Elucidating oncogenic mechanisms in human B cell malignancies

Rebecca Caeser

Darwin College

This dissertation is submitted for the degree of

Doctor of Philosophy

Submission date

November 2018

Summary: Elucidating oncogenic mechanisms in human B cell malignancies

Rebecca Caeser

This study consists of two pieces of work investigating haematological malignancies; Acute Lymphoblastic Leukaemia (ALL) and Diffuse Large B Cell Lymphoma (DLBCL). Firstly, Pre-B ALL represents the most common paediatric malignancy and despite increasingly improved outcomes for patients, ~ 20% of all patients diagnosed with ALL relapse. Activating mutations in the RAS pathway are common (~50%) and result in hyperactivation of the MAPK pathway. I identified Erk negative feedback control via DUSP6 to be crucial for NRAS^{G12D}-mediated pre-B cell transformation and investigated its potential as a therapeutic target. I showed that a small molecule inhibitor of DUSP6 (BCI) selectively induced cell death in patient-derived pre-B ALL cells; with a higher sensitivity observed in relapse pre-B ALL. I also discovered that a high level of Erk activity is required for proliferation of normal pre-B cells, but dispensable in leukemic pre-B ALL cells. In addition, I found that human B cell malignancies can be grouped into three categories that fundamentally differ in their ability to control Erk signalling strength.

Secondly, DLBCL is the most common haematological malignancy and although potentially curable with chemotherapy, 40% of patients still succumb from their disease. Recent exome sequencing studies have identified hundreds of genetic alterations but, for most, their contribution to disease, or their importance as therapeutic targets, remains uncertain. I optimised a novel approach to screen the functional importance of these mutations. This was achieved by reconstituting non-malignant, primary, human germinal centre B cells (GC B cells) with combinations of wildtype and mutant genes to recapitulate the genetic events of DLBCL. When injected into immunodeficient mice, these oncogene-transduced GC B cells gave rise to tumours that closely resemble human DLBCL, reinforcing the biological relevance of this system. To screen potential tumour suppressor mutations in this system in a high throughput fashion, I developed a lymphoma-focused CRISPR library of 692 genes recurrently altered in B cell lymphomas. These experiments identified GNA13 as an unexpectedly potent tumour suppressor in human GC B cells and provided new understanding to its mechanism of action.

These findings provide novel understanding of the complexity of oncogenic mechanisms in human B cell malignancies.

Declaration of Originality

- This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

- It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

- It does not exceed the prescribed word limit for the relevant Degree Committee.

Table of Contents

Acknowledgements

Firstly, I would like to express my thanks and gratitude to my supervisor, Daniel Hodson for giving me the opportunity to join his laboratory and despite this, keeps a sense of humour throughout. His guidance, expertise and patience during my PhD were invaluable. I am hugely appreciative of his support and advice over the past couple of years, which has helped to shape me as a scientist.

Thank you also to Markus Müschen, under whom I began this thesis. His determination, enthusiasm and rigour have been a great source of inspiration for me. I am also hugely appreciative to Bertie Göttgens, in his role as my secondary supervisor. His advice and support during tough times was invaluable.

The members of the Hodson lab have contributed immensely to this thesis. The group has been a source of advice, support and friendship over the past two years. A big thanks goes to Miriam for all her encouragement, support and help but more importantly for her delicious Tiramisu. I would also like to thank Jane for generously taking over some of my lab responsibilities so I could finally complete this thesis. Thanks also to Zelvera for her help and to Jade and Eirini for their support.

Thank you to Nima at UCSF for his support and willingness to answer all of my never-ending questions at the start of my PhD. I would also like to thank Priyanka and Maurizio for creating such a friendly work environment when we first moved to Cambridge. Professionally and personally I owe Priyanka a lot, she has helped me to overcome several challenges to complete this PhD. I would like to thank all members of the Müschen lab for their encouragement and stimulating discussions.

Special thanks goes to everyone on Level 5 for creating such a supportive and friendly work environment. In particular, thanks are due to Mairi for always being entertaining, as well as helping to make some beautiful schematics.

Finally, thank you to my parents for their loving support and to Nick for all of the pasta.

Abbreviations

2-ME	2-Mercaptoethanol
4OHT	4-Hydroxytamoxifen
ABC DLBCL	Activated B-Cell like Diffuse Large B Cell Lymphoma
AIC	Autoimmunity Checkpoints
AKT (PKB)	AMP Activated Protein Kinase
ALL	Acute Lymphoblastic Leukaemia
ALT	Alternative Lengthening of Telomeres
AML	Acute Myeloid Leukaemia
ARE	Adenylate-uridylate-rich elements
AUBPs	ARE-specific binding proteins
BCI	(E/Z)-BCI hydrochloride
BCR	B Cell Receptor
BFP	Blue Fluorescent Protein
Blast	Blasticidin
CAS9	CRISPR associated protein 9
CD40Lg	CD40 ligand
CLL	Chronic Lymphoid Leukaemia
CML	Chronic Myeloid Leukaemia
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CSR	Class Switch Recombination
CTL	Cytotoxic T Lymphocyte
DLBCL	Diffuse Large B Cell Lymphoma
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
Dox	Doxycycline
DSB	Double Strand Break
DUSP6	Dual Specificity Phosphatase 6
EGFR	Epidermal growth factor receptor
ERK	Extracellular-signal-Regulated Kinase
EV	Empty Vector
FACS	Fluorescence-Activated Cell Sorting
FBS	Fetal Bovine Serum
FDC	Follicular Dendritic Cells

FGF	Fibroblast Growth Factor
FGFR	Fibroblast growth factor receptor
FITC	Fluorescein isothiocyanate
FL	Follicular Lymphoma
GaLV	Gibbon Ape Leukaemia Virus
GAP	GTPase-activating proteins
GC B Cells	Germinal Centre B cells
GCB DLBCL	Germinal Centre B-Cell like Diffuse Large B Cell Lymphoma
GDP	Guanosine Diphosphate
GEF	Guanine nucleotide exchange factor
GFP	Green Fluorescent Protein
GPCR	G-protein-coupled receptors
gRNA	guide RNA
GTP	Guanosine Triphosphate
HDR	Homology Directed Repair
Hsp90	Heat Shock Protein 90
IgH	Immunoglobulin heavy chain
IL-X	Interleukin-X
ITAM	Immunoreceptor tyrosine-based activation motif
MACS	Magnetic-activated cell sorting
MAPK	Mitogen Activated Protein Kinase
MEK	Mitogen-activated protein kinase kinase
MEM	Modified Eagle's Medium
MKP	MAPK phosphatases
MoMLV	Moloney murine leukemia virus
MuLV	Murine leukemia virus
NGS	Next generation sequencing
NHEJ	Non-Homologous End Joining
NK	Natural Killer cell
NSG	NOD scid gamma
NSLC	Non-small-cell lung carcinoma
NTC	Non-targeting control
OIS	Oncogene-induced senescence
OS	Overall Survival

PAM	Protospacer adjacent motif
PBMC	Peripheral blood mononuclear cells
PFS	Progression Free Survival
Ph	Philadelphia chromosome
PIP2	Phosphatidylinositol 4,5-bisphosphate
PIP3	Phosphatidylinositol 3,4,5-trisphosphate
PTEN	Phosphatase and Tensin homologue
Puro	Puromycin
P-X	Phospho-X
RAF	Rapidly Accelerated Fibrosarcoma
RAS	Rat Sarcoma
R-CHOP	rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone
RNAi	RNA interference
RPMI	Roswell Park Memorial Institute medium
RTK	Receptor Tyrosine Kinase
SD	Standard Deviation
SHM	Somatic Hypermutation
shRNA	short hairpin RNA
SOS	Son of Sevenless
TALEN	Transcription activator-like effector nucleases
Tfh	Follicular B helper T cells
TKI	Tyrosine Kinase Inhibitor
TNF	Tumour-necrsos factor
UCSF	University of California, San Francisco
ZFN	Zinc Finger Nuclease

Chapter 1 ERK feedback regulation in B cell malignancies

1 Abstract

Pre-B acute lymphoblastic leukaemia (ALL) is the most common paediatric malignancy. Despite increasingly improved outcomes for patients, ~ 20% of all patients diagnosed with ALL relapse. Activating mutations in the RAS pathway are common (~50%) at presentation and lead to hyperactivation of the MAPK pathway. I identified ERK negative feedback control via the Dual Specificity Phosphatase 6 (DUSP6) to be essential for NRAS^{G12D}-mediated transformation of mouse pre-B cells and investigated its potential as a therapeutic target. I showed that a small molecule inhibitor of DUSP6 (BCI) selectively induced cell death in patient-derived pre-B ALL cells; with a higher sensitivity observed in relapsed pre-B ALL. I also discovered that a high level of ERK activity is required for proliferation of normal pre-B cells, but dispensable in leukaemic pre-B ALL cells. In addition, I found that human B cell malignancies can be grouped into three categories that fundamentally differ in their ability to control ERK signalling strength. Albeit counterintuitive, these findings highlight the negative feedback regulator DUSP6 as a potential therapeutic target.

2 Introduction

2.1. Pre-B acute lymphoblastic leukaemia

Pre-B acute lymphoblastic leukaemia (Pre-B ALL) accounts for nearly a quarter of all childhood cancers, representing the most common paediatric malignancy[13]. It is caused by multiple factors including genetic susceptibility, exogenous/endogenous exposures, genetic alterations and chance[14]. Despite the high percentage of children with ALL, survival rates have greatly improved for patients, from 10% to more than 85%, over the last four decades[13]. Yet ALL remains one of the leading causes of person-years of life lost in the US (362,000 years in 2010; National Center of Health Statistics). 5,430 patients were diagnosed in the US in 2008 (*Leukemia and Lymphoma Society*) and 654 patients in the UK in 2011[15]. The 5 year disease-free survival rate is ~85% for children and ~55% for adults[13].

The survival rate in paediatric ALL is now approaching 90% with new advances in combination therapy tailored to patient pharmacodynamics and pharmacogenomics, personalized chemotherapy and improvements in supportive clinical care[14]. The cure rate is likely to increase further due to the discovery of additional molecular lesions that will lead to the development of novel targeted drugs and treatment strategies[16]. Nevertheless, 20% of all patients diagnosed with ALL will relapse and the outlook for those children is poor; only around one-third of children survive long-term after disease recurrence[17]. As such, innovative approaches to improve survival while reducing adverse effects are needed.

2.2. The molecular and genetic landscape of ALL

Acute lymphoblastic leukaemia is a neoplasm of immature lymphoid progenitors. Approximately 80% are B cell in origin[5]. It is caused by transforming genetic events in early progenitors followed by the acquisition of further genetic alterations that often play an important role in cell survival, differentiation and proliferation[18]. Normal B cells derive from blood stem cells in the bone marrow[19]. These cells go through an orderly development, involving a strictly regulated maturation and selection process by means of the expression of cell surface markers, transcription factors and by immunoglobulin gene arrangements[19]. These cells eventually become memory B cells or plasma cells[19,20]. During their orderly development, B cells undergo selection

at several checkpoints based on the expression and antigen-specificity of the B cell receptor (BCR)[9,21,22]. This ensures that autoreactive B cells are eliminated from the B cell repertoire. The majority of B-ALLs in children exhibit a developmental arrest at the pre-pro-B and pre-B stages[18] (Figure 1).
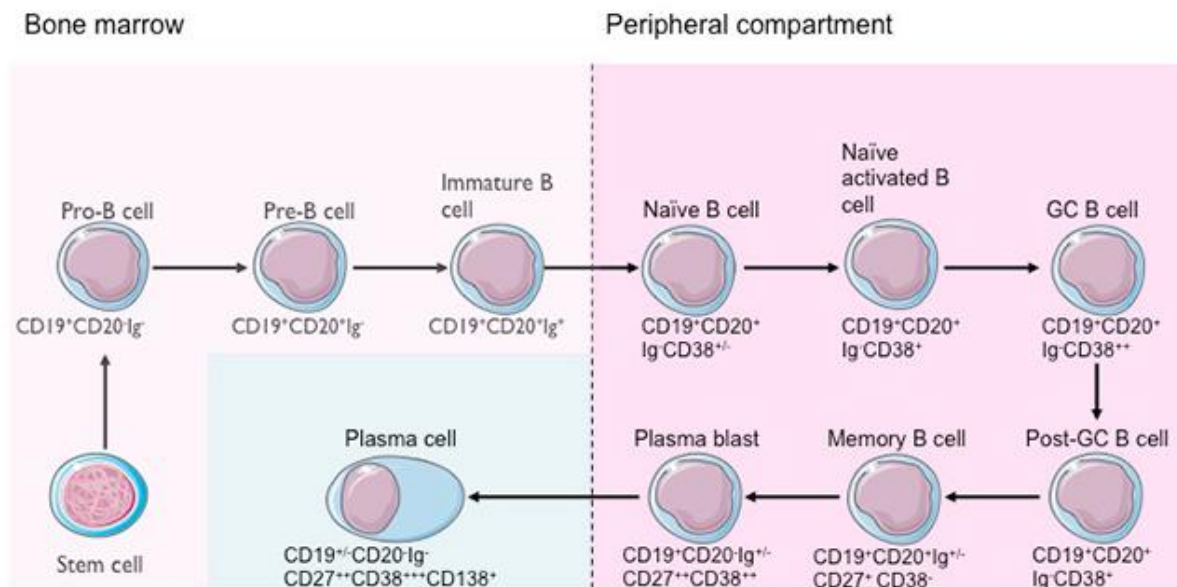
Figure 1 B cell ontogeny *B cell differentiation starts in the bone marrow and progresses from hematopoietic stem cells through to pro-B, pre-B cells and immature pre-B cell stages. During this development, VDJ rearrangements at the immunoglobulin locus leads to surface expression of the pre-B cell receptor (pre-BCR) which comprises of an IgM heavy chain and surrogate light chains (VpreB and λ-5). Signalling from the pre-BCR results in progression of B cell development and eventually the expression of a mature BCR composed of rearranged heavy and light chains. Cells that have successfully passed this checkpoint leave the bone marrow and further differentiate into marginal-zone B cells and mature follicular B cells in the spleen.[10]*

Next-generation sequencing has given us a better understanding of the genetic lesions that characterize ALL. Mullighan argued that many ALL types can be traced back to chromosomal rearrangements, sequence mutations and changes in submicroscopic DNA copy number[5]. In B-ALL, almost 80% exhibit aneuploidy,

hyperdiploidy, hypodiploidy and chromosomal rearrangements such as t(12;21) ETV6-RUNX1 (TEL-AML1), t(1;19) TCF3-PBX1 (E2A-PBX1), t(9;22) BCR-ABL1 and rearrangement of MLL at 11q23 to a number of other fusion partners[5] (Figure 2A).

Such chromosomal rearrangements are most probably acquired early in leukaemogenesis leading to transcriptional and epigenetic dysregulation and aberrant self-renewal[5]. The result of this is an accumulation of immature B cells. However, rearrangements alone are not sufficient for leukaemogenesis and it is now known that leukaemogenesis requires further structural and genetic alterations in genes involved in tumour suppression, cell cycle regulation and chromatin modification[5]. Furthermore, heterogeneity of the diagnosed ALL samples and genetic alterations in minor clones may result in resistance to therapy and disease relapse[5].
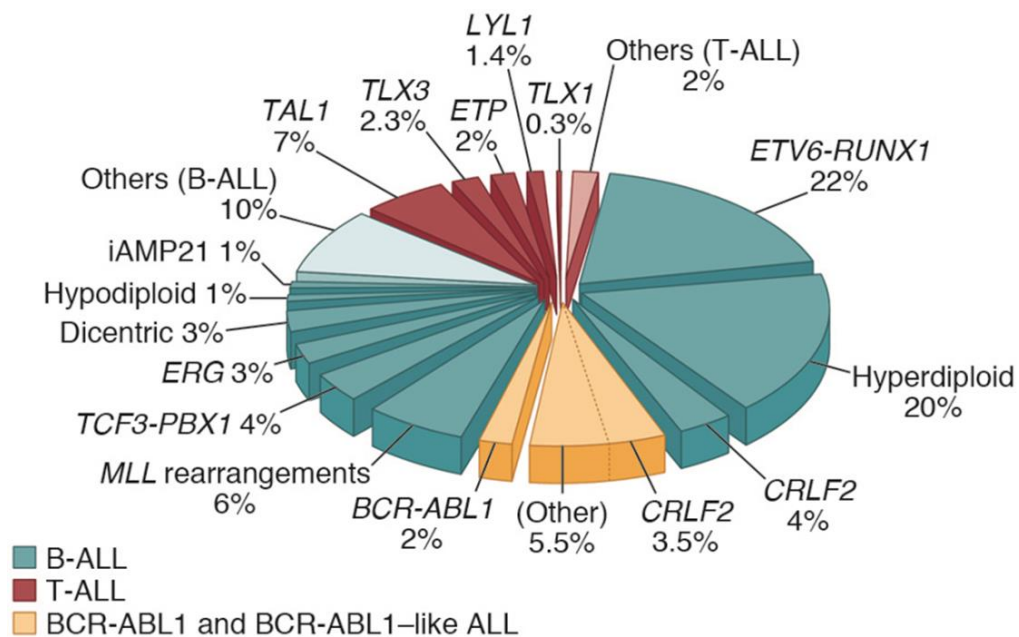
Figure 2 Frequency of cytogenetic subtypes of paediatric ALL Blue and red represent the relative frequency of B-ALL and T-ALL genetic subtypes, respectively. Yellow represents the BCR-ABL1–like subtype and BCR-ABL1–positive ALL.[5]

2.3 RAS/RAF/MAPK pathway signalling

The mitogen-activated protein kinase (MAPK) cascade plays an important role in a number of diverse cellular functions like survival, differentiation, angiogenesis, migration and cell proliferation[23]. The pathway initiates at the cell surface by various receptors such as receptor tyrosine kinases (RTK), GPCRs, T cell receptors or B cell receptors[23]. Ligand binding leads to recruitment of RAS.

RAS belongs to a family of monomeric membrane-associated GTPases, consisting of HRAS, KRAS and NRAS[24]. These proteins cycle between an active GTP bound state and an inactive GDP bound state through the catalyzation of guanine nucleotide exchange factors (GEFs), such as son of sevenless (SOS)[24]. Interaction with upstream activators and downstream effectors occurs at the inner surface of the plasma membrane. RAS is negatively regulated by GTPase activating proteins (GAPs) that stimulate GTP hydrolysis and leave RAS in the inactive GDP bound state[23]. When in its GTP bound form, RAS triggers the activation of RAF (MAPKKK) which subsequently phosphorylates MEK (MAPKK) on serine residues, which further phosphorylates ERK (MAPK) on tyrosine and threonine residues[25] (Figure 3).

The MAP kinases ERK1 and ERK2, also known as MAPK3 and MAPK1 are 44- and 42-kDa serine/threonine kinases that regulate growth factor-responsive targets in the cytosol and also in the nucleus where it regulates gene expression[26]. ERK1 and ERK2 are activated via threonine and tyrosine phosphorylation at Thr202/Tyr204 and Thr185/Tyr187, respectively[26]. It is known that strong ERK activation needs phosphorylation at both sites, with tyrosine being phosphorylated before threonine[26]. ERK1 and ERK2 share many similarities; they are 84% identical at the amino acid level and are ubiquitously expressed[27]. They also share the same activators and substrates with no known agonist able to activate only one of the two kinases[28]. However, $Erk1^{-/-}$ mice are viable and fertile[29] whereas $Erk2^{-/-}$ present with early embryonic death due to placental defects[30].

In resting cells, ERK is mainly found in the cytoplasm. Upon activation, ERK translocates to the nucleus where it phosphorylates a number of transcription factors such as FOS, JUN and MYC, as well as inducing the expression of its own negative feedback regulators at three levels[31]. First, via SPRY2, a sprouty family signalling inhibitor that negatively regulates the activation of RAS. Second, via DUSP6, a dual

specificity phosphatase that dephosphorylates ERK. And third, via ERK itself by cooperating with transcription factors like ETV1, ETV4 and ETV5 that drive transcription of DUSP6 and SPRY2[4].
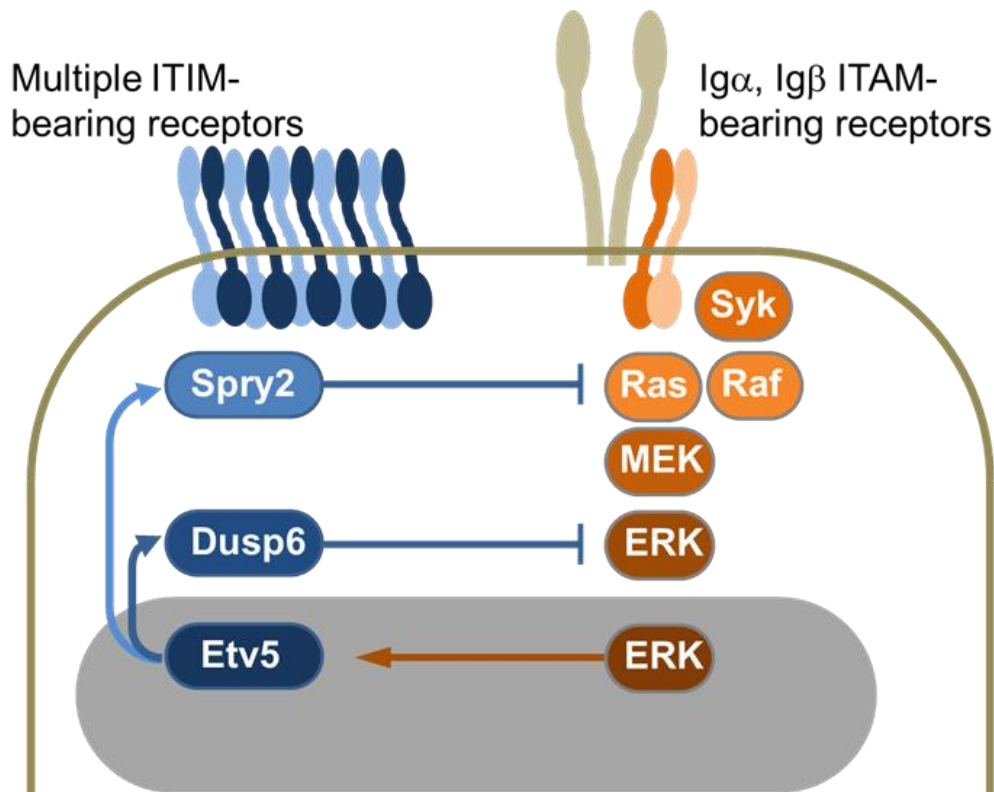
Figure 3 Schematic of the MAPK pathway illustrating negative feedback signalling *In human tumours, the RAS-RAF-MEK-ERK pathway is activated by a number of mechanisms for example binding of a ligand to or mutational activation of Igα, Igβ ITAM-bearing receptors and mutations in RAS or RAF. Activated ERK translocates to the nucleus where it drives transcriptional activation of its own negative feedback regulators (Etv5, Dusp6, Spry2) as well as other targets involved in negative in cell cycle progression and proliferation (modified from [4]).*

2.4 RAS pathway mutations in human cancer

Activated *RAS* genes were first discovered in cancer in 1982[32] and attempts to target RAS pharmaceutically have been the focus of many research groups ever since – with disappointing results to date. It is not surprising that the RAS pathway plays an important role in cancer therapy seeing as it is a key regulator of normal cell growth and malignant transformation. Therefore the rationale behind targeting this pathway is to inhibit the survival, growth, and spread of tumours[33]. *RAS* genes are the most common targets for somatic gain-of-function mutations; occurring in around 30% of human cancers[1]. *KRAS* mutations are most frequent in pancreatic, colorectal, endometrial, biliary tract, lung and cervical cancers; *NRAS* and *HRAS* mutations are prevalent in melanoma and bladder cancer, respectively and *KRAS* and *NRAS* mutations predominate in myeloid malignancies[1] (Figure 4).

Cancer type	HRAS	KRAS	NRAS	BRAF
Biliary tract	0%	33%	1%	14%
Bladder	11%	4%	3%	0%
Breast	0%	4%	0%	2%
Cervix	9%	9%	1%	0%
Colon	0%	32%	3%	14%
Endometrial	1%	15%	0%	1%
Kidney	0%	1%	0%	0%
Liver	0%	8%	10%	3%
Lung	1%	19%	1%	2%
Melanoma	6%	2%	18%	43%
Myeloid leukaemia	0%	5%	14%	1%
Ovarian	0%	17%	4%	15%
Pancreas	0%	60%	2%	3%
Thyroid	5%	4%	7%	27%

The mutation data was obtained from the Sanger Institute Catalogue of Somatic Mutations in Cancer web site[148].

Figure 4 HRAS, KRAS, NRAS and BRAF mutations in human cancer[1].

A high incidence of *RAS* mutations in human cancer has made *RAS* a common target in cancer therapy. Cancer therapeutic agents include Farnesyltransferase inhibitors, antisense oligonucleotides against RAS and RAF, and kinase inhibitors, targeting pathways upstream of RAS and also RAS effector pathways[1]. However, it is still difficult to design drugs with a sufficient therapeutic window that eliminate tumour cells but not normal, healthy cells[1].

2.5 RAS pathway mutations in ALL

Deregulation of the MAPK pathway is not only found in solid tumours but also in children with ALL[24]. This is usually caused by gene deletion, chromosomal translocation and point mutations[24]. Mutations in *NRAS* and *KRAS* are highly prevalent in ALL. *NRAS* mutations are common in ALL and other haematological diseases, while *KRAS* mutations are more common in epithelial malignancies[24]. The majority of NRAS/KRAS mutations identified in ALL occur in codons 12, 13, and 61, which lead to constitutive activation and subsequent hyperactivation of ERK and other MAP kinases (MAPK). Zhang[2] reported a 50% frequency of somatic alterations in the RAS signalling pathway in 187 high-risk childhood B-precursor ALL (Figure 5).
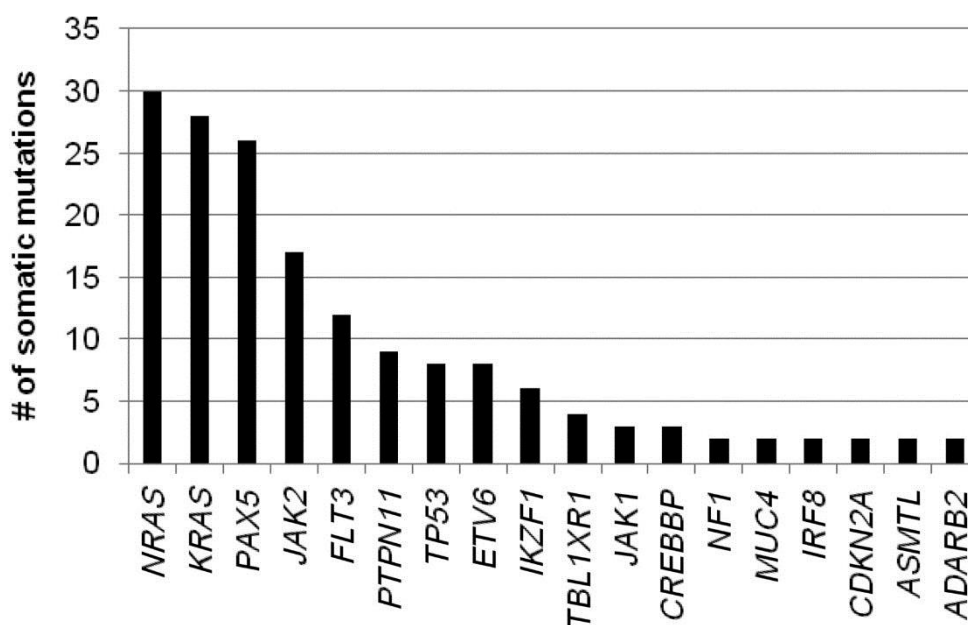
Figure 5 Frequency of key signalling pathway mutations in 187 ALL patients[2].

Moreover, RAS pathway mutations are also highly prevalent in children with relapsed ALL, occurring in 78 out of 206 patients from the ALL-REZ BFM 2002 trial with mutations usually being mutually exclusive[34]. Out of that cohort, 47 matched diagnostic patient samples were studied and it was found that mutations in the *RAS* pathway are often predominate at relapse but may be detected as a subclone at diagnosis[34]. Therefore, MAPK pathway activation can be pharmacologically exploited and may guide new therapies needed for relapsed ALL.

2.6 Negative feedback regulators

Negative feedback regulation is a critical component in many signalling pathways within normal cells[35]. Signalling pathways have to be precisely regulated by positive and negative regulation to drive robust cellular response and to avoid dangerous errors leading to disease[35]. Different mechanisms of negative regulation are responsible for a variety of diverse cellular behaviours such as differentiation and developmental processes[35]. Feedback loops can account for such regulation and are able to induce the expression of self-inhibitors which dampen the signalling duration once a threshold has been reached[35]. This control of signalling protects cells from oncogenic signalling which otherwise might lead to tumour development[4]. In many transformed cells, the level of signalling is increased due to mutations or deletions of negative regulators such as *PTEN*. Interestingly, Pratilas showed that in *BRAF* mutated cancers, increased levels of signalling often results in increased activity of negative feedback regulators[36]. Similar levels of MAPK pathway activation were observed in tumours with mutant *BRAF* and with RTK activation. But only BRAF^{V600E} driven tumours were dependent on ERK signalling for proliferation and its activation resulted in increased levels of MAPK negative feedback regulators such as DUSP6 and SPRTY[36]. Remarkably, in BRAF^{V600E} mutated tumours the high activity of negative feedback did not prevent tumour growth, suggesting that the ERK signalling pathway is insensitive to feedback inhibition[36].

2.6.1 MAPK Negative feedback regulator, DUSP6 (MKP3)

Signalling in the MAPK pathway occurs via phosphorylation hence MAPK phosphatases (MKPs) represent the main form of negative feedback regulation and play a key role in determining the magnitude and duration of kinase activation[37].

Dual specificity phosphatase 6 (DUSP6, also known as MKP3) belongs to the family of 10 mitogen-activated protein kinase (MAPK) phosphatases that are key players in the negative regulation of three major components of the MAPK pathway; ERK1/2, JNK and P38alpha[38-40]. It does so by dephosphorylating either the serine/threonine or tyrosine residue, or both, of the kinase[39,41,42]. DUSP6 negatively regulates signalling from receptor tyrosine kinases (RTKs) such as EGFR[43], FGFR and their downstream

effector RAS[44]. It is regulated at the promoter level by ETS1, a well-known nuclear target of ERK and plays a role in developmental and pathological conditions[43].

While DUSP5, DUSP6 and DUSP7 only dephosphorylate ERK1/2, the phosphatases DUSP1, DUSP4 and DUSP9 target both ERKs, p38 and JNK[37,40]. It appears that DUSP6 serves as a cytoplasmic anchor for ERK, determining its subcellular localization and mediating a spatio-temporal mechanism of ERK signalling regulation[45]. The cytoplasmic localization is mediated by the leucine-rich nuclear export signal located within its N-terminal domain[45].

All DUSPs share a common structure; with a C-terminal phosphatase domain that confers phosphatase activity for both phosphoserine/threonine and phosphotyrosine residues and a N-terminal non-catalytic domain (Figure 6)[8,39].
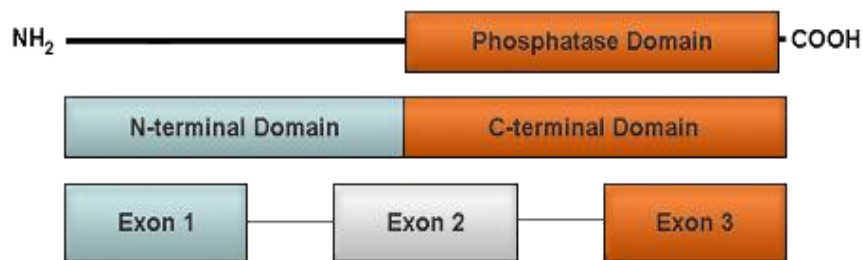
Figure 6 Structure of the DUSP6 (MKP-3) gene
DUSP6 consists of a C-terminal catalytic domain and an
N-terminal non-catalytic domain. Exon 3 and half of Exon
2 encode the functional phosphatase domain.[8]

Dusp6[-/-] mice are viable, fertile and have increased basal ERK1/2 phosphorylation in many organs like heart, spleen, kidney, brain and fibroblasts but do not show a prolonged ERK1/2 activation after stimulation, suggesting that DUSP6 plays a more profound role in basal ERK1/2 signalling tone[46]. There is no change in ERK5, p38 or JNK activtion[46]. It is known that DUSP6 plays a role during otic development; inactivation of DUSP6 leads to partially penetrant postnatal lethality, hearing loss and skeletal malformations[47].

The role of DUSP6 in cancer differs between tumour types and functions in a dual manner; pro-oncogenic or tumour-suppressive[40]. DUSP6 as well as other negative feedback regulators of ERK such as ETV5 and SPRY2 were found to be frequently inactivated in solid tumours and mature B cell lymphoma and have an important role as tumour suppressors in attenuating oncogenic signalling[31,40,48]. In pancreatic cancer, low expression levels of DUSP6 and gain-of-function mutations in *KRAS* might synergize, resulting in constitutive activation of ERK1/2 and thus uncontrolled cell growth[49]. Such behaviour has been observed in non-small lung cancer (NSLC) and ovarian cancer. In NSLC cell lines, DUSP6 expression levels were reduced and restoration of DUSP6 resulted in growth suppression, suggesting anti-tumour effects[50]. In human ovarian cancer cells, loss of DUSP6 mediated by oxidative stress resulted in increased ERK1/2 activity which in turn contributed to tumorigenicity and chemoresistance[51]. DUSP6 has also been reported to have a tumour-suppressive role in oesophageal squamous cell carcinoma and nasopharyngeal carcinoma[40].

In contrast, its pro-oncogenic role can be observed in other cancer types that harbour an activating mutation in the MAPK pathway. For instance, DUSP6 can be found at high levels in melanoma cell lines with a $BRAF^{V600E}$ or $NRAS^{Q61R}$ mutation[52]. Moreover, high levels of DUSP6 expression have also been reported in breast epithelial cells expressing oncogenic *HRAS* and also in myeloma cells harbouring *NRAS* activating mutations[53]. These findings suggest that over-expression of DUSP6 has a compensatory role in the negative feedback control of the MAPK pathway due to activating mutations in *RAS* or *RAF*. The tumour-promoting role of DUSP6 has also been observed in human glioblastoma, endometrial carcinoma and thyroid carcinoma[40]. Interestingly, elevated DUSP6 expression was reported to be correlated with a poor prognosis in NSCLC patients as part of a five gene signature model[54]. However, elevated DUSP6 expression in *KRAS* mutant colon cancer has no effect on survival outcome[55].

Furthermore, colleagues in the Müschen laboratory found that primary samples from patients with pre-B ALL (n=83) showed significantly lower levels of CpG methylation in the promoters of *DUSP6, ETV5* and *SPRY2* as well as higher protein expression levels compared to normal pre-B cells (Figure 7)[4].
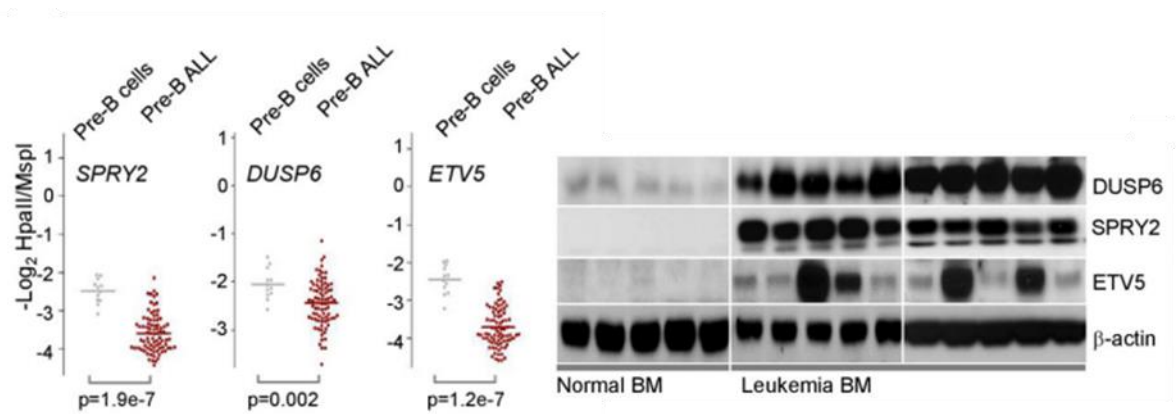
Figure 7 Negative feedback regulators of ERK (DUSP6, ETV5, SPRY2) are selectively activated in human pre-B ALL cells[4].

Conventional treatment regimens for cancer usually focus on agents that will suppress the signalling strength of transforming oncogenes below the minimum level that is needed for survival and proliferation[56]. However, in those cancers that have high levels of DUSP6 expression, inhibiting DUSP6 might be beneficial and open up new treatment options. Molina identified a small-molecule inhibitor of Dusp6 — (*E*)-2-benzylidene-3-(cyclohexylamino)-2,3-dihydro-1*H*-inden-1-one (BCI) — using a transgenic zebrafish chemical screen[44]. It also inhibits DUSP1 at low micromolar concentrations[44].

BCI treatment inhibited DUSP6 activity and resulted in increased FGF signalling leading to an expansion in cardiac progenitors in zebrafish embryos and in cultured cells[44]. Here, DUSP6 regulates heart organ size. *In vitro* studies show that BCI inhibits the catalytic activation of DUSP6 by ERK2 substrate binding[44]. Molina predicted that BCI binds on the DUSP6 phosphatase domain, suggesting an allosteric mechanism of action[44].

Such observations demonstrate the complex role of DUSP6 in cancer and therefore different aspects such as cancer cell type, etiology or disease stage need to be considered before targeting this negative feedback regulator as a therapeutic tool.

2.7 Aims of this Study

I set out to test the hypothesis that in addition to the oncogene itself, other factors also confer permissiveness to oncogenic signalling and allow a normal pre-B cell to tolerate an oncogenic level of signalling strength.

In particular, I examined the role of the ERK negative feedback regulator DUSP6 in RAS-mutant acute lymphoblastic leukaemia and tested its potential as a therapeutic target using genetic mouse models and a small molecule inhibitor targeting DUSP6 (BCI).

3 Materials & Methods

3.1 Materials

3.1.1 Overview of mouse strains used in this study

Mouse strain	Source	Purpose
$Dusp6^{-/-}$	Jackson Laboratories	Genetic loss-of-function experiments
$Mapk1^{fl/fl}$	Dr. Martin McMahon, UCSF	Inducible deletion of $Mapk1^{fl/fl}$
$LSL\text{-}NRAS^{G12D}$	MD Kevin Shannon, UCSF	Inducible hyperactivation of $NRAS$
$BRAF^{F\text{-}V600E}$	Dr. Martin McMahon, UCSF	Inducible hyperactivation of $BRAF$
$Mb\text{-}1$-Cre	Jackson Laboratories	Cre recombinase expression
$CD21$-Cre	Jackson Laboratories	Cre recombinase expression
$NOD/SCID$	Jackson Laboratories	Transplant recipient mice

3.1.2 Overview of genotyping primers used in this study

Primer	Sequence	Band size bp	
$Dusp6$	C_ 5'-CATTGACTCGGAGAGTGATCTGGT-3' WT_ 5'-CAGTCCATCAGCAGCAGCTGTTCG-3' M_ 5'-GGCTCTATGGCTTCTGAGGCGGA-3'	Wild type Mutant	321 250
$Mapk1^{fl/fl}$	F_ 5'-CAGAGCCAACAATCCCAAAC-3' R_ 5'-GGCTGCAACCATCTCACA-3'	Wild type Mutant	278 350
$LSL\text{-}NRAS^{G12D}$	WT F_ 5'-AGACGCGGAGACTTGGCGAGC-3' WT R_ 5'-GCTGGATCGTCAAGGCGCTTTTCC-3' M_ 5'-AGCTAGCCACCATGGCTTGAGTAAGTCTGCA-3'	Wild type Mutant	487 345
$BRAF^{F\text{-}V600E}$	F_ 5'-TGAGTATTTTTGTGGCAACTGC-3' R_ 5'-CTCTGCTGGGAAAGCGGC-3'	Wild type Mutant	185 307
$Mb1$	F_ 5'-TTCAGCCTTCAGTCTAACATC-3' R_ 5'-ATCTGTGAAGACAGGGTGC-3'	Wild type	375
$Mb1$-Cre	F_ 5'-CCCTGTGGATGCCACCTC-3' R_ 5'-GTCCTGGCATCTGTCAGAG-3'	Mutant	450
$CD21$-Cre	F_5'-GCGGTCTGGCAGTAAAAACTATC-3' R_5'-GTGAAACAGCATTGCTGTCACTT-3'	Mutant	100

3.1.3 Retroviral vectors; Constitutive expression

Construct	Overexpression of
MSCV IRES-Neo	Neomycin
MSCV IRES-Puro	Puromycin
MSCV BCR-ABL1 IRES-Neo	BCR-ABL1
MSCV IRES-NRAS^{G12D}-Puro	NRAS^{G12D}
pCGDNsam-IRES- NRAS^{G12D}-Kusabira Orange	NRAS^{G12D}

Retroviral vectors; Inducible expression

Construct	Overexpression of
MSCV ER^{T2} IRES-GFP	GFP
pRetroX-Tet3G	Tet-On
pRetroX-TRE3G-NRAS^{G12D}	Tet-On NRAS^{G12D}
MSCV Cre ER^{T2} IRES-GFP	Cre; GFP
MSCV ER^{T2} IRES-Puro	Puromycin
MSCV Cre ER^{T2} IRES-Puro	Cre; Puromycin

3.1.4 Overview of western blot antibodies used in this study

Antigen	Clone ID	Company
Beta-actin	Polyclonal (ab8227)	Abcam
NRAS	F155	Santa Cruz Biotechnology
DUSP6	G-4	Santa Cruz Biotechnology
ERK1/2	Polyclonal (9102)	Cell Signaling
Phospho -ERK1/2 (T202/Y204)	D13.14.4E	Cell Signaling
BCL6	D8	Santa Cruz Biotechnology
STAT5	Polyclonal (9363)	Cell Signalling
Phospho-STAT5 (Y694)	D47E7	Cell Signalling
Phospho-SHP-1 (Tyr564)	D11G5	Cell Signalling
SHP-1	C14H6	Cell Signalling
p27 Kip1	D69C12	Cell Signalling
P21 Waf1/Cip1	12D1	Cell Signalling
SPRY2	D3G1A	Cell Signalling
Akt	Polyclonal (9272)	Cell Signalling
Phospho-Akt (Ser473)	Polyclonal (9271)	Cell Signalling

3.1.5 Overview of patient-derived pre-B ALL samples studied here

Case	Disease	Karyotype
LAX7	at diagnosis	
LAX7R	at relapse	$KRAS^{G12V}$

Notes: All primary samples are bone marrow biopsies, blast content >80%; LAX, Los Angeles

3.1.6 Overview of cell lines used in this study

Cell line	Type	Source
REH	Pre-B ALL	Müschen
697	Pre-B ALL	Müschen
SMS-SB	Pre-B ALL	Müschen
NALM6	Pre-B ALL	Müschen
Gumbus	Burkitt lymphoma	Dr. D. Hodson
Jeko-1	Mantle Cell Lymphoma	DSMZ
HBL-1	DLBCL	Dr. D. Hodson
DOHH-2	DLBCL	Dr. D. Hodson
JJN3	Multiple Myeloma	Müschen
JSC-1	B-NHL -Primary effusion lymphoma	Dr. J. Jung, USC
L428	Hodgkin's lymphoma	Müschen
KM-H2	Hodgkin's lymphoma	DSMZ

3.1.7 Overview of qRT-PCR Primers used in this study

Primer	Sequence
*Bcl6*_F	5' – CCTGCAACTGGAAGAAGTATAAG – 3'
Bcl6-R	5' – AGTATGGAGGCACATCTCTGTAT – 3'
*Hprt*_F	5' – GGGGGCTATAAGTTCTTTGC - 3'
*Hprt*_R	5' – TCCAACACTTCGAGAGGTCC - 3'

3.2 Methods

3.2.1 Patient samples, cell lines and human cells

Primary cases were obtained in compliance with the Institutional Review Boards of the University of California, San Francisco (UCSF). Primary human ALL samples (LAX7, LAX7R) were cultured on OP9 stroma in Minimum Essential Medium (MEMα, Life Technologies) with GlutaMAX containing 20% FBS (Fetal Bovine Serum), 100 IU ml^{-1} penicillin, 100 µg ml^{-1} streptomycin and 1mM sodium pyruvate at 37°C in a humidified incubator with 5% CO_2. Human cells were cultured in Roswell Park Memorial Institute medium (RPMI-1640, Invitrogen, Carlsbad, CA) with GlutaMAX containing 20% FBS, 100 IU/ml penicillin and 100 µg/ml streptomycin and kept at 37°C in a humidified incubator (5% CO_2 and 95% atmosphere). *See 3.1.5 and 3.1.6 for a list of cell lines and patient samples used in this study.*

3.2.2 Extraction of bone marrow cells from mice

Bone marrow cells were extracted from young (> 6 weeks) transgenic mice as follows: Cavities of femur and tibia were flushed with PBS, filtrated through 40 µM filters and depleted of erythrocytes using lysis buffer (BD PharmLyse, BD Biosciences) before washed cells were either frozen or cultured for experiments. All mouse experiments were subject to institutional approval by University of California, San Francisco (UCSF). *See 3.1.1 for a list of mouse strains used in this study*.

3.2.3 Genotyping and Polymerase Chain Reaction

For molecular genotyping, blood samples from mice were obtained as a source of genomic DNA and used in a Polymerase Chain Reaction (PCR). The Thermo Scientific Phusion Blood Direct PCR Kit was used to amplify DNA from whole blood using a Phusion Hot Start II High-Fidelity DNA Polymerase. Following the PCR, the samples were analysed by agarose gel electrophoresis. This method of tissue collection has been approved by the University of California San Francisco Institutional Animal Care and Use Committee. *See 3.1.2 for a list of genotyping primers used in this study.*

3.2.4 Retroviral production / transduction

Lipofectamine 2000 (Invitrogen) and Opti-MEM media (Invitrogen) were used for transfecting MSCV-based retroviral constructs. 293FT cells plated in a 20 cm^2 dish were co-transfected with 20 μg of plasmid pHIT60 (gag-pol), 5 μg of pHIT123 (ecotropic envelope, provided by D.B. Kohn, University of California, Los Angeles, Los Angeles, CA) and 20 μg of a retroviral construct. Dulbecco's modified Eagle's medium (DMEM, Invitrogen) with Glutamax containing 10% fetal bovine serum, 100 IU/ml penicillin, 100 ug/ml streptomycin, 25mM HEPES, 1mM sodium pyruvate and 0.1 mM non-essential amino acids were used for cultivation. After 16 hours, regular media was replaced by growth media and 10mM sodium butyrate was added. Following 8 h incubation, the media was replaced with regular growth media and 18 h later the virus supernatant was filtered through a 0.45 μM filter. For retroviral transduction, the virus was loaded by centrifugation (2000 x g, 90 min at 32°C) on 50 ug/ml Retronectin (Takara) coated non-tissue 6-plates. 2-3 x 10^6 pre-B cells were loaded on the retronectin coated plate and centrifuged at 600 x g for 30 minutes and maintained at 37°C with 5% CO2 for 2 days before transferring into culture flasks. *See 3.1.3 for a list of retroviral vectors used in this study*.

3.2.5 Transformation of competent cells

50 μl of competent Eschericia coli (E.coli, Invitrogen) cells were mixed with 1 μl DNA, gently mixed by flicking the bottom of the tube and placed on ice for 20-30 min. The tube was then placed into a 42°C water bath for 30-60 sec and immediately put on ice for 2 min. 500 μl of LB media was added and grown in 37°C shaking incubator. After 45 min, the samples were plated onto a pre-warmed 10 cm LB agar plate containing the appropriate antibiotic and left overnight at in 37°C. Next day, bacterial colonies were selected and put into a culture consisting of LB media and the appropriate antibiotic and left for 6-8 h in a 37°C shaking incubator before it was added to a larger overnight culture and left overnight. Next morning, the bacterial plasmids were purified using MAXI or MINI prep (Qiagen).

3.2.6 Plasmid Purification

Bacterial cells from an overnight culture were harvested by centrifugation at 6000 x g for 15 min at 4°C and then plasmids were purified using the QIAGEN Plasmid Maxi Kit following the manufacturer's instructions.

3.2.7 Maintaining Bone Marrow cells in culture with IL7 and generation of inducible NRAS^{G12D} Tet-On 3G Gene Expression Systems

Bone marrow cells from young age-matched mice were harvested and retrovirally transformed with pRetroX-Tet3G (Clontech) retrovirus in the presence of 10ng IL7/ml (Peprotech) and selected for neomycin (1000 µg/ml) for 1-2 weeks. After neomycin selection, cells were further transduced with inducible pRetroX-TRE3G-NRAS^{G12D} (Clontech) retrovirus and selected for puromycin (1 µg/ml) for 1-2 days. Addition of 1µg/ml Doxycycline (Clontech) to the culture medium leads to induction of NRAS^{G12D}. Cells were cultured in Iscove's Modified Dulbecco's Medium (IMDM, Life Technologies) with Glutamax containing 20% FBS, 100 IU ml-1 penicillin, 100 µg/ml streptomycin and 2-ME and kept at 37°C in a humidified incubator (5% CO2 and 95% atmosphere).

3.2.8 FACS (Fluorescence-activated cell sorting)

FACS antibodies were used to stain human and mouse samples according to manufacturer's instructions. FC block was used, as well as respective isotype controls. For cell sorting, cells were collected in a Falcon tube with FBS and were centrifuged at 400 g for 5 min before putting them back in culture.

3.2.9 Dead Cell Apoptosis Analysis

Alexa Fluor® 488 Annexin V/Dead Cell Apoptosis Kit (Life Technologies) was used for the detection of apoptotic cells according to the manufacturer's instructions. Externalization of phosphatidylserine (Annexin V, Alexa Fluor® 488 Conjugate) and DNA content (DAPI, 4',6-diamidino-2-phenylindole) were measured and gating on all cells was used for further analysis.

3.2.10 Western blotting

Cell extracts were prepared as follows: CelLytic buffer (Sigma) supplemented with 1% protease inhibitor cocktail (Roche), 1% Phosphatase inhibitor cocktail (Calbiochem Millipore) and 1mM PMSF (Life Technologies) was used to lyse the cells followed by loading the same amount of protein per sample on NuPAGE (Invitrogen) 4-12% Bis-Tris gradient gels and were further transferred on PVDF membranes (Invitrogen). Primary antibodies were used with the HRP

immunodetection system (Life Technologies/Millipore) to detect mouse and human proteins. Western blot antibodies were obtained from Abcam, Cell Signaling or Santa Cruz Biotechnology. *See 3.1.4 for a list of western blot antibodies used in this study.*

3.2.11 Colony Forming Assay for mouse ALL

10,000 transformed lymphoid cells (or 100,000 for myeloid B cells) were resuspended in murine MethoCult™ GF M3630 (StemCell Technologies) and cultured on 3 cm dishes with an extra water supply to prevent evaporation. Colonies were counted after 7-14 days and photos of the dishes and colonies were taken.

3.2.12 Cell viability assay for human or murine cells

100,000 human derived xenograft cells or NRAS^{G12D}/BCR-ABL transformed mouse ALL cells were plated in a volume of 100 µl medium on Opilux Falcon Clear 96-well plate (BD Biosciences). After 2-3 h incubation at 37°C, agonist/antagonist were diluted in medium and added to each well at the indicated concentration in a total culture volume of 120 µl. After 3 days culturing, 12 µl of Cell Counting Kit 8 solution (CCK-8; Dojindo Inc.) was added to each well, incubated for 1-4 hours at 37°C and absorbance was measured at 450 nm. Baseline values of untreated cells were used as a reference to calculate fold changes (set to 100%).

3.2.13 Extraction of genomic DNA from cultured cells

Genomic DNA from up to 10×10^6 cells was extracted using NucleoSPIN Tissue kit (Macherey-Nagel). Briefly, cells were resuspended in a final volume of 200 µl Buffer T1 and 25 µl proteinase K solution and 200 µl Buffer B3 were added. After incubation at 70°C for 10-15 min, 210 µl ethanol (96-100%) was added to the sample and vigorously vortexed. This was followed by DNA binding to the silica membrane in the columns and pure genomic DNA was eluted in 100 µl Buffer BE.

3.2.14 RNA purification

Total RNA from cells was extracted using NucleoSPIN RNA from Macherey-Nagel and cDNA was produced from 1 µg of total RNA using qScript™ cDNA SuperMix (Quanta Biosciences).

3.2.15 Quantitative real-time PCR (qRT-PCR)

After RNA purification, quantitative real-time PCR was performed with FAST SYBR® Green Master Mix (Applied Biosystems) and the Vii7 real-time PCR system (Applied Biosystems) according to standard PCR conditions. *See 3.1.7 for a list of qRT-PCR Primers used in this study.*

3.2.16 Statistical Analysis

All pairwise comparisons between the means were calculated by a two-tailed t test using GraphPad software.

4 Results

4.1 Examining the role of DUSP6 in RAS-mediated transformation of pre-B cells

4.1.1 DUSP6 deletion reduced colony formation in BRAF^{V600E} driven mouse pre-B cells

The RAS pathway is frequently activated in pre-B ALL[2]. Yet others in the Muschen laboratory have observed surprisingly high expression levels of negative feedback regulators such as *DUSP6*, *ETV5* and *SPRY2*[4]. To test whether this negative feedback might actually be beneficial to leukaemic cells, I took advantage of two existing mouse lines to create an inducible model of BRAF activation with and without DUSP6 deletion. I crossed *Braf*$^{V600E\ CA/CA}$ with *Dusp6*$^{+/+}$ or *Dusp6*$^{-/-}$ mice resulting in the desired genotype after 2 generations (F2 generation, Figure 1A). In *Braf*$^{V600E\ CA}$ knock-in mice, exons 15-18 and a STOP cassette are flanked by *loxP* sites and followed by exons 15-18 containing the activating V600E mutation in exon 15. Following Cre-mediated excision of the *loxP*-flanked region, transcripts are generated using the V600E mutant exon 15 (Figure 1B).

B cell precursors from *Braf*$^{V600E\ CA/+}$ *Dusp6*$^{+/+}$ and *Braf*$^{V600E\ CA/+}$ *Dusp6*$^{-/-}$ mice were cultured in interleukin 7 (IL-7)-containing media and retrovirally transduced with GFP-tagged, 4-hydroxy tamoxifen (4-OHT)-inducible Cre-ER^{T2} (Cre) or empty vector (EV; ER^{T2}). Figure 1C illustrates a schematic of the experimental outline. The Cre recombinase is fused with a mutated form of the oestrogen receptor that is further bound to the heat shock protein (Hsp90)[57]. Upon binding of 4-hydroxytamoxifen (4-OHT) to the mutant estrogen receptor, Hsp90 dissociates from the complex and the Cre recombinase translocates to the nucleus where it excises genomic segments flanked by *loxP* sequences[57].

4-OHT-mediated activation of *Braf*^{V600E} resulted in a strong increase in ERK phosphorylation in *Dusp6*$^{+/+}$ pre-B cells. In *Dusp6*$^{-/-}$ pre-B cells, basal ERK activity was already elevated but increased further upon activation of BRAF^{V600E} (Figure 2A). This confirmed the Cre/*loxP* system was functioning properly.

Next, I performed a competitive fitness assay and colony forming assay to identify whether the oncogene $BRAF^{V600E}$ provides a growth benefit for pre-B cells and whether DUSP6 has a positive or negative effect on $BRAF^{V600E}$ driven pre-B cell growth. IL7-dependent pre-B cells were monitored for GFP^+ cells by flow cytometry over a period of 10 days after induction of $BRAF^{V600E}$. $BRAF^{V600E}$ activation resulted in a growth advantage in pre-B cells, regardless of DUSP6 status (Figure 2B). Deletion of DUSP6 had no effect on cell growth either in the presence or absence of $BRAF^{V600E}$ induction (Figure 2B).

However, in a colony forming assay, induction of $BRAF^{V600E}$ in $Dusp6^{+/+}$ pre-B cells resulted in increased colony formation whereas $Dusp6^{-/-}$ pre-B cells expressing $BRAF^{V600E}$ only moderately induced colony formation (Figure 2C). The EV controls, $Dusp6^{-/-}$ and $Dusp6^{+/+}$ pre-B cells alone failed to transform and induce significant colony formation.

These findings show that in a colony forming assay, deletion of DUSP6 impaired colony formation in $BRAF^{V600E}$ driven pre-B cells compared to $BRAF^{V600E}$ driven $Dusp6^{+/+}$ pre-B cells. However, in a competitive fitness assay, pre-B cells driven by $BRAF^{V600E}$ exert a growth advantage regardless of their DUSP6 status.
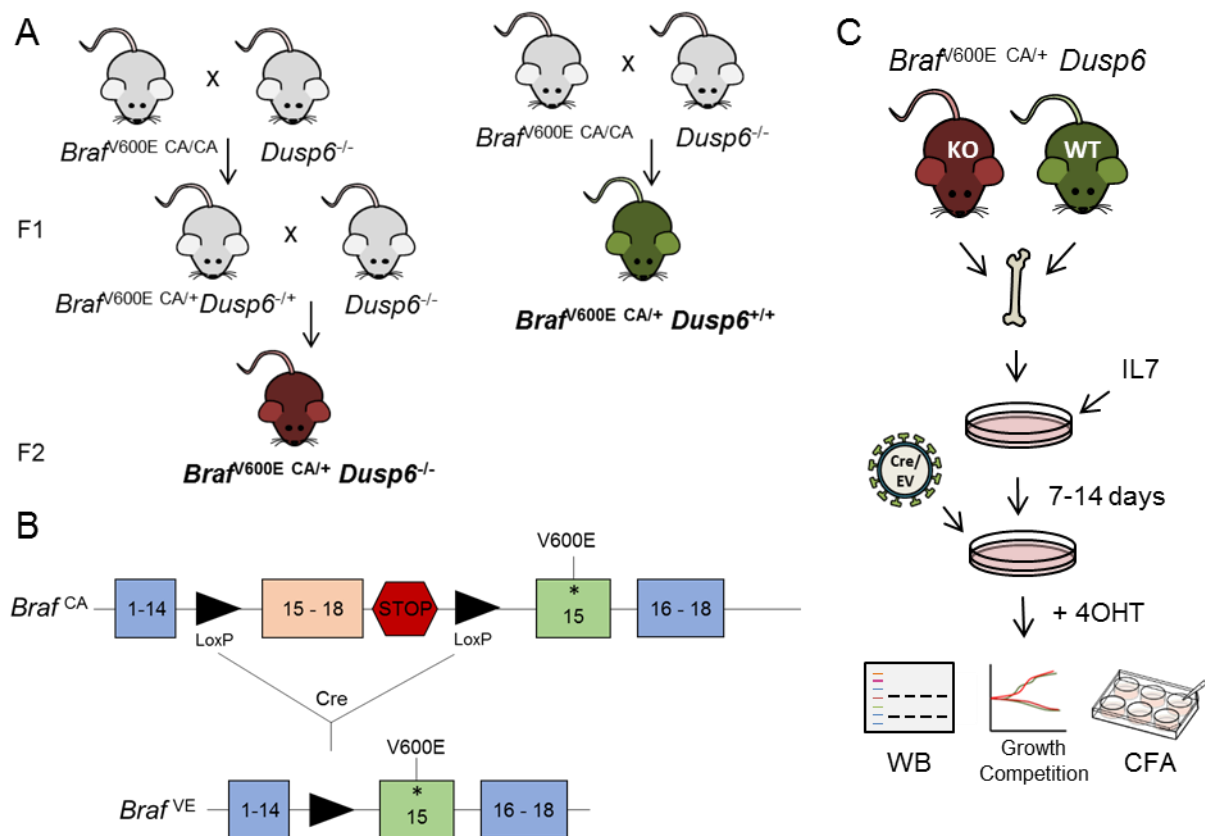
Figure 1 Schematic of mouse crosses and experimental design outline

(A) Schematic diagram of the development of the transgenic mouse $Braf^{V600E\ CA/+}\ Dusp6^{-/-}$ and $Braf^{V600E\ CA/+}\ Dusp6^{+/+}$.

(B) Knock-in mice carry the transcript encoded by endogenous exons 1-14 and *loxP*-flanked human exons 15-18 prior to Cre-mediated recombination. Following Cre-mediated excision, the transcripts are subsequently generated using the mutant exon 15 (which contains constitutively active $Braf^{V600E}$) and the endogenous downstream exons 16-18. Blue boxes: mouse exons, orange box: human exon, green box: mouse exon modified to harbour a V600E amino acid substitution.

(C) Bone marrow from age-matched $Braf^{V600E\ CA/+}\ Dusp6^{-/-}$ and $Braf^{V600E\ CA/+}\ Dusp6^{+/+}$ mice was harvested and pre-B cells were selected with IL7 and cultured for 7-14 days before transduced with GFP-tagged, 4-OHT-inducible Cre or empty vector (EV). After the addition of 4-OHT, Western Blot, Growth Competition Assay and Colony Forming Assay were performed.
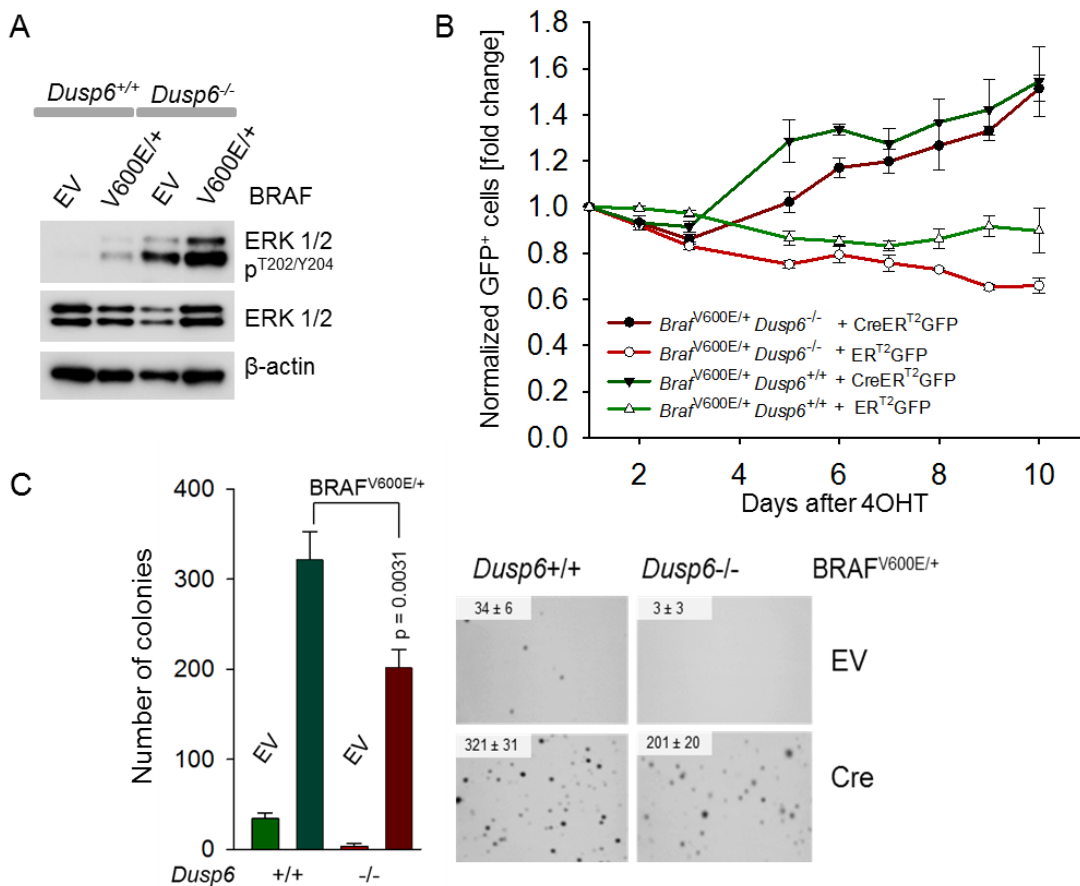
Figure 2 *Dusp6* deletion reduced colony formation in BRAF^{V600E} driven mouse pre-B cells

(A) Western blot analysis on *Dusp6$^{+/+}$* and *Dusp6$^{-/-}$* IL7-dependent pre-B cells in the presence and absence of BRAF^{V600E}. Total and phosphorylated ERK1/2 is shown and β-actin was used as loading control. Representative of > 3 experiments.

(B) *Dusp6$^{+/+}$ Braf$^{V600E/+}$* and *Dusp6$^{-/-}$ Braf$^{V600E/+}$* pre-B cells were transduced with GFP-tagged, 4-OHT inducible Cre or empty vector control (EV). Following induction with 4-OHT, enrichment or depletion of GFP$^+$ cells was monitored by flow cytometry. Illustrated is time course showing the average relative changes (± SD) of GFP$^+$ cells following induction (n = 3). Representative of 2 experiments with 3 replicates/experiment.

(C) Representative images and quantification (mean values ± SD, n = 3) of the colony-forming ability of *Dusp6$^{+/+}$ Braf$^{V600E/+}$* and *Dusp6$^{-/-}$ Braf$^{V600E/+}$* pre-B cells. Following induction with 4-OHT for 48 hours; 10, 000 cells were seeded in semi-solid methylcellulose and monitored for colony formation for 14 days. The p value was calculated from t test. Representative of 3 experiments with 3 replicates/experiment.

4.1.2 DUSP6 is essential for NRAS[G12D]-mediated pre-B cell transformation

To further investigate the significance of DUSP6-mediated negative control of ERK signalling in RAS-driven pre-B ALL, I used a doxycycline-inducible vector system for expression of oncogenic *NRAS*[G12D]. Here, *NRAS*[G12D] is under the control of a TRE3G promoter that only activates transcription when cultured in the presence of doxycycline (Dox), a synthetic tetracycline derivative[58]. Doxycycline binds to the Tet-On 3G transactivator protein which in turn binds the TRE3G promoter, activating transcription[58].

Dusp6[+/+] and *Dusp6*[-/-] IL7-dependent pre-B cells were transduced with the described doxycycline-inducible vector system for regulated gene expression of NRAS[G12D]. Figure 3A illustrates a schematic of the experimental outline. I performed a colony forming assay in methylcellulose to study self-renewal ability in the presence and absence of NRAS[G12D] and found that *Dusp6*[+/+] pre-B cells expressing NRAS[G12D] were able to transform and increase colony formation whereas *Dusp6*[-/-] pre-B cells expressing NRAS[G12D] failed to transform or induce colony formation (Figure 3B). A western blot was performed to detect changes in ERK phosphorylation levels. Induction of NRAS[G12D] in *Dusp6*[+/+] pre-B cells caused a moderate increase in ERK phosphorylation. In *Dusp6*[-/-] pre-B cells, the basal level of ERK is higher and further increased upon induction of NRAS[G12D] (Figure 3C).

Interestingly, I found that in *Dusp6*[-/-] pre-B cells, acute activation of NRAS[G12D] caused acute toxicity, leading to a decrease in cell viability whilst in *Dusp6*[+/+] pre-B cells there was no significant impact on cell viability after 3 days (Figure 3D).

These findings suggest that DUSP6 is required for RAS-driven B cell lineage ALL, presumably through negative feedback control in buffering ERK signal strength. In contrast, in other tumour types (e.g. solid tumours) negative feedback regulation of ERK acts as a tumour suppressor and opposes RAS-driven malignant transformation[31]. Thus, there appears to be a lineage specific requirement for DUSP6 in ERK mediated transformation of precursor B cells.
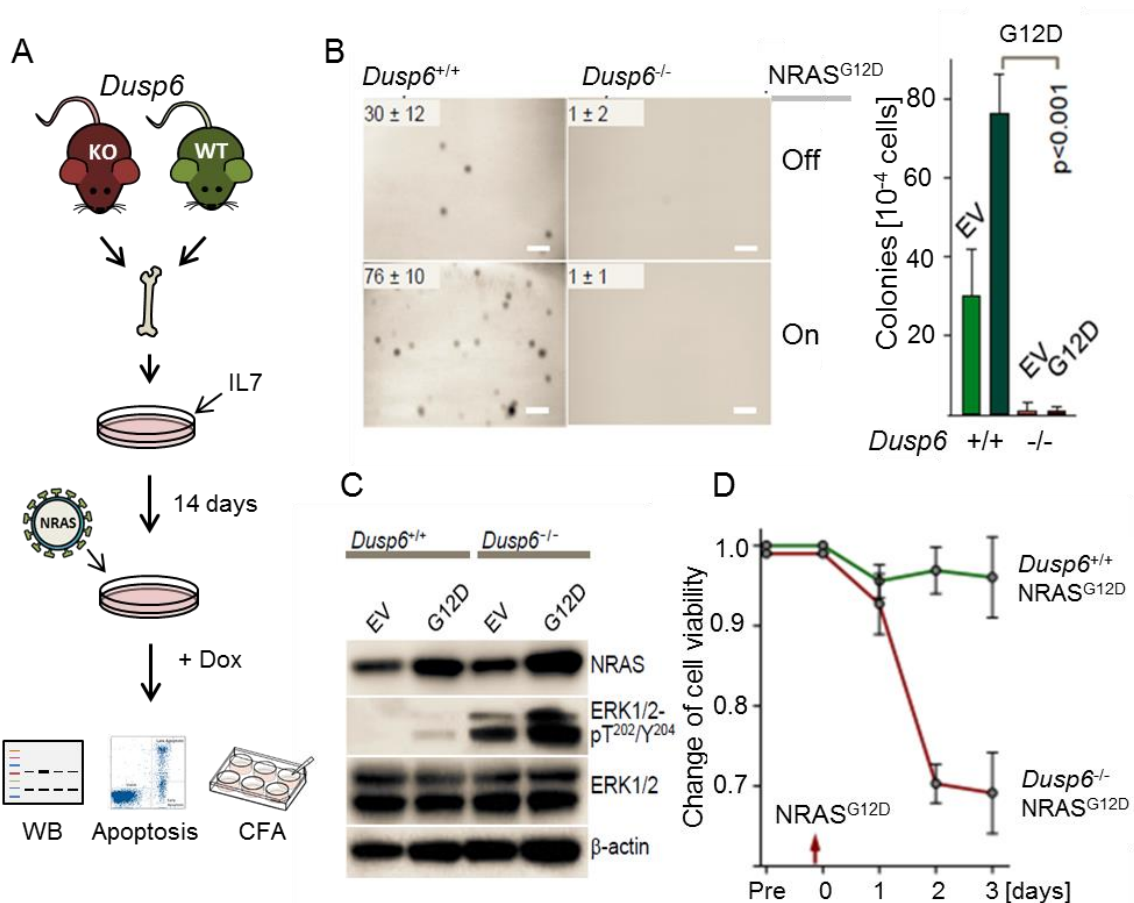
Figure 3 DUSP6 is essential for NRAS^{G12D} - mediated pre-B cell transformation

(A) Bone marrow from age-matched *Dusp6$^{+/+}$* and *Dusp6$^{-/-}$* mice was harvested and IL7 dependent pre-B cells were engineered to carry a doxycycline (Dox)-inducible Tet^{On}-*NRAS^{G12D}* mutant. After the addition of Dox; Western Blot, Apoptosis and Colony Forming Assay were performed.

(B) Representative images and quantification (mean values ± SD, n = 3) of the colony-forming ability of *Dusp6$^{+/+}$* and *Dusp6$^{-/-}$* IL7-dependent pre-B cells transduced with doxycycline inducible NRAS^{G12D}. 10, 000 cells were seeded in semi-solid methylcellulose and monitored for colony formation for 14 days. The scale bar represents 1mm. The p value was calculated from t test.

(C) Western blot analysis on *Dusp6$^{+/+}$* and *Dusp6$^{-/-}$* IL7-dependent pre-B cells in the presence and absence of NRAS^{G12D}. NRAS, total and phosphorylated ERK1/2 are shown and β-actin was used as loading control. Representative of 3 experiments.

(D) Cell Viability of *Dusp6$^{+/+}$* and *Dusp6$^{-/-}$* IL7-dependent pre-B cells upon acute induction of NRAS^{G12D} was measured by Annexin V/DAPI, and analyzed by flow cytometry. Illustrated is time course showing the average relative changes (± SD) of cell viability. Representative of 3 experiments with 3 replicates/experiment.

38

4.1.3 Pharmacological inhibition of DUSP6 in human relapsed pre-B ALL can overcome drug-resistance

To examine whether DUSP6 represents a potential therapeutic target for the treatment of pre-B ALL, I used a recently developed small molecule inhibitor of DUSP6, 2-benzylidene-3-(cyclohexylamino)-2,3-dihydro-1H-inden-1-one (BCI, Figure 4A). BCI acts within the phosphatase domain of DUSP6 to prevent the catalytic stimulation of phosphatase activity induced by ERK2[44].

I studied two patient-derived pre-B ALL cells that were isolated either at time of diagnosis (LAX7) or at the time of relapse (LAX7R) when cells had acquired a $KRAS^{G12V}$ mutation. The patient relapsed 11 months after the initial diagnosis and whole exome sequencing verified the acquisition of the new $KRAS^{G12V}$ mutation (Figure 4B). This provided a system for me to test DUSP6 inhibition in a paired human cell system with and without ERK activation. Both samples were treated with BCI with increasing concentrations and a western blot performed to study ERK phosphorylation levels.

As expected, I found oncogenic activation of ERK in LAX7R due to the acquired RAS mutation whereas LAX7 (*KRAS* wild-type) showed no phosphorylation of ERK (Figure 4C). DUSP6 small molecule inhibition using BCI in the *KRAS* wild-type sample (LAX7) had no biochemical effect. However, in the $KRAS^{G12V}$ mutant LAX7R cells, constitutive ERK activity was strongly increased with increasing concentrations of BCI (Figure 4C). This is expected as inhibiting ERK negative feedback regulation should result in an increase in MAPK signalling.

Typically, relapsed ALL cells are considered to be more resistant to therapy than de novo disease[13]. However, when I performed a cell viability assay on *KRAS* wild-type LAX7 and $KRAS^{G12V}$ mutant LAX7R cells, I found that the relapse ALL cells were more sensitive to BCI than cells from the diagnostic sample (LAX7) (Figure 4D).

Acquisition of mutations in the RAS pathway frequently occur in pre-B ALL and hence this finding highlights a potential strategy that may be of use to overcome drug-resistance in such patients.
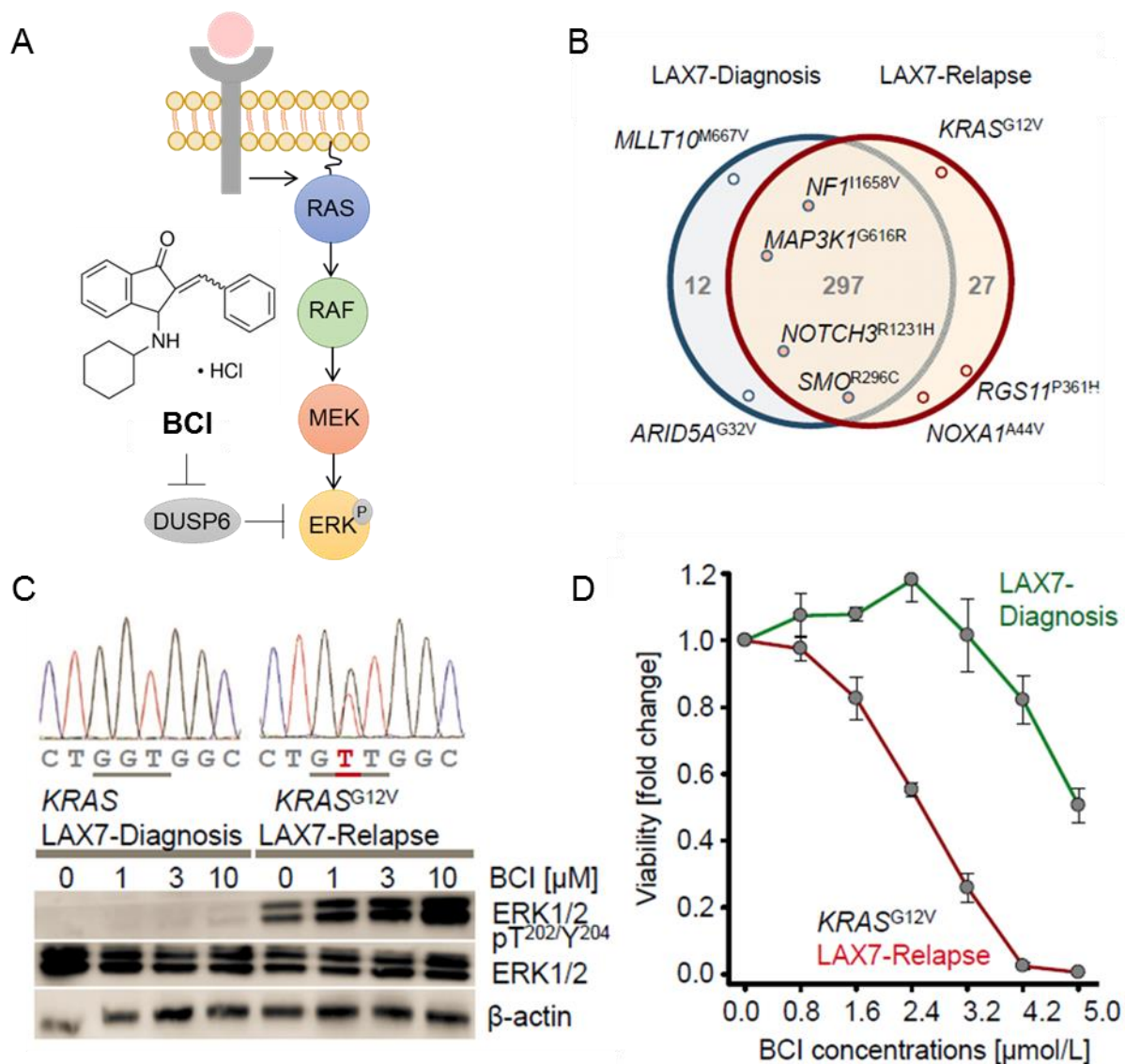
Figure 4 _KRAS_^{G12V} mutant cells are more sensitive to small molecule inhibition of DUSP6 than KRAS WT cells

(A) Chemical structure of DUSP6 inhibitor (E/Z)-BCI hydrochloride.

(B) Whole exome sequencing identified mutations shared between LAX7 and LAX7R as well as unique mutations (Analysis performed by Christian Hurtz).

(C) Chromatogram showing the absence of the _KRAS_^{G12V} mutation in LAX7 and the presence in LAX7R. Western blot analysis showing the biochemical effect BCI has on ERK activity in LAX7 and LAX7R. β-actin was used as loading control. Representative of 2 experiments.

(D) Cell Viability assay with LAX7 and LAX7R when incubated with BCI at increasing concentrations for 72 hours. Illustrated are the average relative changes (± SD) of cell viability. Representative of 2 experiments with 3 technical replicates/experiment.

4.2 Examining the role of DUSP6 in mature B cell malignancies

4.2.1 Mature B cell lymphomas without functional BCR were non-responsive to DUSP6 inhibition whereas pre-B ALLs are highly sensitive

Since we had shown that DUSP6 was highly upregulated and a tumour activator in pre-B ALL (Shojaee and Caeser[4]), I wondered if the same phenomenon would apply to mature B cell malignancies. Previous work had reported epigenetic inactivation of another MAPK pathway inhibitor *Spry2* led to B cell lymphoma progression, suggesting a tumour suppressor role[48]. This prompted me to examine the role of DUSP6 in late B cell development.

Here, I studied the effect of pharmacological inhibition of DUSP6 by the small molecule inhibitor BCI in a range of ALL (REH, 697, SMS-SB, NALM6), lymphoma (Gumbus, Jeko-1, HBL-1, DOHH2, JSC-1, L428, KM-H2) and multiple myeloma (JJN3) cell lines. I performed a cell viability assay with increasing concentrations of BCI and discovered three groups of human B cell malignancies that differ with respect to their sensitivity to experimental perturbation by BCI (Figure 5). These groups were:

1.) Pre-B ALL (REH, 697, SMS-SB, NALM6): highly sensitive to experimental perturbation by BCI at approximately 1 μM
2.) Mature B cell lymphomas (Gumbus, Jeko-1, HBL-1, DOHH2): largely resistant to perturbations by BCI up to 4 μM
3.) Mature B cell lymphomas (JSC-1, L428, KM-H2) and multiple myeloma (JJN3): completely non-responsive to perturbations by BCI up to 4 μM

Interestingly, one factor that distinguished Group 2 and 3 was their B cell receptor status. Group 2 consists of Gumbus (Burkitt lymphoma), Jeko-1 (Mantle cell lymphoma), HBL-1 and DOHH2 (Diffuse large B cell lymphoma) that harbour a functional BCR[9]. Group 3 consists of JJN3 (Multiple myeloma), JSC-1 (Primary effusion lymphoma), L428 and KM-H2 (Hodgkin's lymphoma) that typically lack a functional BCR[9]. The differences seen in respect to their sensitivity to ERK hyperactivation may relate to stage specific differences BCI might have due to different autoimmunity checkpoints such as central and peripheral tolerance and I will further elaborate on this in the Discussion section.

To investigate a potential role of DUSP6 at controlling ERK signalling strength in these cell lines, I treated each cell line with increasing concentrations of BCI for 2 hours and then performed a western blot studying ERK phosphorylation levels (Figure 6). As expected, pre-B ALL cells showed an increase in ERK phosphorylation with increasing concentrations of BCI (Group 1). Interestingly, mature B cell lymphomas that were largely resistant to BCI (Group 2) showed the opposite effect and a decrease in ERK phosphorylation could be seen. This group might have evolved other highly robust feedback control mechanisms such as SPRY2. Mature B cell lymphomas and multiple myeloma in Group 3 showed mostly no change in ERK phosphorylation levels or no clear pattern with increasing BCI concentrations. BCI selectively inhibits DUSP6 and DUSP1 and is not reported to block any other phosphatases[44].

This finding highlights that human B cell malignancies differ with respect to ERK signalling strength and have different sensitivities towards experimental perturbations.
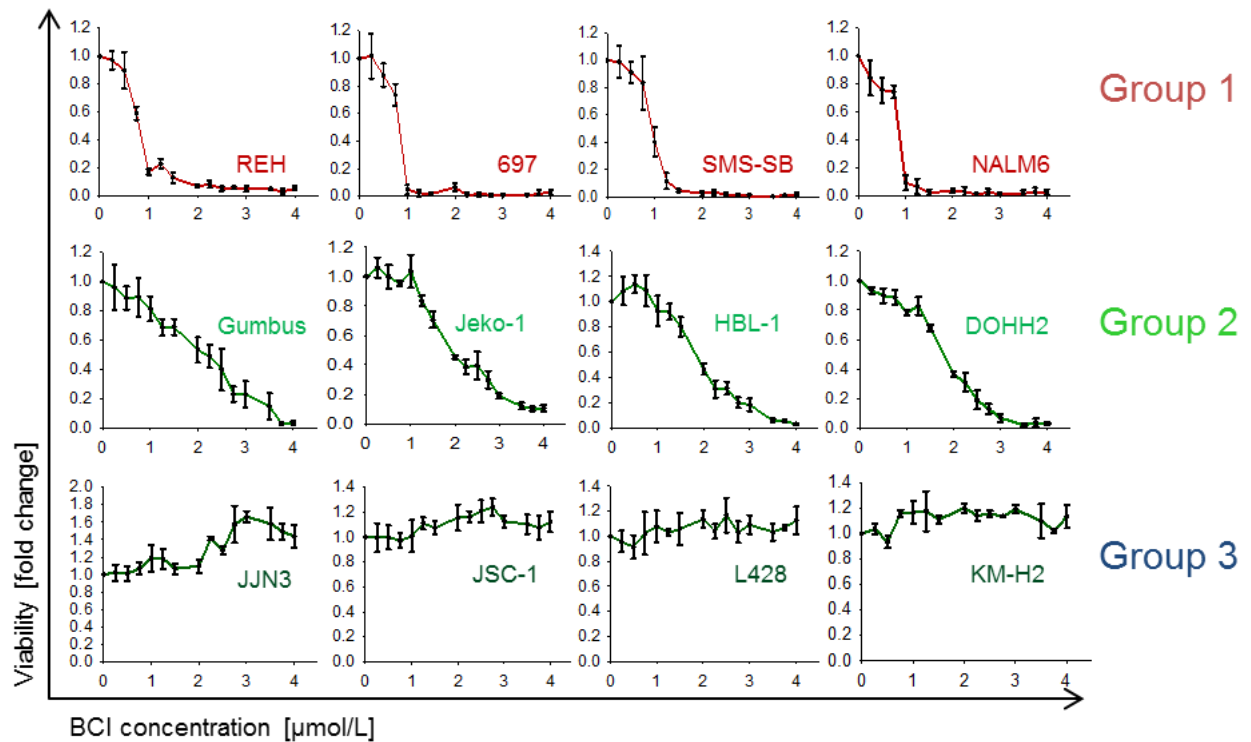
Figure 5 Mature B cell lymphomas and multiple myeloma without a functional BCR are non-responsive to DUSP6 inhibition whereas pre-B ALLs are highly sensitive

Cell Viability assay with pre-B ALL cell lines (REH, 697, SMS-SB, NALM6), mature B cell lymphomas with a functional BCR (Gumbus, Jeko-1, HBL-1, DOHH2) and mature B cell lymphomas and multiple myeloma without a functional BCR (JJN3, JSC-1, L428, KM-H2) when incubated with BCI at increasing concentrations for 72 hours. Illustrated are the average relative changes (± SD) of cell viability (n = 3). Representative of 2 – 3 experiments per cell line with 3 replicates/experiment.
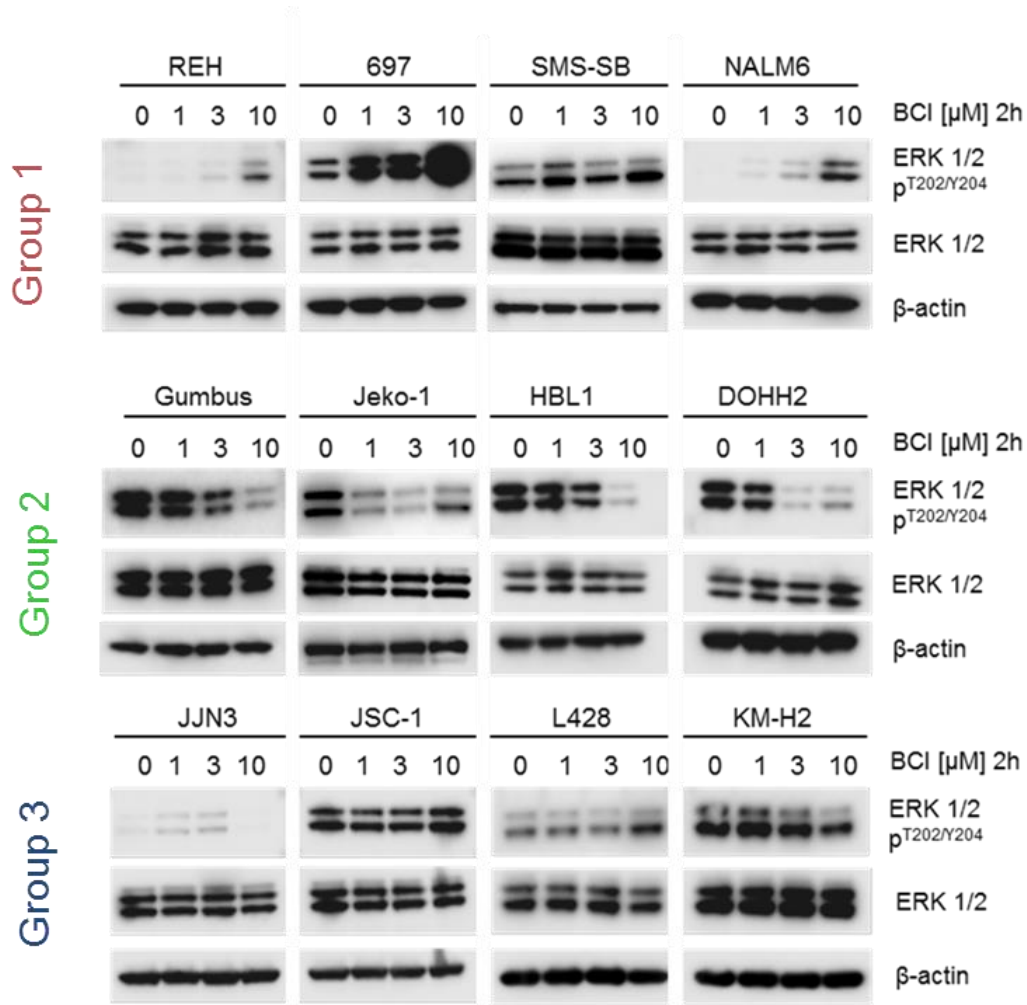
Figure 6 Human B cell malignancies differ with respect to ERK signalling strength

Western blot analysis showing the biochemical effect increasing concentrations of BCI has on ERK activity in all three groups identified. Total and phosphorylated ERK1/2 is shown and β-actin was used as loading control. Representative of 2 – 3 experiments per cell line.

4.3 Examining the consequences of loss of ERK2 function

4.3.1 Loss of ERK2 function resulted in reduced colony forming ability of pre-B cells but has no effect in NRAS^{G12D}- or BCR-ABL1-driven pre-B ALL cells

Given that activating lesions in the RAS pathway resulting in ERK1/2 phosphorylation are found in ~50% of high-risk childhood pre-B ALL at diagnosis[2], I have focused here on the effector kinases ERK1/2. It is not entirely clear yet whether ERK1 and ERK2 have distinct physiological functions or whether they operate redundantly. ERK2 (MAPK1) is known to be crucial in T cell development[59] and here, I investigated the specific role of ERK2 in B cell proliferation.

To do so, I used a 4-hydroxy tamoxifen-inducible Cre-*loxP* Activation system of *Mapk1*$^{fl/fl}$ mutant mice. These *Mapk1*$^{fl/fl}$ mutant mice carry *loxP* sites flanking exon 2 of the *Mapk1* gene that is deleted following cre-mediated recombination.

Progenitor B cells from *Mapk1*$^{fl/fl}$ mice were transduced with either NRAS$^{G12D-MIP}$ or BCR-ABL1^{Neo} (p210-^{Neo}). After respective selection with puromycin or neomycin, BCR-ABL1^{Neo} pre-B cells were further transduced with 4-OHT-inducible Cre-MIP and empty vector control (EV-MIP). NRAS$^{G12D-MIP}$ pre-B cells were transduced with GFP-tagged, 4-OHT-inducible Cre or empty vector control (EV) and sorted for GFP expression. Cre activity was induced using 4-hydroxy tamoxifen (4OHT) resulting in deletion of MAPK1 (ERK2), which was confirmed by western blot (Figure 7A). Interestingly, phosphorylation of ERK1 seems to compensate for the loss of ERK2.

Next, I performed a colony forming assay in methylcellulose to study proliferation ability in the presence and absence of MAPK1/ERK2 in pre-B cells, NRAS^{G12D} and BCR-ABL1 driven ALL cells. Cre-mediated deletion of *Mapk1/Erk2* in *Mapk1*$^{fl/fl}$ pre-B cells resulted in significantly decreased colony formation in comparison to empty vector control (Figure 7B). By contrast, those ALL cells driven by either NRAS^{G12D} or BCR-ABL1 were able to transform and showed strong colony formation regardless of MAPK1 status (Figure 7C & Figure 8A). However, an effect was seen in a serial re-plating assay that is based on their self-renewal ability. Serial re-plating showed that loss of MAPK1 significantly decreased colony formation in comparison to empty vector control (Figure 7C & Figure 8A). Transformed pre-B cells expressing either NRAS^{G12D} or BCR-ABL1 exert impaired self-renewal ability.

Additionally, Cre-mediated deletion of *Mapk1* in NRAS^{G12D} expressing pre-B ALL cells led to depletion of cells in a growth competitive assay over 7 days (Figure 8B). Here, *Mapk1*$^{fl/fl}$ pre-B cells were transduced with NRAS^{G12D}- Kusabira Orange followed by GFP-tagged, 4-OHT-inducible Cre or empty vector control (EV).

These findings suggest that MAPK1/ERK2 activity is required for proliferation of non-malignant cells, however ALL cells driven by an oncogene (*NRAS*^{G12D} and *BCR-ABL1*) do not need ERK2 activity for proliferation. This might be due to ERK1 compensating for the loss of ERK2 however this did not rescue the impairment observed in cell proliferation in pre-B cells.
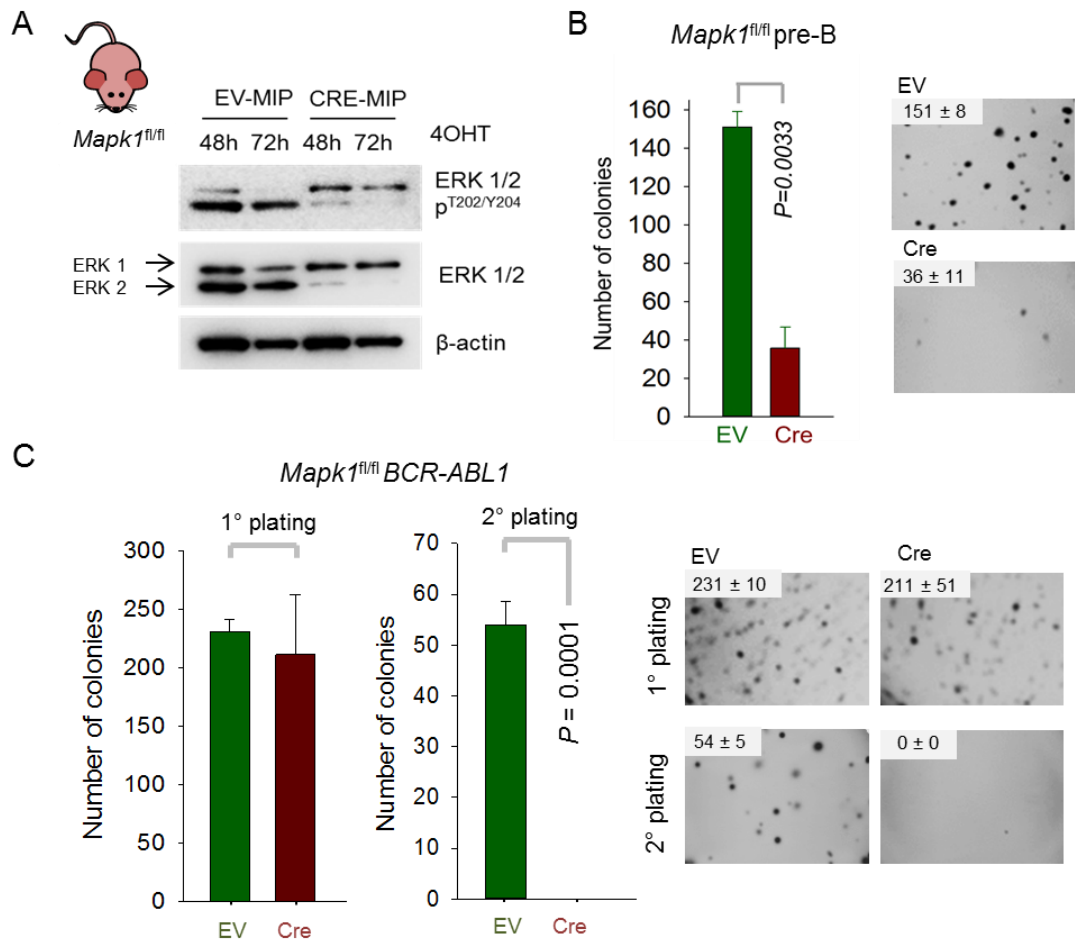
Figure 7 Deletion of *Mapk1* resulted in reduced colony forming ability in pre-B cells but had no effect in *Mapk1*^{fl/fl} *BCR-ABL1*-driven ALL

(A) *Mapk1*^{fl/fl} pre-B cells were transduced with 4-OHT inducible Cre or empty vector control (EV). Following 4-OHT treatment for 48 and 72 hours, western blotting showed successful deletion of *Mapk1/Erk2*. ERK1 and ERK2 are indicated. Representative of > 3 experiments.

(B) *Mapk1*^{fl/fl} pre-B cells were transduced with puromycin tagged, 4-OHT-inducible Cre or empty vector control (EV). After puromycin selection, cells were treated with 4-OHT for 48 hours and 10, 000 cells seeded in semi-solid methylcellulose and monitored for colony formation for 14 days. Representative images and quantification (mean values ± SD, n = 3) are shown. The p value was calculated from t test. Representative of > 3 experiments with 3 replicates/experiment.

(C) Colony forming ability of *Mapk1*^{fl/fl} pre-B cells expressing *BCR-ABL1* was examined by serial re-plating upon deletion of *Mapk1*. The p value was calculated from t test. Representative of 3 experiments with 3 replicates/experiment for 1° plating.
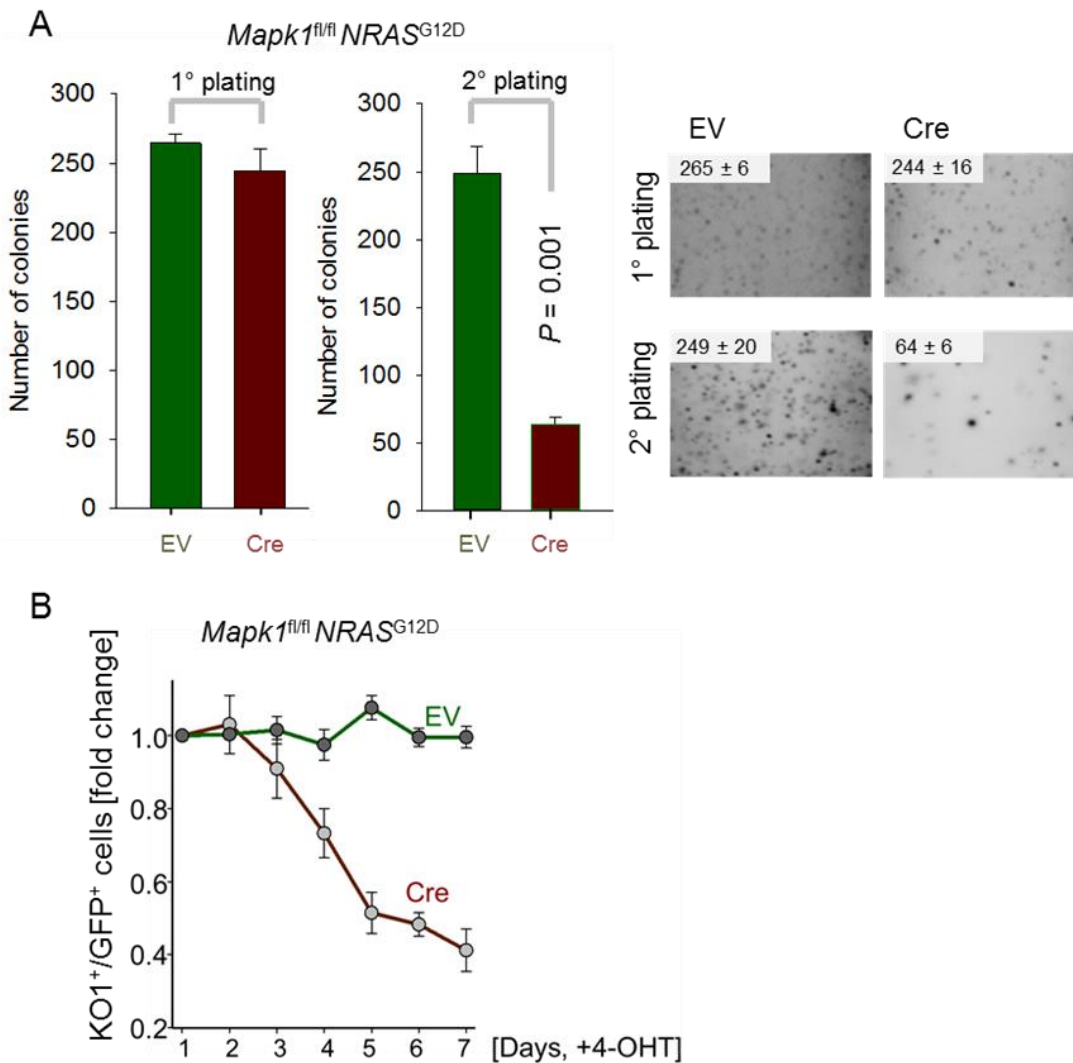
Figure 8 Deletion of *Mapk1* in NRAS^{G12D}-ALL cells had no effect on colony formation but shows a growth disadvantage compared to empty vector control

(A) *Mapk1$^{fl/fl}$-NRAS^{G12D}-ALL* cells were transduced with GFP-tagged, 4-OHT-inducible Cre or empty vector control (EV) and sorted for GFP expression. Cells were treated with 4-OHT for 48 hours and 10, 000 cells were seeded in semi-solid methylcellulose and monitored for colony formation for 14 days. Representative images and quantification (mean values ± SD, n = 3) are shown. The p value was calculated from t test. Representative of 3 experiments with 3 replicates/experiment.

(B) *Mapk1$^{fl/fl}$* pre-B cells expressing KO1 tagged-*NRAS^{G12D}* were transduced with GFP-tagged, 4-OHT-inducible Cre or EV. Following induction with 4-OHT, enrichment or depletion of double positive KO1$^+$/GFP$^+$ cells was monitored by flow cytometry. Illustrated is time course showing the average relative changes (± SD) of GFP$^+$ cells following induction (n=3).

4.3.2 ERK2 activity proved to be redundant for colony formation in myeloid progenitor B cells and in NRAS^{G12D} or BCR-ABL1 CML-like cells

I have shown that MAPK1/ERK2 activity is required for colony formation in pre-B cells but is redundant in pre-B cells driven by an oncogene such as *NRAS*^{G12D} and *BCR-ABL1* (Figure 7+8). Interestingly, Shojaee and Caeser[4] showed that even though robustness of ERK negative control is a predictor of poor clinical outcome in B cell-lineage, it is not in myeloid leukaemia. The question arises whether my finding is also a lineage specific effect. Therefore, I tested potential divergent-specific functions by performing colony forming assays in a myeloid environment for pre-B cells cultured in IL-3/IL-6/SCF and NRAS^{G12D} and BCR-ABL1 CML-like cells.

Bone marrow from *Mapk1*$^{fl/fl}$ mice was harvested and B cell progenitors were cultured in the presence of IL-3/IL-6/SCF before transduction with either NRAS$^{G12D-MIP}$ or BCR-ABL1^{Neo}. After respective selection with Puromycin or Neomycin, cells were transduced with GFP-tagged, 4-OHT inducible Cre or empty vector control (EV) and sorted for GFP expression. I found that Cre-mediated deletion of *Mapk1/Erk2* in *Mapk1*$^{fl/fl}$ myeloid pre-B cells resulted in no significant difference in colony forming ability in comparison to empty vector control (Figure 9A). Myeloid ALL cells driven by NRAS^{G12D} or BCR-ABL1 were able to form colonies regardless of MAPK1 status (Figure 9B&C). However, I observed that those cells transduced with NRAS^{G12D} formed three times more colonies than *BCR-ABL1* transduced cells.

This finding suggests that ERK activity is redundant for colony forming ability in myeloid progenitor B cells and in NRAS^{G12D} or BCR-ABL1 CML-like cells. This is in contrast to the requirement for ERK activity for colony formation in lymphoid progenitor B cells.
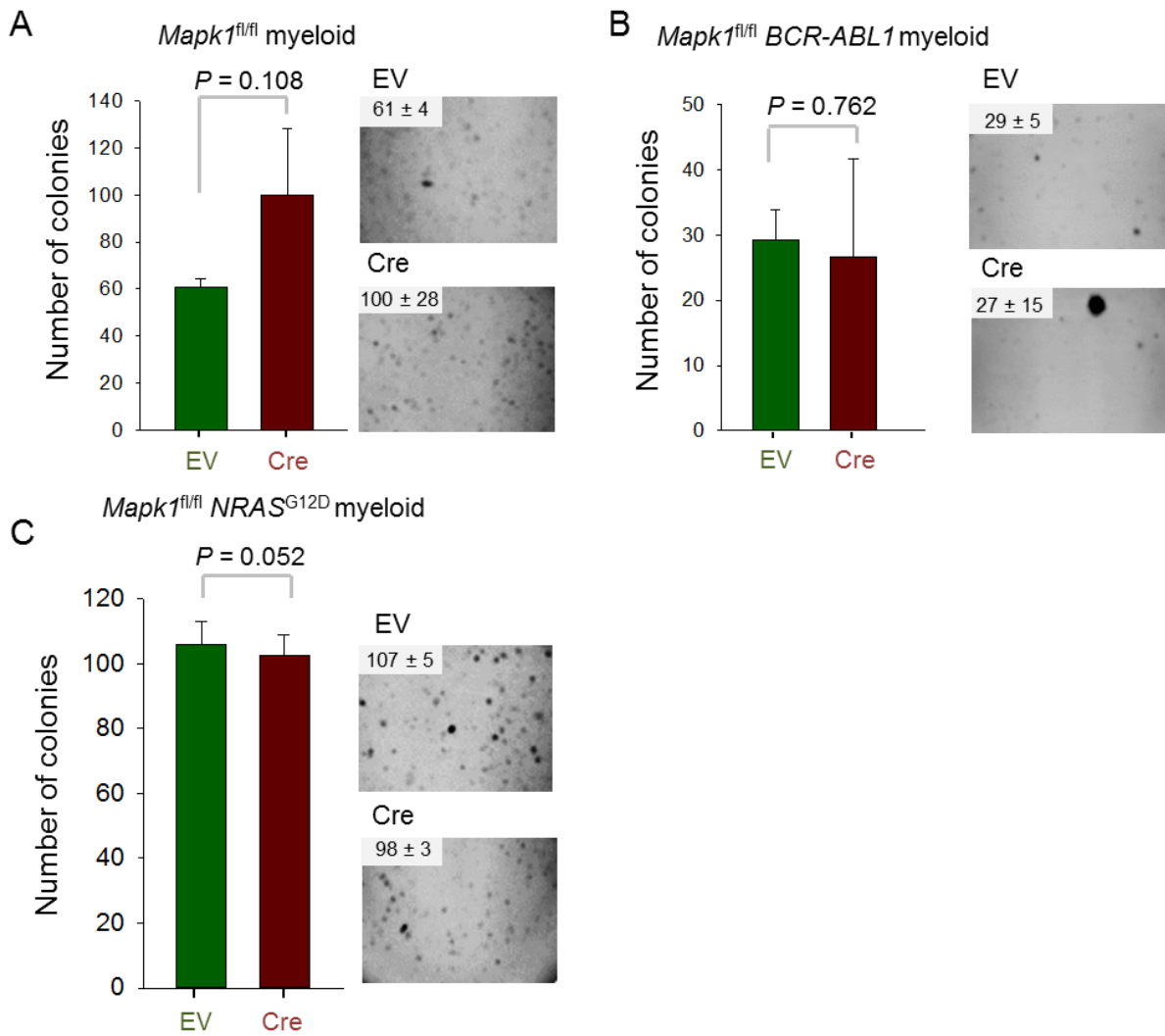
Figure 9 Deletion of *Mapk1* in myeloid progenitor B cells and *NRAS*[G12D] or *BCR-ABL1* CML-like cells had no effect on colony formation ability

Mapk1[fl/fl] pre-B cells were cultured with IL-3/IL-6/SCF (A) and transduced with either BCR-ABL1[Neo] (B) or *NRAS*[G12D-MIP] (C) followed by infection with GFP$^+$ tagged, 4-OHT-inducible Cre or empty vector control (EV). Cells were treated with 4-OHT for 48 hours and 100, 000 cells were seeded in semi-solid methylcellulose and monitored for colony formation for 14 days. Representative images and quantification (mean values ± SD, n = 3) are shown. The p value was calculated from t test.

4.4 Examining the role of BCL6 in RAS-mediated transformation of pre-B cells

Aberrant activation of oncogenic RAS can lead to oncogene-induced senescence (OIS) which is characterized by Arf/p53 induction and irreversible cell cycle arrest in the G_1 phase[60]. The transcriptional repressor and protooncogene *BCL6* was identified as a factor that rescues RAS-induced senescence[61]. BCL6 is also known to cause drug resistance in *Ph*[+] acute lymphoblastic leukemia (ALL)[62], highlighting the importance of developing new therapeutic targets. Given that activating lesions in the RAS pathway are found in ~50% of high-risk childhood pre-B ALL at diagnosis[2], I tested the hypothesis that BCL6 may facilitate RAS-mediated transformation of pre-B cells.

4.4.1 Activation of the RAS-ERK signalling pathway upregulated BCL6 expression

My previous experiments shown here demonstrated that negative feedback regulation of ERK signalling is required for oncogenic transformation. Here, I examined the consequences of acute activation of BRAF[V600E] on BCL6 expression in the presence and absence of DUSP6.

As before (Figure 1A), induction of BRAF[V600E] by 4-OHT in murine pre-B cells led to increased phosphorylation of ERK in *Dusp6*[+/+] pre-B cells. ERK signalling was constitutively active in *Dusp6*[-/-] pre-B cells and was further increased upon induction of BRAF[V600E] (Figure 10A). Interestingly, BCL6 expression was upregulated following BRAF[V600E] activation at the protein (Figure 10A) and mRNA (Figure 10B) level. *Dusp6*[+/+] pre-B cells had higher BCL6 protein expression level than *Dusp6*[-/-] pre-B cells, however this is reversed at the mRNA level suggesting post-transcriptional gene regulation. Consistent with this, elevated BCL6 levels in association with hyperactivation of ERK were observed in a patient-derived *KRAS*[G12V] mutant LAX7R cells but not in the LAX7 diagnostic sample (Figure 10C).

These findings suggest that increased BCL6 protein expression was a consequence of ERK activation.
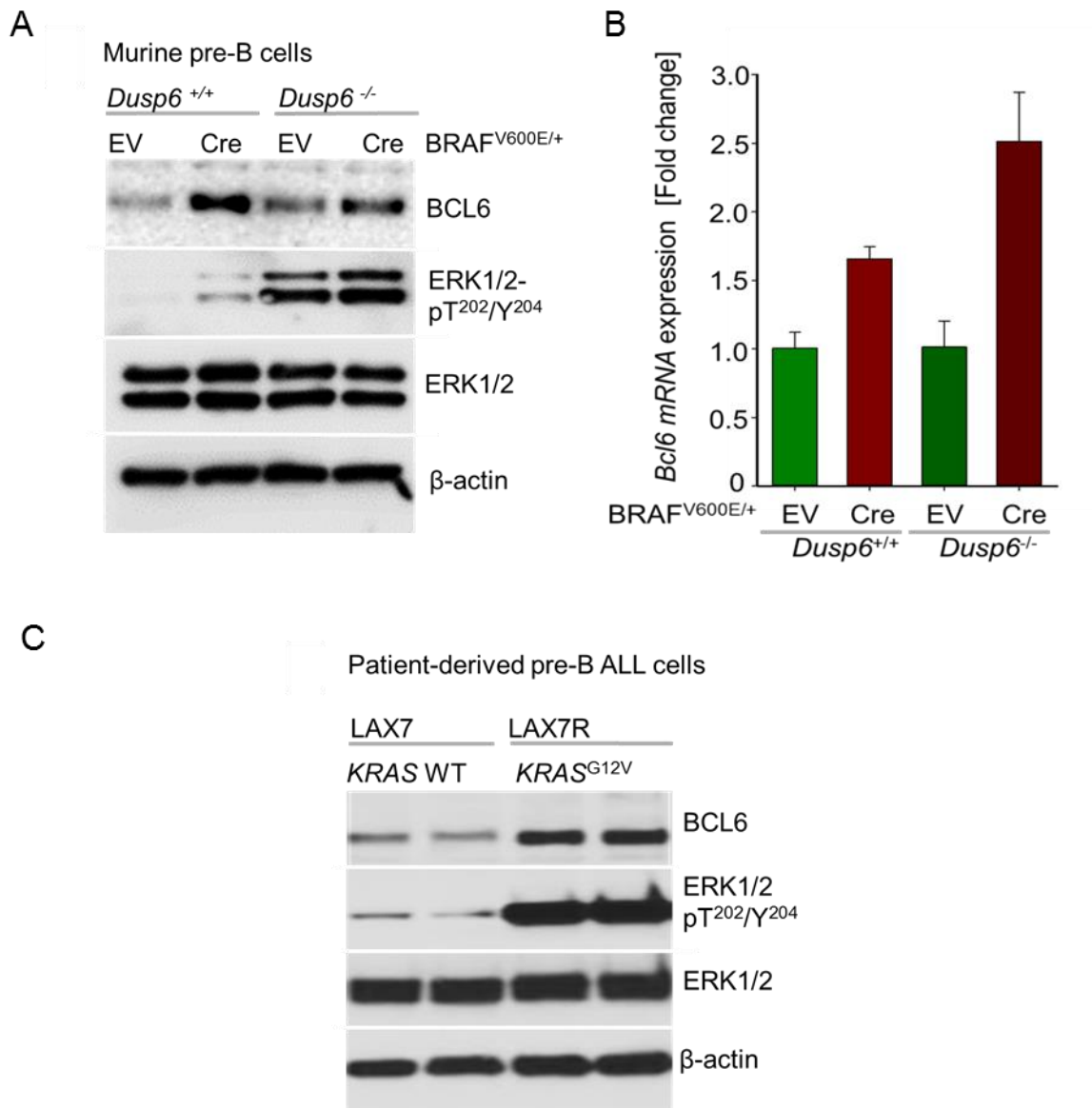
Figure 10 Activation of the RAS-ERK pathway upregulates BCL6 expression

(A) *Braf^V600E/+ Dusp6^+/+* and *Braf^V600E/+ Dusp6^-/-* pre-B cells were transduced with 4-OHT inducible Cre or empty vector control (EV). Following 4-OHT treatment for 48 hours, western blotting was used to examine levels of BCL6, phospho-ERK1/2-T^{202}/Y^{204}, ERK1/2 and β-actin. Representative of 3 experiments.

(B) Expression levels of Bcl6 were also examined following treatment with 4-OHT for 24 hours by qRT-PCR (n=3).

(C) Levels of BCL6, phospho-ERK1/2-T^{202}/Y^{204}, ERK1/2 and β-actin in patient-derived LAX7 (*KRAS* WT) and LAX7R (*KRAS^G12V^*) cells were examined by western blotting. This western blot was performed by Christian Hurtz.

4.4.2 Oncogenic RAS signalling induced BCL6 expression by suppressing STAT5 Activity

Next, I wished to identify the mechanism by which RAS pathway activation induced BCL6 expression. Previous work by Duy[62] showed that STAT5 represses BCL6 in BCR-ABL1-transformed pre-B ALL.

Here, I took advantage of the doxycycline-inducible vector system for expression of oncogenic NRAS^{G12D} (described in 4.1.2) in murine pre-B cells. Interestingly, while overexpression of NRAS^{G12D} resulted in oncogenic activation of ERK in pre-B cells, phosphorylation of STAT5^{Y694} was strongly inhibited (Figure 11A). Similarly, treatment of patient-derived *KRAS*^{G12V} mutant pre-B ALL cells (LAX7R) with increasing concentrations of BCI, (2-benzylidene-3-(cylohexylamino)-1-Indanone hydrochloride), a small molecule inhibitor of ERK - DUSP6, resulted in reduced phosphorylation of STAT5^{Y694} and hyperactivation of ERK (Figure 11B).

It was previously reported that imatinib increases BCL6 expression in BCR-ABL1-transformed pre-B ALL cells[62]. Here, I found that Cre-mediated deletion of *Mapk1/Erk2* suppressed imatinib-mediated induction of BCL6 in BCR-ABL1-driven ALL cells, as well as phosphorylation of STAT5 which was seen to be mutually exclusive with ERK phosphorylation levels (Figure 11C). STAT5 signalling is known to be negatively regulated by a SH2-domain protein tyrosine phosphatase PTPN6[63]. Here, I found that whilst oncogenic NRAS^{G12D} signalling resulted in lower STAT5 activity (Figure 11A&B), expression of NRAS^{G12D} strongly induced the inhibitory protein tyrosine phosphatase PTPN6 in association with ERK hyperactivation (Figure 11D).

Furthermore, BCL6 is known to mediate transcriptional repression of cell cycle checkpoint regulators p21 (CDKN1A) and p27 (CDKN1B)[62] so here I examined the role of such transcriptional repressors upon NRAS^{G12D} induction. I observed that in murine pre-B cells, induction of NRAS^{G12D} following Dox treatment resulted in upregulation of p27, p21 as well as p53 (Figure 11E).

Those findings suggest that activation of oncogenic RAS induced BCL6 expression through PTPN6-mediated dephosporylation of STAT5 and upregulation of transcriptional repressors.
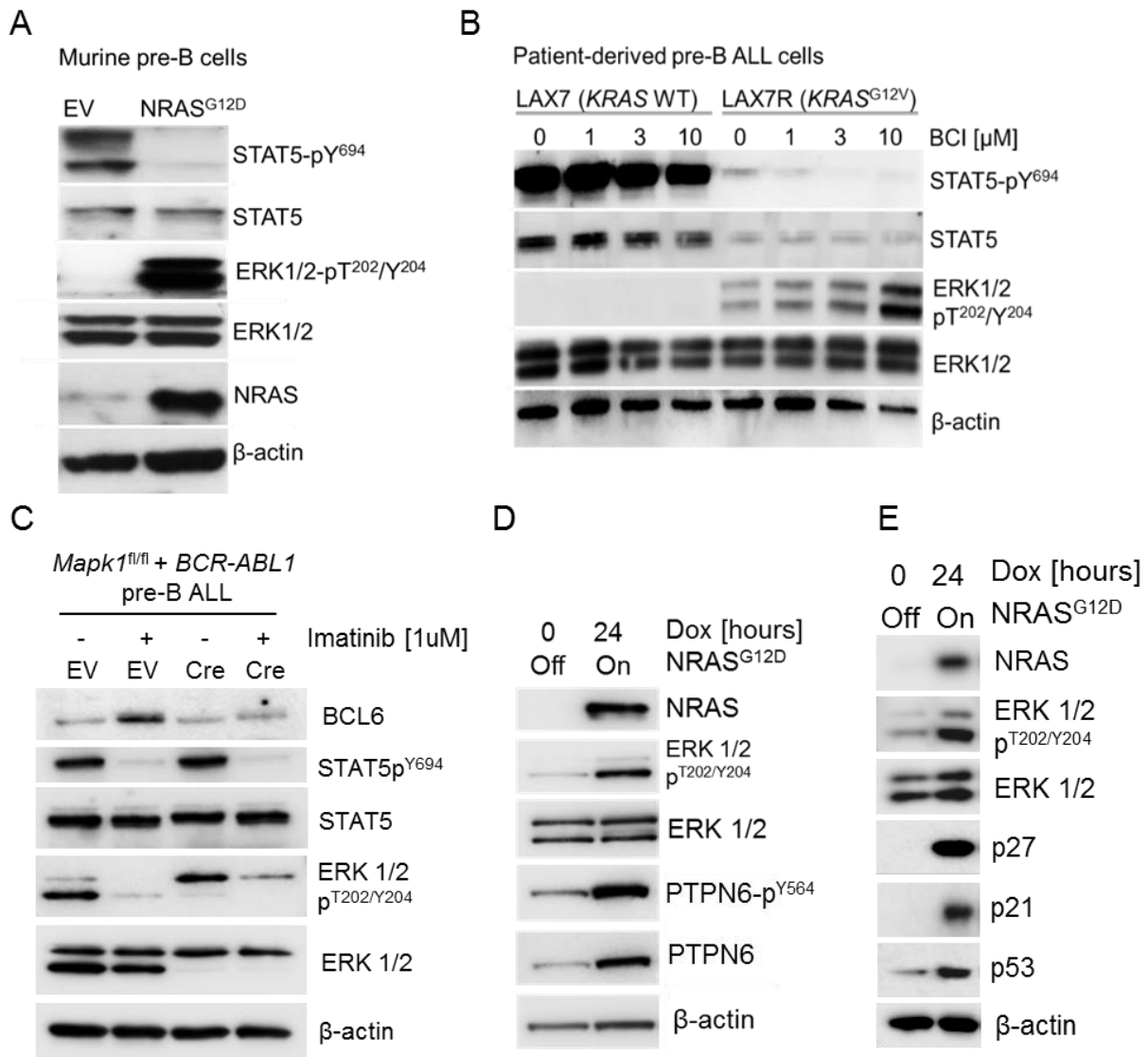
Figure 11 Activation of oncogenic RAS induced BCL6 expression through PTPN6-mediated dephosporylation of STAT5 and upregulation of transcriptional repressors

(A) Levels of phospho-STAT5-Y^{694}, STAT5, phospho-ERK1/2-T^{202}/Y^{204}, ERK1/2, NRAS, and β-Actin were examined in murine pre-B cells expressing EV or NRAS^{G12D}. Western blot was performed by Seyedmehdi Shojaee.

(B) Patient-derived pre-B ALL samples at diagnosis (LAX7; *KRAS* WT) and at relapse (LAX7R; *KRAS*^{G12V}) were treated with increasing concentrations of BCI. Western blotting was used to examine levels of phospho-STAT5-Y^{694}, STAT5, phospho-ERK1/2-T^{202}/Y^{204}, ERK1/2 and β-actin. Representative of 2 experiments.

(C) *Mapk1*[fl/fl] BCR-ABL1-driven pre-B cells were transduced with 4-OHT inducible Cre or empty vector control (EV). Following 4-OHT treatment for 48 hours, cells were treated with imatinib (1µM) or vehicle control for 24 hours. Western blot for BCL6, phospho-STAT5-Y^{694}, STAT5, phospho-ERK1/2-T^{202}/Y^{204} and ERK1/2 was performed using β-actin as loading control. Representative of 3 experiments.

(D) Levels of phospho-PTPN6-Y^{564}, PTPN6, phospho-ERK1/2-T^{202}/Y^{204}, ERK1/2, NRAS, and β-Actin were examined in murine pre-B cells expressing EV or NRAS^{G12D}.

(E) Levels of NRAS, phospho-ERK1/2-T^{202}/Y^{204}, ERK1/2, p27, p21, p53 and β-actin were examined in murine pre-B cells expressing EV or NRAS^{G12D}.

5 Discussion

It is known that in solid tumours, acute hyperactivation of oncogenes such as RAS or BRAF can cause oncogene-induced senescence which functions to limit tumour development[31]. It is believed that the activation of such pathways initially leads to a burst of proliferation followed by cellular senescence and therefore can be described as a tumour suppression mechanism that prevents the progression of benign lesions[31]. The cause of oncogene-induced senescence is suggested to be due to strong activation of negative feedback pathways such as RasGEFs, Sprouty proteins, RasGAPs, and DUSPs rather than hyperactive RAS signalling itself[31]. These negative pathway regulators are designed to inhibit RAS/PI3K signaling which in turn activate mediators of the senescence response such as Rb and p53[31]. Therefore, it is not surprising to see that in advanced BRAF^{V600E}- and RAS-driven solid tumours, negative feedback regulators of ERK are frequently disabled by being deleted, mutated or hypermethylated[36].

In apparent contrast to this paradigm, negative feedback regulators of the MAPK pathway are highly expressed in pre-B ALLs. Using genetic mouse models and a small molecule inhibitor targeting DUSP6, I have shown that in RAS-driven pre-B ALL, ERK-negative feedback control enables malignant transformation; making it a therapeutic target in human pre-B ALL. This might seem counterintuitive to current treatment strategies that aim to inhibit oncogenic signalling as this study proposes the opposite strategy[56]. For instance, I have shown that cells lacking hyperactive ERK signalling (LAX7) were less dependent on ERK-negative feedback control and are spared when treated with the small molecule inhibitor BCI in comparison to RAS-driven ALL cells (LAX7R) that showed a high sensitivity towards BCI. This hypothesis was not supported by the growth competition assay I performed using a BRAF^{V600E} genetic mouse model. Here, pre-B cells driven by BRAF^{V600E} exerted a growth advantage regardless of their DUSP6 status. This experiment was conducted over 10 days and it is possible that initially pre-B cells driven by BRAF^{V600E} get a strong proliferative boost regardless of DUSP6 status and exert growth changes later on.

B cell malignancies that originate from different stages of B cell development are under intense selective pressure and unlike other types of cancer, are uniquely

susceptible to negative selection of self-reactive clones[9]. B cell malignancies remain fully sensitive to autoimmunity checkpoints (AIC), that lymphocytes have evolved during early stages of B cell development to protect against autoimmune disease[9]. An autoreactive BCR that binds self-antigen results in strong BCR signalling above the maximum signalling threshold. This leads to negative selection and cell death at AICs, enforcing central B cell tolerance[64]. Cells that signal below the minimum threshold owing to a non-functional BCR are eliminated by neglect[65]. Positive selection of nonautoreactive immature B cells requires tonic BCR signalling resulting in intermediate BCR signaling strength and survival[66]. According to this scenario, ERK hypersignalling, a distal BCR signaling substrate, would induce cell death equivalent to checkpoint activation for the deletion of autoreactive B cells (Figure 5.1). To avoid elimination of these RAS-driven B cells due to hypersignalling, tumour cells highly express negative feedback signalling to bring signalling below the maximum threshold[4]. This sensitivity of B cells to different ERK signalling strengths makes ERK negative feedback a potential therapeutic target in pre-B ALL. As described in this thesis, DUSP6 inhibition by BCI led to hyperactivation of ERK and subsequent cell death in human pre-B ALL cells – most likely due to AIC activation. Colleagues in the Müschen laboratory have subsequently tested this hypothesis in an *in vivo* setting and demonstrated that BCI treatment induced remission in immunodeficient mice injected with patient derived pre-B ALL cells (Shojaee and Caeser[4]). Furthermore, albeit seemingly paradoxical given the well-known role of PTEN as a tumour suppressor, small molecule inhibition of PTEN, which induced strong AKT hyperactivation, was also toxic in pre-B ALL cells[67]. Acute hyperactivation of BCR signalling can also be achieved by inhibition of phosphatases such as SHP1 and SHIP1 resulting in cell death in pre-B ALL cells[68]. The concept of blocking negative feedback regulators and pushing signalling above the maximum threshold resulting in AIC activation and cell death might represent a promising therapeutic avenue in oncogene-driven ALL and could broaden currently available treatment strategies[9].

In contrast, inhibiting signalling from tyrosine kinase receptors pushes signalling below the minimum signalling level needed for their survival[56] (Figure 5.1). Over the last few decades, new molecular and genomic technologies have been developed which led to enormous progress in oncology. Many tyrosine kinase inhibitors have been approved by the FDA for the use in humans such as imatinib (Gleevec) against

BCR-ABL1-leukaemia[69] or erlotinib against EGFR in non-small cell lung cancer[70]. Furthermore, Irving proposed the use of MEK inhibitors such as selumetinib to attenuate oncogenic ERK signaling in RAS pathway-mutated ALL cells[34].
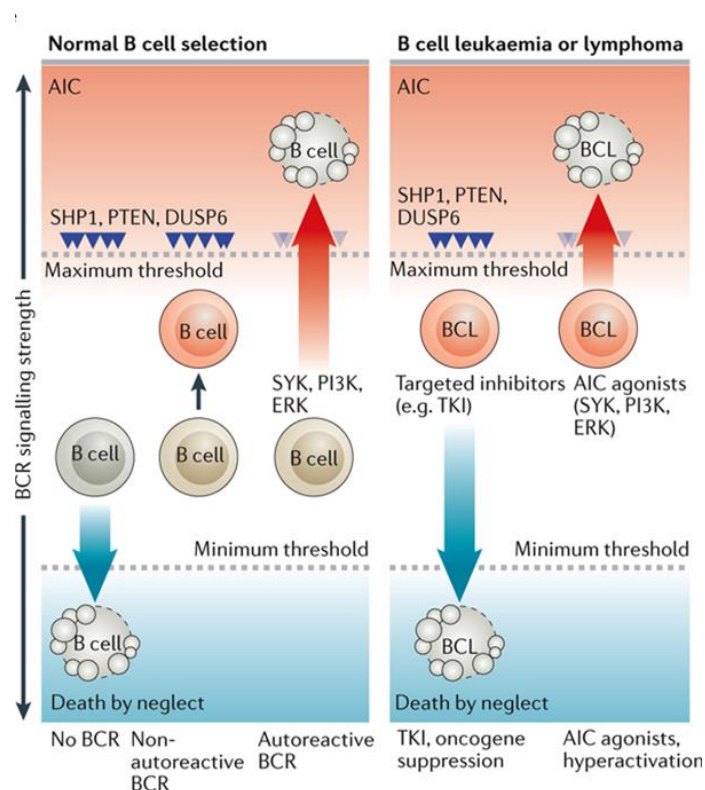
Figure 5.1 Regulation of BCR signalling strength in normal B cells and B cell leukaemia (modified from Markus Müschen[9]).

However, the proposed treatment strategy of pharmacological autoimmunity checkpoint agonists in human pre-B ALL has two main limitations. First, oncogenic signalling from SYK, PI3K and ERK are well-known drivers of several different cancers, and whether prolonged treatment with AIC agonists may lead to the development of new cancers through activation of dormant pre-malignant cells in other organs is unclear[9]. A solution to this might be short pulses of AIC treatment that are potentially sufficient to commit pre-B ALL cells to cell death[9]. Secondly, a major limitation of TKI's and potentially AIC agonists is the development of drug resistance due to newly acquired genomic alterations. Big efforts have been put into combatting such drug resistance and ponatinib, a third-generation TKI, is of particular note because it suppresses all BCR-ABL1 single mutants in Philadelphia chromosome-positive (Ph$^+$) leukaemia, including the drug resistance mediator BCR-ABL1^{T315I} [71].

Hence, as with any targeted therapy, it seems likely that inhibition of DUSP6 will select for subclones resistant to DUSP6 inhibition. In our paper (Shojaee and Caeser[4]) we discuss a strategy to avoid this scenario by alternating therapies so that selective pressures from opposite ends are applied sequentially (Figure 5.2).
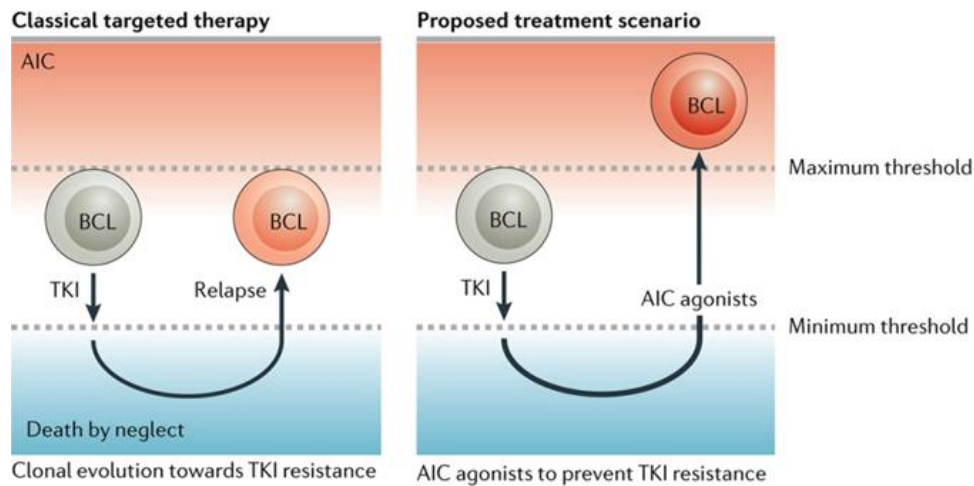
Figure 5.2 Proposed treatment regimen to overcome drug-resistance Treatment with tyrosine kinase inhibitors drives malignant cells into signal starvation; however, a majority of patients develop drug resistance. Subsequent treatment with BCI causes excessive signalling and forces such surviving cells out of their comfort zone leading to cell death. (Markus Müschen[9])

This sequential treatment regimen starts with using TKI or MEK inhibitors to attenuate oncogenic signalling, driving those cells into a ''signal starvation zone''. It is believed that leukaemic cells may adapt to low levels of oncogenic signalling, leaving them in a ''comfort zone''[4]. This is followed by a second round of treatment with an AIC agonist to push signalling in the opposite direction. In this scenario, TKI treatment and selection for resistant clones do not confer adaptive fitness in the subsequent second treatment round[4]. While this treatment strategy is promising and might represent a new therapeutic concept, detailed studies of negative feedback inhibitors such as BCI will be needed in an *in vivo* setting.

The discovery of three groups of human B cell malignancies in this study that fundamentally differ with respect to their ability to control ERK signalling strength

might also be explained on the basis of autoimmunity checkpoints. The first group consisting of pre-B ALL cells have high levels of negative feedback regulators such as DUSP6 and treatment with BCI results in overwhelmingly strong BCR signalling and subsequent cell death at low concentrations of BCI. These cells are subject to central B cell tolerance; triggering negative selection and AICs induced cell death[9]. Signal hyperactivation was seen on the protein level when blotting for phospho-ERK.

The second group consisting of BCR[+] B cell lymphomas are likely to be driven by constitutively active signalling from an autoreactive BCR and have potentially escaped elimination at the central B cell tolerance checkpoint. Following BCI treatment, these cells are now being subject to peripheral tolerance; the second checkpoint of immunological tolerance[22] and undergo cell death at high concentrations of BCI. Following treatment with BCI at increasing concentrations, phospho-ERK levels decrease suggesting these cells have most likely evolved other highly robust feedback control mechanisms and are therefore largely resistant to perturbations.

The third group consists of B cell lymphomas that lack a B cell receptor and are completely non-responsive to BCI and do not exhibit significant activity of ERK feedback control. Mature B cell lymphomas such as classical Hodgkin Lymphomas or Primary Effusion Lymphoma often lack a BCR due to crippling somatic mutations that were acquired during affinity maturation in the germinal centre[9]. This group might already be beyond autoimmunity checkpoints making them resistant to perturbations.

Moreover, I have shown that ERK2 deletion leads to impaired B cell proliferation in murine pre-B cells but is redundant in NRAS[G12D] and BCR-ABL1-driven pre-B ALL. This suggests one functional isoform is sufficient for proliferation in oncogene-driven pre-B ALL but not in normal pre-B cells. ERK1 protein expression increased after ERK2 deletion suggesting a compensatory effect; however, this did not rescue the impairment observed in cell proliferation in pre-B cells. Studies have shown that the loss of ERK2 usually has a greater effect than loss of ERK1, despite them being highly similar[28,72]. The greatest effect can be seen in *Erk2*[-/-] mice that are embryonically lethal[30] whereas *Erk1*[-/-] mice are viable and fertile[29]. Interestingly, transgenic expression of *Erk1* rescues the placental developmental defects seen in *Erk2*[-/-] mice[73]. Furthermore, in neural development, loss of ERK2 resulted in impaired proliferation of neural progenitors and production of fewer neurons resulting in loss of

higher-order functionality in the mature brain[72]. However, loss of ERK1 did not result in an overt learning phenotype[74]. Cell proliferation was also slowed down in fibroblasts after reducing expression of ERK2[28]. Busca hypothesized that the determinant factor to achieve ERK function is not isoform specificity but the global ERK quantity[75] which might have been the determinant factor in my experiments described here. NRAS^{G12D} and BCR-ABL1-driven pre-B ALL mouse cells have stronger signalling through the ERK pathway than normal pre-B cells so loss of ERK2 in these cells had no effect as the overall global ERK quanitity was strong enough to achieve ERK function. However, in normal pre-B cells, loss of ERK2 proved to be detrimental as global ERK quantity was reduced. Despite the vast amount of studies performed on ERK signalling, the functional differences between ERK1 and ERK2 remain controversial[75].

In addition to DUSP6 as a potential therapeutic target in pre-B ALL, I also propose that pharmacological inhibition of BCL6 may provide another therapeutic option in RAS-driven ALL. BCL6 was identified to cause drug resistance in Ph$^+$ ALL and genetic ablation of BCL6 impaired leukaemogenesis[62]. In RAS-driven ALL, activation of ERK upregulated BCL6 expression at the protein and mRNA level. Interestingly, *NRAS*^{G12D} signalling inhibited the negative regulator of BCL6, STAT5 by activating the inhibitory protein tyrosine phosphatase PTPN6. Notably, BCL6 deletion in *NRAS*^{G12D} expressing pre-B cells suppressed the development of leukaemia in immunodeficient mice whereas expression of NRAS^{G12D} in *Bcl6*$^{+/+}$ mice resulted in fatal leukaemia (unpublished data, Muschen laboratory). Interestingly, whilst I showed that *KRAS*^{G12V} mutant LAX7R cells were sensitive to pharmacological inhibition by DUSP6, these cells also showed sensitivity towards a BCL6 inhibitor RI-BPI (unpublished data, Muschen laboratory). These findings suggest BCL6 to be a potential therapeutic target in RAS-driven pre-B ALL. In contrast, during normal germinal centre formation, activation of MAPK signalling results in BCL6 phosphorylation, which in turn leads to its rapid degradation by the ubiquitin/proteasome pathway[76]. Given that germinal centre derived lymphomas express the BCL6 protein[77], the MAPK-mediated degradation of BCL6 might be blocked by genetic alterations suggesting a potential relevance for treatment of B cell lymphoma[76].

In conclusion, I have demonstrated the importance of the MAPK pathway in pre-B

ALL and exploited its negative feedback regulators to treat pre-B ALL. I identified ERK negative feedback control via the Dual Specificity Phosphatase 6 (DUSP6) as well as BCL6 to be crucial for NRAS^{G12D}-mediated pre-B cell transformation, suggesting their potential as therapeutic targets. I also discovered that a high level of ERK activity is required for proliferation of normal pre-B cells, but dispensable in leukaemic pre-B ALL cells. In addition, I found that human B cell malignancies can be grouped into three categories that fundamentally differ in their ability to control ERK signalling strength. These are novel findings shedding light on the functional role of MAPK signalling and its negative regulators in pre-B ALL.

Chapter 2 Synthetic lymphomas as a tool for the functional prioritisation of driver genes

1 Abstract

Diffuse Large B Cell Lymphoma (DLBCL) is the commonest form of non-Hodgkin lymphoma. Although potentially curable with chemotherapy, 40% of patients still succumb to their disease. Recent exome sequencing studies have identified hundreds of recurrently mutated genes. However, for most, their contribution to disease, or their importance as therapeutic targets, remains uncertain. I optimised a novel approach for the functional screening of these mutant genes. This was achieved by reconstituting non-malignant, primary, human germinal centre B cells (GC B cells) with combinations of wildtype and mutant genes to recapitulate the genetic events of DLBCL. When injected into immunodeficient mice, these oncogene-transduced GC B cells gave rise to tumours that closely resemble human DLBCL, reinforcing the biological relevance of this system. To screen potential tumour suppressor mutations in this system in a high throughput fashion, I developed a lymphoma-focused CRISPR library of 692 genes recurrently altered in B cell lymphomas. These experiments identified GNA13 as an unexpectedly potent tumour suppressor in human GC B cells and provided new understanding to its mechanism of action.

2 Introduction

2.1 Diffuse Large B Cell Lymphoma (DLBCL)

Diffuse large B cell lymphoma (DLBCL) is the commonest form of non-Hodgkin lymphoma[78]. Although potentially curable with immunochemotherapy (R-CHOP; rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone), 40% of patients succumb to their disease, many within a year of diagnosis[79]. So far, targeted small molecule inhibitors against distinct molecular pathways have shown disappointing results in clinic. This is mainly due to the high genetic heterogeneity of DLBCL and the lack of understanding of how many of these mutations contribute functionally to the disease[3,80-82].

2.2 Germinal Centre Reaction

The cell of origin of DLBCL is the germinal centre B cell (GC B cell). Germinal centres are transient structures that form in secondary lymphoid organs once a naïve B cell encounters a T-cell dependent antigen in the context of T cell help[12,83]. An outline of the germinal centre reaction can be seen in Figure 1. Antigen-activated B cells differentiate into centroblasts that undergo rapid proliferation and somatic hypermutation (SHM) in the dark zone[12,83]. During the process of SHM, point mutations are introduced into the variable region (IgV) of the heavy and light chain resulting in a change in the amino-acid sequence and altered affinity for an antigen[12,83]. Centroblasts then migrate to the light zone as centrocytes where they compete for survival signals based upon affinity of the (mutant) antigen receptor[83]. Centrocytes may also undergo class-switch recombination (CSR)[12,83]. Successfully selected centrocytes eventually differentiate into antibody-secreting plasma cells and memory B cells that protect against invading pathogens. This process is very efficient as both cell types appear within one week of antigen encounter[83]. Centroblasts and centrocytes are able to cycle between dark and light zone through a process guided by chemokine gradients[12].

Survival signals are provided to centrocytes by follicular dendritic cells (FDC) and co-stimulatory B cell surface receptors like the tumour-necrosis factor (TNF)-receptor family member CD40 and its ligand CD154 expressed by T follicular helper (Tfh) cells[12]. FDCs capture and retain unprocessed antigen as intact Ag-Ab complexes through Fc and complement receptors for antigen presentation to a GC B cell[84-86].

FDCs are intimately associated with GC B cells providing a microenvironment stimulating GC B cell proliferation, activation and viability[84]. They do so by expressing a large number of cofactors such as adhesion molecules (ICAM-1 and VCAM-1) that enhance cell-cell contact, anti-apoptotic molecules (BAFF/BLys) that prevent apoptosis and proliferation stimuli (8D6,IL-15,IL-6)[84-86]. It is suggested that FDCs express many more molecules of importance not yet identified[84]. T follicular helper (Tfh) cells also produce an array of cytokines such as IL-4, IL-10 and IL-21[87]. These cytokines play an important role in GC B cell proliferation, differentiation and antibody class switch[87].
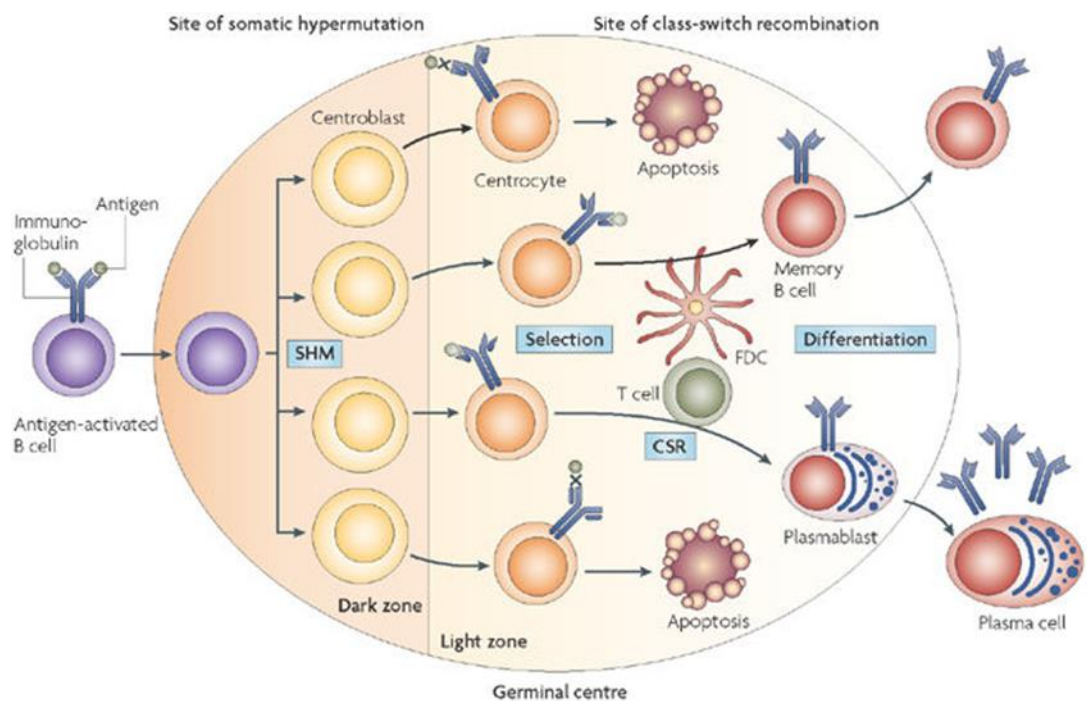
Figure 1 Outline of the germinal centre reaction *Once a B cell encounters an antigen, it moves to the dark zone where it undergoes rapid proliferation and somatic hypermutation. Selection for the BCR with the highest affinity occurs in the light zone before GC B cells differentiate into memory B cells or plasma cells.[12]*

The GC reaction is orchestrated by a large network of transcriptional modulators and its initiation is dependent on a transitory peak in NF-κB, IRF4 and MYC, which are later transcriptionally suppressed in the dark zone[11] (Figure 2). MYC is then temporarily re-expressed in those light zone B cells that are primed to cycle back to

the dark zone[3,88]. During GC initiation, IRF8, IRF4 (through activation by NF-κB) and MEF2B all contribute to expression of BCL6, considered the master regulator of the GC reaction[3]. After the initial peak of MYC expression, BCL6 then transcriptionally represses MYC and multiple other cellular processes that affect many signalling pathways. For instance, BCL6 downregulates the response to DNA damage, prevents premature B cell activation and blocks terminal differentiation[3]. The importance of BCL6 in the GC reaction can be seen in BCL6-deficient mice where GC formation is completely absent[3]. Conversely, translocation of BCL6 leading to constitutive expression results in lymphomagenesis in both humans and mice[3]. Paired box 5 (PAX5) is expressed throughout the entire germinal centre reaction and is only downregulated in cells committed to plasma cell differentiation by IRF4, BLIMP1 and XBP1 expression[3]. NF-κB (which activates IRF4) contributes to GC initiation by BCL6 induction; however, at the GC exit stage this pathway promotes plasma cell differentiation by downregulating BCL6[3]. This in turn releases the repression on *PRDM1* (encoding BLIMP1). BLIMP1 is required for plasma cell differentiation as it inhibits cell proliferation and also makes the switch to a plasma cell-specific transcriptional programme including the activation of XBP1 by repressing PAX5[3].

The importance of the germinal centre reaction can be seen in patients that are unable to form GCs because of inherited immunodeficiency syndromes such as Hyper-IgM syndrome.
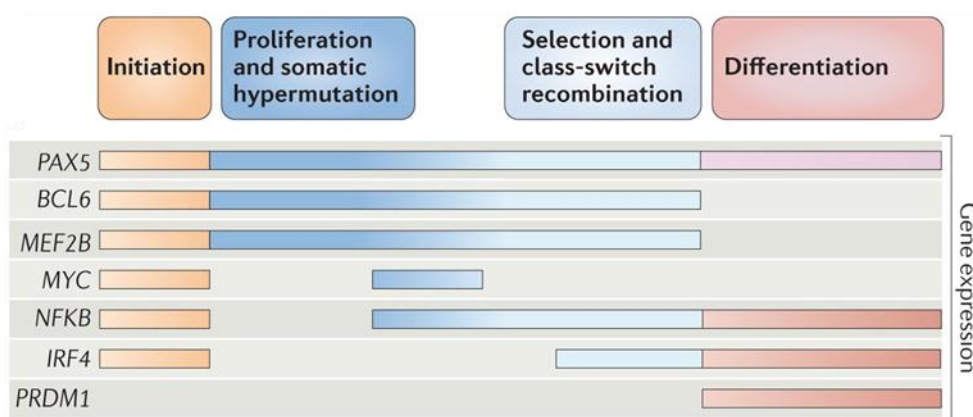
Figure 2 Network of transcriptional modulators in the GC reaction (modified from[3]).

2.3 Mechanisms of genetic lesions in DLBCL

The pathways driving tumourigenesis are often hijacked by genetic aberrations including amplifications, deletions and point mutations leading to gain-of-function mutations that activate signalling effectors and loss-of-function mutations that inactivate negative feedback signalling[89]. Interestingly, DLBCL is also characterized by two further genetic alterations; chromosomal translocations and aberrant SHMs which arise from errors in V(D)J recombination, somatic hypermutation (SHM) and class-switch recombination (CSR)[90]. Those remodelling mechanisms are often hazardous because they take place in an environment where B cells proliferate rapidly and DNA damage checkpoints are silenced by BCL6[3,12]. Chromosomal translocations in DLBCL typically involve recombination between the immunoglobulin loci and a proto-oncogene locus commonly affecting genes that regulate GC B cell development[3,12]. This is notably different from acute leukaemias where chromosomal translocations result in fusion genes[12]. Common chromosomal translocations are the t(14;18) translocation between the immunoglobulin heavy chain locus (IgH) and *BCL2* loci or the t(8;14) translocation between the immunoglobulin and *MYC* loci[3,12].

Aberrant SHM, another hallmark of DLBCL, is the aberrant process of SHM on non-immunoglobulin loci[3,12]. It causes dysregulated expression of proto-oncogenes (for example *MYC* and *PIM1*) by introducing somatic mutations in the 5' regulatory region affecting the correct function of the corresponding gene[91]. BCL6 is often dysregulated through translocation of the intact protein coding sequence to regulatory regions such as promoters by a partner chromosome or as the result of aSHM that prevents the silencing of BCL6, resulting in a block in GC differentiation[92].

2.4 Evolving classification of DLBCL

In the past, gene expression profiling had identified different subgroups of DLBCL which research groups have used to classify DLBCL ever since; the activated B-cell-like (ABC), germinal-centre-B-cell-like (GCB) subgroup and unclassified[93]. ABC and GCB DLBCL are made up of different transcriptional modules consisting of hundreds of genes[93]. The GCB subgroup expresses genes characteristic of germinal centre B cells whereas the ABC subgroup corresponds to a slightly later stage of GC

development and is characterized by chronic active signalling through the B cell receptor[94]. The groups differed in overall survival after R-CHOP with GCB DLBCL having a better prognosis than ABC DLBCL[94]. The two subtypes also showed a differential response to targeted agents such as ibrutinib which is preferentially effective in patients with relapsed or refractory ABC DLBCL[94].

With the advent of next-generation sequencing (NGS), studies have revealed a remarkable high molecular and clinical heterogeneity in DLBCL. Many recurrent genetic alterations are present at low frequency and there is evidence for the existence of at least five molecular subtypes[80-82]. Previously, this high heterogeneity made it difficult to identify these low-frequency mutations due to sample-size limitations[95]. Pan-cancer analysis suggested that to achieve 90% power, 400 tumour-normal DLBCL pairs are necessary to identify those genes mutated at a frequency of 5% or higher[95]. To better understand DLBCL biology, three groups have recently published large genomic studies which together converge on a few hundred genes recurrently mutated genes in DLBCL[6,7,96].

Reddy performed whole-exome and transcriptome sequencing on 1001 DLBCL patients that were treated with rituximab-containing regimens and including 400 paired germline DNAs[7]. They identified 150 genetic drivers of the disease (Figure 3) and sought to describe their clinical significance[7]. The most frequently mutated gene is *KMT2D* (*MLL2)*, followed by *BCL2* and *MYD88*[7]. The sample size provided >90% power to identify mutations at a frequency of 5%[7]. Notably, 27 out of the 150 driver genes (for example *SPEN, KLHL14, MGA*) were not previously associated with DLBCL[7]. Using an unbiased CRISPR screen in cell lines, Reddy identified that the knockout of 35 driver genes (for example *EBF1, IRF4, CARD11* in ABC DLBCL and *XPO1, PTPN6, ZBTB7A* in GCB DLBCL) resulted in a growth disadvantage, classifying them as oncogenes[7]. Caution must be exercised in the interpretation of these screens as several genes that could be considered "positive controls" for these experiments behaved aberrantly. An example is BTK, which is well established to be essential for the survival of ABC cell lines but not in those with mutant CARD11 such as OCI-Ly3[97]. Completely opposite results are presented in the Reddy[7] paper and the explanation for this is not clear. For 9 out of the 35 genes, targeted drugs are available which DLBCL patients that harbour mutations in one of these 9 genes (36%) might benefit from[7]. Existing drug targets include genes against *MTOR, BCL2*

or *SYK*[7]. However, it seems over-simplistic to infer sensitivity to a particular targeted agent purely from the presence of mutation within the target gene.

A genomic risk model provided information on the link between combinations of genetic alterations and clinical outcome[7]. For instance, *MYC* genetic alterations with a high MYC expression are associated with the least favourable prognosis whereas GCB DLBCL with either *CD70* or *MYD88* alterations seemed to have the most favourable prognosis[7].
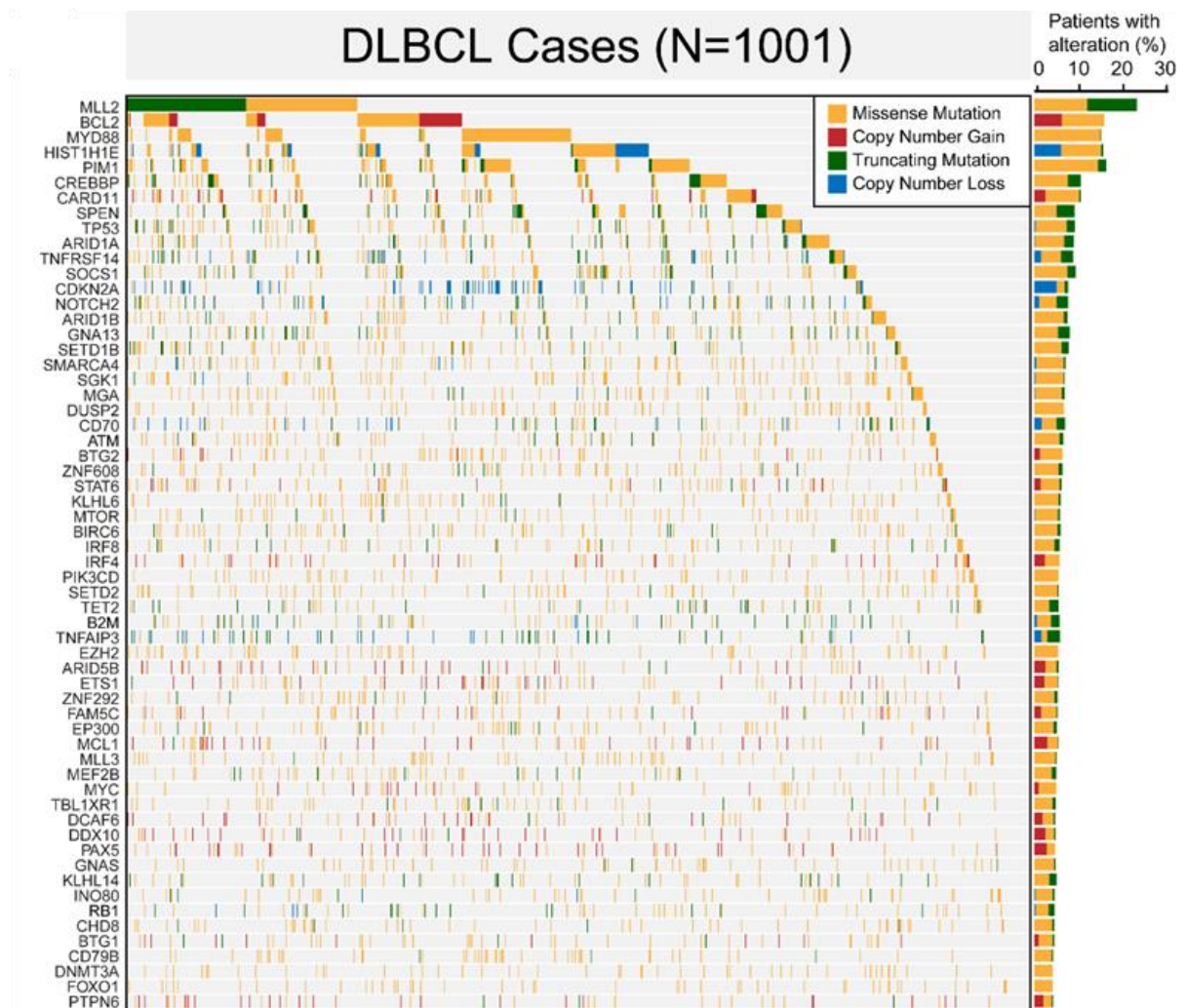
Figure 3 Recurrently altered genes in 1001 DLBCL cases *This heatmap shows the top 60 out of 150 most frequently mutated genes (frequency >5%). Missense mutations are highlighted in yellow, copy number gain in red, truncating mutation in green and copy number loss in blue. The bargraph on the right displays the percentage of gene-level alteration.*[7]

Another informative sequencing study by Chapuy and colleagues was published soon after[96]. This study carried-out whole-exome sequencing on 304 primary DLBCL from newly diagnosed patients (85% having been treated with R-CHOP)[96]. 55% of primary DLBCL samples lacked patient-matched normal samples so a computational approach was used to filter germline variants and artefacts from tumour-only samples[96]. Chapuy identified 40 previously undescribed genetic alterations such as *LYN, HVCN1* and several histone genes[96]. Taking into account co-occurring alterations, this study identified 5 robust clusters with discrete genetic alterations and a further subset (C0) without any detectable genetic alterations[96]. Cluster 1 is characterized by genetic alterations in *BCL6* and *NOTCH2* pathway components as well as *BCL10, TNFAIP3(A20)*, and *FAS*[96]. It also has multiple immune evasion strategies such as mutations in *CD70* and *B2M*[96]. Transcriptional profiling suggests that cluster 1 is mainly ABC-type tumours. These mutations are also often found in marginal zone lymphoma (MZL) suggesting cluster 1 represents a transformed MZL precursor[96]. Cluster 2 presents with inactivation of tumour suppressor genes maintaining chromosome stability such as frequent biallelic inactivation of *TP53* and 17p copy loss[96]. Somatic copy number alterations were observed in cell cycle modifiers like 9p21.13/*CDKN2A* and 13q14.2/*RB1*[96]. This cluster can be associated with both GCB and ABC DLBCL[96]. Cluster 3 mainly has *BCL2* mutations with *BCL2/IgH* rearrangements, mutations in chromatin modifiers such as *KMT2D, CREBBP*, and *EZH2* and also *PTEN* alterations; making it more GCB-like[96]. Cluster 4 is also GCB-like and harbours mutations in several key signalling pathways, such as BCR/PI3K, JAK/STAT, and BRAF pathway components[96]. Genetic alterations were also found in multiple histones and chromatin-modifying enzymes[96]. Clusters 3 and 4 perturb chromatin by distinct mechanisms and both harbour unique targetable genetic alterations. This suggests that they are genetically distinct and might respond differently to targeted therapies[96]. Cluster 5, which is mainly ABC-like, has defining mutations in *CD79B, MYD88, TBL1XR1, PIM1, BTG1*, and *ETV6* and has a distinct genetic mechanism to the ABC-like Cluster 1[96]. Cluster 1 and 5 harbour different types and incidence of MYD88 mutations. Interestingly, cluster 1 has low or absent canonical AID (cAID) activity whereas cluster 5 shows increased cAID activity[96]. Chapuy attributed significant differences in progression-free survival (PFS) and overall survival (OS) across the different clusters[96]. Clusters C0, C1 and C4 DLBCL had more favourable outcomes than C3 and C5[96].

Another study by Schmitz and colleagues identified genetic subtypes of DLBCL that differ from each other by recurrent genetic aberrations[6]. Schmitz performed exome and transcriptome sequencing on 574 DLBCL biopsy samples and deep amplicon sequencing on 372 genes[6]. Based upon genetic abnormalities, four molecular subtypes were identified, termed MCD, BN2, N1 and EZB (Figure 4) that seem to have different outcomes after immunochemotherapy.
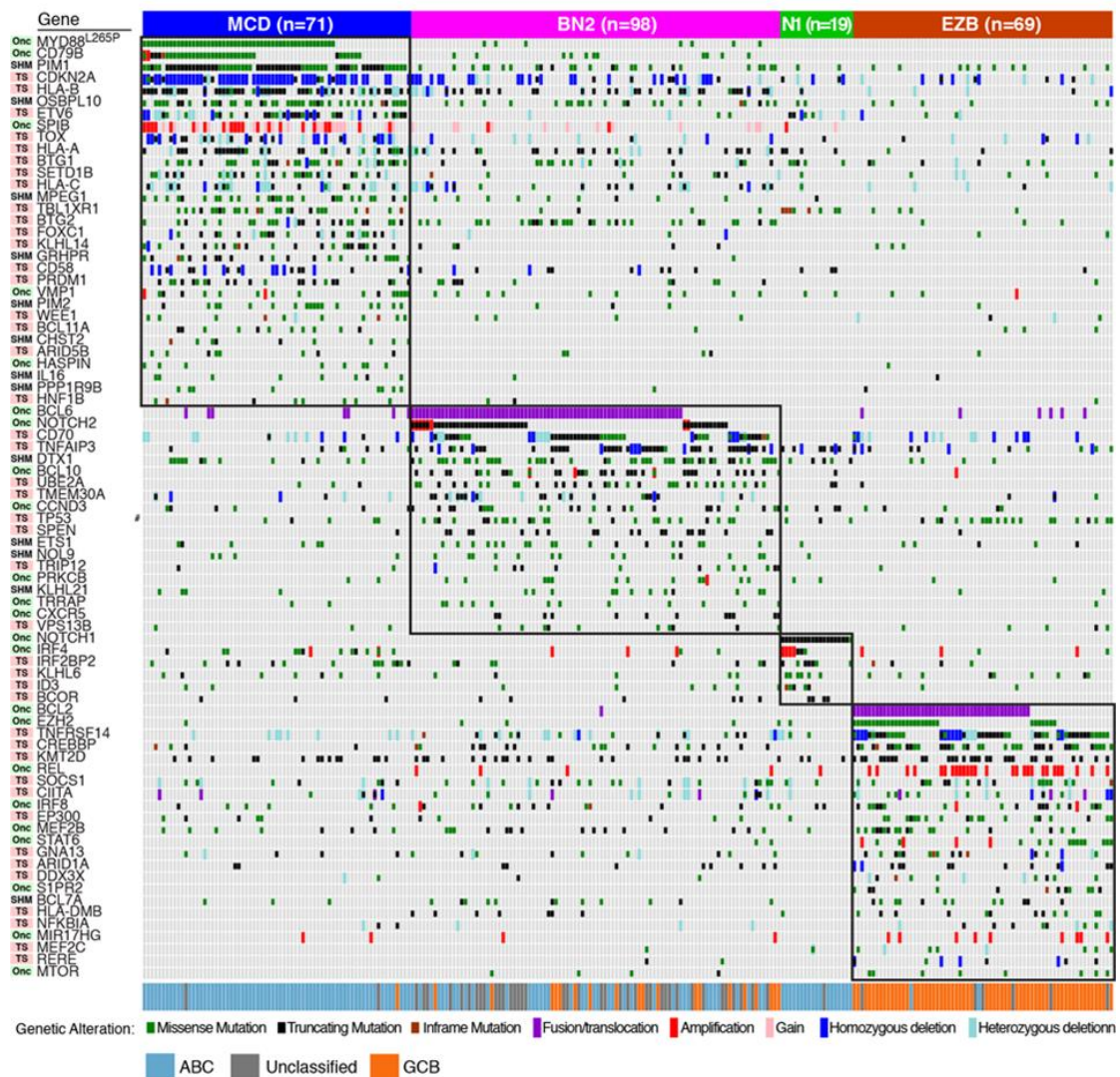
Figure 4 Exome sequencing identifies four new genetic subtypes in DLBCL
Subtype MCD is based on the co-occurrence of MYD88 and CD79 mutations, BN2 is based on BCL6 fusions and NOTCH2 mutations, N1 is characterized by NOTCH1 mutations and EZB is based on EZH2 mutations and BCL2 translocations. (modified from[6])

MCD has a co-occurrence of genetic mutations previously associated with ABC DLBCL; $MYD88^{L265P}$ and $CD79B$[6]. This group is similar to Cluster 5 in Shipp's study. The BN2 subgroup has $BCL6$ fusions and $NOTCH2$ mutations[6]. These are enriched for cases unclassified by gene expression and when compared to Shipp's group shows similarities to Cluster 1. Interestingly, the third subgroup N1 shares a functional similarity with BN2; it has a $NOTCH1$ mutation but yet they differ genetically, phenotypically and clinically[6]. The last subgroup EZB is enriched for $EZH2$ mutations and $BCL2$ translocations that are usually seen in GCB DLBCL and can be attributed to Shipp's Cluster 3. Both BN2 and EZB subgroups have a more favourable outcome after immunochemotherapy than MCD and N1 subgroups[6]. However, Chapuy reported that cluster 3, which correlates to the EZB subgroup, had a less favourable outcome[96].

All three studies suggest the possibility of precision-medicine in DLBCL, with the use of targeted agents in the context of a particular molecular subtype. However, caution is advised due to significant differences between these sequencing studies. Out of the few hundred genetic alterations these studies converged on, there were many mutations that were not shared between them and the same genetic alterations such as $BCL2$ mutations and chromatin modifier mutations (Cluster 3 and EZB subtype) were correlated with different outcomes between two studies. Moreover, the four subgroups suggested by Schmitz are derived from a pre-defined grouping where they used a set of "seed classes" and moved patient cases into and out of the 4 defined subtypes using an algorithm-based approach[6]. Overall, only approximately 50% of samples are classified into one of the 4 subgroups with the remaining 50% being attributed to other ABC, GCB or other unclassified subtypes. Clusters 2 and 4 from Chapuy's study [96] that do not match with any of the 4 subgroups suggested by Schmitz[6] might fall into these remaining 50%. Furthermore, whilst Chapuy[96] and Schmitz[6] grouped genetic alterations into 4 – 5 subtypes, Reddy[7] divided its identified mutations into 39 subtypes based upon their impact on prognosis. This large number makes it very difficult to compare to the other studies and impractical for use in clinic. It is highly likely that further subtypes will be identified in the future as technology develops. Finally, it is important to point out that these clusters were generated computationally. Although the gene expression profiling and differences in outcome support their existence as separate entities there is currently very little functional,

experimental evidence for co-operation between the various genetic alterations found within each cluster.

2.5 DLBCL pathogenesis

The genome of DLBCL is strikingly complex, with a high intra- and inter-tumour variability[80-82]. Some of these mutations converge on common signalling pathways that are already well-characterized; however, for many mutations the functional mechanism and contribution to lymphomagenesis remains unclear[80-82]. As discussed in 2.4, DLBCL most likely consists of several subgroups; however, in this section I will focus on the main genetic lesions found in ABC and GCB DLBCL to date (Figure 5).

GCB- and ABC-DLBCL	%	GCB-DLBCL	%	ABC-DLBCL	%
BCL6 Tx	20-40	BCL2 Tx/M	34	TNFAIP3 M/D	30
MLL2/MLL3 M	32-38	GNA13 M	25	MYD88 M	30
CREBBP/EP300 M/D	32	EZH2 M	22	CDKN2A/B D	30
B2M/CD58 M/D	21-29	BCL6 BSE1 M	15	BCL2 Amp	24-30
TP53 M	20	MYC Tx	10	PRDM1 M/D	25
MEF2B M	11	miR17-92 G	6-12	CD79A/B M	20
FOXO1 M	8	PTEN D	6-11	CARD11 M	9

■ epigenetic modification ■ proliferation ■ BCL6 deregulation ■ NF-κB/BCR signaling ■ DNA damage response
■ immune escape ■ apoptosis ■ terminal differentiation ■ JAK/STAT signaling ■ cell cycle ■ other

Figure 5 Subtype-specific genetic alterations in DLBCL (modified from [11]).

In GCB- and ABC- DLBCL, epigenetic remodelling is disrupted in around 30% of cases[11]. Genetic inactivation of the acetyltransferase *EP300* and/or *CREBBP* results in impaired acetylation of *BCL6* and *p53* leading to constitutive activation of the *BCL6* proto-oncogene and decreased function of *p53*[98]. Mutations in the histone methyltransferase *KMT2D* are likely to have an effect on chromatin regulation by reprogramming the epigenome of the precursor cancer cell[11,99,100].

Genetic alterations affecting the *BCL6* locus occur in roughly 35% of GCB- and ABC- DLBCL patients[11]. Chromosomal translocation of *BCL6*, as described in 2.3, is known to contribute to lymphomagenesis by suppressing DNA damage responses and blocking GC differentiation[11]. aSHM of *BCL6* in the first nonconding

exon deregulates its activity by disrupting a negative autoregulatory loop that normally controls its own transcription and by preventing IRF4 from binding and subsequent transcriptional repression following CD40 signalling[3]. Moreover, gain of function mutations in *MEF2B*, a transcriptional activator of *BCL6* are seen in around 10% of cases and inactivation of *FBXO11*, which leads to impaired proteasomal-mediated degradation of *BCL6* is seen in around 4% of cases[11]. *TP53* being mutated and deleted in around 20% is also shared across both DLBCL subtypes[11].

Interestingly, 60% of lymphoma cells seem to escape immune surveillance by loss of MHC class I expression, making them invisible to cytotoxic T lymphocytes (CTLs) and natural-killer cells (NK)[89,101]. This is due to inactivation of beta-2-microglobulin (*B2M*) which results in the inability to express HLA class I complex on the surface[89,101]. Inactivating mutations are also found in genes encoding *HLA-A, HLA-B, HLA-C*[101].

In GCB DLBCL, common aberrations are chromosomal translocations of *MYC* and *BCL2* in 35% and 15% of cases, respectively[11]. Gain-of-function mutations in *EZH2* are seen in around 20% of cases and have been shown to contribute to lymphomagenesis by repressing *CDKN1A, PRDM1* and *IRF4*[3]. Additionally, mutations in the GNA13 pathway (~30%) result in impaired apoptosis and loss of confinement to the germinal centre[102]. This leads to dissemination to the lymph and bone marrow thereby promoting lymphoma[102]. Here, loss-of-function mutations in *GNA13, S1PR2, ARHGEF1* and *P2RY8* inactivate the GNA13 pathway[11].

ABC DLBCL is best characterized by the constitutive activation of NF-κB and the block of GC B cell differentiation to plasma cells[11,89]. Around 20% of cases harbour mutations in *CD79A* and/or *CD79B*, which contributes to chronic active BCR signalling by preventing endocytosis of the receptor and/or weakening the negative pathway regulator *LYN*[97]. Activating mutations in *CARD11* (~10%) lead to constitutive NF-κB activation via the CARD-11-BCL-10-MALT1 complex[103]. Constitutive NF-κB activation is also achieved by mutations in *MYD88* which forms a complex with *IRAK1* and *IRAK4*, thereby activating NF-κB as well as the JAK/STAT3 pathway[104]. It is therefore not surprising to find the key negative regulator of the NF-κB pathway, *TNFAIP3* (A20) to be mutated in around 30% of ABC DLBCL cases leading to prolonged NF-κB responses[105]. Pharmacological

inhibition of NF-κB by the BTK inhibitor ibrutinib resulted in cell death of ABC DLBCL cell lines, emphasising the importance of the NF-κB pathway[106]. The inability of ABC DLBCL cells to differentiate into plasma cells is partly due to mutations in *PRDM1*, which are observed in around 30% of cases[11].

Moreover, somatic mutations in *CCND3* are found in approximately 14% of ABC DLBCL and 10% of GCB DLBCL[107]. *CCND3*, encoding Cyclin D3, belongs to the highly conserved cyclin family that regulates the CDK kinases driving cell cycle G1/S transition[107]. Mutations are often seen in the threonine residue at position 283 (T283) which stabilizes the Cyclin D3 protein[107]. It does so by changing the phosphorylation motif at Thr283 which is essential for proteosomal degradation[107].

2.6 Genome editing using CRISPR

The emergence of genome-editing technologies has enabled researchers to precisely manipulate the sequence of the human genome; Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) being the most popular one to date[108]. Over the years there have been several attempts to edit the genome including RNA interference (RNAi)[109], zinc finger nucleases (ZFNs)[110] and transcription activator-like effector nucleases (TALENs)[111]. However, these approaches are costly, time-consuming to engineer, and have DNA binding specificity issues[108]. Therefore, CRISPR has evolved as the primary tool for genome editing, due to its simplicity, adaptability and high efficiency as a site-specific nuclease.

The bacterial CRISPR/CAS9 system derived from *Streptococcus pyogenes* has proven to be highly versatile and effective in introducing double-strand breaks (DSBs) into mammalian cells[108,112-114]. The system requires the introduction of an RNA-guided CRISPR-associated endonuclease (CAS9) and a "guide" RNA (gRNA)[108]. The gRNA is a short synthetic RNA consisting of a ''scaffold'' sequence and a user-defined 20 nucleotide region ("spacer") at the 5' end[108]. The "scaffold" sequence is necessary for CAS9 binding whereas the "spacer" is designed so that it can target virtually any genomic location that is present immediately upstream of a Protospacer Adjacent Motif (PAM)[108]. In fact, even if the sequence is fully complementary, CAS9 does not cut without a PAM sequence[115]. Once the CAS9:gRNA complex has annealed to the target DNA in a 3' to 5' direction, RuvC

and HNH, two functional endonuclease domains in CAS9, cut both DNA strands, generating a double-stranded break (DSB) approximately 3-4 nucleotides upstream of the PAM sequence[108]. These DNA breaks are then repaired by either Homology-directed repair (HDR) or Non-homologous end joining (NHEJ)[108]. HDR allows for gene correction or insertion of precise point mutations if a donor template is supplied[108]. In the absence of a homologous repair template, NHEJ results in insertions and/or deletions (indels) at the DSB site, disrupting the targeted locus[116]. Typically, indels give rise to in-frame amino acid deletions, insertions or frameshift mutations resulting in a loss-of-function mutation in the target gene[116]. The CRISPR/CAS9 system takes advantage of this, which makes it highly efficient and versatile.

2.7 Unmet need in Lymphoma

A major goal in cancer research has been to identify a complete list of mutated cancer genes in the hope to elucidate the molecular basis and find potential therapeutic targets[117].

As discussed, exome, transcriptome and whole genome sequencing has revealed a striking degree of molecular and genetic heterogeneity in DLBCL in the last decade [80-82]. Three large genomic studies[6,7,96] that in total have analysed almost 2000 DLBCL patient samples have gained remarkable insight into genetic drivers and different genetic subtypes that may respond differently to targeted therapy. Although some new oncogenes and tumour-suppressor genes have been identified that occur at high frequency, most mutations are found in only a few percent of patients and for many of these mutant genes there is very little functional understanding as to how they contribute to lymphomagenesis and how they act cooperatively with other mutations. Therefore, this striking degree of genetic heterogeneity in DLBCL makes targeted therapy highly challenging. This is compounded by the fact that there is no ideal way to experimentally model all proposed genetic subtypes. Therefore, generating preclinical model systems to capture this genetic heterogeneity is a current research priority[118].

Current tools available for functional analysis of mutations are limited to lymphoma cell lines and transgenic mice. Cell lines represent the late stages of lymphoma, carry a high mutational burden and have often evolved in culture for many years.

Transgenic mice are costly, time-consuming to develop and oncogenic mechanisms in the mouse do not always accurately reflect those in humans. Thus, the available tools to address the most important questions in DLBCL tumour biology are currently inadequate.

2.8 Aims of this study

I set out to develop a novel co-culture system that could be used to examine the functional importance of mutations and combinations of mutations in lymphomagenesis using a non-malignant, primary, human germinal centre B cell (GC B cell).

3 Materials & Methods

3.1 Materials

3.1.1 Overview of mouse strains used in this study

Mouse strain	Source	Purpose
NSG	Jackson Laboratories	Transplant recipient mice

3.1.2 Retroviral vectors; Constitutive expression

Construct	Overexpression of
MSCV-IRES-huDC2	BCL6-t2A-BCL2
MSCV-IRES-huDC2	MYC-t2A-BCL2
MSCV-CAS9-2A-tagBFP	CAS9
MSCV-IRES-Thy1.1	p53dd
MSCV-IRES-LyT2	CCND3 T283A
MSCV-IRES-vLyT2.	Empty
MSCV-IRES-vLyT2.	BCL6 WT
MSCV	MYC
PBABE–TERT.Hygro	hTERT
MSCV-IRES-Thy1.1	CDK4 R24C
MSCV-IRES-vxy-puro	CD40L
MSCV-IRES-LyT2	IL21

3.1.3 Lentiviral vectors; Constitutive expression

Construct	Overexpression of	Source
pKLV2-U6gRNA-Bbsi-PGK-GFP	Lymphoma library and single gRNAs	Modified from Addgene #67979

3.1.4 Overview of Western blot antibodies used in this study

Antigen	Clone ID	Company
Beta-actin	13E5	Cell Signalling
AKT	9272	Cell Signalling
Phospho-AKT (Ser473)	9271	Cell Signalling
GNA13	EPR5436	Abcam
PTEN	138GG	Cell Signalling

3.1.5 Overview of cell lines studied used in this study

Cell line	Type	Source
HBL-1	ABC DLBCL	Dr. D. Hodson
BJAB	GCB DLBCL	Dr. D. Hodson
U2932	ABC DLBCL	Dr. D. Hodson
FDCs	Tonsillar Follicular Dendritic Cells	Chan-Sik Park

3.1.6 Overview of Flow Cytometry Antibodies studied used in this study

Surface antigen	Clone ID	Company
CD38	HB7	BioLegend
CD20	2H7	BioLegend
CD19	HIB19	BioLegend
CD10	97C5	BioLegend
CD2	RPA-2.10	BioLegend
CD90.1 Thy1.1	OX-7	BioLegend
CD154	24-31	BioLegend
CD8a	53-6.7	BioLegend
CD22	HIB22	BioLegend

3.1.7 Overview of PCR primers used in this study

Primer	Sequence 5' - 3'
Zhang_F	GTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTT GTGGAAAGGACGAAACACC
Zhang_R_modified	ACTTTTTCAAGTTGATAACGGACTAGCCTTATTTAAACTTG CTATGCTGTTTCCAGCATAGCTCTTAAAC
gLibrary-HiSeq_50bp-SE-U1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGG AAAGGACGAAACA
gLibrary-HiSeq_50bp-SE-L1	TCGGCATTCCTGCTGAACCGCTCTTCCGATCTCTAAAGCG CATGCTCCAGAC
iPCRtagseq	AAGAGCGGTTCAGCAGGAATGCCGAGACCGATCTC

U6-Illumina-seq2	TCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
Indexing Adapter PE 1.0	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTAC ACGACGCTCTTCCGATC*T
iPCRtagT1	CAAGCAGAAGACGGCATACGAGATAACGTGATGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT2	CAAGCAGAAGACGGCATACGAGATAAACATCGGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT3	CAAGCAGAAGACGGCATACGAGATATGCCTAAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT4	CAAGCAGAAGACGGCATACGAGATAGTGGTCAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT5	CAAGCAGAAGACGGCATACGAGATACCACTGTGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT6	CAAGCAGAAGACGGCATACGAGATACATTGGCGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT7	CAAGCAGAAGACGGCATACGAGATCAGATCTGGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT8	CAAGCAGAAGACGGCATACGAGATCATCAAGTGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT9	CAAGCAGAAGACGGCATACGAGATCGCTGATCGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT10	CAAGCAGAAGACGGCATACGAGATACAAGCTAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT11	CAAGCAGAAGACGGCATACGAGATCTGTAGCCGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT12	CAAGCAGAAGACGGCATACGAGATAGTACAAGGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT13	CAAGCAGAAGACGGCATACGAGATAACAACCAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT14	CAAGCAGAAGACGGCATACGAGATAACCGAGAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT15	CAAGCAGAAGACGGCATACGAGATAACGCTTAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT16	CAAGCAGAAGACGGCATACGAGATAAGACGGAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT17	CAAGCAGAAGACGGCATACGAGATAAGGTACAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT18	CAAGCAGAAGACGGCATACGAGATACACAGAAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT19	CAAGCAGAAGACGGCATACGAGATACAGCAGAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT20	CAAGCAGAAGACGGCATACGAGATACCTCCAAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT21	CAAGCAGAAGACGGCATACGAGATACGCTCGAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT22	CAAGCAGAAGACGGCATACGAGATACGTATCAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT23	CAAGCAGAAGACGGCATACGAGATACTATGCAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT24	CAAGCAGAAGACGGCATACGAGATAGAGTCAAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT25	CAAGCAGAAGACGGCATACGAGATAGATCGCAGAGATCG

	GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT26	CAAGCAGAAGACGGCATACGAGATAGCAGGAAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT27	CAAGCAGAAGACGGCATACGAGATAGTCACTAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT28	CAAGCAGAAGACGGCATACGAGATATCCTGTAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT29	CAAGCAGAAGACGGCATACGAGATATTGAGGAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T
iPCRtagT56	CAAGCAGAAGACGGCATACGAGATGTACGCAAGAGATCG GTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T

3.1.8 Overview of qRT-PCR Primers used in this study

Primer	Sequence
EBV F QP1L	5' – GCCGGTGTGTTCGTATATGG – 3'
EBV R QP2L	5' – CAAAACCTCAGCAAATATATGAG – 3'

3.1.9 G Blocks used for GaLV_WT_RC, GaLV_MTR_RC and GaLV_TR_RC envelope constructs

G Block	Sequence 5' - 3'
5' GaLV_WT	CTAGAGTCGA CCTGCAGGAT ATCGAATCCA CCATGGTATT GCTGCCTGGG TCCATGCTTC TCACCTCAAA CCTGCACCAC CTTCGGCACC AGATGAGTCC TGGGAGCTGG AAAAGACTGA TCATCCTCTT AAGCTGCGTA TTCGGCGGCGGCGGGACGAG TCTGCAAAAT AAGAACCCCC ACCAGCCCAT GACCCTCACT TGGCAGGTAC TGTCCCAAAC TGGAGACGTT GTCTGGGATA CAAAGGCAGT CCAGCCCCCT TGGACTTGGT GGCCCACACT TAAACCTGAT GTATGTGCCT TGGCGGCTAG TCTTGAGTCC TGGGATATCC CGGGAACCGA TGTCTCGTCC TCTAAACGAG TCAGACCTCC GGACTCAGAC TATACTGCCG CTTATAAGCA AATCACCTGG GGAGCCATAG GGTGCAGCTA CCCTCGGGCT AGGACTAGAA TGGCAAGCTC TACCTTCTAC GTATGTCCCC GGGATGGCCG GACCCTTTCA GAAGCTAGAA GGTGCGGGGG GCTAGAATCC CTATACTGTA AAGAATGGGA TTGTGAGACC ACGGGGACCG GTTATTGGCT ATCTAAATCC TCAAAGACC TCATAACTGT AAAATGGGAC CAAAATAGCG AATGGACTCA AAAATTTCAA CAGTGTCACC AGACCGGCTG GTGTAACCCC CTTAAAATAG ATTTCACAGA CAAAGGAAAA TTATCCAAGG ACTGGATAAC GGGAAAAACC TGGGGATTAA GATTCTATGT GTCTGGACAT CCAGGCGTAC AGTTCACCAT TCGCTTAAAA ATCACCAACA TGCCAGCTGT GGCAGTAGGT CCTGACCTCG TCCTTGTGGA ACAAGGACCT CCTAGAACGT CCCTCGCTCT CCCACCTCCT CTTCCCCCAA GGGAAGCGCC ACCGCCATCT CTCCCCGACT CTAACTCCAC AGCCCTGGCG ACTAGTGCAC AAACTCCCAC GGTGAGAAAA ACAATTGTTA CCCTAAACAC TCCGCCTCCC ACCACAGGCG ACAGACTTTT TGATCTTGTG CAGGGGGCCT TCCTAACCTT AAATGCTACC AACCCAGGGG CCACTGAGTC TTGCTGGCTT TGTTTGGCCA TGGGCCCCCC

	TTATTATGAA GCAATAGCCT CATCAGGAGA GGTCGCCTAC TCCACCGACC TTGACCGGTG CCGCTGGGGG ACCCAAGGAA AGCTCACCCT CACTGAGGTC TCAGGACACG GGTTGTGCAT AGGAAAGGTG CCCTTTACCC ATCAGCATCT CTGCAATCAG ACCCTATCCA TCAATTCCTC CGGAGACCAT CAGTATCTGC TCCCCTCCAA CCATAGCTGG TGGGCTTGCA GCACTGGCCT CACCCCTTGC CTCTCCACCT CAGTTTTTAA TCAGACTAGA GATTTCTGTA TCCAGGTCCA GCTGATTCCT CGCATCTATT ACTATCCTGA AGAAGTTTTG TTACAGGCCT ATGACAATTC TCACCCCAGG ACTAAAAGAG AGGCTGTCTC ACTTACCCTA GCTGTTTTAC TGGGGTTGGG AATCACGGCG GGAATAGGTA CTGGTTCAAC TGCCTTAATT AAAGG
3' GaLV_WT	GGGAATAGGT ACTGGTTCAA CTGCCTTAAT TAAAGGACCT ATAGACCTCC AGCAAGGCCT GACAAGCCTC CAGATCGCCA TAGATGCTGA CCTCCGGGCC CTCCAAGACT CAGTCAGCAA GTTAGAGGAC TCACTGACTT CCCTGTCCGA GGTAGTGCTC CAAAATAGGA GAGGCCTTGA CTTGCTGTTT CTAAAAGAAG GTGGCCTCTG TGCGGCCCTA AAGGAAGAGT GCTGTTTTTA CATAGACCAC TCAGGTGCAG TACGGGACTC CATGAAAAAA CTCAAAGAAA AACTGGATAA AAGACAGTTA GAGCGCCAGA AAAGCCAAAA CTGGTATGAA GGATGGTTCA ATAACTCCCC TTGGTTCACT ACCCTGCTAT CAACCATCGC TGGGCCCCTA TTACTCCTCC TTCTGTTGCT CATCCTCGGG CCATGCATCA TCAATAAGTT AGTTCAATTC ATCAATGATA GGATAAGTGC AGTTAAAATT CTGGTCCTTA GACAGAAATA TCAGGCCCTA GAGAACGAAG GTAACCTTTA AGAATTCATT GATCATAATC AGCCATACCA C
3' GaLV_MTR	GGGAATAGGTACTGGTTCAACTGCCTTAATTAAAGGACCTATAGAC CTCCAGCAAGGCCTGACAAGCCTCCAGATCGCCATAGATGCTGAC CTCCGGGCCCTCCAAGACTCAGTCAGCAAGTTAGAGGACTCACTG ACTTCCCTGTCCGAGGTAGTGCTCCAAAATAGGAGAGGCCTTGAC TTGCTGTTTCTAAAAGAAGGTGGCCTCTGTGCGGCCCTAAAGGAA GAGTGCTGTTTTTACATAGACCACTCAGGTGCAGTACGGGACTCC ATGAAAAAACTCAAAGAAAAACTGGATAAAAGACAGTTAGAGCGC CAGAAAAGCCAAAACTGGTATGAAGGATGGTTCAATAACTCCCCTT GGTTTACCACCTTGATATCTACCATTATGGGACCCCTCATTGTACT CCTATTGATTTTGCTCTTCGGACCCTGCATTCTTAATCGATTAGTC CAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTTGA CTCAACAATATCACCAGCTGAAGCCTATAGAGTACGAGCCATAGG AATTCATTGATCATAATCAGCCATACCAC
3' GaLV_TR	GGGAATAGGT ACTGGTTCAA CTGCCTTAAT TAAAGGACCT ATAGACCTCC AGCAAGGCCT GACAAGCCTC CAGATCGCCA TAGATGCTGA CCTCCGGGCC CTCCAAGACT CAGTCAGCAA GTTAGAGGAC TCACTGACTT CCCTGTCCGA GGTAGTGCTC CAAAATAGGA GAGGCCTTGA CTTGCTGTTT CTAAAAGAAG GTGGCCTCTG TGCGGCCCTA AAGGAAGAGT GCTGTTTTTA CATAGACCAC TCAGGTGCAG TACGGGACTC CATGAAAAAA CTCAAAGAAA AACTGGATAA AAGACAGTTA GAGCGCCAGA AAAGCCAAAA CTGGTATGAA GGATGGTTCA ATAACTCCCC TTGGTTCACT ACCCTGCTAT CAACCATCGC TGGGCCCCTA TTACTCCTCC TTCTGTTGCT CATCCTCGGG CCATGCATCA TCAATCGATT AGTCCAATTT GTTAAAGACA GGATATCAGT GGTCCAGGCT CTAGTTTTGA CTCAACAATA TCACCAGCTG AAGCCTATAG AGTACGAGCC ATAGGAATTC ATTGATCATA ATCAGCCATA CCAC

3.2 Methods

3.2.1 Cell lines and human primary GC B cells

Cell lines were cultured in Roswell Park Memorial Institute medium (RPMI-1640, Invitrogen, Carlsbad, CA) and primary human GC B cells were cultured in Advanced Roswell Park Memorial Institute medium (Advanced RPMI-1640, Invitrogen, Carlsbad, CA) with GlutaMAX containing 20% FBS, 100 IU/ml penicillin and 100 µg/ml streptomycin and kept at 37°C in a humidified incubator (5% CO_2 and 95% atmosphere). All cell lines used in this study are negative for mycoplasma contamination.

3.2.2 Isolation of germinal centre B cells from human tonsil tissue

Fresh, tonsil tissue was sourced from the Addenbrooke's ENT Department and processed directly to preserve viability. Ethical approval for the use of human tissue has been granted by the Health Research Authority Cambridgeshire Research Ethics Committee (REC no. 07/MRE05/44). Germinal centre B cells were purified using the human B cell negative selection isolation Kit II (MACS, Miltenyi Biotec). This was modified to include negative selection antibodies IgD-BIOT (SouthernBiotech) and CD44-BIOT (SouthernBiotech). Cells were stained for CD38, CD20, CD19 and CD10 and considered germinal centre B cells when positive (>90%) for all four cell surface markers. *See 3.1.6 for a list of Flow Cytometry Antibodies used in this study.*

3.2.3 EBV screen of isolated GC B Cells using quantitative real-time PCR

GC B cells purified from tonsils were tested for EBV status using a quantitative real-time PCR (qPCR) assay[119]. Serial dilutions of EBV positive genomic DNA (Namalwa) and genomic DNA from an EBV negative cell line (239FT) were used. qPCR was set up using FAST SYBR Green Master Mix (ThermoFisher) with 4 replicates per sample and run on a Vii7 real-time PCR system (Applied Biosystems). The following cycling conditions were used: 95 °C for 20 s, 40 cycles of 95 °C for 1 s, 60 °C for 20 s and the final extension, 60 °C for 1 min. *See 3.1.8 for a list of qPCR primers used in this study.*

3.2.4 Cloning of alternative retroviral and lentiviral envelope constructs used for transducing primary human GC B cells

pHIT123 (provided by D.B. Kohn, University of California, Los Angeles, Los Angeles, CA) containing the retroviral ecotropic envelope, human cytomegalovirus immediate-early promoter and the origin of replication from simian virus 40 was used

as the backbone. To produce the retroviral envelope construct GaLV_WT_RC, the vector backbone (pHIT123) was digested with EcoRI (NEB) at 37°C for 1 hour and PCR-purified (Life Technologies, Qiagen). A 20 µl Gibson ligation reaction (NEB) was performed using 50 ng of the PCR-purified vector (pHIT123) and 30 ng of the 5' GaLV_WT (1585 bp) insert and 10 ng of 3' GaLV_WT (571 bp) to produce the retroviral envelope construct. 5' GaLV_WT and 3' GaLV_WT are the original GaLV envelope sequence taken from NCBI. After incubation for 1 hour at 50°C, 1 µl of the ligated reaction was transformed into 25 µl of electrocompetent cells (C3040, NEB) according to the manufacturer's protocol. G-Blocks were ordered from eurofins.

Next, GaLV_WT_RC was digested with EcoRI-HF (NEB) and PacI (NEB) to cut out the 3' GaLV WT insert and purified using the PCR purification kit (QIAGEN). A 20 µl Gibson ligation reaction (NEB) was performed using 50 ng of digested GaLV_WT_RC and 10 ng of 3' GaLV_MTR (574 bp) or 10 ng of 3' GaLV_TR (574 bp). 3' GaLV_MTR contains the 3' GaLV envelope sequence replaced by the MuLV transmembrane region, cytoplasmic region and R peptide region[120]. 3' GaLV_TR contains the 3' GaLV envelope sequence replaced by the MuLV cytoplasmic region and R peptide region[120]. After incubation for 1 hour at 50°C, 1 µl of the ligated reaction was transformed into 25 ul of electrocompetent cells (C3040, NEB) according to the manufacturer's protocol. GaLV_MTR_RC and GaLV_TR_RC were verified by Sanger sequencing. 3' GaLV_MTR G Block was ordered from IDT and GaLV_TR_RC from eurofins. *See 3.1.9 for a list of G Block sequences used in this study.*

3.2.5 Retroviral and lentiviral production for infecting human primary GC B cells

Using 1 ml of Opti-MEM media (Invitrogen) and 18 µl of TransIT-293 (Mirus), 293FT cells plated in a 10 cm^2 dish were co-transfected with 1 µg pHIT60 (gag-pol), 1 µg GALV WT (envelope, built by Rebecca Caeser) and 4 µg of a retroviral construct. Using 1 ml of Opti-MEM media (Invitrogen) and 33 µl of TransIT-293 (Mirus), 8.3 µg pCMVDeltaR8.91 (gag-pol) (kindly provided by Dan Webster, NCI), 2.8 µg GALV MTR (envelope, built by Rebecca Caeser) and 11 µg of a lentiviral construct were used for a 10 cm^2 dish. Dulbecco's modified Eagle's medium (DMEM, Invitrogen) with Glutamax containing 10% fetal bovine serum, 100 IU/ml penicillin, 100 µg/ml streptomycin were used for cultivation. After 48h, the virus supernatant

was filtered through a 0.45 µM filter. If needed, media was replenished and harvested again at 72 hours.

3.2.6 Retroviral and lentiviral production for infecting cell lines

Virus was produced as in 3.2.5 with 8.3 µg pCMV DeltaR8.91 (gag-pol) and 2.8 µg pUC.MDG (kindly provided by Dan Webster, NCI) encoding the VSVG envelope as well as 11 µg of a lentiviral construct were used for a 10 cm^2 dish. 1 µg pHIT60 (gag-pol) and 1 µg pHITEA3x6 (envelope) as well as 4 µg of a retroviral construct were used for a 10 cm^2 dish.

3.2.7 Retroviral and lentiviral transduction

For retroviral/lentiviral transduction, cells were resuspended with the virus and infected by centrifugation (1500 x g, 90 min at 32°C) with the addition of 10 µg/ml Polybrene (INSIGHT biotechnology) and 25 µM HEPES (ThermoFisher). After the spin, virus supernatant was replaced with fresh media straight away if transducing retroviral constructs or replaced after > 4 hours if transducing lentiviral constructs. Cells were maintained at 37°C with 5% CO2 for 2 - 4 days before analysing by FACS.

3.2.8 Extraction of genomic DNA from cultured cells

Cells (up to 40 x 10^6) were resuspended in a final volume of 1200 µl Lysis Buffer (0.1M Tris pH8.0, 0.2M NaCl, 5mM EDTA, 0.4% SDS) and 30 µl Proteinase K (20mg/ml, final concentration: 0.5mg/ml) solution. After incubation at 60°C overnight or until the pellet is fully dissolved, genomic DNA was precipitated using isopropanol and washed with 70% Ethanol. Genomic DNA was resuspended in TE Buffer and concentration was measured using Nanodrop.

3.2.9 Single gRNA cloning

Oligonucleotides were ordered from IDT with the appropriate overhang sequences (underlined) for cloning into the gRNA expression plasmid pKLV2_U6gRNA_Bbsi_PGK_GFP. The expression plasmid pKLV2_U6gRNA_Bbs1-PGK-tagBFP-2A-GFP-W was purchased from Addgene (#67979) and the tagBFP protein removed, resulting in pKLV2_U6gRNA_Bbsi_PGK_GFP. gRNAs that start with a G nucleotide at the first position result in an increased CAS9-mediated cleavage efficiency when using a U6 promoter [121]. Therefore, the first nucleotide from a 20nt gRNA sequence was replaced with a G nucleotide. The final sequences ordered were as follows:

Forward 5'-<u>CACCG</u>-gRNA 19nt-3'

Reverse 5'-<u>AAAC</u>-gRNA 20nt-3'

Forward and reverse oligos were annealed for 10 min at 95°C and cooled to room temperature. Annealed gRNAs were ligated into the Bbsi-cut expression plasmid using T4 DNA Ligase (NEB) and T4 Buffer (NEB) and incubated at 25°C for 2 hours. After transforming the ligation reaction into DH5α competent cells, positive clones were identified by Sanger sequencing.

3.2.10 Generation of lymphoma-focused CRISPR guideRNA library

gRNA sequences were obtained from Kosuke Yusa's study[122] and David E Root's study[123]. Root's human gRNA list (Brunello) provides a 20nt-gRNA sequence so I appended the first nucleotide with a G nucleotide (as explained in 3.2.9). Yusa's human gRNA list provides a 19nt-gRNA sequence so I appended a G nucleotide to the 5' end. Appropriate overhang sequences (underlined) for cloning into the gRNA expression plasmid pKLV2_U6gRNA_Bbsi_PGK_GFP (modified from Addgene #67979) were appended to all 6000 gRNAs. A 70-mer oligo pool was purchased from TWIST BIOSCIENCE as follows:

5'- <u>TATCTTGTGGAAAGGACGAAACACCG</u>-N_{19}-<u>GTTTAAGAGCTATGCTGGAAACAGC</u>-3'

N_{19} represents each of the 6000 gRNA sequences. The single-stranded oligo pool was converted to double-stranded DNA by PCR amplification using Q5 Hot Start High-Fidelity 2X Master Mix (NEB) with 3 ng of the oligo pool as a template and primers (Zhang_F and Zhang_R_modified). The following PCR conditions were used: 95 °C for 2 min, 10 cycles of 95 °C for 20 s, 60 °C for 20 s and 72 °C for 30 s, and the final extension, 72 °C for 3 min. The 150bp PCR product was gel purified from a 2% Agarose gel using the Gel Extraction kit (Qiagen) and eluted in 20 μl EB Buffer. Four Gibson Assembly reactions were performed using 14.4 ng of the purified 150bp fragment and 200 ng of the BbsI-digested pKLV2_U6gRNA_Bbsi_PGK_GFP with Gibson HiFi DNA Assembly Master Mix (NEB). Gibson Assembly reactions were pooled and column-purified using MinElute PCR purification kit (QIAGEN). Eight electroporations were performed using 1 μl of the purified Gibson reaction and 20 μl of Endura Competent Cells (Lucigen). The mixture was transferred to a 0.1 cm cuvette and electroporated at 1.8KV. Immediately after, 2 ml of prewarmed SOC media was added to each reaction and placed on a shaker at 37°C for one hour. The electroporated cells were combined and plated onto sixteen 24.5 cm^2 LB + ampicillin

agar plates using ColiRollers Plating Beads (Merck Millipore) and left at 30°C overnight. 306 000 resulting bacteria from sixteen plates (51x library complexity) were scraped off and plasmid DNA was purified using a Plasmid Maxi kit (Qiagen). The decanted pellet was weighed as pellets bigger than 0.45g will block the Maxi kit column. The Staudt group have empirically determined that ~ 10-30x coverage is sufficient for 100% gRNA representation in the library (personal communication with James Phelan). *See 3.1.7 for a list of PCR primers user in this study.*

3.2.11 Generation of mutant libraries and screening

GC B cells were transduced with the oncogene cocktail and CAS9-BFP retrovirus until CAS9-BFP reached between 50 and 80%. The number of cells being transduced was dependent on the percentage of CAS9 and the anticipated MOI of double positive cells. For example, for a cell population consisting of 60% CAS9-BFP and an anticipated 30% of double positive cells (CAS9-BFP and CRISPR library-GFP); 31.7×10^6 were transduced with a pre-determined volume of lentiviral CRISPR library to obtain a coverage of 1000x the size of the library. Four days after transduction, BFP and GFP expression was analysed by flow cytometry and at every harvest timepoint going forward. At every passage, the numbers of cells that represent 1000x coverage of the library, taking into account only double positive cells were seeded in a new T75cm^2 tissue culture flask. This was crucial to avoid genetic bottlenecking. Seeing that double positive cells enriched over time, the number of cells to passage became smaller. This also applied to the cell number that was harvested approximately every 14 days. Harvest timepoints may vary between different screens.

3.2.12 Generation of Illumina sequencing library

Genomic DNA extraction was conducted as described in 3.2.8. For sequencing of all gRNAs in the CRISPR library, primers (gLibrary-HiSeq_50bp-SE-U1 and −L1) were used to amplify the region containing the gRNA. For this PCR, it is crucial to use sufficient genomic DNA to capture every gRNA in the cell population. This depends on the complexity of populations to be analysed. For human cells, 10^6 cells contain around 6.6 µg of genomic DNA (assuming normal copy number). Therefore, the number of cells (millions) harvested * 6.6 µg will correspond to the amount of genomic DNA needed in the first PCR. For example, if a cell population was 30% double positive for CAS9-BFP and CRISPR library-GFP, then 20×10^6 cells were harvested to achieve 1000x coverage (6000 guides x 1000 coverage / 0.3). In

this case 131 µg (20 x 6.6) of genomic DNA was used in the first PCR, with a maximum of 10 µg per 50 µl reaction. Therefore, 13 independent PCR reactions were performed using 10 µg of genomic DNA per reaction with Q5 Hot Start High-Fidelity 2× Master Mix. The following PCR conditions were used: 98 °C for 30 s, 20-24 cycles of 95 °C for 10 s, 61 °C for 15 s and 72 °C for 20 s, and the final extension, 72 °C for 2 min. 5 µl from each PCR reaction were run on 2% Agarose gel and PCR was run for a few more cycles if there was no PCR product or PCR bands were still faint. Expected band size was 262 bp. Next, 5 µl from each individual PCR reaction per sample was taken, pooled and purified using QIAquick PCR Purification Kit (Qiagen). DNA was eluted in 50 µl EB buffer (Qiagen) and concentration was quantified on the nanodrop. In the second PCR, nextgen sequencing adaptors (P5, P7) compatible with Illumina's HiSeq4000 and a barcode were added. One nanogram of the purified PCR product was used with NEBNext Q5 Hot Start HiFi PCR Master Mix with the following conditions: 98 °C for 30 s, 9-12 cycles of 98 °C for 10 s, 65 °C for 75 s and the final extension, 65 °C for 5 min. Forward primer named Indexing Adapter PE 1.0 and different reverse indexing primers (iPCRtagT1-56) were used in this second PCR. A different reverse indexing primer was used for each sample. 5 µl from each PCR reaction were run on 2 % agarose gel and checked for visible PCR bands. Expected band size was 262 bp. The PCR products were purified with Agencourt AMPure XP beads in a PCR-product-to-bead ratio of 1:0.7 and eluted in 30 µl EB Buffer (Qiagen). The purified libraries were quantified, pooled and sequenced on Illumina HiSeq4000 by 50-bp single-end sequencing. Two custom sequencing primers were used here; iPCRtagseq which reads through the indices and U6-Illumina-seq2 which reads through the gRNA sequence. Joao Dias provided bioinformatics support to extract read numbers for each gRNA in the library and I further analysed the data. Enriched gRNAs were defined based upon enrichment relative to the plasmid pool counts. All primer sequences were obtained from [122]. *See 3.1.7 for a list of PCR primers used in this study.*

3.2.13 Mouse injection and tumour harvest

Cultured human germinal centre B cells were injected subcutaneously into the left flank of NSG mice. Up to 10 x 10^6 cells were washed and resuspended with Matrigel (Corning) in a 1:1 ratio. Mice were culled when tumours reached 12 mm in size and organs like spleen, lymph nodes and liver checked for metastasis. Tumours were processed as follows: 1.) 3-5 tumour pieces were snap frozen and stored at -

80°C for further biochemical analysis. 2.) 2-3 tumour pieces were emerged in Formalin for further histological analysis. 3.) Immediately after harvest, tumours were mashed through a 70 µM strainer, washed and stained for GC B cell markers (CD38, CD20, CD19, and CD10) and other appropriate markers. Viable cells were frozen in 10% DMSO and a cell pellet frozen for further analysis. *See 3.1.6 for a list of Flow Cytometry Antibodies used in this study.*

3.2.14 Histology

Mouse tumours derived from human GC B cells were fixed in 10% Formalin after harvest and histological analysis performed by Dr Hesham Eldaly at the Addenbrooke's Histology Department.

3.2.15 FACS (Fluorescence-activated cell sorting)

Cells were stained with Fluorophore-labelled antibodies in 2% BSA in PBS for 20 min at 4°C. The stained/or unstained cells were analysed on the LSRII (BD). *See 3.1.6 for a list of Flow Cytometry antibodies used in this study.*

3.2.16 Immunofluorescent staining of intracellular cytokines

Intracellular staining of phosphorylated AKT was performed as follows: Cell suspension and pre-warmed Fixation Buffer (BD Cytofix) was gently mixed in a 1:1 ratio and incubated at 37°C for 15min. Cell suspension was pelleted and washed with PBS twice at 350g for 5 min. Ice-cold True-Phos perm buffer (BD Cytofix) was added dropwise to the cell pellet whilst vortexing, followed by incubation at -20 ^{o}C for at least 60 min. Cells were further washed twice and resuspended in FACS buffer (PBS + 2% FBS) containing the appropriate antibody at a dilution of 1:50 (Phospho-Akt Ser473, Cell Signalling, #11962). After staining for 30min, cells were washed and resuspended in FACS buffer followed by analysis on the LSRII (BD).

3.2.17 Transformation of competent cells

As described in Chapter 1, Methods 3.2.5.

3.2.18 Plasmid Purification

As described in Chapter 1, Methods 3.2.6.

3.2.19 Western blotting

As described in Chapter 1, Methods 3.2.10. *See 3.1.4 for a list of western blot antibodies used in this study.*

3.2.20 RNA purification

Total RNA from cells was extracted using NucleoSPIN RNA from Macherey-Nagel (Bethlehem, PA) and cDNA was produced from 1μg of total RNA using qScript^{TM} cDNA SuperMix (Quanta Biosciences, Beverly, MA).

3.2.21 Cell Cycle Analysis

Click-iT® EdU Alexa Fluor® 647 Imaging Kit (Life Technologies) was used for cell cycle analysis according to the manufacturer's instructions. EdU incorporation (APC-labelled anti-EdU antibodies) and DNA content (7-AAD, 7-aminoactinomycin D) were measured in fixed and permeabilized cells. Viable cells were identified based on scatter morphology and gated cells were further analysed.

3.2.22 Illumina sequencing analysis

Reads in fastq format were demultiplexed by Joao Dias and read counts given to me for further analysis. All raw read counts were normalized to Reads Per Kilobase of transcript per Million mapped reads (RPKM).

To calculate the gRNA score, the following equation was used:

$$gRNA\ score = z - score\left[log2\left(\frac{Timepoint\ X\ normalized\ gRNA\ count}{normalized\ plasmid\ library\ gRNA\ count}\right)\right]$$

To calculate the CRISPR gene score, the following equation was used:

$$CRISPR\ gene\ score = Average\ (gRNA\ score\ for\ gene)$$

To calculate Fold enrichment over time, the following equation was used:

$$Fold\ enrichment\ of\ gene\ X = \left[log2\left(\frac{\frac{Timepoint\ X\ normalized\ gRNA\ count}{normalized\ plasmid\ library\ gRNA\ count}}{Median\ of\ non-targeting\ controls}\right)\right]$$

4 Results

4.1 Optimization of novel co-culture system consisting of follicular dendritic cells and germinal centre B cells (GC B cells)

4.1.1 Follicular dendritic cells expressing CD40Lg and IL21 allow for *in vitro* growth of isolated GC B cells from tonsil

To screen the functional importance of the numerous mutations found in DLBCL using a non-malignant, primary, human germinal centre B cell (GC B cell), we first needed to culture GC B cells *in vitro*, something that has proven notoriously difficult in the past. Several groups have attempted this with slightly different culturing methods. Banchereau and colleagues first pioneered a culture method that stimulates the proliferation of B cells by culturing on mouse fibroblastic L cells stably expressing FcyRII/CDw32 and by the addition of anti-CD40 and cytokines such as IL-4[124,125]. A similar method was presented by Kwakkenbos and colleagues who isolated memory B cells from blood or frozen peripheral blood mononuclear cells (PBMC) and cultured these cells for a limited timespan on irradiated mouse fibroblastic L cells with CD40Lg and IL21[126]. In the context of mouse B cells, Nojima and colleagues developed a culture system in which mouse naïve B cells proliferate and adopt a GC B cell phenotype[127]. This culture system consisted of fibroblasts transfected with CD40Lg and BAFF with the addition of soluble IL-4[127].

Another attractive approach was presented by Kim and colleagues who established a follicular dendritic cell (FDC) – like cell line, HK[128]. HK cells alone did not stimulate GC B cell proliferation but this could be induced with the addition of anti-CD40 to the co-culture[128]. This system mimics to some degree the germinal centre environment in the lymph nodes and allows for growth and survival of human tonsillar B cells. As mentioned earlier, FDCs play an important role in GC development; they are necessary for differentiation and proliferation of GC B cells by providing co-stimulatory functions[128,129]. GC B cells die rapidly in the absence of FDCs. FDCs used by Kim and colleagues had a very slow growth rate and efforts to immortalize these cells had previously been unsuccessful due to difficulty of transduction (personal communication with Chan-Sik Park, University of Ulsan College of Medicine, Seoul, Korea). This meant FDCs had to be freshly isolated from tonsil and grown for individual experiments.

To bypass these problems and to improve on existing techniques, we optimized a new co-culture system based upon human tonsillar FDCs (Figure 1A). These cells were obtained from Chan-Sik Park. Similar to HK cells, these primary FDCs have a very slow growth rate and stop proliferating after approximately 20 population doublings. In order to eliminate the frequent need to re-establish FDC cultures from tonsils and to avoid variations between cultures, we decided to immortalize these FDCs to extend their replicative capacity. This allowed us to use the same FDCs throughout my experiments. We also wished to enhance B cell proliferation by transducing these FDCs to stably express CD40Lg and cytokines.

Several methods for immortalizing cells exist, such as inactivation of tumour suppressor genes (e.g. p53, Rb and others), overexpression of mutants like Ras or CDK4; or through the expression of human Telomerase Reverse Transcriptase protein (hTERT). Studies have shown that co-expression of hTERT and p53 siRNA allows for immortalization of human, primary, ovarian, or epithelial cells[130]. Similarly, co-expression of hTERT and CDK4 can immortalize bronchial epithelial cells[131]. As different cell types seem to require different immortalisation strategies, we transduced primary FDCs with various combinations of immortalising vectors to identify the method that gives the best long-term growth advantage. Instead of p53 siRNA, we used a dominant negative form of p53[132] and the hyperactive mutant CDK4 R24C. The combinations were as follows:

1.) hTERT
2.) CDK4 R24C
3.) P53dd
4.) hTERT + CDK4 R24C
5.) hTERT + P53dd
6.) P53dd + CDK4 R24C
7.) hTERT + P53dd + CDK4 R24C
8.) Untransduced

We identified the combination of hTERT, a dominant negative form of p53 and CDK4 R24C to be able to immortalize FDCs (Figure 1B). FDCs transduced with hTERT alone stopped replicating after a similar timespan to untransduced cells, after approximately 85 days. The other combinations were also not able to immortalize these cells before they reached replicative senescence. This suggests that in this cell type, hTERT, P53dd and CDK4 R24C are required for full immortalization.

Studies have shown that B cells can be stimulated *in vitro* by CD40Lg and cytokines such as IL4, IL2, IL10 or IL21[124,126,133,134]. Unpublished data from Daniel Hodson shows that stimulation with IL21 results in better B cell proliferation than IL4, IL2, or IL10, and hence IL21 was used for the feeder system set up here. Therefore, immortalized FDC were further retrovirally transduced with membrane-bound CD40Lg-Puro and IL21. Transduced FDCs were selected with puromycin and two rounds of transduction with IL21 were necessary for >90% transduction efficiency (data not shown). This created an immortalised FDC-derived feeder system that expresses the dominant T-cell helper molecules CD40Lg and IL21.

When GC B cells were cultured on this feeder system, we observed vigorous cell growth for a period of approximately 11 days. Withdrawal of any individual component of the system, such as feeder cells, IL21 or CD40Lg completely abrogated this proliferation (Figure 2A). Having identified a feeder system that supports B cell growth, the immortalized CD40Lg-IL21 expressing FDCs were further expanded large scale, irradiated (30 Gy) and frozen down in aliquots for future experiments. This allowed for consistency between experiments.

The immortalized follicular dendritic cells expressing CD40Lg and IL21 allowed us to grow primary human GC B cells short term *in vitro*. The purity of human GC B cells was >90%, as assessed by expression of $CD38^+CD20^+CD19^+CD10^+$ (Figure 2B). In order to prolong growth of these cells, I introduced oncogenes into these human GC B cells which allowed for long term growth *in vitro* (as will be described in 4.1.3). A schematic of this approach can be seen in Figure 2B.
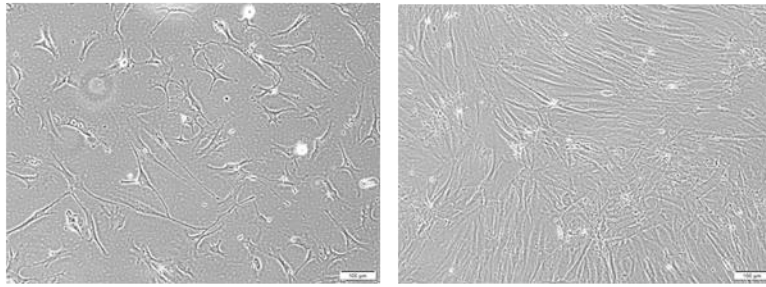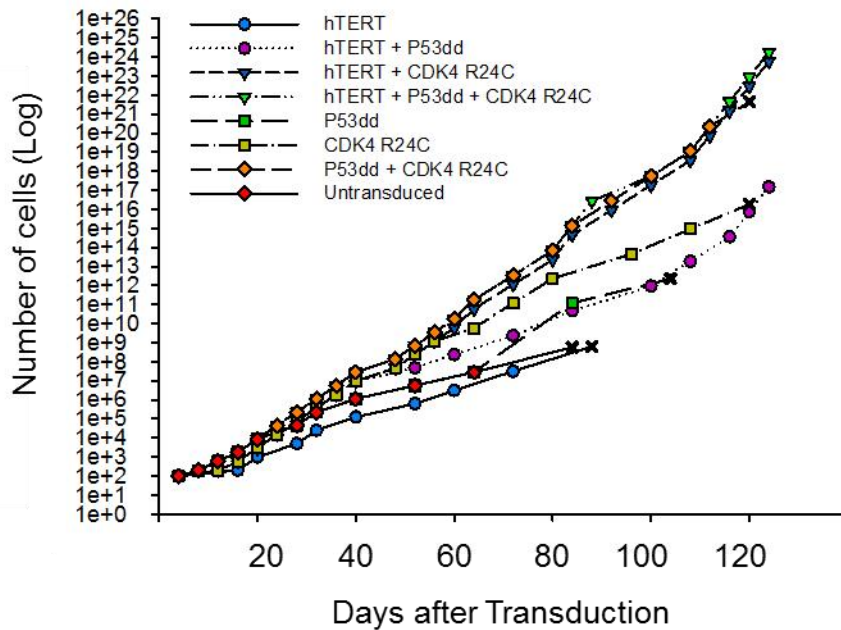
Figure 1 hTERT, P53dd and CDK4 R24C allow full immortalization of human tonsillar follicular dendritic cells

(A) Representative images of adherent follicular dendritic cells at different confluency are shown.

(B) Immortalization of follicular dendritic cells was achieved by transduction with different oncogenes. Cell proliferation rate was monitored by manual cell counting. Illustrated is time course showing the number of cells following transduction. The combination hTERT, P53dd and CDK4 R24C gave the best proliferation advantage closely followed by the combination hTERT and CDK4 R24C. Other single genes and combinations that had stopped proliferating and underwent replicative senescence are marked with an x at their final timepoint. Experiment was performed with Miriam Di Re (Hodson Laboratory).
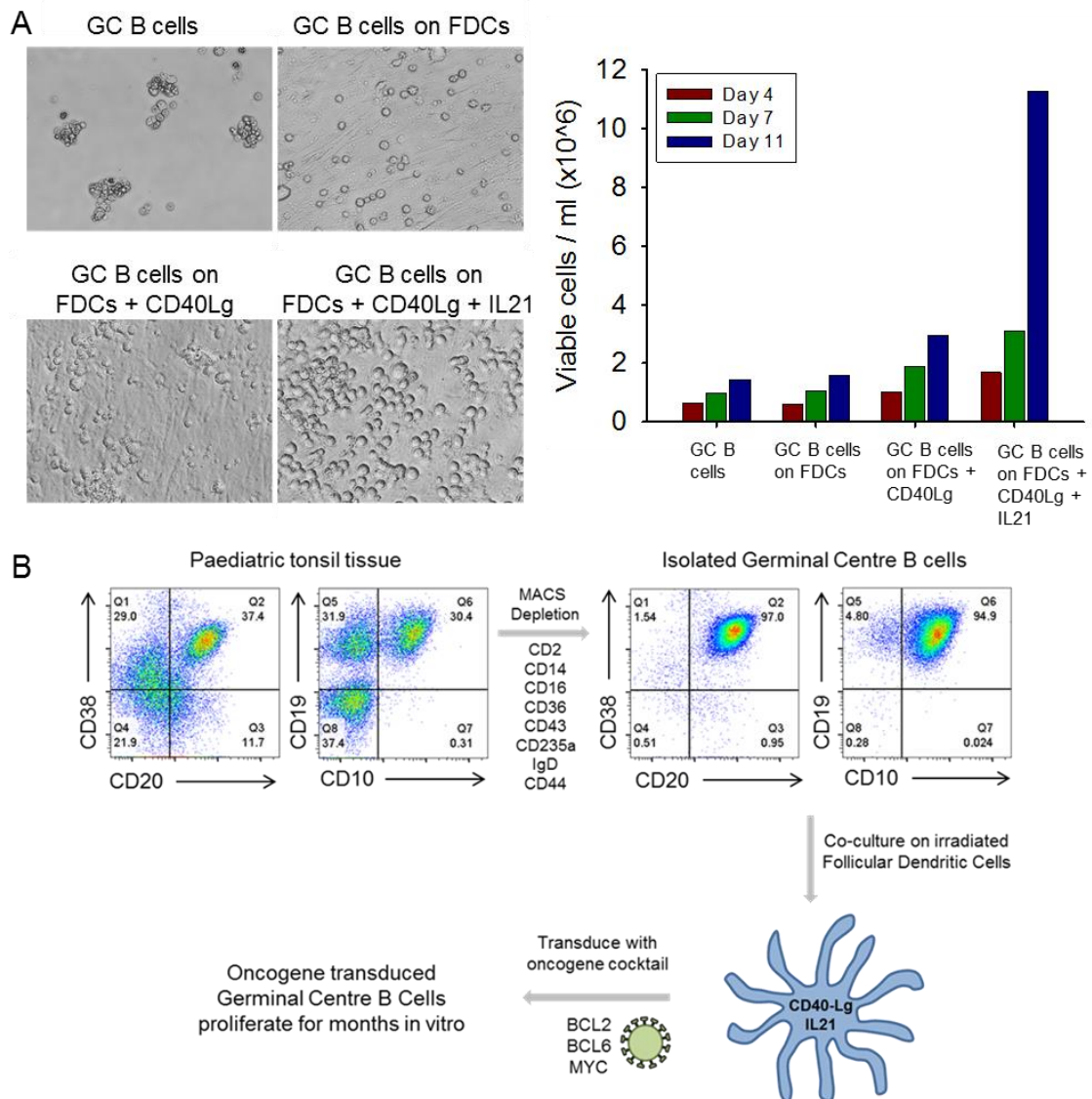
Figure 2 Immortalized follicular dendritic cells expressing CD40Lg and IL21 allow for *in vitro* growth of isolated GC B cells

(A) On the left, representative images are shown of GC B cells, GC B cells on FDCs, GC B cells on FDCs + CD40Lg and GC B cells on FDCs + CD40Lg + IL21. On the right, quantitative analysis of the number of viable cells at three different timepoints; Day 4, Day 7 and Day 11. Experiment was performed with Miriam Di Re (Hodson Laboratory).

(B) GC B cells were isolated from tonsil tissue using a MACS human B cell negative selection kit as well as anti-IgD and anti-CD44. After purification, cells were stained with CD38, CD20, CD19 and CD10. These cells were only used for experiments if the purity was >90%. GC B cells were co-cultured on irradiated follicular dendritic cells expressing CD40Lg and IL21 and further transduced with an oncogene cocktail to allow for long-term growth *in vitro.* Representative plots are shown.

4.1.2 A newly engineered retroviral and lentiviral envelope allows for high transduction efficiency in GC B cells

Transducing primary human GC B cells has been notoriously difficult. Conventional enveloped viruses such as VSV-G and amphotrophic retro and lentivirus are unable to transduce human B cells at high efficiency[135-137]. Pseudotyping, the process of replacing the native envelope with different viral fusion proteins allows vectors to transduce cells that would not be possible with the native envelope[120]. The murine leukaemia virus (MuLV) is one of the most commonly used vectors for gene transfer and often pseudotyped by envelope proteins from other type C mammalian retroviruses such as the Gibbon ape leukaemia virus (GaLV)[120]. GaLV uses an internalization receptor, SLC20A1 that is highly expressed on human B cells and renders them susceptible to retroviral infection[138-140]. In fact, previous studies showed that retrovirus packaging cells containing the Moloney murine leukaemia virus (MoMLV) gag-pol proteins and the Gibbon ape leukaemia virus (GaLV) env protein were able to infect human B cells[141,142].

However, the envelope protein from GaLV does not form functional pseudotyped lentiviral vectors[120]. The reason for this is not fully understood but Christodoulopoulos suggested that this is due to the cytoplasmic tail being incompatible to be incorporated into lentiviral vector particles[120]. The cytoplasmic region (cytoplasmic tail (T) and R peptide (R)) belongs to the transmembrane subunit which also consists of the transmembrane region (M) (Figure 3A). Christodoulopoulos showed that replacing the whole transmembrane region (MTR) or the cytoplasmic region (TR) with the corresponding MuLV region produced high titers of lentivirus vectors[120].

Here, I took advantage of the homology between the GaLV and MuLV envelope proteins and constructed retroviral and lentiviral vectors that are able to efficiently transduce primary human GC B cells (Methods 3.2.4). pHIT123 containing the Mo-MuLV Env protein, human cytomegalovirus immediate-early promoter and the origin of replication from simian virus 40 was used as the backbone. I cloned the GaLV envelope into this backbone with the MTR or TR region replaced by the corresponding MuLV region, creating two lentiviral vectors named GaLV_MTR_RC and GaLV_TR_RC, respectively. A schematic can be seen in Figure 3A. GaLV_MTR_RC and GaLV_TR_RC achieved transduction efficiencies of approximately 80% and 50%, respectively as was determined by expression of GFP

(Figure 3B). As expected, VSV-G, the ecotropic MuLV and the full GaLV envelope (GaLV_WT_RC) were not able to transduce primary human GC B cells (Figure 3B).

Additionally, all three newly engineered constructs were able to achieve high retroviral transduction efficiencies in GC B cells as was determined by expression of GFP (~80%) in these cells (Figure 3C).

The engineered retroviral and lentiviral envelopes described here, allowed us to efficiently transduce human GC B cells which is critical to genetically modify these cells.
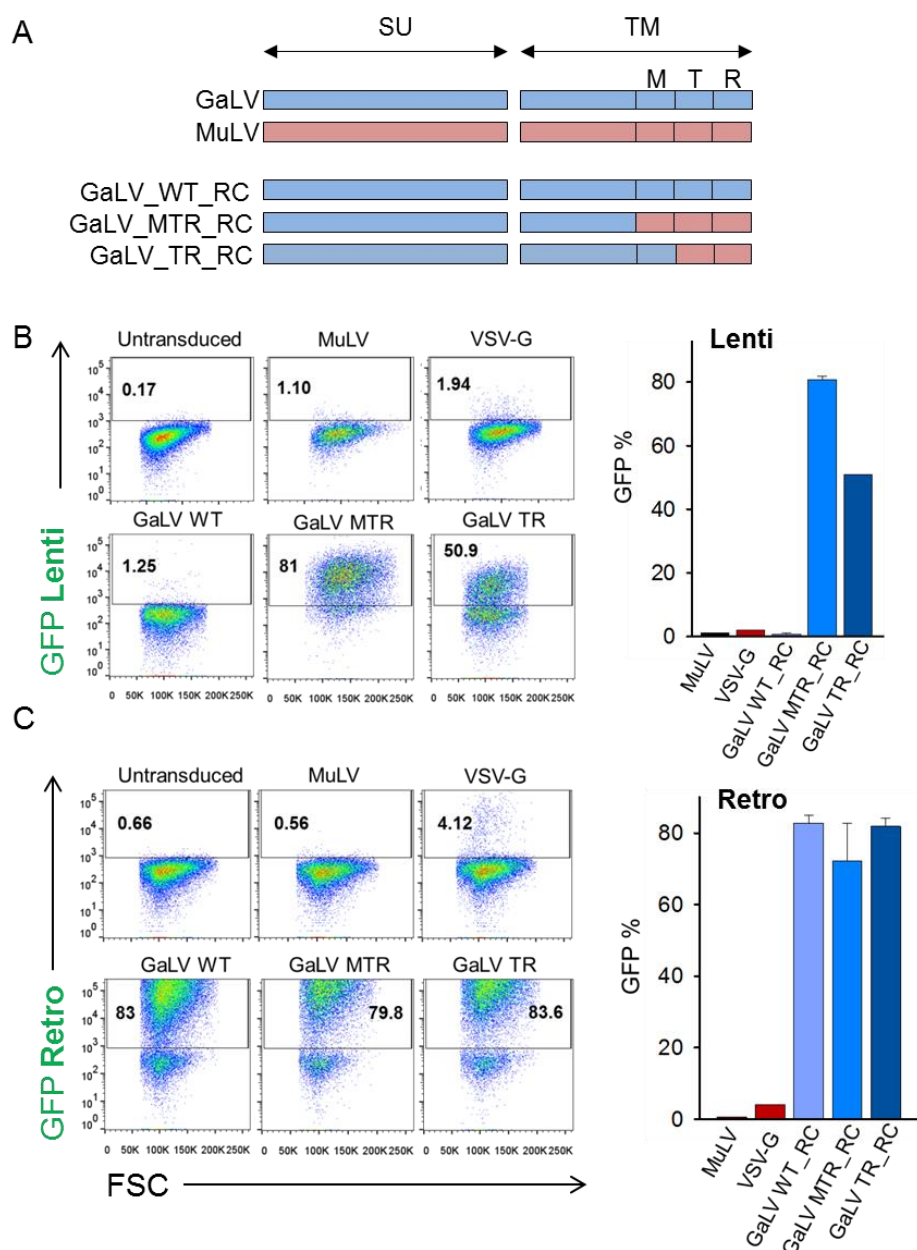
Figure 3 A newly engineered retroviral and lentiviral envelope allows for high transduction efficiency in human GC B cells

(A) Schematic of the retroviral and lentiviral envelopes, GaLV_WT_RC, GaLV_MTR_RC and GaLV_TR_RC. M = transmembrane region, T = cytoplasmic tail, R = R peptide, SU = surface subunit, TM = transmembrane subunit[120].

(B&C) Primary human GC B cells were transduced with a lentiviral control (B) or retroviral control (C) construct using the newly engineered envelope constructs as well as VSV-G and MuLV. Three days after transduction, transduction efficiencies in primary human GC B cells were determined by expression of GFP using a flow cytometer. Representative of 3 experiments is shown when error bar is indicated on the bar graph.

4.1.3 A minimum of two oncogenic hits are necessary for long term survival of human GC B cells

As described, human GC B cells proliferate vigorously on immortalized follicular dendritic cells expressing CD40Lg and IL21 for up to 11 days. In order to prolong growth of these cells *in vitro* and to allow for long term experiments, I introduced some of the most common genetic alterations into these human GC B cells using the viral envelope described in 4.1.2. The platforms were as follows:

1.) BCL2
2.) BCL6
3.) MYC
4.) BCL2 + BCL6
5.) BCL2 + MYC
6.) Untransduced

I identified two oncogene combinations that after transduction, allowed GC B cells to grow and proliferate *in vitro* for months. Successful cocktail combinations were BCL2 + BCL6 and BCL2 + MYC. For combination experiments, I cloned each oncogene into a single t2A linked vector (BCL6-t2A-BCL2 and MYC-t2A-BCL2). These two platforms allowed cell growth with continuous rapid proliferation for >3 months, at which point the cells were discarded due to logistical reason (Figure 4A). Oncogenes on their own; such as BCL2 alone, BCL6 alone or MYC alone were not sufficient to transform cells and cells died rapidly (Figure 4A).

FACS staining for the cell surface marker CD2 linked to the t2A linker showed that > 90% of cells were positively transduced three days after infection (Figure 4B). Even in cases where the transduction efficiency did not reach 90%, GC B cells rapidly self-selected for BCL6-t2A-BCL2 or MYC-t2A-BCL2 within a week.

It is likely that these are not the only combinations that allow for long-term growth of GC B cells and I am currently screening additional combinations.

The ability to grow oncogene transduced GC B cells for months *in vitro* makes them a suitable tool for long-term studies and for further exploitation.

A

B

| Untransduced | MYC-t2A-BCL2 | BCL6-t2A-BCL2 |
| 0.36 | 99.6 | 98.8 |
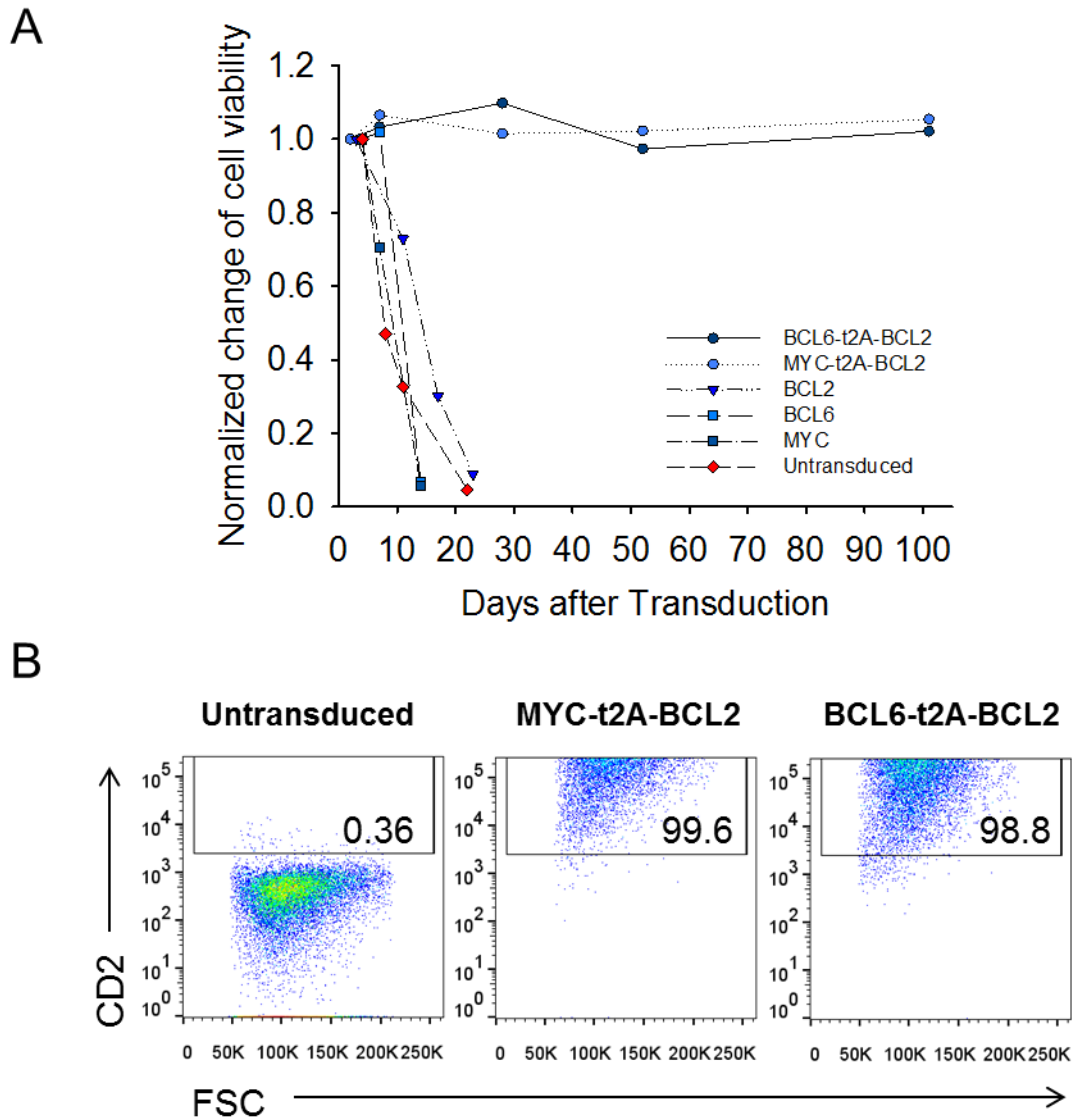
Figure 4 A minimum of two oncogenic hits are necessary for long term survival of human GC B cells

(A) Cell Viability of human GC B cells transduced with different oncogenes was measured. Illustrated is time course showing normalized change in cell viability. Representative of 2-3 experiments.

(B) Human GC B cells transduced with MYC-t2A-BCL2 or BCL6-t2A-BCL2 were stained for its marker CD2 and analysed by flow cytometry three days after transduction. Representative of > 3 experiments.

4.2 Immunodeficient mice injected with transduced GC B cells develop human tumours that closely resemble human DLBCL

I wished to examine whether oncogene transduced human GC B cells, described in this study, would model lymphomagenesis in immunodeficient mice (NOD/SCID/gamma mice) and would therefore provide a tool to study DLBCL genetics.

I performed subcutaneous injections to elucidate the mechanism of local tumour formation and intravenous injection (IV) to learn about metastatic tumour formation. GC B cells (as well as FDCs) transduced with 4 different oncogene cocktails were injected into four immunodeficient mice per cohort. Pilot experiments showed that GC B cells transduced with only two oncogenes (such as *BCL2* and *BCL6)* were insufficient to form tumours and therefore three and four gene combinations were used. The cohorts were as follows:

1.) MYC, BCL2, P53dd,
2.) MYC, BCL2, P53dd, CCND3 T283A
3.) BCL6, BCL2, P53dd,
4.) BCL6, BCL2, P53dd, CCND3 T283A

As described before (2.5), somatic mutations in *CCND3* are found in approximately 14% of ABC DLBCL and 10% of GCB DLBCL[107]. Mutations are often seen in the threonine residue at position 283 (T283) which produces a highly stable isoform. High CCND3 expression has been shown to be a prognostic factor with poor clinical outcome in patients with DLBCL[143]. Hence, CCND3 plays an important role in GC development and DLBCL lymphomagenesis. Unpublished results of a mutant open reading frame (ORF) screen in human GC B cells from our laboratory identified *CCND3 T283A* to allow for increased cell proliferation when overexpressed. To confirm this and further explore the function of the *CCND3 T283A* mutation, I injected cells transduced with MYC/BCL6, BCL2, P53dd +/- CCND3 T283A.

I did not detect any engraftment when cells were injected intravenously (5M cells per mouse, data not shown) but subcutaneous injection of 10M cells (in Matrigel) per mouse gave rise to tumours that closely resembled human DLBCL. Tumours were very fast growing and time from first detectable tumour to reaching its maximum

allowed size of 12 mm was usually within one week. Hence, Kaplan-Meier Curve in Figure 5A is illustrated as survival and not as tumour growth time. Cohorts injected with cocktail 1 had the most aggressive phenotype and mice had to be culled due to tumours reaching its maximum allowed size (12mm) after a median time of 35 days. Cohorts injected with cocktail 2 and 3 had to be culled at similar times with a median survival time of 111 days for both. 2 mice in cohort 3 and 1 mouse in cohort 2 were found dead with no tumour. The survival curve suggests that MYC is a stronger oncogene than BCL6 and CCND3 T283A accelerates tumour formation regardless of whether it is on a MYC or BCL6 background. I did not observe any spleen, liver or lymph node enlargement in any of these cohorts. The fourth cohort with BCL6, BCL2 and P53dd is still alive with no visible tumour formation at > 6 months post injection.

Tumours were processed immediately after harvest and analysed by flow cytometry for their immunophenotype. Figure 5B shows a representative example of surface expression markers; >90% of cells were positive for GC B cell markers CD19, CD10 and CD38 as well as for the oncogene cocktail markers CD2 (MYC-t2A-BCL2 or BCL6-t2A-BCL2), Thy1.1 (P53dd) and LyT2 (CCND3 T238A/Cntrl).

In addition, tumours were evaluated with a panel of immunohistochemical markers (haematoxylin & eosin, CD20, CD10, CD79A, MUM1 (IRF4), PAX5, BCL2, P53 and CD3) which showed that these tumours closely resembled the histological appearances of human DLBCL. Tumours stained positive for CD20, CD10, CD79A, MUM1 (IRF4), PAX5, BCL2, P53 and negative for the T cell marker CD3 (Figure 6); all markers that are currently used to diagnose DLBCL patients in clinic. All tonsil donors were screened for their EBV status with a highly sensitive PCR to rule out expansion of an EBV-infected cell population (Methods 3.2.3). The EBV status was confirmed using EBER in situ hybridization (Figure 6). All donors used in experiments here were EBV negative.

Re-transplantation experiments showed that tumour formation was much faster than in the initial transplantation. For example, cells from cohort 2 resulted in tumour formation at a median time of 36 days instead of 111 days for the initial transplant (data not shown). Immunohistochemistry also confirmed these tumours to be of B cell origin.

This reinforces the biological relevance of this co-culture system as a tool to study DLBCL genetics and allows for user-defined mutational human lymphoma models.

Figure 5 Immunodeficient mice injected with transduced GC B cells develop human tumours that closely resemble human DLBCL

(A) GC B cells (as well as FDCs) transduced with four different oncogene cocktails were injected subcutaneously into NOD/SCID/gamma mice (n = 4 per cohort) and checked regularly for palpable tumours. Overall survival of the recipient mice is plotted as a Kaplan-Meier curve.

(B) Cells isolated from tumours were stained for oncogene cocktail markers CD2 (MYC-t2A-BCL2 or BCL6-t2A-BCL2), Thy1.1 (P53dd), LyT2 (CCND3/Cntrl) and B cell markers CD19, CD20 and CD10 and were analysed by flow cytometry. A representative example is shown.

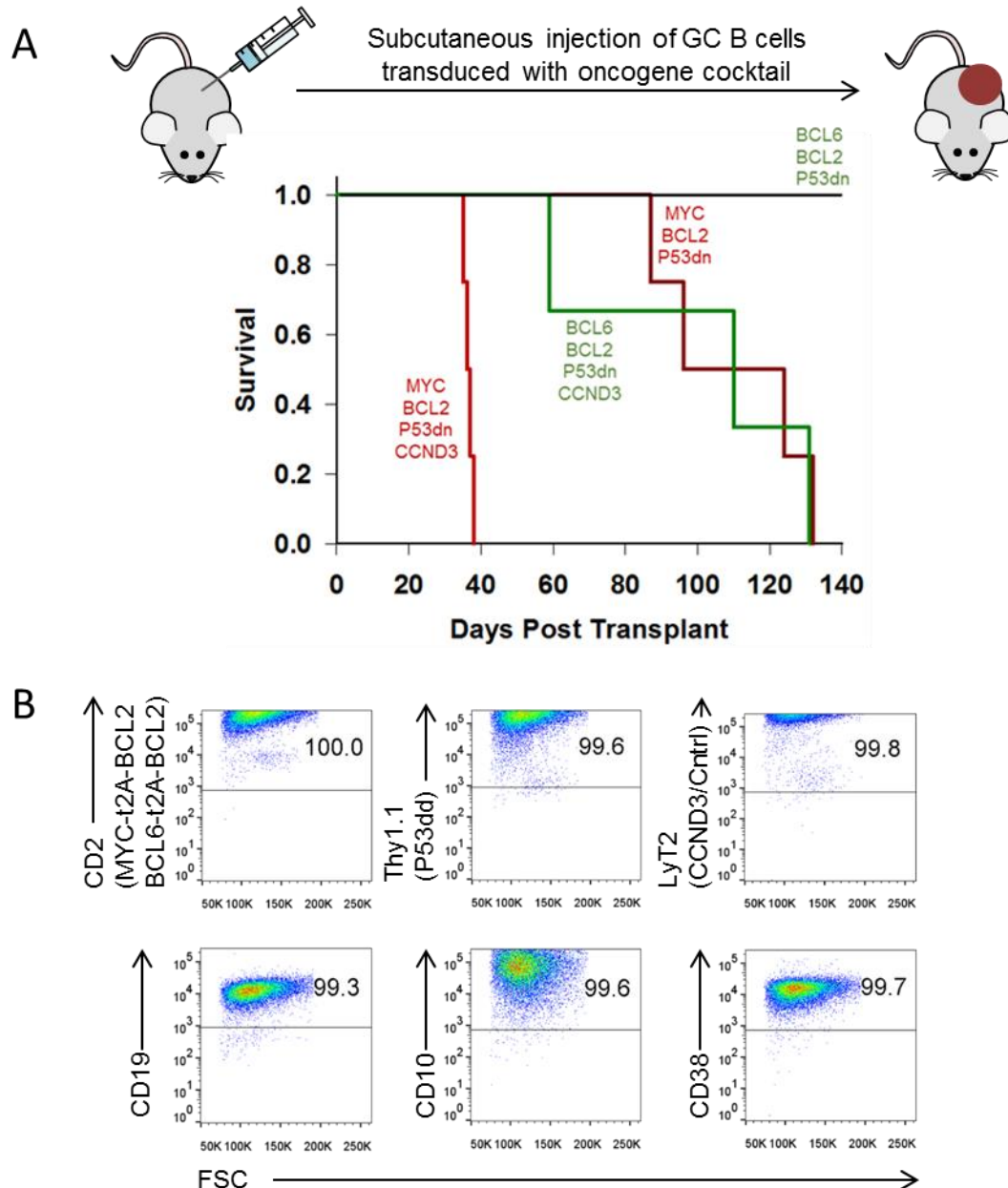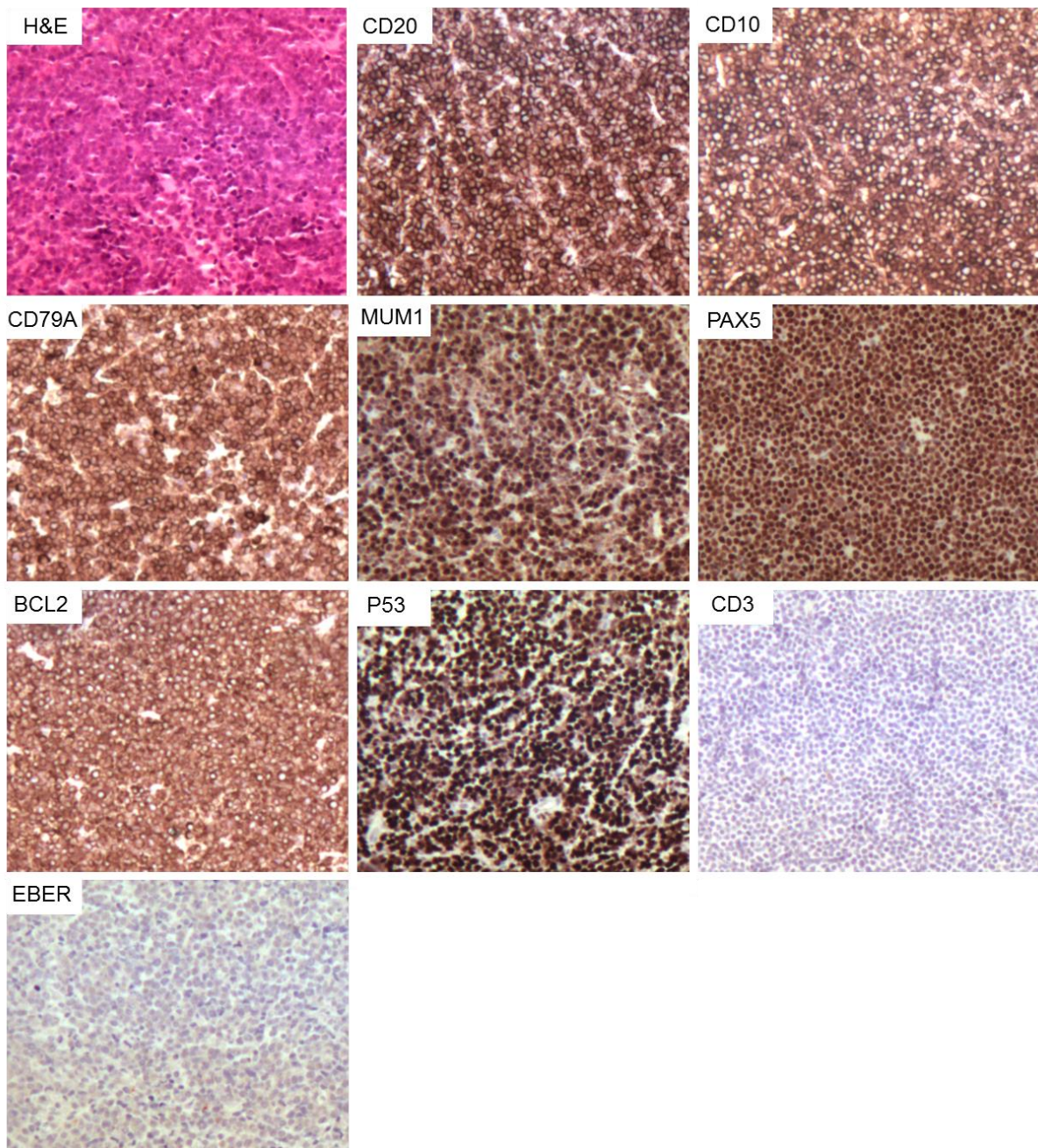**Figure 6 Immunodeficient mice injected with transduced GC B cells develop human tumours that closely resemble human DLBCL**

Immunohistochemistry images for H&E, CD20, CD10, CD79A, MUM1 (IRF4), PAX5, BCL2, P53, CD3 and EBER are shown (Magnification 10x). A representative example is shown. Histology was performed by Dr. Eldaly, Consultant Histopathologist Addenbrooke's Hospital.

4.3 Optimisation of the CRISPR-CAS9 system in primary human GC B cells

4.3.1 Optimising CAS9 expression in primary human GC B cells

I wished to use the described co-culture system to screen tumour suppressor genes in DLBCL using CRISPR. The *Streptoccocus pyogenes* CRISPR system requires the guide RNA and the endonuclease CAS9 to create a double-strand break that will subsequently edit the genomic DNA[108]. The CAS9 component is available in multiple formats and here, I screened three different CAS9 constructs (Retroviral CAS9-Puromycin, Lentiviral CAS9-Blasticidin and Retroviral CAS9-BFP) in order to identify the one that results in the highest CAS9 expression in human GC B cells. CAS9 is a particular challenge to transduce because its large size puts it at the limits of packaging ability for retro and lentivirus. To minimise construct size, we constructed a CAS9-BFP vector that contains only the MSCV backbone with no IRES or structure other than CAS9-t2A-BFP between the 5' and 3' LTR (Miriam Di Re, Hodson laboratory).

Isolated GC B cells were transduced with fresh CAS9 virus or 10x concentrated CAS9 virus (with Lenti-X, Clontech) for all three constructs. Western blot for CAS9 showed the concentrated CAS9-BFP virus to give the strongest protein expression along with unconcentrated CAS9-Puro virus (Figure 7). CAS9-BFP is a smaller retroviral construct that, with its fluorescence marker, allows for convenient tracking of CAS9 over time. This will be important for future experiments. CAS9-Puro and CAS9-Blast did not have a fluorescence marker but flow cytometry showed that concentrated CAS9-BFP resulted in 31.4% BFP$^+$ cells (Figure 7). To increase the percentage of CAS9-BFP$^+$ cells, sequential rounds of transduction with CAS9-BFP were necessary until 50-80% CAS9-BFP$^+$ cells were achieved. Further optimisation for the ideal virus harvest timepoint was performed and I identified concentrated 48h CAS9-BFP virus to give the best transduction efficiency (data not shown).

Successful retroviral infection with CAS9-BFP allowed me to now further transduce these GC B cells with the gRNA to edit genomic DNA.
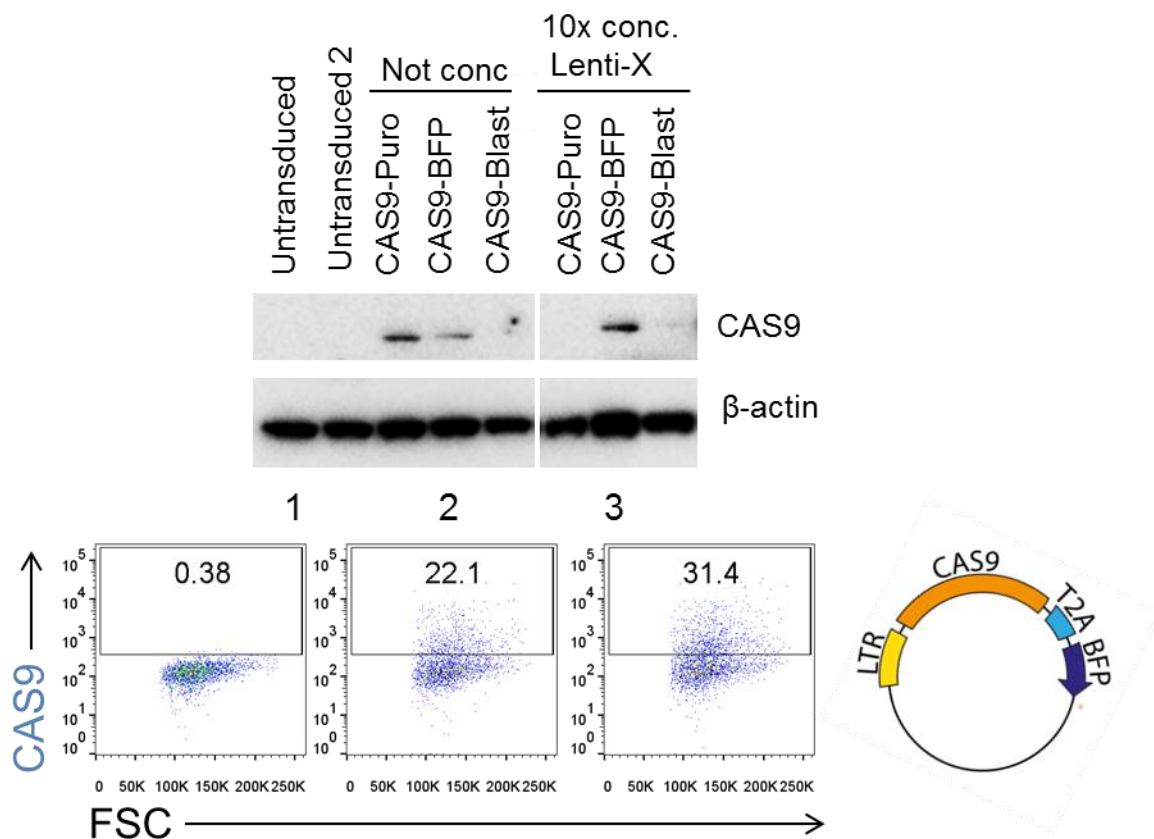
Figure 7 The retroviral CAS9-BFP construct resulted in the highest CAS9 expression in human GC B cells

GC B cells were transduced with unconcentrated or 10x concentrated (with Lenti-X) CAS9-Puro, CAS9-BFP and CAS9-Blast. Western blotting was used to show CAS9 protein expression, using β-actin as a loading control. Flow cytometry gating on BFP+ cells identified 10x concentrated CAS9-BFP to give the highest CAS9 transduction efficiency.

4.3.2 Knock-down of endogenous CD19 and CD22 in GC B cells validates the feasibility of CRISPR in this system and cell type

Before large-scale high-throughput CRISPR experiments could be conducted using the optimised co-culture feeder system, it was necessary to test whether knock-down of single genes using CRISPR was feasible and functional in primary human GC B cells.

To test this, I transduced GC B cells with CAS9-BFP as well as the oncogene cocktail BCL2 and BCL6. Cells were further transduced with two gRNAs against CD19, two gRNAs against CD22 or non-targeting control (NTC). Six days after gRNA transduction, FACS analysis showed efficient knock-down of endogenous CD19 and CD22 in comparison to NTC when gated on double positive CAS9 (BFP) and gRNA (GFP) expressing cells, approximately 30% of the cell population (Figure 8A). At this point, cells were > 90% positive for BCL2 and BCL6 so gating on double positive CAS9 and gRNA expressing cells was sufficient to detect knock-down in this setting.

This confirmed that gene knock-down using CRISPR is feasible in primary human GC B cells and suggested that large-scale CRISPR experiments would be successful in this cell type.

4.3.3 Monitoring the competitive survival of GC B cells after knock-down of tumour suppressor genes validates the feasibility of CRISPR knock-down in this system and cell type

After the CRISPR knock-down feasibility was verified in primary GC B cells, it was necessary to test whether knock-down of tumour suppressor genes would result in a phenotype such as a growth advantage. To do so, I performed a small pilot CRISPR screen in primary human B cells.

I transduced GC B cells with CAS9-BFP as well as the oncogene cocktail BCL2 and BCL6. Cells were further transduced with gRNAs against known tumour suppressor genes *TP53, PTEN, A20*, the cell-essential gene *RPS6* and non-targeting control. The gRNA constructs expressed a trackable GFP marker. Following transduction of gRNAs, percentage of CAS9/gRNA-expressing cells was monitored by flow

cytometry and the fold change relative to the baseline was plotted. At this point, cells were > 90% positive for BCL2 and BCL6. Figure 8B shows that, as expected, GC B cells had a growth advantage if known tumour suppressor genes, such as *TP53*, *PTEN* and *A20* were knocked-down. The most enriched gene was *TP53*. Ribosomal Protein S6 (*RPS6*) seemed to have a growth disadvantage which is consistent with its function as a ribosomal component and a recognised cell-essential gene.

This small pilot CRISPR screen confirmed that human GC B cells can be exploited to study genes that provide a growth advantage or disadvantage using CRISPR.
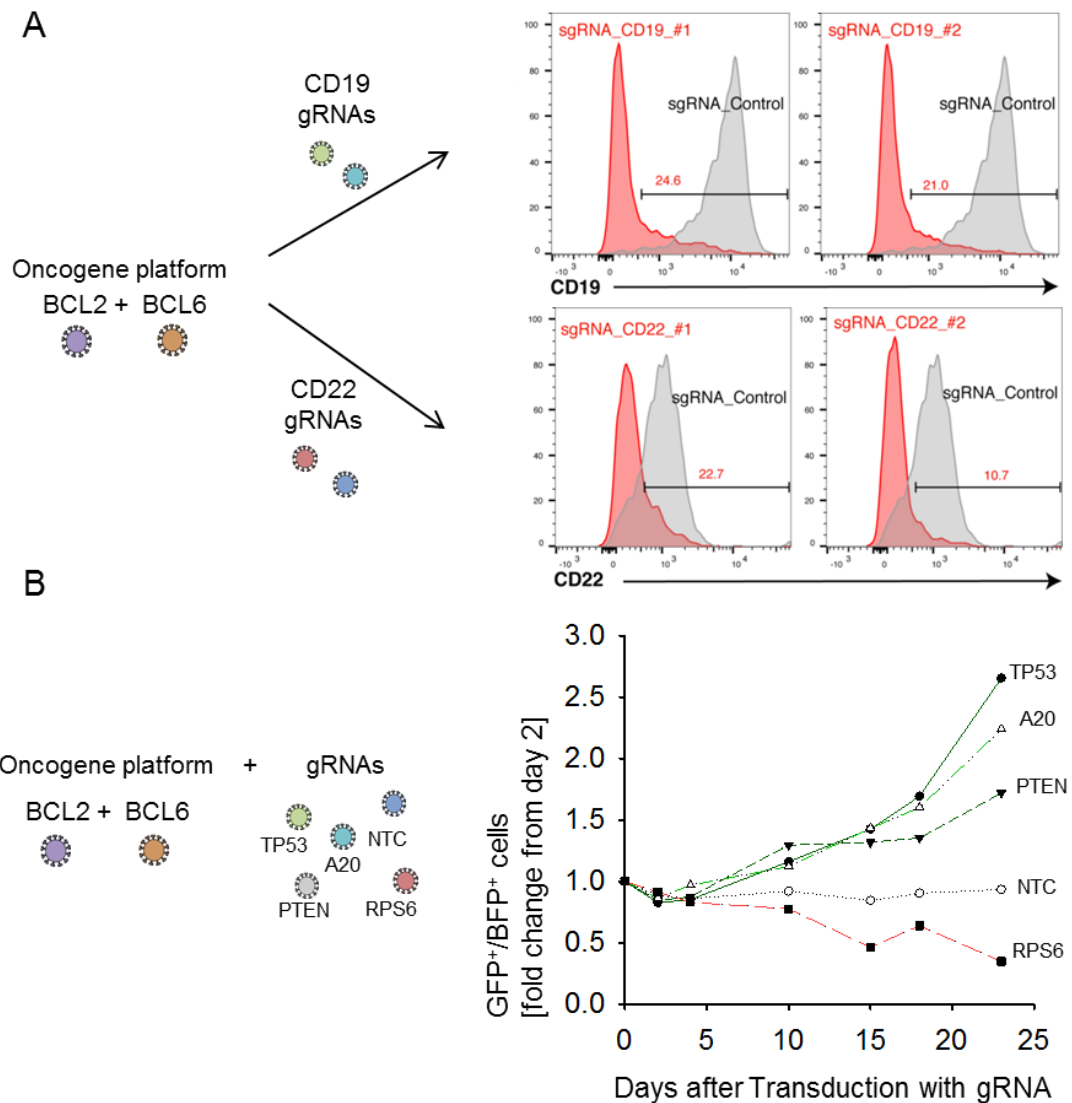
Figure 8 Validation of functional CRISPR knock-down in human primary GC B cells

(A) GC B cells were transduced with BCL2, BCL6 and CAS9-BFP and subsequently with gRNAs against CD19, CD22 and non-targeting control. Staining for CD19 and CD22 was performed 6 days after gRNA transduction and gated on double positive CAS9 (BFP) and gRNA (GFP) expressing cells, Red histograms show CD19/CD22 expression in cells transduced with the indicated gRNA. Grey histograms show expression of CD19/CD22 transduced with a non-targeting control.

(B) GC B cells were transduced with BCL2, BCL6 and CAS9-BFP and subsequently with gRNAs against TP53, PTEN, A20, RPS6 and non-targeting control (NTC). Following transduction with gRNAs, enrichment or depletion of BFP⁺GFP⁺ cells was monitored by flow cytometry. Illustrated is time course showing the relative changes of BFP⁺GFP⁺ cells following transduction.

4.4 Design, construction and validation of a lymphoma-focused CRISPR gRNA library targeting 692 genes

I wished to use the described co-culture system to screen putative tumour suppressor genes in DLBCL using CRISPR in a high-throughput fashion. To do so, I designed and constructed a lymphoma-focused CRISPR gRNA library, which contains 6000 gRNAs targeting 692 genes. Here, I illustrate the design, construction and validation of the lymphoma-focused CRISPR library.

I designed the CRISPR gRNA library to target genes that are recurrently mutated or deleted in DLBCL, Follicular Lymphoma and Burkitt Lymphoma. Manual selections from published transcriptome and exome studies (TCGA, Pan Cancer Atlas and Broad[81]) as well as screening through the literature ([11,80,82,107,144-160]) identified 692 genes to be included in the library. gRNAs for each gene were obtained from two recent whole genome gRNA libraries: 1.) Kosuke Yusa's study used a new design pipeline to identify all possible gRNA target sites, remove unwanted gRNAs, compute design scores and select suitable gRNAs[122]. This approach generated an optimal human gRNA library. 2.) David E Root's study used computational design rules to create a fully optimised human gRNA library that minimizes off-target effects but at the same time maximizes on-target activity[123]. This library is named Brunello. Combining gRNAs from each list resulted in almost all selected genes being targeted by nine gRNAs. The library consisted of 6000 gRNAs in total, including 250 non-targeting controls (NTC) acting as a robust inert control.

To construct the library, forward strand sequences were designed with appropriate overhangs added. These were synthesised (Twist Bioscience) as a pool of 6,000 oligos. The single-stranded oligo pool was converted to double-stranded DNA by PCR amplification. The resulting PCR product was gel extracted and cloned into the BbsI-digested gRNA expression vector, which expressed the gRNA from a U6 promoter. Additionally, the vector contains the green fluorescence protein (GFP) to allow for convenient tracking of the library over time using a flow cytometer. In order to maintain representation of the library, a total of approximately 300 000 bacterial colonies were harvested, representing at least a 50-fold coverage of the library. See Methods 3.2.10 for a more detailed description of the construction of the lentiviral gRNA library. After construction, the library was subjected to next generation

sequencing (NGS) of a PCR product across the gRNA sequences to assess the coverage of gRNAs in the library. Reads in fastq format were demultiplexed by Joao Dias and further analysis was performed by myself.

Illumina sequencing at a depth of >6 x 10^6 mapped reads showed that all 6000 gRNAs were present in the library with 1100 median sequencing reads. Analysis of the gRNA read counts showed that 99% of all read counts were within four times of the mean (Figure 9A). The gRNA sequence reads had a frequency of 99% being four times of the mean, making this a very even distribution (Figure 9B). A shRNA CRISPR library is considered to be adequately represented when the sequence reads have a frequency of 80-90% within a 10-fold range[161]. Hence, the library constructed in this study should be considered to be adequately represented. Analysis of the genes represented in the library shows that 88% of genes were represented by at least 8 or 9 gRNAs (Figure 9C). The remaining genes had less than 8 gRNAs due to difficulty identifying 8 or 9 separate high-quality gRNAs. The non-targeting controls are not represented in the graph but are present with 250 gRNAs. Figure 9D shows the library read counts from two different sequencing runs plotted against each other. It shows that for the majority of gRNAs, the read counts are identical and hardly differ between two separate sequencing runs.

Analysis of the gRNA representation verified that the lymphoma-focused CRISPR library was adequately represented and allowed me to use this library for positive CRISPR screens in human GC B cells which will be described next.
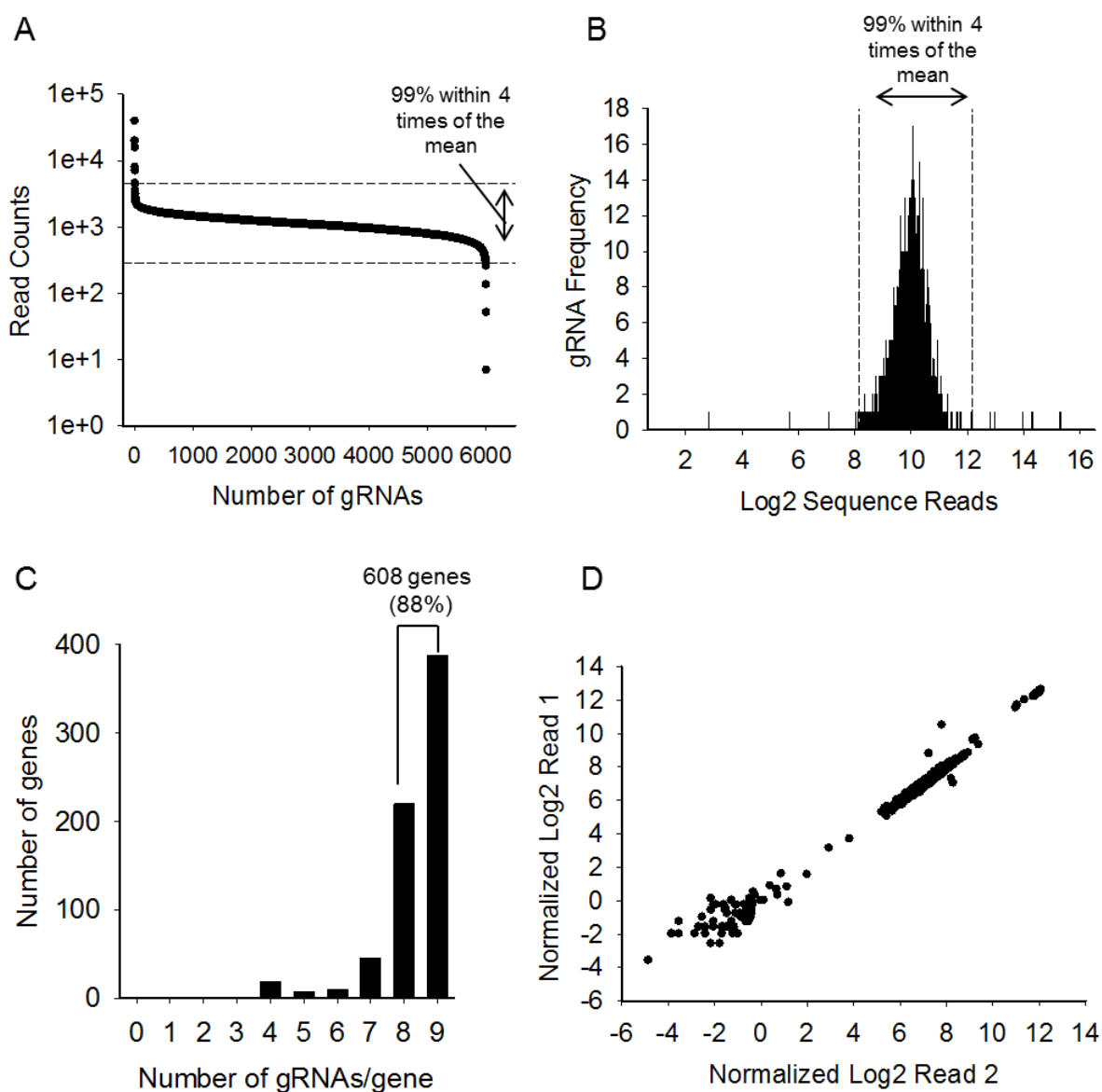
Figure 9 Successful construction of a lymphoma-focused CRISPR library with an adequate representation of all 6000 gRNAs

(A) Illumina sequencing revealed that 99% of read counts were within 4 times of the mean. Illustrated are the CRISPR library read counts vs number of gRNAs.

(B) Illumina sequencing revealed that 99% of sequence reads were within 4 times of the mean. Illustrated is the gRNA frequency vs log2 sequence reads.

(C) Illumina sequencing revealed the number of gRNAs present per gene. The non-targeting control with 250 gRNA is not shown here.

(D) Illumina sequencing revealed the correlation of library read counts between two different sequencing runs.

4.5 Functional high-throughput screening of tumour suppressor gene mutations on a user-defined mutational background in human GC B cells

4.5.1 Positive CRISPR screen in GC B cells on a BCL2 and BCL6 background identified GNA13 as a potent tumour suppressor

I wished to screen tumour suppressor genes in DLBCL using the described co-culture system and the lymphoma-focused CRISPR library to identify these mutations that drive the malignant phenotype. To do so, I conducted a positive-selection CRISPR screen to identify genes, that when knocked-down, result in a growth advantage in GC B cells.

As most cases of DLBCL are not driven by the Epstein Barr Virus (EBV), only EBV negative GC B cells were used for all my experiments. Therefore, every tonsil donor was screened for their EBV status with a highly sensitive PCR that we have optimised[119] (Methods 3.2.3).

Isolated GC B cells were transduced with BCL2 and BCL6 and CAS9 was retrovirally introduced through several rounds of transduction until > 50% of cells expressed CAS9. Using the lymphoma-focused CRISPR library (as described in 4.4), I conducted an enrichment screen to identify genes whose depletion provide a survival advantage in combination with the oncogene platform (here, BCL2 and BCL6). The CAS9, BCL2, BCL6-expressing cell population was transduced with the lymphoma-focused CRISPR library containing 6000 gRNAs including 250 non-targeting controls. Two things had to be taken into consideration here: 1.) To avoid multiple gRNAs integrating into one cell, cells were infected at a low multiplicity of infection (MOI). Approximately 10-30% of the cell population were CAS9/gRNA expressing cells. This assured that the effect seen in a particular cell is due to a single gRNA and not multiple gRNAs ''piggy backing'' on the observed result. 2.) To make sure that every gRNA was adequately represented in the population, cells were infected at 1000 x coverage. These calculations took into account that only around 50-80% of the cell population expressed CAS9. For example, to get a 1000 x coverage with a library the size of 6000 and an estimated 30% of CAS9/gRNA expressing cells = 6000 (gRNAs) x 1000 (representation) x 2 (50% CAS9) x 3.3 (30% CAS9/gRNA expressing cells) = I infected 40×10^6 cells. This cell number was maintained throughout the entire screen to prevent genetic bottlenecking.

This was performed in three biological replicates (Donors KB, AT and AG) and a schematic of the CRISPR screen is shown in Figure 10A. Cells were allowed to proliferate for approximately 70 days and it was expected that cells expressing gRNAs against tumour suppressor genes would be selectively enriched in the cell population. The screen was designed to detect tumour suppressor genes and not oncogenes which would be selectively depleted from the culture. This is because at the outset of the experiment only 50% of cells express CAS9. Therefore, whilst there is ample room for gRNAs to enrich, it will be much harder to detect those that drop out.

After the gRNA library was introduced into the CAS9, BCL2, BCL6-expressing GC B cells, cells were harvested roughly every two weeks (Day 4, 14, 28, 57 and 70) for genomic DNA. To define the enrichment in the screen, abundance of gRNAs present at different timepoints in the cell population was measured from genomic DNA by PCR across the gRNA sequences using a two-step strategy. First, the gRNA sequences were amplified and then the appropriate adaptors (P5 and P7) added for next-generation sequencing. In order to sequence multiple libraries in one sequencing run, I also added a barcode at the 3'end. Following next-generation sequencing, the libraries were de-multiplexed and the abundance of gRNAs in the plasmid pool was compared. The number of cells being harvested at each timepoint was dependent on the percentage of cells that were double positive for CAS9 and the library. This was important to make sure every gRNA is captured in the cell population. Hence, the number of cells collected was 1000 times more complex than the gRNA library, when considering only CAS9/gRNA expressing cells.

Genes were ranked according to their CRISPR gene score (Figure 10B), which is, in essence, the number of standard deviations away from the mean of all gRNAs targeting a specific gene (z-score). The z-score is a useful statistical tool, allowing me to compare scores from different populations/distributions. This will be helpful later when comparing different mutational platform screens. It also provides an exact position of a score in a normal distribution. A positive z-score indicates the data point is above the mean of the population whereas a negative z-score is below the mean. It also reveals the number of standard deviations a data point is above or below the mean.

Illumina sequencing revealed some of the most enriched genes to be, as expected, established tumour suppressors such as *TP53* and *CDKN2A* (Figure 10B). Strikingly, in three replicated experiments, performed in cells from three different tonsil donors, *GNA13* presented with the most potent or among the most potent enrichment, being many log2 folds higher than for any other gene. For example, *GNA13* had a 3.7 log2-fold higher enrichment than *PTEN*, a well-known potent tumour suppressor. Even though this was designed as an enrichment screen and not a drop-out screen, genes that were depleted in the screen included *BCL6, MYC, BCL2, CD79B, CARD11, POU2AF1* – all known B cell oncogenes (Figure 10B). The 250 gRNAs against the non-targeting control had a CRISPR gene score of 0.22.

Enrichment of tumour-suppressor genes *GNA13, TP53, CDKN2A* was consistent between three different donors (KB, AT and AG). The CRISPR gene score looks at the average gRNA score for a particular gene, so I next analysed the gRNA score for each individual gRNA. Wang stated that a gRNA "scores" if the z-score is more than 2 [162]. Figure 10C shows the gRNA score for genes that had more than 2 gRNAs significantly enriched (gRNA score >2) and circles represent gRNAs for the indicated gene. Remarkably, 8 out of 9 gRNAs against *GNA13* had enriched significantly at Day 70 and 6 out of 8 and 5 out of 9 for *TP53* and *CDKN2A*, respectively. Notably, 9 other genes presented with 2 or 3 gRNAs with a score of > 2. This highlighted the variability of biological processes that are altered in this experiment, including cell adhesion (*KANK2* and *ACTG1*), transcription (*TP63*), signalling (*NFKB2, PTEN, GSK3B*), cell cycle (*PTPRO),* chromatin remodelling (*ATRX*) and mitochondrial function (*VPS13C*).

From this list, only *NFKB2* and *PTEN* were listed as recurrently altered genes in Reddy's[7] and Chapuy's[96] sequencing study. NF-κB signalling is essential for B cell growth and survival and in many B cell cancers, is found to be abnormally activated by genetic mutations, promoting lymphomagenesis[163]. However, in this study here, knock-down of *NFKB2*, the nuclear factor NF-kappa-B p100 subunit, is associated with a growth advantage. This may relate to the feeder system used, which includes strong CD40Lg (and thereby NFKB activation) or perhaps to inhibition of plasma cell differentiation. *PTEN*, the widely known negative regulator of PI3K signalling, is one of the most frequently mutated genes in human cancers, including GCB-DLBCL[164]. GSK3, glycogen synthase kinase 3, is a downstream target of AKT, involved in cell survival[165]. PTPRO is a protein tyrosine phosphatase and believed to function as a

tumour suppressor[166]. TP63 is a member of the TP53 family of transcription factors, exerting similar biological functions to those of TP53. However, unlike TP53, TP63 is rarely inactivated in cancer[167]. KANK2 and ACTG1 are both genes involved in cytoskeletal formation and found to be mutated in Burkitt lymphoma and primary central nervous system lymphoma, respectively[107,168]. ATRX is a chromatin remodelling factor and has been identified to cause alternative lengthening of telomeres (ALT) when mutated[169]. Lastly, loss of VPS13C causes mitochondrial dysfunction.

Next, I wished to examine whether the gRNAs for my top three hits, *GNA13, TP53* and *CDKN2A* had a gradual enrichment over time or an acute enrichment at a certain timepoint. I calculated the log2 fold enrichment, followed by division of the median scores of the 250 non-targeting controls for each gRNA for *GNA13, TP53* and *CDKN2A* over six different timepoints (Day 4, 14, 28, 57 and 70). It can be seen that from Day 14, gRNAs gradually enriched over time in donor KB (Figure 11). Consistency between the three donors can be seen in Figure 11. Figure 12 shows the same graph as in Figure 11 with gRNAs being connected by a line for clarity. In all three donors for *GNA13* and *TP53*, a single outlier gRNA can be seen with an enrichment many log2-folds lower than for the other gRNAs.

I next tested whether these *GNA13* gRNAs resulted in efficient knock-down of its protein. To do so, I transduced the CAS9 expressing lymphoma cell line HBL1-CAS9 with nine *GNA13* gRNAs, four *PTEN* gRNAs and four non-targeting control gRNAs that were used in the lymphoma-focused CRISPR library. A western blot after 10 days of transduction confirmed successful knock-down of 8 out of 9 *GNA13* gRNAs and all *PTEN* gRNAs in comparison to the NTC (Figure 13A). Interestingly, gRNA 6 which was the only gRNA that failed to enrich in the CRISPR screen in all three donors (Figure 11 and Figure 12) was highly toxic to HBL1 cells soon after transduction and was therefore not included in the western blot. This could also be seen in a competitive fitness assay where knock-down of gRNA 6 in HBL1 resulted in a growth disadvantage in comparison to non-targeting controls (n=4) (Figure 13B). The knock-down of the remaining *GNA13* and *PTEN* gRNAs showed no effect in cell growth. This supports the fact that the primary GC B cells system provides a more powerful tool to study genetic drivers in DLBCL than cell lines, where the required oncogenic pathways may already be activated through other mechanisms.

All 9 *GNA13* gRNAs are spread throughout the entire length of the gene with gRNA 6 targeting the gene in the switch domain, which is important for its GTP binding affinity. However, two other gRNAs also target the switch domain and therefore it is not clear how gRNA 6 caused cell death, but presumably due to off-target activity. To conclude, 8 out of 9 *GNA13* gRNAs and 4 out of 4 *PTEN* gRNAs were functional.

Here, I identified GNA13 as an unexpectedly potent tumour suppressor in human GC B cells on a BCL2 and BCL6 background.

A

CAS9, BCL2, BCL6
expressing GC B cells

Add gRNA library
(one per cell)

Day 0 → Day 4

Harvest cells
(4, 14, 28, 57, 70 days)

Expected effect of
tumour suppressor

Day 70

CRISPR Screen

Oncogene platform + gRNA library

BCL2 + BCL6   692 genes

$$\text{gRNA score} = \text{z-score}\left[\log 2\left(\frac{\text{Day 70 normalized gRNA count}}{\text{normalized plasmid library gRNA count}}\right)\right]$$

$$\text{CRISPR gene score} = \text{Average (gRNA score for gene)}$$

B

CRISPR gene score Day 70 vs Gene Rank

GNA13, TP53, CDKN2A, PTEN, S1PR2, P2RY8, ARHGEF12, CD79B, BCL6, CARD11, POU2AF1, BCL2, MYC, ARHGEF3, NTC

C

gRNA score Day 70

GNA13, TP53, CDKN2A, VPS13C, NFKB2, GSK3B, PTEN, KANK2, TP63, ACTG1, ATRX, PTPRO, Mean +/- SD 250 NTC

Figure 10 Enrichment CRISPR screen in primary GC B cells transduced with BCL2 and BCL6 identified GNA13 as a potent tumour suppressor

(A) Outline of experimental design and mathematical formulas used.

(B) Genes were ranked from highest to lowest according to their CRISPR gene scores at Day 70 (log2 scale). Selected tumour suppressor genes as well as oncogenes are highlighted in green and red, respectively. Everything above the horizontal line is positively enriched. CRISPR gene score for non-targeting control (NTC) (n=250) was 0.22. Representative of 3 experiments.

(C) gRNA scores for genes with more than 2 gRNAs with a gRNA score >2 are shown at Day 70 (log2 scale). Every circle represents a gRNA for the indicated gene. The mean of all 250 non-targeting controls (NTC) +/- Standard Deviation is shown.
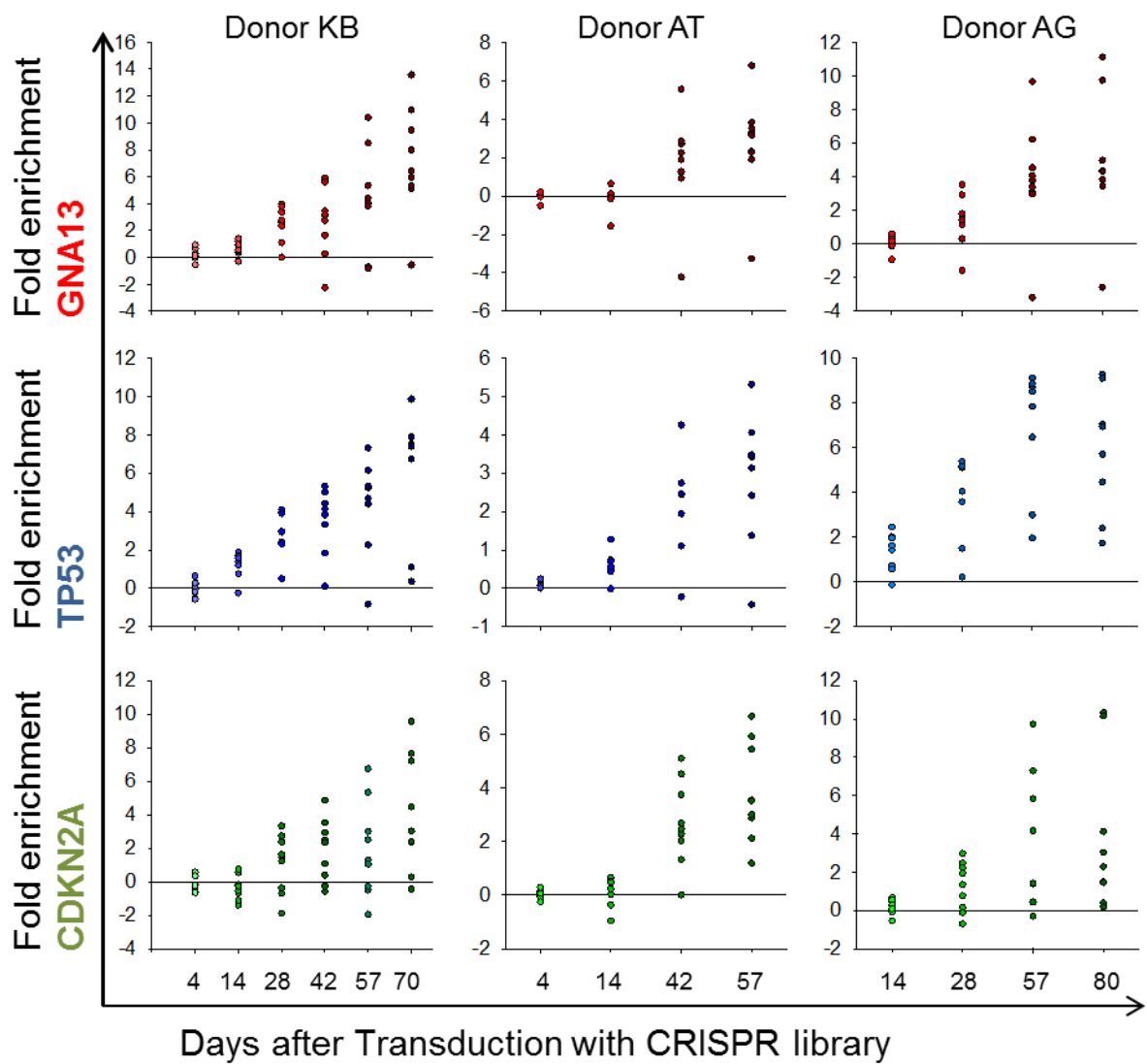
Figure 11 *GNA13, TP53* and *CDKN2A* are top hits in three different biological donors

Normalized log2 fold enrichment for *GNA13, TP53 and CDKN2A* is shown over different timepoints after transduction with the CRISPR library in three different donors (Donor KB, AT and AG). Circles represent individual gRNAs for the indicated gene and everything above the horizontal line is positively enriched.
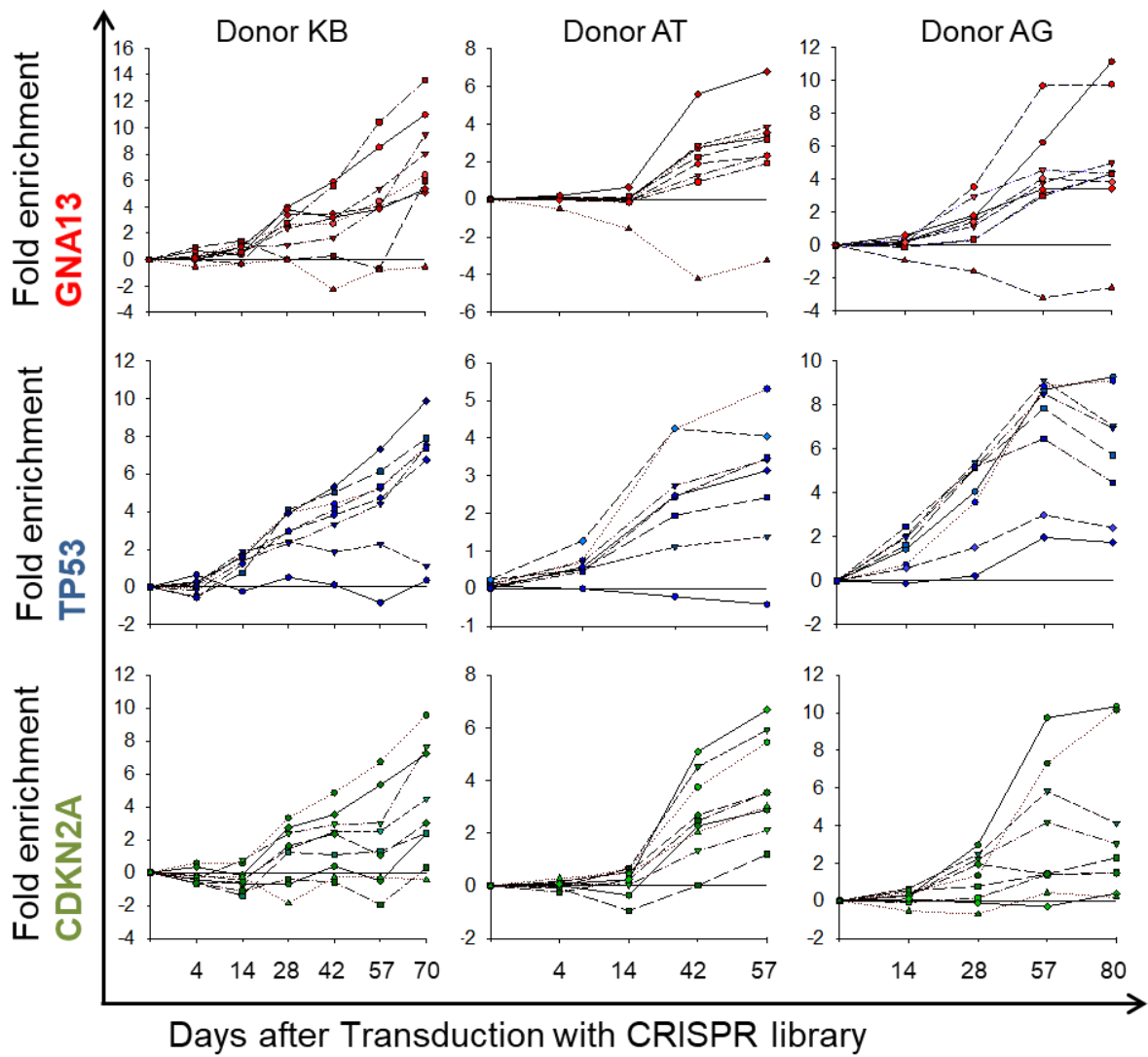
Figure 12 *GNA13, TP53* and *CDKN2A* are top hits in three different biological donors

Same as in Figure 11. Dots representing individual gRNAs are connected for clarity.
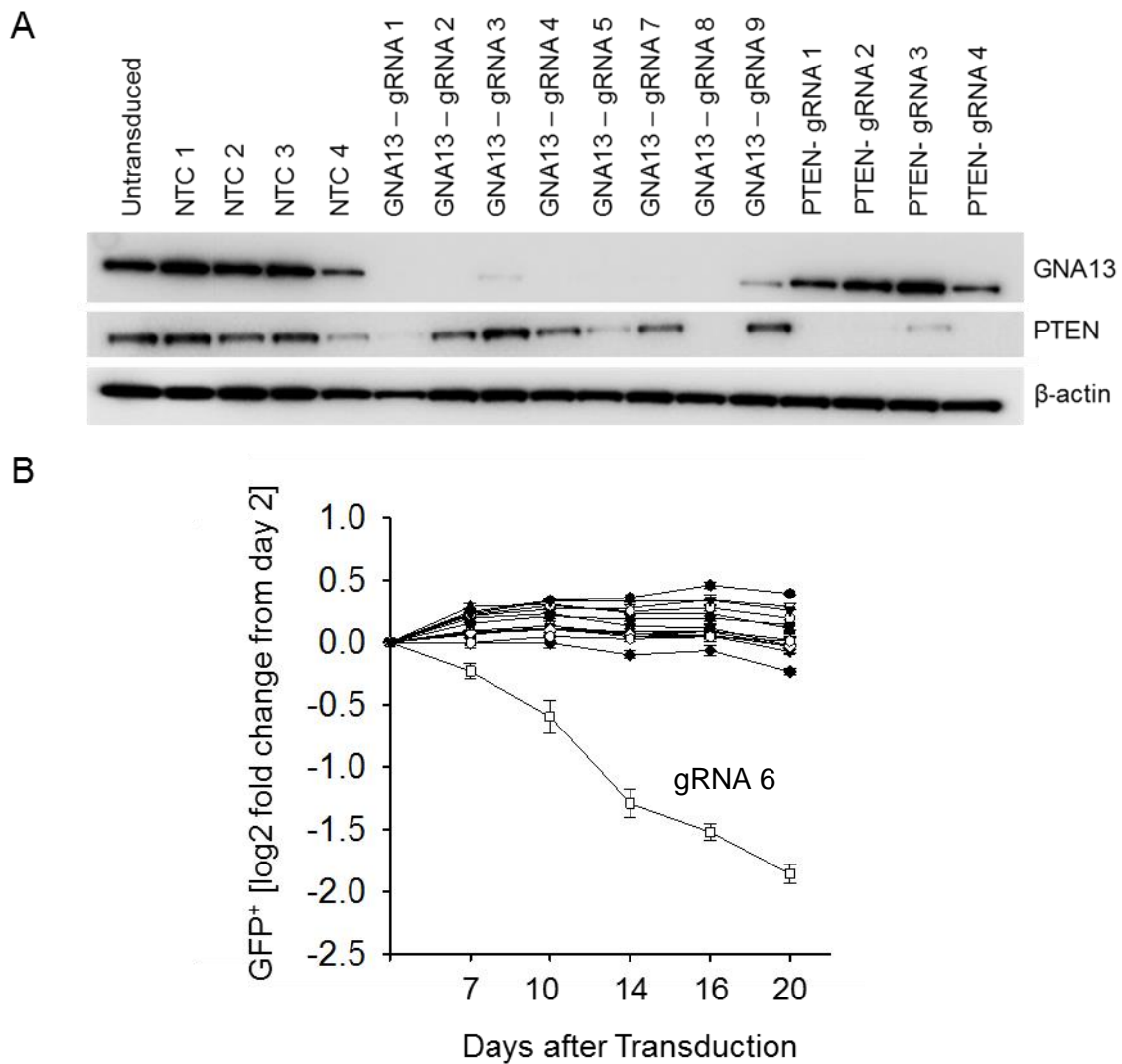
Figure 13 Validation of *GNA13* gRNAs functionality in the ABC-DLBCL cell line HBL1

(A) HBL1-CAS9 cells were transduced with 8 gRNAs against *GNA13,* 4 against *PTEN* and 4 against non-targeting control. Cells were harvested 10 days after transduction and a western blot performed to validate knock-down. β-actin was used as a loading control. Representative of > 3 experiments.

(B) HBL1 cells were transduced with 9 *GNA13*, 4 *PTEN* and 4 non-targeting control gRNAs and enrichment or depletion of GFP$^+$ cells was monitored by flow cytometry. Illustrated is time course showing the log2 fold-change relative to baseline (± Standard Deviation) of GFP$^+$ cells following transduction (n = 3). Representative of 3 experiments with 3 replicates/experiment.

4.5.1.1 GNA13 may provide a survival advantage to human GC B cells that is independent of its canonical functions

The identification of GNA13 as an unexpectedly potent tumour-suppressor prompted us to investigate its mechanism. GNA13 is a G protein alpha subunit downstream of the receptors S1PR2/P2RY8 and upstream of RhoA. When activated, GNA13 is known to have two mechanisms; 1.) Local confinement of germinal centre B cells and 2.) Growth regulation by suppressing phospho-AKT[102]. In our model system of cells being confined to a tissue culture flask, there is no potential for migration and hence, mechanism one seems unlikely to be the cause of the enrichment. Whilst the effect on pAKT signalling is an attractive explanation, it is pertinent that GNA13 had a considerably higher enrichment score than S1PR2, P2RY8 or PTEN (Figure 10B) despite them all having similar proposed effects on AKT phosphorylation.

To further investigate whether GNA13 knock-down results in an increase in phospho-AKT, we performed phospho-flow cytometry for the analysis of AKT activation in primary GC B cells. To do so, GC B cells were transduced with CAS9, BCL2, BCL6 and 9 individual gRNAs against *GNA13,* 4 against *PTEN* and 4 against non-targeting controls (NTC). PTEN served as a positive control as knock-down of PTEN, a negative regulator of the AKT pathway, will result in an increase in phospho-AKT levels. After gRNA transduction, intracellular staining of phospho-AKT (Ser473) was performed and analysed on a flow cytometer. FACS analysis showed, as expected, a robust increase in phospho-AKT levels for all 4 *PTEN* gRNAs in comparison to NTC when gated on double positive CAS9 and gRNA expressing cells (Figure 14). GNA13 knock-down, on the other hand, did not lead to an increase in phospho-AKT compared to NTCs. This was the case for all 9 *GNA13* gRNAs.

The experiment suggests that loss of GNA13 may provide a proliferative or survival advantage to human germinal center B cells through a mechanism independent of its canonical functions.
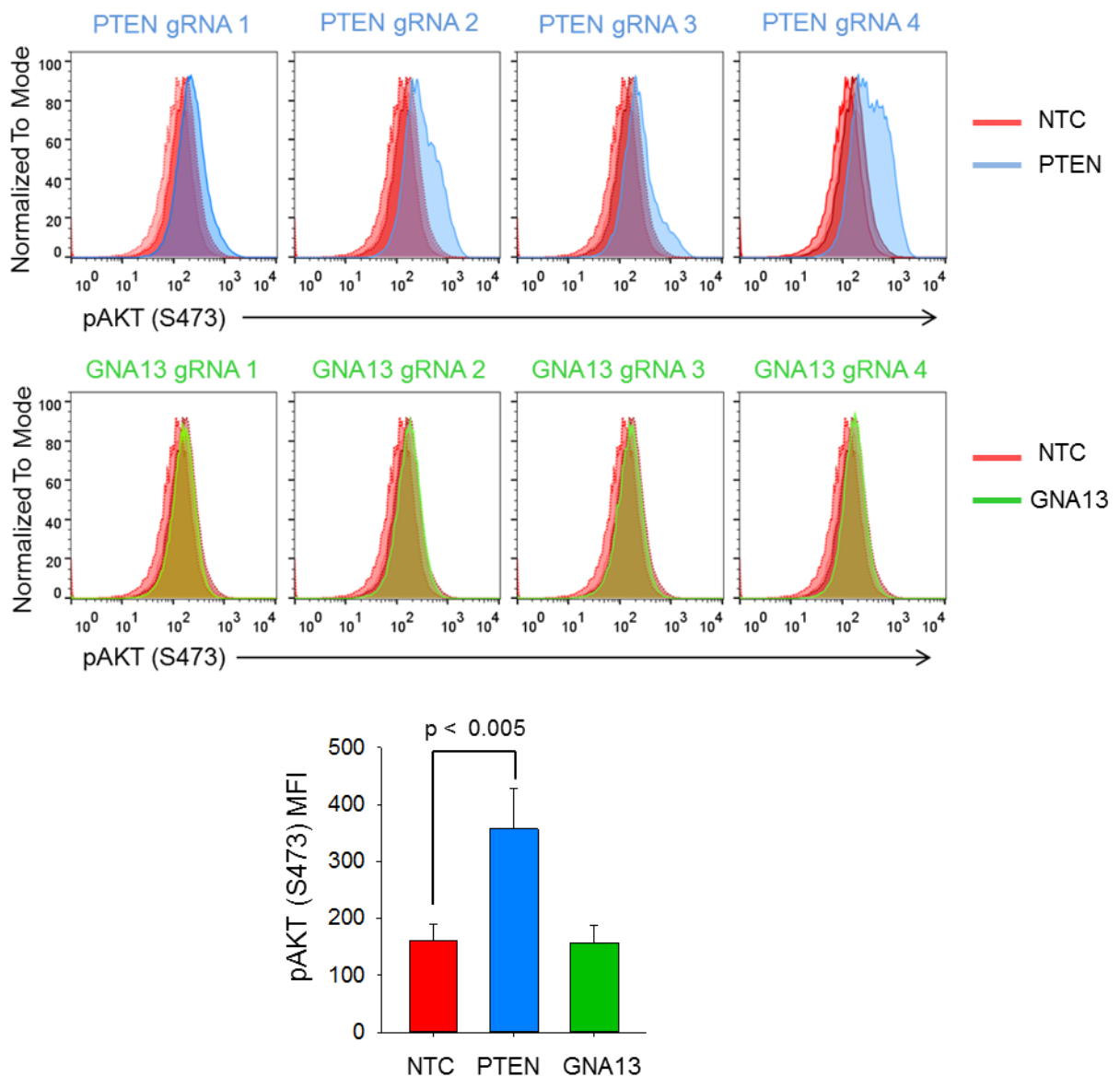
Figure 14 GNA13 knock-down does not activate AKT in primary GC B cells

Intracellular FACS staining for pAKT (S473) in GC B cells transduced with CAS9, BCL2, BCL6 and 4 gRNAs against *PTEN*, 9 gRNAs against *GNA13* and 3 gRNAs against non-targeting control (NTC). Shown are 4 *PTEN* gRNAs (Blue) against 3 NTC gRNAs (Red shades) and a representative example of 4 *GNA13* gRNAs (Green) against 3 NTC gRNAs. Barchart illustrates the mean fluorescence intensity of all gRNAs for the indicated gene (± SD). The p value was calculated from t test. This experiment was performed with Jie Gao, Hodson laboratory.

4.5.2 Positive CRISPR screen in GC B cells on a BCL2 and MYC background

To identify how different mutations cooperate with each other and identify potent tumour suppressor genes on a different mutational background, I transduced GC B cells with CAS9, BCL2 and MYC and further transduced them with the lymphoma-focused CRISPR library. The experiment was conducted as described in 4.5.1.

This screen was performed in three biological replicates (Donor AG, TADB and WK) and cells were allowed to proliferate for approximately 70 days. Illumina sequencing revealed the most enriched gene to be *TP53* followed by *CDKN1B* (Figure 15A). The 250 gRNAs against the non-targeting control have a CRISPR gene score of 0.05. With this mutational background, *GNA13* only presents with a CRISPR gene score of 1.3 instead of 4.9 which was seen on a BCL2 and BCL6 background. It is also interesting to note that the highest enriched genes on a BCL2 and BCL6 background have a CRISPR gene score several times higher than genes on a BCL2 and MYC background. Figure 15B shows that there are only 3 genes that have more than 2 gRNAs significantly enriched (gRNA score >2). 6 out of 8 gRNAs significantly enriched for *TP53*. 3 out of 6 and 3 out of 8 gRNAs were significantly enriched for *CDKN1B* and *ZFP36L1*, respectively.

Despite there being less enrichment than in the BCL2 and BCL6 background, several interesting results emerged from this screen. Members of the Zinc Finger Protein (ZFP36) family, such as ZFP36, ZFP36L1 and ZFP36L2 enriched with a CRISPR gene score of approximately 1.5. ZFP36 and its family members are RNA-binding proteins and control gene expression by regulating mRNA turnover[170]. It is known that MYC directly suppresses the tumour suppressor ZFP36 to induce tumourigenesis[170]. ZFP36 suppression is often seen in malignancies with MYC but enforced expression of ZFP36 impairs lymphomagenesis in an E-mu MYC mouse model[170].

The CRISPR screen on a MYC and BCL2 platform does not show as much enrichment as with the BCL2-BCL6 screen and suggests that MYC, in combination with BCL2, is already a powerful oncogenic hit with less "oncogenic space" to be filled by knock down of other tumour suppressors.
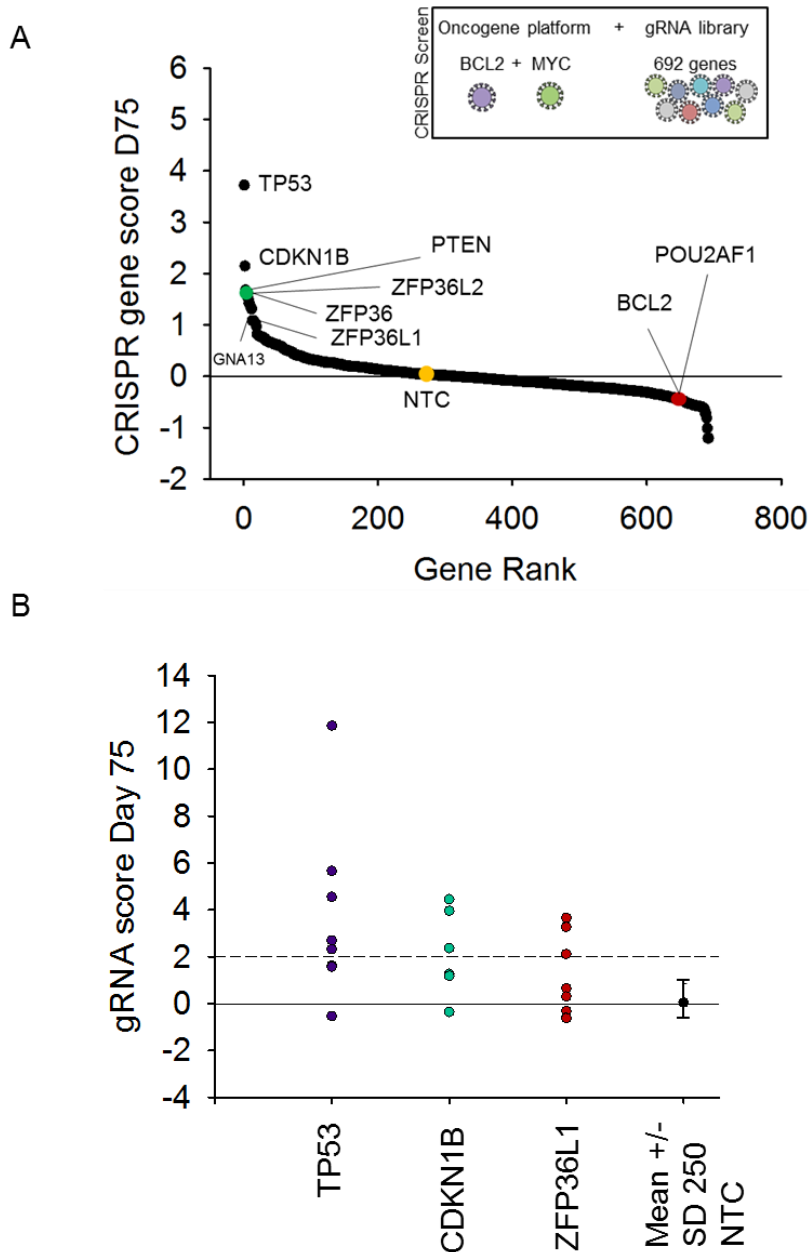
Figure 15 Enrichment CRISPR screen in primary GC B cells transduced with BCL2 and MYC

(A) Genes were ranked from highest to lowest according to their CRISPR gene scores at Day 75 (log2 scale). Selected tumour suppressor genes as well as oncogenes are highlighted in green and red, respectively. Everything above the horizontal line is positively enriched. CRISPR gene score for non-targeting control (n=250) is 0.05. Representative of 3 experiments.

(B) gRNA scores for genes with more than 2 gRNAs with a gRNA score >2 are shown at Day 75 (log2 scale). The mean of all 250 non-targeting controls ± Standard Deviation is shown.

4.5.3 Positive CRISPR screen in GC B cells on a background with MYC alone

I have shown a minimum of two oncogenic hits are necessary for long term survival of human GC B cells on the described feeder system. However, GC B cells transduced with one oncogene alone, such as MYC do not allow for long term growth in-vitro. Cells die after approximately two weeks in culture (4.1.3, Figure 4A). However, subsequently transducing these cells with the CRISPR library rescued this effect and cells were able to grow in culture for weeks.

I transduced GC B cells with CAS9 and MYC and further transduced them with the lymphoma-focused CRISPR library. The experiment was conducted as described in 4.5.1 unless otherwise stated here. CRISPR screens with a "one oncogene" platform were more difficult to manage. Initially, cells were slower at proliferation than when transduced with a "two oncogenes" platform. Therefore, cells could often not be harvested at a 1000x gRNA representation in the cell pool. This meant either skipping timepoints to allow for an increased cell number at a later timepoint or lowering the gRNA representation to 500x. Cells were never harvested at a representation less than 500x. The MYC-alone screen was performed in one tonsil donor (JJ) and cells were allowed to proliferate for approximately 40 days. The screen was stopped here as the GFP$^+$ percentage had already reached almost 100%. Illumina sequencing revealed some of the most enriched genes to be *NCOA1, NFKBIE* and *ZFP36L1* (Figure 16A). *ZFP36L1* scored before in the BCL2+MYC screen (4.5.2) and here, it significantly enriched with 4 out of 8 gRNAs. Interestingly, *ZFP36L1* has a higher CRISPR gene score here of 1.6 than in the BCL2+MYC screen (CRISPR gene score of 1). *ZFP36* and *ZFP36L2* did not enrich in the MYC alone screen here. The top hit *NCOA1* is a transcriptional activator and has histone acetyltransferase activity. Out of 8 gRNAs, it only enriched with 1 gRNA, suggesting this to be an off-target effect. *NFKBIE*, a known tumour suppressor, inhibits NF-κB signalling but again, only enriches with 1 out of 8 gRNAs. In this screen, *GNA13* and *TP53* were not among the most potent tumour suppressor genes and have similar CRISPR gene scores than in the BCL2-MYC screen (4.5.2).

As mentioned before, the CRISPR gene score takes into account the average of all the gRNA scores for a particular gene. Given that several genes in this screen, especially the top hits such as *NCOA1,* only had one gRNA significantly enriched; the question arises whether taking the median of gRNA scores for a particular gene is

more suitable in this case. This would avoid single values that are most likely off-targets effects, skewing the results. Hence, for this screen I also calculated the CRISPR gene scores according to the median of gRNA scores for a particular gene (Figure 16B). Using this formula, *NCOA1* and *NFKBIE* now have a CRISPR gene score similar to the non-targeting control. *ZFP36L1* emerged as the top hit followed by *SETD2*, which enriched with 3 out of 8 gRNAs. *SETD2* plays a role in chromatin modification and has been reported to be somatically mutated in DLBCL[80]. Using this formula for screens where top hits have most gRNAs enriched for a particular gene does not make a difference. For the BCL2-BCL6 screen described in 4.5.1, the results were very similar with *GNA13/TP53/CDKN2A* emerging as the top hits (data not shown). This suggests that taking into account the median for the CRISPR gene score might be more suitable for screens that emerge with several genes having only one highly enriched gRNA.

Interestingly, both formulas showed an enrichment of *ZFP36L1* which was also enriched in the BCL2-MYC screen (4.5.2). This suggests that *ZFP36L1* and *MYC* might be cooperating mutations resulting in a survival advantage regardless of BCL2 status.
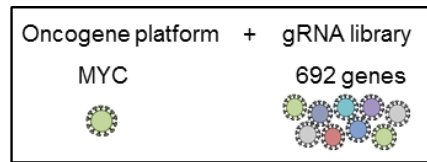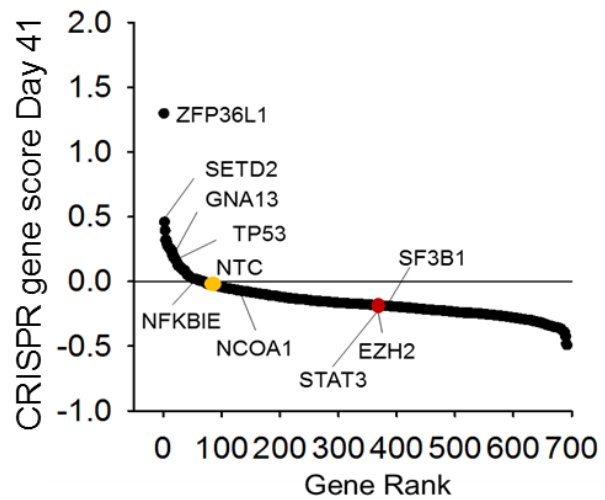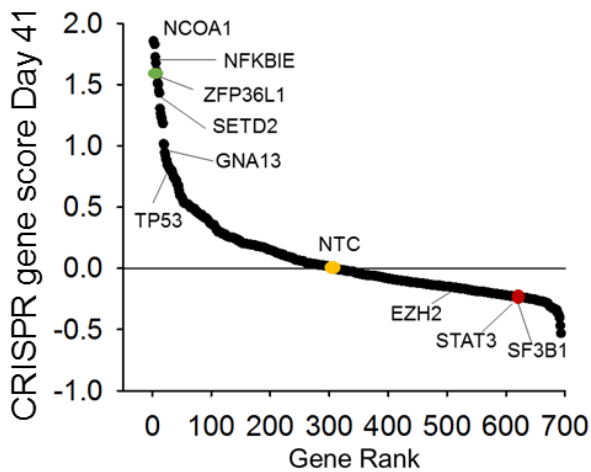
Figure 16 Enrichment CRISPR screen in primary GC B cells transduced with MYC alone

GC B cells transduced with CAS9, MYC and the CRISPR library. CRISPR gene scores were calculated using either the average of all gRNA scores for a particular gene (A) or the median of all gRNA scores for a particular gene (B) at Day 41 and ranked from highest to lowest (log2 scale). Selected tumour suppressor genes as well as oncogenes are highlighted in green and red, respectively.

4.5.4 Positive CRISPR screen in the ABC-DLBCL cell line HBL1 shows that the primary GC B cells system is a more powerful tool for the functional characterization of tumour suppressor, driver genes

As described earlier, cell lines represent the late stages of lymphoma, carry a high mutational burden and have often evolved in culture for many years. Knock-down of tumour suppressor genes in cell lines therefore seems unlikely to render them "more oncogenic" than they already are. For example, GNA13 knock-down in the ABC-DLBCL cell line HBL1 did not result in a survival advantage (Figure 13B) whereas in primary GC B cells it was the most potent tumour suppressor identified. I hypothesized that the defined genetics of the primary GC B cell system is superior to cell lines for screening potential tumour suppressor genes. To test this, I performed a parallel, positive selection CRISPR screen in the cell line HBL1.

HBL1 stably expressing CAS9 was transduced with the CRISPR library and allowed to proliferate for 70 days, a timepoint similar to most other screens. This allowed me to compare these screens with each other. Here, the cell population is 100% CAS9$^+$ so only GFP$^+$ (CRISPR library) cells were monitored over time. Illumina sequencing revealed the most enriched gene to be *CBL*, an E3 ubiquitin-protein ligase with a CRISPR gene score of 2.2 (Figure 17). 5 out of 9 gRNAs for *CBL* had a gRNA score of >2, considering them significantly enriched. *CBL* was not listed as frequently mutated in the three large genomic studies[6,7,96] recently published. However, in Phelan's[171] CRISPR screen in HBL1, *CBL* ranked 23rd out of 19117 genes with a CRISPR screen score of 1.37 at Day 21. Moreover, *CBLB*, a member of the *CBL* gene family, was among the 150 most frequently mutated genes in Reddy's[7] 1001 DLBCL sequencing study. *CBL* was followed by *PTEN* with a CRISPR gene score of 1.8. To put this into context, the highest enriched gene with the BCL2-BCL6 platform (*GNA13*) had a CRISPR gene score of 4.9 at Day 70 (Figure 10B). This is more than twice as high as for the highest enriched gene in HBL1 at the same timepoint. Moreover, in a cell line, *GNA13* shows up with a CRISPR gene score of 0.7 at Day 70. This is 7 log2-fold less than that seen for *GNA13* in the BCL-BCL6 screen.

Moreover, it is noteworthy that in the cell line experiment, most genes (427 out of 692 genes) have a CRISPR gene score of < 0. In contrast, the majority of genes in screens performed in primary GC B cells had a CRISPR gene score of > 0. This confirms that the primary GC B cell system is a more powerful tool than cell lines to screen potential tumour suppressor genes in lymphoma.

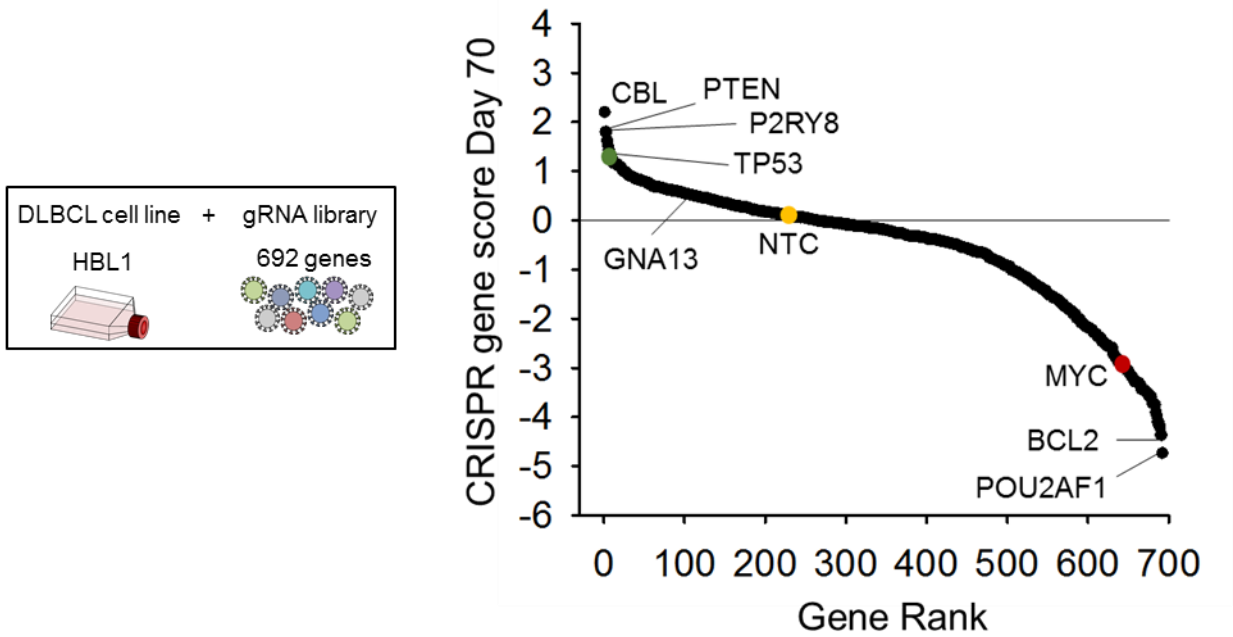**Figure 17 Enrichment CRISPR screen in the ABC-DLBCL cell line HBL1**

HBL1 expressing CAS9 was transduced with the CRISPR library. Genes are ranked from highest to lowest according to their CRISPR gene scores at Day 70 (log2 scale). Selected tumour suppressor genes as well as oncogenes are highlighted in green and red, respectively. Everything above the horizontal line is positively enriched.

5 Discussion

Despite DLBCL being heterogenous clinically, morphologically and molecularly, patients are currently grouped together as a single disease and receive the same treatment. This heterogeneity has been highlighted by recent mutation analysis papers that now suggest at least several genetic subtypes. What these genetic subtypes mean for the selection of optimal therapy remains unclear. The development of novel preclinical models of lymphoma, and its genetic subtypes is paramount at this stage. The model system described here has the potential to be a starting point for functional high-throughput screening of tumour suppressor gene mutations on a user-defined mutational background. An unlimited number of genetic combinations could be tested with this system. This will help to better understand the importance of the numerous somatic mutations in DLBCL and subsequently their implications for treatment. Moreover, as shown, this system also provides a reliable model of disease *in vivo*. I show that injecting human GC B cells transduced with oncogenes into immunodeficient mice is feasible and an effective tool to generate genetically customized lymphoma models. There are currently many new drugs in development and testing all of these, in combinations, in patients and in an empiric way will not be possible. Pre-clinical model systems to nominate the most effective and safest drug combinations are needed and I suspect this to be a promising system that might aid drug development in the future.

It was striking to see the rapid tumour onset in mice injected with human GC B cells that were transduced with MYC, BCL2, P53dd and CCND3 T283A. Replacing MYC with BCL6 in this group slowed down tumour onset of injected mice by 76 days (median time). This group presented with a similar tumour onset time to mice injected with MYC, BCL2, P53dd. It suggests that MYC is a strong oncogene that is necessary for tumour formation. However, the combination of MYC and CCND3 T283A resulted in a more aggressive phenotype. The oncogene platform BCL6, BCL2, P53dd only formed tumours when CCND3 T283A was added. This highlights the role of CCND3 T283A in DLBCL lymphomagenesis. Its highly important role in Burkitt lymphoma is already known[107]. Furthermore, I have collected viable cells, cell pellets and frozen tumour pieces from each mouse in order to perform a comprehensive biochemical analysis such as western blotting for downstream genes

and signalling pathways and RNA-sequencing.

Exploiting the unlimited number of different mutational backgrounds that can be introduced into human GC B cells, I also created a lymphoma model expressing BCL2, BCL6 and MYC. Injecting human GC B cells transduced with these oncogenes into immunodeficient mice resulted in a median survival time of 108 days (data not shown). These tumours stained positive for B cell markers (CD38, CD19 and CD10). The *in vivo* work described here reinforces the biological relevance of this system and allows for user-defined mutational human lymphoma models.

Moreover, I hypothesized that the selective pressures *in vivo* and *in vitro* are different. The enrichment of GNA13 *in vitro*, which was described in this study, might not reflect of what is needed for tumour formation *in vivo*. Therefore, I performed *in vivo* positive selection CRISPR screens to identify those gRNAs associated with tumour formation. I performed screens on a BCL2-MYC and BCL2-BCL6 platform and these human GC B cells were injected into immunodeficient mice seven days after transduction with the CRISPR library (n = 4 per cohort). Some mice have started to form tumours and subsequent work will focus on analysing these and identifying those gRNAs that were needed for tumour formation. Mice with only the BCL2-MYC and BCL2-BCL6 platform alone were also injected, serving as a control. It might be necessary to redesign the library and remove any especially potent tumour-suppressing genes such as *TP53* that would dominate tumour formation and mask weaker tumour suppressor genes.

In this study, the guanine nucleotide-binding protein subunit alpha-13 (GNA13) demonstrated unexpected potency as a genetic driver in DLBCL. These subunits are part of so-called guanine nucleotide-binding proteins (G proteins) that function as signal transducers[172]. G proteins and their G protein coupled receptors (GPCRs) represent the largest class of cell-surface molecules[173]. GPCRs are crucial for transducing signals involved in multiple processes such as cell growth, proliferation, survival or motility[173]. Once an extracellular ligand binds to a GPCR, it undergoes conformational change and initiates signal transduction[173]. It does so by coupling to and acting as a guanine nucleotide exchange factor (GEF) to a GDP-bound heterotrimeric G protein consisting of α, β and γ subunits[173]. The Gα subunit can be divided into four main families: $G\alpha_s$, $G\alpha_i$, $G\alpha_q$ and $G\alpha_{12}$ (including $G\alpha_{13}$ which is

encoded by *GNA13*)[173]. GPCRs are overexpressed in many different cancer types with Gα$_s$ being the most mutated G protein in cancer[172]. GNA13 is reported to be highly upregulated in several different cancers such as hepatocellular carcinoma[174], prostate cancer cells[175] and breast cancer cells[176]. High *GNA13* mRNA levels, but interestingly not *GNA12*, are associated with poor survival and metastases in patients with head and neck, ovarian, lung and gastric cancers[177]. In head and neck squamous cell carcinoma (HNSCC), GNA13 induces drug resistance and modulates a Tumour-Initiating Cell (TIC) phenotype *in vitro* and *in vivo* via NFκB and MAPK signalling pathways[177]. Blockade of GNA13 or downstream pathway effectors with small molecule inhibitors re-sensitized cells to chemotherapy[177]. These findings make GNA13 a potential prognostic biomarker for tumour progression in many solid tumours. Understanding the mechanistic action of GNA13 is critical to identifying how it might be targeted therapeutically. GNA13 and GNA12 signal through multiple key signalling pathways such as Rho, c-jun N-terminal kinase (JNK), extracellular signal-regulated kinase (ERK), p38, ERK5/6 or NFκB but these events have yet to be fully elucidated[177,178].

Although high expression of G proteins and GPCRs have been linked to poor outcome in solid tumours, my data and other studies[102,179,180] suggest that in lymphoma, GNA13 has a tumour suppressor role. Inactivating mutations of *GNA13* are found in approximately 25% of GCB-DLBCL and 30% of Burkitt Lymphoma (BL)[11]. The mechanism by which GNA13 mutations promote lymphomagenesis is not fully elucidated. Studies have shown that GNA13 antagonizes AKT phosphorylation, which protects against cell death, and promotes germinal centre confinement[102,180]. I did not see this in my experiments. In my model system, I am unable to observe how loss of homing to the germinal centre might provide a competitive advantage. If GNA13 worked through the AKT pathway, I would have expected for PTEN to also score at least as strongly in the enrichment screen. This was not the case (Results 4.5.1). This was further confirmed by intracellular staining for phospho-AKT (Ser473) in GNA13 depleted human GC B cells transduced with BCL2 and BCL6. Whilst phospho-AKT increased in PTEN-deleted GC B cells, AKT signalling did not significantly change in GNA13 depleted cells (Results 4.5.1.1). This was supported by Stelling's study that showed TGF-β -induced S1PR2 expression in SU-DHL-6 had no effect on AKT activation, as assessed by western blotting for phospho-AKT (Ser473)[181]. These findings suggest that GNA13 might work through a mechanism

independent of its canonical functions in lymphoma. RhoA, downstream effector of GNA13, is also found to be recurrently mutated in DLBCL and BL but surprisingly, I did not see enrichment of RhoA gRNAs. In my system, deletion of S1PR2 and P2RY8, both part of the Gα$_{13}$/RhoA axis, did not provide a growth advantage as strongly as GNA13. Performing a phosphoproteomics experiment in BCL2/BCL6-transduced GC B cells to elucidate which pathway(s) are targeted by GNA13 in my system would be highly interesting and I am currently pursuing this. Whether a S1PR2/P2RY8 agonist in GNA13 depleted BCL2/BCL6 transduced GC B cells would rescue the effect and result in a growth disadvantage, would he highly informative too. This approach is rather counterintuitive to treatment strategies in solid tumours where GNA13 is upregulated and treatment aims to inhibit the GNA13 pathway.

Furthermore, the potency of GNA13 as a tumour suppressor seems to be limited to the BCL2 and BCL6 platform. Whilst GNA13 did enrich in the BCL2 and MYC screen with a CRISPR gene score of 1.3 at a comparable timepoint, it was enriched 3.8 log2-fold higher in the BCL2-BCL6 transduced GC B cells. I suspect this is because the BCL2-BCL6 platform mimics more of a GCB-type DLBCL and GNA13 mutations are frequently found in GCB-DLBCL whereas the BCL2-MYC platform resembles more of an ABC-like DLBCL. Calado showed that MYC positive mature GC B cells had a phenotype of recently activated lymphocytes, which displays a later stage of development[182]. Moreover, Healy reported that GNA13 loss in combination with MYC overexpression resulted in lymphoma development in mice[102]. Additionally, loss of S1PR2 in mice harbouring a c-MYC transgene accelerated lymphomagenesis compared to S1PR2$^{+/+}$ c-MYC transgenic mice[181]. This discrepancy might be due to different selective pressures in vivo and in vitro.

Furthermore, whilst the CRISPR screen on a BCL2-MYC background did not show as much enrichment as with the BCL2-BCL6 background, it was interesting to see the ZFP36 family members enrich specifically with BCL2-MYC and ZFP36L1 also with MYC alone as a platform. As mentioned, ZFP36 and its family members ZFP36L1 and ZFP36L2 regulate gene expression by controlling mRNA turnover[170]. These AU-binding proteins (AUBPs) do so by marking AU-rich elements (AREs), found within many transcripts, for deadenylation and degradation by mRNA decay enzymes[183]. Interestingly, MYC is known to target AUBPs such as ZFP36 and directly suppresses its transcription which is often seen in cancers with MYC involvement[170]. It was therefore interesting to see a similar result in my CRISPR

screen; knock-down of ZFP36, ZFP36L1 and ZFP36L2 resulted in a survival advantage in BCL2-MYC transduced GC B cells. Reddy[7] and Schmitz[6] also identified ZFP36L1 to be recurrently mutated in DLBCL but this was not reported in Chapuy's[96] study. Numerous studies have shown that the ZFP36 protein family is associated with multiple other malignancies. ZFP36 is a negative regulator of several proinflammatory cytokines such as TNF-α and a decrease in ZFP36 expression can therefore result in the development of immune-related malignancies such as ulcerative colitis or rheumatoid arthritis[184]. Reduced ZFP36 expression levels are reported in aggressive prostate and breast cancer and predict poor outcome[185]. In addition, inactivation of ZFP36L1 and ZFP36L2 in mice during thymopoiesis leads to T cell acute lymphoblastic leukemia[186]. A study by Campbell and colleagues highlighted increased rates of inactivating mutations in ZFP36L1 and ZFP36L2 in several cancers such as bladder, breast and colorectal cancer[187]. These findings highlight that the ZFP36 family members are implicated in tumourigenesis and suggest a tumour suppressor role.

Moreover, positive CRISPR screens performed on single platform backgrounds such as MYC were only performed using one biological donor and will have to be repeated. Therefore, I am cautious to interpret too much into these screens. What I can conclude is that human GC B cells transduced with BCL6 alone, MYC alone or BCL2 alone die rapidly but adding a second oncogene or the CRISPR library to these cells allows for long term growth. A minimum of two oncogenic hits are necessary for long term survival of human GC B cells grown on the feeder system described here.

I recognise that the screens performed here are not exhaustive. GC B cells were always grown on Follicular Dendritic Cells (FDC) expressing CD40Lg and IL21, though necessary for the survival and proliferation of GC B cells, our results may be limited to this feeder set up. It is possible that FDCs expressing other B cell activation factors like BAFF would give slightly different results in the positive CRISPR screens. Therefore, future screens might withdraw components of the feeder system such as CD40Lg or IL21 and include other components such as BAFF. The huge cell numbers needed for the screens (due to only a small number of cells being infected with CAS9 and the CRISPR library) did not allow performing screens in parallel on different cytokine expressing feeders. Since the design of my lymphoma-focused CRISPR library targeting 692 genes, three large sequencing studies[6,7,96] have narrowed down these recurrent DLBCL driver genes to approximately 150. A library

designed today could perhaps restrict to only these 150 driver genes (as well as non-targeting controls). While having a bigger library like the one used in this study is not necessarily bad, a smaller library makes experiments more flexible in terms of number of cells to transduce and allows for more screens being conducted at the same time. Moreover, in a big library, weaker tumour suppressor genes could be masked by stronger ones such as TP53. Therefore, an alternative approach could be to have two individual libraries to separate strong and weak tumour suppressor genes from each other, similar to the approach used in Kim's mutant ORF screen[188]. Another limitation is that screens on different mutational backgrounds should ideally be performed using the same tonsil donor where GC B cells were isolated from. It is often difficult to get big cell numbers from tonsils so performing more than one screen on cells from a donor is often not feasible.

Whilst I have isolated GC B cells from tonsils using a MACS human B cell negative selection kit as well as anti-IgD and anti-CD44 and consider them GC B cells when they are CD38, CD20, CD19 and CD10 positive, I recognise that there are other approaches that research groups have used in the past. Klein and colleagues also used magnetic cell separation to isolate GC B cells and distinguished between centroblasts and centrocytes[189]. Purified centroblasts were $CD77^+$ and $CD38^{high}$ and centrocytes were $CD10^+$, $CD38^{high}$, $CD77^-$, $CD39^-$ and $CD3^-$ [189]. Kwakkenbos and colleagues considered tonsil germinal centre cells to be $CD38^+CD20^+$ [133]. They later reported that mature B cells can be cultured with CD40Lg and cytokines like IL-4, IL-10 and IL-21 which activate STATs, especially STAT3 and STAT5[134]. STAT3 signalling, in particular, results in plasma cell differentiation and a block in proliferation by increasing BLIMP1 and XBP1 expression[134]. Similarly, IL21 also induces expression of these two transcription factors. Introducing BCL6 into these cells cultured with IL-4, IL-10 or IL-21 inhibits terminal differentiation into plasma cells but many cells are dying[134]. Expressing the antiapoptotic factor BCL-XL rescued these cells from cell death and allows cells to expand rapidly[134]. The most robust proliferation of BCL6/BCL2-transduced cells was seen with the CD40Lg/IL-21 culture system [134] and I therefore used this culture system set up for my experiments.

Apart from the BCL2 and BCL6 background that allows for rapid proliferation, I have also used a BCL2 and MYC background as described in Results 4.5.2. The study from Kwakkenbos suggested that initially, cells with a BCL2-MYC background would proliferate rapidly but then undergo terminal differentiation and cell cycle arrest[126].

Interestingly, I did not observe this phenomenon. BCL2 and MYC-transduced cells proliferated for months without undergoing cell cycle arrest. Immunohistochemistry showed that this background in combination with P53dd still expressed the B cell marker CD20 (Results 4.2). CD20 is not seen in plasma cells and suggests that the BCL2/MYC platform does not generate a plasma cell-like phenotype. Interestingly, I did observe a phenotype for the two platforms used. BCL2/BCL6-transduced GC B cells seem to be feeder dependent; cells tightly attach to feeders and are generally not in suspension. On the other hand, BCL2/MYC-transduced GC B cells seem to be feeder independent; they are mostly in suspension. It would be interesting to see whether after transduction with BCL2 and MYC, they still maintain a high proliferation rate without the help of the feeder system. I noticed that cells isolated from mouse tumours (as described in Results 4.2) and put back in culture with feeders and without feeders did not show a difference in viability and proliferation rate.

Comparing the positive CRISPR screens performed in human GC B cells to the screen in the ABC-DLBCL cell line HBL1 shows that the primary GC B cell system provides a more powerful tool for the functional prioritisation of driver genes. The CRISPR gene scores of the small number of genes that did enrich in the cell line screen were not comparable to what was seen in human GC B cells. The highest enriched gene with the BCL2-BCL6 platform (*GNA13*) had a log2 CRISPR gene score more than twice as high as for the highest enriched gene in HBL1 at a similar timepoint. Most genes in the cell line screen were depleted suggesting that cell lines might be more suitable for drop out screens instead of enrichment screens. The dropout of known oncogenes such as *BCL2* and the enrichment of known tumour suppressor genes such as *PTEN* confirm the CRISPR-CAS9 system to be functional in this screen.

Interestingly, many genes that are known to be recurrently mutated in DLBCL, such as *CREBBP* or *EP300* did not enrich in my CRISPR screens described here. Inactivating mutations and deletions of *CREBBP* and less frequently *EP300* are reported in approximately 30% of DLBCL cases and 60% of FL[11,98,145]. Mutations in *CREBBP* and *EP300* are mainly monoallelic, mutually exclusive and impair its ability to acetylate histones as well as transcriptional activators such as *TP53* and acetylate and inactivate *BCL6*[98]. Interestingly, *CREBBP* deletion alone in pro-B cells or GC B cells is insufficient for the development of lymphoma; however, when combined with BCL2 deregulation, accelerated tumour formation[190]. This suggests that additional

oncogenic events are needed in *CREBBP* deleted/mutated cells to undergo clonal expansion[190]. My screens were performed on a BCL2 and either BCL6 or MYC background and this shows, that this cannot be the limiting factor as to why I did not see enrichment of *CREBBP/EP300* gRNAs in my CRISPR screens. However, one reason that I did not see enrichment for either *CREBBP* or *EP300*, might be that these two tumour suppressor genes have a potentially early role in tumour clonal expansion and do not promote lymphomagenesis if acquired later on in B cell development such as in the GC B cell[145,148,191]. For example, *CREBBP* deletion upon initiation of the GC resulted in a less pronounced phenotype than when *CREBBP* was deleted in early B cell development[190]. A similar phenomenon was also observed with *KMT2D,* a methyltransferase which is recurrently mutated in DLBCL and FL[11,148]. Deletion of *KMT2D* in mice during early B cell development led to an increase in GC B cells and increased B cell proliferation[99]. This was not observed when *KMT2D* was deleted after initiation of the GC reaction. Moreover, phylogenetic analysis has shown that *CREBBP* or *KMT2D* were found to be present in a precursor clone before FL progressed to DLBCL[145,148,191]. All my CRISPR screens were performed after initiation of the GC reaction and not in early B cells which is why deletion of *CREBBP/EP300* or *KMT2D* might not have an effect anymore. This suggests that these mutations are important in the early stages of DLBCL pathogenesis, so-called founder mutations and others in disease progression. Furthermore, *CREBBP* deletion impairs H3K27 enhancer acetylation and silences genes that are required for terminal differentiation and immune response in mature B cells[192]. Normally, *CREBBP* regulated enhancers are counter-regulated by BCL6/SMRT/HDAC3 complexes via H3K27 deacetylation, which bind to MHC class II loci[192]. However, when *CREBBP* is deleted, the complex is disturbed resulting in loss of immune surveillance by hiding from T-cells through silencing of MHC class II expression[192]. In my co-culture system, there is no immune surveillance as no T-cells or natural-killer cells are present. It is possible that in this setting, there is no need to hide from T cells so *CREBBP* deletion is not advantageous to GC B cells. This is consistent with the fact that none of the HLAs (A, B and C), encoding the MHC class I proteins or HLAs (DMA and DRB1) encoding MHC class II proteins showed an enrichment in my CRISPR screens.

The challenge with next generation sequencing being widely available is to correlate the emerging sequencing studies and carefully interpret them in terms of therapeutic

implications. I have discussed here three large genomic studies[6,7,96] that have gained remarkable insight into genetic drivers and different genetic subtypes that may respond differently to targeted therapy. However, these sequencing studies show discrepancies between each other in terms of what genes were identified to be most significantly mutated in DLBCL. The three large genomic studies converge on approximately 200 frequently mutated genes with many genes being unique to each study and not shared between them. These studies also propose different genetic subtypes with a different clinical outcome. How any of these subtypes can be translated into the clinic is currently unclear. Hence, pre-clinical model systems to capture these different genetic subtypes and test targeted treatment strategies are a major current research priority.

The primary culture system described here provides a technique to escape our reliance of cell lines. This study highlights its potential for functional high-throughput screening of tumour suppressor gene mutations on a user-defined mutational background. This information could be used to target relevant pathways in the context of individual tumours. I have demonstrated this by combining the novel primary culture system with the powerful tool CRISPR. This approach identified *GNA13* as a potent tumour suppressor, which therefore represents a potential role as a therapeutic target. It is paramount to understand the functional role of the vast amount of somatic mutations found in DLBCL and the implications for therapy before advances in treatment strategies can be made. I anticipate that the primary culture system, described here, provides a powerful tool that is complimentary to cell lines and mouse models for the analysis of DLBCL tumour biology and will contribute to the development of targeted therapy.

Bibliography

1   Schubbert, S., Shannon, K. & Bollag, G. Hyperactive Ras in developmental disorders and cancer. *Nature reviews. Cancer* **7**, 295-308, doi:10.1038/nrc2109 (2007).

2   Zhang, J. *et al.* Key pathways are frequently mutated in high-risk childhood acute lymphoblastic leukemia: a report from the Children's Oncology Group. *Blood* **118**, 3080-3087, doi:10.1182/blood-2011-03-341412 (2011).

3   Basso, K. & Dalla-Favera, R. Germinal centres and B cell lymphomagenesis. *Nature reviews. Immunology* **15**, 172-184, doi:10.1038/nri3814 (2015).

4   Shojaee, S. *et al.* Erk Negative Feedback Control Enables Pre-B Cell Transformation and Represents a Therapeutic Target in Acute Lymphoblastic Leukemia. *Cancer cell* **28**, 114-128, doi:10.1016/j.ccell.2015.05.008 (2015).

5   Mullighan, C. G. The molecular genetic makeup of acute lymphoblastic leukemia. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program* **2012**, 389-396, doi:10.1182/asheducation-2012.1.389 (2012).

6   Schmitz, R. *et al.* Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N Engl J Med* **378**, 1396-1407, doi:10.1056/NEJMoa1801445 (2018).

7   Reddy, A. *et al.* Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. *Cell* **171**, 481-494 e415, doi:10.1016/j.cell.2017.09.027 (2017).

8   Zhenfeng Zhang, B. H. *DUSP6 (dual specificity phosphatase 6)*, 2012).

9   Muschen, M. Autoimmunity checkpoints as therapeutic targets in B cell malignancies. *Nature reviews. Cancer* **18**, 103-116, doi:10.1038/nrc.2017.111 (2018).

10   Inc., S. B. *B cell CD antigens*, <https://www.sinobiological.com/b-cell-cd-antigens.html> (

11   Pasqualucci, L. & Dalla-Favera, R. The genetic landscape of diffuse large B-cell lymphoma. *Semin Hematol* **52**, 67-76, doi:10.1053/j.seminhematol.2015.01.005 (2015).

12   Klein, U. & Dalla-Favera, R. Germinal centres: role in B-cell physiology and malignancy. *Nature reviews. Immunology* **8**, 22-33, doi:10.1038/nri2217 (2008).

13   Bhojwani, D. & Pui, C. H. Relapsed childhood acute lymphoblastic leukaemia. *The Lancet. Oncology* **14**, e205-217, doi:10.1016/S1470-2045(12)70580-6 (2013).

14   Inaba, H., Greaves, M. & Mullighan, C. G. Acute lymphoblastic leukaemia. *Lancet* **381**, 1943-1955, doi:10.1016/S0140-6736(12)62187-4 (2013).

15   NHS.   *Acute   lymphoblastic   leukaemia*, <http://www.nhs.uk/conditions/leukaemia-acute-lymphoblastic/Pages/Introduction.aspx> (2015).

16   Pui, C. H., Mullighan, C. G., Evans, W. E. & Relling, M. V. Pediatric acute lymphoblastic leukemia: where are we going and how do we get there? *Blood* **120**, 1165-1174, doi:10.1182/blood-2012-05-378943 (2012).

17   Raetz, E. A. & Bhatla, T. Where do we stand in the treatment of relapsed acute lymphoblastic leukemia? *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program* **2012**, 129-136, doi:10.1182/asheducation-2012.1.129 (2012).

18   Teachey, D. T. & Hunger, S. P. Predicting relapse risk in childhood acute lymphoblastic leukaemia. *British journal of haematology* **162**, 606-620, doi:10.1111/bjh.12442 (2013).

19   Hardy, R. R. & Hayakawa, K. B cell development pathways. *Annu Rev Immunol* **19**, 595-621, doi:10.1146/annurev.immunol.19.1.595 (2001).

20   Pieper, K., Grimbacher, B. & Eibel, H. B-cell biology and development. *J Allergy Clin Immunol* **131**, 959-971, doi:10.1016/j.jaci.2013.01.046 (2013).

21   Melchers, F. Checkpoints that control B cell development. *J Clin Invest* **125**, 2203-2210, doi:10.1172/JCI78083 (2015).

22   Ren, W. *et al.* Surrogate light chain is required for central and peripheral B-cell tolerance and inhibits anti-DNA antibody production by marginal zone B cells. *Eur J Immunol* **45**, 1228-1237, doi:10.1002/eji.201444917 (2015).

23   Dhillon, A. S., Hagan, S., Rath, O. & Kolch, W. MAP kinase signalling pathways in cancer. *Oncogene* **26**, 3279-3290, doi:10.1038/sj.onc.1210421 (2007).

24   Knight, T. & Irving, J. A. Ras/Raf/MEK/ERK Pathway Activation in Childhood Acute Lymphoblastic Leukemia and Its Therapeutic Targeting. *Frontiers in oncology* **4**, 160, doi:10.3389/fonc.2014.00160 (2014).

25   Lodish H., B. A., Zipursky S L., Matsudaira P. , Baltimore D. , and Darnell J. *Molecular Cell Biology*. Vol. 4th edition (W. H. Freeman, 2000).

26   Ferrell, J. E., Jr. & Bhatt, R. R. Mechanistic studies of the dual phosphorylation of mitogen-activated protein kinase. *The Journal of biological chemistry* **272**, 19008-19016 (1997).

27   Boulton, T. G. *et al.* ERKs: a family of protein-serine/threonine kinases that are activated and tyrosine phosphorylated in response to insulin and NGF. *Cell* **65**, 663-675 (1991).

28   Lefloch, R., Pouyssegur, J. & Lenormand, P. Single and combined silencing of ERK1 and ERK2 reveals their positive contribution to growth signaling depending on their expression levels. *Mol Cell Biol* **28**, 511-527, doi:10.1128/MCB.00800-07 (2008).

29   Nekrasova, T. *et al.* ERK1-deficient mice show normal T cell effector function and are highly susceptible to experimental autoimmune encephalomyelitis. *J Immunol* **175**, 2374-2380 (2005).

30   Hatano, N. *et al.* Essential role for ERK2 mitogen-activated protein kinase in placental development. *Genes Cells* **8**, 847-856 (2003).

31   Courtois-Cox, S. *et al.* A negative feedback signaling network underlies oncogene-induced senescence. *Cancer cell* **10**, 459-472, doi:10.1016/j.ccr.2006.10.003 (2006).

32   Cox, A. D., Fesik, S. W., Kimmelman, A. C., Luo, J. & Der, C. J. Drugging the undruggable RAS: Mission possible? *Nat Rev Drug Discov* **13**, 828-851, doi:10.1038/nrd4389 (2014).

33   Downward, J. Targeting RAS signalling pathways in cancer therapy. *Nature reviews. Cancer* **3**, 11-22, doi:10.1038/nrc969 (2003).

34   Irving, J. *et al.* Ras pathway mutations are prevalent in relapsed childhood acute lymphoblastic leukemia and confer sensitivity to MEK inhibition. *Blood* **124**, 3420-3430, doi:10.1182/blood-2014-04-531871 (2014).

35   Freeman, M. Feedback control of intercellular signalling in development. *Nature* **408**, 313-319, doi:10.1038/35042500 (2000).

36   Pratilas, C. A. *et al.* (V600E)BRAF is associated with disabled feedback inhibition of RAF-MEK signaling and elevated transcriptional output of the pathway. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 4519-4524, doi:10.1073/pnas.0900780106 (2009).

37   Owens, D. M. & Keyse, S. M. Differential regulation of MAP kinase signalling by dual-specificity protein phosphatases. *Oncogene* **26**, 3203-3213, doi:10.1038/sj.onc.1210412 (2007).

38   Bagnyukova, T. V. *et al.* DUSP6 regulates drug sensitivity by modulating DNA damage response. *Br J Cancer* **109**, 1063-1071, doi:10.1038/bjc.2013.353 (2013).

39   Theodosiou, A. & Ashworth, A. MAP kinase phosphatases. *Genome Biol* **3**, REVIEWS3009 (2002).

40   Ahmad, M. K., Abdollah, N. A., Shafie, N. H., Yusof, N. M. & Razak, S. R. A. Dual-specificity phosphatase 6 (DUSP6): a review of its molecular characteristics and clinical relevance in cancer. *Cancer Biol Med* **15**, 14-28, doi:10.20892/j.issn.2095-3941.2017.0107 (2018).

41   Camps, M., Nichols, A. & Arkinstall, S. Dual specificity phosphatases: a gene family for control of MAP kinase function. *FASEB J* **14**, 6-16 (2000).

42   Keyse, S. M. Protein phosphatases and the regulation of mitogen-activated protein kinase signalling. *Curr Opin Cell Biol* **12**, 186-192 (2000).

43   Zhang, Z. *et al.* Dual specificity phosphatase 6 (DUSP6) is an ETS-regulated negative feedback mediator of oncogenic ERK signaling in lung cancer cells. *Carcinogenesis* **31**, 577-586, doi:10.1093/carcin/bgq020 (2010).

44   Molina, G. *et al.* Zebrafish chemical screening reveals an inhibitor of Dusp6 that expands cardiac cell lineages. *Nature chemical biology* **5**, 680-687, doi:10.1038/nchembio.190 (2009).

45   Karlsson, M., Mandl, M. & Keyse, S. M. Spatio-temporal regulation of mitogen-activated protein kinase (MAPK) signalling by protein phosphatases. *Biochemical Society transactions* **34**, 842-845, doi:10.1042/BST0340842 (2006).

46   Maillet, M. *et al.* DUSP6 (MKP3) null mice show enhanced ERK1/2 phosphorylation at baseline and increased myocyte proliferation in the heart affecting disease susceptibility. *The Journal of biological chemistry* **283**, 31246-31255, doi:10.1074/jbc.M806085200 (2008).

47   Urness, L. D., Li, C., Wang, X. & Mansour, S. L. Expression of ERK signaling inhibitors Dusp6, Dusp7, and Dusp9 during mouse ear development. *Dev Dyn* **237**, 163-169, doi:10.1002/dvdy.21380 (2008).

48   Frank, M. J. *et al.* Expression of sprouty2 inhibits B-cell proliferation and is epigenetically silenced in mouse and human B-cell lymphomas. *Blood* **113**, 2478-2487, doi:10.1182/blood-2008-05-156943 (2009).

49   Furukawa, T., Sunamura, M., Motoi, F., Matsuno, S. & Horii, A. Potential tumor suppressive pathway involving DUSP6/MKP-3 in pancreatic cancer. *The*

American journal of pathology **162**, 1807-1815, doi:10.1016/S0002-9440(10)64315-5 (2003).

50 Okudela, K. *et al.* Down-regulation of DUSP6 expression in lung cancer: its mechanism and potential role in carcinogenesis. *The American journal of pathology* **175**, 867-881, doi:10.2353/ajpath.2009.080489 (2009).

51 Chan, D. W. *et al.* Loss of MKP3 mediated by oxidative stress enhances tumorigenicity and chemoresistance of ovarian cancer cells. *Carcinogenesis* **29**, 1742-1750, doi:10.1093/carcin/bgn167 (2008).

52 Bloethner, S. *et al.* Effect of common B-RAF and N-RAS mutations on global gene expression in melanoma cell lines. *Carcinogenesis* **26**, 1224-1232, doi:10.1093/carcin/bgi066 (2005).

53 Warmka, J. K., Mauro, L. J. & Wattenberg, E. V. Mitogen-activated protein kinase phosphatase-3 is a tumor promoter target in initiated cells that express oncogenic Ras. *The Journal of biological chemistry* **279**, 33085-33092, doi:10.1074/jbc.M403120200 (2004).

54 Chen, H. Y. *et al.* A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med* **356**, 11-20, doi:10.1056/NEJMoa060096 (2007).

55 De Roock, W., Biesmans, B., De Schutter, J. & Tejpar, S. Clinical biomarkers in oncology: focus on colorectal cancer. *Mol Diagn Ther* **13**, 103-114, doi:10.2165/01250444-200913020-00004 (2009).

56 Weinstein, I. B. & Joe, A. K. Mechanisms of disease: Oncogene addiction--a rationale for molecular targeting in cancer therapy. *Nat Clin Pract Oncol* **3**, 448-457, doi:10.1038/ncponc0558 (2006).

57 Feil, R., Wagner, J., Metzger, D. & Chambon, P. Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. *Biochem Biophys Res Commun* **237**, 752-757, doi:10.1006/bbrc.1997.7124 (1997).

58 Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 5547-5551 (1992).

59 D'Souza, W. N., Chang, C. F., Fischer, A. M., Li, M. & Hedrick, S. M. The Erk2 MAPK regulates CD8 T cell proliferation and survival. *J Immunol* **181**, 7617-7629 (2008).

60   Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88**, 593-602 (1997).

61   Shvarts, A. *et al.* A senescence rescue screen identifies BCL6 as an inhibitor of anti-proliferative p19(ARF)-p53 signaling. *Genes & development* **16**, 681-686, doi:10.1101/gad.929302 (2002).

62   Duy, C. *et al.* BCL6 enables Ph+ acute lymphoblastic leukaemia cells to survive BCR-ABL1 kinase inhibition. *Nature* **473**, 384-388, doi:10.1038/nature09883 (2011).

63   Neel, B. G., Gu, H. & Pao, L. The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. *Trends Biochem Sci* **28**, 284-293, doi:10.1016/S0968-0004(03)00091-4 (2003).

64   Pelanda, R. & Torres, R. M. Central B-cell tolerance: where selection begins. *Cold Spring Harb Perspect Biol* **4**, a007146, doi:10.1101/cshperspect.a007146 (2012).

65   Lam, K. P., Kuhn, R. & Rajewsky, K. In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death. *Cell* **90**, 1073-1083 (1997).

66   Nemazee, D. Mechanisms of central tolerance for B cells. *Nature reviews. Immunology* **17**, 281-294, doi:10.1038/nri.2017.19 (2017).

67   Shojaee, S. *et al.* PTEN opposes negative selection and enables oncogenic transformation of pre-B cells. *Nat Med* **22**, 379-387, doi:10.1038/nm.4062 (2016).

68   Chen, Z. *et al.* Signalling thresholds and negative B-cell selection in acute lymphoblastic leukaemia. *Nature* **521**, 357-361, doi:10.1038/nature14231 (2015).

69   Hughes, T. P. *et al.* Frequency of major molecular responses to imatinib or interferon alfa plus cytarabine in newly diagnosed chronic myeloid leukemia. *N Engl J Med* **349**, 1423-1432, doi:10.1056/NEJMoa030513 (2003).

70   Wang, Y., Schmid-Bindert, G. & Zhou, C. Erlotinib in the treatment of advanced non-small cell lung cancer: an update for clinicians. *Ther Adv Med Oncol* **4**, 19-29, doi:10.1177/1758834011427927 (2012).

71   Zabriskie, M. S. *et al.* BCR-ABL1 compound mutations combining key kinase domain positions confer clinical resistance to ponatinib in Ph chromosome-

positive leukemia. *Cancer cell* **26**, 428-442, doi:10.1016/j.ccr.2014.07.006 (2014).

72   Samuels, I. S. *et al.* Deletion of ERK2 mitogen-activated protein kinase identifies its key roles in cortical neurogenesis and cognitive function. *J Neurosci* **28**, 6983-6995, doi:10.1523/JNEUROSCI.0679-08.2008 (2008).

73   Fremin, C., Saba-El-Leil, M. K., Levesque, K., Ang, S. L. & Meloche, S. Functional Redundancy of ERK1 and ERK2 MAP Kinases during Development. *Cell Rep* **12**, 913-921, doi:10.1016/j.celrep.2015.07.011 (2015).

74   Selcher, J. C., Nekrasova, T., Paylor, R., Landreth, G. E. & Sweatt, J. D. Mice lacking the ERK1 isoform of MAP kinase are unimpaired in emotional learning. *Learn Mem* **8**, 11-19, doi:10.1101/lm.37001 (2001).

75   Busca, R., Pouyssegur, J. & Lenormand, P. ERK1 and ERK2 Map Kinases: Specific Roles or Functional Redundancy? *Front Cell Dev Biol* **4**, 53, doi:10.3389/fcell.2016.00053 (2016).

76   Niu, H., Ye, B. H. & Dalla-Favera, R. Antigen receptor signaling induces MAP kinase-mediated phosphorylation and degradation of the BCL-6 transcription factor. *Genes & development* **12**, 1953-1961 (1998).

77   Cattoretti, G. *et al.* BCL-6 protein is expressed in germinal-center B cells. *Blood* **86**, 45-53 (1995).

78   Smith, A., Howell, D., Patmore, R., Jack, A. & Roman, E. Incidence of haematological malignancy by sub-type: a report from the Haematological Malignancy Research Network. *Br J Cancer* **105**, 1684-1692, doi:10.1038/bjc.2011.450 (2011).

79   Swerdlow SH, C. E., Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J. *WHO classification of tumors of haematopoietic and lymphoid tissues*. 4th edn, Vol. 2 (IARC, 2008).

80   Zhang, J. *et al.* Genetic heterogeneity of diffuse large B-cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 1398-1403, doi:10.1073/pnas.1205299110 (2013).

81   Lohr, J. G. *et al.* Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 3879-3884, doi:10.1073/pnas.1121343109 (2012).

82   Morin, R. D. *et al.* Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood* **122**, 1256-1265, doi:10.1182/blood-2013-02-483727 (2013).

83   De Silva, N. S. & Klein, U. Dynamics of B cells in germinal centres. *Nature reviews. Immunology* **15**, 137-148, doi:10.1038/nri3804 (2015).

84   Kosco-Vilbois, M. H. Are follicular dendritic cells really good for nothing? *Nature reviews. Immunology* **3**, 764-769, doi:10.1038/nri1179 (2003).

85   Allen, C. D. & Cyster, J. G. Follicular dendritic cell networks of primary follicles and germinal centers: phenotype and function. *Semin Immunol* **20**, 14-25, doi:10.1016/j.smim.2007.12.001 (2008).

86   Park, C. S. & Choi, Y. S. How do follicular dendritic cells interact intimately with B cells in the germinal centre? *Immunology* **114**, 2-10, doi:10.1111/j.1365-2567.2004.02075.x (2005).

87   King, C. J. a. C. Cytokines in the Germinal Center Niche. *antibodies* **5**, doi:0.3390/antib5010005 (2016).

88   Dominguez-Sola, D. *et al.* The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. *Nat Immunol* **13**, 1083-1091, doi:10.1038/ni.2428 (2012).

89   Shaffer, A. L., 3rd, Young, R. M. & Staudt, L. M. Pathogenesis of human B cell lymphomas. *Annu Rev Immunol* **30**, 565-610, doi:10.1146/annurev-immunol-020711-075027 (2012).

90   Kuppers, R. & Dalla-Favera, R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene* **20**, 5580-5594, doi:10.1038/sj.onc.1204640 (2001).

91   Pasqualucci, L. *et al.* Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* **412**, 341-346, doi:10.1038/35085588 (2001).

92   Pasqualucci, L. *et al.* BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 11816-11821 (1998).

93   Alizadeh, A. A. *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503-511, doi:10.1038/35000501 (2000).

94   Wilson, W. H. *et al.* Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat Med* **21**, 922-926, doi:10.1038/nm.3884 (2015).

95   Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495-501, doi:10.1038/nature12912 (2014).

96   Chapuy, B. *et al.* Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med* **24**, 679-690, doi:10.1038/s41591-018-0016-8 (2018).

97   Davis, R. E. *et al.* Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* **463**, 88-92, doi:10.1038/nature08638 (2010).

98   Pasqualucci, L. *et al.* Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* **471**, 189-195, doi:10.1038/nature09730 (2011).

99   Zhang, J. *et al.* Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med* **21**, 1190-1198, doi:10.1038/nm.3940 (2015).

100   Lunning, M. A. & Green, M. R. Mutation of chromatin modifiers; an emerging hallmark of germinal center B-cell lymphomas. *Blood Cancer J* **5**, e361, doi:10.1038/bcj.2015.89 (2015).

101   Challa-Malladi, M. *et al.* Combined genetic inactivation of beta2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. *Cancer cell* **20**, 728-740, doi:10.1016/j.ccr.2011.11.006 (2011).

102   Healy, J. A. *et al.* GNA13 loss in germinal center B cells leads to impaired apoptosis and promotes lymphoma in vivo. *Blood* **127**, 2723-2731, doi:10.1182/blood-2015-07-659938 (2016).

103   Lenz, G. *et al.* Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science* **319**, 1676-1679, doi:10.1126/science.1153629 (2008).

104   Ngo, V. N. *et al.* Oncogenically active MYD88 mutations in human lymphoma. *Nature* **470**, 115-119, doi:10.1038/nature09671 (2011).

105   Kato, M. *et al.* Frequent inactivation of A20 in B-cell lymphomas. *Nature* **459**, 712-716, doi:10.1038/nature07969 (2009).

106   Davis, R. E., Brown, K. D., Siebenlist, U. & Staudt, L. M. Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. *J Exp Med* **194**, 1861-1874 (2001).

107   Schmitz, R. *et al.* Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature* **490**, 116-120, doi:10.1038/nature11378 (2012).

108   Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262-1278, doi:10.1016/j.cell.2014.05.010 (2014).

109   Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-811, doi:10.1038/35888 (1998).

110   Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**, 636-646, doi:10.1038/nrg2842 (2010).

111   Joung, J. K. & Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* **14**, 49-55, doi:10.1038/nrm3486 (2013).

112   Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**, 230-232, doi:10.1038/nbt.2507 (2013).

113   Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823, doi:10.1126/science.1231143 (2013).

114   Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826, doi:10.1126/science.1232033 (2013).

115   Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62-67, doi:10.1038/nature13011 (2014).

116   Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet* **16**, 299-311, doi:10.1038/nrg3899 (2015).

117   Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome. *Nature* **458**, 719-724, doi:10.1038/nature07943 (2009).

118   Weinstock, D. M. *et al.* A roadmap for discovery and translation in lymphoma. *Blood* **125**, 2175-2177, doi:10.1182/blood-2015-01-623777 (2015).

119   Stevens, S. J. *et al.* Monitoring of epstein-barr virus DNA load in peripheral blood by quantitative competitive PCR. *J Clin Microbiol* **37**, 2852-2857 (1999).

120   Christodoulopoulos, I. & Cannon, P. M. Sequences in the cytoplasmic tail of the gibbon ape leukemia virus envelope protein that prevent its incorporation

into lentivirus vectors. *J Virol* **75**, 4129-4138, doi:10.1128/JVI.75.9.4129-4138.2001 (2001).

121 Koike-Yusa, H., Li, Y., Tan, E. P., Velasco-Herrera Mdel, C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol* **32**, 267-273, doi:10.1038/nbt.2800 (2014).

122 Tzelepis, K. *et al.* A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. *Cell Rep* **17**, 1193-1205, doi:10.1016/j.celrep.2016.09.079 (2016).

123 Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184-191, doi:10.1038/nbt.3437 (2016).

124 Banchereau, J., de Paoli, P., Valle, A., Garcia, E. & Rousset, F. Long-term human B cell lines dependent on interleukin-4 and antibody to CD40. *Science* **251**, 70-72 (1991).

125 Rousset, F., Garcia, E. & Banchereau, J. Cytokine-induced proliferation and immunoglobulin production of human B lymphocytes triggered through their CD40 antigen. *J Exp Med* **173**, 705-710 (1991).

126 Kwakkenbos, M. J. *et al.* Genetic manipulation of B cells for the isolation of rare therapeutic antibodies from the human repertoire. *Methods* **65**, 38-43, doi:10.1016/j.ymeth.2013.07.002 (2014).

127 Nojima, T. *et al.* In-vitro derived germinal centre B cells differentially generate memory B or plasma cells in vivo. *Nat Commun* **2**, 465, doi:10.1038/ncomms1475 (2011).

128 Kim, H. S., Zhang, X., Klyushnenkova, E. & Choi, Y. S. Stimulation of germinal center B lymphocyte proliferation by an FDC-like cell line, HK. *J Immunol* **155**, 1101-1109 (1995).

129 Kosco, M. H., Pflugfelder, E. & Gray, D. Follicular dendritic cell-dependent adhesion and proliferation of B cells in vitro. *J Immunol* **148**, 2331-2339 (1992).

130 Yang, G. *et al.* Knockdown of p53 combined with expression of the catalytic subunit of telomerase is sufficient to immortalize primary human ovarian surface epithelial cells. *Carcinogenesis* **28**, 174-182, doi:10.1093/carcin/bgl115 (2007).

131 Ramirez, R. D. *et al.* Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. *Cancer Res* **64**, 9027-9034, doi:10.1158/0008-5472.CAN-04-3703 (2004).

132 Hahn, W. C. *et al.* Enumeration of the simian virus 40 early region elements necessary for human cell transformation. *Mol Cell Biol* **22**, 2111-2123 (2002).

133 Kwakkenbos, M. J. *et al.* Generation of stable monoclonal antibody-producing B cell receptor-positive human memory B cells by genetic programming. *Nat Med* **16**, 123-128, doi:10.1038/nm.2071 (2010).

134 Kwakkenbos, M. J., van Helden, P. M., Beaumont, T. & Spits, H. Stable long-term cultures of self-renewing B cells and their applications. *Immunol Rev* **270**, 65-77, doi:10.1111/imr.12395 (2016).

135 Serafini, M., Naldini, L. & Introna, M. Molecular evidence of inefficient transduction of proliferating human B lymphocytes by VSV-pseudotyped HIV-1-derived lentivectors. *Virology* **325**, 413-424, doi:10.1016/j.virol.2004.04.038 (2004).

136 Janssens, W. *et al.* Efficiency of onco-retroviral and lentiviral gene transfer into primary mouse and human B-lymphocytes is pseudotype dependent. *Hum Gene Ther* **14**, 263-276, doi:10.1089/10430340360535814 (2003).

137 Mock, U., Thiele, R., Uhde, A., Fehse, B. & Horn, S. Efficient lentiviral transduction and transgene expression in primary human B cells. *Hum Gene Ther Methods* **23**, 408-415, doi:10.1089/hgtb.2012.160 (2012).

138 Hunter, E. in *Retroviruses* (eds J. M. Coffin, S. H. Hughes, & H. E. Varmus) (1997).

139 Miller, A. D. Cell-surface receptors for retroviruses and implications for gene transfer. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 11407-11413 (1996).

140 O'Hara, B. *et al.* Characterization of a human gene conferring sensitivity to infection by gibbon ape leukemia virus. *Cell Growth Differ* **1**, 119-127 (1990).

141 Miller, A. D. *et al.* Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus. *J Virol* **65**, 2220-2224 (1991).

142 Wilson, C., Reitz, M. S., Okayama, H. & Eiden, M. V. Formation of infectious hybrid virions with gibbon ape leukemia virus and human T-cell leukemia virus retroviral envelope glycoproteins and the gag and pol proteins of Moloney murine leukemia virus. *J Virol* **63**, 2374-2378 (1989).

143   Filipits, M. *et al.* Cyclin D3 is a predictive and prognostic factor in diffuse large B-cell lymphoma. *Clinical cancer research : an official journal of the American Association for Cancer Research* **8**, 729-733 (2002).

144   Vater, I. *et al.* The mutational pattern of primary lymphoma of the central nervous system determined by whole-exome sequencing. *Leukemia* **29**, 677-685, doi:10.1038/leu.2014.264 (2015).

145   Pasqualucci, L. *et al.* Genetics of follicular lymphoma transformation. *Cell Rep* **6**, 130-140, doi:10.1016/j.celrep.2013.12.027 (2014).

146   Pasqualucci, L. *et al.* Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet* **43**, 830-837, doi:10.1038/ng.892 (2011).

147   Okosun, J. *et al.* Recurrent mTORC1-activating RRAGC mutations in follicular lymphoma. *Nat Genet* **48**, 183-188, doi:10.1038/ng.3473 (2016).

148   Okosun, J. *et al.* Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat Genet* **46**, 176-181, doi:10.1038/ng.2856 (2014).

149   Morin, R. D. *et al.* Genetic Landscapes of Relapsed and Refractory Diffuse Large B-Cell Lymphomas. *Clinical cancer research : an official journal of the American Association for Cancer Research* **22**, 2290-2300, doi:10.1158/1078-0432.CCR-15-2123 (2016).

150   Morin, R. D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298-303, doi:10.1038/nature10351 (2011).

151   de Miranda, N. F. *et al.* DNA repair genes are selectively mutated in diffuse large B cell lymphomas. *J Exp Med* **210**, 1729-1742, doi:10.1084/jem.20122842 (2013).

152   de Miranda, N. F. *et al.* Exome sequencing reveals novel mutation targets in diffuse large B-cell lymphomas derived from Chinese patients. *Blood* **124**, 2544-2553, doi:10.1182/blood-2013-12-546309 (2014).

153   Melchardt, T. *et al.* Clonal evolution in relapsed and refractory diffuse large B-cell lymphoma is characterized by high dynamics of subclones. *Oncotarget* **7**, 51494-51502, doi:10.18632/oncotarget.9860 (2016).

154   Mareschal, S. *et al.* Whole exome sequencing of relapsed/refractory patients expands the repertoire of somatic mutations in diffuse large B-cell lymphoma. *Genes Chromosomes Cancer* **55**, 251-267, doi:10.1002/gcc.22328 (2016).

155   Kretzmer, H. *et al.* DNA methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic

mutation and transcriptional control. *Nat Genet* **47**, 1316-1325, doi:10.1038/ng.3413 (2015).

156   Juskevicius, D. *et al.* Distinct genetic evolution patterns of relapsing diffuse large B-cell lymphoma revealed by genome-wide copy number aberration and targeted sequencing analysis. *Leukemia* **30**, 2385-2395, doi:10.1038/leu.2016.135 (2016).

157   Jiang, Y. *et al.* Deep sequencing reveals clonal evolution patterns and mutation events associated with relapse in B-cell lymphomas. *Genome Biol* **15**, 432, doi:10.1186/s13059-014-0432-0 (2014).

158   Braggio, E. *et al.* Genome-Wide Analysis Uncovers Novel Recurrent Alterations in Primary Central Nervous System Lymphomas. *Clinical cancer research : an official journal of the American Association for Cancer Research* **21**, 3986-3994, doi:10.1158/1078-0432.CCR-14-2116 (2015).

159   Todorovic Balint, M. *et al.* Gene Mutation Profiles in Primary Diffuse Large B Cell Lymphoma of Central Nervous System: Next Generation Sequencing Analyses. *Int J Mol Sci* **17**, doi:10.3390/ijms17050683 (2016).

160   Novak, A. J. *et al.* Whole-exome analysis reveals novel somatic genomic alterations associated with outcome in immunochemotherapy-treated diffuse large B-cell lymphoma. *Blood Cancer J* **5**, e346, doi:10.1038/bcj.2015.69 (2015).

161   Cellecta, I. *Cellecta Pooled shRNA Library Screening Reference Manual*, 2015).

162   Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80-84, doi:10.1126/science.1246981 (2014).

163   De Silva, N. S. *et al.* Transcription factors of the alternative NF-kappaB pathway are required for germinal center B-cell development. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 9063-9068, doi:10.1073/pnas.1602728113 (2016).

164   Pfeifer, M. *et al.* PTEN loss defines a PI3K/AKT pathway-dependent germinal center subtype of diffuse large B-cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 12420-12425, doi:10.1073/pnas.1305656110 (2013).

165   Uddin, S. *et al.* Role of phosphatidylinositol 3'-kinase/AKT pathway in diffuse large B-cell lymphoma survival. *Blood* **108**, 4178-4186, doi:10.1182/blood-2006-04-016907 (2006).

166   Chen, L., Juszczynski, P., Takeyama, K., Aguiar, R. C. & Shipp, M. A. Protein tyrosine phosphatase receptor-type O truncated (PTPROt) regulates SYK phosphorylation, proximal B-cell-receptor signaling, and cellular proliferation. *Blood* **108**, 3428-3433, doi:10.1182/blood-2006-03-013821 (2006).

167   Petitjean, A., Hainaut, P. & Caron de Fromentel, C. TP63 gene in stress response and carcinogenesis: a broader role than expected. *Bull Cancer* **93**, E126-135 (2006).

168   Bruno, A. *et al.* Mutational analysis of primary central nervous system lymphoma. *Oncotarget* **5**, 5065-5075, doi:10.18632/oncotarget.2080 (2014).

169   Lovejoy, C. A. *et al.* Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. *PLoS Genet* **8**, e1002772, doi:10.1371/journal.pgen.1002772 (2012).

170   Rounbehler, R. J. *et al.* Tristetraprolin impairs myc-induced lymphoma and abolishes the malignant state. *Cell* **150**, 563-574, doi:10.1016/j.cell.2012.06.033 (2012).

171   Phelan, J. D. *et al.* A multiprotein supercomplex controlling oncogenic signalling in lymphoma. *Nature*, doi:10.1038/s41586-018-0290-0 (2018).

172   O'Hayre, M. *et al.* The emerging mutational landscape of G proteins and G-protein-coupled receptors in cancer. *Nature reviews. Cancer* **13**, 412-424, doi:10.1038/nrc3521 (2013).

173   Dorsam, R. T. & Gutkind, J. S. G-protein-coupled receptors and cancer. *Nature reviews. Cancer* **7**, 79-94, doi:10.1038/nrc2069 (2007).

174   Xu, Y. *et al.* High expression of GNA13 is associated with poor prognosis in hepatocellular carcinoma. *Sci Rep* **6**, 35948, doi:10.1038/srep35948 (2016).

175   Rasheed, S. A., Teo, C. R., Beillard, E. J., Voorhoeve, P. M. & Casey, P. J. MicroRNA-182 and microRNA-200a control G-protein subunit alpha-13 (GNA13) expression and cell invasion synergistically in prostate cancer cells. *The Journal of biological chemistry* **288**, 7986-7995, doi:10.1074/jbc.M112.437749 (2013).

176   Rasheed, S. A. *et al.* MicroRNA-31 controls G protein alpha-13 (GNA13) expression and cell invasion in breast cancer cells. *Mol Cancer* **14**, 67, doi:10.1186/s12943-015-0337-x (2015).

177 Rasheed, S. A. K. *et al.* GNA13 expression promotes drug resistance and tumor-initiating phenotypes in squamous cell cancers. *Oncogene* **37**, 1340-1353, doi:10.1038/s41388-017-0038-6 (2018).

178 Goldsmith, Z. G. & Dhanasekaran, D. N. G protein regulation of MAPK networks. *Oncogene* **26**, 3122-3142, doi:10.1038/sj.onc.1210407 (2007).

179 O'Hayre, M. *et al.* Inactivating mutations in GNA13 and RHOA in Burkitt's lymphoma and diffuse large B-cell lymphoma: a tumor suppressor function for the Galpha13/RhoA axis in B cells. *Oncogene* **35**, 3771-3780, doi:10.1038/onc.2015.442 (2016).

180 Muppidi, J. R. *et al.* Loss of signalling via Galpha13 in germinal centre B-cell-derived lymphoma. *Nature* **516**, 254-258, doi:10.1038/nature13765 (2014).

181 Stelling, A. *et al.* The tumor suppressive TGF-beta/SMAD1/S1PR2 signaling axis is recurrently inactivated in diffuse large B-cell lymphoma. *Blood* **131**, 2235-2246, doi:10.1182/blood-2017-10-810630 (2018).

182 Calado, D. P. *et al.* The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat Immunol* **13**, 1092-1100, doi:10.1038/ni.2418 (2012).

183 Suk, F. M. *et al.* ZFP36L1 and ZFP36L2 inhibit cell proliferation in a cyclin D-dependent and p53-independent manner. *Sci Rep* **8**, 2742, doi:10.1038/s41598-018-21160-z (2018).

184 Liang, J., Song, W., Tromp, G., Kolattukudy, P. E. & Fu, M. Genome-wide survey and expression profiling of CCCH-zinc finger family reveals a functional module in macrophage activation. *PLoS One* **3**, e2880, doi:10.1371/journal.pone.0002880 (2008).

185 Brennan, S. E. *et al.* The mRNA-destabilizing protein tristetraprolin is suppressed in many cancers, altering tumorigenic phenotypes and patient prognosis. *Cancer Res* **69**, 5168-5176, doi:10.1158/0008-5472.CAN-08-4238 (2009).

186 Hodson, D. J. *et al.* Deletion of the RNA-binding proteins ZFP36L1 and ZFP36L2 leads to perturbed thymic development and T lymphoblastic leukemia. *Nat Immunol* **11**, 717-724, doi:10.1038/ni.1901 (2010).

187 Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029-1041 e1021, doi:10.1016/j.cell.2017.09.042 (2017).

188   Kim, E. *et al.* Systematic Functional Interrogation of Rare Cancer Variants Identifies Oncogenic Alleles. *Cancer Discov* **6**, 714-726, doi:10.1158/2159-8290.CD-16-0160 (2016).

189   Klein, U. *et al.* Transcriptional analysis of the B cell germinal center reaction. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 2639-2644, doi:10.1073/pnas.0437996100 (2003).

190   Zhang, J. *et al.* The CREBBP Acetyltransferase Is a Haploinsufficient Tumor Suppressor in B-cell Lymphoma. *Cancer Discov* **7**, 322-337, doi:10.1158/2159-8290.CD-16-1417 (2017).

191   Green, M. R. *et al.* Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E1116-1125, doi:10.1073/pnas.1501199112 (2015).

192   Jiang, Y. *et al.* CREBBP Inactivation Promotes the Development of HDAC3-Dependent Lymphomas. *Cancer Discov* **7**, 38-53, doi:10.1158/2159-8290.CD-16-0975 (2017).