*-value**
TNFA signalling via NFκB	1.35E-12
Hypoxia	6.16E-05
Inflammatory response	0.003176
Apoptosis	0.008663
P53 pathway	0.017729

Table 30. All canonical and hallmark pathways enriched in the 72 genes down-regulated by BILF1 in DG-75 cells

Canonical pathway	***p*-value**
EGFR SMRTE pathway	0.013
CBL pathway	0.016
Cardiac EGF pathway	0.022
SPRY pathway	0.022
TEL pathway	0.022
TFF pathway	0.025
HER2 pathway	0.026
MCALPAIN pathway	0.030
ERK pathway	0.033
EGF pathway	0.037
AT1R pathway	0.040
AGR pathway	0.043
Hallmark pathway	***p*-value**
Hypoxia	0.019
Glycolysis	0.019
Apical junction	0.020
Hedgehog signalling	0.037

4.4.4 Overlap between genes regulated by BILF1 in DG-75 cells and genes differentially expressed in BILF1-positive eBL

I next wanted to compare the extent to which the BILF1-targets I had identified in DG-75 cells were also concordantly differentially expressed in primary eBL. To do this, I re-analysed RNA sequencing data previously published by Abate et al., derived from 20 endemic BL samples (SRA accession PRJNA292327) and RNA sequencing data from 4 germinal centre B cell (GC B) samples (GSE45982) published by Béguelin et al., 2013 (Abate et al., 2015, Béguelin et al., 2013) using EBV annotations I refined as described in section 2.2.20 of the Materials and Methods. The number of mapped sequencing reads for BILF1 in each of these samples is shown in Table 31. I next compared cellular gene expression in the 17 BILF1-expressing eBL with that in the 3 GC B cell samples which do not show BILF1 expression. 4030 genes were up-regulated and 3217 genes were down-regulated in BILF-expressing eBL samples compared to GC B cells (fold change ≥1.5 and $p<0.05$).

I observed no significant enrichment of genes downregulated by BILF1 in DG-75 cells in the genes either up-regulated or down-regulated in BILF1-positive eBL (OR=1.00; $p=$ 0.99 and OR=1.16; $p=$ 0.67), respectively (Figure 39). Furthermore, I observed no significant enrichment of genes up-regulated by BILF1 in DG-75 cells in the genes up-regulated in BILF1-positive eBL cases (OR=1.05; $p=$ 0.84). However, I did observe a significant enrichment of genes up-regulated by BILF1 in the genes down-regulated in BILF1-positive eBL cases (OR=2.72; $p<$ 0.0001) (Figure 39A).

Table 31. Number of reads in endemic Burkitt lymphoma samples mapped to human reference genome (hg19).

Sample	Reads mapped to hg19	Total BILF1 reads	BILF1 reads (CPM)
eBL81	27951602	63	2.25
eBL22	43088688	77	1.79
eBL84	23470600	34	1.45
eBL30	17223517	18	1.05
eBL19	21106346	22	1.04
eBL50	36587265	16	0.44
eBL35	31537781	11	0.35
eBL60	29521251	9	0.3
eBL80	33066993	9	0.27
eBL23	26199978	6	0.23
eBL69	30622512	7	0.23
eBL49	23333219	4	0.17
eBL40	35807489	6	0.17
eBL27	27812284	4	0.14
eBL45	17284461	2	0.12
eBL62	33414770	2	0.06
eBL15	25893798	1	0.04
GCB 4	112690263	1	0.01
eBL20	31850521	0	0
eBL43	25776475	0	0
eBL48	26847694	0	0
GCB 1	110098531	0	0
GCB 2	83663270	0	0
GCB 3	73605144	0	0

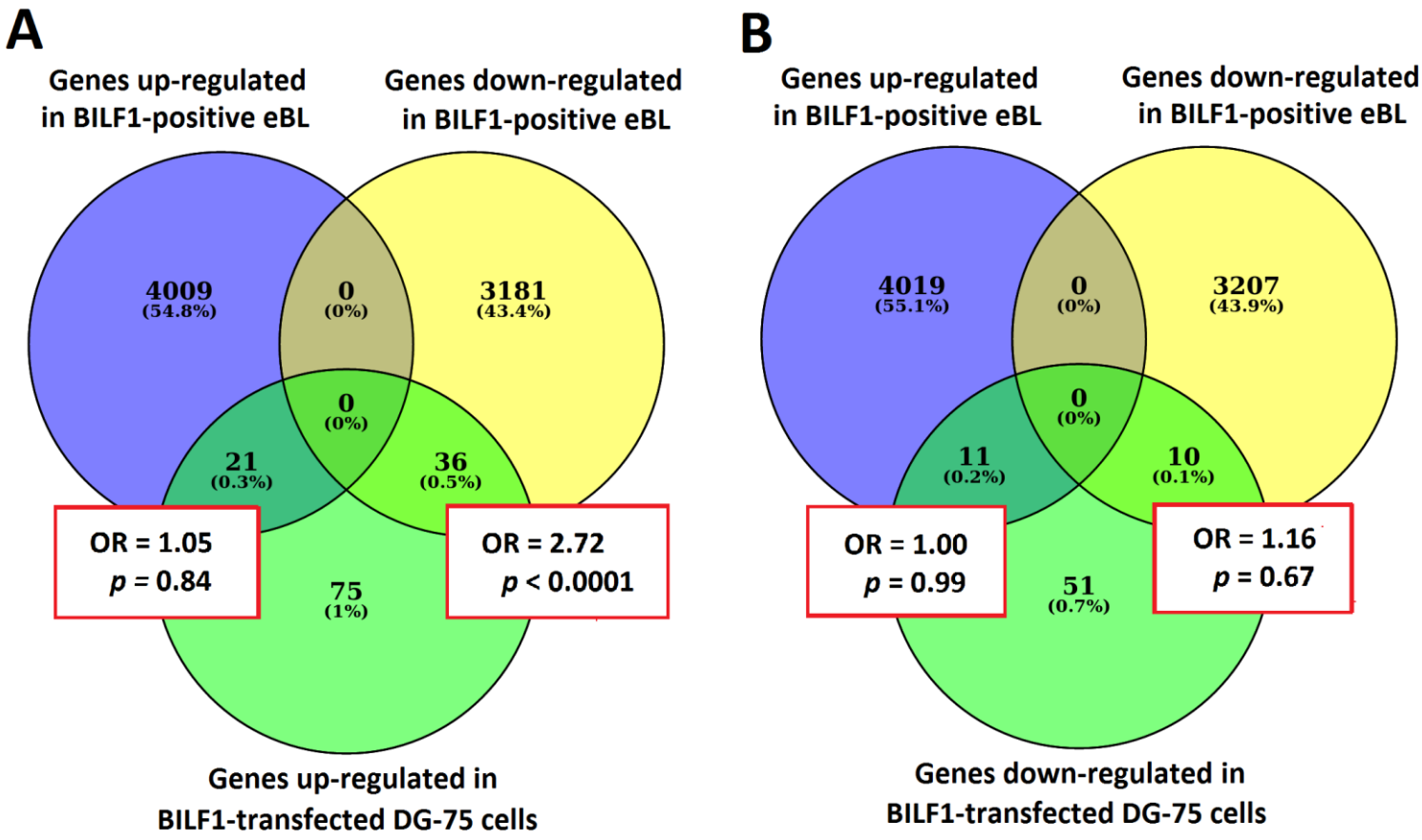

Figure 39. Overlap between the genes regulated by BILF1 in DG-75 cells and BILF1-positive endemic Burkitt lymphoma. A. Significant enrichment of genes up-regulated by BILF1 in DG-75 cells was only observed in genes down-regulated by BILF1 in BILF1-positive eBL. B. No significant enrichment of BILF1-down-regulated genes in DG-75 cells was observed in differentially expressed genes in BILF1-positive eBL.

4.5 BILF1 regulated genes in primary human germinal centre B cells

In light of the disappointing overlap between genes differentially expressed following BILF1 expression in DG75 cells and those concordantly differentially expressed in eBL, I decided to adopt an alternative approach in which primary human germinal centre B cells transfected with BILF1 were subjected to RNA sequencing. Transfection was performed using well established protocols (sections 2.2.13-2.2.17 of the Materials and Methods) previously used to elucidate the transcriptional program of individual viral genes in GC B cells (Vockerodt et al., 2008).

RNA sequencing data obtained for BILF1-transfected and vector-transfected GC B cells was normalized as described earlier. The number of BILF1 gene counts per million reads in BILF1-transfected GC B cells (samples T25-T27) was 9386, 23208 and 8123, respectively. Low levels of BILF1 expression were observed in empty vector transfected GC B samples (11, 4.8, and 6 CPM in samples T25, T26, and T27, respectively).

4.5.1 Genes regulated by BILF1 in primary human germinal centre B cells

463 genes were up-regulated (Appendix 20) and 356 down-regulated (Appendix 21) in BILF1-transfected GC B samples compared to empty vector controls (fold change ≥1.5 and $p<0.05$). A list of 50 genes (25 up, 25 down) with the highest fold change in expression are shown in Table 32.

Table 32. 25 most differentially expressed genes in BILF1-transfected GC B cells.

Genes up-regulated by BILF1 in GC B cells		Genes down-regulated by BILF1 in GC B cells	
Gene	**Fold Change**	**Gene**	**Fold Change**
PNMA2	6.76	IL26	-7.41
LINC01269	6.61	LOC100129924	-6.96
CNIH3	6.33	WDR38	-5.56
TCONS 00029157	5.58	KSR2	-5.12
DOK2	4.77	TRHR	-4.15
FMR1-AS1	4.01	NAALADL2	-4.06
CAMK2A	3.97	GRPR	-4.06
TIAF1	3.74	LOC339807	-3.77
CST7	3.49	LAG3	-3.71
LOC102724919	3.40	CD28	-3.54
SLC37A2	3.40	HIST1H3H	-3.45
GNA14	3.40	ELOVL3	-3.42
SPTLC3	3.39	TMEM59L	-3.34
SLC19A1	3.32	LOC442497	-3.32
CACNG7	3.32	TRPV4	-3.28
WNT1	3.30	LOC102724484	-3.27
NPR1	3.18	ULBP2	-3.25
NACAD	3.10	LOC101060553	-3.23
MTVR2	3.06	JHDM1D-AS1	-3.21
TRAPPC2L	2.91	NOD2	-3.17
RCAN2	2.90	MIR4324	-3.14
GREB1	2.89	SMCO3	-3.12
IL24	2.86	CCNI2	-3.12
CDC14B	2.84	MIR554	-3.05
OSTCP1	2.82	WNK2	-3.04

4.5.2 Gene ontology analysis of BILF1-regulated genes in germinal centre B cells

An ontology analysis revealed that genes up-regulated by BILF1 in GC B cells were enriched for several biological process potentially relevant to lymphomagenesis, including 'oxidative phosphorylation' (GO:0006119; ***p***= 1.81E^{-15}), 'ATP metabolic process' (GO: 0046034; ***p***= 1.45E^{-13}), 'viral transcription' (GO: 0019083; *p*= 1.16E^{-07}), 'viral life cycle' (GO: 0019058; ***p***= 8.27E^{-05}), 'tumour necrosis factor-mediated signalling pathway' (GO: 0033209; ***p***= 0.0011), and 'regulation of B cell receptor signalling pathway' (GO: 0050855; ***p***= 0.0011) (Table 33 and Appendix 22 showing 125 most significantly enriched biological processes). The 356 genes down-regulated by BILF1 in GC B cells, were enriched for 'regulation of phosphorylation' (GO: 0042325; ***p***= 4.01E^{-06}), 'MAPK cascade' (GO: 0000165; ***p***= 7.$^{23E-05}$), 'negative regulation of angiogenesis' (GO: 0016525; ***p***= 3.45E^{-03}), 'lymphocyte aggregation' (GO: 0071593; ***p***= 0.0189), and 'chemokine production' (GO: 0032602; ***p***= 0.0195) (Table 34 and Appendix 23).

Table 33. Selected biological processes enriched in 463 genes up-regulated by BILF1 in GC B cells.

Gene ontology term	*p*-value
GO:0042773~ATP synthesis coupled electron transport	6.53E-16
GO:0006119~oxidative phosphorylation	1.81E-15
GO:0046034~ATP metabolic process	1.45E-13
GO:0019083~viral transcription	1.16E-07
GO:0019080~viral gene expression	2.84E-07
GO:0019058~viral life cycle	8.27E-05
GO:0033209~tumour necrosis factor-mediated signalling pathway	0.0011
GO:0050855~regulation of B cell receptor signalling pathway	0.0374
GO:0042769~DNA damage response, detection of DNA damage	0.0450

Table 34. Selected biological processes enriched in 356 genes down-regulated by BILF1 in GC B cells.

Gene ontology term	*p*-value
GO:0042325~regulation of phosphorylation	4.01E-06
GO:0043549~regulation of kinase activity	6.65E-05
GO:0000165~MAPK cascade	7.23E-05
GO:1901343~negative regulation of vasculature development	1.30E-04
GO:0009967~positive regulation of signal transduction	4.08E-04
GO:0008283~cell proliferation	5.76E-04
GO:0016525~negative regulation of angiogenesis	3.45E-03
GO:0071593~lymphocyte aggregation	0.0189
GO:0032602~chemokine production	0.0195
GO:0051249~regulation of lymphocyte activation	0.0248
GO:0045786~negative regulation of cell cycle	0.0261
GO:0050864~regulation of B cell activation	0.0291
GO:0042094~interleukin-2 biosynthetic process	0.0294
GO:0032660~regulation of interleukin-17 production	0.0352
GO:0070373~negative regulation of ERK1 and ERK2 cascade	0.0437

4.5.3 Pathway analysis of genes regulated by BILF1in GC B cells

I next performed a pathway analysis using gene sets from MSigDB database. Analysis of genes up-regulated by BILF1in GC B cells revealed the enrichment of several pathways, including 'TCA cycle and respiratory electron transport' (***p***= $9E^{-28}$), 'Influenza viral RNA transcription and replication' (***p***= $3E^{-24}$), 'DNA repair' (***p***= $2E^{-4}$), signalling by NOTCH1 (***p***= $2E^{-4}$), 'Vif-mediated degradation of APOBEC3G' (***p***= $7E^{-4}$), 'p53 dependent G1 DNA damage response' (***p***= $9E^{-4}$), 'TRIF mediated TLR3 signalling' (***p***= $3E^{-3}$), and 'G-alpha Q signalling pathway' (***p***= $3E^{-3}$) (Table 35 and complete list in Appendix 24). Hallmark pathway analysis for these genes revealed a significant enrichment in pathways including 'oxidative phosphorylation' (***p***= $4.68E^{-13}$), 'p53 pathway' (***p***= 0.00023), 'MYC target genes' (***p***= 0.0066), and 'mitotic spindle' (***p***= 0.037) (Table 36).

Genes down-regulated by BILF1in GC B cells were enriched for several pathways including 'G-protein coupled receptor ligand binding' (***p***= $4E^{-6}$), and 'chemokine receptors bind chemokines' (***p***= $4E^{-2}$) (Table 37 and complete list in Appendix 25). Hallmark pathway analysis revealed enrichment of several pathways, including 'KRAS signalling up-regulated' (***p***= $4.76E^{-5}$), 'TNFA signalling via NF-κB' (***p***= 0.001), and 'inflammatory response' (***p***= 0.0031) (Table 38).

Table 35. Selected canonical and all Hallmark pathways enriched in 463 genes up-regulated by BILF1 in GC B cells.

Canonical pathway	***p*-value**
TCA cycle and respiratory electron transport	3E-24
Influenza viral RNA transcription and replication	5E-17
DNA repair	2E-04
Signalling by NOTCH1	2E-04
Vif mediated degradation of APOBEC3G	7E-04
P53 dependent G1 DNA damage response	9E-04
Cell cycle checkpoints	2E-03
NFκB and map kinases activation mediated by TLR4 signalling repertoire	2E-03
TRIF mediated TLR3 signalling	3E-03
G alpha q signalling events	3E-03
Traf6 mediated induction of NFκB and MAP kinases upon TLR7 8 or 9 activation	3E-03
G alpha I signalling events	4E-03
G alpha Z signalling events	5E-03
Signalling by ERBB4	5E-03
Viral messenger RNA synthesis	5E-03
P53 independent G1 S DNA damage checkpoint	7E-03
PD1 signalling	9E-03
Signalling by EGFR in cancer	1E-02
Regulation of apoptosis	1E-02
Signalling by the B cell receptor BCR	2E-02
Hallmark pathway	***p*-value**
Oxidative phosphorylation	4.68E-13
P53 pathway	0.000237
Adipogenesis	0.000727
MYC targets V2	0.006656
DNA repair	0.009433
MTORC1 signalling	0.021359
Allograft rejection	0.022619
Unfolded protein response	0.026873
NOTCH signalling	0.027355
Mitotic spindle	0.037863
Apoptosis	0.047205

Table 36. Selected canonical and all Hallmark pathways enriched in 356 genes down-regulated by BILF1 in GC B cells.

Canonical pathway	***p*-value**
GPCR ligand binding	4E-06
Signalling by GPCR	1E-04
Activated TLR4 signalling	2E-03
PI3K cascade	8E-03
GPCR downstream signalling	9E-03
Signalling by PDGF	3E-02
MAP kinase activation in TLR cascade	3E-02
Chemokine receptors bind chemokines	4E-02
Hallmark pathway	***p*-value**
Estrogen response late	4.76E-05
KRAS signalling up	0.00086
TNFA signalling via NFκB	0.000993
Spermatogenesis	0.007149
Inflammatory response	0.012757
Estrogen response early	0.013423
TGF beta signalling	0.030734
Androgen response	0.039133
Interferon gamma response	0.040076
UV response up	0.046988

4.5.4 Overlap between genes regulated by BILF1 in both DG-75 and GC B cells

I next compared the extent to which the genes regulated by BILF1 in DG-75 and GC B cells overlapped. Significant overlap was seen for the genes down-regulated by BILF1 in both DG-75 and GC B cells ($p < 0.0001$), but not in the genes up-regulated by BILF1 in both DG-75 and GC B cells ($p = 0.078$) (Figure 40).

4.5.5 Overlap between genes regulated by BILF1 in GC B cells and genes differentially expressed in BILF1-positive eBL

I next studied the extent to which genes regulated by BILF1 in GC B cells overlapped those differentially expressed in BILF1-positive primary eBL. I observed a significant enrichment of genes up-regulated by BILF1 in GC B cells among those genes up-regulated in BILF1-positive eBL (OR=3.90; $p<0.0001$). Furthermore, I observed a significant enrichment of genes down-regulated by BILF1 in GC B cells among genes up-regulated in BILF1-positive eBL cases (OR=1.67; $p<0.0001$) (Figure 41).

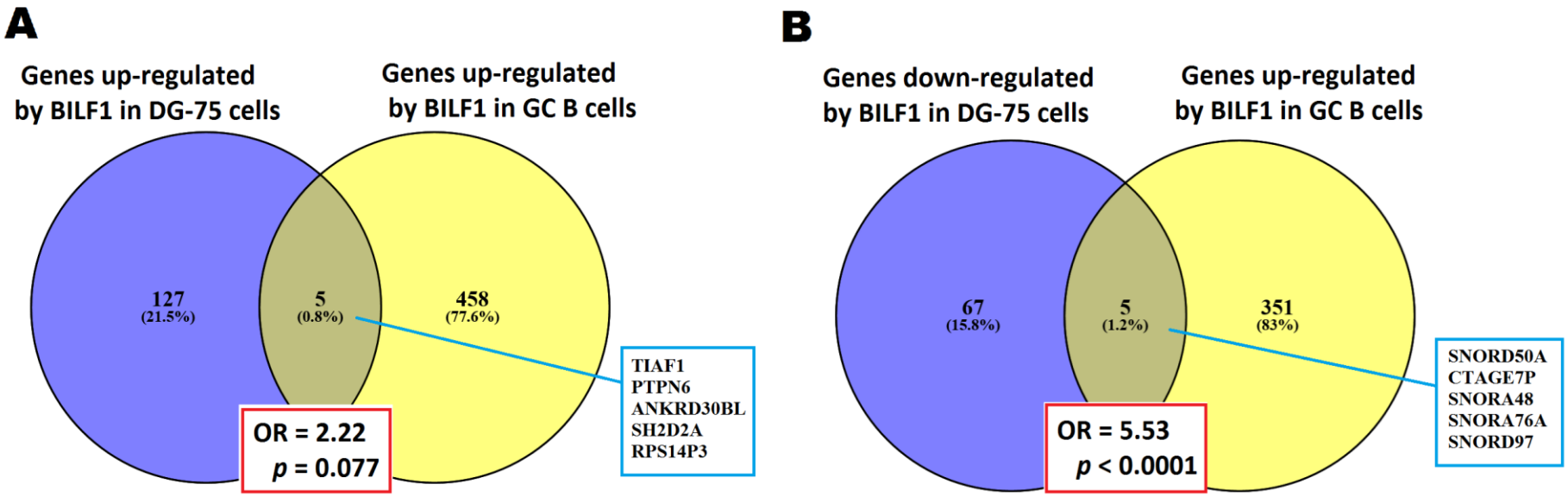

Figure 40. Genes concordantly regulated by BILF1 in DG-75 and GC B cells. A. No statistically significant overlap was observed between the genes up-regulated by BILF1 in DG-75 and GC B samples. B. Overlap between BILF1 down-regulated genes in DG-75 and GC B samples. Statistically significant overlap was observed for genes down- regulated by BILF1 in DG-75 and GC B cells.

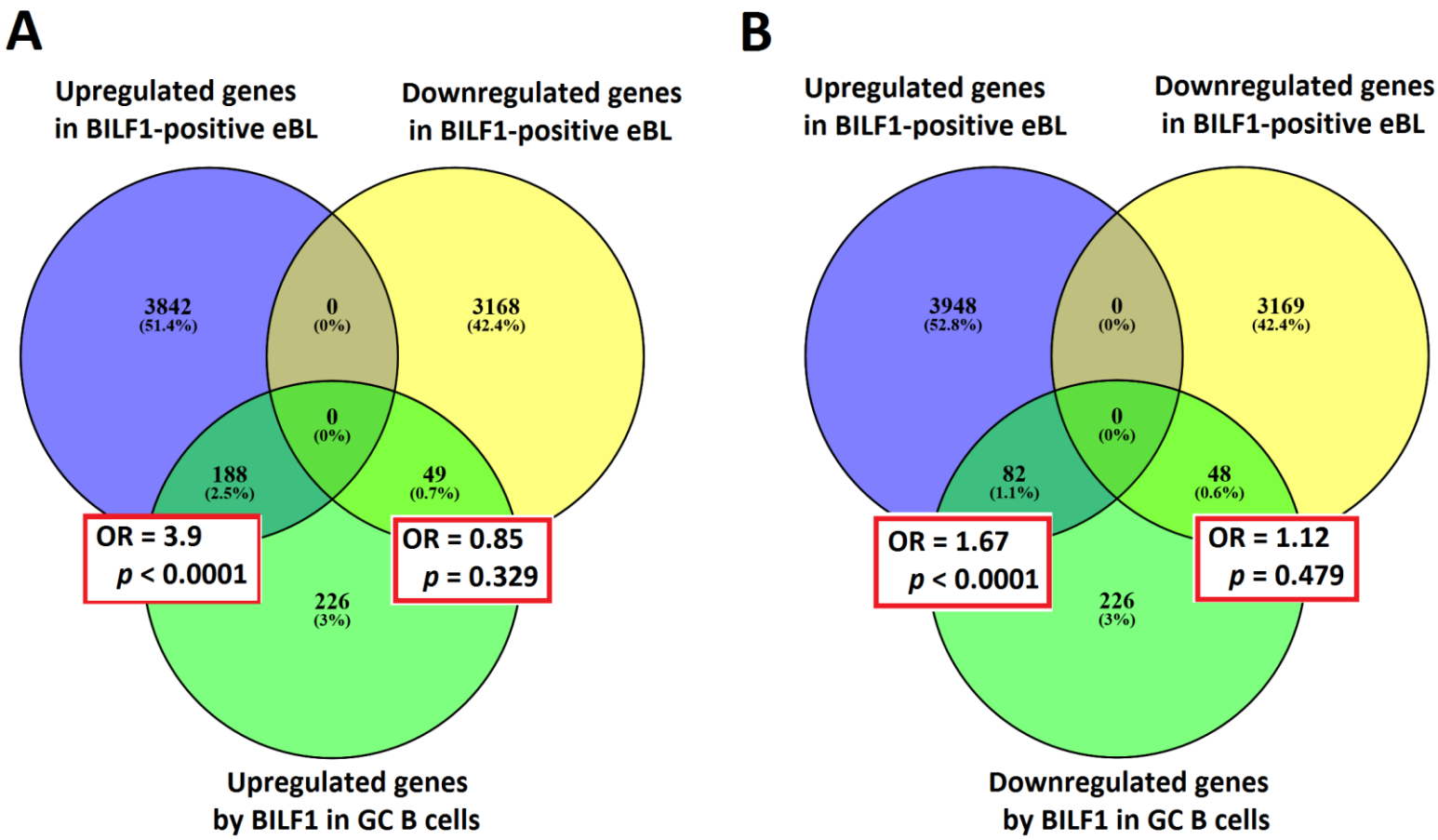

Figure 41. Overlap between the genes regulated by BILF1 in GC B cells and BILF1-positive endemic Burkitt lymphoma. A. Significant enrichment of BILF1-up-regulated genes in GC B cells was observed in the up-regulated genes in BILF1-positive eBL. B. No significant enrichment of BILF1-down-regulated genes in GC B cells was observed in the genes down-regulated in BILF1-positive eBL, however significant enrichment of down-regulated genes by BILF1 in GC B cells was observed in the up-regulated genes in BILF1-positive eBL.

4.6 MYC-target genes are enriched in genes regulated by BILF1 in GCB genes

Having shown that the genes up-regulated by BILF1 were enriched in 'MYC target genes' (***p***= 0.0066), I next wanted to further explore the potential overlap between the transcriptional programmes controlled by BILF1 and MYC. I first used a previously published dataset which had reported the transcriptional changes that follow the expression of MYC in primary GC B cells (Schrader et al., 2012). To further confirm a relationship between the transcriptional programmes of BILF1 and MYC I used a second published dataset which reported the gene expression changes that follow the induction of a tet-regulatable MYC gene in P493-6 cells, a variant of the lymphoblastoid cell line, EREB2-5 (Li et al., 2005).

4.6.1 Overlap between genes regulated by BILF1 in GC B cells and genes regulated by MYC in primary GC B cells

The MYC targets were taken from an experiment in which MYC was over-expressed in primary GC B cells (Schrader et al., 2012) this array platform had 20219 unique genes. There were 19423 genes that were common between the Myc targets in primary GC B cells array platform and the RNA sequencing genome build (total genes 26438).

346 genes were down-regulated by MYC in GC B cells. 340 are on both arrays. All 62 genes up-regulated by MYC were on both arrays. For obvious reasons I excluded MYC from this analysis, leaving 61 up-regulated genes and reducing the number of genes on both platforms to 19422. There were 447 genes up-regulated and 347 genes down-regulated by BILF1 in GC B cells; 396 of the up-regulated genes and 280 of the down-regulated genes were on both arrays.

Four genes were up-regulated in both BILF1-transfected GCB and Myc-transfected primary GC B cells (OR = 3.39; ***p***= 0.0135) (Figure 42A). Four genes down-regulated by BILF1 were also present in genes up-regulated by MYC in primary GC B cells (OR = 4.85; ***p***= 0.0009)

(Figure 42A). Furthermore, two genes were common in genes down-regulated in both data sets (OR = 0.40; ***p***= 0.19) (Figure 42B). Eleven genes up-regulated by BILF1 in GC B cells were also downregulated by MYC in primary GC B cells (OR = 1.62; ***p***= 0.12) (Figure 42B).

4.6.2 Overlap between genes regulated by BILF1 in GC B cells and genes regulated by MYC in P493-6 cells

Li et al used P493-6 cells, a variant of the lymphoblastoid cell line EREB2-5, which is stably transfected with a tetracyclin-regulatable MYC gene.

4115 genes were up-regulated by MYC in P493-6 cells, and 3413 genes were down-regulated by MYC in these cells; 3371 up-regulated genes and 2822 down-regulated genes were on both arrays. As before, I excluded the MYC gene from these analyses, the denominator was therefore 14124, and the number of MYC up-regulated genes, 3370. 339 genes up-regulated and 196 genes down-regulated by BILF1 in GC B cells were present on both arrays.

I observed a significant enrichment of genes up-regulated by BILF1 in GC B cells in the genes up-regulated by MYC in P493-6 cells (OR = 1.64; ***p***< 0.0001) (Figure 43A). No significant enrichment of genes up or down-regulated by BILF1 was observed in the genes down-regulated by MYC in P493-6 cells (OR = 1.19; ***p***= 0.213 and OR = 0.69; ***p***= 0.104) (Figure 43B). I also observed no significant enrichment of genes down-regulated by BILF1 in the genes up-regulated by MYC in P493-6 cells (OR = 0.68; ***p***= 0.085) (Figure 43A).

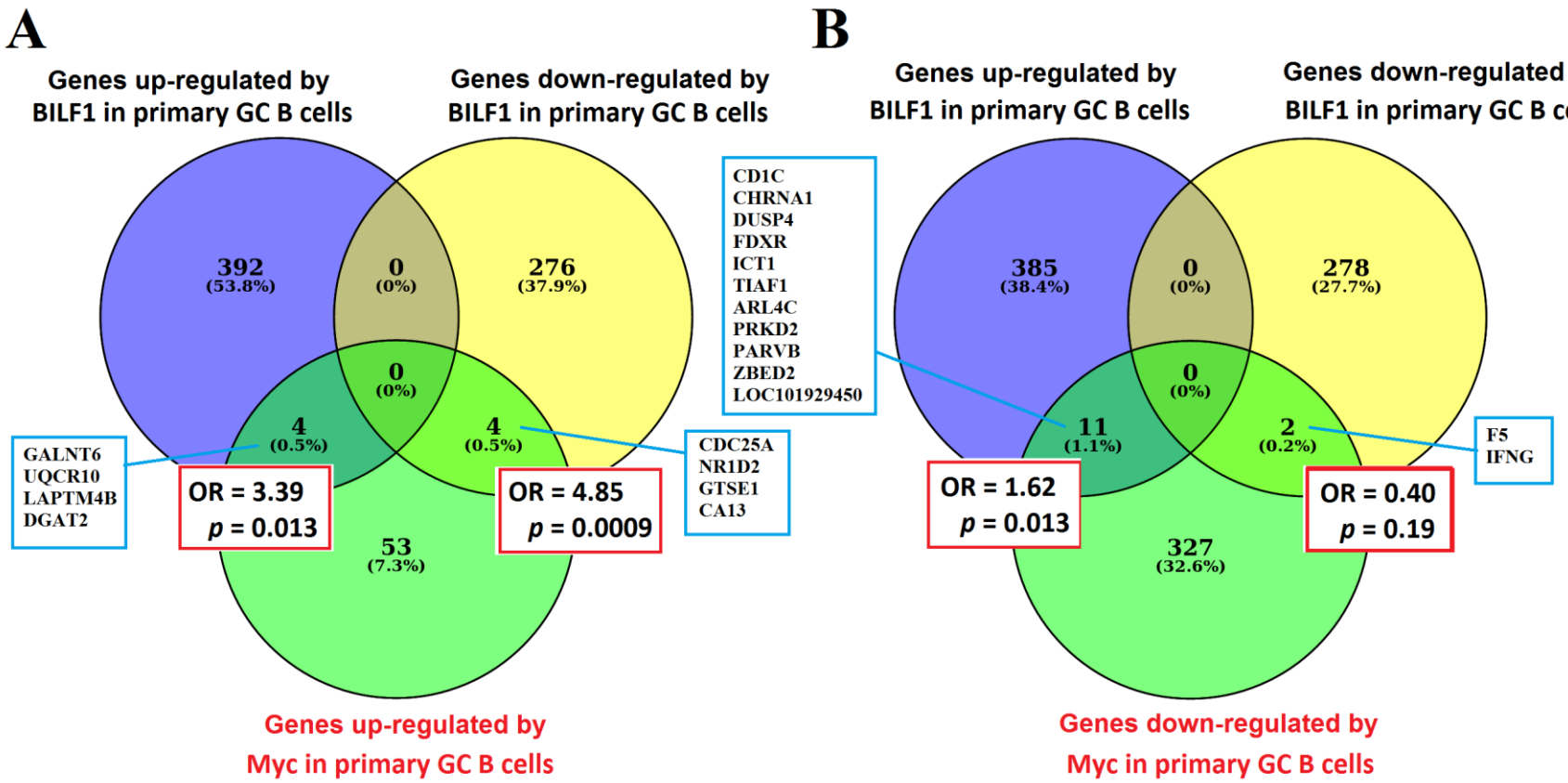

Figure 42. Overlap between differentially expressed genes between BILF1-transfected GCB and Myc-target genes from Shrader et al., 2012. A. Significant overlap of up-regulated genes by Myc in GC B cells was observed in genes regulated by BILF1 in GC B cells. B. However, no significant overlap of down-regulated genes by Myc in the down-regulated genes by BILF1 in GC B cells. Significant enrichment of down-regulated genes by Myc was observed in the up-regulated genes by BILF1 in GC B cells.

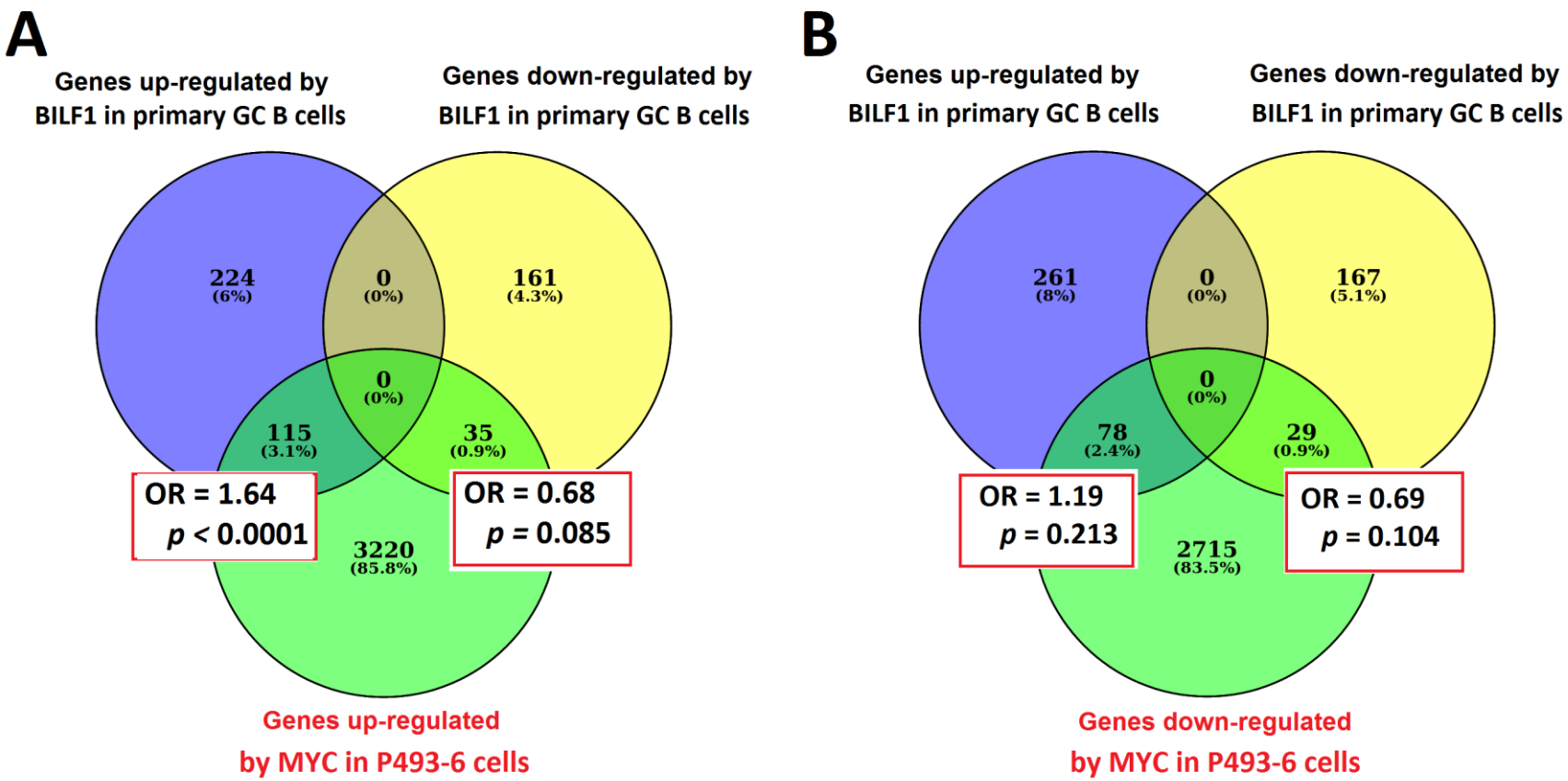

Figure 43. Overlap between differentially expressed genes between BILF1-transfected GCB and Myc-target genes in P493-6.

4.7 Enrichment of metabolic signatures in the genes regulated by BILF1 in GC B cells

One of the striking observations I made during the gene ontology/ pathway analyses of genes was the significant enrichment of genes up-regulated by BILF1 in GC B cells in the genes involved in oxidative phosphorylation (***p***<0.0001). This is of interest because recent studies have shown that the malignant transformation of mutated cells is accompanied by profound alterations in cellular metabolism that fulfil the energy requirements of cancer cell growth and proliferation (Hanahan and Weinberg, 2011). Furthermore, Birkenmeier et al., suggested that the HRS cells increase oxidative phosphorylation capacity by inducing mitochondrial proliferation, thereby promoting cell survival and tumour progression. This group also suggested that NFκB promotes tumour progression and survival by increasing oxidative phosphorylation capacity even in the presence of an optimal glucose supply (Birkenmeier et al., 2016).

To further explore the potential regulation by BILF1 of genes with function in oxidative phosphorylation I re-analysed the transcriptional signatures of 96 metabolic pathways reported by Gaude E and Frezza C et al (Gaude and Frezza, 2016). I performed gene set enrichment analysis on the overlapping gene sets. I found a striking enrichment of genes up-regulated by BILF1 among genes comprising the 'oxidative phosphorylation' signature (OR= 50.24; ***p***<0.0001) (Figure 44A). I also observed weaker but statistically significant enrichment of genes involved in 'Thiamine metabolism', 'Pyrimidine biosynthesis', 'ROS detoxification', 'Cyclic nucleotides metabolism', Folate metabolism', 'Lysine metabolism', 'Glutathione metabolism', and 'Lysosomal transport' among genes up-regulated by BILF1.

I also found significant enrichment of 'glycolysis and gluconeogenesis' signature (OR=3.25; ***p***=0.0367) in the genes down-regulated by BILF1 in GC B cells (Figure 44B).

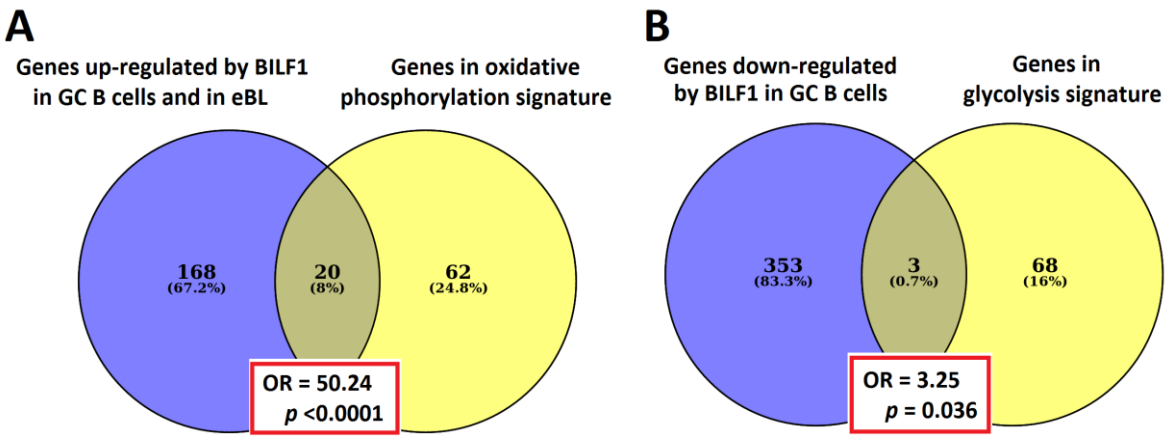

Figure 44. Metabolic signatures enriched in genes regulated by BILF1 in GC B cells. Significant enrichment of genes in oxidative phosphorylation signature was observed in the genes up-regulated by BILF1 in GC B cells and eBL (A). Glycolysis signatures were enriched in genes downregulated by BILF1 in GC B cells (B).

4.8 DISCUSSION

The majority of the studies that have investigated the contribution of individual EBV genes have focussed almost entirely on the latent genes. However, more recently it has become apparent that some genes formerly thought to be expressed only during EBV lytic cycle and may also be expressed during latency. The prototypic example of this is the viral encoded BHRF1 gene, which is a homolog of the cellular anti-apoptotic gene, BCL2. BHRF1 has been shown to be expressed during latency in a subset of Burkitt lymphomas carrying an EBV genome with an EBNA2 deletion and in which viral latent transcripts, including BHRF1, are expressed from the alternative latency promoter Wp. BHRF1 has been shown to provide protection from apoptosis in BL cells (Kelly et al., 2009a).

The high-throughput qPCR analysis showed high levels of expression of a number of so called lytic cycle genes, these included BILF1, BALF4 and less well characterized RNAs LF1, LF2, and LF3. LF1, LF2, and LF3 were poorly characterized as they were deleted in the B95-8 prototypic strain. Previous studies on the function of BALF4 (gp 100 precursor) reported that BALF4 gene is not required for DNA lytic replication and lytic protein production (Neuhierl et al., 2009). However, BILF1, another gene originally thought to be exclusively expressed in the virus lytic cycle, was previously reported to be important for viral pathogenesis (Beisser et al., 2005, Nijmeijer et al., 2010a, Zuo et al., 2009)

BILF1 is a member of a distinct GPCR-like family whose sequences and genomic positions are conserved among all of the currently known γ-1-herpesviruses and in a subset of γ-2-herpesvirus species (Beisser et al., 2005). BILF1-family genes are positioned at the 5’ end of a conserved gene block that includes the viral DNA polymerase gene, except the ORF74-like genes that are positioned near another conserved gene block that includes the large tegument protein gene (Figure 45). Beisser et al., showed that BILF1-encoded protein increases CRE-mediated transcription in EBV-positive Burkitt lymphoma cells and lymphoblastoid B cells

transfected with BILF1 through Gi/Go coupling. They also showed a reduction in the levels of phosphorylated RNA-dependent antiviral protein kinase R (PKR) in COS-7 and Burkitt lymphoma cells transfected with BILF1. These data suggest that BILF1 encodes a constitutively active GPCR capable of modulating various intracellular signalling pathways, as the above observed effects do not require a ligand to interact with the BILF1 gene product (Beisser et al., 2005, Nijmeijer et al., 2010a).

Paulsen, S.J., et al., showed that BILF1 is heavily glycosylated protein (~50 K-Da) localized on the plasma membrane similar to KSHV-encoded GPCR, ORF74, suggesting that BILF1 and ORF74 have similar functions during γ-herpesvirus replication. They also showed that BILF1 inhibited forskolin-triggered CREB activation via stimulation of endogenous G proteins thus suggesting that EBV may use BILF1 to regulate Gi-activated pathways during viral lytic replication, thereby affecting disease progression (Paulsen et al., 2005). Canonical pathways analysis on the up-regulated genes by BILF1 in GC B cells in my study also showed enrichment for genes involved in 'Gastrin-CREB signalling pathway via PKC and MAPK' ($\boldsymbol{p}$=0.005), supporting the previously reported functions of BILF1.

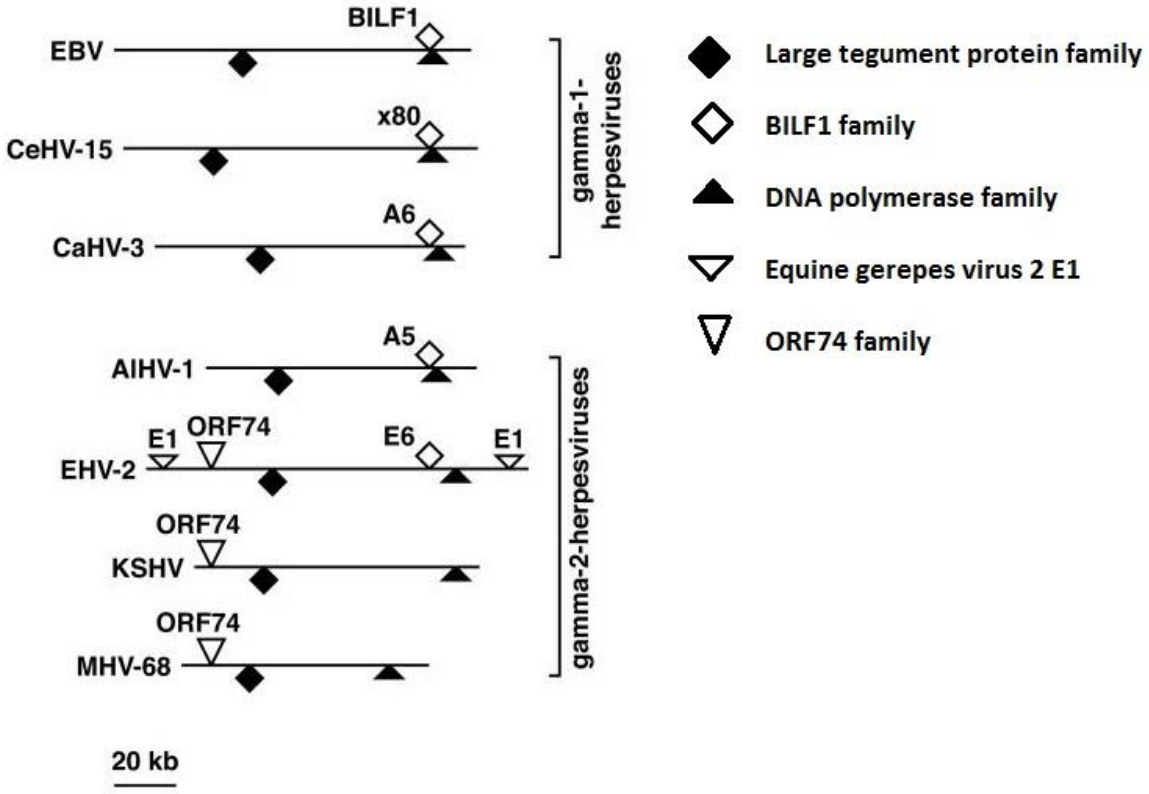

Figure 45. BILF1-like genes conserved in γ-1 and 2 herepesviruses. EBV BILF1 and KSHV ORF74 belong to different viral GPCR gene families. The lines represent herpesvirus genomes. The symbols indicate gene positions, which were derived from GenBank accession nos. AY037858, NC 001345, NC 001650, NC 001826, NC 002531, NC 003409, and NC 004367. Open symbols represent viral GPCR genes. The names of viral GPCR genes are indicated above their corresponding symbols. Genomes and gene positions are drawn to scale. EBV, Epstein-Barr virus; CeHV15, cercopithecine herpesvirus 15; CaHV-3, callitrichine herpesvirus 3; AlHV-1, alcephaline herpesvirus 1; EHV2, equine herpesvirus 2; KSHV, Kaposi's sarcoma-associated herpesvirus; MHV-68, murine gammaherpesvirus 68 **(**This figure is reproduced from Beisser et al., 2005)**.**

Vischer et al., demonstrated the ability of BILF1 to hetero-dimerize with chemokine receptors CXCR4, CXCR3, CCR9, and CC10 altering the responsiveness of B lymphocytes to chemokines there by altering homing and homeostasis of infected B lymphocytes and might be essential for EBV dissemination and/or be involved in EBV-induced pathogenesis (Vischer et al., 2008). The chemokine receptor CXCR4 is expressed on tumour cells, including B cell lymphoma and EBV-associated nasopharyngeal carcinoma, and plays a key role in B cell homing (Kunkel and Butcher, 2003). CXCR4 expression allows tumour cells to metastasize toward CXCL12-secreting organs (Hu, 2005, Niu et al., 2012). BILF1 forms hetero-oligomeric complexes with CXCR4 and inhibits CXCL12 binding to this receptor and attenuates migration of BILF1-expressing plasma B cells to CXCL12 gradients allowing these cells to home and/or remain at sites that are most optimal for viral replication and dissemination (Nijmeijer et al., 2010b).

Previous studies have shown the ability of EBV to evade immune mechanisms during lytic replication by targeting the MHC class I antigen processing pathway at the level of translation through BGLF5 (Rowe et al., 2007)and peptide transport through BNLF2a (Hislop et al., 2007b). Several studies have shown that BILF1 can down-regulate surface HLA class I expression and thus preventing the presentation of viral antigens and cytotoxic T cell recognition of the infected cell (Zuo et al., 2009, Griffin et al., 2013). BGLF5 and BNLF2a are both products of early lytic cycle genes (Zuo et al., 2009, Yuan et al., 2006). Previous studies showed conflicting views on the early/late status of BILF1 (Paulsen et al., 2005, Beisser et al., 2005, Yuan et al., 2006), however present study indicates that BILF1 can also be expressed during latent infection in B cells. The high-throughput qPCR revealed a high level of BILF1 expression in BL samples, whereas levels of BZLF1 expression were substantially lower. BZLF1 is required for the entry of B cells into the viral lytic cycle and is therefore a good marker of cells undergoing the replicative cycle. The possibility that BILF1

is expressed in virus latency is supported by the observation that Akata cells lacking BZLF1 contained low levels of BILF1 transcripts (in the RNA in situ hybridization experiment). Furthermore, it has been shown that BILF1 is expressed in lymphoblastoid cell lines generated by infection of B cells with a recombinant virus lacking a BZLF1 gene (Rowe and Zuo, 2010). The RNAScope on primary tissues provided confirmation of BILF1 expression in the majority of tumour cells in EBV positive cases, which might be expected were BILF1 a latent gene. Taking together these observations suggest that BILF1 is expressed during virus latency in B cells.

Because BILF1 is a constitutively active GPCR, I reasoned that its effects are likely to be manifest in altered transcription. For this reason I studied the transcriptional changes that followed BILF1 expression. My initial experiments were focussed on a well utilized model, DG-75, an EBV-negative BL cell line. However, comparison of those transcriptional changes that followed by BILF1 expression in DG-75 with those observed when BILF1-positive eBL were compared with normal GC B cells yielded no significant overlap in concordantly expressed genes. With hindsight I realized that DG-75 may not have been the most appropriate model to identify BILF1-transcriptional targets as it is a fully transformed cell line, presumably having acquired the necessary changes without the requirement for EBV or BILF1. Therefore, I next studied the transcriptional changes that followed the expression of BILF1 in normal primary human germinal centre B cells. Our group has extensively used these cells as a model to identify the transcriptional consequences of the over-expression of viral cellular genes relevant to lymphomagenesis. This model would seem to be particularly appropriate to the study of BL, which is a tumour of the GC stages of B cell differentiation. In contrast to the DG-75 system I found that BILF1 induced more transcriptional changes in primary GC B cells many of which overlapped with genes differentially expressed in BILF1-positive eBL.

The pathway analysis using Hallmark gene sets showed that genes up-regulated by BILF1 in GC B cells were enriched in MYC-targets. This was of particular interest because MYC is over-expressed in BL as a consequence of the consistent presence of one of three translocations involving the c-Myc gene on chromosome 8. Furthermore, I also observed a striking overlap between genes up-regulated by BILF1 and those up-regulated by MYC in two different experimental settings; one in which MYC had been over-expressed in GC B cells and another in which MYC expression was induced in the P493-6 lymphoblastoid cell line. However, MYC itself was not itself apparently a transcriptional target of BILF1 in either DG-75 or GC B cells, so at this time it is unclear if MYC expression is modified directly by BILF1. Alternatively, MYC and BILF1 may engage similar cell signalling pathways resulting in similar transcriptional end points. In this respect, it is interesting to note that MYC is a potent inducer of oxidative phosphorylation.

Previous studies have established the link between MYC and the upregulation of glucose metabolism. In particular, one of its target genes, lactate dehydrogenase A (LDH-A) is known to convert pyruvate through glycolysis (Shim et al., 1997). Subsequent studies have reported that several other genes involved in glucose metabolism are directly regulated by MYC, including glucose transporter GLUT1, hexokinase 2 (HK2), phosphofructokinase (PFKM), and enolase 1 (ENO1) (Kim et al., 2004, Kim et al., 2007, Osthus et al., 2000). An interplay between MYC and HIF (hypoxia inducible factor) is also well documented (Figure 16) where glycolytic genes may be regulated by both transcription factors (Dang et al., 2009).

MYC and HIF-1 regulate genes involved in glycolysis (glucose transporter Glut1, hexokinase 2 (HK2), pyruvate kinase M2 (PKM2), lactate dehydrogenase A (LDHA), and pyruvate dehydrogenase kinase 1 (PDK1)), converting glucose to lactate (Figure 46). Myc can also stimulate glutamine metabolism through the regulation of transporters (SLC1A5) and glutaminase (GLS), where glutamine is converted to α-ketoglutarate (α-KG) for catabolism

through the tricarboxylic acid (TCA) cycle to malate, which is transported into the cytoplasm and converted to pyruvate and then to lactate (glutaminolysis) (Dang et al., 2009).

The pathway analysis of genes up-regulated by BILF1 in my study showed a significant enrichment of genes involved in the 'HIF pathway' (***p***=0.003), 'Regulation of hypoxia inducible factor HIF by oxygen' (***p***=0.02), 'EPO signalling pathway' (***p***= $9E^{-08}$), and 'TCA cycle and respiratory electron transport' (***p***= 9E-28), along with 'MYC target genes' (***p***= 0.0066).

The pathway analysis of genes up-regulated by BILF1 in GC B cells showed a significant enrichment of genes involved in oxidative phosphorylation (***p***<0.0001). Nearly a century ago Otto Warburg first described the dysregulation of cellular metabolism in cancer cells, where metabolism of cancer cells relies mostly on glycolysis even in the presence of oxygen, whereas normal cells fully oxidize glucose in the mitochondria (Warburg, 1924). Recent advances in the technologies that enable a more comprehensive examination of the intricacies of cancer metabolism show that the metabolic reprogramming of cancer goes beyond activation of glycolysis (Hu et al., 2013). Furthermore, a recent study on metabolic profiling of NCI-60 cell lines found that increased glycine uptake strongly correlates with proliferation rates of cancer cells, suggesting an increased requirement of building blocks for nucleotide biosynthesis in cancer cells (Jain et al., 2012). Genomic (Monti et al., 2005, Park et al., 2016) and metabolomic (Caro et al., 2012) studies on DLBCL showed enrichment of genes involved in oxidative phosphorylation in OxPhos-DLBCL subset.

Gene set enrichment analysis of transcriptional signatures of 96 metabolic pathways in the genes up-regulated by BILF1 in GC B cells and up-regulated genes in BILF1-positive eBL revealed a striking enrichment of genes up-regulated by BILF1 among genes comprising the 'oxidative phosphorylation' signature (OR= 50.24; ***p***<0.0001) (Gaude and Frezza, 2016).

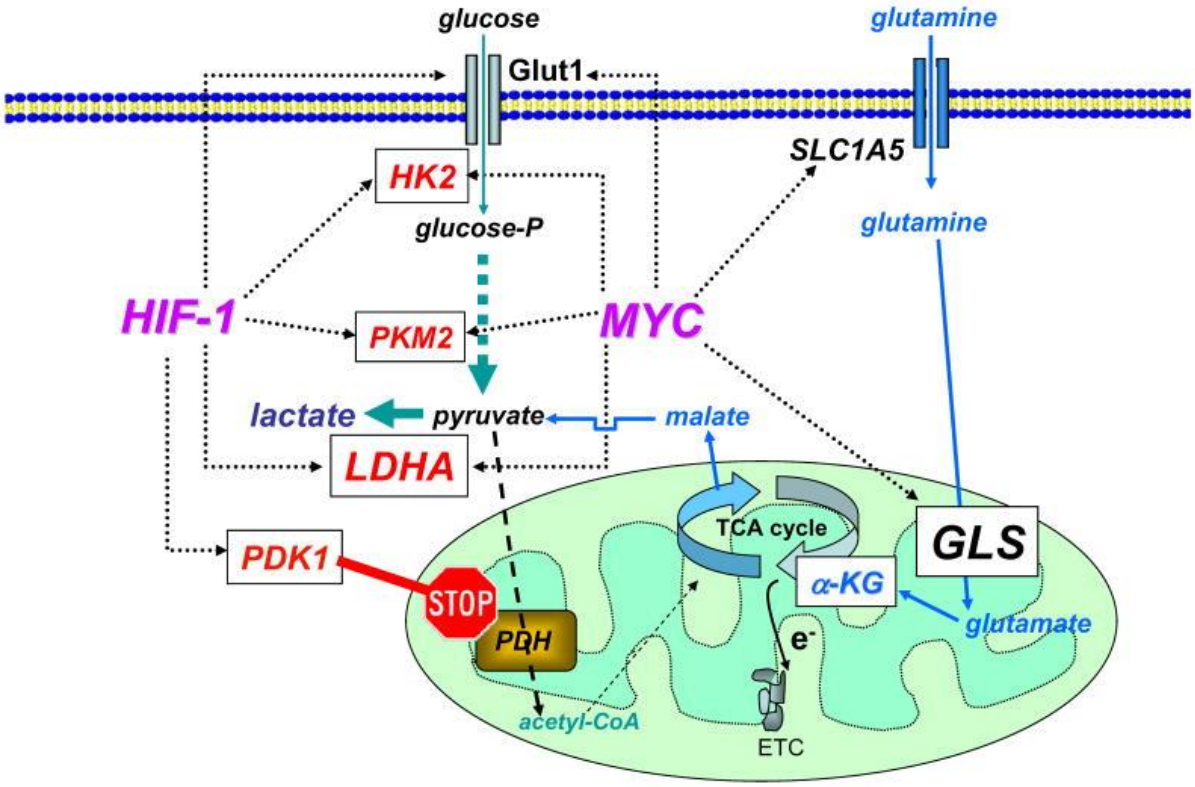

Figure 46. Myc and HIF-1 regulate glucose metabolism. Myc and HIF-1 are shown to regulate (dotted lines) genes involved in glucose metabolism (glucose transporter Glut1, hexokinase 2 (HK2), pyruvate kinase M2 (PKM2), lactate dehydrogenase A (LDHA), and pyruvate dehydrogenase kinase 1 (PDK1)), favouring the conversion of glucose to lactate (glycolysis). Myc is also shown to stimulate glutamine metabolism through the regulation of transporters (SLC1A5) and glutaminase (GLS). Glutamine is shown converted to α-ketoglutarate (α-KG) and then to malate through the tricarboxylic acid (TCA) cycle. Malate is then transported into the cytoplasm and converted to pyruvate and then to lactate (glutaminolysis). PDH = pyruvate dehydrogenase. PDH = pyruvate dehydrogenase (reproduced from Dang CV. et al, 2009).

CHAPTER 5

CONCLUSIONS

AND FUTURE WORK

5. CONCLUSIONS AND FUTURE WORK

This thesis has explored the pathogenesis of two childhood B cell lymphomas, Hodgkin lymphoma (HL) and Burkitt lymphoma (BL). Both of these tumours present particular clinical problems. In the case of HL, cure rates, especially in children, are high. However, toxicity associated with the use of combination chemotherapy and radiotherapy is problematic, resulting in excess morbidity and mortality in HL survivors. Furthermore, there are no biomarkers available which can detect the small proportion of HL patients, who will either relapse or have refractory disease. This is mainly because the biological factors that underpin the more aggressive forms of HL are largely unknown.

Deficiencies in our understanding of the biology of HL are in part due to a paucity of knowledge of the genetic abnormalities that characterize this tumour. This is in turn a consequence of the low abundance of tumour cells in affected tissues which is compromised genetic studies. For this reason to date there has only been one study which has reported the global genome of the HRS cell, this study included only ten cases and importantly relied on the isolation of the tumour cell population by flow sorting (Reichel et al., 2015). This is not an approach that can be employed for the analysis of archived tumour samples and for this reason it is unlikely to be a practical solution to the analysis of HRS genomes in retrospective studies, including samples from clinical trials.

In the first results chapter, I present an entirely new approach to the analysis of HRS cell exomes based upon the amplification of DNA from HRS cells microdissected from snap frozen tumour tissues. This analysis allowed the description of the exomes of 13 cases of paediatric HL. As might have been anticipated from the analysis of only 150 cells per case,

coverage was somewhat lower than has been reported for previous exome studies in other tumour settings using many more cells as a starting point. This low coverage was presumably responsible for a somewhat lower sensitivity. However, the sequencing of two separate biological replicates from each patient together with the use of stringent variant calling and filtering criteria ensured a high degree of specificity reflected in the high concordance between exome sequencing and Sanger sequencing in the validation steps. Work ongoing in the laboratory is investigating approaches to improve the sensitivity of the exome sequencing including the microdissection of more tumour cells. Furthermore, although, my initial experiments with cells microdissected from Formalin Fixed Paraffin Embedded (FFPE) samples yielded smaller amplification products than obtained from frozen material, it is not clear if these would be suited to exome sequencing and also requires further work.

My analysis of the paediatric exomes revealed some novel insights into the biology of HL. First, I observed striking heterogeneity between samples reflected by the relatively small numbers of recurrently mutated genes. However, a pathway analysis revealed that many of the affected genes converge on a relatively small number of pathogenic pathways, highlighting the importance of considering pathway level analysis, rather than simply gene level analysis, when exploring whole exome data especially in tumours that display considerable heterogeneity at the gene level. In particular, I observed enrichment of genes with mitotic spindle functions among genes with protein-altering mutations. This is of particular importance given that HRS cells are characterized by chromosomal instability. Furthermore, the mitotic spindle checkpoint is a target of a number of chemotherapeutic drugs that are used routinely to treat patients with HL (Dominguez-Brauer et al., 2015, Hande, 2008). At present it is expected that once optimized the approach adopted in this thesis could be applied to a larger cohort of patients with paediatric HL such as those that will be enrolled to EuroNet-PHL-C2 (European Network-Paediatric Hodgkin Lymphoma Study

Group Second International Inter-Group Study for Classical Hodgkin Lymphoma in Children and Adolescents) trial about to open in the UK (ClinicalTrials.gov identifier: NCT02684708). This international, multicentre, randomised controlled trial aims to reduce the indication for radiotherapy in newly diagnosed patients with cHL without compromising cure rates. It also aims to investigate a chemotherapy intensification randomisation in intermediate and advanced cHL to compensate for reduction in radiotherapy. Given that in this trial, patients are assessed by PET scan after two cycles of chemotherapy, the genomic analysis of this cohort would provide an ideal opportunity to explore the impact of the HRS genome on chemotherapy responses.

A second important observation was that mutations were significantly more frequent in EBV-negative HL cases than in EBV-positive HL. This observation is in keeping with other studies of viral associated malignancies including EBV-positive and EBV-negative gastric cancers and HPV-positive and HPV-negative head and neck cancers (Network., 2014, Riaz et al., 2014), where it has been shown that viral negative forms of the same cancer have more mutations than viral positive forms. A further striking observation in head and neck cancer is that patients with HPV-positive cancer have significantly longer survival than patients with HPV-negative tumours. Analysis of the impact of EBV on survival in paediatric HL shows a significantly longer progression-free survival for patients with EBV-positive disease compared with those with EBV-negative disease (Powell et al., unpublished work). However, it remains to be seen if these survival differences are a consequence of the lower mutation rates observed in virus-positive tumours.

I also observed that EBV-negative HL is characterized by higher frequency of mutations in mitotic spindle associated genes compared with EBV-positive HL. I found that EBV could transcriptionally downregulate mitotic spindle associated genes, suggesting that the inactivation of the mitotic spindle is likely to be critical for the development of HL, and also

that it could be achieved by distinct mechanisms; transcriptional downregulation in the presence of EBV, and by mutations in the absence of EBV. These mechanisms are unlikely to be entirely mutually exclusive.

The molecular biology of BL, including the EBV-positive endemic forms, is already well described. So, in the case of BL, I turned my attention to a relatively unexplored area, that of viral gene expression and its potential contribution to BL pathogenesis. It is already been shown that the EBV-encoded BHRF1 gene, formerly thought to be exclusively expressed during the viral lytic cycle, is a *bona fide* latent gene. Therefore, I considered it likely that BHRF1 was not the only lytic cycle gene to be expressed in latently infected BL cells. Using a combination of high-throughput qPCR (Fluidigm), the re-analysis of RNA sequencing data, and RNAScope experiments, I showed BILF1 is expressed in a subset of primary eBL, where it appears to be present in most if not all of the tumour cells.

The over-expression of BILF1 in primary GC B cells revealed that BILF1 could partially recapitulate the transcriptional programme of primary BL. Gene set enrichment analysis of BILF1 target genes suggested a potential role of BILF1 in oxidative phosphorylation. Future studies will explore the extent to which BILF1 can re-programme the metabolome of B cells and how this might contribute to the pathogenesis of BL. I also observed striking overlap between the transcriptional programme induced by BILF1 and that induced by MYC. However, I did not observe that MYC was a transcriptional target of BILF1 in GC B cells and so several other possibilities remain to explain this overlap, including the potential regulation of BILF1 by MYC.

My observations are consistent with a potent role of BILF1 and pathogenesis of lymphoma, which is different from its established role in the immune escape of lytically infected cells mediated by the downregulation of surface HLA class I (Zuo et al., 2009). These effects are likely to be mediated by the functions of BILF1 as a constitutively active GPCR. Future

studies will investigate the signalling pathways engaged by BILF1 in B cells and the extent to which these can explain the transcriptional effects observed in this study. Furthermore, as a GPCR BLF1 is an attractive therapeutic target. Once important phenotypic effects of BILF1 in B cells are confirmed, then the potential therapeutic targeting of BILF1 could be explored. It is noteworthy that the primary observation made here that BILF1 is also expressed in a subset of NPC is now being followed up in collaboration with colleagues in Southeast Asia and a manuscript is already in preparation. In addition to BILF1, I also identified a number of other highly expressed 'lytic cycle' genes about which a little or nothing is known, including the LF family These are obvious candidates to be taken forward in future studies.

Future work

The work presented here suggests several specific avenues for future study. A subset of recurrently mutated genes in my series and/or in Reichel and Liu series were selected to perform targeted sequencing (Tables 37 and 39) in a larger cohort of paediatric HLs. Additionally, we also selected a subset (Table 39) of mitotic spindle associated genes mutated in paediatric HLs that were also downregulated by EBV in GC B cells (Leonard et al., 2011)and /or downregulated in HRS cells (Brune et al., 2008) for validation by q-PCR and IHC methods.

Furthermore, a subset of MYC-target genes (Schrader et al., 2012) that are also differentially expressed in GC B cells regulated by BILF1 (Table 38) have been selected for validation by high-throughput q-PCR (Fluidigm) method in a larger cohort of BL samples. I am also designing an experiment to obtain metabolomics data from a larger cohort of BL samples for a subset of genes involved in oxidative phosphorylation that are upregulated by BILF1 in GC B cells and that are also upregulated in BILF1-positive eBLs (Table 40).

I am currently preparing manuscripts for publication which will include these new data.

Table 37. Genes with protein-altering mutations selected for targeted sequencing.

Gene symbol	Number of samples mutated in combined Reichel and Liu series (n=17)	Number of samples mutated in paediatric HL (n=13)
TNFAIP	9	1
HIST1H	8	0
B2M	7	0
CSF2RB	7	0
DNAH1	7	0
IGLL5	7	0
ITPKB	7	0
PCLO	7	0
SOCS1	7	0
FAT4	6	0
GPR112	6	0
GPR98	6	1
IGHE	6	0
LRRN3	6	0
MUC16	6	1
ABCA1	5	0
ACTB	5	0
C2orf16	5	0
DNAH1	5	0
FAT1	5	0
SMG1	0	5
TPTE2	0	4
OR4C5	1	4
ZNF717	1	3
SREK1	1	3
SGOL2	0	3
SF3B2	0	3
PROM1	0	3
HCFC2	1	3
F8	1	3
F5	1	3
CSMD3	3	3
CCDC8	0	3
RANBP	4	2
ITPR2	4	2
USP15	0	2
UGT1A	0	2

Table 38. Genes selected for metabolomics profiling in a larger cohort of BL samples.

ATP5E	COX5B	COX6C	NDUFA1	NDUFB2	NDUFB10	NDUFA13
ATP5G1	COX6A1	COX7B	NDUFA7	NDUFB7	NDUFS8	NDUFA11
ATP5I	COX6B1	COX7C	NDUFA8	NDUFB9	UQCR10	

Table 39. Selected mitotic spindle associated genes mutated in paediatric HLs for validation by q-PCR and IHC methods.

Mitotic spindle associated genes mutated in paediatric HLs	Downregulated by EBV infection (Leonard et al., 2011)	Downregulated in HRS cells (Brune et al., 2008)	Validation by IHC
AGO4	AGO4	AGO4	AGO4
ALMS1	ALMS1	ALMS1	ALMS1
ATM		ATM	
BIRC6	BIRC6		
CDK5RAP2	CDK5RAP2		
CEP57	CEP57	CEP57	CEP57
CSE1L		CSE1L	
CTCF	CTCF	CTCF	CTCF
ECT2	ECT2	ECT2	ECT2
EFHC1	EFHC1		
NCOR1		NCOR1	
NIN		NIN	
PDS5B	PDS5B	PDS5B	PDS5B
RANBP2	RANBP2	RANBP2	RANBP2
RB1		RB1	
RMDN2	RMDN2		
SGOL2	SGOL2		SGOL2
SMC1A	SMC1A	SMC1A	SMC1A
TOP2B		TOP2B	
TOPBP1	TOPBP1	TOPBP1	TOPBP1
XPO1	XPO1		

Table 40. MYC-target genes regulated by BILF1 in GC B cells for validation.

Genes regulated by MYC in GC B cells	Genes regulated by BILF1 in GC B cells
CAMKK1	UP
CANX	UP
CD320	UP
CHID1	UP
COX6A1	UP
MBNL2	DOWN
NME1	UP
PDE4A	UP
PFDN2	UP
PHB	UP
PRDX1	UP
PSMD2	UP
TRIAP1	UP
TXN	UP

CHAPTER 6

APPENDIX

6. APPENDIX

Appendix 1: Paediatric HL samples submitted for exome sequencing

Sample	Exome capture method	260/280	260/230	Quantity (ng)
S1 Normal	Nextera	1.78	2.2	400
S1 1 Tumour (Sigma)	Nextera	1.84	2.46	400
S1 2 Tumour (Sigma)	Nextera	1.87	2.65	400
S2 Normal	Nextera	1.9	2	400
S2 1 Tumour (Sigma)	Nextera	1.85	2.15	400
S2 2 Tumour (Sigma)	Nextera	1.95	2.49	400
S3 Normal	Nextera	1.86	2.3	400
S3 1 Tumour (Sigma)	Nextera	1.82	2.36	400
S3 2 Tumour (Sigma)	Nextera	1.89	2.7	400
S4 Normal	Nimblegen	2.01	2.28	2270
S4 1 Tumour (Sigma)	Nextera	1.9	2.24	400
S4 2 Tumour (Sigma)	Nextera	1.88	2.4	400
S5 Normal	Nextera	1.86	2.4	400
S5 1 Tumour (Sigma)	Nextera	1.85	2.4	400
S5 2 Tumour (Sigma)	Nextera	1.92	2.6	400
S6 Normal	Nimblegen	1.92	2.37	2145
S6 1 Tumour (Sigma)	Nimblegen	1.96	2.18	1135
S6 2 Tumour (Sigma)	Nextera	1.9	2.6	400
S7 Normal	Nextera	2	2.3	400
S7 1 Tumour (Sigma)	Nextera	1.83	2.5	400
S7 2 Tumour (Sigma)	Nextera	1.96	2.7	400
S8 Normal	Nextera	1.9	2.6	400
S8 1 Tumour (Sigma)	Nextera	1.8	2.29	400
S8 2 Tumour (Sigma)	Nextera	1.86	2.6	400
S9 Normal	Nimblegen	2.00	1.87	1000
S9 1 Tumour (Sigma)	Nimblegen	1.94	2.34	1037
S9 1 Tumour (Sigma)	Nextera	1.8	2.4	400
S10 Normal	Nimblegen	1.88	2.16	1457
S10 1 Tumour (Sigma)	Nimblegen	1.94	2.30	1125
S11 Normal	Nextera	1.93	2.27	400
S11 1 Tumour (Sigma)	Nextera	1.9	2.6	400
S11 2 Tumour (Sigma)	Nextera	1.86	2.6	400
S12 Normal	Nextera	1.88	2.29	400
S12 1 Tumour (Sigma)	Nextera	1.9	2.3	400
S12 2 Tumour (Sigma)	Nextera	1.87	2.49	400
S13 Normal	Nextera	1.93	2.28	400
S13 1 Tumour (Sigma)	Nextera	1.9	2.4	400
S13 2 Tumour (Sigma)	Nextera	1.92	2.79	400
S1 Normal	Nextera	1.78	2.2	400
S1 1 Tumour (Sigma)	Nextera	1.84	2.46	400
S1 2 Tumour (Sigma)	Nextera	1.87	2.65	400
S2 Normal	Nextera	1.9	2	400
S2 1 Tumour (Sigma)	Nextera	1.85	2.15	400
S2 2 Tumour (Sigma)	Nextera	1.95	2.49	400

Appendix 2: PCR and sequencing primers for house-keeping genes

Name	Sequence (5' to 3')	PCR-product (bp)
ATM forward	CTACCCCAACAGCGACATGG	461
ATM reverse	TCTGCAGGCTGACCCAGTAA	
BCR forward	GGCTGCCCTACATTGATGACT	403
BCR reverse	AGGCTATGAAGCAAAGGGACC	
BRAF forward	CCTCTGTTTGTTGGGCAGGA	478
BRAF reverse	TGGAATCTCTGGGGAACGGA	
CDKN2A forward	GTGAGATCTGGAAGGAGGTGG	422
CDKN2A reverse	TCTAAGCATGGCTGTTCGAGG	
CYLD forward	AGAATGCAGCGTTACAGACAAAC	596
CYLD reverse	TCCGGAGACTGTCCTCTCTG	
ETV6 forward	ATGCCCTGCTGCCAACTTAT	477
ETV6 reverse	GCAGGCCTTAGGAAGTCTGG	
HRAS forward	GCCACAGAGCCGAAGAAAAC	492
HRAS reverse	GGCCTGGCATGAGGTATGTC	
JAK2 forward	TTCGAGCAAAGATCCAAGACTATCA	451
JAK2 reverse	ACTGAATTCCACCGTTTCCAGT	
KDM6A forward	TTTCCCCAGAACACTCAGGC	328
KDM6A reverse	GCTCCACGCCTCCTCAATAC	
NPM1 forward	AAGAAATGGATTTGCACAGTAAGT	518
NPM1 reverse	CCTGGAAAGATAGCTATGCCACA	
NRAS forward	TCAGGAATGTTGGGTCTCAGTT	403
NRAS reverse	ACTTCACACTTGCACCGAAAT	
PCM1 forward	ATTCCCAAGGTAGAGCAACAGC	301
PCM1 reverse	CTCCATTGCAGGTTCTCCCC	
Pik3ca forward	AGAACAATGCCTCCACGACC	236
Pik3ca reverse	TTCTTCACGGTTGCCTACTGG	
PTEN forward	ACCATAACCCACCACAGCTAGA	252
PTEN reverse	TCTAGGGCCTCTTGTGCCTT	
SOCS1 forward	CACAAAGTGCTCAAACGCGA	273
SOCS1 reverse	GGTCCCCCTGGTTGTTGTAG	
SUFU forward	GAGTCTCACCCACCGAGTCC	200
SUFU reverse	GATAGCGGTAACCTGGAGCG	
TNFAIP3 ex8forward	ACTGTCAGCATCTCTGTATCGGTG	188
TNFAIP3 ex8reverse	ATGAGGAGACAGAACCTGGCAGAG	
TP53 forward	CGTGTTCCTAAAATGAGGGGGA	265
TP53 reverse	AGGCCAGCTCTGCTACACTA	
ATM internal-Reverse	TGGCTCCAAGTAAGCCAAAGT	204
BCR internal-Reverse	CTCCACGCTCAGTTCAAAGC	328
BCR internal-Forward	TTGTGGTGACTTGGGGTTGC	161
CDKN2A internal-Forward	ATCCAGAGTACTTGGGGTGTG	355

CYLD internal-Forward	ATGCCCTGCTGCCAACTTAT	254
ETV6 internal-Forward	CAGCCTTGACAGTGGTTGTG	350
HRAS internal-Reverse	ATAGCCTTTGGGAGGGGGAG	368
JAK2 internal-Forward	GCGAATAAGGTACAGATTTCGCAG	212
KDM6A internal-Forward	GGAAGGTTACACAACTGAGCC	538
NPM1 internal-Forward	CACAGTAAGTAGAGCAGGCAC	463
NRAS internal-Forward	GCGCCAATTGTACCATGTCC	344
PCM1 internal-Forward	AGTGAAAACCGAAAGCCCTTCA	193
PI3KCA internal-Forward	GAACTGTGGGGCATCCACTT	271
PTEN internal-Reverse	TGACCAATGGCTAAGTGAAGATGA	135
SOCS1 internal-Reverse	AACTCGCACCTCCTACCTCT	401
SUF internal-Forward	CCTCTCCAGTTCCCCCAGT	209
TP53 internal-Reverse	GGCAGACATGTTCCCTACCC	359

Appendix 3: Summary of exome sequences mapped to hg19

Sample	Number of total reads	Number of reads mapped to hg19	Number of de-duplicated reads mapped to hg19	Percentage of reads de-duplicated reads mapped to hg19
S1_1T	52982418	49388813	39098426	79.16
S1_2T	18457784	15109232	10853837	71.84
S1_Normal	43459734	42723024	34844738	81.56
S2_1T	45170576	39295454	29607223	75.35
S2_2T	30160728	25468631	18986498	74.55
S2_Normal	44016574	43407422	37090820	85.45
S3_1T	30792566	26354149	20099545	76.27
S3_2T	26970450	23097204	18128259	78.49
S3_Normal	26543008	26106778	23075364	88.39
S4_1T	13688500	11876421	10165504	85.59
S4_2T	24437198	20975247	17344225	82.69
S4_Normal	1.73E+08	1.68E+08	1.44E+08	85.68
S5_1T	38075914	34515232	27658719	80.13
S5_2T	31270904	25394165	19875062	78.27
S5_Normal	36918370	36379083	31788530	87.38
S6_1T	47525522	20632679	17873878	86.63
S6_2T	40490738	36266864	27927501	77.01
S6_Normal	1.28E+08	1.24E+08	1.06E+08	85.86
S7_1T	18233490	13406234	10741397	80.12
S7_2T	27383540	21722212	17010090	78.31
S7_Normal	71048628	70131419	63635432	90.74
S8_1T	27180052	23998343	19382890	80.77
S8_2T	39987020	32195493	25088072	77.92
S8_Normal	44108778	43390475	39337216	90.66
S9_1T	62436382	25162228	21823801	86.73
S9_2T	24198644	21035810	17101118	81.3
S9_Normal	1.05E+08	83559457	64064749	76.67
S10_1T	78022084	34099910	29012204	85.08
S10_2T	23838192	17298606	14682956	84.88
S10_Normal	1.57E+08	1.27E+08	93608826	73.64
S11_1T	33858398	27558083	20844841	75.64
S11_2T	28138442	22537498	17495344	77.63
S11_Normal	60749026	59737127	50348182	84.28
S12_1T	21905612	17792530	14178459	79.69
S12_2T	32410952	26554604	20323830	76.54
S12_Normal	50020384	49311261	43679423	88.58
S13_1T	25248462	21351419	16920239	79.25
S13_2T	34537668	27326475	20521960	75.1
S13_Normal	62109966	61217081	54985887	89.82

*Samples failed in exome sequencing, as confirmed in data analysis, were marked in Red. 1T=first tumour replicate; 2T=second tumour replicate.

Appendix 4: List of most frequently mutated genes in my series

KLF6	ZNF292	SMC1B	PDE4DIP	GAPVD1	CLOCK
SMG1	ZFR	SMARCC1	PCCA	GALC	CKLF−CMTM1
TTN−AS1	ZBTB24	SMARCAL1	NOL12	FUS−RP11−388M20.1	CHD2
TTN	Z85986.1	SMARCAD1	NDUFB5	FUS	CEP41
TPTE2	YTHDF3	SLC36A1	NAV3	FIP1L1	CDK5RAP2
OR4C5	XIST	SLC35A1	NAA16	FGFR4	CDH3
CSMD3	WDR35	SEC31A	MYO6	ETNK1	CDC27
ZNF717	USP15	RYR2	MUC19	EPN2	CCT6B
STAG3L2	UPF2	RP11−888D10.4	MUC16	EPG5	CCIN
SREK1	UGT1A9	RP11−799N11.1	MICU3	EML4	CCDC73
snoU13	UGT1A8	RP11−403P17.2	MICAL1	EIF3H	C6orf165
SGOL2	UGT1A7	RP11−388M20.1	MATR3	EFCAB7	C20orf194
SF3B2	UGT1A6	RP11−382A20.7	LNX1	EDRF1−AS1	C1orf87
SBF2	UGT1A5	RP11−382A20.6	LIPE−AS1	EDRF1	C12orf45
RP11−1167A19.2	UGT1A10	RP11−382A20.5	LAMA4	EAPP	BTBD1
RGS3	TSIX	RP11−21J18.1	LAMA3	DYNC2H1	BSPH1
R3HDM2	TRIP13	RP11−135L13.4	KLHL7	DPYD	BRD9
PSMG2	TRIOBP	RP11−123K3.4	KLHL1	DPCR1	BNIP2
PROM1	TNFRSF9	RP11−115F18.1	KIF3A	DNAH12	BBS4
PACS1	TMEM57	RNF7P1	KIAA1009	DMTF1	BBS2
NF1	TMEM131	RNF213	JAKMIP1	DLEU2L	ATXN2
MUC4	TEX15	RN7SKP64	ITPR2	DLEC1	ATAD2
MUC17	TEX10	RMDN2−AS1	ITPR1	DIAPH3	APIP
MIR4273	TAF2	RMDN2	ITGB3BP	DFNB59	ANKRD44
JMJD1C	SYNE1	RIMS2	IPP	DCBLD2	ANKRD12
HMCN1	STK38	RANBP2	IGHV3−71	DAG1	AGPAT9
HCFC2	STK31	RALGPS2	IGHV2−70	CYSLTR1	ADAMTS12
F8	STAMBPL1	RAD50	HIVEP1	CTD−2047H16.4	AC004237.1
F5	ST3GAL6	QSER1	HECTD1	CTC−297N7.11	ABCC2
CHD4	SRRM2	PROS1	HCG21	COL9A1	ABCA1
CEP76	SOX5	PRKRA	GTF3C1	CMTM2	
CCDC80	SORBS2	PRKDC	GNE	CMTM1	
ARFGEF2	SNRNP200	PPHLN1	GLTSCR1L	CMPK1	
ZNF644	SMC6	PKHD1	GARS	CLTA	

Appendix 5: List of most frequently mutated genes (protein-altering) in my series

SMG1	ZNF292	RALGPS2	KIF3A	DPYD	BTBD1
TPTE2	ZFR	RAD50	KIAA1009	DNAH12	BSPH1
OR4C5	USP15	QSER1	JMJD1C	DMTF1	BRD9
ZNF717	UGT1A8	PROS1	ITPR2	DLEU2L	BNIP2
TTN	TRIOBP	PRKRA	ITPR1	CYSLTR1	BBS2
SREK1	TMEM57	PDE4DIP	IPP	COL9A1	ATAD2
SGOL2	SYNE1	NF1	IGHV2−70	CMTM1	APIP
SF3B2	STK38	NDUFB5	HECTD1	CMPK1	ANKRD12
PROM1	ST3GAL6	NAV3	GTF3C1	CKLF−CMTM1	AGPAT9
MUC17	SRRM2	MYO6	GNE	CHD4	ADAMTS12
HCFC2	SNRNP200	MUC4	GARS	CHD2	ABCA1
F8	SMC1B	MICU3	EPN2	CEP76	
F5	SMARCC1	LAMA4	EPG5	CCT6B	
CSMD3	SBF2	LAMA3	EML4	CCDC73	
CCDC80	RYR2	KLHL7	EIF3H	C20orf194	
ZNF644	RANBP2	KLHL1	EAPP	C12orf45	

Appendix 6: Biological processes enriched in 42 genes common in all 3 HL studies

Gene ontology term	*p*-value
GO:0042592~homeostatic process	1.1E-06
GO:0001894~tissue homeostasis	4.1E-05
GO:0060249~anatomical structure homeostasis	5.9E-05
GO:0032088~negative regulation of NF-kappaB transcription factor activity	3.7E-04
GO:0048871~multicellular organismal homeostasis	3.7E-04
GO:0044085~cellular component biogenesis	7.1E-04
GO:0048878~chemical homeostasis	8.2E-04
GO:0022607~cellular component assembly	8.9E-04
GO:0045184~establishment of protein localization	1.2E-03
GO:0007010~cytoskeleton organization	1.4E-03
GO:0055088~lipid homeostasis	1.9E-03
GO:0043433~negative regulation of sequence-specific DNA binding transcription factor activity	2.5E-03
GO:0032886~regulation of microtubule-based process	2.9E-03
GO:1902589~single-organism organelle organization	3.3E-03
GO:0042060~wound healing	4.0E-03
GO:0007049~cell cycle	4.2E-03
GO:0010803~regulation of tumour necrosis factor-mediated signalling pathway	4.6E-03
GO:0043124~negative regulation of I-kappaB kinase/NF-kappaB signalling	4.9E-03
GO:0032845~negative regulation of homeostatic process	5.0E-03
GO:0019725~cellular homeostasis	5.3E-03
GO:0033036~macromolecule localization	5.8E-03
GO:0008104~protein localization	5.8E-03
GO:0016266~O-glycan processing	6.3E-03
GO:0060113~inner ear receptor cell differentiation	6.3E-03
GO:0048646~anatomical structure formation involved in morphogenesis	6.5E-03
GO:0007017~microtubule-based process	7.0E-03
GO:0042490~mechanoreceptor differentiation	7.1E-03
GO:0051641~cellular localization	7.7E-03
GO:0009611~response to wounding	8.5E-03
GO:0030031~cell projection assembly	9.1E-03
GO:0000819~sister chromatid segregation	9.7E-03
GO:0000226~microtubule cytoskeleton organization	9.8E-03
GO:0032844~regulation of homeostatic process	1.2E-02
GO:1902532~negative regulation of intracellular signal transduction	1.2E-02
GO:0055082~cellular chemical homeostasis	1.2E-02
GO:0071705~nitrogen compound transport	1.3E-02
GO:0022402~cell cycle process	1.5E-02
GO:0048514~blood vessel morphogenesis	1.6E-02
GO:0007155~cell adhesion	1.6E-02

GO:0022610~biological adhesion	1.6E-02
GO:0048739~cardiac muscle fibre development	1.8E-02
GO:0006493~protein O-linked glycosylation	1.8E-02
GO:0098813~nuclear chromosome segregation	1.9E-02
GO:0010927~cellular component assembly involved in morphogenesis	2.0E-02
GO:0070507~regulation of microtubule cytoskeleton organization	2.4E-02
GO:0015031~protein transport	2.6E-02
GO:0001568~blood vessel development	2.6E-02
GO:0061041~regulation of wound healing	2.7E-02
GO:0007062~sister chromatid cohesion	2.7E-02
GO:0007059~chromosome segregation	2.8E-02
GO:0016925~protein sumoylation	2.8E-02
GO:0044872~lipoprotein localization	2.9E-02
GO:0042953~lipoprotein transport	2.9E-02
GO:1903047~mitotic cell cycle process	3.0E-02
GO:0007596~blood coagulation	3.0E-02
GO:0001944~vasculature development	3.1E-02
GO:0001959~regulation of cytokine-mediated signalling pathway	3.3E-02
GO:0007605~sensory perception of sound	3.3E-02
GO:0007599~hemostasis	3.3E-02
GO:0050817~coagulation	3.3E-02
GO:1903034~regulation of response to wounding	3.4E-02
GO:0060759~regulation of response to cytokine stimulus	3.5E-02
GO:0072358~cardiovascular system development	3.7E-02
GO:0072359~circulatory system development	3.7E-02
GO:0032703~negative regulation of interleukin-2 production	3.7E-02
GO:0051090~regulation of sequence-specific DNA binding transcription factor activity	3.7E-02
GO:0051649~establishment of localization in cell	3.8E-02
GO:0070925~organelle assembly	3.9E-02
GO:0050954~sensory perception of mechanical stimulus	3.9E-02
GO:0002221~pattern recognition receptor signalling pathway	3.9E-02
GO:0030030~cell projection organization	4.0E-02
GO:0000278~mitotic cell cycle	4.0E-02
GO:0030168~platelet activation	4.0E-02
GO:0006897~endocytosis	4.1E-02
GO:0006874~cellular calcium ion homeostasis	4.2E-02
GO:0090181~regulation of cholesterol metabolic process	4.3E-02
GO:0032692~negative regulation of interleukin-1 production	4.5E-02
GO:0055074~calcium ion homeostasis	4.5E-02
GO:0051642~centrosome localization	4.7E-02
GO:0071379~cellular response to prostaglandin stimulus	4.7E-02
GO:0050796~regulation of insulin secretion	4.7E-02
GO:0072503~cellular divalent inorganic cation homeostasis	4.8E-02
GO:0048839~inner ear development	4.8E-02

Appendix 7: Biological processes enriched in 103 genes shared between my study and Reichel et al., 2015

Gene ontology term	*p*-value
GO:1902589~single-organism organelle organization	4.06E-04
GO:0022607~cellular component assembly	4.29E-04
GO:0042592~homeostatic process	5.24E-04
GO:0001894~tissue homeostasis	5.26E-04
GO:0044085~cellular component biogenesis	9.67E-04
GO:0048871~multicellular organismal homeostasis	0.001177
GO:0007010~cytoskeleton organization	0.001856
GO:0060249~anatomical structure homeostasis	0.002138
GO:0016266~O-glycan processing	0.003563
GO:0060113~inner ear receptor cell differentiation	0.003563
GO:0051170~nuclear import	0.004258
GO:0042490~mechanoreceptor differentiation	0.004275
GO:0016925~protein sumoylation	0.004783
GO:0032088~negative regulation of NF-kappaB transcription factor activity	0.005716
GO:0000226~microtubule cytoskeleton organization	0.006233
GO:0006913~nucleocytoplasmic transport	0.009622
GO:0060122~inner ear receptor stereocilium organization	0.010212
GO:0051169~nuclear transport	0.010504
GO:0034504~protein localization to nucleus	0.01064
GO:0071705~nitrogen compound transport	0.011529
GO:0032845~negative regulation of homeostatic process	0.012582
GO:0007017~microtubule-based process	0.013514
GO:0006606~protein import into nucleus	0.014286
GO:0044744~protein targeting to nucleus	0.014286
GO:1902593~single-organism nuclear import	0.014488
GO:0032446~protein modification by small protein conjugation	0.014712
GO:0051641~cellular localization	0.015249
GO:0070647~protein modification by small protein conjugation or removal	0.015618
GO:0006493~protein O-linked glycosylation	0.016154
GO:0007160~cell-matrix adhesion	0.016814
GO:0034329~cell junction assembly	0.016814
GO:0048878~chemical homeostasis	0.017438
GO:0031589~cell-substrate adhesion	0.020174
GO:0030031~cell projection assembly	0.020231
GO:0042060~wound healing	0.020725
GO:0060119~inner ear receptor cell development	0.022123
GO:0046822~regulation of nucleocytoplasmic transport	0.023946
GO:0044257~cellular protein catabolic process	0.024257

GO:0045184~establishment of protein localization	0.025295
GO:0055088~lipid homeostasis	0.025592
GO:0017038~protein import	0.026554
GO:0002433~immune response-regulating cell surface receptor signalling pathway involved in phagocytosis	0.027215
GO:0038096~Fc-gamma receptor signalling pathway involved in phagocytosis	0.027215
GO:0032844~regulation of homeostatic process	0.027767
GO:0061041~regulation of wound healing	0.027768
GO:0050810~regulation of steroid biosynthetic process	0.027938
GO:0010803~regulation of tumour necrosis factor-mediated signalling pathway	0.027938
GO:0007059~chromosome segregation	0.028094
GO:1902532~negative regulation of intracellular signal transduction	0.028558
GO:0002431~Fc receptor mediated stimulatory signalling pathway	0.028891
GO:0038094~Fc-gamma receptor signalling pathway	0.029462
GO:0043124~negative regulation of I-kappaB kinase/NF-kappaB signalling	0.029999
GO:0048646~anatomical structure formation involved in morphogenesis	0.031882
GO:0010035~response to inorganic substance	0.032733
GO:0043433~negative regulation of sequence-specific DNA binding transcription factor activity	0.033013
GO:0032886~regulation of microtubule-based process	0.037425
GO:0006464~cellular protein modification process	0.038411
GO:0036211~protein modification process	0.038411
GO:0034330~cell junction organization	0.038602
GO:1903034~regulation of response to wounding	0.039404
GO:0048514~blood vessel morphogenesis	0.039549
GO:0043933~macromolecular complex subunit organization	0.039765
GO:0002757~immune response-activating signal transduction	0.041895
GO:0032790~ribosome disassembly	0.045029
GO:0048739~cardiac muscle fibre development	0.045029
GO:0051649~establishment of localization in cell	0.045288
GO:0075733~intracellular transport of virus	0.045988
GO:1902581~multi-organism cellular localization	0.047228
GO:1902583~multi-organism intracellular transport	0.047228
GO:0009611~response to wounding	0.047843
GO:0007155~cell adhesion	0.048319
GO:1903779~regulation of cardiac conduction	0.04848
GO:0070925~organelle assembly	0.04916
GO:0022610~biological adhesion	0.049521
GO:0046794~transport of virus	0.049745

Appendix 8: 377 protein-altering genes in my series that were also downregulated in Brune or Steidl array

ABCA1	CEBPZ	GIGYF2	MRPS30	QSER1	TAB2
ACBD3	CEP192	GLUL	MSANTD2	RAB14	TAF2
AGO4	CEP57	GNB2L1	MSH2	RAD50	TANK
AKAP8	CEP76	GNE	MSL2	RAD51C	TARDBP
AKT2	CEP78	GNL2	MTDH	RALGPS2	TBC1D1
ALDH18A1	CHD1L	GPCPD1	MTRR	RANBP2	TBC1D7
ALDH3A2	CHD9	GTF2I	MUT	RASGRP2	TBCEL
ALMS1	CLCN3	HBS1L	MX1	RB1	TBP
AMY2B	CMPK1	HCFC2	MYCBP2	RBBP7	TEP1
ANAPC5	CMTR2	HEATR3	MYEF2	RBM48	TEX10
ANKHD1	CNTN1	HEATR5A	MYO6	RC3H1	TFDP1
ANKMY2	COG1	HEATR5B	MYO9A	RFC3	TMEM131
ANKRD12	CPNE3	HECTD1	NAA15	RIF1	TMEM156
ANO5	CSE1L	HEG1	NAA16	RMDN2	TMEM183A
ANXA2	CTCF	HLA-DQA1	NASP	RNASEH2A	TMEM50A
APIP	CUX1	HMCN1	NAT10	RNF138	TMTC3
APPL1	CWC27	HNRNPC	NBN	RNF7	TNKS2
ARFGAP3	CYB5R4	HNRNPDL	NCAPG2	RPL9	TOP2B
ARFGEF2	DDX20	HSD17B4	NCOR1	RSPRY1	TOPBP1
ARHGAP25	DDX6	HSPA13	NDUFB5	RUFY1	TPP1
ARID4A	DHX38	HSPA8	NEMF	SC5D	TRAPPC8
ARL13B	DIAPH3	ICA1	NF1	SCAI	TRIM26
ARL14EP	DMTF1	IFRD1	NFKB1	SCO1	TRIM66
ARNT	DNAJC10	IGHV1-69	NIN	SCOC	TRIOBP
ARV1	DOCK7	IKZF5	NOL11	SEC23B	TRIP12
ASCC3	DOCK9	IMPACT	NUP50	SEC31A	TTC12
ATAD2	DUS2	INADL	OCIAD2	SEL1L	TTN
ATF7	DYNC1LI1	INSIG2	ODC1	SENP2	UBE3A
ATM	EAPP	INVS	OFD1	SETD2	UGGT2
ATPAF1	ECT2	IPO7	PANK1	SETD7	USH1C
ATXN10	EDRF1	IPP	PCOLCE2	SETX	USP15
AVL9	EFHC1	IRF8	PDCD2L	SF3B3	USP46
B3GALNT1	EIF3H	ITGB8	PDCD6	SGOL1	VAMP4
BAZ2B	EIF4G2	ITPR1	PDE7A	SGOL2	VCAM1
BBS12	EIF4G3	ITPR2	PDS5B	SLC12A2	VEGFB
BBS2	EML6	JMJD1C	PDSS1	SLC1A1	VPS13A
BCLAF1	EP300	KIAA0125	PGM1	SLC25A12	XPNPEP1
BCR	EPC1	KIAA0922	PHF23	SLC25A13	XPO1
BIRC6	ERCC6	KIAA1551	PIBF1	SLC25A3	XRCC4
BNIP2	ERCC6L2	KIAA2026	PIGK	SLC25A33	YKT6

BTBD1	ETF1	KLHL7	PIK3CB	SLC35A1	YTHDF3
C10orf76	EVI2B	KMT2A	PITPNC1	SLC4A1AP	ZBTB21
C11orf30	EWSR1	LAMA3	PIWIL4	SLC4A7	ZCRB1
C11orf54	FAF1	LBR	PLSCR1	SLC7A6	ZFR
C1S	FAM111B	LIFR	PMS1	SMARCC1	ZIK1
C20orf24	FAM114A2	LPGAT1	PNRC1	SMC1A	ZNF12
C5	FAM120B	LRBA	POLQ	SMC6	ZNF143
C5orf22	FAM210A	LRPPRC	PPFIA1	SMG1	ZNF23
C6orf89	FAM83D	LSM5	PPP3CA	SNAP29	ZNF253
C7orf31	FAR1	LYSMD3	PPP3CB	SNAPC1	ZNF292
CARF	FAS	MAN1A2	PPP6C	SNRNP200	ZNF33B
CASD1	FAT3	MBD4	PREPL	SON	ZNF430
CASP2	FBXO45	MBNL2	PRKDC	SP4	ZNF439
CCDC113	FBXO7	MCCC1	PRKRA	SPG11	ZNF567
CCNL1	FCRL5	MCM6	PROSER1	SREK1	ZNF625
CCT6B	FNBP1	MCM9	PRPF3	SRPR	ZNF644
CDC7	FUNDC1	MCPH1	PRRC2C	SSB	ZNF682
CDCA2	GALC	MCTP2	PSME3	STK17B	ZNF700
CDK13	GAPT	MEF2A	PTEN	STK38	ZNF708
CDK5R1	GAPVD1	MFN1	PTK2	STK38L	ZNF717
CDK5RAP2	GARS	MIER3	PTPN14	SUCLA2	ZNF91
CDR2	GBA3	MRPL15	PTPRC	SUGP2	ZNF93
CDYL	GCNT2	MRPS18B	PUS10	SYNE1	

Appendix 9: 141 protein-altering genes in my series that were also upregulated in Brune or Steidl array

ABI3BP	CD109	FAT3	LMLN	PGAP1	SRXN1
ACACA	CD96	FBXO40	LRPPRC	PRRC2C	ST8SIA4
ACVR1	CDH8	FERMT1	MACF1	PSME3	STAT5B
ADAMTS6	CEACAM1	FGF7	MED13L	PTPRD	SVEP1
ADAMTSL4	CHD2	FMNL2	MESDC2	RBFOX2	SYNPO
AKAP13	CIT	FNDC3B	MKS1	RELN	TARS
AKT3	CLEC12A	GALNT10	MLEC	RNF145	TMEM136
ALDH3A2	CLOCK	GLDC	MLLT4	RNF213	TMEM50A
ALK	CSMD3	GLUL	MOGS	RNF217	TMEM97
ALPK2	CTSA	GNB2L1	MTTP	RNF7	TMTC2
ANAPC5	CYP1B1	GSS	MUC4	RPL9	TNFAIP3
ANKRD36	CYSLTR1	GTF3C1	MYO10	RTCB	TRPS1
ANKS1B	DIAPH1	HEPH	NAV3	RYR2	TSPAN5
ANXA2	DNAH12	HTR2A	NEDD4L	S100B	USH1C
ARHGAP10	DNAH6	IFIT3	NEK11	SAMHD1	UTS2B
BCS1L	DOCK10	IL12RB2	NFE2L1	SEC61A2	VAV2
C12orf45	DPCR1	JAG1	OCIAD2	SIM1	VCAM1
C1S	DYNC2H1	JAKMIP1	OTUD7A	SLC20A1	WDR43
C20orf194	EHBP1	KCNA1	PAFAH1B3	SLC28A2	ZBED6
C6orf89	ELOVL7	KIF5A	PBX4	SLC4A4	ZCRB1
CACHD1	EPN2	KNOP1	PCCA	SLC5A6	ZNF462
CALCOCO2	ERAP1	LAMA4	PDCD2L	SNCAIP	
CAMSAP2	F5	LAMTOR2	PDE4DIP	SQLE	
CCDC80	FAM98A	LIMCH1	PDLIM5	SRRM2	

Appendix 10: 125 biological processes that are significantly enriched in protein-altering genes that were also down-regulated in Brune or Steidl array

Gene ontology term	*p*-value
GO:0034645~cellular macromolecule biosynthetic process	1.3E-08
GO:0006259~DNA metabolic process	3.8E-08
GO:1903047~mitotic cell cycle process	1.0E-07
GO:0022402~cell cycle process	2.5E-07
GO:0010467~gene expression	3.2E-07
GO:0051276~chromosome organization	3.5E-07
GO:0000278~mitotic cell cycle	3.5E-07
GO:0051171~regulation of nitrogen compound metabolic process	3.9E-07
GO:0006974~cellular response to DNA damage stimulus	5.1E-07
GO:0007049~cell cycle	6.4E-07
GO:0006281~DNA repair	9.0E-07
GO:0044770~cell cycle phase transition	1.7E-06
GO:0033554~cellular response to stress	2.1E-06
GO:2000112~regulation of cellular macromolecule biosynthetic process	2.7E-06
GO:0010468~regulation of gene expression	3.6E-06
GO:0019219~regulation of nucleobase-containing compound metabolic process	6.0E-06
GO:0010564~regulation of cell cycle process	8.3E-06
GO:0010556~regulation of macromolecule biosynthetic process	1.1E-05
GO:0051053~negative regulation of DNA metabolic process	1.2E-05
GO:0016070~RNA metabolic process	1.3E-05
GO:0044772~mitotic cell cycle phase transition	1.5E-05
GO:0007059~chromosome segregation	2.0E-05
GO:0006302~double-strand break repair	3.7E-05
GO:0060249~anatomical structure homeostasis	4.0E-05
GO:0000723~telomere maintenance	4.6E-05
GO:0007093~mitotic cell cycle checkpoint	4.8E-05
GO:0045930~negative regulation of mitotic cell cycle	7.1E-05
GO:0051592~response to calcium ion	7.8E-05
GO:0006464~cellular protein modification process	8.5E-05
GO:0036211~protein modification process	8.5E-05
GO:0051726~regulation of cell cycle	8.6E-05
GO:0044839~cell cycle G2/M phase transition	9.3E-05
GO:0006261~DNA-dependent DNA replication	9.4E-05
GO:0034654~nucleobase-containing compound biosynthetic process	9.6E-05
GO:0018130~heterocycle biosynthetic process	1.0E-04
GO:0032200~telomere organization	1.1E-04
GO:0032205~negative regulation of telomere maintenance	1.1E-04
GO:0019438~aromatic compound biosynthetic process	1.2E-04
GO:0006260~DNA replication	1.3E-04

GO:0006310~DNA recombination	1.4E-04
GO:1901991~negative regulation of mitotic cell cycle phase transition	1.6E-04
GO:0000724~double-strand break repair via homologous recombination	1.7E-04
GO:0000725~recombinational repair	1.8E-04
GO:0000086~G2/M transition of mitotic cell cycle	1.8E-04
GO:0044774~mitotic DNA integrity checkpoint	1.9E-04
GO:0010948~negative regulation of cell cycle process	1.9E-04
GO:0044843~cell cycle G1/S phase transition	2.2E-04
GO:0006351~transcription, DNA-templated	2.9E-04
GO:1901988~negative regulation of cell cycle phase transition	3.1E-04
GO:0000819~sister chromatid segregation	3.3E-04
GO:0010605~negative regulation of macromolecule metabolic process	3.4E-04
GO:0043933~macromolecular complex subunit organization	3.4E-04
GO:0051246~regulation of protein metabolic process	3.6E-04
GO:0098813~nuclear chromosome segregation	3.7E-04
GO:0000075~cell cycle checkpoint	4.1E-04
GO:0032268~regulation of cellular protein metabolic process	4.3E-04
GO:2000134~negative regulation of G1/S transition of mitotic cell cycle	4.5E-04
GO:0000726~non-recombinational repair	5.2E-04
GO:0044773~mitotic DNA damage checkpoint	5.9E-04
GO:1902807~negative regulation of cell cycle G1/S phase transition	5.9E-04
GO:1902806~regulation of cell cycle G1/S phase transition	6.4E-04
GO:0051252~regulation of RNA metabolic process	8.2E-04
GO:0018193~peptidyl-amino acid modification	8.6E-04
GO:0016310~phosphorylation	8.9E-04
GO:1902589~single-organism organelle organization	9.1E-04
GO:0031325~positive regulation of cellular metabolic process	1.0E-03
GO:1901796~regulation of signal transduction by p53 class mediator	1.0E-03
GO:2001251~negative regulation of chromosome organization	1.1E-03
GO:0009892~negative regulation of metabolic process	1.1E-03
GO:0006355~regulation of transcription, DNA-templated	1.1E-03
GO:0000082~G1/S transition of mitotic cell cycle	1.1E-03
GO:0033044~regulation of chromosome organization	1.2E-03
GO:0051172~negative regulation of nitrogen compound metabolic process	1.2E-03
GO:1903506~regulation of nucleic acid-templated transcription	1.4E-03
GO:0044085~cellular component biogenesis	1.4E-03
GO:0097659~nucleic acid-templated transcription	1.4E-03
GO:2000045~regulation of G1/S transition of mitotic cell cycle	1.4E-03
GO:0090329~regulation of DNA-dependent DNA replication	1.5E-03
GO:2001141~regulation of RNA biosynthetic process	1.6E-03
GO:0010212~response to ionizing radiation	1.6E-03
GO:0006312~mitotic recombination	1.7E-03
GO:0051052~regulation of DNA metabolic process	1.8E-03
GO:1901987~regulation of cell cycle phase transition	1.8E-03

GO:0042770~signal transduction in response to DNA damage	1.8E-03
GO:0018205~peptidyl-lysine modification	1.9E-03
GO:0006303~double-strand break repair via nonhomologous end joining	1.9E-03
GO:0032774~RNA biosynthetic process	1.9E-03
GO:0010604~positive regulation of macromolecule metabolic process	1.9E-03
GO:0045934~negative regulation of nucleobase-containing compound metabolic process	2.0E-03
GO:0051173~positive regulation of nitrogen compound metabolic process	2.1E-03
GO:0010038~response to metal ion	2.1E-03
GO:1901990~regulation of mitotic cell cycle phase transition	2.4E-03
GO:0010639~negative regulation of organelle organization	2.5E-03
GO:0035556~intracellular signal transduction	2.7E-03
GO:0006796~phosphate-containing compound metabolic process	2.7E-03
GO:0045935~positive regulation of nucleobase-containing compound metabolic process	2.8E-03
GO:0006793~phosphorus metabolic process	2.9E-03
GO:0022607~cellular component assembly	2.9E-03
GO:0031399~regulation of protein modification process	3.0E-03
GO:0031570~DNA integrity checkpoint	3.0E-03
GO:0031324~negative regulation of cellular metabolic process	3.0E-03
GO:0032508~DNA duplex unwinding	3.1E-03
GO:0000722~telomere maintenance via recombination	3.3E-03
GO:0006468~protein phosphorylation	3.3E-03
GO:1904354~negative regulation of telomere capping	3.4E-03
GO:0045786~negative regulation of cell cycle	3.5E-03
GO:0016925~protein sumoylation	3.7E-03
GO:0010608~posttranscriptional regulation of gene expression	3.7E-03
GO:0032392~DNA geometric change	4.2E-03
GO:0016071~mRNA metabolic process	4.2E-03
GO:0042325~regulation of phosphorylation	4.7E-03
GO:0009893~positive regulation of metabolic process	4.7E-03
GO:0000070~mitotic sister chromatid segregation	5.4E-03
GO:0007346~regulation of mitotic cell cycle	5.5E-03
GO:0051649~establishment of localization in cell	6.3E-03
GO:0010557~positive regulation of macromolecule biosynthetic process	6.3E-03
GO:0000729~DNA double-strand break processing	6.8E-03
GO:0000077~DNA damage checkpoint	6.9E-03
GO:0006278~RNA-dependent DNA biosynthetic process	7.0E-03
GO:0048285~organelle fission	7.1E-03
GO:0000280~nuclear division	7.1E-03
GO:0006913~nucleocytoplasmic transport	7.2E-03
GO:0006396~RNA processing	7.3E-03
GO:0051129~negative regulation of cellular component organization	7.4E-03
GO:0071103~DNA conformation change	7.7E-03

Appendix 11: Gene ontology terms enriched in protein-altering genes that were also upregulated in Brune or Steidl array

Gene ontology term	*p*-value
GO:0022603~regulation of anatomical structure morphogenesis	0.001315
GO:0072359~circulatory system development	0.003426
GO:0072358~cardiovascular system development	0.003426
GO:0006768~biotin metabolic process	0.003671
GO:0001568~blood vessel development	0.004505
GO:0001944~vasculature development	0.006681
GO:0006928~movement of cell or subcellular component	0.012059
GO:0022604~regulation of cell morphogenesis	0.012232
GO:0061443~endocardial cushion cell differentiation	0.013051
GO:0007155~cell adhesion	0.013913
GO:0022610~biological adhesion	0.014402
GO:2000026~regulation of multicellular organismal development	0.015774
GO:0003013~circulatory system process	0.018823
GO:0001525~angiogenesis	0.018848
GO:0043616~keratinocyte proliferation	0.019647
GO:0002715~regulation of natural killer cell mediated immunity	0.021954
GO:0045824~negative regulation of innate immune response	0.025613
GO:0048813~dendrite morphogenesis	0.025814
GO:0071947~protein deubiquitination involved in ubiquitin-dependent protein catabolic process	0.025932
GO:0032787~monocarboxylic acid metabolic process	0.030389
GO:0045088~regulation of innate immune response	0.030998
GO:0030029~actin filament-based process	0.032166
GO:0034394~protein localization to cell surface	0.032214
GO:0030334~regulation of cell migration	0.035697
GO:0051270~regulation of cellular component movement	0.037331
GO:0042592~homeostatic process	0.038726
GO:0044255~cellular lipid metabolic process	0.04003
GO:0040011~locomotion	0.040617
GO:0030001~metal ion transport	0.040644
GO:0035249~synaptic transmission, glutamatergic	0.040909
GO:0048514~blood vessel morphogenesis	0.042396
GO:0019752~carboxylic acid metabolic process	0.044154
GO:0006643~membrane lipid metabolic process	0.045081
GO:0043436~oxoacid metabolic process	0.045676
GO:0035051~cardiocyte differentiation	0.046032
GO:0030856~regulation of epithelial cell differentiation	0.046965
GO:0009605~response to external stimulus	0.047611

Appendix 12: Biological processes enriched in 857 genes with protein-altering mutations in my series

Gene ontology term	*p*-value
GO:0060249~anatomical structure homeostasis	0.000
GO:0001894~tissue homeostasis	0.000
GO:0045494~photoreceptor cell maintenance	0.000
GO:0051276~chromosome organization	0.000
GO:0051053~negative regulation of DNA metabolic process	0.001
GO:0006281~DNA repair	0.001
GO:0007049~cell cycle	0.001
GO:0032205~negative regulation of telomere maintenance	0.001
GO:1903047~mitotic cell cycle process	0.001
GO:0001895~retina homeostasis	0.001
GO:0098656~anion transmembrane transport	0.001
GO:0022402~cell cycle process	0.001
GO:0044770~cell cycle phase transition	0.001
GO:0006259~DNA metabolic process	0.001
GO:0006974~cellular response to DNA damage stimulus	0.001
GO:0042592~homeostatic process	0.001
GO:0048871~multicellular organismal homeostasis	0.002
GO:0007059~chromosome segregation	0.002
GO:0006464~cellular protein modification process	0.002
GO:0036211~protein modification process	0.002
GO:0032508~DNA duplex unwinding	0.002
GO:0006302~double-strand break repair	0.002
GO:0042787~protein ubiquitination involved in ubiquitin-dependent protein catabolic process	0.002
GO:0051592~response to calcium ion	0.002
GO:0044772~mitotic cell cycle phase transition	0.002
GO:0000723~telomere maintenance	0.002
GO:0010038~response to metal ion	0.002
GO:0000278~mitotic cell cycle	0.002
GO:0044839~cell cycle G2/M phase transition	0.002
GO:0006820~anion transport	0.003
GO:0032392~DNA geometric change	0.003
GO:0000086~G2/M transition of mitotic cell cycle	0.003
GO:0000724~double-strand break repair via homologous recombination	0.003
GO:0000725~recombinational repair	0.003
GO:0098813~nuclear chromosome segregation	0.004
GO:0042384~cilium assembly	0.004
GO:0060113~inner ear receptor cell differentiation	0.004
GO:0035825~reciprocal DNA recombination	0.004

GO:0007131~reciprocal meiotic recombination	0.004
GO:0044085~cellular component biogenesis	0.005
GO:0010035~response to inorganic substance	0.005
GO:0032200~telomere organization	0.005
GO:0044782~cilium organization	0.005
GO:0044179~hemolysis in other organism	0.006
GO:0052331~hemolysis in other organism involved in symbiotic interaction	0.006
GO:0051801~cytolysis in other organism involved in symbiotic interaction	0.006
GO:0019836~hemolysis by symbiont of host erythrocytes	0.006
GO:0001897~cytolysis by symbiont of host cells	0.006
GO:0051715~cytolysis in other organism	0.006
GO:1902589~single-organism organelle organization	0.006
GO:0022607~cellular component assembly	0.006
GO:0042490~mechanoreceptor differentiation	0.006
GO:0044774~mitotic DNA integrity checkpoint	0.007
GO:0044257~cellular protein catabolic process	0.007
GO:0010564~regulation of cell cycle process	0.007
GO:0033044~regulation of chromosome organization	0.008
GO:0050885~neuromuscular process controlling balance	0.008
GO:0051603~proteolysis involved in cellular protein catabolic process	0.008
GO:0032204~regulation of telomere maintenance	0.009
GO:0007093~mitotic cell cycle checkpoint	0.010
GO:0032845~negative regulation of homeostatic process	0.010
GO:0034645~cellular macromolecule biosynthetic process	0.010
GO:0006261~DNA-dependent DNA replication	0.011
GO:0006310~DNA recombination	0.011
GO:0019752~carboxylic acid metabolic process	0.011
GO:0007130~synaptonemal complex assembly	0.012
GO:0060271~cilium morphogenesis	0.012
GO:0044248~cellular catabolic process	0.012
GO:0006511~ubiquitin-dependent protein catabolic process	0.012
GO:0003006~developmental process involved in reproduction	0.013
GO:0060119~inner ear receptor cell development	0.013
GO:0043436~oxoacid metabolic process	0.013
GO:0007017~microtubule-based process	0.013
GO:0032006~regulation of TOR signalling	0.014
GO:1904353~regulation of telomere capping	0.014
GO:0033554~cellular response to stress	0.014
GO:0044265~cellular macromolecule catabolic process	0.014
GO:0019941~modification-dependent protein catabolic process	0.015
GO:0035058~nonmotile primary cilium assembly	0.015
GO:0044380~protein localization to cytoskeleton	0.015
GO:0045930~negative regulation of mitotic cell cycle	0.015
GO:0070192~chromosome organization involved in meiotic cell cycle	0.016

GO:2000780~negative regulation of double-strand break repair	0.017
GO:0006835~dicarboxylic acid transport	0.017
GO:0019218~regulation of steroid metabolic process	0.017
GO:0006629~lipid metabolic process	0.017
GO:0001907~killing by symbiont of host cells	0.017
GO:0001915~negative regulation of T cell mediated cytotoxicity	0.017
GO:1904354~negative regulation of telomere capping	0.017
GO:0044004~disruption by symbiont of host cell	0.017
GO:0044702~single organism reproductive process	0.018
GO:0008202~steroid metabolic process	0.018
GO:0043632~modification-dependent macromolecule catabolic process	0.018
GO:1901990~regulation of mitotic cell cycle phase transition	0.018
GO:0010927~cellular component assembly involved in morphogenesis	0.018
GO:0070925~organelle assembly	0.019
GO:0015711~organic anion transport	0.019
GO:0070193~synaptonemal complex organization	0.019
GO:0030163~protein catabolic process	0.019
GO:0007127~meiosis I	0.019
GO:0051641~cellular localization	0.020
GO:0032787~monocarboxylic acid metabolic process	0.020
GO:2001251~negative regulation of chromosome organization	0.020
GO:0006768~biotin metabolic process	0.021
GO:0035864~response to potassium ion	0.021
GO:0043933~macromolecular complex subunit organization	0.021
GO:0051726~regulation of cell cycle	0.021
GO:1901987~regulation of cell cycle phase transition	0.022
GO:0048285~organelle fission	0.022
GO:0000075~cell cycle checkpoint	0.023
GO:0006888~ER to Golgi vesicle-mediated transport	0.023
GO:0048193~Golgi vesicle transport	0.023
GO:0010212~response to ionizing radiation	0.023
GO:0051321~meiotic cell cycle	0.024
GO:0033036~macromolecule localization	0.024
GO:0014896~muscle hypertrophy	0.024
GO:0000732~strand displacement	0.025
GO:0034454~microtubule anchoring at centrosome	0.025
GO:0045738~negative regulation of DNA repair	0.025
GO:0007507~heart development	0.025
GO:0048813~dendrite morphogenesis	0.025
GO:0031589~cell-substrate adhesion	0.027
GO:0070085~glycosylation	0.028
GO:0000002~mitochondrial genome maintenance	0.028
GO:0044773~mitotic DNA damage checkpoint	0.030
GO:0048858~cell projection morphogenesis	0.031

GO:0045132~meiotic chromosome segregation	0.032
GO:0015721~bile acid and bile salt transport	0.032
GO:0006664~glycolipid metabolic process	0.033
GO:0050905~neuromuscular process	0.034
GO:0000280~nuclear division	0.034
GO:0051052~regulation of DNA metabolic process	0.034
GO:0010389~regulation of G2/M transition of mitotic cell cycle	0.034
GO:0048814~regulation of dendrite morphogenesis	0.034
GO:0048646~anatomical structure formation involved in morphogenesis	0.035
GO:2000779~regulation of double-strand break repair	0.035
GO:0000226~microtubule cytoskeleton organization	0.035
GO:1902806~regulation of cell cycle G1/S phase transition	0.035
GO:0072358~cardiovascular system development	0.036
GO:0072359~circulatory system development	0.036
GO:0006082~organic acid metabolic process	0.036
GO:0000819~sister chromatid segregation	0.036
GO:0043413~macromolecule glycosylation	0.037
GO:0006486~protein glycosylation	0.037
GO:0006687~glycosphingolipid metabolic process	0.037
GO:1901991~negative regulation of mitotic cell cycle phase transition	0.037
GO:0051640~organelle localization	0.037
GO:0006493~protein O-linked glycosylation	0.037
GO:0051128~regulation of cellular component organization	0.038
GO:0008206~bile acid metabolic process	0.038
GO:0070647~protein modification by small protein conjugation or removal	0.038
GO:1903509~liposaccharide metabolic process	0.038
GO:0007018~microtubule-based movement	0.039
GO:0044843~cell cycle G1/S phase transition	0.039
GO:0000902~cell morphogenesis	0.039
GO:0045922~negative regulation of fatty acid metabolic process	0.040
GO:0072698~protein localization to microtubule cytoskeleton	0.040
GO:0010467~gene expression	0.041
GO:0007160~cell-matrix adhesion	0.041
GO:0010468~regulation of gene expression	0.042
GO:0046939~nucleotide phosphorylation	0.042
GO:0031023~microtubule organizing center organization	0.042
GO:0071103~DNA conformation change	0.042
GO:0031570~DNA integrity checkpoint	0.043
GO:2000045~regulation of G1/S transition of mitotic cell cycle	0.044
GO:0009812~flavonoid metabolic process	0.044
GO:0032990~cell part morphogenesis	0.044
GO:0000726~non-recombinational repair	0.044
GO:0072393~microtubule anchoring at microtubule organizing center	0.045
GO:0032210~regulation of telomere maintenance via telomerase	0.045

Appendix 13: Genes with protein-altering mutations in EBV-positive cases in my series

ABCA1	CCDC73	FAM228A	MCTP2	PTK2	ST3GAL6
ABCA5	CCNL1	FANCL	MED24	PTPRC	STAG1
ABCC5	CCT6B	FAR1	MFN1	RAB14	STAMBPL1
ADAMTS12	CD34	FAT3	MICU3	RAD50	STAT5B
ADAMTSL4	CDC27	FBXO40	MKS1	RAD51C	STOX1
AGO4	CDC7	FBXW10	MLLT4	RAPGEF6	SUCLA2
AGPAT9	CDCA2	FGF7	MROH8	RASGEF1B	SUGP2
ALB	CDH3	FRG1B	MRPL15	RASGRP2	SUN3
ANAPC5	CDK5RAP2	FUNDC1	MRPS18B	RB1	SYNE1
ANK2	CDYL	GALC	MRPS30	RBM44	TAB2
ANKMY2	CEACAM8	GALNT10	MRRF	RBM48	TAF15
ANKRD12	CEP192	GAPVD1	MSL2	RC3H1	TARS
ANKRD36	CEP76	GARS	MTDH	RDH12	TEX11
ANKS1B	CHD2	GPR128	MTMR7	RFC3	TMEM196
APIP	CHD4	GSS	MTRR	RNF138	TMEM50A
APOB	CLCA4	GTF2I	MUC12	ROBO2	TMEM57
ARFGAP3	CLCN3	HBS1L	MUC19	ROS1	TMPRSS6
ARHGAP10	CLEC12A	HCFC2	MUC4	RP11-53I6.2	TNXB
ARID4A	CMPK1	HEATR3	MYO6	RPL9	TOP2B
ARL14EP	CMTR2	HEATR5B	MYO9A	RPS6KA6	TPP1
ARV1	CNTN4	HLF	MYOM2	RRH	TPTE2
ATAD2	COL9A1	HMCN1	NAT10	SAMHD1	TRDN
ATL1	CSMD3	HNRNPC	NAV3	SBF2	TRIM49C
ATM	CTC-432M15.3	HSPA13	NCOR1	SCN9A	TRIOBP
ATXN10	CWC27	HSPA8	NDUFB5	SEC31A	TRIP12
ATXN2	CYSLTR1	IGLC3	NEDD4L	SEC61A2	TTBK2
B3GALNT1	DCDC1	IKZF5	NF1	SF3B2	TTN
BBS2	DEFB115	IL12RB2	NSUN2	SF3B3	TUBA1C
BCLAF1	DIAPH1	INSIG2	OCIAD2	SGOL2	TULP3
BCS1L	DIAPH3	IPP	OFD1	SLC1A1	UBE3A
BEND5	DLEC1	ITPR2	OR10A6	SLC20A1	UBR4
BRD9	DLEU2L	JMJD1C	OR1N1	SLC25A12	UGT2B17
BSPH1	DNAH12	KCNA1	OR4C5	SLC25A13	USH1C
BTBD1	DNAH6	KCNB2	PDCD11	SLC25A3	USH2A
BTBD9	DOCK10	KIAA0895	PDE4DIP	SLC30A3	VAV2
C10orf76	DOCK7	KIAA0922	PDE7A	SLC35A1	VWDE
C11orf54	DUS2	KIAA1009	PDS5B	SLC5A1	WDR35
C12orf45	DYNC1LI1	KIAA1462	PDZD2	SLC5A6	WDR64
C14orf23	DYNC2H1	KIAA1551	PGM1	SLC6A14	XPNPEP1
C17orf85	EHBP1	KIAA1715	PHF23	SLC9A2	ZCRB1
C1orf27	EIF3H	KIF3A	PIK3CB	SLCO1C1	ZNF292
C20orf194	EML4	KLHL1	PLCZ1	SLCO6A1	ZNF321P

C20orf24	EPC1	LAMA3	PMS1	SMARCAL1	ZNF416
C5orf22	EPG5	LAMA4	POLQ	SMARCC1	ZNF439
C6orf89	EPN2	LBR	PPFIA1	SMG1	ZNF625
C7orf31	ERCC4	LIMCH1	PPIL6	SNCAIP	ZNF625-ZNF20
C9	ERCC6	LMAN1	PPP1R10	SNRNP200	ZNF714
CACNG3	ERCC6L2	LMLN	PPP6C	SON	ZNF717
CAMSAP2	ERVMER34-1	LRP6	PREX2	SPAG17	ZNF816
CARF	F5	MAN1A2	PRKRA	SPATS1	ZNF91
CASD1	F8	MBD4	PROM1	SPTAN1	ZNF93
CASP2	FAF1	MCCC1	PROS1	SREK1	ZNF99
CCDC175	FAIM3	MCHR2	PRPF3	SRPR	
CCDC30	FAM114A2	MCM6	PSG3	SRXN1	

Appendix 14: Genes with protein-altering mutations in EBV-negative cases in my series

ABCA1	CEP78	GK5	MUC17	R3HCC1L	TBL3
ABCC2	CHD1	GLDC	MUC7	R3HDM2	TBP
ABHD14A-ACY1	CHD1L	GLT6D1	MUT	RAB3GAP2	TCOF1
ABI3BP	CHD2	GLUD2	MX1	RAD50	TCTN2
AC012215.1	CHD9	GLUL	MYCBP2	RALGPS2	TEP1
AC112715.2	CIT	GNB2L1	MYEF2	RANBP2	TEX10
ACACA	CKLF-CMTM1	GNE	MYH4	RBBP7	TEX15
ACAD10	CLOCK	GNL2	MYH8	RBFOX2	TFDP1
ACBD3	CLSPN	GPCPD1	MYO10	RBM27	TIGD7
ACVR1	CMTM1	GPR98	MYO18B	RELN	TLN2
ADAM9	CNTN1	GTF3C1	MYO5B	RGPD3	TMEM131
ADAMTS12	COG1	HCFC2	MYOM1	RGS6	TMEM136
ADAMTS3	CORO7	HEATR5A	NAA15	RIF1	TMEM156
ADAMTS6	CORO7-PAM16	HECTD1	NAA16	RMDN2	TMEM183A
ADAT2	CPNE3	HEG1	NASP	RNASEH2A	TMEM97
AHNAK	CSE1L	HEPH	NAV3	RNF145	TMTC2
AHNAK2	CSMD3	HIATL2	NBN	RNF169	TMTC3
AKAP13	CTCF	HIVEP1	NCAPG2	RNF213	TNFAIP3
AKAP8	CTCFL	HKDC1	NDUFB5	RNF217	TNKS2
AKT2	CTD-2370N5.3	HLA-DQA1	NEK11	RNF7	TOB1
AKT3	CTSA	HNRNPDL	NEMF	RRAGB	TOPBP1
ALDH18A1	CTTNBP2	HSD17B4	NF1	RSPRY1	TP53BP2
ALDH3A2	CUBN	HTR2A	NFE2L1	RTCB	TPTE2
ALK	CUX1	ICA1	NFKB1	RUFY1	TRAPPC8
ALMS1	CYB5R4	IFIT3	NHS	RYR2	TRHR
ALPK2	CYLD	IFRD1	NIN	S100B	TRIM26
ALPK3	CYP1B1	IGHV1-18	NKRF	SARDH	TRIM66
AMY2B	CYP7B1	IGHV1-69	NOL11	SASS6	TRIM72
ANKHD1	CYSLTR1	IGHV2-70	NOL4	SC5D	TRIOBP
ANKHD1-EIF4EBP3	DDX20	IGHV2OR16-5	NPIPA5	SCAI	TRPS1
ANKRD12	DDX6	IGKV1-6	NPIPB5	SCO1	TSPAN5
ANO5	DDX60	IL17RB	NR1H4	SCOC	TTC12
ANP32D	DEPDC7	IMPACT	NUDT19	SEC23B	TTN
ANXA2	DHX38	INADL	NUGGC	SEL1L	TXNRD3
APIP	DHX9	INVS	NUP210L	SENP2	UBE4B
APOL5	DLEU2L	IPO7	NUP50	SETD2	UBN1
APPL1	DLG2	IPP	OCA2	SETD7	UBR5
ARFGEF2	DMTF1	IRF8	ODC1	SETX	UGGT2
ARHGAP25	DNAJC10	ITGA10	OR4C5	SF3B2	UGT1A3

ARL13B	DOCK5	ITGB7	OR4F4	SGOL1	UGT1A5
ARMC10	DOCK9	ITGB8	OR4Q2	SGOL2	UGT1A8
ARNT	DPCR1	ITIH1	OR52E2	SIM1	UNC13C
ARSF	DPRX	ITPR1	OR52R1	SLC12A2	USP15
ASCC3	DPYD	ITPR2	OR5AR1	SLC14A2	USP26
ATAD2	DSPP	JAG1	OR6C75	SLC22A14	USP45
ATF7	DTHD1	JAKMIP1	OTUD4	SLC25A14	USP46
ATG16L1	DYX1C1	JMJD1C	OTUD7A	SLC25A33	UTS2B
ATPAF1	EAPP	KDM8	OVCH1	SLC26A5	VAMP4
AVL9	ECT2	KDR	PAFAH1B3	SLC26A7	VCAM1
BAZ2B	EDRF1	KIAA0125	PALB2	SLC28A2	VEGFB
BBS12	EFHC1	KIAA1009	PANK1	SLC35F4	VPS13A
BBS2	EGFLAM	KIAA1199	PAQR5	SLC44A4	VRK2
BBS4	EIF3H	KIAA2026	PARP14	SLC4A1AP	WBP11
BCR	EIF4G2	KIF24	PBX4	SLC4A4	WDR43
BHMT	EIF4G3	KIF3A	PCCA	SLC4A7	WDR87
BIRC6	ELF5	KIF5A	PCDHAC1	SLC7A6	WNT2B
BIVM-ERCC5	ELOVL7	KLHL1	PCOLCE2	SLCO1B3	XPO1
BNIP2	EML6	KLHL7	PDCD2L	SMARCA2	XRCC4
BRD9	EMR1	KLRD1	PDCD6	SMARCC1	YKT6
BSCL2	ENPP3	KMT2A	PDCL2	SMC1A	YME1L1
BSPH1	ENTHD1	KNDC1	PDLIM5	SMC1B	YTHDF3
BTBD1	EP300	KNOP1	PDSS1	SMC6	ZBED6
C10ORF68	EPN2	KRT76	PGAP1	SMEK2	ZBTB21
C11orf30	ERAP1	LAMA3	PGBD1	SMG1	ZDBF2
C12orf45	ERCC5	LAMTOR2	PIBF1	SMG5	ZFR
C1orf110	ESRP1	LEPREL1	PIGK	SNAP29	ZIK1
C1orf87	ETF1	LIFR	PITPNC1	SNAPC1	ZNF12
C1S	EVI2B	LNX1	PIWIL3	SNRNP200	ZNF143
C20orf194	EWSR1	LPA	PIWIL4	SOGA3	ZNF157
C2orf66	EXD1	LPGAT1	PKHD1	SORBS2	ZNF23
C5	F5	LPPR5	PKLR	SP4	ZNF253
C6orf10	FADS2	LRBA	PLCH1	SPAG9	ZNF263
C7orf25	FAHD2B	LRPPRC	PLS3	SPG11	ZNF292
C8orf37	FAM111B	LRRC9	PLSCR1	SQLE	ZNF33B
CACHD1	FAM120B	LRRIQ1	PNRC1	SREK1	ZNF343
CALCOCO2	FAM210A	LRTOMT	POLG	SRRM2	ZNF384
CASP10	FAM83D	LSM5	PPARGC1A	SSB	ZNF430
CATSPERD	FAM98A	LST3	PPFIA4	ST3GAL6	ZNF462
CCDC113	FAS	LYSMD3	PPP3CA	ST8SIA4	ZNF530
CCDC129	FBXO45	MACF1	PPP3CB	ST8SIA6	ZNF549
CCDC73	FBXO7	MAP3K4	PREPL	STAT2	ZNF560
CCDC80	FCRL5	MBNL2	PRKDC	STK17B	ZNF567

CCIN	FER1L6	MCM9	PRKRA	STK31	ZNF577
CCT6B	FERMT1	MCPH1	PROM1	STK38	ZNF644
CD109	FHL5	MED13L	PROS1	STK38L	ZNF660
CD96	FHOD1	MEF2A	PROSER1	STT3A	ZNF662
CDH10	FMNL2	MESDC2	PRRC2C	SVEP1	ZNF682
CDH23	FNBP1	MFAP5	PSMC5	SYCP1	ZNF69
CDH8	FNDC3B	MICU3	PSME3	SYCP2	ZNF700
CDHR5	FNDC7	MIER3	PTEN	SYNE1	ZNF708
CDK13	FREM2	MLEC	PTPN14	SYNPO	ZNF717
CDK5R1	FSIP1	MOGS	PTPN21	TAAR8	ZNF724P
CDKL4	GABRB1	MSANTD2	PTPRD	TAF2	ZNF729
CDR2	GAPT	MSH2	PTPRQ	TANK	ZNF793
CEACAM1	GARS	MSTN	PTX3	TARDBP	ZNF799
CEBPZ	GBA3	MTBP	PUS10	TBC1D1	ZNF804B
CEP57	GCNT2	MTTP	PUS7	TBC1D7	ZSCAN32
CEP76	GIGYF2	MUC16	QSER1	TBCEL	

Appendix 15: List of most frequently mutated genes in 23 primary HL cases

Gene	Number of patients mutated	EBV status	Gene	Number of patients mutated	EBV status
F8	3	POS	**NUP50**	2	NEG
TPTE2	3	POS	**PCLO**	2	NEG
AGPAT9	2	POS	**SLIT2**	2	NEG
CCDC175	2	POS	**VPS13C**	2	NEG
CHD4	2	POS	**ZNF430**	2	NEG
CMPK1	2	POS	**EGFLAM**	2	NEG
COL9A1	2	POS	**IGHV1-24**	2	NEG
DNAH12	2	POS	**IGLV3-10**	2	NEG
EML4	2	POS	**LAMA3**	2	NEG
EPG5	2	POS	**MFSD5**	2	NEG
ITPR2	2	POS	**NPAS3**	2	NEG
LAMA4	2	POS	**TRAPPC8**	2	NEG
MUC4	2	POS	**ZBTB21**	2	NEG
MYO6	2	POS	**ADSSL1**	2	NEG
OR4C5	2	POS	**CTTNBP2**	2	NEG
PDE4DIP	2	POS	**EBF1**	2	NEG
SBF2	2	POS	**FNDC7**	2	NEG
SGOL2	2	POS	**IGHG4**	2	NEG
SMG1	2	POS	**MUC2**	2	NEG
TMEM57	2	POS	**MYCBP2**	2	NEG
TRDN	2	POS	**PCDHA5**	2	NEG
TTN	2	POS	**POLE**	2	NEG
ZNF717	2	POS	**RALGPS2**	2	NEG
IGHV2-70	8	NEG	**RYR2**	2	NEG
TNFAIP3	7	NEG	**SF3B2**	2	NEG
IGHE	5	NEG	**SLC14A2**	2	NEG
TTN	5	NEG	**UBR5**	2	NEG
IGLV3-1	5	NEG	**UNC13C**	2	NEG
MUC17	5	NEG	**ZFHX3**	2	NEG
HIST1H1E	5	NEG	**ZFP36L1**	2	NEG
MUC4	4	NEG	**ANXA1**	2	NEG
B2M	4	NEG	**BNIP2**	2	NEG
IGLL5	4	NEG	**CTD-2370N5.3**	2	NEG
ITPKB	4	NEG	**DNAH11**	2	NEG
EEF1A1	4	NEG	**FLNC**	2	NEG
SOCS1	4	NEG	**HECW2**	2	NEG
ITPR1	3	NEG	**IGLV2-23**	2	NEG
LRRN3	3	NEG	**INVS**	2	NEG

CSMD3	3	NEG	**PROCA1**	2	NEG
HECTD1	3	NEG	**SEC23IP**	2	NEG
SMG1	3	NEG	**TRIP11**	2	NEG
UBE2A	3	NEG	**ABL1**	2	NEG
DNAH14	3	NEG	**ADAMTS12**	2	NEG
SRRM2	3	NEG	**CCDC60**	2	NEG
C1orf87	3	NEG	**CKLF-CMTM1**	2	NEG
CCDC80	3	NEG	**CTD-3088G3.8**	2	NEG
GPR98	3	NEG	**EAPP**	2	NEG
IGHJ5	3	NEG	**F5**	2	NEG
CHD1	3	NEG	**FAT4**	2	NEG
MUC16	3	NEG	**GNA13**	2	NEG
RANBP2	3	NEG	**GTF3C1**	2	NEG
ACTG1	2	NEG	**HCFC2**	2	NEG
LPHN3	2	NEG	**IGHV1-69**	2	NEG
CMTM1	2	NEG	**LRP1B**	2	NEG
ANXA2	2	NEG	**NBPF10**	2	NEG
ARIH2	2	NEG	**NHS**	2	NEG
MACF1	2	NEG	**PTPN21**	2	NEG
PTPRD	2	NEG	**SETDB1**	2	NEG
ACTB	2	NEG	**SMC1B**	2	NEG
CYLD	2	NEG	**SREK1**	2	NEG
SOX6	2	NEG	**SYCP2**	2	NEG
EFCAB5	2	NEG	**TICRR**	2	NEG
NOL4	2	NEG	**USP15**	2	NEG
PROM1	2	NEG	**WDR87**	2	NEG
ALMS1	2	NEG	**WEE1**	2	NEG
KLHL7	2	NEG	**ZFC3H1**	2	NEG
PLS3	2	NEG	**ZNF292**	2	NEG
ANKHD1	2	NEG	**ZNF644**	2	NEG
DMD	2	NEG	**ALK**	2	NEG
GNE	2	NEG	**CUBN**	2	NEG
CDH23	2	NEG	**GPATCH2**	2	NEG
COL4A6	2	NEG	**IGKV1D-8**	2	NEG
CPM	2	NEG	**ITPR2**	2	NEG
DMTF1	2	NEG	**PKHD1**	2	NEG
HEPH	2	NEG	**SLCO1B7**	2	NEG
NR1H4	2	NEG	**TMEM235**	2	NEG
BCL7A	2	NEG	**ABCA13**	2	NEG
CRYM	2	NEG	**ALOX12**	2	NEG
FSIP2	2	NEG	**ANKHD1-EIF4EBP3**	2	NEG
GPR112	2	NEG	**CCDC73**	2	NEG
NEK1	2	NEG	**DDX20**	2	NEG
NF1	2	NEG	**DPYD**	2	NEG

RNF145	2	NEG	**FLG**	2	NEG
SPTB	2	NEG	**GLUD2**	2	NEG
TBC1D15	2	NEG	**IGKJ2**	2	NEG
XRCC4	2	NEG	**IGKV2-24**	2	NEG
ARHGAP25	2	NEG	**IGLV7-46**	2	NEG
ECT2	2	NEG	**LRRIQ1**	2	NEG
IGKV4-1	2	NEG	**LST3**	2	NEG
PRRC2C	2	NEG	**MUC19**	2	NEG
PRSS3	2	NEG	**MYO5B**	2	NEG
ZNF217	2	NEG	**OR4C5**	2	NEG
ACACA	2	NEG	**OR52E2**	2	NEG
CHD1L	2	NEG	**PIM2**	2	NEG
CSF2RB	2	NEG	**QSER1**	2	NEG
DGKI	2	NEG	**SMCR8**	2	NEG
FNBP1	2	NEG	**STAB2**	2	NEG
HNRNPDL	2	NEG	**STK38**	2	NEG
IGLV3-25	2	NEG	**UGT1A8**	2	NEG
MYO18B	2	NEG	**ZFR**	2	NEG
NPIPB5	2	NEG	**ZNF670**	2	NEG

Appendix 16: Biological processes enriched in EBV-negative tumours my series

Gene ontology term	*p*-value
GO:0051276~chromosome organization	0.002
GO:0060249~anatomical structure homeostasis	0.002
GO:0006464~cellular protein modification process	0.003
GO:0036211~protein modification process	0.003
GO:0014896~muscle hypertrophy	0.003
GO:0051171~regulation of nitrogen compound metabolic process	0.004
GO:0051603~proteolysis involved in cellular protein catabolic process	0.005
GO:0048814~regulation of dendrite morphogenesis	0.005
GO:0044257~cellular protein catabolic process	0.005
GO:0007507~heart development	0.006
GO:0010468~regulation of gene expression	0.006
GO:0010564~regulation of cell cycle process	0.007
GO:0032508~DNA duplex unwinding	0.008
GO:0000002~mitochondrial genome maintenance	0.008
GO:0044248~cellular catabolic process	0.008
GO:0048813~dendrite morphogenesis	0.009
GO:0060113~inner ear receptor cell differentiation	0.009
GO:0006664~glycolipid metabolic process	0.009
GO:0006281~DNA repair	0.010
GO:1903509~liposaccharide metabolic process	0.010
GO:0032392~DNA geometric change	0.011
GO:0060119~inner ear receptor cell development	0.011
GO:0044265~cellular macromolecule catabolic process	0.011
GO:0007049~cell cycle	0.012
GO:0071103~DNA conformation change	0.012
GO:0016358~dendrite development	0.012
GO:0042490~mechanoreceptor differentiation	0.012
GO:0034454~microtubule anchoring at centrosome	0.012
GO:0006629~lipid metabolic process	0.013
GO:2000112~regulation of cellular macromolecule biosynthetic process	0.013
GO:0034645~cellular macromolecule biosynthetic process	0.013
GO:0006511~ubiquitin-dependent protein catabolic process	0.015
GO:0044774~mitotic DNA integrity checkpoint	0.015
GO:0033036~macromolecule localization	0.015
GO:0019219~regulation of nucleobase-containing compound metabolic process	0.015
GO:0045444~fat cell differentiation	0.016
GO:0007286~spermatid development	0.016
GO:0042787~protein ubiquitination involved in ubiquitin-dependent protein catabolic process	0.016
GO:0003006~developmental process involved in reproduction	0.017

GO:0019941~modification-dependent protein catabolic process	0.017
GO:0045494~photoreceptor cell maintenance	0.018
GO:0035058~nonmotile primary cilium assembly	0.018
GO:0044380~protein localization to cytoskeleton	0.018
GO:0006974~cellular response to DNA damage stimulus	0.018
GO:0022402~cell cycle process	0.019
GO:0030163~protein catabolic process	0.019
GO:0044782~cilium organization	0.019
GO:0016070~RNA metabolic process	0.020
GO:0006643~membrane lipid metabolic process	0.020
GO:0048515~spermatid differentiation	0.020
GO:0043632~modification-dependent macromolecule catabolic process	0.020
GO:0016053~organic acid biosynthetic process	0.021
GO:0006687~glycosphingolipid metabolic process	0.021
GO:0050885~neuromuscular process controlling balance	0.021
GO:0034453~microtubule anchoring	0.021
GO:0006351~transcription, DNA-templated	0.021
GO:0007281~germ cell development	0.021
GO:0042384~cilium assembly	0.021
GO:0006302~double-strand break repair	0.021
GO:0032006~regulation of TOR signalling	0.022
GO:0072393~microtubule anchoring at microtubule organizing center	0.022
GO:0031589~cell-substrate adhesion	0.022
GO:0010038~response to metal ion	0.024
GO:0006913~nucleocytoplasmic transport	0.024
GO:0019752~carboxylic acid metabolic process	0.024
GO:0044702~single organism reproductive process	0.024
GO:0010035~response to inorganic substance	0.025
GO:0051592~response to calcium ion	0.025
GO:0010556~regulation of macromolecule biosynthetic process	0.026
GO:0051128~regulation of cellular component organization	0.026
GO:0051641~cellular localization	0.026
GO:0046394~carboxylic acid biosynthetic process	0.027
GO:0043436~oxoacid metabolic process	0.027
GO:1904353~regulation of telomere capping	0.027
GO:0008104~protein localization	0.027
GO:0051169~nuclear transport	0.028
GO:0052697~xenobiotic glucuronidation	0.028
GO:0060537~muscle tissue development	0.028
GO:0032886~regulation of microtubule-based process	0.029
GO:0032787~monocarboxylic acid metabolic process	0.030
GO:0044773~mitotic DNA damage checkpoint	0.030
GO:0033365~protein localization to organelle	0.032
GO:0045930~negative regulation of mitotic cell cycle	0.032

GO:0050773~regulation of dendrite development	0.033
GO:0003300~cardiac muscle hypertrophy	0.033
GO:0000077~DNA damage checkpoint	0.034
GO:0042592~homeostatic process	0.034
GO:0022603~regulation of anatomical structure morphogenesis	0.034
GO:0014706~striated muscle tissue development	0.035
GO:0006259~DNA metabolic process	0.035
GO:0055006~cardiac cell development	0.035
GO:0008610~lipid biosynthetic process	0.036
GO:0022604~regulation of cell morphogenesis	0.036
GO:0001894~tissue homeostasis	0.036
GO:0033554~cellular response to stress	0.037
GO:0014897~striated muscle hypertrophy	0.038
GO:0045833~negative regulation of lipid metabolic process	0.038
GO:0007276~gamete generation	0.039
GO:0072359~circulatory system development	0.039
GO:0072358~cardiovascular system development	0.039
GO:0048646~anatomical structure formation involved in morphogenesis	0.039
GO:0021987~cerebral cortex development	0.040
GO:0044770~cell cycle phase transition	0.040
GO:1902806~regulation of cell cycle G1/S phase transition	0.041
GO:0001573~ganglioside metabolic process	0.042
GO:1903047~mitotic cell cycle process	0.042
GO:0015701~bicarbonate transport	0.042
GO:0070646~protein modification by small protein removal	0.042
GO:0006355~regulation of transcription, DNA-templated	0.043
GO:0071705~nitrogen compound transport	0.045
GO:0016310~phosphorylation	0.046
GO:0009062~fatty acid catabolic process	0.046
GO:0021543~pallium development	0.047
GO:0010565~regulation of cellular ketone metabolic process	0.048
GO:0031570~DNA integrity checkpoint	0.048
GO:0000723~telomere maintenance	0.048
GO:0010225~response to UV-C	0.048
GO:0007093~mitotic cell cycle checkpoint	0.050
GO:1903506~regulation of nucleic acid-templated transcription	0.050
GO:0009247~glycolipid biosynthetic process	0.050

Appendix 17: Biological processes enriched in EBV-positive tumours in my series

Gene ontology term	*p*-value
GO:0060249~anatomical structure homeostasis	1.1E-06
GO:1902589~single-organism organelle organization	4.3E-05
GO:0043933~macromolecular complex subunit organization	4.9E-05
GO:0001894~tissue homeostasis	9.0E-05
GO:0001895~retina homeostasis	1.3E-04
GO:0044085~cellular component biogenesis	1.4E-04
GO:0045494~photoreceptor cell maintenance	2.0E-04
GO:0022607~cellular component assembly	3.0E-04
GO:0051053~negative regulation of DNA metabolic process	5.8E-04
GO:0032205~negative regulation of telomere maintenance	7.4E-04
GO:0042592~homeostatic process	9.8E-04
GO:0007131~reciprocal meiotic recombination	1.0E-03
GO:0035825~reciprocal DNA recombination	1.0E-03
GO:0051304~chromosome separation	1.1E-03
GO:0000280~nuclear division	1.2E-03
GO:0006820~anion transport	1.6E-03
GO:2001251~negative regulation of chromosome organization	1.9E-03
GO:0048871~multicellular organismal homeostasis	1.9E-03
GO:0000724~double-strand break repair via homologous recombination	2.2E-03
GO:0032845~negative regulation of homeostatic process	2.2E-03
GO:0006888~ER to Golgi vesicle-mediated transport	2.3E-03
GO:0000725~recombinational repair	2.3E-03
GO:0007017~microtubule-based process	2.5E-03
GO:0048285~organelle fission	2.7E-03
GO:0007010~cytoskeleton organization	3.0E-03
GO:0007059~chromosome segregation	3.1E-03
GO:0098656~anion transmembrane transport	4.6E-03
GO:0010639~negative regulation of organelle organization	4.9E-03
GO:1903047~mitotic cell cycle process	4.9E-03
GO:0051276~chromosome organization	5.1E-03
GO:0000723~telomere maintenance	6.0E-03
GO:0098813~nuclear chromosome segregation	7.0E-03
GO:0000278~mitotic cell cycle	7.4E-03
GO:0051235~maintenance of location	7.5E-03
GO:0007127~meiosis I	8.0E-03
GO:0015810~aspartate transport	8.6E-03
GO:0070925~organelle assembly	9.6E-03
GO:0000086~G2/M transition of mitotic cell cycle	9.7E-03
GO:0000226~microtubule cytoskeleton organization	9.8E-03

GO:0032200~telomere organization	9.8E-03
GO:0007067~mitotic nuclear division	1.0E-02
GO:0051301~cell division	1.2E-02
GO:0045132~meiotic chromosome segregation	1.2E-02
GO:0015740~C4-dicarboxylate transport	1.3E-02
GO:0044839~cell cycle G2/M phase transition	1.5E-02
GO:0048193~Golgi vesicle transport	1.5E-02
GO:0044770~cell cycle phase transition	1.6E-02
GO:0065003~macromolecular complex assembly	1.6E-02
GO:0070192~chromosome organization involved in meiotic cell cycle	1.7E-02
GO:0048208~COPII vesicle coating	1.7E-02
GO:0048207~vesicle targeting, rough ER to cis-Golgi	1.7E-02
GO:0089711~L-glutamate transmembrane transport	1.8E-02
GO:0006259~DNA metabolic process	1.8E-02
GO:0033044~regulation of chromosome organization	1.9E-02
GO:0006901~vesicle coating	1.9E-02
GO:0022402~cell cycle process	1.9E-02
GO:0071822~protein complex subunit organization	1.9E-02
GO:0044772~mitotic cell cycle phase transition	2.0E-02
GO:0008589~regulation of smoothened signalling pathway	2.1E-02
GO:0048199~vesicle targeting, to, from or within Golgi	2.1E-02
GO:0090114~COPII-coated vesicle budding	2.1E-02
GO:0032366~intracellular sterol transport	2.1E-02
GO:0006974~cellular response to DNA damage stimulus	2.1E-02
GO:0032204~regulation of telomere maintenance	2.4E-02
GO:0006281~DNA repair	2.5E-02
GO:0006310~DNA recombination	2.7E-02
GO:0032984~macromolecular complex disassembly	2.7E-02
GO:0000712~resolution of meiotic recombination intermediates	2.7E-02
GO:0071826~ribonucleoprotein complex subunit organization	2.8E-02
GO:0007049~cell cycle	2.9E-02
GO:0006811~ion transport	2.9E-02
GO:0000819~sister chromatid segregation	2.9E-02
GO:0007091~metaphase/anaphase transition of mitotic cell cycle	3.0E-02
GO:0030301~cholesterol transport	3.0E-02
GO:0007018~microtubule-based movement	3.0E-02
GO:0032844~regulation of homeostatic process	3.1E-02
GO:0043603~cellular amide metabolic process	3.1E-02
GO:0070894~regulation of transposon integration	3.2E-02
GO:0043490~malate-aspartate shuttle	3.2E-02
GO:0070895~negative regulation of transposon integration	3.2E-02
GO:0070893~transposon integration	3.2E-02
GO:0007224~smoothened signalling pathway	3.3E-02
GO:0015918~sterol transport	3.3E-02

GO:0010971~positive regulation of G2/M transition of mitotic cell cycle	3.3E-02
GO:0015813~L-glutamate transport	3.3E-02
GO:1903046~meiotic cell cycle process	3.4E-02
GO:0032508~DNA duplex unwinding	3.4E-02
GO:0006835~dicarboxylic acid transport	3.4E-02
GO:0006903~vesicle targeting	3.6E-02
GO:1902751~positive regulation of cell cycle G2/M phase transition	3.7E-02
GO:0051052~regulation of DNA metabolic process	3.8E-02
GO:0006312~mitotic recombination	3.8E-02
GO:0051298~centrosome duplication	4.0E-02
GO:0051307~meiotic chromosome separation	4.1E-02
GO:0015800~acidic amino acid transport	4.1E-02
GO:0006897~endocytosis	4.1E-02
GO:0032392~DNA geometric change	4.2E-02
GO:0051592~response to calcium ion	4.2E-02
GO:0045995~regulation of embryonic development	4.2E-02
GO:0015711~organic anion transport	4.2E-02
GO:0051650~establishment of vesicle localization	4.3E-02
GO:0048259~regulation of receptor-mediated endocytosis	4.3E-02
GO:0010965~regulation of mitotic sister chromatid separation	4.4E-02
GO:0000729~DNA double-strand break processing	4.4E-02
GO:0016071~mRNA metabolic process	4.5E-02
GO:0051129~negative regulation of cellular component organization	4.5E-02
GO:0044784~metaphase/anaphase transition of cell cycle	4.6E-02
GO:0060271~cilium morphogenesis	4.7E-02
GO:0019836~hemolysis by symbiont of host erythrocytes	4.8E-02
GO:0051715~cytolysis in other organism	4.8E-02
GO:0001897~cytolysis by symbiont of host cells	4.8E-02
GO:0044179~hemolysis in other organism	4.8E-02
GO:0052331~hemolysis in other organism involved in symbiotic interaction	4.8E-02
GO:0051801~cytolysis in other organism involved in symbiotic interaction	4.8E-02
GO:0045787~positive regulation of cell cycle	4.8E-02
GO:0051306~mitotic sister chromatid separation	4.9E-02
GO:0010927~cellular component assembly involved in morphogenesis	4.9E-02
GO:0042384~cilium assembly	5.0E-02
GO:0006302~double-strand break repair	5.0E-02

Appendix 18: Biological processes enriched in 132 genes up-regulated by BILF1 in DG-75 cells

Gene ontology term	*p*-value
GO:1903900~regulation of viral life cycle	0.002
GO:0035914~skeletal muscle cell differentiation	0.004
GO:0012501~programmed cell death	0.005
GO:0001938~positive regulation of endothelial cell proliferation	0.005
GO:0051384~response to glucocorticoid	0.006
GO:0006915~apoptotic process	0.007
GO:0031960~response to corticosteroid	0.009
GO:0008219~cell death	0.009
GO:1903901~negative regulation of viral life cycle	0.011
GO:0050792~regulation of viral process	0.011
GO:0007519~skeletal muscle tissue development	0.011
GO:0043900~regulation of multi-organism process	0.013
GO:0048566~embryonic digestive tract development	0.013
GO:0060538~skeletal muscle organ development	0.013
GO:0051591~response to cAMP	0.014
GO:0007339~binding of sperm to zona pellucida	0.015
GO:0001936~regulation of endothelial cell proliferation	0.015
GO:0043903~regulation of symbiosis, encompassing mutualism through parasitism	0.015
GO:0048525~negative regulation of viral process	0.018
GO:0051924~regulation of calcium ion transport	0.021
GO:0001935~endothelial cell proliferation	0.021
GO:0034097~response to cytokine	0.021
GO:0051592~response to calcium ion	0.023
GO:0035556~intracellular signal transduction	0.024
GO:0035036~sperm-egg recognition	0.025
GO:1990441~negative regulation of transcription from RNA polymerase II promoter in response to endoplasmic reticulum stress	0.025
GO:0002763~positive regulation of myeloid leukocyte differentiation	0.026
GO:0050790~regulation of catalytic activity	0.027
GO:0010646~regulation of cell communication	0.027
GO:0008277~regulation of G-protein coupled receptor protein signalling pathway	0.027
GO:0071277~cellular response to calcium ion	0.029
GO:0009725~response to hormone	0.030
GO:0023051~regulation of signalling	0.032
GO:0050673~epithelial cell proliferation	0.032
GO:0046683~response to organophosphorus	0.032
GO:0002521~leukocyte differentiation	0.034
GO:0009628~response to abiotic stimulus	0.036
GO:0000122~negative regulation of transcription from RNA polymerase II promoter	0.037

GO:0071495~cellular response to endogenous stimulus	0.038
GO:0070372~regulation of ERK1 and ERK2 cascade	0.039
GO:0006816~calcium ion transport	0.040
GO:0006470~protein dephosphorylation	0.042
GO:0007249~I-kappaB kinase/NF-kappaB signalling	0.043
GO:0014074~response to purine-containing compound	0.043
GO:0043547~positive regulation of GTPase activity	0.043
GO:1902533~positive regulation of intracellular signal transduction	0.043
GO:0043618~regulation of transcription from RNA polymerase II promoter in response to stress	0.044
GO:0051345~positive regulation of hydrolase activity	0.044
GO:0009988~cell-cell recognition	0.046
GO:0070588~calcium ion transmembrane transport	0.046
GO:0043901~negative regulation of multi-organism process	0.046
GO:0045682~regulation of epidermis development	0.047
GO:0043549~regulation of kinase activity	0.048
GO:0043902~positive regulation of multi-organism process	0.048
GO:0010243~response to organonitrogen compound	0.048
GO:0050679~positive regulation of epithelial cell proliferation	0.049
GO:0070509~calcium ion import	0.049
GO:0043620~regulation of DNA-templated transcription in response to stress	0.050
GO:0097201~negative regulation of transcription from RNA polymerase II promoter in response to stress	0.050
GO:0002923~regulation of humoral immune response mediated by circulating immunoglobulin	0.050

Appendix 19. Canonical pathways enriched in 132 up-regulated by BILF1 in DG-75 cells

Pathway	*p*-value
EPO pathway	9.00E-08
ARENRF2 pathway	3.00E-06
IL3 pathway	5.00E-06
CDMAC pathway	6.00E-06
PPARA pathway	9.00E-06
IL2 pathway	2.00E-05
GLEEVEC pathway	2.00E-05
TPO pathway	2.00E-05
EGF pathway	5.00E-05
PDGF pathway	5.00E-05
IL2RB pathway	8.00E-05
Keratinocyte pathway	2.00E-04
DREAM pathway	4.00E-04
IL10 pathway	6.00E-04
Cardiac EGF pathway	7.00E-04
CCR5 pathway	7.00E-04
ETS pathway	7.00E-04
NGF pathway	7.00E-04
MAPK pathway	1.00E-03
IGF1 pathway	1.00E-03
IL6 pathway	1.00E-03
Insulin pathway	1.00E-03
NTHI pathway	1.00E-03
Stress pathway	1.00E-03
GH pathway	2.00E-03
TNFR1 pathway	2.00E-03
IL1R pathway	2.00E-03
BCR pathway	3.00E-03
GPCR pathway	3.00E-03
MET pathway	3.00E-03
TOLL pathway	3.00E-03
FCER1 pathway	3.00E-03
Biopeptides pathway	4.00E-03
TCR pathway	5.00E-03
Neurotransmitters pathway	1.00E-02
FREE pathway	2.00E-02
SODD pathway	2.00E-02
CDK5 pathway	2.00E-02
Fibrinolysis pathway	3.00E-02

D4GDI pathway	3.00E-02
Extrinsic pathway	3.00E-02
Granulocytes pathway	3.00E-02
Platelet app pathway	3.00E-02
RANKL pathway	3.00E-02
CD40 pathway	3.00E-02
HIF pathway	3.00E-02
HSP27 pathway	3.00E-02
B cell survival pathway	3.00E-02
IL22BP pathway	3.00E-02
LAIR pathway	3.00E-02
RELA pathway	3.00E-02
41BB pathway	4.00E-02
IL7 pathway	4.00E-02
PML pathway	4.00E-02
TNFR2 pathway	4.00E-02
TID pathway	4.00E-02
AMI pathway	4.00E-02
ATM pathway	4.00E-02
Cytokine pathway	4.00E-02
NK cells pathway	4.00E-02
IL12 pathway	5.00E-02
SPPA pathway	5.00E-02
Intrinsic pathway	5.00E-02
NFκB pathway	5.00E-02

Appendix 20: 463 genes up-regulated by BILF1 in GC B cells

Gene	Fold change	Gene	Fold change	Gene	Fold change
PNMA2	6.76	**LOC101927780**	1.86	**ZNF444**	1.63
LINC01269	6.61	**ZNF576**	1.86	**PDE1B**	1.63
CNIH3	6.33	**MYOF**	1.86	**NDUFA7**	1.63
TCONS 00029157	5.58	**IZUMO4**	1.86	**JAGN1**	1.62
DOK2	4.77	**NDUFB11**	1.85	**TSPO**	1.62
FMR1-AS1	4.01	**CCDC167**	1.85	**CKLF**	1.62
CAMK2A	3.97	**NDUFS8**	1.85	**PHLDA3**	1.62
TIAF1	3.74	**ZNF787**	1.85	**PHPT1**	1.62
CST7	3.49	**MAP2K6**	1.84	**ITPA**	1.62
LOC102724919	3.40	**ELOF1**	1.84	**RPS21**	1.62
SLC37A2	3.40	**BIN2**	1.84	**TIMM8B**	1.62
GNA14	3.40	**RPL23P8**	1.84	**REXO4**	1.62
SPTLC3	3.39	**C19orf24**	1.84	**HSF5**	1.61
SLC19A1	3.32	**PYM1**	1.83	**CORO1A**	1.61
CACNG7	3.32	**C1orf145**	1.82	**TOMM7**	1.61
WNT1	3.30	**ZNF593**	1.82	**NEDD8**	1.61
NPR1	3.18	**SSSCA1**	1.82	**NR2C2AP**	1.61
NACAD	3.10	**VWCE**	1.82	**RNF26**	1.61
MTVR2	3.06	**EXOSC4**	1.82	**LIPE**	1.61
TRAPPC2L	2.91	**ATP6V1F**	1.82	**DYNLRB1**	1.61
RCAN2	2.90	**RAB13**	1.81	**PNOC**	1.61
GREB1	2.89	**LOC100507346**	1.81	**GCHFR**	1.61
IL24	2.86	**FAM96B**	1.81	**KIAA1024**	1.61
CDC14B	2.84	**LOC100506136**	1.81	**FIBP**	1.61
OSTCP1	2.82	**NDUFB7**	1.81	**RBX1**	1.61
RAB33A	2.79	**CENPW**	1.81	**TWF2**	1.60
PSENEN	2.78	**HIST1H1E**	1.81	**ZNF485**	1.60
C9orf142	2.78	**COX7B**	1.81	**CDIPT**	1.60
GABRR1	2.77	**LMO2**	1.80	**CARD8-AS1**	1.60
ARMCX5-GPRASP2	2.75	**ALDOC**	1.80	**LSM7**	1.60
MIR650	2.71	**ATP13A2**	1.80	**TRMT61A**	1.60
NKD1	2.70	**RPS19**	1.80	**PEX26**	1.59
PGPEP1L	2.69	**CCR6**	1.79	**EDF1**	1.59
SLC24A3	2.68	**FLJ44635**	1.79	**CFAP43**	1.59
KDM4D	2.67	**CPXM1**	1.79	**SSR4**	1.59
PLAC8L1	2.61	**APOL3**	1.78	**TRIAP1**	1.59
SPAG8	2.60	**ABRACL**	1.78	**ATP5E**	1.59
KLK4	2.59	**SIT1**	1.78	**TXN**	1.59

ATP8B4	2.59	**SNRPD2**	1.78	**MPEG1**	1.59
LRRTM2	2.58	**PFDN2**	1.78	**PHB**	1.59
PDCD1	2.58	**GEMIN6**	1.78	**FEZ1**	1.59
MROH6	2.57	**ZNF582**	1.78	**MRPL22**	1.58
TNFRSF11A	2.54	**CD52**	1.77	**RPL18**	1.58
PXT1	2.51	**CCDC51**	1.77	**ZNF513**	1.58
MAMLD1	2.51	**SLC26A11**	1.77	**GALNT6**	1.58
PLG	2.50	**POLR2I**	1.77	**CXorf65**	1.58
CACNA1E	2.48	**TMEM258**	1.77	**MRPS11**	1.58
PDE4A	2.46	**POLR2J**	1.77	**MRPL17**	1.58
EPS8L2	2.45	**TMEM25**	1.77	**ALS2CR12**	1.58
C1orf101	2.45	**WDR18**	1.76	**YDJC**	1.58
NTN4	2.42	**DHCR24**	1.76	**RPL30**	1.58
TMEM160	2.39	**CDK5R1**	1.76	**TPRG1**	1.58
ABHD1	2.39	**ANXA2R**	1.76	**RNF181**	1.57
FAM195B	2.36	**PSMB9**	1.76	**LAPTM4B**	1.57
VWA7	2.35	**PSMB3**	1.76	**ZNF229**	1.57
ADGRA2	2.35	**PRDX5**	1.75	**RRP9**	1.57
THAP7	2.33	**RPL35**	1.75	**SH3GL1**	1.57
DOK6	2.32	**FAM43A**	1.75	**MIR142**	1.57
LAMTOR2	2.31	**XKR8**	1.75	**PQLC1**	1.57
ARHGEF37	2.31	**HSD17B10**	1.75	**TP53I11**	1.57
MT2A	2.31	**VAMP8**	1.74	**FOSL2**	1.57
C22orf23	2.30	**MOB2**	1.74	**NOP10**	1.57
PRKX-AS1	2.28	**RPS14P3**	1.74	**NLGN3**	1.57
TBXAS1	2.28	**MRPS23**	1.74	**ATP5D**	1.56
TFR2	2.26	**FAM98C**	1.74	**PDCL3P4**	1.56
PARVB	2.25	**ANAPC11**	1.74	**PTRHD1**	1.56
MDP1	2.24	**MRPL14**	1.74	**TNFRSF13B**	1.56
PDE6C	2.24	**MCAM**	1.73	**AGAP5**	1.56
ITGAX	2.23	**FAM162A**	1.73	**NDUFAF4**	1.56
UQCR11	2.22	**EMP3**	1.73	**PRKCH**	1.56
TMEM74B	2.21	**CACNA2D3**	1.73	**CD320**	1.56
LINC01279	2.20	**RPL36**	1.73	**IFITM2**	1.56
FAM86FP	2.20	**RPLP2**	1.73	**NDUFA1**	1.55
MRPL12	2.20	**NDUFA11**	1.72	**SLC17A9**	1.55
C11orf98	2.19	**LINC01560**	1.72	**RBKS**	1.55
RNF144A	2.19	**RPL37**	1.72	**PTGS1**	1.55
PLEKHA6	2.19	**PRDX1**	1.72	**MYL12A**	1.55
TBC1D30	2.16	**USMG5**	1.72	**CLIC1**	1.55
SDC4	2.16	**APRT**	1.72	**SQRDL**	1.55
ALKBH7	2.15	**STK11**	1.72	**SBNO2**	1.55
DGAT2	2.14	**C1orf21**	1.72	**MED27**	1.55
CTNNA1	2.13	**CD1C**	1.72	**PRKD2**	1.55

HTR3B	2.13	**OST4**	1.72	**TMEM261**	1.55
PTPN6	2.12	**ITGB3**	1.72	**GAS7**	1.55
MGMT	2.11	**SSNA1**	1.71	**UQCRHL**	1.55
SPATA21	2.10	**NDUFB2**	1.71	**FAU**	1.55
FARSA	2.09	**HIST2H2AB**	1.71	**RPS14**	1.55
CD63	2.08	**SUFU**	1.71	**CLIC5**	1.54
CAMKK1	2.08	**IL4I1**	1.71	**COMMD1**	1.54
NRROS	2.08	**MRPL51**	1.71	**MAD2L2**	1.54
LINC00987	2.08	**PFDN6**	1.71	**HIRIP3**	1.54
CHMP6	2.07	**SLC22A13**	1.70	**NDUFB10**	1.54
C6orf99	2.07	**UBA52**	1.70	**LTB**	1.54
MPST	2.06	**RPL38**	1.70	**RHOBTB3**	1.54
ZNF774	2.06	**SYNGR2**	1.70	**LIX1L**	1.54
FGR	2.05	**FIS1**	1.70	**NDUFB9**	1.54
CATSPER3	2.05	**LINC00589**	1.70	**RPS18**	1.54
SLC24A2	2.05	**YIF1A**	1.69	**SHFM1**	1.54
FRMD3	2.04	**DDRGK1**	1.69	**ARFGAP1**	1.54
HSD17B1	2.04	**ORAOV1**	1.69	**MAGEF1**	1.53
CRIP1	2.02	**LOC101243545**	1.69	**RAB1B**	1.53
AMZ1	2.01	**RGS9**	1.69	**MRPL27**	1.53
LOC101927472	2.01	**FRAT1**	1.69	**HSD17B7**	1.53
TMEM53	2.01	**ITGB2**	1.69	**SPRYD7**	1.53
ZNF491	2.00	**TRMT112**	1.69	**ATP5I**	1.53
ATP10A	2.00	**CHID1**	1.69	**ATP5G1**	1.53
CYP1A1	2.00	**RGS19**	1.68	**CCDC12**	1.53
NDUFA8	2.00	**CMTM3**	1.68	**ASNA1**	1.53
TBC1D9	2.00	**SLC25A1**	1.68	**LMNA**	1.53
SLC16A13	1.99	**TMSB10**	1.68	**CDKN1A**	1.53
PPP1R14B	1.99	**PBX4**	1.68	**PLOD3**	1.53
TMSB4X	1.99	**COMMD5**	1.68	**UBL5**	1.53
CRYM	1.99	**HTR3A**	1.68	**PSMG3**	1.53
TSTD1	1.98	**SH3BGRL3**	1.68	**UMPS**	1.53
ISG15	1.98	**IQCD**	1.68	**COX6B1**	1.53
MIR663AHG	1.97	**LHFPL3-AS2**	1.67	**FDXR**	1.53
PET117	1.97	**PTPRCAP**	1.67	**OR5B21**	1.53
ATAD3A	1.96	**MRPL24**	1.67	**MRPS21**	1.53
ZBED2	1.96	**RPL27**	1.67	**LOC101929450**	1.52
BRMS1	1.96	**LIN37**	1.67	**PTMA**	1.52
NFKBIL1	1.95	**NOL7**	1.67	**PPP1R35**	1.52
ECEL1	1.95	**G6PC3**	1.67	**USE1**	1.52
HIST1H2AI	1.95	**PEPD**	1.67	**MYEOV2**	1.52
LTA	1.94	**SNRNP25**	1.66	**GSTP1**	1.52
ARL14EPL	1.94	**POU6F1**	1.66	**COX7C**	1.52
RABAC1	1.94	**CD70**	1.66	**MRPL38**	1.52

HIST1H3I	1.93	**C7orf50**	1.66	**GPR132**	1.52
HIST1H2AL	1.93	**RFXANK**	1.66	**TALDO1**	1.52
NDUFA13	1.93	**RPLP1**	1.66	**UQCC2**	1.52
IL27RA	1.93	**DTX2**	1.65	**POMP**	1.52
KRTCAP2	1.93	**ST8SIA5**	1.65	**C19orf53**	1.51
CHCHD1	1.92	**UQCR10**	1.65	**N4BP3**	1.51
NME1	1.92	**MYDGF**	1.65	**SLCO4A1**	1.51
TNFRSF9	1.91	**RPS27L**	1.65	**ADPRHL2**	1.51
CD96	1.91	**RPS19BP1**	1.65	**CSTB**	1.51
IFI30	1.90	**ZBTB32**	1.65	**TBC1D10C**	1.51
COX5B	1.90	**MYL6**	1.65	**PSKH1**	1.51
ANKRD30BL	1.90	**ZFPL1**	1.65	**HIST2H2AC**	1.51
CRYM-AS1	1.90	**LOC101929470**	1.65	**COX8A**	1.51
MRPL52	1.90	**GNB2**	1.65	**SF3B5**	1.51
APOA1-AS	1.90	**KLK1**	1.65	**PPP1CA**	1.51
RIN3	1.90	**MIR6087**	1.65	**PFDN5**	1.50
TRAPPC5	1.90	**RPL37A**	1.65	**ALOX5AP**	1.50
SH2D2A	1.90	**RINL**	1.64	**MRPS18B**	1.50
DLG5	1.89	**ACTA1**	1.64	**RNH1**	1.50
MRPL36	1.88	**SMAGP**	1.64	**TNFRSF18**	1.50
CCT6B	1.88	**CHRNA1**	1.64	**SPCS1**	1.50
MTFP1	1.88	**ZNF213-AS1**	1.64	**PADI2**	1.50
AHRR	1.88	**SLC25A11**	1.64	**RNF187**	1.50
ARL4C	1.88	**TMA7**	1.64	**ICT1**	1.50
ZNF71	1.88	**MRPS12**	1.63	**COA6**	1.50
AP2S1	1.88	**DUSP4**	1.63	**COX6C**	1.50
STK31	1.87	**RPS16**	1.63	**ERP27**	1.50
TRABD2A	1.87	**COX6A1**	1.63		
ATP5EP2	1.86	**APP**	1.63		

Appendix 21: 356 genes down-regulated by BILF1 in GC B cells

Gene	Fold change	Gene	Fold change	Gene	Fold change
IL26	-7.41	**ERICH6**	-2.09	**LOC285074**	-1.71
LOC100129924	-6.96	**OLMALINC**	-2.09	**TENM1**	-1.71
WDR38	-5.56	**LINC01140**	-2.09	**ARC**	-1.70
KSR2	-5.12	**PHOSPHO2**	-2.08	**STAG3L5P**	-1.69
TRHR	-4.15	**NFIX**	-2.06	**FAM221A**	-1.69
NAALADL2	-4.06	**RDH10-AS1**	-2.06	**KCTD6**	-1.69
GRPR	-4.06	**MOB4**	-2.05	**DEPDC1**	-1.69
LOC339807	-3.77	**CTSV**	-2.05	**NKX3-1**	-1.69
LAG3	-3.71	**SNORD144**	-2.05	**STARD4-AS1**	-1.69
CD28	-3.54	**ARHGAP33**	-2.04	**KLF5**	-1.69
HIST1H3H	-3.45	**SUGCT**	-2.03	**INCA1**	-1.69
ELOVL3	-3.42	**ADRA2B**	-2.03	**TXNDC16**	-1.68
TMEM59L	-3.34	**MYOZ3**	-2.03	**C6orf163**	-1.68
LOC442497	-3.32	**ABCG1**	-2.02	**OSBPL10-AS1**	-1.67
TRPV4	-3.28	**KCNH1**	-2.02	**MIR133A1HG**	-1.67
LOC102724484	-3.27	**LRRC45**	-2.02	**GPR22**	-1.67
ULBP2	-3.25	**SNORA29**	-2.02	**PCDH9**	-1.66
LOC101060553	-3.23	**FZD7**	-2.01	**ACACB**	-1.66
JHDM1D-AS1	-3.21	**P2RY13**	-2.01	**BRE-AS1**	-1.66
NOD2	-3.17	**MLLT10P1**	-2.01	**TMC5**	-1.66
MIR4324	-3.14	**KCCAT333**	-2.00	**KIF20A**	-1.65
SMCO3	-3.12	**RYR1**	-2.00	**CASD1**	-1.65
CCNI2	-3.12	**HAVCR1P1**	-1.99	**GSTM3**	-1.65
MIR554	-3.05	**EPB41L4B**	-1.99	**TC2N**	-1.65
WNK2	-3.04	**LINC01504**	-1.98	**VASN**	-1.65
DNMBP-AS1	-3.03	**TMEM120A**	-1.98	**CETN3**	-1.64
JPH1	-3.03	**CPED1**	-1.97	**HOXA6**	-1.64
GPX8	-3.02	**CHPT1**	-1.97	**PLCE1**	-1.64
CCL27	-2.96	**SNORD29**	-1.97	**ASTL**	-1.64
OSBPL6	-2.91	**MAF**	-1.97	**PRKCA-AS1**	-1.64
TRNP1	-2.91	**ZNF704**	-1.97	**KBTBD11**	-1.63
ALDH1A3	-2.83	**LOC100289495**	-1.96	**ASAP1-IT2**	-1.63
PYCARD-AS1	-2.82	**PVRL1**	-1.96	**WNT16**	-1.63
NCKAP5	-2.80	**GAPDHS**	-1.96	**CFL2**	-1.62
C1QTNF6	-2.75	**PRKAR2B**	-1.96	**ASPN**	-1.62
ST7-OT3	-2.73	**LARGE-AS1**	-1.95	**SNORD28**	-1.62
CD14	-2.72	**LOC730101**	-1.95	**SMIM4**	-1.62
ZC2HC1B	-2.71	**RASAL2**	-1.94	**KCNQ5-AS1**	-1.62
LOC100132078	-2.71	**MIR4639**	-1.93	**CCDC171**	-1.62

FMN2	-2.70	**ZNF385C**	-1.93	**HYMAI**	-1.62
IPCEF1	-2.70	**TAS2R19**	-1.93	**AHNAK**	-1.62
ALPL	-2.69	**LECT1**	-1.92	**STX1A**	-1.61
NPHP1	-2.69	**COL2A1**	-1.92	**NUP50-AS1**	-1.61
SNORD50A	-2.67	**TAS2R46**	-1.92	**ZPBP2**	-1.61
EIF1AX-AS1	-2.67	**SNORA48**	-1.92	**SGCE**	-1.61
AZIN2	-2.63	**TTLL13P**	-1.92	**MAFB**	-1.61
GAREM	-2.62	**LOC645434**	-1.92	**MORC4**	-1.61
STARD13	-2.62	**SYCE1L**	-1.91	**FOS**	-1.60
PTGER3	-2.61	**KIF9-AS1**	-1.91	**ALG6**	-1.60
DSEL	-2.60	**PHF24**	-1.91	**CCNE2**	-1.60
FNIP2	-2.59	**PIWIL4**	-1.91	**LOC101927543**	-1.60
PTPN13	-2.57	**PYGO1**	-1.90	**LOC400958**	-1.60
TP63	-2.57	**UBAC2-AS1**	-1.90	**NR1D2**	-1.60
SPRY1	-2.56	**GRHL1**	-1.90	**DEPDC1B**	-1.59
LOC400997	-2.55	**PDE3B**	-1.90	**KCNQ1OT1**	-1.59
MYO7B	-2.55	**SNORA31**	-1.89	**SNAI1**	-1.59
ASMTL-AS1	-2.54	**CFAP46**	-1.88	**FLT1**	-1.59
TRPC5	-2.51	**PDCD4-AS1**	-1.88	**TBC1D2B**	-1.59
LYPLAL1-AS1	-2.50	**FAM227A**	-1.88	**FOSB**	-1.59
ZACN	-2.49	**FYB**	-1.87	**LMLN**	-1.59
SEPT3	-2.49	**SNORD97**	-1.86	**KLB**	-1.59
SNX24	-2.48	**TMEM9B-AS1**	-1.86	**PAXBP1-AS1**	-1.58
FRK	-2.46	**PYHIN1**	-1.86	**SNORA4**	-1.58
FNDC1	-2.46	**TJP2**	-1.86	**DPY19L2**	-1.58
CAMK2N1	-2.45	**CTAGE7P**	-1.86	**YPEL2**	-1.58
FBXO43	-2.45	**SNORA33**	-1.85	**NPTX2**	-1.58
CLIC6	-2.43	**PCDHGC4**	-1.85	**E2F8**	-1.57
FER1L4	-2.41	**MCTP1**	-1.85	**HSPA1L**	-1.57
CACNA1D	-2.41	**BTLA**	-1.84	**SLC9A9**	-1.57
IFNG	-2.39	**ARHGAP32**	-1.84	**EML6**	-1.57
SEMA3B	-2.39	**SEPN1**	-1.83	**WNK3**	-1.57
ZNF404	-2.39	**LOC100131635**	-1.82	**FZD5**	-1.57
ZNF365	-2.38	**HCK**	-1.82	**SNORD71**	-1.57
BAMBI	-2.36	**ACVR1**	-1.82	**LOC100507156**	-1.57
MFI2-AS1	-2.35	**BPGM**	-1.82	**CFAP54**	-1.57
ICOS	-2.33	**SLC16A10**	-1.81	**CLEC2B**	-1.57
GPR137C	-2.33	**F5**	-1.81	**RORA-AS1**	-1.56
RASGRP4	-2.31	**HBEGF**	-1.81	**SNORD99**	-1.56
SNORD50B	-2.31	**ZNF80**	-1.80	**MTRNR2L5**	-1.56
SGCD	-2.30	**LOC101927190**	-1.80	**ZBTB20-AS4**	-1.56
KIAA1656	-2.30	**LOC145474**	-1.80	**SH3D19**	-1.56
RHBDF1	-2.30	**DDR2**	-1.80	**KIF13A**	-1.56
SNORD101	-2.29	**SMIM18**	-1.80	**CCDC150**	-1.56

LINC00861	-2.29	**IRS2**	-1.79	**ADRB2**	-1.55
WBP1	-2.28	**LOC102723704**	-1.79	**CLSPN**	-1.55
IL13RA1	-2.27	**KDELC2**	-1.79	**DNAJC3-AS1**	-1.55
BAIAP3	-2.26	**TJP3**	-1.79	**TIMM8A**	-1.55
MIR186	-2.26	**CDC20B**	-1.79	**KIAA1211**	-1.55
INPP4B	-2.25	**RIMS2**	-1.78	**GTSE1**	-1.55
SYNM	-2.24	**LRRC70**	-1.78	**SLC38A6**	-1.55
SBSPON	-2.22	**NUDT9P1**	-1.78	**SLC30A1**	-1.55
OR2B2	-2.22	**DLGAP5**	-1.78	**NANOS1**	-1.55
RHEBL1	-2.21	**PLN**	-1.78	**SENCR**	-1.55
AVPI1	-2.20	**GAS2L3**	-1.77	**LINC01184**	-1.55
CEP70	-2.20	**SLC2A12**	-1.77	**ANGPTL1**	-1.54
DIP2C	-2.20	**SHANK2**	-1.77	**TMEM81**	-1.54
MTCL1	-2.19	**NAALAD2**	-1.77	**C1GALT1C1**	-1.54
TIPARP-AS1	-2.18	**TSIX**	-1.77	**KCNA3**	-1.54
NUP62CL	-2.18	**C1QL3**	-1.76	**DLEU1**	-1.54
ATP1B2	-2.18	**TIGIT**	-1.76	**C15orf41**	-1.53
FERMT2	-2.17	**NPY6R**	-1.76	**CACYBP**	-1.53
SNORD12	-2.17	**IFI44L**	-1.76	**CDC25A**	-1.53
SLC4A9	-2.17	**DTHD1**	-1.75	**EYS**	-1.53
AKAP6	-2.16	**SCARNA9L**	-1.75	**FAM98B**	-1.53
ANKRD36BP2	-2.15	**SNORA76A**	-1.75	**STRIP2**	-1.53
MIR4742	-2.15	**LOC100507557**	-1.75	**MBNL2**	-1.52
TBX21	-2.15	**PELI2**	-1.75	**IFNG-AS1**	-1.52
ANKDD1B	-2.14	**CCDC149**	-1.74	**RAB23**	-1.52
IL18	-2.14	**P2RY14**	-1.74	**ID2**	-1.51
BMP7	-2.13	**RPS16P5**	-1.74	**TUFT1**	-1.51
LPL	-2.11	**POU3F1**	-1.73	**NUDT4**	-1.51
MIR597	-2.11	**FAM83D**	-1.73	**SPRY4**	-1.50
CD200	-2.10	**SLC16A2**	-1.73	**DYNC2H1**	-1.50
PTGS2	-2.10	**MLLT3**	-1.73	**BACE2**	-1.50
CA13	-2.10	**KLF4**	-1.73	**FRY**	-1.50
THBS1	-2.10	**MLF1**	-1.73	**RNF103**	-1.50
TOB2P1	-2.10	**IFI44**	-1.73	**TRIM2**	-1.50
PFN2	-2.10	**CAV1**	-1.71	**CREG1**	-1.50
CXCL10	-2.09	**CSRP2**	-1.71		

Appendix 22: Biological processes enriched in 463 genes up-regulated by BILF1in GC B cells

Gene ontology term	*p*-value
GO:0042773~ATP synthesis coupled electron transport	6.53E-16
GO:0006119~oxidative phosphorylation	1.81E-15
GO:0042775~mitochondrial ATP synthesis coupled electron transport	7.62E-15
GO:0022904~respiratory electron transport chain	2.64E-14
GO:0022900~electron transport chain	3.88E-14
GO:0009205~purine ribonucleoside triphosphate metabolic process	4.31E-14
GO:0009199~ribonucleoside triphosphate metabolic process	7.95E-14
GO:0009144~purine nucleoside triphosphate metabolic process	8.75E-14
GO:0046034~ATP metabolic process	1.45E-13
GO:0009161~ribonucleoside monophosphate metabolic process	7.74E-13
GO:0009141~nucleoside triphosphate metabolic process	7.74E-13
GO:0009167~purine ribonucleoside monophosphate metabolic process	1.48E-12
GO:0009126~purine nucleoside monophosphate metabolic process	1.62E-12
GO:0009123~nucleoside monophosphate metabolic process	2.63E-12
GO:0007005~mitochondrion organization	3.55E-12
GO:0046128~purine ribonucleoside metabolic process	8.58E-11
GO:0042278~purine nucleoside metabolic process	1.06E-10
GO:0045047~protein targeting to ER	1.24E-10
GO:0009119~ribonucleoside metabolic process	1.40E-10
GO:0072599~establishment of protein localization to endoplasmic reticulum	2.29E-10
GO:0045333~cellular respiration	2.93E-10
GO:0006415~translational termination	3.24E-10
GO:0006614~SRP-dependent cotranslational protein targeting to membrane	3.24E-10
GO:0070125~mitochondrial translational elongation	5.06E-10
GO:0070126~mitochondrial translational termination	6.01E-10
GO:0009116~nucleoside metabolic process	8.25E-10
GO:0006613~cotranslational protein targeting to membrane	9.53E-10
GO:0032543~mitochondrial translation	1.08E-09
GO:0006412~translation	1.09E-09
GO:0000184~nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	1.83E-09
GO:0019693~ribose phosphate metabolic process	1.83E-09
GO:1901657~glycosyl compound metabolic process	2.98E-09
GO:0009259~ribonucleotide metabolic process	3.22E-09
GO:0070972~protein localization to endoplasmic reticulum	3.41E-09
GO:0043043~peptide biosynthetic process	3.48E-09
GO:1902600~hydrogen ion transmembrane transport	3.85E-09
GO:0009150~purine ribonucleotide metabolic process	5.36E-09
GO:0006518~peptide metabolic process	5.94E-09
GO:0006612~protein targeting to membrane	8.65E-09

GO:0006414~translational elongation	8.71E-09
GO:0033108~mitochondrial respiratory chain complex assembly	1.06E-08
GO:0006163~purine nucleotide metabolic process	1.92E-08
GO:0009117~nucleotide metabolic process	3.36E-08
GO:0006364~rRNA processing	3.52E-08
GO:0015980~energy derivation by oxidation of organic compounds	4.04E-08
GO:0006753~nucleoside phosphate metabolic process	4.61E-08
GO:0043241~protein complex disassembly	4.84E-08
GO:0043604~amide biosynthetic process	5.64E-08
GO:0016072~rRNA metabolic process	5.66E-08
GO:0072521~purine-containing compound metabolic process	6.64E-08
GO:0032984~macromolecular complex disassembly	9.67E-08
GO:0006091~generation of precursor metabolites and energy	1.03E-07
GO:0006402~mRNA catabolic process	1.10E-07
GO:0019083~viral transcription	1.16E-07
GO:0055086~nucleobase-containing small molecule metabolic process	1.18E-07
GO:0010257~NADH dehydrogenase complex assembly	1.20E-07
GO:0032981~mitochondrial respiratory chain complex I assembly	1.20E-07
GO:0097031~mitochondrial respiratory chain complex I biogenesis	1.20E-07
GO:0042254~ribosome biogenesis	1.28E-07
GO:0000956~nuclear-transcribed mRNA catabolic process	1.71E-07
GO:0006401~RNA catabolic process	1.76E-07
GO:0006120~mitochondrial electron transport, NADH to ubiquinone	2.11E-07
GO:1902582~single-organism intracellular transport	2.82E-07
GO:0019080~viral gene expression	2.84E-07
GO:0043603~cellular amide metabolic process	2.95E-07
GO:0015992~proton transport	2.97E-07
GO:0044033~multi-organism metabolic process	3.10E-07
GO:0006818~hydrogen transport	3.55E-07
GO:0043624~cellular protein complex disassembly	3.58E-07
GO:0071822~protein complex subunit organization	5.65E-07
GO:0006413~translational initiation	1.09E-06
GO:0006605~protein targeting	2.55E-06
GO:0090150~establishment of protein localization to membrane	2.56E-06
GO:1901566~organonitrogen compound biosynthetic process	2.96E-06
GO:0055114~oxidation-reduction process	3.03E-06
GO:0006123~mitochondrial electron transport, cytochrome c to oxygen	3.39E-06
GO:0034655~nucleobase-containing compound catabolic process	4.46E-06
GO:0046700~heterocycle catabolic process	5.41E-06
GO:0034470~ncRNA processing	5.41E-06
GO:0072594~establishment of protein localization to organelle	5.62E-06
GO:0044270~cellular nitrogen compound catabolic process	7.31E-06
GO:0022411~cellular component disassembly	7.73E-06
GO:0019439~aromatic compound catabolic process	9.05E-06

GO:0022613~ribonucleoprotein complex biogenesis	9.66E-06
GO:0043933~macromolecular complex subunit organization	1.38E-05
GO:0061024~membrane organization	2.06E-05
GO:1901361~organic cyclic compound catabolic process	2.12E-05
GO:0046907~intracellular transport	3.05E-05
GO:0072657~protein localization to membrane	3.72E-05
GO:0044802~single-organism membrane organization	4.74E-05
GO:0034660~ncRNA metabolic process	7.51E-05
GO:0019058~viral life cycle	8.27E-05
GO:0034220~ion transmembrane transport	2.01E-04
GO:0015985~energy coupled proton transport, down electrochemical gradient	3.99E-04
GO:0015986~ATP synthesis coupled proton transport	3.99E-04
GO:0006811~ion transport	4.11E-04
GO:1902580~single-organism cellular localization	4.14E-04
GO:0019637~organophosphate metabolic process	4.60E-04
GO:0033365~protein localization to organelle	4.77E-04
GO:0070887~cellular response to chemical stimulus	5.07E-04
GO:0051649~establishment of localization in cell	6.05E-04
GO:0006886~intracellular protein transport	6.31E-04
GO:0009057~macromolecule catabolic process	6.79E-04
GO:0044085~cellular component biogenesis	7.06E-04
GO:0098660~inorganic ion transmembrane transport	7.92E-04
GO:0098655~cation transmembrane transport	0.0010
GO:0000028~ribosomal small subunit assembly	0.0011
GO:0042776~mitochondrial ATP synthesis coupled proton transport	0.0011
GO:0016071~mRNA metabolic process	0.0011
GO:0033209~tumour necrosis factor-mediated signalling pathway	0.0011
GO:0043623~cellular protein complex assembly	0.0011
GO:0015031~protein transport	0.0011
GO:0051641~cellular localization	0.0014
GO:0055085~transmembrane transport	0.0015
GO:0044265~cellular macromolecule catabolic process	0.0019
GO:0034613~cellular protein localization	0.0019
GO:0006396~RNA processing	0.0021
GO:0033036~macromolecule localization	0.0021
GO:0070727~cellular macromolecule localization	0.0023
GO:0009206~purine ribonucleoside triphosphate biosynthetic process	0.0025
GO:1901135~carbohydrate derivative metabolic process	0.0026
GO:0009145~purine nucleoside triphosphate biosynthetic process	0.0027
GO:0008104~protein localization	0.0029
GO:0034622~cellular macromolecular complex assembly	0.0030
GO:0009156~ribonucleoside monophosphate biosynthetic process	0.0034
GO:0006839~mitochondrial transport	0.0037
GO:0042274~ribosomal small subunit biogenesis	0.0039

GO:0009201~ribonucleoside triphosphate biosynthetic process	0.0039
GO:0098662~inorganic cation transmembrane transport	0.0040
GO:0006754~ATP biosynthetic process	0.0045
GO:0045184~establishment of protein localization	0.0048
GO:1990542~mitochondrial transmembrane transport	0.0049
GO:0042455~ribonucleoside biosynthetic process	0.0050
GO:0009124~nucleoside monophosphate biosynthetic process	0.0058
GO:0006812~cation transport	0.0060
GO:0015672~monovalent inorganic cation transport	0.0061
GO:0009168~purine ribonucleoside monophosphate biosynthetic process	0.0068
GO:0009127~purine nucleoside monophosphate biosynthetic process	0.0068
GO:0042451~purine nucleoside biosynthetic process	0.0068
GO:0046129~purine ribonucleoside biosynthetic process	0.0068
GO:0071310~cellular response to organic substance	0.0078
GO:0016310~phosphorylation	0.0081
GO:0071345~cellular response to cytokine stimulus	0.0081
GO:0031397~negative regulation of protein ubiquitination	0.0081
GO:0009163~nucleoside biosynthetic process	0.0081
GO:0006796~phosphate-containing compound metabolic process	0.0081
GO:0006793~phosphorus metabolic process	0.0086
GO:0009142~nucleoside triphosphate biosynthetic process	0.0086
GO:0000462~maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	0.0088
GO:1901659~glycosyl compound biosynthetic process	0.0088
GO:0002532~production of molecular mediator involved in inflammatory response	0.0106
GO:0071493~cellular response to UV-B	0.0126
GO:1903321~negative regulation of protein modification by small protein conjugation or removal	0.0128
GO:0050854~regulation of antigen receptor-mediated signalling pathway	0.0137
GO:0044419~interspecies interaction between organisms	0.0144
GO:0044403~symbiosis, encompassing mutualism through parasitism	0.0144
GO:0016032~viral process	0.0152
GO:0044764~multi-organism cellular process	0.0167
GO:0065003~macromolecular complex assembly	0.0172
GO:0006461~protein complex assembly	0.0192
GO:0070271~protein complex biogenesis	0.0195
GO:0042989~sequestering of actin monomers	0.0196
GO:0030490~maturation of SSU-rRNA	0.0215
GO:0044248~cellular catabolic process	0.0218
GO:0019221~cytokine-mediated signalling pathway	0.0227
GO:0060252~positive regulation of glial cell proliferation	0.0236
GO:0070527~platelet aggregation	0.0262
GO:0010467~gene expression	0.0291
GO:0010821~regulation of mitochondrion organization	0.0319
GO:0046364~monosaccharide biosynthetic process	0.0324

Appendix 23: 125 biological processes most significantly enriched in 356 genes down-regulated by BILF1in GC B cells

Gene ontology term	*p*-value
GO:0051049~regulation of transport	2.71E-06
GO:0042325~regulation of phosphorylation	4.01E-06
GO:0001932~regulation of protein phosphorylation	4.83E-06
GO:0019220~regulation of phosphate metabolic process	1.03E-05
GO:0051174~regulation of phosphorus metabolic process	1.04E-05
GO:0035556~intracellular signal transduction	1.28E-05
GO:0006468~protein phosphorylation	1.41E-05
GO:0023014~signal transduction by protein phosphorylation	2.11E-05
GO:0031399~regulation of protein modification process	2.30E-05
GO:0034762~regulation of transmembrane transport	3.90E-05
GO:1902531~regulation of intracellular signal transduction	4.65E-05
GO:0042327~positive regulation of phosphorylation	6.17E-05
GO:0043549~regulation of kinase activity	6.65E-05
GO:0001934~positive regulation of protein phosphorylation	7.16E-05
GO:0000165~MAPK cascade	7.23E-05
GO:0072358~cardiovascular system development	7.83E-05
GO:0072359~circulatory system development	7.83E-05
GO:0045937~positive regulation of phosphate metabolic process	8.10E-05
GO:0010562~positive regulation of phosphorus metabolic process	8.10E-05
GO:0034765~regulation of ion transmembrane transport	8.99E-05
GO:0065009~regulation of molecular function	1.22E-04
GO:0051050~positive regulation of transport	1.24E-04
GO:1901343~negative regulation of vasculature development	1.30E-04
GO:0014066~regulation of phosphatidylinositol 3-kinase signalling	1.35E-04
GO:0045859~regulation of protein kinase activity	1.47E-04
GO:0019216~regulation of lipid metabolic process	1.49E-04
GO:0016310~phosphorylation	1.77E-04
GO:0051338~regulation of transferase activity	2.27E-04
GO:0051246~regulation of protein metabolic process	2.73E-04
GO:0043408~regulation of MAPK cascade	3.59E-04
GO:0048646~anatomical structure formation involved in morphogenesis	4.04E-04
GO:0009967~positive regulation of signal transduction	4.08E-04
GO:0010647~positive regulation of cell communication	4.84E-04
GO:0006796~phosphate-containing compound metabolic process	5.31E-04
GO:0023056~positive regulation of signalling	5.33E-04
GO:0043269~regulation of ion transport	5.47E-04
GO:0006793~phosphorus metabolic process	5.60E-04
GO:0008283~cell proliferation	5.76E-04

GO:1904062~regulation of cation transmembrane transport	6.54E-04
GO:1904951~positive regulation of establishment of protein localization	6.82E-04
GO:0070201~regulation of establishment of protein localization	7.52E-04
GO:0031401~positive regulation of protein modification process	7.80E-04
GO:0023051~regulation of signalling	8.22E-04
GO:2000026~regulation of multicellular organismal development	8.64E-04
GO:0048015~phosphatidylinositol-mediated signalling	9.07E-04
GO:0032412~regulation of ion transmembrane transporter activity	9.44E-04
GO:0060341~regulation of cellular localization	9.45E-04
GO:0022603~regulation of anatomical structure morphogenesis	9.94E-04
GO:0048017~inositol lipid-mediated signalling	1.02E-03
GO:0032268~regulation of cellular protein metabolic process	1.08E-03
GO:0051093~negative regulation of developmental process	1.09E-03
GO:0022898~regulation of transmembrane transporter activity	1.14E-03
GO:0009966~regulation of signal transduction	1.15E-03
GO:0050790~regulation of catalytic activity	1.19E-03
GO:0009893~positive regulation of metabolic process	1.20E-03
GO:0051223~regulation of protein transport	1.22E-03
GO:0051241~negative regulation of multicellular organismal process	1.32E-03
GO:1902533~positive regulation of intracellular signal transduction	1.45E-03
GO:0010959~regulation of metal ion transport	1.49E-03
GO:0032270~positive regulation of cellular protein metabolic process	1.49E-03
GO:0060537~muscle tissue development	1.51E-03
GO:0051222~positive regulation of protein transport	1.52E-03
GO:0010646~regulation of cell communication	1.56E-03
GO:0051240~positive regulation of multicellular organismal process	1.69E-03
GO:0031325~positive regulation of cellular metabolic process	1.71E-03
GO:0019935~cyclic-nucleotide-mediated signalling	1.84E-03
GO:0032409~regulation of transporter activity	1.86E-03
GO:0051924~regulation of calcium ion transport	1.86E-03
GO:0006928~movement of cell or subcellular component	1.89E-03
GO:0010604~positive regulation of macromolecule metabolic process	1.97E-03
GO:0048870~cell motility	2.00E-03
GO:0051674~localization of cell	2.00E-03
GO:0071900~regulation of protein serine/threonine kinase activity	2.11E-03
GO:0001568~blood vessel development	2.17E-03
GO:0031214~biomineral tissue development	2.31E-03
GO:0016477~cell migration	2.71E-03
GO:0030335~positive regulation of cell migration	2.86E-03
GO:0001837~epithelial to mesenchymal transition	2.88E-03
GO:0014706~striated muscle tissue development	2.96E-03
GO:0048584~positive regulation of response to stimulus	3.11E-03
GO:0010038~response to metal ion	3.21E-03
GO:0042127~regulation of cell proliferation	3.32E-03

GO:0032844~regulation of homeostatic process	3.38E-03
GO:0007166~cell surface receptor signalling pathway	3.40E-03
GO:0016525~negative regulation of angiogenesis	3.45E-03
GO:0044057~regulation of system process	3.62E-03
GO:0048659~smooth muscle cell proliferation	3.63E-03
GO:0001503~ossification	3.63E-03
GO:0008015~blood circulation	3.77E-03
GO:2000147~positive regulation of cell motility	3.84E-03
GO:2000181~negative regulation of blood vessel morphogenesis	3.87E-03
GO:0051247~positive regulation of protein metabolic process	3.88E-03
GO:0003013~circulatory system process	3.95E-03
GO:0001944~vasculature development	3.96E-03
GO:0010628~positive regulation of gene expression	4.04E-03
GO:0001525~angiogenesis	4.17E-03
GO:0030855~epithelial cell differentiation	4.22E-03
GO:0007267~cell-cell signalling	4.25E-03
GO:0001667~ameboidal-type cell migration	4.26E-03
GO:0006811~ion transport	4.27E-03
GO:0048660~regulation of smooth muscle cell proliferation	4.34E-03
GO:0048596~embryonic camera-type eye morphogenesis	4.50E-03
GO:0051128~regulation of cellular component organization	4.50E-03
GO:0006816~calcium ion transport	4.51E-03
GO:0019932~second-messenger-mediated signalling	4.52E-03
GO:0042592~homeostatic process	0.0046
GO:0060429~epithelium development	0.0047
GO:0051272~positive regulation of cellular component movement	0.0048
GO:0048304~positive regulation of isotype switching to IgG isotypes	0.0048
GO:0040011~locomotion	0.0048
GO:0043405~regulation of MAP kinase activity	0.0050
GO:0040017~positive regulation of locomotion	0.0050
GO:0050729~positive regulation of inflammatory response	0.0051
GO:0044093~positive regulation of molecular function	0.0051
GO:0007167~enzyme linked receptor protein signalling pathway	0.0052
GO:0032880~regulation of protein localization	0.0057
GO:0098662~inorganic cation transmembrane transport	0.0057
GO:2000021~regulation of ion homeostasis	0.0057
GO:0034220~ion transmembrane transport	0.0058
GO:0007179~transforming growth factor beta receptor signalling pathway	0.0059
GO:0001775~cell activation	0.0060
GO:0034764~positive regulation of transmembrane transport	0.0060
GO:2000242~negative regulation of reproductive process	0.0061
GO:0048291~isotype switching to IgG isotypes	0.0061
GO:1903867~extraembryonic membrane development	0.0061

Appendix 24: Canonical pathways enriched in genes up-regulated by BILF1 in GC B cells

Pathway	*p*-value
Respiratory electron transport ATP synthesis by chemi-osmotic coupling and heat production by uncoupling proteins	9E-28
TCA cycle and respiratory electron transport	3E-24
Respiratory electron transport	3E-23
Influenza viral RNA transcription and replication	5E-17
SRP dependent cotranslational protein targeting to membrane	1E-16
Influenza life cycle	2E-15
Peptide chain elongation	2E-15
Translation	6E-15
Metabolism of RNA	1E-14
3 UTR mediated translational regulation	2E-14
Nonsense mediated decay enhanced by the exon junction complex	2E-14
Metabolism of proteins	3E-14
Metabolism of mRNA	6E-14
Hemostasis	4E-06
Formation of ATP by chemiosmotic coupling	5E-06
Formation of the ternary complex and subsequently the 43S complex	2E-05
Immune system	4E-05
Activation of the mRNA upon binding of the cap binding complex and EIFS and subsequent binding to 43S	4E-05
Ion channel transport	6E-05
DNA repair	2E-04
Signalling by NOTCH1	2E-04
Platelet activation signalling and aggregation	2E-04
MRNA SPLICING MINOR pathway	3E-04
Transmembrane transport of small molecules	4E-04
DIABETES pathways	5E-04
Response to elevated platelet cytosolic Ca2	5E-04
Ligand gated ion channel transport	6E-04
Transmission across chemical synapses	6E-04
Cell surface interactions at the vascular wall	6E-04
VIF mediated degradation of APOBEC3G	7E-04
Neurotransmitter receptor binding and downstream transmission in the postsynaptic cell	7E-04
SCFSKP2 mediated degradation of P27 P21	8E-04
HIV infection	9E-04
P53 dependent G1 DNA damage response	9E-04
Signalling by NOTCH	1E-03
Activated NOTCH1 transmits signal to the nucleus	1E-03
Autodegradation of CDH1 by CDH1 APC C	1E-03
Prefoldin mediated transfer of substrate to CCT TRIC	1E-03

Cytokine signalling in immune system	1E-03
Interferon signalling	1E-03
Cyclin e associated events during G1 s transition	1E-03
Signalling by WNT	1E-03
Regulation of the Fanconi anaemia pathway	2E-03
APC C CDH1 mediated degradation of CDC20 and other APC C CDH1 targeted proteins in late mitosis early G1	2E-03
ORC1 removal from chromatin	2E-03
APC C CDC20 mediated degradation of mitotic proteins	2E-03
Cell cycle checkpoints	2E-03
Gluconeogenesis	2E-03
Integration of energy metabolism	2E-03
Metabolism of nucleotides	2E-03
NFκB and map kinases activation mediated by TLR4 signalling repertoire	2E-03
TRIF mediated TLR3 signalling	3E-03
PECAM1 interactions	3E-03
Adaptive immune system	3E-03
G alpha Q signalling events	3E-03
Platelet homeostasis	3E-03
TRAF6 mediated induction of NFκB and map kinases upon TLR7 8 or 9 activation	3E-03
Unfolded protein response	3E-03
Host interactions of HIV factors	3E-03
The NLRP3 inflammasome	3E-03
Regulation of mitotic cell cycle	3E-03
Developmental biology	3E-03
G alpha I signalling events	4E-03
Hormone sensitive lipase HSL mediated triacylglycerol hydrolysis	4E-03
MYD88 mal cascade initiated on plasma membrane	4E-03
Regulation of MRNA stability by proteins that bind au rich elements	4E-03
Metabolism of lipids and lipoproteins	4E-03
Mitotic G1 G1-S phases	4E-03
Cell cycle	5E-03
G alpha Z signalling events	5E-03
GASTRIN CREB signalling pathway via PKC and MAPK	5E-03
Signalling by ERBB4	5E-03
Viral messenger RNA synthesis	5E-03
Activation of chaperone genes by XBP1S	5E-03
Notch1 intracellular domain regulates transcription	5E-03
nucleotide binding domain leucine rich repeat containing receptor NLR signalling pathways	5E-03
Innate immune system	5E-03
Synthesis of DNA	6E-03
Apoptosis	6E-03
CDK mediated phosphorylation and removal of CDC6	6E-03
Activated TLR4 signalling	6E-03

NRIF signals cell death from the nucleus	6E-03
Neuronal system	6E-03
Downstream signalling events of B cell receptor BCR	6E-03
Autodegradation of the E3 ubiquitin ligase COP1	7E-03
Mitochondrial protein import	7E-03
P53 independent G1 S DNA damage checkpoint	7E-03
Signalling by NGF	7E-03
Inflammasomes	7E-03
Platelet sensitization by LDL	7E-03
SCF beta TRCP mediated degradation of EMI1	7E-03
Destabilization of mRNA by AUF1 HNRNP D0	7E-03
Protein folding	7E-03
MicroRNA miRNA biogenesis	9E-03
Oxygen dependent proline hydroxylation of hypoxia inducible factor alpha	9E-03
PD1 signalling	9E-03
Synthesis and interconversion of nucleotide di and triphosphates	9E-03
CDT1 association with the cdc6 orc origin complex	9E-03
mRNA splicing	9E-03
Abortive elongation of HIV1 transcript in the absence of tat	9E-03
Signalling by EGFR in cancer	1E-02
S phase	1E-02
Regulation of apoptosis	1E-02
Fanconi anaemia pathway	1E-02
G1 s transition	1E-02
Metabolism of carbohydrates	1E-02
Cell cycle mitotic	1E-02
APC CCDC20 mediated degradation of cyclin B	1E-02
Cholesterol biosynthesis	1E-02
Regulatory RNA pathways	1E-02
ER phagosome pathway	1E-02
Interferon gamma signalling	1E-02
Downregulation of TGF beta receptor signalling	1E-02
Toll receptor cascades	1E-02
Activation of NF kappa-B in B cells	1E-02
Interferon alpha beta signalling	1E-02
APC CDC20 mediated degradation of NEK2A	1E-02
Thromboxane signalling through tp receptor	1E-02
Assembly of the pre replicative complex	1E-02
Glucose metabolism	1E-02
Signalling by the B cell receptor BCR	2E-02
ADP signalling through P2RY1	2E-02
EGFR downregulation	2E-02
Regulation of hypoxia inducible factor HIF by oxygen	2E-02
TGF beta receptor signalling activates SMADS	2E-02

Phase1 functionalization of compounds	2E-02
Membrane trafficking	2E-02
Endosomal sorting complex required for transport ESCRT	2E-02
Myogenesis	2E-02
Trafficking of AMPA receptors	2E-02
RIG I MDA5 mediated induction of IFN ALPHA BETA pathways	2E-02
Elongation arrest and recovery	2E-02
Antigen processing cross presentation	2E-02
Processing of capped intron containing pre mRNA	2E-02
NGF signalling via TRKA from the plasma membrane	2E-02
Formation of the hiv1 early elongation complex	2E-02
Formation of transcription coupled NER TC NER repair complex	2E-02
mRNA capping	2E-02
Antigen processing ubiquitination proteasome degradation	2E-02
Opioid signalling	2E-02
Biological oxidations	2E-02
Integrin cell surface interactions	2E-02
M G1 transition	2E-02
NEG regulators of RIG I MDA5 signalling	2E-02
Signal amplification	2E-02
Amyloids	3E-02
Thrombin signalling through proteinase activated receptors pars	3E-02
Ion transport by p type ATPases	3E-02
Post NMDA receptor activation events	3E-02
Signal transduction by L1	3E-02
Triglyceride biosynthesis	3E-02
mRNA processing	3E-02
Regulation of insulin secretion	3E-02
Activation of NMDA receptor upon glutamate binding and postsynaptic events	3E-02
G1 phase	3E-02
Nuclear signalling by ERBB4	4E-02
IIL1 signalling	4E-02
Formation of RNA polymerase II elongation complex	4E-02
RNA polymerase II transcription pre initiation and promoter opening	4E-02
Fatty acid triacylglycerol and ketone body metabolism	4E-02
Late phase of HIV life cycle	4E-02
Signalling by ERBB2	4E-02
Axon guidance	4E-02
Class I MHC mediated antigen processing presentation	4E-02
Transcription coupled NER TC NER	5E-02
Lipid digestion mobilization and transport	5E-02
Generic transcription pathway	5E-02

Appendix 25: Canonical pathways enriched in genes down-regulated by BILF1 in GC B cells

Pathway	*p*-value
GPCR ligand binding	4E-06
Class A1 rhodopsin like receptors	7E-05
Signalling by GPCR	1E-04
Immune system	4E-04
G alpha I signalling events	1E-03
Hemostasis	2E-03
P2Y receptors	2E-03
Activated TLR4 signalling	2E-03
Lipid digestion mobilization and transport	2E-03
Transmembrane transport of small molecules	3E-03
Signalling by ILS	4E-03
Nucleotide like purinergic receptors	4E-03
Toll receptor cascades	5E-03
Cytokine signalling in immune system	5E-03
Co-stimulation by the CD28 family	6E-03
Metabolism of lipids and lipoproteins	8E-03
PI3K cascade	8E-03
GPCR downstream signalling	9E-03
NFκB and MAP kinases activation mediated by TLR4 signalling repertoire	9E-03
Transcriptional regulation of white adipocyte differentiation	9E-03
G0 and early G1	1E-02
TRAF6 mediated induction of NFκB and MAP kinases upon TLR7 8 or 9 activation	1E-02
Mitotic G2 G2-M phases	1E-02
Basigin interactions	1E-02
Lipoprotein metabolism	1E-02
The role of NEF in HIV1 replication and disease pathogenesis	1E-02
MYD88 MAL cascade initiated on plasma membrane	1E-02
Class B 2 secretin family receptors	1E-02
Insulin receptor signalling cascade	1E-02
SLC mediated transmembrane transport	1E-02
Transport of inorganic cations anions and amino acids oligopeptides	2E-02
Amine ligand binding receptors	2E-02
Peptide ligand binding receptors	2E-02
IL1 signalling	2E-02
Signalling by insulin receptor	3E-02
Voltage gated potassium channels	3E-02
Developmental biology	3E-02
Integration of energy metabolism	3E-02

Cell-cell communication	3E-02
Signalling by PDGF	3E-02
MAP kinase activation in TLR cascade	3E-02
Adaptive immune system	4E-02
Chemokine receptors bind chemokines	4E-02
Cell cycle mitotic	4E-02
Phospholipase C mediated cascade	4E-02
Factors involved in megakaryocyte development and platelet production	4E-02
Loss of NLP from mitotic centrosomes	4E-02
Prostanoid ligand receptors	5E-02

CHAPTER 6

REFERENCES

7. REFERENCES

ABATE, F., AMBROSIO, M. R., MUNDO, L., LAGINESTRA, M. A., FULIGNI, F., ROSSI, M., ZAIRIS, S., GAZANEO, S., DE FALCO, G., LAZZI, S., BELLAN, C., ROCCA, B. J., AMATO, T., MARASCO, E., ETEBARI, M., OGWANG, M., CALBI, V., NDEDE, I., PATEL, K., CHUMBA, D., PICCALUGA, P. P., PILERI, S., LEONCINI, L. & RABADAN, R. 2015. Distinct Viral and Mutational Spectrum of Endemic Burkitt Lymphoma. *PLOS Pathogens,* 11**,** e1005158.

ABBOT, S. D., ROWE, M., CADWALLADER, K., RICKSTEN, A., GORDON, J., WANG, F., RYMO, L. & RICKINSON, A. B. 1990. Epstein-Barr virus nuclear antigen 2 induces expression of the virus-encoded latent membrane protein. *J Virol,* 64**,** 2126-34.

ADAMSON, A. L., DARR, D., HOLLEY-GUTHRIE, E., JOHNSON, R. A., MAUSER, A., SWENSON, J. & KENNEY, S. 2000. Epstein-Barr virus immediate-early proteins BZLF1 and BRLF1 activate the ATF2 transcription factor by increasing the levels of phosphorylated p38 and c-Jun N-terminal kinases. *J Virol,* 74**,** 1224-33.

ALFIERI, C., BIRKENBACH, M. & KIEFF, E. 1991. Early events in Epstein-Barr virus infection of human B lymphocytes. *Virology,* 181**,** 595-608.

ALLAN, G. J., INMAN, G. J., PARKER, B. D., ROWE, D. T. & FARRELL, P. J. 1992. Cell growth effects of Epstein-Barr virus leader protein. *J Gen Virol,* 73 (Pt 6)**,** 1547-51.

ALTMANN, M. & HAMMERSCHMIDT, W. 2005. Epstein-Barr virus provides a new paradigm: a requirement for the immediate inhibition of apoptosis. *PLoS Biol,* 3**,** e404.

ANDERTON, J. A., BOSE, S., VOCKERODT, M., VRZALIKOVA, K., WEI, W., KUO, M., HELIN, K., CHRISTENSEN, J., ROWE, M., MURRAY, P. G. & WOODMAN, C. B. 2011. The H3K27me3 demethylase, KDM6B, is induced by Epstein-Barr virus and over-expressed in Hodgkin's Lymphoma. *Oncogene,* 30**,** 2037-43.

ANGEL, P., IMAGAWA, M., CHIU, R., STEIN, B., IMBRA, R. J., RAHMSDORF, H. J., JONAT, C., HERRLICH, P. & KARIN, M. 1987. Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. *Cell,* 49**,** 729-39.

APPELBAUM, F. R., DEISSEROTH, A. B., GRAW, R. G., HERZIG, G. P., LEVINE, A. S., MAGRATH, I. T., PIZZO, P. A., POPLACK, D. G. & ZIEGLER, J. L. 1978. Prolonged complete remission following high dose chemotherapy of burkitt's lymphoma in relapse. *Cancer,* 41**,** 1059-1063.

ARRAND, J. R. & RYMO, L. 1982. Characterization of the major Epstein-Barr virus-specific RNA in Burkitt lymphoma-derived cells. *J Virol,* 41**,** 376-89.

ARRAND, J. R., WALSH-ARRAND, J. E. & RYMO, L. 1983. Cytoplasmic RNA from normal and malignant human cells shows homology to the DNAs of Epstein-Barr virus and human adenoviruses. *EMBO J,* 2**,** 1673-83.

ARTAVANIS-TSAKONAS, S., MATSUNO, K. & FORTINI, M. E. 1995. Notch signaling. *Science,* 268**,** 225-32.

BABCOCK, G. J., DECKER, L. L., FREEMAN, R. B. & THORLEY-LAWSON, D. A. 1999. Epstein-Barr virus–infected resting memory B cells, not proliferating lymphoblasts, accumulate in the peripheral blood of immunosuppressed patients. *Journal of Experimental Medicine,* 190**,** 567-576.

BABCOCK, G. J. & THORLEY-LAWSON, D. A. 2000. Tonsillar memory B cells, latently infected with Epstein–Barr virus, express the restricted pattern of latent genes

previously found only in Epstein–Barr virus-associated tumors. *Proceedings of the National Academy of Sciences,* 97**,** 12250-12255.

BAER, R., BANKIER, A. T., BIGGIN, M. D., DEININGER, P. L., FARRELL, P. J., GIBSON, T. J., HATFULL, G., HUDSON, G. S., SATCHWELL, S. C., SÉGUIN, C., TUFFNELL, P. S. & BARRELL, B. G. 1984. DNA sequence and expression of the B95-8 Epstein—Barr virus genome. *Nature,* 310**,** 207-211.

BAJAJ, B. G., MURAKAMI, M., CAI, Q., VERMA, S. C., LAN, K. & ROBERTSON, E. S. 2008. Epstein-Barr Virus Nuclear Antigen 3C Interacts with and Enhances the Stability of the c-Myc Oncoprotein. *Journal of Virology,* 82**,** 4082-4090.

BARGOU, R. C., EMMERICH, F., KRAPPMANN, D., BOMMERT, K., MAPARA, M. Y., ARNOLD, W., ROYER, H. D., GRINSTEIN, E., GREINER, A., SCHEIDEREIT, C. & DÖRKEN, B. 1997. Constitutive nuclear factor-kappaB-RelA activation is required for proliferation and survival of Hodgkin's disease tumor cells. *Journal of Clinical Investigation,* 100**,** 2961-2969.

BARTH, S., PFUHL, T., MAMIANI, A., EHSES, C., ROEMER, K., KREMMER, E., JAKER, C., HOCK, J., MEISTER, G. & GRASSER, F. A. 2008. Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. *Nucleic Acids Res,* 36**,** 666-75.

BARTON, E. S., WHITE, D. W., CATHELYN, J. S., BRETT-MCCLELLAN, K. A., ENGLE, M., DIAMOND, M. S., MILLER, V. L. & VIRGIN, H. W. 2007. Herpesvirus latency confers symbiotic protection from bacterial infection. *Nature,* 447**,** 326-329.

BAUMFORTH, K. R., BIRGERSDOTTER, A., REYNOLDS, G. M., WEI, W., KAPATAI, G., FLAVELL, J. R., KALK, E., PIPER, K., LEE, S., MACHADO, L., HADLEY, K., SUNDBLAD, A., SJOBERG, J., BJORKHOLM, M., PORWIT, A. A., YAP, L. F., TEO, S., GRUNDY, R. G., YOUNG, L. S., ERNBERG, I., WOODMAN, C. B. & MURRAY, P. G. 2008. Expression of the Epstein-Barr virus-encoded Epstein-Barr virus nuclear antigen 1 in Hodgkin's lymphoma cells mediates Up-regulation of CCL20 and the migration of regulatory T cells. *Am J Pathol,* 173**,** 195-204.

BÉGUELIN, W., POPOVIC, R., TEATER, M., JIANG, Y., BUNTING, K. L., ROSEN, M., SHEN, H., YANG, S. N., WANG, L. & EZPONDA, T. 2013. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer cell,* 23**,** 677-692.

BEISSER, P. S., VERZIJL, D., GRUIJTHUIJSEN, Y. K., BEUKEN, E., SMIT, M. J., LEURS, R., BRUGGEMAN, C. A. & VINK, C. 2005. The Epstein-Barr virus BILF1 gene encodes a G protein-coupled receptor that inhibits phosphorylation of RNA-dependent protein kinase. *J Virol,* 79**,** 441-9.

BELLO-FERNANDEZ, C., PACKHAM, G. & CLEVELAND, J. L. 1993. The ornithine decarboxylase gene is a transcriptional target of c-Myc. *Proc Natl Acad Sci U S A,* 90**,** 7804-8.

BELLOWS, D. S., HOWELL, M., PEARSON, C., HAZLEWOOD, S. A. & HARDWICK, J. M. 2002. Epstein-Barr virus BALF1 is a BCL-2-like antagonist of the herpesvirus antiapoptotic BCL-2 proteins. *J Virol,* 76**,** 2469-79.

BENTZ, G. L., MOSS, C. R., WHITEHURST, C. B., MOODY, C. A. & PAGANO, J. S. 2015. LMP1-Induced Sumoylation Influences the Maintenance of Epstein-Barr Virus Latency through KAP1. *Journal of Virology,* 89**,** 7465-7477.

BENTZ, G. L., WHITEHURST, C. B. & PAGANO, J. S. 2011. Epstein-Barr Virus Latent Membrane Protein 1 (LMP1) C-Terminal-Activating Region 3 Contributes to LMP1-Mediated Cellular Migration via Its Interaction with Ubc9. *Journal of Virology,* 85**,** 10144-10153.

BERNHARDT, K., HAAR, J., TSAI, M.-H., POIREY, R., FEEDERLE, R. & DELECLUSE, H.-J. 2016. A Viral microRNA Cluster Regulates the Expression of PTEN, p27 and of a bcl-2 Homolog. *PLOS Pathogens,* 12**,** e1005405.

BHAT, R. A. & THIMMAPPAYA, B. 1985. Construction and analysis of additional adenovirus substitution mutants confirm the complementation of VAI RNA function by two small RNAs encoded by Epstein-Barr virus. *J Virol,* 56**,** 750-6.

BIECHE, I., VACHER, S., LALLEMAND, F., TOZLU-KARA, S., BENNANI, H., BEUZELIN, M., DRIOUCH, K., ROULEAU, E., LEREBOURS, F., RIPOCHE, H., CIZERON-CLAIRAC, G., SPYRATOS, F. & LIDEREAU, R. 2011. Expression analysis of mitotic spindle checkpoint genes in breast carcinoma: role of NDC80/HEC1 in early breast tumorigenicity, and a two-gene signature for aneuploidy. *Mol Cancer,* 10**,** 23.

BIGGIN, M., BODESCOT, M., PERRICAUDET, M. & FARRELL, P. 1987. Epstein-Barr virus gene expression in P3HR1-superinfected Raji cells. *J Virol,* 61**,** 3120-32.

BILELLO, J. P., LANG, S. M., WANG, F., ASTER, J. C. & DESROSIERS, R. C. 2006. Infection and Persistence of Rhesus Monkey Rhadinovirus in Immortalized B-Cell Lines. *Journal of Virology,* 80**,** 3644-3649.

BIRKENMEIER, K., DRÖSE, S., WITTIG, I., WINKELMANN, R., KÄFER, V., DÖRING, C., HARTMANN, S., WENZ, T., REICHERT, A. S., BRANDT, U. & HANSMANN, M.-L. 2016. Hodgkin and Reed-Sternberg cells of classical Hodgkin lymphoma are highly dependent on oxidative phosphorylation. *International Journal of Cancer,* 138**,** 2231-2246.

BODESCOT, M., BRISON, O. & PERRICAUDET, M. 1986. An Epstein-Barr virus transcription unit is at least 84 kilobases long. *Nucleic Acids Res,* 14**,** 2611-20.

BODESCOT, M., PERRICAUDET, M. & FARRELL, P. J. 1987. A promoter for the highly spliced EBNA family of RNAs of Epstein-Barr virus. *J Virol,* 61**,** 3424-30.

BOLGER, A. M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* 30**,** 2114-20.

BONIZZI, G. & KARIN, M. 2004. The two NF-κB activation pathways and their role in innate and adaptive immunity. *Trends in Immunology,* 25**,** 280-288.

BRÄUNINGER, A., SCHMITZ, R., BECHTEL, D., RENNÉ, C., HANSMANN, M.-L. & KÜPPERS, R. 2005. Molecular biology of Hodgkin's and Reed/Sternberg cells in Hodgkin's lymphoma. *International Journal of Cancer,* 118**,** 1853-1861.

BRAUNINGER, A., SCHMITZ, R., BECHTEL, D., RENNE, C., HANSMANN, M. L. & KUPPERS, R. 2006. Molecular biology of Hodgkin's and Reed/Sternberg cells in Hodgkin's lymphoma. *Int J Cancer,* 118**,** 1853-61.

BRODEUR, S. R., CHENG, G., BALTIMORE, D. & THORLEY-LAWSON, D. A. 1997. Localization of the major NF-kappaB-activating site and the sole TRAF3 binding site of LMP-1 defines two distinct signaling motifs. *J Biol Chem,* 272**,** 19777-84.

BROOKS, L., YAO, Q. Y., RICKINSON, A. B. & YOUNG, L. S. 1992. Epstein-Barr virus latent gene transcription in nasopharyngeal carcinoma cells: coexpression of EBNA1, LMP1, and LMP2 transcripts. *J Virol,* 66**,** 2689-97.

BROOKS, L. A., LEAR, A. L., YOUNG, L. S. & RICKINSON, A. B. 1993. Transcripts from the Epstein-Barr virus BamHI A fragment are detectable in all three forms of virus latency. *J Virol,* 67**,** 3182-90.

BRUNE, V., TIACCI, E., PFEIL, I., DORING, C., ECKERLE, S., VAN NOESEL, C. J., KLAPPER, W., FALINI, B., VON HEYDEBRECK, A., METZLER, D., BRAUNINGER, A., HANSMANN, M. L. & KUPPERS, R. 2008. Origin and pathogenesis of nodular lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis. *J Exp Med,* 205**,** 2251-68.

BURKITT, D. 1958. A sarcoma involving the jaws in African children. *Br J Surg,* 46**,** 218-23.

BURKITT, D. 1962a. A children's cancer dependent on climatic factors. *Nature,* 194**,** 232-4.

BURKITT, D. 1962b. Determining the climatic limitations of a children's cancer common in Africa. *Br Med J,* 2**,** 1019-23.

CABANNES, E., KHAN, G., AILLET, F., JARRETT, R. F. & HAY, R. T. 1999. Mutations in the IkBa gene in Hodgkin's disease suggest a tumour suppressor role for IkappaBalpha. *Oncogene,* 18**,** 3063-70.

CABRAS, G., DECAUSSIN, G., ZENG, Y., DJENNAOUI, D., MELOULI, H., BROULLY, P., BOUGUERMOUH, A. M. & OOKA, T. 2005. Epstein-Barr virus encoded BALF1 gene is transcribed in Burkitt's lymphoma cell lines and in nasopharyngeal carcinoma's biopsies. *J Clin Virol,* 34**,** 26-34.

CAI, X., SCHAFER, A., LU, S., BILELLO, J. P., DESROSIERS, R. C., EDWARDS, R., RAAB-TRAUB, N. & CULLEN, B. R. 2006. Epstein-Barr virus microRNAs are evolutionarily conserved and differentially expressed. *PLoS Pathog,* 2**,** e23.

CALDWELL, R. G., BROWN, R. C. & LONGNECKER, R. 2000. Epstein-Barr virus LMP2A-induced B-cell survival in two unique classes of EmuLMP2A transgenic mice. *J Virol,* 74**,** 1101-13.

CALDWELL, R. G., WILSON, J. B., ANDERSON, S. J. & LONGNECKER, R. 1998. Epstein-Barr virus LMP2A drives B cell development and survival in the absence of normal B cell receptor signals. *Immunity,* 9**,** 405-11.

CARBONE, A., GLOGHINI, A., GATTEI, V., ALDINUCCI, D., DEGAN, M., DE PAOLI, P., ZAGONEL, V. & PINTO, A. 1995a. Expression of functional CD40 antigen on Reed-Sternberg cells and Hodgkin's disease cell lines. *Blood,* 85**,** 780-9.

CARBONE, A., GLOGHINI, A., GRUSS, H. J. & PINTO, A. 1995b. CD40 ligand is constitutively expressed in a subset of T cell lymphomas and on the microenvironmental reactive T cells of follicular lymphomas and Hodgkin's disease. *Am J Pathol,* 147**,** 912-22.

CARBONE, P. P., KAPLAN, H. S., MUSSHOFF, K., SMITHERS, D. W. & TUBIANA, M. 1971. Report of the Committee on Hodgkin's Disease Staging Classification. *Cancer Res,* 31**,** 1860-1.

CARO, P., KISHAN, A. U., NORBERG, E., STANLEY, I. A., CHAPUY, B., FICARRO, S. B., POLAK, K., TONDERA, D., GOUNARIDES, J., YIN, H., ZHOU, F., GREEN, M. R., CHEN, L., MONTI, S., MARTO, J. A., SHIPP, M. A. & DANIAL, N. N. 2012. Metabolic signatures uncover distinct targets in molecular subsets of diffuse large B cell lymphoma. *Cancer Cell,* 22**,** 547-60.

CARPENTER, L. M., NEWTON, R., CASABONNE, D., ZIEGLER, J., MBULAITEYE, S., MBIDDE, E., WABINGA, H., JAFFE, H. & BERAL, V. 2008. Antibodies against malaria and Epstein-Barr virus in childhood Burkitt lymphoma: a case-control study in Uganda. *Int J Cancer,* 122**,** 1319-23.

CARUCCIO, N. 2011. Preparation of next-generation sequencing libraries using Nextera™ technology: simultaneous DNA fragmentation and adaptor tagging by in vitro transposition. *High-Throughput Next Generation Sequencing: Methods and Applications***,** 241-255.

CHANG, D. W., CLAASSEN, G. F., HANN, S. R. & COLE, M. D. 2000. The c-Myc transactivation domain is a direct modulator of apoptotic versus proliferative signals. *Mol Cell Biol,* 20**,** 4309-19.

CIBULSKIS, K., LAWRENCE, M. S., CARTER, S. L., SIVACHENKO, A., JAFFE, D., SOUGNEZ, C., GABRIEL, S., MEYERSON, M., LANDER, E. S. & GETZ, G.

2013. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol,* 31**,** 213-9.

CLEMENS, M. J., LAING, K. G., JEFFREY, I. W., SCHOFIELD, A., SHARP, T. V., ELIA, A., MATYS, V., JAMES, M. C. & TILLERAY, V. J. 1994. Regulation of the interferon-inducible eIF-2 alpha protein kinase by small RNAs. *Biochimie,* 76**,** 770-8.

COHEN, J. I., WANG, F., MANNICK, J. & KIEFF, E. 1989. Epstein-Barr virus nuclear protein 2 is a key determinant of lymphocyte transformation. *Proceedings of the National Academy of Sciences,* 86**,** 9558-9562.

COLLER, H. A., GRANDORI, C., TAMAYO, P., COLBERT, T., LANDER, E. S., EISENMAN, R. N. & GOLUB, T. R. 2000. Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. *Proc Natl Acad Sci U S A,* 97**,** 3260-5.

CONNELLY, R. R. & CHRISTINE, B. W. 1974. A cohort study of cancer following infectious mononucleosis. *Cancer Res,* 34**,** 1172-8.

CORDIER, M., CALENDER, A., BILLAUD, M., ZIMBER, U., ROUSSELET, G., PAVLISH, O., BANCHEREAU, J., TURSZ, T., BORNKAMM, G. & LENOIR, G. M. 1990. Stable transfection of Epstein-Barr virus (EBV) nuclear antigen 2 in lymphoma cells containing the EBV P3HR1 genome induces expression of B-cell activation molecules CD21 and CD23. *J Virol,* 64**,** 1002-13.

COUNTRYMAN, J. & MILLER, G. 1985. Activation of expression of latent Epstein-Barr herpesvirus after gene transfer with a small cloned subfragment of heterogeneous viral DNA. *Proc Natl Acad Sci U S A,* 82**,** 4085-9.

CRAWFORD, D. H., RICKINSON, A. & JOHANNESSEN, I. 2014. *Cancer virus: the story of Epstein-Barr virus,* Oxford University Press.

D'SOUZA, B. N., EDELSTEIN, L. C., PEGMAN, P. M., SMITH, S. M., LOUGHRAN, S. T., CLARKE, A., MEHL, A., ROWE, M., GELINAS, C. & WALLS, D. 2004. Nuclear factor kappa B-dependent activation of the antiapoptotic bfl-1 gene by the Epstein-Barr virus latent membrane protein 1 and activated CD40 receptor. *J Virol,* 78**,** 1800-16.

DAMBAUGH, T., HENNESSY, K., CHAMNANKIT, L. & KIEFF, E. 1984. U2 region of Epstein-Barr virus DNA may encode Epstein-Barr nuclear antigen 2. *Proc Natl Acad Sci U S A,* 81**,** 7632-6.

DANG, C. V., LE, A. & GAO, P. 2009. MYC-Induced Cancer Cell Energy Metabolism and Therapeutic Opportunities. *Clinical Cancer Research,* 15**,** 6479-6483.

DAVISON, A. J., EBERLE, R., EHLERS, B., HAYWARD, G. S., MCGEOCH, D. J., MINSON, A. C., PELLETT, P. E., ROIZMAN, B., STUDDERT, M. J. & THIRY, E. 2009. The order Herpesvirales. *Arch Virol,* 154**,** 171-7.

DE-THÉ, G., GESER, A., DAY, N. E., TUKEI, P. M., WILLIAMS, E. H., BERI, D. P., SMITH, P. G., DEAN, A. G., BORNKAMM, G. W., FEORINO, P. & HENLE, W. 1978. Epidemiological evidence for causal relationship between Epstein-Barr virus and Burkitt's lymphoma from Ugandan prospective study. *Nature,* 274**,** 756-761.

DE JESUS, O., SMITH, P. R., SPENDER, L. C., ELGUETA KARSTEGL, C., NILLER, H. H., HUANG, D. & FARRELL, P. J. 2003. Updated Epstein-Barr virus (EBV) DNA sequence and analysis of a promoter for the BART (CST, BARF0) RNAs of EBV. *J Gen Virol,* 84**,** 1443-50.

DE MARTEL, C., FERLAY, J., FRANCESCHI, S., VIGNAT, J., BRAY, F., FORMAN, D. & PLUMMER, M. 2012. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol,* 13**,** 607-15.

DEACON, E. M. 1993. Epstein-Barr virus and Hodgkin's disease: transcriptional analysis of virus latency in the malignant cells. *Journal of Experimental Medicine,* 177**,** 339-349.

DEACON, E. M., PALLESEN, G., NIEDOBITEK, G., CROCKER, J., BROOKS, L., RICKINSON, A. B. & YOUNG, L. S. 1993. Epstein-Barr virus and Hodgkin's disease: transcriptional analysis of virus latency in the malignant cells. *J Exp Med,* 177**,** 339-49.

DECAUSSIN, G., LECLERC, V. & OOKA, T. 1995. The lytic cycle of Epstein-Barr virus in the nonproducer Raji line can be rescued by the expression of a 135-kilodalton protein encoded by the BALF2 open reading frame. *J Virol,* 69**,** 7309-14.

DECAUSSIN, G., SBIH-LAMMALI, F., DE TURENNE-TESSIER, M., BOUGUERMOUH, A. & OOKA, T. 2000. Expression of BARF1 gene encoded by Epstein-Barr virus in nasopharyngeal carcinoma biopsies. *Cancer Res,* 60**,** 5584-8.

DEN HOLLANDER, J., RIMPI, S., DOHERTY, J. R., RUDELIUS, M., BUCK, A., HOELLEIN, A., KREMER, M., GRAF, N., SCHEERER, M., HALL, M. A., GOGA, A., VON BUBNOFF, N., DUYSTER, J., PESCHEL, C., CLEVELAND, J. L., NILSSON, J. A. & KELLER, U. 2010. Aurora kinases A and B are up-regulated by Myc and are essential for maintenance of the malignant state. *Blood,* 116**,** 1498-1505.

DESGRANGES, C., WOLF, H., DE-THÉ, G., SHANMUGARATNAM, K., CAMMOUN, N., ELLOUZ, R., KLEIN, G., LENNERT, K., MUÑOZ, N. & ZUR HAUSEN, H. 1975. Nasopharyngeal carcinoma. X. Presence of Epstein-Barr genomes in separated epithelial cells of tumours in patients from Singapore, Tunisia and Kenya. *International Journal of Cancer,* 16**,** 7-15.

DIEPSTRA, A., NIENS, M., VELLENGA, E., VAN IMHOFF, G. W., NOLTE, I. M., SCHAAPVELD, M., VAN DER STEEGE, G., VAN DEN BERG, A., KIBBELAAR, R. E., TE MEERMAN, G. J. & POPPEMA, S. 2005. Association with HLA class I in Epstein-Barr-virus-positive and with HLA class III in Epstein-Barr-virus-negative Hodgkin's lymphoma. *The Lancet,* 365**,** 2216-2224.

DING, Y., QIN, L., KOTENKO, S. V., PESTKA, S. & BROMBERG, J. S. 2000. A single amino acid determines the immunostimulatory activity of interleukin 10. *J Exp Med,* 191**,** 213-24.

DING, Y., QIN, L., ZAMARIN, D., KOTENKO, S. V., PESTKA, S., MOORE, K. W. & BROMBERG, J. S. 2001. Differential IL-10R1 expression plays a critical role in IL-10-mediated immune regulation. *J Immunol,* 167**,** 6884-92.

DIRMEIER, U., NEUHIERL, B., KILGER, E., REISBACH, G., SANDBERG, M. L. & HAMMERSCHMIDT, W. 2003. Latent membrane protein 1 is critical for efficient growth transformation of human B cells by epstein-barr virus. *Cancer Res,* 63**,** 2982-9.

DOLAN, A., ADDISON, C., GATHERER, D., DAVISON, A. J. & MCGEOCH, D. J. 2006. The genome of Epstein-Barr virus type 2 strain AG876. *Virology,* 350**,** 164-70.

DOLYNIUK, M., PRITCHETT, R. & KIEFF, E. 1976a. Proteins of Epstein-Barr virus. I. Analysis of the polypeptides of purified enveloped Epstein-Barr virus. *J Virol,* 17**,** 935-49.

DOLYNIUK, M., WOLFF, E. & KIEFF, E. 1976b. Proteins of Epstein-Barr Virus. II. Electrophoretic analysis of the polypeptides of the nucleocapsid and the glucosamine- and polysaccharide-containing components of enveloped virus. *J Virol,* 18**,** 289-97.

DOMINGUEZ-BRAUER, C., THU, KELSIE L., MASON, JACQUELINE M., BLASER, H., BRAY, MARK R. & MAK, TAK W. 2015. Targeting Mitosis in Cancer: Emerging Strategies. *Molecular Cell,* 60**,** 524-536.

DONAGHY, G. & JUPP, R. 1995. Characterization of the Epstein-Barr virus proteinase and comparison with the human cytomegalovirus proteinase. *J Virol,* 69**,** 1265-70.

DOZZO, M., CAROBOLANTE, F., DONISI, P. M., SCATTOLIN, A., MAINO, E., SANCETTA, R., VIERO, P. & BASSAN, R. 2016. Burkitt lymphoma in adolescents

and young adults: management challenges. *Adolescent Health, Medicine and Therapeutics,* Volume 8**,** 11-29.

DUTTON, A., O'NEIL, J. D., MILNER, A. E., REYNOLDS, G. M., STARCZYNSKI, J., CROCKER, J., YOUNG, L. S. & MURRAY, P. G. 2004. Expression of the cellular FLICE-inhibitory protein (c-FLIP) protects Hodgkin's lymphoma cells from autonomous Fas-mediated death. *Proceedings of the National Academy of Sciences,* 101**,** 6611-6616.

DUTTON, A., REYNOLDS, G. M., DAWSON, C. W., YOUNG, L. S. & MURRAY, P. G. 2005. Constitutive activation of phosphatidyl-inositide 3 kinase contributes to the survival of Hodgkin's lymphoma cells through a mechanism involving Akt kinase and mTOR. *J Pathol,* 205**,** 498-506.

DUTTON, A., WOODMAN, C. B., CHUKWUMA, M. B., LAST, J. I., WEI, W., VOCKERODT, M., BAUMFORTH, K. R., FLAVELL, J. R., ROWE, M., TAYLOR, A. M., YOUNG, L. S. & MURRAY, P. G. 2007. Bmi-1 is induced by the Epstein-Barr virus oncogene LMP1 and regulates the expression of viral target genes in Hodgkin lymphoma cells. *Blood,* 109**,** 2597-603.

EBERLE, F. C., MANI, H. & JAFFE, E. S. 2009. Histopathology of Hodgkin's lymphoma. *Cancer J,* 15**,** 129-37.

EDSON, C. M. & THORLEY-LAWSON, D. A. 1981. Epstein-Barr virus membrane antigens: characterization, distribution, and strain differences. *J Virol,* 39**,** 172-84.

ELIOPOULOS, A. G., BLAKE, S. M., FLOETTMANN, J. E., ROWE, M. & YOUNG, L. S. 1999. Epstein-Barr virus-encoded latent membrane protein 1 activates the JNK pathway through its extreme C terminus via a mechanism involving TRADD and TRAF2. *J Virol,* 73**,** 1023-35.

ELIOPOULOS, A. G., CAAMANO, J. H., FLAVELL, J., REYNOLDS, G. M., MURRAY, P. G., POYET, J. L. & YOUNG, L. S. 2003. Epstein-Barr virus-encoded latent infection membrane protein 1 regulates the processing of p100 NF-kappaB2 to p52 via an IKKgamma/NEMO-independent signalling pathway. *Oncogene,* 22**,** 7557-69.

ELIOPOULOS, A. G., STACK, M., DAWSON, C. W., KAYE, K. M., HODGKIN, L., SIHOTA, S., ROWE, M. & YOUNG, L. S. 1997. Epstein-Barr virus-encoded LMP1 and CD40 mediate IL-6 production in epithelial cells via an NF-kappaB pathway involving TNF receptor-associated factors. *Oncogene,* 14**,** 2899-916.

EMMERICH, F., THEURICH, S., HUMMEL, M., HAEFFKER, A., VRY, M. S., DOHNER, K., BOMMERT, K., STEIN, H. & DORKEN, B. 2003. Inactivating I kappa B epsilon mutations in Hodgkin/Reed-Sternberg cells. *J Pathol,* 201**,** 413-20.

EPSTEIN, A. 1999. On the discovery of Epstein-Barr virus: a memoir. *Epstein Barr Virus Report,* 6**,** 58-63.

EPSTEIN, M. 1979. CITATION CLASSIC-VIRUS-PARTICLES IN CULTURED LYMPHOBLASTS FROM BURKITTS-LYMPHOMA. *Current Contents/Life Sciences***,** 10-10.

EPSTEIN, M. A., ACHONG, B. G. & BARR, Y. M. 1964. VIRUS PARTICLES IN CULTURED LYMPHOBLASTS FROM BURKITT'S LYMPHOMA. *The Lancet,* 283**,** 702-703.

EPSTEIN, M. A. & BARR, Y. M. 1964. CULTIVATION IN VITRO OF HUMAN LYMPHOBLASTS FROM BURKITT'S MALIGNANT LYMPHOMA. *Lancet,* 1**,** 252-3.

EVANS, T. J., FARRELL, P. J. & SWAMINATHAN, S. 1996. Molecular genetic analysis of Epstein-Barr virus Cp promoter function. *J Virol,* 70**,** 1695-705.

FAHMI, H., COCHET, C., HMAMA, Z., OPOLON, P. & JOAB, I. 2000. Transforming growth factor beta 1 stimulates expression of the Epstein-Barr virus BZLF1

immediate-early gene product ZEBRA by an indirect mechanism which requires the MAPK kinase pathway. *J Virol,* 74**,** 5810-8.

FAHRAEUS, R., FU, H. L., ERNBERG, I., FINKE, J., ROWE, M., KLEIN, G., FALK, K., NILSSON, E., YADAV, M., BUSSON, P. & ET AL. 1988. Expression of Epstein-Barr virus-encoded proteins in nasopharyngeal carcinoma. *Int J Cancer,* 42**,** 329-38.

FALZETTI, D., CRESCENZI, B., MATTEUCI, C., FALINI, B., MARTELLI, M. F., VAN DEN BERGHE, H. & MECUCCI, C. 1999. Genomic instability and recurrent breakpoints are main cytogenetic findings in Hodgkin's disease. *Haematologica,* 84**,** 298-305.

FEEDERLE, R., KOST, M., BAUMANN, M., JANZ, A., DROUET, E., HAMMERSCHMIDT, W. & DELECLUSE, H. J. 2000. The Epstein-Barr virus lytic program is controlled by the co-operative functions of two transactivators. *EMBO J,* 19**,** 3080-9.

FELSHER, D. W. & BISHOP, J. M. 1999. Transient excess of MYC activity can elicit genomic instability and tumorigenesis. *Proc Natl Acad Sci U S A,* 96**,** 3940-4.

FENG, W. H., ISRAEL, B., RAAB-TRAUB, N., BUSSON, P. & KENNEY, S. C. 2002. Chemotherapy induces lytic EBV replication and confers ganciclovir susceptibility to EBV-positive epithelial cell tumors. *Cancer Res,* 62**,** 1920-6.

FERRY, J. A. 2006. Burkitt's Lymphoma: Clinicopathologic Features and Differential Diagnosis. *The Oncologist,* 11**,** 375-383.

FINGEROTH, J. D., WEIS, J. J., TEDDER, T. F., STROMINGER, J. L., BIRO, P. A. & FEARON, D. T. 1984. Epstein-Barr virus receptor of human B lymphocytes is the C3d receptor CR2. *Proc Natl Acad Sci U S A,* 81**,** 4510-4.

FIUMARA, P., SNELL, V., LI, Y., MUKHOPADHYAY, A., YOUNES, M., GILLENWATER, A. M., CABANILLAS, F., AGGARWAL, B. B. & YOUNES, A. 2001. Functional expression of receptor activator of nuclear factor κB in Hodgkin disease cell lines. *Blood,* 98**,** 2784-2790.

FLEISHER, G., HENLE, W., HENLE, G., LENNETTE, E. T. & BIGGAR, R. J. 1979. Primary infection with Epstein-Barr virus in infants in the United States: clinical and serologic observations. *J Infect Dis,* 139**,** 553-8.

FLEMINGTON, E. & SPECK, S. H. 1990. Identification of phorbol ester response elements in the promoter of Epstein-Barr virus putative lytic switch gene BZLF1. *J Virol,* 64**,** 1217-26.

FLEMINGTON, E. K., GOLDFELD, A. E. & SPECK, S. H. 1991. Efficient transcription of the Epstein-Barr virus immediate-early BZLF1 and BRLF1 genes requires protein synthesis. *J Virol,* 65**,** 7073-7.

FOGHSGAARD, L. & JAATTELA, M. 1997. The ability of BHRF1 to inhibit apoptosis is dependent on stimulus and cell type. *J Virol,* 71**,** 7509-17.

FORBES, S. A., BEARE, D., GUNASEKARAN, P., LEUNG, K., BINDAL, N., BOUTSELAKIS, H., DING, M., BAMFORD, S., COLE, C., WARD, S., KOK, C. Y., JIA, M., DE, T., TEAGUE, J. W., STRATTON, M. R., MCDERMOTT, U. & CAMPBELL, P. J. 2015. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res,* 43**,** D805-11.

FOX, C. P., HAIGH, T. A., TAYLOR, G. S., LONG, H. M., LEE, S. P., SHANNON-LOWE, C., O'CONNOR, S., BOLLARD, C. M., IQBAL, J., CHAN, W. C., RICKINSON, A. B., BELL, A. I. & ROWE, M. 2010. A novel latent membrane 2 transcript expressed in Epstein-Barr virus-positive NK- and T-cell lymphoproliferative disease encodes a target for cellular immunotherapy. *Blood,* 116**,** 3695-704.

FUKUDA, M. & LONGNECKER, R. 2004. Latent membrane protein 2A inhibits transforming growth factor-beta 1-induced apoptosis through the phosphatidylinositol 3-kinase/Akt pathway. *J Virol,* 78**,** 1697-705.

FUKUDA, M. & LONGNECKER, R. 2005. Epstein-Barr virus (EBV) latent membrane protein 2A regulates B-cell receptor-induced apoptosis and EBV reactivation through tyrosine phosphorylation. *J Virol,* 79**,** 8655-60.

GAHN, T. A. & SUGDEN, B. 1995. An EBNA-1-dependent enhancer acts from a distance of 10 kilobase pairs to increase expression of the Epstein-Barr virus LMP gene. *J Virol,* 69**,** 2633-6.

GALAKTIONOV, K., CHEN, X. & BEACH, D. 1996. Cdc25 cell-cycle phosphatase as a target of c-myc. *Nature,* 382**,** 511-7.

GANDHI, M. K., MOLL, G., SMITH, C., DUA, U., LAMBLEY, E., RAMUZ, O., GILL, D., MARLTON, P., SEYMOUR, J. F. & KHANNA, R. 2007. Galectin-1 mediated suppression of Epstein-Barr virus specific T-cell immunity in classic Hodgkin lymphoma. *Blood,* 110**,** 1326-1329.

GAUDE, E. & FREZZA, C. 2016. Tissue-specific and convergent metabolic transformation of cancer correlates with metastatic potential and patient survival. *Nature Communications,* 7**,** 13041.

GENE ONTOLOGY, C. 2015. Gene Ontology Consortium: going forward. *Nucleic Acids Res,* 43**,** D1049-56.

GESER, A., BRUBAKER, G. & DRAPER, C. C. 1989. EFFECT OF A MALARIA SUPPRESSION PROGRAM ON THE INCIDENCE OF AFRICAN BURKITT'S LYMPHOMA. *American Journal of Epidemiology,* 129**,** 740-752.

GESER, A., DE THE, G., LENOIR, G., DAY, N. E. & WILLIAMS, E. H. 1982. Final case reporting from the Ugandan prospective study of the relationship between EBV and Burkitt's lymphoma. *Int J Cancer,* 29**,** 397-400.

GEWURZ, B. E., TOWFIC, F., MAR, J. C., SHINNERS, N. P., TAKASAKI, K., ZHAO, B., CAHIR-MCFARLAND, E. D., QUACKENBUSH, J., XAVIER, R. J. & KIEFF, E. 2012. Genome-wide siRNA screen for mediators of NF-kappaB activation. *Proc Natl Acad Sci U S A,* 109**,** 2467-72.

GIRES, O., KOHLHUBER, F., KILGER, E., BAUMANN, M., KIESER, A., KAISER, C., ZEIDLER, R., SCHEFFER, B., UEFFING, M. & HAMMERSCHMIDT, W. 1999. Latent membrane protein 1 of Epstein-Barr virus interacts with JAK3 and activates STAT proteins. *EMBO J,* 18**,** 3064-73.

GLIMELIUS, I. & DIEPSTRA, A. 2017. Novel treatment concepts in Hodgkin lymphoma. *J Intern Med,* 281**,** 247-260.

GOORMACHTIGH, G., OUK, T. S., MOUGEL, A., TRANCHAND-BUNEL, D., MASY, E., LE CLORENNEC, C., FEUILLARD, J., BORNKAMM, G. W., AURIAULT, C., MANET, E., FAFEUR, V., ADRIAENSSENS, E. & COLL, J. 2006. Autoactivation of the Epstein-Barr virus oncogenic protein LMP1 during type II latency through opposite roles of the NF-kappaB and JNK signaling pathways. *J Virol,* 80**,** 7382-93.

GORUPPI, S., GUSTINCICH, S., BRANCOLINI, C., LEE, W. M. & SCHNEIDER, C. 1994. Dissection of c-myc domains involved in S phase induction of NIH3T3 fibroblasts. *Oncogene,* 9**,** 1537-44.

GRADOVILLE, L., GROGAN, E., TAYLOR, N. & MILLER, G. 1990. Differences in the extent of activation of Epstein-Barr virus replicative gene expression among four nonproducer cell lines stably transformed by oriP/BZLF1 plasmids. *Virology,* 178**,** 345-54.

GRANDORI, C., MAC, J., SIEBELT, F., AYER, D. E. & EISENMAN, R. N. 1996. Myc-Max heterodimers activate a DEAD box gene and interact with multiple E box-related sites in vivo. *EMBO J,* 15**,** 4344-57.

GRASSER, F. A., MURRAY, P. G., KREMMER, E., KLEIN, K., REMBERGER, K., FEIDEN, W., REYNOLDS, G., NIEDOBITEK, G., YOUNG, L. S. & MUELLER-LANTZSCH, N. 1994. Monoclonal antibodies directed against the Epstein-Barr virus-encoded nuclear antigen 1 (EBNA1): immunohistologic detection of EBNA1 in the malignant cells of Hodgkin's disease. *Blood,* 84**,** 3792-8.

GRATAMA, J. W., ZUTTER, M. M., MINAROVITS, J., OOSTERVEER, M. A., THOMAS, E. D., KLEIN, G. & ERNBERG, I. 1991. Expression of Epstein-Barr virus-encoded growth-transformation-associated proteins in lymphoproliferations of bone-marrow transplant recipients. *Int J Cancer,* 47**,** 188-92.

GREGORY, C. D., MURRAY, R. J., EDWARDS, C. F. & RICKINSON, A. B. 1988. Downregulation of cell adhesion molecules LFA-3 and ICAM-1 in Epstein-Barr virus-positive Burkitt's lymphoma underlies tumor cell escape from virus-specific T cell surveillance. *J Exp Med,* 167**,** 1811-24.

GRIFFIN, B. D., GRAM, A. M., MULDER, A., VAN LEEUWEN, D., CLAAS, F. H. J., WANG, F., RESSING, M. E. & WIERTZ, E. 2013. EBV BILF1 Evolved To Downregulate Cell Surface Display of a Wide Range of HLA Class I Molecules through Their Cytoplasmic Tail. *The Journal of Immunology,* 190**,** 1672-1684.

GRIFFITHS-JONES, S., GROCOCK, R. J., VAN DONGEN, S., BATEMAN, A. & ENRIGHT, A. J. 2006. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res,* 34**,** D140-4.

GUNAWARDANA, J., CHAN, F. C., TELENIUS, A., WOOLCOCK, B., KRIDEL, R., TAN, K. L., BEN-NERIAH, S., MOTTOK, A., LIM, R. S., BOYLE, M., ROGIC, S., RIMSZA, L. M., GUITER, C., LEROY, K., GAULARD, P., HAIOUN, C., MARRA, M. A., SAVAGE, K. J., CONNORS, J. M., SHAH, S. P., GASCOYNE, R. D. & STEIDL, C. 2014. Recurrent somatic mutations of PTPN1 in primary mediastinal B cell lymphoma and Hodgkin lymphoma. *Nat Genet,* 46**,** 329-35.

GUTENSOHN, N. & COLE, P. 1980. Epidemiology of Hodgkin's disease. *Semin Oncol,* 7**,** 92-102.

HAMMERSCHMIDT, W. & SUGDEN, B. 1988. Identification and characterization of oriLyt, a lytic origin of DNA replication of Epstein-Barr virus. *Cell,* 55**,** 427-33.

HAMMERSCHMIDT, W. & SUGDEN, B. 1989. Genetic analysis of immortalizing functions of Epstein-Barr virus in human B lymphocytes. *Nature,* 340**,** 393-7.

HANAHAN, D. & WEINBERG, ROBERT A. 2011. Hallmarks of Cancer: The Next Generation. *Cell,* 144**,** 646-674.

HANDE, K. R. 2008. Topoisomerase II inhibitors. *Update on Cancer Therapeutics,* 3**,** 13-26.

HARADA, S. & KIEFF, E. 1997. Epstein-Barr virus nuclear protein LP stimulates EBNA-2 acidic domain-mediated transcriptional activation. *J Virol,* 71**,** 6611-8.

HARDWICK, J. M., LIEBERMAN, P. M. & HAYWARD, S. D. 1988. A new Epstein-Barr virus transactivator, R, induces expression of a cytoplasmic early antigen. *J Virol,* 62**,** 2274-84.

HARTMANN, S., MARTIN-SUBERO, J. I., GESK, S., HUSKEN, J., GIEFING, M., NAGEL, I., RIEMKE, J., CHOTT, A., KLAPPER, W., PARRENS, M., MERLIO, J. P., KUPPERS, R., BRAUNINGER, A., SIEBERT, R. & HANSMANN, M. L. 2008. Detection of genomic imbalances in microdissected Hodgkin and Reed-Sternberg cells of classical Hodgkin's lymphoma by array-based comparative genomic hybridization. *Haematologica,* 93**,** 1318-26.

HASENCLEVER, D. & DIEHL, V. 1998. A prognostic score for advanced Hodgkin's disease. International Prognostic Factors Project on Advanced Hodgkin's Disease. *N Engl J Med,* 339**,** 1506-14.

HECHT, D. J., DAVIS, A. J., BROOKS, C. F. & JOHNSON, P. A. 2000. Molecular cloning and expression analysis of the complementary deoxyribonucleic acid for chicken inhibin/activin beta(B) subunit. *Biol Reprod,* 62**,** 1128-34.

HECHT, J. L. & ASTER, J. C. 2000. Molecular Biology of Burkitt's Lymphoma. *Journal of Clinical Oncology,* 18**,** 3707-3721.

HENDERSON, S., HUEN, D., ROWE, M., DAWSON, C., JOHNSON, G. & RICKINSON, A. 1993. Epstein-Barr virus-coded BHRF1 protein, a viral homologue of Bcl-2, protects human B cells from programmed cell death. *Proc Natl Acad Sci U S A,* 90**,** 8479-83.

HENDERSON, S., ROWE, M., GREGORY, C., CROOM-CARTER, D., WANG, F., LONGNECKER, R., KIEFF, E. & RICKINSON, A. 1991. Induction of bcl-2 expression by Epstein-Barr virus latent membrane protein 1 protects infected B cells from programmed cell death. *Cell,* 65**,** 1107-15.

HENLE, G., HENLE, W., CLIFFORD, P., DIEHL, V., KAFUKO, G. W., KIRYA, B. G., KLEIN, G., MORROW, R. H., MUNUBE, G. M., PIKE, P., TUKEI, P. M. & ZIEGLER, J. L. 1969. Antibodies to Epstein-Barr virus in Burkitt's lymphoma and control groups. *J Natl Cancer Inst,* 43**,** 1147-57.

HENLE, W., DIEHL, V., KOHN, G., ZUR HAUSEN, H. & HENLE, G. 1967. Herpes-type virus and chromosome marker in normal leukocytes after growth with irradiated Burkitt cells. *Science,* 157**,** 1064-5.

HERMEKING, H., WOLF, D. A., KOHLHUBER, F., DICKMANNS, A., BILLAUD, M., FANNING, E. & EICK, D. 1994. Role of c-myc in simian virus 40 large tumor antigen-induced DNA synthesis in quiescent 3T3-L1 mouse fibroblasts. *Proc Natl Acad Sci U S A,* 91**,** 10412-6.

HERROLD, R. E., MARCHINI, A., FRUEHLING, S. & LONGNECKER, R. 1996. Glycoprotein 110, the Epstein-Barr virus homolog of herpes simplex virus glycoprotein B, is essential for Epstein-Barr virus replication in vivo. *J Virol,* 70**,** 2049-54.

HERTLE, M. L., POPP, C., PETERMANN, S., MAIER, S., KREMMER, E., LANG, R., MAGES, J. & KEMPKES, B. 2009. Differential gene expression patterns of EBV infected EBNA-3A positive and negative human B lymphocytes. *PLoS Pathog,* 5**,** e1000506.

HIATT, J. B., PATWARDHAN, R. P., TURNER, E. H., LEE, C. & SHENDURE, J. 2010. Parallel, tag-directed assembly of locally derived short sequence reads. *Nat Methods,* 7**,** 119-22.

HICKABOTTOM, M., PARKER, G. A., FREEMONT, P., CROOK, T. & ALLDAY, M. J. 2002. Two nonconsensus sites in the Epstein-Barr virus oncoprotein EBNA3A cooperate to bind the co-repressor carboxyl-terminal-binding protein (CtBP). *J Biol Chem,* 277**,** 47197-204.

HIRIART, E., BARDOUILLET, L., MANET, E., GRUFFAT, H., PENIN, F., MONTSERRET, R., FARJOT, G. & SERGEANT, A. 2003. A region of the Epstein-Barr virus (EBV) mRNA export factor EB2 containing an arginine-rich motif mediates direct binding to RNA. *J Biol Chem,* 278**,** 37790-8.

HISLOP, A. D., RESSING, M. E., VAN LEEUWEN, D., PUDNEY, V. A., HORST, D., KOPPERS-LALIC, D., CROFT, N. P., NEEFJES, J. J., RICKINSON, A. B. & WIERTZ, E. J. 2007a. A CD8+ T cell immune evasion protein specific to Epstein-Barr virus and its close relatives in Old World primates. *J Exp Med,* 204**,** 1863-73.

HISLOP, A. D., RESSING, M. E., VAN LEEUWEN, D., PUDNEY, V. A., HORST, D., KOPPERS-LALIC, D., CROFT, N. P., NEEFJES, J. J., RICKINSON, A. B. & WIERTZ, E. J. H. J. 2007b. A CD8+T cell immune evasion protein specific to Epstein-Barr virus and its close relatives in Old World primates. *The Journal of Experimental Medicine,* 204**,** 1863-1873.

HJALGRIM, H. 2012. On the aetiology of Hodgkin lymphoma. *Dan Med J,* 59**,** B4485.

HJALGRIM, H., EKSTROM SMEDBY, K., ROSTGAARD, K., MOLIN, D., HAMILTON-DUTOIT, S., CHANG, E. T., RALFKIAER, E., SUNDSTROM, C., ADAMI, H. O., GLIMELIUS, B. & MELBYE, M. 2007. Infectious Mononucleosis, Childhood Social Environment, and Risk of Hodgkin Lymphoma. *Cancer Research,* 67**,** 2382-2388.

HOLLEY-GUTHRIE, E. A., SEAMAN, W. T., BHENDE, P., MERCHANT, J. L. & KENNEY, S. C. 2005. The Epstein-Barr virus protein BMRF1 activates gastrin transcription. *J Virol,* 79**,** 745-55.

HOLTICK, U., VOCKERODT, M., PINKERT, D., SCHOOF, N., STÜRZENHOFECKER, B., KUSSEBI, N., LAUBER, K., WESSELBORG, S., LÖFFLER, D. & HORN, F. 2005. STAT3 is essential for Hodgkin lymphoma cell proliferation and is a target of tyrphostin AG17 which confers sensitization for apoptosis. *Leukemia,* 19**,** 936-944.

HONG, M., MURAI, Y., KUTSUNA, T., TAKAHASHI, H., NOMOTO, K., CHENG, C. M., ISHIZAWA, S., ZHAO, Q. L., OGAWA, R., HARMON, B. V., TSUNEYAMA, K. & TAKANO, Y. 2006. Suppression of Epstein-Barr nuclear antigen 1 (EBNA1) by RNA interference inhibits proliferation of EBV-positive Burkitt's lymphoma cells. *J Cancer Res Clin Oncol,* 132**,** 1-8.

HORIE, R., WATANABE, T., MORISHITA, Y., ITO, K., ISHIDA, T., KANEGAE, Y., SAITO, I., HIGASHIHARA, M., MORI, S., KADIN, M. E. & WATANABE, T. 2002. Ligand-independent signaling by overexpressed CD30 drives NF-κB activation in Hodgkin–Reed-Sternberg cells. *Oncogene,* 21**,** 2493-2503.

HOWE, J. G. & STEITZ, J. A. 1986. Localization of Epstein-Barr virus-encoded small RNAs by in situ hybridization. *Proc Natl Acad Sci U S A,* 83**,** 9006-10.

HSU, D. H., DE WAAL MALEFYT, R., FIORENTINO, D. F., DANG, M. N., VIEIRA, P., DE VRIES, J., SPITS, H., MOSMANN, T. R. & MOORE, K. W. 1990. Expression of interleukin-10 activity by Epstein-Barr virus protein BCRF1. *Science,* 250**,** 830-2.

HU, C., WEI, W., CHEN, X., WOODMAN, C. B., YAO, Y., NICHOLLS, J. M., JOAB, I., SIHOTA, S. K., SHAO, J.-Y., DERKAOUI, K. D., AMARI, A., MALONEY, S. L., BELL, A. I., MURRAY, P. G., DAWSON, C. W., YOUNG, L. S. & ARRAND, J. R. 2012. A Global View of the Oncogenic Landscape in Nasopharyngeal Carcinoma: An Integrated Analysis at the Genetic and Expression Levels. *PLoS ONE,* 7**,** e41055.

HU, J. 2005. The Expression of Functional Chemokine Receptor CXCR4 Is Associated with the Metastatic Potential of Human Nasopharyngeal Carcinoma. *Clinical Cancer Research,* 11**,** 4658-4665.

HU, J., LOCASALE, J. W., BIELAS, J. H., O'SULLIVAN, J., SHEAHAN, K., CANTLEY, L. C., HEIDEN, M. G. V. & VITKUP, D. 2013. Heterogeneity of tumor-induced gene expression changes in the human metabolic network. *Nature Biotechnology,* 31**,** 522-529.

HUANG DA, W., SHERMAN, B. T. & LEMPICKI, R. A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc,* 4**,** 44-57.

HUEN, D. S., HENDERSON, S. A., CROOM-CARTER, D. & ROWE, M. 1995. The Epstein-Barr virus latent membrane protein-1 (LMP1) mediates activation of NF-kappa B and cell surface phenotype via two effector regions in its carboxy-terminal cytoplasmic domain. *Oncogene,* 10**,** 549-60.

HUMME, S., REISBACH, G., FEEDERLE, R., DELECLUSE, H. J., BOUSSET, K., HAMMERSCHMIDT, W. & SCHEPERS, A. 2003. The EBV nuclear antigen 1 (EBNA1) enhances B cell immortalization several thousandfold. *Proceedings of the National Academy of Sciences,* 100**,** 10989-10994.

IKEDA, J., MAMAT, S., TIAN, T., WANG, Y., RAHADIANI, N., AOZASA, K. & MORII, E. 2010. Tumorigenic potential of mononucleated small cells of Hodgkin lymphoma cell lines. *Am J Pathol,* 177**,** 3081-8.

IZUMI, K. M., CAHIR MCFARLAND, E. D., TING, A. T., RILEY, E. A., SEED, B. & KIEFF, E. D. 1999. The Epstein-Barr virus oncoprotein latent membrane protein 1 engages the tumor necrosis factor receptor-associated proteins TRADD and receptor-interacting protein (RIP) but does not induce apoptosis or require RIP for NF-kappaB activation. *Mol Cell Biol,* 19**,** 5759-67.

IZUMI, K. M. & KIEFF, E. D. 1997. The Epstein-Barr virus oncogene product latent membrane protein 1 engages the tumor necrosis factor receptor-associated death domain protein to mediate B lymphocyte growth transformation and activate NF-kappaB. *Proc Natl Acad Sci U S A,* 94**,** 12592-7.

ILHE, J. N. 2001. The Stat family in cytokine signaling. Curr Opin Cell Biol. Apr; 13(2): 211–217.

JACOBS, J. J., SCHEIJEN, B., VONCKEN, J. W., KIEBOOM, K., BERNS, A. & VAN LOHUIZEN, M. 1999. Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. *Genes Dev,* 13**,** 2678-90.

JACOBSON, C. & LACASCE, A. 2014. How I treat Burkitt lymphoma in adults. *Blood,* 124**,** 2913-2920.

JAIN, M., NILSSON, R., SHARMA, S., MADHUSUDHAN, N., KITAMI, T., SOUZA, A. L., KAFRI, R., KIRSCHNER, M. W., CLISH, C. B. & MOOTHA, V. K. 2012. Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science,* 336**,** 1040-4.

JAYACHANDRA, S., LOW, K. G., THLICK, A.-E., YU, J., LING, P. D., CHANG, Y. & MOORE, P. S. 1999. Three unrelated viral transforming proteins (vIRF, EBNA2, and E1A) induce the MYC oncogene through the interferon-responsive PRF element by using different transcription coadaptors. *Proceedings of the National Academy of Sciences,* 96**,** 11566-11571.

JIN, X. W. & SPECK, S. H. 1992. Identification of critical cis elements involved in mediating Epstein-Barr virus nuclear antigen 2-dependent activity of an enhancer located upstream of the viral BamHI C promoter. *J Virol,* 66**,** 2846-52.

JOHANNSEN, E., KOH, E., MOSIALOS, G., TONG, X., KIEFF, E. & GROSSMAN, S. R. 1995. Epstein-Barr virus nuclear protein 2 transactivation of the latent membrane protein 1 promoter is mediated by J kappa and PU.1. *J Virol,* 69**,** 253-62.

JOHANNSEN, E., LUFTIG, M., CHASE, M. R., WEICKSEL, S., CAHIR-MCFARLAND, E., ILLANES, D., SARRACINO, D. & KIEFF, E. 2004. Proteins of purified Epstein-Barr virus. *Proc Natl Acad Sci U S A,* 101**,** 16286-91.

JOHANNSEN, E., MILLER, C. L., GROSSMAN, S. R. & KIEFF, E. 1996. EBNA-2 and EBNA-3C extensively and mutually exclusively associate with RBPJkappa in Epstein-Barr virus-transformed B lymphocytes. *J Virol,* 70**,** 4179-83.

JOHNSON, S. A., PLEIMAN, C. M., PAO, L., SCHNERINGER, J., HIPPEN, K. & CAMBIER, J. C. 1995. Phosphorylated immunoreceptor signaling motifs (ITAMs) exhibit unique abilities to bind and activate Lyn and Syk tyrosine kinases. *J Immunol,* 155**,** 4596-603.

JONES, R. M., BRANDA, J., JOHNSTON, K. A., POLYMENIS, M., GADD, M., RUSTGI, A., CALLANAN, L. & SCHMIDT, E. V. 1996. An essential E box in the promoter of

the gene encoding the mRNA cap-binding protein (eukaryotic initiation factor 4E) is a target for activation by c-myc. *Mol Cell Biol,* 16**,** 4754-64.

JOOS, S., GRANZOW, M., HOLTGREVE-GREZ, H., SIEBERT, R., HARDER, L., MARTIN-SUBERO, J. I., WOLF, J., ADAMOWICZ, M., BARTH, T. F., LICHTER, P. & JAUCH, A. 2003. Hodgkin's lymphoma cell lines are characterized by frequent aberrations on chromosomes 2p and 9p including REL and JAK2. *Int J Cancer,* 103**,** 489-95.

JOOS, S., KUPPER, M., OHL, S., VON BONIN, F., MECHTERSHEIMER, G., BENTZ, M., MARYNEN, P., MOLLER, P., PFREUNDSCHUH, M., TRUMPER, L. & LICHTER, P. 2000. Genomic imbalances including amplification of the tyrosine kinase gene JAK2 in CD30+ Hodgkin cells. *Cancer Res,* 60**,** 549-52.

JUNGNICKEL, B., STARATSCHEK-JOX, A., BRAUNINGER, A., SPIEKER, T., WOLF, J., DIEHL, V., HANSMANN, M. L., RAJEWSKY, K. & KUPPERS, R. 2000. Clonal deleterious mutations in the IkappaBalpha gene in the malignant cells in Hodgkin's lymphoma. *J Exp Med,* 191**,** 395-402.

JUSZCZYNSKI, P., OUYANG, J., MONTI, S., RODIG, S. J., TAKEYAMA, K., ABRAMSON, J., CHEN, W., KUTOK, J. L., RABINOVICH, G. A. & SHIPP, M. A. 2007. The AP1-dependent secretion of galectin-1 by Reed Sternberg cells fosters immune privilege in classical Hodgkin lymphoma. *Proceedings of the National Academy of Sciences,* 104**,** 13134-13139.

KAISER, C., LAUX, G., EICK, D., JOCHNER, N., BORNKAMM, G. W. & KEMPKES, B. 1999. The proto-oncogene c-myc is a direct target gene of Epstein-Barr virus nuclear antigen 2. *J Virol,* 73**,** 4481-4.

KALLA, M. & HAMMERSCHMIDT, W. 2012. Human B cells on their route to latent infection–early but transient expression of lytic genes of Epstein-Barr virus. *European journal of cell biology,* 91**,** 65-69.

KANG, M. S., LU, H., YASUI, T., SHARPE, A., WARREN, H., CAHIR-MCFARLAND, E., BRONSON, R., HUNG, S. C. & KIEFF, E. 2005. Epstein-Barr virus nuclear antigen 1 does not induce lymphoma in transgenic FVB mice. *Proc Natl Acad Sci U S A,* 102**,** 820-5.

KAPP, U., WOLF, J., HUMMEL, M., PAWLITA, M., VON KALLE, C., DALLENBACH, F., SCHWONZEN, M., KRUEGER, G. R., MULLER-LANTZSCH, N., FONATSCH, C. & ET AL. 1993. Hodgkin's lymphoma-derived tissue serially transplanted into severe combined immunodeficient mice. *Blood,* 82**,** 1247-56.

KAPP, U., YEH, W.-C., PATTERSON, B., ELIA, A. J., KÄGI, D., HO, A., HESSEL, A., TIPSWORD, M., WILLIAMS, A., MIRTSOS, C., ITIE, A., MOYLE, M. & MAK, T. W. 1999. Interleukin 13 Is Secreted by and Stimulates the Growth of Hodgkin and Reed-Sternberg Cells. *The Journal of Experimental Medicine,* 189**,** 1939-1946.

KARN, J., WATSON, J. V., LOWE, A. D., GREEN, S. M. & VEDECKIS, W. 1989. Regulation of cell cycle duration by c-myc levels. *Oncogene,* 4**,** 773-87.

KASHKAR, H., HAEFS, C., SHIN, H., HAMILTON-DUTOIT, S. J., SALVESEN, G. S., KRÖNKE, M. & JÜRGENSMEIER, J. M. 2003. XIAP-mediated Caspase Inhibition in Hodgkin's Lymphoma–derived B Cells. *The Journal of Experimental Medicine,* 198**,** 341-347.

KATO, M., SANADA, M., KATO, I., SATO, Y., TAKITA, J., TAKEUCHI, K., NIWA, A., CHEN, Y., NAKAZAKI, K., NOMOTO, J., ASAKURA, Y., MUTO, S., TAMURA, A., IIO, M., AKATSUKA, Y., HAYASHI, Y., MORI, H., IGARASHI, T., KUROKAWA, M., CHIBA, S., MORI, S., ISHIKAWA, Y., OKAMOTO, K., TOBINAI, K., NAKAGAMA, H., NAKAHATA, T., YOSHINO, T., KOBAYASHI,

Y. & OGAWA, S. 2009. Frequent inactivation of A20 in B-cell lymphomas. *Nature,* 459**,** 712-716.

KAYE, K. M., IZUMI, K. M. & KIEFF, E. 1993. Epstein-Barr virus latent membrane protein 1 is essential for B-lymphocyte growth transformation. *Proceedings of the National Academy of Sciences,* 90**,** 9150-9154.

KELLY, G., BELL, A. & RICKINSON, A. 2002a. Epstein-Barr virus-associated Burkitt lymphomagenesis selects for downregulation of the nuclear antigen EBNA2. *Nat Med,* 8**,** 1098-104.

KELLY, G., BELL, A. & RICKINSON, A. 2002b. Epstein–Barr virus–associated Burkitt lymphomagenesis selects for downregulation of the nuclear antigen EBNA2. *Nature Medicine,* 8**,** 1098-1104.

KELLY, G. L., LONG, H. M., STYLIANOU, J., THOMAS, W. A., LEESE, A., BELL, A. I., BORNKAMM, G. W., MAUTNER, J., RICKINSON, A. B. & ROWE, M. 2009a. An Epstein-Barr Virus Anti-Apoptotic Protein Constitutively Expressed in Transformed Cells and Implicated in Burkitt Lymphomagenesis: The Wp/BHRF1 Link. *PLoS Pathogens,* 5**,** e1000341.

KELLY, G. L., LONG, H. M., STYLIANOU, J., THOMAS, W. A., LEESE, A., BELL, A. I., BORNKAMM, G. W., MAUTNER, J., RICKINSON, A. B. & ROWE, M. 2009b. An Epstein-Barr virus anti-apoptotic protein constitutively expressed in transformed cells and implicated in burkitt lymphomagenesis: the Wp/BHRF1 link. *PLoS Pathog,* 5**,** e1000341.

KELLY, G. L., MILNER, A. E., TIERNEY, R. J., CROOM-CARTER, D. S., ALTMANN, M., HAMMERSCHMIDT, W., BELL, A. I. & RICKINSON, A. B. 2005. Epstein-Barr virus nuclear antigen 2 (EBNA2) gene deletion is consistently linked with EBNA3A, -3B, and -3C expression in Burkitt's lymphoma cells and with increased resistance to apoptosis. *J Virol,* 79**,** 10709-17.

KELLY, G. L. & RICKINSON, A. B. 2007. Burkitt lymphoma: revisiting the pathogenesis of a virus-associated malignancy. Hematology Am Soc Hematol Educ Program.277–284

KENNEDY, G., KOMANO, J. & SUGDEN, B. 2003. Epstein-Barr virus provides a survival factor to Burkitt's lymphomas. *Proc Natl Acad Sci U S A,* 100**,** 14269-74.

KERR, B. M., LEAR, A. L., ROWE, M., CROOM-CARTER, D., YOUNG, L. S., ROOKES, S. M., GALLIMORE, P. H. & RICKINSON, A. B. 1992. Three transcriptionally distinct forms of Epstein-Barr virus latency in somatic cell hybrids: cell phenotype dependence of virus promoter usage. *Virology,* 187**,** 189-201.

KHAN, G. & HASHIM, M. J. 2014. Global burden of deaths from Epstein-Barr virus attributable malignancies 1990-2010. *Infectious agents and cancer,* 9**,** 38.

KHANIM, F., DAWSON, C., MESEDA, C. A., DAWSON, J., MACKETT, M. & YOUNG, L. S. 1997. BHRF1, a viral homologue of the Bcl-2 oncogene, is conserved at both the sequence and functional level in different Epstein-Barr virus isolates. *J Gen Virol,* 78 (Pt 11)**,** 2987-99.

KIEFF, E. 1996. Epstein-Barr virus and its replication. In: Fields Virology. Philadelphia, PA: Lippincott–Raven, pp. 2343–2396

KIEFF, E. & RICKINSON, A. 2001. Fields virology. *Knipe DM, Howley PM, editors,* 2001**,** 2511-75.

KIESER, A. 2008. Pursuing different 'TRADDes': TRADD signaling induced by TNF-receptor 1 and the Epstein-Barr virus oncoprotein LMP1. *Biol Chem,* 389**,** 1261-71.

KIESER, A., KILGER, E., GIRES, O., UEFFING, M., KOLCH, W. & HAMMERSCHMIDT, W. 1997. Epstein-Barr virus latent membrane protein-1 triggers AP-1 activity via the c-Jun N-terminal kinase cascade. *EMBO J,* 16**,** 6478-85.

KILGER, E., KIESER, A., BAUMANN, M. & HAMMERSCHMIDT, W. 1998. Epstein-Barr virus-mediated B-cell proliferation is dependent upon latent membrane protein 1, which simulates an activated CD40 receptor. *EMBO J,* 17**,** 1700-9.

KIM, J. W., GAO, P., LIU, Y. C., SEMENZA, G. L. & DANG, C. V. 2007. Hypoxia-inducible factor 1 and dysregulated c-Myc cooperatively induce vascular endothelial growth factor and metabolic switches hexokinase 2 and pyruvate dehydrogenase kinase 1. *Mol Cell Biol,* 27**,** 7381-93.

KIM, J. W., ZELLER, K. I., WANG, Y., JEGGA, A. G., ARONOW, B. J., O'DONNELL, K. A. & DANG, C. V. 2004. Evaluation of myc E-box phylogenetic footprints in glycolytic genes by chromatin immunoprecipitation assays. *Mol Cell Biol,* 24**,** 5923-36.

KING, W., DAMBAUGH, T., HELLER, M., DOWLING, J. & KIEFF, E. 1982. Epstein-Barr virus DNA XII. A variable region of the Epstein-Barr virus genome is included in the P3HR-1 deletion. *J Virol,* 43**,** 979-86.

KIS, L. L., SALAMON, D., PERSSON, E. K., NAGY, N., SCHEEREN, F. A., SPITS, H., KLEIN, G. & KLEIN, E. 2010. IL-21 imposes a type II EBV gene expression on type III and type IB cells by the repression of C-and activation of LMP-1-promoter. *Proceedings of the National Academy of Sciences,* 107**,** 872-877.

KNIPE, D.M., HOWLEY, P.M., GRIIFIN, D.E.,LAMB, R.A., MARTIN, M.A., ROIZMAN, B., STRUS, S.E. 2007. Fields virology, 5th ed Lippincott Williams & Wilkins, Philadelphia, PA

KNUTSON, J. C. 1990. The level of c-fgr RNA is increased by EBNA-2, an Epstein-Barr virus gene required for B-cell immortalization. *J Virol,* 64**,** 2530-6.

KOBOLDT, D. C., LARSON, D. E. & WILSON, R. K. 2013. Using VarScan 2 for Germline Variant Calling and Somatic Mutation Detection. *Curr Protoc Bioinformatics,* 44**,** 15 4 1-17.

KOVALCHUK, A. L., QI, C.-F., TORREY, T. A., TADDESSE-HEATH, L., FEIGENBAUM, L., PARK, S. S., GERBITZ, A., KLOBECK, G., HOERTNAGEL, K., POLACK, A., BORNKAMM, G. W., JANZ, S. & MORSE, H. C. 2000. Burkitt Lymphoma in the Mouse. *The Journal of Experimental Medicine,* 192**,** 1183-1190.

KOZOMARA, A. & GRIFFITHS-JONES, S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res,* 42**,** D68-73.

KREM, M. M., LUO, P., ING, B. I. & HORWITZ, M. S. 2012. The kelch protein KLHDC8B guards against mitotic errors, centrosomal amplification, and chromosomal instability. *J Biol Chem,* 287**,** 39083-93.

KREUZ, S., SIEGMUND, D., SCHEURICH, P. & WAJANT, H. 2001. NF- B Inducers Upregulate cFLIP, a Cycloheximide-Sensitive Inhibitor of Death Receptor Signaling. *Molecular and Cellular Biology,* 21**,** 3964-3973.

KROENLEIN, H., SCHWARTZ, S., REINHARDT, R., RIEDER, H., MOLKENTIN, M., GOKBUGET, N., HOELZER, D., THIEL, E. & BURMEISTER, T. 2012. Molecular analysis of the t(2;8)/MYC-IGK translocation in high-grade lymphoma/leukemia by long-distance inverse PCR. *Genes Chromosomes Cancer,* 51**,** 290-9.

KRØIGÅRD, A. B., THOMASSEN, M., LÆNKHOLM, A.-V., KRUSE, T. A. & LARSEN, M. J. 2016. Evaluation of Nine Somatic Variant Callers for Detection of Somatic Mutations in Exome and Targeted Deep Sequencing Data. *PLOS ONE,* 11**,** e0151664.

KRYSIAK, K., GOMEZ, F., WHITE, B. S., MATLOCK, M., MILLER, C. A., TRANI, L., FRONICK, C. C., FULTON, R. S., KREISEL, F., CASHEN, A. F., CARSON, K. R., BERRIEN-ELLIOTT, M. M., BARTLETT, N. L., GRIFFITH, M., GRIFFITH, O. L. & FEHNIGER, T. A. 2017. Recurrent somatic mutations affecting B-cell receptor signaling pathway genes in follicular lymphoma. *Blood,* 129**,** 473-483.

KUBE, D. 2001. STAT3 is constitutively activated in Hodgkin cell lines. *Blood,* 98**,** 762-770.

KUBE, D., VOCKERODT, M., WEBER, O., HELL, K., WOLF, J., HAIER, B., GRASSER, F. A., MULLER-LANTZSCH, N., KIEFF, E., DIEHL, V. & TESCH, H. 1999. Expression of epstein-barr virus nuclear antigen 1 is associated with enhanced expression of CD25 in the Hodgkin cell line L428. *J Virol,* 73**,** 1630-6.

KUDOH, A., FUJITA, M., KIYONO, T., KUZUSHIMA, K., SUGAYA, Y., IZUTA, S., NISHIYAMA, Y. & TSURUMI, T. 2003. Reactivation of lytic replication from B cells latently infected with Epstein-Barr virus occurs with high S-phase cyclin-dependent kinase activity while inhibiting cellular DNA replication. *J Virol,* 77**,** 851-61.

KUNKEL, E. J. & BUTCHER, E. C. 2003. Plasma-cell homing. *Nature Reviews Immunology,* 3**,** 822-829.

KUPPERS, R. 2003. Somatic hypermutation and B cell receptor selection in normal and transformed human B cells. *Ann N Y Acad Sci,* 987**,** 173-9.

KUPPERS, R. 2009. Molecular biology of Hodgkin lymphoma. *Hematology Am Soc Hematol Educ Program***,** 491-6.

KUPPERS, R., ENGERT, A. & HANSMANN, M. L. 2012. Hodgkin lymphoma. *J Clin Invest,* 122**,** 3439-47.

KUPPERS, R., SCHMITZ, R., DISTLER, V., RENNE, C., BRAUNINGER, A. & HANSMANN, M. L. 2005. Pathogenesis of Hodgkin's lymphoma. *Eur J Haematol Suppl***,** 26-33.

KUPPERS, R., SCHWERING, I., BRAUNINGER, A., RAJEWSKY, K. & HANSMANN, M. L. 2002. Biology of Hodgkin's lymphoma. *Ann Oncol,* 13 Suppl 1**,** 11-8.

LAHERTY, C. D., HU, H. M., OPIPARI, A. W., WANG, F. & DIXIT, V. M. 1992. The Epstein-Barr virus LMP1 gene product induces A20 zinc finger protein expression by activating nuclear factor kappa B. *J Biol Chem,* 267**,** 24157-60.

LAICHALK, L. L. & THORLEY-LAWSON, D. A. 2005. Terminal differentiation into plasma cells initiates the replicative cycle of Epstein-Barr virus in vivo. *J Virol,* 79**,** 1296-307.

LAKE, A., SHIELD, L. A., CORDANO, P., CHUI, D. T., OSBORNE, J., CRAE, S., WILSON, K. S., TOSI, S., KNIGHT, S. J., GESK, S., SIEBERT, R., HAY, R. T. & JARRETT, R. F. 2009. Mutations of NFKBIA, encoding IkappaB alpha, are a recurrent finding in classical Hodgkin lymphoma but are not a unifying feature of non-EBV-associated cases. *Int J Cancer,* 125**,** 1334-42.

LAM, N. & SUGDEN, B. 2003. CD40 and its viral mimic, LMP1: similar means to different ends. *Cell Signal,* 15**,** 9-16.

LAUX, G., DUGRILLON, F., ECKERT, C., ADAM, B., ZIMBER-STROBL, U. & BORNKAMM, G. W. 1994. Identification and characterization of an Epstein-Barr virus nuclear antigen 2-responsive cis element in the bidirectional promoter region of latent membrane protein and terminal protein 2 genes. *J Virol,* 68**,** 6947-58.

LAUX, G., ECONOMOU, A. & FARRELL, P. J. 1989. The terminal protein gene 2 of Epstein-Barr virus is transcribed from a bidirectional latent promoter region. *Journal of general virology,* 70**,** 3079-3084.

LAUX, G., PERRICAUDET, M. & FARRELL, P. J. 1988. A spliced Epstein-Barr virus gene expressed in immortalized lymphocytes is created by circularization of the linear viral genome. *The EMBO journal,* 7**,** 769.

LEE, M. A. & YATES, J. L. 1992. BHRF1 of Epstein-Barr virus, which is homologous to human proto-oncogene bcl2, is not essential for transformation of B cells or for virus replication in vitro. *J Virol,* 66**,** 1899-906.

LEE, S., SAKAKIBARA, S., MARUO, S., ZHAO, B., CALDERWOOD, M. A., HOLTHAUS, A. M., LAI, C. Y., TAKADA, K., KIEFF, E. & JOHANNSEN, E. 2009. Epstein-Barr Virus Nuclear Protein 3C Domains Necessary for Lymphoblastoid Cell Growth: Interaction with RBP-J Regulates TCL1. *Journal of Virology,* 83**,** 12368-12377.

LEE, S. P., CONSTANDINOU, C. M., THOMAS, W. A., CROOM-CARTER, D., BLAKE, N. W., MURRAY, P. G., CROCKER, J. & RICKINSON, A. B. 1998. Antigen presenting phenotype of Hodgkin Reed-Sternberg cells: analysis of the HLA class I processing pathway and the effects of interleukin-10 on epstein-barr virus-specific cytotoxic T-cell recognition. *Blood,* 92**,** 1020-30.

LEONARD, P. G., BEZAR, I. F., SIDOTE, D. J. & STOCK, A. M. 2012. Identification of a hydrophobic cleft in the LytTR domain of AgrA as a locus for small molecule interactions that inhibit DNA binding. *Biochemistry,* 51**,** 10035-43.

LEONARD, S., WEI, W., ANDERTON, J., VOCKERODT, M., ROWE, M., MURRAY, P. G. & WOODMAN, C. B. 2011. Epigenetic and transcriptional changes which follow Epstein-Barr virus infection of germinal center B cells and their relevance to the pathogenesis of Hodgkin's lymphoma. *J Virol,* 85**,** 9568-77.

LEVY, D. E. & DARNELL, J. E. 2002. Signalling: STATs: transcriptional control and biological impact. *Nature Reviews Molecular Cell Biology,* 3**,** 651-662.

LI, F., WANG, Y., ZELLER, K. I., POTTER, J. J., WONSEY, D. R., O'DONNELL, K. A., KIM, J. W., YUSTEIN, J. T., LEE, L. A. & DANG, C. V. 2005. Myc Stimulates Nuclearly Encoded Mitochondrial Genes and Mitochondrial Biogenesis. *Molecular and Cellular Biology,* 25**,** 6225-6234.

LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics,* 25**,** 1754-60.

LI, H. P. & CHANG, Y. S. 2003. Epstein-Barr virus latent membrane protein 1: structure and functions. *J Biomed Sci,* 10**,** 490-504.

LI, L. H., NERLOV, C., PRENDERGAST, G., MACGREGOR, D. & ZIFF, E. B. 1994. c-Myc represses transcription in vivo by a novel mechanism dependent on the initiator element and Myc box II. *EMBO J,* 13**,** 4070-9.

LIAO, Y., SMYTH, G. K. & SHI, W. 2013. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res,* 41**,** e108.

LIBERZON, A., BIRGER, C., THORVALDSDÓTTIR, H., GHANDI, M., MESIROV, JILL P. & TAMAYO, P. 2015. The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Systems,* 1**,** 417-425.

LISTER, T. A., CROWTHER, D., SUTCLIFFE, S. B., GLATSTEIN, E., CANELLOS, G. P., YOUNG, R. C., ROSENBERG, S. A., COLTMAN, C. A. & TUBIANA, M. 1989. Report of a committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds meeting. *J Clin Oncol,* 7**,** 1630-6.

LIU, P. & SPECK, S. H. 2003. Synergistic autoactivation of the Epstein-Barr virus immediate-early BRLF1 promoter by Rta and Zta. *Virology,* 310**,** 199-206.

LIU, Y., ABDUL RAZAK, F. R., TERPSTRA, M., CHAN, F. C., SABER, A., NIJLAND, M., VAN IMHOFF, G., VISSER, L., GASCOYNE, R., STEIDL, C., KLUIVER, J., DIEPSTRA, A., KOK, K. & VAN DEN BERG, A. 2014. The mutational landscape of Hodgkin lymphoma cell lines determined by whole-exome sequencing. *Leukemia,* 28**,** 2248-2251.

LO, A. K., TO, K. F., LO, K. W., LUNG, R. W., HUI, J. W., LIAO, G. & HAYWARD, S. D. 2007. Modulation of LMP1 protein expression by EBV-encoded microRNAs. *Proc Natl Acad Sci U S A,* 104**,** 16164-9.

LOHR, J. G., STOJANOV, P., LAWRENCE, M. S., AUCLAIR, D., CHAPUY, B., SOUGNEZ, C., CRUZ-GORDILLO, P., KNOECHEL, B., ASMANN, Y. W., SLAGER, S. L., NOVAK, A. J., DOGAN, A., ANSELL, S. M., LINK, B. K., ZOU, L., GOULD, J., SAKSENA, G., STRANSKY, N., RANGEL-ESCARENO, C., FERNANDEZ-LOPEZ, J. C., HIDALGO-MIRANDA, A., MELENDEZ-ZAJGLA, J., HERNANDEZ-LEMUS, E., SCHWARZ-CRUZ Y CELIS, A., IMAZ-ROSSHANDLER, I., OJESINA, A. I., JUNG, J., PEDAMALLU, C. S., LANDER, E. S., HABERMANN, T. M., CERHAN, J. R., SHIPP, M. A., GETZ, G. & GOLUB, T. R. 2012. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A,* 109**,** 3879-84.

LONGNECKER, M. P., HARBAK, K., KISSLING, G. E., HOPPIN, J. A., EGGESBO, M., JUSKO, T. A., EIDE, J. & KOCH, H. M. 2013. The concentration of bisphenol A in urine is affected by specimen collection, a preservative, and handling. *Environ Res,* 126**,** 211-4.

LONGNECKER, R. & KIEFF, E. 1990. A second Epstein-Barr virus membrane protein (LMP2) is expressed in latent infection and colocalizes with LMP1. *J Virol,* 64**,** 2319-26.

LONGNECKER, R. & KIEFF, E. 2013. Epstein-Barr virus, p 1898–1959. Fields virology, 6th ed, vol 2 Lippincott Williams & Wilkins, Philadelphia, PA.

LONGNECKER, R., MILLER, C. L., MIAO, X. Q., MARCHINI, A. & KIEFF, E. 1992. The only domain which distinguishes Epstein-Barr virus latent membrane protein 2A (LMP2A) from LMP2B is dispensable for lymphocyte infection and growth transformation in vitro; LMP2A is therefore nonessential. *J Virol,* 66**,** 6461-9.

LONGNECKER, R., MILLER, C. L., MIAO, X. Q., TOMKINSON, B. & KIEFF, E. 1993a. The last seven transmembrane and carboxy-terminal cytoplasmic domains of Epstein-Barr virus latent membrane protein 2 (LMP2) are dispensable for lymphocyte infection and growth transformation in vitro. *J Virol,* 67**,** 2006-13.

LONGNECKER, R., MILLER, C. L., TOMKINSON, B., MIAO, X. Q. & KIEFF, E. 1993b. Deletion of DNA encoding the first five transmembrane domains of Epstein-Barr virus latent membrane proteins 2A and 2B. *J Virol,* 67**,** 5068-74.

LOVE, C., SUN, Z., JIMA, D., LI, G., ZHANG, J., MILES, R., RICHARDS, K. L., DUNPHY, C. H., CHOI, W. W., SRIVASTAVA, G., LUGAR, P. L., RIZZIERI, D. A., LAGOO, A. S., BERNAL-MIZRACHI, L., MANN, K. P., FLOWERS, C. R., NARESH, K. N., EVENS, A. M., CHADBURN, A., GORDON, L. I., CZADER, M. B., GILL, J. I., HSI, E. D., GREENOUGH, A., MOFFITT, A. B., MCKINNEY, M., BANERJEE, A., GRUBOR, V., LEVY, S., DUNSON, D. B. & DAVE, S. S. 2012. The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet,* 44**,** 1321-5.

LU, C. C., JENG, Y. Y., TSAI, C. H., LIU, M. Y., YEH, S. W., HSU, T. Y. & CHEN, M. R. 2006. Genome-wide transcription program and expression of the Rta responsive gene of Epstein-Barr virus. *Virology,* 345**,** 358-72.

LU, F., CHEN, H.-S., KOSSENKOV, A. V., DEWISPELEARE, K., WON, K.-J. & LIEBERMAN, P. M. 2016. EBNA2 Drives Formation of New Chromosome Binding Sites and Target Genes for B-Cell Master Regulatory Transcription Factors RBP-jκ and EBF1. *PLOS Pathogens,* 12**,** e1005339.

LUKES, R. J. & BUTLER, J. J. 1966. The pathology and nomenclature of Hodgkin's disease. *Cancer Res,* 26**,** 1063-83.

LUPTON, S. & LEVINE, A. J. 1985. Mapping genetic elements of Epstein-Barr virus that facilitate extrachromosomal persistence of Epstein-Barr virus-derived plasmids in human cells. *Molecular and cellular biology,* 5**,** 2533-2542.

LYNCH, D. T., ZIMMERMAN, J. S. & ROWE, D. T. 2002. Epstein-Barr virus latent membrane protein 2B (LMP2B) co-localizes with LMP2A in perinuclear regions in transiently transfected cells. *J Gen Virol,* 83**,** 1025-35.

LYNGAA, R., NØRREGAARD, K., KRISTENSEN, M., KUBALE, V., ROSENKILDE, M. M. & KLEDAL, T. N. 2010. Cell transformation mediated by the Epstein–Barr virus G protein-coupled receptor BILF1 is dependent on constitutive signaling. *Oncogene,* 29**,** 4388-4398.

MAGGIO, E. M., VAN DEN BERG, A., DE JONG, D., DIEPSTRA, A. & POPPEMA, S. 2003. Low Frequency of FAS Mutations in Reed-Sternberg Cells of Hodgkin's Lymphoma. *The American Journal of Pathology,* 162**,** 29-35.

MAGRATH, I. 1990. The pathogenesis of Burkitt's lymphoma. *Adv Cancer Res,* 55**,** 133-270.

MAGRATH, I. 2012. Towards Curative Therapy in Burkitt Lymphoma: The Role of Early African Studies in Demonstrating the Value of Combination Therapy and CNS Prophylaxis. *Advances in Hematology,* 2012**,** 1-7.

MANI, H. & JAFFE, E. S. 2009. Hodgkin lymphoma: an update on its biology with new insights into classification. *Clin Lymphoma Myeloma,* 9**,** 206-16.

MARCHINI, A., TOMKINSON, B., COHEN, J. I. & KIEFF, E. 1991. BHRF1, the Epstein-Barr virus gene with homology to Bc12, is dispensable for B-lymphocyte transformation and virus replication. *J Virol,* 65**,** 5991-6000.

MARCU, K. B., BOSSONE, S. A. & PATEL, A. J. 1992. myc function and regulation. *Annu Rev Biochem,* 61**,** 809-60.

MARECHAL, V., DEHEE, A., CHIKHI-BRACHET, R., PIOLOT, T., COPPEY-MOISAN, M. & NICOLAS, J. C. 1999. Mapping EBNA-1 domains involved in binding to metaphase chromosomes. *J Virol,* 73**,** 4385-92.

MARR-BELVIN, A. K., CARVILLE, A. K., FAHEY, M. A., BOISVERT, K., KLUMPP, S. A., OHASHI, M., WANG, F., O'NEIL, S. P. & WESTMORELAND, S. V. 2008. Rhesus lymphocryptovirus type 1-associated B-cell nasal lymphoma in SIV-infected rhesus macaques. *Vet Pathol,* 45**,** 914-21.

MARSHALL, N. A. 2004. Immunosuppressive regulatory T cells are abundant in the reactive lymphocytes of Hodgkin lymphoma. *Blood,* 103**,** 1755-1762.

MARSHALL, W. L., YIM, C., GUSTAFSON, E., GRAF, T., SAGE, D. R., HANIFY, K., WILLIAMS, L., FINGEROTH, J. & FINBERG, R. W. 1999. Epstein-Barr virus encodes a novel homolog of the bcl-2 oncogene that inhibits apoptosis and associates with Bax and Bak. *J Virol,* 73**,** 5181-5.

MARTHA, S., REDDY, K. H., BISWAL, N. & PARIDA, K. 2012. Facile synthesis of InGaZn mixed oxide nanorods for enhanced hydrogen production under visible light. *Dalton Trans,* 41**,** 14107-16.

MARTIN-SUBERO, J. I., KLAPPER, W., SOTNIKOVA, A., CALLET-BAUCHU, E., HARDER, L., BASTARD, C., SCHMITZ, R., GROHMANN, S., HOPPNER, J., RIEMKE, J., BARTH, T. F. E., BERGER, F., BERND, H. W., CLAVIEZ, A., GESK, S., FRANK, G. A., KAPLANSKAYA, I. B., MOLLER, P., PARWARESCH, R. M., RUDIGER, T., STEIN, H., KUPPERS, R., HANSMANN, M. L. & SIEBERT, R. 2006. Chromosomal Breakpoints Affecting Immunoglobulin Loci Are Recurrent in Hodgkin and Reed-Sternberg Cells of Classical Hodgkin Lymphoma. *Cancer Research,* 66**,** 10332-10338.

MARUO, S., WU, Y., ITO, T., KANDA, T., KIEFF, E. D. & TAKADA, K. 2009. Epstein-Barr virus nuclear protein EBNA3C residues critical for maintaining lymphoblastoid cell growth. *Proceedings of the National Academy of Sciences,* 106**,** 4419-4424.

MATEYAK, M. K., OBAYA, A. J. & SEDIVY, J. M. 1999. c-Myc regulates cyclin D-Cdk4 and -Cdk6 activity but affects cell cycle progression at multiple independent points. *Mol Cell Biol,* 19**,** 4672-83.

MATHAS, S. 2002. Aberrantly expressed c-Jun and JunB are a hallmark of Hodgkin lymphoma cells, stimulate proliferation and synergize with NF-kappaB. *The EMBO Journal,* 21**,** 4104-4113.

MATHAS, S., LIETZ, A., ANAGNOSTOPOULOS, I., HUMMEL, F., WIESNER, B., JANZ, M., JUNDT, F., HIRSCH, B., JÖHRENS-LEDER, K., VORNLOCHER, H.-P., BOMMERT, K., STEIN, H. & DÖRKEN, B. 2004. c-FLIP Mediates Resistance of Hodgkin/Reed-Sternberg Cells to Death Receptor–induced Apoptosis. *The Journal of Experimental Medicine,* 199**,** 1041-1052.

MCCARTHY, N. J., HAZLEWOOD, S. A., HUEN, D. S., RICKINSON, A. B. & WILLIAMS, G. T. 1996. The Epstein-Barr virus gene BHRF1, a homologue of the cellular oncogene Bcl-2, inhibits apoptosis induced by gamma radiation and chemotherapeutic drugs. *Adv Exp Med Biol,* 406**,** 83-97.

MCKENNA, A., HANNA, M., BANKS, E., SIVACHENKO, A., CIBULSKIS, K., KERNYTSKY, A., GARIMELLA, K., ALTSHULER, D., GABRIEL, S., DALY, M. & DEPRISTO, M. A. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res,* 20**,** 1297-303.

MEINHARDT, A., BURKHARDT, B., ZIMMERMANN, M., BORKHARDT, A., KONTNY, U., KLINGEBIEL, T., BERTHOLD, F., JANKA-SCHAUB, G., KLEIN, C., KABICKOVA, E., KLAPPER, W., ATTARBASCHI, A., SCHRAPPE, M., REITER, A. & BERLIN-FRANKFURT-MUNSTER, G. 2010. Phase II window study on rituximab in newly diagnosed pediatric mature B-cell non-Hodgkin's lymphoma and Burkitt leukemia. *J Clin Oncol,* 28**,** 3115-21.

MEITINGER, C., STROBL, L. J., MARSCHALL, G., BORNKAMM, G. W. & ZIMBER-STROBL, U. 1994. Crucial sequences within the Epstein-Barr virus TP1 promoter for EBNA2-mediated transactivation and interaction of EBNA2 with its responsive element. *J Virol,* 68**,** 7497-506.

MERCHANT, M., CALDWELL, R. G. & LONGNECKER, R. 2000. The LMP2A ITAM is essential for providing B cells with development and survival signals in vivo. *J Virol,* 74**,** 9115-24.

MERCHANT, M. & LONGNECKER, R. 2001. LMP2A survival and developmental signals are transmitted through Btk-dependent and Btk-independent pathways. *Virology,* 291**,** 46-54.

METKAR, S. S., NARESH, K. N., REDKAR, A. A., SOMAN, C. S., ADVANI, S. H. & NADKARNI, J. J. 1999. Expression of Fas and Fas Ligand in Hodgkin's Disease. *Leukemia & Lymphoma,* 33**,** 521-530.

MINK, S., MUTSCHLER, B., WEISKIRCHEN, R., BISTER, K. & KLEMPNAUER, K. H. 1996. A novel function for Myc: inhibition of C/EBP-dependent gene activation. *Proc Natl Acad Sci U S A,* 93**,** 6635-40.

MITCHELL, T. & SUGDEN, B. 1995. Stimulation of NF-kappa B-mediated transcription by mutant derivatives of the latent membrane protein of Epstein-Barr virus. *J Virol,* 69**,** 2968-76.

MIZUNO, T., ZHONG, X. & ROTHSTEIN, T. L. 2003. *APOPTOSIS,* 8**,** 451-460.

MOLYNEUX, E. M., ROCHFORD, R., GRIFFIN, B., NEWTON, R., JACKSON, G., MENON, G., HARRISON, C. J., ISRAELS, T. & BAILEY, S. 2012. Burkitt's lymphoma. *The Lancet,* 379**,** 1234-1244.

MONTI, S., SAVAGE, K. J., KUTOK, J. L., FEUERHAKE, F., KURTIN, P., MIHM, M., WU, B., PASQUALUCCI, L., NEUBERG, D., AGUIAR, R. C., DAL CIN, P.,

LADD, C., PINKUS, G. S., SALLES, G., HARRIS, N. L., DALLA-FAVERA, R., HABERMANN, T. M., ASTER, J. C., GOLUB, T. R. & SHIPP, M. A. 2005. Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. *Blood,* 105**,** 1851-61.

MOORMANN, A. M., CHELIMO, K., SUMBA, P. O., TISCH, D. J., ROCHFORD, R. & KAZURA, J. W. 2007. Exposure to holoendemic malaria results in suppression of Epstein-Barr virus-specific T cell immunosurveillance in Kenyan children. *J Infect Dis,* 195**,** 799-808.

MOORMANN, A. M., SNIDER, C. J. & CHELIMO, K. 2011. The company malaria keeps: how co-infection with Epstein-Barr virus leads to endemic Burkitt lymphoma. *Curr Opin Infect Dis,* 24**,** 435-41.

MOORTHY, R. K. & THORLEY-LAWSON, D. A. 1993. All three domains of the Epstein-Barr virus-encoded latent membrane protein LMP-1 are required for transformation of rat-1 fibroblasts. *Journal of Virology,* 67**,** 1638-1646.

MORGAN, D. G., NIEDERMAN, J. C., MILLER, G., SMITH, H. W. & DOWALIBY, J. M. 1979. Site of Epstein-Barr virus replication in the oropharynx. *Lancet,* 2**,** 1154-7.

MORIN, R. D., ASSOULINE, S., ALCAIDE, M., MOHAJERI, A., JOHNSTON, R. L., CHONG, L., GREWAL, J., YU, S., FORNIKA, D., BUSHELL, K., NIELSEN, T. H., PETROGIANNIS-HALIOTIS, T., CRUMP, M., TOSIKYAN, A., GRANDE, B. M., MACDONALD, D., ROUSSEAU, C., BAYAT, M., SESQUES, P., FROMENT, R., ALBUQUERQUE, M., MONCZAK, Y., OROS, K. K., GREENWOOD, C., RIAZALHOSSEINI, Y., ARSENEAULT, M., CAMLIOGLU, E., CONSTANTIN, A., PAN-HAMMARSTROM, Q., PENG, R., MANN, K. K. & JOHNSON, N. A. 2016. Genetic Landscapes of Relapsed and Refractory Diffuse Large B-Cell Lymphomas. *Clin Cancer Res,* 22**,** 2290-300.

MOSIALOS, G., BIRKENBACH, M., YALAMANCHILI, R., VANARSDALE, T., WARE, C. & KIEFF, E. 1995. The Epstein-Barr virus transforming protein LMP1 engages signaling proteins for the tumor necrosis factor receptor family. *Cell,* 80**,** 389-99.

MURPHY, S. B. 1980. Classification, staging and end results of treatment of childhood non-Hodgkin's lymphomas: dissimilarities from lymphomas in adults. *Semin Oncol,* 7**,** 332-9.

MURRAY, P. G., CONSTANDINOU, C. M., CROCKER, J., YOUNG, L. S. & AMBINDER, R. F. 1998. Analysis of major histocompatibility complex class I, TAP expression, and LMP2 epitope sequence in Epstein-Barr virus–positive Hodgkin's disease. *Blood,* 92**,** 2477-2483.

MURRAY, P. G., YOUNG, L. S., ROWE, M. & CROCKER, J. 1992a. Immunohistochemical demonstration of the Epstein-Barr virus-encoded latent membrane protein in paraffin sections of Hodgkin's disease. *J Pathol,* 166**,** 1-5.

MURRAY, R. J., KURILLA, M. G., BROOKS, J. M., THOMAS, W. A., ROWE, M., KIEFF, E. & RICKINSON, A. B. 1992b. Identification of target antigens for the human cytotoxic T cell response to Epstein-Barr virus (EBV): implications for the immune control of EBV-positive malignancies. *J Exp Med,* 176**,** 157-68.

MUSCHEN, M., RAJEWSKY, K., BRAUNINGER, A., BAUR, A. S., OUDEJANS, J. J., ROERS, A., HANSMANN, M. L. & KUPPERS, R. 2000a. Rare occurrence of classical Hodgkin's disease as a T cell lymphoma. *J Exp Med,* 191**,** 387-94.

MUSCHEN, M., RE, D., BRAUNINGER, A., WOLF, J., HANSMANN, M. L., DIEHL, V., KUPPERS, R. & RAJEWSKY, K. 2000b. Somatic mutations of the CD95 gene in Hodgkin and Reed-Sternberg cells. *Cancer Res,* 60**,** 5640-3.

MUTALIMA, N., MOLYNEUX, E., JAFFE, H., KAMIZA, S., BORGSTEIN, E., MKANDAWIRE, N., LIOMBA, G., BATUMBA, M., LAGOS, D., GRATRIX, F.,

BOSHOFF, C., CASABONNE, D., CARPENTER, L. M. & NEWTON, R. 2008. Associations between Burkitt Lymphoma among Children in Malawi and Infection with HIV, EBV and Malaria: Results from a Case-Control Study. *PLoS ONE,* 3**,** e2505.

NAKAGOMI, H., DOLCETTI, R., BEJARANO, M. T., PISA, P., KIESSLING, R. & MASUCCI, M. G. 1994. The Epstein-Barr virus latent membrane protein-1 (LMP1) induces interleukin-10 production in Burkitt lymphoma lines. *Int J Cancer,* 57**,** 240-4.

NEMEROW, G. R., HOUGHTEN, R. A., MOORE, M. D. & COOPER, N. R. 1989. Identification of an epitope in the major envelope protein of Epstein-Barr virus that mediates viral binding to the B lymphocyte EBV receptor (CR2). *Cell,* 56**,** 369-77.

NERI, A., BARRIGA, F., INGHIRAMI, G., KNOWLES, D. M., NEEQUAYE, J., MAGRATH, I. T. & DALLA-FAVERA, R. 1991. Epstein-Barr virus infection precedes clonal expansion in Burkitt's and acquired immunodeficiency syndrome-associated lymphoma. *Blood,* 77**,** 1092-5.

NETWORK., C. G. A. R. 2014. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature,* 513**,** 202-9.

NEUHIERL, B., FEEDERLE, R., ADHIKARY, D., HUB, B., GELETNEKY, K., MAUTNER, J. & DELECLUSE, H. J. 2009. Primary B-Cell Infection with a BALF4 Epstein-Barr Virus Comes to a Halt in the Endosomal Compartment yet Still Elicits a Potent CD4-Positive Cytotoxic T-Cell Response. *Journal of Virology,* 83**,** 4616-4623.

NEWCOM, S. R. & GU, L. 1995. Transforming growth factor beta 1 messenger RNA in Reed-Sternberg cells in nodular sclerosing Hodgkin's disease. *Journal of Clinical Pathology,* 48**,** 160-163.

NGOMA, T., ADDE, M., DUROSINMI, M., GITHANG'A, J., AKEN'OVA, Y., KAIJAGE, J., ADEODOU, O., RAJAB, J., BROWN, B. J., LEONCINI, L., NARESH, K., RAPHAEL, M., HURWITZ, N., SCANLAN, P., ROHATINER, A., VENZON, D. & MAGRATH, I. 2012. Treatment of Burkitt lymphoma in equatorial Africa using a simple three-drug combination followed by a salvage regimen for patients with persistent or recurrent disease. *British Journal of Haematology,* 158**,** 749-762.

NIEDOBITEK, G., AGATHANGGELOU, A., ROWE, M., JONES, E. L., JONES, D. B., TURYAGUMA, P., ORYEMA, J., WRIGHT, D. H. & YOUNG, L. S. 1995. Heterogeneous expression of Epstein-Barr virus latent proteins in endemic Burkitt's lymphoma. *Blood,* 86**,** 659-65.

NIEDOBITEK, G., KREMMER, E., HERBST, H., WHITEHEAD, L., DAWSON, C. W., NIEDOBITEK, E., VON OSTAU, C., ROONEY, N., GRASSER, F. A. & YOUNG, L. S. 1997. Immunohistochemical detection of the Epstein-Barr virus-encoded latent membrane protein 2A in Hodgkin's disease and infectious mononucleosis. *Blood,* 90**,** 1664-72.

NIJMEIJER, S., LEURS, R., SMIT, M. J. & VISCHER, H. F. 2010a. The Epstein-Barr virus-encoded G protein-coupled receptor BILF1 hetero-oligomerizes with human CXCR4, scavenges Galphai proteins, and constitutively impairs CXCR4 functioning. *J Biol Chem,* 285**,** 29632-41.

NIJMEIJER, S., LEURS, R., SMIT, M. J. & VISCHER, H. F. 2010b. The Epstein-Barr Virus-encoded G Protein-coupled Receptor BILF1 Hetero-oligomerizes with Human CXCR4, Scavenges GαiProteins, and Constitutively Impairs CXCR4 Functioning. *Journal of Biological Chemistry,* 285**,** 29632-29641.

NITSCHE, F., BELL, A. & RICKINSON, A. 1997. Epstein-Barr virus leader protein enhances EBNA-2-mediated transactivation of latent membrane protein 1 expression: a role for the W1W2 repeat domain. *J Virol,* 71**,** 6619-28.

NIU, X., WATTS, V. L., CINGOLANI, O. H., SIVAKUMARAN, V., LEYTON-MANGE, J. S., ELLIS, C. L., MILLER, K. L., VANDEGAER, K., BEDJA, D., GABRIELSON, K. L., PAOLOCCI, N., KASS, D. A. & BAROUCH, L. A. 2012. Cardioprotective effect of beta-3 adrenergic receptor agonism: role of neuronal nitric oxide synthase. *J Am Coll Cardiol,* 59**,** 1979-87.

NODA, C., MURATA, T., KANDA, T., YOSHIYAMA, H., SUGIMOTO, A., KAWASHIMA, D., SAITO, S., ISOMURA, H. & TSURUMI, T. 2011. Identification and Characterization of CCAAT Enhancer-binding Protein (C/EBP) as a Transcriptional Activator for Epstein-Barr Virus Oncogene Latent Membrane Protein 1. *Journal of Biological Chemistry,* 286**,** 42524-42533.

NONKWELO, C., SKINNER, J., BELL, A., RICKINSON, A. & SAMPLE, J. 1996. Transcription start sites downstream of the Epstein-Barr virus (EBV) Fp promoter in early-passage Burkitt lymphoma cells define a fourth promoter for expression of the EBV EBNA-1 protein. *J Virol,* 70**,** 623-7.

NOZAWA, Y., WAKASA, H. & ABE, M. 1998. Costimulatory molecules (CD80 and CD86) on Reed-Sternberg cells are associated with the proliferation of background T cells in Hodgkin's disease. *Pathology International,* 48**,** 10-14.

O'GRADY, T., WANG, X., HÖNER ZU BENTRUP, K., BADDOO, M., CONCHA, M. & FLEMINGTON, E. K. 2016. Global transcript structure resolution of high gene density genomes through multi-platform data integration. *Nucleic Acids Research,* 44**,** e145-e145.

ORIOL, A., RIBERA, J. M., BERGUA, J., GIMÉNEZ MESA, E., GRANDE, C., ESTEVE, J., BRUNET, S., MORENO, M. J., ESCODA, L., HERNANDEZ-RIVAS, J. M. & HOELZER, D. 2008. High-dose chemotherapy and immunotherapy in adult Burkitt lymphoma. *Cancer,* 113**,** 117-125.

OSTHUS, R. C., SHIM, H., KIM, S., LI, Q., REDDY, R., MUKHERJEE, M., XU, Y., WONSEY, D., LEE, L. A. & DANG, C. V. 2000. Deregulation of glucose transporter 1 and glycolytic gene expression by c-Myc. *J Biol Chem,* 275**,** 21797-800.

OUDEJANS, J., JIWA, N., KUMMER, J., HORSTMAN, A., VOS, W., BAAK, J., KLUIN, P. M., VAN DER VALK, P., WALBOOMERS, J. & MEIJER, C. 1996. Analysis of major histocompatibility complex class I expression on Reed-Sternberg cells in relation to the cytotoxic T-cell response in Epstein-Barr virus-positive and-negative Hodgkin's disease. *Blood,* 87**,** 3844-3851.

OWEN, T. J., D O'NEIL, J., DAWSON, C. W., HU, C., CHEN, X., YAO, Y., WOOD, V. H., MITCHELL, L. E., WHITE, R. J. & YOUNG, L. S. 2010. Epstein-Barr virus-encoded EBNA1 enhances RNA polymerase III-dependent EBER expression through induction of EBER-associated cellular transcription factors. *Molecular cancer,* 9**,** 241.

OZAWA, M. G., BHADURI, A., CHISHOLM, K. M., BAKER, S. A., MA, L., ZEHNDER, J. L., LUNA-FINEMAN, S., LINK, M. P., MERKER, J. D., ARBER, D. A. & OHGAMI, R. S. 2016. A study of the mutational landscape of pediatric-type follicular lymphoma and pediatric nodal marginal zone lymphoma. *Mod Pathol,* 29**,** 1212-20.

PAJIC, A., SPITKOVSKY, D., CHRISTOPH, B., KEMPKES, B., SCHUHMACHER, M., STAEGE, M. S., BRIELMEIER, M., ELLWART, J., KOHLHUBER, F., BORNKAMM, G. W., POLACK, A. & EICK, D. 2000. Cell cycle activation by c-myc in a Burkitt lymphoma model cell line. *International Journal of Cancer,* 87**,** 787-793.

PALLESEN, G., HAMILTON-DUTOIT, S. J., ROWE, M., LISSE, I., RALFKIAER, E., SANDVEJ, K. & YOUNG, L. S. 1991a. Expression of Epstein-Barr virus replicative proteins in AIDS-related non-Hodgkin's lymphoma cells. *J Pathol,* 165**,** 289-99.

PALLESEN, G., HAMILTON-DUTOIT, S. J., ROWE, M. & YOUNG, L. S. 1991b. Expression of Epstein-Barr virus latent gene products in tumour cells of Hodgkin's disease. *The Lancet,* 337**,** 320-322.

PANG, M.-F., LIN, K.-W. & PEH, S.-C. 2009. The signaling pathways of Epstein-Barr virus-encoded latent membrane protein 2A (LMP2A) in latency and cancer. *Cellular and Molecular Biology Letters,* 14.

PARK, H. Y., LEE, S. B., YOO, H. Y., KIM, S. J., KIM, W. S., KIM, J. I. & KO, Y. H. 2016. Whole-exome and transcriptome sequencing of refractory diffuse large B-cell lymphoma. *Oncotarget,* 7**,** 86433-86445.

PARKER, G. A., CROOK, T., BAIN, M., SARA, E. A., FARRELL, P. J. & ALLDAY, M. J. 1996. Epstein-Barr virus nuclear antigen (EBNA)3C is an immortalizing oncoprotein with similar properties to adenovirus E1A and papillomavirus E7. *Oncogene,* 13**,** 2541-9.

PASQUALUCCI, L., KHIABANIAN, H., FANGAZIO, M., VASISHTHA, M., MESSINA, M., HOLMES, A. B., OUILLETTE, P., TRIFONOV, V., ROSSI, D., TABBO, F., PONZONI, M., CHADBURN, A., MURTY, V. V., BHAGAT, G., GAIDANO, G., INGHIRAMI, G., MALEK, S. N., RABADAN, R. & DALLA-FAVERA, R. 2014. Genetics of follicular lymphoma transformation. *Cell Rep,* 6**,** 130-40.

PAULSEN, S. J., ROSENKILDE, M. M., EUGEN-OLSEN, J. & KLEDAL, T. N. 2005. Epstein-Barr virus-encoded BILF1 is a constitutively active G protein-coupled receptor. *J Virol,* 79**,** 536-46.

PEARSON, G. R., VROMAN, B., CHASE, B., SCULLEY, T., HUMMEL, M. & KIEFF, E. 1983. Identification of polypeptide components of the Epstein-Barr virus early antigen complex with monoclonal antibodies. *J Virol,* 47**,** 193-201.

PELICCI, P. G., KNOWLES, D. M., 2ND, ARLIN, Z. A., WIECZOREK, R., LUCIW, P., DINA, D., BASILICO, C. & DALLA-FAVERA, R. 1986. Multiple monoclonal B cell expansions and c-myc oncogene rearrangements in acquired immune deficiency syndrome-related lymphoproliferative disorders. Implications for lymphomagenesis. *J Exp Med,* 164**,** 2049-60.

PICCALUGA, P. P., AGOSTINELLI, C., GAZZOLA, A., TRIPODO, C., BACCI, F., SABATTINI, E., SISTA, M. T., MANNU, C., SAPIENZA, M. R., ROSSI, M., LAGINESTRA, M. A., SAGRAMOSO-SACCHETTI, C. A., RIGHI, S. & PILERI, S. A. 2011. Pathobiology of hodgkin lymphoma. *Adv Hematol,* 2011**,** 920898.

PIMIENTA, G., FOK, V., HASLIP, M., NAGY, M., TAKYAR, S. & STEITZ, J. A. 2015. Proteomics and Transcriptomics of BJAB Cells Expressing the Epstein-Barr Virus Noncoding RNAs EBER1 and EBER2. *PLOS ONE,* 10**,** e0124638.

PIRIOU, E. R., VAN DORT, K., NANLOHY, N. M., MIEDEMA, F., VAN OERS, M. H. & VAN BAARLE, D. 2004. Altered EBV viral load setpoint after HIV seroconversion is in accordance with lack of predictive value of EBV load for the occurrence of AIDS-related non-Hodgkin lymphoma. *J Immunol,* 172**,** 6931-7.

PLEIMAN, C. M., ABRAMS, C., GAUEN, L. T., BEDZYK, W., JONGSTRA, J., SHAW, A. S. & CAMBIER, J. C. 1994. Distinct p53/56lyn and p59fyn domains associate with nonphosphorylated and phosphorylated Ig-alpha. *Proc Natl Acad Sci U S A,* 91**,** 4268-72.

POLACK, A., DELIUS, H., ZIMBER, U. & BORNKAMM, G. W. 1984. Two deletions in the Epstein-Barr virus genome of the Burkitt lymphoma nonproducer line Raji. *Virology,* 133**,** 146-57.

POPE, J. H., HORNE, M. K. & SCOTT, W. 1968. Transformation of foetal human keukocytes in vitro by filtrates of a human leukaemic cell line containing herpes-like virus. *Int J Cancer,* 3**,** 857-66.

PORTIS, T., DYCK, P. & LONGNECKER, R. 2003. Epstein-Barr Virus (EBV) LMP2A induces alterations in gene transcription similar to those observed in Reed-Sternberg cells of Hodgkin lymphoma. *Blood,* 102**,** 4166-78.

PROCHOWNIK, E. V., KUKOWSKA, J. & RODGERS, C. 1988. c-myc antisense transcripts accelerate differentiation and inhibit G1 progression in murine erythroleukemia cells. *Mol Cell Biol,* 8**,** 3683-95.

PUSCH, O., SOUCEK, T., HENGSTSCHLAGER-OTTNAD, E., BERNASCHEK, G. & HENGSTSCHLAGER, M. 1997. Cellular targets for activation by c-Myc include the DNA metabolism enzyme thymidine kinase. *DNA Cell Biol,* 16**,** 737-47.

RABSON, M., GRADOVILLE, L., HESTON, L. & MILLER, G. 1982. Non-immortalizing P3J-HR-1 Epstein-Barr virus: a deletion mutant of its transforming parent, Jijoye. *J Virol,* 44**,** 834-44.

RADKOV, S. A., TOUITOU, R., BREHM, A., ROWE, M., WEST, M., KOUZARIDES, T. & ALLDAY, M. J. 1999. Epstein-Barr virus nuclear antigen 3C interacts with histone deacetylase to repress transcription. *J Virol,* 73**,** 5688-97.

RAGOCZY, T. & MILLER, G. 1999. Role of the epstein-barr virus RTA protein in activation of distinct classes of viral lytic cycle genes. *J Virol,* 73**,** 9858-66.

RAGOCZY, T. & MILLER, G. 2001. Autostimulation of the Epstein-Barr virus BRLF1 promoter is mediated through consensus Sp1 and Sp3 binding sites. *J Virol,* 75**,** 5240-51.

RAINEY, J. J., MWANDA, W. O., WAIRIUMU, P., MOORMANN, A. M., WILSON, M. L. & ROCHFORD, R. 2007a. Spatial distribution of Burkitt's lymphoma in Kenya and association with malaria risk. *Trop Med Int Health,* 12**,** 936-43.

RAINEY, J. J., OMENAH, D., SUMBA, P. O., MOORMANN, A. M., ROCHFORD, R. & WILSON, M. L. 2007b. Spatial clustering of endemic Burkitt's lymphoma in high-risk regions of Kenya. *Int J Cancer,* 120**,** 121-7.

RAMASUBRAMANYAN, S., OSBORN, K., AL-MOHAMMAD, R., NARANJO PEREZ-FERNANDEZ, I. B., ZUO, J., BALAN, N., GODFREY, A., PATEL, H., PETERS, G., ROWE, M., JENNER, R. G. & SINCLAIR, A. J. 2015. Epstein-Barr virus transcription factor Zta acts through distal regulatory elements to directly control cellular gene expression. *Nucleic Acids Research,* 43**,** 3563-3577.

RASTI, N., FALK, K. I., DONATI, D., GYAN, B. A., GOKA, B. Q., TROYE-BLOMBERG, M., AKANMORI, B. D., KURTZHALS, J. A., DODOO, D., CONSOLINI, R., LINDE, A., WAHLGREN, M. & BEJARANO, M. T. 2005. Circulating epstein-barr virus in children living in malaria-endemic areas. *Scand J Immunol,* 61**,** 461-5.

RATHORE, B. & KADIN, M. E. 2010. Hodgkin's lymphoma therapy: past, present, and future. *Expert Opin Pharmacother,* 11**,** 2891-906.

RE, D., HOFMANN, A., WOLF, J., DIEHL, V. & STARATSCHEK-JOX, A. 2000. Cultivated H-RS cells are resistant to CD95L-mediated apoptosis despite expression of wild-type CD95. *Experimental Hematology,* 28**,** 31-35.

RECHSTEINER, M. P., BERGER, C., ZAUNER, L., SIGRIST, J. A., WEBER, M., LONGNECKER, R., BERNASCONI, M. & NADAL, D. 2008. Latent membrane protein 2B regulates susceptibility to induction of lytic Epstein-Barr virus infection. *J Virol,* 82**,** 1739-47.

REEDMAN, B. M. & KLEIN, G. 1973. Cellular localization of an Epstein-Barr virus (EBV)-associated complement-fixing antigen in producer and non-producer lymphoblastoid cell lines. *International Journal of Cancer,* 11**,** 499-520.

REICHEL, J., CHADBURN, A., RUBINSTEIN, P. G., GIULINO-ROTH, L., TAM, W., LIU, Y., GAIOLLA, R., ENG, K., BRODY, J., INGHIRAMI, G., CARLO-STELLA,

C., SANTORO, A., RAHAL, D., TOTONCHY, J., ELEMENTO, O., CESARMAN, E. & ROSHAL, M. 2015. Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. *Blood,* 125**,** 1061-72.

RENGSTL, B., NEWRZELA, S., HEINRICH, T., WEISER, C., THALHEIMER, F. B., SCHMID, F., WARNER, K., HARTMANN, S., SCHROEDER, T., KUPPERS, R., RIEGER, M. A. & HANSMANN, M. L. 2013. Incomplete cytokinesis and re-fusion of small mononucleated Hodgkin cells lead to giant multinucleated Reed-Sternberg cells. *Proceedings of the National Academy of Sciences,* 110**,** 20729-20734.

RIAZ, N., MORRIS, L. G., LEE, W. & CHAN, T. A. 2014. Unraveling the molecular genetics of head and neck cancer through genome-wide approaches. *Genes & Diseases,* 1**,** 75-86.

RICHTER, J., SCHLESNER, M., HOFFMANN, S., KREUZ, M., LEICH, E., BURKHARDT, B., ROSOLOWSKI, M., AMMERPOHL, O., WAGENER, R., BERNHART, S. H., LENZE, D., SZCZEPANOWSKI, M., PAULSEN, M., LIPINSKI, S., RUSSELL, R. B., ADAM-KLAGES, S., APIC, G., CLAVIEZ, A., HASENCLEVER, D., HOVESTADT, V., HORNIG, N., KORBEL, J. O., KUBE, D., LANGENBERGER, D., LAWERENZ, C., LISFELD, J., MEYER, K., PICELLI, S., PISCHIMAROV, J., RADLWIMMER, B., RAUSCH, T., ROHDE, M., SCHILHABEL, M., SCHOLTYSIK, R., SPANG, R., TRAUTMANN, H., ZENZ, T., BORKHARDT, A., DREXLER, H. G., MÖLLER, P., MACLEOD, R. A. F., POTT, C., SCHREIBER, S., TRÜMPER, L., LOEFFLER, M., STADLER, P. F., LICHTER, P., EILS, R., KÜPPERS, R., HUMMEL, M., KLAPPER, W., ROSENSTIEL, P., ROSENWALD, A., BRORS, B. & SIEBERT, R. 2012. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nature Genetics,* 44**,** 1316-1320.

RICKINSON, A.B. &KIEFF, E. 2013. Epstein-Barr Virus. Philadelphia: Lippincott Williams and Wilkins; 2655–2700 p

ROBERTSON, K. D. & AMBINDER, R. F. 1997. Mapping promoter regions that are hypersensitive to methylation-mediated inhibition of transcription: application of the methylation cassette assay to the Epstein-Barr virus major latency promoter. *J Virol,* 71**,** 6445-54.

ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-40.

ROCHFORD, R., CANNON, M. J. & MOORMANN, A. M. 2005. Opinion — tropical infectious diseases: Endemic Burkitt's lymphoma: a polymicrobial disease? *Nature Reviews Microbiology,* 3**,** 182-187.

ROONEY, C. M., ROWE, D. T., RAGOT, T. & FARRELL, P. J. 1989. The spliced BZLF1 gene of Epstein-Barr virus (EBV) transactivates an early EBV promoter and induces the virus productive cycle. *J Virol,* 63**,** 3109-16.

ROSA, M. D., GOTTLIEB, E., LERNER, M. R. & STEITZ, J. A. 1981. Striking similarities are exhibited by two small Epstein-Barr virus-encoded ribonucleic acids and the adenovirus-associated ribonucleic acids VAI and VAII. *Molecular and Cellular Biology,* 1**,** 785-796.

ROSDAHL, N., LARSEN, S. O. & CLEMMESEN, J. 1974. Hodgkin's Disease in Patients with Previous Infectious Mononucleosis: 30 Years' Experience. *BMJ,* 2**,** 253-256.

ROSENWALD, I. B., RHOADS, D. B., CALLANAN, L. D., ISSELBACHER, K. J. & SCHMIDT, E. V. 1993. Increased expression of eukaryotic translation initiation factors eIF-4E and eIF-2 alpha in response to growth induction by c-myc. *Proc Natl Acad Sci U S A,* 90**,** 6175-8.

ROUGHAN, J. E. & THORLEY-LAWSON, D. A. 2009. The intersection of Epstein-Barr virus with the germinal center. *J Virol,* 83**,** 3968-76.

ROVEDO, M. & LONGNECKER, R. 2007. Epstein-barr virus latent membrane protein 2B (LMP2B) modulates LMP2A activity. *J Virol,* 81**,** 84-94.

ROWE, M., FITZSIMMONS, L. & BELL, A. I. 2014a. Epstein-Barr virus and Burkitt lymphoma. *Chin J Cancer,* 33**,** 609-19.

ROWE, M., GLAUNSINGER, B., VAN LEEUWEN, D., ZUO, J., SWEETMAN, D., GANEM, D., MIDDELDORP, J., WIERTZ, E. J. H. J. & RESSING, M. E. 2007. Host shutoff during productive Epstein-Barr virus infection is mediated by BGLF5 and may contribute to immune evasion. *Proceedings of the National Academy of Sciences,* 104**,** 3366-3371.

ROWE, M., KELLY, G. L., BELL, A. I. & RICKINSON, A. B. 2009. Burkitt's lymphoma: the Rosetta Stone deciphering Epstein-Barr virus biology. *Semin Cancer Biol,* 19**,** 377-88.

ROWE, M., RAITHATHA, S. & SHANNON-LOWE, C. 2014b. Counteracting effects of cellular Notch and Epstein-Barr virus EBNA2: implications for stromal effects on virus-host interactions. *J Virol,* 88**,** 12065-76.

ROWE, M., ROWE, D., GREGORY, C. D., YOUNG, L. S., FARRELL, P., RUPANI, H. & RICKINSON, A. 1987a. Differences in B cell growth phenotype reflect novel patterns of Epstein-Barr virus latent gene expression in Burkitt's lymphoma cells. *The EMBO journal,* 6**,** 2743.

ROWE, M., ROWE, D. T., GREGORY, C. D., YOUNG, L. S., FARRELL, P. J., RUPANI, H. & RICKINSON, A. B. 1987b. Differences in B cell growth phenotype reflect novel patterns of Epstein-Barr virus latent gene expression in Burkitt's lymphoma cells. *EMBO J,* 6**,** 2743-51.

ROWE, M. & ZUO, J. 2010. Immune responses to Epstein–Barr virus: molecular interactions in the virus evasion of CD8+ T cell immunity. *Microbes and Infection,* 12**,** 173-181.

RU, H. Y., CHEN, R. L., LU, W. C. & CHEN, J. H. 2002. hBUB1 defects in leukemia and lymphoma cells. *Oncogene,* 21**,** 4673-4679.

RUF, I. K., LACKEY, K. A., WARUDKAR, S. & SAMPLE, J. T. 2005. Protection from interferon-induced apoptosis by Epstein-Barr virus small RNAs is not mediated by inhibition of PKR. *J Virol,* 79**,** 14562-9.

RUF, I. K., RHYNE, P. W., YANG, C., CLEVELAND, J. L. & SAMPLE, J. T. 2000. Epstein-Barr Virus Small RNAs Potentiate Tumorigenicity of Burkitt Lymphoma Cells Independently of an Effect on Apoptosis. *Journal of Virology,* 74**,** 10223-10228.

SÁEZ, C., PEREDA, T., BORRERO, J. J., ESPINA, A., ROMERO, F., TORTOLERO, M., PINTOR-TORO, J. A., SEGURA, D. I. & JAPÓN, M. A. 2002. Expression of hpttg proto-oncogene in lymphoid neoplasias. *Oncogene,* 21**,** 8173-8177.

SAHA, A., HALDER, S., UPADHYAY, S. K., LU, J., KUMAR, P., MURAKAMI, M., CAI, Q. & ROBERTSON, E. S. 2011. Epstein-Barr Virus Nuclear Antigen 3C Facilitates G1-S Transition by Stabilizing and Enhancing the Function of Cyclin D1. *PLoS Pathogens,* 7**,** e1001275.

SAHA, A., MURAKAMI, M., KUMAR, P., BAJAJ, B., SIMS, K. & ROBERTSON, E. S. 2009. Epstein-Barr Virus Nuclear Antigen 3C Augments Mdm2-Mediated p53 Ubiquitination and Degradation by Deubiquitinating Mdm2. *Journal of Virology,* 83**,** 4652-4669.

SAMPLE, J., BROOKS, L., SAMPLE, C., YOUNG, L., ROWE, M., GREGORY, C., RICKINSON, A. & KIEFF, E. 1991. Restricted Epstein-Barr virus protein expression in Burkitt lymphoma is due to a different Epstein-Barr nuclear antigen 1 transcriptional initiation site. *Proc Natl Acad Sci U S A,* 88**,** 6343-7.

SAMPLE, J., LIEBOWITZ, D. & KIEFF, E. 1989. Two related Epstein-Barr virus membrane proteins are encoded by separate genes. *J Virol,* 63**,** 933-7.

SAN LUCAS, F. A., WANG, G., SCHEET, P. & PENG, B. 2012. Integrated annotation and analysis of genetic variants from next-generation sequencing studies with variant tools. *Bioinformatics,* 28**,** 421-2.

SAUNDERS, C. T., WONG, W. S., SWAMY, S., BECQ, J., MURRAY, L. J. & CHEETHAM, R. K. 2012. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics,* 28**,** 1811-7.

SCHAEFER, B. C., STROMINGER, J. L. & SPECK, S. H. 1995. Redefining the Epstein-Barr virus-encoded nuclear antigen EBNA-1 gene promoter and transcription initiation site in group I Burkitt lymphoma cell lines. *Proceedings of the National Academy of Sciences,* 92**,** 10565-10569.

SCHMITZ, R., HANSMANN, M. L., BOHLE, V., MARTIN-SUBERO, J. I., HARTMANN, S., MECHTERSHEIMER, G., KLAPPER, W., VATER, I., GIEFING, M., GESK, S., STANELLE, J., SIEBERT, R. & KUPPERS, R. 2009. TNFAIP3 (A20) is a tumor suppressor gene in Hodgkin lymphoma and primary mediastinal B cell lymphoma. *J Exp Med,* 206**,** 981-9.

SCHMITZ, R., YOUNG, R. M., CERIBELLI, M., JHAVAR, S., XIAO, W., ZHANG, M., WRIGHT, G., SHAFFER, A. L., HODSON, D. J., BURAS, E., LIU, X., POWELL, J., YANG, Y., XU, W., ZHAO, H., KOHLHAMMER, H., ROSENWALD, A., KLUIN, P., MULLER-HERMELINK, H. K., OTT, G., GASCOYNE, R. D., CONNORS, J. M., RIMSZA, L. M., CAMPO, E., JAFFE, E. S., DELABIE, J., SMELAND, E. B., OGWANG, M. D., REYNOLDS, S. J., FISHER, R. I., BRAZIEL, R. M., TUBBS, R. R., COOK, J. R., WEISENBURGER, D. D., CHAN, W. C., PITTALUGA, S., WILSON, W., WALDMANN, T. A., ROWE, M., MBULAITEYE, S. M., RICKINSON, A. B. & STAUDT, L. M. 2012. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature,* 490**,** 116-20.

SCHRADER, A., BENTINK, S., SPANG, R., LENZE, D., HUMMEL, M., KUO, M., ARRAND, J. R., MURRAY, P. G., TRUMPER, L., KUBE, D. & VOCKERODT, M. 2012. High Myc activity is an independent negative prognostic factor for diffuse large B cell lymphomas. *Int J Cancer,* 131**,** E348-61.

SEARS, J., UJIHARA, M., WONG, S., OTT, C., MIDDELDORP, J. & AIYAR, A. 2004. The amino terminus of Epstein-Barr Virus (EBV) nuclear antigen 1 contains AT hooks that facilitate the replication and partitioning of latent EBV genomes by tethering them to cellular chromosomes. *J Virol,* 78**,** 11487-505.

SETO, E., MOOSMANN, A., GRÖMMINGER, S., WALZ, N., GRUNDHOFF, A. & HAMMERSCHMIDT, W. 2010. Micro RNAs of Epstein-Barr Virus Promote Cell Cycle Progression and Prevent Apoptosis of Primary Human B Cells. *PLoS Pathogens,* 6**,** e1001063.

SHANNON-LOWE, C., ADLAND, E., BELL, A. I., DELECLUSE, H.-J., RICKINSON, A. B. & ROWE, M. 2009. Features distinguishing Epstein-Barr virus infections of epithelial cells and B cells: viral genome expression, genome maintenance, and genome amplification. *Journal of virology,* 83**,** 7749-7760.

SHARP, T. V., SCHWEMMLE, M., JEFFREY, I., LAING, K., MELLOR, H., PROUD, C. G., HILSE, K. & CLEMENS, M. J. 1993. Comparative analysis of the regulation of the interferoninducible protein kinase PKR by Epstein - Barr virus RNAs EBER-1 and EBER-2 and adenovirus VA, RNA. *Nucleic Acids Research,* 21**,** 4483-4490.

SHIBATA, D. & WEISS, L. M. 1992. Epstein-Barr virus-associated gastric adenocarcinoma. *Am J Pathol,* 140**,** 769-74.

SHIM, H., DOLDE, C., LEWIS, B. C., WU, C. S., DANG, G., JUNGMANN, R. A., DALLA-FAVERA, R. & DANG, C. V. 1997. c-Myc transactivation of LDH-A: implications for tumor metabolism and growth. *Proc Natl Acad Sci U S A,* 94**,** 6658-63.

SHIRAMIZU, B., BARRIGA, F., NEEQUAYE, J., JAFRI, A., DALLA-FAVERA, R., NERI, A., GUTTIEREZ, M., LEVINE, P. & MAGRATH, I. 1991. Patterns of chromosomal breakpoint locations in Burkitt's lymphoma: relevance to geography and Epstein-Barr virus association. *Blood,* 77**,** 1516-26.

SINCLAIR, A. J., BRIMMELL, M., SHANAHAN, F. & FARRELL, P. J. 1991. Pathways of activation of the Epstein-Barr virus productive cycle. *J Virol,* 65**,** 2237-44.

SINCLAIR, A. J. & FARRELL, P. J. 1995. Host cell requirements for efficient infection of quiescent primary B lymphocytes by Epstein-Barr virus. *J Virol,* 69**,** 5461-8.

SINCLAIR, A. J., PALMERO, I., PETERS, G. & FARRELL, P. 1994. EBNA-2 and EBNA-LP cooperate to cause G0 to G1 transition during immortalization of resting human B lymphocytes by Epstein-Barr virus. *The EMBO journal,* 13**,** 3321.

SIXBEY, J. W., SHIRLEY, P., CHESNEY, P. J., BUNTIN, D. M. & RESNICK, L. 1989. Detection of a second widespread strain of Epstein-Barr virus. *Lancet,* 2**,** 761-5.

SKARE, J., FARLEY, J., STROMINGER, J., FRESEN, K., CHO, M. & ZUR HAUSEN, H. 1985. Transformation by Epstein-Barr virus requires DNA sequences in the region of BamHI fragments Y and H. *Journal of virology,* 55**,** 286-297.

SKINNIDER, B. F. 2001. Interleukin 13 and interleukin 13 receptor are frequently expressed by Hodgkin and Reed-Sternberg cells of Hodgkin lymphoma. *Blood,* 97**,** 250-255.

SKINNIDER, B. F. 2002. Signal transducer and activator of transcription 6 is frequently activated in Hodgkin and Reed-Sternberg cells of Hodgkin lymphoma. *Blood,* 99**,** 618-626.

SPANGGORD, R. J., VUYISICH, M. & BEAL, P. A. 2002. Identification of Binding Sites for Both dsRBMs of PKR on Kinase-Activating and Kinase-Inhibiting RNA Ligands†. *Biochemistry,* 41**,** 4511-4520.

SPECK, P., KLINE, K. A., CHERESH, P. & LONGNECKER, R. 1999. Epstein-Barr virus lacking latent membrane protein 2 immortalizes B cells with efficiency indistinguishable from that of wild-type virus. *J Gen Virol,* 80 (Pt 8)**,** 2193-203.

SPECK, S. H., CHATILA, T. & FLEMINGTON, E. 1997. Reactivation of Epstein-Barr virus: regulation and function of the BZLF1 gene. *Trends Microbiol,* 5**,** 399-405.

SPECK, S. H. & STROMINGER, J. L. 1985. Analysis of the transcript encoding the latent Epstein-Barr virus nuclear antigen I: a potentially polycistronic message generated by long-range splicing of several exons. *Proc Natl Acad Sci U S A,* 82**,** 8305-9.

SPEIR, M. L., ZWEIG, A. S., ROSENBLOOM, K. R., RANEY, B. J., PATEN, B., NEJAD, P., LEE, B. T., LEARNED, K., KAROLCHIK, D., HINRICHS, A. S., HEITNER, S., HARTE, R. A., HAEUSSLER, M., GURUVADOO, L., FUJITA, P. A., EISENHART, C., DIEKHANS, M., CLAWSON, H., CASPER, J., BARBER, G. P., HAUSSLER, D., KUHN, R. M. & KENT, W. J. 2016. The UCSC Genome Browser database: 2016 update. *Nucleic Acids Res,* 44**,** D717-25.

SPURGEON, S. L., JONES, R. C. & RAMAKRISHNAN, R. 2008. High Throughput Gene Expression Measurement with Real Time PCR in a Microfluidic Dynamic Array. *PLoS ONE,* 3**,** e1662.

STEIDL, C., DIEPSTRA, A., LEE, T., CHAN, F. C., FARINHA, P., TAN, K., TELENIUS, A., BARCLAY, L., SHAH, S. P., CONNORS, J. M., VAN DEN BERG, A. & GASCOYNE, R. D. 2012. Gene expression profiling of microdissected Hodgkin Reed-Sternberg cells correlates with treatment outcome in classical Hodgkin lymphoma. *Blood,* 120**,** 3530-3540.

STUART, E. T. & GRUSS, P. 1995. PAX genes: what's new in developmental biology and cancer? *Hum Mol Genet,* 4 Spec No**,** 1717-20.

SUGANO, N., CHEN, W., ROBERTS, M. L. & COOPER, N. R. 1997. Epstein-Barr virus binding to CD21 activates the initial viral promoter via NF-kappaB induction. *J Exp Med,* 186**,** 731-7.

SWAMINATHAN, S., TOMKINSON, B. & KIEFF, E. 1991. Recombinant Epstein-Barr virus with small RNA (EBER) genes deleted transforms lymphocytes and replicates in vitro. *Proc Natl Acad Sci U S A,* 88**,** 1546-50.

SWART, R., RUF, I. K., SAMPLE, J. & LONGNECKER, R. 2000. Latent membrane protein 2A-mediated effects on the phosphatidylinositol 3-Kinase/Akt pathway. *J Virol,* 74**,** 10838-45.

SWERDLOW, S. H., CAMPO, E., PILERI, S. A., HARRIS, N. L., STEIN, H., SIEBERT, R., ADVANI, R., GHIELMINI, M., SALLES, G. A., ZELENETZ, A. D. & JAFFE, E. S. 2016. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood,* 127**,** 2375-90.

SZYMANOWSKA, N., KLAPPER, W., GESK, S., KÜPPERS, R., MARTÍN-SUBERO, J. I. & SIEBERT, R. 2008. BCL2 and BCL3 are recurrent translocation partners of the IGH locus. *Cancer Genetics and Cytogenetics,* 186**,** 110-114.

T, H. 1832. On some Morbid Appearances of the Absorbent Glands and Spleen. *Medico-Chirurgical Transactions,* 17.

TAKADA, K. & ONO, Y. 1989. Synchronous and sequential activation of latently infected Epstein-Barr virus genomes. *J Virol,* 63**,** 445-9.

THANGAVELU, M. & LE BEAU, M. M. 1989. Chromosomal abnormalities in Hodgkin's disease. *Hematol Oncol Clin North Am,* 3**,** 221-36.

THOMAS, D. A., FADERL, S., O'BRIEN, S., BUESO-RAMOS, C., CORTES, J., GARCIA-MANERO, G., GILES, F. J., VERSTOVSEK, S., WIERDA, W. G., PIERCE, S. A., SHAN, J., BRANDT, M., HAGEMEISTER, F. B., KEATING, M. J., CABANILLAS, F. & KANTARJIAN, H. 2006. Chemoimmunotherapy with hyper-CVAD plus rituximab for the treatment of adult Burkitt and Burkitt-type lymphoma or acute lymphoblastic leukemia. *Cancer,* 106**,** 1569-1580.

THORLEY-LAWSON, D. A. & POODRY, C. A. 1982. Identification and isolation of the main component (gp350-gp220) of Epstein-Barr virus responsible for generating neutralizing antibodies in vivo. *J Virol,* 43**,** 730-6.

TIERNEY, R. J., SHANNON-LOWE, C. D., FITZSIMMONS, L., BELL, A. I. & ROWE, M. 2015. Unexpected patterns of Epstein–Barr virus transcription revealed by a High throughput PCR array for absolute quantification of viral mRNA. *Virology,* 474**,** 117-130.

TORRISI, M. R., CIRONE, M., PAVAN, A., ZOMPETTA, C., BARILE, G., FRATI, L. & FAGGIONI, A. 1989. Localization of Epstein-Barr virus envelope glycoproteins on the inner nuclear membrane of virus-producing cells. *J Virol,* 63**,** 828-32.

TOVEY, M. G., LENOIR, G. & BEGON-LOURS, J. 1978. Activation of latent Epstein-Barr virus by antibody to human IgM. *Nature,* 276**,** 270-2.

TSAI, C. N., LEE, C. M., CHIEN, C. K., KUO, S. C. & CHANG, Y. S. 1999. Additive effect of Sp1 and Sp3 in regulation of the ED-L1E promoter of the EBV LMP 1 gene in human epithelial cells. *Virology,* 261**,** 288-94.

TSIMBOURI, P., DROTAR, M. E., COY, J. L. & WILSON, J. B. 2002. bcl-xL and RAG genes are induced and the response to IL-2 enhanced in EmuEBNA-1 transgenic mouse lymphocytes. *Oncogene,* 21**,** 5182-7.

TSURUMI, T., FUJITA, M. & KUDOH, A. 2005. Latent and lytic Epstein-Barr virus replication strategies. *Rev Med Virol,* 15**,** 3-15.

TSURUMI, T., KOBAYASHI, A., TAMAI, K., YAMADA, H., DAIKOKU, T., YAMASHITA, Y. & NISHIYAMA, Y. 1996. Epstein-Barr virus single-stranded DNA-binding protein: purification, characterization, and action on DNA synthesis by the viral DNA polymerase. *Virology,* 222**,** 352-64.

USHMOROV, A., LEITHAUSER, F., SAKK, O., WEINHAUSEL, A., POPOV, S. W., MOLLER, P. & WIRTH, T. 2006. Epigenetic processes play a major role in B-cell-specific gene silencing in classical Hodgkin lymphoma. *Blood,* 107**,** 2493-500.

VAN ROOSBROECK, K., COX, L., TOUSSEYN, T., LAHORTIGA, I., GIELEN, O., CAUWELIER, B., DE PAEPE, P., VERHOEF, G., MARYNEN, P., VANDENBERGHE, P., DE WOLF-PEETERS, C., COOLS, J. & WLODARSKA, I. 2011. JAK2 rearrangements, including the novel SEC31A-JAK2 fusion, are recurrent in classical Hodgkin lymphoma. *Blood,* 117**,** 4056-4064.

VERBEKE, C. S., WENTHE, U., GROBHOLZ, R. & ZENTGRAF, H. 2001. Fas Ligand Expression in Hodgkin Lymphoma. *The American Journal of Surgical Pathology,* 25**,** 388-394.

VETTRAINO, M., MANERBA, M., GOVONI, M. & DI STEFANO, G. 2013. Galloflavin suppresses lactate dehydrogenase activity and causes MYC downregulation in Burkitt lymphoma cells through NAD/NADH-dependent inhibition of sirtuin-1. *Anticancer Drugs,* 24**,** 862-70.

VIEIRA, P., DE WAAL-MALEFYT, R., DANG, M. N., JOHNSON, K. E., KASTELEIN, R., FIORENTINO, D. F., DEVRIES, J. E., RONCAROLO, M. G., MOSMANN, T. R. & MOORE, K. W. 1991. Isolation and expression of human cytokine synthesis inhibitory factor cDNA clones: homology to Epstein-Barr virus open reading frame BCRFI. *Proc Natl Acad Sci U S A,* 88**,** 1172-6.

VISCHER, H. F., NIJMEIJER, S., SMIT, M. J. & LEURS, R. 2008. Viral hijacking of human receptors through heterodimerization. *Biochemical and Biophysical Research Communications,* 377**,** 93-97.

VITRE, B. D. & CLEVELAND, D. W. 2012. Centrosomes, chromosome instability (CIN) and aneuploidy. *Current Opinion in Cell Biology,* 24**,** 809-815.

VOCKERODT, M., MORGAN, S. L., KUO, M., WEI, W., CHUKWUMA, M. B., ARRAND, J. R., KUBE, D., GORDON, J., YOUNG, L. S., WOODMAN, C. B. & MURRAY, P. G. 2008. The Epstein-Barr virus oncoprotein, latent membrane protein-1, reprograms germinal centre B cells towards a Hodgkin's Reed-Sternberg-like phenotype. *J Pathol,* 216**,** 83-92.

VOCKERODT, M., SOARES, M., KANZLER, H., KUPPERS, R., KUBE, D., HANSMANN, M. L., DIEHL, V. & TESCH, H. 1998. Detection of clonal Hodgkin and Reed-Sternberg cells with identical somatically mutated and rearranged VH genes in different biopsies in relapsed Hodgkin's disease. *Blood,* 92**,** 2899-907.

VOCKERODT, M., WEI, W., NAGY, E., PROUZOVA, Z., SCHRADER, A., KUBE, D., ROWE, M., WOODMAN, C. B. & MURRAY, P. G. 2013. Suppression of the LMP2A target gene, EGR-1, protects Hodgkin's lymphoma cells from entry to the EBV lytic cycle. *J Pathol,* 230**,** 399-409.

WANG, D., LIEBOWITZ, D. & KIEFF, E. 1985. An EBV membrane protein expressed in immortalized lymphocytes transforms established rodent cells. *Cell,* 43**,** 831-840.

WANG, F., FLANAGAN, J., SU, N., WANG, L.-C., BUI, S., NIELSON, A., WU, X., VO, H.-T., MA, X.-J. & LUO, Y. 2012. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *The Journal of Molecular Diagnostics,* 14**,** 22-29.

WANG, F., GREGORY, C., SAMPLE, C., ROWE, M., LIEBOWITZ, D., MURRAY, R., RICKINSON, A. & KIEFF, E. 1990a. Epstein-Barr virus latent membrane protein

(LMP1) and nuclear proteins 2 and 3C are effectors of phenotypic changes in B lymphocytes: EBNA-2 and LMP1 cooperatively induce CD23. *Journal of virology,* 64**,** 2309-2318.

WANG, F., GREGORY, C., SAMPLE, C., ROWE, M., LIEBOWITZ, D., MURRAY, R., RICKINSON, A. & KIEFF, E. 1990b. Epstein-Barr virus latent membrane protein (LMP1) and nuclear proteins 2 and 3C are effectors of phenotypic changes in B lymphocytes: EBNA-2 and LMP1 cooperatively induce CD23. *J Virol,* 64**,** 2309-18.

WARBURG, O. 1924. ber den Stoffwechsel der Carcinomzelle. *Die Naturwissenschaften,* 12**,** 1131-1137.

WARNER, N., BURBERRY, A., FRANCHI, L., KIM, Y. G., MCDONALD, C., SARTOR, M. A. & NUNEZ, G. 2013. A genome-wide siRNA screen reveals positive and negative regulators of the NOD2 and NF-kappaB signaling pathways. *Sci Signal,* 6**,** rs3.

WASIL, L. R., TOMASZEWSKI, M. J., HOJI, A. & ROWE, D. T. 2013. The Effect of Epstein-Barr Virus Latent Membrane Protein 2 Expression on the Kinetics of Early B Cell Infection. *PLoS ONE,* 8**,** e54010.

WATANABE, A., INOKUCHI, K., YAMAGUCHI, H., MIZUKI, T., TANOSAKI, S., SHIMADA, T. & DAN, K. 2004. Near-triploidy and near-tetraploidy in hematological malignancies and mutation of the p53 gene. *Clin Lab Haematol,* 26**,** 25-30.

WATANABE, M., OGAWA, Y., ITO, K., HIGASHIHARA, M., KADIN, M. E., ABRAHAM, L. J., WATANABE, T. & HORIE, R. 2003. AP-1 Mediated Relief of Repressive Activity of the CD30 Promoter Microsatellite in Hodgkin and Reed-Sternberg Cells. *The American Journal of Pathology,* 163**,** 633-641.

WAXMAN, I.M., HOCHBERG, J., CAIRO, M.S. 2011. Non-Hodgkin Lymphoma. Nelson Textbook of Pediatrics. 19th ed. New York, NY: Elsevier Inc; 2011. pp. 1739–1745.

WEBER-MATTHIESEN, K., DEERBERG, J., POETSCH, M., GROTE, W. & SCHLEGELBERGER, B. 1995. Clarification of dubious karyotypes in Hodgkin’s disease by simultaneous fluorescence immunophenotyping and interphase cytogenetics (FICTION). *Cytogenetic and Genome Research,* 70**,** 243-245.

WENIGER, M. A., BARTH, T. F. & MOLLER, P. 2006a. Genomic alterations in Hodgkin's lymphoma. *Int J Hematol,* 83**,** 379-84.

WENIGER, M. A., MELZNER, I., MENZ, C. K., WEGENER, S., BUCUR, A. J., DORSCH, K., MATTFELDT, T., BARTH, T. F. E. & MÖLLER, P. 2006b. Mutations of the tumor suppressor gene SOCS-1 in classical Hodgkin lymphoma are frequent and associated with nuclear phospho-STAT5 accumulation. *Oncogene,* 25**,** 2679-2684.

WHITE, R. E., RÄMER, P. C., NARESH, K. N., MEIXLSPERGER, S., PINAUD, L., ROONEY, C., SAVOLDO, B., COUTINHO, R., BÖDÖR, C., GRIBBEN, J., IBRAHIM, H. A., BOWER, M., NOURSE, J. P., GANDHI, M. K., MIDDELDORP, J., CADER, F. Z., MURRAY, P., MÜNZ, C. & ALLDAY, M. J. 2012. EBNA3B-deficient EBV promotes B cell lymphomagenesis in humanized mice and is found in human tumors. *Journal of Clinical Investigation,* 122**,** 1487-1502.

WHITTLE, H. C., BROWN, J., MARSH, K., GREENWOOD, B. M., SEIDELIN, P., TIGHE, H. & WEDDERBURN, L. 1984. T-cell control of Epstein-Barr virus-infected B cells is lost during P. falciparum malaria. *Nature,* 312**,** 449-50.

WILLE, C. K., NAWANDAR, D. M., PANFIL, A. R., KO, M. M., HAGEMEIER, S. R. & KENNEY, S. C. 2012. Viral Genome Methylation Differentially Affects the Ability of BZLF1 versus BRLF1 To Activate Epstein-Barr Virus Lytic Gene Expression and Viral Replication. *Journal of Virology,* 87**,** 935-950.

WILLIAMS, S. A. & GREENE, W. C. 2005. Host factors regulating post-integration latency of HIV. *Trends Microbiol,* 13**,** 137-9.

WILSON, J. B., BELL, J. L. & LEVINE, A. J. 1996. Expression of Epstein-Barr virus nuclear antigen-1 induces B cell neoplasia in transgenic mice. *EMBO J,* 15**,** 3117-26.

WOELLMER, A., ARTEAGA-SALAS, J. M. & HAMMERSCHMIDT, W. 2012. BZLF1 Governs CpG-Methylated Chromatin of Epstein-Barr Virus Reversing Epigenetic Repression. *PLoS Pathogens,* 8**,** e1002902.

WOISETSCHLAEGER, M., STROMINGER, J. L. & SPECK, S. H. 1989. Mutually exclusive use of viral promoters in Epstein-Barr virus latently infected lymphocytes. *Proceedings of the National Academy of Sciences,* 86**,** 6498-6502.

WU, L., NAKANO, H. & WU, Z. 2006a. The C-terminal activating region 2 of the Epstein-Barr virus-encoded latent membrane protein 1 activates NF-kappaB through TRAF6 and TAK1. *J Biol Chem,* 281**,** 2162-9.

WU, W., FAN, Y. H., KEMP, B. L., WALSH, G. & MAO, L. 1998. Overexpression of cdc25A and cdc25B is frequent in primary non-small cell lung cancer but is not associated with overexpression of c-myc. *Cancer Res,* 58**,** 4082-5.

WU, Y., MARUO, S., YAJIMA, M., KANDA, T. & TAKADA, K. 2007. Epstein-Barr virus (EBV)-encoded RNA 2 (EBER2) but not EBER1 plays a critical role in EBV-induced B-cell growth transformation. *J Virol,* 81**,** 11236-45.

WU, Y. M., YAN, J., CHEN, L. L., SUN, W. L. & GU, Z. Y. 2006b. Infection frequency of Epstein-Barr virus in subgingival samples from patients with different periodontal status and its correlation with clinical parameters. *J Zhejiang Univ Sci B,* 7**,** 876-83.

XERRI, L., CARBUCCIA, N., PARC, P., HASSOUN, J. & BIRG, F. 1995. Frequent expression of FAS/APO-1 in Hodgkin's disease and anaplastic large cell lymphomas FAS/APO-1 in Hodgkin's disease and anaplastic large cell lymphomas. *Histopathology,* 27**,** 235-241.

XING, L. & KIEFF, E. 2007. Epstein-Barr virus BHRF1 micro- and stable RNAs during latency III and after induction of replication. *J Virol,* 81**,** 9967-75.

YAJIMA, M., KANDA, T. & TAKADA, K. 2005. Critical role of Epstein-Barr Virus (EBV)-encoded RNA in efficient EBV-induced B-lymphocyte growth transformation. *J Virol,* 79**,** 4298-307.

YAMAMOTO, R., NISHIKORI, M., KITAWAKI, T., SAKAI, T., HISHIZAWA, M., TASHIMA, M., KONDO, T., OHMORI, K., KURATA, M., HAYASHI, T. & UCHIYAMA, T. 2008. PD-1-PD-1 ligand interaction contributes to immunosuppressive microenvironment of Hodgkin lymphoma. *Blood,* 111**,** 3220-3224.

YATES, J. L., WARREN, N. & SUGDEN, B. 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. *Nature,* 313**,** 812-5.

YE, J., COULOURIS, G., ZARETSKAYA, I., CUTCUTACHE, I., ROZEN, S. & MADDEN, T. L. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics,* 13**,** 134.

YI, Z., WALLIS, A. M. & BISHOP, G. A. 2015. Roles of TRAF3 in T cells: many surprises. *Cell Cycle,* 14**,** 1156-1163.

YOKOYAMA, S., TSURUOKA, M., TAKEDA, O., DOZONO, H., NIIMI, S., TADA, A., ARIHIRO, T., TERAJIMA, Y., ITO, H., FUJITA, K. & ET AL. 1987. [Detection of c-myc mRNA in developing human placental villi by in situ hybridization with a biotin-labelled probe]. *Nihon Sanka Fujinka Gakkai Zasshi,* 39**,** 1651-2.

YOUNG, L., ALFIERI, C., HENNESSY, K., EVANS, H., O'HARA, C., ANDERSON, K. C., RITZ, J., SHAPIRO, R. S., RICKINSON, A., KIEFF, E. & ET AL. 1989.

Expression of Epstein-Barr virus transformation-associated genes in tissues of patients with EBV lymphoproliferative disease. *N Engl J Med,* 321**,** 1080-5.

YOUNG, L. S., DAWSON, C. W., CLARK, D., RUPANI, H., BUSSON, P., TURSZ, T., JOHNSON, A. & RICKINSON, A. B. 1988. Epstein--Barr Virus Gene Expression in Nasopharyngeal Carcinoma. *Journal of General Virology,* 69**,** 1051-1065.

YOUNG, L. S., DAWSON, C. W. & ELIOPOULOS, A. G. 2000. The expression and function of Epstein-Barr virus encoded latent genes. *Mol Pathol,* 53**,** 238-47.

YOUNG, L. S., LAU, R., ROWE, M., NIEDOBITEK, G., PACKHAM, G., SHANAHAN, F., ROWE, D. T., GREENSPAN, D., GREENSPAN, J. S., RICKINSON, A. B. & ET AL. 1991. Differentiation-associated expression of the Epstein-Barr virus BZLF1 transactivator protein in oral hairy leukoplakia. *J Virol,* 65**,** 2868-74.

YOUNG, L. S. & RICKINSON, A. B. 2004. Epstein-Barr virus: 40 years on. *Nat Rev Cancer,* 4**,** 757-68.

YOUNG, L. S., YAO, Q. Y., ROONEY, C. M., SCULLEY, T. B., MOSS, D. J., RUPANI, H., LAUX, G., BORNKAMM, G. W. & RICKINSON, A. B. 1987. New type B isolates of Epstein-Barr virus from Burkitt's lymphoma and from normal individuals in endemic areas. *J Gen Virol,* 68 (Pt 11)**,** 2853-62.

YU, R., LU, W., CHEN, J., MCCABE, C. J. & MELMED, S. 2003. Overexpressed Pituitary Tumor-Transforming Gene Causes Aneuploidy in Live Human Cells. *Endocrinology,* 144**,** 4991-4998.

YUAN, J., CAHIR-MCFARLAND, E., ZHAO, B. & KIEFF, E. 2006. Virus and cell RNAs expressed during Epstein-Barr virus replication. *J Virol,* 80**,** 2548-65.

ZALANI, S., HOLLEY-GUTHRIE, E. & KENNEY, S. 1996. Epstein-Barr viral latency is disrupted by the immediate-early BRLF1 protein through a cell-specific mechanism. *Proc Natl Acad Sci U S A,* 93**,** 9194-9.

ZHANG, J., GRUBOR, V., LOVE, C. L., BANERJEE, A., RICHARDS, K. L., MIECZKOWSKI, P. A., DUNPHY, C., CHOI, W., AU, W. Y., SRIVASTAVA, G., LUGAR, P. L., RIZZIERI, D. A., LAGOO, A. S., BERNAL-MIZRACHI, L., MANN, K. P., FLOWERS, C., NARESH, K., EVENS, A., GORDON, L. I., CZADER, M., GILL, J. I., HSI, E. D., LIU, Q., FAN, A., WALSH, K., JIMA, D., SMITH, L. L., JOHNSON, A. J., BYRD, J. C., LUFTIG, M. A., NI, T., ZHU, J., CHADBURN, A., LEVY, S., DUNSON, D. & DAVE, S. S. 2013. Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A,* 110**,** 1398-403.

ZHANG, Q., HONG, Y., DORSKY, D., HOLLEY-GUTHRIE, E., ZALANI, S., ELSHIEKH, N. A., KIEHL, A., LE, T. & KENNEY, S. 1996. Functional and physical interactions between the Epstein-Barr virus (EBV) proteins BZLF1 and BMRF1: Effects on EBV transcription and lytic replication. *J Virol,* 70**,** 5131-42.

ZIMBER-STROBL, U., KREMMER, E., GRASSER, F., MARSCHALL, G., LAUX, G. & BORNKAMM, G. W. 1993. The Epstein-Barr virus nuclear antigen 2 interacts with an EBNA2 responsive cis-element of the terminal protein 1 gene promoter. *EMBO J,* 12**,** 167-75.

ZIMBER, U., ADLDINGER, H. K., LENOIR, G. M., VUILLAUME, M., KNEBEL-DOEBERITZ, M. V., LAUX, G., DESGRANGES, C., WITTMANN, P., FREESE, U. K., SCHNEIDER, U. & ET AL. 1986. Geographical prevalence of two types of Epstein-Barr virus. *Virology,* 154**,** 56-66.

ZINDY, F., EISCHEN, C. M., RANDLE, D. H., KAMIJO, T., CLEVELAND, J. L., SHERR, C. J. & ROUSSEL, M. F. 1998. Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization. *Genes Dev,* 12**,** 2424-33.

ZUO, J., CURRIN, A., GRIFFIN, B. D., SHANNON-LOWE, C., THOMAS, W. A., RESSING, M. E., WIERTZ, E. J. H. J. & ROWE, M. 2009. The Epstein-Barr Virus G-Protein-Coupled Receptor Contributes to Immune Evasion by Targeting MHC Class I Molecules for Degradation. *PLoS Pathogens,* 5**,** e1000255.

ZUR HAUSEN, A., BRINK, A. A., CRAANEN, M. E., MIDDELDORP, J. M., MEIJER, C. J. & VAN DEN BRULE, A. J. 2000. Unique transcription pattern of Epstein-Barr virus (EBV) in EBV-carrying gastric adenocarcinomas: expression of the transforming BARF1 gene. *Cancer Research,* 60**,** 2745-2748.

ZUR HAUSEN, H. & DE VILLIERS, E.-M. 2015. Reprint of: Cancer "Causation" by Infections—Individual Contributions and Synergistic Networks. *Seminars in Oncology,* 42**,** 207-222.

ZYSS, D. & GERGELY, F. 2009. Centrosome function in cancer: guilty or innocent? *Trends in Cell Biology,* 19**,** 334-346.