Identification of cooperating oncogenic lesions in Myc-driven lymphoma

Marcus Patrick Henry Lefebure

Submitted in total fulfillment of the requirements of the degree of Doctor of Philosophy

September 2017

Sir Peter MacCallum Department of Oncology, University of Melbourne

Peer reviewed publications arising from this thesis

Lefebure, M., Tothill, R.W., Kruse, E., Hawkins, E.D., Shortt, J., Matthews, G.M., Gregory, G.P., Martin, B.P., Kelly, M.J., Todorovski, I.*, et al.* (2017). Genomic characterisation of Emu-Myc mouse lymphomas identifies Bcor as a Myc co-operative tumour-suppressor gene. Nature communications *8*, 14581.

Abstract

MYC is a potent oncogene that is deregulated in nearly 50% of all human malignancies and as such, is considered an attractive molecular target for inhibition. However, MYC is rarely mutated, has no enzymatic activity that can be pharmacologically exploited and is expressed by normal cells leading to the current view that MYC is “undruggable” and indeed, its pharmacological inhibition has proved elusive. Therefore, discovering genes and pathways that interact with oncogenic MYC signalling and identifying them as potential therapeutic targets in cancers with ectopic MYC expression is of high clinical importance. The E*μ-Myc* mouse has been utilised extensively as a faithful model of MYC-driven B cell lymphomagenesis. The E*μ-Myc* transgene mimics the t(8;14) translocation apparent in Burkitt’s lymphoma, where ectopic Myc expression is driven by the E*μ*- (*IGH*) promoter elements. Despite being driven by a single oncogene, E*μ-Myc* lymphomas demonstrate remarkable heterogeneity indicating that the pathway to frank clonal neoplasia relies on oncogenic Myc signalling and the acquisition of at least one other mutation that cooperates with Myc. Hence, the E*μ-Myc* model is a powerful tool in identifying MYC-cooperative genes and pathways. However, to date there has only been partial characterisation of the secondary, tertiary and quaternary mutations that can cooperate with Myc in driving E*μ-Myc* lymphomagenesis.

To identify somatically acquired Myc-cooperative lesions, massive-parallel sequencing was applied to spontaneous E*μ-Myc* B cell lymphomas. Whole genome sequencing

was used to map three copies of the E*μ-Myc* transgene to chromosome 19 in the germline corresponding with an adjacent chromosome 19 segmental copy number gain. The chromosome 19 amplicon is in a region syntenic to an oncogenic region frequently amplified in human B cell malignancies. The chromosome 19 amplicon was demonstrated to undergo additional somatic gain in 50% of E*μ-Myc* lymphoma. In addition to the identification of mutations in genes already implicated in E*μ-Myc* lymphoma (*Trp53*, *Cdkn2a*, *Nras*, *Kras*), whole exome sequencing identified high frequency protein truncating mutations in *Bcl6*-co-repressor (*Bcor*). Furthermore, co-occurring tertiary driver lesions involving *Cdkn2a* ($p19^{ARF}$) deletion and either *Bcor* or *Ras* mutations were identified in clonal E*μ-Myc* lymphomas. RNAi and CRISPR-Cas9 mediated knockdown/knockout of *Bcor* in E*μ-Myc* foetal liver hematopoietic stem cells reconstituted into recipient mice demonstrated significantly reduced latency of disease onset, validating *Bcor* as a tumor suppressor gene in the Eμ-*Myc* model. Gene-expression profiling of these E*μ-Myc* tumours with forced *Bcor*-loss identified a reliable signature of *Bcor* loss that was distinct to *Trp53* mutation signatures and was redolent of Tgfβ-pathway activation signature.

The E*μ-Myc* model of lymphoma has been heavily utilised but never fully genomically characterised until now. By applying next generation sequencing technology to a first generation animal model of cancer, this thesis challenges several persisting assumptions made about E*μ-Myc* lymphoma. Firstly, data herein suggests that both oncogenic Myc expression along with the chromosome 19 amplification is the initiating driver event in E*μ-Myc* lymphoma. This has obvious implications for the conclusions

drawn in many publications predicated on the assumption that ectopic Myc expression is the exclusive initiating oncogenic lesion in E*μ-Myc* lymphoma. Secondly, the discovery that homozygous deletion of *Cdkn2a* does not totally attenuate selective pressure for the acquisition of tertiary driver mutations indicates the significance of *Cdkn2a* deletion in E*μ-Myc* lymphoma is overestimated. Thus, deductions made about the cooperative mechanism between *CDKN2A* and how it opposes proliferative MYC-signalling in human neoplastic transformation may need to be revisited. Finally, the identification of biologically functional high frequency *Bcor* mutations in E*μ-Myc* lymphoma has defined a novel pathway that is potentially capable of restraining oncogenic MYC activity. That Bcor inactivation is reminiscent of Tgfβ-pathway enhancement is suggestive of perturbation of the oncogene induced senescence pathway. If this is the pathway through which Bcor exerts its tumour suppressive activity then it is feasible that dissection of this pathway will lead to the identification of novel therapeutic targets that can be selectively exploited in human malignancies in which MYC is oncogenically deregulated.

Declaration

This is to certify that:

A) This thesis comprises only my original work towards this PhD except where indicated in the preface and the text.
B) Due acknowledgement has been made in the text to all other material used.
C) This thesis is less than 100,000 words in length, exclusive of tables, figures and bibliographies.

Marcus Lefebure

The Sir Peter MacCallum Department of Oncology

Peter MacCallum Cancer Centre,

Parkville, Victoria, Australia

And

The University of Melbourne

Parkville, Victoria, Australia

Preface

This thesis acknowledges contributions from the following persons who performed experiments as listed below:

Figure 3.03 – FISH performed by Dr. Meghan Wall

Figure 3.06 – FACS analysis of Pd-l1 expression performed by Dr. Gareth Gregory

Figure 3.07 – Ruxolitinib treatment of cell lines and western blot performed by Dr. Gareth Gregory

Figure 3.16A – E*μ-Myc* cell lines transduced and sorted by Mr. Stuart Craig

Figure 4.01 – Design and testing of *Bcor*-targeting shRNA performed by Dr. Elizabeth Kruse

Figure 4.06 – E*μ-Myc* foetal liver cells transduced and validated by Dr. Lev Kats

Figure 4.08 – *in vitro* competition assay and western blots performed by Mr. Ben Martin and Ms. Madison Kelly

Next generation sequencing – Dr. Gisela Mir Arnau, Mr. Timothy Semple, Mr. Timothy Holloway and the genomics core facility at The Peter MacCallum Cancer Centre performed the library preparation and re-sequencing in all cases.

Acknowledgements

I would like to thank Dr. Edwin Hawkins and Prof. Ricky Johnstone for supervising me throughout my PhD research and during the write-up phase. It is not an understatement to say that without the support of Ricky, Edwin and Libby this thesis would not have been finished. In particular, the help Edwin and Libby provided me over skype from London, in their offices at WEHI and in their lounge-room during weeknights has been greatly received and it is something that I will not forget. I would like to thank Dr. Richard Tothill and Dr. Geoff Matthews for their supervisorial support throughout my PhD.

My PhD was funded by the Leukaemia Foundation of Australia and Arrow Bone Marrow.

I am exceedingly grateful for the technical assistance, scientific advice and friendship that I have received from members of the Johnstone Laboratory: Mark Bishton, Leonie Cluse, Ben Martin, Andrea Newbold, Eva Vidacs, Gareth Gregory, Simon Hogg, Lev Kats, Madison Kelly, Sang-Kyu Kim, Deborah Roseingrave and Stephin Vervoort. I would like to thank the flow cytometry team, genomics core, bioinformatics group and animal house staff for their technical assistance that went into producing this thesis.

Thanks must also go to my partner Priya Stewart and the friends and family who provided moral support throughout my PhD and especially over the last six months.

List of abbreviations used

Human and mouse genes and proteins are differentiated in accordance with Nature Publishing Group guidelines. Briefly, human genes/proteins are entirely capitalised and mouse genes/proteins have only the first letter capitalised. Genes are italicised and proteins are not. For example:

	Human	Mouse
Protein	XYZ	Xyz
Gene	*XYZ*	*Xyz*

Gene and amino acid abbreviations are not listed here, as the abbreviations used are considered standard. Gene abbreviations used in this thesis can be found here.

http://www.informatics.jax.org/mgihome/nomen/

http://www.gene.ucl.ac.uk/nomenclature

Abbreviation	Definition
-/-	knockout/null
+/-	heterozygous
ABC	activated B cell-like
ACK	ammonium/chloride/potassium
ADTEx	aberrant detection in tumour exome
AKMDMEM	Anne Kelso's modified Dulbecco's modified Eagle's medium
AML	acute myeloid leukaemia
ANOVA	analysis of variance
BAC	bacterial artificial chromosome
B-ALL	B cell acute lymphoblastic leukaemia
BFP	blue fluorescent protein

BL	Burkitt's lymphoma
BLAST	Basic local alignment search tool
bona fide	(Latin) in good faith
bp	base pair
C	Celsius
CD	cluster of differentiation
cDNA	complementary DNA
CDS	coding DNA sequence
CGC	cancer gene census
Chr	chromosome
CLL	chronic lymphoblastic leukaemia
CMML	chronic myelomonocytic leukaemia
CN-	cytogenetically normal
CNV	copy number variation(s)
COSMIC	catalogue of somatic mutations in cancer
CSHL	Cold Spring Harbor Laboratories
dbSNP	single nucleotide polymorphism database
DDR	DNA damage response
de novo	(Latin) of new
DH	double hit
DLBCL	diffuse large B cell lymphoma
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
et al	(Latin) and others
FACS	fluorescent-activated cell sorting
FC	fold change
FDR	false discovery rate
FISH	fluorescent *in situ* hybridisation
g	gram(s)
g	gravity
GC	germinal centre
GCB	germinal centre B cell-like

GEMM	genetically engineered mouse model
GFP	green fluorescent protein
HSC	hematopoietic stem cell(s)
Ig	immunoglobulin
IGV	integrated genome viewer
in situ	(Latin) in place
in vitro	(Latin) in glass
in vivo	(Latin) in life/organism
InDel	small insertion(s) or deletion(s)
IQR	inter-quartile range
KD	knockdown
KO	knockout
L	litre
LB	lysogeny broth
LN	lymph node
logFC	log fold change
LOH	loss of heterozygosity
M	molar
m	milli
MAPK	mitogen activated protein kinase
MDS	myelodysplatic syndrome
MEF	mouse embryonic fibroblast
miRNA	micro RNA
ml	millilitre
MOI	multiplicity of infection
Mo-MLB	Moloney murine leukaemia virus
MPS	massive parallel sequencing
mRNA	messenger RNA
MSCV	murine stem cell virus
NGS	next generation sequencing
OFCD	oculofaciocardiodental
OIS	oncogene induced senescence
p-	phospho-

PAM	protospacer-adjacent motif
PBS	phosphate buffered saline
PcG	polycomb group
PCR	polymerase chain reaction
Ph+ B-ALL	Philadelphia positive B-ALL
PMBCL	primary mediastinal B cell lymphoma
PRC	polycomb repressive complex
QRT-PCR	quantitative real-time polymerase chain reaction
RFP	red fluorescent protein
RIPA	radioimmunoprecipitation assay
RNA	ribonucleic acid
RNAi	RNA interference
RNA-seq	RNA sequencing
ROS	reactive oxygen species
SEM	standard error of the mean
sgRNA	single-guide RNA
sh	short-hairpin
SH	single hit
shRNA	short-hairpin RNA
SNP	single nucleotide polymorphism
SNV	single nucleotide variant(s)
T-ALL	T cell acute lymphoblastic leukaemia
TAM-seq	targeted amplicon sequencing
TBE	tris/borate/EDTA solution
UTR	untranslated region
VAF	variant allele frequency
WB	western blot
WBC	white blood cell(s)
WBCC	WBC count
WEHI	Walter and Eliza Hall Institute
WES	whole exome sequencing
WGS	whole genome sequencing
WHO	World Health Organisation

WT	wild type
μ	micro
ρ	pico

Table of Contents

1 Introduction

1.1 Haematological malignancies

Haematological malignancies represent a broad spectrum of diseases linked by their association to normal cellular counterparts arising in the haematopoetic and lymphoid tissues (Sabattini et al., 2010). Combined, haematological malignancies contribute to approximately 9% of all cancers and affect the population in economically developed countries without prejudice, as they are the fourth most common cancer diagnosis in both men and women (Smith et al., 2011). These malignancies are commonly grouped into four main categories as defined by the USA's Surveillance Epidemiology and End Results Program, The UK's National Cancer Intelligence Network and the World Health Organisation (WHO); Hodgkin's lymphoma (HL), non-Hodgkin's lymphoma (NHL), leukaemia and multiple myeloma (MM) (Jemal et al., 2011; Jemal et al., 2010; Westlake, 2009). MM are neoplastic disorders characterised by malignant antibody producing cells (known as plasma cells). MM is associated with osteoporosis, hypocalcaemia, renal dysfunction, peripheral neuropathy and pan-cytopaenias (Hideshima et al., 2007; Kyle and Rajkumar, 2008). Leukaemia is characterised by abnormal oncogenic lymphoblastic or myeloblastic infiltrate leading to pan-cytopaenias, increased risk of infection and elevated peripheral white blood cell counts (Estey, 2009; Smith et al., 2011; Zenz et al., 2010). Leukaemia can be further stratified into four main tumour subtypes; acute lymphoblastic leukaemia (ALL), chronic lymphoblastic leukaemia (CLL), acute myeloid leukaemia (AML) and chronic myeloid leukaemia (CML). ALL and CLL are malignant forms of the lymphoid lineage while AML and CML originate from the myeloid lineage (Smith et al., 2011). HL is a disease of the lymphatic system affecting specific B lymphocytes termed Hodgkin and Reed-Sternberg (HRS)

cells (Kuppers, 2009). NHL is made up of a broad spectrum of diseases that directly effect the lymphatic system and have little in common with one another except that they cannot be classified as HL due to the absence of pathogenic HRS cells (Smith et al., 2011). NHL can include T cell lymphomas, NK cell lymphomas and the most common type of haematological malignancy in adults, B cell lymphoma. Despite being the fourth most common cancer diagnosis in economically developed nations, haematological malignancies are over-represented in mortality rates, as they account for the second most common cause of cancer death (Howlader N, 2016; Smith et al., 2011). Chronic diseases such as MM, CLL, CML and the more indolent NHL subtypes are considered incurable. These diseases slowly progress over long periods of time despite management with conventional chemotherapies and are capable of transforming to more aggressive neoplasms (Coiffier, 2001; Rossi et al., 2008; Sehn, 2016). Typically, an aggressive regimen of multi-treatment chemotherapy is the gold-standard therapy for high-grade acute lymphoma/leukaemia, and while this is considered potentially curative, the rate of relapse is high (Coiffier, 2001). Primary chemotherapy resistant or relapsed haematological malignancies are associated with an incredibly poor prognosis as the treatment options thereafter become increasingly limited and potentially toxic (Bassan and Hoelzer, 2011; Pui et al., 2008; Roschewski et al., 2014). In all cases, there is a clear unmet medical need for the development of novel therapeutics that can improve upon current rates of remission, relapse and resistance.

Due to the high prevalence of these diseases and the subsequent health burden they account for, significant effort has been directed towards dissecting key

mutations/oncogenic drivers associated with haematological malignancies. This has included early forms of mutation analysis such as 2-D fractionation (Sanger et al., 1965) and the capillary-based Sanger sequencing (Sanger et al., 1977) to current techniques that include next generation sequencing approaches (discussed later in section 1.6). These analyses have revealed a range of frequently occurring types of mutations such as missense mutations, insertions or deletions (InDels), copy number variations (CNV) and translocations. Using these genomic analysis techniques, significant genomic and genetic heterogeneity is observed in haematological malignancies. Mutations in the *Myc* oncogene and components of this pathway are one of the most frequently observed oncogenic drivers in haematological malignancies.

1.2 Cancer Genetics

Cancer is, in essence, a genetic disease. It is widely understood that cancers develop from the accumulation of multiple mutations that work provide a selective advantage to the affected cell, thereby contributing the initiation and progression of malignancy (Hartwell, 1992; Loeb, 1991; Stratton, 2011). However, the majority of mutations acquired in a cancerous cell are inconsequential, termed ‘passenger’ mutations, with only a minority of mutations capable of endowing a selective advantage to the cell (termed ‘driver’ mutations) (Bozic et al., 2010; Vogelstein et al., 2013). Driver mutations affecting oncogenes and tumour suppressor genes are responsible for tumourigenesis.

1.2.1 Oncogenes

Oncogenes are mutated in ways that render them constitutively active or active under conditions in which the wild type gene is not. Typically, activating mutations in oncogenes occur through chromosomal translocation (Konopka et al., 1985; Tsujimoto et al., 1985b), gene amplifications (Bartlett et al., 2001) or from subtle mutations affecting crucial residues that increase activity or increase stability of the gene product (Rodenhuis, 1992). Mutations in oncogenes are considered dominant as a gain of function mutation in only a single allele is required for activation and/or loss of regulation of the oncogene product (Rivenbark, 2017).

1.2.2 Tumour suppressor genes

Tumour suppressor genes are targeted by mutations that invariably lead to the inability of the encoded protein to perform its function. Inactivating events can occur through missense mutations that are deleterious to the activity and/or stability of the gene product, mutations that result in a premature truncation of the gene product and from framseshift insertions and deletions (Bozic et al., 2010; Rivenbark, 2017; Vogelstein et al., 2013). Tumour suppressor loss of function generally relies on mutational inactivation or loss of both alleles and are resultantly thought of as recessive (Knudson, 2002).

1.3 *MYC*

Perhaps more than any other gene, aberrations in the *MYC* gene is most closely associated with haematological neoplastic development. The *MYC*-transcription factor family is comprised of 4 genes – *c-MYC, L-MYC, N-MYC* and *S-MYC* with a fifth gene, *B-MYC*, encoding a MYC protein with N-terminal homology to the other MYC-family members, but lacking critical C-terminal elements. MYC family genes are ubiquitously expressed, with MYC proteins present at physiological levels in most proliferating cells (Felsher and Bishop, 1999b; Flores et al., 2004; Jain et al., 2002; Pelengaris et al., 1999; Sodir et al., 2011; Soucek et al., 2008; Verbeek et al., 1991). Of the *MYC*-family genes, *N-MYC, L-MYC* and *c-MYC* have been implicated in tumourigenesis, with *c-MYC* perhaps being the most well studied and from hereon is referred to as *MYC* (Adhikary and Eilers, 2005). *MYC* aberrations occur frequently in haematological malignancies (Adhikary and Eilers, 2005; Meyer and Penn, 2008; Vita and Henriksson, 2006). Enhanced MYC expression is involved in almost every aspect of cancer cell biology by promoting; cell proliferation and growth (Eilers et al., 1991; Iritani and Eisenman, 1999), vasculogenesis and angiogenesis (Baudino et al., 2002), genomic instability (Felsher and Bishop, 1999b) and metastasis (Pelengaris et al., 2002). Given that unchecked MYC expression is a potent oncogenic stimulus, MYC expression is stringently regulated at every possible checkpoint from transcriptional control to protein stabilisation. In addition to its oncogenic effects, MYC intrinsically promotes cancer protective mechanisms such as activation of cellular senescence (Reimann et al., 2010) and activation of apoptosis (Dansen et al., 2006). MYC has also recently been

demonstrated to be exquisitely linked to the number of divisions that B cell populations progress through before immune responses cease (Heinzel et al., 2017).

1.3.1 MYC driven transcriptional regulation

MYC-family proteins have been estimated to bind to 25,000 sites in the human genome indicating that a significant portion of all genes are activated or repressed by MYC-proteins (Adhikary and Eilers, 2005). The number of MYC binding sites far outnumbers the number of MYC molecules in a cell, given that proliferating fibroblast cells are estimated to contain 1000 MYC molecules (Moore et al., 1987). In normal cells, the occupation of these binding sites must therefore be regulated during cell cycle, dependent on signals and/or in cooperation with other transcription factors (Fernandez et al., 2003; Orian et al., 2003; Waters et al., 1991). Additionally, MYC-binding sites also lie within non-coding regions or genes that do not show evidence of MYC-mediated transcriptional regulation (Fernandez et al., 2003). Taken together, this explains the relatively low number of directly MYC-regulated protein-coding genes, which account for 10-15% of the genome (Adhikary and Eilers, 2005). Moreover, cells that express higher levels of MYC do not show a change to their gene expression profile, but rather see an amplification of the pre-existing transcriptional profile (Lin et al., 2012; Nie et al., 2012). This suggests that MYC activity is heavily context dependent and explains the vastly different gene-expression profiles between different tumour cells that each express high levels of MYC.

MYC-targets are implicated in most aspects of cellular physiology; however, certain aspects are particularly vulnerable to MYC because of the feedback mechanisms and pathways associated with MYC. These include cell cycle progression, protein biosynthesis, energy metabolism and mitochondrial function, which are all biological aspects that are capable of contributing in tumourogenic transformation when deregulated (Luscher and Vervoorts, 2012). Pol II is responsible for transcribing many MYC-target genes, but MYC is known to directly regulate ribosome biogenesis and upregulation of translational machinery through transcriptional activation of Pol I and Pol III, respectively (Gomez-Roman et al., 2003; Grandori et al., 2005). A second tier of MYC target regulation is a complex micro RNA (miRNA) network. MYC has a predominantly suppressive effect on the target miRNA, which includes putative tumour suppressor genes such as *miR-15a/16-1*, *miR34a* and *let-7* family members (Chang et al., 2008). However, MYC is still capable of transcriptionally upregulating oncogenic miRNA as well, for example, the *miR-17-92* cluster products are proliferative and commonly overexpressed in B cell lymphomas (He et al., 2005).

1.3.2 MYC and malignancy

The effects of MYC on cell cycle, metabolism and ribosomal biogenesis make it an obvious candidate oncogene capable of driving malignant transformation. However, *MYC* overexpression alone is insufficient in initiating tumourigenesis (Dang, 1999; Evan et al., 1992; Spencer and Groudine, 1991). This infers that there is a requirement for acquiring cooperating secondary oncogenic lesions in addition to *MYC* to realise full neoplastic potential. In addition to the regulation of MYC itself, MYC-targets are often

co-regulated by independent opposing factors outside of MYC-driven cellular processes (Meyer and Penn, 2008). MYC is known to have a pro-apoptotic effect as well as generating reactive oxygen species (ROS) and replicative stresses that can strongly induce a cancer-protective DNA damage response (DDR) (Hermeking and Eick, 1994; Ray et al., 2006; Vafa et al., 2002). An emerging neoplastic clone must circumvent these inherent MYC-driven "fail safes" to harness the full oncogenic potential of MYC.

1.3.2.1 The role of MYC in apoptosis

MYC has been demonstrated to have context dependent roles in apoptosis as a means to safeguard against malignant transformation that must be overcome to initiate neoplasia. Firstly, the apoptotic activity of MYC is largely underpinned by its ability to initiate the DDR and engage the $p14^{ARF}$-MDM2-TP53 tumour suppressor axis (Lowe et al., 2004). Proposed mechanisms for Myc in activating DDR includes increased amounts of ROS and formation of aberrant DNA replication intermediates (Vafa et al., 2002). MYC promotes $p14^{ARF}$ (an alternate open reading frame protein product, encoded by *CDKN2A*) expression, which cooperates with ATM and inhibits MDM2, which it itself is a TP53 antagonist, thus releasing TP53 (Kamijo et al., 1998; Quelle et al., 1995; Zindy et al., 1998) (Figure 1.01). Increased TP53 activity provides a strong tumour suppressor response capable of counteracting the proliferative effect of MYC by activating pro-apoptotic effector molecules like PUMA and NOXA (Nakano and Vousden, 2001; Villunger et al., 2003). As a result, the $p14^{ARF}$-MDM2-TP53 pathway has substantive selective pressure placed on it to become deregulated during MYC-driven neoplastic progression (Dang et al., 2005). Additionally, BIM, a BH3-only protein,

is upregulated by MYC in a TP53-independent manner (Egle et al., 2004). Loss of BIM was shown to reduce pressure for $p14^{ARF}$/TP53 inactivation during Myc-driven malignant transformation. Genetic investigations into spontaneously arising experimental models of MYC-transformed tumours and human lymphoma often reveal accompanying perturbations in the $p14^{ARF}$-MDM2-TP53 pathways as a putative means to bypass MYC-induced apoptosis (Eischen et al., 1999). In support of this observation, genetic crosses of *Myc*-transgenic mice with *Trp53*$^{-/-}$, *Cdkn2a*$^{-/-}$ (encoding $p19^{ARF}$, the mouse orthologue of $p14^{ARF}$) or *Atm*-deficient mice resulted in accelerated tumourigenesis, likely by overcoming Myc-induced apoptosis (Reimann et al., 2007; Schmitt et al., 1999). Moreover, knockdown or knockout of the *Trp53*-dependent apoptosis effectors *Puma*, and to a lesser extent *Noxa,* also accelerated *Myc*-driven tumourigenesis (Hemann et al., 2004; Michalak et al., 2009). Conversely, manipulation of the Trp53-independent pathways through overexpression of anti-apoptotic *Bcl2* family members or deletion of *Bim* in pre-neoplastic *Myc*-overexpressing cells augmented the latent oncogenic activity of *Myc*, consequently reducing the pressure to mutate *Trp53* or delete $p19^{ARF}$ in *Myc*-driven tumourigenesis (Egle et al., 2004).

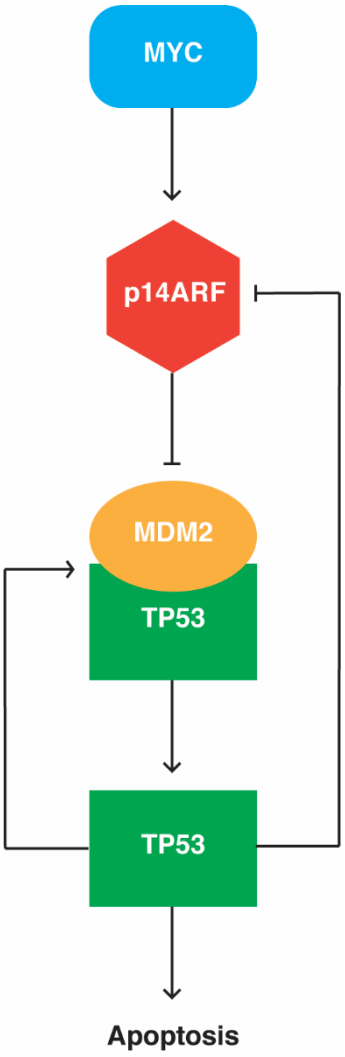

MYC
p14ARF
MDM2
TP53
TP53
Apoptosis

Figure 1.01 – MYC-induced apoptotic pathway.

MYC expression induces expression of $p14^{ARF}$ ($p19^{ARF}$ in mice). $p14^{ARF}$ binds to MDM2 neutralising it and releasing TP53. TP53 can now mediate TP53-dependent transcription resulting in apoptosis. In turn, MDM2 is a TP53 responsive gene, while TP53 also negatively regulates $p14^{ARF}$ expression. This pathway is under significant selective pressure to become inactivated in response to oncogenic MYC-signalling, thus allowing the pre-neoplasm to circumvent the apoptotic response triggered by unchecked MYC expression.

1.3.2.2 The role of MYC in senescence

Oncogene induced senescence (OIS) provides yet another barrier to MYC-driven transformation (Braig et al., 2005). OIS is the permanent shift of a cell to an irreversible state of cell cycle arrest accompanied by characteristic epigenetic, biochemical and morphological alterations (Braig et al., 2005). MYC-induced DDR promotes OIS that is dependent on the histone methyltransferase Suv39h1 that epigenetically silences S-phase genes (Reimann et al., 2010). As expected, *Suv39h1* deletion accelerated *Myc*-driven transformation by perturbing the in-built senescence program (Reimann et al., 2010).

1.3.3 The role of MYC in B cell malignancies

MYC overexpression is a hallmark of Burkitt's lymphoma (BL) and defines a subset of diffuse large B cell lymphoma (DLBCL) and B cell ALL (B-ALL), tethering the progression and pathology of these B lineage malignancies to dysfunctional MYC pathways. Through translocation, rearrangement or amplification, *MYC* is overexpressed in 100% of BL (Boxer and Dang, 2001; Frost et al., 2004), 10-25% of DLBCL (Akasaka et al., 2000; Frost et al., 2004; Kramer et al., 1998; Vitolo et al., 1998) and 47-52% of B-ALL (Burmeister et al., 2005; Miranda Peralta et al., 1991). BL, DLBCL, and B-ALL represent B lineage tumours that are among the most commonly diagnosed haematological tumours in both adolescents and adults (Miller et al., 1995; Morton et al., 2006). These B lineage neoplasms are almost always treated with intensive chemotherapy and/or radiotherapy with curative intent. However, the diseases commonly relapse as refractory cancers with an unfavourable prognosis

(Brugiatelli et al., 1997; Pui and Evans, 1998; Pui et al., 2008; Roschewski et al., 2014). This differential response rate within and between tumour subtypes is strongly suggestive of heterogeneous disease, despite a large number of them presenting with ectopic MYC expression. Supporting this, outside of aberrant MYC expression, each specified B lineage tumour harbours a unique genomic landscape, shaped by mutation events and genomic instability (Alexandrov et al., 2013a; Alexandrov et al., 2013b; Alizadeh et al., 2000; Lohr et al., 2012; Love et al., 2012b; Mullighan et al., 2008). For the development of effective treatments capable of eliciting durable responses, both the intra- and intertumour heterogeneity of B lineage malignancies should be better understood. As demonstrated in models of "oncogene addiction", identifying the crucial mutations in a cancer also identifies a potential weakness that is capable of being exploited by targeted therapy (Sharma and Settleman, 2007; Weinstein, 2000, 2002). As discussed below, understanding of the unique genomic drivers of these malignancies has provided insight into recurrently activated or suppressed genes and pathways critical for B cell transformation and tumour maintenance. Identification of unique genes implicated in B lineage malignancies provides a suite of novel molecular targets that actively drive tumourigenesis and can therefore be amenable to targeted therapeutic inhibition.

1.3.3.1 MYC aberrations and secondary mutations in BL

B cell neoplasms are extremely diverse as defined by their genomic landscapes, biology and histology. As tissue is readily available from patients with haematological malignancies the underlying biology, histopathological and molecular pathogenesis of B

cell lineage cancers is relatively well characterised and defined by The WHO (Campo et al., 2011). Indeed, an increasing number of B cell lineage malignancies are identified based on recurrent cytogenetic and molecular abnormalities, such as chromosomal translocations (Campbell, 2005). Not only does this facilitate the correct identification and prognostication of individual tumours, but also provides evidence of the driving oncogenic events required for neoplastic transformation. An example of this is exemplified in BL, an aggressive germinal centre B cell lymphoid malignancy (Schmitz et al., 2014; Schmitz et al., 2012). BL is defined cytogenetically by chromosomal translocations that juxtapose *MYC* (8q24) to either *IGH* (14q32) or immunoglobulin light chain promoters Igλ (22q11) or Igκ (2p12). Thus, ectopic expression of *MYC* is driven by an aberrant promoter region (Boerma et al., 2009). BL cases that present without identifiable *MYC* rearrangements are rare. However, these cases retain a BL-typical gene-expression profile that indicate cryptic *MYC* rearrangements or deregulation by other means (Ferry, 2006). Normally, BL manifests with a classically defined t(8;14)(q24;q32) karyotype. However, secondary and tertiary mutations and cooperating oncogenic lesions have also been identified (Love et al., 2012a). Extracted *MYC* DNA from BL samples revealed hotspot mutations that directly and indirectly increased the oncogenic potential by stabilising MYC (Bahram et al., 2000; Hemann et al., 2005). Mutations around Thr58 prevent phosphorylation by GSK-3b resulting in increased protein stability (Bahram et al., 2000). Thr58 mutants also reduce the capacity of p53 to induce apoptosis by downregulating MYC induced BIM expression (Hemann et al., 2005). Molecular evasion of TP53-mediated apoptosis is important in BL, as TP53 mutations are relatively uncommon at diagnosis.

Additionally, at least 70 other recurrent somatically mutated genes have been identified in BL including known oncogenes and tumour suppressors including *ID3, GNA13, RET, PIK3R1, ARID1A* and *SMARCA4.* This indicates that in addition to the initiating *MYC* translocation in BL, cooperating mutations are likely required for malignant transformation (Love et al., 2012a).

1.3.3.2 MYC aberrations and secondary mutations in DLBCL

Sharing many of the histopathological, cytogenetic and biological signatures of BL is DLBCL. DLBCL is the most common form of lymphoma in adults (Morton et al., 2006). DLBCL patients respond variably to standard chemotherapy regimens. Initial cure rates are approximately 35-40%, while the majority of relapsed patients succumb to disease (Coiffier, 2001). Gene expression profiling and histopathological analysis has identified at least three sub-categories of DLBCL known as; germinal centre B cell-like (GCB) DLBCL, activated B cell-like (ABC) DLBCL and primary mediastinal (thymic) B cell lymphoma (PMBCL), that are related to the maturation stage of the pathogenic B cells (Alizadeh et al., 2000). In addition, DLBCL subtypes are highly heterogeneous at the genomic level. *MYC* translocations and *MYC* copy number gains are common events that predicted poor prognosis in DLBCL and are correlated with low proliferation rates, distinguishing them from the same *MYC* translocations found in BL (Yoon et al., 2008). *BCL2/IGH* translocations are observed commonly in GCB DLBCL but rarely observed in ABC DLBCL (Huang et al., 2002). In GCB DLBCL the *BCL2/IGH* translocation co-occurred at high frequency with *MYC* aberrations, highlighting the ability of *BCL2* and *MYC* to cooperatively drive multistep lymphomagenesis (Yoon et al., 2008). Unlike

common translocations, single nucleotide variants (SNV) and small insertions or deletions (InDels) are extremely variable in DLBCL. Recently, a large study identified 322 candidate cancer genes in patients with DLBCL, including a large number of rare, non-recurrent novel mutations (Zhang et al., 2013). The population-associated heterogeneity of DLBCL is highlighted by the modest overlap (approximately 10-20% of somatically acquired mutations) observed in four studies that investigated the genomic landscape of DLBCL (Lohr et al., 2012; Morin et al., 2013; Pasqualucci et al., 2011; Zhang et al., 2013). DLBCL and BL share many morphological features and cytogenetic abnormalities as both tumour types display high proliferative indexes, *MYC* rearrangements and complex karyotypes (So et al., 2013). Indeed, "unclassifiable B cell lymphoma with features intermediate between DLBCL and BL" is a recognised category of lymphoma (Campo et al., 2011). However, comprehensive genomic investigations have demonstrated that although clinically DLBCL and BL are similar, there is a large amount of heterogeneity between the two diseases as evidenced by a plethora of mutually exclusive recurrent mutations (Lohr et al., 2012; Love et al., 2012a).

1.3.3.3 Targeting the apoptotic pathway in MYC driven B cell lymphomas

Where the cytogenetics of BL and DLBCL do overlap is in the presence of dual *MYC* and *BCL2* rearrangements (Aukema et al., 2011). Through co-overexpression of *MYC* and *BCL2* lymphoma cells had acquired the proliferative signal and the countermeasure that inhibits the apoptotic response to unchecked mitogenesis (Swerdlow, 2014). 34% of DLBCL and BL treated with frontline aggressive induction chemotherapy exhibit overexpression of both *MYC* and *BCL2* at relapse (Hu et al., 2013). This genomic trait

is associated with chemotherapy resistance and poor survival prognosis as the “double hit” (DH) lymphomas demonstrate 30% five-years overall survival rate compared to 75% in “single hit” (SH) lymphomas or lymphomas without any *MYC* or *BCL2* rearrangements (Johnson et al., 2009). *BCL2* is a potent anti-apoptotic gene that is bound and neutralised by BH3-only proteins (Cheng et al., 2001; Lindsten et al., 2000; Willis et al., 2007; Youle and Strasser, 2008). The rearrangement and ectopic expression of *BCL2* in lymphoma was shown to confer multi-drug chemotherapy resistance by cooperating with co-occurring *MYC* overexpression (Schmitt et al., 2000). This discovery has led to the development of BH3 mimetics that inhibit *BCL2* and its related family members. This strategy has demonstrated pre-clinical efficacy in previously multi-drug resistant lymphomas (Mason et al., 2008; Roberts et al., 2016; Vandenberg and Cory, 2013; Whitecross et al., 2009). In BL, *MYC* rearrangement is the founding event; however, the epigenetic silencing of the pro-apoptotic BH3-family member protein *BIM*, is apparent in 46% of human BL and is associated with chemotherapy resistance (Richter-Larrea et al., 2010). Pre-clinically, chemotherapy sensitivity was restored by re-expression of *BIM* either through genetic manipulation or pharmacological intervention (Richter-Larrea et al., 2010).

1.3.3.4 MYC aberrations and secondary mutations in B-ALL

Unlike BL and DLBCL, B-ALL is a disease of immature B cells, as typified by the non-expression of several developmental features and markers that are characteristic of mature B cells (Campo et al., 2011; Vogler et al., 1978). B-ALL is observed in both children and adults, but the incidence of disease is highest between ages two and five

years, with approximately 60% of cases occurring in patients aged younger than 20 (Campo et al., 2011; Inaba et al., 2013). Childhood B-ALL exhibits a genotype of *MYC* t(8;14), t(2;8) or t(8;22) at a frequency of 2%, which rises to 4% in adult onset B-ALL (Pui and Evans, 1998). However, as previously mentioned, *MYC* is overexpressed in 47-52% of cases by means other than translocation, differentiating B-ALL from BL and DLBCL (Burmeister et al., 2005; Miranda Peralta et al., 1991). Traditionally, standard cytogenetic analysis demonstrated that the majority of B-ALL cases harbour gross chromosomal aberrations (Harrison, 2009). These characteristic chromosomal rearrangements include high hyperdiploidy with non-random gain of at least 5 chromosomes (from chrX, chr4, chr6, chr10, chr14, chr17, chr18 and chr21) and hypodiploidy resulting in fewer than 44 chromosomes (Harrison, 2013). As B-ALL subtypes are characterised by the acquisition of a large number of chromosomal aberrations, SNV and InDels that collaborate to drive the leukaemic clone, the roles of many genes mutated in B-ALL remain unknown due to the paucity of mouse models that faithfully recapitulate human B-ALL.

The heterogeneity discussed above highlights the complexity of the disease manifestation of BL, DLBCL and B-ALL. Thus, understanding mutations and genomic aberrations throughout B cell lineage malignancies could potentially identify drug-amenable mutations that are susceptible to targeted therapeutic intervention (Witzig and Gupta, 2010). The success of this approach is exemplified by identification of the *BCR-ABL* fusion and its activity as an oncogenic diver in poor-risk Philadelphia positive B-ALL (PH+ B-ALL). Ph+ B-ALL that was refractory to first line chemotherapy is now

considered treatable with a targeted combination of Imatinib, a targeted kinase inhibitor against the ABL fusion protein and chemotherapy (Foa et al., 2011; Talpaz et al., 2006). As introduced earlier, in DLBCL and BL *MYC* rearrangements are considered a common oncogenic, or primary, driver event. However, *MYC* is in fact mutated or oncogenically deregulated in approximately 50% of human malignancies (Klapproth and Wirth, 2010; Meyer and Penn, 2008). Importantly, in B cell lineage lymphomas/leukaemias with *MYC* deregulation there is direct evidence for the requirement of secondary, tertiary and quaternary mutations as seen in BL and DLBCL. Thus, identifying how clonal diversity is driven by the acquisition and function of these additionally acquired mutations in *MYC* driven B cell lineage malignancies could therefore yield not only new biological insight into oncogenic pathways involved with B cell lymphomagenesis but also discovery of new therapeutic targets that could be potentially translated to the treatment of a wide range of malignancies.

1.4 Clonal diversity of tumours

The sequential selection of primary, secondary, tertiary and quaternary mutations discussed earlier underpins oncogenesis. It is widely accepted that cancer progression can evolve through Darwinian processes (Nowell, 1976). In its simplest form, neoplastic tumourigenesis can be described as a 'normal' cell of origin acquiring the means by which to outgrow and outcompete its surrounding, non-mutated entities (Nowell, 1976). At a genetic level, cancer progression is initiated by the acquisition of multiple mutations capable of activating oncogenic pathways and/or silencing tumour suppressor networks (Vogelstein et al., 2013) (Figure 1.02). The acquisition of mutations in the founding

clone must be potent enough to provide a selective advantage otherwise the pre-neoplastic cell faces suppression and eradication by inherent tumour-protective failsafe mechanisms. This process is exemplified by the presence of covert pre-malignant lesions observed in patients that is suggestive of consistent immune surveillance (Greaves and Maley, 2012b; Mori et al., 2002). Cancer initiating mutations can result from intrinsic error in the DNA repair processes, or can be associated with genotoxic insults such as ultra-violet light, cigarette smoke, ionising radiation and chemotherapies that have all been strongly linked to cancer development (Greaves and Maley, 2012a; Stratton, 2011). Though mutagenic processes are not generally biologically tolerated, there are some exceptions: in B cells there is a functional mutagenic process where enzymes target recombination in immunoglobulin genes to enhance antibody diversity and affinity. This process is fundamental for production of unique variable domains in the B cell receptor of individual B cell clones that underlies function of the immune system (Burnet, 1976; Tsai et al., 2008). If initiating mutations or subsequent mutations involve a gene implicated in DNA stability and repair, then within that clone there is an increased rate of subsequent gene mutations, potentially accelerating the neoplastic transformation (Bardelli et al., 2001; Cahill et al., 1999). Aberrant *MYC* expression through translocation or the acquisition of stabilising mutations is one such genetic insult capable of leading to increased mutagenesis due to increased DNA damage (Meyer and Penn, 2008). The increased mutation rates apparent in cancers and the vast number of proliferative cycles required for the formation of a macroscopic tumour allows for gross intratumoural heterogeneity through genetic diversity as defined by subclonal populations harbouring unique arrays of mutations (Marusyk et al., 2012).

Intratumour heterogeneity has been demonstrated in B cell leukaemia, where 25% of cases had more than one representative tumour clone in the sample (Campbell et al., 2008). Intratumour heterogeneity has been demonstrated temporally. Leukaemic cells harbouring an *ETV6-RUNX1* fusion were followed over serial transplantations in immunocompromised mice. These tumours demonstrated non-linear and complex branching evolutionary paths, with copy number alterations in the subclones occurring independently and reiteratively (Anderson et al., 2011). This is also the case for BL and DLBCL, which harbour a high number of founding mutation events such as *MYC* translocations at high variant allele frequency (VAF) and many low frequency secondary, tertiary and quaternary mutations of both known and unknown significance (Lohr et al., 2012; Love et al., 2012a; Morin et al., 2011; Pasqualucci et al., 2011; Zhang et al., 2013). In the case of BL and *MYC*-translocated DLBCL, if the acquisition of oncogenic MYC were sufficient in driving malignant transformation, then the heterogeneity of subclonally represented mutations would likely not present so profoundly.

1.4.1.1 Clonal diversity, resistance and relapse

By definition, greater intratumour heterogeneity increases the frequency of subclonal populations harbouring mutations that can potentially confer chemotherapy resistance. This premise is a central problem in cancer treatment (Merlo et al., 2006). Dominant clonal outgrowth from a heterogeneous population in response to therapeutic challenge has been demonstrated in both solid and malignant tumours (Marusyk and Polyak, 2010) (Figure 1.02). Recently, cancer cell lines tagged with unique DNA “barcodes”

were exposed to chemotherapy and tracked over time to investigate clonal selection. These experiments demonstrated that the majority of resistant clones were detectable as small subclonal populations that selectively escaped and survived therapeutic challenge (Bhang et al., 2015). This is evident in AML, where subclones harbouring additional mutations were present prior to therapy, survived throughout treatment and expanded rapidly at relapse, becoming the dominant clone (Ding et al., 2012).

As previously discussed, the identification of the oncogenic *BCR-ABL* fusion gene drove the development of the ABL kinase inhibitor, imatinib, which rapidly improved outcome in patients with Ph+ leukaemia (Buchdunger et al., 1996; Zimmermann et al., 1996). However, a significant proportion of these patients suffered relapse and of note, the majority of the relapsed tumours presented as a clone that contained a single amino acid change in the kinase domain of *BCR-ABL*, rendering imatinib ineffective (Deininger et al., 2005; Gorre et al., 2001). This kinase domain mutation was present in leukaemia subclones in patients prior to treatment with Imatinib (Deininger et al., 2005; Gorre et al., 2001). Therefore, imatinib therapy is thought to positively select for the resistant subclones leading at first to remission (loss of sensitive dominant clones) and eventual relapse of the newly resistant disease (Hofmann et al., 2003; Roche-Lestienne et al., 2002; Shah et al., 2002). In fact, in the majority of samples from relapsed childhood B-ALL, patients harboured some genetic aberrations present at diagnosis (Mullighan et al., 2008). This indicates that as a result of intratumoural heterogeneity, subclonal populations have persisted and even acquired further mutations and cytogenetic changes (Mullighan et al., 2008). Therapy resistance driven by clonal diversity is not

limited to haematological malignancies, as mutations conferring therapy resistance have been identified in subclones prior to treatment in solid tumours. For example, *KIT* mutations conferring imatinib resistance have been identified in gastrointestinal stromal tumours (Corless and Heinrich, 2008). In addition to directly attenuating pharmacological activity of a drug, mutations occurring in tumour subclones have been shown to activate drug-targeted oncogenic signalling pathways through alternative means. This is exemplified by *BRAF*-mutant melanomas that relapse as *BRAF*-inhibitor resistant clones due to the outgrowth of subclones harbouring *KRAS* mutations, reactivating the BRAF pathway downstream of BRAF (Nazarian et al., 2010).

MYC driven haematopoietic malignancies, typified by BL and aggressive DLBCL, are often found to be profoundly chemotherapy refractory due the suite of oncogenic insults they acquire over time and across therapies (Landau et al., 2014). However, unlike the *BCR-ABL* translocation, targeted inhibition of MYC has been elusive (Meyer and Penn, 2008). In murine models of *MYC* driven haematological malignancies, tumours were shown to regress following *Myc*–inactivation through genetic engineering (Felsher and Bishop, 1999a; Giuriato et al., 2006; Karlsson et al., 2003; Soucek et al., 2013). Soucek *et al* demonstrated that short-term systemic *Myc* silencing had a negligible toxicity profile; however, the authors posited that sustained *Myc* silencing would lead to severe side effects (Soucek et al., 2013). These genetic processes mimicked potential pharmacological inhibition strategies of MYC. However, under these circumstances tumours exhibiting novel chromosomal changes that are impervious to the loss of Myc signalling evolve as a result of intratumoural heterogeneity (Karlsson et al., 2003).

Similarly, when ectopic *Myc*-expression was abrogated in primary murine haematopoietic tumours driven by *Myc* overexpression, the secondary relapsed tumours demonstrated clonal expansion of pre-existing malignant subclones that lacked both the tumour suppressor gene *Trp53* and a dependency on Myc (Giuriato et al., 2006). These findings were supported by the high frequencies of novel chromosomal rearrangements and *TP53* mutations in adult human relapsed BL (Love et al., 2012a; Schmitz et al., 2012).

Identifying the dominant oncogenic driver event is critical to developing targeted therapeutics, as demonstrated by the discovery of *BCR-ABL* in CML and the contextual efficacy of imatinib. Similarly, the ability to induce tumour remission in *Myc*-driven malignancies by *Myc*-silencing using murine models demonstrates the importance of targeting the dominant driver mutation. However, even with selective disruption of the initiating oncogenic lesion, relapse and resistance may be conferred by the selection of treatment-resistant subclones present in the tumour population prior to treatment induction (Bhang et al., 2015; McGranahan and Swanton, 2017; Turke et al., 2010) (Figure 1.02). Therefore, a better understanding of intratumoural heterogeneity, subclonal mutations and pathways to resistance may lead to improvements in the rates of cancer relapse and disease progression.

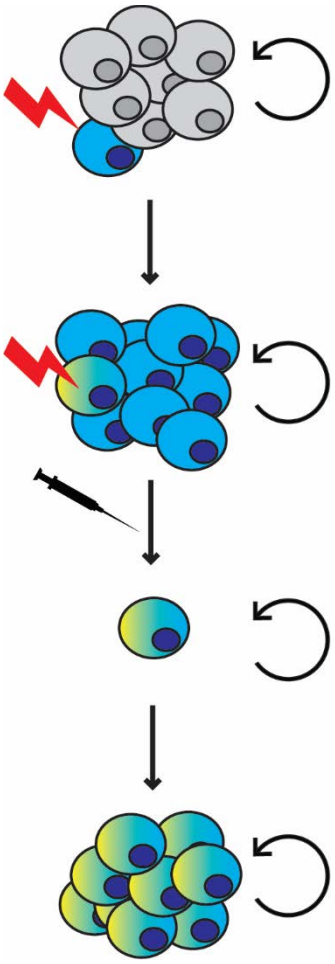


Figure 1.02 – An example of intratumour heterogeneity, clonal selection and resistance

Schematic demonstrating that at first, a ‘normal’ cell has acquired a mutation (red bolt) allowing it to outcompete its surrounding tissue and becoming a malignant entity. This initially acquired driver mutation will be present in all descendent cells (blue cells). Subclonal populations exist due to the random acquisition of mutations, some of which can confer therapy resistance (yellow-blue gradient). Therapy is able to treat the bulk of the tumour, however, the pre-existing resistant subclonal population is selected for, and as a result the tumour relapses as the therapy refractory clone.

1.4.1.2 Understanding clonal diversity to overcome resistance and relapse

Therapy resistant subclones usually result in tumour relapse as certain subclonal populations harbour mutations allowing them to directly circumvent targeted therapeutic mechanisms (Deininger et al., 2005; Gorre et al., 2001) or through activation of complementary and cooperative oncogenic pathways, attenuating the reliance of the tumour on the clonal driver mutation (Giuriato et al., 2006; Karlsson et al., 2003; Nazarian et al., 2010). Understanding and predicting the pathways to resistance in response to therapies has yielded new and efficacious targeted therapies. This is highlighted by the identification of *MET*-amplified subclones present in a human lung carcinoma cell line that are resistant to erlotinib treatment and are selectively advantaged under erlotinib treatment (Bhang et al., 2015). Here, it was demonstrated that treatment of lung carcinoma cell lines was more effective with a combination of erlotinib and the MET-inhibitor, crizotinib (Bhang et al., 2015). This same study demonstrated the existence of ABL-inhibitor resistant CML cells. Subclonal resistance was partially overcome by treatment with GNF-2, an allosteric inhibitor of ABL, and as such is not affected by the amino acid change at the kinase domain that confers resistance to traditional ABL kinase domain inhibitors (Bhang et al., 2015). Indeed, potent next-generation allosteric inhibitors of ABL1 are currently under development and in early phase clinical trials in the treatment of tumours resistant to classical kinase inhibitors (ClinicalTrials.gov, NCT02081378).

By understanding intratumour heterogeneity, and how subclonal populations can cooperate with primary driver mutations, chemotherapy resistance could potentially be

overcome in a targeted manner. If so, remission and relapse rates can potentially be improved. Given that effective targeting of *MYC* has remained elusive and *MYC* is oncogenically deregulated in the majority of BL and DLBCL, there is a clear unmet clinical need for further understanding the clonal diversity and inconspicuous subclonal mutations that concomitantly occur and cooperate with ectopic *MYC* expression.

1.5 Mouse models of malignancy

Cross-species comparisons between human and animal models of disease have proven to be a fertile ground for discovery, validation and characterisation of new cancer genes and oncogenic processes (Van Dyke and Jacks, 2002; van Miltenburg and Jonkers, 2012). Genetic manipulations including single oncogene insertion or inactivation of tumour suppressor genes in the germline of recombinant mice can predispose mice to the development of cancer (Adams and Cory, 1991). For example; the SV40 virus containing the oncogenic T antigen element elicited brain tumours in mice when it integrated into mouse germline (Brinster et al., 1984), insertion of a fusion gene made up of a hormonally inducible mouse mammary tumour promoter virus with *Myc* lacking its normal promoter region resulted in the development of spontaneous mammary adenocarcinomas (Stewart et al., 1984) and B cell specific expression of *Myc* by way of chromosomal translocation gave rose to the development of a B cell lymphoma within the transgenic mice (Adams et al., 1985). These models have allowed investigations into the causal links between mutations and their effect on critical cellular processes that leads to cancer initiation and maintenance (Macleod and Jacks, 1999). Sporadic murine mutation-driven tumours (usually in a cell-specific manner), often show

predictable histopathological and molecular progression compared to the analogous human diseases (Robles-Espinoza and Adams, 2013). However, in many cases, the founding alteration to a cancer gene in a GEMM is not sufficient to drive malignant transformation (Andrechek and Nevins, 2010). Importantly, this trait of GEMM provides a platform with which to study the process of tumour evolution and clonal diversity (previously discussed in 1.4). As tumours driven by a single oncogenic lesion will ostensibly progress with the acquisition of secondary mutations and/or epigenetic changes intra- and intertumoural heterogeneity and clonal diversity is generated (Andrechek and Nevins, 2010). This is exemplified in TgT_{121} mice, which cell-specifically inactivate the pRb pathway (Saenz Robles et al., 1994). These mice are predisposed to a focal emergence of tumours, with 38% of resulted tumours having somatically acquired an inactivating *Trp53* mutation, circumventing the Trp53-dependent apoptosis apparent in the early progression of these tumours (Lu et al., 2001). Thus, elucidating the cooperative effects of *Tp53* and *Rb* mutations that are frequently co-mutated in human malignancies (Williams et al., 1994). These dynamic GEMM allow the investigation of somatically acquired mutations capable of cooperating with initial germline mutations to drive tumourigenesis, and eventually progression to malignancy. This approach has yielded insights into clinically relevant mutations present in human malignancies that were not previously thought to play a role in tumourigenesis.

1.5.1 The E*μ-Myc* transgenic mouse model of Myc-driven B cell lymphoma

As alluded to above, GEMM have aided in understanding cryptic genomic events, such as gene translocations. At first, the 'recombination mistake' associated with BL was a perplexing notion as other oncogenes identified, such as HRAS, were variants of normal cellular genes that had become activated following somatic mutations (Reddy et al., 1982). Around the same time the *BCR-ABL* fusion was mapped to the translocation in the Philadelphia chromosome with this abnormality generating a novel fusion product (de Klein et al., 1982). These mutations, and novel fusion proteins, are conspicuously absent when referencing MYC dysfunction. In the case of MYC, overexpression of a non-mutated gene is adequate for tumour initiation, which at the time was unexpected (Meyer and Penn, 2008). To better understand the role of translocated *MYC* in B cell malignancies, Adams *et al* generated the E*μ-Myc* murine model of lymphoma, a GEMM, which mimicked the t(8;14) translocation apparent in BL (Adams et al., 1985). E*μ-Myc* transgenic mice were generated by introducing a DNA construct obtained from a spontaneous mouse plasmacytoma, where *Myc* was coupled to immunoglobulin (Ig) heavy chain enhancer elements (E*μ*) into fertilized ova (Harris et al., 1988). These transgenic mice developed Burkitt's-like leukaemia/lymphoma with 100% penetrance. E*μ-Myc* mice underwent a smouldering pre-malignant polyclonal B cell expansion followed by onset of clonal lymphoid neoplasia ranging from a pre-B to naïve-B phenotype (Adams et al., 1985; Sidman et al., 1993). Leukocytes in pre-malignant E*μ-Myc* mice had a large nuclear volume and overall size indicating higher proliferation rates than WT leukocytes (Strasser et al., 1990). However, during the pre-malignant

phase, the total numbers of leukocytes between E*μ-Myc* and WT mice were remarkably similar despite B cells in the E*μ-Myc* mice proliferating at a higher rate (Strasser et al., 1990). This phenomenon accounted for the 'remission' phase in E*μ-Myc* mice where tumour suppressive pathways were activated in response to unchecked Myc expression, regulating the proliferation of B cells in the bone marrow (Strasser et al., 1990). Pre-leukaemic E*μ-Myc* mice exhibited increased B cell proliferation and death until a clone spontaneously developed a means to circumvent the self-protective pro-apoptotic pathways that normally regulated B cell expansion, resulting in progressive lymphoma/leukaemia (Sidman et al., 1993).

E*μ-Myc* lymphoma is driven by a single, constitutively active proto-oncogene, Myc, on a genetically homozygous background. However, despite this E*μ-Myc* mice exhibit a high degree of heterogeneity as illustrated by gene expression profile analysis, latency of disease onset and cell surface antigen phenotype (Mori et al., 2008). Much like human malignancies (discussed in 1.4), it is widely acknowledged that lymphomas in the E*μ-Myc* setting arise by further somatic mutations that augment the oncogenic potential conferred by *Myc* overexpression. This reliance on the acquisition of secondary, tertiary and quaternary mutations in E*μ-Myc* lymphomas gives rise to the intertumoural heterogeneity apparent in the resultant cancers and thus can be utilised to study the heterogeneity of parallel human malignancies (Mori et al., 2008). The E*μ-Myc* model has been an incredibly valuable tool in the identification of oncogenic and tumour suppressive networks that are capable of cooperating with *Myc* in driving malignant transformation. Given the high prevalence of *Myc* dysfunction in human malignancies,

these discoveries have been crucial in shaping our current understanding of *Myc* driven neoplastic progression.

1.5.1.1 Eμ-Myc lymphoma and the p19^{ARF}-MDM2-Tp53 tumour suppressor axis

Mutations in *TP53* and *CDKN2A* have been frequently implicated in human malignancies (Sherr and Weber, 2000). TP53 is a potent tumour suppressor gene that induces cell cycle arrest and apoptosis in response to cellular stresses such as DNA damage and mitogenic oncogene expression (Hermeking and Eick, 1994; Kastan et al., 1991; Lowe, 1999). The *CDKN2A* gene encodes p16^{INK4A}, a cyclin dependent kinase inhibitor that leads to activation of *Rb* and subsequent cell cycle arrest (Serrano et al., 1993), and p14^{ARF} (p19^{ARF} in mice), that releases TP53 by binding the TP53 negative regulator, MDM2 (Kamijo et al., 1998). In mouse embryonic fibroblasts (MEF) constitutive *Myc* activation increased cellular levels of p19^{ARF} and Trp53, leading to replicative crisis as the cells undergo increased apoptosis. Furthermore, MEF lacking either p19^{ARF} or Trp53 were resistant to Myc-induced apoptosis (Zindy et al., 1998). Similarly, in juvenile E*μ-Myc* transgenic mice, B cells are more rapidly dividing but the increased proliferation is offset by increased apoptosis (Jacobsen et al., 1994; Strasser et al., 1990). As seen in MEF, it was hypothesised that ectopic *Myc* expression driven by the E*μ* promoter accounts for both the high proliferative and high apoptotic rates in B cells, and that for E*μ-Myc* lymphomagenesis to occur, secondary, tertiary and quaternary mutations that allow B cells to circumvent the apoptotic effects of Myc must be acquired (Eischen et al., 1999). Studies in which E*μ-Myc* transgenic mice were crossed with either *Tp53*$^{-/-}$ or *Tp53*$^{+/-}$, or *Cdkn2a*$^{-/-}$ or *Cdkn2a*$^{+/-}$ (encoding p19^{ARF}) mice

demonstrated markedly accelerated rates of lymphomagenesis compared to E*μ-Myc* animals crossed onto WT background (Eischen et al., 1999; Schmitt et al., 1999). There is redundancy shared between these genes as the rate of Myc-driven lymphoma is not accelerated by deletions in both *Cdkn2a* and *Trp53* (Schmitt et al., 1999). The loss of *Cdkn2a* or *Trp53* reduced the rate of apoptosis in E*μ-Myc* driven lymphoma (Eischen et al., 1999; Schmitt et al., 1999), and rendered the resulting lymphomas refractory to cytotoxic therapy such as standard first-line chemotherapy (Eischen et al., 1999). Selective breeding demonstrated the importance of $p19^{ARF}$ and Trp53 in restraining Myc activity. However, inactivation of the $p19^{ARF}$-Mdm2-Trp53 tumour suppressor axis occurs spontaneously in approximately 50% of E*μ-Myc* lymphomas (Eischen et al., 1999). These high frequency *de novo* inactivating mutations in *Cdkn2a* and *Trp53* suggest that the $p19^{ARF}$-Mdm2-Trp53 pathway is placed under significant selective pressure when Myc is driving lymphomagenesis (Eischen et al., 1999) (Figure 1.01). Although the identification of spontaneous *Cdkn2a* and *Trp53* mutations in E*μ-Myc* lymphoma was among the first descriptions of the genetic events that are responsible for driving malignant transformation from a pre-cancerous B cell, many additional genes have been demonstrated to have oncogenic properties in the context of *Myc*. In BL, inactivating *TP53* mutations and biallelic *CDKN2A* deletions occur independent of one another at high frequencies (Lindstrom et al., 2001; Love et al., 2012b). Thus, the consistency in tumour evolution and competing mutations in the E*μ-Myc* mouse model of lymphoma faithfully recapitulates the genetic mechanisms that underpin MYC-driven oncogenesis in parallel human malignancies.

1.5.1.2 Eμ-Myc lymphoma and Bcl-2 family members

BCL2 family proteins further influence apoptosis in the E*μ-Myc* model independently of the aforementioned p19^{ARF}-Mdm2-Trp53 pathway. This includes both anti-apoptotic proteins (*BCL-2, BCL-X$_L$* and *MCL1*), pro-apoptotic (*BAK* and *BAX*) and BH3-only pro-apoptotic members (*BIM, BID, NOXA* and *PUMA*) (Kvansakul et al., 2008; Martinou and Youle, 2011). *BCL2* was first implicated in haematological malignancies as a transcriptional element associated with the t(14;18) translocation (Tsujimoto et al., 1987; Tsujimoto et al., 1985a; Tsujimoto et al., 1985b). Overexpression of *Bcl2* in a transgenic mouse model that mimicked the t(14;18) translocation demonstrated accumulation of small pre-B cell and B cell lymphocytes attributable to an increase in viability but not proliferative capacity (McDonnell et al., 1989; McDonnell et al., 1990). However, when crossed with E*μ-Myc* transgenic mice, *Bcl2*-translocated transgenic mice generated B cell lymphomas at an accelerated rate, highlighting the synergistic capability of *Myc* and *Bcl2* in driving lymphomagenesis (Strasser et al., 1990). Co-expression of *Bcl2* and *Myc* in E*μ-Myc* lymphomas drove malignant transformation in the absence of loss-of-function mutations in either *Trp53* or *Cdkn2a* (Schmitt et al., 2002). Furthermore, in approximately half of spontaneous E*μ-Myc* lymphomas, levels of *Bcl2* and the related anti-apoptotic family member *Bcl-x$_L$* were elevated irrespective of *Cdkn2a* and *Trp53* functional status (Eischen et al., 2001). These results indicate that *Bcl2* can cooperate with *Myc* to drive lymphomagenesis by providing a critical anti-apoptotic signal while relieving the selective pressure placed on the p19^{ARF}-Mdm2-Trp53 pathway (Eischen et al., 2001). *Bcl2* and other anti-apoptotic *Bcl2*-family members are antagonised by the BH3-only protein, *Bim* (O'Connor et al., 1998). E*μ-*

Myc mice were crossed with *Bim*$^{+/-}$ mice resulting in accelerated lymphomagenesis compared to E*μ-Myc* littermate controls (Egle et al., 2004). These experiments demonstrated for the first time the potential of oncogenic *Myc* expression combined with either *Bcl2* overexpression or attenuated *Bim* in driving B cell lymphomagenesis. Clinically, these findings are important as BL and *MYC*-rearranged DLBCL regularly present with concurrent *BCL2* overexpression, which is considered to be of a prognostically unfavourable genetic profile (Valera et al., 2013). ABT-737 (Navitoclax) is a BH3-only mimetic and pan-antagonist of anti-apoptotic Bcl2-family proteins, including Bcl2 and Bcl-x_L (Tse et al., 2008). Treatment of mice transplanted with E*μ-Myc;Bcl2* lymphoma cells with ABT-737 significantly improved survival rates compared to vehicle treated controls. Furthermore, combination treatment with ABT-737 and the chemotherapeutic cyclophosphamide resulted in durable disease free survival in the transplanted mice (Mason et al., 2008). Next generation BH3-only mimetics that do not have the same toxicity profile as ABT-737, such as ABT-199 (Venetoclax) are in clinical use (Cang et al., 2015; Roberts et al., 2016). The E*μ-Myc* mouse model of lymphoma has been critical in identifying the cooperative oncogenic effects of *Myc* and *Bcl2* overexpression, *Myc* loss of *Bim* and the opposing relationship between *Bcl2* and *Bim.* This has ultimately led to the development of a targeted therapy that promises efficacy in patients with notoriously poor outcome in the refractory cases (Cang et al., 2015; clinicaltrials.gov, NCT01328626). This highlights the historical importance and value of the E*μ-Myc* mouse model of lymphoma as a discovery tool in dissecting the tumour-suppressive and tumour-promoting network of genes that interact with *Myc*.

1.5.1.3 Forward genetic screens in Eμ-Myc lymphoma

In addition to reverse genetic-screens using GEMM (whereby mutations are discovered in spontaneously developed malignancies, as detailed above) insertional mutagenesis using viral technology, siRNA screens and DNA transposons are efficient methods to perform forward genetic screens to study genes implicated in cancer (McIntyre et al., 2012). Forward genetic screening systems rely on the accumulation of mutations in cells, where an eventual activating mutation(s) in a proto-oncogene and/or inactivating mutation(s) in a tumour suppressor gene occurs that bestows a selective advantage and promotes clonal outgrowth (Kool and Berns, 2009). Following mutagenesis and tumourigenesis, sequencing the insertion sites across a panel of tumours for recurrently mutated loci provides evidence to suggest that the associated genes are implicated in tumourigenesis (de Jong et al., 2011; Uren et al., 2009). Importantly, this approach has shown remarkable overlap between oncogenes/tumour suppressors in human malignancies and orthologous genes in mouse mutagenesis screens (Mattison et al., 2009). Indeed, forward genetic screens have identified cancer genes in GEMM before they were understood to be relevant in human disease (Bender et al., 2010; Cuypers et al., 1984; Dreyfus et al., 1990). Mutagenesis screens in GEMM of cancer have also been utilised in the understanding of critical cooperative oncogenic networks. Most notably, mutagenesis screens in E*μ-Myc* mice have identified clinically relevant genes capable of cooperating with oncogenic *Myc* to drive lymphomagenesis (Haupt et al., 1991; van Lohuizen et al., 1991). Infection of E*μ-Myc* transgenic animals with the Moloney murine leukaemia virus (Mo-MLV) resulted in accelerated B cell lymphomagenesis compared to the non-infected E*μ-Myc* mice (Haupt et al., 1991).

Through interrogation of the Mo-MLV-induced E*μ-Myc* lymphomas, the provirus insertion site was identified within the *Bmi-1* gene (a polycomb group (PcG) gene) in up to 50% of the tumours studied (Haupt et al., 1991; van Lohuizen et al., 1991). PcG genes were originally recognised as epigenetic silencers and several members were implicated in human malignancies (Jacobs and van Lohuizen, 2002; Raaphorst, 2005). It was demonstrated that the provirus insertion resulted in increased *Bmi-1* expression, providing strong evidence that it acts as an oncogene in E*μ-Myc* lymphoma (Haupt et al., 1991; Verbeek et al., 1991). Bmi-1 was later demonstrated to be negative regulator of Myc-induced expression of $p19^{ARF}$ and its downstream apoptotic response (Jacobs et al., 1999). Following the identification of *Bmi-1* in forward genetic screens in E*μ-Myc* lymphoma, ectopic overexpression in *BMI-1* has been implicated in several human malignancies (Raaphorst, 2005), used as a prognostic factor (Chowdhury et al., 2007) and has been shown to correlate with drug resistance (Bhattacharyya et al., 2012). To more broadly identify tumour suppressor genes that oppose the oncogenic effects of *Myc* expression, genome-wide *in vivo* loss of function RNAi screens were performed in the E*μ-Myc* model of lymphoma (Bric et al., 2009). Here, E*μ-Myc* haematopoietic stem cells were transduced with complex shRNA pools and injected into recipient mice with resultant lymphomas being sequenced for shRNA integration. Enrichment of specific shRNA in the lymphomas resulted in the hypothesis that the genes targeted by the enriched shRNA were tumour suppressive in the context of E*μ-Myc* lymphoma. The resulting screens identified over 10 potential tumour suppressor genes, including *Mek1* and *Rad17*, which were extensively validated (Bric et al., 2009). *Mek1* is an activator of the MAPK pathway, which is oncogenic in many malignancies (Dhillon et al., 2007).

However, it was demonstrated to have anti-oncogenic effects supporting previous studies that demonstrated that Mek1 inhibition can destabilise Myc and therefore be, contextually, tumour suppressive (Bric et al., 2009; Sears et al., 2000). Attenuation of *Rad17* in E*μ-Myc* lymphoma resulted in dysfunction of the Myc-induced DDR (Bric et al., 2009). Following its identification as a tumour suppressor in E*μ-Myc* lymphoma, a significant fraction of unfavourable DLBCL were found to demonstrate decreased expression of *RAD17* (Bric et al., 2009). Thus, provirus insertional mutagenesis and high throughput RNAi screens in the E*μ-Myc* model of lymphoma have proved effective in identifying potentially physiologically relevant genes that when activated or attenuated can cooperate with *Myc* to drive lymphomagenesis.

1.5.1.4 Ras activation in Eμ-Myc lymphoma

The E*μ-Myc* model of lymphoma has also been utilised to identify cryptic mutations that cooperate with *Myc* outside the classic tumour suppressive pathways highlighted above. Activating mutations in *RAS* family genes (*NRAS, KRAS* and *HRAS*) have long been known to occur in human malignancy (Der et al., 1982; Malumbres and Barbacid, 2003; Shih and Weinberg, 1982). RAS proteins, which have significantly overlapping function, control pathways associated with cell activation and proliferation. Thus, constitutively active RAS proteins contribute to many aspects of oncogenesis including increased proliferation and invasiveness (Lowy and Willumsen, 1993; Malumbres and Barbacid, 2003). Activating *Ras* mutations have also been described in E*μ-Myc* lymphoma. Two out of 14 spontaneously derived E*μ-Myc* lymphomas demonstrated *Kras* and *Nras* point mutations at codon 61, leading to constitutively active Ras proteins (Alexander et al.,

1989; Buhrman et al., 2011a). Point mutations in *Kras* and *Nras* at codon 61 have been identified frequently in human malignancies, with 1% of all *Kras* and 61% of all *Nras* mutations occurring at this hotspot (Cox and Der, 2010). Retroviral expression of constitutively active Nras in the E*μ-Myc* system led to rapid malignant transformation *in vivo* validating for the first time the oncogenic cooperation between Myc and Nras (Alexander et al., 1989). The mechanism of mutant Nras and Kras cooperation with Myc in driving lymphomagenesis is unclear. However it is postulated that constitutively activated Nras or Kras can stabilise Myc protein and/or regulate key *Myc* repressors (Alexander et al., 1989).

Despite being driven by a single oncogene on a homozygous genetic background, E*μ-Myc* lymphomas demonstrate a remarkable degree of inter-tumoural heterogeneity that is attributable to the secondary mutations acquired that occur during transformation. The E*μ-Myc* model of lymphoma has demonstrated its value as a discovery tool used to identify genes or pathways of clinical importance as their loss- or gain-of-function is able to cooperate with Myc leading to malignant transformation. However, only $p19^{ARF}$ loss, Tp53 mutation and Ras activating mutations have been described *de novo* in E*μ-Myc* lymphoma. From a clinical perspective, the identification of MYC co-operating genes and pathways is paramount. Genes and pathways that are spontaneously mutated during MYC-driven tumourigenesis must be under selective pressure suggesting they play a physiological role in restraining the proliferative effects of MYC. The physiological role of MYC is believed to be a global transcription factor. Thus, it is not readily amenable to direct therapeutic pharmacological inhibition. Identifying new genes

that co-operate with MYC might therefore yield not only new biological insight into oncogenic pathways involved with B cell lymphomagenesis but also aid in the discovery of new therapeutic targets that could be potentially translated to the treatment of a wide range of malignancies.

1.6 Next generation sequencing

For over 25 years, identification of cancer mutations and genes altered by those mutations has been central to understanding tumourigenesis and the normal cellular function of those genes. Over 1.6% of the human coding genome has been identified as being recurrently mutated in cancer and therefore likely to contribute to tumourigenesis (Stratton et al., 2009). Most of the mutations in these 'cancer genes' were discovered through identification of their physical location in the genome by low-resolution genomic screens such as cytogenetic analysis in leukaemia, transforming activity assays and forward genetic screens (Stratton et al., 2009). Higher resolution sequencing technology, such as Sanger sequencing, at the DNA level, identified substantial complexity in cancer genomes and has led to the discovery of many alterations in critical genes and pathways in cancer (Meyerson et al., 2010). For genomic analysis, automated Sanger sequencing was the gold-standard approach for over two decades (Metzker, 2010). Notably, Sanger sequencing led to the completion of the first complete human genome sequence (Venter et al., 2001). However, there has been a fundamental shift away from 'first generation' sequencing technology with the development, continued improvement and falling costs of the 'next generation' sequencing (NGS) platforms (Metzker, 2010). Breakthroughs in NGS technology has

led to a flood of information relating to somatically acquired transforming mutations. These discoveries, through use of whole genome sequencing (WGS), whole exome/exon sequencing (WES) and/or RNA sequencing (RNA-seq), are exemplified in seminal publications describing novel genetic insults in a wide range of haematological malignancies (Chin and Gray, 2008). The number of genes known to be implicated in haematological neoplastic transformation is constantly increasing as more studies are applying massive parallel sequencing (MPS) to cancer cohorts as the cost of NGS technologies rapidly decline (International Cancer Genome et al., 2010). A high frequency of transforming mutations found by MPS involved in lymphocyte activation, apoptosis and differentiation are consistent with existing knowledge, while other genes are poorly understood. Recent MPS studies have provided compelling evidence for the significance of high frequency recurrent mutations and their role in haematological malignancies, with these attracting an immediate research focus. However, many genes are mutated at a much lower frequency and will take significant research effort to dissect their contribution to cancer.

1.6.1 Whole exome sequencing

WES (or exome-capture) is an NGS technology that allows for targeted genomic capture. By using specific genomic hybridisation and capture techniques, WES allows for unbiased analysis of known protein coding regions in the genome (Teer and Mullikin, 2010). WES allows for cost effective discovery of unknown mutations in the coding regions of genes that drive genetically dependent disease, such as cancer, as exonic regions account for 1% of the human genome but harbour 85% of mutations with known

disease effects (Choi et al., 2009). Because WES targets such a small proportion of the genome, vastly deeper coverage across the target bases can be achieved with considerably less cost and raw sequence than one would require to sequence the whole genome (Meyerson et al., 2010). Recently, WES has been utilised to define mutations that occur in human DLBCL. Genes implicated in cancer development are observed as mutated somatically mutated across a cohort of malignancies. One such study identified 322 recurrently mutated genes in 73 DLBCL patient samples and 21 well defined DLBCL cell lines (Zhang et al., 2013). Of these 322 genes, 270 had not previously been associated with DLBCL. Furthermore, targeted inhibition of at least one of these 'unknown' genes had anti-lymphoma efficacy in the samples harbouring that mutation (Zhang et al., 2013). Separate independent studies investigating the mutations of DLBCL have identified a similar amount of known and unknown genes in DLBCL (Lohr et al., 2012; Morin et al., 2011; Pasqualucci et al., 2011; Zhang et al., 2013). These studies highlight the ability of WES to identify novel previously undescribed gene mutations in haematological malignancies.

1.6.2 RNA sequencing

Massive parallel cDNA sequencing, transcriptome sequencing, or RNA sequencing (RNA-seq), allows for characterisation and quantification of transcriptomes through interrogating messenger RNA (mRNA) transcripts (Ozsolak and Milos, 2011). Development in RNA-seq technology allows for start site mapping, strand-specific measurements, gene fusion detection and detection of alternative splicing events (Mardis and Wilson, 2009; Wang et al., 2009). Unlike earlier platforms such as micro-

array techniques, RNA-seq is not dependent on *in silico* prediction, but is a tool that can discover novel transcripts rather than relying solely on known mRNA transcripts (Mardis and Wilson, 2009). This attribute of RNA-seq has identified both known and novel fusion proteins in primary human prostate cancer samples (Maher et al., 2009; Trapnell et al., 2009). As RNA-seq interrogates transcripts, it can be used to assess allelic imbalance where mutations are heterzygous. In NHL, 16.8% of heterozygous mutated genes display prefernetial expression of the mutant allele. Of interest, this subset contained genes with obvious links to cancer including *BCL2, EZH2, TP53* and *IRF4* (Morin et al., 2011). Thus, RNA-seq is a crucial tool that will facilitate our understanding of tumour evolution and selection of clones that escape therapy and generate relapse.

1.6.3 Whole genome sequencing

Of all the NGS technology discussed herein, WGS provides the most complete genomic characterisation. However, it is laborious and costly compared to other NGS technologies discussed previously in chapter 1.6. Unlike RNA-seq or WES, WGS does not target regions of DNA or mRNA transcripts, but instead offers complete sequencing of all coding and non-coding regions. This provides sufficient resolution for the discovery of nucleotide substitutions, structural alterations and copy number alterations (Meyerson et al., 2010). WGS allows for the identification of chromosomal rearrangements, which were previously only identified in haematological malignancies by cytogenetic investigation (Meyerson, 2007; Rowley, 2001). Additionally, WGS can detect mutations that are unobservable using 'capture' methods (such as WES or RNA-seq) as they reside in non-coding regions including promoters, enhancers and non-

coding RNA (Meyerson et al., 2010). The power of WGS in blood cancers was exemplified by recent work in CLL. Here, WGS in adult CLL identified mutations in non-coding regions such as the 3' UTR of *NOTCH1* and *PAX5* enhancer regions, resulting in marked deregulation of these genes (Puente et al., 2015). *PAX5* and *NOTCH1* are transcription factors with well-known importance in driving the formation of leukaemia and other malignancies (Lobry et al., 2011; O'Brien et al., 2011). Mutations in the *NOTCH1* gene are linked with poor-prognosis in CLL (Puente et al., 2011; Villamor et al., 2013); however, mutations have not been previously described in the non-coding domain of *NOTCH1.* Thus, in the context of CLL, WGS gives a more complete picture of the role and possible sources of *NOTCH1* deregulation because functional mutations in non-coding areas previously ignored by capture methods are examined. In addition to identifying SNV and InDel mutations, WGS is a powerful tool for identifying large structural variations such as chromosomal rearrangements and fusion genes. WGS performed in colorectal adenocarcinoma demonstrated an average of 75 somatic rearrangements per tumour and identified novel, functional fusion proteins that had remained unidentified (Bass et al., 2011).

As we move towards a better understanding of genes and pathways implicated in cancer, we move towards a new era in personalised medicine. The rapid increase in availability of cost-effective NGS technology has led to a new frontier in medicine via unprecedented insight into the cancer genome. Integrated NGS investigations have proved to be effective in identifying and defining the mutational landscapes of human malignancies, improving clinical evaluation and interventions.

1.7 Hypothesis and aims

MYC overexpression is a hallmark of B cell malignancies, and it is well understood that at least one secondary mutation is required for neoplastic transformation. This 'second hit' must be an inactivating or activating mutation in one of the many genes that restrain or augment oncogenic MYC signalling, respectively. Defining these secondary mutations in MYC-driven malignancies is of clinical importance, as they can potentially be therapeutically targeted. The E*μ-Myc* model of B cell lymphoma/leukaemia faithfully recapitulates parallel human disease and has been used as a powerful tool in identifying secondary mutations that cooperate with MYC. However, the concomitant Myc-cooperative mutations required for spontaneous E*μ-Myc* lymphomagenesis have only been identified in up to 50% of all cases (Eischen et al., 1999; Schmitt et al., 1999).

The hypothesis of this thesis is that using next generation sequencing technology in the context of E*μ-Myc* will describe novel secondary mutations that cooperate with Myc in driving B cell lymphomagenesis. In testing this hypothesis, the first comprehensive characterisation of the genomic landscape of E*μ-Myc* lymphoma was performed; identifying novel putative cancer genes implicated in Myc-driven B cell lymphomagenesis.

2 Materials and Methods

2.1 Reagents

2.1.1 Solutions and buffers

The solution and buffers utilised in this study are summarized in Table 2-1.

2.1.2 Cell lines and culture media

The cell lines and respective culture mediums used in this study are highlighted in Table 2-2.

2.1.3 Vectors and vector inserts

A full list of vectors and vector inserts used in this study are available in Table 2-3.

2.1.4 List of primers used

A complete list of oligonucleotide primers utilized in this study is summarized in Table 2-4. Oligonucleotide primers used in this study were purchased from IDT Technologies, Coralville, IA, USA – unless otherwise specified.

2.1.5 PCR cycles

A full list of PCR cycles used in this study is available in Table 2-5.

2.1.6 List of antibodies used

A complete list of antibodies utilised in this study is summarised in Table 2-6.

2.2 Cell culture

2.2.1 Tissue culture methods

2.2.1.1 Culture of Eμ-Myc lymphomas

E*μ-Myc* mice that were sacrificed due to the signs and symptoms of progressive lymphoma had their spleens removed at necropsy. Under aseptic conditions, the spleens of E*μ-Myc* mice were homogenized by mechanical abrasion between two frosted microscope slides in the presence of sterilised phosphate buffered saline (PBS). The splenic cell suspension was then centrifuged at 423g for 10 minutes and the pellet was re-suspended in Anne Kelso's modified Dulbecco's modified Eagle's medium (AKMDMEM) with additional supplements highlighted in Table 2-2, which is referred to as E*μ-Myc* media herein. Generation of E*μ-Myc* cell lines was attempted from primary E*μ-Myc* splenocytes while the majority were cryogenically stored for later use. Previously generated E*μ-Myc* lymphoma cell lines were obtained from sources summarised in Table 2-2. Lymphomas were cultured in 6-well tissue culture plates (CELLSTAR®, Greiner Bio-One, Kremsmünster, Austria) in E*μ-Myc* media. Established E*μ-Myc* lymphoma-derived cell lines were passaged at a 1:10 dilution 3-times weekly, whereas primary E*μ-Myc* lymphomas were passaged at a slower rate of 1-time weekly and at a larger dilution of 1:5 until a stable E*μ-Myc* cell line was established. Once a stable Eμ-*Myc* cell line was established, several small aliquots were cryogenically stored in freeze media (Table 2-1).

2.2.1.2 Adherent cell lines and cultures

Phoenix-Eco packaging cells, referred to as Phoenix-E cells herein, (a gift from Marc Pellegrini, Walter and Eliza Hall Institute [WEHI], Melbourne, VIC, Australia) were cultured in DMEM with additional supplements as highlighted in Table 2-2. Phoenix-E cells were cultured in T-75 or T-25 cell culture flasks (CELLSTAR®, Greiner Bio-One) and were passaged twice-weekly by aspiration of the media, removing non-adherent cells and cell debris. Sterile PBS was added to the flasks, which were then mechanically agitated to release adherent cells from the flask and into suspension. The suspension of Phoenix-E cells was collected, centrifuged at 423g for 4 minutes and resuspended in 10ml of fresh Phoenix-E media before using 1-2ml of the suspension to seed a new T-75 or T-25 cell culture flask. Mouse embryonic fibroblasts (MEF) cell lines used in this study were parental NIH/3T3 (ATCC, Manassas VA, USA) and transduced NIH/3T3. These cell types were cultured in DMEM and additional supplements highlighted in Table 2-2 (herein referred to as MEF media). MEF cell lines were passaged three-times weekly and were cultured in T-75 or T-25 cell culture flasks. To passage MEFs, the media was aspirated to remove non-adherent cells and cell debris. Following media removal, 5ml of TrypLE™ Express (Gibco®, Invitrogen Life Technologies Corp.) was added and the flask was incubated at 37°C for 5 minutes. Adherent MEFs were then mechanically agitated to suspend the previously adherent cells. 10ml of sterile 4°C PBS was added to MEF flasks, inhibiting the TrypLE™ reaction. The PBS/MEF cell line suspension was centrifuged at 423g for four minutes upon which the cell pellet was re-suspended in 10ml of fresh MEF cell lines media and 1-2ml of the suspension was used to seed a new T-75 or T-25 flask.

Table 2-1 Buffers and solutions

Buffer	Components
Standard Buffers and solutions	
PBS	137mM NaCl 2.7mM KCl 10mM Na_2HPO_4 1.8mM KH_2PO_4 dH_2O Autoclave-sterilised
TBE – pH 8.0	89mM Tris-HCl 89mM Boric Acid 2mM EDTA
LB	0.1%w/v tryptone 0.5%w/v yeast extract 0.1%w/v NaCl dH_2O Autoclave-sterilised
Lysis Buffers	
ACK lysis buffer	0.15M NH_4Cl 10mM $KHCO_3$ 0.1mM EDTA H_2O
RIPA lysis Buffer – pH 7.4	150mM NaCl 1%v/v NP-40 0.5%w/v sodium deocycholate 0.1%w/v SDS 50mM Tris-HCl 1mM EDTA
Retroviral transduction buffers	
2x HeBS – pH 7.05	50mM HEPES 280mM NaCl 1.5mM Na_2HPO_4 anhydrous dH_2O filter sterilised

HEPES water	1.25mM HEPES dH_2O filter sterilised
$CaCl_2$ solution	7.3%w/v $CaCl_2$ dH_2O filter sterilised
Western blotting buffers	
5x loading buffer – pH 6.8	250mM Tris-HCl 50%v/v glycerol 10%w/v SDS 2.5% v/v β-mercaptoethanol 0.001w/v bromophenol blue
SDS running buffer – pH 8.3	25mM Tris-HCl 129mM glycine 0.1% w/v SDS dH_2O
Wet transfer buffer – pH 8.3	25mM Tris-HCl 192mM glycine 15%v/v methanol dH_2O
Semi-dry transfer buffer	48mM Tris-HCl 39mM glycine 0.04%w/v SDS 10%v/v methanol dH_2O
Gel buffers	
Agarose gel	1%w/v agarose 0.002%v/v ethidium bromide TBE
Resolving gel buffer – pH 8.8	1M Tris-HCl dH_2O
Polyacrylamide Gel - Resolving	2ml resolving gel buffer 1.6-2.4ml 40% acrylamide 80ul of 10%w/v SDS

	made up to 7.980ml with dH_2O 75µl 10% APS 25µl TEMED
Stacking gel buffer – pH 6.8	0.5M Tris-HCl dH_2O
Polyacrylamide gel - stacking	1.5ml stacking gel buffer 645µl 40% acrylamide 80µl 10%w/v SDS made up to 3.05ml with dH_2O 30µl 10% APS 10µl TEMED
Blocking buffers	
Skim-Milk blocking buffer	5%w/v skim milk PBS
BSA blocking buffer	5%w/v BSA PBS
Miscellaneous buffers	
Freeze media	10%v/v DMSO FCS
FACS buffer	1%v/v FCS 0.1%w/v sodium acetate 2mM EDTA PBS Filter sterilised
Hypotonic KCl	75mM KCl dH_2O
MACS buffer	2mM EDTA 0.5%w/v BSA PBS Filter sterilised

2.2.2 Isolation of primary tissues

2.2.2.1 Isolation of C57BL/6 MEF

For the generation of primary C57BL6 MEF, mice were plug-mated and pregnant females were sacrificed by cervical dislocation 13.5-14.5 days post coitum. Under aseptic conditions uterine horns were removed and placed on a sterile 10cm^2 tissue culture dish (CELLSTAR®, Greiner Bio-One, Kremsmünster, Austria), on ice. Embryos were carefully dissected from embryonic sacs and placed into 1ml ice cold PBS (Invitrogen Life Technologies Corp.) in individual wells of a 24 well tissue culture plate (CELLSTAR®, Greiner Bio-One). In a 10cm^2 culture dish the placental tissue was removed from each embryo and the embryos were sacrificed by decapitation and foetal livers were discarded. The remaining embryonic tissue was then suspended in 3mL of TrypLE™ Express (Gibco®, Invitrogen Life Technologies Corp.) per embryo in a 10-50mL conical tube and incubated at room temperature (RT) for 5-15 minutes. Embryonic tissue was homogenised by forced passage through a 20-gauge needle and filtered through a 70µm strainer (BD Falcon^{TM}, BD Biosciences, Bedford, MA, USA), which was washed with 10ml DMEM. Then, the embryonic tissue suspension was centrifuged for 4 minutes at 423g. Following centrifugation, the supernatant was aspirated and the cell pellet washed in a further 10-30ml of DMEM. MEF cell pellets were then resuspended in primary MEF media (Table 2-2), transferred into a T-75 or T-25 cell culture flask (CELLSTAR®, Greiner Bio-One) and cultured overnight. Culture media, debris and non-adherent cells were aspirated from the flask, 7-15ml primary MEF media was added and cells were cultured until a confluent monolayer formed. When MEFs were 75-90% confluent they were expanded to larger flasks and cultured

until confluent. At this point primary MEFs were cryogenically stored (in FBS containing 10% Dimethyl sulfoxide (DMSO)) or propagated further for experimental use.

2.2.2.2 Isolation of C57BL/7 foetal liver stem cells

For the generation of primary Eμ-*Myc* foetal liver stem cells, 6+ weeks old male Eμ-*Myc* mice were paired and plug-mated with 6+ weeks old WT C57BL/6 females. Pregnant females were culled by cervical dislocation at 13.5-14.5 days post-coitum. Under aseptic conditions, the uterine horns were removed and placed on ice in a sterile 10cm^2 tissue culture dish (CELLSTAR®, Greiner Bio-One). The embryos were dissected from the embryonic sacs and placed in 1ml of 4°C PBS (Invitrogen Life Technologies Corp.) in wells of a 24-well tissue culture plate (CELLSTAR®, Greiner Bio-One). The placental sacs were removed from the embryos, which were then transferred to a new 10cm^2 tissue culture dish and were sacrificed by decapitation. The head of the embryo was then placed in a DNase-free 1.5ml tube (Eppendorf, Hamburg, Germany) and snap frozen for genotype analysis. Using two forceps, the foetal livers were removed from the embryo intact and placed in a 2ml cryopreservation tube (Sigma-Aldrich, St. Louis, MO, USA). 1ml of freeze media (FBS containing 10% DMSO) (Table 2-1) was added to the foetal livers and using a P1000 pipette set to 1ml aspiration volume, the foetal livers were homogenised by vigorous mechanical agitation. At this point, the foetal livers were cryogenically stored for use at a later date – pending genotyping results (where WT or foetal livers of ambiguous transgene status were discarded), or were used immediately for experiments.

2.2.2.3 Extraction and isolation of bone marrow-derived cells

Mice were culled by cervical dislocation and, under aseptic conditions, their hind-legs were removed and placed in PBS in a 10cm^2 tissue culture dish (CELLSTAR®, Greiner Bio-One) kept at 4°C. Using scissors and forceps; the skin, muscle connective tissue and deep fascia were removed from each hind-leg, exposing a completely intact femur, tibia and fibula. The bones were then cleaned with a dry clean tissue to remove any remaining tissue. The femur was cut marginally below the ball of the hip and marginally above the kneecap upon which, using a 26 gauge syringe, the bone marrow cells were extracted from the femur by forced flushing of 2-3ml of PBS through the femoral lumen. The fibula was removed from the lower hind-leg and the tibia was cut marginally below the kneecap and above the ankle, upon which the bone marrow cells were extracted by forced flushing with a 26-gauge needle and 2-3ml of PBS through the lumen of the fibula bone. The suspension was homogenised by mechanically agitation and flushed through a 70µm filter using a 26-gauge syringe. The single cell BM cell suspension was then centrifuged at 423g for 4 minutes. The cell pellet was re-suspended in 5ml ACK lysis buffer (Table 2-1) incubated at 4°C for 30 minutes to specifically deplete erythrocytes. Following erythrocyte depletion, the cells were centrifuged at 423g for 4 minutes, and the remaining ACK lysis buffer was aspirated. The cell pellet was then resuspended in 20ml PBS and centrifuged at 423g for 4 minutes to remove excess ACK lysis buffer. Cells were used immediately for experimental purposes or cryogenically stored at -80°C in freeze media (Table 2-1) for later use.

2.2.2.4 Extraction and isolation of splenocytes

Mice were culled by cervical dislocation and, under aseptic conditions, spleen was removed and placed in PBS in a 10cm^2 tissue culture dish (CELLSTAR®, Greiner Bio-One) and kept at 4°C. Splenocytes were extracted into PBS by mechanical shearing of the intact spleen using two frosted glass microscope slides, centrifuged at 423g for 4 minutes and resuspended in ACK lysis buffer (Table 2-1) for 30 minutes at 4°C. Following erythrocyte depletion, the cells were centrifuged at 423g for 4 minutes, and the remaining ACK lysis buffer (Table 2-1) was aspirated. The cell pellet was then resuspended in 20ml PBS and centrifuged at 423g for 4 minutes to remove excess ACK lysis buffer (Table 2-1). Cells were used immediately for experimental purposes or cryogenically stored at -80°C in freeze media (Table 2-1) for later use.

2.2.2.5 Extraction and isolation of lymph node lymphocytes

Mice were culled by cervical dislocation and the axillary, brachial, inguinal and mesenteric lymph nodes were aseptically removed and placed in PBS in a 10cm^2 tissue culture dish (CELLSTAR®, Greiner Bio-One) and incubated on ice. Lymphocytes were isolated from lymph nodes and stored as in 2.2.2.4 above.

Table 2-2 Media and cell lines

Culture media – supplements – $C0_2$	Cell type	Genetic manipulation	Source
MEF media			
DMEM - 10% FBS, 2mM L-glutamine, 50μM β2-ME - 10% CO_2	NIH/3T3	*Wild type*	Purchased from ATCC, Manassas, VA, USA
		pLMS-*Bcor*.sh1 transfected	Transfected by Dr. Libby Kruse, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia
		pLMS-*Bcor*.sh2 transfected	
		pLMS-*Bcor*.sh3 transfected	
		pLMS-*Bcor*.sh4 transfected	
		pLMS-*Bcor*.sh5 transfected	
		pLMS-*Bcor*.sh6 transfected	
		pLMS-*Bcor*.sh7 transfected	
		pLMS-*Bcor*.sh8 transfected	
		pLMS-*Bcor*.sh9 transfected	
		pQCIG-*BcorG2* transfected	Transfected by Dr. Lev Kats, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia
		pQCIG-*Trp53*(b) transfected	
		pQCIG-Scram transfected	
Primary MEF media			
DMEM - 10% FBS, 2mM L-glutamine, 2 mM pen/strep, 50μM β2-ME - 10% CO_2	Primary C57BL/6 MEF	*Wild type*	
Phoenix-E media			
DMEM - 10% FBS, 2mM L-glutamine, 2 mM pen/strep - 10% CO_2	Phoenix-E	*Wild type*	Nolan Laboratory, Stamford University, CA
		pLMS.scram transfected	
		pLMS-*Bcor*.sh9 transfected	
		pLMS-*Trp53*.1224 transfected	
		MSCV transfected	Transfected by Mr. Stuart Craig, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia
		MSCV-Index transduced	

		MSCV-*Bcor* transfected	Transfected by Mr. Ben Martin, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia
		pQCIG-*Bcor*G2 transduced	Transfected by Dr. Lev Kats, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia
		pQCIG-*Trp53*(b) transduced	
		pQCIG-Scram transduced	
Eμ-*Myc* media			
Anne Kelso modified DMEM – 10% FBS, 2mM L-asparagine, 2nM pen/strep, 50μM 2-ME – 10% CO_2	E*μ-Myc* lymphoma cell lines	*Wild type*	
		pLMS.scram transduced	
		pLMS-*Bcor*.sh9 transduced	
		pLMS-*Trp53*.1224 transduced	
		MSCV-Index transduced	Transduced by Mr. Stuart Craig, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia
		MSCV transduced	Transduced by Mr. Ben Martin, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia
		MSCV-*Bcor* transduced	
+ 2ng/mL IL-3, 2ng/mL IL-6 and 10ng/mL SCF	E*μ-Myc* primary foetal liver derived cells	pLMS.scram transduced	
		pLMS-*Bcor*.sh9 transduced	
		pLMS-*Trp53*.1224 transduced	
		pQCIG-*Bcor*G2 transduced	Transduced by Dr. Lev Kats, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia
		pQCIG-*Trp53*(b) transduced	
		pQCIG-Scram transduced	
Human cell line media			
RPMI 1640 - 10% FBS, 2mM L-glutamine, 1.5g/L sodium bicarbonate, 4.5g/L glucose, 10nM HEPES,	NAMALWA	*Wild type*	Transduced by Mr. Ben Martin, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia
		MSCV	

1mM sodium pyruvate, 2mM pen/strep - 10% CO_2		MSCV.*Bcor*	

2.3 Retroviral transduction

2.3.1 Vectors

2.3.1.1 shRNA design

cDNA sequences in FASTA format were imported into a long oligo design tool (http://biodev.cea.fr/DSIR/DSIR.html). RNA design was set to siRNA 21nt, the score threshold was set to 90 and 4 or more nucleotide runs and immunostimulatory motifs were filtered out. The target sequence for each oligo was found with the base immediately 5' of the target sequence noted and its complementary base added immediately 3' of the oligo. A mismatched base was then added immediately 5' of the target sequence oligo. The components of the short-hairpin RNA (shRNA) were concatenated using Microsoft Excel and were ordered from Integrated DNA Technologies (IDT) (Coralville, IA, USA). A list of shRNA oligos can be found in table Table 2-4.

2.3.1.2 Cloning shRNA into vectors

The shRNA had cloning sites introduced by EcoR1 and XhoI (Promega, Madison, WI, USA) and in a 1:5 vector-insert molar ratio were ligated into a vector backbone containing ampicillin resistance motif using T4 ligase (Pomega, Madison, WI, USA), as per manufacturer's instructions, and incubated overnight at RT. Vectors used in this study are described in table Table 2-3.

2.3.1.3 Transformation

10µl of plasmid DNA was added to 90µl of competent *E. Coli* (TOP10F) (Life Technologies, ThermoFisher Scientific, Waltham, MA, USA) suspension and incubated at 4°C for 30 minutes. The bacterial suspension was then incubated at 42°C for 90 seconds, upon which 500µl of lysogeny broth (LB) was added, followed by incubation at 37°C under slight agitation for 30 minutes. The bacterial cultures were centrifuged at 13,000g for 1 minute and the pellet was resuspended in 200µl of LB of which 100µl was seeded evenly across an LB-based bacterial growth plate that had been treated with ampicillin. The plates were incubated over night at 37°C.

2.3.1.4 Generation of high-copy plasmids from E. coli

Colonies were harvested from the ampicillin-containing LB-based bacterial growth plate and were transferred to 3ml LB/0.1% ampicillin and incubated overnight at 37°C under gentle agitation. The bacterial culture (1 mL) was utilised for copy number amplification by mini-preparation using the NucleoSpin® plasmid kit (Macherey-Nagle, Düren, Germany) as per manufacturers instruction. The resultant plasmid DNA for the colony was then prepared for Sanger sequencing using BigDye® terminator v3.1 chemistry (Life Technologies, Thermo Fisher, Scientific, Waltham, MA, USA) (Table 2-5) to confirm correct integration of shRNA into the vector backbone.

2.3.1.5 Sanger Sequencing

Sanger sequencing was used primarily to show correct integration of shRNA into a vector backbone; however, it was used in several other instances to show a mutation in

genomic DNA where next-generation sequencing results were ambiguous. A list of sequencing primers used in Sanger sequencing reactions is shown in Table 2-4. 5µl of DNA template (~100-600ng of DNA) was combined with 3.2pmol/µl of sequencing oligo, 2µl of BigDye® terminator, 4µl of BigDye® buffer (Life Technologies, Thermo Fisher, Scientific, Waltham, MA, USA) and 5.8µl of H_2O. The reaction was performed at the conditions highlighted in Table 2-5 on an MJ Research PTC-200 thermal cycler (GMI, Ramsey, MN, USA) and following completion, 16µl of H_2O was added to the reactions. In a separate 1.5ml tube (Eppendorf, Hamburg, Germany) 10µl of 3.0M sodium acetate (NaA) [pH 4.8-5.2] and 54µl of 94%v/v ethanol (EtOH) was combined. The PCR reactions were transferred to the 1.5ml tube and the samples were vigorously mixed by bench-top vortex and incubated at 4°C for 30 minutes. The samples were centrifuged for 20 minutes at 13,000g, the supernatant was discarded and the pellets resuspended in 70%v/v EtOH. The samples were centrifuged at 13,000g for 5 minutes and the supernatant was removed and air-dried to remove any residual EtOH. Samples were sent to the Melbourne Translational Genomics Platform (University of Melbourne, Melbourne, AUS) to undergo capillary-based sequencing. The sequence was checked with 4peaks V 1.7.2 (4Peaks by A. Griekspoor and Tom Groothuis, mekentosj.com.) to ensure correct insertion of the shRNA or the cDNA construct.

2.3.2 Retrovirus production and infection

2.3.2.1 Calcium phosphate transfection of packaging cell line

To generate retroviral supernatants, Phoenix-E (Table 2-2) packaging cells were seeded at $7.5x10^5$ in a $10cm^2$ tissue culture dish (CELLSTAR®, Greiner Bio-One) in

10mL of Phoenix-E media (Table 2-2) and were grown to 80% confluence. 3 hours prior to transfection, the media was aspirated and fresh media was applied without disturbing the adherent layer of Phoenix-E cells. Transfection reagent was prepared by diluting 20µg of plasmid DNA in 250µl of HEPES-buffered H_2O (Table 2-1) and the vector constructs were added to 250µl of 0.5M calcium chloride ($CaCl_2$) solution (Table 2-1). In a separate tube air was slowly bubbled through 500µl of 2xHEBS (Table 2-1) as the DNA-$CaCl_2$ solution was slowly added in a drop-wise manner. When the DNA-$CaCl_2$ solution has been added to the 2×HEBS solution, the mixture was incubated at RT for 20 minutes. The complete (1ml) transfection media was then distributed evenly across the Phoenix-E culture and the plate was gently agitated. The Phoenix-E media was replenished 24hrs following addition of complete transfection reagent.

Table 2-3 List of vectors used

Gene ID	Vector backbone	Insert	Construct ID
shRNA constructs			
Bcor	pLMS	TGCTGTTGACAGTGAGCGACGAGAAGAAGCAGACACTAAATAGTGAAGCCACAGAT GTATTTAGTGTCTGCTTCTTCTCGGTGCCTACTGCCTCGGA	pLMS-*Bcor*.sh9
Trp53	pLMS	CTCGAGAAGGTATATTGCTGTTGACAGTGAGCGCCCACTACAAGTACATGTGTAATAGTGAAGCCAC AGATGTATTACACATGTACTTGTAGTGGATGCCTACTGCCTCGGAATTC	pLMS-*Trp53*.1224
CRISP/Cas9 constructs			
Bcor	pQCIG	CAGTGGCTGGGCCAAGCCGT	pQCIG-*Bcor*G2
Trp53	pQCIG	GAAGTCACAGCACATGACGG	pQCIG-*Trp53*(b)

2.3.2.2 Coating plates with retronectin

Under aseptic conditions, retronectin (ClonTech, Mountain View, CA, USA) was added to each well in a 6-well non-treated plate (Corning, Corning, NY, USA) at a final concentration of 30µg/ml in PBS. The retronectin-coated plates were incubated at RT for 2 hours under gentle agitation and the retronectin was aspirated. The plates were washed with PBS and were replenished with 1ml of a blocking agent that was made up of sterile 2% bovine serum albumin (BSA) fraction V (Roche, Basel, Switzerland) in PBS (Table 2-1) and stored at 4°C overnight. Immediately prior to use, the retronectin-coated plates were washed in PBS to remove the blocking agent.

2.3.2.3 Harvesting viral supernatant

48 hours following transfection the viral supernatant was removed from the Phoenix-E cell culture and passed through a 0.45µm filter using a 50ml syringe, removing any cell debris from the media. The Phoenix-E cell culture that was previously transfected was replenished with fresh media and was further propagated. The filtered supernatant was placed in Amicon Ultra-15 Centrifugal Filter Units (Merck Millipore, Billerica, MA, USA) and centrifuged at 3000g for 12 minutes to concentrate the virus. Concentrated virus was made up to 1.5ml with Phoenix-E media and was transferred to the previously retronectin-coated plates. The plates were sealed and centrifuged at 2000g for 1 hour upon which the media is aspirated. This procedure is repeated with fresh retronectin-coated plates on the following day.

2.3.2.4 Retroviral transduction

Approximately 1-5x10^6 of the target cells were resuspended in 1ml of their appropriate media (for this study, E*μ-Myc* foetal liver cells, Eμ-*Myc* tumour derived cell lines and human lymphoma cell lines were used as per Table 2-2) and were added to the wells that have been preloaded with retronectin and virus. The plates were centrifuged at 2000g for 1 hour and were incubated at 37°C for 4-6 hours before being made up to a total volume of 5ml using the relevant fresh media. This process was repeated on the following day using fresh plates that had been preloaded with retronectin and virus. 24 hours following consecutive retroviral infections the cells were removed from the plates containing the virus and were centrifuged at 423g for 4 minutes, resuspended in 5ml of fresh media and incubated in the appropriate conditions where they were propagated for at least two days prior to use. Transduction efficiency was assessed by flow cytometry and was based on expression of fluorescent markers that are expressed only in transduced cells (GFP, Venus, dsRED and BFP were used in this study).

2.4 Biochemical techniques

2.4.1 Ribonucleic acid

RNA extractions, polymerase chain reaction (PCR) and quantitative real-time polymerase chain reaction (QRT-PCR) were carried out using diethyl pyrocarbonate (DEPC) treated H_2O, aerosol resistant pipette tips and RNase and pyrogen free plastics. DNA extractions and preparations were carried out using sterile

UltraPure™ DNase/RNase-Free Distilled Water (Life Technologies, Waltham, MA, USA), aerosol resistant sterile pipette tips and, where relevant, were conducted in the allocated pre- or post-PCR laboratory areas.

2.4.1.1 RNA extraction using TRIzol®

For RNA extractions using TRIzol® (Life Technologies, Waltham, MA, USA), prepared cell pellets were washed in 4°C PBS and centrifuged at 423g for 4 minutes before being resuspended in 1ml TRIzol® reagent. The samples were incubated for 5 minutes at RT and, at this point, samples could be stored at -80°C for later extraction or used immediately. 200µl of chloroform was added to the samples, which were mixed by vigorously passing the solution through an 18 gauge blunt needle with a 3ml syringe and were then allowed to incubate for a further 2-3 minutes. The samples were then centrifuged at 12,000g for 15 minutes and the aqueous phase was collected into a 1.5ml tube (Eppendorf, Hamburg, Germany) where 500µl of 100% isopropanol was added and the samples incubated for 15 minutes at RT. After centrifugation at 12,000g for 10 minutes, the supernatant was removed, H_2O/70%v/v ethanol was added and the samples were vigorously agitated and then centrifuged at 8,000g for 5 minutes. Supernatant was removed, the pellets were air dried and then dissolved in DEPC treated H_2O following which samples were stored at -80°C or RNA concentrations and purity were assessed by NanoDrop spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA) at 260/280nM.

2.4.1.2 RNA extraction using RNeasy Kits

RNA extractions using the RNeasy blood and tissue kit (Qiagen, Hilden, Germany) were performed followed the manufacturer's instructions; however, during the final step, the dry RNA pellets were dissolved in DEPC treated H_2O rather than the recommended elution buffer. Prior to samples being stored -80°C, RNA concentrations and purity were assessed by spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA) at 260/280nM.

2.4.1.3 DNA extraction using DNeasy Kits

DNA was extracted from cells using the DNeasy blood and tissue kit (Qiagen, Hilden, Germany) according to manufacturer's instructions. Briefly, cells were twice washed in PBS followed by centrifugation at 423g for 4 minutes. The cell pellets were resuspended in 200µl PBS and the DNeasy blood and tissue kit was used for the lysis and collection of DNA. Importantly, RNase A (Sigma-Aldrich, St. Louis, MO, USA) was added at the step where Proteinase K was added to the samples and finally, the DNA was eluted into DNase/RNase-Free Distilled Water (Life Technologies, Waltham, MA, USA) as opposed to the provided elution buffer. The DNA was stored at -20°C or was immediately assessed by spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA) at 260/280nM.

2.4.1.4 Polymerase chain reaction (PCR)

Oligonucleotide primers specific to genes of interest (Table 2-4) were designed using Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0/) and were checked for

target-specificity using UCSC *in silico* PCR tool (http://genome.ucsc.edu/). PCR reactions were performed in twin.tec® PCR Plates (Eppendorf, Hamburg, Germany) using 10ng genomic DNA template, 10ul of 2 x Phusion® high-fidelity PCR master-mix (New England Biolabs, Ipswich, UK), 2.5µl of 4µM forward and reverse oligonucleotide primers and made up to a total reaction volume of 20µl with H_2O. PCR reactions were performed using Tetrad2® thermo-cycler (Bio-Rad, Hercules, CA, USA) at conditions summarised Table 2-5. To demonstrate that the PCR reactions generated a DNA product, 5µl of product was mixed with 1µl of 6x DNA loading dye (Thermo Fisher Scientific, Waltham, MA, USA), loaded into TBE/1%w/v agarose gel containing 2µl of ethidium bromide solution (Sigma-Aldrich, St. Louis, MO, USA) and separated by gel-electrophoresis in TBE buffer. GeneRuler 100bp DNA ladder (Thermo Fisher Scientific, Waltham, MA, USA) was simultaneously run to estimate sizes of PCR product bands for each PCR reaction.

2.4.1.5 Bead purification of DNA

PCR reaction products in twin.tec® PCR Plates (Eppendorf, Hamburg, Germany) were incubated with Agencourt AMPure XP magnetic beads (Beckman Coulter, Pasadena, CA, USA) in a ratio of 1:0.9 at RT for 10 minutes, allowing the magnetic beads to bind DNA fragments that are greater than 200bp. The PCR plates were then placed on a magnetic block and incubated for 5 minutes at RT. The bead-free supernatant was removed, leaving the DNA-bound beads magnetically bound in the plate. While still on the magnetic block, each well was then twice washed with 70%v/v ethanol, upon which the bead pellet is air-dried of residual ethanol. The

plate was then removed from the magnetic block and the dry bead pellets were resuspended in 20µl H_2O and incubated for 10 minutes at RT while under slight agitation. The plate was then placed back on a magnetic block separating the now DNA-unbound beads and H_2O. The DNA-H_2O solution was aspirated and collected for gel electrophoresis to validate size of DNA products.

2.4.1.6 Quantitative real time polymerase chain reaction (QRT-PCR)

Oligonucleotide primers specific to genes of interest (Table 2-4) were designed using Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0/) and were checked for target-specificity using UCSC *in silico* PCR tool (http://genome.ucsc.edu/). In a MicroAmp® fast optical 96-well reaction plate (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) 2.5pmol of the target gene specific primer was added to each well in triplicate. 10ng of genomic DNA and 5µl of SYBR® green PCR master-mix (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) was added to each well containing the target gene specific primers. Each well was made up to 10µl of total reaction volume with H_2O and the plate was sealed and briefly centrifuged. The PCR was performed on StepOnePlus Real Time PCR System (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) using the StepOne Software v2.1 and conditions detailed in Table 2-5. Data was analysed in Microsoft Excel using the $2^{-\Delta\Delta C_T}$ method (Livak and Schmittgen, 2001) with the control gene being the ribosomal gene *L32*. All primers used were validated against *L32* and gave a single-peak dissociation curve, indicating that the target gene-specific primers were single gene specific.

Table 2-4 Primers used

Sample	Chr:Base	Gene	FWD PRIMER	REV PRIMER
Targeted amplicon sequencing primers				
ML362	4:53062556	Abca1	ACACTGACGACATGGTTCTACAGATTTTGTCACTGCTACACTGGC	TACGGTAGCAGAGACTTGGTCTCCCCTTTCAGGAGAGTCTAGAAC
ML52	16:20333612	Abcc5	ACACTGACGACATGGTTCTACACGGGAACATGAAGATGACAGGTA	TACGGTAGCAGAGACTTGGTCTAGACTTACTGATCCAGGAGACCA
ML73	13:8757281	Adarb2	ACACTGACGACATGGTTCTACACTGAACACACTCTCTTCTTCCCA	TACGGTAGCAGAGACTTGGTCTAGTAGAAACTGATCTTGCTCCGG
ML52	2:93677449	Alx4	ACACTGACGACATGGTTCTACAGTCTCCGACTTCTTGAGTGTCTC	TACGGTAGCAGAGACTTGGTCTTTCTGTCATGTGGCCCAAGAGAT
ML358	13:19125060	Amph	ACACTGACGACATGGTTCTACAGGTTGTCCAGAATTCTGCAATGT	TACGGTAGCAGAGACTTGGTCTTCCTTTGGCTCCTCCATCTCTAG
ML61	10:69932398	Ank3	ACACTGACGACATGGTTCTACAGCCTGTTCTTAAAATTGCATGCC	TACGGTAGCAGAGACTTGGTCTTTACCTGCTCTTTGGATGGTACC
ML73	18:36560930	Ankhd1	ACACTGACGACATGGTTCTACACGTGAGGCTCTTTGGGGAG	TACGGTAGCAGAGACTTGGTCTTCCTCGTCGCTCCCAGAG
ML21	X:12048475	Bcor	ACACTGACGACATGGTTCTACACTTCAGTGATGTTAGTCCCCTGG	TACGGTAGCAGAGACTTGGTCTTGTAATGAGCTGACCAACCTGAA
ML33	X:12048474	Bcor	ACACTGACGACATGGTTCTACACTTCAGTGATGTTAGTCCCCTGG	TACGGTAGCAGAGACTTGGTCTTGTAATGAGCTGACCAACCTGAA
ML352	X:12047721	Bcor	ACACTGACGACATGGTTCTACATCTTTAATTTTCTGCTGTTTGGCA	TACGGTAGCAGAGACTTGGTCTGTAACGCCTGCTTCTTCTGC
ML43	X:12040487	Bcor	ACACTGACGACATGGTTCTACAGGGTCAGCACCATAGGAAAG	TACGGTAGCAGAGACTTGGTCTGGGTGCTCACTGTAGGTGGT
ML39	7:3339009	Cacng7	ACACTGACGACATGGTTCTACATTCCCTATGGTCAGTCTCTTCCT	TACGGTAGCAGAGACTTGGTCTAAGAAGGTTCTTCCCCTAGCTCC
ML358	2:155979186	Cep250	ACACTGACGACATGGTTCTACAATTCTAAGTCCAGACAAGCTCCC	TACGGTAGCAGAGACTTGGTCTTTCCATTTCTGTCTGCTGCTCTC
ML33	6:125110108	Chd4	ACACTGACGACATGGTTCTACAATCTTCCTTTGCTTCTGTCACCA	TACGGTAGCAGAGACTTGGTCTTTCTCAGTCACTAGTTGTCCTGC
ML73	9:10144172	Cntn5	ACACTGACGACATGGTTCTACAAAGAGTGGCTTCTCGACTAAGAA	TACGGTAGCAGAGACTTGGTCTTAAGATGGCTTCGGAATGGAACT
ML21	15:55448833	Col14a1	ACACTGACGACATGGTTCTACATATCTGTCTCACCACTGTAGGGT	TACGGTAGCAGAGACTTGGTCTTGTATGACTGACTTTAGCATCTTGC
ML63	6:4539595	Col1a2	ACACTGACGACATGGTTCTACACAAGGAAATGGCAACTCAGCTC	TACGGTAGCAGAGACTTGGTCTTCGACTAGGACAGAGTAGGTGAA
ML43	1:45376784	Col5a2	ACACTGACGACATGGTTCTACAAGTCGTTTGACCCTTTAAGGACC	TACGGTAGCAGAGACTTGGTCTCCAGTCTCCTAACACAGCTATCA
ML362	4:141034522	Crocc	ACACTGACGACATGGTTCTACATGCCAGCTTCACTGCCAG	TACGGTAGCAGAGACTTGGTCTCCAAGGCTCCTGCTCATTGTC
ML43	4:141020458	Crocc	ACACTGACGACATGGTTCTACAGCTGCAGCTCCAGTTTGATG	TACGGTAGCAGAGACTTGGTCTTCACACAGTCCTTGAGGCTAAC
ML52	5:52166753	Dhx15	ACACTGACGACATGGTTCTACATGAGTGACGGTATCTCATAGGTT	TACGGTAGCAGAGACTTGGTCTAAGCGTGAAGTTGATGATTTGGG
ML30	2:180689463	Dido1	ACACTGACGACATGGTTCTACAGACACCGGCACATTCTTTTTCC	TACGGTAGCAGAGACTTGGTCTCCACCAGTAAGGAGTTCAGGAAA

ML358	8:125088279	Disc1	ACACTGACGACATGGTTCTACACAGGCTTGGCGGACTTGG	TACGGTAGCAGAGACTTGGTCTCTCCCATTCCCTGAGCAATGTAT
ML20	14:31266689	Dnahc1	ACACTGACGACATGGTTCTACAATGGGAAGGAGGGTAAGCTGATA	TACGGTAGCAGAGACTTGGTCTATTGTCAGTAATGCCAAGATGCG
ML20	11:34267958	Dock2	ACACTGACGACATGGTTCTACACAACAGTTAGCCTTGCTTGTCTC	TACGGTAGCAGAGACTTGGTCTTGGAGAAGTTGCTAGATTACCGG
ML52	9:107179149	Dock3	ACACTGACGACATGGTTCTACAATAGGCTCACACTTACCTTCACA	TACGGTAGCAGAGACTTGGTCTATTCAGTGCTTAATTTAGTGACTCTCA
ML20	9:21960597	Epor	ACACTGACGACATGGTTCTACACGATCTAAGTCAGACCTCAGGTG	TACGGTAGCAGAGACTTGGTCTCTAGTTCCTCGGGGTGCAG
ML33	6:47530602	Ezh2	ACACTGACGACATGGTTCTACATCAGAGAGAAGGGGAAGAGGTAG	TACGGTAGCAGAGACTTGGTCTCGGGTGTGTGTATGTTAGAGACA
ML63	19:44952170	Fam178a	ACACTGACGACATGGTTCTACAAGGCTTCCTTACCACTGCATATC	TACGGTAGCAGAGACTTGGTCTATTCCTTCACTCCTGAGCCATTT
ML363	16:36969962	Fbxo40	ACACTGACGACATGGTTCTACAGTCTTTCCAGAACACCATCTTGC	TACGGTAGCAGAGACTTGGTCTTGTTTTAACAGGTTCCTTGGGGA
ML52	18:25090683	Fhod3	ACACTGACGACATGGTTCTACAACCTTCCTGGGTTTGCCAC	TACGGTAGCAGAGACTTGGTCTCCGGATCGTCTTCTTTTTCTTGG
ML58	18:25090683	Fhod3	ACACTGACGACATGGTTCTACAACCTTCCTGGGTTTGCCAC	TACGGTAGCAGAGACTTGGTCTCCGGATCGTCTTCTTTTTCTTGG
ML63	18:25090683	Fhod3	ACACTGACGACATGGTTCTACAACCTTCCTGGGTTTGCCAC	TACGGTAGCAGAGACTTGGTCTCCGGATCGTCTTCTTTTTCTTGG
ML33	3:53549751	Frem2	ACACTGACGACATGGTTCTACAAATTCCCTCCAGTATTGGTCGTC	TACGGTAGCAGAGACTTGGTCTTTTTTAATTTAGTCCCTATGTAAACAATGT
ML52	11:53000205	Fstl4	ACACTGACGACATGGTTCTACACACAAGAAGTTAAGGCCTGTGTG	TACGGTAGCAGAGACTTGGTCTCTGTCGTCCTATTGAGCAAGCA
ML39	3:80860235	Glrb	ACACTGACGACATGGTTCTACAACCAGTGCCTTTATAGTATTTTGTAC	TACGGTAGCAGAGACTTGGTCTTGATATGCATATTTTCCTCTAACTTGT
ML352	5:110181880	Golga3	ACACTGACGACATGGTTCTACAGAAAGCAAGTCCAGCAGTGATG	TACGGTAGCAGAGACTTGGTCTAAAGTTCACGTCCTCCTCCAAAA
ML358	2:21648994	Gpr158	ACACTGACGACATGGTTCTACATGCTGTTCTTAGGAATGAGGCAT	TACGGTAGCAGAGACTTGGTCTAACTATGGACAATCAAGACAGGT
ML69	15:76698591	Gpt	ACACTGACGACATGGTTCTACAATTCATTCAAGAAGGTGCTCACG	TACGGTAGCAGAGACTTGGTCTCCAGGAGAGCATGAGTTGAGG
ML352	6:64429755	Grid2	ACACTGACGACATGGTTCTACATGTCCAAGCAAACAGATATCCCT	TACGGTAGCAGAGACTTGGTCTAGTACCTTTTGAATGCCTGCTTG
ML363	6:64429755	Grid2	ACACTGACGACATGGTTCTACATGTCCAAGCAAACAGATATCCCT	TACGGTAGCAGAGACTTGGTCTAGTACCTTTTGAATGCCTGCTTG
ML61	11:75169452	Hic1	ACACTGACGACATGGTTCTACACTGGAGACTCAGATATCGGACCT	TACGGTAGCAGAGACTTGGTCTCTGGGTGCTATGCTAGGATGATT
ML63	11:75169441	Hic1	ACACTGACGACATGGTTCTACACTGGAGACTCAGATATCGGACCT	TACGGTAGCAGAGACTTGGTCTCTGGGTGCTATGCTAGGATGATT
ML69	11:75167080	Hic1	ACACTGACGACATGGTTCTACAGAGTAGCAAGCCTGGATGACG	TACGGTAGCAGAGACTTGGTCTCTACCTGCAGATCCCTGACCTC
ML52	5:35391883	Hmx1	ACACTGACGACATGGTTCTACAGGAGACCGGCGAGGAGAT	TACGGTAGCAGAGACTTGGTCTGTCTTCTTCCTGCGGCCA
ML27	17:32011205	Hsf2bp	ACACTGACGACATGGTTCTACAACTCCAGTGTGAACATTGTCCTT	TACGGTAGCAGAGACTTGGTCTTTTTCAACATAACTGGCCAGACG
ML20	13:105445458	Htr1a	ACACTGACGACATGGTTCTACACCTACCTTTCTGTGAGAGCAGTT	TACGGTAGCAGAGACTTGGTCTCACTACCACCACCATCATCATCA
ML363	16:34220177	Kalrn	ACACTGACGACATGGTTCTACACACCTGCTCAGAGGTCTTGTAAA	TACGGTAGCAGAGACTTGGTCTCTCTTTCATGCCACTTCCTTGC

ML358	6:145246771	Kras	ACACTGACGACATGGTTCTACAGAGCAGCGTTACCTCTATCGTAG	TACGGTAGCAGAGACTTGGTCTTGATAATCTTGTGTGAGACATGTTCT
ML363	6:145234351	Kras	ACACTGACGACATGGTTCTACATCACATGCCAACTTTCTTATTCAACT	TACGGTAGCAGAGACTTGGTCTTCTCAGGACTCCTACAGGAAACA
ML58	6:145234355	Kras	ACACTGACGACATGGTTCTACATCACATGCCAACTTTCTTATTCAACT	TACGGTAGCAGAGACTTGGTCTTCTCAGGACTCCTACAGGAAACA
ML367	6:145234355	Kras	ACACTGACGACATGGTTCTACATCACATGCCAACTTTCTTATTCAACT	TACGGTAGCAGAGACTTGGTCTTCTCAGGACTCCTACAGGAAACA
ML58	6:145234355	Kras	ACACTGACGACATGGTTCTACATCACATGCCAACTTTCTTATTCAACT	TACGGTAGCAGAGACTTGGTCTTCTCAGGACTCCTACAGGAAACA
ML63	15:101859872	Krt77	ACACTGACGACATGGTTCTACATTGGGAAAAGATAGGTGAGGGAG	TACGGTAGCAGAGACTTGGTCTTAGTTACGTCTCCAGCAGCAGAA
ML43	15:101532226	Krt84	ACACTGACGACATGGTTCTACATGGATGAGCTCTTTTTGAAGGGA	TACGGTAGCAGAGACTTGGTCTCAGCTGTAACTGTTAACCAGAGC
ML21	2:180198502	Lama5	ACACTGACGACATGGTTCTACACAAGGTAGGGCTACAAACTGCA	TACGGTAGCAGAGACTTGGTCTCTGGATCAGCAGTGTGGAGTG
ML30	2:180198585	Lama5	ACACTGACGACATGGTTCTACACTACATTCCTGACAAGTGGCACC	TACGGTAGCAGAGACTTGGTCTGCAGAGAATTAGAGCAAGGAGGA
ML73	16:24762155	Lpp	ACACTGACGACATGGTTCTACACCTCCAACCAAGGACGCTATTAT	TACGGTAGCAGAGACTTGGTCTGATATAGGTCTTCTTGGGGCCAC
ML43	2:157379189	Manbal	ACACTGACGACATGGTTCTACACTTAGAAGTTGGTGCTGGAGAGT	TACGGTAGCAGAGACTTGGTCTCTTACCGCCTCATGGGACTTG
ML30	3:82383336	Map9	ACACTGACGACATGGTTCTACAAAGGAAGCAAAGAGAATAGCTGC	TACGGTAGCAGAGACTTGGTCTAAAAATCCTGCATTCAAGCCCAG
ML362	10:37136594	Marcks	ACACTGACGACATGGTTCTACAGTCGCTCCCTCTGCTTCAG	TACGGTAGCAGAGACTTGGTCTGTCCGCCTCCTCCACGTC
ML58	10:37136871	Marcks	ACACTGACGACATGGTTCTACAAGCCTCATCCTTTTCGGCC	TACGGTAGCAGAGACTTGGTCTAGGAGAATGGCCACGTAAAAGT
ML20	11:74836441	Mnt	ACACTGACGACATGGTTCTACATTCTGCTAAGACCCTTGTGTGTT	TACGGTAGCAGAGACTTGGTCTCGGTGGAGACAAAGGCAGT
ML63	5:37824276	Msx1	ACACTGACGACATGGTTCTACACTCCTCATCTGTGCCCATGG	TACGGTAGCAGAGACTTGGTCTGATCCAGGGCTGTCTCGAG
ML39	4:148491684	Mtor	ACACTGACGACATGGTTCTACATTCTGTGCTTTACTCTGTAGGGC	TACGGTAGCAGAGACTTGGTCTATGATAACAAGTGTCCATGGCCT
ML352	11:67364919	Myh13	ACACTGACGACATGGTTCTACAATCACTCCTGGTTTCCCTTGAAG	TACGGTAGCAGAGACTTGGTCTGTCCGCTCTGTCTGCTCC
ML61	11:67361360	Myh13	ACACTGACGACATGGTTCTACACTAGGAAAAAGCTGGCTCAGAGG	TACGGTAGCAGAGACTTGGTCTGGGAAAGACCTTGTCGAAGTTTC
ML363	18:74644048	Myo5b	ACACTGACGACATGGTTCTACATGCTAGGAATAGAACACAGCCAG	TACGGTAGCAGAGACTTGGTCTACAATCCAGGAGAACAGCTGAG
ML39	3:27226123	Nceh1	ACACTGACGACATGGTTCTACAGGCCACCTTTGTCGTAATTCATT	TACGGTAGCAGAGACTTGGTCTACACACTCTTCTGGTTTCTGTGT
ML30	11:98328076	Neurod2	ACACTGACGACATGGTTCTACACCTTGGTCATCTTGCGTTTCTTC	TACGGTAGCAGAGACTTGGTCTTGGAGGAGAAGAGATCCCTGAAC
ML30	19:46081352	Nolc1	ACACTGACGACATGGTTCTACATTGCTTTTGTTTCTCAGGTACAC	TACGGTAGCAGAGACTTGGTCTCCTGCTACAAAGTTACAGCCCTA
ML27	3:103060270	Nras	ACACTGACGACATGGTTCTACAAAAGATTGAGCCTGTCCTTGTCC	TACGGTAGCAGAGACTTGGTCTTGATGGCAAATACACAGAGGAAC
ML20	12:89187318	Nrxn3	ACACTGACGACATGGTTCTACACGACTATGTTAACCTGGCTCTGA	TACGGTAGCAGAGACTTGGTCTACTTTCCTGCCCCACAAAGTATC
ML362	2:86829079	Olfr1094	ACACTGACGACATGGTTCTACACCGGCTTCACAATCCAATGTATT	TACGGTAGCAGAGACTTGGTCTACAAAGCGATCATAAGCCATTGC

ML352	7:105368397	Olfr692	ACACTGACGACATGGTTCTACAGAACAATATCCCCATGAGACCCA	TACGGTAGCAGAGACTTGGTCTCTAGTAGGGCCAGCAGATACATG
ML358	14:50681366	Olfr747	ACACTGACGACATGGTTCTACACCATTACTGCCAGGAAGAGACAT	TACGGTAGCAGAGACTTGGTCTCACGTCTACATACCCCCATGTAC
ML73	10:129321677	Olfr780	ACACTGACGACATGGTTCTACAACTTTTTCTGCAACTTGCTCTGG	TACGGTAGCAGAGACTTGGTCTAGGTGAAATGTGGCATAAACTGG
ML30	4:138898051	Otud3	ACACTGACGACATGGTTCTACAGCAGTATCTGAGCATGGGTACTA	TACGGTAGCAGAGACTTGGTCTGTTCCAGATGCTTCATCAAGACG
ML358	17:26198143	Pdia2	ACACTGACGACATGGTTCTACACCTACTCCACACCTGTGTACTTG	TACGGTAGCAGAGACTTGGTCTCAAAGAACTAGCTCCCGAGTACA
ML52	4:43019849	Pigo	ACACTGACGACATGGTTCTACATGTGGAGTAGAAGGTCTGGGTAG	TACGGTAGCAGAGACTTGGTCTGAAGTTTGGGGTAGCACGTGA
ML63	2:60553094	Pla2r1	ACACTGACGACATGGTTCTACAGGTTCCTGGATGTGAGTCAAATC	TACGGTAGCAGAGACTTGGTCTAAGAGTGGTAGCTTCTCTTGCC
ML363	15:76172181	Plec	ACACTGACGACATGGTTCTACAAAAAAGAAGACCCCAGTGCCTC	TACGGTAGCAGAGACTTGGTCTTCTTCCTCCTACTCTTCCTCAGG
ML20	17:12418768	Plg	ACACTGACGACATGGTTCTACATGATTGAGAACAAGGTGTGCAAC	TACGGTAGCAGAGACTTGGTCTTGTGTTGTCTCTCTGTGGGATTC
ML52	10:79741779	Polrmt	ACACTGACGACATGGTTCTACATGAAACAGGCTGTCATGACTTCT	TACGGTAGCAGAGACTTGGTCTAAGCAGATGATGGAAGAAGGCTT
ML21	4:22487697	Pou3f2	ACACTGACGACATGGTTCTACAGAAAACAGAGGTCCCAGGAAAGA	TACGGTAGCAGAGACTTGGTCTAAAGAGAAAAGGATGACCCCTCC
ML73	4:22487697	Pou3f2	ACACTGACGACATGGTTCTACAGAAAACAGAGGTCCCAGGAAAGA	TACGGTAGCAGAGACTTGGTCTAAAGAGAAAAGGATGACCCCTCC
ML358	4:143896971	Pramef6	ACACTGACGACATGGTTCTACAGCCTATCCCAAGGTATGCAGATT	TACGGTAGCAGAGACTTGGTCTGTGATCTGTGAGAAGCTGGAGTT
ML33	2:92975407	Prdm11	ACACTGACGACATGGTTCTACAGCAGCTTTCCAGACTTAAGCATG	TACGGTAGCAGAGACTTGGTCTCACTGGAAGATGAAGAGGAGGAG
ML358	7:105482092	Prkcdbp	ACACTGACGACATGGTTCTACATGTATGCGACGCACGGAG	TACGGTAGCAGAGACTTGGTCTTCAGAGATCATGGGGGAGAGC
ML367	2:155754341	Procr	ACACTGACGACATGGTTCTACATAGTTCCTCTCACTGTTAGCTGC	TACGGTAGCAGAGACTTGGTCTGTTGTAGGCATTGAGCTGCTTC
ML21	7:127019838	Prrt2	ACACTGACGACATGGTTCTACACCTGTACTTTCTGTCAGGACCTC	TACGGTAGCAGAGACTTGGTCTGCATCCAAACCAGATGTGAACAG
ML43	6:29171026	Prrt4	ACACTGACGACATGGTTCTACAGACAGGAAGGCAGCGAGAAAG	TACGGTAGCAGAGACTTGGTCTCTGGCTTGCCTGCTCCTG
ML358	4:148252775	Ptchd2	ACACTGACGACATGGTTCTACACACATCTGTATGACCCTTACCGA	TACGGTAGCAGAGACTTGGTCTTAGGTGTATAGAGCGCCTTTTGG
ML58	15:73320835	Ptk2	ACACTGACGACATGGTTCTACAGTTAACTCAAGAGTCTGCATGGG	TACGGTAGCAGAGACTTGGTCTGCCTTTCCTAGTTGTCTCTCCTG
ML43	3:103886315	Ptpn22	ACACTGACGACATGGTTCTACAGCCTCTTCTACAACCTCCTTCTT	TACGGTAGCAGAGACTTGGTCTAAAGAAGCAGAGGAAGAGAAGGG
ML39	4:75957197	Ptprd	ACACTGACGACATGGTTCTACACTTAAATTCGAGCTCCATTCCGG	TACGGTAGCAGAGACTTGGTCTGCCCAGAGGAATTACATGGTACA
ML352	4:129084144	Rnf19b	ACACTGACGACATGGTTCTACAAGACATCACCTCAGATGAGTGTG	TACGGTAGCAGAGACTTGGTCTGATGCAGTTACAACAAGTGTGCT
ML352	X:74271704	Rpl10	ACACTGACGACATGGTTCTACATGGATTCAGAGCTGGTTTAGTCC	TACGGTAGCAGAGACTTGGTCTAAAGTTCTCGAGAAGTGGGGAAT
ML39	4:136052666	Rpl11	ACACTGACGACATGGTTCTACACTTTCCAGATCCACTCCTCTCTG	TACGGTAGCAGAGACTTGGTCTTTGATGGCTGTGTCCCTCTTC
ML39	18:89103865	Rttn	ACACTGACGACATGGTTCTACACTTTCCAGATCCACTCCTCTCTG	TACGGTAGCAGAGACTTGGTCTTTGATGGCTGTGTCCCTCTTC

ML52	5:124393962	Sbno1	ACACTGACGACATGGTTCTACATAGAGCTCTCATAGTCGCTTTCG	TACGGTAGCAGAGACTTGGTCTCACACACACACCTCAAAACAT
ML52	11:113874603	Sdk2	ACACTGACGACATGGTTCTACAAGGCAGCTAGACTTTGCTCAAA	TACGGTAGCAGAGACTTGGTCTACGTCACCTCTTCTCAGATGTTG
ML63	6:82939530	Sema4f	ACACTGACGACATGGTTCTACAGGACAAGGGTTCCTGCATCTTA	TACGGTAGCAGAGACTTGGTCTCAGAGATGCTGGCCAGGG
ML362	6:82939530	Sema4f	ACACTGACGACATGGTTCTACAGGACAAGGGTTCCTGCATCTTA	TACGGTAGCAGAGACTTGGTCTCAGAGATGCTGGCCAGGG
ML20	5:123158704	Setd1b	ACACTGACGACATGGTTCTACAGTAGTGGCTATGGCAGAAGAGAG	TACGGTAGCAGAGACTTGGTCTCTGCATCTCAGGGACTTCGTC
ML362	11:72253399	Slc13a5	ACACTGACGACATGGTTCTACAACTCCTCGTTCAGCACTTTGTAG	TACGGTAGCAGAGACTTGGTCTGGAGAGAGGAGCTGATGGAAATC
ML20	17:57053552	Slc25a23	ACACTGACGACATGGTTCTACAAGCCCTTGTGTTTACTATCTGCT	TACGGTAGCAGAGACTTGGTCTCAAAGTCCAACCGGCTTAACATT
ML52	14:111681332	Slitrk5	ACACTGACGACATGGTTCTACATGGAGGATTACAAAGACCTGCAC	TACGGTAGCAGAGACTTGGTCTGCTCCTCAAATGGTGGCTTTC
ML30	11:96035030	Snf8	ACACTGACGACATGGTTCTACAGGCGACAGAACCTGGGAC	TACGGTAGCAGAGACTTGGTCTATTCCGGGAGAATTCAGTGTCTC
ML73	2:72971051	Sp3	ACACTGACGACATGGTTCTACACATTGTCTGAGAACTTCCCGAGA	TACGGTAGCAGAGACTTGGTCTTCCAGAATCTCATACCACAGACTG
ML362	10:121925886	Srgap1	ACACTGACGACATGGTTCTACACAGGGTGTATGGAGTCTACCTTT	TACGGTAGCAGAGACTTGGTCTAAACCTGTTGTCTCCAGTGAACT
ML63	5:135835279	Srrm3	ACACTGACGACATGGTTCTACAGAGCCCGGCTTGGTGAAG	TACGGTAGCAGAGACTTGGTCTAAGTGCTGAACAGTTTGATCTCG
ML63	1:39386007	Tbc1d8	ACACTGACGACATGGTTCTACAGTGACTACGGCACACACAATAC	TACGGTAGCAGAGACTTGGTCTTGACAGAGGAGATAGAGAGGGAC
ML367	7:30226919	Tbcb	ACACTGACGACATGGTTCTACAGTCACAGTGACAAAAAGACACGT	TACGGTAGCAGAGACTTGGTCTGTTCGATCCTTCATGAAACGCAG
ML58	1:19234346	Tfap2b	ACACTGACGACATGGTTCTACAGTGACTACGGCACACACAATAC	TACGGTAGCAGAGACTTGGTCTCCTCCAAAGCCAATCCAATATGT
ML39	16:4552012	Tfap4	ACACTGACGACATGGTTCTACACTTGAACAACACCACTAACAGGC	TACGGTAGCAGAGACTTGGTCTTCTCCTGAACCTACTCCTGTTCT
ML21	8:13370913	Tfdp1	ACACTGACGACATGGTTCTACATTCTCTCCATCTGTGTGGGGAAT	TACGGTAGCAGAGACTTGGTCTGCCGATCCATTTGATCTCCTTCT
ML61	8:13370913	Tfdp1	ACACTGACGACATGGTTCTACATTCTCTCCATCTGTGTGGGGAAT	TACGGTAGCAGAGACTTGGTCTGCCGATCCATTTGATCTCCTTCT
ML73	8:13370913	Tfdp1	ACACTGACGACATGGTTCTACATTCTCTCCATCTGTGTGGGGAAT	TACGGTAGCAGAGACTTGGTCTGCCGATCCATTTGATCTCCTTCT
ML20	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG
ML21	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG
ML27	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG
ML30	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG
ML39	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG
ML52	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG
ML61	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG
ML69	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG

ML73	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG
ML352	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG
ML353	4:43689482	Tmem8b	ACACTGACGACATGGTTCTACAAGGTGTGAATTTCCAAGCCATTG	TACGGTAGCAGAGACTTGGTCTAGACTAACCTGCTTGATCACAGG
ML73	7:19514441	Trappc6a	ACACTGACGACATGGTTCTACATTCCTGCTGACAAGACCAAGG	TACGGTAGCAGAGACTTGGTCTCGTCCATGTGCTTCTGGAACAT
ML30	11:69588388	Trp53	ACACTGACGACATGGTTCTACAACCTTGACACCTGATCGTTACTC	TACGGTAGCAGAGACTTGGTCTTGTGACTTCTTGTAGATGGCCAT
ML362	11:69589547	Trp53	ACACTGACGACATGGTTCTACAATATGACAAGAGGGGTTGGGAAC	TACGGTAGCAGAGACTTGGTCTAATTTTCTTCTTCTGTACGGCGG
ML61	11:69588505	Trp53	ACACTGACGACATGGTTCTACACCTGTGCAGTTGTGGGTCA	TACGGTAGCAGAGACTTGGTCTGAATAAGTCAGAAGCCGGGAGAT
ML73	11:69588401	Trp53	ACACTGACGACATGGTTCTACAACCTTGACACCTGATCGTTACTC	TACGGTAGCAGAGACTTGGTCTTGTGACTTCTTGTAGATGGCCAT
ML362	2:35979552	Ttll11	ACACTGACGACATGGTTCTACAGTCCTTGGGCTGCTCCTC	TACGGTAGCAGAGACTTGGTCTCAGCAGCCCGGAGAAGAAG
ML39	2:76784591	Ttn	ACACTGACGACATGGTTCTACAGCTGGGATTCTGAGAGTCTTC	TACGGTAGCAGAGACTTGGTCTACTTCAAGTTTTCTAGGGGGAGA
ML73	2:76749137	Ttn	ACACTGACGACATGGTTCTACACTTCACAGACTCAGCTGTTTCG	TACGGTAGCAGAGACTTGGTCTCGAGAGTCCACAAGAAAGGCATA
ML21	5:130269166	Tyw1	ACACTGACGACATGGTTCTACACCATTTGCATGTTTGTGGATTGT	TACGGTAGCAGAGACTTGGTCTCCATGTTTGCTCTGAACTGCATT
ML27	5:130269166	Tyw1	ACACTGACGACATGGTTCTACACCATTTGCATGTTTGTGGATTGT	TACGGTAGCAGAGACTTGGTCTCCATGTTTGCTCTGAACTGCATT
ML33	5:130269166	Tyw1	ACACTGACGACATGGTTCTACACCATTTGCATGTTTGTGGATTGT	TACGGTAGCAGAGACTTGGTCTCCATGTTTGCTCTGAACTGCATT
ML39	5:130269166	Tyw1	ACACTGACGACATGGTTCTACACCATTTGCATGTTTGTGGATTGT	TACGGTAGCAGAGACTTGGTCTCCATGTTTGCTCTGAACTGCATT
ML43	5:130269166	Tyw1	ACACTGACGACATGGTTCTACACCATTTGCATGTTTGTGGATTGT	TACGGTAGCAGAGACTTGGTCTCCATGTTTGCTCTGAACTGCATT
ML52	5:130269166	Tyw1	ACACTGACGACATGGTTCTACACCATTTGCATGTTTGTGGATTGT	TACGGTAGCAGAGACTTGGTCTCCATGTTTGCTCTGAACTGCATT
ML58	5:130269166	Tyw1	ACACTGACGACATGGTTCTACACCATTTGCATGTTTGTGGATTGT	TACGGTAGCAGAGACTTGGTCTCCATGTTTGCTCTGAACTGCATT
ML61	5:130269166	Tyw1	ACACTGACGACATGGTTCTACACCATTTGCATGTTTGTGGATTGT	TACGGTAGCAGAGACTTGGTCTCCATGTTTGCTCTGAACTGCATT
ML352	5:130269166	Tyw1	ACACTGACGACATGGTTCTACACCATTTGCATGTTTGTGGATTGT	TACGGTAGCAGAGACTTGGTCTCCATGTTTGCTCTGAACTGCATT
ML358	17:31213788	Ubash3a	ACACTGACGACATGGTTCTACATATGACAGGGTTTTCTGTTGCAG	TACGGTAGCAGAGACTTGGTCTAGACCTCATGTGCTCTGTTCTTT
ML21	11:23471554	Usp34	ACACTGACGACATGGTTCTACATTTGCAGATGTGAAACAGGCATT	TACGGTAGCAGAGACTTGGTCTACAAAACAAAAAGCCAAACAAACA
ML30	11:23471554	Usp34	ACACTGACGACATGGTTCTACATTTGCAGATGTGAAACAGGCATT	TACGGTAGCAGAGACTTGGTCTACAAAACAAAAAGCCAAACAAACA
ML33	11:23471554	Usp34	ACACTGACGACATGGTTCTACATTTGCAGATGTGAAACAGGCATT	TACGGTAGCAGAGACTTGGTCTACAAAACAAAAAGCCAAACAAACA
ML39	11:23471554	Usp34	ACACTGACGACATGGTTCTACATTTGCAGATGTGAAACAGGCATT	TACGGTAGCAGAGACTTGGTCTACAAAACAAAAAGCCAAACAAACA
ML58	11:23471554	Usp34	ACACTGACGACATGGTTCTACATTTGCAGATGTGAAACAGGCATT	TACGGTAGCAGAGACTTGGTCTACAAAACAAAAAGCCAAACAAACA
ML63	11:23471554	Usp34	ACACTGACGACATGGTTCTACATTTGCAGATGTGAAACAGGCATT	TACGGTAGCAGAGACTTGGTCTACAAAACAAAAAGCCAAACAAACA

ML73	11:23471554	Usp34	ACACTGACGACATGGTTCTACATTTGCAGATGTGAAACAGGCATT	TACGGTAGCAGAGACTTGGTCTACAAAACAAAAAGCCAAACAAACA
ML352	11:23471554	Usp34	ACACTGACGACATGGTTCTACATTTGCAGATGTGAAACAGGCATT	TACGGTAGCAGAGACTTGGTCTACAAAACAAAAAGCCAAACAAACA
ML363	11:23471554	Usp34	ACACTGACGACATGGTTCTACATTTGCAGATGTGAAACAGGCATT	TACGGTAGCAGAGACTTGGTCTACAAAACAAAAAGCCAAACAAACA
ML358	11:23471554	Usp34	ACACTGACGACATGGTTCTACATTTGCAGATGTGAAACAGGCATT	TACGGTAGCAGAGACTTGGTCTACAAAACAAAAAGCCAAACAAACA
ML21	2:73447859	Wipf1	ACACTGACGACATGGTTCTACATTGGACTACGCAGCAAAAAGC	TACGGTAGCAGAGACTTGGTCTTCTAATGCACCAATGTGTTCTGT
ML58	7:99911809	Xrra1	ACACTGACGACATGGTTCTACAAAGCAATCTGAACAATGCCAGAC	TACGGTAGCAGAGACTTGGTCTCTCCTTCCGAGGCATCAACATAT
ML43	1:33772989	Zfp451	ACACTGACGACATGGTTCTACATGAAAGCCCCAAATAAACGTTCC	TACGGTAGCAGAGACTTGGTCTTGAACGGTTCTTACTGCATTCAT
ML363	7:24585730	Zfp575	ACACTGACGACATGGTTCTACACCAGTTTGGAAGGAAACGAGAAC	TACGGTAGCAGAGACTTGGTCTGGCCACCGTTCCAAACTGG
ML39	2:82256568	Zfp804a	ACACTGACGACATGGTTCTACAATTGGCTTCTCCTTTGCATTTCC	TACGGTAGCAGAGACTTGGTCTTACCTTTCTCCTCAGAGCTCAGA
ML52	2:167044192	Znfx1	ACACTGACGACATGGTTCTACACTTTCTTCTTCGAGGTCTCACCA	TACGGTAGCAGAGACTTGGTCTATCCTTTTCCTACCTGCTTCCTC
Sanger sequencing primers				
4242	X:12057555-12057982	Bcor	CCACATGCTGGATGACTGAT	N/A
6066	3: 103060270	Nras	N/A	TGGCAAATACACAGAGGAACC
pLMS	N/A	Mir30	N/A	ATGGTGGCGAATAAGTACGC
4242 A1	N/A	MSCV-Index	AATGATACGGCGACCACCGAGATCTACACTGACGACATGGTTCTACA	N/A
4242 A4	N/A	MSCV-Index	AATGATACGGCGACCACCGAGATCTACACTGACGACATGGTTCTACA	N/A
4242 C2	N/A	MSCV-Index	AATGATACGGCGACCACCGAGATCTACACTGACGACATGGTTCTACA	N/A
6066 D2	N/A	MSCV-Index	AATGATACGGCGACCACCGAGATCTACACTGACGACATGGTTCTACA	N/A
6066 D11	N/A	MSCV-Index	AATGATACGGCGACCACCGAGATCTACACTGACGACATGGTTCTACA	N/A
6066 F6	N/A	MSCV-Index	AATGATACGGCGACCACCGAGATCTACACTGACGACATGGTTCTACA	N/A
sHairpin	N/A	Mir30 5’	TGTTTGAATGAGGCTTCAGTAC	
QRT-PCR primers				
4242 A1	4:89276752-89276949	Cdkn2a	CTGCTCAACTACGGTGCAGA	TCGCACGATGTCTTGATGTC
4242 A4	4:89276752-89276949	Cdkn2a	CTGCTCAACTACGGTGCAGA	TCGCACGATGTCTTGATGTC
4242 C2	4:89276752-89276949	Cdkn2a	CTGCTCAACTACGGTGCAGA	TCGCACGATGTCTTGATGTC

6066 D2	4:89276752-89276949	Cdkn2a	CTGCTCAACTACGGTGCAGA	TCGCACGATGTCTTGATGTC
6066 D11	4:89276752-89276949	Cdkn2a	CTGCTCAACTACGGTGCAGA	TCGCACGATGTCTTGATGTC
6066 F6	4:89276752-89276949	Cdkn2a	CTGCTCAACTACGGTGCAGA	TCGCACGATGTCTTGATGTC
Control	chr6:115805514-115808743	L32	TTCCTGGTCCACAATGTCAAG	TGTGAGCGATCTCAGCAC

2.4.2 Protein techniques

2.4.2.1 Preparation of protein lysates

For preparation of total protein lysates from *in vitro* cell lines, 1-5x10^6 cells were harvested and twice washed in 4°C PBS by centrifugation at 423g for 4 minutes. The supernatant was aspirated and the dry pellets were either utilised immediately or were snap frozen in liquid nitrogen for storage at -80°C. Primary tissue was harvested as previously described (2.2.2) and were freshly prepared into a single cell suspension and twice washed in 4°C PBS by centrifugation at 423g for 4 minutes. Dry pellets were used fresh or were snap frozen in liquid nitrogen for storage at -80°C. Frozen and/or fresh cell pellets were lysed by resuspension in RIPA buffer (Table 2-1) with complete EDTA-free protease inhibitor cocktail (Roche Diagnostics Australia Pty. Ltd., Castle Hill, NSW, Australia] and phosSToP inhibitor cocktail [Roche Diagnostics Australia Pty. Ltd., Castle Hill, NSW, Australia) at a volume of 20µl of RIPA buffer for every 10^6 cells. The cells were lysed on ice for 30-60 minutes, lysates were then centrifuged at 13,000g for 15 minutes and quantified for total protein concentration using the bicinchoninic acid (BCA) assay (Thermofisher Scientific, Waltham, MA, USA) as per manufacturer's instructions. Lysates were then used immediately or stored at -80^{o}C prior to use.

2.4.2.2 Immunoblotting (western blotting)

Following protein concentration estimation, 5× loading buffer (Table 2-1) was added to the protein lysates. The samples were incubated at 95°C for 5 minutes and microcentrifuged briefly (3-5 seconds). 10-30µg of protein from each sample was

loaded into a 4-15% gradient Mini-PROTEAN®**TGX** polyacrylamide gel (Bio-Rad Laboratories, Hercules, CA, USA) or in laboratory-prepared polyacrylamide gel (Table 2-1) and separated in SDS running buffer (Table 2-1). Pre-stained molecular weight standards (10µL, PageRuler™ or PageRuler™ Plus Prestained Protein Ladder, Fermentas Life Sciences, Burlington, Ontario, Canada) were run on each gel to ascertain the molecular weight of proteins of interest. Gels were electroblotted onto Immobilon-P PVDF membrane (Millipore, Bedford, MA, USA) in either western transfer buffer (Table 2-1) in a wet transfer apparatus (Biorad, Hercules, CA, USA) at 100V for 90 minutes at 4^0C, or in semi-dry transfer buffer (Table 2-1) in a semi-dry transfer apparatus (Bio-Rad Laboratories, Hercules, CA, USA) at 0.07A for 90 minutes at RT. Membranes were then blocked in PBS/5%w/v skim milk powder for 1 hour at RT while under slight agitation. Blocked membranes were incubated with the appropriate antibody solution (Table 2-6) in PBS/5%w/v BSA at 4°C for 24 hours or RT for 1 hour. After incubation with primary antibody, the membranes were washed x3 in PBS for 15 minutes at RT and incubated in the appropriate HRP-conjugated secondary antibody (Table 2-6) in PBS/5%w/v skim milk powder for 1 hour at RT. Membranes were washed ×3 in PBS for 15 minutes at RT and incubated for 5 minutes at RT with enhanced chemoluminescence reagents (Amersham™ ECL™ or Amersham™ ECL™ Plus, GE Healthcare, Uppsala, Sweden), detected by FUJIFILM Super RX (FujiFilm, Brookvale, NSW, Australia) and developed with an Agfa CP1000 developer (Agfa, Mortsel, Belgium). All membranes were re-probed with an appropriate primary antibody to be used as a loading control parameter (Table 2-6).

Table 2-5 List of PCR protocols used

Temperature	Time	Cycles
Amplicon sequencing PCR 1		
95°C	2 minutes	X1
95°C	30 seconds	X9
68°C Decrease 1°C every cycle	30 seconds	
72°C	30 seconds	
95°C	30 seconds	X32
61°C	30 seconds	
72°C	30 seconds	
72°C	5 minutes	X1
12°C	*	
Amplicon sequencing PCR 2		
95°C	10 minutes	X1
95°C	15 seconds	X15
60°C	30 seconds	
72°C	1 minute	
72°C	3 minutes	X1
12°C	*	
Generic PCR		
95°C	5 minutes	X1
95°C	30 seconds	X35
60°C	30 seconds	
72°C	1 minute	
72°C	3 minutes	X1
12°C	*	
BigDye® incorporating PCR for Sanger Sequencing		
96°C	5 minutes	X1
96°C	10 seconds	X30
50°C	5 seconds	
60°C	4 minutes	
4°C	*	

2.5 Next generation sequencing

2.5.1 Preparation of samples, sequencing and analysis

All NGS work was performed in the Molecular Genomics Core Facility and the pathology department at the PMCC.

2.5.1.1 FICOLL purification and enrichment of lymphocytes

All E*μ-Myc* cells used for whole exon sequencing (WES) were sourced from freshly frozen lymph node or splenic suspensions that had not undergone *in vitro* passaging (2.2.2). Freshly thawed lymphocyte preparations were purified by FICOLL gradient. Briefly, the thawed lymphocytes were washed ×2 by centrifugation at 423g for 4 minutes and cell pellet resuspension in 3ml of fresh E*μ-Myc* media. Using a 3ml transfer pipette (Copan, Murrieta, CA, USA) the samples were gently dispensed onto 5ml of FICOLL paque (GE Healthcare, Little Chalfront, UK) contained in a 15ml Falcon™ conical centrifuge tube (ThermoFisher Scientific, Waltham, MA, USA). The tubes were centrifuged at 800g for 10 minutes under the slowest acceleration and deceleration settings. The interphase layer was transferred to a new tube, enriched with E*μ-Myc* media (Table 2-2) and washed ×2 by centrifugation at 423g for 4 minutes. The cells were resuspended in Eμ-*Myc* media and were prepared as described in 2.4.1.1, 2.4.1.2 and 2.4.1.3.

2.5.1.2 Whole genome sequencing

Cells were prepared as described in 2.5.1.1 and DNA was extracted using the methods detailed in 2.4.1.3. Approximately 1µg of gDNA was prepared for fragment-sequencing libraries and then processed for whole genome capture according to standard protocols used for TruSeq chemistry (Illumina, San Diego, CA, USA). LN88 and its germline control – tail88 – were sequenced with each sample occupying three lanes of an HiSeq™ 2000 (illumina, San Diego, CA, USA) resulting in deep, paired-end sequencing with an average of 30-fold read coverage across the entire genome. Medium coverage genomes were prepared as above but were sequenced with each sample occupying a single lane on an HiSeq™ 2000 (Illumina, San Diego, CA, USA) giving an average read depth of 10-fold across the entire genome. Low coverage genomes were sequenced in an indexed pool of 3 samples occupying a single lane of the HiSeq™ (Illumina, San Diego, CA, USA) resulting in 3-fold average coverage across the genome. Reads were aligned using Bowtie2 v2.1.0 (Johns Hopkins University, Baltimore, MD, USA) to mouse reference genome mm10 (Dec. 2011 GRCm38/mm10).

2.5.1.2.1 Transgene assembly topology

The topology of the E*µ-Myc* transgene was determined in collaboration with Assoc. Prof. Tony Papenfuss and Dr. Jan Schroeder who performed analysis of the WGS data generated from chapter 2.5.1.2. Briefly, duplicates were identified and removed from the aligned data using Picard's Mark Duplicates v1.89 (Broad Institute, Boston, MA, USA). BreakDancer v1.3.5 (Chen et al., 2009), CREST v0.0.1 (Wang et al., 2011) and Socrates v0.9.5 (Schröder et al., 2014) were employed to identify structural variants in

the data. Full whole genome *de novo* assembly was undertaken using Gossamer (Conway et al., 2012) to compare resulting contigs with the genomic sequence relevant to the E*μ-Myc* transgene.

2.5.1.3 Whole exon sequencing

Cells were prepared as described in 2.5.1.1 and DNA was extracted as detailed in 2.4.1.3. Approximately 1μg of gDNA was prepared for fragment-sequencing libraries and then processed for exome capture according to standard protocols used for SureSelect Mouse All Exon chemistry (Agilent Technologies, Santa Clara, CA, USA). Each sample was sequenced as part of an indexed pool of 3 samples on a single lane of a HiSeq™ 2000 (Illumina, San Diego, CA, USA) using a paired-end sequencing strategy giving approximately $6x10^7$ 100bp reads per sample. This approach gave a mean sequence depth of 100-fold and at least 10-fold coverage in 90% of targeted bases. By combining WES data from tumour samples and an appropriate germline control that was representative of the genetic background of the subject animal(s), alignment of short read data and calling of sequence variants including single base changes, small insertions and deletions was done using a somatic pipeline. The variant calling methods were based on Genome Analysis Tool Kit (GATK) (DePristo et al., 2011) and muTect (Cibulskis et al., 2013), with the variants annotated using the ENSEMBL database (Hubbard et al., 2002).

2.5.1.3.1 Inferring copy number variations from whole exome sequencing data

The Aberration Detection in Tumour Exome (ADTex) (Amarasinghe et al., 2014) tool was used to detect copy number variations and to resolve polyploidy/aneuploidy from WES data. This allowed for a cost effective analysis, as usually to generate this type of data WGS is required. From the 16 tumours in the prospective E*μ-Myc* cohort, ML33, ML353 and ML52 had undergone LC-WGS and were selected to act as controls to generate a reference that the 13 remaining tumours were compared to. This generated a consensus CNV plot for each case, which was then combined for all cases to generate a cumulative plot showing regions of loss and gain across the genome.

2.5.1.4 RNA-sequencing

Cells were prepared as discussed in 2.5.1.1 and RNA was extracted using the methods highlighted in 2.4.1.1. 1μg of RNA was used generate polyA enriched cDNA libraries using TruSeq sample preparation kit (Illumina, San Diego, CA, USA) and paired-end RNA-sequencing (2x150bp) was performed on the HiSeq 2000 sequencing platform (Illumina, San Diego, CA, USA), generating, on average, ~$4x10^7$ reads per sample. RNA-sequencing reads were aligned using the short read aligner software – TopHat (Trapnell et al., 2012). The raw RNA-seq data was converted to feature counts using HTSeq software package allowing generation of expression matrices collating read counts per gene based on ENSEMBL annotation (Hubbard et al., 2002). EdgeR v3.0 (Bioconductor, Fred Hutchinson Cancer Research Center, Seattle, WA, USA) statistical software was used in the statistical programming tool R (free publicly available software

http://www.r-project.org/) to normalise gene expression and perform statistical analysis by linear regression (Law et al., 2014).

2.5.1.5 Targeted amplicon sequencing

2.5.1.5.1 Filtering candidate genes from exome capture resequencing data

A filtered gene list was generated based on the exome sequencing data for each prospective E*μ-Myc* mouse that was included in the sequencing study. WES was performed on the E*μ-Myc* tumours as described in chapter 2.5.1.3. Once the sequencing data was analysed, the results were filtered based on the consequence of the mutation effecting the canonical gene in a predicted deleterious manner and whether the mutated gene had a human orthologue that is found to be mutated in cancer cases based on its presence in the COSMIC database (COSMIC, The Wellcome trust Sanger Institute, London, UK). From this gene-list, all genes that are recurrently mutated, regardless of variant allele frequency or read depth, were included whereas mutations in genes that were found to be non-recurrent were only included if they had a variant allele frequency (VAF) of ≥0.2. Once a gene list was finalised, oligonucleotide primers flanking the variant base by 100-200 bases were developed using primer3 software and UCSC *in silico* PCR tool as previously described in 2.4.1.4. The oligonucleotide primers were then modified by addition of a universal forward tag (5'-ACACTGACGACATGGTTCTACA-3') at the 3' end of the primer and a universal reverse tag (5'-TACGGTAGCAGAGACTTGGTCT-3') immediately 5' of the respective oligonucleotide primers.

2.5.1.5.2 Targeted amplicon two-stage PCR

A PCR was performed to incorporate the forward and reverse universal tags into the PCR product for each amplicon. Targeted amplicon PCR were completed using oligonucleotide primers with the universal forward and reverse tags as per section 2.5.1.5.1, with PCR cycles detailed in Table 2-5. The PCR products were validated and purified as per sections 2.4.1.4 and 2.4.1.5, respectively. The purified and validated PCR products were diluted 1:100 with H_2O and pooled together provided all samples were predicted to have different mutations. The pooled amplicons were plated out in twin.tec® PCR Plates (Eppendorf, Hamburg, Germany) using 1µl of the previously generated DNA template, 10ul of 2 x Phusion® high-fidelity PCR master-mix (New England Biolabs, Ipswich, UK), 2.5µl of 4µM forward and reverse oligonucleotide primers and made up to contain a total reaction volume of 20µl with H_2O. The oligonucleotide primers (Illumina, Sand Diego, CA, USA) that were used in the second PCR contain a motif that recognizes and binds to the forward and reverse universal tags in the PCR products, but also contained a unique barcode sequence that integrated into the product arising from this PCR (Figure 2.01). The PCR was performed at conditions highlighted Table 2-5, the products were purified a second time as per section 2.4.1.5 and the generation of products was validated as described in 2.4.1.4, with an expected size shift of +59bp correlating with the size of the incorporated barcodes (Figure 2.01).

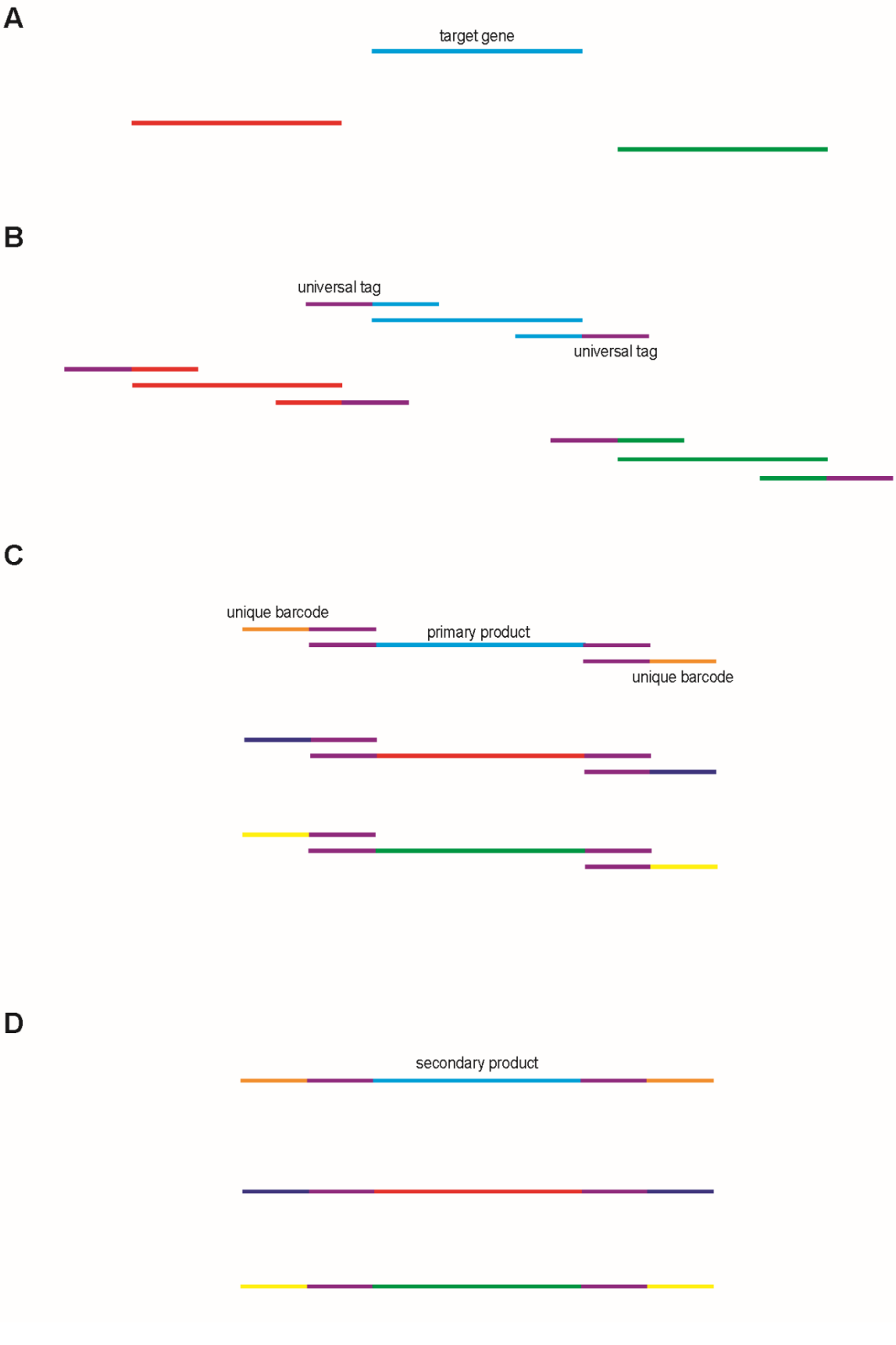

A
target gene
B
universal tag
universal tag
C
unique barcode
primary product
unique barcode
D
secondary product

Figure 2.01 – schematic of the two-step PCR process required for targeted amplicon sequencing

A) Three target genes are represented as blue, red and green lines
B) PCR1. Gene-specific primers with a universal tag are used to generate PCR products containing the universal tag sequence
C) PCR2. Products from PCR1 can now be mixed so long as each pool contains samples with unique mutations. Primers that recognize the universal sequence but contain a unique barcode are used in this reaction
D) Final products are generated. Each product has a 3' and 5' unique barcode, universal sequence and the gene product. The unique barcodes allow sequencing on MiSeq (Illumina, San Diego, CA, USA).

2.5.1.5.3 Analysis of targeted amplicon sequencing

The products from 2.5.1.5.2 were combined to create a pooled amplicon library that was diluted 1:100 with H_2O. To confirm the average fragment size of the library, the pooled library was prepared using the Agilent DNA 1000 Kit (Agilent Technologies, Santa Clara, CA, USA) as per manufacturer's instructions and was analysed using the Agilent 2100 Bioanalyser (Agilent Technologies, Santa Clara, CA, USA). An equimolar pool of libraries was generated and diluted to 2nM, which was confirmed using the Agilent 2100 Bioanalyser and also Qubit® Fluorometric Quantitation (Life Technologies, Thermo Fisher, Scientific, Waltham, MA, USA), as per manufacturer's instructions. The library was loaded onto the MiSeq (Illumina, San Diego, CA, USA) at 6ρM and produced a yield of approximately 4.5-5Gb. Custom ports on the MiSeq cartridge were used to incorporate the custom sequence tagged (CS) fluidigm primers CS1 (5'-ACACTGACGACATGGTTCTACA-3') (port 18, read 1), CS2RC (5'-AGACCAAGTCTCTGCTACCGTA-3') (port 19, index read) and CS2 (5'-TACGGTAGCAGAGACTTGGTCT-3') (port 20, read 2), which were used at a final concentration of 0.5μM. Variants were identified using the bioinformatics pipeline previously described (2.5.1.5.1).

2.6 *In vitro* experiments

2.6.1 Immunofluorescent staining and flow cytometry

All flow cytometry analysis was performed using the BD LSR II, BD FACSCanto™ II, BD FACSCanto™ II Loader and the BD LSRFortessa™ X-20 (BD Biosciences, East

Rutherford, NJ, USA). Where possible, at least 10,000 events were recorded for each sample. Acquisitions of events were recorded using BD FACSDiva™ software (BD Biosciences, East Rutherford, NJ, USA) and the analysis was performed using FlowJo v9.4.1 (for Macintosh) (Tree Star Inc., Ashland, OR, USA). Fluorescence activated cell sorting (FACS) was performed using the BD FACSAria™ II and the BD FACSAria™ Fusion (BD Biosciences, East Rutherford, NJ, USA).

2.6.1.1 GFP, dsRED and BFP expression

To assess maintenance of green fluorescent protein (GFP), red fluorescent protein (dsRED) and blue fluorescent protein (BFP) as surrogate markers for transduction efficiency in E*μ-Myc* lymphomas were harvested as previously described (sections 2.2.1.1 and 2.2.1.2). Cells were washed ×2 in PBS by centrifugation at 423g for 4 minutes. Cells were then resuspended in 100-500μl FACS buffer (Table 2-1), transferred to FACS tubes and analysed on the relevant channel (FITC, PE, PacificBlue) using flow cytometry.

2.6.1.2 Cell surface staining

For cell surface antigen staining, cells were washed in PBS by centrifugation at 423g for 4 minutes and the supernatant was discarded. Cells were resuspended in 200μl FACS buffer (Table 2-1), transferred to 96-well plate (CELLSTAR®, Greiner Bio-One, Kremsmünster, Austria) and centrifuged at 423g for 4 minutes upon which the supernatant was discarded. The cells were then incubated in 50-100μl of FACS buffer containing appropriately diluted primary antibodies (Table 2-6) at 4°C for 30 minutes.

The cells were centrifuged at 423g for 4 minutes and the supernatant was removed, upon which the cells were washed ×2 with 100µl of FACS buffer (Table 2-1) by centrifugation at 423g for 4 minutes. Finally, the cells were resuspended in 100µl of FACS buffer (Table 2-1), transferred to FACS tubes and analysed. In the cases where a biotin-conjugated primary antibody is required, following the final wash cells were incubated with an appropriate streptavidin-fluorochrome conjugated antibody for 30 minutes at 4°C. Cells were then resuspended and prepared for flow cytometric analysis as in 2.6.1.

2.6.1.3 Quantification of cell numbers by flow cytometry

Quantification of cell numbers was performed by flow cytometry as previously described (2.6.1) (Hawkins et al., 2007). Unlabelled BD Calibrite™ 2 beads (BD Biosciences, East Rutherford, NJ, USA) were prepared at a concentration of 10^6 beads/ml in FACS buffer (Table 2-1). 20µl (20,000 beads) of the calibration bead solution was added to the samples following incubation with the primary antibodies, ensuring that all samples received the same amount of beads. Samples were transferred to a FACS tube and analysed by flow cytometry; however, in the first FSC/SSC gate two populations were highlighted – viable cells *and* the calibration bead population. Samples were acquired normally, with the post-acquisition analysis including a bead:cell ratio, which was used to determine total cell counts in a sample and can be extrapolated out to large samples.

2.6.2 Fluorescence *in situ* hybridisation (FISH)

Bone marrow lymphocytes were isolated from 4-week-old pre-malignant E*μ-Myc* animals and their wild type littermate controls according to protocols highlighted in 2.2.2.3. Lymphocytes were transported at RT in E*μ-Myc* media (Table 2-2) to Dr. Meaghan Wall (The Victorian Cancer Cytogenetics Service, St. Vincent's pathology, Melbourne, AUS) who performed fluorescence *in situ* hybridization (FISH). Briefly, cytogenetic suspensions of the cultured lymphocytes were generated by treatment with hypotonic KCl (Table 2-1) and fixed in a 3:1 solution of methanol-to-acetic acid. Mouse BACs targetting the *Myc* locus (307D14 labelled with spectrum green) and *Jak2* locus (324L2 labelled with spectrum orange) were sourced from The Centre For Applied Genomics (Toronto, Ontario, CAN) and were validated for chromosomal location prior to use. Suspensions were dropped onto a microscope slide and incubated for 15 minutes at 100°C. The probes were co-denatured for 2 minutes at 78°C and then hybridised overnight at 37°C prior to being added to the microscope slides. Slides were washed according to the Vysis rapid wash protocol, air dried and mounted with DAPI/antifade (Abbott Laboratories, Chicago, IL, USA). Metaphases were captured and analysed using MetaSystems Isis software interfaced with a Zeiss Axioplan Z fluorescence microscope.

Table 2-6 List of antibodies used

Antibody	Species	Manufacturer	Catalogue #	Dilution/diluent
Flow cytometry antibodies				
CD117 (c-Kit)	Rat IgG2b, k	BD Pharmingen	553869	1:200
CD19	Rat IgG2a, κ	EBioscience	51-0193-80	1:600
CD19	Rat IgG2a, κ	EBioscience	35-0192-81	1:600
CD45.1	Mouse IgG2a, κ	EBioscience	35-0453-81	1:200
CD45.1	Mouse IgG2a, κ	EBioscience	12-0453-81	1:200
CD45.2	Mouse IgG2a, k	EBioscience	47-0454-82	1:200
CD45.2	Mouse IgG2a, κ	BD Pharmingen	553772	1:200
CD45.2	Mouse IgG2a, κ	EBioscience	17-0454-81	1:200
CD45R (B220)	Rat IgG2a, κ	EBioscience	17-0452-81	1:200
CD45R (B220)	Rat IgG2a, k	BD Pharmingen	552094	1:200
CD45R (B220)	Rat IgG2a, k	EBioscience	48-0452-82	1:200
CD45R (B220)	Rat IgG2a, k	EBioscience	35-0452-81	1:200
CD45R (B220)	Rat IgG2a, k	BD Pharmingen	553088	1:200
CD45R (B220)	Rat IgG2a, κ	EBioscience	12-0452-82/12-0452-85	1:200
IgD	Rat IgG2a, k	BD Pharmingen	560869	1:500
IgM	Rat IgG2a, κ	EBioscience	35-5790-81	1:100
Streptavidin		EBioscience	17-4317-82	1:500
Streptavidin		DiagnDiagnostics	10-4317-82	1:500
Streptavidin		BD Pharmingen	554062	1:500
Streptavidin		EBioscience	15-4317-82	1:500
Streptavidin		EBioscience	25-4317-82	1:500
Streptavidin		BD Pharmingen	554064	1:500
Streptavidin		BD Pharmingen	551419	1:500
Streptavidin		BD Pharmingen	554060	1:500
Streptavidin		EBioscience	12-4317-87	1:500
Protein immunoblotting (western blotting) antibodies				
anti-mouse-HRP	Rabbit polyclonal	Dako	P 0260	1:2000/5% skim milk

anti-rabbit HRP	Swine polyclonal	Dako	P 0399	1:2000/5% skim milk
Bcor	Rabbit polyclonal	Gift from Dr. Micah Gearhart	n/a	1:3000/5% skim milk
p19 ^{ARF}	Rabbit polyclonal	Santa Cruz Biotech Inc.	sc-1063	1:1000/5% skim milk
Trp53	Mouse monoclonal	Novocastra	NCL-p53-D07	1:500/5% skim milk
B-Actin	Mouse monoclonal	Sigma Aldrich	A2228	1:10000/5% skim milk or 5% BSA

2.7 Experimental animals

All experiments involving animals were conducted in compliance with the guidelines of the PMCC Experimental Animal Ethics Committee under the animal ethics approval numbers of E472 and conducted in compliance with the Australian code of practice for the care and use of animals for scientific purposes. All experimental animals were held in pathogen free facilities at the PMCC, East Melbourne.

2.7.1 Animal strains and sources

E*μ-Myc* C57BL/6 transgenic mice were originally generated at the Walter and Eliza Hall Institute (WEHI, Melbourne, AUS) as previously described (Adams et al., 1985). E*μ-Myc* mice were bred at the PMCC transgenic facility with C57BL/6 female mice that were purchased from WEHI. B6.SJL-Ptprc^{a} Pep3^{b}/BoyJ (referred to as *Ptprc^{a}* mice hereafter) were bred at the PMCC transgenic facility. The severely immunocompromised mice NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ (referred to as NSG hereafter) were either purchased from WEHI or bred at the PMCC transgenic facility and were kept in isolator boxes to reduce the chance of infections. C57BL/6;*Cdkn2a*$^{-/-}$ mice were a gift from Assoc. Prof. Ross Dickens, WEHI. C57BL/6;*Cdkn2a*$^{-/-}$ female mice were crossed with E*μ-Myc* male transgenic mice to generate E*μ-Myc;Cdkn2a*$^{+/-}$ transgenic mice. *Loxp Bcor Loxp* (*Bcor*$^{fl/x,\ fl/y}$) transgenic mice on a C57BL/6 background were a gift from Dr. Micah Gearhart and Prof. Vivian Bardwell (University of Minnesota, MI, USA). *MB1-Cre* C57BL/6 mice were a gift from Prof. Michael Reth (The

Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany). *Bcor*$^{fl/x\ fl/y}$ were crossed onto *MB1-Cre* C57BL/6 transgenic animals to generate *Bcor*$^{fl/y}$ *MB1-Cre* transgenic animals.

2.7.2 Temporal analysis of mutations

Transgenic E*μ-Myc* mice and WT C57BL/6 littermate controls were genotyped by PCR for the E*μ-Myc* transgene at 2 weeks of age. These mice were subjected to fortnightly bleeds and thrice weekly assessment of tumour burden from 4 weeks of age until obvious signs of progressive lymphoma.

2.7.2.1 Fortnightly tissue collection

Fortnightly, from 4-weeks of age onwards, each transgenic E*μ-Myc* mouse and littermate controls had 50μl of blood collected via the retro-orbital sinus using a heparin coated capillary tube (Thermo Fisher Scientific, Waltham, MA, USA). Blood was collected into a 1.5ml tube (Eppendorf, Hamburg, Germany) that had been pre-loaded with EDTA. The blood sample was agitated and 15μl of whole blood was transferred into a separate 1.5ml tube containing 285μl PBS. ACK lysis buffer (Table 2-1) was added to the whole blood, which was incubated at 4°C for 30 minutes to selectively lyse the erythrocyte fraction.

2.7.2.2 Preparation of whole blood and isolation of B cell fraction

Following erythrocyte lysis, the samples were centrifuged at 423g for 4 minutes and the supernatant removed. The cell pellet was again resuspended in ACK lysis buffer (Table 2-1) and incubated for 30 minutes at 4°C to lyse any residual erythrocytes from the suspension. The sample was centrifuged at 423g for 4 minutes and the cell pellet was then washed ×2 by resuspension in PBS followed by centrifugation at 423g for 4 minutes to remove any residual ACK lysis buffer (Table 2-1). The cell pellet was resuspended in 45µl MACS buffer (Table 2-1) and 15µl CD45R (B220) MicroBeads (Miltenyi Biotech, Bergisch Gladbach, Germany) and was incubated for 30 minutes at 4°C. Samples were centrifuged at 423g for 4 minutes, the supernatant was removed, the cell pellet was resuspended in 500µl MACS buffer (Table 2-1) and the sample was placed on ice until further use. Meanwhile, MS Collumns (Miltenyi Biotech, Bergisch Gladbach, Germany) were loaded onto a magnetic stand and were pre-wetted with 500µl of MACS buffer (Table 2-1), which was allowed to drip entirely through the column. The 500µl samples were added to the pre-wetted MS columns and allowed to drip through entirely before 500µl of MACS buffer (Table 2-1) was twice added to wash any residual $B220^-$ cells from the MS column. Finally 500µl of MACS buffer (Table 2-1) was added to the MS column, upon which the column was removed from the magnet and the solution containing $B220^+$ cells was mechanically forced using the provided plunger into a 1.5ml tube (Eppendorf, Hamburg, Germany). The samples were centrifuged at 423g for 4 minutes and resuspended in 250µl of PBS. 50µl of the sample was used to validate the efficiency B220 positive selection and to immune-phenotype

the cells by flow cytometry as described (2.6.1). A DNA extraction (as described in 2.4.1.3) was performed on the remaining 200µl of the sample.

2.7.2.3 Full blood count and analysis

The 15µl of whole blood that was removed early from the sample (section 2.7.2.1) was diluted at 1/20 in PBS and analysed on an Advia 120® automated haematology analyser (Seimens Healthcare Diagnostics, Deerfield, IL, USA). Blood counts for each sample were then corrected for the dilution and represented as 10^9cells/litre.

2.7.3 Transplantation of Eµ-*Myc* lymphomas

All E*µ-Myc* cells used *in vivo* were sourced from freshly frozen lymph node or splenic suspensions that had not undergone *in vitro* passaging. E*µ-Myc* lymphoma cell suspensions were thawed in E*µ-Myc* media at 37°C and washed ×2 in PBS by centrifugation at 423g for 4 minutes. The cells were resuspended in 10ml PBS, 5µl of the suspension was diluted at 1/25 in trypan blue stain. The viable cells were counted using a bench-top microscope. Cells were then re-suspended in PBS at a density of $5x10^6$ viable cells/ml. Recipient mice were injected with 200µl (10^6 cells) via *intra venous* (iv) injection through the lateral tail vein. Transplanted mice were monitored for the development of lymphoma by lymph node palpation and periodic retro-orbital bleeds for haematological analysis.

2.7.4 Eμ-*Myc* foetal liver derived tumours

2.7.4.1 Eμ-Myc foetal liver transplantation

Following 48-72 hours of propagation in virus free media, transduced foetal liver cells (section 2.3.2.4) were prepared for injection by 2×washing the cells in PBS by centrifugation at 423g for 4 minutes prior to resuspension in PBS at a concentration of $5x10^6$ cells/ml. Primary recipient animals were conditioned with two fractions of 5Gy (Gammacell® 40 caesium source, Atomic Energy of Canada Ltd., Ontario, Canada), four hours apart. Conditioned recipients received 200μl (10^6 cells) via iv injection through the lateral tail vein. Mice were re-housed and provided with additional supplements (Ensure®, Abbott, Abbott Park, IL, USA) and drinking water containing antibiotics (Neomycin sulphate 2.5g/L and Polymyxin B 1.3g/L; both from Sigma-Aldrich®, St Louis, MO, USA) for one month following receipt of cells. Mice were monitored from 4 weeks post-transplant as previously described in 2.7.3.

2.7.4.2 Survival studies in Eμ-Myc lymphoma

Mice were provided with Ensure® supplements and were monitored daily as described in 2.7.4.1. Additionally, mice were monitored weekly from 6 weeks post-transplant for tumour burden by WBC counts and flow cytometry on samples obtained from the retro-orbital sinus. Mice exhibiting signs of tachypnoea, cachexia, general failure to thrive distress, weight loss, bloating, ruffling of fur and/or hind limb paralysis were humanely

sacrificed by cervical dislocation and necropsied for signs of disease burden. In each case, cause of death and spleen weight was documented.

2.8 Statistical analyses

Where applicable, statistical analyses – 1-way ANOVA, 2-way ANOVA, student's t-test and correlation coefficients – and graphical representations were performed using GraphPad Prism® software for Mac OS X Version 5.0d, GraphPad Software Inc., La Jolla, CA, USA). For the entirety of this study, significance was assumed when $\rho<0.05$.

3 The mutational landscape of E*μ-Myc* lymphoma

3.1 Introduction

Since its development almost 30 years ago, the E*μ-Myc* transgenic mouse, produced to model the t(8;14) translocation seen in human Burkitt's lymphoma (BL), has been utilised to identify genes that cooperate with, or suppress, *Myc*-driven malignant transformation (Adams et al., 1985; Harris et al., 1988). E*μ-Myc* mice develop acute B cell leukaemia/lymphoma with 100% penetrance. The mice undergo a period of pre-malignant B cell expansion along with self-limiting remission and onset of clonal lymphoid neoplasia of pre-B cell to naïve B cell phenotype (Harris et al., 1988; Sidman et al., 1993). Many studies have been published using the E*μ-Myc* model of lymphoma to interrogate the efficacy and mechanisms of action of diverse anticancer agents and to study interactions between *Myc* and candidate cancer driver, or suppressor, genes. The E*μ-Myc* model has been used to comprehensively dissect the $p19^{ARF}$-Mdm2-Trp53 tumour suppressor pathway at the biological, genetic and biochemical level to demonstrate its role in mediating *Myc*-driven apoptosis and lymphomagenesis (Eischen et al., 1999; Schmitt et al., 1999). Investigations of the perturbations in the $p19^{ARF}$-Mdm2-Tp53 pathway in E*μ-Myc* lymphoma have profound clinical applicability as *TP53* mutations are found to co-occur with ectopic *MYC* expression in both DLBCL and BL (Lohr et al., 2012; Love et al., 2012b). Clearly, the intrinsic apoptotic pathway is implicated in Myc-driven lymphomagenesis as forced expression of pro-survival Bcl2-family proteins or deletion of pro-apoptotic *BH3*-only genes accelerate lymphomagenesis in the E*μ-Myc* model (Egle et al., 2004; Frenzel et al., 2010; Hemann et al., 2005). The E*μ-Myc* model has also been studied, interrogated and manipulated to definitively demonstrate that Bcl2 and Myc can functionally cooperate to drive

lymphomagenesis (Strasser et al., 1990). Co-overexpression of *BCL2* and *MYC* by rearrangement, stabilising mutations or other cryptic mechanisms represent common clinical challenges in the treatment of DLBCL and BL as these genetic alterations confer multi-drug chemoresistance and poor overall prognosis (Hu et al., 2013; Johnson et al., 2009; Schmitt et al., 2000). The E*μ-Myc* model was used as a pre-clinical tool in the development of BH3-only mimetic agents that work to inhibit pro-survival *Bcl2* family members, which has been translated now to early phase clinical trials that show promise in the setting of refractory DLBCL (Cang et al., 2015; Mason et al., 2008; Vandenberg and Cory, 2013; clinicaltrials.gov, NCT01328626). Spontaneous activating *Kras* and *Nras* mutations have been identified in *de novo* E*μ-Myc* lymphomas at codon 61, permanently activating the proteins with constitutively activated Ras having an important cooperative role in *Myc*-driven transformation (Alexander et al., 1989). The same activating mutations were later described to occur at high frequency in human malignancies (Cox and Der, 2010). The E*μ-Myc* model has been used to demonstrate that oncogenic Ras signalling may play an important co-operative role in Myc transformation either through stabilisation of Myc or by regulating key repressors of Myc target genes (Alexander et al., 1989).

While the E*μ-Myc* model has proved useful in interrogating pathways and genes that interact with Myc in driving lymphomagenesis, there remains a lack of understanding of all of the genes that are spontaneously mutated in this model. Despite E*μ-Myc* lymphomagenesis being driven by ectopic expression of a single oncogene on a defined genetic background there is a large degree of inter-tumoural heterogeneity. Gene

expression signatures of E*μ-Myc* lymphomas demonstrated that several different forms of B cell lymphomas can be identified, with tumour type associated with tumour latency (Mori et al., 2008). The subclinical course in E*μ-Myc* mice is protracted as the animals display marked B cell proliferation in B cell compartments such as the bone marrow and peripheral blood; however, the increased proliferative rate of these cells is initially offset by their increased apoptotic index (Jacobsen et al., 1994). For frank clonal E*μ-Myc* lymphomagenesis, the pre-malignant B cells must circumvent their increased apoptotic rate whilst maintaining the high proliferation index. Overcoming the pro-apoptotic effect of *Myc* overexpression in E*μ-Myc* lymphoma is crucial for malignant transformation. Firstly, Myc induces expression of $p19^{ARF}$, an inhibitor of Mdm2 E3 ligase, which in turn leads to stabilisation of Trp53. Homozygous deletion of the *Cdkn2a* locus coding for $p19^{ARF}$ or disruption of *Trp53* by mutation or deletion occurs as a somatic event in approximately 50% of Eμ-*Myc* lymphomas (Eischen et al., 1999; Schmitt et al., 1999). Spontaneous *Ras*-family member mutations have also been described in *de novo* E*μ-Myc* lymphomas with oncogenic Ras clearly demonstrated to be capable of cooperating with Myc in driving lymphomagenesis (Alexander et al., 1989). Despite being potent cooperating lesions acting in concert with *Myc* to drive malignancy in E*μ-Myc* lymphoma, mutations in *Bcl2* family members or *BH3*-only genes remained unreported in *de novo* E*μ-Myc* lymphoma.

The aforementioned intrinsic apoptotic and Ras signalling pathways are clearly implicated in Eμ-*Myc* lymphomagenesis; however, it appears that the genetic lesions required to co-operate with *Myc* remain unknown in at least half of all cases. It is

therefore plausible that additional gene mutations independent to those pathways already implicated are required for lymphoma progression. Next generation sequencing technologies were employed to catalog the mutations that occur in E*μ-Myc* lymphomas with an aim to define the secondary, tertiary and quaternary mutations that have remained undefined. In this chapter, whole genome sequencing (WGS), whole exome sequencing (WES) and targeted amplicon sequencing (TAM-seq) methodologies were used to comprehensively characterise the E*μ-Myc* model at a deep molecular level. Studies were employed to define transgene architecture and identify somatic mutations and structural alterations that occur in E*μ-Myc* lymphoma. In addition, through retrospective sequencing of stored DNA, the temporal acquisition of mutations in E*μ-Myc* lymphomas were detected and tracked. Herein, previously undefined mutations in the polycomb repressive complex 1 (PRC-1) -like component and *Bcl6-co-repressor* gene (*Bcor*) in E*μ-Myc* lymphoma were identified. Unexpectedly, contemporaneous multigenic lesions involving *Cdkn2a* deletions and other driver mutations including *Nras, Kras* and *Bcor* were identified providing important information on the proposed tumor suppressor function of $p19^{ARF}$. Results in this chapter provide extensive information on the mutational landscape underlying lymphomagenesis in the E*μ-Myc* model, identify Bcor as a putative tumor suppressor gene and challenges the long-held paradigm that E*μ-Myc* lymphoma is a simple ‘two-hit’ model of lymphoma.

3.2 Results

3.2.1 The architecture of the E*μ-Myc* transgene at its insertion site

3.2.1.1 Location of the Eμ-Myc transgene insertion site

Using WGS, the genetic architecture of the *Eμ-Myc* transgene and its insertion site was examined in detail allowing for comprehensive characterisation of the transgenic model. Deep WGS (30× coverage) was applied to DNA extracted from a spontaneous E*μ-Myc* lymphoma (LN88) and matching germline DNA extracted from the tail of mouse #88 (taill88). The E*μ-Myc* transgene sequence was included with the mouse reference, which enabled mapping of breakpoints within the transgene and elucidation of transgene copy number in both tail88 and LN88, allowing visualisation of the transgene, the insertion site architecture in the germline and the identification of any somatic changes (Figure 3.01). LN88 had a total of 748491546 reads with 715746134 mapped reads (95.63% mapping) and tail88 had 927194551 total reads and 891783026 mapped reads (96.18% mapping). LN88 and tail88 had acceptable whole genome coverage with 93.02% and 95.27% of target bases with ≥10-fold coverage, respectively. The E*μ-Myc* transgene was designed to mimic the BL translocation between the *IGH* and *MYC* loci and as such includes *IgH* enhancer elements (chr12) and *Myc* (chr15) in addition to DNA elements including the pUC12 vector and viral phiX sequences (Adams et al., 1985). The expected DNA elements contained within the E*μ-Myc* transgene cassette, including the breakpoint between the cloned translocation of chr12 (*IgH*) and chr15 (*Myc*) in addition to the flanking clone sites within the pUC12 vector and small phiX fragments were identified (Figure 3.01). By using the known pUC12 and phiX DNA as a

genomic “waypoint”, a breakpoint within the tail88 genome containing chr19 and pUC12 vector elements at the 5’ end of the E*μ-Myc* transgene was identified (Figure 3.01). This indicated that chr19 was the founding insertion site for the E*μ-Myc* transgene cassette in this animal, recapitulating data published in a contemporary study (Fusello et al., 2013).

3.2.1.2 High coverage whole genome deep sequencing identified copy number variations spanning the Eμ-Myc transgene cassette

Copy number analysis was applied to the WGS data set for both LN88 and tail88 to characterise the large structural changes that occurred spanning the E*μ-Myc* transgenic cassette. Copy number analysis revealed the presence of three copies of phiX in tail88 and four copies in LN88 and three copies of pUC12 vector elements in tail88 and four copies in LN88 (Figure 3.01). pUC12 and phiX are not present endogenously in mice and therefore must have originated from the transgene. Five copies of *Myc* and *IgH* enhancer elements were present in the hemizygous E*μ-Myc* germline, corresponding to three transgenic copies and two endogenous copies of each gene (Figure 3.01). Six copies of *Myc* and *IgH* enhancer elements were present in the somatic sample, LN88, corresponding to two endogenous copies and four transgenic copies (Figure 3.01). Importantly, all changes in copy number were supported by a corresponding breakpoint (Figure 3.01).

3.2.1.3 High coverage WGS characterised CNV at the Eμ-Myc transgene insertion site

The WGS data set was further mined to characterise copy number variations at the E*μ-Myc* transgene insertion site to elucidate whether the copy number gains in the transgene cassette were apparent in the endogenous region surrounding the insertion site. Three copies of a 3Mb amplicon on chr19 proximal to the site of transgene insertion were identified in tail88 (Figure 3.02A), recapitulating studies performed concurrent to this investigation (Fusello et al., 2013). This chr19 amplicon is endogenous and is not part of the E*μ-Myc* transgenic cassette. This same amplicon was present at four copies in LN88 (Figure 3.02B) indicating that this amplicon underwent an additional somatic amplification in this animal as when copy number in the LN88 amplicon is subtracted from copy number changes in the tail88 amplicon there remains a +1 copy number gain (Figure 3.02C). To assess whether this event is a recurrent event during lymphoma transformation in the E*μ-Myc* model, medium coverage WGS on three other independently derived E*μ-Myc* lymphomas LN218, LN219 and LN299 was performed. 97.13% of reads were mapped (total reads, 81772788) in sample LN218 with 87.53% of target bases with ≥1-fold coverage. 96.9% of reads were mapped in LN219 (82440626 total reads) with 87.95% of bases having ≥1-fold coverage. LN299 had 96.7% of reads mapped (90451542 total reads) with 89.11% if target bases having ≥1-fold coverage. LN218 and LN219 demonstrated an extra copy of the chr19 amplicon above baseline (four total copies) while the chr19 amplicon for sample LN299 remained copy number neutral compared to the E*μ-Myc* germline control (Figure 3.02D).

3.2.1.4 Biological validation of the Eμ-Myc transgene insertion site

Biological validation of the E*μ-Myc* transgenic insertion into chr19 was undertaken using fluorescent *in situ* hybridization (FISH). Fluorescently labelled BAC probes that hybridised to a region on chr15 containing *Myc* and chr19 were utilised so that co-staining of transgenic *Myc* and chr19 would be apparent if the hypothesis driven by the WGS data was correct. Indeed, FISH performed on cells from a hemizygous E*μ-Myc* mouse demonstrated co-localisation of the two BAC probes (Figure 3.03A), whereas FISH performed on cells from a WT C57BL/6 mouse show no co-staining of the chr15 and chr19 probes (Figure 3.03B), confirming juxtaposition of the normally distal chr15 and chr19 elements on 'derivative 19' (Der(19)). The relative intensity of each probe on Der(19) indicated that chr15 elements (*Myc*) have been inserted into chr19 validating the observations made by mapping the E*μ-Myc* transgene insertion using WGS methods and a previous study that ran concurrently to this investigation (Fusello et al., 2013).

3.2.1.5 Establishing the Eμ-Myc chr19 amplicon as a founding event

Given the E*μ-Myc* colony used in this study was one of many E*μ-Myc* colonies housed globally, whether the chr19 amplicon proximal to the transgene insertion site was a conserved copy number variation or was specific to the Peter MacCallum Cancer Centre colony needed to be established. E*μ-Myc* germline tail DNA from the Walter and Eliza Hall Institute (WEHI, Melbourne, Australia) and Cold Spring Harbor Laboratories (CSHL, New York, USA) was obtained for low coverage WGS with the aim of assessing copy number was performed (Figure 3.04). The WEHI-E*μ-Myc* genome had 86.61% of

reads mapped with 45.67% of bases having ≥1-fold coverage, while CSHL-E*μ-Myc* had 93.45% of reads mapped with 59.41% of bases having ≥1-fold coverage. WGS of germline DNA from the CSHL-E*μ-Myc* and the WEHI-E*μ-Myc* animals showed amplification of the chr19 amplicon (Figure 3.04C and Figure 3.04D, respectively). This indicated that these structural rearrangements were founding events that likely occurred following the pronuclear injection of the original transgenic vector.

3.2.1.6 Architecture of the Eμ-Myc transgene

Using WGS, the arrangement of the E*μ-Myc* transgene and the chr19 amplicon proximal to the insertion site was resolved (Figure 3.05A). There was evidence for additional breaks between transgenic copies in the germline indicating that the transgenic copies were arranged as concatenated repeats and occur in sequence (Figure 3.05B). In some cases, the chr19 amplicon and the transgene were somatically duplicated in sequence and were mapped following the concatenated repeat region and is separated from the germline transgene cassette by a region of unmappable, repetitive DNA (Figure 3.05C).

3.2.1.7 Cd274 and Jak-2 are among genes identified in the Eμ-Myc chr19 amplicon

The chr19 amplicon found in E*μ-Myc* animals is syntenic to the human 9p24.1 region, which is frequently amplified in B cell malignancies such as nodular sclerosing Hodgkin lymphoma and primary mediastinal B cell lymphoma (Green et al., 2010). Contained within this region of gain on chr19 are 44 known genes (Table 3-1) including *Jak-2*

(*JAK-2*) and *Cd274* (*CD274*) (encoding Pd-l1), which have been previously implicated in human and murine haematological malignancies (Chen et al., 2013; Green et al., 2010; Mullighan, 2012, 2013; Waibel et al., 2013). Two lymphomas from the prospective E*μ-Myc* cohort, ML52 and ML353, were analysed by medium-coverage WGS and found to have three copies of the chr19 amplicon, while ML73 and ML33 were found to have four copies of the chr19 amplicon. To investigate whether *Cd274* gene dosage, or copy number, correlates with Pd-l1 expression each lymphoma was analysed by flow cytometry for surface Pd-l1 (Figure 3.06). Lymphomas ML52 and ML353 (three copies of *Cd274*) had a geometric mean fluorescence intensity for surface Pd-l1 of 10.7 and 5.3, respectively, while ML73 and ML33 (four copies of *Cd274*) had a geometric mean fluorescence intensity of 1.3 and 5.3, respectively (Figure 3.06). Despite the low number of biological repeats, these data indicated that the number of copies of *Cd274* does not correlate with Pd-l1 expression levels, differing from previously published conclusions (Fusello et al., 2013).

Ectopic JAK-2 activity has been implicated in several human malignancies and has been shown to be a potent driver of malignant transformation (Mullighan et al., 2009; Van Roosbroeck et al., 2011). In acute B cell lymphomas there is data demonstrating that targeted JAK-2 inhibition in a subtype of tumours is a promising therapeutic pathway (Roberts et al., 2012). E*μ-Myc* lymphoma cell lines 4242 and 6066 were treated with increasing concentrations of FDA-approved JAK-2-inhibitor, ruxolitinib, *in vitro* with no increase in apoptosis between the non-treated cells to the cells treated with 2500nm ruxolitinib for 24h (Figure 3.07A). The SET2 cell line, which was used as a

positive control of Jak-2 inhibition, is a human megakaryoblastic leukemic cell line that harbours an activating *JAK-2* mutation that sensitises the cells to JAK-2 inhibition by ruxolitinib treatment (Quentmeier et al., 2006; Uozumi et al., 2000). The SET2 cell line demonstrated dose-dependent cell death in response to ruxolitinib (Figure 3.07A). Furthermore, when the Jak-2 signalling pathway was analysed in 4242 cells compared to a ruxolitinib-sensitive cell line (SET2) there was no Jak-2 signalling activity in either the treated or untreated 4242 cells (Figure 3.07B). Ruxolitinib had no effect on the cells E*μ-Myc*, which when taken together with the total lack of activity in the Jak-2 signalling pathway in E*μ-Myc* lymphoma cells, indicated that the amplification in *Jak-2* in the chr19 amplicon is non-functional, at least in E*μ-Myc* lymphoma maintenance.

Table 3-1 Genes present in the chr19 amplicon

Ensembl Gene ID	Associated Gene Name	Description
ENSMUSG00000093236	Gm24108	predicted gene, 24108 [Source:MGI Symbol;Acc:MGI:5453885]
ENSMUSG00000065556	Mir101b	microRNA 101b [Source:MGI Symbol;Acc:MGI:3618696]
ENSMUSG00000097735	D930032P07Rik	RIKEN cDNA D930032P07 gene [Source:MGI Symbol;Acc:MGI:2443529]
ENSMUSG00000097855	A930007I19Rik	RIKEN cDNA A930007I19 gene [Source:MGI Symbol;Acc:MGI:1925029]
ENSMUSG00000024785	Rcl1	RNA terminal phosphate cyc2.lase-like 1 [Source:MGI Symbol;Acc:MGI:1913275]
ENSMUSG00000024815	Trpd52l3	tumor protein D52-like 3 [Source:MGI Symbol;Acc:MGI:1913995]
ENSMUSG00000024863	Mbl2	mannose-binding lectin (protein C) 2 [Source:MGI Symbol;Acc:MGI:96924]
ENSMUSG00000053536	Cstf2t	cleavage stimulation factor, 3' pre-RNA subunit 2, tau [Source:MGI Symbol;Acc:MGI:1932622]
ENSMUSG00000068466	Gm5518	predicted gene 5518 [Source:MGI Symbol;Acc:MGI:3648673]
ENSMUSG00000024827	Gldc	glycine decarboxylase [Source:MGI Symbol;Acc:MGI:1341155]
ENSMUSG00000024780	Cdc37l1	cell division cycle 37-like 1 [Source:MGI Symbol;Acc:MGI:1914322]
ENSMUSG00000024789	JAK-2	Janus kinase 2 [Source:MGI Symbol;Acc:MGI:96629]
ENSMUSG00000058816	Ppp1r2-ps3	protein phosphatase 1, regulatory (inhibitor) subunit 2, pseudogene 3 [Source:MGI Symbol;Acc:MGI:3645534]
ENSMUSG00000024806	Mlana	melan-A [Source:MGI Symbol;Acc:MGI:108454]
ENSMUSG00000084432	Gm23426	predicted gene, 23426 [Source:MGI Symbol;Acc:MGI:5453203]
ENSMUSG00000087937	Gm26046	predicted gene, 26046 [Source:MGI Symbol;Acc:MGI:5455823]
ENSMUSG00000089016	Gm24883	predicted gene, 24883 [Source:MGI Symbol;Acc:MGI:5454660]
ENSMUSG00000063754	Gm10136	predicted pseudogene 10136 [Source:MGI Symbol;Acc:MGI:3704242]
ENSMUSG00000100075	1700018L02Rik	RIKEN cDNA 1700018L02 gene [Source:MGI Symbol;Acc:MGI:1914579]
ENSMUSG00000024782	Ak3	adenylate kinase 3 [Source:MGI Symbol;Acc:MGI:1860835]
ENSMUSG00000050957	Insl6	insulin-like 6 [Source:MGI Symbol;Acc:MGI:1351595]
ENSMUSG00000039097	Rln1	relaxin 1 [Source:MGI Symbol;Acc:MGI:97931]
ENSMUSG00000016495	Plgrkt	plasminogen receptor, C-terminal lysine transmembrane protein [Source:MGI Symbol;Acc:MGI:1915009]
ENSMUSG00000016496	Cd274	CD274 antigen [Source:MGI Symbol;Acc:MGI:1926446]
ENSMUSG00000016498	Pdcd1lg2	programmed cell death 1 ligand 2 [Source:MGI Symbol;Acc:MGI:1930125]
ENSMUSG00000038658	Ric1	RAB6A GEF complex partner 1 [Source:MGI Symbol;Acc:MGI:1924893]
ENSMUSG00000052942	Glis3	GLIS family zinc finger 3 [Source:MGI Symbol;Acc:MGI:2444289]
ENSMUSG00000089952	4933413C19Rik	RIKEN cDNA 4933413C19 gene [Source:MGI Symbol;Acc:MGI:1921702]
ENSMUSG00000046324	Ermp1	endoplasmic reticulum metallopeptidase 1 [Source:MGI Symbol;Acc:MGI:106250]
ENSMUSG00000024935	Slc1a1	solute carrier family 1 (neuronal/epithelial high affinity glutamate transporter, system Xag), member 1 [Source:MGI Symbol;Acc:MGI:105083]
ENSMUSG00000089993	Gm5822	predicted gene 5822 [Source:MGI Symbol;Acc:MGI:3648836]
ENSMUSG00000040105	Ppapdc2	phosphatidic acid phosphatase type 2 domain containing 2 [Source:MGI Symbol;Acc:MGI:1921661]
ENSMUSG00000064202	4430402I18Rik	RIKEN cDNA 4430402I18 gene [Source:MGI Symbol;Acc:MGI:1918036]
ENSMUSG00000046138	9930021J03Rik	RIKEN cDNA 9930021J03 gene [Source:MGI Symbol;Acc:MGI:2444398]
ENSMUSG00000074909	Ranbp6	RAN binding protein 6 [Source:MGI Symbol;Acc:MGI:2683212]
ENSMUSG00000067351	Rps15a-ps2	ribosomal protein S15A, pseudogene 2 [Source:MGI Symbol;Acc:MGI:3643316]
ENSMUSG00000090737	Gm6788	predicted gene 6788 [Source:MGI Symbol;Acc:MGI:3646480]
ENSMUSG00000098069	Gm26942	predicted gene, 26942 [Source:MGI Symbol;Acc:MGI:5504057]

ENSMUSG00000024810	Il33	interleukin 33 [Source:MGI Symbol;Acc:MGI:1924375]
ENSMUSG00000024817	Uhrf2	ubiquitin-like, containing PHD and RING finger domains 2 [Source:MGI Symbol;Acc:MGI:1923718]
ENSMUSG00000058607	Rpl31-ps20	ribosomal protein L31, pseudogene 20 [Source:MGI Symbol;Acc:MGI:3644512]
ENSMUSG00000024868	Dkk1	dickkopf homolog 1 (Xenopus laevis) [Source:MGI Symbol;Acc:MGI:1329040]
ENSMUSG00000052920	Prkg1	protein kinase, cGMP-dependent, type I [Source:MGI Symbol;Acc:MGI:108174]

3.2.2 Somatic copy number alterations in E*μ-Myc* lymphoma

To better understand the somatically acquired large structural alterations across the E*μ-Myc* cohort, WES from a number of lymphomas (n=9) and low-to-medium coverage WGS (n=13) was performed. Samples that underwent low-coverage WGS had approximately 3× coverage across the genome and samples that underwent medium-coverage WGS had approximately 10× coverage spanning the genome. To analyse somatically acquired copy number gains and losses in the samples that had not been whole-genome sequenced, the aberrant detection in tumour exome (ADTEx) tool was utilised (Amarasinghe et al., 2014). ADTEx resolved polyploidy/aneuploidy from WES data by using three prospective E*μ-Myc* tumours (ML33, ML353 and ML52) for which there was WGS data available as control reference genomes. For example, medium coverage WGS revealed that ML33 had a large gain on chr3 (Figure 3.08A) while ML353 and ML52 remained copy number neutral at the same region (data not shown). ML362 is of unknown copy number at chr3; however, when the total reads revealed from WES was compared to lymphomas ML33, ML353 and ML52 of known copy number it was shown to align with ML352 and ML52 but not with ML33, indicating that ML362 was copy number neutral at chr3 (Figure 3.08B). The average copy number for a segment of DNA was solved from a window size of 10,000bp (adjacent windows with the same copy number were considered a single CNV). Segments of DNA for which there could be no consensus on copy number were omitted from this analysis.

While each genome varied in the total number of copy number losses and gains, it was apparent there were common regions of recurrent copy number gain or copy number

loss (Figure 3.08C). Copy number information for each E*μ-Myc* tumour was combined to form a cumulative plot showing only the regions of recurrent somatic copy number alterations as a percentage of total cases (Figure 3.08C). Most of these somatically acquired copy number alterations were characterised by large segmental amplifications or whole chromosome duplications, with focal low level gains being less common (Figure 3.08C). This under representation of smaller, focal somatically acquired copy number alterations is likely due to the large window size that was employed in the analysis, which favoured detection of large copy number variations over detection of smaller amplicons. The genes within the recurrent regions of gain or loss in the E*μ-Myc* cohort were cross-referenced against the genes listed in the Cancer Gene Census (CGC), and were filtered so that only genes that are implicated in human malignancies were selected (Table 3-2). 5/22 cases demonstrated large segmental gains or whole chromosome duplication of chr3, including regions containing the *Kras* (Cox and Der, 2010) and *Bcl9* oncogenes (Takada et al., 2012) (Figure 3.08C). Chr6 was duplicated or gained in a region containing the leukaemia-associated oncogene, *Nras* (Cox and Der, 2010), in 2/22 lymphomas (Figure 3.08C). Duplications and large segments of copy number gain on chr10, encompassing the *Trp53* inhibitor *Mdm2* (Eischen et al., 1999) were present in lymphomas ML52 and ML43 (2/22) (Figure 3.08C). Lymphoma ML52 harboured a *Cdkn2a* deletion (as per section 3.2.3.6), indicating that any gain in Mdm2 would be non-functional; however, E*μ-Myc* tumours lacking *Cdkn2a* have been demonstrated to harbour concomitant Mdm2 overexpression (Eischen et al., 1999). ML43 did not harbour any other perturbation in the $p19^{ARF}$-Mdm2-Trp53 pathway (as per section 3.2.3.4) apart from the gain in Mdm2. Chr12, containing the oncogenes

Akt1 (Wendel et al., 2004) and *Mycn* (Bordow et al., 1998) and the DNA methyltransferase *Dnmt3a* (Ley et al., 2010) that is recurrently mutated in human AML, was duplicated in 2/22 cases (Figure 3.08C). 4/22 cases exhibited whole chromosome duplication of chr15, containing *Myc*, the gene proposed to be driving E*μ-Myc* lymphoma (Adams et al., 1985) (Figure 3.08C). 3/22 cases demonstrated whole chromosome duplication of chr18 containing the oncogene *Apc* (Morin et al., 1996) (Figure 3.08C). 3/22 cases showed monosomy or deletion of part of chr13 (Figure 3.08C). Focal somatically acquired copy number alterations included low level gains on chr1 including the oncogene *Cxcr7* (Miao et al., 2007) in 6/22 cases and chr7 containing the oncogene *Fgfr2* (Katoh and Katoh, 2003) in 3/22 cases (Figure 3.08C). Focal regions of hemizygous loss were apparent on chr4 in 7/22 cases, chr9 in 4/22 cases and chr11 in 5/22 cases (Figure 3.08C). The somatic amplification of the E*μ-Myc* transgene and the chr19 amplicon proximal to the transgenic insertion site found originally in LN88 and validated in LN218 and LN219 (discussed in section 3.2.1) was present in a total of 12 out of 22 E*μ-Myc* lymphomas from the prospective and retrospective cohorts (Figure 3.08C). There was no difference in survival (in days) between parental E*μ-Myc* mice that harboured lymphomas that had somatically acquired another copy of chr19 amplicon compared to those mice that exhibited lymphomas that retained the germline levels of the chr19 amplicon (data not shown). This indicated that the additional somatic gain of chr19 did not result in a more aggressive lymphoma, as measured by survival.

Copy number analysis across the entire prospective cohort (16 lymphomas) using ADTEx showed variable copy number variation events between samples. Total number of copy number loss in the prospective cohort was 54 while the total number of copy number gain events was 61 (Figure 3.09A). CNV per lymphoma, including both gains and losses, ranged from 4 to 13, highlighting the heterozygous presentation of E*μ-Myc* lymphomas (Figure 3.09A) (Mori et al., 2008). In some prostate and colorectal cancers a general increase in copy number variations across the genome preceded transformation, predicted relapse to therapy and decreased prognosis (Wang et al., 2016; Yu et al., 2012) while in follicular lymphoma, increases in copy number variations precedes transformation to aggressive DLBCL (Ross et al., 2007). Given this information it is intuitive to predict that in E*μ-Myc* lymphoma increased copy number variations would correlate with reduced lifespan of the E*μ-Myc* mice or, conversely, one might make the assumption that over a longer lifespan, a lymphoma may naturally acquire more CNVs. However, total number of copy number variations per lymphoma did not significantly correlate with changes in survival time (linear regression shows $R = -0.182$, slope not significantly non-zero, $p > 0.05$) (Figure 3.09B).

Table 3-2 Recurrent CNVs found the retrospective and prospective E*μ-Myc* cohorts and their intersection with human genes described in the Cancer Gene Census (CGC)

Gene ID	Gene Symbol	Human orthologue	Chr	Exon start	Exon end	% cases altered
ENSMUSG00000044337	Gain-Cxcr7	ACKR3	1	90200000	90250000	0.272727273
ENSMUSG00000015522	Gain-Arnt	ARNT	3	94885459	95879340	0.227272727
ENSMUSG00000033161	Gain-Atp1a1	ATP1A1	3	98130985	102920791	0.227272727
ENSMUSG00000028191	Gain-Bcl10	BCL10	3	145595921	146134183	0.181818182
ENSMUSG00000038256	Gain-Bcl9	BCL9	3	96643386	98111915	0.227272727
ENSMUSG00000044468	Gain-Fam46c	FAM46C	3	98130985	102920791	0.227272727
ENSMUSG00000028086	Gain-Fbxw7	FBXW7	3	71460000	88000000	0.181818182
ENSMUSG00000044167	Gain-Foxo1	FOXO1	3	35200000	53260819	0.136363636
ENSMUSG00000028034	Gain-Fubp1	FUBP1	3	148822762	152226413	0.181818182
ENSMUSG00000027823	Gain-Gmps	GMPS	3	56100000	64250000	0.136363636
ENSMUSG00000040289	Gain-Hey1	HEY1	3	3508089	9950000	0.136363636
ENSMUSG00000027720	Gain-Il2	IL2	3	35200000	53260819	0.136363636
ENSMUSG00000048332	Gain-Lhfp	LHFP	3	35200000	53260819	0.136363636
ENSMUSG00000028063	Gain-Lmna	LMNA	3	88072869	89207608	0.227272727
ENSMUSG00000027684	Gain-Mecom	MECOM	3	28950000	35150000	0.136363636
ENSMUSG00000048416	Gain-Mlf1	MLF1	3	64650000	71460000	0.136363636
ENSMUSG00000053192	Gain-Mllt11	MLLT11	3	94885459	95879340	0.227272727
ENSMUSG00000042784	Gain-Muc1	MUC1	3	89210507	89245955	0.227272727
ENSMUSG00000027878	Gain-Notch2	NOTCH2	3	96643386	98111915	0.227272727
ENSMUSG00000027852	Gain-Nras	NRAS	3	102935356	103062722	0.227272727
ENSMUSG00000028072	Gain-Ntrk1	NTRK1	3	71460000	88000000	0.181818182
ENSMUSG00000038170	Gain-Pde4dip	PDE4DIP	3	96643386	98111915	0.227272727
ENSMUSG00000027665	Gain-Pik3ca	PIK3CA	3	28950000	35150000	0.136363636
ENSMUSG00000004895	Gain-Prcc	PRCC	3	71460000	88000000	0.181818182
ENSMUSG00000028149	Gain-Rap1gds1	RAP1GDS1	3	136934856	139200000	0.181818182
ENSMUSG00000048109	Gain-Rbm15	RBM15	3	106684868	107931476	0.227272727
ENSMUSG00000074637	Gain-Sox2	SOX2	3	28950000	35150000	0.136363636
ENSMUSG00000027630	Gain-Tbl1xr1	TBL1XR1	3	18950000	22900000	0.136363636
ENSMUSG00000040943	Gain-Tet2	TET2	3	131400000	133473849	0.181818182
ENSMUSG00000033014	Gain-Trim33	TRIM33	3	103073865	104464472	0.227272727
ENSMUSG00000027803	Gain-Wwtr1	WWTR1	3	56100000	64250000	0.136363636
ENSMUSG00000028484	Loss-Psip1	PSIP1	4	83453383	83456808	0.318181818
ENSMUSG00000000184	Gain-Ccnd2	CCND2	6	126900000	129550000	0.090909091
ENSMUSG00000003031	Gain-Cdkn1b	CDKN1B	6	131100000	145950000	0.090909091
ENSMUSG00000030057	Gain-Cnbp	CNBP	6	82000000	103550000	0.090909091
ENSMUSG00000031865	Gain-Dctn1	DCTN1	6	82000000	103550000	0.090909091
ENSMUSG00000030172	Gain-Erc1	ERC1	6	115967628	123300000	0.090909091
ENSMUSG00000030275	Gain-Etnk1	ETNK1	6	131100000	145950000	0.090909091
ENSMUSG00000030199	Gain-Etv6	ETV6	6	131100000	145950000	0.090909091
ENSMUSG00000034023	Gain-Fancd2	FANCD2	6	113337907	115844913	0.090909091
ENSMUSG00000030067	Gain-Foxp1	FOXP1	6	82000000	103550000	0.090909091
ENSMUSG00000015053	Gain-Gata2	GATA2	6	82000000	103550000	0.090909091
ENSMUSG00000030180	Gain-Kdm5a	KDM5A	6	115967628	123300000	0.090909091

ENSMUSG00000030265	Gain-Kras	KRAS	6	131100000	145950000	0.090909091
ENSMUSG00000035158	Gain-Mitf	MITF	6	82000000	103550000	0.090909091
ENSMUSG00000000440	Gain-Pparg	PPARG	6	113337907	115844913	0.090909091
ENSMUSG00000016487	Gain-Ppfibp1	PPFIBP1	6	146050000	149550000	0.090909091
ENSMUSG00000000441	Gain-Raf1	RAF1	6	113337907	115844913	0.090909091
ENSMUSG00000030110	Gain-Ret	RET	6	115967628	123300000	0.090909091
ENSMUSG00000030062	Gain-Rpn1	RPN1	6	82000000	103550000	0.090909091
ENSMUSG00000030257	Gain-Srgap3	SRGAP3	6	107850000	113334095	0.090909091
ENSMUSG00000030094	Gain-Xpc	XPC	6	82000000	103550000	0.090909091
ENSMUSG00000038346	Gain-Zfp384	ZNF384	6	123650000	126850000	0.090909091
ENSMUSG00000030849	Gain-Fgfr2	FGFR2	7	130900000	130950000	0.136363636
ENSMUSG00000063903	Gain-Klk1	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000063133	Gain-Klk1b1	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000044485	Gain-Klk1b11	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000038968	Gain-Klk1b16	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000066516	Gain-Klk1b21	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000060177	Gain-Klk1b22	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000063713	Gain-Klk1b24	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000053719	Gain-Klk1b26	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000063177	Gain-Klk1b27	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000066515	Gain-Klk1b3	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000066513	Gain-Klk1b4	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000066512	Gain-Klk1b5	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000063089	Gain-Klk1b8	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000059042	Gain-Klk1b9	KLK2	7	39400000	46500000	0.090909091
ENSMUSG00000009471	Gain-Myod1	MYOD1	7	39400000	46500000	0.090909091
ENSMUSG00000036712	Gain-Cyld	CYLD	8	88700000	88750000	0.272727273
ENSMUSG00000031537	Gain-Ikbkb	IKBKB	8	22600000	22700000	0.227272727
ENSMUSG00000031592	Gain-Pcm1	PCM1	8	41250000	41300000	0.181818182
ENSMUSG00000031799	Gain-Tpm4	TPM4	8	72100000	72150000	0.227272727
ENSMUSG00000032598	Loss-Nckipsd	NCKIPSD	9	108800000	108850000	0.181818182
ENSMUSG00000006728	Gain-Cdk4	CDK4	10	119250000	130000000	0.090909091
ENSMUSG00000025408	Gain-Ddit3	DDIT3	10	119250000	130000000	0.090909091
ENSMUSG00000018166	Gain-Erbb3	ERBB3	10	119250000	130000000	0.090909091
ENSMUSG00000056758	Gain-Hmga2	HMGA2	10	119250000	130000000	0.090909091
ENSMUSG00000020105	Gain-Lrig3	LRIG3	10	119250000	130000000	0.090909091
ENSMUSG00000020184	Gain-Mdm2	MDM2	10	116632086	119200000	0.090909091
ENSMUSG00000019982	Gain-Myb	MYB	10	20370000	21290000	0.090909091
ENSMUSG00000025402	Gain-Nab2	NAB2	10	119250000	130000000	0.090909091
ENSMUSG00000020154	Gain-Ptprb	PTPRB	10	112050000	116581720	0.090909091
ENSMUSG00000002147	Gain-Stat6	STAT6	10	119250000	130000000	0.090909091
ENSMUSG00000020218	Gain-Wif1	WIF1	10	119250000	130000000	0.090909091
ENSMUSG00000069769	Loss-Msi2	MSI2	11	88400000	88450000	0.227272727
ENSMUSG00000001729	Gain-Akt1	AKT1	12	109300000	113429684	0.090909091
ENSMUSG00000051721	Gain-BC068281	C2orf44	12	3176259	18600000	0.090909091
ENSMUSG00000041415	Gain-Dicer1	DICER1	12	89050000	107350000	0.090909091

ENSMUSG00000020661	Gain-Dnmt3a	DNMT3A	12	3176259	18600000	0.090909091
ENSMUSG00000004151	Gain-Etv1	ETV1	12	37100000	54701786	0.090909091
ENSMUSG00000035451	Gain-Foxa1	FOXA1	12	56350000	59750000	0.090909091
ENSMUSG00000021908	Gain-Gm6768	NCOA4	12	117800000	120018754	0.090909091
ENSMUSG00000021192	Gain-Golga5	GOLGA5	12	89050000	107350000	0.090909091
ENSMUSG00000047454	Gain-Gphn	GPHN	12	75950000	82650000	0.090909091
ENSMUSG00000021270	Gain-Hsp90aa1	HSP90AA1	12	109300000	113429684	0.090909091
ENSMUSG00000059436	Gain-Max	MAX	12	75950000	82650000	0.090909091
ENSMUSG00000037169	Gain-Mycn	MYCN	12	3176259	18600000	0.090909091
ENSMUSG00000020647	Gain-Ncoa1	NCOA1	12	3176259	18600000	0.090909091
ENSMUSG00000021068	Gain-Nin	NIN	12	68350000	75900000	0.090909091
ENSMUSG00000001496	Gain-Nkx2-1	NKX2-1	12	56350000	59750000	0.090909091
ENSMUSG00000059060	Gain-Rad51b	RAD51B	12	75950000	82650000	0.090909091
ENSMUSG00000041359	Gain-Tcl1	TCL1A	12	89050000	107350000	0.090909091
ENSMUSG00000021188	Gain-Trip11	TRIP11	12	89050000	107350000	0.090909091
ENSMUSG00000020963	Gain-Tshr	TSHR	12	89050000	107350000	0.090909091
ENSMUSG00000021377	Loss-Dek	DEK	13	46500000	48400000	0.181818182
ENSMUSG00000041417	Loss-Pik3r1	PIK3R1	13	97150000	104900000	0.181818182
ENSMUSG00000000078	Loss-Klf6	KLF6	13	3400000	12800000	0.136363636
ENSMUSG00000021326	Loss-Trim27	TRIM27	13	17050000	25300000	0.136363636
ENSMUSG00000021356	Loss-Irf4	IRF4	13	27300000	31200000	0.136363636
ENSMUSG00000048368	Loss-Omd	OMD	13	48400000	53050000	0.136363636
ENSMUSG00000021457	Loss-Syk	SYK	13	48400000	53050000	0.136363636
ENSMUSG00000005320	Loss-Fgfr4	FGFR4	13	54490000	59900000	0.136363636
ENSMUSG00000021488	Loss-Nsd1	NSD1	13	54490000	59900000	0.136363636
ENSMUSG00000021461	Loss-Fancc	FANCC	13	61250000	68150000	0.136363636
ENSMUSG00000021466	Loss-Ptch1	PTCH1	13	61250000	68150000	0.136363636
ENSMUSG00000021611	Loss-Tert	TERT	13	73500000	74650000	0.136363636
ENSMUSG00000021754	Loss-Map3k1	MAP3K1	13	110850000	114000000	0.136363636
ENSMUSG00000021756	Loss-Il6st	IL6ST	13	110850000	114000000	0.136363636
ENSMUSG00000033237	Gain-Arid2	ARID2	15	95540000	96432022	0.181818182
ENSMUSG00000023027	Gain-Atf1	ATF1	15	98911288	101677798	0.181818182
ENSMUSG00000022483	Gain-Col2a1	COL2A1	15	96576520	98908141	0.181818182
ENSMUSG00000014313	Gain-Cox6c	COX6C	15	34900000	38650000	0.181818182
ENSMUSG00000022336	Gain-Eif3e	EIF3E	15	40800000	44600000	0.181818182
ENSMUSG00000055024	Gain-Ep300	EP300	15	74950000	86450000	0.181818182
ENSMUSG00000061731	Gain-Ext1	EXT1	15	52500000	57250000	0.181818182
ENSMUSG00000001656	Gain-Hoxc11	HOXC11	15	101678782	103354505	0.181818182
ENSMUSG00000001655	Gain-Hoxc13	HOXC13	15	101678782	103354505	0.181818182
ENSMUSG00000003882	Gain-Il7r	IL7R	15	9250000	11700000	0.181818182
ENSMUSG00000054263	Gain-Lifr	LIFR	15	4650000	8319715	0.181818182
ENSMUSG00000042292	Gain-Mkl1	MKL1	15	74950000	86450000	0.181818182
ENSMUSG00000048154	Gain-Mll2	KMT2D	15	96576520	98908141	0.181818182
ENSMUSG00000022346	Gain-Myc	MYC	15	61850000	62107652	0.181818182
ENSMUSG00000022443	Gain-Myh9	MYH9	15	74950000	86450000	0.181818182
ENSMUSG00000005125	Gain-Ndrg1	NDRG1	15	62219210	68350000	0.181818182

ENSMUSG00000000489	Gain-Pdgfb	PDGFB	15	74950000	86450000	0.181818182
ENSMUSG00000022314	Gain-Rad21	RAD21	15	44650000	52450000	0.181818182
ENSMUSG00000033762	Gain-Recql4	RECQL4	15	74950000	86450000	0.181818182
ENSMUSG00000051920	Gain-Rspo2	RSPO2	15	40800000	44600000	0.181818182
ENSMUSG00000023018	Gain-Smarcd1	SMARCD1	15	98911288	101677798	0.181818182
ENSMUSG00000037487	Gain-Ubr5	UBR5	15	34900000	38650000	0.181818182
ENSMUSG00000068036	Loss-Mllt4	MLLT4	17	13750000	13800000	0.272727273
ENSMUSG00000052397	Loss-Ezr	EZR	17	6550000	6900000	0.090909091
ENSMUSG00000005871	Gain-Apc	APC	18	33250000	44300000	0.136363636
ENSMUSG00000036452	Gain-Arhgap26	ARHGAP26	18	33250000	44300000	0.136363636
ENSMUSG00000024610	Gain-Cd74	CD74	18	57200000	68600000	0.136363636
ENSMUSG00000024382	Gain-Ercc3	ERCC3	18	31000000	33200000	0.136363636
ENSMUSG00000006740	Gain-Kif5b	KIF5B	18	3150000	7300000	0.136363636
ENSMUSG00000032688	Gain-Malt1	MALT1	18	57200000	68600000	0.136363636
ENSMUSG00000024620	Gain-Pdgfrb	PDGFRB	18	57200000	68600000	0.136363636
ENSMUSG00000024548	Gain-Setbp1	SETBP1	18	70500000	79600000	0.136363636
ENSMUSG00000024515	Gain-Smad4	SMAD4	18	70500000	79600000	0.136363636
ENSMUSG00000037013	Gain-Ss18	SS18	18	10181030	15600000	0.136363636
ENSMUSG00000024420	Gain-Zfp521	ZNF521	18	10181030	15600000	0.136363636
ENSMUSG00000016496	Gain-Cd274	CD274	19	29358566	29372252	0.545454545
ENSMUSG00000024789	Gain-JAK-2	JAK-2	19	29134210	29289006	0.545454545
ENSMUSG00000016498	Gain-Pdcd1lg2	PDCD1LG2	19	29385610	29422170	0.5
ENSMUSG00000024947	Loss-Men1	MEN1	19	5550000	6390123	0.090909091
ENSMUSG00000024639	Loss-Gnaq	GNAQ	19	15550000	28200000	0.090909091
ENSMUSG00000013663	Loss-Pten	PTEN	19	31302101	34150000	0.090909091
ENSMUSG00000024778	Loss-Fas	FAS	19	34250000	34650000	0.090909091
ENSMUSG00000025215	Loss-Tlx1	TLX1	19	34700000	46068262	0.090909091
ENSMUSG00000025225	Loss-Nfkb2	NFKB2	19	46078489	52200000	0.090909091
ENSMUSG00000025231	Loss-Sufu	SUFU	19	46078489	52200000	0.090909091
ENSMUSG00000025041	Loss-Nt5c2	NT5C2	19	46078489	52200000	0.090909091
ENSMUSG00000024983	Loss-Vti1a	VTI1A	19	52900000	61227527	0.090909091
ENSMUSG00000024985	Loss-Tcf7l2	TCF7L2	19	52900000	61227527	0.090909091
ENSMUSG00000041362	Loss-4930506M07Rik	KIAA1598	19	52900000	61227527	0.090909091

3.2.3 Whole exome sequencing reveals novel genetic drivers of E*μ-Myc* lymphoma

E*μ-Myc* lymphomas demonstrate a remarkable level of heterogeneity in their behavior, appearance and aggressiveness (Mori et al., 2008). Ectopic Myc expression in E*μ-Myc* mice drives both proliferation and apoptosis, with the apoptotic pathway putatively needing to be circumvented before a clinical manifestation of lymphoma (Eischen et al., 1999; Schmitt et al., 1999). In at least 50% of sporadic E*μ-Myc* lymphomas, *Cdkn2a* or *Trp53* were shown to be mutated (Eischen et al., 1999). *De novo* activating mutations in *Ras*-family proteins have also been implicated in E*μ-Myc* at lower frequencies (Alexander et al., 1989). In the remaining 50% of E*μ-Myc* lymphomas the secondary Myc-cooperating mutation remains unknown. It was hypothesised that WES would reveal previously undefined mutations in genes that can cooperate with Myc in driving lymphomagenesis in the E*μ-Myc* lymphoma. WES was performed on a total of 23 E*μ-Myc* lymphomas to investigate the type, frequency and number of somatic mutations that sporadically occur in this model of disease. 16 lymphomas were derived prospectively from E*μ-Myc* transgenic mice with the aim of sequencing them and seven cases were taken retrospectively from lymphomas previously generated in the Johnstone Laboratory and were utilised as validation lymphomas. Tail DNA from two animals, tail88 and ML62 (WT), underwent WES and were used to filter germline polymorphisms from the data, with tail88 being more closely related to the retrospective cohort and ML62 being more closely related to the prospective cohort. dbSNP annotated variants were filtered out of the analysis as were any SNV or InDel that was found to be recurring in two or more closely related animals as these variants were most

likely polymorphisms or recurrent sequencing artifacts. After filtering, the gene variant list in the prospective cohort totalled 82 (Table 3-3). The lymphomas from the prospective cohort each had between 3-26 SNV or InDels predicted to truncate or alter the translated amino acid sequence of proteins encoded by annotated genes (Figure 3.10A). A subset of single nucleotide variants and InDels were further validated using targeted amplicon sequencing (TAM-seq), indicating a high specificity of 95.1% (78/82) of the variant calling pipeline employed. The variation in the retrospective cohort in terms of numbers of variants per sample was much greater, with each lymphoma harbouring between three and 36 SNV or InDels predicted to truncate or alter the translated amino acid sequence encoded by annotated genes (Figure 3.10A). In both cohorts, the number of SNV was much greater than the total number of frame-shift InDels. Aggressive B cell lymphomas demonstrate a greater mutational index than indolent B cell lymphomas (Vaque et al., 2014). In E*μ-Myc* lymphoma the total number of mutations, both single nucleotide variants and frame-shift InDels, did not correlate with aggressiveness of E*μ-Myc* lymphoma in the prospective cohort as assessed by time until sacrifice ($R = 0.0298$, not significantly non-zero, $p > 0.05$) (Figure 3.10B).

3.2.3.1 Trp53 mutations

Prospective E*μ-Myc* lymphomas ML30 (99.4% of reads mapped, with 98.03% of bases having ≥10-fold coverage from a total of 127338506 reads) ML73 (99.51% of reads mapped, with 97.80% of bases having ≥10-fold coverage from a total of 107057758 reads) and ML362 (99.45% of reads mapped, with 96.38% of bases having ≥10-fold coverage from a total of 86860920 reads) and retrospective E*μ-Myc* lymphoma #13

harboured a frame-shift InDel or a missense SNV in the *Trp53* gene (4/23 cases) (Figure 3.11). ML30 contained an F to L amino acid change at Trp53 position 131, ML73 contained an A to V missense mutation at Trp53 position 135, ML362 harboured a 34bp deletion at amino acid position 261 and #13 contained a 1bp deletion resulting in a splice donor variant (reference: ENSMUSG00000059552) (Figure 3.12A). *TRP53* is a known tumour suppressor gene and has been implicated in many human malignancies as well as being identified previously as a frequently mutated gene in E*μ-Myc* lymphoma (Eischen et al., 1999; Olivier et al., 2010; Schmitt et al., 1999).

3.2.3.2 Nras mutations

Prospective E*μ-Myc* lymphoma ML27 (99.31% of reads mapped, with 97.84% of bases having ≥10-fold coverage from a total of 114937656 reads) and retrospective E*μ-Myc* lymphoma #6066 (99.52% of reads mapped, with 86.34% of bases having ≥10-fold coverage from a total of 54874574 reads) both exhibited an *Nras* C to A missense mutation at CDS position 181 corresponding to a Q to K amino acid substitution at codon 61 (reference: ENSMUSG00000027852) (Figure 3.11, Figure 3.12B). The $Nras^{Q61K}$ missense mutation has been previously described in spontaneous E*μ-Myc* lymphoma (Alexander et al., 1989) and is acknowledged as a driving oncogenic mutation in human malignancy (Buhrman et al., 2011b).

3.2.3.3 Kras mutations

In the prospective E*μ-Myc* lymphomas ML353 (99.38% of reads mapped, with 98.11% of bases having ≥10-fold coverage from a total of 113193116 reads), ML58 (99.36% of

reads mapped, with 98.15% of bases having ≥10-fold coverage from a total of 131685208 reads), ML358 (99.47% of reads mapped, with 97.54% of bases having ≥10-fold coverage from a total of 106433066 reads) and ML367 (99.42% of reads mapped, with 98.2% of target bases having ≥10-fold coverage from a total of 127447478 reads) demonstrated point mutations in the *Kras* gene (Figure 3.11). ML367 and ML58 both contained Q61H single nucleotide variants, recapitulating published works demonstrating that these mutations occur *de novo* in E*μ-Myc* lymphoma (Alexander et al., 1989). Lymphoma ML353 contained an E63K substitution and lymphoma ML358 contained a G12D substitution (reference: ENSMUSG00000030265) (Figure 3.12C). *KRAS* mutations at these codons lead to permanent activation of KRAS and result in ectopic MAPK signaling, which is an oncogenic pathway in human malignancy, particularly in colorectal and lung carcinomas (Brose et al., 2002; Lievre et al., 2006).

3.2.3.4 Bcor mutations

The most frequently mutated gene was the *Bcl6-corepressor* (*Bcor*), mutated in seven E*μ-Myc* lymphomas from both cohorts at a frequency of 32% (Figure 3.11). This is the first time mutations in *Bcor* have been demonstrated in E*μ-Myc* lymphoma. ML21 (99.39% of reads mapped, with 97.87% of bases having ≥10-fold coverage from a total of 116678540 reads) harboured a 2bp insertion in the *Bcor* at cDNA position 4458-4459 (Figure 3.12D). ML33 (99.44% of reads mapped, with 98.4% of bases having ≥10-fold coverage from a total of 129895014 reads) contained a 1bp insertion in *Bcor* at cDNA position 4459-4460 (Figure 3.12D). ML43 (99.25% of reads mapped, with 98.29% of

bases having ≥10-fold coverage from a total of 131612702 reads) contained a C to T substitution at CDS position 4651, corresponding to a R1551* amino acid substitution resulting in a premature stop codon in *Bcor* (Figure 3.12D). ML352 (99.53% of reads mapped, with 98.05% of bases having ≥10-fold coverage from a total of 122286806 reads) contained a 7bp deletion at cDNA position 4652-4658 of the *Bcor* gene (Figure 3.12D). Retrospective sample 299 (99.62% of reads mapped, with 94.89% of bases having ≥10-fold coverage from a total of 77042612 reads) contained the same mutation found in prospective lymphoma ML33. Retrospective sample LN88 (99.51% of reads mapped, with 94.89% of bases having ≥10-fold coverage from a total of 83836218 reads) contained a 2bp insertion at CDS position 4606-4607 in the *Bcor* gene (Figure 3.12D). Retrospective lymphoma 4242 displayed no *Bcor* mutation of consequence by WES; however, RNA sequencing demonstrated a 217bp deletion in exon 4 that was too large for WES to detect. All the mutations identified in *Bcor* were frame-shift, InDels or stop gained mutations that are predicted to cause premature protein truncation (reference: ENSMUSG00000040363), providing strong evidence to suggest that *Bcor* is a tumour suppressor gene in the context of E*μ-Myc* lymphoma (Vogelstein et al., 2013).

3.2.3.5 Low frequency mutations identified in Eμ-Myc lymphoma

In addition to a frame-shift InDel in *Bcor,* lymphoma ML33 harboured a splice site variant in *Ezh2*, a PRC-2 component previously shown to function as a tumour suppressor gene in Eμ-Myc lymphoma (reference: ENSMUSP00000080419) (Lee et al., 2013) (Figure 3.11). ML39 (99.40% of reads mapped, with 98.15% of bases having ≥10-fold coverage from a total of 120656576 reads) harboured a mutation in the gene

Mtor at a residue conserved between mice and humans and was predicted to be damaging using the SWIFT algorithm (reference: ENSMUSP00000099510) (Figure 3.11). Deregulation of the Akt-Mtor-Eif4e pathway has been shown to cooperate with Myc in E*μ-Myc* lymphomagenesis (Grabiner et al., 2014; Wendel et al., 2004) and activating *MTOR* mutations have previously been described in several human malignances including B cell lymphomas (Zhang et al., 2013). Lymphoma LN88 contained a mutation in *mir142*, which is reported in approximately 20% of human DLBCL (Figure 3.11) (Kwanhian et al., 2012). ML352 and 299 had mutations in *Rpl10*, a ribosomal protein that is recurrently mutated in human T-ALL (reference: ENSMUSP00000082055) (De Keersmaecker et al., 2013) (Figure 3.11). Lymphomas 218 (99.32% of reads mapped, with 97.43% of bases having ≥10-fold coverage from a total of 116813120 reads), 219 (99.23% of reads mapped, with 97.52% of bases having ≥10-fold coverage from a total of 125443086 reads) and ML20 (99.37% of reads mapped, with 97.58% of bases having ≥10-fold coverage from a total of 106496552 reads) did not harbour SNVs or InDels in genes that have an obvious link to cancer (Figure 3.11).

3.2.3.6 Cdkn2a mutations

WES and subsequent variant calling did not immediately reveal any *Cdkn2a* mutations in E*μ-Myc* lymphomas in either the retrospective or prospective cohorts. *Cdkn2a* (encoding $p19^{ARF}$) is mutated in spontaneous E*μ-Myc* lymphoma at a frequency of approximately 24% (Eischen et al., 1999), so the absence of any mutation or deletion events in the sequenced lymphomas was puzzling. Using IGV, reads spanning the

Cdkn2a locus were visualised manually for each E*μ-Myc* lymphoma under study. E*μ-Myc* lymphomas 4242, 6066, ML52, ML63 and ML353 were found to contain a focal deletion including *Cdkn2a*, which encodes $p19^{ARF}$ (Figure 3.11). Three lymphomas from the prospective cohort demonstrated fewer reads spanning the *Cdkn2a* locus compared to WT littermate control. ML52 (108740340 total reads with 99.38% of reads mapped and 97.62% of target bases with ≥10-fold coverage) and ML353 both had near total loss of reads spanning the *Cdkn2a* locus (Figure 3.12E, Figure 3.13A). ML63 (124637028 total reads with 99.34% of reads mapped and 98.09% of target bases with ≥10-fold coverage) had no reads spanning *Cdkn2a* exons 1α, exon 2 and exon 3 with exon 1β remaining intact (Figure 3.12E, Figure 3.13A).

The apparent loss of reads spanning the *Cdkn2a* locus in lymphomas ML52, ML63 and ML353 was validated using an orthogonal platform. QRT-PCR validated the focal deletions in *Cdkn2a* apparent in lymphoma ML52, ML63 and ML353 (Figure 3.13B). RNA-seq highlighted the presence of focal deletions in *Cdkn2a* in retrospective E*μ-Myc* samples 4242 and 6066, which were supported by a lack of reads across the *Cdkn2a* locus as viewed in IGV (data not shown).

Table 3-3 SNVs found in the retrospective and prospective E*μ-Myc* lymphomas by WES

Sample	Chr	Position	Gene symbol	WES VAF	Human orthologue	# of samples with variant	# of samples with variant in this gene	CGC
ML43	1	33772989	Zfp451	0.41	ZNF451	1	1	.
ML43	1	45376784	Col5a2	0.07	COL5A2	1	2	.
4242	1	34302863	Dst	0.34	DST	1	1	.
4242	1	149901132	Pla2g4a	0.17	PLA2G4A	1	1	.
4242	1	53207367	Pms1	0.39	PMS1	1	1	
6066	1	150412897	Tpr	0.43	TPR	1	1	yes
6066	1	68040416	Erbb4	0.29	ERBB4	1	1	.
218	1	32333510	Khdrbs2	0.41	KHDRBS2	1	1	.
218	1	87099522	Alpi	0.4	ALPI	1	1	.
13	1	93046768	Kif1a	0.14	KIF1A	1	1	.
13	1	17746761	Crispld1	0.42	CRISPLD1	1	1	.
ML39	1	104934342	Cdh20	0.09	CDH20	1	1	.
ML52	1	153799938	Rgsl1	0.07	RGSL1	1	1	.
ML73	2	72971051	Sp3	0.06	SP3	1	2	.
ML39	2	76784591	Ttn	0.2	TTN	1	3	.
ML73	2	76749137	Ttn	0.07	TTN	1	3	.
4242	2	21213238	Thnsl1	0.41	THNSL1	2	2	.
6066	2	21213238	Thnsl1	0.94	THNSL1	2	2	.
ML21	2	73447859	Wipf1	0.42	WIPF1	1	1	.
ML30	2	180689463	Dido1	0.44	DIDO1	1	1	.
ML33	2	92975407	Prdm11	0.44	PRDM11	1	1	.
ML358	2	21648994	Gpr158	0.4	GPR158	1	1	.
ML362	2	86829079	Olfr1094	0.34	OR5T1	1	1	.
ML367	2	155754341	Procr	0.34	PROCR	1	1	.
ML39	2	82256568	Zfp804a	0.42	ZNF804A	1	1	.
ML43	2	157379189	Manbal	0.39	MANBAL	1	1	.
ML52	2	93677449	Alx4	0.49	ALX4	1	1	.
4242	2	138280398	Btbd3	0.13	BTBD3	1	1	.
4242	2	152093547	Scrt2	0.45	SCRT2	1	1	.
6066	2	93813813	Ext2	0.1	EXT2	1	1	
6066	2	153151105	Hck	0.42	HCK	1	1	.
13	2	33569272	Lmx1b	0.09	LMX1B	1	1	.
ML33	2	24997329	Pnpla7	0.18	PNPLA7	1	1	.
4242	2	37790431	Crb2	0.23	CRB2	1	2	.
ML33	3	53549751	Frem2	0.07	FREM2	1	2	.
ML27	3	103060270	Nras	0.54	NRAS	2	2	yes
ML30	3	82383336	Map9	0.32	MAP9	1	1	.
ML39	3	80860235	Glrb	0.24	GLRB	1	1	.

ML39	3	27226123	Nceh1	0.27	NCEH1	1	1	.
ML43	3	103886315	Ptpn22	0.32	PTPN22	1	1	.
6066	3	103060270	Nras	0.71	NRAS	2	2	yes
4242	3	60615689	Mbnl1	0.22	MBNL1	1	1	.
4242	3	59209659	P2ry13	0.2	P2RY13	1	1	.
6066	3	138104958	Mttp	0.6	MTTP	1	1	.
218	3	107458694	Kcnc4	0.12	KCNC4	1	1	.
219	3	144957179	Clca6	0.07	CLCA4	1	1	.
ML21	3	49756648	Pcdh18	0.09	PCDH18	1	1	.
ML27	3	46446171	Pabpc4l	0.06	PABPC4L	1	1	.
ML353	3	54266402	Trpc4	0.05	TRPC4	1	1	.
ML43	3	80801719	Gria2	0.14	GRIA2	1	1	.
6066	3	94764538	Cgn	0.16	CGN	1	2	.
6066	3	88436949	Sema4a	0.12	SEMA4A	1	2	.
4242	4	126165567	Thrap3	0.2	THRAP3	2	2	yes
4242	4	126165496	Thrap3	0.17	THRAP3	2	2	yes
ML358	4	143896971	Pramef6	0.42	PRAMEF12	1	1	.
ML358	4	148252775	Ptchd2	0.34	PTCHD2	1	1	.
ML39	4	148491684	Mtor	0.43	MTOR	1	1	.
ML39	4	75957197	Ptprd	0.44	PTPRD	1	1	.
ML39	4	136052666	Rpl11	0.22	RPL11	1	1	.
ML352	4	129084144	Rnf19b	0.18	RNF19B	1	2	.
ML43	4	141020458	Crocc	0.51	CROCC	1	2	.
299	4	86576310	Rraga	0.46	RRAGA	1	1	.
4242	4	11780099	Cdh17	0.39	CDH17	1	1	.
4242	4	63171423	Kif12	0.39	KIF12	1	1	.
219	4	59806130	Snx30	0.44	SNX30	1	1	.
219	4	118436133	Cdc20	0.75	CDC20	1	1	.
219	4	154282025	Arhgef16	0.29	ARHGEF16	1	1	.
4242	4	126165542	Thrap3	0.22	THRAP3	1	2	yes
4242	4	126165536	Thrap3	0.21	THRAP3	1	2	yes
ML52	5	52166753	Dhx15	0.45	DHX15	1	1	.
ML52	5	124393962	Sbno1	0.47	SBNO1	1	1	.
ML63	5	37824276	Msx1	0.48	MSX1	1	2	.
299	5	135386855	Pom121	0.04	POM121	1	1	.
4242	5	90550205	Afm	0.19	AFM	1	1	.
4242	5	77407551	Igfbp7	0.21	IGFBP7	1	1	.
4242	5	129022093	Ran	0.14	RAN	1	1	.
6066	5	73012300	Slc10a4	0.12	SLC10A4	1	1	.
6066	5	124761080	Dnahc10	0.47	DNAH10	1	1	.
218	5	21303275	Ccdc146	0.38	CCDC146	1	1	.
88	5	129789096	Cct6a	0.12	CCT6A	1	1	.
ML21	5	57720305	Pcdh7	0.1	PCDH7	1	1	.

ML362	5	6770629	Zfp804b	0.1	ZNF804B	1	1	.
219	5	14718868	Pclo	0.09	PCLO	1	2	.
ML352	6	64429755	Grid2	0.5	GRID2	2	2	.
ML353	6	64429755	Grid2	0.28	GRID2	2	2	.
ML367	6	145234355	Kras	0.56	KRAS	2	4	yes
ML58	6	145234355	Kras	0.77	KRAS	2	4	yes
ML33	6	125110108	Chd4	0.54	CHD4	1	1	.
ML33	6	47530602	Ezh2	0.4	EZH2	1	1	yes
ML63	6	4539595	Col1a2	0.48	COL1A2	1	1	.
ML73	6	40571803	Olfr460	0.39	OR9A4	1	1	.
ML353	6	145234351	Kras	0.62	KRAS	1	4	yes
ML358	6	145246771	Kras	0.98	KRAS	1	4	yes
4242	6	51451093	Nfe2l3	0.21	NFE2L3	1	1	.
218	6	48044434	Zfp777	0.17	ZNF777	1	1	.
219	6	107568260	Lrrn1	0.14	LRRN1	1	1	.
ML353	6	23246623	Fezf1	0.18	FEZF1	1	1	.
ML43	6	141563093	Slco1c1	0.08	SLCO1C1	1	1	.
ML352	7	105368397	Olfr692	0.47	OR52W1	2	2	.
ML21	7	127019838	Prrt2	0.51	PRRT2	1	1	.
ML353	7	24585730	Zfp575	0.23	ZNF575	1	1	.
ML367	7	30226919	Tbcb	0.54	TBCB	1	1	.
ML39	7	3339009	Cacng7	0.38	CACNG7	1	1	.
ML58	7	99911809	Xrra1	0.49	XRRA1	1	1	.
ML73	7	19514441	Trappc6a	0.42	TRAPPC6A	1	1	.
299	7	78460326	Ntrk3	0.07	NTRK3	1	1	yes
ML21	7	135172513	Dock1	0.17	DOCK1	1	1	.
ML21	7	103187116	Olfr592	0.07	OR52J3	1	1	.
ML52	7	13007062	Zbtb45	0.13	ZBTB45	1	1	.
299	8	122908752	Ankrd11	0.06	ANKRD11	1	1	.
218	8	26110811	Hook3	0.42	HOOK3	1	1	yes
219	8	57837287	Galntl6	0.23	GALNTL6	1	1	.
ML21	8	13061054	Proz	0.16	PROZ	1	1	.
ML39	8	25724109	Ppapdc1b	0.07	PPAPDC1B	1	1	.
ML52	8	72320905	Klf2	0.1	KLF2	1	1	.
ML52	8	107070475	Terf2	0.15	TERF2	1	1	.
299	8	109623617	Pkd1l3	0.12	PKD1L3	1	2	.
4242	9	106650582	Grm2	0.91	GRM2	2	2	.
ML52	9	107179149	Dock3	0.3	DOCK3	1	1	.
ML73	9	10144172	Cntn5	0.26	CNTN5	1	1	.
4242	9	108559361	Qrich1	0.03	QRICH1	1	1	.
88	9	40314988	Gramd1b	0.33	GRAMD1B	1	1	.
ML52	10	79741779	Polrmt	0.54	POLRMT	1	1	.
299	10	80008425	Abca7	0.35	ABCA7	1	1	.

4242	10	78748361	Ccdc105	0.19	CCDC105	1	1	.
219	10	80325797	Pcsk4	0.08	PCSK4	1	1	.
13	10	27368982	Lama2	0.74	LAMA2	1	1	.
ML20	10	91096336	Ikbip	0.12	IKBIP	1	1	.
ML353	10	59352090	P4ha1	0.18	P4HA1	1	1	.
ML353	10	52050880	Ros1	0.08	ROS1	1	1	yes
ML73	10	79829396	9130017N09Rik	0.1	C19orf21	1	1	.
4242	10	127578119	Lrp1	0.17	LRP1	1	2	.
ML20	11	34267958	Dock2	0.04	DOCK2	2	3	.
ML52	11	53000205	Fstl4	0.5	FSTL4	1	1	.
ML52	11	113874603	Sdk2	0.37	SDK2	1	1	.
ML362	11	72253399	Slc13a5	0.91	SLC13A5	1	2	.
ML30	11	69588388	Trp53	0.93	TP53	1	5	yes
ML362	11	69589547	Trp53	0.2	TP53	1	5	yes
ML73	11	69588401	Trp53	0.05	TP53	1	5	yes
4242	11	34267958	Dock2	0.2	DOCK2	2	3	.
6066	11	98453248	Grb7	0.1	GRB7	1	1	.
219	11	103504283	Lrrc37a	0.1	LRRC37A3	1	1	.
ML21	11	33633165	Kcnip1	0.05	KCNIP1	1	1	.
ML43	11	17945996	Etaa1	0.19	ETAA1	1	1	.
299	11	43420405	Pttg1	0.24	PTTG1	1	2	.
299	11	96039367	Snf8	0.53	SNF8	1	2	.
219	11	9275162	Abca13	0.07	ABCA13	1	2	.
218	11	34688501	Dock2	0.07	DOCK2	1	3	.
13	11	69587279	Trp53	0.91	TP53	1	5	yes
ML20	12	89187318	Nrxn3	0.28	NRXN3	1	1	.
299	12	108158320	Setd3	0.22	SETD3	1	1	.
4242	12	40844404	Dock4	0.2	DOCK4	1	1	.
219	12	81315710	Slc8a3	0.44	SLC8A3	1	1	.
ML358	13	19125060	Amph	0.37	AMPH	1	1	.
ML73	13	8757281	Adarb2	0.57	ADARB2	1	1	.
299	13	70601097	BC018507	0.38	KIAA0947	1	1	.
4242	13	73328436	Ndufs6	0.95	NDUFS6	1	1	.
ML21	13	92439679	Ankrd34b	0.05	ANKRD34B	1	1	.
13	13	105444959	Htr1a	0.88	HTR1A	1	2	.
ML20	14	31266689	Dnahc1	0.43	DNAH1	1	1	.
ML358	14	50681366	Olfr747	0.39	OR11H4	1	1	.
218	14	75125221	Lrrc63	0.23	LRRC63	1	1	.
13	14	68084942	Nefl	0.48	NEFL	1	1	.
ML367	14	32835661	Fam170b	0.07	FAM170B	1	1	.
4242	14	20779245	Camk2g	0.39	CAMK2G	1	2	.
ML353	15	76172181	Plec	0.21	PLEC	1	1	.
ML21	15	55448833	Col14a1	0.21	COL14A1	1	1	.

ML43	15	101532226	Krt84	0.27	KRT84	1	1	.
ML58	15	73320835	Ptk2	0.46	PTK2	1	1	.
4242	15	102154269	Soat2	0.17	SOAT2	1	1	.
ML353	16	36969962	Fbxo40	0.25	FBXO40	1	1	.
ML353	16	34220177	Kalrn	0.44	KALRN	1	1	.
ML52	16	20333612	Abcc5	0.38	ABCC5	1	1	.
ML73	16	24762155	Lpp	0.4	LPP	1	1	yes
ML39	16	4552012	Tfap4	0.45	TFAP4	1	2	.
ML39	16	17281111	Pi4ka	0.14	PI4KA	1	1	.
ML358	17	31213788	Ubash3a	0.23	UBASH3A	1	1	.
218	17	56294300	Arrdc5	0.43	ARRDC5	2	2	.
219	17	56294300	Arrdc5	0.5	ARRDC5	2	2	.
ML20	17	12418768	Plg	0.29	PLG	1	1	.
ML20	17	57053552	Slc25a23	0.42	SLC25A23	1	1	.
ML27	17	32011205	Hsf2bp	0.39	HSF2BP	1	1	.
ML358	17	26198143	Pdia2	0.54	PDIA2	1	1	.
299	17	74619819	Birc6	0.17	BIRC6	1	1	.
4242	17	45568375	Hsp90ab1	0.2	HSP90AB1	1	1	yes
4242	17	45568364	Hsp90ab1	0.22	HSP90AB1	1	1	yes
219	17	43624176	Tdrd6	0.51	TDRD6	1	1	.
219	17	74255291	Memo1	0.05	MEMO1	1	1	.
13	17	55956666	Ebi3	0.2	EBI3	1	1	.
ML21	17	37225983	Olfr96	0.19	OR11A1	1	1	.
4242	17	12237004	Map3k4	0.22	MAP3K4	1	2	.
219	17	37070050	Gabbr1	0.05	GABBR1	1	2	.
ML58	18	25090683	Fhod3	0.43	FHOD3	3	4	.
ML63	18	25090683	Fhod3	0.57	FHOD3	3	4	.
ML52	18	25090683	Fhod3	0.46	FHOD3	3	4	.
ML39	18	89103865	Rttn	0.52	RTTN	1	1	.
ML353	18	74644048	Myo5b	0.07	MYO5B	1	2	.
13	18	68407737	Mc2r	0.58	MC2R	1	1	.
ML58	18	60838861	Tcof1	0.07	TCOF1	1	1	.
ML63	19	44952170	Fam178a	0.5	FAM178A	1	1	.
299	19	27323216	Kcnv2	0.15	KCNV2	1	1	.
4242	19	42604857	Loxl4	0.04	LOXL4	1	1	.
218	19	46551671	Arl3	0.11	ARL3	1	1	.
ML63	19	47253800	Neurl1a	0.06	NEURL	1	1	.
ML73	19	47896517	Itprip	0.13	ITPRIP	1	1	.
ML73	19	4602258	Pcx	0.13	PC	1	1	.
ML33	X	12048474	Bcor	0.37	BCOR	2	8	yes
ML352	X	12047721	Bcor	0.12	BCOR	1	8	yes
ML352	X	74271704	Rpl10	0.17	RPL10	1	2	yes
ML21	X	12048475	Bcor	0.4	BCOR	1	8	yes

ML43	X	12040487	Bcor	0.25	BCOR	1	8	yes
299	X	12048474	Bcor	0.9	BCOR	2	8	yes
88	X	12048327	Bcor	0.91	BCOR	1	8	yes
4242	X	12058550	Bcor	0.87	BCOR	1	8	yes
4242	X	12057624	Bcor	0.93	BCOR	1	8	yes
4242	X	74022612	Irak1	0.45	IRAK1	1	1	.
ML33	X	157350469	Phex	0.05	PHEX	1	1	.
ML39	X	73921324	Naa10	0.15	NAA10	1	1	.
ML52	X	7900968	Timm17b	0.15	TIMM17B	1	1	.
299	X	74272200	Rpl10	0.46	RPL10	1	2	yes

3.2.4 TAM-seq validated WES data

WES data had relatively low coverage across known coding DNA, resulting in mutations being called where the supporting reads were in low number. To validate the SNVs characterized in the previous section (section 3.2.3) a PCR-based orthogonal method, TAM-seq, was utilised. TAM-seq generates thousands of supporting reads across only targeted regions allowing for high confidence variant detection and variant allele frequency (VAF) estimations. VAF by WES correlated strongly with VAF estimate by targeted amplicon sequencing methods (linear regression shows $R = 0.4849$, slope significantly non-zero, $p < 0.05$); however, VAF as estimated by TAM-seq had a narrower range (Figure 3.14A-B). This is likely due to the TAM-seq methodology generating a 10,000-fold increase in read depth compared to WES methodologies. All the lymphomas in the prospective E*μ-Myc* cohort contain variants that clustered at 50% VAF as detected by the highly sensitive TAM-seq method (Figure 3.14B). Most variants clustered at a VAF of 0.5 as would be expected for a monoclonal sample where a heterozygous somatic mutation lies in a diploid chromosomal region. In addition to the dominant clones that harboured mutations at 50% VAF, many of the prospective E*μ-Myc* lymphomas had variants with variant allele frequencies of significantly less than 0.5, indicating that they were possibly sub-clones within a clonal lymphoma population (Figure 3.15). Examples of putative polyclonal E*μ-Myc* lymphomas include ML73 and ML352 where the driver gene mutations (*Trp53* and *Bcor,* respectively) had a VAF significantly less than 0.5.

3.2.5 Copy number information can correct for VAF deviations.

In a monoclonal, diploid population variant allele frequencies would cluster around 50% as with every 'variant' call corresponding to the variant allele there is the 'non-variant call' corresponding to the WT allele. In certain lymphomas, and as is primarily the case with candidate driver mutations, there were frequencies higher than 50% VAF (Figure 3.14A, Figure 3.14B, Figure 3.15). These cases were mostly demonstrated in variants of known oncogenic potential, such as *Trp53*, *Nras* and *Kras.* ML30 harboured a mutation in *Trp53* with a VAF of 100% (Figure 3.15). This likely indicates that at chr11, where *Trp53* is located, there is loss of heterozygosity of the wild type *Trp53* allele or, if there is no deletion event, each copy of *Trp53* would contain the exact same mutation meaning that 100% of the reads across the gene were called as variant reads. No large deletions were found spanning chr11 in lymphoma ML30; however, it is extremely unlikely for identical mutations to occur in each copy of a gene, suggesting that a deletion even that was too small for the window size (<10,000 bases) employed in the CNV analysis to detect has occurred. ML27 harboured an activating *Nras* mutation with a VAF of 67% (Figure 3.15). The *Nras* locus is located on chr6, which is amplified in 2/22 cases, including in ML27. The VAF of 67% coupled with the amplification suggests that the variant copy of *Nras* was amplified and not the wild type copy giving two variant calls for every non-variant call. ML58 and ML353 contained a *Kras* mutation at 67% VAF (Figure 3.15). This can be explained by a segmental gain spanning the *Kras* locus on chr6 in both lymphomas. Finally, ML358 and ML367 harboured *Kras* mutations at 100% VAF (Figure 3.15). This was likely a result of a deletion across the WT *Kras* locus in both cases meaning 100% of the reads across *Kras* will cover the variant allele.

In the cases where a VAF of 67% or 100% did not correlate with the CNV data at that mutation it was likely due to the large window size selected (10,000 bases) to perform the CNV analysis aimed at increasing the specificity and decreasing the amounts of false-positives detected in the data set. This increased specificity meant that the ability to detect smaller amplifications and deletions was omitted.

3.2.6 Co-occurring driver mutations

3.2.6.1 Spontaneously co-occurring driver mutations occur at high variant allele frequencies indicating they simultaneously exist in the dominant clone

E*μ-Myc* lymphomagenesis has been described as a "two-hit" model with the first oncogenic insult being ectopically expressed *Myc* combined with a second driver mutation in a gene such as *Trp53* or *Cdkn2a*, which allows the cells to circumvent the apoptotic properties of Myc (Eischen et al., 1999; Schmitt et al., 1999). In this study, in several E*μ-Myc* lymphomas, instances where there was more than one putative driver mutation detected in a lymphoma. ML353 harboured an activating *Kras*^{Q61H} mutation at 67% VAF in addition to a homozygous deletion event of *Cdkn2a* with very low genomic reads by QRT-PCR, with residual reads likely to be slight contamination by non-cancerous tissue (Figure 3.13A, Figure 3.13B, Figure 3.15). In ML33, there is a loss of function mutation in both *Ezh2* and *Bcor*, with variant allele frequencies of 40% and 41%, respectively, indicating that mutations co-occur in a clonal, diploid tumour (Figure 3.15). Lymphoma 6066 harboured a homozygous deletion of *Cdkn2a* concomitant with an activating *Kras*^{Q61H} mutation (Figure 3.11). In lymphoma 4242, both a homozygous deletion in *Cdkn2a* and an out-of-frame deletion in *Bcor* were detected (Figure 3.11).

These concomitant mutations had comparative VAF in the case of SNVs and near total loss of *Cdkn2a* reads by QRT-PCR indicating that the mutations were likely to be present in the dominant tumour clone as opposed to existing in separate subclones.

3.2.6.2 Single cell bar-coding to demonstrate co-occurring driver mutations can occur in the same cell

To unequivocally demonstrate that two driver mutations can arise in one E*μ-Myc* cell, E*μ-Myc* lymphoma cell lines, 4242 (*Cdkn2a* deletion and *Bcor* deletion) and 6066 (*Cdkn2a* deletion and *Nras*^{Q61K} activating mutation), were "barcoded" with a unique, inert DNA sequence allowing tracking of single-cell derived colonies and subsequent mutational analysis in each colony. To design a complex barcode pool, 198bp oligonucleotide (FL-index) containing a 16bp variable barcode sequence was designed, resulting in a theoretical library size of 4,194,304 unique barcodes. The barcode inserts were cloned into an inert MSCV construct that was also endowed with a BFP cassette. In the context of cellular barcoding, it is necessary to transduce the target cells with a low multiplicity of infection (MOI), lowering the probability of multiple barcodes being introduced into the same cell. A target MOI of 0.1 was chosen as, at this fraction, only 4.9% of the BFP-positive population will theoretically contain two or more barcodes and 95.1% of the transduced population will contain one barcode. Lymphoma cell lines 4242 and 6066 were transduced with MSCV.Index at the desired MOI, generating 4242-MSCV.Index and 6066-MSCV.Index cell lines. From the transduced pool, the BFP positive populations were single cell sorted into individual wells and were allowed to propagate *in vitro*. Sanger sequencing demonstrated the each 6066-MSCV.Index and

4242-MSCV.Index colonies harboured a single, unique barcode confirming that the colonies had proliferated from a single cell (Figure 3.16A). QRT-PCR was performed on each of the 4242 and 6066 single cell-derived colonies and each colony demonstrated near-zero reads across *Cdkn2a* compared to WT DNA (Figure 3.16B). Lymphoma cell line 4242 is known to contain a *Bcor* deletion of 217bp in exon 4. PCR primers were designed flanking this region, with the resultant DNA products Sanger-sequenced; demonstrating that the predicted 217bp deletion in *Bcor* was present in each of the single-cell derived 4242-MSCV.Index colonies (Figure 3.16C). Lymphoma 6066 is known to harbour an activating $Nras^{Q61K}$ mutation, correlating to a C-to-A amino acid change at coding position 181. Sanger sequencing revealed the presence of the activating $Nras^{Q61K}$ mutation in each of the three single cell-derived 6066-MSCV.Index colonies (Figure 3.16D). Taken together, these results indicate that both the *Cdkn2a* deletion and the *Bcor* deletion or activating $Nras^{Q61K}$ mutations were present in the same cell for lymphoma 4242 or 6066, respectively.

3.2.7 Activating mutations in the context of *Cdkn2a*

A surprising finding from the sequencing of E*μ-Myc* lymphomas was the identification of additional candidate driving mutations in the context of *Cdkn2a* loss, namely for lymphomas 4242 (217bp deletion in *Bcor* exon 4), 6066 ($Nras^{Q61K}$) and ML353 ($Kras^{Q61H}$). To further determine if *Cdkn2a* deletion can occur in the context of other lymphoma-driving mutations and at what frequency, Eμ-*Myc* mice were crossed to $Cdkn2a^{-/-}$ mice to generate heterozygous E*μ-Myc*;$Cdkn2a^{+/-}$ mice, time until sacrifice due to lymphoma burden was assessed and the resultant lymphomas were sequenced.

It has previously been shown that E*μ-Myc*;*Cdkn2a*$^{+/-}$ mice invariably lost their WT allele (at a frequency of 88.2%) resulting in what were ostensibly *Cdkn2a*$^{-/-}$ E*μ-Myc* lymphomas (Schmitt et al., 1999). Given the previous data shown to suggest that additional driver mutations can occur in the context of *Cdkn2a* deletions, the question of whether these E*μ-Myc*;*Cdkn2a*$^{+/-}$ mice that have additional selective pressure to develop loss of heterozygosity at the remaining WT *Cdkn2a* allele, will develop tertiary mutations to candidate lymphomagenesis-driver genes. Given the pressure to develop loss of heterozygosity (LOH) at the *Cdkn2a* allele, E*μ-Myc*;*Cdkn2a*$^{+/-}$ mice (n=6) demonstrated accelerated lymphomagenesis compared to the E*μ-Myc* (prospective cohort, n=16) transgenic animals, with a median survival time of 69.5 days compared to 103 days ($p<0.05$, log-rank (Mantel-Cox) test), respectively (Figure 3.17A, Table 3-4). WES on the spontaneous E*μ-Myc*;*Cdkn2a*$^{+/-}$ lymphomas was performed. As expected, compared to WT, there was a clear loss in reads across *Cdkn2a* exon 1β, which encodes the p19^{ARF} protein, in all of the E*μ-Myc*;*Cdkn2a*$^{+/-}$ lymphomas, indicating LOH in 6/6 cases investigated (Figure 3.17B, Table 3-4). Further analysis of the WES data revealed that 1/6 of the E*μ-Myc*;*Cdkn2a*$^{-/+}$ (E*μ-Myc*;*Cdkn2a*$^{+/-}$ #3) lymphomas harboured a heterozygous pathogenic *Kras*^{Q61H} mutation (Figure 3.18A, Table 3-4). The *Kras*^{Q61H} mutation was present at a VAF of 67%, which is interestingly seen at the same VAF in the prospective cohort (ML58 and ML353); indicating that this lymphoma gained an extra copy of the variant *Kras* allele on chr3 (Figure 3.18A). To functionally validate the presence of the *Kras* mutation, the signaling pathway downstream of the Ras proteins was dissected. *Nras*^{Q61K} is a constitutively active variant of *Nras*, causing constant stimulation MAP kinase pathway and leads to increased levels of phospho-Erk (p-Erk)

(Avruch et al., 2001). *Kras*^{Q61H} operates functionally similar to *Nras*^{Q61K} in this pathway (Misale et al., 2012). Immunoblotting assays of 6066 (known *Nras*^{Q61K} mutant), 4242 (known *Nras* WT) and E*μ-Myc*;*Cdkn2a*$^{+/-}$ #3 demonstrated increased p-Erk signals in lymphoma 6066 and E*μ-Myc*;*Cdkn2a*$^{+/-}$ #3 compared to lymphoma 4242, indicating a more active Ras-signaling pathway biologically confirming the sequencing result (Figure 3.18B). Finally, to prove LOH of the remaining *Cdkn2a* allele in this lymphoma, QRT-PCR was performed and demonstrated loss of *Cdkn2a* exon 1β (encoding p19^{ARF}) in E*μ-Myc*;*Cdkn2a*$^{+/-}$ #3 compared to WT (Figure 3.18C). Collectively, these data suggest that heterozygous loss of *Cdkn2a* creates a selective pressure resulting in the loss of the wild type *Cdkn2a* allele; however, concomitant tertiary and quaternary driver mutations in cancer-associated genes such as *Kras, Nras* and *Bcor* can, and do, occur in E*μ-Myc* lymphoma.

Table 3-4 SNVs found in the E*μ-Myc;Cdkn2a*$^{+/-}$ lymphomas by WES

Sample	Chr	Position	Gene symbol	WES VAF	# of samples with variant	# of samples with variant in this gene
Eu-Myc-p19ARF-4	1	31204097	Lags	0.06	1	2
Eu-Myc-p19ARF-4	1	93076316	Kif1a	0.12	1	3
Eu-Myc-p19ARF-4	1	191354602	Ppp2r5a	0.08	1	1
Eu-Myc-p19ARF-5	1	107536082	SERPINB10	0.11	1	1
Eu-Myc-p19ARF-5	1	107536082	Serpinb10	0.11	1	1
Eu-Myc-p19ARF-5	1	185407086	APRs	0.12	1	3
Eu-Myc-p19ARF-2	1	153762124	Nasal	0.12	1	1
Eu-Myc-p19ARF-2	1	174012264	Ifi205	0.02	1	1
Eu-Myc-p19ARF-7	1	57943210	Spats2l	0.11	1	2
Eu-Myc-p19ARF-7	1	91105214	Lrrfip1	0.06	1	1
Eu-Myc-p19ARF-4	2	117298503	Rasgrp1	0.07	1	1
Eu-Myc-p19ARF-6	2	180538822	Ntsr1	0.12	1	1
Eu-Myc-p19ARF-2	2	32256052	Uck1	0.13	1	1
Eu-Myc-p19ARF-2	2	36377442	Olfr338	0.09	1	1
Eu-Myc-p19ARF-2	2	126754958	Usp8	0.06	1	3
Eu-Myc-p19ARF-3	2	109137429	Mettl15	0.07	1	1
Eu-Myc-p19ARF-3	2	164649807	Wfdc9	0.13	1	1
Eu-Myc-p19ARF-7	2	28839881	Gtf3c4	0.11	1	1
Eu-Myc-p19ARF-7	2	181991343	Gm14496	0.13	1	2
Eu-Myc-p19ARF-4	3	62340518	Arhgef26	0.05	1	2
Eu-Myc-p19ARF-4	4	21678854	Prdm13	0.18	1	1
Eu-Myc-p19ARF-5	4	126578328	Clasp	0.17	1	2
Eu-Myc-p19ARF-6	4	133681020	Arid1a	0.08	1	2
Eu-Myc-p19ARF-2	4	139779625	Pax7	0.55	2	3
Eu-Myc-p19ARF-3	4	139779625	Pax7	0.51	2	3
Eu-Myc-p19ARF-3	4	155842607	Mxra8	0.11	1	3
Eu-Myc-p19ARF-7	4	15976046	Nan	0.12	1	2
Eu-Myc-p19ARF-7	4	46196056	Xpc	0.14	1	3
Eu-Myc-p19ARF-4	5	24391785	Atg9b	0.11	1	3
Eu-Myc-p19ARF-4	5	139758567	Ints1	0.4	1	1
Eu-Myc-p19ARF-4	5	150070368	Rxfp2	0.12	1	3
Eu-Myc-p19ARF-4	5	151586871	Vmn2r18	0.12	1	1
Eu-Myc-p19ARF-6	5	129844866	Cct6a	0.5	1	2
Eu-Myc-p19ARF-2	5	136911128	Rabl5	0.05	1	2
Eu-Myc-p19ARF-2	5	146948906	Gtf3a	0.12	1	2
Eu-Myc-p19ARF-3	5	24598604	Smarcd3	0.17	1	1
Eu-Myc-p19ARF-3	5	115885013	Cit	0.12	1	3
Eu-Myc-p19ARF-7	5	119680782	Tbx3	0.12	1	1
Eu-Myc-p19ARF-4	6	132700320	Tas2r121	0.27	1	1
Eu-Myc-p19ARF-3	6	145234355	Kras	0.67	1	2

Eu-Myc-p19ARF-7	6	122711575	Mango	0.14	1	1
Eu-Myc-p19ARF-4	7	111051024	Ctr9	0.12	1	2
Eu-Myc-p19ARF-5	7	114562943	Cyp2r1	0.14	1	1
Eu-Myc-p19ARF-5	7	122078492	Ubfd1	0.6	1	1
Eu-Myc-p19ARF-6	7	79464988	Plg	0.39	1	1
Eu-Myc-p19ARF-6	7	120407502	Abca15	0.04	1	3
Eu-Myc-p19ARF-2	7	25914893	Cyp2b10	0.28	1	1
Eu-Myc-p19ARF-3	7	24303439	Zfp94	0.13	1	1
Eu-Myc-p19ARF-3	7	28299077	Dll3	0.15	1	1
Eu-Myc-p19ARF-7	7	4993583	Zfp579	0.38	1	3
Eu-Myc-p19ARF-4	8	105620284	Fam65a	0.43	2	2
Eu-Myc-p19ARF-5	8	110957730	St3gal2	0.55	1	2
Eu-Myc-p19ARF-2	8	4213732	BC068157	0.33	1	1
Eu-Myc-p19ARF-2	8	71508003	Plan	0.26	1	1
Eu-Myc-p19ARF-2	8	105620284	Fam65a	0.46	2	2
Eu-Myc-p19ARF-4	9	8004192	Yap1	0.12	1	2
Eu-Myc-p19ARF-4	9	38930732	Olfr930	0.41	1	1
Eu-Myc-p19ARF-4	9	110617846	Setd2	0.08	1	2
Eu-Myc-p19ARF-5	9	14593284	Amotl1	0.15	1	2
Eu-Myc-p19ARF-5	9	45001070	Cd3e	0.12	1	1
Eu-Myc-p19ARF-5	9	58312549	Loxl1	0.1	1	1
Eu-Myc-p19ARF-5	9	69999104	Bnip2	0.11	1	2
Eu-Myc-p19ARF-6	9	20878196	Angptl6	0.22	1	1
Eu-Myc-p19ARF-6	9	39426969	Olfr952	0.21	1	1
Eu-Myc-p19ARF-7	9	100553592	Slc35g2	0.13	1	1
Eu-Myc-p19ARF-3	10	29147076	Soga3	0.17	1	1
Eu-Myc-p19ARF-7	10	53750769	Fam184a	0.07	1	3
Eu-Myc-p19ARF-7	10	81145048	Zbtb7a	0.12	1	1
Eu-Myc-p19ARF-4	11	62131899	Specc1	0.11	1	2
Eu-Myc-p19ARF-4	11	69122798	Hes7	0.25	1	1
Eu-Myc-p19ARF-4	11	102690322	Ccdc43	0.13	1	3
Eu-Myc-p19ARF-5	11	69369143	Chd3	0.15	1	1
Eu-Myc-p19ARF-2	11	52121909	Ppp2ca	0.19	1	1
Eu-Myc-p19ARF-3	11	86512908	Rps6kb1	0.11	1	1
Eu-Myc-p19ARF-3	11	120786358	Gps1	0.12	1	2
Eu-Myc-p19ARF-7	11	4740503	Cabp7	0.13	1	3
Eu-Myc-p19ARF-7	11	71098959	Nlrp1a	0.05	1	2
Eu-Myc-p19ARF-3	12	98658326	Spata7	0.48	1	3
Eu-Myc-p19ARF-7	12	51950043	Heatr5a	0.15	1	2
Eu-Myc-p19ARF-4	13	14316548	Hecw1	0.43	1	1
Eu-Myc-p19ARF-4	14	57525585	Il17d	0.18	1	2
Eu-Myc-p19ARF-6	14	47530597	Fbxo34	0.11	1	1
Eu-Myc-p19ARF-2	14	51009612	Rnase10	0.07	1	1

Eu-Myc-p19ARF-3	14	75328115	Zc3h13	0.2	1	3
Eu-Myc-p19ARF-7	14	62292929	Dleu7	0.12	1	2
Eu-Myc-p19ARF-4	15	88780741	Zbed4	0.06	1	1
Eu-Myc-p19ARF-2	15	34083407	Math	0.33	1	2
Eu-Myc-p19ARF-3	15	76890742	Zfp7	0.12	1	1
Eu-Myc-p19ARF-3	15	81904831	Aco2	0.15	1	1
Eu-Myc-p19ARF-4	17	12928618	Acat3	0.03	1	2
Eu-Myc-p19ARF-4	17	34716435	Tunb	0.03	1	2
Eu-Myc-p19ARF-4	17	80424185	Sos1	0.1	1	3
Eu-Myc-p19ARF-6	17	28330566	Rpl10a	0.28	1	1
Eu-Myc-p19ARF-2	17	32202361	Brd4	0.11	1	1
Eu-Myc-p19ARF-3	17	32798651	Zfp811	0.47	1	1
Eu-Myc-p19ARF-3	17	47699546	Frs3	0.11	1	2
Eu-Myc-p19ARF-4	18	76137309	Zbtb7c	0.17	1	1
Eu-Myc-p19ARF-6	18	22120589	Ccdc178	0.6	2	2
Eu-Myc-p19ARF-6	18	40218821	Yipf5	0.07	1	1
Eu-Myc-p19ARF-2	18	9917186	Gm5501	0.13	1	1
Eu-Myc-p19ARF-7	18	22120589	Ccdc178	0.53	2	2
Eu-Myc-p19ARF-2	19	8775446	Taf6l	0.51	1	1
Eu-Myc-p19ARF-7	19	10478261	Ppp1r32	0.03	1	1
Eu-Myc-p19ARF-4	X	7664938	Prickle3	0.11	1	3
Eu-Myc-p19ARF-4	X	79351175	Gm8787	0.17	1	1
Eu-Myc-p19ARF-5	X	101265020	Il2rg	0.15	1	1
Eu-Myc-p19ARF-5	X	101265020	Gm20489	0.15	1	1
Eu-Myc-p19ARF-5	X	120905969	Pcdh11x	0.13	1	2
Eu-Myc-p19ARF-6	X	8893412	B630019K06Rik	0.25	1	1
Eu-Myc-p19ARF-6	X	73794424	Pdzd4	0.13	1	1
Eu-Myc-p19ARF-6	X	152660010	Shroom2	0.19	1	2
Eu-Myc-p19ARF-2	X	18146364	Dusp21	0.12	1	1
Eu-Myc-p19ARF-2	X	134691377	Armcx4	0.13	1	2
Eu-Myc-p19ARF-3	X	73570400	Atp2b3	0.41	1	3
Eu-Myc-p19ARF-7	X	7922502	Pcsk1n	0.2	1	2
Eu-Myc-p19ARF-7	X	73919536	Naa10	0.05	1	1
Eu-Myc-p19ARF-7	X	136246444	Wbp5	0.29	1	3
Eu-Myc-p19ARF-7	Y	1314598	Usp9y	0.14	1	1

3.2.8 Monitoring temporal acquisition of mutations in E*μ-Myc* lymphomas

3.2.8.1 Relationship between variant and overall survival

E*μ-Myc* mice undergo predictable premalignant to malignant leukemic phases during lymphomagenesis therefore providing a progression model to study the temporal acquisition of somatic mutations leading to lymphomagenesis (Harris et al., 1988; Sidman et al., 1993). Serial sampling of B cells ($B220^{+}$) from the blood of individual animals four weeks after birth and then at fortnightly intervals until the mice were sacrificed due to progressive lymphoma was performed. TAM-seq was performed retrospectively on the serial B cell DNA isolates targetting a panel of mutations identified from WES with high VAF representative of likely or putative driver mutations within the end-stage lymphomas. This gene panel consisted of *Bcor, Trp53, Nras, Kras, Ezh2, Mtor, Cntn5* and *Adarb2*. Ideally, the gene panel would have also included *Cdkn2a*; however, the breakpoints in the deletions could not be mapped and there was not enough material to perform QRT-PCR on the DNA from B cell isolates. Given the apoptotic effects of Myc are mediated by activation of the $p19^{ARF}$-Mdm2-Trp53 apoptotic pathway, it was predicted that any mutations attenuating this pathway would be the most potent oncogenic 'second hit'. Similarly, given *Ras*-family proteins are implicated in E*μ-Myc* lymphomagenesis, it was hypothesised that activating mutations in *Nras* and *Kras* would also be superiorly potent in driving lymphomagenesis when compared to *Bcor* mutations. To test these predictions, the prospective E*μ-Myc* cohort was grouped by mutation (*Bcor, N/Kras, Trp53, Cdkn2a* or other/unknown) and their time until

sacrifice due to progressive lymphoma was measured. There was no correlation between the type of candidate driver gene that was mutated and the total survival time of the animal ($p>0.05$, Log-rank (Mantel-Cox) test) (Figure 3.19A). Mice that developed a *Bcor* mutation had a median survival time of 111 days from birth (n=4). Mice that were found to have a *Cdkn2a* deletion had a median survival time of 103 days (n=3). Mice that developed either an *Nras* or *Kras* mutations had a median survival time of 63 days from birth (n=5). Mice that developed loss-of-function mutations in *Trp53* had a median survival time of 95 days (n=3). Mice ML20 and ML39 had a median survival time of 146 days (Figure 3.19A). While there was a trend demonstrating that the *Nras* and *Kras* tumours developed at an accelerated rate compared to the others, larger numbers of E*μ-Myc* mice would be needed in the study to show a statistically significant hierarchy of oncogenic potential among the common driver mutations.

3.2.8.2 Time from variant detection in the peripheral blood to sacrifice

The E*μ-Myc* model has a well-documented pre-leukaemic phase preceding any signs of tumour burden in which the B cells rapid rate of division and growth is counteracted by their rapid rate of apoptosis. Given this information about the E*μ-Myc* model, it was hypothesized that the type of variant will not affect overall survival of the transgenic mice, but that the type of variant will affect survival time from the point of variant detection in the peripheral blood. To test this hypothesis, a retrospective survival analysis was performed with animals from the prospective E*μ-Myc* cohort grouped by variant type (*Bcor, Trp53,* or *Nras* and *Kras*) and time until sacrifice from the time at which the driver variant was identified in the peripheral blood was measured. There

was no correlation between mutation type or candidate mutation and survival time once that specific mutation is detectable in the isolated peripheral blood B cells ($p > 0.05$, Log-rank (Mantel-Cox) test) (Figure 3.19B). Once *Bcor* mutations identified in the lymphomas were detectable in the B cells isolated from peripheral blood of E*μ-Myc* transgenic mice, the median survival time was 21 days (n=4) compared to 14 days and 11 days for *Trp53* (n=3) mutations and either *Nras* or *Kras* mutations (n=5), respectively (Figure 3.19B). E*μ-Myc* mice that had mutations in unknown drivers could not be included in this analysis as there was no target variant we could track in the peripheral blood B cells. Similarly, mice with deletions in *Cdkn2a* could not be included in this analysis as to measure the presence of a deletion by PCR requires either a) knowledge of the exact breakpoints or b) a guaranteed 100% pure population, neither of which were possible to obtain.

3.2.8.3 Tracking mutation- and disease-burden in the prospective Eμ-Myc cohort

The prospective E*μ-Myc* lymphoma cohort was set up so that over time, blood samples were collected and analysed for signs of leukaemia and so that B cells could be isolated, have their DNA extracted and be stored for later use. Mutations discovered in end-stage E*μ-Myc* lymphomas could then be tracked using TAM-seq throughout each corresponding E*μ-Myc* animal's history and correlated with the first signs of leukaemic burden, which in this study is defined as a WBC count of $\geq 14 \times 10^9$ cells/L. At 4-weeks of age the prospective cohort of E*μ-Myc* mice were not leukaemic and the candidate driver mutations present at end-stage for each E*μ-Myc* lymphoma were either not present, or present at a very low VAF in the peripheral blood B cells of each mouse (Figure 3.20).

This indicated that the mutations identified by WES were somatically acquired and were not germline variants. In most cases a rapid expansion of WBC in the peripheral blood of affected mice correlated with the increase in relative VAF of the candidate driver mutations identified in each lymphoma. The *Bcor* variant found in lymphoma ML21 was detectable in the peripheral blood B cells of the parental mouse at 12-weeks of age with a VAF of 0.12, correlating with a WBC count of 13.4×10^9 cells/L (Figure 3.21). At 14-weeks and 16-weeks of age the VAF of the ML21 *Bcor* mutation had increased to 0.29 and 0.46, respectively, while the WBC count had initially decreased at week 14 to 10.2×10^9 cells/L before an increase at week 16 to 24.6×10^9 cells/L (Figure 3.21). E*μ-Myc* transgenic mouse ML21 demonstrated moribund lymphoma at week 17 with a *Bcor* VAF of 0.51 and a WBC count of 34.2×10^9 cells/L. Lymphoma ML27 harboured an activating $Nras^{Q61K}$ mutation that is first detectable in the blood of E*μ-Myc* transgenic mouse ML27 at 6-weeks of age with a normal WBC count of 9×10^9 cells/L (Figure 3.21). At week 7.3 E*μ-Myc* mouse ML27 had terminal lymphoma burden and was culled with a peripheral blood $Nras^{Q61K}$ mutation present at a frequency of 0.65 and a WBC count of 121.6×10^9 cells/L (Figure 3.21). Lymphoma ML30 harboured a *Trp53* mutation, which was first detected in the peripheral blood B cells of E*μ-Myc* mouse ML30 at week 12 and was detected at a VAF of 0.68 with a corresponding WBC count of 26×10^9 cells/L (Figure 3.21). E*μ-Myc* mouse ML30 was sacrificed at week 13.6 with a corresponding *Trp53* VAF of 0.99 in the peripheral blood B cells and a WBC count of 49.4×10^9 cells/L (Figure 3.21). It was hypothesised that lymphoma ML33 was driven by *Bcor* and *Ezh2* mutations, both of which were first identified in the peripheral blood B cells of mouse ML33 at week 12 with a VAF of 0.34 and 0.44, respectively (Figure 3.21).

Corresponding to the first detection of the *Bcor* and *Ezh2* mutations, E*μ-Myc* mouse ML33 was demonstrably leukaemic at week 12 with a WBC count of 29.4×10^9 cells/L (Figure 3.21). By week 14 the VAF of *Bcor* and *Ezh2* mutations in peripheral blood B cells had increased to 0.38 and 0.44, respectively, corresponding to a concurrent increase in WBC count to 71.6×10^9 cells/L in mouse ML33 (Figure 3.21). *Eμ-Myc* mouse ML33 was sacrificed at week 14.7 with progressive lymphoma and a variant frequency in the peripheral blood B cells of 0.43 for the *Bcor* variant and 0.49 for the *Ezh2* variant. At the time of sacrifice mouse ML33 presented with a WBC count of 64.8×10^9 cells/L, which had decreased slightly since week 14 but was still considered leukaemic (Figure 3.21). A mutation in *Mtor* was identified in lymphoma ML39 and was hypothesised to be the secondary oncogenic insult required for lymphomagenesis. The *Mtor* variant identified in lymphoma ML39 was first detected in the peripheral blood B cells of E*μ-Myc* mouse ML39 at week 22 at a VAF of 0.35 with a corresponding WBC count of 8.4×10^9 cells/L, which was pre-leukaemic (Figure 3.21). At week 24, mouse ML39 demonstrated signs and symptoms of progressive lymphoma and was sacrificed with an *Mtor* VAF of 0.51 in the peripheral blood B cells and a WBC count of 45.2×10^9 cells/L (Figure 3.21). Lymphoma ML42 contained a *Bcor* mutation that was first detected in the peripheral blood B cells of E*μ-Myc* mouse ML42 at week 8 with a VAF of 0.31 (Figure 3.21). At week 8, mouse ML43 was pre-leukaemic with a WBC count 0f 9.8×10^9 cells/L (Figure 3.21). The *Bcor* VAF increased to 0.42 by week 10, which corresponded with a concurrent increase in the WBC count to 27.4×10^9 cells/L (Figure 3.21). At week 11.3, E*μ-Myc* mouse ML43 was culled with progressive disease and had a *Bcor* VAF of 0.50 in the peripheral blood B cells and a WBC count of 50.4×10^9 cells/L

(Figure 3.21). Lymphoma ML58 harboured a *Kras*^{Q61H} mutation that was first identified in peripheral blood B cells of E*μ-Myc* mouse ML58 at 6-weeks of age at a VAF of 0.06 and a corresponding WBC count of 41.2×10^9 cells/L (Figure 3.21). At week 8 the *Kras*^{Q61H} VAF had mildly increased to 0.09 while the WBC count was measured at 55.6×10^9 cells/L in E*μ-Myc* mouse ML58 (Figure 3.21). At 9-weeks of age, ML58 was culled with signs of progressive lymphoma, a *Kras*^{Q61H} mutation at a VAF of 0.68 in the peripheral blood B cells and a WBC count of 158.8×10^9 cells/L (Figure 3.21). In the case of lymphoma ML73 there appeared to be multiple clones that harboured either *Trp53, Cntn5* or *Adarb2* mutations given the apparent lack of correlation between the variant allele frequencies of each of the variants (Figure 3.21). The earliest point at which mutations in lymphoma ML73 were found in peripheral blood B cells of mouse ML73 was at week 10 with the *Adarb2* mutation identified with a VAF of 0.18 and a correlating peripheral WBC count of 52.4x10^9 cells/L, which was the first blood count indicating the animal was leukaemic (Figure 3.21). The VAF of the *Adarb2* mutation spontaneously regressed at week 12 despite increased peripheral blood WBC count (76.8×10^9/L) (Figure 3.21). In circulating B cells at week 14 the mutation in *Trp53* was detectable at a VAF of 0.35, the *Cntn5* mutation was detectable at VAF of 0.1 and the *Adarb2* mutation was detectable at a VAF of 0.44, however, no reading could be estimated for total peripheral blood WBC count due to machine error (Figure 3.21). At week 16.28 mouse ML73 was moribund, which correlated with an unexpected decline in the *Trp53* mutation VAF to 0.08, a steady *Adarb2* mutation VAF of 0.44 and a rise in the VAF of *Cntn5* to 0.24 with a correlating peripheral blood WBC count of 145.2x10^9 cells/L. Lymphoma ML352 contained a *Bcor* mutation, which was first identified in the

peripheral blood B cells of mouse ML352 at week 18 and at a VAF of 0.04 (Figure 3.21). At week 18.6, mouse ML352 was sacrificed with signs of progressive lymphoma; however, at this time-point the *Bcor* VAF was 0.14 and the WBC count was 5.0×10^9 cells/L (Figure 3.21). Lymphoma ML352 also harboured an *Rpl9* mutation that was not tracked in the stored B cell DNA. Lymphoma ML353 harboured a *Kras*^{E63K} mutation that was first detected in the peripheral blood B cells of mouse ML353 at a VAF of 0.48 at week 30, corresponding to a WBC count of 5.4×10^9 cells/L (Figure 3.21). At week 32 the *Kras*^{E63K} VAF had increased to 0.69 while the WBC count remained relatively stable at 11.6×10^9 cells/L (Figure 3.21). E*μ-Myc* mouse ML353 was culled at week 34 with signs and symptoms of progressive lymphoma corresponding with an increased WBC count of 36.8×10^9 cells/L and a *Kras*^{E63K} VAF of 0.65 (Figure 3.21). The *Kras*^{G12D} mutation identified in lymphoma ML358 was first detected in the peripheral blood B cells of mouse ML358 at a VAF of 0.39 by week 6, corresponding to a WBC count of 25.6×10^9 cells/L (Figure 3.21). E*μ-Myc* mouse ML358 was culled at week 8 with a WBC burden of 42.2×10^9 cells/L and a *Kras*^{G12D} VAF of 0.99 in the peripheral blood B cells (Figure 3.21). The *Trp53* mutation identified in lymphoma ML362 was first detected in mouse ML362 at the time of sacrifice due to progressive lymphoma. At week 8 variant *Trp53* was detected at an allele frequency of 0.48, corresponding to a WBC count of 65.8×10^9 cells/L (Figure 3.21). Lymphoma ML367 harboured a *Kras*^{Q61H} mutation that was first detected in the peripheral blood B cells of mouse ML367 at week 22 with a concurrent WBC count of 34.2×10^9 cells/L (Figure 3.21). At week 22.6, E*μ-Myc* mouse ML367 was culled with mortal disease burden correlating with a *Kras*^{Q61H} VAF of 0.90 and a WBC count of 94.0×10^9 cells/L (Figure 3.21). Lymphomas ML52, ML63 and

ML353 harboured a *Cdkn2a* deletion, which was not included in this analysis due to the breakpoints remaining undefined and the limited material available at each time point. No candidate driver SNV was identified in lymphoma ML20 and as such it was not included in the temporal acquisition of mutations analysis.

3.3 Discussion

Many human cancers are characterised by aberrant MYC expression through structural rearrangements, copy number increases, stabilising mutations and other cryptic genomic events (Meyer and Penn, 2008). The E*μ-Myc* transgenic mouse was developed as a model of MYC-driven malignancy, genomically mimicking the t(8;14) translocation in which *MYC* is juxtaposed to *IGH* enhancer elements, which is characteristic of BL (Adams et al., 1985). Endogenous MYC drives S-phase entry in the cell cycle, and therefore, MYC can confer a selective growth advantage to cells by providing ectopic proliferation signals when overexpression is forced (Eilers et al., 1991; Mateyak et al., 1997). However, aside from driving increased proliferation, ectopic MYC overexpression drives cellular apoptotic programs (Askew et al., 1991; Evan et al., 1992). This is consistent with studies demonstrating markedly accelerated lymphomagenesis in E*μ-Myc* mice crossed with mice that have deregulated apoptotic pathways, such as E*μ-Bcl2* transgenic and *Trp53*$^{+/-}$ mice (Hsu et al., 1995; Schmitt et al., 1999; Strasser et al., 1990). The clinical course, latency and gene expression profile of E*μ-Myc* lymphomas are remarkably heterogeneous despite being driven by a single oncogene on a homozygous genetic background, indicating that secondary, tertiary and quaternary mutations must be acquired to drive lymphomagenesis (Mori et al., 2008). Mutations that disrupt Myc-induced tumour suppressive pathways such as inactivating mutations in *Cdkn2a* (encoding p19^{ARF}) and *Trp53*, allowing the growth promoting effects of Myc to predominate above the apoptotic signals, occur *de novo* in up to half of all E*μ-Myc* lymphomas (Eischen et al., 1999; Schmitt et al., 1999). The E*μ-Myc* model has also been used to implicate the intrinsic apoptotic response in Myc-

driven oncogenesis, as over-expression of pro-survival genes such as *Bcl2*, *Bcl$_{xl}$* and *Mcl1* or deletion of pro-apoptotic BH3-only genes such as *Bim, Bmf* and *Bad* can also expedite lymphoma onset in the E*μ-Myc* model (Egle et al., 2004; Frenzel et al., 2010; Michalak et al., 2009; Strasser et al., 1990; Whitecross et al., 2009). However, unlike perturbations to the p19^{ARF}-Mdm2-Tp53 tumour suppressor pathway, *de novo* mutations in *Bcl2*-family genes have yet to be detected in spontaneously arising E*μ-Myc* lymphoma. Despite evidence that naturally occurring mutations in *Cdkn2a*, *Trp53* and *Nras* can cooperate with Myc in driving E*μ-Myc* lymphomagenesis, other, potentially novel and clinically relevant, cooperating lesions remain undefined in up to half of all cases (Alexander et al., 1989; Eischen et al., 1999). Through application of next generation sequencing technologies, the mutational landscape of E*μ-Myc* lymphoma was defined in this chapter leading to the identification of *Bcl6*-co-repressor (*Bcor*) as the most recurrently mutated gene in E*μ-Myc* lymphoma and the discovery that multiple cooperative oncogenic events may be required to allow for malignant transformation.

3.3.1 E*μ-Myc* transgene topology and insertion site

To date, this was the first comprehensive genome-wide sequencing analysis of lymphoid malignancies arising in the E*μ-Myc* transgenic mouse model. WGS revealed there were three copies of the *Myc* transgene in the hemizygous E*μ-Myc* germline, associated with a germline amplicon on chr19. These data recapitulated findings published in a contemporary study conducted by Fusello *et al* that also identified the germline chr19 amplicon and its association with the E*μ-Myc* transgene insertion site (Fusello et al., 2013). In this thesis, the observation of the germline chr19 amplicon was

extended to include two distinct E*μ-Myc* colonies, confirming that the amplification of the region on chr19 was likely a founding event in the generation of the E*μ-Myc* model rather than an institute-dependent genetic artifact. This finding suggested that E*μ-Myc* transgenic mice contain a higher *Myc*-gene dosage than previously assumed.

The amplified 3Mb region of DNA proximal to the E*μ-Myc* transgene on chr19 is syntenic to the human 9p24.1 region, which is frequently amplified in B cell malignancies. This region, both in murine and in human, contains the tumour promoting genes *JAK-2* and *CD274* (PD-L1) (Green et al., 2010). The possibility that the extra copies of transgenic *Myc* and the amplification of a region containing *Jak-2* and *Pd-l1* as a founding germline event can potentially explain why gene expression profiles of E*μ-Myc* lymphomas have been found to resemble large B cell lymphomas rather than Burkitt's Lymphoma (Green et al., 2010; Mori et al., 2008). Fusello *et al* recognised the chr19 amplicon would harbour *Cd274* (encoding Pd-l1) and demonstrated that Pd-l1 expression was significantly enhanced on circulating lymphocytes in E*μ-Myc* mice compared to WT animals, suggesting that in this case gene dosage affects expression (Fusello et al., 2013). Contrary to that finding, in this chapter it was demonstrated that there was additional somatic gain of the chr19 amplicon and the E*μ-Myc* transgene in approximately half of the E*μ-Myc* lymphomas tested, and that despite copy number differences in *Cd274*, Pd-l1 expression remained stable between tumours. These data indicated that Pd-l1 expression was not driven exclusively by gene dosage, and that positive selection of E*μ-Myc* lymphoma clones with enhanced Pd-l1 expression is likely. Furthermore, the extra somatic gain of the chr19 amplicon is a high frequency recurrent

copy number gain; however, not all E*μ-Myc* lymphomas demonstrated extra somatic gain of this region suggesting it is not absolutely required for malignant transformation of E*μ-Myc* lymphoma. *Jak-2* amplifications in E*μ-Myc* lymphoma have been previously described and thought to cooperate with Myc to drive lymphomagenesis (Fusello et al., 2013). Contrary to the hypothesis made by Fusello *et al*, results detailed in this chapter demonstrated that E*μ-Myc* lymphomas were insensitive to Jak-2 inhibition as they lacked appreciable Jak-2-pathway activation as assessed by phosphorylation of downstream proteins. This indicated that germline or somatic amplification of *Jak-2* was unlikely to contribute to E*μ-Myc* oncogenesis. However, the chr19 amplicon contains another 41 genes such as *Il33*, which is known to promote gastric cancer cell invasion (Yu et al., 2015) and *Pdcd1lg2* (Pd-l2), which is implicated as an oncogene in primary mediastinal large B cell lymphoma (Twa et al., 2014) suggesting that there is much potential for further investigations dissecting the importance of this region in cooperating with oncogenically expressed Myc in driving E*μ-Myc* lymphomagenesis. E*μ-Myc* mimics the t(8;14) translocation characteristic of BL; however, discoveries made in this investigation highlight that with further germline and somatic amplification of a region containing known drivers of lymphoma such as *Jak-2*, *Cd274, Il33* and *Pdcd1lg2* in addition to 39 other genes comes the implication that *Myc* is potentially not the exclusive driver of lymphomagenesis that was introduced when the E*μ-Myc* model was created. These results have implication for the large body of drug screens and forward genetic screens conducted in the E*μ-Myc* model of lymphoma based on the premise of *Myc* being the sole initiator of lymphomagenesis.

3.3.2 Clonal concomitant driver mutations in E*μ-Myc* lymphoma

An overarching paradigm established from previous studies using the Eμ-*Myc* model was that a single co-operating lesion in the p19^{ARF}-Mdm2-Trp53 pathway or the Ras pathway would be sufficient cooperate with oncogenic Myc and drive tumor development (Alexander et al., 1989; Eischen et al., 1999; Schmitt et al., 2002). However, data presented in this thesis challenges this hypothesis, at least with regards to the role of *Cdkn2a* (encoding p19^{ARF}), as more than one putative "driver lesion" could be detected in some lymphomas. This was exemplified in lymphomas ML353, 6066 and 4242 where *Kras*^{Q61H}, *Nras*^{Q61K} and *Bcor* were mutated, respectively, in addition to *Cdkn2a* loss. Further supporting this was the observation that in 1/6 E*μ-Myc;Cdkn2a*$^{+/-}$ lymphoma, there was loss of heterozygosity at the remaining WT *Cdkn2a* allele with a contemporaneous *Kras*^{Q61H} mutation present at a high VAF. Importantly, through high VAF suggesting that InDels were clonal and single-cell barcoding experiments it was demonstrated that the co-occurrence of driver mutations in the E*μ-Myc* lymphoma samples were indeed two mutations in a single clone as opposed to two mutations in two sub-clones. Taken together, these data indicate that while heterozygous loss of *Cdkn2a* creates a selective pressure that resulted in the loss of the second wild type *Cdkn2a* allele, it did not fully ameliorate selective pressure for gain- or loss-of-function mutations in other genes outside of the p19^{ARF}-Mdm2-Trp53 tumour suppressor axis and that perhaps the importance of *Cdkn2a* loss is overstated in E*μ-Myc* lymphoma. Indeed, these observations are consistent with a previous report made by Bertwistle *et al* indicating that on its own, *Cdkn2a* inactivation leading to loss of p19^{ARF} was insufficient in cooperating with Myc and

driving lymphomagenesis in E*μ-Myc* mice (Bertwistle and Sherr, 2007). Bertwistle *et al* posited that there must be additional "cryptic" mutations other than *Cdkn2a* loss accompanying the conversion of pre-malignant E*μ-Myc* B cells in the setting of E*μ-Myc*;*Cdkn2a*$^{+/-}$ lymphoma. The identification activating mutations in the *Ras* family and loss-of-function mutations in *Bcor* in this thesis can be defined as two such "cryptic" events.

3.3.3 Whole exome sequencing revealed novel *de novo* mutations in E*μ-Myc* lymphoma

In addition to expected mutations in *Trp53, Cdkn2a, Nras* and *Kras,* WES identified novel *de novo* mutations that have previously been unreported in E*μ-Myc* lymphoma and are represented in human malignancies. A high VAF mutation in *Mtor* was identified in E*μ-Myc* lymphoma ML39 at a residue conserved in the human gene. Previously, deregulation of the *Mtor* signalling pathway has been shown to cooperate with Myc in driving E*μ-Myc* lymphomagenesis (Wendel et al., 2004) and activating *MTOR* mutations have been described in DLBCL (Grabiner et al., 2014); however, this study is the first to identify the spontaneous acquisition of a potentially activating *Mtor* mutation in E*μ-Myc* lymphoma. An essential splice site variant in the gene *Ezh2* was identified in E*μ-Myc* lymphoma ML33. Ezh2 is a subunit of PRC-2 and has been implicated as a tumour suppressor gene in E*μ-Myc* lymphoma (Lee et al., 2013). Through genetic knockdown of Ezh2, Lee *et al* demonstrated that E*μ-Myc* lymphomas developed at an accelerated rate in the absence of functional Ezh2 and, resultantly

PRC-2, leading to the hypothesis that the PRC-2 complex is tumour suppressive in E*μ-Myc* lymphoma (Lee et al., 2013). Data presented in this chapter is the first to highlight the occurrence of a *de novo* inactivating mutation of *Ezh2* in E*μ-Myc* lymphoma, supporting the hypothesis previously made. Similarly, a mutation in *mir142*, which is deregulated in up to 20% of DLBCL (Kwanhian et al., 2012), was demonstrated to occur spontaneously in lymphoma LN88. Taken together, these spontaneously occurring low frequency mutations indicated that these are pathways that are placed under selective pressure, driven by ectopic Myc expression, to become deregulated. This supported the hypothesis that clinically relevant Myc-cooperative mutations remained undiscovered in E*μ-Myc* lymphoma. Further genomic studies utilising greater numbers of E*μ-Myc* lymphomas are required to define whether mutations in these genes are recurrent and at what frequency, or whether the number of low frequency mutations identified will increase linearly with increasing experimental numbers, as seen with DLBCL (Lohr et al., 2012).

Bcor was identified as the most frequently mutated gene in spontaneous E*μ-Myc* lymphoma, having been acquired in 32% of cases tested. All mutations observed in *Bcor* were predicted loss of function mutations by frame-shift InDels or stop-codon introduction by missense mutations. Using the “20/20” rule promoted by Vogelstein *et al,* as more than 20% of the mutations in *Bcor* were inactivating mutations, *Bcor* likely functions as tumour suppressor gene in E*μ-Myc* lymphoma (Vogelstein et al., 2013). Similar protein truncating *BCOR* mutations have also been recently reported in both solid and liquid (Table 3-5) human malignancies strongly supporting a tumour-

suppressive function in most cases. Furthermore, *Bcor* mutations increased from low VAF at first detection in circulating peripheral blood B cells to high VAF in parallel with increasing white blood cell counts, indicating that there is positive selection, clonal outgrowth and progressive lymphomagenesis following *Bcor* inactivation. BCOR is a subunit of a variant polycomb group (PcG) PRC-1 distinguished by its association with PCGF1 and the histone modifying subunits RING1B and KDM2B (Gearhart et al., 2006; Ghetu et al., 2008; Junco et al., 2013). Importantly, the variant BCOR/PRC-1 complex is functionally distinct from the canonical PRC-1 complex containing the oncogene BMI, first identified through an E*μ-Myc* retroviral mutagenesis screen (Haupt et al., 1991; Haupt et al., 1993; van Lohuizen et al., 1991). As previously discussed Lee *et al*, posits that PRC-1 and PRC-2 complexes have opposing activity in E*μ-Myc* lymphoma, with destabilisation of the PRC-1 complexes through *Bmi1* knockdown leading to delayed tumourigenesis while destabilisation of the PRC-2 complex through *Ezh2* knockdown accelerated disease onset (Haupt et al., 1993; Jacobs et al., 1999; Lee et al., 2013). However, in *Eμ-Myc* lymphoma, mutations affecting *Bcor* were exclusively predicted to be loss-of-function mutations, suggesting that Bcor-containing-PRC-1 complexes are tumour suppressive and act distinctly from oncogenic Bmi1-containing-PRC-1 complexes in E*μ-Myc* lymphoma. Interestingly, E*μ-Myc* lymphoma ML33 harboured a high VAF *Ezh2* loss-of-function mutation in addition to a high VAF *Bcor* inactivating mutation, indicating that the tumour suppressive processes mediated by PRC-1 and PRC-2 are likely to be independent of one another. While the E*μ-Myc* model has been useful as a tool for understanding the interactions between the intrinsic apoptotic pathway, activating mutations in pro-survival *Bcl2*-family members and inactivating

mutations in pro-apoptotic *BH3*-only proteins are not observed, *de novo*. This indicates that the activation or inactivation of these pathways do not need to be overcome for E*μ-Myc* lymphomagenesis to occur, suggesting that their respective promotion or restraint of the Myc driven proliferative response is limited. The identification of novel *de novo* mutations in E*μ-Myc* lymphoma has yielded valuable information about genes and pathways that positively or negatively influence the proto-oncogenic effects of Myc.

Table 3-5 BCOR mutations in human haematological malignancies

Disease (COSMIC v73)	Mutation count	Cases screened	Category
Acute lymphoblastic B cell leukaemia		106	Acute, non-myeloid leukaemia
Acute leukaemia		1	Acute, non-myeloid leukaemia
Acute leukaemia of ambiguous lineage		2	Acute, non-myeloid leukaemia
Acute lymphoblastic leukaemia	2	51	Acute, non-myeloid leukaemia
Acute lymphoblastic T cell leukaemia	4	214	Acute, non-myeloid leukaemia
B cell lymphoma unspecified		9	Aggressive B-NHL
Burkitt's lymphoma		64	Aggressive B-NHL
Diffuse large B cell lymphoma		252	Aggressive B-NHL
Follicular lymphoma		35	Aggressive B-NHL
Primary central nervous system lymphoma		9	Aggressive B-NHL
Acute myeloid leukaemia	36	923	AML/MDS
Acute myeloid leukaemia associated with MDS	5	44	AML/MDS
Acute myeloid leukaemia therapy related		27	AML/MDS
Blast phase chronic myeloid leukaemia		7	AML/MDS
Chronic myelomonocytic leukaemia	3	77	AML/MDS
Myelodysplastic syndrome	42	882	AML/MDS
Myelodysplastic-myeloproliferative neoplasm-unclassifiable		16	AML/MDS
Chronic myeloid leukaemia		28	Chronic myeloproliferative neoplasms
Myeloproliferative neoplasm		2	Chronic myeloproliferative neoplasms
Chronic eosinophilic leukaemia-hypereosinophilic syndrome		1	Chronic myeloproliferative neoplasms
Chronic neutrophilic leukaemia		1	Chronic myeloproliferative neoplasms
Essential thrombocythaemia		61	Chronic myeloproliferative neoplasms
Mast cell neoplasm		3	Chronic myeloproliferative neoplasms
Myelofibrosis		45	Chronic myeloproliferative neoplasms
Polycythaemia vera		48	Chronic myeloproliferative neoplasms
Hodgkin lymphoma		6	Hodgkin lymphoma
Hairy cell leukaemia	1	24	Low-grade B cell neoplasia
Mantle cell lymphoma		41	Low-grade B cell neoplasia
Marginal zone lymphoma		15	Low-grade B cell neoplasia
Chronic lymphocytic leukaemia-small lymphocytic lymphoma	5	425	Low-grade B cell neoplasia
Plasma cell myeloma		45	Low-grade B cell neoplasia
Anaplastic large cell lymphoma		2	Mature T-cell malignancies
Adult T cell lymphoma-leukaemia		1	Mature T-cell malignancies
Angioimmunoblastic T cell lymphoma		7	Mature T-cell malignancies
Enteropathy type T cell lymphoma		1	Mature T-cell malignancies
NK-T cell lymphoma		68	Mature T-cell malignancies
Peripheral T cell lymphoma unspecified		34	Mature T-cell malignancies
T cell large granular lymphocytic leukaemia		2	Mature T-cell malignancies
Mycosis fungoides-Sezary syndrome	1	13	Mature T-cell malignancies
Blastic plasmacytoid dendritic cell neoplasm		3	Other
Langerhans cell histiocytosis		3	Other

3.3.4 Chapter conclusions

E*μ-Myc* lymphoma has for a long time been considered heterogeneous in nature; however, investigations have failed in demonstrating the occurrence of *de novo* mutations outside of *Cdkn2a*, *Trp53* and *Ras*-family members. Additionally, only recently has the complexity of the transgene insertion site been demonstrated (Fusello et al., 2013), with this investigation highlighting further somatically acquired complexities surrounding the E*μ-Myc* transgene insertion and chr19 amplicon and suggesting that there is selective pressure placed on acquiring further gains of transgenic *Myc* and/or the chr19 amplicon, which contains a suite of genes implicated in malignancy. The identification of multiple known driver mutations within a clonal tumour indicates that individual clones evolve over time, requiring multiple cooperative events to enable for malignant transformation. Furthermore, *Cdkn2a* appears to be the common gene co-mutated with additional drivers, indicating that perhaps the importance of *Cdkn2a* in restraining oncogenic Myc activity is overstated. Despite extensive use of the E*μ-Myc* model to identify Myc-cooperative pathways and cancer gene discovery through forward genetic screens, this study being the first to comprehensively sequence E*μ-Myc* lymphomas, identified. The identification of novel *de novo* mutations in E*μ-Myc* is important, as it is suggestive of the selective pressure placed on these genes and the role they must play counteracting the proliferative effect when Myc signalling is unchecked. Strikingly, the most frequently mutated gene in E*μ-Myc* lymphoma is *Bcor*, which appears to be a tumour suppressor gene as mutations in *Bcor* are exclusively inactivating. In addition to being identified in this study, *BCOR* inactivating mutations have been described at low frequencies in human malignancies; however, biological

validation of the tumour suppressive activity of *BCOR* has never been explored. The E*μ-Myc* model provides a powerful platform to explore the hypothesis that *Bcor* is a tumour suppressor gene, capable of restraining oncogenic Myc activity.

3.4 Chapter 3 figures and figure legends

The corresponding figure legend can be found on the page following each figure.

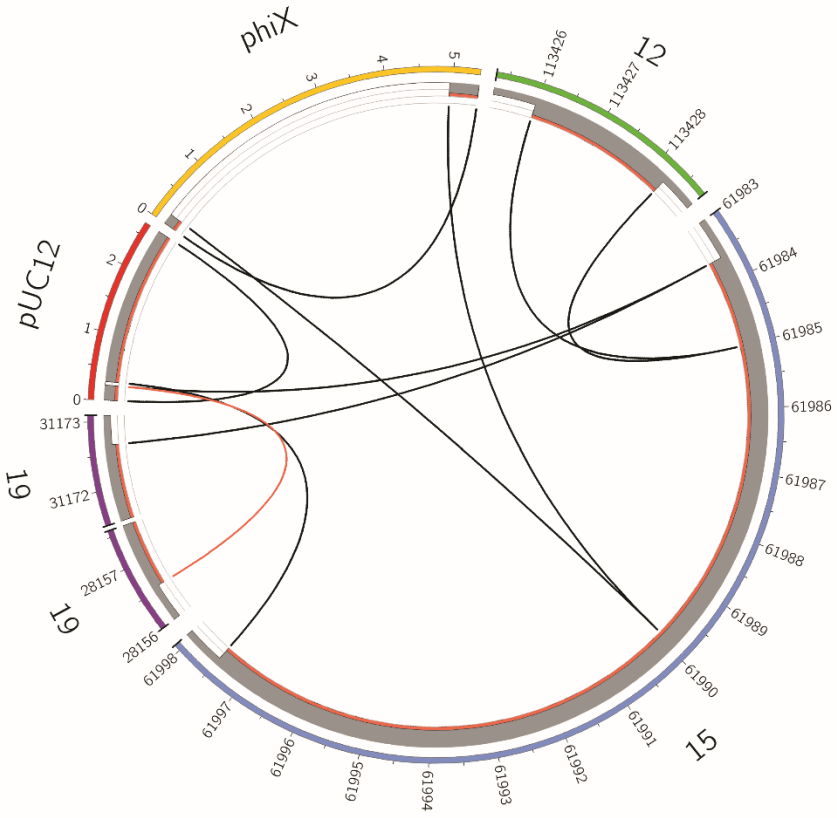

phiX
pUC12
19
19
21
15
0
1
2
3
4
5
0
1
2
31173
31172
28157
28156
61998
61997
61996
61995
61994
61993
61992
61991
61990
61989
61988
61987
61986
61985
61984
61983
113428
113427
113426

Figure 3.01 – Transgene insertion site and associated genomic alterations in E*μ-Myc* lymphoma

Circos plot showing the E*μ-Myc* transgenic cassette (pUC12, phiX, chr12 (E*μ*), chr15 (*Myc*)), identified breakpoints and estimated copy number. Outer coloured bars depict chromosome and transgene segments. Grey bars and links represent germline copies with two copies per increment. Red bars represent additional somatic copy-number gain and breakpoints.

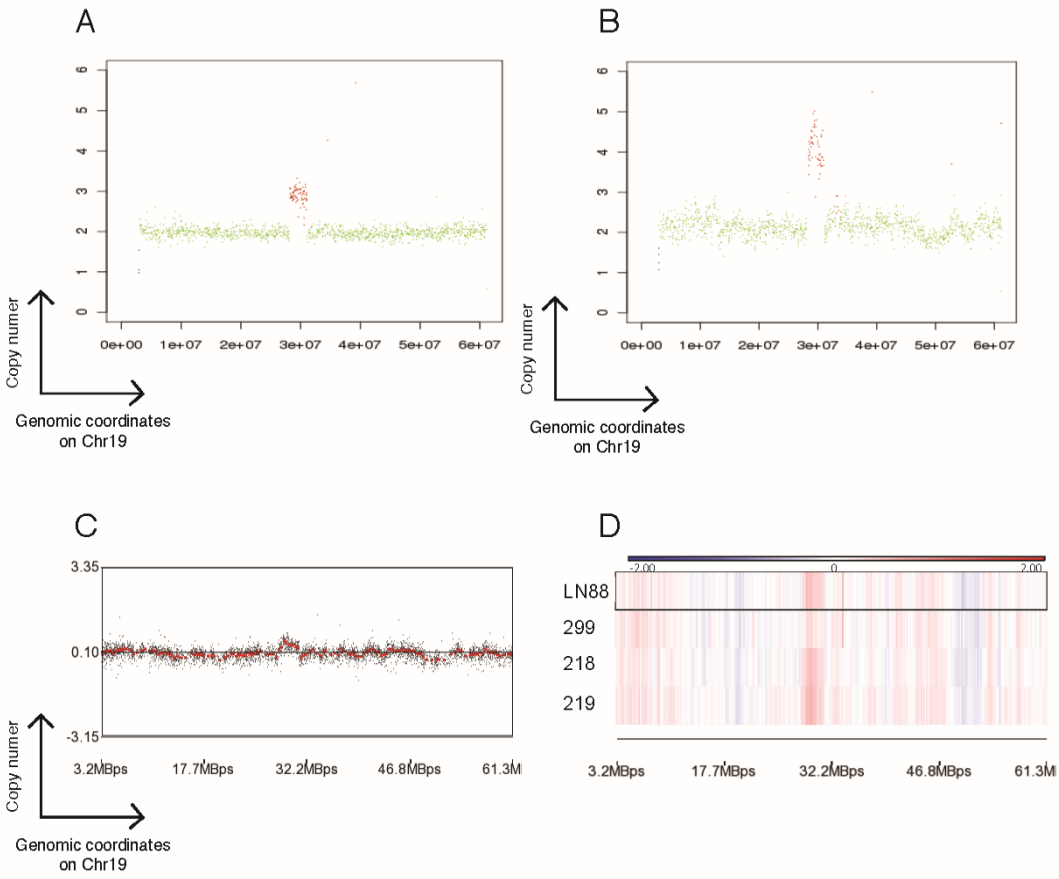
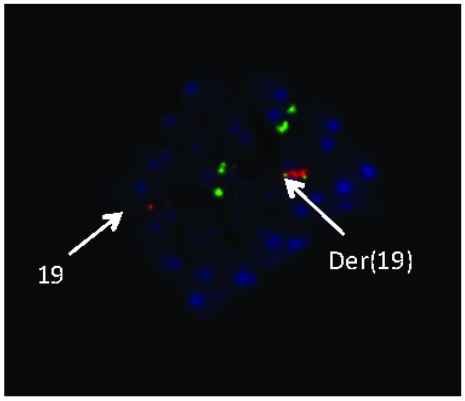

A
B
C
D
Copy numer
Genomic coordinates
on Chr19
0e+00
1e+07
2e+07
3e+07
4e+07
5e+07
6e+07
3.35
0.10
-3.15
3.2MBps
17.7MBps
32.2MBps
46.8MBps
61.3M
-2.00
0
2.00
LN88
299
218
219

Figure 3.02 – Copy number analysis of the chr19 amplicon

A) WGS of tail88 demonstrated three copies of the chr19 amplicon.
B) WGS of LN88 demonstrated four copies of the chr19 amplicon.
C) Copy number information gained from WGS of tail88 was subtracted from LN88, indicating that an additional copy of the chr19 amplicon was somatically gained.
D) Germline copy number information was subtracted from LN88, lymphoma 299, lymphoma 218 and lymphoma 219, demonstrating that there has been somatic gain of an extra copy of the chr19 amplicon in LN88, lymphoma 218 and lymphoma 219.

A



B

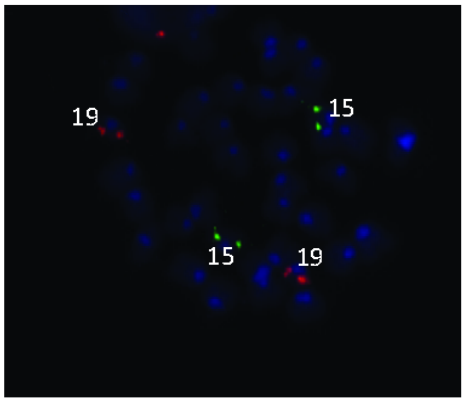


Figure 3.03 – BAC metaphase FISH on E*μ-Myc* transgenic bone marrow cultured cells and wild type controls.

A) Eμ-*Myc* bone marrow cultured cells: BAC RP23-307D14 (spectrum green) spans the *Myc* locus on chromosome 15. BAC RP23-324L2 (spectrum orange) spans *Jak2* on chromosome 19. There was an extra, half strength green signal (*Myc*) co-localising with a red signal (*Jak2*) on the derivative chromosome 19.

B) Wild type cultured cells: Localisation of BAC RP23-307D14 (spectrum green) to each copy of chromosome 15 and localisation of a BAC RP23-324L2 (spectrum orange) to each copy of chromosome 19. Spectrum green and spectrum orange signals were of equal strength and there was no co-localisation of *Myc* and *Jak2.*

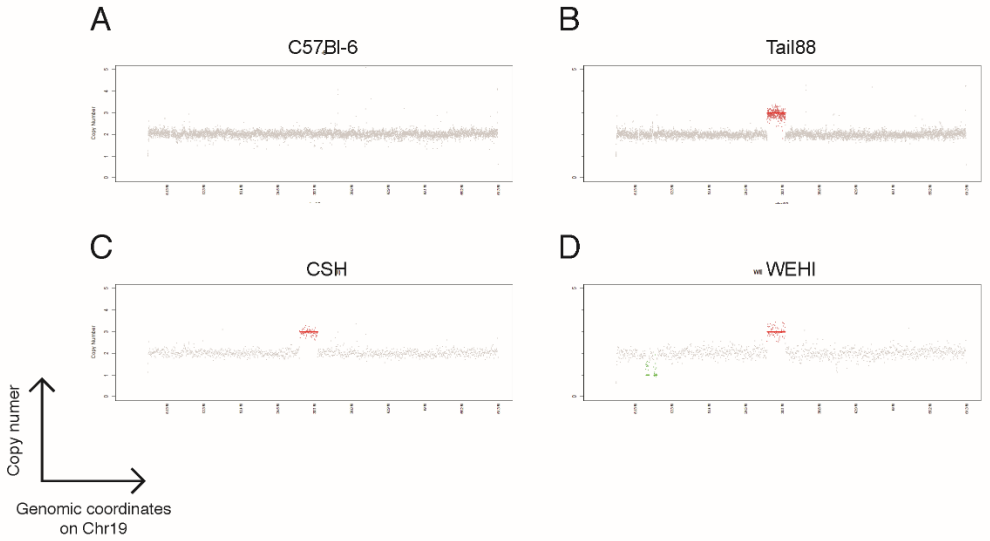
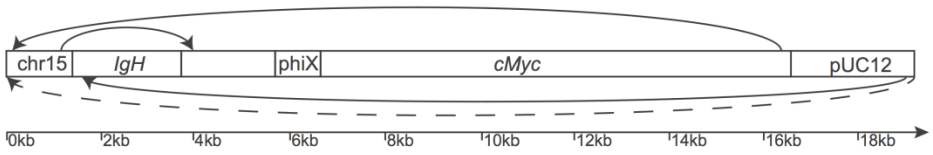

A
C57Bl-6
B
Tail88
C
CSH
D
WEHI
Copy numer
Genomic coordinates
on Chr19

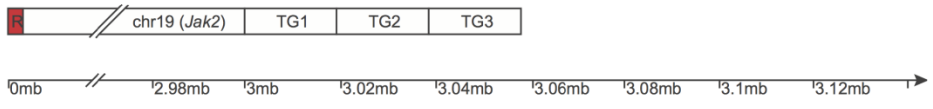
Figure 3.04 - Copy number analysis spanning chromosome 19 in three independently derived Eμ-*Myc* germline DNA samples

WGS and subsequent copy-number analysis highlighted germline amplification of the Chr19 segment proximal to, and including, the E*u-Myc* transgene insertion in three independently derived E*μ-Myc* colonies. A) C57BL/6 WT control B) E*u-Myc* Peter MacCallum Cancer C) E*u-Myc* Cold Spring Harbor D) E*u-Myc* WEHI.

A


chr15
IgH
phiX
cMyc
pUC12
0kb
2kb
4kb
6kb
8kb
10kb
12kb
14kb
16kb
18kb

B


R
chr19 (Jak2)
TG1
TG2
TG3
0mb
2.98mb
3mb
3.02mb
3.04mb
3.06mb
3.08mb
3.1mb
3.12mb

C

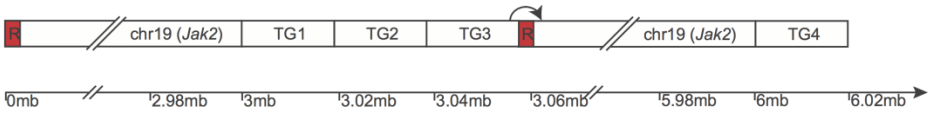

R
chr19 (Jak2)
TG1
TG2
TG3
R
chr19 (Jak2)
TG4
0mb
2.98mb
3mb
3.02mb
3.04mb
3.06mb
5.98mb
6mb
6.02mb

Figure 3.05 - E*μ-Myc* transgenic architecture

A) Schematic demonstrating the arrangement of the E*μ-Myc* transgenic cassette in the germline.
B) Schematic demonstrating the insertion and arrangement of the three in-sequence repeats of the E*μ-Myc* transgene (TG1-3) in the germline sample; tail88.
C) Schematic highlighting the extra gain of the transgene and the segment of Chr19 in somatic sample; LN88.

'R' in the red box indicates unmappable repetitive DNA elements.

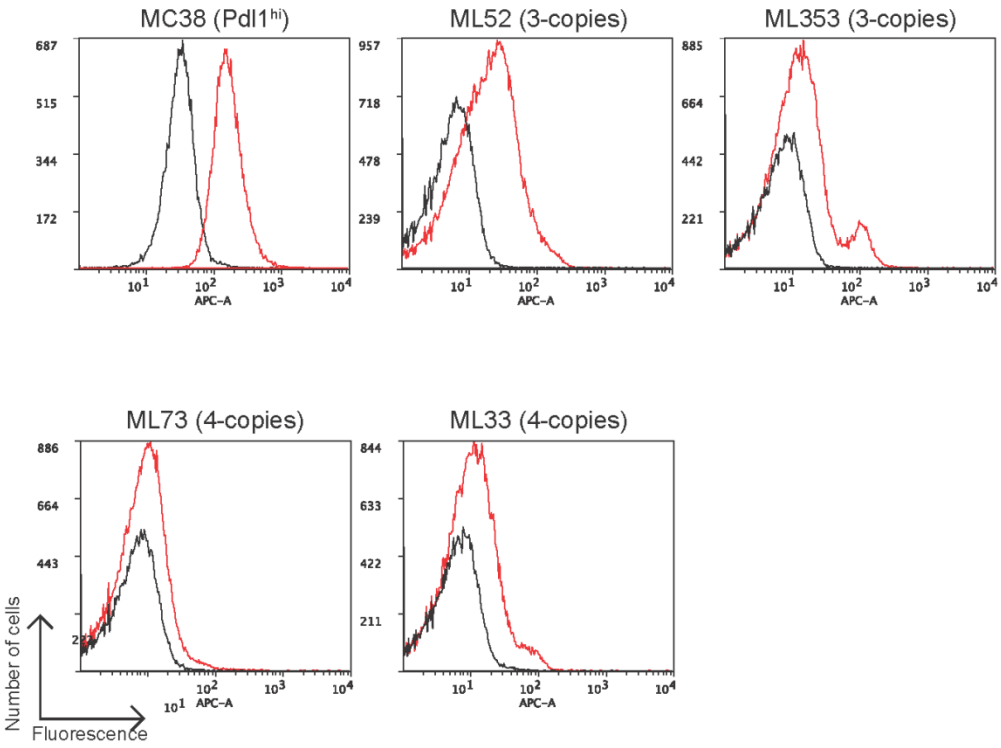
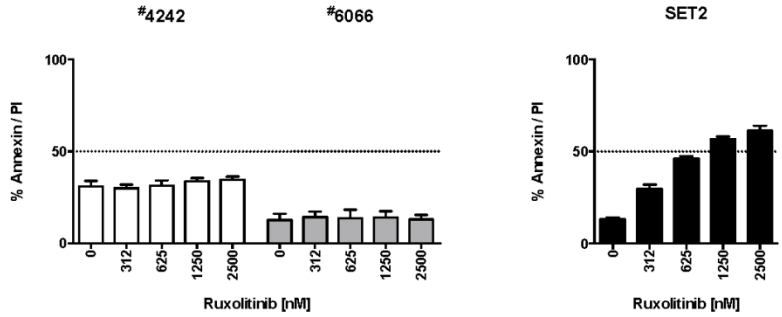

MC38 (Pdl1hi)
ML52 (3-copies)
ML353 (3-copies)
ML73 (4-copies)
ML33 (4-copies)
687
515
344
172
957
718
478
239
885
664
442
221
886
664
443
844
633
422
211
APC-A
Number of cells
Fluorescence

Figure 3.06 – Analysing Pd-l1 expression in E*μ-Myc* lymphomas

FACS analysis of Pd-l1 expression in E*μ-Myc* lymphomas was shown to have varying *Cd274* (Pd-l1) genomic copy-number. MC38 is a murine colorectal carcinoma cell line and Pd-l1(Hi) control. Lymphomas ML52 and ML353 harboured 3-copies of the chromosome 19 amplicon (including *Cd274*), lymphomas ML73 and ML33 harboured 4 copies of the chromosome 19 amplicon (including *Cd274*). ML52 and ML353 had a mean fluorescence intensity for Pd-l1 of 107 and 5.3, respectively, while ML73 and ML33 demonstrated a Pd-l1 mean fluorescence intensity of 1.3 and 5.3, respectively. Viable cells were gated based on morphology and CD220 positivity (not shown). Rat IgG2α isotype control (black), PD-L1 streptavidin APC - biotin APC Cy7 (red).

A


#4242
#6066
SET2
% Annexin / PI
100
50
0
0
312
625
1250
2500
Ruxolitinib [nM]

B

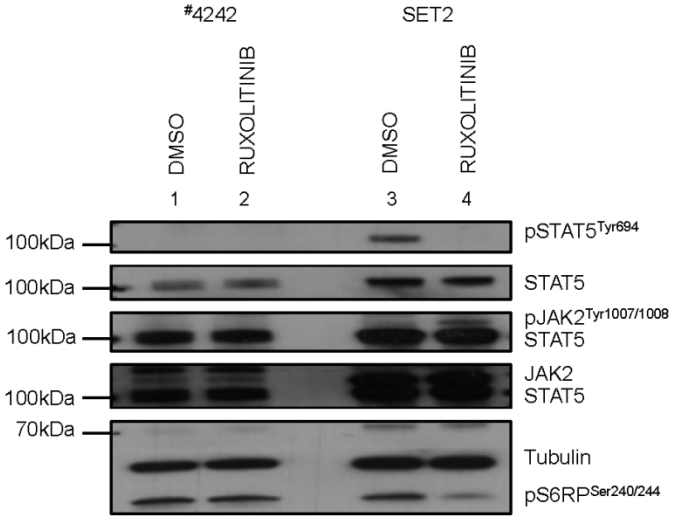
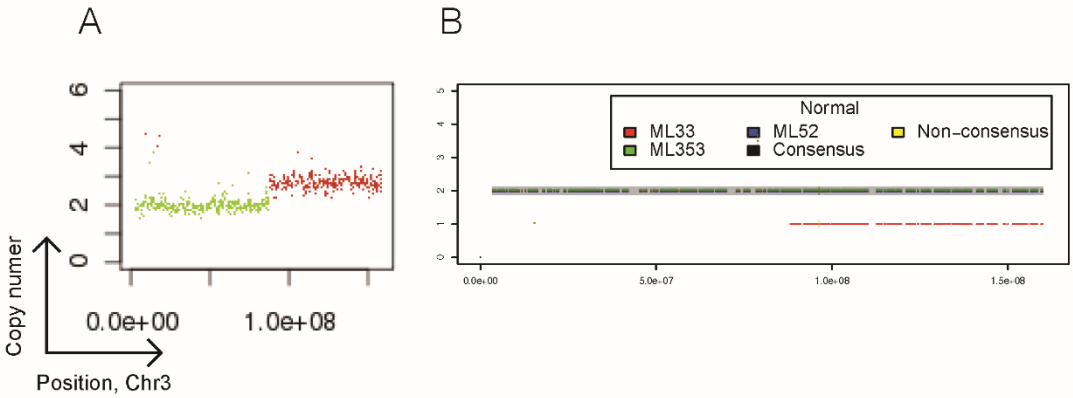

#4242
SET2
DMSO
RUXOLITINIB
1
2
3
4
100kDa
pSTAT5Tyr694
STAT5
pJAK2Tyr1007/1008
JAK2
70kDa
Tubulin
pS6RPSer240/244

Figure 3.07 –E*μ-Myc* lymphoma cell lines treated with ruxolitinib

A) E*μ-Myc* lymphomas #4242 and #6066 (left) were cultured in the presence of doubling dilutions of ruxolitinib for 24 hours prior to flow cytometric analysis for annexin-V / propidium iodide uptake. Human SET2(V617F) cell line (right) was used as positive control for ruxolitinib, cultured in RPMI media with doubling dilutions of ruxolitinib for 48 hours prior to flow cytometric analysis for annexin-V / propidium iodide uptake.

B) E*μ-Myc* lymphoma #4242 and human SET2(V617F) cell lines were exposed to 1uM ruxolitinib or DMSO vehicle control for three hours prior to protein extraction and separation with SDS-PAGE. Immunoblotting was performed to show expression of pSTAT5, STAT5, pJAK2, JAK2, pS6RP and tubulin loading control.


A
B
Copy numer
0
2
4
6
0.0e+00
1.0e+08
Position, Chr3
Normal
ML33
ML353
ML52
Consensus
Non−consensus
0.0e+00
5.0e+07
1.0e+08
1.5e+08

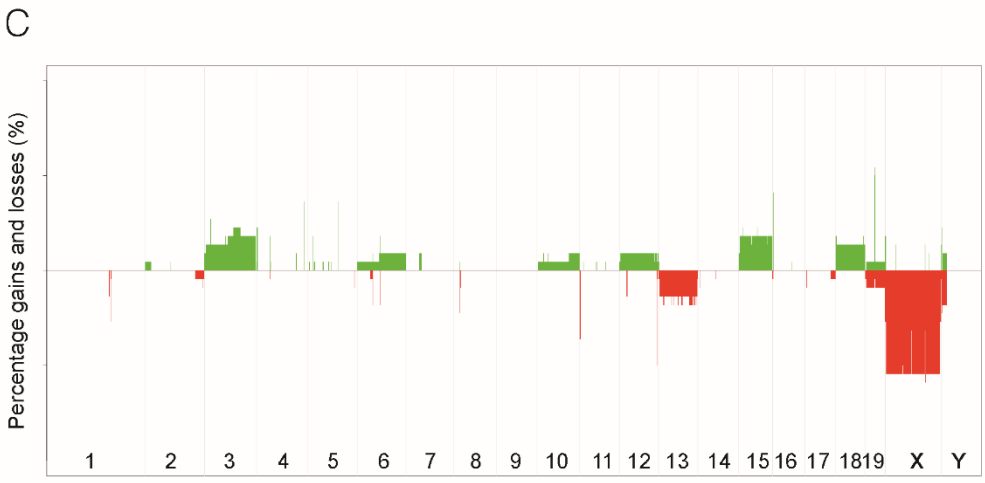

C
Percentage gains and losses (%)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
X
Y

Figure 3.08 – Consensus CNV inferred from WGS and ADTEx

A) Representative Freec plot demonstrating somatic copy number alterations that spanned chr3 in E*μ-Myc* lymphoma ML33. The green colouring represents copy number neutral (copy number of 2) regions and the red region represents a copy number gain.

B) Representative ADTEx plot demonstrating somatic copy number alterations in ML362 inferred from WES spanning chr3 when compared to three whole genome-sequenced controls. In this case, ML362 had the same copy number as ML353 and ML352 but demonstrated proportionately one copy less than ML33 (corresponding to the red line, shown in Figure 8A).

C) Somatic copy-number alterations detected in *Eμ-Myc* lymphomas by combining WES or WGS information across the prospective and retrospective cohorts. Overlapping regions of gain (green) or loss (red) are shown as a percentage of cases (22 cases analysed) and were plotted against genome position (chromosome number).

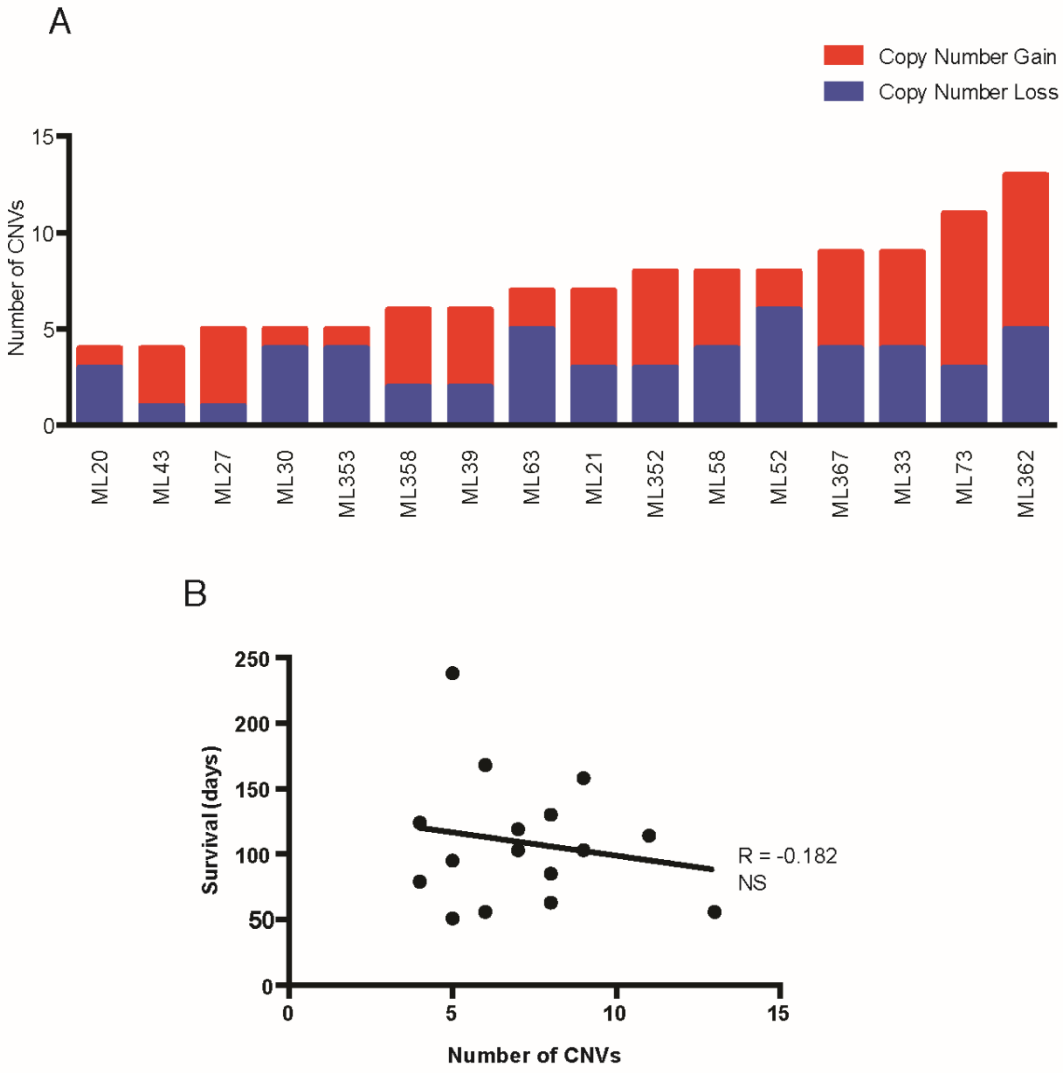
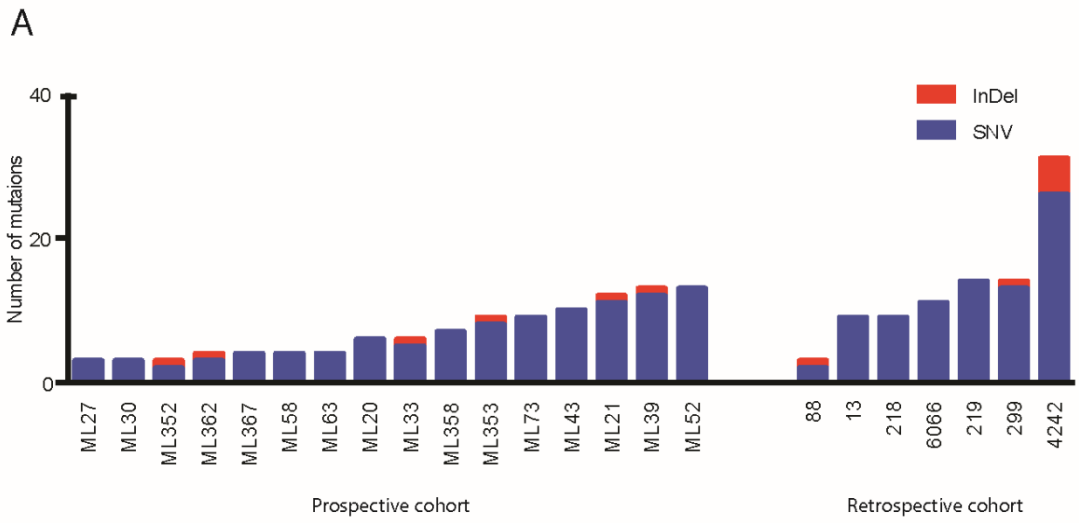

A
Copy Number Gain
Copy Number Loss
Number of CNVs
15
10
5
0
ML20
ML43
ML27
ML30
ML353
ML358
ML39
ML63
ML21
ML352
ML58
ML52
ML367
ML33
ML73
ML362
B
Survival (days)
250
200
150
100
50
0
0
5
10
15
Number of CNVs
R = -0.182
NS

Figure 3.09 – CNV information derived from WES

A) The prospective E*μ-Myc* lymphomas are arranged in ascending order of total number of CNV with copy number gains in red and copy number losses in blue. CNV in each case ranged from a minimum of four to a maximum of 13.

B) A comparison between total number of CNV and survival time was made. Linear regression analysis was performed and there was no correlation between survival and total number of CNV ($R = -0.182$, with the slope not significantly non-zero, $p > 0.05$).


A
40
20
0
Number of mutaions
InDel
SNV
ML27
ML30
ML352
ML362
ML367
ML58
ML63
ML20
ML33
ML358
ML353
ML73
ML43
ML21
ML39
ML52
88
13
218
6066
219
299
4242
Prospective cohort
Retrospective cohort

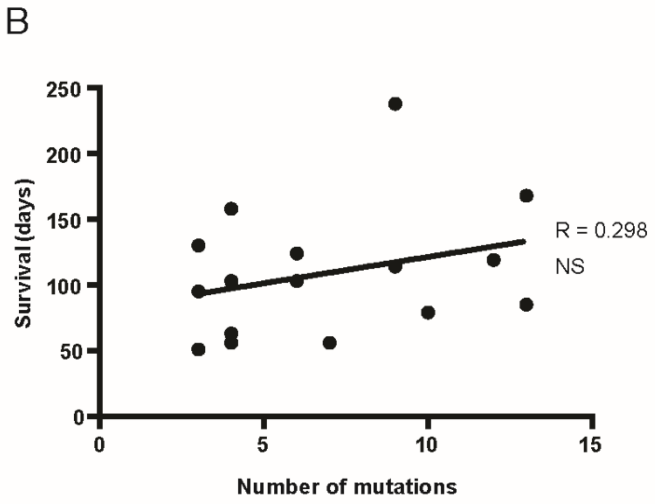

B
250
200
150
100
50
0
Survival (days)
0
5
10
15
Number of mutations
R = 0.298
NS

Figure 3.10 – SNV and InDel information derived from WES

A) Mutation frequency across 23 E*μ-Myc* lymphomas, encompassing both the prospective and retrospective cohorts. For each cohort, lymphomas are arranged in ascending order of total number of SNV ranging from four to 36. InDels are represented by red shading and SNV are represented by blue shading.

B) A comparison between total number of SNV and survival time was made. Linear regression analysis was performed and there was no correlation between survival and total number of SNV ($R = 0.298$, with the slope not significantly non-zero, $p > 0.05$).

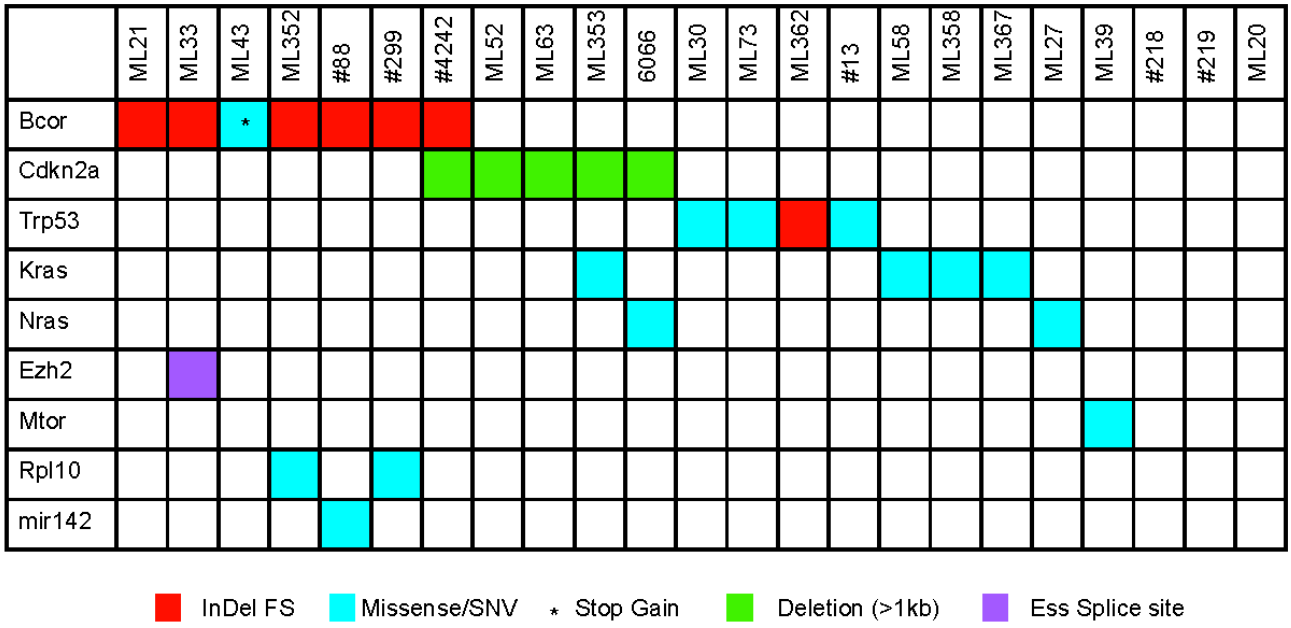

ML21
ML33
ML43
ML352
#88
#299
#4242
ML52
ML63
ML353
6066
ML30
ML73
ML362
#13
ML58
ML358
ML367
ML27
ML39
#218
#219
ML20
Bcor
Cdkn2a
Trp53
Kras
Nras
Ezh2
Mtor
Rpl10
mir142
InDel FS
Missense/SNV
Stop Gain
Deletion (>1kb)
Ess Splice site

Figure 3.11 – Somatic mutations detected by WES in the prospective and retrospective E*μ-Myc* lymphomas

Checkerboard demonstrating putative mutations affecting known cancer genes in E*μ-Myc* lymphomas. Red shading represents frame-shifted InDels, blue shading signifies missense SNV (* denoting that a stop codon was gained), green represents large deletions that could not be detected by whole-exome sequencing and purple represents essential splice site variants.

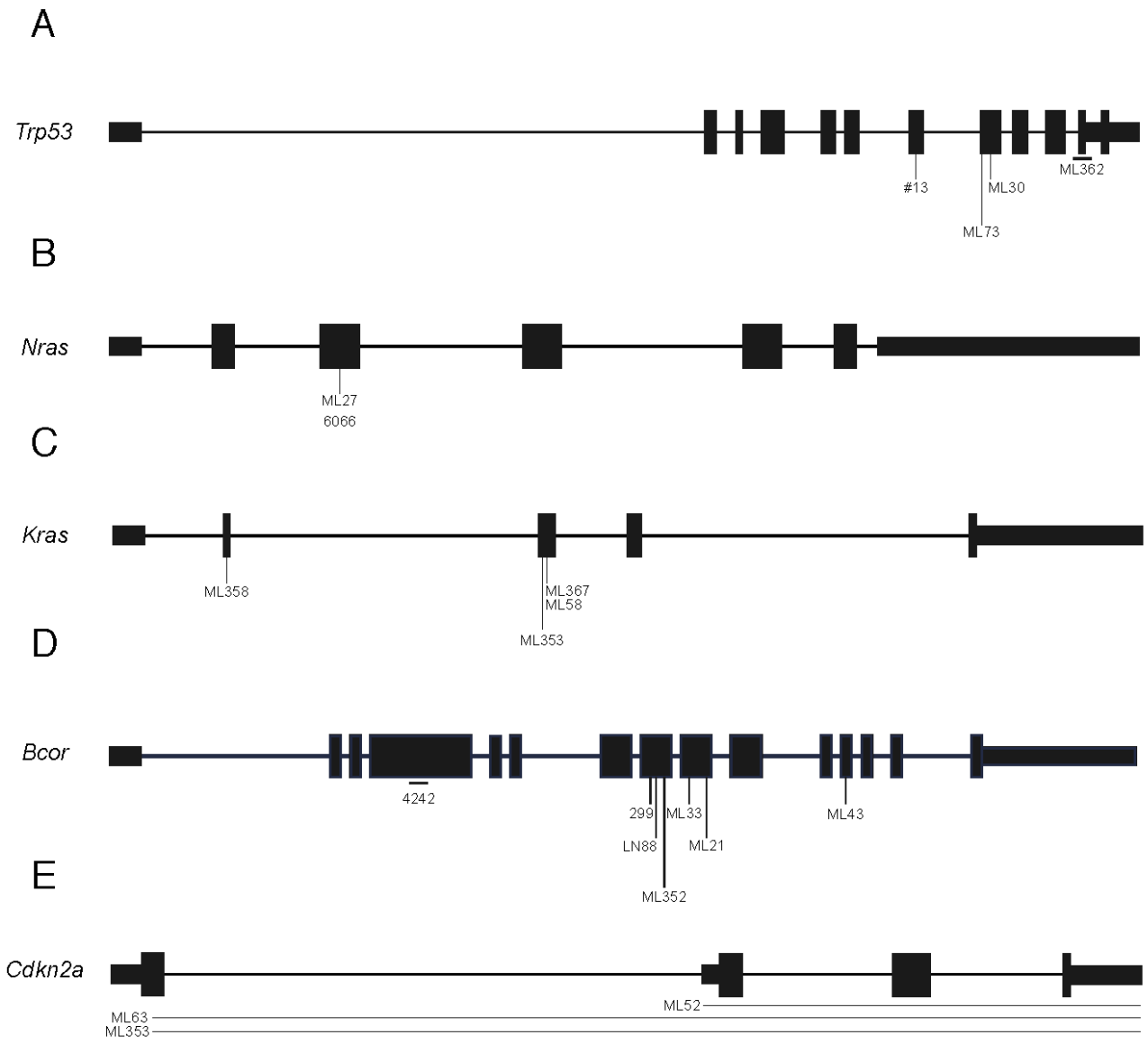
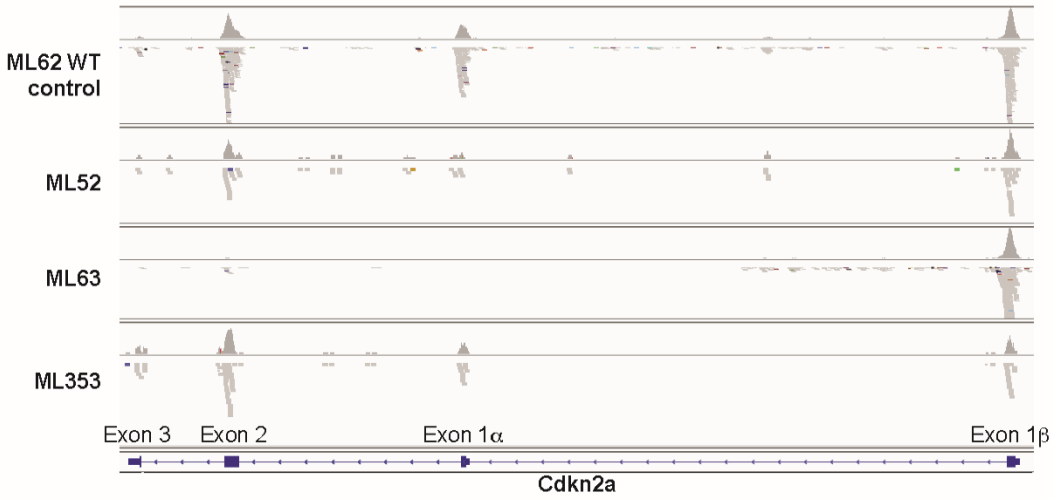

A
Trp53
#13
ML30
ML362
ML73
B
Nras
ML27
6066
C
Kras
ML358
ML367
ML58
ML353
D
Bcor
4242
299
ML33
ML43
LN88
ML21
ML352
E
Cdkn2a
ML52
ML63
ML353

Figure 3.12 – Schematic of the mutations identified in the recurrently mutated candidate driver genes in the E*μ-Myc* lymphoma cohorts

A) A schematic of *Trp53* demonstrating the SNVs identified in lymphomas ML30, ML73 and #13 and the deletion identified in lymphoma ML362.
B) A schematic of *Nras* demonstrating the point mutation identified in lymphomas ML27 and 6066.
C) A schematic of *Kras* demonstrating the mutations identified in lymphomas ML358, ML353, ML367 and ML58.
D) Schematic of *Bcor* demonstrating the deletion identified in lymphoma 4242 and the SNVs identified in lymphomas 299, LN88, ML352, ML33, ML21 and ML43.
E) Schematic of *Cdkn2a* highlighting the deletions found in lymphomas ML63, ML52 and ML353.

A


ML62 WT control
ML52
ML63
ML353
Exon 3
Exon 2
Exon 1α
Exon 1β
Cdkn2a

B

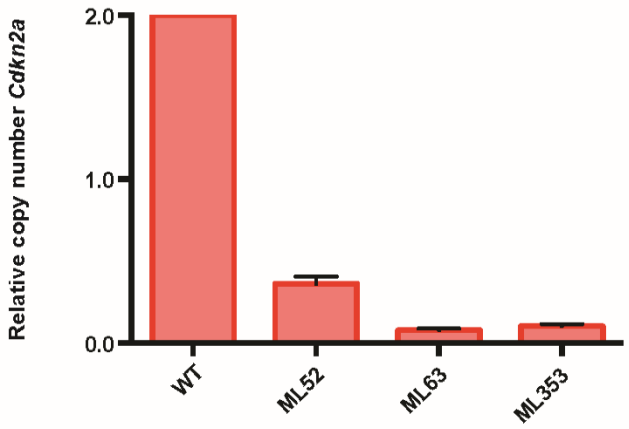

Relative copy number Cdkn2a
2.0
1.0
0.0
WT
ML52
ML63
ML353

Figure 3.13 – Large deletions of *Cdkn2a* as detected by WES and QRT-PCR

A) WES analysis of reads spanning *Cdkn2a* as visualised in IGV. Potential *Cdkn2a* deletion was demonstrated in lymphomas ML52, ML63 and ML535 compared to WT littermate control (ML62).
B) QRT-PCR displaying relative copy number for *Cdkn2a*. *Cdkn2a* deletion was confirmed in ML52, ML63 and ML353 compared to WT littermate control (ML62). Data shown is mean ± SEM in 3 replicates

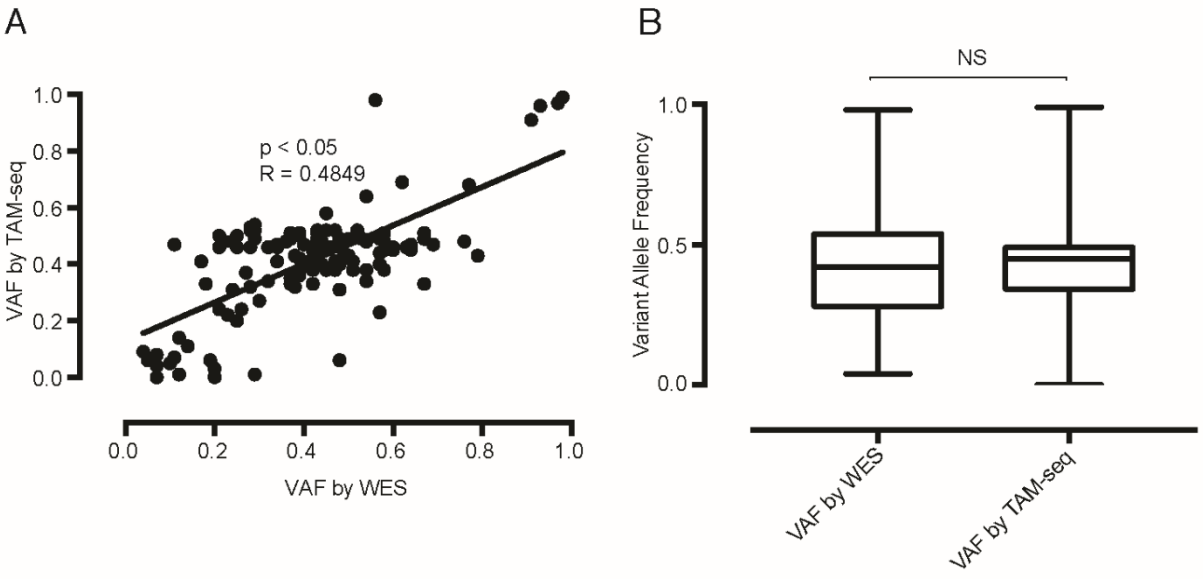

A
1.0
0.8
0.6
0.4
0.2
0.0
VAF by TAM-seq
p < 0.05
R = 0.4849
0.0
0.2
0.4
0.6
0.8
1.0
VAF by WES
B
NS
1.0
0.5
0.0
Variant Allele Frequency
VAF by WES
VAF by TAM-seq

Figure 3.14 – VAF as assessed by WES and TAM-seq

A) VAF for validated mutations by WES (x-axis) plotted against VAF for the same validated mutations as assessed by TAM-seq. (y-axis). Linear regression analysis demonstrated a positive correlation between the VAF estimate for WES data and TAM-seq data (R=0.4849, with the slope significantly non-zero $p<0.05$).

B) Box and whisker plot demonstrating the spread of VAF for both the WES and TAM-seq methodologies. The mean VAF for WES methodology and TAM-seq methodology was not statistically significant at 0.4156 and 0.4115, respectively. However, there was a trend to indicate that the VAF as assessed by TAM-seq demonstrated a narrower inter-quartile range (IQR), with the 25th percentile and 75th percentile being 0.3425 and 0.4900, respectively, compared to 0.2800 and 0.5375, respectively. Data shown is mean ± SEM.

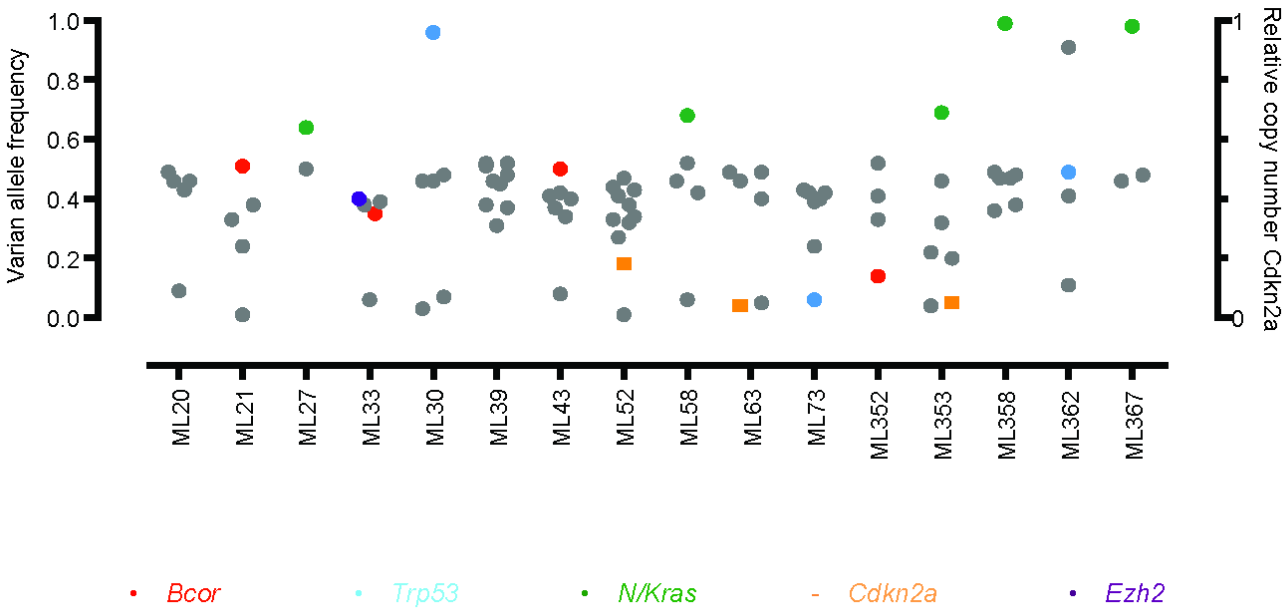

Varian allele frequency
1.0
0.8
0.6
0.4
0.2
0.0
Relative copy number Cdkn2a
1
0
ML20
ML21
ML27
ML33
ML30
ML39
ML43
ML52
ML58
ML63
ML73
ML352
ML353
ML358
ML362
ML367
Bcor
Trp53
N/Kras
Cdkn2a
Ezh2

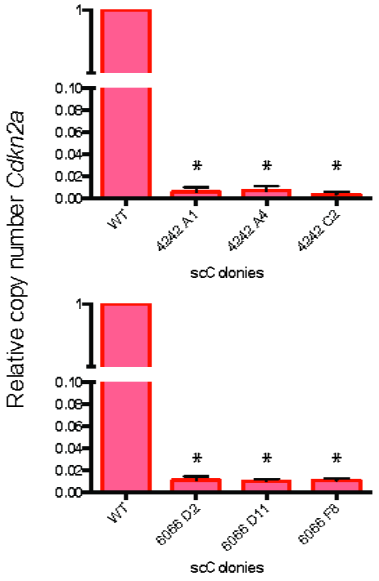
Figure 3.15 – VAF of each validated SNV identified in the prospective E*μ-Myc* cohort

TAM-seq was performed on the prospective E*μ-Myc* tumour cohort to validate mutations from the WES screen. VAF is shown on the left y-axis for the mutations that were validated, with candidate driver mutations highlighted (red = *Bcor*, blue = *Trp53*, green = *Nras* and *Kras* and purple = *Ezh2*). The right y-axis displays relative copy number of *Cdkn2a* in the samples where *Cdkn2a* loss is suspected based on WES read depth and QRT-PCR analysis. Relative copy number of *Cdkn2a* is represented as orange boxes for E*μ-Myc* lymphoma ML52, ML63 and ML353.

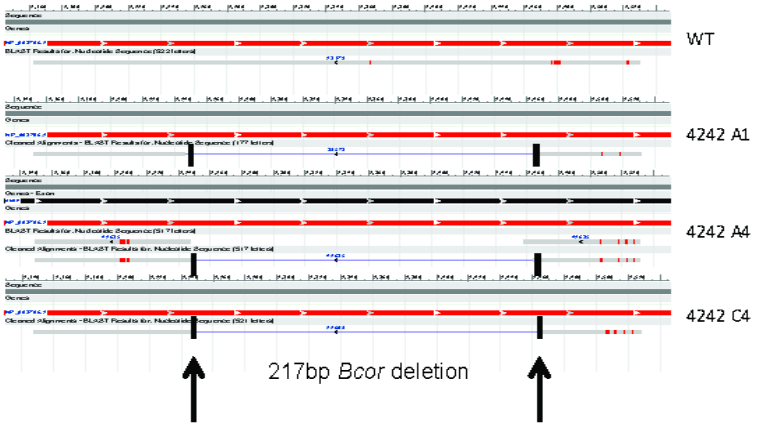
A

Well	Cell line	Barcode
A1	4242	GTCTGAATGAAGTCGA
A4	4242	CTCTGTGCGAGGAATT
C2	4242	CTAGCTGTCTGCTGCG
D2	6066	CAATCAGTGAGGTAGA
D11	6066	CACTGTATCAAAGATA
F8	6066	GTCGCTATGACACATA

B



C



D

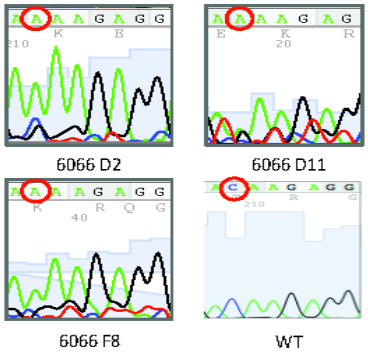
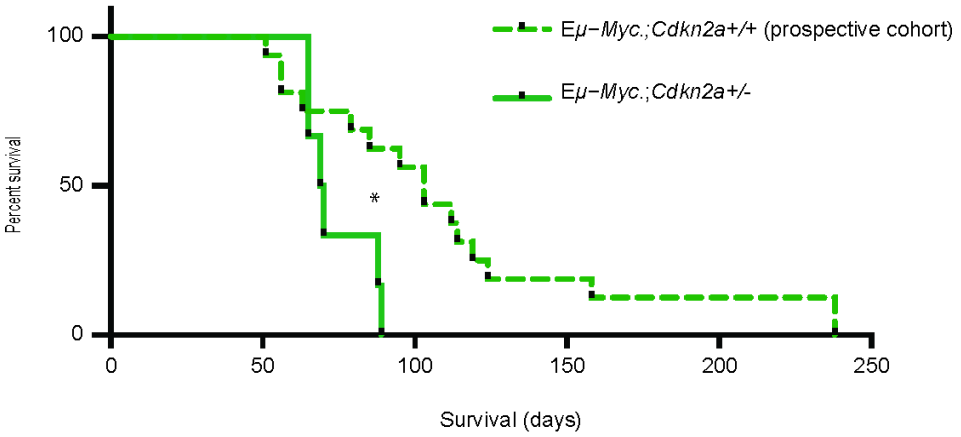


Figure 3.16 – Validation of coalescing driver mutations in retrospective E*μ-Myc* lymphomas 4242 and 6066.

4242 and 6066 cell lines were retrovirally transduced with an inert DNA "barcode" and BFP fluorescent reporter at an MOI of 1. Colonies were grown from a single-cell of 4242 or 6066 origin based on BFP expression.

A) Each single-cell derived colony was Sanger sequenced to show the presence of just one barcode, thus indicating the colony is truly derived from a single cell (clone). This table highlights the barcode corresponding to each single-cell derived colony.
B) QRT-PCR was performed on each single-cell derived colony to demonstrate that *Cdkn2a* was lost compared to WT control. Each QRT-PCR was performed in triplicate, with the mean relative copy number shown ± SEM, * = $p < 0.05$ compared to WT, 1-way ANOVA.
C) Sanger sequencing was performed on the 4242 single-cell derived colonies. A BLAST readout of the sequences is represented and highlights the expected 217bp deletion in *Bcor* apparent in 4242 compared to WT DNA.
D) Sanger sequencing spanning *Nras* confirmed C-to-A mutation corresponding to Q61K point mutation in each of the 6066 single-cell-derived colonies compared to WT DNA.

A


Eμ−Myc.;Cdkn2a+/+ (prospective cohort)
Eμ−Myc.;Cdkn2a+/-
Percent survival
100
50
0
*
0
50
100
150
200
250
Survival (days)

B

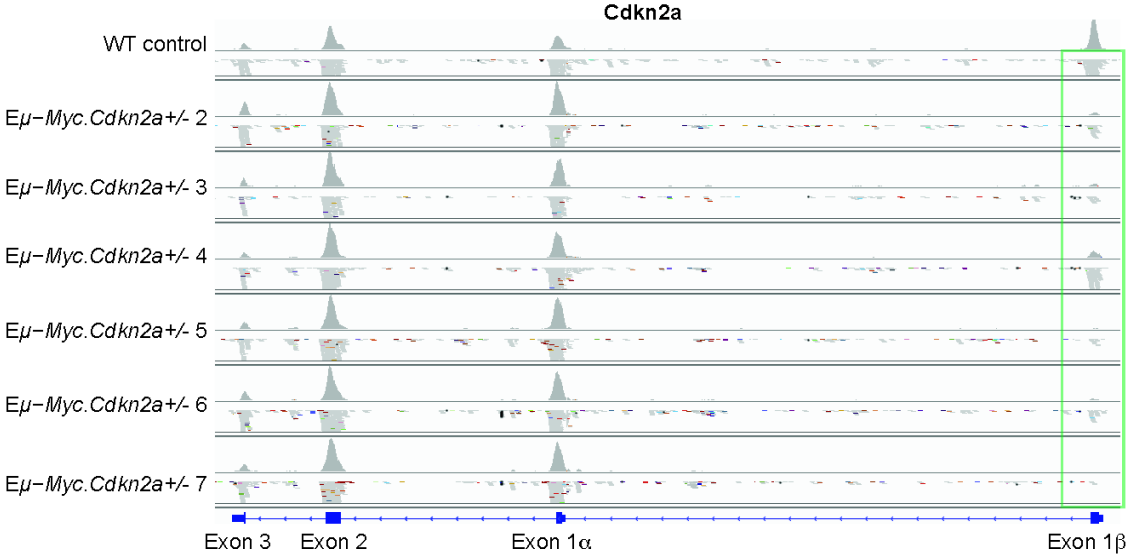

Cdkn2a
WT control
Eμ−Myc.Cdkn2a+/- 2
Eμ−Myc.Cdkn2a+/- 3
Eμ−Myc.Cdkn2a+/- 4
Eμ−Myc.Cdkn2a+/- 5
Eμ−Myc.Cdkn2a+/- 6
Eμ−Myc.Cdkn2a+/- 7
Exon 3
Exon 2
Exon 1α
Exon 1β

Figure 3.17 – Overall survival of E*μ-Myc*.*Cdkn2a*$^{+/-}$ mice

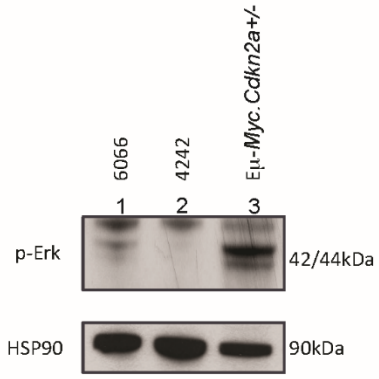
A) E*μ-Myc*.*Cdkn2a*$^{+/-}$ (n=6) transgenic mice were culled at obvious sign of disease progression. Kaplan-Meier curve demonstrating transgenic E*μ-Myc*.*Cdkn2a*$^{+/-}$ mice (solid line) and E*μ-Myc*.*Cdkn2a*$^{+/+}$ (prospective cohort, dashed line), with a median survival time of 69.5 days and 103 days, respectively. * = $p<0.05$ by Log-rank (Mantel-Cox) test

B) WES analysis of reads spanning *Cdkn2a* locus as viewed in IGV. There is clear reduced depth in exon 1β (encoding p19^{ARF}) in the E*μ-Myc*.*Cdkn2a*$^{+/-}$ lymphoma samples, which is indicative of loss-of-heterozygosity.

A

GENE	CHR	POS	REF	ALT	CONSEQUENCE	REF DEPTH	ALT DEPTH	VAF
Kras	6	123124	T	G	Q61H	58	37	.67

B



C

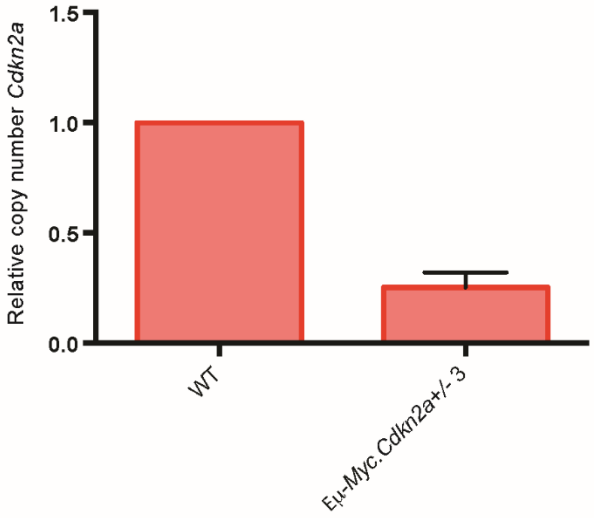
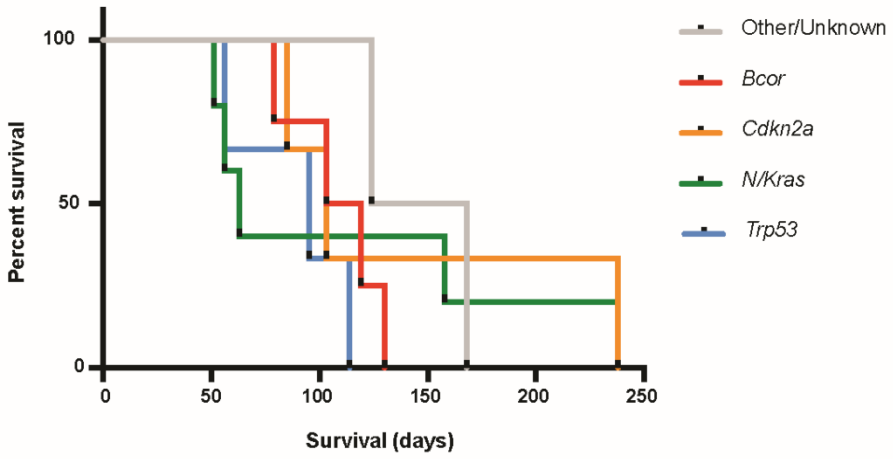


Figure 3.18 – Loss-of-heterozygosity occurred with concomitant activating *Kras* mutation in E*μ-Myc.Cdkn2a*$^{+/-}$ 3

A) WES readout described the somatically acquired activating Q61H *Kras* mutation found in E*u-Myc*.*Cdkn2a*$^{+/-}$ 3.
B) Western blot analysis demonstrated elevated phosphorylated Erk (p-Erk) in *Nras* mutant 6066 and E*u-Myc*.*Cdkn2a*$^{+/-}$ 3 compared to N/Kras WT 4242 lymphoma.
C) QRT-PCR demonstrated relative copy number for *Cdkn2a* confirms loss-of-heterozygosity in E*u-Myc*.*Cdkn2a*$^{+/-}$ 3 compared to WT control.

A


100
50
0
Percent survival
0
50
100
150
200
250
Survival (days)
Other/Unknown
Bcor
Cdkn2a
N/Kras
Trp53

B

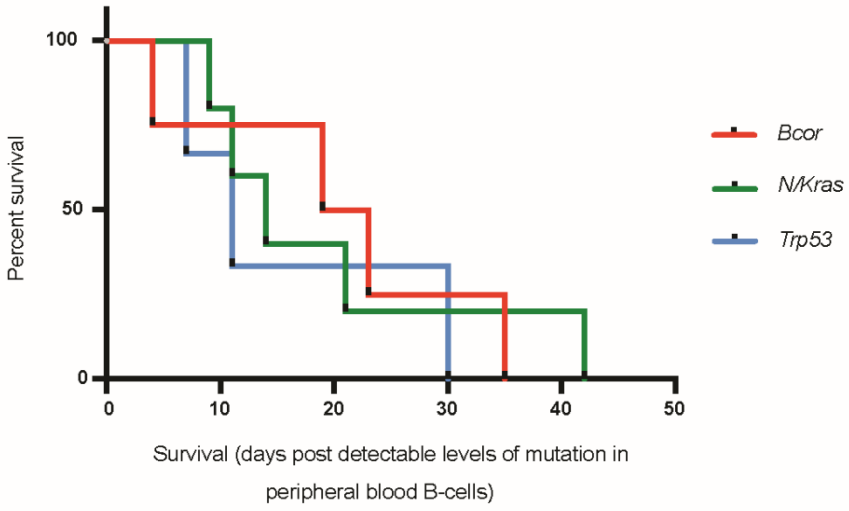

100
50
0
Percent survival
0
10
20
30
40
50
Survival (days post detectable levels of mutation in peripheral blood B-cells)
Bcor
N/Kras
Trp53

Figure 3.19 – Survival of the prospective E*μ-Myc* cohort grouped per candidate driver mutation

A) Kaplan-Meier curve demonstrating overall survival of all prospective E*μ-Myc* mice stratified into groups based on the identified candidate driver mutation in each lymphoma. The other/unknown group (grey) includes samples ML39 (*Mtor*) and ML20 (no known candidate driver mutation) and had a median survival time of 146 days (n=2). The E*μ-Myc* mice that exhibited *Bcor* mutations (red) had a median survival time of 111 days (n=4). E*μ-Myc* mice found to harbour *Cdkn2a* deletions (orange) demonstrated a median survival time of 103 days (n=3). E*μ-Myc* mice harbouring *Nras* or *Kras* mutations displayed a median survival time of 63 days (n=5). E*μ-Myc* mice with mutations in *Trp53* had median survival time of 95 (n=3). The survival curves were not statistically significant, $p > 0.05$, log-rank (Mantel-Cox) test.

B) Kaplan-Meier curve demonstrating survival (days) post-detectable levels of candidate driver in the peripheral blood for animals in the prospective E*μ-Myc* cohort. E*μ-Myc* mice with no known candidate driver mutations were not included in this analysis, along with *Cdkn2a*-deleted lymphomas. Once mutations in *Bcor* (red) were detected in the peripheral blood of E*μ-Myc* mice, the median survival time was 21 days (n=4). When an *Nras* or *Kras* mutation was detected in circulating peripheral blood B cells, the median survival time for the animal was 14 days (n=5). Once *Trp53* mutations were detected in the peripheral blood of the prospective E*μ-Myc* mice, the median survival time was 11 days (n=3). The survival post-detectable levels of the driver mutation in the peripheral blood were not statistically significant between the groups, $p > 0.05$, log-rank (Mantel-Cox) test.

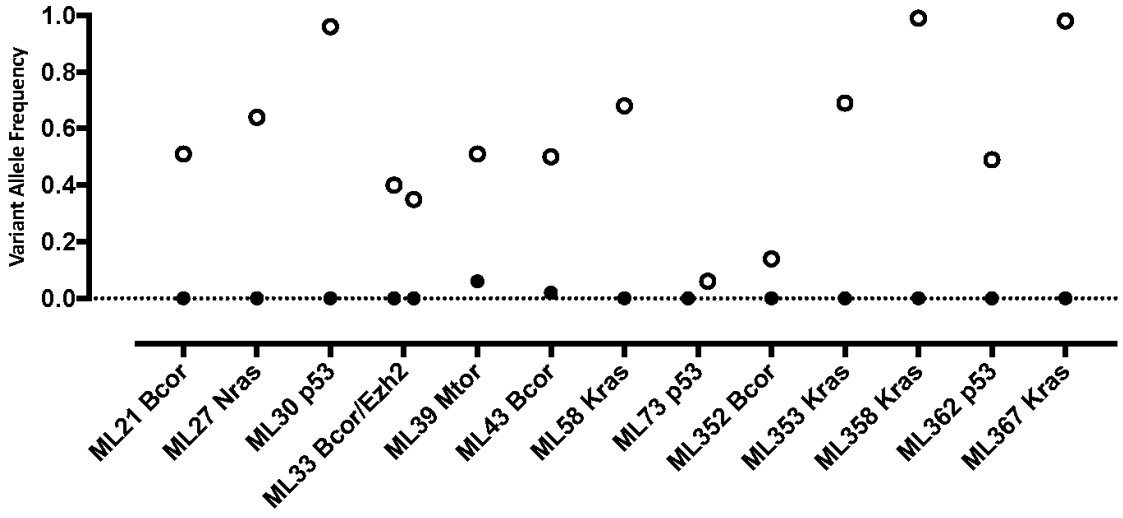

Variant Allele Frequency
1.0
0.8
0.6
0.4
0.2
0.0
ML21 Bcor
ML27 Nras
ML30 p53
ML33 Bcor/Ezh2
ML39 Mtor
ML43 Bcor
ML58 Kras
ML73 p53
ML352 Bcor
ML353 Kras
ML358 Kras
ML362 p53
ML367 Kras

Figure 3.20 – TAM-seq demonstrated recurrent driver mutations are somatically acquired

Targeted amplicon sequencing was performed on DNA extracted from peripheral blood at 4-weeks of age (filled circles) and at end-stage (empty-circles) to validate somatic mutations in E*μ-Myc* lymphomas. Candidate driver mutations were either undetectable or present at very low frequencies (<0.1) at 4-weeks of age and in each case were found at increased levels at end-point.

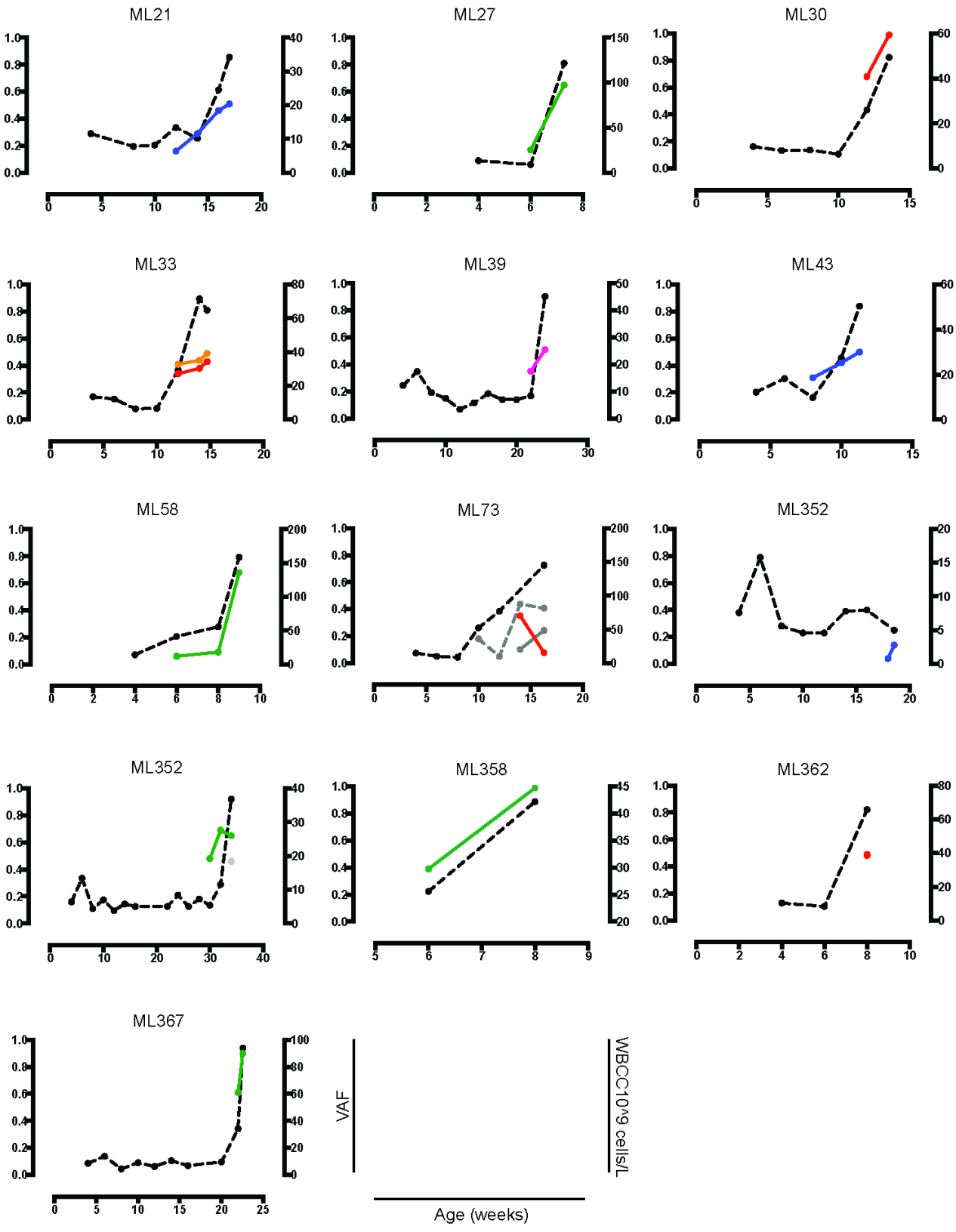

ML21
ML27
ML30
ML33
ML39
ML43
ML58
ML73
ML352
ML352
ML358
ML362
ML367
VAF
WBCC10^9 cells/L
Age (weeks)
Bcor
Trp53
N/Kras
Ezh2
Mtor
Cntn5
Adarb2
WBCC

Figure 3.21 – Temporal analysis of allelic burden of putative driver mutations in the prospective *Eμ-Myc* cohort.

Fortnightly, the prospective E*μ-Myc* mice were bled and their circulating B cells were isolated and used for retrospective TAM-seq. Mutations in the identified candidate driver genes in each lymphoma were sequenced in the circulating peripheral blood B cells of each corresponding mouse, allowing for a temporal analysis of allelic burden of the candidate driver mutation. Age (weeks) is plotted against VAF for specific mutations in *Bcor* (blue), *Trp53* (red), *Nras* (green), *Kras* (green), *Ezh2* (orange), *Mtor* (violet), *Cntn5* (grey) and *Adarb2* (grey). The variants are only displayed when they were detected by TAM-seq. For each sample WBC counts (x10^9cells/L) are represented by a dashed black line and is plotted against the right-y-axis. WBC counts are shown at all time-points. Tumours that harboured only a *Cdkn2a* (ML62, ML63) mutation or had no candidate driver mutation (ML20) are not shown.

4 Biological validation of *Bcor* as a tumour suppressor gene in E*μ-Myc* lymphoma

4.1 Introduction

In chapter 3 somatic loss-of-function mutations in the Bcl6-co-repressor gene (*Bcor*) were identified at high frequency in murine E*μ-Myc* lymphomas. *Bcor* is located on chromosome X and is highly conserved between mice and humans. In mice, the predominant "long" isoform encodes a 1759 amino acid protein. BCOR is expressed widely across many tissue types and is crucial for early embryonic development (Wamstad and Bardwell, 2007). X-linked dominant inheritance of BCOR mutations in humans causes the rare familial condition known as oculofaciocardiodental (OFCD) disorder (Ng et al., 2004). Truncating BCOR loss of function mutations are embryonic lethal in males, and cause severe developmental defects in females such as dental abnormalities, congenital cataracts, craniofacial defects and congenital heart disease (Cox et al., 2010; Ng et al., 2004).

SNV, frameshift InDel mutations and larger structural changes in *BCOR* have been identified in both solid and haematological malignancies with a reported frequency of up to 15% in any individual cancer type (COSMIC v67) (Table 4-1). The majority of mutations in *BCOR* induce premature protein truncations, which is a characteristic of tumour suppressor genes (Vogelstein et al., 2013; Zender et al., 2008). In haematological malignancies, *BCOR* mutations appear to be disease or subtype defined. *BCOR* is mutated in acute myeloid leukaemia (AML), myelodysplastic syndrome (MDS) and chronic myelomonocytic leukaemia (CMML) (Damm et al., 2013; Grossmann et al., 2011). In AML, *BCOR* mutations occur predominantly within cytogenetically normal cases (CN-AML), with a mutation frequency of 4% in the CN-

AML subtype (Grossmann et al., 2011). *BCOR* mutations in AML are associated with poorer prognosis (Damm et al., 2013). In B cell lineage leukaemias and lymphomas, the occurrence of *BCOR* mutations is likely dependent on disease type or stage of B cell differentiation within the malignancy. *BCOR* mutations have not been observed in post germinal centre B cell malignancies, such as diffuse large B cell lymphoma (DLBCL) and Burkitt's lymphoma (BL) (Lohr et al., 2012; Love et al., 2012a) but have been found in CLL (Quesada et al., 2011) and B cell acute lymphoblastic leukaemia (B-ALL) at a frequency of 1-2% (Papaemmanuil et al., 2014). CLL is characterized as either pre- or post-germinal centre depending on the presence of somatic hypermutation in the *IGV* locus. B-ALL is a pre-GC malignancy where the pathogenic cells resemble early B cell progenitors of the Pro/Pre-B stage of development, which is phenotypically mimicked in E*μ-Myc* lymphoma. All mutations identified in the E*μ-Myc* lymphomas profiled in chapter 3 were frameshift InDels resulting in premature protein truncation or were missense incorporating an early stop-codon. This suggests somatic *BCOR* mutations in human and murine cancers results in the loss of function through premature protein truncation and supports the hypothesis that *BCOR* is a putative tumour suppressor gene.

Table 4-1 Frequency of somatic *BCOR* mutations found in human malignancies

Cancer Site	Samples	Mutations	% Of samples with *Bcor* mutation
Breast	990	8	0.81%
Central nervous system	575	9	1.57%
Cervix	14	1	7.14%
Endometrium	308	48	15.58%
Eye	80	5	6.25%
Haematopoietic and lymphoid	1592	35	2.20%
Kidney	664	1	0.15%
Large intestine	**616**	38	6.17%
Liver	337	2	0.59%
Lung	874	32	3.66%
Oesophagus	173	2	1.16%
Ovary	**502**	3	0.60%
Pancreas	366	1	0.27%
Prostate	385	1	0.26%
Skin	328	6	1.83%
Upper aerodigestve tract	169	1	0.59%
Urinary tract	105	3	2.86%

BCL6 is a potent oncogene that is frequently ectopically expressed through translocation events in DLBCL (Ye et al., 1995). Normally, BCL6 functions to dampen the DNA damage response to somatic hypermutation in germinal centre B cells (Cattoretti et al., 1995); however, as BCL6 represses potent tumour suppressor genes such as *TP53* and *CDKN2A,* overexpression or oncogenic deregulation of BCL6 can contribute to oncogenesis in post-GC B cell lymphomas (Phan and Dalla-Favera, 2004). BCOR can interact with, and potentiate BCL6-mediated repression in GC B cells (Ghetu et al., 2008; Huynh et al., 2000). Furthermore, disruption of the BCL6 BTB/POZ (BCOR) binding domain results in prevention of GC formation and the killing of DLBCL cells suggesting that *BCOR* could be acting as an oncogene in cells where BCL6 is expressed (Polo et al., 2004). *Bcor*-inactivating mutations described in chapter 3, in the context of *Eμ-Myc* lymphomagenesis, challenges the suggestion that Bcor acts as an oncoprotein. E*μ-Myc* lymphomas with a short latency (<300 days), such as the lymphomas sequenced in chapter 3, demonstrate a gene expression profile indicative of pre-GC B cells and lack expression of *Bcl6* (Mori et al., 2008). Additionally, BCOR is ubiquitously expressed in tissues whereas BCL6 expression is much more tissue-specific indicating that BCOR may perform additional functions independent of BCL6 in cells lacking BCL6 (Huynh et al., 2000). A question that arises from the studies described in chapter 3 is does mutation of *Bcor* affect Bcl6 to cooperate with Myc in driving lymphomagenesis or is Bcor functioning in a Bcl6-independent manner in the context of E*μ-Myc* lymphomas? It is currently unknown if BCOR and BCL6 cooperate during early lymphocyte development. BCL6 is expressed in some early pre-B cell types, and is required to repress the DNA damage response due to immunoglobulin (Ig)

light chain rearrangement (Duy et al., 2010). If Bcor acts through its canonical pathway as a Bcl6 co-repressor in E*μ-Myc* progenitor B cells then it is possible that truncation and loss of function of Bcor inhibits B cell development and GC transition by blocking the molecular processes required for assembly of a functional B cell receptor. Conversely, it is apparent that the non-canonical BCOR complex, the BCOR-PRC1-like complex, is functional within cell types that do not express BCL6 (Gearhart et al., 2006). Additionally, BCOR is mutated in many cancer types that do not express BCL6 (COSMIC v67) (Table 4-1). Taken together, this indicates that BCOR has BCL6-independent functions and binding partners, which if disrupted can contribute to malignant progression through a currently unknown mechanism.

The data described in chapter 3 underpins the hypothesis that inactivation of *Bcor* cooperates with oncogenic Myc expression to drive lymphomagenesis. Data presented in this chapter functionally validates *Bcor* as a tumour suppressor gene in Myc-driven lymphoma. Subsequent analyses demonstrate that lymphomas driven by Bcor-loss are of an immature B cell phenotype and present with a unique gene expression profile that is distinct to E*μ-Myc* lymphomas that have lost *Trp53* or harbour an oncogenic *Nras* mutation. Taken together, this suggests that the *Bcor* tumour suppressor mechanism acts independently of the pathways known to be deregulated in E*μ-Myc* lymphomagenesis.

4.2 Results

4.2.1 *Bcor*-targeting short-hairpin RNA design and selection

Chapter 3 describes *Bcor* as the most recurrently mutated gene in E*μ-Myc* lymphoma, affecting 7/23 sequenced cases. All *Bcor* mutations identified were predicted loss-of-function mutations, defining *Bcor* as a tumour suppressor gene in E*μ-Myc* lymphoma (Vogelstein et al., 2013). The loss-of-function mutations described in chapter 3 appear to occur randomly throughout the coding region of the gene, with a lack of any clear mutational 'hotspot' (Figure 4-1A). To mimic the loss of function mutations described in chapter 3 and to test the hypothesis that *Bcor* acts as a tumour suppressor gene in E*μ-Myc* lymphoma, nine short-hairpin RNA (shRNA) targeting *Bcor* for genetic depletion were designed (Figure 4-1A). shRNA were designed spanning *Bcor* from exon 4 to the 3' untranslated region (3'-UTR) (Figure 4-1A). *Bcor* mutations were found to occur in exon 4, 8, 9, and 12 (Figure 4-1A). To assess the effectiveness of each shRNA in knocking down *Bcor* expression, QRT-PCR on 3T3 transfected with each shRNA was performed (Figure 4-1B). The shRNA capable of mediating the greatest decrease in *Bcor* mRNA was shRNA9 with a ratio of 0.34-to-1 when compared to the mock-transfected control cells (Figure 4-1B). Prior to measuring levels of knockdown, the cells were not sorted based on GFP expression, which would imply that knockdown efficiency could have been underestimated using this method. shRNA9 demonstrated the greatest level of *Bcor* knockdown and as such was cloned into the pLMS-GFP knockdown vector (pLMS-*Bcor*.sh9) and was then utilised for subsequent knockdown experiments, along with a non-targeting shRNA (pLMS-Scram) as a negative control of

gene knockdown and a *Trp53*-targeting shRNA (pLMS-*Trp53*.1224) as a positive control of tumour suppressor gene knockdown.

4.2.2 Transfecting packaging cell lines

To generate retrovirus containing pLMS-*Trp53*.1224, pLMS-*Bcor*.sh9 and pLMS-Scram, phoenix-E packaging cells were transfected with each vector and the transfection efficiency assessed by way of measuring GFP positivity. Each vector was successfully transfected into Phoenix-E cells (Figure 4-2A-D**)**. Phoenix-E cells transfected with the pLMS-*Bcor*.sh9 construct demonstrated a 35.5% GFP positive population from at least 10,000 analysed cells (Figure 4-2B). Phoenix-E cells transfected with the pLMS-Scram construct demonstrated a 39.8% GFP positive population from at least 10,000 cells analysed (Figure 4-2C). Transfecting the Phoenix-E cells with the pLMS-*Trp53*.1224 construct demonstrated a 47.8% positive population (Figure 4-2D). Retroviral supernatant produced by the Phoenix-E cells was harvested for transduction of E*μ-Myc* foetal liver cells.

4.2.3 Transducing E*μ-Myc* foetal liver cells

To generate haematopoietic stem cells with mock-knockdown (Scram) or knockdown of *Bcor* or *Tp53*, E*μ-Myc* foetal liver cells were cultured in the presence of the retroviral supernatant containing either pLMS-*Bcor*.sh9, pLMS-*Trp53*.1224 or pLMS-Scram to create E*μ-Myc*-*Bcor*.sh9, E*μ-Myc*-*Trp53.*1224 and E*μ-Myc*-Scram foetal liver cells. GFP positivity of the cultured E*μ-Myc* foetal liver cells was used as an indicator of

transduction efficiency for each cell population (Figure 4-2E-H). Transducing E*μ-Myc* foetal liver cells with the retrovirus expressing the pLMS-*Bcor*.sh9 construct led to an 84.2% GFP positive population among at least 10,000 analysed cells (Figure 4-2F). Transducing E*μ-Myc* foetal liver cells with the pLMS-Scram construct resulted in a transduction efficiency of 79.9% by GFP positivity (Figure 4-2G). Transducing E*μ-Myc* foetal liver cells with the pLMS-*Trp53*.1224 construct was most efficient under the conditions described, leading to a viable GFP positive population of 92.7% from at least 10,000 analysed cells (Figure 4-2H). Quantifying the transduction efficiency of pLMS-*Bcor*.sh9, pLMS-*Trp53*.1224 or pLMS-Scram into the E*μ-Myc* cells allowed for accurate transplantation of 1×10^6 cells/μL into lethally irradiated recipient mice. The transduced cells generated in this section are referred to as E*μ-Myc-Bcor*.sh9, E*μ-Myc-Trp53.*1224 and E*μ-Myc*-Scram foetal liver cells herein.

4.2.4 Transplanting mice with E*μ-Myc-Bcor*.sh9, E*μ-Myc-Trp53*.1224 or E*μ-Myc*-Scram foetal liver cells

4.2.4.1 Temporal analysis of GFP positive circulating B cells in mice transplanted with Eμ-Myc-Bcor.sh9, Eμ-Myc-Trp53.1224 or Eμ-Myc-Scram foetal liver cells

Mice reconstituted with E*μ-Myc* foetal liver cells or E*μ-Myc* foetal liver cells transduced with an empty or non-targeting have been shown to demonstrate disease characteristics similar to transgenic *Eμ-Myc* in terms of heterogeneous disease onset (Kelly et al., 2007; Schmitt et al., 2002). Given the nature of the *Bcor* mutations identified in the

sequenced E*μ-Myc* lymphomas (defined in chapter 3), it was hypothesised that mice transplanted with E*μ-Myc-Bcor*.sh9 foetal liver cells would demonstrate accelerated onset of disease compared to mice reconstituted with E*μ-Myc*-Scram foetal liver cells. *Trp53* knockdown in E*μ-Myc* foetal liver cells was shown to accelerate lymphomagenesis in recipient mice (Schmitt et al., 2002) and hence was used as a positive control in these experiments. At five weeks post-transplant, mice that received E*μ-Myc*-Scram foetal liver cells (n=13) demonstrated GFP positivity in 13.92% (± 4.5 SEM, n=13) of the circulating peripheral blood B cells (Figure 4-3A). The percentage of GFP positive cells remained stable compared to baseline (week five) for the duration of the experiment in this cohort with GFP positive cells representing of 13.35% (± 3.2 SEM, n=13) of the circulating B cell population at week six, 13.8% (± 4.1 SEM, n=13) at week seven, 15.8% (± 4.8 SEM, n=13) by week eight and 17.8% (± 6.4 SEM, n=13) at week nine (Figure 4-3A).

In mice that were reconstituted with E*μ-Myc-Trp53*.1224 foetal liver cells (n=12) 20.00% (± 4.3 SEM, n=12) of the viable circulating peripheral blood B cells were GFP positive at five weeks post-transplant, which was not significantly different to that of mice reconstituted with E*μ-Myc*-scram foetal liver cells at the same time point ($p > 0.05$, 2-way ANOVA) (Figure 4-3A). By week six, mice reconstituted with E*μ-Myc-Trp53*.1224 foetal liver cells demonstrated that 55.60% (± 5.7 SEM, n=7) of the circulating peripheral blood B cells were GFP positive, which was significantly greater than that of the same cohorts baseline measurement (week five) ($p < 0.05$, 2-way ANOVA) and that of mice reconstituted with E*μ-Myc*-scram foetal liver cells at the same time point (week six) ($p <$

0.05, 2-way ANOVA) (Figure 4-3A). This trend continued, with 51.80% (± 51.8 SEM, n=2) of circulating peripheral blood B cells being GFP positive in mice reconstituted with E*μ-Myc-Trp53*.1224 foetal liver cells by week six, which was significantly greater than the baseline levels for the same group ($p < 0.05$ 2-way ANOVA) and the negative control at the same time point ($p < 0.05$ 2-way ANOVA) (Figure 4-3A).

At week five, 13.4% (± 2.4 SEM, n=13) of circulating peripheral blood B cells were GFP positive in mice transplanted with E*μ-Myc-Bcor*.sh9 foetal liver cells, which was not significantly different to mice reconstituted with E*μ-Myc*-Scram foetal liver cells ($p > 0.05$, 2-way ANOVA) (Figure 4-3A). By week seven, 32.0% (± 4.3 SEM, n=13) of the circulating peripheral blood B cells were GFP positive in mice transplanted with E*μ-Myc-Bcor*.sh9 foetal liver cells displayed, which was significantly greater than the negative control at week seven ($p < 0.05$, 2-way ANOVA) but not compared to baseline levels ($p > 0.05$, 2-way ANOVA) (Figure 4-3A). GFP positive B cells increased to represent 37.9% (± 10.0 SEM, n=11) of the circulating B cell population in mice reconstituted with E*μ-Myc-Bcor*.sh9 foetal liver cells at week eight, which was significantly greater than both that of baseline levels ($p < 0.05$, 2-way ANOVA) and the negative control at the same time point ($p < 0.05$, 2-way ANOVA) (Figure 4-3A). These results indicate that, as expected, E*μ-Myc-Trp53*.1224 foetal liver cells had a growth advantage, as the GFP-positive population was shown to expand over time. This was also paralleled in the E*μ-Myc-Bcor*.sh9 foetal liver cells, where the circulating GFP positive population increased over time.

4.2.4.2 Temporal analysis of peripheral WBC counts in mice transplanted with Eμ-Myc-Bcor.sh9, Eμ-Myc-Trp53.1224 or Eμ-Myc-Scram foetal liver cells

In addition to investigating the percentage of GFP positive cells in the peripheral blood, circulating blood white blood cell counts were also utilised as a marker of disease burden in recipient mice transplanted with transduced E*μ-Myc* foetal liver cells. Clinically and experimentally, a high WBC count in the periphery is a hallmark of lymphoma and leukaemia (Daley et al., 1990; Harris et al., 1988; Lenz et al., 2008; Muller-Hermelink et al., 2001; Shaffer Iii et al., 2011). Increased WBC count is due to increasing disease burden in the tumour niche such as lymphoid tissue and/or bone marrow, forcing cells into the peripheral blood stream (Lenz and Staudt, 2010; Muller-Hermelink et al., 2001). Normal (pre-leukaemic) counts in C57BL/6 mice have been defined between 3.90×10^9 cells/L to 14×10^9 cells/L with a mean WBC count of $8.5\times10\times^9$ cells/L (Charles River Laboratories, 2012). Any WBC count greater than 14×10^9 cells/L was judged to be indicative of leukaemia in this study. Both cohorts of mice that received E*μ-Myc*-scram (n=12) or E*μ-Myc*-Bcor.sh9 (n=12) foetal liver cells demonstrated an average total WBC count at five weeks post-transplant of $4.2\text{x}10^9$ cells/L (± 0.3 SEM) and 4.6×10^9 cells/L (± 5.8 SEM), respectively, correlating with each cohort being, on average, pre-leukaemic ($\leq14\times10^9$ cells/L). This was significantly less than that of mice transplanted with E*μ-Myc*-p53.1224 foetal liver cells (n=12) that had an average white blood cell count of 23.9×10^9 cells/L ($p < 0.05$, 2-way ANOVA) at the same time point (Figure 4-3B). Similarly, by week six mice transplanted with E*μ-Myc*-Scram foetal liver cells (5.0×10^9 cells/L, n=12) or E*μ-Myc-Bcor*.sh9 foetal liver cells (6.3×10^9 cells/L, n=12) demonstrated significantly less tumour burden compared to mice

that received E*μ-Myc*-p53.1224 foetal liver cells (23.9×10^9 cells/L, n=7) ($p < 0.05$, 2-way ANOVA) (Figure 4-3B). At eight weeks post-reconstitution, mice that received E*μ-Myc*-Bcor.sh9 foetal liver cells displayed significantly increased tumour burden compared to their baseline tumour burden at five weeks post-transplant with an average white blood cell count of 20.8×10^9 cells/L (± 10.8 SEM, n=11) ($p < 0.05$, 2-way ANOVA) (Figure 4-3B). The cohort of mice that had received E*μ-Myc*-Scram foetal liver cells did not reach a point in the experiment where their disease burden significantly exceeded the baseline measurement at five weeks post-transplant ($p > 0.05$, 2-way ANOVA), albeit with one mouse clearly becoming leukaemic over this time (Figure 4-3B). These data indicate that mice transplanted with E*μ-Myc-Bcor.*sh9 or E*μ-Myc-Trp53* foetal liver cells developed signs of leukaemia at an accelerated rate compared to mice that received E*μ-Myc*-Scram foetal liver cells.

4.2.4.3 Temporal analysis of donor cell populations in mice transplanted with Eμ-Myc-Bcor.sh9, Eμ-Myc-Trp53.1224 or Eμ-Myc-Scram foetal liver cells

E*μ-Myc* transgenic mice were bred on a C57BL/6 background such that they expressed the pan-leukocyte marker CD45.2, whereas the transplant recipient mice used in this study were the syngeneic *Ptprca*^{a} breed of mice expressing CD45.1. This system allowed for injection of unsorted (GFP positive and GFP negative) E*μ-Myc* foetal liver cells (CD45.2 positive) into mice that were CD45.1 positive, enabling tracking of the donor cell compartment in the hosts. At weekly intervals, commencing at five weeks post-transplant, the recipient mice were bled, the peripheral blood B cells were isolated and the cells were analysed for CD45.2 positivity corresponding to the donor

compartment. At five weeks post-transplant, mice that had received either E*μ-Myc*-scram, E*μ-Myc*-*Trp53*.1224 or E*μ-Myc*-Bcor.sh9 foetal liver cells displayed large populations of CD45.2 positive donor B cells in the peripheral blood (Figure 4-3C). Mice that received E*μ-Myc*-scram foetal liver cells had a CD45.2 positive donor population of 52.3% (± 4.8 SEM, n=13) of circulating B cells five weeks post-transplant (Figure 4-3C). Mice that received E*μ-Myc*-*Trp53*.1224 foetal liver cells had a significantly larger CD45.2 positive donor cell population compared to the E*μ-Myc*-Scram cohort with 81.9% (± 2.0 SEM, n=12) of circulating, viable single cells being donor-derived at week five ($p < 0.05$, 2-way ANOVA) (Figure 4-3C). Mice that received E*μ-Myc-Bcor*.sh9 foetal liver cells exhibited a CD45.2 positive donor population of 67.5% (± 6.3 SEM, n=13) of circulating B cells five weeks post-transplant (Figure 4-3C). For the remainder of the experiment, no cohorts of recipient mice demonstrated significantly increased CD45.2 positive populations compared to either the E*μ-Myc*-Scram control group or the baseline (week five) reading for each group ($p > 0.05$, 2-way ANOVA for both inter-time point analysis and intra-time point analysis) (Figure 4-3C). The E*μ-Myc*-Scram cohort of recipient mice averaged demonstrated an average donor cell population of 57.6% (±1.7 SEM) across the duration of the experiment with a range of 52.3% (± 4.8 SEM) at five weeks-post transplant to 62.3% (± 4.7 SEM) at nine weeks post-transplant (Figure 4-3C). Mice that received E*μ-Myc*-*Trp53*.1224 foetal liver cells had an average CD45.2 expression in circulating B cells of 78.1% (± 1.9 SEM) and a range of 75.8% (± 9.0 SEM) seven weeks post-transplant to 81.9% (± 2.0 SEM) five weeks post-transplant over the period for which the recipient mice were monitored (Figure 4-3C). Similarly, mice transplanted with E*μ-Myc-Bcor*.sh9 foetal liver cells demonstrated minimal

fluctuation in the average expression of CD45.2 in circulating B cells for the period in which the animals were observed, with an average donor population representing 68.4% (± 1.4 SEM) of the circulating B cells and range of 63.8% (± 6.4 SEM) six weeks post-transplant to 71.5% (± 9.7 SEM) nine weeks post-transplant (Figure 4-3C). These results indicated that the total donor cell compartment did not substantially change in the recipient mice. However, when taken with the data showcasing the expanded GFP positive population in the mice that received E*μ-Myc-Trp53*.1224 or E*μ-Myc-Bcor*.sh9, the results indicated that ratio of transduced cells to non-transduced cells within the donor compartment increased. Importantly, GFP positive cells were always CD45.2 positive (Figure 4-3D).

4.2.4.4 Temporal analysis of surface IgM and IgD profiles in mice transplanted with Eμ-Myc-Bcor.sh9, Eμ-Myc-Trp53.1224 or Eμ-Myc-Scram foetal liver cells

In addition to tracking donor derived and transduced populations, temporal changes in surface Ig markers were also assessed. Surface IgM and IgD profile can be used as an indicator of differentiation stage of circulating peripheral blood B cells (Nagasawa, 2006). Common lymphocyte progenitors can give rise to small pre-pro-B cells, which are surface B220 positive ($B220^+$) CD19 negative ($CD19^-$), IgM negative (IgM^+) and IgD negative (IgD^-) (Hardy et al., 1991). Pre-pro-B cells give rise to pro-B cells, which express B220 and CD19 on their surface but not IgM or IgD (Li et al., 1996). Pro-B cells develop into pre-B cells, which have a surface expression profile consisting of $B220^+$, $CD19^+$, IgM^- and IgD^- (Allman et al., 1999). Pre-B cells develop into immature

lymphocytes with a surface expression profile of $B220^+$, $CD19^+$, IgM^+ and IgD^-, which, in turn, develop into mature B cells, which are B220 positive, CD19 positive, IgM positive and IgD positive (Hardy et al., 1991; Nagasawa, 2006). In this experiment, mice that received E*μ-Myc-Bcor.*sh9, E*μ-Myc*-*Trp53.*1224 or E*μ-Myc*-Scram foetal liver cells were bled at weekly intervals so that the peripheral blood B cells could be assessed for changing surface expression of IgM and IgD over time. All circulating B cells were $B220^+$ and $CD19^+$.

Circulating peripheral blood B cells in mice that were transplanted with E*μ-Myc*-Scram foetal liver cells demonstrated a stable surface IgM and IgD expression profile from five weeks post-transplant to nine weeks post-transplant (Figure 4-3E). In this cohort, dual IgM and IgD positive cells were the most represented population, making up an average of 43.0% to 52.5% of the circulating B cells at each time point, followed by the IgM^+/IgD^- (13.1% to 32.1%), IgM^-/IgD^- (11.3% to 27.5%) and IgM^-/IgD^+ (5.8% to 9.2%) (Figure 4-3E).

B cells in mice transplanted with E*μ-Myc*-*Trp53*.1224 foetal liver cells demonstrated two dominant populations of cells that consisted of IgM^+/IgD^- and IgM^-/IgD^- surface expression profiles, which remained stable at five weeks and six weeks post-transplant (Figure 4-3F). At seven weeks post-transplant the dominant B cell population had a surface Ig expression profile of IgM^+/IgD^-; however, only two mice remained alive at this time point (Figure 4-3F). In mice that had received E*μ-Myc.Trp53*.1224 foetal liver cells, the dominant circulating B cell population was IgM^+/IgD^-, comprising of 42.0% to 84.3%

of the circulating B cells at each time point (Figure 4-3F). Cells with a surface immunophenotype of IgM^-/IgD^- or IgM^+/IgD^+ represented 14.0% to 33.8% and 1.6% to 21.7% of the circulating B cells measured at each time point, respectively (Figure 4-3F).

Five weeks post-transplant, 60.3% of the GFP positive B cells in the mice that received E*μ-Myc*.*Bcor.*sh9 were IgM^-/IgD^- (±6.9 SEM, n=13) (Figure 4-3G). B cells with a surface immunophenotype of IgM^+/IgD^-, IgM^+/IgD^+ or IgM^-/IgD^+ were represented as 21.6% (±6.3 SEM, n=13), 13.8% (±4.3 SEM, n=13) and 4.2% (±0.7 SEM, n=13) of the total population, respectively, at five weeks post-transplant (Figure 4-3G). From six weeks post-transplant to eight weeks post-transplant there is a trend indicating that as the lymphomas develop, the IgM/IgD double negative population expands (Figure 4-3G). This trend is lost by nine-weeks post-transplant; however, this is likely due to smaller numbers of experimental animals due to disease progression.

These results indicate that circulating E*μ-Myc*-Scram foetal liver derived B cells were much more heterogeneous in their surface marker expression and, therefore, their maturation stage (Figure 4-3H). The IgM/IgD surface expression profiles on circulating E*μ-Myc*-*Trp53*.1224 foetal liver derived B cells are likely to be representative of immature B cells (IgM^+/IgD^-), pre-B cells or pro-B cells (IgM^-/IgD^-) (Figure 4-3H). Circulating E*μ-Myc*.*Bcor.*sh9 foetal liver derived B cells appeared more likely to present as either pre-B cell or pro-B cell as indicated by IgM^-/IgD^- surface Ig expressing profile of the dominant B cell population (Figure 4-3H). All cells analysed were $GFP^+/B220^+/CD19^+$.

4.2.5 Disease latency and presentation in mice transplanted with Eμ-*Myc-Bcor*.sh9, Eμ-*Myc-Trp53*.1224 or Eμ-*Myc*-Scram foetal liver cells

*4.2.5.1 Monitoring disease latency in mice transplanted with E*μ-Myc-Bcor.*sh9, E*μ-Myc-Trp53.*1224 or E*μ-Myc-*Scram*

Time until sacrifice due to lymphoma burden was measured for each cohort of mice transplanted with E*μ-Myc-Bcor.*sh9, E*μ-Myc-Trp53.*1224 or E*μ-Myc*-Scram foetal liver cells. Previous studies showed that knockdown of *Trp53* in E*μ-Myc* foetal liver cells resulted in rapid tumour development in the recipient mice (Bric et al., 2009). All *Bcor* mutations identified in the WES screen were predicted loss of function mutations and the outgrowth of GFP positive cells in the blood of the animals transplanted with E*μ-Myc-Bcor.*sh9 foetal liver cells (detailed in section 4.2.4.1). Therefore, it was hypothesised that genetic depletion of *Bcor* transcript would lead to accelerated lymphomagenesis.

As expected, mice that received E*μ-Myc-Trp53*.1224 foetal liver cells displayed rapid tumour development with 100% penetrance (n=12) (Figure 4-4A). Mice that received E*μ-Myc-Trp53*.1224 foetal liver cells displayed a median survival time of 45 days, which was significantly lower than the groups of mice that received E*μ-Myc-Bcor.*sh9 and E*μ-Myc*-Scram foetal liver cells ($p < 0.05$, log-rank (Mantel-Cox) test) (Figure 4-4A). Mice that received E*μ-Myc-Bcor.*sh9 foetal liver cells demonstrated accelerated lymphomagenesis compared to mice that received E*μ-Myc*-Scram foetal liver cells, with a median survival time of 80 days post-transplant, with 12 out of the 13 mice

succumbing to lymphoma ($p < 0.05$, log-rank (Mantel-Cox) test) (Figure 4-4A). As expected, and as is previously described (Bric et al., 2009), mice transplanted with E*μ-Myc*-Scram foetal liver cells did develop disease over time with delayed kinetics of disease progression compared to mice transplanted with E*μ-Myc-Trp53*.1224 or E*μ-Myc-Bcor.*sh9 foetal liver cells (Figure 4-4A). This cohort had a median survival time of 151 days post-transplant with only seven out of 13 mice succumbing to disease at the conclusion of the experiment ($p > 0.05$, log-rank (Mantel-Cox) test) (Figure 4-4A).

4.2.5.2 Presentation of disease in mice transplanted with Eμ-Myc-Bcor.sh9, Eμ-Myc-Trp53.1224 or Eμ-Myc-Scram

At sacrifice due to signs of progressive lymphoma, the tumour burden as a total WBC count in the peripheral blood was measured in each animal reconstituted with either E*μ-Myc-Trp53.*1224*,* E*μ-Myc-Bcor*.sh9 or E*μ-Myc*-Scram foetal liver cells. Mice that received E*μ-Myc*-Scram foetal liver cells had an average terminal WBC count of 49.57×10^9 cells/L ($\pm 14.23\times10^9$ cells/L, SEM) (Figure 4-4B). At sacrifice, mice that were transplanted with E*μ-Myc-Bcor*.sh9 foetal liver cells demonstrated an average WBC count of 73.82×10^9 cells/L ($\pm 8.69\times10^9$ cells/L, SEM) (Figure 4-4B). Mice reconstituted with E*μ-Myc-Trp53.*1224 foetal liver cells had an average WBC count of 82.04×10^9 cells/L ($\pm 11.37\times10^9$ cells/L, SEM) at time of death (Figure 4-4B). All cohorts of mice were leukaemic at time of death with a mean WBC count that exceeded 14×10^9 cells/L; however, no group of mice had a significantly greater terminal peripheral WBC count compared to the other groups ($p > 0.05$, 2-way ANOVA) (Figure 4-4B). These data

indicated that although time until sacrifice differed, all mice died with the clinical symptoms of leukaemia/lymphoma.

In addition to increased WBC count, splenomegaly is another clinical indicator of lymphoma burden (Dasgupta et al., 1965; Davey et al., 1973). At time of sacrifice, a necropsy on each mouse transplanted with E*μ-Myc-Trp53.*1224*,* E*μ-Myc-Bcor*.sh9 or E*μ-Myc*-Scram foetal liver cells was assessed as a means of measuring disease burden. Mice reconstituted with E*μ-Myc-Trp53.*1224 foetal liver cells demonstrated a mean spleen weight of 0.54g (n=12) (Figure 4-4C). Mice transplanted with E*μ-Myc*-Scram foetal liver cells displayed a mean spleen weight of 0.46g (n=12) (Figure 4-4C). Mice that received E*μ-Myc-Bcor*.sh9 demonstrated an average spleen weight of 0.64g (N=12) at time of necropsy (Figure 4-4C). A WT C57BL/6 mouse spleen weighs approximately 0.1g (data not shown) indicating that each cohort of transplant recipient animals demonstrated lymphoma burden as assessed by spleen weights (Figure 4-4C). There was no significant difference in average spleen weights between the cohorts of recipient mice ($p > 0.05$, 2-way ANOVA) (Figure 4-4C). At time of euthanasia, spleens were harvested from six representative animals per cohort and analysed for GFP positivity. 41.07% of splenic B cells harvested from mice transplanted with E*μ-Myc-Trp53*.1224 demonstrated GFP positivity (Figure 4-4D). This was a significantly larger GFP positive splenic B cell population compared to mice transplanted with E*μ-Myc*-Scram (20.78% of splenic B cells were GFP positive), but not to mice transplanted with E*μ-Myc-Bcor*.sh9 (29.75% of splenic B cells were GFP positive) (Figure 4-4D). Consistent GFP positivity in the spleens of E*μ-Myc-Trp53.*1224 or E*μ-Myc-Bcor*.sh9

foetal liver cell-derived lymphoid compartments indicated that there was a competitive advantage harboured in the pLMS-*Trp53.*1224 or pLMS-*Bcor*.sh9 vectors.

4.2.5.3 Knockdown efficiency in Eμ-Myc-p53.1224 and Eμ-Myc-Bcor.sh9 foetal liver derived lymphomas

Lymphomas derived from E*μ-Myc-Bcor*.sh9 and E*μ-Myc-Trp53*.1224 foetal liver cells were GFP positive (D); however, the gene-knockdown efficiency was unknown. Transfection of 3T3 cells with pLMS-*Bcor*.sh9 elicited incomplete knockdown, with *Bcor* mRNA reduced to 34% compared to non-transfected cells (Figure 4-1B). It was hypothesised that E*μ-Myc-Bcor*.sh9 foetal liver cells that had more robust *Bcor* knockdown would outcompete cells with incomplete *Bcor* knockdown, to give rise to *Bcor*-depleted lymphomas in the recipient mice.

At time of death, splenic cells from the E*μ-Myc-Trp53*.1224 and E*μ-Myc-Bcor*.sh9 cohorts were harvested and sorted by flow cytometry based on GFP expression. Protein lysates from GFP positive cells from six representative E*μ-Myc-Bcor*.sh9 lymphomas and from lymphoma cell lines 6066 and 4242, which were defined in chapter 3.2.3 as $Bcor^{WT}$ and $Bcor^{MUT}$, respectively were used for protein immunoblotting to detect Bcor. E*μ-Myc-Bcor*.sh9 foetal liver derived lymphoma cells show complete knockdown of the Bcor protein in 2/6 cases (lanes 4 and 6) and downregulation of Bcor in the remaining cases (lanes 3, 5, 7 and 8), while lymphoma cell line 4242 demonstrated total loss of Bcor (lane 1) (Figure 4-4E). Lymphoma cell line 6066 demonstrated retention of the Bcor signal (lane 2) (Figure 4-4E). Consistent actin

expression indicated that differential protein loading did not account for the differences in Bcor expression in lymphomas and cell lines assessed (Figure 4-4E). The depletion of Bcor in the E*μ-Myc-Bcor*.sh9 lymphomas demonstrated that the lymphomas developed in the absence of Bcor, which was an on-target effect on the pLMS-*Bcor*.sh9 vector transduced into the foetal liver cells.

Trp53 is a negative regulator of $p19^{ARF}$ and in E*μ-Myc* lymphomas when *Trp53* is deleted or mutated; $p19^{ARF}$ expression is increased, serving as useful surrogate biomarker of Trp53 depletion (Eischen et al., 1999; Zindy et al., 1998). Protein lysates from GFP positive cells from six representative E*μ-Myc-Trp53*.1224 lymphomas and E*μ-Myc* lymphoma cell lines 4242, which was genomically defined as $Trp53^{WT}$ in section 3.2.3 and 3391 (which has been demonstrated as $Trp53^{MUT}$ (Shortt et al., 2013)) were used for western blotting to detect Trp53 and $p19^{ARF}$ (Figure 4-4F). No detectable Trp53 was observed in lymphoma cell line 3391 (lane 1) and E*μ-Myc-Trp53*.1224 lymphoma cells (lanes 3 to 8), while lymphoma cell line 4242 expressed Trp53 (lane 2) (Figure 4-4F). Supporting the loss of Trp53 signal in the lymphoma cell line 3391 (lane 1) and E*μ-Myc-Trp53*.1224 lymphoma cells (lanes 3 to 8) were that these samples demonstrated robust $p19^{ARF}$ expression compared to the lymphoma cell line 4242 (lane 2), which exhibited no $p19^{ARF}$ signal indicative of functional Trp53 (Figure 4-4F). Actin expression was approximately equivalent between the cell lines and primary lymphomas confirming that the lanes were each loaded with similar amounts of total protein (Figure 4-4F). The absence of Trp53 and enhanced expression of $p19^{ARF}$

indicated that the E*μ-Myc*-*Trp53*.1224 lymphomas have developed in the absence of Trp53 due to on-target effects of the pLMS-*Trp53*.1224 vector.

4.2.5.4 Analysis of surface IgM and IgD on splenic Eμ-Myc foetal liver cell derived lymphomas

Temporal analysis of changing surface IgM/IgD phenotype in circulating B cells revealed that the populations of B cells were not significantly different between E*μ-Myc*-Scram, E*μ-Myc*-*Trp53*.1224 and E*μ-Myc*-*Bcor*.sh9 lymphomas (section 4.2.4.4) (Figure 4-3E-G). It was hypothesised that the surface IgM/IgD phenotype of splenic E*μ-Myc* foetal liver derived lymphomas would support the trend that was discussed previously (section 4.2.4.4), with the presentation of circulating E*μ-Myc*-*Trp53*.1224 lymphoma cells as IgM^+/IgD^- (mature B cells) and E*μ-Myc*-*Bcor*.sh9 circulating cells as IgM^-/IgD^- (pro-B cells or pre-B cells). Splenic tissue was harvested from healthy WT C57BL/6 mice (n=3) and mice that were euthanised due to signs of progressive lymphoma following transplant with E*μ-Myc*-Scram (n=6), E*μ-Myc*-*Trp53*.1224 (n=12) and E*μ-Myc*-*Bcor*.sh9 (n=12) foetal liver cells. Splenocytes were flow cytometrically sorted based on GFP positivity and were analysed for surface Ig (IgM/IgD) profile after being filtered on B220/CD19 positivity. WT C57BL/6 splenic cells displayed significantly higher percentage of IgM/IgD double positive cells (41.67% ± 2.77% SEM) compared to E*μ-Myc*-Scram, E*μ-Myc*-*Trp53*.1224 and E*μ-Myc*-*Bcor*.sh9 lymphomas, which all displayed absence of the IgM^+/IgD^+ population, represented by 2.73% (± 2.04% SEM), 3.69% (± 1.06% SEM) and 1.10% (± 0.26% SEM) of the splenic B cells, respectively ($p < 0.05$, 2-way ANOVA) (Figure 4-5A-E). E*μ-Myc*-Scram and E*μ-Myc*-*Trp53*.1224 lymphomas

presented with significantly larger population of cells consisting of an IgM^+/IgD^- surface immunoglobulin phenotype (50.53% ± 16.89% SEM and 44.10% ± 8.95% SEM, respectively), compared to C57BL/6 splenocytes or E*μ-Myc-Bcor*.sh9 lymphomas (17.22% ± 6.51% SEM) ($p < 0.05$, 2-way ANOVA) (Figure 4-5A-E). E*μ-Myc-Bcor*.sh9 lymphomas exhibited a significantly larger population of IgM^-/IgD^- B cells (81.56% ± 6.56% SEM) compared to WT C57BL/6 (38.90% ± 6.93% SEM), E*μ-Myc*.Scram lymphomas (46.62% ± 18.03% SEM) and E*μ-Myc*.p53.1224 lymphomas (51.85% ± 9.60% SEM) ($p < 0.05$, 2-way ANOVA) (Figure 4-5A-E). These data indicated that compared to WT C57BL/6 splenic B cells, E*μ-Myc* foetal liver derived lymphomas can be classified immunophenotypically as immature B cells, with almost total loss of the mature IgM^+/IgD^+ compartment (Figure 4-5A-E). The data also suggested that E*μ-Myc-Bcor.*sh9 lymphomas were largely pro-B- or pre-B cell tumours, and were more immature than the mostly IgM^+/IgD^- E*μ-Myc-Trp53*.1224 lymphomas.

This finding was extended to include sporadic E*μ-Myc* lymphomas with known genetic lesions. Surface immunoglobulin expression analysis was performed on the sporadic E*μ-Myc* tumours detailed in section 3.2.3 (12 $Bcor^{WT}$ E*μ-Myc* lymphomas and four $Bcor^{MUT}$ E*μ-Myc* lymphomas), six E*μ-Myc-Bcor*.sh9 lymphomas and six E*μ-Myc*.p53.1224 lymphomas. 90% of $Bcor^{MUT}$ tumours (be it sporadically or through shRNA-mediated knockdown) presented as an IgM and IgD double negative B cell lymphoma (Figure 4-5F). This is a significant increase compared to the $Bcor^{WT}$ group in which only 44.4% of the E*μ-Myc* tumours were predominantly IgM and IgD negative ($p < 0.05$, Chi-squared test) (Figure 4-5F). These data indicated that E*μ-Myc* lymphomas in

which Bcor function is lost by mutation or genetic manipulation are generally more immature compared to tumours that developed in the presence of WT Bcor.

4.2.6 CRISPR/Cas9-mediated deletion of *Bcor in vivo*

4.2.6.1 CRISPR/Cas9 construct

To further validate the discovery that *Bcor* deletion cooperates with oncogenic *Myc*-expression to accelerate lymphomagenesis, an orthogonal method of genetically deleting *Bcor* in E*μ-Myc* cells was utilised. The CRISPR/Cas9 technology has been developed for use in mammalian genome editing, generating new possibilities to drive genetic engineering in animal models of cancer (Mali et al., 2013). The pQ-based retroviral construct was designed to drive expression of Cas9, a GFP reporter and the gene-targeting guide RNA (sgRNA) (Figure 4-6A). The sgRNAs were designed to be either non-targeting (Scram), *Trp53*-targeted or *Bcor*-targeted. An sgRNA targeting *Bcor* (pQCIG-*Bcor*G2) was designed to be complementary to nucleotide 1388-1410 (exon 4), which is proximal to the deletion identified in lymphoma 4242, and was integrated into the pQCIG-trigger construct (Figure 4-6B). pQCIG-*Trp53*(b) and pQCIG-scram were constructs kindly provided by Abba Malina and Jerry Pelletier designed to target *Trp53* and to be non-targeting, respectively. 3T3 cells were transfected with pQCIG-*Bcor*G2 as a proof of concept. A DNA extraction was performed on six different cultures of GFP-positive 3T3 cells. Following DNA extractions, a PCR was performed on each representative culture, with the primers designed to flank the region in which the *Bcor*-targeting sgRNA binds to and PCR products were run on a DNA chip to assess size differences (Figure 5-6C). WT DNA (lane 8) demonstrated a single *Bcor*

band at the expected fragment size (262 bases) with no background signal (Figure 4-6C). Two DNA samples (lane 2 and lane 6) demonstrated markedly reduced product size and a large number of non-specific signals compared to WT DNA (Figure 4-6C). Three DNA samples (lane 3, lane 4 and lane 5) showed subtle reductions in product size and an increase in the non-specific, background signal compared to WT DNA (Figure 4-6C). DNA from lane 5 demonstrated an increase in the PCR product size compared to WT DNA (Figure 4-6C). These data indicate that the pQCIG-*Bcor*G2 constructs in the 3T3 cells introduced structurally disruptive variations including deletions and insertions at the targeted *Bcor* locus.

4.2.6.2 Monitoring disease latency in mice transplanted with Eμ-Myc.pQCIG-Trp53(b), Eμ-Myc.pQCIG-BcorG2 or Eμ-Myc.pQCIG-scram foetal liver cells

To recapitulate the experiments shown in 4.2.5 using shRNA to knockdown Bcor or Trp53, the Cas9 constructs were co-transduced into E*μ-Myc* foetal livers with pQCIG-*Trp53*(b), pQCIG-*Bcor*G2 or pQCIG-scram creating E*μ-Myc*.pQCIG-*Trp53*(b), E*μ-Myc*.pQCIG-*Bcor*G2 or E*μ-Myc*.pQCIG-scram foetal liver cells that were then injected into sub-lethally irradiated syngeneic recipients. Time until sacrifice due to signs and symptoms of progressive lymphoma was measured for each cohort (Figure 4-7A). Mice that received E*μ-Myc*.pQCIG-scram foetal liver cells (n=6) did not reach median survival time in the defined 160 days, with a single death occurring in this group (Figure 4-7A). Mice transplanted with E*μ-Myc*.pQCIG-*Trp53*(b) (n=6) or E*μ-Myc*.pQCIG-*Bcor*G2 (n=6) foetal liver cells demonstrated a median survival time of 34 and 70 days post-transplant, respectively (Figure 4-7A). Mice transplanted with E*μ-Myc*.pQCIG-*Trp53*(b) foetal liver

cells demonstrated significantly accelerated lymphomagenesis compared to mice that received E*μ-Myc*.pQCIG-*Bcor*G2 or E*μ-Myc*.pQCIG-scram foetal liver cells ($p < 0.05$, Log-rank (Mantel-Cox) test) (Figure 4-7A). Recapitulating what was demonstrated using shRNA constructs in section 4.2.5, mice that received E*μ-Myc*.pQCIG-*Bcor*G2 demonstrated significantly accelerated time to sacrifice due to progressive lymphoma compared to the negative control ($p < 0.05$, Log-rank (Mantel-Cox) test) (Figure 4-7A). These data confirm that Bcor loss cooperated with Myc to drive and/or initiate lymphomagenesis in the E*μ-Myc* model of lymphoma.

4.2.6.3 Efficiency of CRISPR/Cas9-mediated deletion of Bcor

Necropsies were performed on the mice that had succumbed to progressive lymphoma following transplantation with E*μ-Myc*.pQCIG-*Bcor*G2 foetal liver cells. The spleens were removed from the animals and were cell-sorted based on GFP expression and protein, DNA and RNA was collected and used for further analysis. To demonstrate that Bcor expression was lost in E*μ-Myc*.pQCIG-*Bcor*G2 lymphomas, Bcor immunoblotting was performed on the protein lysates from the GFP-positive lymphoma cells. As defined by WES in section 3.2.3, lymphoma cell lines 4242 (*Bcor*^{MUT}) and *6066* (*Bcor*^{WT}) were used as a positive and negative control of Bcor loss (Figure 4-7B). Each of the five representative primary E*μ-Myc*-pQCIG-*Bcor*G2 lymphomas and lymphoma cell line 4242 demonstrated total lack of Bcor protein compared to lymphoma cell line 6066 cells (Figure 4-7B). Levels of actin protein were consistent between samples, indicating that the differences in Bcor expression could not be attributed to total protein loading inconsistencies (Figure 4-7B).

RNA sequencing was performed on all six E*μ-Myc.*pQCIG-*Bcor*G2 lymphomas and sequence reads across *Bcor* were assessed (discussed below, section 4.2.10) (Figure 4-7C). E*μ-Myc.*pQCIG-*Bcor*G2 lymphomas 1, 2, 4 and 5 demonstrated small deletion events at the sgRNA target sequence in the *Bcor* transcript (Figure 4-7C). E*μ-Myc.*pQCIG-*Bcor*G2 lymphomas 3 and 6 demonstrated large regions of reduced read depth spanning the *Bcor* transcript, indicating that the pQCIG-*Bcor*G2 construct had likely caused a large deletion event (Figure 4-7C). Taken together, these results demonstrated that the pQCIG-*Bcor*G2 construct correctly targeted *Bcor*, leading to protein disruption and resulting in accelerated lymphomagenesis in the E*μ-Myc* model of lymphoma.

4.2.7 *Bcor* re-expression in *Bcor*-null cell lines

The results to date indicate that genetic depletion of Bcor cooperated with Myc to accelerate lymphomagenesis in the *Eμ-Myc* model. To complement this finding, WT *Bcor* was forcibly expressed in lymphoma cell line 4242, defined in section 3.2.3 as a $Bcor^{MUT}$ lymphoma. It was hypothesised that forced WT *Bcor* expression in an E*μ-Myc* lymphoma cell line that harboured a high VAF *Bcor* loss-of-function mutation would confer a competitive disadvantage to that cell line. The MSCV expression vector co-expressed WT Bcor and a GFP reporter protein. Lymphoma cell line 4242 was transduced with MSCV-*Bcor* and an empty MSCV vector, to be used as a negative control. The cells were sorted based on GFP expression, which resulted in pure transduced populations of 4242-MSCV.*Bcor* and 4242-MSCV cell lines. Each transduced cell line was mixed with the parental non-transduced, 4242 lymphoma cell

line in a ratio of 1:1. A temporal analysis of the ratio of transduced cells to non-transduced cells, by way of GFP positivity, was assessed by flow cytometry. The ratio of 4242-MSCV and parental 4242 cells remained at approximately 50% until day 13 post culture, where a trend of increasing ratio of transduced cells to non-transduced cells became apparent. At day 22, GFP positive cells made up approximately 60% of the population (Figure 4-8A). In culture, the ratio of 4242-MSCV.*Bcor* to parental 4242 cells declined by day 9, where 4242-MSCV.*Bcor* accounted for 40% of the cells in culture (Figure 4-8A). By day 22 of the experiment, 4242-MSCV.*Bcor* cells represented approximately 22% of the cell population (Figure 4-8A). Whole cell lysates from parental 4242 cells, 4242-MSCV cells and 4242-MSCV.*Bcor* cells were analysed by immunoblotting assay for Bcor expression to demonstrate that Bcor had been restored in the 4242-MSCV.*Bcor* cells. Parental 4242 cells and 4242-MSCV demonstrated no Bcor expression whereas 4242-MSCV.*Bcor* cells expressed Bcor protein (Figure 4-8B). The actin immunoblot demonstrated that each sample is represented by an equal amount of total protein (Figure 4-8B). Taken together, these *in vitro* longitudinal proliferation studies demonstrated that restoration of *Bcor* in *Bcor*^{MUT} E*μ-Myc* lymphoma cell line 4242 confers a competitive disadvantage within the cells.

Similar longitudinal competitive proliferation assays were performed in NAMALWA cells, which are of human BL origin (Henderson et al., 1983). NAMALWA cells were chosen for this assay due to their inherently low levels of *Bcor* expression (data not shown). NAMALWA cells were transduced with either an empty MSCV vector or an MSCV.*Bcor* expression vector. The cells were sorted based on GFP expression and cells

transduced with MSCV alone or with MSCV.*Bcor* were co-cultured with non-transduced parental NAMALWA cells. The ratio of NAMALWA-MSCV to parental NAMALWA cells increased to approximately 53% to 47% by the third day of the experiment, while at conclusion of the experiment NAMALWA-MSCV cells contributed approximately 60% of the co-culture (Figure 4-8C). The ratio of NAMALWA-MSCV.*Bcor* to parental NAMALWA cells decreased over time in the co-culture (Figure 4-8C). At six days post-sort, NAMALWA-MSCV.*Bcor* cells represented approximately 42% of the co-culture, which continued to drop further to approximately 32% by day 11 (Figure 4-8C). As expected, NAMALWA-MSCV cells displayed no Bcor expression by immunoblot; however, transduction with MSCV.*Bcor* restored Bcor expression in the NAMALWA cells (Figure 4-8D). Actin expression demonstrated that protein loading was consistent for each cell line (Figure 4-8D). These data indicated that *Bcor* restoration in *Bcor*^{MUT} lymphomas or in *Bcor* non-expressing NAMALWA BL cell lines drives loss of tumour cell representation.

4.2.8 Gene expression profile of E*μ-Myc* foetal liver derived lymphomas

Currently, there is no published evidence pointing to a role for Bcor in directly regulating tumor cell proliferation or survival so exactly how Bcor loss co-operates with Myc to drive tumor progression remains unclear. To begin to elucidate the underlying biology of *Bcor*-mutant lymphomas, RNA sequencing analysis of Eμ-*Myc*-*Bcor*.sh9 (n=6), Eμ-*Myc*-*Trp53.*1224 (n=6) lymphomas or Eμ-Myc lymphomas overexpressing *Nras*^{Q61K} (Eμ-*Myc*-*Nras*^{Q61K}, n=5) was performed. Briefly, E*μ-Myc-Nras*^{Q61K} lymphomas were

generated in the Johnstone Laboratory by cloning full length *Nras* harbouring a C to A missense mutation at CDS position 181 leading to a Q61K amino acid substitution. The *Nras*^{Q61K} mutation occurs spontaneously in E*μ-Myc* lymphoma cell line 6066 and ML27, as defined in section 3.2.3, and has been identified as a clinically relevant hotspot mutation in human malignancies (Buhrman et al., 2011b). Injecting E*μ-Myc-Nras*^{Q61K} foetal liver cells into recipient mice resulted in the rapid formation of lymphomas (data not shown), recapitulating published data (Alexander et al., 1989). RNA was extracted from GFP positive lymphomas and RNA-seq was performed. The transcriptional snap shot encompassed expression data for 1112 genes, filtered from a list of 16132 as shown (Table 4-2). Unsupervised clustering was performed on Eμ-*Myc-Bcor*.sh9, Eμ-*Myc-Trp53.*1224 and Eμ-*Myc-Nras*^{Q61K} using the filtered 1112 genes (Figure 4-9A). The E*μ-Myc* tumours clustered in groups, with the Eμ-*Myc-Bcor*.sh9 tumours clustering together, the Eμ-*Myc-Trp53.*1224 tumours clustering together and the Eμ-*Myc-Nras*^{Q61K} clustering together based on gene expression data of the 1112 genes in the gene list (Figure 4-9A). Recapitulating this is the two-dimensional principal component analysis that also demonstrated that the tumours clustered together dependent on the introduced genetic lesion (Figure 4-9B). These results indicated that there are distinct molecular signals and gene expression profiles for each tumour based on the introduced driver mutation.

Table 4-2 Filters applied to gene list prior to unsupervised clustering analysis

Filter	Value
% Present	80
SD (gene vector)	2.0
Observation	≥1.0
Absolute value	≥2.0

4.2.9 Unique gene expression signature in *Bcor*^{MUT} tumours

To generate a *Bcor*^{MUT} gene expression signature, the gene expression profile of Eμ-*Myc*-*Bcor*.sh9 lymphomas (n=6) was contrasted against RNA sequencing data from Eμ-*Myc*-*Trp53*.1224 (n=6) and E*μ*-*Myc*-*Nras*^{Q61K} (n=5) lymphomas, which were pooled together to form the *Bcor*^{WT} gene set. WES of spontaneous *Eμ-Myc* lymphomas in chapter 3 failed to demonstrate co-occurrence of mutations in the *Ras* genes, *Trp53* or *Bcor*, leading to the assumption that neither Eμ-*Myc*-*Trp53*.1224 or E*μ*-*Myc*-*Nras*^{Q61K} lymphomas harboured spontaneously acquired *Bcor* mutations. The gene expression signature of *Bcor*^{WT} E*μ*-*Myc* lymphomas (consisting of Eμ-*Myc*-*Trp53*.1224 and E*μ*-*Myc*-*Nras*^{Q61K} lymphomas) was subtracted from the *Bcor*^{MUT} E*μ*-*Myc* lymphoma gene signature (Figure 4-10). This approach identified 393 genes that were significantly differentially expressed by at least a factor of two in *Bcor*^{MUT} lymphomas compared to *Bcor*^{WT} lymphomas (adjusted p-value ≤ 0.05, Log fold-change ≥1 or ≤-1) (Table 4-3). *Bcl6* was one of the most downregulated genes in the *Bcor*^{MUT} tumours compared to *Bcor*^{WT} lymphomas with a Log fold change (LogFC) of -3.3 (adjusted p-value < 0.05) (Table 4-3). None of the other genes carrying mutations identified in chapter 3.2.3 (*Kras, Nras*, *Trp53* or *Ezh2)* were differentially expressed in this analysis except for *Cdkn2a*, which was downregulated in the *Bcor*^{MUT} group with a LogFC of -3.9 (adjusted p-value < 0.05) (Table 4-3). This change in *Cdkn2a* expression was likely representative of inactive Trp53 in the Eμ-*Myc*-*Trp53*.1224 lymphomas in the *Bcor*^{WT} group, rather than true loss of expression in *Bcor*^{MUT} lymphomas.

Table 4-3 Significantly differentially expressed genes in *Bcor*^{MUT} E*μ-Myc* lymphomas

Mouse gene	Human gene	Log fold-change	Adjusted p-value
Bahcc1	BAHCC1	8.604006521	2.06E-08
Cpn1	CPN1	7.073555889	0.001858205
Gm11772		6.596025126	2.06E-08
Cited1	CITED1	5.277396694	0.001806948
Robo1	ROBO1	5.253332875	0.015741734
Ebf3	EBF3	5.23260973	0.004428867
Islr2	ISLR2	5.228926584	3.07E-05
2900052L18Rik		5.083420064	6.09E-05
Rassf6	RASSF6	5.049037097	0.011664748
Islr	ISLR	4.761537881	3.07E-05
Prox1	PROX1	4.551686784	0.00526891
1700018A04Rik		4.459337285	0.001848225
Fras1	FRAS1	4.374569406	0.004320425
Gm11620		4.265853519	0.002422343
Podxl		4.208025352	0.045862252
Gm11377		4.088370759	0.001920843
Arsi	ARSI	4.059724919	0.001858205
A430105I19Rik	C15orf52	4.033326742	0.002990867
Krt18	KRT18	4.026290796	0.016459419
AI848285		4.023921621	0.000188944
Slco5a1	SLCO5A1	3.799037564	0.002831583
Zfhx3	ZFHX3	3.797113253	0.002377807
Plekha7	PLEKHA7	3.78527021	0.013235101
Prkag3	PRKAG3	3.73183529	0.001858205
Trerf1	TRERF1	3.71228115	0.012035464
Gm9847		3.497628377	0.001264058
Ppp1r26	PPP1R26	3.425846663	0.002254495
Kbtbd11	KBTBD11	3.422168424	0.002193755
Cdsn		3.389475081	0.001783641
Srgap1	SRGAP1	3.353374043	0.003895048
Gm11744		3.295575898	0.049335153
Nefh	NEFH	3.214143752	0.001550446
Acacb	ACACB	3.201588201	0.023033254
Apobec2	APOBEC2	3.117108054	0.004238537
Cbr3	CBR3	3.093637592	0.022511627
Espn	ESPN	3.089078325	0.007986864
C430049B03Rik		3.085930511	0.007548572
Nes		3.036423947	0.04660056
Foxf2	FOXF2	2.985058403	0.004960083

Soga1	SOGA1	2.966578904	0.041382082
Ptrf	PTRF	2.921857856	0.015783285
A930004D18Rik		2.919121411	0.000290509
Arhgap42	ARHGAP42	2.901068858	0.026380051
Gcgr	GCGR	2.867915529	0.048491925
Nlgn2	NLGN2	2.797136625	0.01675833
Mapk12	MAPK12	2.707820915	0.000116344
Socs2	SOCS2	2.695542672	0.000498751
Skida1	SKIDA1	2.685718458	0.013721414
8430427H17Rik		2.684876532	0.001806948
Gls2	GLS2	2.665009307	0.005913835
Tmcc2	TMCC2	2.633583947	0.009960402
Stac	STAC	2.564888566	0.045192386
Cuedc1	CUEDC1	2.52656988	0.006459757
Eps8l2	EPS8L2	2.522022104	0.001297554
Lrp5	LRP5	2.513554799	0.000534399
Lamb2	LAMB2	2.491144816	0.003159707
Gadd45b	GADD45B	2.487227725	0.004671744
Tgfb2	TGFB2	2.448804354	0.041046851
Plekhh3	PLEKHH3	2.429826561	0.041046851
Zfp629	ZNF629	2.370206776	0.006938858
Ltbp4	LTBP4	2.32722693	0.04765159
AC166052.1		2.318645383	0.004065969
Arhgef10	ARHGEF10	2.288919574	0.034222833
Bag3	BAG3	2.286594472	0.049037206
Trio	TRIO	2.28342288	0.034222833
Arhgef10l	ARHGEF10L	2.260937251	0.022511627
Aes	AES	2.258239126	0.016459419
Nid2	NID2	2.249605268	0.022495712
Prodh	PRODH	2.245536784	0.010690477
Gm10451		2.240469441	0.002599521
Fbxl2	FBXL2	2.237895668	0.019471813
Pkn3	PKN3	2.179963254	0.000309124
Erbb3	ERBB3	2.125065157	0.034298813
Tbc1d30	TBC1D30	2.075375041	0.004479794
Klrg2	KLRG2	2.073056598	0.015214874
Plekha5	PLEKHA5	2.063848391	0.001297554
Bhlha15	BHLHA15	2.05017317	0.01740441
Kif7	KIF7	2.03720846	0.014423889
Nt5dc2	NT5DC2	2.020146216	0.004065969
BC064078		2.013294125	0.003255224
Stom	STOM	1.97094645	0.004479794
Pard6g	PARD6G	1.969512537	0.02010567

Jup	JUP	1.908041187	0.025937301
B4galt2	B4GALT2	1.874376598	0.024169529
Pcgf2	PCGF2	1.865235161	0.002451835
Tubb2a	TUBB2A	1.846929652	0.001550446
Ets2	ETS2	1.816166724	0.024833239
Gm410		1.809616262	0.030994183
3425401B19Rik	C10orf71	1.801768153	0.011664748
Ass1	ASS1	1.778703011	0.010818323
Slc6a9	SLC6A9	1.761794346	0.027637946
Angptl4	ANGPTL4	1.75469144	0.020580932
Lzts2	LZTS2	1.751544476	0.04829211
Epn2	EPN2	1.737616161	0.041809364
Lrfn1	LRFN1	1.733357374	0.021399331
Lphn1	LPHN1	1.725208742	0.03421227
Tmem8	TMEM8A	1.694488971	0.035482199
Mb21d1	MB21D1	1.686083658	0.00614767
Nfix	NFIX	1.674203516	0.00526891
Fkbp11	FKBP11	1.666215349	0.0217707
Ryk	RYK	1.662039477	0.02778769
Gas2l1	GAS2L1	1.654108595	0.037446499
Zbtb8a	ZBTB8A	1.635285103	0.001297554
Akap7	AKAP7	1.633267272	0.005903234
Xylb	XYLB	1.631974041	0.000715231
Tmem59l	TMEM59L	1.627429188	0.020520631
Gm5424		1.60639419	0.003310848
B930041F14Rik	C1orf233	1.575501945	0.041382082
Atn1	ATN1	1.557911879	0.021028898
Itga9	ITGA9	1.535207834	0.012167864
Enc1	ENC1	1.500210315	0.034298813
Cisd3		1.495566353	0.037721016
Kank3	KANK3	1.493829923	0.004428867
1190002N15Rik	C3orf58	1.489002432	0.016223014
Gm13196		1.480557678	0.017673493
Gareml	GAREML	1.472827105	0.018135229
Frat2	FRAT2	1.457907529	0.010413824
Rfx2	RFX2	1.450301843	0.034298813
1110059E24Rik	C9orf85	1.444077166	0.013342391
Dclre1a	DCLRE1A	1.439948482	0.015684578
Inpp5a	INPP5A	1.431023187	0.011240875
Zhx2	ZHX2	1.419027108	0.001412495
Cenpv	CENPV	1.384815909	0.001478662
Acvr2b	ACVR2B	1.37051788	0.031145633
Glipr2	GLIPR2	1.36889859	0.021048403

Gkap1	GKAP1	1.35095237	0.011664748
Kcnk5	KCNK5	1.347551727	0.001297554
2900008C10Rik		1.347468007	0.04866966
Itsn1	ITSN1	1.343687427	0.007548572
Casp7	CASP7	1.3264408	0.001550446
Tspan4	TSPAN4	1.325648914	0.035482199
Tmem191c		1.31770248	0.011664748
Zfp623	ZNF623	1.311940645	0.001297554
Atg9b	ATG9B	1.301802278	0.007786583
Cln6	CLN6	1.28027671	0.004428867
Mxi1	MXI1	1.279415745	0.002038755
Baiap2	BAIAP2	1.277239829	0.002337887
Sepn1	SEPN1	1.274693708	0.024915817
Selenbp2	SELENBP1	1.269231313	0.021655329
Tubb2b	TUBB2B	1.263540701	0.020847687
Selenbp1		1.254428749	0.010653394
St3gal4	ST3GAL4	1.233891124	0.001550446
Thbd	THBD	1.227773627	0.017118008
Cltb	CLTB	1.21443962	0.041809364
Cacna1a	CACNA1A	1.203607952	0.022495712
Luzp1	LUZP1	1.188420518	0.005015437
Snx33	SNX33	1.187529144	0.016858758
Bambi	BAMBI	1.180284159	0.04064368
Pcdhgc4	PCDHGC4	1.170200277	0.009960402
Naf1	NAF1	1.156780946	0.004960083
Ppp1r13l	PPP1R13L	1.137805849	0.032279824
Ppp2r5b	PPP2R5B	1.095149811	0.037721016
Btrc	BTRC	1.079808021	0.007874525
Slc27a1	SLC27A1	1.078810196	0.049512812
Rab31	RAB31	1.053784923	0.017168556
Smad3	SMAD3	1.053451434	0.020719271
Fbxl4	FBXL4	1.052148836	0.034933058
Erf	ERF	1.040873878	0.031212501
Bag2	BAG2	1.02111258	0.006837195
Lrrn4	LRRN4	1.012035611	0.034222833
Pitpnc1		1.009473835	0.031792655
Dnmt3b	DNMT3B	1.00797726	0.00526891
Elk3	ELK3	-1.004791841	0.04829211
Tec	TEC	-1.008312661	0.03829012
Gng2	GNG2	-1.02441136	0.030391933
Ube2l6	UBE2L6	-1.03764133	0.019707551
Dyrk1b	DYRK1B	-1.04160488	0.014423889
AA414768		-1.054096197	0.04660056

Abhd8	ABHD8	-1.056924822	0.005812773
Gimap5	GIMAP1-GIMAP5	-1.059770612	0.016743218
Pbxip1	PBXIP1	-1.06497945	0.006456759
2610018G03Rik		-1.079302802	0.003979406
Rdh12	RDH12	-1.084676076	0.022495712
Soat1	SOAT1	-1.102311915	0.043223093
Btg1	BTG1	-1.105932069	0.015518648
Cd53	CD53	-1.108934205	0.023314567
H2-M3	HLA-G	-1.122346398	0.036648145
Bpgm	BPGM	-1.131076093	0.01027547
Arhgap25	ARHGAP25	-1.149428616	0.016888468
2610524H06Rik		-1.150261454	0.018955672
Pgam2	PGAM2	-1.153999982	0.041809364
Dhx40	DHX40	-1.169339853	0.017500242
Trim34a		-1.189191927	0.007490409
Fmo1	FMO1	-1.1910521	0.032641874
Ifi203		-1.195638456	0.004428867
Slc2a9	SLC2A9	-1.199207022	0.017165741
Gm16340		-1.212190317	0.034298813
Plac8	PLAC8	-1.212518371	0.012167864
Intu	INTU	-1.214226635	0.021571293
Aim1	AIM1	-1.218431141	0.011970455
Dpy19l3	DPY19L3	-1.228203651	0.021584103
Srgn		-1.231323704	0.002377807
Mrgpre	MRGPRE	-1.23416813	0.013777362
Pecam1	PECAM1	-1.249872417	0.004113285
AC123702.2		-1.262647292	0.049389047
Gstm4	GSTM4	-1.274106766	0.027637946
Gramd3	GRAMD3	-1.302677837	0.019372364
Snx29	SNX29	-1.312803836	0.049070427
Kctd7	KCTD7	-1.321382098	0.005015437
Csgalnact2	CSGALNACT2	-1.328333605	0.021227197
AC142191.1		-1.345607669	0.02453399
Fam64a	FAM64A	-1.35051651	0.032641874
A530040E14Rik		-1.35082199	0.045192386
CT010433.2		-1.370378173	0.042652695
Dcaf12	DCAF12	-1.374363761	0.002422343
Itih5	ITIH5	-1.37619062	0.049091401
Map3k5	MAP3K5	-1.387282757	0.039106544
Sh3pxd2a	SH3PXD2A	-1.394321822	0.02504033
Rph3al	RPH3AL	-1.401350461	0.005465385
Gbp5	GBP5	-1.403720168	0.039607174
Klhl6	KLHL6	-1.41554343	0.002337887

Foxo1	FOXO1	-1.434646611	0.000721731
4930431F12Rik		-1.434763289	0.032315186
Ccdc88b	CCDC88B	-1.447277087	0.018527772
Lpp	LPP	-1.4510491	0.04920026
Trim36	TRIM36	-1.455146135	0.025937301
Zfp820		-1.474280905	0.02866495
Ssx2ip	SSX2IP	-1.486718414	0.044726902
Paip2b	PAIP2B	-1.489041681	0.001412495
Hrh2	HRH2	-1.491771783	0.026380051
Sh2d2a	SH2D2A	-1.495206148	0.036491005
Cd93	CD93	-1.501191294	0.002340278
Lrrk2	LRRK2	-1.525382586	0.005968472
Dennd2d	DENND2D	-1.545942739	0.008813404
Cyth4	CYTH4	-1.552213113	0.049070427
Sfxn5	SFXN5	-1.561938282	0.034398208
Arl6	ARL6	-1.58354807	0.039106544
H2-Ob	HLA-DOB	-1.587395385	0.002350744
Abtb2	ABTB2	-1.58878214	0.037721016
Atp6v0a1	ATP6V0A1	-1.628845384	0.022356701
Eid1	EID1	-1.659980405	0.004954567
Abcb1b		-1.667012253	0.011240875
Tjp2	TJP2	-1.667535074	0.040078699
Mxd4	MXD4	-1.669103729	0.012981255
Ikbke	IKBKE	-1.681163345	0.013342391
Evl	EVL	-1.695357891	0.010012114
Btnl2	BTNL2	-1.72804581	0.0217707
Mier2	MIER2	-1.776853013	0.034298813
Fam213a	FAM213A	-1.779691226	0.031223895
Trim5	TRIM5	-1.800156751	0.002451835
Slc37a2	SLC37A2	-1.80502491	0.00318503
Pik3cg	PIK3CG	-1.81211396	0.001412495
Slamf7	SLAMF7	-1.813292712	0.001297554
Pyhin1	IFI16	-1.821697221	0.019539686
Gm10392		-1.843239345	0.038213566
2610035D17Rik		-1.869046707	0.03400491
Pde4dip	PDE4DIP	-1.878355523	0.00318503
Dcbld2	DCBLD2	-1.883093589	0.0217707
Tmem108	TMEM108	-1.897084612	0.049261661
Grhl1	GRHL1	-1.916032408	0.03421227
Tanc2	TANC2	-1.926356876	0.04678005
Kif3a	KIF3A	-1.955700259	0.00318503
Dnajb5	DNAJB5	-1.955796347	0.024915817
Cd244	CD244	-1.96104245	0.010568762

Abhd5	ABHD5	-1.972260582	0.023428815
Cpm	CPM	-1.986144661	0.002350744
Snx10	SNX10	-1.989343486	0.009726156
Tmem2	TMEM2	-2.012235101	0.004263103
Celf4	CELF4	-2.012878174	0.045736554
Stk17b	STK17B	-2.031229884	0.001264058
Sell	SELL	-2.034728882	0.004954567
Lax1	LAX1	-2.065231337	0.005441137
Tctn2	TCTN2	-2.073168667	0.04829211
Cd2	CD2	-2.077043477	0.008068408
Lef1	LEF1	-2.088372222	0.016850088
Evi5l		-2.093196377	0.049818261
Cyp4f18	CYP4F2	-2.114404295	0.004479794
Syne1	SYNE1	-2.12680065	0.049091401
Cd52		-2.141186453	0.003159707
Ankrd50		-2.184918129	0.004479794
Acer3	ACER3	-2.21508712	0.00830263
Slc14a1	SLC14A1	-2.229388528	0.01027547
2610008E11Rik		-2.233401864	0.008526074
Rab37	RAB37	-2.234464248	0.030621596
Gimap6	GIMAP6	-2.267347139	0.001297554
Zfp248	ZNF248	-2.26826938	0.003255224
Stap1	STAP1	-2.271918803	0.001341229
Rasal2	RASAL2	-2.283110999	0.010449009
Mitf	MITF	-2.315830848	0.026380051
Ms4a6b		-2.321660683	0.009726156
Tmem255a	TMEM255A	-2.332028186	0.009531273
H2-Ab1	HLA-DQB1	-2.332362846	0.023726186
Dip2a	DIP2A	-2.341141751	0.003895048
Slfn2		-2.357225773	0.048870191
Fam92a	FAM92A1	-2.361076763	0.011664748
Abcb1a	ABCB1	-2.363810973	0.034933058
Map3k9	MAP3K9	-2.392460827	0.001297554
Dap	DAP	-2.398168674	0.002451835
Megf9	MEGF9	-2.432327185	0.014297383
Gm8369		-2.44299906	0.037087297
Sptbn4	SPTBN4	-2.445715711	0.02778955
Dennd2c	DENND2C	-2.445963676	0.041346122
Hip1	HIP1	-2.456536678	0.017163788
Sh2b2	SH2B2	-2.507876982	0.032518628
Gimap8	GIMAP8	-2.520305601	0.015684578
Endou	ENDOU	-2.529663362	0.048870191
Camkk1	CAMKK1	-2.537115905	0.049639106

Ms4a4b		-2.54567826	0.027680845
Cldn12	CLDN12	-2.549961439	0.006044317
Smpdl3a	SMPDL3A	-2.554514079	0.049196524
Tmem38b	TMEM38B	-2.579793663	0.004428867
9130008F23Rik	C6orf141	-2.61987393	0.017056517
Myo10	MYO10	-2.655222046	0.015116167
Samsn1	SAMSN1	-2.670380149	0.001297554
Gpr183	GPR183	-2.678061632	0.042727824
Rasa4	RASA4	-2.681182262	0.045970694
Med12l	MED12L	-2.682537863	0.021537391
Qrfp	QRFP	-2.682808204	0.025977916
Lamb3	LAMB3	-2.690421154	0.006615529
Vopp1	VOPP1	-2.702368431	0.015146727
Fcrl1	FCRL1	-2.702724125	0.000534399
Kcnb1	KCNB1	-2.710041679	0.009572946
Cdkn2b	CDKN2B	-2.726829888	0.013455926
Rhobtb1	RHOBTB1	-2.733898302	0.001412495
Colq	COLQ	-2.739712702	0.040144913
Map3k19	MAP3K19	-2.74347283	0.009268281
Fbxo30	FBXO30	-2.762388674	0.003159707
Ephx1	EPHX1	-2.767023449	0.029985012
Gsap	GSAP	-2.83296523	0.039454491
Rragb	RRAGB	-2.882514429	0.038372398
Ly75	LY75	-2.887045621	0.039651387
Lrrc49	LRRC49	-2.899758627	0.005012567
Magi3	MAGI3	-2.936188333	0.006402884
Tmtc1	TMTC1	-2.970142249	0.045192386
Pik3r3	PIK3R3	-2.975710173	0.038297706
Gm4955		-2.9886528	0.038828657
Klf8	KLF8	-3.033056577	0.030391933
Gbp9	GBP6	-3.04725693	0.019707551
H2-Eb1	HLA-DRB5	-3.057401892	0.033606914
Gimap3		-3.061468767	0.017044543
H2-DMb2	HLA-DMB	-3.075266962	0.049037206
Ankrd6	ANKRD6	-3.083422313	0.048870191
Cmya5		-3.108741403	0.016342476
Cd74	CD74	-3.11523775	0.022495712
Cr2	CR2	-3.125648517	0.033589945
Flnb	FLNB	-3.135553796	0.037159127
Hsd11b1	HSD11B1	-3.143443483	0.002213036
Gng12	GNG12	-3.161065054	0.037087297
Ccdc112	CCDC112	-3.187311336	0.004671744
Inpp1	INPP1	-3.20686442	0.009726156

Scimp	SCIMP	-3.227233933	0.025947222
Rnf125		-3.231542901	0.022495712
Lyz1		-3.250015787	0.04660056
Ppfibp1	PPFIBP1	-3.257240462	0.049052709
Arsb	ARSB	-3.315589987	0.015087986
Bcl6	BCL6	-3.334017056	0.003580665
Klrb1		-3.369535745	0.020902391
Pydc4		-3.372857805	0.010449009
Gpr174	GPR174	-3.378241503	0.012034902
Mfsd6	MFSD6	-3.413786493	0.001858205
St3gal6	ST3GAL6	-3.417953586	0.008578311
Ubash3b	UBASH3B	-3.471829561	0.001592303
Ipcef1	IPCEF1	-3.486086835	0.009531273
Cd63	CD63	-3.528357661	0.021320121
Gbp4		-3.605876805	0.049037206
Bbs7	BBS7	-3.665425066	0.035872246
Hdac9	HDAC9	-3.677006638	0.001264109
Itm2a	ITM2A	-3.683327261	0.004671744
Zfp677		-3.684558398	0.032641874
Eif4e3	EIF4E3	-3.68730399	0.004671744
Axin2	AXIN2	-3.692253002	0.002781193
Hmgn3		-3.719381763	0.003255224
Tbx19	TBX19	-3.727095369	0.02329649
Pik3r6	PIK3R6	-3.729029727	0.012167864
Maoa	MAOA	-3.741783232	0.003255224
Tnfrsf21	TNFRSF21	-3.75073147	0.032315186
Atp1b1	ATP1B1	-3.793364485	0.002081825
Pydc3		-3.831803467	0.003895048
Cybb	CYBB	-3.850358199	0.001297554
Hivep3	HIVEP3	-3.884863546	0.004542176
Cyp4v3	CYP4V2	-3.900221873	0.000715231
Cdkn2a	CDKN2A	-3.927894507	0.04765159
Gria3	GRIA3	-4.105645606	0.048491925
Tnfrsf26		-4.108451347	0.00399835
Ext1	EXT1	-4.179576825	0.007986864
Rgs18	RGS18	-4.196535311	0.021571293
St3gal5	ST3GAL5	-4.218659359	0.028223815
Zc2hc1a	ZC2HC1A	-4.350293064	0.001962669
Gabarapl1	GABARAPL1	-4.391512407	0.02453399
Plcxd2	PLCXD2	-4.52066083	0.010829826
Mylk	MYLK	-4.667055985	0.032126294
Slamf1	SLAMF1	-4.817137054	0.009531273
2810025M15Rik		-4.840750286	0.045161631

Pla2g4a	PLA2G4A	-4.864080998	0.002451835
Pkp4	PKP4	-5.099617514	0.028223815
Rundc3b	RUNDC3B	-5.397705118	0.013973937
Uchl1	UCHL1	-5.524175091	0.004428867
Scn8a	SCN8A	-5.593577641	0.005012567
Zfp2	ZFP2	-5.63488634	0.006459757
Slc30a4	SLC30A4	-5.675321282	0.00526891
Prelid2	PRELID2	-5.986662803	0.022631453
Map9	MAP9	-6.206019969	0.0075295
Npr3	NPR3	-6.704019886	0.008830301

4.2.10 Supervised clustering analysis using *Bcor*^{MUT} gene expression signature

RNA sequencing was also performed on E*μ-Myc*-pQCIG-*Bcor*G2 lymphomas characterised in chapter 3 and 12 sporadic E*μ-Myc* lymphomas, which were genomically defined in chapter 3.2.3 (ML21, ML33, 4242, ML43, ln88, and 6066) or were archival E*μ-Myc* lymphomas that had been generated in the Johnstone Laboratory (X190, X204, X220, X240, X028, X298) with unknown genomic profiles. To test whether *Bcor* loss-of-function gene expression signature drives a reproducible transcriptional pattern in E*μ-Myc* lymphomas, gene expression data using the *Bcor*^{MUT} signature gene set (defined in section 4.2.9) was used to interrogate gene sets derived from E*μ-Myc*-pQCIG-*Bcor*G2 lymphomas and sporadic tumours in which the mutational status of *Bcor* is known or unknown. E*μ-Myc*.pQCIG-*Bcor*G2 lymphomas and sporadic known *Bcor*-mutant lymphomas (ML21, ML33, ML43, ln88 (X88), 299 (X299) and #4242) clustered with the Eμ-*Myc-Bcor*.sh9 tumours indicating a shared gene expression pattern when Bcor is experimentally depleted, deleted or spontaneously lost (Figure 4.11). Eμ-*Myc-Trp53.*1224, Eμ-*Myc-Nras*^{Q61K} and sporadic known-*Bcor*^{WT} E*μ-Myc* lymphomas (6066) did not cluster with the *Bcor*^{MUT} tumours (Figure 4.11). To determine the predictive strength of the *Bcor*^{MUT} gene expression signature sporadic E*μ-Myc* tumours with an unknown *Bcor* mutational status (X190, X204, X220, X240, X028 and X298) were introduced into the supervised clustering analysis (Figure 4-11). Lymphoma X190 was the only sporadic E*μ-Myc* lymphoma with unknown *Bcor*-mutational status that clustered with the *Bcor*^{MUT} tumours. To determine if lymphoma X190 harboured a *Bcor* inactivating mutation, the reads spanning the *Bcor* transcript

were interrogated and compared to E*μ-Myc* lymphomas with known mutational profiles (as defined in section 3.2.3); 4242 and 6066 (Figure 4-12A-D). Lymphoma X190 demonstrated significantly less reads spanning the *Bcor* transcript compared to lymphoma 6066 (*Bcor*^{WT}) and 4242, in which the known small deletion in *Bcor* exon 4 was apparent (Figure 4-12A). Reads spanning the *Atp6ap2* gene, which is proximal to *Bcor* on the X-chromosome, were approximately equal between lymphomas 4242, X190 and 6066, demonstrating that the diminished *Bcor* reads apparent in lymphoma X190 is specific to *Bcor* and is not due to general read depth loss in sample X190 (Figure 4-12B). Reads spanning *Cdkn2a* and *Nras* confirm known mutations in lymphomas 4242 and 6066 and suggest that lymphoma X190 is WT at *Nras* and *Cdkn2a* (Figure 4-12C and Figure 4-12D). Taken together, these results strongly suggested that E*μ-Myc* lymphoma X190 had acquired a *de novo Bcor* deletion, imbuing it with a reproducible and unique *Bcor*^{MUT} gene expression signature.

4.2.11 Pathway analysis of the *Bcor* loss-of-function signature

To determine if the *Bcor* loss-of-function gene signature provided any molecular insight into the functional interaction between *Myc* and *Bcor* and how they can cooperate to drive lymphomagenesis in the E*μ-Myc* model of lymphoma, pathway analysis was performed. Pathway mapping of significant genes (adjusted p-value ≤ 0.05) was undertaken using the PANTHER pathway analysis software tool (Mi and Thomas, 2009). PANTHER pathway analysis identified upregulation of the *Tgfβ* pathway as the most affected pathway given the 393 significantly altered genes apparent when Eμ-*Myc*-*Bcor.*sh9 tumours are contrasted against a group comprised of Eμ-*Myc*-*Trp53.*1224 and

Eμ-*Myc*-*Nras*^{Q61K} (Bonferroni corrected p = 0.0058) (Table 4-4) (Figure 4-13A-F). *Tgfβ* pathway members *Cited1, Bambi, Acvr2b, Smad3, Mapk12 and Tgfb2* were positively enhanced in the comparison between Eμ-*Myc*-*Bcor*.sh9 tumours and the Eμ-*Myc*-*Trp53.*1224 *and* Eμ-*Myc*-*Nras*^{Q61K} tumours (data not shown). When Eμ-*Myc* lymphomas with depletion, deletion or mutation of *Bcor* were compared to Eμ-*Myc* lymphomas with no known perturbation of *Bcor* function, the expression of *Cited1, Avcr2b, Smad3, Mapk12* and *Tgfb2* was significantly enhanced.

Cited1 positively enhances *Tgfβ* signaling through its positive regulation of the Cited1/Smad3/p300Cbp transcriptional co-activator complex and demonstrated significantly enhanced expression in the *Bcor*^{MUT} lymphomas compared to *Bcor*^{WT} (log_2 expression of 4.2 compared to –2.3, paired student's t-test, p < 0.0001) (Derynck et al., 2001; Yahata et al., 2000) (Figure 4-13A). *Acvr2b* is an activin belonging to the *Tgfβ* superfamily of structurally related receptors (Ebner et al., 1993). When activated, Avcvr2 leads to the phosphorylation of Smad2 and Smad3 and potentiates *Tgfβ* signalling (Attisano et al., 1996). *Avcvr2* is expressed at significantly higher levels in *Bcor*^{MUT} lymphomas compared to *Bcor*^{WT} lymphomas with a log_2 expression of 2.4 compared to 0.8 (Paired student's t-test, p < 0.0001) (Figure 4-13C). Smad3 is an intranuclear signal transducer and is transcriptionally activated by *Tgfβ* signaling and is overexpressed in *Bcor*^{MUT} lymphomas compared to *Bcor*^{WT} lymphomas (log_2 expression of 6.1 compared to 5.3, paired student's t-test, p < 0.01) (Zhang et al., 1998) (Figure 4-13D). *Mapk12* log_2 expression in *Bcor*^{MUT} tumours was 4.9, which was significantly greater than the log_2 expression in *Bcor*^{WT} lymphomas (student's t-test, p < 0.0001)

(Figure 4-13E). Mapk12 is known to potentiate *Tgfβ* signaling by inhibiting the BMP-pathway, allowing Smad4 to associate with Smad3 and lead to its accumulation in the nucleus (Massague, 1998). The *Tgfβ* pathway ligand, *Tgfb2*, was more highly expressed in *Bcor*^{MUT} lymphomas compared to *Bcor*^{WT} lymphomas with a $\log_2$ expression of 1.0 and -2.1, respectively (student's t-test, $p < 0.001$) (Figure 4-13F).

Table 4-4 Pathways affected by *Bcor* loss of function using PANTHER

PANTHER pathway	RELIST	Input	Input (fold enrichment)	Input (adjusted p-value)
Tgfβ signaling pathway	93	6	>5	$3.30E^{-5}$
Interferon γ signaling pathway	31	3	>5	$1.08E^{-3}$
Arginine biosynthesis	8	2	>5	$1.24E^{-3}$
Axon guidance mediated by Slit/Robo	24	2	>5	$1.05E^{-2}$

4.3 Discussion

BCOR loss-of-function mutations have been identified in both haematological and solid human malignancies (Damm et al., 2013; Grossmann et al., 2011; Pugh et al., 2012). Vogelstein *et al* classifies *BCOR* as a tumour suppressor as 70% of *BCOR* mutations spanning 21 mutated tumour types, including both solid and liquid cancers, were truncating mutations with no common missense mutations or identical in-frame InDels (Vogelstein et al., 2013). The data presented in chapter 3 identifies *Bcor* loss-of-function mutations as the most frequent mutations in E*μ-Myc* lymphoma. To date, the tumour suppressive capabilities of Bcor have never been validated. This study is the first to functionally validate *Bcor* as a potent tumour suppressor in *Myc*-driven lymphoma and demonstrated that lymphomas with *Bcor* loss are immature B cells with distinct a gene expression profile.

4.3.1 Bcor-inactivation cooperates with Myc to drive E*μ-Myc* lymphomagenesis

RNA-mediated knockdown of *Bcor* significantly accelerated lymphomagenesis in the E*μ-Myc* model and was shown to be efficient by protein immunoblotting. Similarly, CRISPR/Cas9 mediated targeting of Bcor resulted in Bcor-protein truncations that significantly accelerated E*μ-Myc* lymphomagenesis. These experiments demonstrated for the first time that *Bcor* functions as a tumour suppressor gene and cooperates with oncogenic Myc in driving lymphomagenesis. Complementing this critical finding was the demonstration that forced ectopic expression of WT *Bcor* in E*μ-Myc* cells in which

Bcor is mutated and human NAMALWA cells in which *BCOR* is weakly expressed resulted in a competitive proliferative disadvantage to these cells. In this setting, Bcor re-expression is likely to be disrupting at least one of the “hallmarks of cancer”, such as resistance to cell death, sustained proliferative signaling, growth suppressor evasion and replicative immortality (Hanahan and Weinberg, 2011).

BCOR is able to bind to BCL6 through the BTB/POZ domain, localising the BCL6-BCOR complex to BCL6 target genes and potentiating BCL6-mediated gene repression (Gearhart et al., 2006). BCL6 functions as a potent oncogene in post-GC B cell malignancies suggesting that co-repressors augmenting the function of BCL6, such as BCOR, are also oncogenic in this context (Hatzi and Melnick, 2014). However, this did not align with the data presented herein demonstrating a tumor suppressor role for Bcor. E*μ-Myc* lymphomas are pre-GC B cell lymphomas that express Bcl6 at relatively low levels (Figure 4-14) (Mori et al., 2008), suggesting that in this context Bcor functions independently of Bcl6. In the absence of Bcl6, Bcor is known to form a PRC1-like complex to mediate transcriptional repression of target genes (Choi et al., 2013). Gene expression profiling of *Bcor*^{MUT} E*μ-Myc* lymphomas revealed significantly less Bcl6 expression compared to E*μ-Myc* lymphomas harbouring other common sporadic mutations (Table 4-3), further highlighting that Bcor is likely to function independently of Bcl6 in this context. To fully understand how Bcor acts as a tumour suppressor, the profile of genes that Bcor represses independently of Bcl6 must be defined. In normal myeloid cells, Bcor restricts proliferation and differentiation through targeted repression of *Hox* genes (Cao et al., 2016). *Bcor* deletion in both normal bone marrow myeloid

cells and human AML cell lines resulted in increased proliferation, profound myeloid cell differentiation and reduced protein levels of other PRC1 complex components (Cao et al., 2016). The differentially expressed genes that are apparent in *Bcor*-deleted myeloid cells are not significantly altered in *Bcor*^{MUT} E*μ-Myc* lymphomas compared *to Bcor*^{WT} E*μ-Myc* lymphomas harbouring other common sporadic mutations. This indicated that the repressive profile and tumour suppressive mechanisms of Bcor are likely cell-type dependent.

4.3.2 Bcor-mutant E*μ-Myc* lymphomas are of immature B cell origin

90% of *Bcor*^{MUT} sporadic E*μ-Myc* tumours and E*μ-Myc-Bcor*.sh9 lymphomas presented as a dominant B cell clone with an IgM$^-$/IgD$^-$ surface profile. Conversely, *Bcor*^{WT} sporadic tumours and E*μ-Myc*.p53.1224 lymphomas presented equally as IgM$^-$/IgD$^-$ or IgM$^+$/IgD$^-$ lymphomas. Dual negative surface expression of IgM and IgD correlates to the surface Ig expression profile of pre-pro-B cells (CD19$^-$), pro-B cells and pre-B cells (Allman et al., 1999; Hardy et al., 1991; Li et al., 1996; Loffert et al., 1994). The predictable IgM$^-$/IgD$^-$ surface expression profile in *Eμ-Myc* lymphomas that harboured *Bcor* inactivating mutations or genetic depletions indicated that *Bcor* loss drives lymphomagenesis in pro-B cells or pre-B cells. These data further strengthen the argument that Bcor acts independently of Bcl6 in this setting, as pro-B cells or pre-B cells demonstrate relatively low Bcl6 expression (Figure 4-14). This raised the question as to whether there is a role for Bcor in normal B-lymphocyte development. 90-100% of mature, circulating B cells in patients with OFCD, a disease driven by germline *BCOR* truncations, demonstrate epigenetic silencing of the X-chromosome carrying the

mutated allele (males with germline *BCOR* mutations do not exist, and are thought to perish early in development) (Ng et al., 2004; Wamstad and Bardwell, 2007). This indicates that germline truncating *BCOR* mutations confer a selective disadvantage in normal B cells and suggests that BCOR contributes in some way to the development of normal mature, circulating B-lymphocytes. Further experiments are required to identify and dissect the mechanism through which *Bcor* acts both as a tumour suppressor gene in B cells and how it contributes to normal B cell development. To this end, the Johnstone Laboratory has attained a transgenic mouse strain with a conditional *Bcor* allele, $Bcor^{FL}$ to further interrogate the involvement of Bcor in B cell development.

4.3.3 Bcor inactivation drives lymphomagenesis via a distinct pathway

Bcor loss-of-function, either by sporadic inactivating mutation, depletion or deletion drives a reproducible and unique gene expression profile in E*μ-Myc* lymphoma. Given that *BCOR*-loss in MDS and related disorders is associated with poor prognosis (Damm et al., 2013), this unique gene expression profile has potential for clinical utility as a predictor or patient outcome. Interestingly, the $Bcor^{MUT}$ gene expression signature is distinct from the gene expression signature obtained through loss of *Trp53* in E*μ-Myc* lymphomas. Disruption of the $p19^{ARF}$-Mdm2-Trp53 tumour suppressor axis was proposed as a major contributor to Myc-driven malignancy, with approximately half of sporadic E*μ-Myc* lymphomas harbouring inactivating mutations or deletions in *Trp53* or *Cdkn2a* (encoding $p19^{ARF}$) (Eischen et al., 1999; Schmitt et al., 1999). Furthermore, data in chapter 3 demonstrated the co-occurrence of *Cdkn2a* deletions and *Bcor*

inactivating mutations in clonal malignancies, suggesting that *Bcor* and *Cdkn2a* act independently of each other. The gene expression profile of E*μ-Myc-Bcor*.sh9 lymphomas was distinct to that of E*μ-Myc-Trp53*.1224 and E*μ-Myc-Nras*^{Q61K} lymphomas, indicating that the pathway to Myc-driven lymphomagenesis that is taken by these oncogenic insults diverges. *Bcor* was identified as the most frequently mutated gene in E*μ-Myc* lymphomas in chapter 3, and has been functionally validated as a potent tumour suppressor gene in chapter 3, suggesting that the importance of the p19^{ARF}-Mdm2-Trp53 pathway is likely overstated in E*μ-Myc* lymphomagenesis.

PANTHER pathway analysis works by comparing an experimental gene expression profile to a suite of already published gene expression profiles, thereby limiting the analysis to what has been previously published. Analysis of the unique gene expression signature associated with *Bcor*-loss revealed upregulation of genes involved in *Tgfβ* signaling as the solitary significantly differentially expressed pathway. The importance of the *Tgfβ* signaling pathway in E*μ-Myc* lymphomagenesis has been previously reported, where tumour microenvironment macrophage-derived cytokines enhanced the OIS pathway (Reimann et al., 2010). In that study it was demonstrated that in E*μ-Myc* lymphomas, knockout of the senescence-associated histone methyltransferase *Suv39h1* significantly accelerated MYC-driven tumourigenesis. The resultant lymphomas possessed an augmented *Tgfβ* signature suggestive of the same signature observed in the *Bcor*^{MUT} E*μ-Myc* lymphomas profiled herein. In E*μ-Myc* lymphoma it is proposed that *Tgfβ* drives *Myc*-driven senescence in a *Suv39h1*-dependent manner raising the interesting potential that *Bcor* regulates this unique

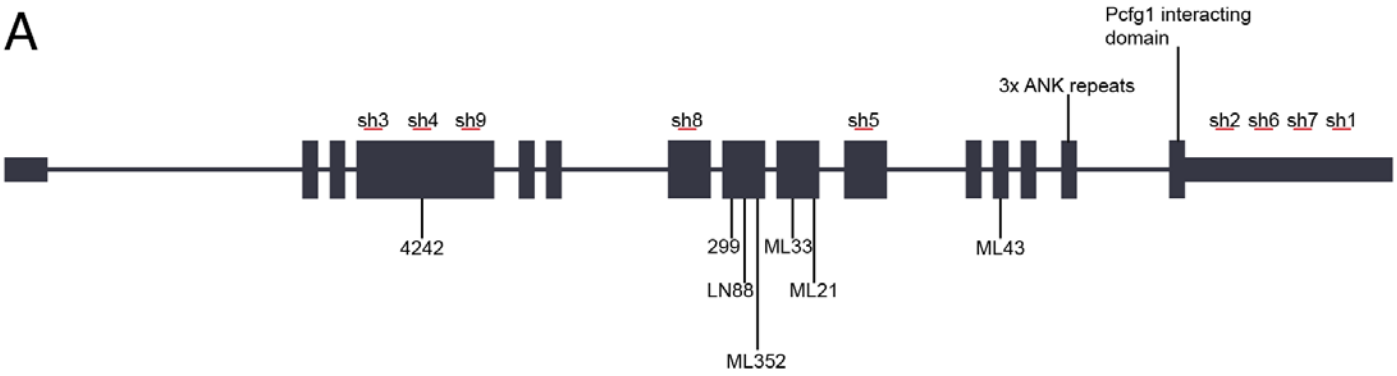
tumour suppressive mechanism in a similar mode. Further experiments are required to delineate the interaction *Bcor* has with the *Tgfβ* signaling pathway and how, through this pathway, *Bcor* acts as a potent tumour suppressor gene.

4.3.4 Chapter conclusions

Results presented in this chapter are the first to validate *Bcor* as a *bona fide* tumour suppressor gene (this has since been published (Lefebure et al., 2017)). Bcor has been demonstrated to potently cooperate with Myc overexpression in the E*μ-Myc* model of lymphoma, therefore, further dissection of the tumour suppressive pathway in which Bcor acts may yield novel therapeutic targets capable of being exploited in Myc-driven haematological malignancies. Currently, further projects are underway in the Johnstone Laboratory aimed at identifying the exact mechanism through which Bcor restrains the oncogenic effects of Myc overexpression.

4.4 Chapter 4 figures and figure legends

The corresponding figure legend can be found on the page following each figure.


A
Pcfg1 interacting domain
3x ANK repeats
sh3 sh4 sh9
sh8
sh5
sh2 sh6 sh7 sh1
4242
299
ML33
ML43
LN88
ML21
ML352

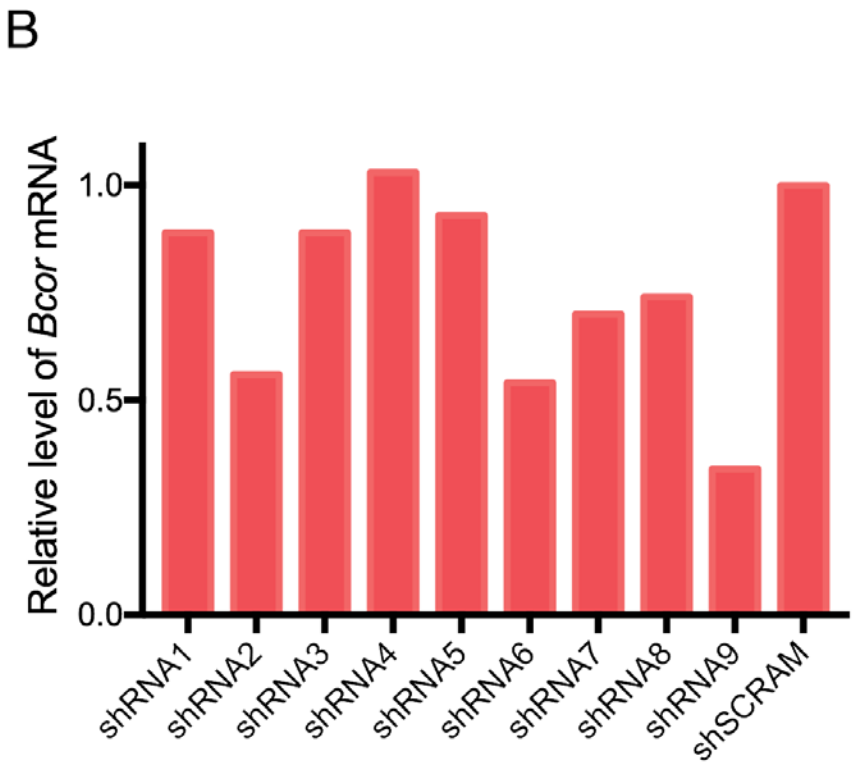

B
Relative level of Bcor mRNA
1.0
0.5
0.0
shRNA1
shRNA2
shRNA3
shRNA4
shRNA5
shRNA6
shRNA7
shRNA8
shRNA9
shSCRAM

Figure 4.01 – *Bcor*-targeting shRNA design and testing

A) Schematic of the *Bcor* transcript showing location of shRNA (sh1-to-sh9), the location of the *Bcor* mutations identified in the retrospective and prospective E*μ-Myc* cohorts, as defined in chapter 3 and the 3 ankyring repeasts and Pcfg1 interacting domain.
B) Each hairpin was tested by transfecting 3T3 cells and then assessing the levels of *Bcor* transcript by QRT-PCR. shRNA9 demonstrated the most robust decrease in *Bcor* transcript compared to the control group (0.34 compared to shRNASCR).

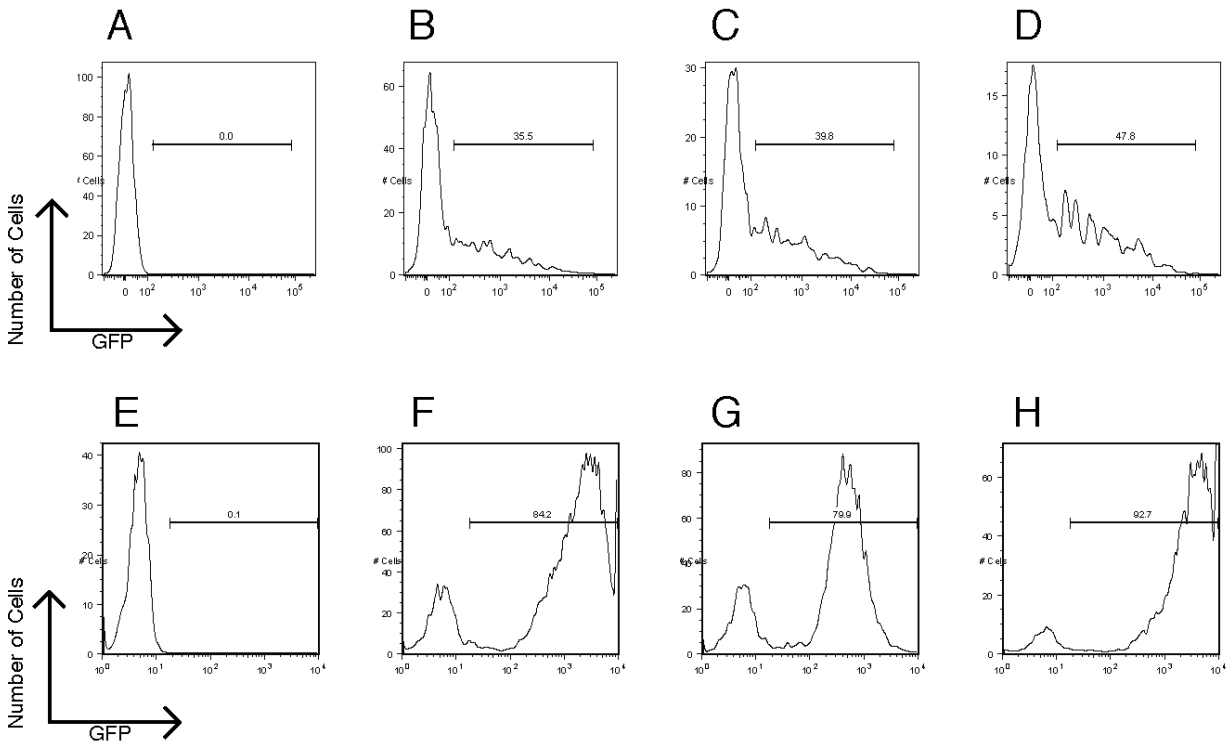

A
B
C
D
E
F
G
H
0.0
35.5
39.8
47.8
0.1
84.2
79.9
92.7
Cells
Number of Cells
GFP

Figure 4.02 – Assessing transfection and transduction efficiency in phoenix-E cells and E*μ-Myc* foetal liver cells, respectively

Mock-transfected, pLMS-*Bcor*.sh9-transfected, pLMS-Scram-transfected and pLMS-*Trp53*.1224-transfected phoenix-E cells were analysed by flow cytometry for GFP expression and are represented by panels A, B, C and D, respectively. Flow cytometry was performed 24h-post transfection. Non-transduced, E*μ-Myc-Bcor*.sh9, E*μ-Myc*-Scram, or E*μ-Myc-Trp53*.1224 foetal liver cells are represented by panels E, F, G and H, respectively. The foetal liver cells were analysed by flow cytometry 72h post-second viral hit for GFP expression. Analysis was performed on at least 10,000, viable single cells.

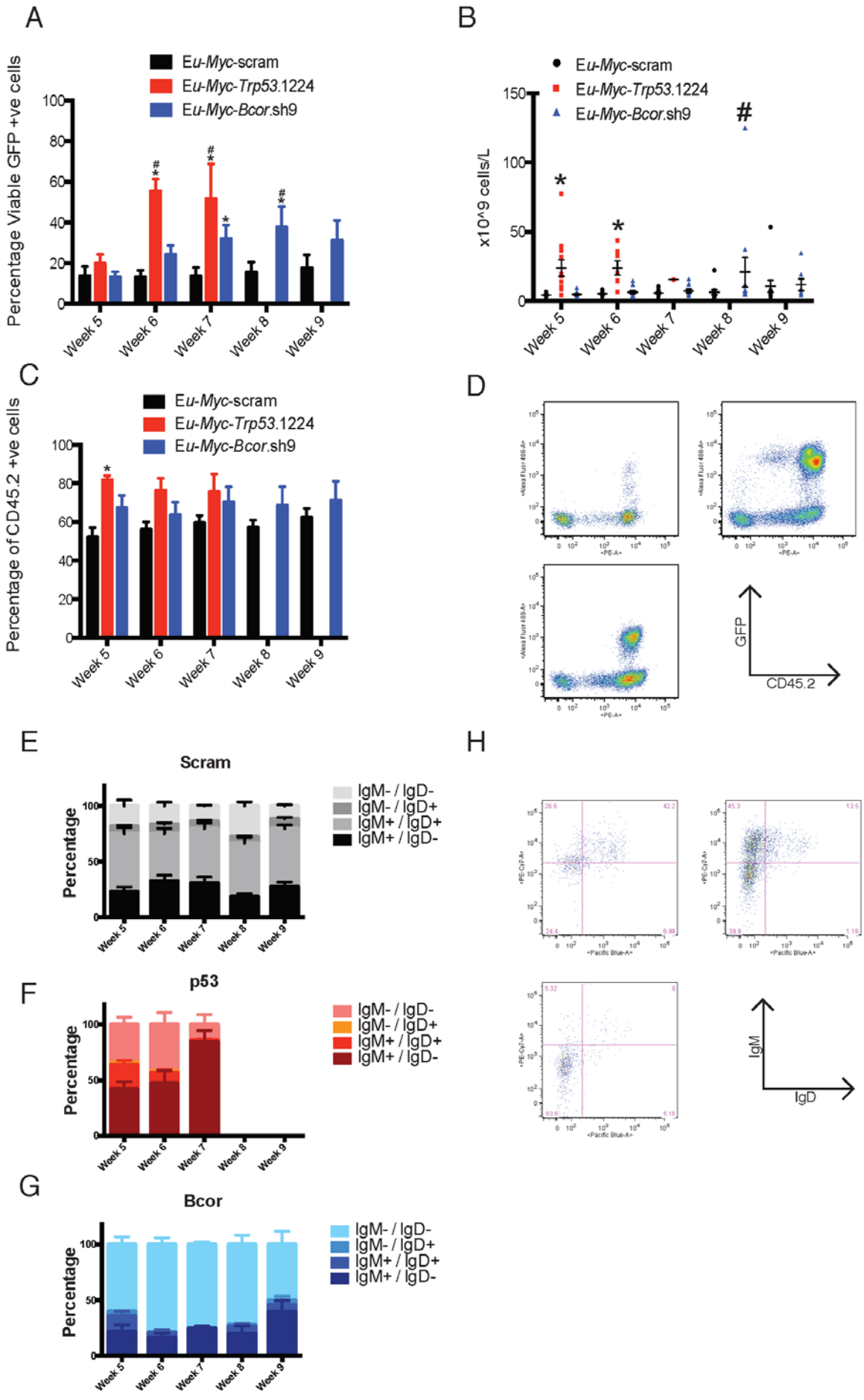

A
Eu-Myc-scram
Eu-Myc-Trp53.1224
Eu-Myc-Bcor.sh9
Percentage Viable GFP +ve cells
Week 5
Week 6
Week 7
Week 8
Week 9
B
x10^9 cells/L
C
Percentage of CD45.2 +ve cells
D
GFP
CD45.2
E
Scram
Percentage
IgM- / IgD-
IgM- / IgD+
IgM+ / IgD+
IgM+ / IgD-
F
p53
G
Bcor
H
IgM
IgD

Figure 4.03 – Temporal analysis of circulating peripheral blood B cells in mice that had received Eμ-*Myc*-*Bcor*.sh9, Eμ-*Myc*-*Trp53*.1224 or Eμ-*Myc*-scram foetal liver cells

A) Temporal analysis of GFP expression in circulating peripheral blood B cells of mice that had received transduced foetal liver cells. Total GFP positivity was calculated as a percentage of morphologically viable, single cells. Recipient mice transplanted with Eμ-*Myc*-*Trp53*.1224 foetal liver cells demonstrated rapid increase in GFP expression over time, whereas mice transplanted with Eμ-*Myc*-*Bcor*.sh9 foetal liver cells demonstrated a more gradual increase in GFP positivity. Mice that received Eμ-*Myc*-scram demonstrated no temporal shifts in GFP expression. Data are mean ± SEM, with * = $p < 0.05$ compared to GFP expression in recipients of Eμ-*Myc*-scram at that time point and # = $p < 0.05$ compared to that groups GFP expression at week 5 (2-way ANOVA).

B) Automated WBC counts of mice transplanted with Eμ-*Myc*-scram, Eμ-*Myc*-*Trp53*.1224 or Eμ-*Myc*-*Bcor*.sh9 foetal liver cells. Recipients of Eμ-*Myc*-*Trp53*.1224 demonstrated rapid increase in total WBC count with Eμ-*Myc*-*Bcor*.sh9-transplanted mice progressing at a slower rate. Mice that received Eμ-*Myc*-scram showing no temporal increase in WBC count. Data are mean ± SEM, with * = $p < 0.05$ compared to GFP expression in recipients of Eμ-*Myc*-scram at that time point and # = $p < 0.05$ compared to that groups GFP expression at week 5 (2-way ANOVA).

C) Temporal analysis of CD45.2 expression in mice transplanted with transduced foetal liver cells. CD45.2 positivity was calculated as a percentage of morphologically viable, single cells. Cohorts transplanted with Eμ-*Myc*-scram or Eμ-*Myc*-*Bcor*.sh9 foetal liver cells demonstrated no change in CD45.2 positivity from 5 weeks post-transplant to 9 weeks post-transplant. At five weeks post-transplant, mice that received Eμ-*Myc*-*Trp53*.1224 foetal liver cells displayed increased amounts of donor cells compared to control. Data are mean ± SEM with * = $p < 0.05$ compared to GFP expression in recipients of Eμ-*Myc*-scram at that time point (2-way ANOVA).

D) Representative flow cytometry plots demonstrating CD45.2 and GFP expression in circulating peripheral blood B cells of mice transplanted with transduced foetal liver cells. Mice transplanted with Eμ-*Myc*-scram, Eμ-*Myc*-*Trp53*.1224 or Eμ-*Myc*-*Bcor*.sh9 foetal liver cells are represented by the top left, top right and bottom left panels, respectively. Cells are gated as morphologically viable, single cells.

E) Temporal changes in surface IgM and IgD on circulating peripheral blood B cells in mice transplanted with Eμ-*Myc*-scram foetal liver cells. Eμ-*Myc*-scram-transplanted mice demonstrated consistent surface IgM and IgD expression on their circulating B cells, with the IgM+/IgD+ population being the most represented population, making up, on average, 43.0% to 52.5% of the population of B cells across each time point followed by IgM+/IgD- (13.1%-32.1%), IgM-/IgD- (11.3%-27.5%) and IgM-/IgD+ (5.8%-9.2%) populations. Data are mean ± SEM for each time point.

F) Temporal changes in surface IgM and IgD on circulating peripheral blood B cells in mice transplanted with Eμ-*Myc*-*Trp53*.1224 foetal liver cells. Mice transplanted with Eμ-*Myc*-p53.1224 demonstrated consistent surface IgM and IgD expression on the on circulating B cells until week 7 where a limited number of experimental animals remained alive. B cells with a surface immunophenotype consisting of IgM+/IgD- or IgM-/IgD- were the most represented populations across the time points measured (42.0%-84.3% and 14.0%-33.8, respectively) followed by IgM+/IgD+ (1.6%-21.7%) and IgM-/IgD+ (0.2%-2.6%) populations. Data are mean ± SEM for each time point.

G) Surface IgM and IgD expression on circulating B cells in mice transplanted with Eμ-*Myc*-*Bcor*.sh9 remained consistent with IgM-/IgD- being the most represented population (50.2%-78.7%) followed by IgM+/IgD- (15.9%-39.2%), IgM+/IgD+ (0.1%-13.8%) and IgM-/IgD+ (0.2%-4.4%) populations. Data are mean ± SEM for each time point.

H) Representative flow cytometry plot displaying IgD and IgM expression in mice transplanted with Eμ-*Myc*-scram (top left), Eμ-*Myc*-*Trp53*.1224 (top right) or Eμ-*Myc*-*Bcor*.sh9 (bottom left). Analysis was performed on morphologically viable, single cells.

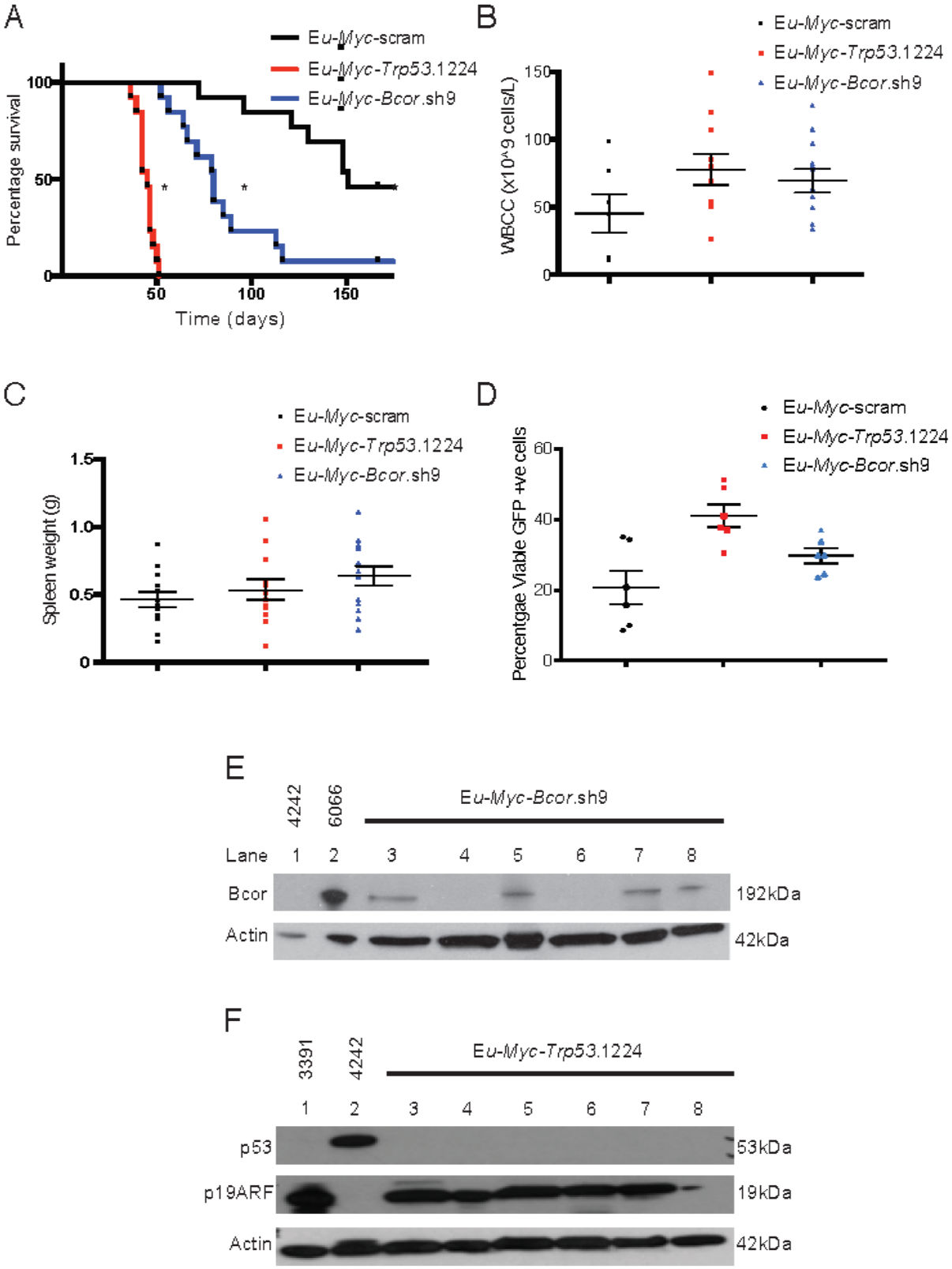

A
Eu-Myc-scram
Eu-Myc-Trp53.1224
Eu-Myc-Bcor.sh9
Percentage survival
100
50
0
50
100
150
Time (days)
*
*
B
Eu-Myc-scram
Eu-Myc-Trp53.1224
Eu-Myc-Bcor.sh9
WBCC (X10^9 cells/L)
150
100
50
0
C
Eu-Myc-scram
Eu-Myc-Trp53.1224
Eu-Myc-Bcor.sh9
Spleen weight (g)
1.5
1.0
0.5
0
D
Eu-Myc-scram
Eu-Myc-Trp53.1224
Eu-Myc-Bcor.sh9
Percentgae Viable GFP +ve cells
60
40
20
0
E
4242
6066
Eu-Myc-Bcor.sh9
Lane 1 2 3 4 5 6 7 8
Bcor
192kDa
Actin
42kDa
F
3391
4242
Eu-Myc-Trp53.1224
1 2 3 4 5 6 7 8
p53
53kDa
p19ARF
19kDa
Actin
42kDa

Figure 4.04 – Disease latency and presentation in mice transplanted with E*μ-Myc-Bcor*.sh9, E*μ-Myc-Trp53*.1224 or Eμ-*Myc*-scram foetal liver cells.

A) Kaplan-Meier curve showing survival of mice transplanted with transduced E*μ-Myc*-Scram (black line, n=13), E*μ-Myc-Trp53.*1224 (red line, n=12) or E*μ-Myc-Bcor*.sh9 (blue line, n=13) foetal liver cells. Mice transplanted with E*μ-Myc-Bcor*.sh9 (median survival time of 80 days) and E*μ-Myc-Trp53.*1224 (median survival time of 45 days) foetal liver cells demonstrated significantly accelerated lymphomagenesis compared to mice transplanted with E*μ-Myc*-Scram (median survival time of 151 days). * = $p < 0.05$ compared to the control group (log-rank Mantel-Cox test).

B) Terminal WBC counts in mice bearing either E*μ-Myc-Bcor*.sh9 (blue, mean WBC count of 73.82×10^9 cells/L, n=11), E*μ-Myc-Trp53.*1224 (red, mean WBC count of 82.04×10^9 cells/L, n=10) or E*μ-Myc*-Scram (black WBC count of 49.57×10^9 cells/L, n=6) foetal liver cells. The cohorts of recipient mice were not significantly distinct in terms of WBC counts ($p > 0.05$, 1-way ANOVA). Data represents individual mice with the groups mean ± SEM.

C) Spleen weight in mice bearing either E*μ-Myc-Bcor*.sh9 (blue, mean spleen weight of 0.64g, n=12), E*μ-Myc-Trp53.*1224 (red, mean spleen weight of 0.54g, n=12) or E*μ-Myc*-Scram (black, mean spleen weight of 0.46g, n=12). The cohorts of recipient mice did not differ significantly in terms of spleen weights ($p > 0.05$, 1-way ANOVA). Data represents individual mice with the groups mean ± SEM.

D) Percentage of GFP positive splenic B cells in mice transplanted with transduced E*μ-Myc* foetal liver cells. 20.78%, 41.07% and 29.75% of splenic B cells in mice transplanted with E*μ-Myc*-Scram (black), E*μ-Myc-Trp53.*1224 (red) or E*μ-Myc-Bcor.*sh9 (blue), respectively, were GFP positive. Data shown for six representative mice from each cohort and the calculated mean ± SEM.

E) Protein immunoblot demonstrating levels of Bcor in six representative E*μ-Myc-Bcor.*sh9 foetal liver cell-derived tumours compared to lymphoma cell lines 4242 (*Bcor*^{MUT}) and 6066 (*Bcor*^{WT}). Western blot analysis was performed with antibodies specific to BCOR. Equivalent protein loading was confirmed by probing for β-Actin.

F) Protein immunoblot demonstrating levels of Trp53 and $p19^{ARF}$ protein in six representative E*μ-Myc-Trp53.*1224 foetal liver cell-derived tumours compared to 4242 (*Trp53*^{WT}, $p19^{ARF}$ deleted) and lymphoma 3391 (*Trp53*^{MUT}, $p19^{ARF\ WT}$). Western blot analysis was performed with antibodies specific to Trp53 and $p19^{ARF}$, respectively. Equivalent protein loading was confirmed by probing for β-Actin.

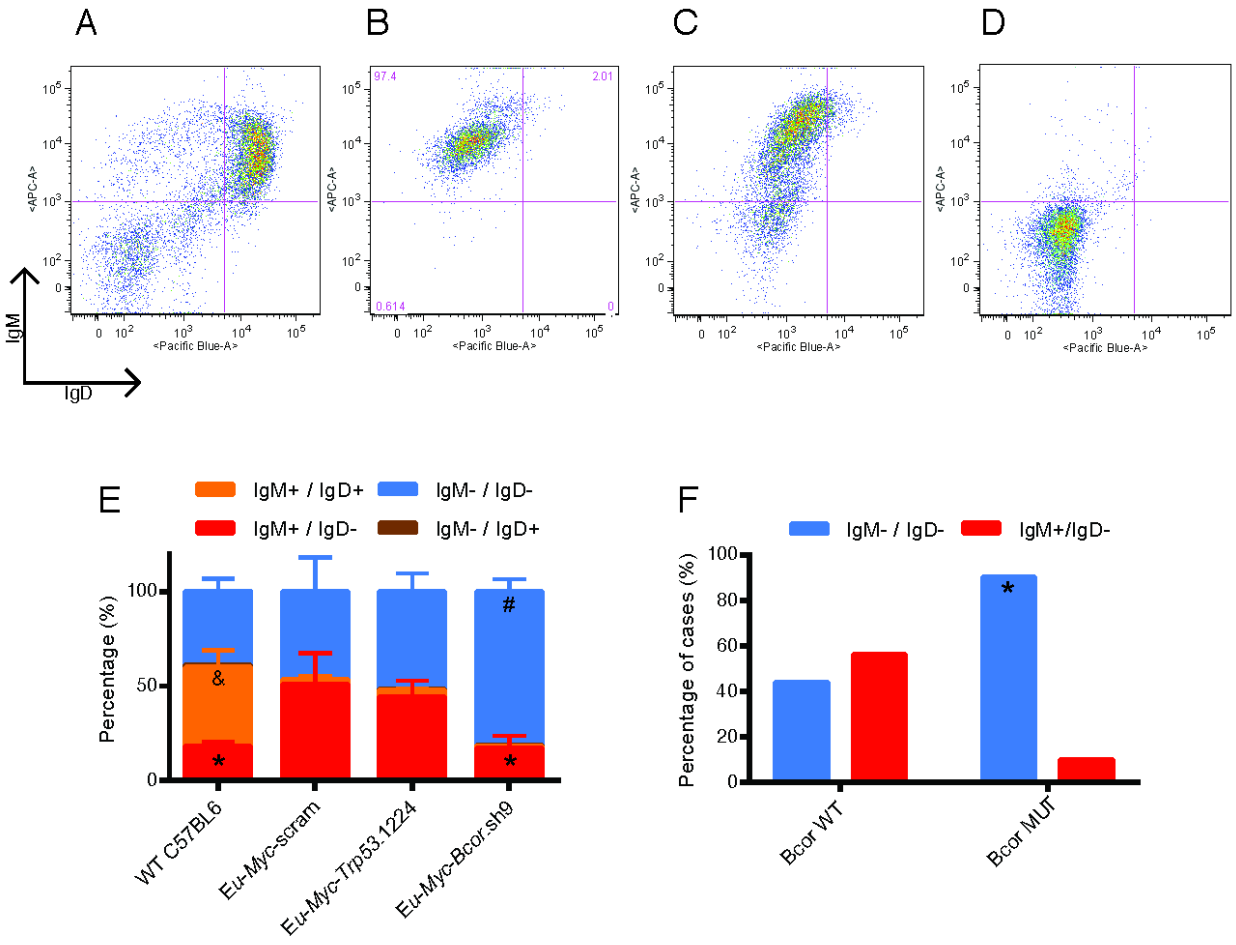
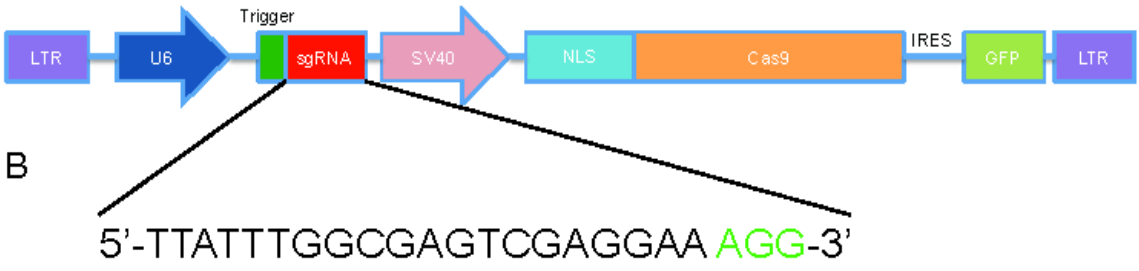

A
B
C
D
IgM
IgD
<APC-A>
<Pacific Blue-A>
97.4
2.01
0.614
0
E
IgM+ / IgD+
IgM- / IgD-
IgM+ / IgD-
IgM- / IgD+
Percentage (%)
100
50
0
&
*
#
*
WT C57BL6
Eu-Myc-scram
Eu-Myc-Trp53.1224
Eu-Myc-Bcor.sh9
F
IgM- / IgD-
IgM+/IgD-
Percentage of cases (%)
100
80
60
40
20
0
*
Bcor WT
Bcor MUT

Figure 4.05 – Characterisation of surface immunoglobulin profile at end stage in E*μ-Myc-Bcor*.sh9, E*μ-Myc-Trp53*.1224 or Eμ-*Myc*-scram foetal liver cell-derived lymphomas

A) Representative flow cytometry plot demonstrating surface IgD and IgM expression in B cell splenocytes derived from non-transplanted WT C57BL/6 mice. Cells are gated as morphologically viable, single cells.
B) Representative flow cytometry plot demonstrating surface IgD and IgM expression in B cell splenocytes derived from moribund mice that were transplanted with E*μ-Myc*-Scram. Cells are gated as morphologically viable, single cells.
C) Representative flow cytometry plot demonstrating surface IgD and IgM expression in B cell splenocytes from moribund mice that were transplanted with E*μ-Myc-Trp53.*1224. Cells are gated as morphologically viable, single cells.
D) Representative flow cytometry plot demonstrating surface IgD and IgM expression in B cell splenocytes from moribund mice that were transplanted with E*μ-Myc-Bcor*.sh9. Cells are gated as morphologically viable, single cells.
E) Quantified surface IgD and IgM expression on B cell splenocytes from mice transplanted with E*μ-Myc*-Scram E*μ-Myc-Trp53.*1224 or E*μ-Myc-Bcor*.sh9 foetal liver cells and WT C57BL/6 splenocytes. E*μ-Myc-Bcor*.sh9 foetal liver-derived lymphomas demonstrated significantly larger populations of splenic B cells with IgM^-/IgD^- phenotype compared to the other three groups. WT C57BL/6 splenocytes and E*μ-Myc-Bcor*.sh9 foetal liver-derived lymphomas demonstrated a significantly larger population of B cells with an IgM^+/IgD^- phenotype compared to E*μ-Myc*-Scram and E*μ-Myc-Trp53.*1224 foetal liver cell-derived lymphomas. WT C57BL/6 splenocytes were the only group that demonstrated a significant population of IgM^+/IgD^+ B cells. Data are mean ± SEM, with * = $p < 0.05$ compared to GFP expression in recipients of E*μ-Myc*-Scram and E*μ-Myc-Trp53.*1224 foetal liver cells, # = $p < 0.05$ compared to C57BL/6 B cells, and recipients of E*μ-Myc*-Scram and E*μ-Myc-Trp53.*1224 foetal liver cells, and & = $p < 0.05$ compared to recipients of E*μ-Myc*-Scram, E*μ-Myc-Trp53.*1224 and E*μ-Myc-Bcor*.sh9 foetal liver cells (2-way ANOVA).
F) Surface IgM and IgD on B cells were assessed in the prospective E*μ-Myc* cohort (four *Bcor* mutant tumours and 12 *Bcor* WT tumours as defined in chapter 3) and 12 foetal liver derived E*μ-Myc* tumours (six *Bcor* mutant tumours (*Bcor* KD) and six *Bcor* WT tumours (p53 KD)). 90% of lymphomas with inactivated *Bcor* presented as IgM^-/IgD^- B cell lymphomas while 10% presented as IgM^+/IgD^- B cell lymphomas. 44.4% of *Bcor* WT tumours presented as IgM^-/IgD^- B cell lymphomas while 56.6% presented as IgM^+/IgD^- B cell lymphomas. *Bcor* mutations are associated with an IgM^-/IgD^- profile (Chi-squared test, $p < 0.05$).

A



B

5'-TTATTTGGCGAGTCGAGGAA AGG-3'

C

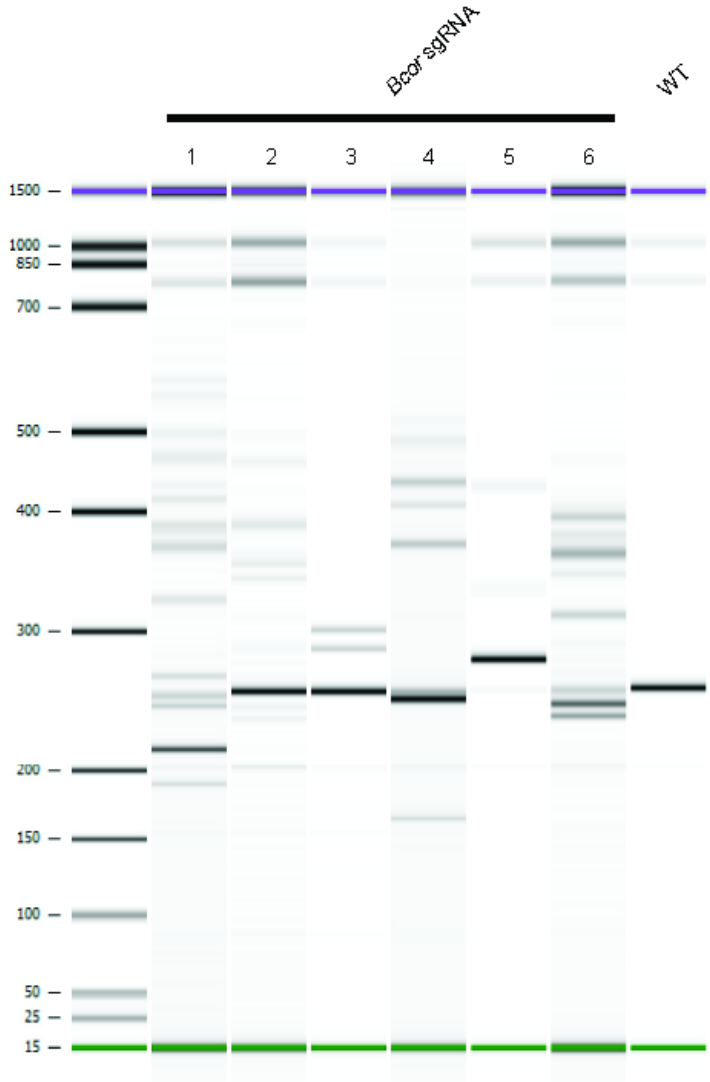


Figure 4.06 – Cas9 mediated editing of *Bcor* in transfected 3T3 cells.

A) Schematic of the pQ-based retroviral construct driving expression of Cas9, GFP and the *Bcor*-targeted guide RNA (sgRNA). The retroviral construct was kindly provided as a gift from Abba Malina and Jerry Pelletier.
B) Sequence of the 20-nucleotide trigger that formed part of the sgRNA, with the green bases representing the protospacer-adjacent motif (PAM). This trigger is complementary to the *Bcor* transcript from nucleotide 1388-1410, which is located in exon 4.
C) DNA chip readout of 3T3 cells that were transfected with pQCIG-*Bcor*G2 construct. DNA was extracted from the cells and the region in *Bcor* exon 4 where the sgRNA binds to was amplified by PCR. The PCR products were analysed on a DNA-bioanalyser with a digital readout shown here. Lanes 1, 2, 3, 4 and 6 demonstrated reductions in size of the product compared to WT. Lane 5 demonstrated an increase in product size compared to WT DNA.

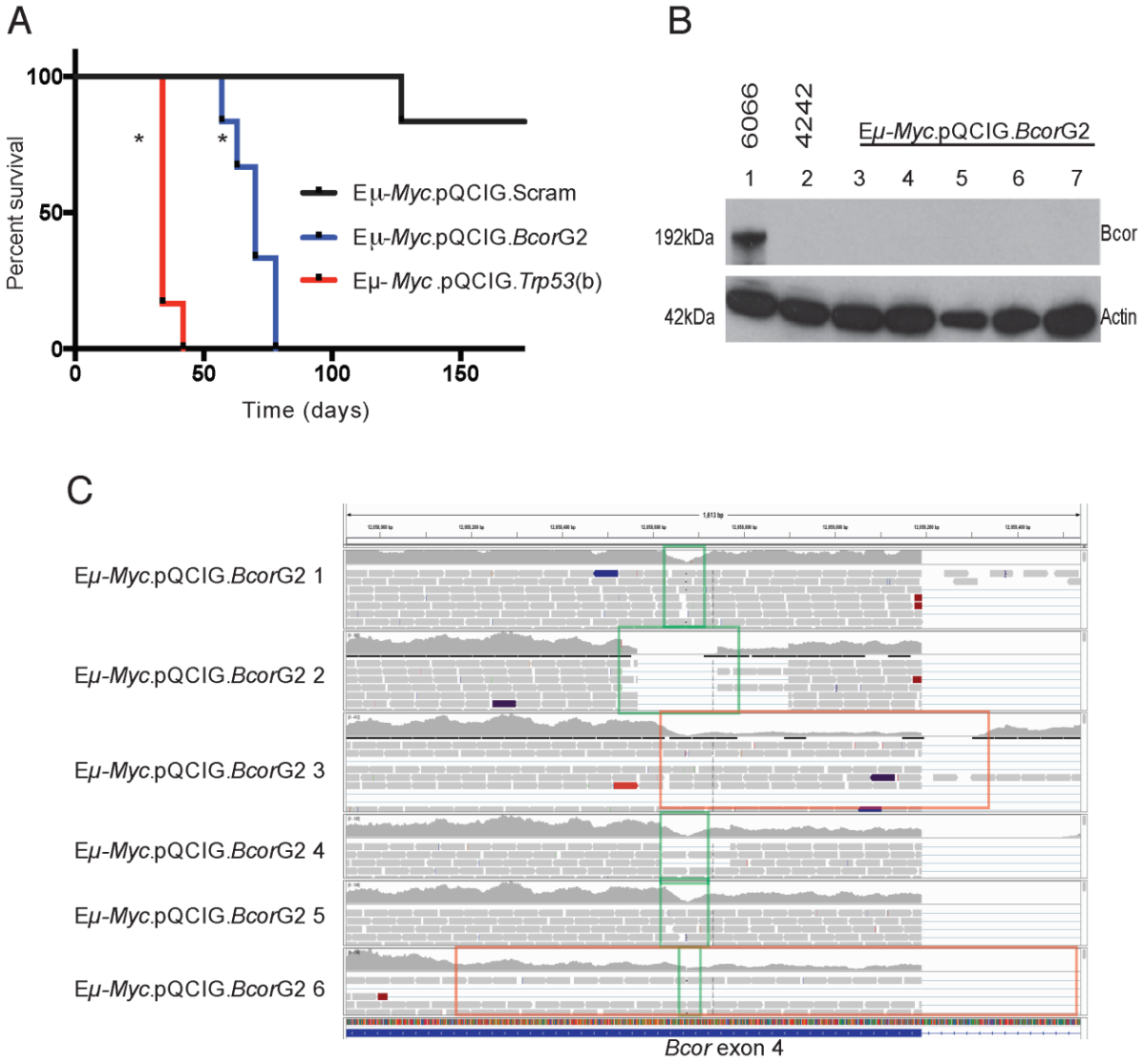

A
100
50
0
Percent survival
0
50
100
150
Time (days)
*
*
Eμ-Myc.pQCIG.Scram
Eμ-Myc.pQCIG.BcorG2
Eμ-Myc.pQCIG.Trp53(b)
B
6066
4242
Eμ-Myc.pQCIG.BcorG2
1
2
3
4
5
6
7
192kDa
Bcor
42kDa
Actin
C
Eμ-Myc.pQCIG.BcorG2 1
Eμ-Myc.pQCIG.BcorG2 2
Eμ-Myc.pQCIG.BcorG2 3
Eμ-Myc.pQCIG.BcorG2 4
Eμ-Myc.pQCIG.BcorG2 5
Eμ-Myc.pQCIG.BcorG2 6
Bcor exon 4

Figure 4.07 – Monitoring disease latency and efficiency of targeted-genomic-editing in mice transplanted with E*μ-Myc*-pQCIG.*Bcor*G2, E*μ-Myc*-pQCIG.*Trp53*(b) or E*μ-Myc*-pQCIG.Scram

A) Kaplan-Meier survival curve demonstrating terminal disease latency in mice transplanted with E*μ-Myc*-pQCIG.*S*cram (black line), E*μ-Myc*-pQCIG.*Bcor*G2 (blue line) or E*μ-Myc*-pQCIG.*Trp53*(b) (red line) foetal liver cells. Mice transplanted with E*μ-Myc*-pQCIG.*Bcor*G2 or E*μ-Myc*-pQCIG.*Trp53*(b) foetal liver cells demonstrated a median survival time of 70 or 34 days post-transplant, respectively. Mice that received E*μ-Myc*-pQCIG.Scram did not reach median survival time. n=6 for each cohort, * = $p < 0.05$ (log-rank Mantel-Cox test) compared to control group.

B) Immunoblot showing levels of Bcor protein in five representative E*μ-Myc*-pQCIG.*Bcor*G2 foetal liver-derived tumours compared to 4242 (*Bcor*^{MUT}) and 6066 (*Bcor*^{WT}). Western blot analysis was performed with antibodies specific to BCOR. Equivalent protein loading was confirmed by probing for β-Actin.

C) RNA-sequencing analysis of reads in IGV spanning exon 4 of the *Bcor* transcript. E*μ-Myc*-pQCIG.*Bcor*G2 1, E*μ-Myc*-pQCIG.*Bcor*G2 2, E*μ-Myc*-pQCIG.*Bcor*G2 4 and E*μ-Myc*-pQCIG.*Bcor*G2 5 demonstrated small deletion events (green boxes). E*μ-Myc*-pQCIG.*Bcor*G2 3 and E*μ-Myc*-pQCIG.*Bcor*G2 6 showed reduced read depth indicative of a large deletion event (red boxes). All events occur around the coordinates for which the sgRNA was designed against.

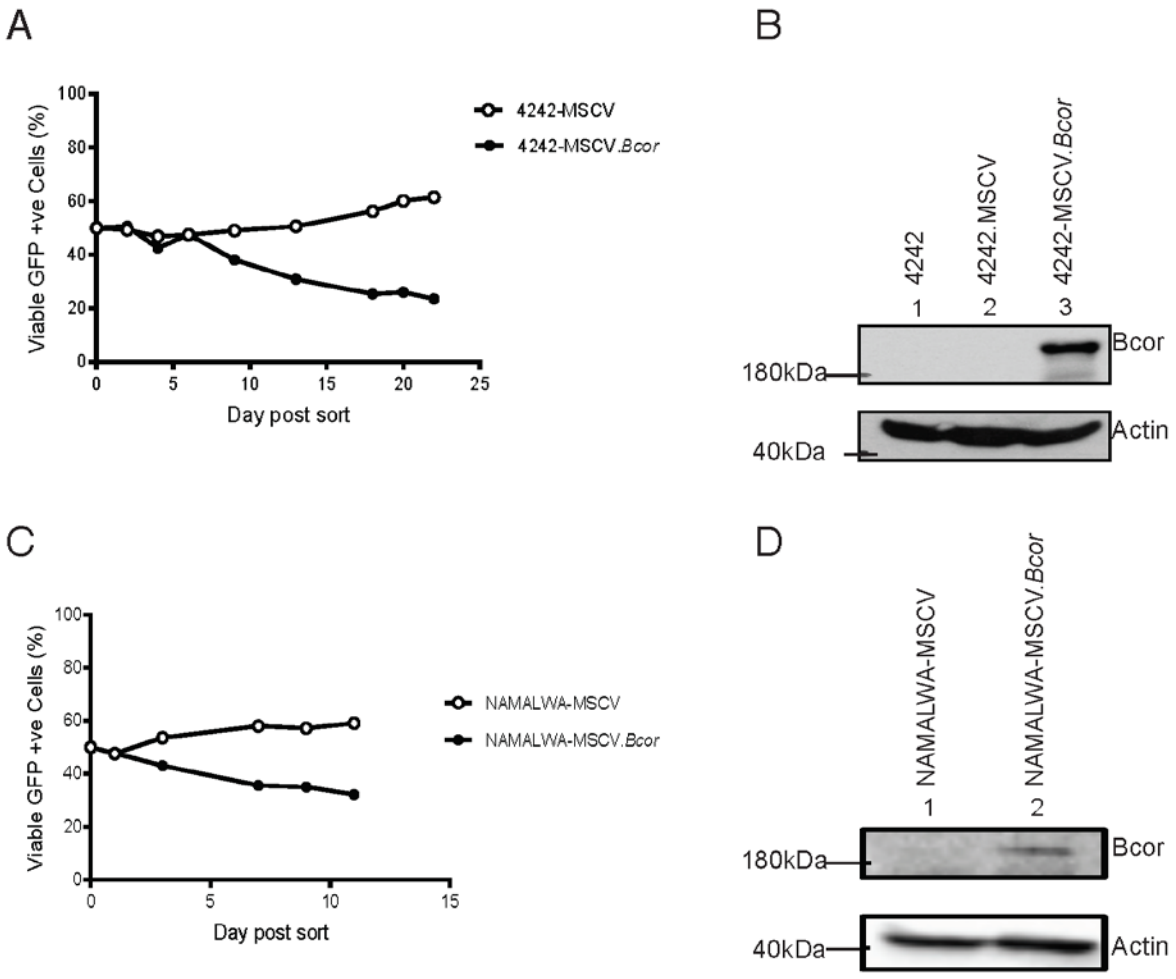

A
Viable GFP +ve Cells (%)
100
80
60
40
20
0
0
5
10
15
20
25
Day post sort
4242-MSCV
4242-MSCV.Bcor
B
4242
4242.MSCV
4242-MSCV.Bcor
1
2
3
180kDa
Bcor
40kDa
Actin
C
Viable GFP +ve Cells (%)
100
80
60
40
20
0
0
5
10
15
Day post sort
NAMALWA-MSCV
NAMALWA-MSCV.Bcor
D
NAMALWA-MSCV
NAMALWA-MSCV.Bcor
1
2
180kDa
Bcor
40kDa
Actin

Figure 4.08 – *Bcor* re-expression in a *Bcor*^{MUT} E*μ-Myc* cell line and a *Bcor*-low human B cell lymphoblastic cell line

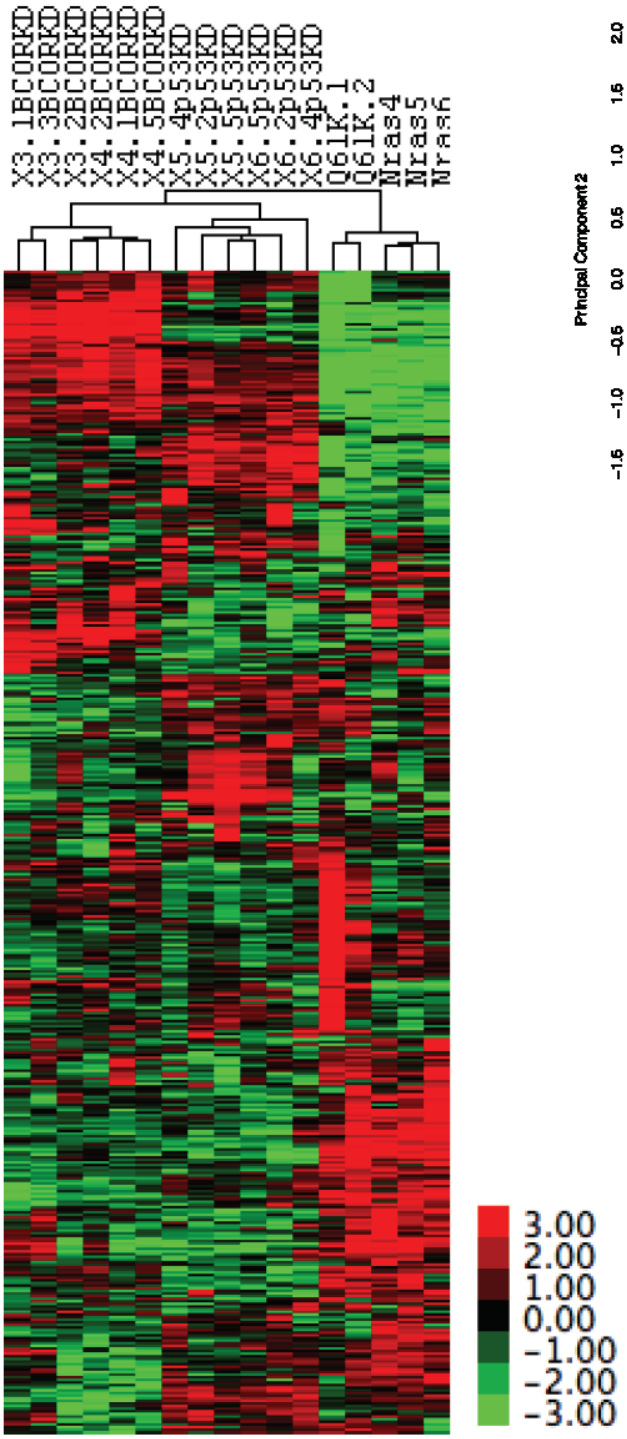
A) 4242 cells (*Bcor*^{MUT}) were transduced with either MSCV empty vector (4242-MSCV) or MSCV.*Bcor* (4242-MSCV.*Bcor*) and were sorted into a ratio of 50:50 with the non-transduced parental cell line. The GFP population was monitored over time and demonstrated that forced *Bcor* expression in 4242 cells was competitively disadvantageous. This experiment was performed once.

B) Protein immunoblot analysis demonstrating Bcor re-expression in the 4242-MSCV.*Bcor* cell line compared to both the cells transduced with the inert vector and the parental cell line.

C) Human B cell line, NAMALWA (*Bcor*-low) was transduced with either MSCV inert vector (NAMALWA-MSCV) or MSCV.*Bcor* (NAMALWA-MSCV.*Bcor*) and were sorted into a ratio of 50:50 with the non-transduced parental cell line (n=1). The GFP population was monitored over time and demonstrated that forced *Bcor* re-expression in NAMALWA cells was competitively disadvantageous. This experiment was performed once.

D) Protein immunoblot analysis demonstrating Bcor re-expression in the NAMALWA cell line that was transduced with MSCV.*Bcor* compared to the cells transduced with the empty vector. In this instance, the chemiluminescent images were captured using the BioRad Chemidock and were not produced on film.

A


X3.1BCORKD
X3.3BCORKD
X3.2BCORKD
X4.2BCORKD
X4.1BCORKD
X4.5BCORKD
X5.4p53KD
X5.2p53KD
X5.5p53KD
X6.2p53KD
X6.4p53KD
Q61K.1
Q61K.2
Nras4
Nras5
Nras6
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00

B

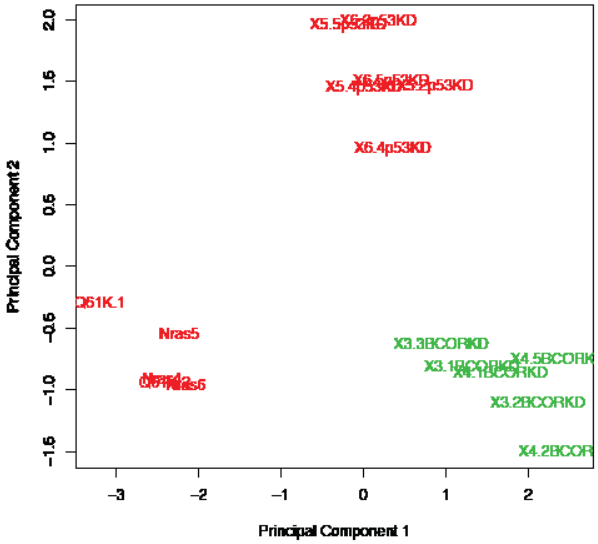

Principal Component 2
Principal Component 1
Q61K.1
Nras5
X3.3BCORKD
X3.2BCORKD
X6.4p53KD

Figure 4.09 – Unsupervised hierarchical clustering based on gene expression profiles derived from Eμ-*Myc-Bcor*.sh9 , Eμ-*Myc-Trp53*.1224 and Eμ-*Myc-Nras*^{Q61K} lymphomas

A) Gene expression profiling and unsupervised hierarchical clustering of 1112 genes from the RNA-sequencing of Eμ-*Myc-Bcor*.sh9 (n=6), Eμ-*Myc-Trp53*.1224 (n=6) and Eμ-*Myc-Nras*^{Q61K} (n=5) lymphomas. The unsupervised clustering analysis demonstrated that lymphomas clustered based on the engineered mutation.

B) 2-dimensional principle component analysis was performed on the total gene expression profile from each representative tumour type, with each tumour grouping with other tumours carrying the same introduced genetic insult.

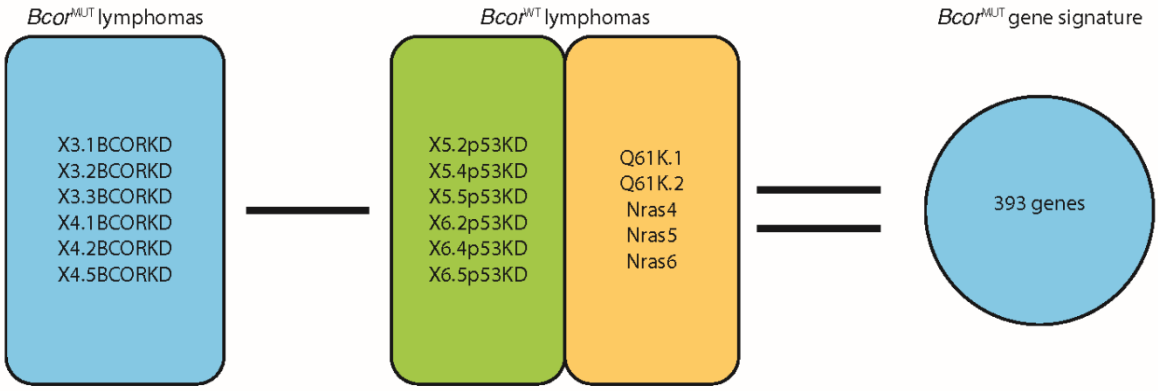

Bcor^MUT lymphomas
X3.1BCORKD
X3.2BCORKD
X3.3BCORKD
X4.1BCORKD
X4.2BCORKD
X4.5BCORKD
Bcor^WT lymphomas
X5.2p53KD
X5.4p53KD
X5.5p53KD
X6.2p53KD
X6.4p53KD
X6.5p53KD
Q61K.1
Q61K.2
Nras4
Nras5
Nras6
Bcor^MUT gene signature
393 genes

Figure 4.10 – Defining a *Bcor*^{MUT} gene set

RNA-sequencing was performed on E*μ-Myc-Bcor*.sh9 foetal liver derived lymphomas as a representative set of *Bcor*^{MUT} tumours and E*μ-Trp53*.1224 and E*μ-Myc-Nras*^{Q61K} foetal liver derived lymphomas as representative *Bcor*^{WT} tumours. This comparison identified 393 significantly differentially expressed genes. This gene set is representative of *Bcor* loss of function, or, *Bcor*^{MUT}.

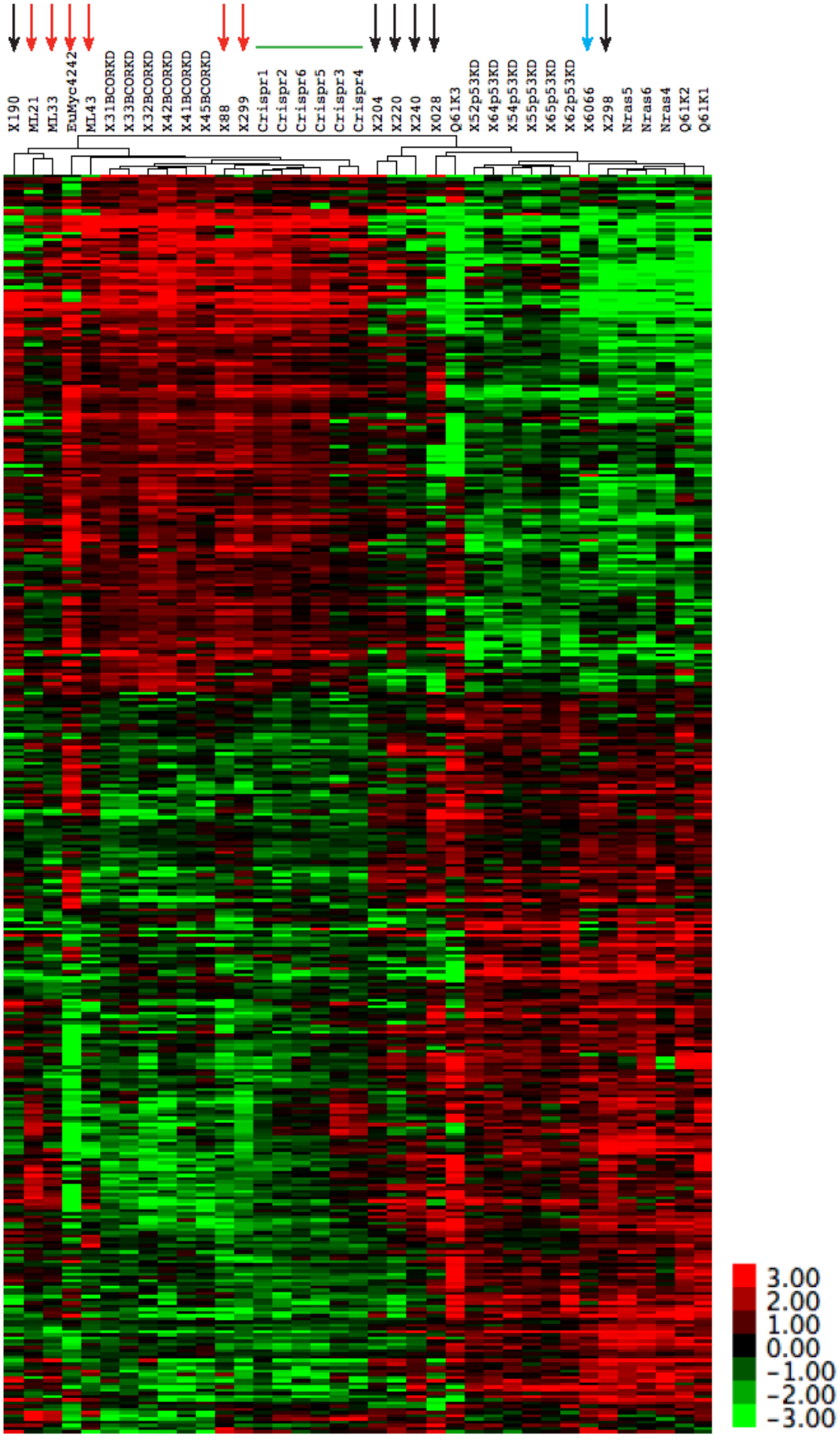

X190
ML21
ML33
EuMyc4242
ML43
X31BCORKD
X33BCORKD
X32BCORKD
X42BCORKD
X41BCORKD
X45BCORKD
X88
X299
Crispr1
Crispr2
Crispr6
Crispr5
Crispr3
Crispr4
X204
X220
X240
X028
Q61K3
X52p53KD
X64p53KD
X54p53KD
X55p53KD
X65p53KD
X62p53KD
X6066
X298
Nras5
Nras6
Nras4
Q61K2
Q61K1
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00

Figure 4.11 – Supervised hierarchical clustering of *Bcor*^{MUT} and *Bcor*^{WT} lymphomas using the *Bcor*^{MUT} gene set

RNA-sequencing analysis was first used to identify 393 significantly differentially expressed genes (FDR-corrected < 0.05 and log_2 fold-change ≥ 2) between E*μ-Myc-Bcor*.sh9 and combined E*μ-Myc.Trp53.*1224 and E*μ-Myc-Ras*^{Q61K} foetal liver derived tumours. RNA-sequencing data for shRNA foetal liver-derived lymphomas plus CRISPR-Cas9 (E*μ-Myc*-pQCIG-*Bcor*G2) and sporadic E*μ-Myc* lymphomas was then clustered against the defined 393 genes. All E*μ-Myc*-pQCIG-*Bcor*G2 (green bar) and sporadic *Bcor* mutant lymphomas (red arrows) clustered with the E*μ-Myc-Bcor*.sh9 foetal liver derived lymphomas. The sporadic *Nras*^{Q61K} mutant cell line #6066 (blue arrow) also clustered correctly with the E*μ-Myc-Ras*^{Q61K} foetal liver-derived lymphomas. Sporadic E*μ-Myc* lymphomas with unknown *Bcor* status (black arrow) are interspersed throughout, with the exception of X190 which clusters with *Bcor*^{MUT} tumours. Heat-map and scale bar represents median normalised log_2-fold gene-expression.

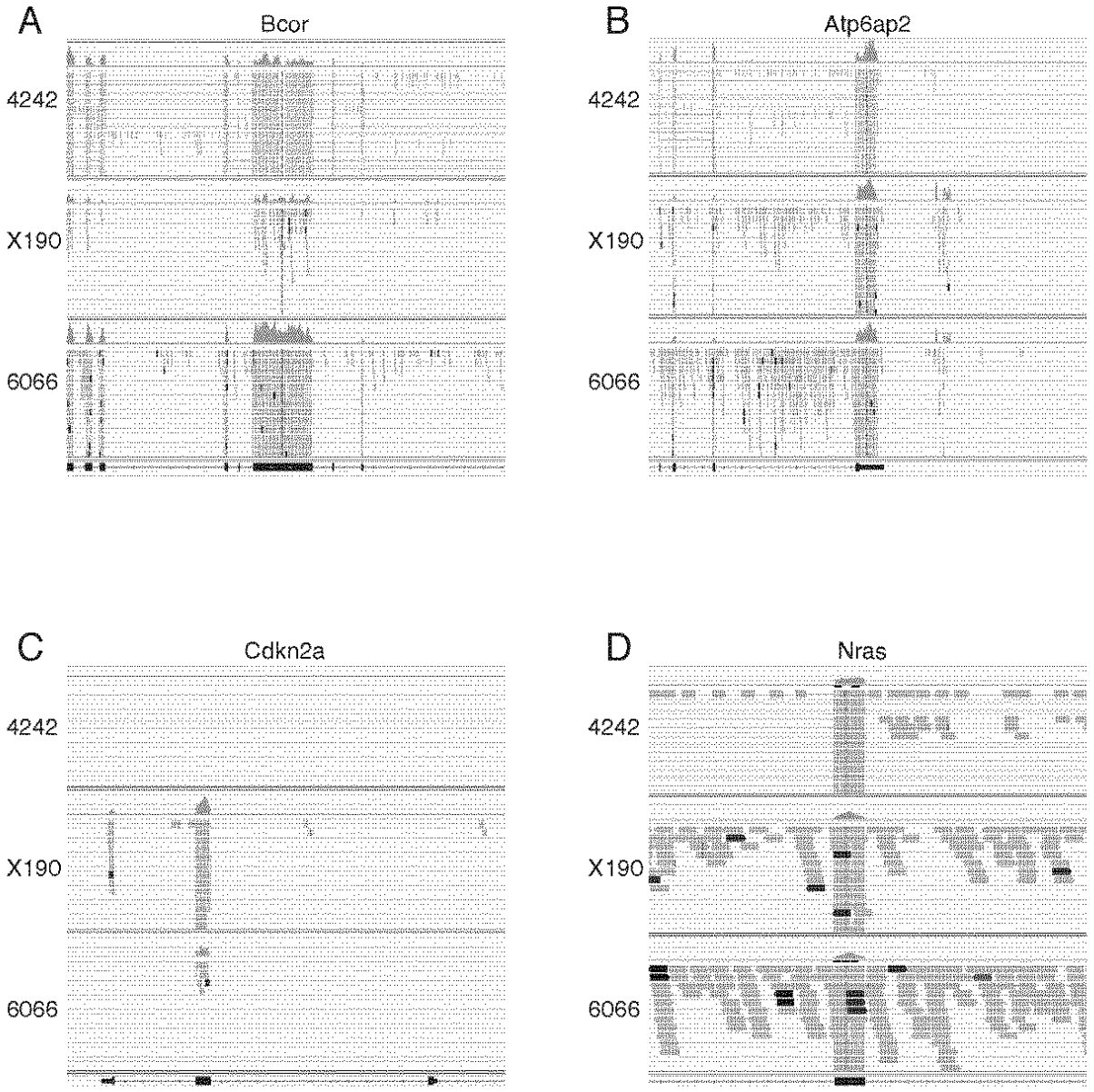

A
Bcor
4242
X190
6066
B
Atp6ap2
4242
X190
6066
C
Cdkn2a
4242
X190
6066
D
Nras
4242
X190
6066

Figure 4.12 – Supervised clustering analysis predicts sporadic *Bcor* mutation in E*μ-Myc* lymphoma X190

A) IGV reads spanning *Bcor* exon 4 in lymphoma 4242, which harboured a known small *Bcor* deletion at exon 4, lymphoma X190 that clustered with $Bcor^{MUT}$ lymphomas and lymphoma 6066 which did not contain a *Bcor* mutation. The predicted small deletion in *Bcor* is visible in lymphoma 4242, while lymphoma X190 has significantly less reads spanning *Bcor* compared to lymphoma 6066 and 4242.
B) IGV reads spanning *Atp6ap2*, a gene located on chromosome X and is proximal to *Bcor.* Lymphoma 4242, X190 and 6066 demonstrated approximately equivalent read counts spanning *Atp6ap2*.
C) IGV reads spanning *Cdkn2a*. Lymphoma 4242 and 6066 harboured *Cdkn2a* deletions, which is demonstrated by the lack of IGV reads. By comparison, lymphoma X190 appeared to contain an intact *Cdkn2a* locus.
D) IGV reads spanning *Nras*. Lymphoma 6066 contained an $Nras^{Q61K}$ mutation, which is not apparent in lymphomas 4242 and X190.

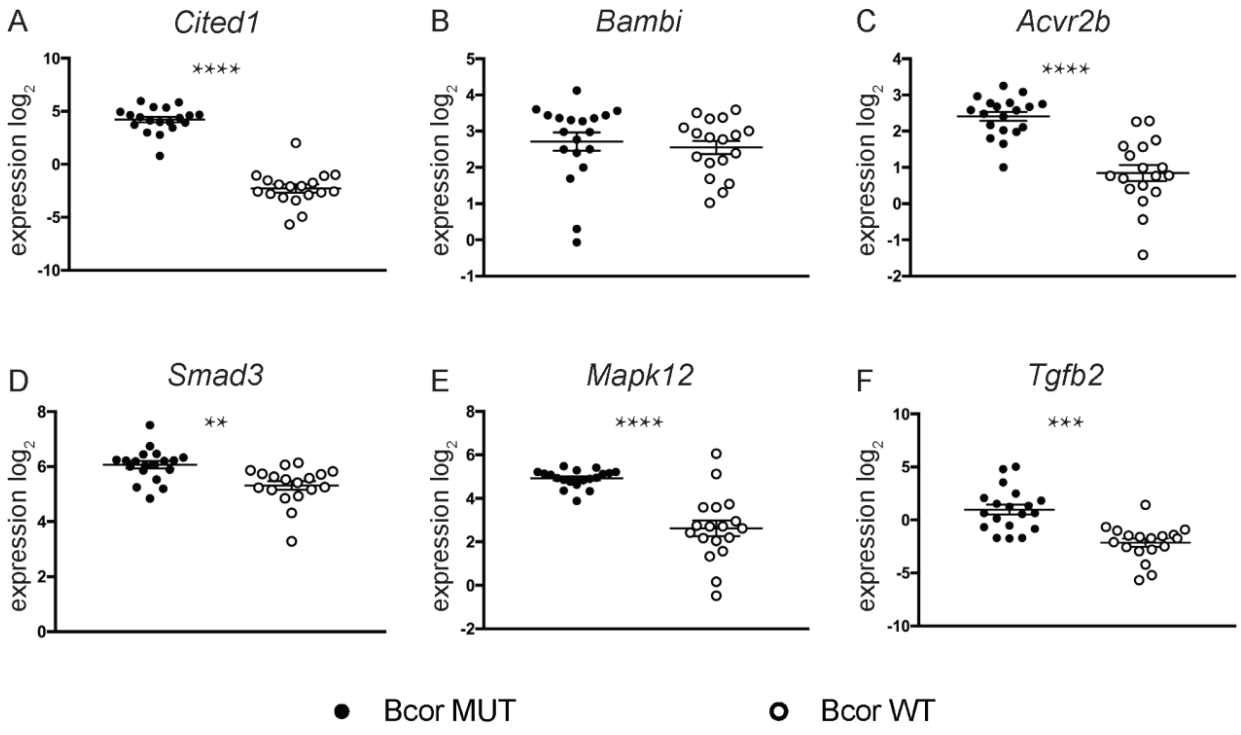

A
Cited1
****
expression log2
10
5
0
-5
-10
B
Bambi
expression log2
5
4
3
2
1
0
-1
C
Acvr2b
****
expression log2
4
3
2
1
0
-1
-2
D
Smad3
**
expression log2
8
6
4
2
0
E
Mapk12
****
expression log2
8
6
4
2
0
-2
F
Tgfb2
***
expression log2
10
5
0
-5
-10
Bcor MUT
Bcor WT

Figure 4.13 – Expression of *Tgfβ* pathway members in *Bcor*^{MUT} and *Bcor*^{WT} E*μ-Myc* lymphomas.

Expression (log_2) of *Tgfβ* pathway family members (A) *Cited1*, (B) *Bambi*, (C) *Acvr2b*, (D) *Smad3*, (E) *Mapk12* and (F) *Tgfβ2* shown for *Bcor*^{MUT} (n=19) and *Bcor*^{WT} (n=18) E*μ-Myc* lymphomas. Average *Cited1* expression (Log_2) is 4.2 in the *Bcor*^{MUT} compared to -2.3 in the *Bcor*^{WT} group, average *Bambi* expression (Log_2) is 2.7 in the *Bcor*^{MUT} compared to 2.6 in the *Bcor*^{WT} group, average *Acvr2b* expression (Log_2) is 2.4 in the *Bcor*^{MUT} compared to 0.8 in the *Bcor*^{WT} group, average *Smad3* expression (Log_2) is 6.1 in the *Bcor*^{MUT} compared to 5.3 in the *Bcor*^{WT} group, average *Mapk12* expression (Log_2) is 4.9 in the *Bcor*^{MUT} compared to 2.6 in the *Bcor*^{WT} group and average *Tgfb2* expression (Log_2) is 1.0 in the *Bcor*^{MUT} compared to -2.1 in the *Bcor*^{WT} group. Data shown is each expression for each individual lymphoma and mean of each subset ± SEM. ** = $p<0.01$, *** = $p<0.0005$, **** = $p<0.0001$ (paired student's t-test).

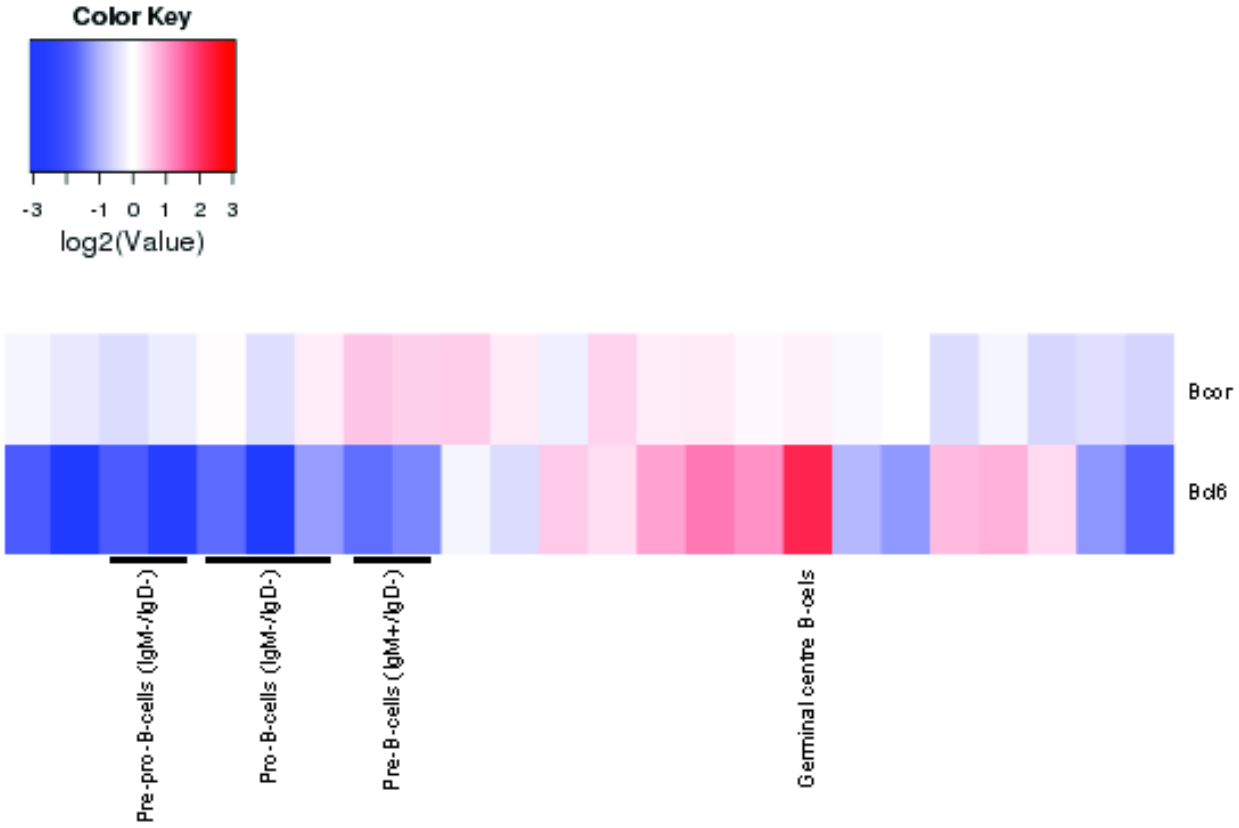

Color Key
-3
-1
0
1
2
3
log2(Value)
Bcor
Bcl6
Pre-pro-B-cells (IgM-/IgD-)
Pro-B-cells (IgM-/IgD-)
Pre-B-cells (IgM+/IgD-)
Germinal centre B-cells

Figure 4.14 – Differential expression of *Bcor* and *Bcl6* in normal B cells of different maturation state.

A heat-map demonstrating the differential expression of *Bcor* and *Bcl6* spanning the B cell lineage from the common lymphocyte progenitor to circulating plasma cells. Pre-pro-B cells, pro-B cells and pre-B cells express very low levels of *Bcl6* and increasing amounts of *Bcor*. The Heat-map was generated using the My Gene Set function provided by the Immunological Genome Project (http://www.immgen.org/).

5 Summary and conclusions

5.1 Genomic characterisation of E*μ-Myc* lymphoma

The E*μ-Myc* model of B cell lymphoma/leukaemia has become one of the most utilised mouse models of cancer since its generation (Adams et al., 1985). Despite its utility as a model of lymphoma, to date the E*μ-Myc* transgenic mouse has not been fully genomically characterised. Ostensibly, the E*μ-Myc* model is described as being a model of BL as it is driven by the same t(8;14) translocation apparent in its human equivalent (Corcoran et al., 1985). However, despite sharing the same initiating oncogenic lesion, the molecular and histopathological characteristics between E*μ-Myc* and BL do not significantly overlap. Data presented in this thesis challenges central paradigms associated with E*μ-Myc* lymphomagenesis.

5.1.1 The potential for multiple founding mutations in E*μ-Myc* lymphoma

A throwaway line preceding many publications that have utilised the E*μ-Myc* model of lymphoma (including chapter 1 in this thesis) is to state that E*μ-Myc* lymphomas are driven by a single oncogenic insult – Myc. However, this thesis, and a study running contemporary to it (Fusello et al., 2013), identified for the first time that the E*μ-Myc* transgene insertion site is a cancer-associated genomic region. At the time of pro-nuclear injection the E*μ-Myc* transgene inserted three times into chromosome 19, resulting in an amplification of a proximal region that is syntenic to human 9p24.1, which is commonly amplified in Hodgkin's lymphoma and primary mediastinal B cell lymphoma (PMBCL) (Green et al., 2010). This provides an explanation as to why E*μ-Myc*

lymphoma gene expression signatures can resemble DLBCL such as PMBCL rather that BL (Mori et al., 2008). Supporting the published data that the 9p24.1 region is oncogenic is that there does appear to be selective pressure placed on the chr19 amplicon during E*μ-Myc* tumourigenesis, as approximately half of the sequenced E*μ-Myc* lymphomas demonstrated further somatic gain of the chr19 amplicon. The chr19 amplicon contains *Jak-2*; however, the amplification of *Jak-2* was shown to be non-functional in E*μ-Myc* lymphoma. *Cd274* (Pdl-1) is also amplified with the ch19 amplicon and while it was demonstrated that an extra somatic gain of *Cd274* did not result in increased expression of Pd-l1, its functional importance was never assessed in this investigation. Recently, it has been demonstrated that E*μ-Myc* lymphomas are sensitive to down-regulation of Pd-l1 (Hogg et al., 2017). Furthermore, in human and murine leukaemia and lymphomas, *CD274* has been identified as a MYC target, through which MYC can initiate and maintain tumourigenesis (Casey et al., 2016). These investigations are suggestive of a cooperative interaction between Pd-l1 upregulation and Myc overexpression and provide a hypothesis as to why the chr19 amplicon is further somatically amplified during lymphomagenesis. Despite this, E*μ-Myc* tumours that had additional gains of the chr19 amplicon were not more aggressive as measured by survival in the parental mice. In addition to *Cd274* and *Jak-2*, the chr19 amplicon harboured another 41 genes, such as *Pdcd1lg2* (Pd-l2) that has been implicated in malignancy but not investigated herein. This provides strong evidence to suggest that within the chr19 amplicon, and by extension human 9p24.1, resides potentially undiscovered tumour-promoting genes capable of cooperating with Myc during malignant transformation. A multiplexed cell death assay using RNAi would be

useful to test the other candidate genes in this region and their oncogenic potential in initiating and maintaining E*μ-Myc* lymphomagenesis (Falkenberg et al., 2014; Falkenberg et al., 2016). The E*μ-Myc* model has been used extensively to study how MYC promotes lymphomagenesis; however, as the chr19 amplification is a founding event and is frequently selectively gained during lymphomagenesis, some of these processes attributed to Myc dysfunction may need to be re-evaluated.

5.1.2 The importance of $p19^{ARF}$ in E*μ-Myc* lymphoma may be overestimated

Much like the assumption that a single oncogenic lesion drives E*μ-Myc lymphoma*, the assumption that for malignant transformation to occur a mutation in either *Cdkn2a* or *Trp53* is required and sufficient was demonstrated to be simplistic. The $p19^{ARF}$-Mdm2-Trp53 pathway, which is activated by aberrant Myc signalling, is thought to be the major safeguard protecting the cell against Myc-driven proliferation (Eischen et al., 1999; Schmitt et al., 1999). Inactivation of this tumour suppressive axis has been identified in up to 50% of spontaneous E*μ-Myc* lymphoma, indicating this pathway is under significant selective pressure in response to Myc-overexpression (Eischen et al., 1999). Trp53 is the effector in this pathway, and as such its direct inactivation or inactivation upstream of Trp53 in $p19^{ARF}$ is viewed as being sufficient to allow the cell to circumvent Myc-driven apoptosis and progress to frank clonal lymphoma/leukaemia (Eischen et al., 1999; Lowe, 1999; Schmitt et al., 2002). However, mutations involving homozygous *Cdkn2a* deletions concomitant with the acquisition of *bona fide* oncogenic mutations

such as activating *Kras* or *Nras* mutations or inactivating *Bcor* mutations were identified for the first time in clonal E*μ-Myc* lymphomas. This result was surprising, given that *Cdkn2a* deletion would result in loss of $p19^{ARF}$ expression, releasing Mdm2 to antagonise Trp53 functions and should therefore not require the acquisition of additional driving mutations for malignant transformation. Yet in 3/5 spontaneous E*μ-Myc* lymphomas that harboured *Cdkn2a* deletion there was concomitant mutation in *Nras*, *Kras* or *Bcor.* Notably, *Trp53* mutations never co-occurred with another obvious oncogenic lesion in the spontaneous E*μ-Myc* lymphomas sequenced in this study. Taken together, these results indicate that inactivation of *Trp53* alone is sufficient in providing the secondary hit required for E*μ-Myc* lymphomagenesis, but *Cdkn2a* deletion is not. Therefore, additional $p19^{ARF}$-independent regulatory stimuli must be incumbent upon Trp53 otherwise $p19^{ARF}$ loss alone would be sufficient with cooperating with Myc and driving lymphomagenesis. These findings build on published data that demonstrated $p19^{ARF}$ inactivation alone was insufficient in driving E*μ-Myc* lymphomagenesis, and that cryptic mutational events accompanied LOH at the *Cdkn2a* allele during malignant transformation in the E*μ-Myc*;$Cdkn2a^{+/-}$ disease (Bertwistle and Sherr, 2007). This thesis uncovers three such cryptic mutational events with an expectation that further mining of the SNV and CNV data in the remaining *Cdkn2a*-deleted lymphomas will uncover novel inconspicuous cooperative genomic lesions. Further uncoupling $p19^{ARF}$ loss from Trp53 inactivation is the identification that $Trp53^{MUT}$, but not *Cdkn2a*-deleted, E*μ-Myc* lymphomas are resistant to chemotherapy (Shortt et al., 2013). Co-occurring *Cdkn2a* deletions and *Trp53* mutations have not been described, indicating that there is redundancy within the pathway; however,

Cdkn2a loss does not totally ameliorate selective pressure for mutations outside of the p19^{ARF}-Mdm2-Trp53 axis. Data presented in this thesis suggests that the importance of p19^{ARF} in restraining oncogenic Myc signalling is overstated. This is of clinical significance given the high prevalence of, and importance placed on, *CDKN2A* deletions in human malignancies.

5.1.3 Novel mutations have been described in E*μ-Myc* lymphoma

In addition to expected *Cdkn2a* deletions, *Trp53* loss-of-function mutations and activating mutations in *Ras*-family genes, WES of E*μ-Myc* lymphomas revealed spontaneous mutations in genes that, until this study, had not been described in E*μ-Myc* lymphoma. An inactivating mutation in *Ezh2* was identified, supporting data published in a contemporary investigation that demonstrated that forced knockdown of Ezh2 (and other PRC2 components) accelerated E*μ-Myc* lymphomagenesis (Lee et al., 2013). Thus, *de novo* inactivation of Ezh2 is likely to cooperate with Myc in driving lymphomagenesis in E*μ-Myc* lymphoma. This is contrary to studies in human germinal centre B cell malignancies such as follicular lymphoma, mantle cell lymphoma and GCB DLBCL where *EZH2* is oncogenic (Béguelin et al., 2013; Morin et al., 2010; Visser et al., 2001). Furthermore, targeted EZH2 inhibition has proved clinically efficacious in germinal centre B cell malignancies that harbour ectopic EZH2 activity (Mccabe et al., 2012). Identifying *Ezh2* as a tumour suppressor gene that when inactivated is capable of cooperating with Myc in early pre-B/pro-B lymphomas raises concerns about the long term utility of EZH2 inhibition in the treatment of lymphoma. Similar to the discussed *Ezh2* mutations, a potential activating *Mtor* mutation was identified in E*μ-Myc*

lymphoma, supporting published investigations demonstrating that selective inhibition of Mtor complex 1 (Mtorc1) with everolimus delayed lymphomagenesis by restoring OIS (Shortt et al., 2013; Wall et al., 2013). Other mutations occurring in *Rpl10*, previously described in T-ALL (De Keersmaecker et al., 2013) and *mir142*, known to be frequently mutated in DLBCL (Kwanhian et al., 2012), were both identified for the first time to be spontaneously mutated in E*μ-Myc* lymphoma. The identification of these low frequency, putative cancer associated genes, demonstrates that the E*μ-Myc* model of lymphoma still has utility as a cancer gene discovery tool. In human DLBCL, the number or rare somatic mutations increased linearly with increased number of sequenced tumours, suggesting that continual sequencing of DLBCL will lead to further identification of novel, functional cancer-associated variants (Zhang et al., 2013). Indeed, four large MPS studies in DLBCL identified minimal overlap in the number of rare somatic events identified (Lohr et al., 2012; Morin et al., 2011; Pasqualucci et al., 2011; Zhang et al., 2013). This indicates that continued sequencing of E*μ-Myc* lymphomas would further identify novel cancer-associated genes capable of cooperating with Myc in B cell lymphomagenesis.

5.2 The role of Bcor in E*μ-Myc* lymphomagenesis

This thesis defined, for the first time, *Bcor* loss-of-function mutations in E*μ-Myc* lymphoma. Furthermore, *Bcor* was the most frequently mutated gene in E*μ-Myc* lymphoma. As previously discussed, *BCOR*-inactivating mutations have been described, at low frequency, in a range of human malignancies including B-ALL, AML, CLL and CMML (Damm et al., 2013; Grossmann et al., 2011; Papaemmanuil et al.,

2014; Quesada et al., 2011; Tiacci et al., 2012), therefore the discovery of parallel mutations in E*μ-Myc* lymphoma is of clinical importance.

5.2.1 *Bcor* is under significant selective pressure for mutation in E*μ-Myc* lymphomagenesis

As discussed, *Bcor* is frequently inactivated during E*μ-Myc* lymphomagenesis. The inactivation of *Bcor* occurred at a higher frequency than both *Trp53* loss of function mutations and *Cdkn2a* deletions (even co-occurring with the latter). This is surprising, given the historical importance placed on the $p19^{ARF}$-Mdm2-Trp53 pathway in restraining the oncogenic effects of Myc (Eischen et al., 1999; Schmitt et al., 2002; Schmitt et al., 1999). A high rate of mutation frequency in a gene is proportional to the selective pressure exerted on it to become activated in the case of an oncogene or inactivated in the case of a tumour suppressor gene (Sjoblom et al., 2006). Therefore, it appears that during E*μ-Myc* lymphomagenesis, *Bcor* is selectively preferentially mutated ahead of other pathways and genes. Thus, *Bcor* is one of the most critical tumour suppressive genes activated at in response to aberrant Myc expression in immature B cells. With MYC deregulation featuring in 50% of human malignancies, identification of a potential novel gene and pathway that works to restrict Myc activity is of profound clinical significance. As tumour suppressor genes are mutated in such a way that they lose function, therapeutically targeting tumour suppressor genes is not a feasible approach. Instead, identifying the genes th at Bcor negatively regulates will potentially yield therapy amenable novel targets involved in positively regulating Myc activity or augmenting Myc effects. Much like the identification of the cooperative

interaction between pro-survival BCL2-family members and MYC in studies utilising E*μ-Myc* lymphoma that resulted in the development of effective targeted therapies (Mason et al., 2008; Roberts et al., 2016; Strasser et al., 1990), the identification of a BCOR tumour suppressive pathway could yield novel molecular targets that are amenable to inhibition in MYC-driven malignancy.

5.2.2 *Bcor* is a *bona fide* tumour suppressor gene

As previously discussed, recurrent protein truncating mutations in *BCOR* in human malignancies, a recent *in vitro* study demonstrating that BCOR regulates myeloid cell proliferation and differentiation (Cao et al., 2016) and observations of highly recurrent Bcor inactivating mutations in E*μ-Myc* suggest a tumour suppressor function. However, this has not been demonstrated *in vivo*. This thesis functionally demonstrated that Bcor loss accelerated Myc driven lymphomagenesis *in vivo* and that forced Bcor-re-expression in *Bcor*^{MUT} cell lines induced an anti-oncogenic response. Bcor-loss resulted in predominantly pre-B or pro-B cell lymphomas that were generally less differentiated compared to other E*μ-Myc* lymphomas harbouring mutations in other genes. Canonically, Bcor binds to Bcl6 and potentiates Bcl6 mediated repression; however, in pre-B and pro-B E*μ-Myc* lymphomas, Bcl6 expression is low (Mori et al., 2008). Independent of Bcl6 expression, Bcor forms a non-canonical PRC1-like complex (Sánchez et al., 2007). Many distinct PRC1 complexes are present in hematopoietic cells, likely with important differences in terms of gene regulatory functions and target specificity (van den Boom et al., 2013; Vidal and Starowicz, 2017). Given that *Bcor* inactivating mutations drive tumourigenesis in the immature B cell compartment and WT BCOR is critical for the development of circulating mature B lymphocytes in humans (Ng

et al., 2004) it is proposed the Bcor plays a role in B cell differentiation as well as restraining oncogenic Myc signalling.

5.2.3 Mechanism through which Bcor represses Myc-driven oncogenesis

The mechanistic pathway through which Bcor exerts its tumour suppressive properties in pre-B and pro-B B cell lymphoma remains undefined. In myeloid cells, BCOR has been shown to control proliferation and differentiation through tightly regulating *Hox*-family genes (Cao et al., 2016); however, there was no overlap between the gene expression profile associated with Bcor inactivation in E*μ-Myc* lymphoma and what has been previously published in myeloid cells and AML cell lines. Thus, Bcor functions appeared to be cell type and context dependent. The *Bcor*^{MUT} gene expression profile was distinct to the gene expression profile of Trp53-deleted and constitutively activated Nras lymphomas, demonstrating that this thesis has uncovered a novel tumour suppressive pathway in E*μ-Myc* lymphoma. *Bcor*-inactivation resulted in enhanced signalling redolent of Tgfβ-pathway signalling, which was the only significantly upregulated pathway identified. Non-cell autonomous Tgfβ signalling from stromal macrophages has been identified as tumour-protective in E*μ-Myc* lymphoma by driving Myc-induced cellular senescence responses in a Suv39h1-dependent manner (Reimann et al., 2010). However, Suv39h1 is also a cell autonomous effector of OIS in response to aberrant Myc signalling and DDR in E*μ-Myc* lymphoma (Reimann et al., 2010; Reimann et al., 2007). In E*μ-Myc* lymphoma knockout of Suv39h1 accelerated lymphomagenesis with the resultant lymphomas presenting with a similarly enhanced Tgfβ signature comparable to the signature observed in this study. The gene

expression profile similarities between *Bcor*-inactivated and *Suv39h1*-deleted E*μ-Myc* lymphomas raises the intriguing possibility that BCOR correspondingly induces senescence in response to MYC signalling and DDR in a similar manner to Suv39h1. This has obvious therapeutic implications. If Bcor, which has been shown to potently prohibit oncogenic Myc signalling in E*μ-Myc* lymphoma, is preferentially inactivated allowing the cells to escape OIS in response to Myc-induced DDR, then treatment with senescence inducing agents, such as PARPi, may prove therapeutically viable.

5.3 Conclusions

The E*μ-Myc* mouse has been used as representative model of MYC-driven lymphoma for 32 years at the time of writing this thesis; however, for the last 15 years it has been used less as a tool for MYC-cooperative gene discovery. Through the application of NGS technology, this study has redefined E*μ-Myc* lymphoma and has identified previously un-described novel mutations capable of cooperating with Myc and driving lymphomagenesis. The genomic “archeology” which worked to define the topology of the transgenic insertion revealed inherent germline amplification in a potent oncogenic genomic region, suggesting that the E*μ-Myc* model has potentially more than one founding driver mutation, having obvious implications for conclusions made under the assumption that Myc is the sole initiator in E*μ-Myc* lymphoma. Furthermore, data presented herein provides important information about the role of *Cdkn2a* in restraining Myc activity and that perhaps, the role of $p19^{ARF}$ has been overstated. Importantly, that Bcor is the most commonly mutated gene in E*μ-Myc* lymphoma indicates that its role as a tumour suppressor gene has been underappreciated. With MYC deregulation playing

a critical part in almost all forms of tumourigenesis, yet its direct inhibition remaining an unappealing therapeutic strategy, the identification of a potent MYC-restraining gene is of significant clinical importance. More work is required to dissect the pathway through which Bcor exerts its cancer-protective activity in response to Myc signalling, however, this thesis provides a solid foundation on which to conduct further research.

6 Bibliography

Adams, J.M., and Cory, S. (1991). Transgenic models of tumor development. Science *254*, 1161-1167.

Adams, J.M., Harris, A.W., Pinkert, C.A., Corcoran, L.M., Alexander, W.S., Cory, S., Palmiter, R.D., and Brinster, R.L. (1985). The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. Nature *318*, 533-538.

Adhikary, S., and Eilers, M. (2005). Transcriptional regulation and transformation by Myc proteins. Nature reviews Molecular cell biology *6*, 635-645.

Akasaka, T., Akasaka, H., Ueda, C., Yonetani, N., Maesako, Y., Shimizu, A., Yamabe, H., Fukuhara, S., Uchiyama, T., and Ohno, H. (2000). Molecular and clinical features of non-Burkitt's, diffuse large-cell lymphoma of B-cell type associated with the c-MYC/immunoglobulin heavy-chain fusion gene. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *18*, 510-518.

Alexander, W.S., Bernard, O., Cory, S., and Adams, J.M. (1989). Lymphomagenesis in E mu-myc transgenic mice can involve ras mutations. Oncogene *4*, 575-581.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.-L.*, et al.* (2013a). Signatures of mutational processes in human cancer. Nature.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Campbell, P.J., and Stratton, M.R. (2013b). Deciphering signatures of mutational processes operative in human cancer. Cell reports *3*, 246-259.

Alizadeh, A.A., Eisen, M.B., Davis, R.E., Ma, C., Lossos, I.S., Rosenwald, A., Boldrick, J.C., Sabet, H., Tran, T., Yu, X.*, et al.* (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature *403*, 503-511.

Allman, D., Li, J., and Hardy, R.R. (1999). Commitment to the B lymphoid lineage occurs before DH-JH recombination. J Exp Med *189*, 735-740.

Amarasinghe, K.C., Li, J., Hunter, S.M., Ryland, G.L., Cowin, P.A., Campbell, I.G., and Halgamuge, S.K. (2014). Inferring copy number and genotype in tumour exome data. BMC genomics *15*, 732.

Anderson, K., Lutz, C., van Delft, F.W., Bateman, C.M., Guo, Y., Colman, S.M., Kempski, H., Moorman, A.V., Titley, I., Swansbury, J.*, et al.* (2011). Genetic variegation of clonal architecture and propagating cells in leukaemia. Nature *469*, 356-361.

Andrechek, E.R., and Nevins, J.R. (2010). Mouse models of cancers: opportunities to address heterogeneity of human cancer and evaluate therapeutic strategies. Journal of molecular medicine *88*, 1095-1100.

Askew, D.S., Ashmun, R.A., Simmons, B.C., and Cleveland, J.L. (1991). Constitutive c-myc expression in an IL-3-dependent myeloid cell line suppresses cell cycle arrest and accelerates apoptosis. Oncogene *6*, 1915-1922.

Attisano, L., Wrana, J.L., Montalvo, E., and Massague, J. (1996). Activation of signalling by the activin receptor complex. Mol Cell Biol *16*, 1066-1073.

Aukema, S.M., Siebert, R., Schuuring, E., van Imhoff, G.W., Kluin-Nelemans, H.C., Boerma, E.J., and Kluin, P.M. (2011). Double-hit B-cell lymphomas. Blood *117*, 2319-2331.

Avruch, J., Khokhlatchev, A., Kyriakis, J.M., Luo, Z., Tzivion, G., Vavvas, D., and Zhang, X.F. (2001). Ras activation of the Raf kinase: tyrosine kinase recruitment of the MAP kinase cascade. Recent progress in hormone research *56*, 127-155.

Bahram, F., von der Lehr, N., Cetinkaya, C., and Larsson, L.G. (2000). c-Myc hot spot mutations in lymphomas result in inefficient ubiquitination and decreased proteasome-mediated turnover. Blood *95*, 2104-2110.

Bardelli, A., Cahill, D.P., Lederer, G., Speicher, M.R., Kinzler, K.W., Vogelstein, B., and Lengauer, C. (2001). Carcinogen-specific induction of genetic instability. Proc Natl Acad Sci U S A *98*, 5770-5775.

Bartlett, J.M., Going, J.J., Mallon, E.A., Watters, A.D., Reeves, J.R., Stanton, P., Richmond, J., Donald, B., Ferrier, R., and Cooke, T.G. (2001). Evaluating HER2 amplification and overexpression in breast cancer. J Pathol *195*, 422-428.

Bass, A.J., Lawrence, M.S., Brace, L.E., Ramos, A.H., Drier, Y., Cibulskis, K., Sougnez, C., Voet, D., Saksena, G., Sivachenko, A.*, et al.* (2011). Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. Nat Genet *43*, 964-968.

Bassan, R., and Hoelzer, D. (2011). Modern therapy of acute lymphoblastic leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *29*, 532-543.

Baudino, T.A., McKay, C., Pendeville-Samain, H., Nilsson, J.A., Maclean, K.H., White, E.L., Davis, A.C., Ihle, J.N., and Cleveland, J.L. (2002). c-Myc is essential for vasculogenesis and angiogenesis during development and tumor progression. Genes & development *16*, 2530-2543.

Béguelin, W., Popovic, R., Teater, M., Jiang, Y., Bunting, K.L., Rosen, M., Shen, H., Yang, S.N., Wang, L., Ezponda, T.*, et al.* (2013). EZH2 Is Required for Germinal Center Formation and Somatic EZH2 Mutations Promote Lymphoid Transformation. Cancer Cell *23*, 677-692.

Bender, A.M., Collier, L.S., Rodriguez, F.J., Tieu, C., Larson, J.D., Halder, C., Mahlum, E., Kollmeyer, T.M., Akagi, K., Sarkar, G.*, et al.* (2010). Sleeping beauty-mediated somatic mutagenesis implicates CSF1 in the formation of high-grade astrocytomas. Cancer Res *70*, 3557-3565.

Bertwistle, D., and Sherr, C.J. (2007). Regulation of the Arf tumor suppressor in Emicro-Myc transgenic mice: longitudinal study of Myc-induced lymphomagenesis. Blood *109*, 792-794.

Bhang, H.E., Ruddy, D.A., Krishnamurthy Radhakrishna, V., Caushi, J.X., Zhao, R., Hims, M.M., Singh, A.P., Kao, I., Rakiec, D., Shaw, P.*, et al.* (2015). Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. Nat Med *21*, 440-448.

Bhattacharyya, J., Mihara, K., Ohtsubo, M., Yasunaga, S., Takei, Y., Yanagihara, K., Sakai, A., Hoshi, M., Takihara, Y., and Kimura, A. (2012). Overexpression of BMI-1 correlates with drug resistance in B-cell lymphoma cells through the stabilization of survivin expression. Cancer science *103*, 34-41.

Boerma, E.G., Siebert, R., Kluin, P.M., and Baudis, M. (2009). Translocations involving 8q24 in Burkitt lymphoma and other malignant lymphomas: a historical review of cytogenetics in the light of todays knowledge. Leukemia *23*, 225-234.

Bordow, S.B., Norris, M.D., Haber, P.S., Marshall, G.M., and Haber, M. (1998). Prognostic significance of MYCN oncogene expression in childhood neuroblastoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *16*, 3286-3294.

Boxer, L.M., and Dang, C.V. (2001). Translocations involving c-myc and c-myc function. Oncogene *20*, 5595-5610.

Bozic, I., Antal, T., Ohtsuki, H., Carter, H., Kim, D., Chen, S., Karchin, R., Kinzler, K.W., Vogelstein, B., and Nowak, M.A. (2010). Accumulation of driver and passenger mutations during tumor progression. Proc Natl Acad Sci U S A *107*, 18545-18550.

Braig, M., Lee, S., Loddenkemper, C., Rudolph, C., Peters, A.H., Schlegelberger, B., Stein, H., Dorken, B., Jenuwein, T., and Schmitt, C.A. (2005). Oncogene-induced senescence as an initial barrier in lymphoma development. Nature *436*, 660-665.

Bric, A., Miething, C., Bialucha, C.U., Scuoppo, C., Zender, L., Krasnitz, A., Xuan, Z., Zuber, J., Wigler, M., Hicks, J.*, et al.* (2009). Functional identification of tumor-suppressor genes through an in vivo RNA interference screen in a mouse lymphoma model. Cancer Cell *16*, 324-335.

Brinster, R.L., Chen, H.Y., Messing, A., van Dyke, T., Levine, A.J., and Palmiter, R.D. (1984). Transgenic mice harboring SV40 T-antigen genes develop characteristic brain tumors. Cell *37*, 367-379.

Brose, M.S., Volpe, P., Feldman, M., Kumar, M., Rishi, I., Gerrero, R., Einhorn, E., Herlyn, M., Minna, J., Nicholson, A.*, et al.* (2002). BRAF and RAS mutations in human lung cancer and melanoma. Cancer Res *62*, 6997-7000.

Brugiatelli, M., Claisse, J.F., Lenormand, B., Morabito, F., Callea, V., Malloum, K., Chevret, S., Binet, J.L., Dighiero, G., and Travade, P. (1997). Long-term clinical outcome of B-cell chronic lymphocytic leukaemia patients in clinical remission phase evaluated at phenotypic level. British journal of haematology *97*, 113-118.

Buchdunger, E., Zimmermann, J., Mett, H., Meyer, T., Muller, M., Druker, B.J., and Lydon, N.B. (1996). Inhibition of the Abl protein-tyrosine kinase in vitro and in vivo by a 2-phenylaminopyrimidine derivative. Cancer Res *56*, 100-104.

Buhrman, G., Kumar, V.S., Cirit, M., Haugh, J.M., and Mattos, C. (2011a). Allosteric modulation of Ras-GTP is linked to signal transduction through RAF kinase. J Biol Chem *286*, 3323-3331.

Buhrman, G., O'Connor, C., Zerbe, B., Kearney, B.M., Napoleon, R., Kovrigina, E.A., Vajda, S., Kozakov, D., Kovrigin, E.L., and Mattos, C. (2011b). Analysis of binding site hot spots on the surface of Ras GTPase. Journal of molecular biology *413*, 773-789.

Burmeister, T., Schwartz, S., Horst, H.A., Rieder, H., Gokbuget, N., Hoelzer, D., and Thiel, E. (2005). Molecular heterogeneity of sporadic adult Burkitt-type

leukemia/lymphoma as revealed by PCR and cytogenetics: correlation with morphology, immunology and clinical features. Leukemia *19*, 1391-1398.

Burnet, F.M. (1976). A modification of Jerne's theory of antibody production using the concept of clonal selection. CA: a cancer journal for clinicians *26*, 119-121.

Cahill, D.P., Kinzler, K.W., Vogelstein, B., and Lengauer, C. (1999). Genetic instability and darwinian selection in tumours. Trends in cell biology *9*, M57-60.

Campbell, L.J. (2005). Cytogenetics of lymphomas. Pathology *37*, 493-507.

Campbell, P.J., Pleasance, E.D., Stephens, P.J., Dicks, E., Rance, R., Goodhead, I., Follows, G.A., Green, A.R., Futreal, P.A., and Stratton, M.R. (2008). Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. Proc Natl Acad Sci U S A *105*, 13081-13086.

Campo, E., Swerdlow, S.H., Harris, N.L., Pileri, S., Stein, H., and Jaffe, E.S. (2011). The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. Blood *117*, 5019-5032.

Cang, S., Iragavarapu, C., Savooji, J., Song, Y., and Liu, D. (2015). ABT-199 (venetoclax) and BCL-2 inhibitors in clinical development. Journal of hematology & oncology *8*, 129.

Cao, Q., Gearhart, M.D., Gery, S., Shojaee, S., Yang, H., Sun, H., Lin, D.C., Bai, J.W., Mead, M., Zhao, Z.*, et al.* (2016). BCOR regulates myeloid cell proliferation and differentiation. Leukemia *30*, 1155-1165.

Casey, S.C., Tong, L., Li, Y., Do, R., Walz, S., Fitzgerald, K.N., Gouw, A.M., Baylot, V., Gutgemann, I., Eilers, M.*, et al.* (2016). MYC regulates the antitumor immune response through CD47 and PD-L1. Science *352*, 227-231.

Cattoretti, G., Chang, C.C., Cechova, K., Zhang, J., Ye, B.H., Falini, B., Louie, D.C., Offit, K., Chaganti, R.S., and Dalla-Favera, R. (1995). BCL-6 protein is expressed in germinal-center B cells. Blood *86*, 45-53.

Chang, T.C., Yu, D., Lee, Y.S., Wentzel, E.A., Arking, D.E., West, K.M., Dang, C.V., Thomas-Tikhonenko, A., and Mendell, J.T. (2008). Widespread microRNA repression by Myc contributes to tumorigenesis. Nat Genet *40*, 43-50.

Chen, B.J., Chapuy, B., Ouyang, J., Sun, H.H., Roemer, M.G., Xu, M.L., Yu, H., Fletcher, C.D., Freeman, G.J., Shipp, M.A.*, et al.* (2013). PD-L1 expression is characteristic of a subset of aggressive B-cell lymphomas and virus-associated malignancies. Clin Cancer Res *19*, 3462-3473.

Chen, K., Wallis, J.W., McLellan, M.D., Larson, D.E., Kalicki, J.M., Pohl, C.S., McGrath, S.D., Wendl, M.C., Zhang, Q., Locke, D.P.*, et al.* (2009). BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nature methods *6*, 677-681.

Cheng, E.H., Wei, M.C., Weiler, S., Flavell, R.A., Mak, T.W., Lindsten, T., and Korsmeyer, S.J. (2001). BCL-2, BCL-X(L) sequester BH3 domain-only molecules preventing BAX- and BAK-mediated mitochondrial apoptosis. Mol Cell *8*, 705-711.

Chin, L., and Gray, J.W. (2008). Translating insights from the cancer genome into clinical practice. Nature *452*, 553-563.

Choi, M., Scholl, U.I., Ji, W., Liu, T., Tikhonova, I.R., Zumbo, P., Nayir, A., Bakkaloglu, A., Ozen, S., Sanjad, S.*, et al.* (2009). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A *106*, 19096-19101.

Choi, W.-I., Jeon, B.-N., Yoon, J.-H., Koh, D.-I., Kim, M.-H., Yu, M.-Y., Lee, K.-M., Kim, Y., Kim, K., Hur, S.S.*, et al.* (2013). The proto-oncoprotein FBI-1 interacts with MBD3 to recruit the Mi-2/NuRD-HDAC complex and BCoR and to silence p21WAF/CDKN1A by DNA methylation. Nucleic Acids Research *41*, 6403-6420.

Chowdhury, M., Mihara, K., Yasunaga, S., Ohtaki, M., Takihara, Y., and Kimura, A. (2007). Expression of Polycomb-group (PcG) protein BMI-1 predicts prognosis in patients with acute myeloid leukemia. Leukemia *21*, 1116-1122.

Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E.S., and Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nature biotechnology *31*, 213-219.

Coiffier, B. (2001). Diffuse large cell lymphoma. Current opinion in oncology *13*, 325-334.

Conway, T., Wazny, J., Bromage, A., Zobel, J., and Beresford-Smith, B. (2012). Gossamer--a resource-efficient de novo assembler. Bioinformatics *28*, 1937-1938.

Corcoran, L.M., Cory, S., and Adams, J.M. (1985). Transposition of the immunoglobulin heavy chain enhancer to the myc oncogene in a murine plasmacytoma. CELL *40*, 71-79.

Corless, C.L., and Heinrich, M.C. (2008). Molecular pathobiology of gastrointestinal stromal sarcomas. Annual review of pathology *3*, 557-586.

Cox, A.D., and Der, C.J. (2010). Ras history: The saga continues. Small GTPases *1*, 2-27.

Cox, B.J., Vollmer, M., Tamplin, O., Lu, M., Biechele, S., Gertsenstein, M., van Campenhout, C., Floss, T., Kuhn, R., Wurst, W*., et al.* (2010). Phenotypic annotation of the mouse X chromosome. Genome Res *20*, 1154-1164.

Cuypers, H.T., Selten, G., Quint, W., Zijlstra, M., Maandag, E.R., Boelens, W., van Wezenbeek, P., Melief, C., and Berns, A. (1984). Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region. Cell *37*, 141-150.

Daley, G.Q., Van Etten, R.A., and Baltimore, D. (1990). Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. Science *247*, 824-830.

Damm, F., Chesnais, V., Nagata, Y., Yoshida, K., Scourzic, L., Okuno, Y., Itzykson, R., Sanada, M., Shiraishi, Y., Gelsi-Boyer, V.*, et al.* (2013). BCOR and BCORL1 mutations in myelodysplastic syndromes and related disorders. Blood *122*, 3169-3177.

Dang, C.V. (1999). c-Myc target genes involved in cell growth, apoptosis, and metabolism. Mol Cell Biol *19*, 1-11.

Dang, C.V., O'donnell, K.A., and Juopperi, T. (2005). The great MYC escape in tumorigenesis. Cancer Cell *8*, 177-178.

Dansen, T.B., Whitfield, J., Rostker, F., Brown-Swigart, L., and Evan, G.I. (2006). Specific requirement for Bax, not Bak, in Myc-induced apoptosis and tumor suppression in vivo. J Biol Chem *281*, 10890-10895.

Dasgupta, T., Coombes, B., and Brasfield, R.D. (1965). Primary Malignant Neoplasms of the Spleen. Surgery, gynecology & obstetrics *120*, 947-960.

Davey, F.R., Skarin, A.T., and Moloney, W.C. (1973). Pathology of splenic lymphoma. American journal of clinical pathology *59*, 95-103.

de Jong, J., de Ridder, J., van der Weyden, L., Sun, N., van Uitert, M., Berns, A., van Lohuizen, M., Jonkers, J., Adams, D.J., and Wessels, L.F. (2011). Computational identification of insertional mutagenesis targets for cancer gene discovery. Nucleic Acids Res *39*, e105.

De Keersmaecker, K., Atak, Z.K., Li, N., Vicente, C., Patchett, S., Girardi, T., Gianfelici, V., Geerdens, E., Clappier, E., Porcu, M.*, et al.* (2013). Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia. Nat Genet *45*, 186-190.

de Klein, A., van Kessel, A.G., Grosveld, G., Bartram, C.R., Hagemeijer, A., Bootsma, D., Spurr, N.K., Heisterkamp, N., Groffen, J., and Stephenson, J.R. (1982). A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. Nature *300*, 765-767.

Deininger, M., Buchdunger, E., and Druker, B.J. (2005). The development of imatinib as a therapeutic agent for chronic myeloid leukemia. Blood *105*, 2640-2653.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M.*, et al.* (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet *43*, 491-498.

Der, C.J., Krontiris, T.G., and Cooper, G.M. (1982). Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and Kirsten sarcoma viruses. Proc Natl Acad Sci U S A *79*, 3637-3640.

Derynck, R., Akhurst, R.J., and Balmain, A. (2001). TGF-beta signaling in tumor suppression and cancer progression. Nat Genet *29*, 117-129.

Dhillon, A.S., Hagan, S., Rath, O., and Kolch, W. (2007). MAP kinase signalling pathways in cancer. Oncogene *26*, 3279-3290.

Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., Mclellan, M.D.*, et al.* (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature *481*, 506-510.

Dreyfus, F., Sola, B., Fichelson, S., Varlet, P., Charon, M., Tambourin, P., Wendling, F., and Gisselbrecht, S. (1990). Rearrangements of the Pim-1, c-myc, and p53 genes in Friend helper virus-induced mouse erythroleukemias. Leukemia *4*, 590-594.

Duy, C., Yu, J.J., Nahar, R., Swaminathan, S., Kweon, S.-M., Polo, J.M., Valls, E., Klemm, L., Shojaee, S., Cerchietti, L.*, et al.* (2010). BCL6 is critical for the development of a diverse primary B cell repertoire. Journal of Experimental Medicine *207*, 1209-1221.

Ebner, R., Chen, R.H., Lawler, S., Zioncheck, T., and Derynck, R. (1993). Determination of type I receptor specificity by the type II receptors for TGF-beta or activin. Science *262*, 900-902.

Egle, A., Harris, A.W., Bouillet, P., and Cory, S. (2004). Bim is a suppressor of Myc-induced mouse B cell leukemia. Proceedings of the National Academy of Sciences of the United States of America *101*, 6164-6169.

Eilers, M., Schirm, S., and Bishop, J.M. (1991). The MYC protein activates transcription of the alpha-prothymosin gene. EMBO J *10*, 133-141.

Eischen, C.M., Weber, J.D., Roussel, M.F., Sherr, C.J., and Cleveland, J.L. (1999). Disruption of the ARF-Mdm2-p53 tumor suppressor pathway in Myc-induced lymphomagenesis. Genes & development *13*, 2658-2669.

Eischen, C.M., Woo, D., Roussel, M.F., and Cleveland, J.L. (2001). Apoptosis triggered by Myc-induced suppression of Bcl-X(L) or Bcl-2 is bypassed during lymphomagenesis. Mol Cell Biol *21*, 5063-5070.

Estey, E.H. (2009). Treatment of acute myeloid leukemia. Haematologica *94*, 10-16.

Evan, G.I., Wyllie, A.H., Gilbert, C.S., Littlewood, T.D., Land, H., Brooks, M., Waters, C.M., Penn, L.Z., and Hancock, D.C. (1992). Induction of apoptosis in fibroblasts by c-myc protein. Cell *69*, 119-128.

Falkenberg, K.J., Gould, C.M., Johnstone, R.W., and Simpson, K.J. (2014). Genome-wide functional genomic and transcriptomic analyses for genes regulating sensitivity to vorinostat. Scientific data *1*, 140017.

Falkenberg, K.J., Newbold, A., Gould, C.M., Luu, J., Trapani, J.A., Matthews, G.M., Simpson, K.J., and Johnstone, R.W. (2016). A genome scale RNAi screen identifies GLI1 as a novel gene regulating vorinostat sensitivity. Cell Death Differ *23*, 1209-1218.

Felsher, D.W., and Bishop, J.M. (1999a). Reversible tumorigenesis by MYC in hematopoietic lineages. Mol Cell *4*, 199-207.

Felsher, D.W., and Bishop, J.M. (1999b). Transient excess of MYC activity can elicit genomic instability and tumorigenesis. Proc Natl Acad Sci U S A *96*, 3940-3944.

Fernandez, P.C., Frank, S.R., Wang, L., Schroeder, M., Liu, S., Greene, J., Cocito, A., and Amati, B. (2003). Genomic targets of the human c-Myc protein. Genes & development *17*, 1115-1129.

Ferry, J.A. (2006). Burkitt's lymphoma: clinicopathologic features and differential diagnosis. The oncologist *11*, 375-383.

Flores, I., Murphy, D.J., Swigart, L.B., Knies, U., and Evan, G.I. (2004). Defining the temporal requirements for Myc in the progression and maintenance of skin neoplasia. Oncogene *23*, 5923-5930.

Foa, R., Vitale, A., Vignetti, M., Meloni, G., Guarini, A., De Propris, M.S., Elia, L., Paoloni, F., Fazi, P., Cimino, G.*, et al.* (2011). Dasatinib as first-line treatment for adult patients with Philadelphia chromosome-positive acute lymphoblastic leukemia. Blood *118*, 6521-6528.

Frenzel, A., Labi, V., Chmelewskij, W., Ploner, C., Geley, S., Fiegl, H., Tzankov, A., and Villunger, A. (2010). Suppression of B-cell lymphomagenesis by the BH3-only proteins Bmf and Bad. Blood *115*, 995-1005.

Frost, M., Newell, J., Lones, M.A., Tripp, S.R., Cairo, M.S., and Perkins, S.L. (2004). Comparative immunohistochemical analysis of pediatric Burkitt lymphoma and diffuse large B-cell lymphoma. American journal of clinical pathology *121*, 384-392.

Fusello, A., Horowitz, J., Yang-Iott, K., Brady, B.L., Yin, B., Rowh, M.A., Rappaport, E., and Bassing, C.H. (2013). Histone H2AX suppresses translocations in lymphomas of Emu-c-Myc transgenic mice that contain a germline amplicon of tumor-promoting genes. Cell cycle *12*, 2867-2875.

Gearhart, M.D., Corcoran, C.M., Wamstad, J.A., and Bardwell, V.J. (2006). Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets. Molecular and Cellular Biology *26*, 6880-6889.

Ghetu, A.F., Corcoran, C.M., Cerchietti, L., Bardwell, V.J., Melnick, A., and Privé, G.G. (2008). Structure of a BCOR corepressor peptide in complex with the BCL6 BTB domain dimer. Molecular cell *29*, 384-391.

Giuriato, S., Ryeom, S., Fan, A.C., Bachireddy, P., Lynch, R.C., Rioth, M.J., van Riggelen, J., Kopelman, A.M., Passegue, E., Tang, F.*, et al.* (2006). Sustained regression of tumors upon MYC inactivation requires p53 or thrombospondin-1 to reverse the angiogenic switch. Proc Natl Acad Sci U S A *103*, 16266-16271.

Gomez-Roman, N., Grandori, C., Eisenman, R.N., and White, R.J. (2003). Direct activation of RNA polymerase III transcription by c-Myc. Nature *421*, 290-294.

Gorre, M.E., Mohammed, M., Ellwood, K., Hsu, N., Paquette, R., Rao, P.N., and Sawyers, C.L. (2001). Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. Science *293*, 876-880.

Grabiner, B.C., Nardi, V., Birsoy, K., Possemato, R., Shen, K., Sinha, S., Jordan, A., Beck, A.H., and Sabatini, D.M. (2014). A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity. Cancer discovery *4*, 554-563.

Grandori, C., Gomez-Roman, N., Felton-Edkins, Z.A., Ngouenet, C., Galloway, D.A., Eisenman, R.N., and White, R.J. (2005). c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. Nat Cell Biol *7*, 311-318.

Greaves, M., and Maley, C.C. (2012a). Clonal evolution in cancer. Nature *481*, 306-313.

Greaves, M., and Maley, C.C. (2012b). Clonal evolution in cancer. Nature *481*, 306-313.

Green, M.R., Monti, S., Rodig, S.J., Juszczynski, P., Currie, T., O'Donnell, E., Chapuy, B., Takeyama, K., Neuberg, D., Golub, T.R.*, et al.* (2010). Integrative analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. Blood *116*, 3268-3277.

Grossmann, V., Tiacci, E., Holmes, A.B., Kohlmann, A., Martelli, M.P., Kern, W., Spanhol-Rosseto, A., Klein, H.-U., Dugas, M., Schindela, S.*, et al.* (2011). Whole-exome sequencing identifies mutations of BCOR in acute myeloid leukemia with normal karyotype. Blood, 1-43.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646-674.

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

Harris, A.W., Pinkert, C.A., Crawford, M., Langdon, W.Y., Brinster, R.L., and Adams, J.M. (1988). The E mu-myc transgenic mouse. A model for high-incidence spontaneous lymphoma and leukemia of early B cells. J Exp Med *167*, 353-371.

Harrison, C.J. (2009). Cytogenetics of paediatric and adolescent acute lymphoblastic leukaemia. British journal of haematology *144*, 147-156.

Harrison, C.J. (2013). Targeting signaling pathways in acute lymphoblastic leukemia: new insights. Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program *2013*, 118-125.

Hartwell, L. (1992). Defects in a cell cycle checkpoint may be responsible for the genomic instability of cancer cells. Cell *71*, 543-546.

Hatzi, K., and Melnick, A. (2014). Breaking bad in the germinal center: how deregulation of BCL6 contributes to lymphomagenesis. Trends in molecular medicine *20*, 343-352.

Haupt, Y., Alexander, W.S., Barri, G., Klinken, S.P., and Adams, J.M. (1991). Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in E mu-myc transgenic mice. CELL *65*, 753-763.

Haupt, Y., Bath, M.L., Harris, A.W., and Adams, J.M. (1993). bmi-1 transgene induces lymphomas and collaborates with myc in tumorigenesis. Oncogene *8*, 3161-3164.

Hawkins, E.D., Hommel, M., Turner, M.L., Battye, F.L., Markham, J.F., and Hodgkin, P.D. (2007). Measuring lymphocyte proliferation, survival and differentiation using CFSE time-series data. Nature protocols *2*, 2057-2067.

He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J.*, et al.* (2005). A microRNA polycistron as a potential human oncogene. Nature *435*, 828-833.

Heinzel, S., Binh Giang, T., Kan, A., Marchingo, J.M., Lye, B.K., Corcoran, L.M., and Hodgkin, P.D. (2017). A Myc-dependent division timer complements a cell-death timer to regulate T cell and B cell responses. Nat Immunol *18*, 96-103.

Hemann, M.T., Bric, A., Teruya-Feldstein, J., Herbst, A., Nilsson, J.A., Cordon-Cardo, C., Cleveland, J.L., Tansey, W.P., and Lowe, S.W. (2005). Evasion of the p53 tumour surveillance network by tumour-derived MYC mutants. Nature *436*, 807-811.

Hemann, M.T., Zilfou, J.T., Zhao, Z., Burgess, D.J., Hannon, G.J., and Lowe, S.W. (2004). Suppression of tumorigenesis by the p53 target PUMA. Proceedings of the National Academy of Sciences of the United States of America *101*, 9333-9338.

Henderson, A., Ripley, S., Heller, M., and Kieff, E. (1983). Chromosome site for Epstein-Barr virus DNA in a Burkitt tumor cell line and in lymphocytes growth-transformed in vitro. Proc Natl Acad Sci U S A *80*, 1987-1991.

Hermeking, H., and Eick, D. (1994). Mediation of c-Myc-induced apoptosis by p53. Science *265*, 2091-2093.

Hideshima, T., Mitsiades, C., Tonon, G., Richardson, P.G., and Anderson, K.C. (2007). Understanding multiple myeloma pathogenesis in the bone marrow to identify new therapeutic targets. Nature reviews Cancer *7*, 585-598.

Hofmann, W.K., Komor, M., Wassmann, B., Jones, L.C., Gschaidmeier, H., Hoelzer, D., Koeffler, H.P., and Ottmann, O.G. (2003). Presence of the BCR-ABL mutation Glu255Lys prior to STI571 (imatinib) treatment in patients with Ph+ acute lymphoblastic leukemia. Blood *102*, 659-661.

Hogg, S.J., Vervoort, S.J., Deswal, S., Ott, C.J., Li, J., Cluse, L.A., Beavis, P.A., Darcy, P.K., Martin, B.P., Spencer, A.*, et al.* (2017). BET-Bromodomain Inhibitors Engage the Host Immune System and Regulate Expression of the Immune Checkpoint Ligand PD-L1. Cell Rep *18*, 2162-2174.

Howlader N, N.A., Krapcho M, Miller D, Bishop K, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (2016). SEER Cancer Statistics Review, 1975-2014 (https://seer.cancer.gov/csr/1975_2014/: National Cancer Institute. Bethesda, MD).

Hsu, B., Marin, M.C., el-Naggar, A.K., Stephens, L.C., Brisbay, S., and McDonnell, T.J. (1995). Evidence that c-myc mediated apoptosis does not require wild-type p53 during lymphomagenesis. Oncogene *11*, 175-179.

Hu, S., Xu-Monette, Z.Y., Tzankov, A., Green, T., Wu, L., Balasubramanyam, A., Liu, W.M., Visco, C., Li, Y., Miranda, R.N.*, et al.* (2013). MYC/BCL2 protein coexpression contributes to the inferior survival of activated B-cell subtype of diffuse large B-cell lymphoma and demonstrates high-risk gene expression signatures: a report from The International DLBCL Rituximab-CHOP Consortium Program. Blood *121*, 4021-4031; quiz 4250.

Huang, J.Z., Sanger, W.G., Greiner, T.C., Staudt, L.M., Weisenburger, D.D., Pickering, D.L., Lynch, J.C., Armitage, J.O., Warnke, R.A., Alizadeh, A.A.*, et al.* (2002). The t(14;18) defines a unique subset of diffuse large B-cell lymphoma with a germinal center B-cell gene expression profile. Blood *99*, 2285-2290.

Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T.*, et al.* (2002). The Ensembl genome database project. Nucleic Acids Res *30*, 38-41.

Huynh, K.D., Fischle, W., Verdin, E., and Bardwell, V.J. (2000). BCoR, a novel corepressor involved in BCL-6 repression. Genes & Development *14*, 1810-1823.

Inaba, H., Greaves, M., and Mullighan, C.G. (2013). Acute lymphoblastic leukaemia. Lancet *381*, 1943-1955.

International Cancer Genome, C., Hudson, T.J., Anderson, W., Artez, A., Barker, A.D., Bell, C., Bernabe, R.R., Bhan, M.K., Calvo, F., Eerola, I.*, et al.* (2010). International network of cancer genome projects. Nature *464*, 993-998.

Iritani, B.M., and Eisenman, R.N. (1999). c-Myc enhances protein synthesis and cell size during B lymphocyte development. Proc Natl Acad Sci U S A *96*, 13180-13185.

Jacobs, J.J., Scheijen, B., Voncken, J.W., Kieboom, K., Berns, A., and van Lohuizen, M. (1999). Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. Genes & development *13*, 2678-2690.

Jacobs, J.J., and van Lohuizen, M. (2002). Polycomb repression: from cellular memory to cellular proliferation and cancer. Biochimica et biophysica acta *1602*, 151-161.

Jacobsen, K.A., Prasad, V.S., Sidman, C.L., and Osmond, D.G. (1994). Apoptosis and macrophage-mediated deletion of precursor B cells in the bone marrow of E mu-myc transgenic mice. Blood *84*, 2784-2794.

Jain, M., Arvanitis, C., Chu, K., Dewey, W., Leonhardt, E., Trinh, M., Sundberg, C.D., Bishop, J.M., and Felsher, D.W. (2002). Sustained loss of a neoplastic phenotype by brief inactivation of MYC. Science *297*, 102-104.

Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. CA: a cancer journal for clinicians *61*, 69-90.

Jemal, A., Siegel, R., Xu, J., and Ward, E. (2010). Cancer statistics, 2010. CA: a cancer journal for clinicians *60*, 277-300.

Johnson, N.A., Savage, K.J., Ludkovski, O., Ben-Neriah, S., Woods, R., Steidl, C., Dyer, M.J., Siebert, R., Kuruvilla, J., Klasa, R.*, et al.* (2009). Lymphomas with concurrent BCL2 and MYC translocations: the critical factors associated with survival. Blood *114*, 2273-2279.

Junco, S.E., Wang, R., Gaipa, J.C., Taylor, A.B., Schirf, V., Gearhart, M.D., Bardwell, V.J., Demeler, B., Hart, P.J., and Kim, C.A. (2013). Structure of the Polycomb Group Protein PCGF1 in Complex with BCOR Reveals Basis for Binding Selectivity of PCGF Homologs. Structure/Folding and Design *21*, 665-671.

Kamijo, T., Weber, J.D., Zambetti, G., Zindy, F., Roussel, M.F., and Sherr, C.J. (1998). Functional and physical interactions of the ARF tumor suppressor with p53 and Mdm2. Proc Natl Acad Sci U S A *95*, 8292-8297.

Karlsson, A., Giuriato, S., Tang, F., Fung-Weier, J., Levan, G., and Felsher, D.W. (2003). Genomically complex lymphomas undergo sustained tumor regression upon MYC inactivation unless they acquire novel chromosomal translocations. Blood *101*, 2797-2803.

Kastan, M.B., Onyekwere, O., Sidransky, D., Vogelstein, B., and Craig, R.W. (1991). Participation of p53 protein in the cellular response to DNA damage. Cancer Res *51*, 6304-6311.

Katoh, M., and Katoh, M. (2003). FGFR2 and WDR11 are neighboring oncogene and tumor suppressor gene on human chromosome 10q26. International journal of oncology *22*, 1155-1159.

Kelly, P.N., Puthalakath, H., Adams, J.M., and Strasser, A. (2007). Endogenous bcl-2 is not required for the development of Emu-myc-induced B-cell lymphoma. Blood *109*, 4907-4913.

Klapproth, K., and Wirth, T. (2010). Advances in the understanding of MYC-induced lymphomagenesis. British journal of haematology *149*, 484-497.

Knudson, A.G. (2002). Cancer genetics. American journal of medical genetics *111*, 96-102.

Konopka, J.B., Watanabe, S.M., Singer, J.W., Collins, S.J., and Witte, O.N. (1985). Cell lines and clinical isolates derived from Ph1-positive chronic myelogenous leukemia patients express c-abl proteins with a common structural alteration. Proc Natl Acad Sci U S A *82*, 1810-1814.

Kool, J., and Berns, A. (2009). High-throughput insertional mutagenesis screens in mice to identify oncogenic networks. Nature reviews Cancer *9*, 389-399.

Kramer, M.H., Hermans, J., Wijburg, E., Philippo, K., Geelen, E., van Krieken, J.H., de Jong, D., Maartense, E., Schuuring, E., and Kluin, P.M. (1998). Clinical relevance of BCL2, BCL6, and MYC rearrangements in diffuse large B-cell lymphoma. Blood *92*, 3152-3162.

Kuppers, R. (2009). The biology of Hodgkin's lymphoma. Nature reviews Cancer *9*, 15-27.

Kvansakul, M., Yang, H., Fairlie, W.D., Czabotar, P.E., Fischer, S.F., Perugini, M.A., Huang, D.C., and Colman, P.M. (2008). Vaccinia virus anti-apoptotic F1L is a novel Bcl-2-like domain-swapped dimer that binds a highly selective subset of BH3-containing death ligands. Cell Death Differ *15*, 1564-1571.

Kwanhian, W., Lenze, D., Alles, J., Motsch, N., Barth, S., Doll, C., Imig, J., Hummel, M., Tinguely, M., Trivedi, P.*, et al.* (2012). MicroRNA-142 is mutated in about 20% of diffuse large B-cell lymphoma. Cancer medicine *1*, 141-155.

Kyle, R.A., and Rajkumar, S.V. (2008). Multiple myeloma. Blood *111*, 2962-2972.

Landau, D.A., Carter, S.L., Getz, G., and Wu, C.J. (2014). Clonal evolution in hematological malignancies and therapeutic implications. Leukemia *28*, 34-43.

Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome biology *15*, R29.

Lee, S.C., Phipson, B., Hyland, C.D., Leong, H.S., Allan, R.S., Lun, A., Hilton, D.J., Nutt, S.L., Blewitt, M.E., Smyth, G.K.*, et al.* (2013). Polycomb repressive complex 2 (PRC2) suppresses Emu-myc lymphoma. Blood *122*, 2654-2663.

Lefebure, M., Tothill, R.W., Kruse, E., Hawkins, E.D., Shortt, J., Matthews, G.M., Gregory, G.P., Martin, B.P., Kelly, M.J., Todorovski, I.*, et al.* (2017). Genomic characterisation of Emu-Myc mouse lymphomas identifies Bcor as a Myc co-operative tumour-suppressor gene. Nature communications *8*, 14581.

Lenz, G., and Staudt, L.M. (2010). Aggressive lymphomas. The New England journal of medicine *362*, 1417-1429.

Lenz, G., Wright, G.W., Emre, N.C.T., Kohlhammer, H., Dave, S.S., Davis, R.E., Carty, S., Lam, L.T., Shaffer, A.L., Xiao, W.*, et al.* (2008). Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. Proceedings of the National Academy of Sciences of the United States of America *105*, 13520-13525.

Ley, T.J., Ding, L., Walter, M.J., McLellan, M.D., Lamprecht, T., Larson, D.E., Kandoth, C., Payton, J.E., Baty, J., Welch, J.*, et al.* (2010). DNMT3A mutations in acute myeloid leukemia. N Engl J Med *363*, 2424-2433.

Li, Y.S., Wasserman, R., Hayakawa, K., and Hardy, R.R. (1996). Identification of the earliest B lineage stage in mouse bone marrow. Immunity *5*, 527-535.

Lievre, A., Bachet, J.B., Le Corre, D., Boige, V., Landi, B., Emile, J.F., Cote, J.F., Tomasic, G., Penna, C., Ducreux, M.*, et al.* (2006). KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. Cancer Res *66*, 3992-3995.

Lin, C.Y., Loven, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. Cell *151*, 56-67.

Lindsten, T., Ross, A.J., King, A., Zong, W.X., Rathmell, J.C., Shiels, H.A., Ulrich, E., Waymire, K.G., Mahar, P., Frauwirth, K.*, et al.* (2000). The combined functions of proapoptotic Bcl-2 family members bak and bax are essential for normal development of multiple tissues. Mol Cell *6*, 1389-1399.

Lindstrom, M.S., Klangby, U., and Wiman, K.G. (2001). p14ARF homozygous deletion or MDM2 overexpression in Burkitt lymphoma lines carrying wild type p53. Oncogene *20*, 2171-2177.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

Lobry, C., Oh, P., and Aifantis, I. (2011). Oncogenic and tumor suppressor functions of Notch in cancer: it's NOTCH what you think. J Exp Med *208*, 1931-1935.

Loeb, L.A. (1991). Mutator phenotype may be required for multistage carcinogenesis. Cancer Res *51*, 3075-3079.

Loffert, D., Schaal, S., Ehlich, A., Hardy, R.R., Zou, Y.R., Muller, W., and Rajewsky, K. (1994). Early B-cell development in the mouse: insights from mutations introduced by gene targeting. Immunological reviews *137*, 135-153.

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.*, et al.* (2012). Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proceedings of the National Academy of Sciences, 1-6.

Love, C., Sun, Z., Jima, D., Li, G., Zhang, J., Miles, R., Richards, K.L., Dunphy, C.H., Choi, W.W., Srivastava, G.*, et al.* (2012a). The genetic landscape of mutations in Burkitt lymphoma. Nat Genet *44*, 1321-1325.

Love, C., Sun, Z., Jima, D., Li, G., Zhang, J., Miles, R., Richards, K.L., Dunphy, C.H., Choi, W.W.L., Srivastava, G.*, et al.* (2012b). The genetic landscape of mutations in Burkitt lymphoma. Nature Genetics.

Lowe, S.W. (1999). Activation of p53 by oncogenes. Endocrine-related cancer *6*, 45-48.

Lowe, S.W., Cepero, E., and Evan, G. (2004). Intrinsic tumour suppression. Nature *432*, 307-315.

Lowy, D.R., and Willumsen, B.M. (1993). Function and regulation of ras. Annual review of biochemistry *62*, 851-891.

Lu, X., Magrane, G., Yin, C., Louis, D.N., Gray, J., and Van Dyke, T. (2001). Selective inactivation of p53 facilitates mouse epithelial tumor progression without chromosomal instability. Mol Cell Biol *21*, 6017-6030.

Luscher, B., and Vervoorts, J. (2012). Regulation of gene transcription by the oncoprotein MYC. Gene *494*, 145-160.

Macleod, K.F., and Jacks, T. (1999). Insights into cancer from transgenic mouse models. J Pathol *187*, 43-60.

Maher, C.A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., Sam, L., Barrette, T., Palanisamy, N., and Chinnaiyan, A.M. (2009). Transcriptome sequencing to detect gene fusions in cancer. Nature *458*, 97-101.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.

Malumbres, M., and Barbacid, M. (2003). RAS oncogenes: the first 30 years. Nature reviews Cancer *3*, 459-465.

Mardis, E.R., and Wilson, R.K. (2009). Cancer genome sequencing: a review. Hum Mol Genet *18*, R163-168.

Martinou, J.C., and Youle, R.J. (2011). Mitochondria in apoptosis: Bcl-2 family members and mitochondrial dynamics. Dev Cell *21*, 92-101.

Marusyk, A., Almendro, V., and Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? Nature reviews Cancer *12*, 323-334.

Marusyk, A., and Polyak, K. (2010). Tumor heterogeneity: causes and consequences. Biochimica et biophysica acta *1805*, 105-117.

Mason, K.D., Vandenberg, C.J., Scott, C.L., Wei, A.H., Cory, S., Huang, D.C., and Roberts, A.W. (2008). In vivo efficacy of the Bcl-2 antagonist ABT-737 against aggressive Myc-driven lymphomas. Proc Natl Acad Sci U S A *105*, 17961-17966.

Massague, J. (1998). TGF-beta signal transduction. Annual review of biochemistry *67*, 753-791.

Mateyak, M.K., Obaya, A.J., Adachi, S., and Sedivy, J.M. (1997). Phenotypes of c-Myc-deficient rat fibroblasts isolated by targeted homologous recombination. Cell growth &

differentiation : the molecular biology journal of the American Association for Cancer Research *8*, 1039-1048.

Mattison, J., van der Weyden, L., Hubbard, T., and Adams, D.J. (2009). Cancer gene discovery in mouse and man. Biochimica et biophysica acta *1796*, 140-161.

Mccabe, M.T., Ott, H.M., Ganji, G., Korenchuk, S., Thompson, C., Aller, G.S.V., Liu, Y., Graves, A.P., Iii, A.D.P., Diaz, E.*, et al.* (2012). EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. Nature, 1-7.

McDonnell, T.J., Deane, N., Platt, F.M., Nunez, G., Jaeger, U., McKearn, J.P., and Korsmeyer, S.J. (1989). bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation. Cell *57*, 79-88.

McDonnell, T.J., Nunez, G., Platt, F.M., Hockenberry, D., London, L., McKearn, J.P., and Korsmeyer, S.J. (1990). Deregulated Bcl-2-immunoglobulin transgene expands a resting but responsive immunoglobulin M and D-expressing B-cell population. Mol Cell Biol *10*, 1901-1907.

McGranahan, N., and Swanton, C. (2017). Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell *168*, 613-628.

McIntyre, R.E., van der Weyden, L., and Adams, D.J. (2012). Cancer gene discovery in the mouse. Current Opinion in Genetics & Development *22*, 14-20.

Merlo, L.M., Pepper, J.W., Reid, B.J., and Maley, C.C. (2006). Cancer as an evolutionary and ecological process. Nature reviews Cancer *6*, 924-935.

Metzker, M.L. (2010). Sequencing technologies - the next generation. Nature reviews Genetics *11*, 31-46.

Meyer, N., and Penn, L.Z. (2008). Reflecting on 25 years with MYC. Nature reviews Cancer *8*, 976-990.

Meyerson, M. (2007). Cancer: broken genes in solid tumours. Nature *448*, 545-546.

Meyerson, M., Gabriel, S., and Getz, G. (2010). Advances in understanding cancer genomes through second-generation sequencing. Nature reviews Genetics *11*, 685-696.

Mi, H., and Thomas, P. (2009). PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. Methods in molecular biology *563*, 123-140.

Miao, Z., Luker, K.E., Summers, B.C., Berahovich, R., Bhojani, M.S., Rehemtulla, A., Kleer, C.G., Essner, J.J., Nasevicius, A., Luker, G.D.*, et al.* (2007). CXCR7 (RDC1) promotes breast and lung tumor growth in vivo and is expressed on tumor-associated vasculature. Proc Natl Acad Sci U S A *104*, 15735-15740.

Michalak, E.M., Jansen, E.S., Happo, L., Cragg, M.S., Tai, L., Smyth, G.K., Strasser, A., Adams, J.M., and Scott, C.L. (2009). Puma and to a lesser extent Noxa are suppressors of Myc-induced lymphomagenesis. Cell death and differentiation *16*, 684-696.

Miller, R.W., Young, J.L., Jr., and Novakovic, B. (1995). Childhood cancer. Cancer *75*, 395-405.

Miranda Peralta, E.I., Valles Ayoub, Y., Hernandez Mendoza, L., Rangel Ramirez, L.M., Castrejon Rojas, A., Collazo-Jaloma, J., Gutierrez Romero, M., Gonzalez Constance, R., and Gariglio Vidal, P. (1991). [MYC protein and proteins antigenically related with MYC in acute lymphoblastic leukemia]. Revista de investigacion clinica; organo del Hospital de Enfermedades de la Nutricion *43*, 139-145.

Misale, S., Yaeger, R., Hobor, S., Scala, E., Janakiraman, M., Liska, D., Valtorta, E., Schiavo, R., Buscarino, M., Siravegna, G.*, et al.* (2012). Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. Nature *486*, 532-536.

Moore, J.P., Hancock, D.C., Littlewood, T.D., and Evan, G.I. (1987). A sensitive and quantitative enzyme-linked immunosorbence assay for the c-myc and N-myc oncoproteins. Oncogene research *2*, 65-80.

Mori, H., Colman, S.M., Xiao, Z., Ford, A.M., Healy, L.E., Donaldson, C., Hows, J.M., Navarrete, C., and Greaves, M. (2002). Chromosome translocations and covert

leukemic clones are generated during normal fetal development. Proc Natl Acad Sci U S A *99*, 8242-8247.

Mori, S., Rempel, R.E., Chang, J.T., Yao, G., Lagoo, A.S., Potti, A., Bild, A., and Nevins, J.R. (2008). Utilization of pathway signatures to reveal distinct types of B lymphoma in the Emicro-myc model and human diffuse large B-cell lymphoma. Cancer Research *68*, 8525-8534.

Morin, P.J., Vogelstein, B., and Kinzler, K.W. (1996). Apoptosis and APC in colorectal tumorigenesis. Proc Natl Acad Sci U S A *93*, 7950-7954.

Morin, R.D., Johnson, N.A., Severson, T.M., Mungall, A.J., An, J., Goya, R., Paul, J.E., Boyle, M., Woolcock, B.W., Kuchenbauer, F.*, et al.* (2010). Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet *42*, 181-185.

Morin, R.D., Mendez-Lago, M., Mungall, A.J., Goya, R., Mungall, K.L., Corbett, R.D., Johnson, N.A., Severson, T.M., Chiu, R., Field, M.*, et al.* (2011). Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature, 1-6.

Morin, R.D., Mungall, K., Pleasance, E., Mungall, A.J., Goya, R., Huff, R.D., Scott, D.W., Ding, J., Roth, A., Chiu, R.*, et al.* (2013). Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. Blood *122*, 1256-1265.

Morton, L.M., Wang, S.S., Devesa, S.S., Hartge, P., Weisenburger, D.D., and Linet, M.S. (2006). Lymphoma incidence patterns by WHO subtype in the United States, 1992-2001. Blood *107*, 265-276.

Muller-Hermelink, H.K., Zettl, A., Pfeifer, W., and Ott, G. (2001). Pathology of lymphoma progression. Histopathology *38*, 285-306.

Mullighan, C.G. (2012). Molecular genetics of B-precursor acute lymphoblastic leukemia. The Journal of clinical investigation *122*, 3407-3415.

Mullighan, C.G. (2013). Genome sequencing of lymphoid malignancies. Blood *122*, 3899-3907.

Mullighan, C.G., Phillips, L.A., Su, X., Ma, J., Miller, C.B., Shurtleff, S.A., and Downing, J.R. (2008). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. Science *322*, 1377-1380.

Mullighan, C.G., Zhang, J., Harvey, R.C., Collins-Underwood, J.R., Schulman, B.A., Phillips, L.A., Tasian, S.K., Loh, M.L., Su, X., Liu, W*., et al.* (2009). JAK mutations in high-risk childhood acute lymphoblastic leukemia. Proc Natl Acad Sci U S A *106*, 9414-9418.

Nagasawa, T. (2006). Microenvironmental niches in the bone marrow required for B-cell development. Nature reviews Immunology *6*, 107-116.

Nakano, K., and Vousden, K.H. (2001). PUMA, a novel proapoptotic gene, is induced by p53. Mol Cell *7*, 683-694.

Nazarian, R., Shi, H., Wang, Q., Kong, X., Koya, R.C., Lee, H., Chen, Z., Lee, M.K., Attar, N., Sazegar, H.*, et al.* (2010). Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. Nature *468*, 973-977.

Ng, D., Thakker, N., Corcoran, C.M., Donnai, D., Perveen, R., Schneider, A., Hadley, D.W., Tifft, C., Zhang, L., Wilkie, A.O.M.*, et al.* (2004). Oculofaciocardiodental and Lenz microphthalmia syndromes result from distinct classes of mutations in BCOR. Nature Genetics *36*, 411-416.

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R.*, et al.* (2012). c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. Cell *151*, 68-79.

Nowell, P.C. (1976). The clonal evolution of tumor cell populations. Science *194*, 23-28.

O'Brien, P., Morin, P., Jr., Ouellette, R.J., and Robichaud, G.A. (2011). The Pax-5 gene: a pluripotent regulator of B-cell differentiation and cancer disease. Cancer Res *71*, 7345-7350.

O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M., Cory, S., and Huang, D.C. (1998). Bim: a novel member of the Bcl-2 family that promotes apoptosis. EMBO J *17*, 384-395.

Olivier, M., Hollstein, M., and Hainaut, P. (2010). TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harbor perspectives in biology *2*, a001008.

Orian, A., van Steensel, B., Delrow, J., Bussemaker, H.J., Li, L., Sawado, T., Williams, E., Loo, L.W., Cowley, S.M., Yost, C.*, et al.* (2003). Genomic binding by the Drosophila Myc, Max, Mad/Mnt transcription factor network. Genes & development *17*, 1101-1114.

Ozsolak, F., and Milos, P.M. (2011). RNA sequencing: advances, challenges and opportunities. Nature reviews Genetics *12*, 87-98.

Papaemmanuil, E., Rapado, I., Li, Y., Potter, N.E., Wedge, D.C., Tubio, J., Alexandrov, L.B., Van Loo, P., Cooke, S.L., Marshall, J.*, et al.* (2014). RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. Nat Genet *46*, 116-125.

Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V.A., Grunn, A., Messina, M., Elliot, O.*, et al.* (2011). Analysis of the coding genome of diffuse large B-cell lymphoma. Nature Genetics.

Pelengaris, S., Khan, M., and Evan, G.I. (2002). Suppression of Myc-induced apoptosis in beta cells exposes multiple oncogenic properties of Myc and triggers carcinogenic progression. Cell *109*, 321-334.

Pelengaris, S., Littlewood, T., Khan, M., Elia, G., and Evan, G. (1999). Reversible activation of c-Myc in skin: induction of a complex neoplastic phenotype by a single oncogenic lesion. Mol Cell *3*, 565-577.

Phan, R.T., and Dalla-Favera, R. (2004). The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. Nature *432*, 635-639.

Polo, J.M., Dell'Oso, T., Ranuncolo, S.M., Cerchietti, L., Beck, D., Da Silva, G.F., Privé, G.G., Licht, J.D., and Melnick, A. (2004). Specific peptide interference reveals BCL6 transcriptional and oncogenic mechanisms in B-cell lymphoma cells. Nature medicine *10*, 1329-1335.

Puente, X.S., Bea, S., Valdes-Mas, R., Villamor, N., Gutierrez-Abril, J., Martin-Subero, J.I., Munar, M., Rubio-Perez, C., Jares, P., Aymerich, M*., et al.* (2015). Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature *526*, 519-524.

Puente, X.S., Pinyol, M., Quesada, V., Conde, L., Ordonez, G.R., Villamor, N., Escaramis, G., Jares, P., Bea, S., Gonzalez-Diaz, M*., et al.* (2011). Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature *475*, 101-105.

Pugh, T.J., Weeraratne, S.D., Archer, T.C., Pomeranz Krummel, D.A., Auclair, D., Bochicchio, J., Carneiro, M.O., Carter, S.L., Cibulskis, K., Erlich, R.L.*, et al.* (2012). Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. Nature *488*, 106-110.

Pui, C.H., and Evans, W.E. (1998). Acute lymphoblastic leukemia. N Engl J Med *339*, 605-615.

Pui, C.H., Robison, L.L., and Look, A.T. (2008). Acute lymphoblastic leukaemia. Lancet *371*, 1030-1043.

Quelle, D.E., Zindy, F., Ashmun, R.A., and Sherr, C.J. (1995). Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. Cell *83*, 993-1000.

Quentmeier, H., MacLeod, R.A., Zaborski, M., and Drexler, H.G. (2006). JAK2 V617F tyrosine kinase mutation in cell lines derived from myeloproliferative disorders. Leukemia *20*, 471-476.

Quesada, V., Conde, L., Villamor, N., Ordonez, G.R., Jares, P., Bassaganyas, L., Ramsay, A.J., Bea, S., Pinyol, M., Martinez-Trillos, A.*, et al.* (2011). Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet *44*, 47-52.

Raaphorst, F.M. (2005). Deregulated expression of Polycomb-group oncogenes in human malignant lymphomas and epithelial tumors. Hum Mol Genet *14 Spec No 1*, R93-R100.

Ray, S., Atkuri, K.R., Deb-Basu, D., Adler, A.S., Chang, H.Y., Herzenberg, L.A., and Felsher, D.W. (2006). MYC can induce DNA breaks in vivo and in vitro independent of reactive oxygen species. Cancer Research *66*, 6598-6605.

Reddy, E.P., Reynolds, R.K., Santos, E., and Barbacid, M. (1982). A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature *300*, 149-152.

Reimann, M., Lee, S., Loddenkemper, C., Dorr, J.R., Tabor, V., Aichele, P., Stein, H., Dorken, B., Jenuwein, T., and Schmitt, C.A. (2010). Tumor stroma-derived TGF-beta limits myc-driven lymphomagenesis via Suv39h1-dependent senescence. Cancer Cell *17*, 262-272.

Reimann, M., Loddenkemper, C., Rudolph, C., Schildhauer, I., Teichmann, B., Stein, H., Schlegelberger, B., Dörken, B., and Schmitt, C.A. (2007). The Myc-evoked DNA damage response accounts for treatment resistance in primary lymphomas in vivo. Blood *110*, 2996-3004.

Richter-Larrea, J.A., Robles, E.F., Fresquet, V., Beltran, E., Rullan, A.J., Agirre, X., Calasanz, M.J., Panizo, C., Richter, J.A., Hernandez, J.M.*, et al.* (2010). Reversion of

epigenetically mediated BIM silencing overcomes chemoresistance in Burkitt lymphoma. Blood *116*, 2531-2542.

Rivenbark, A.G. (2017). An Overview of Cancer Genes. In The Molecular Basis of Human Cancer, W.B. Coleman, and G.J. Tsongalis, eds. (New York, NY: Springer New York), pp. 121-142.

Roberts, A.W., Davids, M.S., Pagel, J.M., Kahl, B.S., Puvvada, S.D., Gerecitano, J.F., Kipps, T.J., Anderson, M.A., Brown, J.R., Gressick, L.*, et al.* (2016). Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia. N Engl J Med *374*, 311-322.

Roberts, K.G., Morin, R.D., Zhang, J., Hirst, M., Zhao, Y., Su, X., Chen, S.C., Payne-Turner, D., Churchman, M.L., Harvey, R.C.*, et al.* (2012). Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. Cancer Cell *22*, 153-166.

Robles-Espinoza, C.D., and Adams, D.J. (2013). Cross-Species Analysis of Mouse and Human Cancer Genomes. Cold Spring Harbor protocols.

Roche-Lestienne, C., Soenen-Cornu, V., Grardel-Duflos, N., Lai, J.L., Philippe, N., Facon, T., Fenaux, P., and Preudhomme, C. (2002). Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. Blood *100*, 1014-1018.

Rodenhuis, S. (1992). ras and human tumors. Seminars in cancer biology *3*, 241-247.

Roschewski, M., Staudt, L.M., and Wilson, W.H. (2014). Diffuse large B-cell lymphoma-treatment approaches in the molecular era. Nature reviews Clinical oncology *11*, 12-23.

Ross, C.W., Ouillette, P.D., Saddler, C.M., Shedden, K.A., and Malek, S.N. (2007). Comprehensive analysis of copy number and allele status identifies multiple chromosome defects underlying follicular lymphoma pathogenesis. Clin Cancer Res *13*, 4777-4785.

Rossi, D., Cerri, M., Capello, D., Deambrogi, C., Rossi, F.M., Zucchetto, A., De Paoli, L., Cresta, S., Rasi, S., Spina, V.*, et al.* (2008). Biological and clinical risk factors of chronic lymphocytic leukaemia transformation to Richter syndrome. British journal of haematology *142*, 202-215.

Rowley, J.D. (2001). Chromosome translocations: dangerous liaisons revisited. Nature reviews Cancer *1*, 245-250.

Sabattini, E., Bacci, F., Sagramoso, C., and Pileri, S.A. (2010). WHO classification of tumours of haematopoietic and lymphoid tissues in 2008: an overview. Pathologica *102*, 83-87.

Saenz Robles, M.T., Symonds, H., Chen, J., and Van Dyke, T. (1994). Induction versus progression of brain tumor development: differential functions for the pRB- and p53-targeting domains of simian virus 40 T antigen. Mol Cell Biol *14*, 2686-2698.

Sánchez, C., Sánchez, I., Demmers, J.A.A., Rodriguez, P., Strouboulis, J., and Vidal, M. (2007). Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor. Molecular & cellular proteomics : MCP *6*, 820-834.

Sanger, F., Brownlee, G.G., and Barrell, B.G. (1965). A two-dimensional fractionation procedure for radioactive nucleotides. Journal of molecular biology *13*, 373-398.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A *74*, 5463-5467.

Schmitt, C.A., Fridman, J.S., Yang, M., Baranov, E., Hoffman, R.M., and Lowe, S.W. (2002). Dissecting p53 tumor suppressor functions in vivo. Cancer Cell *1*, 289-298.

Schmitt, C.A., McCurrach, M.E., de Stanchina, E., Wallace-Brodeur, R.R., and Lowe, S.W. (1999). INK4a/ARF mutations accelerate lymphomagenesis and promote chemoresistance by disabling p53. Genes & Development *13*, 2670-2677.

Schmitt, C.A., Rosenthal, C.T., and Lowe, S.W. (2000). Genetic analysis of chemoresistance in primary murine lymphomas. Nat Med *6*, 1029-1035.

Schmitz, R., Ceribelli, M., Pittaluga, S., Wright, G., and Staudt, L.M. (2014). Oncogenic mechanisms in Burkitt lymphoma. Cold Spring Harbor perspectives in medicine *4*.

Schmitz, R., Young, R.M., Ceribelli, M., Jhavar, S., Xiao, W., Zhang, M., Wright, G., Shaffer, A.L., Hodson, D.J., Buras, E.*, et al.* (2012). Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. Nature.

Schröder, J., Hsu, A., Boyle, S.E., Macintyre, G., Cmero, M., Tothill, R.W., Johnstone, R.W., Shackleton, M., and Papenfuss, A.T. (2014). Socrates: identification of genomic rearrangements in tumour genomes by re-aligning soft clipped reads. Bioinformatics (Oxford, England).

Sears, R., Nuckolls, F., Haura, E., Taya, Y., Tamai, K., and Nevins, J.R. (2000). Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability. Genes & development *14*, 2501-2514.

Sehn, L.H. (2016). Introduction to a review series: the paradox of indolent B-cell lymphoma. Blood *127*, 2045-2046.

Serrano, M., Hannon, G.J., and Beach, D. (1993). A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. Nature *366*, 704-707.

Shaffer Iii, A.L., Young, R.M., and Staudt, L.M. (2011). Pathogenesis of Human B Cell Lymphomas. Annual review of immunology.

Shah, N.P., Nicoll, J.M., Nagar, B., Gorre, M.E., Paquette, R.L., Kuriyan, J., and Sawyers, C.L. (2002). Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. Cancer Cell *2*, 117-125.

Sharma, S.V., and Settleman, J. (2007). Oncogene addiction: setting the stage for molecularly targeted cancer therapy. Genes & development *21*, 3214-3231.

Sherr, C.J., and Weber, J.D. (2000). The ARF/p53 pathway. Current opinion in genetics & development *10*, 94-99.

Shih, C., and Weinberg, R.A. (1982). Isolation of a transforming sequence from a human bladder carcinoma cell line. Cell *29*, 161-169.

Shortt, J., Martin, B.P., Newbold, A., Hannan, K.M., Devlin, J.R., Baker, A.J., Ralli, R., Cullinane, C., Schmitt, C.A., Reimann, M.*, et al.* (2013). Combined inhibition of PI3K-related DNA damage response kinases and mTORC1 induces apoptosis in MYC-driven B-cell lymphomas. Blood *121*, 2964-2974.

Sidman, C.L., Shaffer, D.J., Jacobsen, K., Vargas, S.R., and Osmond, D.G. (1993). Cell populations during tumorigenesis in Eu-myc transgenic mice. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK *7*, 887-895.

Sjoblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., Mandelker, D., Leary, R.J., Ptak, J., Silliman, N.*, et al.* (2006). The consensus coding sequences of human breast and colorectal cancers. Science *314*, 268-274.

Smith, A., Howell, D., Patmore, R., Jack, A., and Roman, E. (2011). Incidence of haematological malignancy by sub-type: a report from the Haematological Malignancy Research Network. British journal of cancer *105*, 1684-1692.

So, C.C., Yung, K.H., Chu, M.L., and Wan, T.S. (2013). Diagnostic challenges in a case of B cell lymphoma unclassifiable with features intermediate between diffuse large B-cell lymphoma and Burkitt lymphoma. International journal of hematology *98*, 478-482.

Sodir, N.M., Swigart, L.B., Karnezis, A.N., Hanahan, D., Evan, G.I., and Soucek, L. (2011). Endogenous Myc maintains the tumor microenvironment. Genes & development *25*, 907-916.

Soucek, L., Whitfield, J., Martins, C.P., Finch, A.J., Murphy, D.J., Sodir, N.M., Karnezis, A.N., Swigart, L.B., Nasi, S., and Evan, G.I. (2008). Modelling Myc inhibition as a cancer therapy. Nature *455*, 679-683.

Soucek, L., Whitfield, J.R., Sodir, N.M., Masso-Valles, D., Serrano, E., Karnezis, A.N., Swigart, L.B., and Evan, G.I. (2013). Inhibition of Myc family proteins eradicates KRas-driven lung cancer in mice. Genes & development *27*, 504-513.

Spencer, C.A., and Groudine, M. (1991). Control of c-myc regulation in normal and neoplastic cells. Advances in cancer research *56*, 1-48.

Stewart, T.A., Pattengale, P.K., and Leder, P. (1984). Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. Cell *38*, 627-637.

Strasser, A., Harris, A.W., Bath, M.L., and Cory, S. (1990). Novel primitive lymphoid tumours induced in transgenic mice by cooperation between myc and bcl-2. Nature *348*, 331-333.

Stratton, M.R. (2011). Exploring the genomes of cancer cells: progress and promise. Science *331*, 1553-1558.

Stratton, M.R., Campbell, P.J., and Futreal, P.A. (2009). The cancer genome. Nature *458*, 719-724.

Swerdlow, S.H. (2014). Diagnosis of 'double hit' diffuse large B-cell lymphoma and B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and Burkitt lymphoma: when and how, FISH versus IHC. Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program *2014*, 90-99.

Takada, K., Zhu, D., Bird, G.H., Sukhdeo, K., Zhao, J.J., Mani, M., Lemieux, M., Carrasco, D.E., Ryan, J., Horst, D.*, et al.* (2012). Targeted disruption of the BCL9/beta-catenin complex inhibits oncogenic Wnt signaling. Science translational medicine *4*, 148ra117.

Talpaz, M., Shah, N.P., Kantarjian, H., Donato, N., Nicoll, J., Paquette, R., Cortes, J., O'Brien, S., Nicaise, C., Bleickardt, E.*, et al.* (2006). Dasatinib in imatinib-resistant Philadelphia chromosome-positive leukemias. N Engl J Med *354*, 2531-2541.

Teer, J.K., and Mullikin, J.C. (2010). Exome sequencing: the sweet spot before whole genomes. Hum Mol Genet *19*, R145-151.

Tiacci, E., Grossmann, V., Martelli, M.P., Kohlmann, A., Haferlach, T., and Falini, B. (2012). The corepressors BCOR and BCORL1: two novel players in acute myeloid leukemia. Haematologica *97*, 3-5.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols *7*, 562-578.

Tsai, A.G., Lu, H., Raghavan, S.C., Muschen, M., Hsieh, C.L., and Lieber, M.R. (2008). Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity. Cell *135*, 1130-1142.

Tse, C., Shoemaker, A.R., Adickes, J., Anderson, M.G., Chen, J., Jin, S., Johnson, E.F., Marsh, K.C., Mitten, M.J., Nimmer, P.*, et al.* (2008). ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. Cancer Res *68*, 3421-3428.

Tsujimoto, Y., Bashir, M.M., Givol, I., Cossman, J., Jaffe, E., and Croce, C.M. (1987). DNA rearrangements in human follicular lymphoma can involve the 5' or the 3' region of the bcl-2 gene. Proc Natl Acad Sci U S A *84*, 1329-1331.

Tsujimoto, Y., Cossman, J., Jaffe, E., and Croce, C.M. (1985a). Involvement of the bcl-2 gene in human follicular lymphoma. Science *228*, 1440-1443.

Tsujimoto, Y., Gorham, J., Cossman, J., Jaffe, E., and Croce, C.M. (1985b). The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. Science *229*, 1390-1393.

Turke, A.B., Zejnullahu, K., Wu, Y.L., Song, Y., Dias-Santagata, D., Lifshits, E., Toschi, L., Rogers, A., Mok, T., Sequist, L.*, et al.* (2010). Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. Cancer Cell *17*, 77-88.

Twa, D.D., Chan, F.C., Ben-Neriah, S., Woolcock, B.W., Mottok, A., Tan, K.L., Slack, G.W., Gunawardana, J., Lim, R.S., McPherson, A.W*., et al.* (2014). Genomic rearrangements involving programmed death ligands are recurrent in primary mediastinal large B-cell lymphoma. Blood *123*, 2062-2065.

Uozumi, K., Otsuka, M., Ohno, N., Moriyama, T., Suzuki, S., Shimotakahara, S., Matsumura, I., Hanada, S., and Arima, T. (2000). Establishment and characterization of a new human megakaryoblastic cell line (SET-2) that spontaneously matures to megakaryocytes and produces platelet-like particles. Leukemia *14*, 142-152.

Uren, A.G., Mikkers, H., Kool, J., van der Weyden, L., Lund, A.H., Wilson, C.H., Rance, R., Jonkers, J., van Lohuizen, M., Berns, A*., et al.* (2009). A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. Nature protocols *4*, 789-798.

Vafa, O., Wade, M., Kern, S., Beeche, M., Pandita, T.K., Hampton, G.M., and Wahl, G.M. (2002). c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability. Molecular cell *9*, 1031-1044.

Valera, A., Lopez-Guillermo, A., Cardesa-Salzmann, T., Climent, F., Gonzalez-Barca, E., Mercadal, S., Espinosa, I., Novelli, S., Briones, J., Mate, J.L*., et al.* (2013). MYC protein expression and genetic alterations have prognostic impact in patients with

diffuse large B-cell lymphoma treated with immunochemotherapy. Haematologica *98*, 1554-1562.

van den Boom, V., Rozenveld-Geugien, M., Bonardi, F., Malanga, D., van Gosliga, D., Heijink, A.M., Viglietto, G., Morrone, G., Fusetti, F., Vellenga, E.*, et al.* (2013). Nonredundant and locus-specific gene repression functions of PRC1 paralog family members in human hematopoietic stem/progenitor cells. Blood *121*, 2452-2461.

Van Dyke, T., and Jacks, T. (2002). Cancer modeling in the modern era: progress and challenges. Cell *108*, 135-144.

van Lohuizen, M., Verbeek, S., Scheijen, B., Wientjens, E., van der Gulden, H., and Berns, A. (1991). Identification of cooperating oncogenes in E mu-myc transgenic mice by provirus tagging. CELL *65*, 737-752.

van Miltenburg, M.H., and Jonkers, J. (2012). Using genetically engineered mouse models to validate candidate cancer genes and test new therapeutic approaches. Current opinion in genetics & development *22*, 21-27.

Van Roosbroeck, K., Cox, L., Tousseyn, T., Lahortiga, I., Gielen, O., Cauwelier, B., De Paepe, P., Verhoef, G., Marynen, P., Vandenberghe, P.*, et al.* (2011). JAK2 rearrangements, including the novel SEC31A-JAK2 fusion, are recurrent in classical Hodgkin lymphoma. Blood *117*, 4056-4064.

Vandenberg, C.J., and Cory, S. (2013). ABT-199, a new Bcl-2-specific BH3 mimetic, has in vivo efficacy against aggressive Myc-driven mouse lymphomas without provoking thrombocytopenia. Blood *121*, 2285-2288.

Vaque, J.P., Martinez, N., Batlle-Lopez, A., Perez, C., Montes-Moreno, S., Sanchez-Beato, M., and Piris, M.A. (2014). B-cell lymphoma mutations: improving diagnostics and enabling targeted therapies. Haematologica *99*, 222-231.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A.*, et al.* (2001). The sequence of the human genome. Science *291*, 1304-1351.

Verbeek, S., van Lohuizen, M., van der Valk, M., Domen, J., Kraal, G., and Berns, A. (1991). Mice bearing the E mu-myc and E mu-pim-1 transgenes develop pre-B-cell leukemia prenatally. Molecular and Cellular Biology *11*, 1176-1179.

Vidal, M., and Starowicz, K. (2017). Polycomb complexes PRC1 and their function in hematopoiesis. Experimental hematology *48*, 12-31.

Villamor, N., Conde, L., Martinez-Trillos, A., Cazorla, M., Navarro, A., Bea, S., Lopez, C., Colomer, D., Pinyol, M., Aymerich, M.*, et al.* (2013). NOTCH1 mutations identify a genetic subgroup of chronic lymphocytic leukemia patients with high risk of transformation and poor outcome. Leukemia *27*, 1100-1106.

Villunger, A., Michalak, E.M., Coultas, L., Mullauer, F., Bock, G., Ausserlechner, M.J., Adams, J.M., and Strasser, A. (2003). p53- and drug-induced apoptotic responses mediated by BH3-only proteins puma and noxa. Science *302*, 1036-1038.

Visser, H.P., Gunster, M.J., Kluin-Nelemans, H.C., Manders, E.M., Raaphorst, F.M., Meijer, C.J., Willemze, R., and Otte, A.P. (2001). The Polycomb group protein EZH2 is upregulated in proliferating, cultured human mantle cell lymphoma. British journal of haematology *112*, 950-958.

Vita, M., and Henriksson, M. (2006). The Myc oncoprotein as a therapeutic target for human cancer. Seminars in cancer biology *16*, 318-330.

Vitolo, U., Gaidano, G., Botto, B., Volpe, G., Audisio, E., Bertini, M., Calvi, R., Freilone, R., Novero, D., Orsucci, L.*, et al.* (1998). Rearrangements of bcl-6, bcl-2, c-myc and 6q deletion in B-diffuse large-cell lymphoma: clinical relevance in 71 patients. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO *9*, 55-61.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr., and Kinzler, K.W. (2013). Cancer genome landscapes. Science *339*, 1546-1558.

Vogler, L.B., Crist, W.M., Bockman, D.E., Pearl, E.R., Lawton, A.R., and Cooper, M.D. (1978). Pre-B-cell leukemia. A new phenotype of childhood lymphoblastic leukemia. N Engl J Med *298*, 872-878.

Waibel, M., Solomon, V.S., Knight, D.A., Ralli, R.A., Kim, S.K., Banks, K.M., Vidacs, E., Virely, C., Sia, K.C., Bracken, L.S.*, et al.* (2013). Combined targeting of JAK2 and Bcl-2/Bcl-xL to cure mutant JAK2-driven malignancies and overcome acquired resistance to JAK2 inhibitors. Cell Rep *5*, 1047-1059.

Wall, M., Poortinga, G., Stanley, K.L., Lindemann, R.K., Bots, M., Chan, C.J., Bywater, M.J., Kinross, K.M., Astle, M.V., Waldeck, K.*, et al.* (2013). The mTORC1 inhibitor everolimus prevents and treats Emu-Myc lymphoma by restoring oncogene-induced senescence. Cancer discovery *3*, 82-95.

Wamstad, J.A., and Bardwell, V.J. (2007). Characterization of Bcor expression in mouse development. Gene Expression Patterns *7*, 550-557.

Wang, H., Liang, L., Fang, J.Y., and Xu, J. (2016). Somatic gene copy number alterations in colorectal cancer: new quest for cancer drivers and biomarkers. Oncogene *35*, 2011-2019.

Wang, J., Mullighan, C.G., Easton, J., Roberts, S., Heatley, S.L., Ma, J., Rusch, M.C., Chen, K., Harris, C.C., Ding, L.*, et al.* (2011). CREST maps somatic structural variation in cancer genomes with base-pair resolution. Nature methods *8*, 652-654.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nature reviews Genetics *10*, 57-63.

Waters, C.M., Littlewood, T.D., Hancock, D.C., Moore, J.P., and Evan, G.I. (1991). c-myc protein expression in untransformed fibroblasts. Oncogene *6*, 797-805.

Weinstein, I.B. (2000). Disorders in cell circuitry during multistage carcinogenesis: the role of homeostasis. Carcinogenesis *21*, 857-864.

Weinstein, I.B. (2002). Cancer. Addiction to oncogenes--the Achilles heal of cancer. Science *297*, 63-64.

Wendel, H.-G., De Stanchina, E., Fridman, J.S., Malina, A., Ray, S., Kogan, S., Cordon-Cardo, C., Pelletier, J., and Lowe, S.W. (2004). Survival signalling by Akt and eIF4E in oncogenesis and cancer therapy. Nature *428*, 332-337.

Westlake, S. (2009). Cancer incidence and mortality in the United Kingdom and constituent countries, 2004-06. Health statistics quarterly, 56-62.

Whitecross, K.F., Alsop, A.E., Cluse, L.A., Wiegmans, A., Banks, K.M., Coomans, C., Peart, M.J., Newbold, A., Lindemann, R.K., and Johnstone, R.W. (2009). Defining the target specificity of ABT-737 and synergistic antitumor activities in combination with histone deacetylase inhibitors. Blood *113*, 1982-1991.

Williams, B.O., Remington, L., Albert, D.M., Mukai, S., Bronson, R.T., and Jacks, T. (1994). Cooperative tumorigenic effects of germline mutations in Rb and p53. Nat Genet *7*, 480-484.

Willis, S.N., Fletcher, J.I., Kaufmann, T., van Delft, M.F., Chen, L., Czabotar, P.E., Ierino, H., Lee, E.F., Fairlie, W.D., Bouillet, P.*, et al.* (2007). Apoptosis initiated when BH3 ligands engage multiple Bcl-2 homologs, not Bax or Bak. Science *315*, 856-859.

Witzig, T.E., and Gupta, M. (2010). Signal transduction inhibitor therapy for lymphoma. Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program *2010*, 265-270.

Yahata, T., de Caestecker, M.P., Lechleider, R.J., Andriole, S., Roberts, A.B., Isselbacher, K.J., and Shioda, T. (2000). The MSG1 non-DNA-binding transactivator binds to the p300/CBP coactivators, enhancing their functional link to the Smad transcription factors. J Biol Chem *275*, 8825-8834.

Ye, B.H., Chaganti, S., Chang, C.C., Niu, H., Corradini, P., Chaganti, R.S., and Dalla-Favera, R. (1995). Chromosomal translocations cause deregulated BCL6 expression by promoter substitution in B cell lymphoma. EMBO J *14*, 6209-6217.

Yoon, S.O., Jeon, Y.K., Paik, J.H., Kim, W.Y., Kim, Y.A., Kim, J.E., and Kim, C.W. (2008). MYC translocation and an increased copy number predict poor prognosis in adult diffuse large B-cell lymphoma (DLBCL), especially in germinal centre-like B cell (GCB) type. Histopathology *53*, 205-217.

Youle, R.J., and Strasser, A. (2008). The BCL-2 protein family: opposing activities that mediate cell death. Nature reviews Molecular cell biology *9*, 47-59.

Yu, X.X., Hu, Z., Shen, X., Dong, L.Y., Zhou, W.Z., and Hu, W.H. (2015). IL-33 Promotes Gastric Cancer Cell Invasion and Migration Via ST2-ERK1/2 Pathway. Digestive diseases and sciences *60*, 1265-1272.

Yu, Y.P., Song, C., Tseng, G., Ren, B.G., LaFramboise, W., Michalopoulos, G., Nelson, J., and Luo, J.H. (2012). Genome abnormalities precede prostate cancer and predict clinical relapse. The American journal of pathology *180*, 2240-2248.

Zender, L., Xue, W., Zuber, J., Semighini, C.P., Krasnitz, A., Ma, B., Zender, P., Kubicka, S., Luk, J.M., Schirmacher, P.*, et al.* (2008). An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer. Cell *135*, 852-864.

Zenz, T., Mertens, D., Kuppers, R., Dohner, H., and Stilgenbauer, S. (2010). From pathogenesis to treatment of chronic lymphocytic leukaemia. Nature reviews Cancer *10*, 37-50.

Zhang, J., Grubor, V., Love, C.L., Banerjee, A., Richards, K.L., Mieczkowski, P.A., Dunphy, C., Choi, W., Au, W.Y., Srivastava, G.*, et al.* (2013). Genetic heterogeneity of diffuse large B-cell lymphoma. Proceedings of the National Academy of Sciences of the United States of America.

Zhang, Y., Feng, X.H., and Derynck, R. (1998). Smad3 and Smad4 cooperate with c-Jun/c-Fos to mediate TGF-beta-induced transcription. Nature *394*, 909-913.

Zimmermann, J., Caravatti, G., Mett, H., Meyer, T., Muller, M., Lydon, N.B., and Fabbro, D. (1996). Phenylamino-pyrimidine (PAP) derivatives: a new class of potent and selective inhibitors of protein kinase C (PKC). Archiv der Pharmazie *329*, 371-376.

Zindy, F., Eischen, C.M., Randle, D.H., Kamijo, T., Cleveland, J.L., Sherr, C.J., and Roussel, M.F. (1998). Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization. Genes & development *12*, 2424-2433.