INVESTIGATING LEUKAEMIC PROPAGATION IN CHILDHOOD ACUTE LYMPHOBLASTIC LEUKAEMIA

SIMON BOMKEN

A thesis submitted in part requirement for the degree of Doctor of Philosophy from the Faculty of Medical Sciences

Newcastle University

Newcastle upon Tyne

October 2012

Leukaemia Stem Cell Laboratory
Northern Institute for Cancer Research
Newcastle University

Abstract

Childhood acute lymphoblastic leukaemia (ALL) does not possess a propagating cell hierarchy, at least as defined by B-cell precursor immunophenotype. Indeed, many, or even all, leukaemic blasts may have the potential to propagate the disease. This unusual characteristic mirrors the substantial capacity for clonal expansion demonstrated by fully differentiated normal lymphoid cells. This Fellowship aimed to investigate the genetic programmes underlying the propagation of acute lymphoblastic leukaemia.

An initial candidate approach confirmed the expression of *PIWIL2*, a gene critical to the maintenance of germline stem cells, in both cell line and primary ALL. Knockdown of *PIWIL2* resulted in reduced cellular proliferation and significant prolongation of doubling time in two ALL cell lines, SEM (*MLL/AF4*) and 697 (*E2A/PBX1*). Unexpectedly, *PIWIL2* was also found to be expressed in peripheral lymphoid cells from healthy donors, but not terminally differentiated cells of myeloid origin, suggesting that *PIWIL2* may have a previously unidentified function in both normal and malignant lymphoid cells.

A second project has developed an *in vitro* genome-wide RNAi screen to identify candidate genes involved in the clonal propagation of ALL. This project has assessed a serial re-plating assay using feeder cell co-culture to provide a surrogate niche environment. Initial results have demonstrated the feasibility of such an approach. The benefit of using a co-culture re-plating assay, as compared to a standard suspension culture approach, remains under investigation.

Finally, this Fellowship developed a protocol for the lentiviral transduction of patient-derived leukaemic blasts and cloned and validated a novel lentiviral vector capable of *in vitro* analysis, *in vivo* disease monitoring and RNAi. With these, it will now be possible to validate candidate leukaemic propagation genes *in vivo*, using primary leukaemic material. The results of these studies will provide candidates for the development of novel therapeutic agents for children with ALL.

Acknowledgements

I would like to acknowledge the Medical Research Council for believing that even a clinician might have something to offer in the way of basic science. I hope their faith was well placed. However, the work described in this thesis was only possible because of the tireless support and assistance offered to me by members of the Leukaemia Stem Cell laboratory who “taught me everything I know”, especially the extremely talented Patricia GarridoCastro, Frida Ponthan and Lars Buechler. I must also thank my colleague and bench mate Klaus Rehe for being such a great friend over the last three years.

I would like to thank a number of people who contributed directly to my projects. Klaus, Frida and Helen Blair performed xeno-transplants and bone marrow aspirates for me. Hesta McNeil spent many hours sorting cells on the FacsVantage. Svetlana Myssina performed early *PIWIL2* experiments, providing the pilot data for my MRC application. Nana Anim-Addo and Cara Hernon took on the Gateway cloning project with me, making an excellent job of their BSc projects. Jen Jackson and Chris Bacon performed and analysed the immunohistochemistry on murine xenograft samples.

To my supervisors, Josef and Olaf, I owe a huge debt of gratitude. Josef has provided me with unbelievable support and opportunity at the beginning of my clinical academic career. Olaf has tried his hardest to teach me just a fraction of his immense knowledge of the world of biology. I could not have chosen a better supervisory team.

Finally, my family – my parents, Charlotte and my two beautiful boys, Nick and Alex. Three years has flown by. It is impossible to imagine all that has happened in that time – there was only one of you at the start! Thank you so much for all your love and support. This is for you.

Table of Contents

Table of Figures

Table of Tables

Abbreviations

αMEM	Alpha modified Minimal Essential Medium
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
APC	Allophycocyanin
BTIC	Brain tumour initiating cell
cDNA	Complementary DNA
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate
CSC	Cancer stem cell
DNA	Deoxyribonucleic acid
DMSO	Dimethylsulphoxide
DS-ALL	Down syndrome associated acute lymphoblastic leukaemia
DTT	DL-Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
FACS	Fluorescence activated cell sorting
FCS	Fetal calf serum
FITC	Fluorescein isothiocyanate
GFP	Green fluorescent protein
GMP	Granulocyte-macrophage precursor
GTP	Guanosine triphosphate

Gy	Gray
HMGA2	High mobility group protein A2
HSC	Haematopoietic stem cell
IL	Interleukin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LMPP	Lymphoid-primed multipotential progenitor
LSC	Leukaemia stem cell
MCS	Multiple cloning site
miRNA	MicroRNA
mir30	MicroRNA 30
MRD	Minimal residual disease
PAC	Puromycin N-acetyl-transferase
PCR	Polymerase chain reaction
PBS	Phosphate buffered saline
PE	Phycoerythrin
piRNA	PIWI-interacting RNA
PPD	PIWI PAZ Domain
Pre-miRNA	Precursor miRNA
Pri-miRNA	Primary miRNA
PVDF	Polyvinylidene difluoride
qRT PCR	Quantitative reverse transcribed polymerase chain reaction
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid

RNAi	RNA interference
RTE	Relative transduction efficiency
SDS	Sodium dodecyl sulfate
s.e.m.	Standard error of the mean
SFEM	Serum free expansion medium
shRNA	Short hairpin RNA
siRNA	Small interfering RNA
TEMED	N,N,N′,N′-Tetramethylethylenediamine
TERT	Telomerase reverse transcriptase
tGFP	Turbo green fluorescent protein
Tris	Tris(hydroxymethyl)aminomethane
TTRAP	Transformation/transcription-domain associated protein

Suppliers

Axis-Shield plc
Luna Place
The Technology Park
Dundee
DD2 1XA
www.axis-shield.com

BD Biosciences
Edmund Halley Road
Oxford Science Park
OX4 4DQ
Oxford
www.bdbiosciences.com

Beckman Coulter (UK) Ltd
Oakley Court
Kingsmead Business Park
London Road
High Wycombe
HP11 1JU
www.beckmancoulter.com

BioRad Laboratories Ltd
Bio-Rad House
Maxted Road
Hemel Hempstead
HP2 7DX
www.bio-rad.com

Caliper Life Sciences
1 Wellfield
Preston Brook
Runcorn
WA7 3AZ
www.caliperls.com

Charles River UK Ltd
Manston Road
Margate
Kent
CT9 4LT
www.criver.com

Colenta Labortechnik
Neunkirchner Str. 117
A-2700, Weiner Neustadt
Austria
www.colenta.at

DBS Genomics
School of Biological & Biomedical Sciences
South Road Science Site
Durham
DH1 3LE
www.dur.ac.uk/biosciences

Eurogentec S.A.
LIEGE Science Park
Rue du Bois Saint Jean 5
4102 Seraing
BELGIUM
www.eurogentec.com

Fermentas Life Sciences
Sheriff House
Sheriff Hutton Industrial Park
York
Y060 6RZ
www.fermentas.de

Fisher Scientific UK
Bishop Meadow Road
Loughborough
LE11 ORG
www.fisher.co.uk

Functional Genomics & Proteomics Facility
School of Biosciences
University of Birmingham
Edgbaston
B15 2TT
www.genomics.bham.ac.uk

GE Healthcare UK Ltd
Amersham Place
Little Chalfont
HP7 9NA
www.gelifesciences.com

Gulmay Ltd
St. Ann's House
St. Ann's Road
Chertsey
KT16 9EH
www.gulmay.co.uk

Hawksley
Marlborough Road
Lancing Business Park
Lancing
BN15 8TN
www.hawksley.co.uk

HyTest Ltd
Intelligate 6th Floor
Joukahaisenkatu 6
20520 Turku
Finland
www.hytest.fi

Invitrogen Ltd
3 Fountain Drive
Inchinnan Business Park
Paisley
PA4 9RF
www.invitrogen.com

The Jackson Laboratory
600 Main Street
Bar Harbor, Maine 04609
USA
www.jax.org

Labtech International Ltd
Acorn House
The Broyle
Ringmer
BN8 5NN
www.labtech.co.uk

Leica Microsystems (UK)
Davy Avenue
Knowlhill
Milton Keynes
MK5 8LB
www.leica-microsystems.com

Macherey-Nagel GmBH
Neumann Neander Str. 6-8
D-52355 Düren
Germany
www.mn-net.com

Merck Chemicals Ltd
Boulevard Industrial Park
Padge Road
Beeston
Nottingham
NG9 2JR
www.merck-chemicals.com

Millipore (U.K.) Ltd
Suite 3 & 5
Building 6
Croxley Green Business Park
Watford
WD18 8YH
www.millipore.com

New England Biolabs
75/77 Knowl Piece
Wilbury Way
Hitchin
SG4 0TY
www.neb.com

Olympus Ltd
KeyMed House
Stock Road
Southend-on-Sea
SS2 5QH
www.olympus.co.uk

Open Biosystems Products
601 Genome Way
Suite 2100
Huntsville
AL 35806
USA
www.openbiosystems.com

PAA Laboratories Ltd
Termare Close
Houndstone Business Park
Yeovil
BA22 8YG
www.paa.com

PEQLAB Ltd
Building 34
Universal Marina
Crableck Lane
Sarisbury Green
SO31 7ZN
www.peqlab.co.uk

Pfizer Ltd
Sandwich Road
Kent
CT13 9NJ
www.pfizer.co.uk

Promega
Delta House
Southampton Science Park
Southampton
SO16 7NS
www.promega.com

QIAGEN Ltd
Fleming Way
Crawley
RH10 9NQ
www.qiagen.com

R&D Systems Europe Ltd.
19 Barton Lane
Abingdon Science Park
Abingdon
OX14 3NB
www.rndsystems.com

Sigma-Aldrich Company Ltd.
The Old Brickyard
New Road
Gillingham
SP8 4XT
www.sigmaaldrich.com

SPOT Imaging Solutions
6540 Burroughs
Sterling Heights
MI 48314
USA
www.spotimaging.com

StemCell Technologies
Northern European Sales Office
Unit 2F
Rutherford House
Pencroft Way
Manchester
M15 6SZ
www.stemcell.com

Thermo Scientific
ABgene House
Blenheim Road
Epsom
KT19 9AP
www.thermoscientific.com

Ultra-Violet Products Ltd
Unit 1
Trinity Hall Farm Estate
Nuffield Road
Cambridge
CB4 1TG
www.uvp.com

Carl Zeiss Ltd
509 Coldhams Lane
Cambridge
CB1 3JS
www.zeiss.co.uk

Chapter 1

Introduction

1.1 Childhood B precursor acute lymphoblastic leukaemia

1.1.1 Introduction

Acute lymphoblastic leukaemia (ALL) is the single most common malignancy of childhood, representing approximately 25% of malignant diagnoses between birth and 15 years of age (CRUK 2011). It provides a paradigm for the improvement in children's cancer therapies with survival improving from less than 30% fifty years ago, to 80% 5 year event free survival and 85% 5 year overall survival with current intensive, multi-agent chemotherapy regimens combined with improvements in supportive care (Pui, Relling et al. 2004). Despite these advances, the relatively high incidence means that it is still one of the most frequent causes of death from malignancy in childhood. In addition, intensive therapies carry a burden of both acute and, particularly, long-term toxicity, which result in a significant increase in mortality for survivors of childhood leukaemia compared to age-matched controls, (see 1.1.6). Furthermore, despite the good overall survival, substantial difficulties remain in identifying and managing those children who have an increased risk of relapse and who have a much poorer chance of cure.

Within childhood acute lymphoblastic leukaemia, leukaemias derived from the B lineage represent over 90% of diagnoses (O'Neill, Bunch et al. 2012). This lineage has quite distinct biology from T lineage disease and so for the remainder of this thesis only B lineage leukaemia will be considered.

1.1.2 Epidemiology

In Caucasian populations the incidence of childhood ALL is approximately 35 cases per million children per year. In the United Kingdom this equates to approximately 370 new diagnoses per year. Boys are slightly more frequently affected than girls with a sex ratio of 1.2:1 (CCRG 2010).

The only established causative exposure for childhood ALL is ionizing radiation, either in the form of *in utero* exposure to diagnostic radiation (Doll and Wakeford 1997; Pearce, Salotti et al. 2012) or environmental exposure, such as following the atomic bombing of Japan in 1945 (Preston, Kusumi et al. 1994). Many other environmental risk factors have been investigated with regard to the risk of developing ALL, mostly without identifying good evidence of association. In particular, despite substantial public anxiety surrounding exposure to environmental radon (UKCCSInvestigators 2002) or neonatal vitamin K injections (Fear, Roman et al. 2003), no evidence of an association between these factors and development of ALL has been found. Another area of concern, exposure to electromagnetic fields from overhead or domestic power cables (Coghill, Steward et al. 1996; Ahlbom, Day et al. 2000), is supported by data which may indicate a moderate excess risk in exposed children (UKCCSInvestigators 1999), but has not been supported by a plausible biological explanation to date.

In contrast, substantial epidemiological evidence exists to support the theory that a delayed exposure or abnormal response to childhood infection may facilitate malignant transformation of a pre-malignant clone, (see 1.1.3). Notably, studies of large UK and US cohorts identified regular attendance at day-care before the age of 1 year, a proxy for early exposure to infections, to be the only factor associated with a decreased risk of developing ALL (Ma, Buffler et al. 2002; Gilham, Peto et al. 2005).

A number of genetic syndromes are well recognised to predispose to the development of ALL. Most of these are rare, with the exception of trisomy 21, Down syndrome, which carries an approximately 20 fold increased risk of ALL (Hasle, Clemmensen et al. 2000). The underlying genetic properties of Down syndrome associated ALL (DS-ALL) are not the same as non DS-ALL, with low incidence of *MLL* rearrangements and t(9;22) (Pui, Raimondi et al. 1993; Forestier, Izraeli et al. 2008; Maloney, Carroll et al. 2010) and an excess representation in the group of patients with deregulation of CRLF2 (Mullighan, Collins-Underwood et al. 2009; Russell, Capasso et al. 2009). Outcome for children treated for DS-ALL has been seen to vary between different trial groups. However, there is a general consensus that children with Down syndrome experience greater treatment related toxicity, benefiting from more intensive supportive care and possibly from certain dose modifications (Maloney 2011). However, some groups have identified poorer outcome following dose modifications (Bohnstedt, Taskinen et al. 2009; Goto, Inukai et al. 2011) and the optimal balance between dose intensity and treatment toxicity remains to be determined.

1.1.3 Origins of the disease

The origin of ALL has been a subject of intense research over many years. Pioneering work by Mel Greaves has demonstrated the development of a pre-malignant clone *in utero*. Clonal populations harbouring the most common genetic aberrations, high hyperdiploidy and the *ETV6/RUNX1* translocation, as well as translocations involving the *MLL* gene, have been identified in monozygotic twins concordant for ALL, implying a common fetal origin to their disease (Mahmoud, Ridge et al. 1995; Ford, Bennett et al. 1998; Maia, van der Velden et al. 2003). Furthermore, analysis of Guthrie blood spots has demonstrated the presence of *ETV6/RUNX1* translocations as frequently as in 1% of newborn blood samples (Mori, Colman et al. 2002). This frequency is 100 fold that of the development of overt *ETV6/RUNX1* positive ALL, implying that the additional genetic changes required for full malignant transformation are relatively rare. Furthermore, in a fascinating recent study of a single pair of monozygotic twins discordant for ALL, Hong *et al* demonstrated the presence of a persistent *ETV6/RUNX1* positive pre-malignant clone in the unaffected twin (Hong, Gupta et al. 2008). Whilst the translocation resulted in an abnormal immunophenotype, believed to represent a “pre-malignant stem cell”, it lacked the additional copy number changes seen in the ALL of the affected twin and failed to undergo malignant transformation over an eighteen month period of monitoring.

The nature of the additional genetic changes required to convert a pre-malignant cell into overt leukaemia has attracted a substantial amount of attention, particularly in recent studies. Some secondary mutations are well recognised, such as the deletion of the second, un-translocated *ETV6* gene in *ETV6/RUNX1* positive ALL (Golub, Barker et al. 1995; Raynaud, Cave et al. 1996), deletion of *CDKN2A* (Okuda, Shurtleff et al. 1995; Zhou, Gu et al. 1995), *CCNC* (Li, Lahti et al. 1996; van Delft, Horsley et al. 2011) or *ATM* (Raimondi, Frestedt et al. 1995). Another group of mutations are found within the pathways determining B precursor development. Recent studies have identified frequent mutations of the transcription factors which govern the process of normal B precursor development, with 40% of B precursor leukaemias harbouring a

mutation of one or more of these genes. Most frequent are the largely monoallelic loss of function mutations within *PAX5*, a master regulator of B lineage commitment and development, and *IKZF1* which encodes the transcription factor IKAROS (Mullighan, Goorha et al. 2007). A final group of common mutations are those which result in activation of key proliferative pathways such as the RAS/RAF/MAPK/ERK pathway (Case, Matheson et al. 2008; Nicholson, Knight et al. 2012).

1.1.4 Cytogenetic subgroups

Childhood acute lymphoblastic leukaemia is characterised by the recurrent finding of a relatively small number of well-defined cytogenetic abnormalities, reviewed in (Moorman 2012). These subgroups are found to correlate strongly with age, white count at diagnosis and immunophenotype, as well as with prognosis. In particular they identify a number of patients with substantially worse prognosis and consequently allow escalation in their treatment. Broadly the abnormalities can be grouped into: translocations - which result in the creation of a fusion oncogene which codes for an abnormal fusion protein with leukaemogenic potential; abnormal ploidy - both hyperdiploid and hypodiploid subgroups exist and carry good and poor risk respectively; other – this group contains a variety of biological subtypes which carry a range of prognoses.

Two aberrations, high hyperdiploidy (defined as 51-65 chromosomes) and t(12;21)(p13;q22) which results in the fusion oncogene *ETV6/RUNX1* (previously known as *TEL/AML1*) are identified in approximately 38% and 25% of B precursor ALL respectively (Moorman, Ensor et al. 2010). They carry a good overall prognosis although, given their relatively high prevalence, still represent a substantial proportion of cases which go on to relapse. Other well defined abnormalities are much less frequent, the most common amongst which are (Moorman, Ensor et al. 2010): t(1;19)(q23;p13) - *TCF3/PBX1* (previously *E2A/PBX1*) present in 4% and carries an intermediate to good prognosis on modern protocols; t(9;22)(q34;q11) – *BCR/ABL* present in 3% and carrying a very poor prognosis with standard chemotherapy but recently greatly improved with the increasing use of tyrosine kinase inhibitors with early results showing 3 year event free survival of approximately 80%, even without allogeneic bone marrow transplantation (Schultz, Bowman et al. 2009; Rives, Estella et al. 2011); intrachromasomal amplification of chromosome 21 – iAMP21 a relatively newly identified abnormality in which multiple complex processes result in the amplification of a common region of 5Mb which includes the *AML1* gene (Harewood, Robinson et al. 2003; Strefford, van Delft et al. 2006). This abnormality affects 2% of cases and whilst it originally held an extremely poor prognosis (Moorman, Richards et al. 2007), identification of these cases and

subsequent escalation to intensive chemotherapy appears to have resulted in improvement in outcome (Heerema, Raetz et al. 2011; Moorman 2012); rearrangements of the *MLL* gene are uncommon outside of infancy but continue to carry a poor prognosis, especially in younger children. Overall they are present in 2% of cases but each individual rearrangement is rather rarer, making sub-classification by translocation partner difficult.

1.1.5 Current management strategies

Current management of childhood B precursor acute lymphoblastic leukaemia in highly developed countries is based on intensive multi-agent chemotherapy regimens, which are applied in a stratified and response adjusted manner. Stratification within the UK is based on the NCI risk criteria whereby a child will receive more intensive chemotherapy if they are greater than 10 years old or if their white blood cell count is greater than 50×10^9/L at any point prior to starting treatment. Within the current UK treatment guidelines, this would involve escalation from regimen A to regimen B. Having started treatment, further escalation in therapy to regimen C would be applied either if cytogenetic analysis demonstrated a high risk abnormality, namely hypodiploidy, t(17;19), *MLL* rearrangement or iAMP21, or if there was a poor response to treatment judged by either failure to achieve morphological remission or high risk minimal residual disease (MRD) monitoring. Patients identified as having t(9;22) Philadelphia chromosome would move to a separate, BCR/ABL specific protocol.

Therapy is divided in individual courses, namely remission induction, consolidation and delayed intensification interspersed with and followed by maintenance therapy (Figure 1-1). The initial courses are intensive and aimed at achieving a molecular remission, as determined by PCR based analysis of clonal B/T cell receptor rearrangements. Subsequent maintenance therapy consists principally of oral chemotherapy aimed at prolonged suppression of residual leukaemic cells. Central nervous system prophylaxis is delivered throughout therapy with repeated intrathecal administration of methotrexate.

Figure 1-1. UK therapy for B precursor ALL. Each regimen contains the same courses, delivered in the same order. The number and delivery of drugs differs between regimens. Maintenance phase is timed to give two years of therapy for girls and three years for boys, taken from the beginning of interim maintenance (not shown, delivered between consolidation and delayed intensification).

Further modifications to therapy are limited, but include the use of tyrosine kinase inhibitors for patients whose leukaemia has t(9;22), or haematopoietic stem cell transplantation in first complete remission for patients with high risk cytogenetic features and poor response to therapy, as determined by analysis of MRD. This strategy is otherwise withheld for relapsed disease.

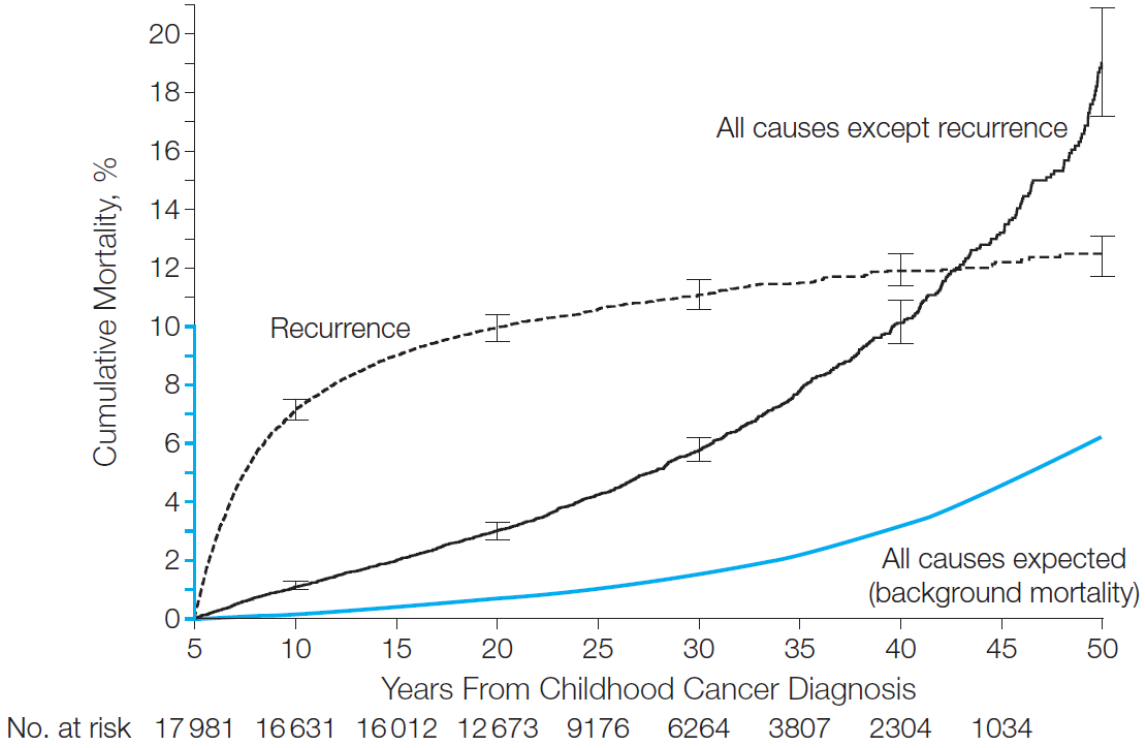
Recent modifications to UK therapy have aimed at using the improved risk stratification offered by MRD analysis to reduce treatment in patients showing a good response to therapy, thereby reducing toxicity. This step is based on the extremely good survival seen in both of these groups of patients - >95% overall survival in the recent UKALL2003 trial. The current UK trial, UKALL2011, is aimed at further reducing the treatment related morbidity and mortality, especially that related to dexamethasone use, see 1.1.6.

1.1.6 Long term effects of treatment

As the overall survival from childhood B precursor ALL has improved to beyond 80%, so the treatment related toxicity has become an increasing concern. As longitudinal studies mature, we are developing a clearer idea of the impact that intensive multi-agent chemotherapy has on long-term health and well-being. Young adults treated for childhood malignancy, including leukaemias, have a significant excess mortality compared to age matched controls (Figure 1-2) (Armstrong, Liu et al. 2009). The risk of death from relapse has dropped as treatment intensity increases, instead, death from second malignancy and cardiac disease now dominates. Not only is this important for the individual survivor, but with one in 640 young adults in the US a survivor of childhood malignancy, it also represents an important strategic healthcare concern.

Current clinical trials continue to attempt to reduce toxicity, focusing on further stratifying patients and reducing treatment intensity to the best risk groups. In addition, it is hoped that altered dosing schedules may provide a degree of protection from toxicity. An example of this is the modification of dexamethasone administration in UKALL 2011, where shorter course of dexamethasone, albeit at a higher dose, are being compared with the current continuous administration during remission induction.

Figure 1-2. Late mortality amongst 5-year survivors of childhood malignancy in the US Childhood Cancer Survivor Study (Armstrong, Liu et al. 2009).



1.2 Cancer stem cell theory

1.2.1 Introduction

That both solid and haematological tumours demonstrate extensive heterogeneity has been known for many years. More recently, knowledge of this physical heterogeneity has been complemented by an increasing understanding of variation in biological and functional characteristics. Much of this work has been aimed at identifying the cell population within a malignancy which is able to maintain the disease and potentially initiate relapse. This cell population was believed to sit at the apex of a hierarchy, where its capacity for indefinite divisions allowed it to populate a tumour. The heterogeneity of that tumour was derived from the differentiation occurring as cells expanded and differentiated. As this situation mirrored normal adult stem cells, in particular the well characterised haematopoietic stem cell, the term cancer stem cell was adopted to describe this population. The identification of cancer stem cells (CSCs) in a broad range of malignancies followed, as it was hoped that targeting this population specifically might offer a novel therapeutic approach. As our understanding of the CSC has developed, this goal seems increasingly difficult to achieve. Nevertheless, understanding the principle of self-renewal within a malignancy may improve the understanding of a tumour's biology.

1.2.2 Definition

As the concept of the cancer stem cell developed, it became important to ensure that a common understanding existed as to what this new term meant. This was particularly true as the term itself led to confusion, with some believing that it implied the cell must have arisen from a normal adult stem cell. Similarly, as *in vivo* engraftment assays were frequently used to define these cells, the alternative name of tumour-initiating cell led some to believe that the CSC was synonymous with the cell of origin of a malignancy. Neither of these assumptions is true, although either may hold for a particular malignancy. In response to the growing interest in CSCs and in an attempt to avoid these confusions, the American Association of Cancer Researchers convened a workshop at which they discussed the need for a definition of the CSC (Clarke, Dick et al. 2006). Whilst their report includes discussion of the relatively infrequency of CSCs and how they may sit within a hierarchy analogous to normal adult stem cells, the definition at which they arrived was:

> *"a cell within a tumor that possess (sic) the capacity to self-renew and to cause the heterogeneous lineages of cancer cells that comprise the tumor."*

A recent review of the state of CSC understanding, performed by a 2011 Working Conference on CSC, has further sought to develop a unified approach to the definition, experimental isolation and functional investigation of CSCs (Valent, Bonnet et al. 2012). In particular, this Conference promoted the detailed reporting of experimental parameters to aid the development of understanding of CSC biology.

1.2.3 Identification of cancer stem cells

Much of the pioneering work on cancer stem cells was performed by John Dick's laboratory in Toronto. His work identifying stem cells in acute myeloid leukaemia provided the first evidence of a rare stem cell with a defined immunophenotype. Using fluorescence activated cell sorting (FACS) and xenotransplantation studies in NOD/scid mice, Dick's laboratory demonstrated enrichment of blasts with engraftment potential in the CD34++CD38- population. The frequency of stem cells in rose from 0.2-200 per million cells up to 10-50,000 per million cells in this population, but not in $CD34^{+}CD38^{+}$ or $CD34^{-}$ cells (Lapidot, Sirard et al. 1994; Bonnet and Dick 1997). Subsequent studies by this group and others have confirmed the presence of a defined and rare Leukaemia Propagating Cell (LPC) (Hope, Jin et al. 2004; Goardon, Marchi et al. 2011). However, experimental limitations may mean that this population is not as rare as was once believed (Taussig, Vargaftig et al. 2010).

In addition to the identification of the AML stem cell, the techniques developed in Dick's laboratory came to define the gold standard for functional assessment of "stemness". These and similar techniques were used to enrich for tumour-initiating capacity in a range of malignancies, including breast cancer (Al-Hajj, Wicha et al. 2003), colon cancer (Dalerba, Dylla et al. 2007; O'Brien, Pollett et al. 2007; Ricci-Vitiani, Lombardi et al. 2007), brain tumours (Hemmati, Nakano et al. 2003; Singh, Clarke et al. 2003), pancreatic cancer (Hermann, Huber et al. 2007; Li, Heidt et al. 2007), ovarian cancer (Zhang, Balch et al. 2008; Alvero, Chen et al. 2009; Curley, Therrien et al. 2009), prostate cancer (Collins, Berry et al. 2005), hepatic cancer (Ma, Chan et al. 2007), lung cancer (Ho, Ng et al. 2007; Eramo, Lotti et al. 2008) and gastric cancer (Fukuda, Saikawa et al. 2009; Takaishi, Okumura et al. 2009).

The importance of the experimental conditions has become increasingly well recognised, even within highly immunocompromised species. Syngeneic transplantation studies of murine leukaemia and lymphoma suggested that inter-species barriers may result in a falsely low CSC frequency (Kelly, Dakic et al. 2007). Key amongst the conditions to be optimised are use of the most immunocompromised mouse strains, particularly NSG mice (McDermott, Eppert

et al. 2010; Taussig, Vargaftig et al. 2010), and direct/orthotopic injection to avoid clearance of transplanted cells from peripheral blood, an important factor even in NSG mice (Taussig, Vargaftig et al. 2010). In addition, the gender of recipient mice can affect engraftment and so must be considered when allocating recipient groups (McDermott, Eppert et al. 2010).

Optimising these conditions can dramatically alter the frequency of tumour-initiating cells. A number of studies have questioned the degree to which the immunophenotype of cancer stem cells is restricted, particularly with regard to expression of CD133 (Joo, Kim et al. 2008; Ogden, Waziri et al. 2008; Wang, Sakariassen et al. 2008). In human melanoma, early studies using less developed xenotransplantation strategies identified a rare stem cell, enriched for by expression of ABCB5 (Schatton, Murphy et al. 2008) or CD271 (Boiko, Razorenova et al. 2010). However, by optimising the assay conditions, tumorigenic cells can be demonstrated at a frequency of 1 in 4 cells, irrespective of their immunophenotype (Quintana, Shackleton et al. 2008; Quintana, Shackleton et al. 2010).

Consequently, whilst the cancer stem cell model seems likely to hold true for a number of conditions, it will not describe the mechanism of tumour maintenance and initiation of relapse in all malignancies. Instead, other models of tumour heterogeneity may be more relevant for particular diseases.

Most recently, three studies have used lineage tracing techniques to demonstrate the presence of distinct populations responsible for populating both benign/pre-malignant (Driessens, Beck et al. 2012; Schepers, Snippert et al. 2012) and malignant (Chen, Li et al. 2012; Driessens, Beck et al. 2012) tumours. These three studies provide the first identification of a cancer stem cell *in situ*, rather than relying on xeno-transplantation studies. Chen *et al* were further able to demonstrate that the glioma CSC compartment was responsible for disease re-growth following tumour de-bulking with standard cytotoxic drug therapy (Chen, Li et al. 2012). Regrowth was reduced by combining cytotoxic drug therapy and “suicide gene” targeting of the CSC compartment, offering initial validity to the concept of targeting the CSC to prevent relapse.

1.2.4 Models of heterogeneity

1.2.4.1 Hierarchical versus Stochastic heterogeneity

For many, the existence of a cancer stem cell has come to imply the existence of a hierarchy. However, the mutually exclusive "hierarchical" and "stochastic" models are intended to describe possible origins of tumour heterogeneity, not stemness. As described in 1.2.1, the hierarchical model mirrors the normal tissue architecture, with a rare stem cell, such as the haematopoietic stem cell, sat at the apex of a hierarchy. Division and maturation of that stem cell results in progenitor/precursor cells with increasing degrees of differentiation but a loss of capacity for self-renewal/clonal expansion. It should be possible to identify and isolate the stem cell in this situation, whilst other populations would lack self-renewal. The stochastic model, however, describes a situation where all populations have an equal potential to self-renew, although most cells will not be capable of doing so at any one time. Instead, randomly applied factors, both intrinsic (genetic, epigenetic, transcriptional) and extrinsic (niche factors, immune response) result in stemness being a functional phenotype, rather than an innate one. As such, tumour-initiating capacity cannot be enriched for.

1.2.4.2 Clonal evolution

The acquisition of increasing numbers of genetic abnormalities throughout tumour initiation and progression has been known about for many years. Furthermore, the capacity for clonal evolution to underlie the development of therapy resistance has long been predicted:

> *"With variants being continually produced, and even increasing in frequency with tumor progression, the neoplasm possesses a marked capacity for generating mutant sublines, resistant to whatever therapeutic modality the physician introduces."(Nowell 1976)*

However, there has been increasing interest in this model recently, as modern techniques such as fluorescence in situ hybridisation (FISH) and next generation sequencing permit mutational analysis either at the single cell level or quantitatively within a mixed population. Such approaches have been used to develop phylogenetic trees, which in turn can predict the order in which mutations were accrued. These data may yield the answers to questions such as: 1) what is the cell of origin; 2) which are driver and which passenger mutations; 3) which mutations are key to therapy resistance.

Two recent studies have used differing approaches to dissecting the clonal evolution of ALL. Using multiplexed FISH analysis of diagnostic, relapsed and xenografted leukaemias, Mel Greaves' and Tariq Enver's laboratories have been able to analyse clonal complexity at a single cell level (Anderson, Lutz et al. 2011). They have shown not only simple linear evolution of clones, but much more complex branching patterns, with complex clonality at diagnosis. Relapse was initiated by either major or minor clones from diagnosis, and on occasion, by more than one (potentially unrelated) clone. In the absence of therapeutic pressure, serial engraftment in NSG mice largely recapitulated the clonal complexity of the diagnostic specimen.

The groups of John Dick and Jim Downing used an unbiased SNP array approach (Notta, Mullighan et al. 2011). They too demonstrated both linear and more complex patterns of clonal evolution, but their evidence suggested, at least for Philadelphia chromosome positive ALL, that a proportion of transplants are populated largely by minor diagnostic clones.

The common essential finding of both of these papers, however, was that leukaemia-propagating cells share in the genetic diversity displayed by the bulk leukaemia. This finding allows both the cancer stem cell and clonal evolution hypotheses to co-exist and interact. Whilst the hypothesis of clonal evolution does not dictate the presence or absence of a hierarchy, the greater diversity found in the absence of a hierarchy would provide a larger genetic pool for clonal evolution in ALL.

1.2.5 The genetics of leukaemia stem cells

The vast majority of work looking at the genetic determinants of leukaemic stem cell biology has been conducted in myeloid malignancy. In part this is due to the much clearer definition of the LSC in AML as compared to ALL, but also, the phenotypic similarity between normal haematopoietic stem cells (HSC) and AML stem cells (CD34+CD38-). One hypothesis is that the AML stem cell is directly derived from an HSC and as such, it is reasonable to consider that they may share a common stemness programme. Whilst this sounds straightforward it is worth noting that the determinants of HSC stemness are only beginning to be clearly defined. Furthermore, much of the underpinning work, as well as a number of studies modelling AML, have been conducted in mice and increasingly we are aware of substantial differences between the control of murine and human haematopoiesis, for review see (Doulatov, Notta et al. 2012).

Nevertheless, there is a substantial body of evidence demonstrating the importance of several pathways in HSC biology, including Wnt which acts in many different ways in what appears to be a highly dose dependent fashion (Luis, Ichii et al. 2012), TGFβ/Smad acting to maintain stem cell quiescence (Blank and Karlsson 2011) and Hedgehog (Mar, Amakye et al. 2011). However, to look at these pathways in isolation undoubtedly misses the detail provided by the interplay between them. A recent study grouped a large number of pathways involved in all stages of haematopoiesis into modules (Novershtern, Subramanian et al. 2011). Interestingly, this study showed that a number of modules were important in multiple and quite diverse branches of the haematopoietic hierarchy. This argues in favour of the importance of interactions, not only between pathways but probably also with the niche/cytokine milieu etc, in providing an overall environment which has been artificially compartmentalised into modules by the approach taken.

Despite the apparent difficulties, a number of candidate leukaemic stem cell programmes have recently been described. In a murine model of *MLL* induced leukaemia, stem cells expressed a genetic signature which was more similar to normal myeloid precursors and embryonic stem cells than to HSCs (Somervaille, Matheny et al. 2009). More recently, Goardon *et al* have examined LSC populations in primary human CD34+ AML (Goardon, Marchi et al. 2011). This study demonstrated hierarchically organised LSC potential in both lymphoid-primed multipotential progenitor-like (LMPP-like, CD38-CD45RA+) and granulocyte-macrophage progenitor-like (GMP-like) populations. Global expression profiles derived from these populations most closely matched the expression profiles derived from the equivalent populations in normal bone marrow, rather than normal HSC populations. However, these populations were enriched for the LSC "self-renewal" expression profile which has previously been determined in a murine model of *MLL-AF9* induced AML (Krivtsov, Twomey et al. 2006). In this setting, global expression profiles most closely matched that of normal GMP cells, but a 363 gene "self-renewal associated" signature was enriched in populations with self-renewal potential, namely HSCs and leukaemic GMPs .

A study contemporaneous to Goardon *et al* and again using primary human AML, found a negative correlation between the expression profile of functionally determined LSCs and that of comparable normal lineage committed populations (Eppert, Takenaka et al. 2011). Instead, Eppert *et al* found that the LSC profile correlated with that of normal cord blood HSCs. The differences in findings between these two studies may derive from the different immunophenotypic definitions used, or the alternative source of normal comparator material (bone marrow v cord blood). However, this study was also able to determine a limited (44 gene) expression signature which was common to both leukaemic and normal haematopoietic stem cells and was highly enriched for genes functioning in stem cell regulation or oncogenesis. Furthermore, they demonstrated a correlation between expression of this signature and clinical outcome, providing evidence of the relevance of the LSC to patients.

Whilst no comparable study has yet been published in ALL, our laboratory has recently performed an analysis of data from sorted CD34+ and CD34- lymphoblastic leukaemic populations, using the Eppert signature (Rehe, Wilson et al. 2012)(Appendix D). We have found that neither CD34+ nor CD34- ALL populations cluster with the CD34+CD38- AML cells or HSCs from Eppert *et al* using Principle Component Analysis. These data are consistent with our current model of ALL propagation which predicts that propagation is mediated by a lymphoid "self-renewal" programme rather than an HSC programme, see 1.2.6.

1.2.6 Leukaemia stem cells in acute lymphoblastic leukaemia

Unlike the clear demonstration of a hierarchy in AML, the nature of the leukaemia propagating cell (LPC) in ALL has been rather more controversial. Initial studies showed that in both high and standard risk leukaemias, populations with an immature $CD34^{+}CD19^{-}$ immunophenotype provided the only source of LPCs (Cobaleda, Gutierrez-Cianca et al. 2000; Cox, Evely et al. 2004). Subsequent studies, however, have failed to confirm these findings. Indeed, two groups have demonstrated in both high risk Philadelphia positive and standard risk ETV6/RUNX1 positive leukaemia, the population expressing the B lymphoid differentiation marker CD19, is the only one to harbour LPCs (Castor, Nilsson et al. 2005; Hong, Gupta et al. 2008). In the study conducted by Tariq Enver's laboratory, leukaemia propagating potential was restricted to an abnormal leukaemia associated immunophenotype, $CD34^{+}CD38^{-/low}CD19^{+}$ (Hong, Gupta et al. 2008).

More recently it has been demonstrated that a number of assay variables can have significant effects on the likelihood of transplantation. In particular, human cells which have been antibody labelled and sorted can be removed by the residual immune system of even the most immunocompromised strains currently available (Taussig, Vargaftig et al. 2010). This constraint may be bypassed to some extent by direct injection into the bone marrow, thereby avoiding peripheral immune effector cells (Taussig, Vargaftig et al. 2010). Additional factors which may reduce rates of engraftment include the strain of mice and specific antibodies used for sorting (Taussig, Vargaftig et al. 2010) and male gender (McDermott, Eppert et al. 2010). Failure to engraft may therefore have explanations other than an inherent lack of leukaemia propagating potential.

We have previously shown that using NSG mice, the most immunocompromised strain currently available, combined with direct intrafemoral injection, leukaemia propagation is not limited to $CD34^+CD19^-$, $CD34^+CD19^+$ or $CD34^-CD19^+$ populations (le Viseur, Hotfilder et al. 2008). This has been confirmed in a separate study by Kong *et al*, who again used NSG mice but transplanted via facial vein injection (Kong, Yoshida et al. 2008). We have also demonstrated that cells from the more mature $CD34^-CD19^+$ population were able to repopulate the entire leukaemia in recipient mice, including the "immature" haematopoietic stem cell like phenotype, $CD34^+CD19^-$. We called this phenomenon malleability in order to differentiate it from the "plasticity" associated with changes in lineage specificity.

Using a murine model of BCR/ABL1 positive $Arf^{-/-}$ ALL, Williams *et al* demonstrated the leukaemia propagating potential of (perhaps) every transformed pre-B cell (Williams, den Besten et al. 2007). These findings have been supported by studies of human leukaemias which have again shown a high frequency of LPCs in unsorted primary (Morisot, Wayne et al. 2010) and primografted (Schmitz, Breithaupt et al. 2011) leukaemia specimens. The high frequency of LPCs, combined with the presence of leukaemia propagation in a range of immunophenotypes argues against the presence of a stem cell hierarchy in ALL. However, the studies published to date have left a number of factors still to be addressed. In particular, the lack of evidence of sorted primary populations transplanted at limiting dilution leaves the possibility of a rare but potent hierarchical stem cell in ALL. Furthermore, studies have focused on high risk cytogenetic subgroups and therefore not determined whether these findings may apply more widely in B precursor ALL.

We have recently studied sorted patient derived blasts, both primary and passaged (primografted), from 13 patients with a variety of standard and high risk cytogenetic backgrounds (Rehe, Wilson et al. 2012)(Appendix D). Transplantation of cells at limiting dilution has demonstrated frequent LPCs, present in primary material between 1:40-1:2,900, increasing to as frequent as 1:6 in primografted samples. Furthermore, there was no difference in the LPC frequency or engraftment kinetics of blast populations sorted for CD10, CD20 or CD34.

In addition to studying patterns of engraftment, expression microarray analysis and quantitative reverse transcriptase PCR of sorted $CD34^+$ and $CD34^-$ blasts populations have demonstrated different transcriptomic patterns which mirror the corresponding normal B-precursor population including expression of IRF4, MS4A1/CD20, IgH constant locus, IgL κ and λ loci (le Viseur, Hotfilder et al. 2008). More recent analysis of the same data has applied putative leukaemia stem cell signatures derived from AML. If, as recent data suggest, there is no hierarchy in ALL and blasts from all populations have equal leukaemia propagating potential then such signatures should not be able to differentiate between these populations. Indeed, we found that whilst the signatures accurately separate immature CD34+CD38- AML blasts from more mature populations, CD34+ and CD34- ALL blasts clustered together, demonstrating an uncoupling of leukaemia propagating potential from differentiation in ALL (Rehe, Wilson et al. 2012)(Appendix D).

The results of the most recent studies, using more highly immunocompromised xenograft assays, strongly suggest that there is no hierarchy in acute lymphoblastic leukaemia. As discussed in 1.2.4.2, clonal evolution is not dependent on either the presence or absence of a hierarchy. Indeed, a situation in which new clones evolve across all immunophenotypes, each of which possesses leukaemia propagating capacity (Anderson, Lutz et al. 2011), argues in favour of a disease which lacks a hierarchy and in which LPCs are not-infrequent. This would provide an ideal milieu for the development of therapy resistant clones and so it is surprising perhaps that ALL, at least in childhood, is remarkably treatable.

1.2.7 The therapeutic relevance of cancer stem cells

Just as the cancer stem cell (hierarchical versus stochastic) and clonal evolution models are not mutually exclusive, so other biological properties add further levels of complexity into the behaviour of a cancer/leukaemia propagating cell, as recently reviewed in (Magee, Piskounova et al. 2012). Lineage plasticity, epithelial-mesenchymal/mesenchymal to epithelial transition, bidirectional interconvertibility (Gupta, Chaffer et al. 2009), malleability (le Viseur, Hotfilder et al. 2008), differences in cell of origin and differing kinetics to each of these factors make the number of potential disease, and even patient, specific cancer propagating properties so large, that for many cancers it seems unlikely that the original hope of specifically targeting the CSC therapeutically will ever be realised. Nevertheless, ensuring that novel therapies do effectively include CSCs in their target profiles, when such a population is present, will be an essential component of developing novel therapies.

The exception to this may be in malignancies where the maintenance of the disease relies upon a programme innate to the parent lineage, and not derived from a particular combination of genetic/epigenetic mutations. Lymphoid malignancy may present such a scenario, as lymphoid maturation is not accompanied by an obligate loss of self-renewal potential. In contrast to most other tissues, both progenitor and fully differentiated lymphoid cells retain the ability to clonally expand: progenitors following successful expression of the pre-B receptor at the cell surface; differentiated lymphocytes following exposure, and re-exposure, to antigen. It may therefore be possible to identify novel therapeutic targets from amongst the pathways governing lymphoid clonal expansion and its counterpart, clonal deletion. Furthermore, it might be hoped that these targets may hold more generalizable therapeutic potential for other diseases involving the lymphoid system, notably chronic lymphocytic leukaemia, myeloma and lymphoma.

One approach to identifying important elements of the clonal expansion programme in malignant lymphoid populations is to apply loss of function analyses. Experimentally, these can be performed by manipulating the endogenous post-transcriptional control mechanism known as RNA interference. This approach, described in 1.3, can either be used to investigate a candidate target or applied in a screening format, described in 1.4.

1.3 Small RNAs and RNA interference

1.3.1 Introduction to small RNAs

A substantial proportion of the non-protein coding genome is transcribed, resulting in an increasingly complex variety of small, and long, non-coding RNA species. Many of these species are expressed across a wide range of organisms and mediate diverse functions in a sequence-directed manner, for review see (Farazi, Juranek et al. 2008). The most well understood function of small, non-coding RNAs is to act as guide strands for RNA-mediated post-transcriptional gene silencing, also known as RNA interference (RNAi). This endogenous process, in mammals exemplified by microRNAs (miRNAs), allows the post-transcriptional control of gene expression by targeting complementary mRNA for translational repression or degradation (depending on the degree of complementarity between guide strand and target mRNA).

Just two endogenous RNA species have been identified in mammals - miRNAs and, more recently, PIWI-interacting RNAs (piRNAs). These two non-coding RNAs have distinct structures, see Table 1-1, and associate with different members of the Argonaute family of proteins which are also highly-conserved across species.

Table 1-1. Characteristics of miRNA and piRNA (adapted from (Farazi, Juranek et al. 2008)

	miRNAs	piRNAs
Expression	Ubiquitous	Germline
Precursor structure	Hairpin structure	?single stranded
Length	20-23	28-33
Modification	Unmodified	2'-O-methylation at 3' terminus
Function	mRNA degradation Translational repression	mRNA cleavage Control of methylation ?Histone modifications
Effector protein partners	Argonaute proteins	PIWI proteins

The Argonaute family comprises proteins with highly conserved domains known as PIWI/PAZ Domain (PPD) proteins, see Figure 1-3. The family is divided into two distinct clades, based on sequence homology: Argonaute clade proteins associate with miRNAs and represent core effectors within the well-studied classical RNA interference pathways; PIWI clade proteins associate with piRNAs and function via alternative pathways which are still not well understood.

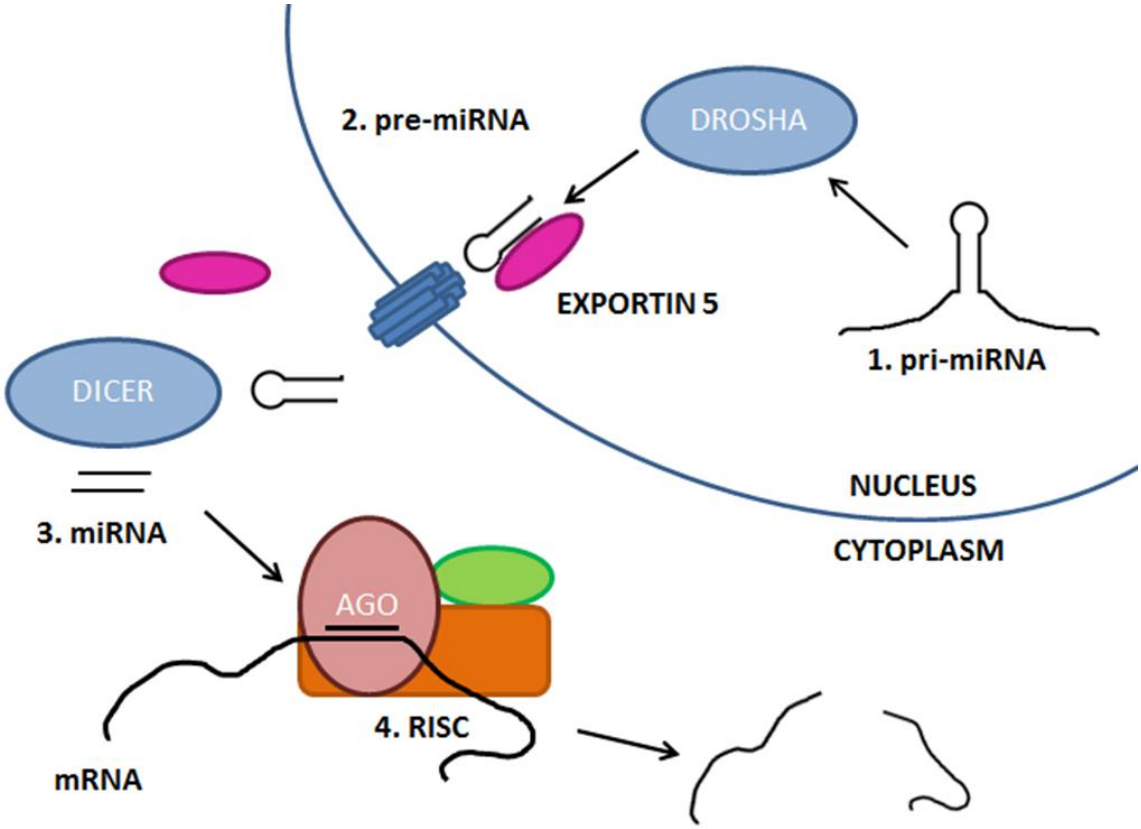
Figure 1-3. The structure of PIWI/PAZ Domain (PPD) proteins. PAZ domain – single stranded RNA binding pocket (Lingel, Simon et al. 2004; Ma, Ye et al. 2004); Mid domain – 5'-phosphate binding; PIWI – RNase H endonuclease domain (Song, Smith et al. 2004). The 34 amino acid PIWI Box (PB) is the most highly conserved region (Cox, Chao et al. 1998).

1.3.2 Argonaute clade proteins and microRNAs

The single stranded RNA guides which associate with Argonaute proteins in mammals are predominantly derived from miRNAs. The human genome contains approaching 500 miRNA encoding genes which are transcribed to give a primary miRNA (pri-miRNA) containing hairpin loops which will be processed to give between 1 and 6 mature miRNAs (Landgraf, Rusu et al. 2007) (Figure 1-4).

The pri-miRNA is bound by the double stranded RNA binding domain of the nuclear protein DGCR8, which subsequently associates with the RNase III protein DROSHA to form the nuclear Microprocessor complex (Gregory, Yan et al. 2004). DROSHA cleaves the pri-miRNA to produce a hairpin with a 2 nucleotide 3' overhang and 5' phosphorylation known as a precursor miRNA (pre-miRNA) (Lee, Jeon et al. 2002; Basyuk, Suavet et al. 2003; Lee, Ahn et al. 2003). The cleaved pre-miRNA is shuttled into the cytoplasm by the GTP dependent nucleocytoplasmic transporter, EXPORTIN 5 (Yi, Qin et al. 2003). In the cytoplasm the hairpin loop is cleaved by another RNase III protein, DICER (Hutvagner, McLachlan et al. 2001), again resulting in a 2 nucleotide 3' overhang and 5' phosphorylation (Elbashir, Lendeckel et al. 2001). These modifications allow binding to the PAZ and Mid domains of Argonaute proteins (see 1.3.1). One strand of the resultant mature miRNA, now 21-23 nucleotides in length, is loaded into a protein complex which includes an Argonaute protein, called the RNA-induced silencing complex, RISC (Hammond, Bernstein et al. 2000). Here, it acts as a guide strand for the binding of single stranded target mRNA.

Figure 1-4. Schema of miRNA processing. 1) Primary miRNA transcripts are generated from genomic regions and processed in the nucleus by the RNase III enzyme DROSHA to give precursor miRNAs (2). Precursor miRNAs are transported into the cytoplasm by EXPORTIN 5 where the loop is removed by another RNase III, DICER (3), to give a mature miRNA. The miRNA is loaded into the RNA Induced Silencing Complex (RISC, 4) where it acts as a guide strand for an Argonaute protein. Depending on the degree of complementarity, this may result in transcript cleavage (as shown) or sequestration and prevention of translation.



In humans, each miRNA may target as many as 100 transcripts (Brennecke, Stark et al. 2005; Lim, Lau et al. 2005), with a 6-8 nucleotide sequence at the 5' end of an miRNA, known as the seed region, targeting the 3' untranslated region of their respective mRNAs by (Lewis, Burge et al. 2005). Only rarely would an endogenous human miRNA have full complementarity to its target. The predominant method of miRNA modulation in humans is therefore inhibition of translation, rather than mRNA degradation, which relies on near complementarity and is seen more frequently in plants. mRNA bound into RISC is then prevented from being translated, or transported to mRNA processing bodies – P bodies, where it can either be degraded or returned for cytoplasmic translation (Parker and Sheth 2007).

1.3.3 PIWI clade proteins and PIWI-interacting RNAs

The second clade of PPD proteins consists of homologues of the Drosophila *P*-element *i*nduced *wi*mpy testis protein - Piwi. First identified in 1997, constitutional knockout studies showed that Piwi mutants had reduced numbers of adult germ cells, despite normal numbers of embryonic germline stem cells (Lin and Spradling 1997). It has subsequently been shown that this protein is essential for the maintenance of germline stem cells (Cox, Chao et al. 1998), and that this function is likely to be reliant on their role in suppressing the expression of retrotransposons, thereby protecting from integration associated mutational events (Aravin, Hannon et al. 2007).

1.3.3.1 PIWI clade proteins in mammals

Piwi homology has been described in a number of mammals. Amongst these are the murine *Miwi* (mouse Piwi), *Mili* (Miwi like) (Kuramochi-Miyagawa, Kimura et al. 2001; Deng and Lin 2002) and *Miwi2* (Carmell, Girard et al. 2007), and their human homologues *PIWIL1* (*PIWI-Like 1, HIWI*) , *PIWIL2 (HILI)*, *PIWIL4 (HIWI2)* and the additional *PIWIL3 (Sharma, Nelson et al. 2001; Sasaki, Shiohama et al. 2003)*. The most extensively studied amongst these are the murine homologues which, as in flies, are expressed in a developmentally controlled manner in testes (but not in ovaries) (Aravin, Sachidanandam et al. 2007). Knockout models again show varying blocks in spermatogenesis without an effect on testicular primordial germ cell numbers or ovarian gametogenesis (Deng and Lin 2002; Kuramochi-Miyagawa, Kimura et al. 2004; Carmell, Girard et al. 2007).

Analysis of the expression of PIWI-Like genes in humans is less thorough. Expression of *PIWIL4* seems to be ubiquitous (Sugimoto, Kage et al. 2007), whilst expression of *PIWIL2* and *3* is largely restricted to the testis (Sasaki, Shiohama et al. 2003). The best studied family member, at least in health, is *PIWIL1*. Within the testis, expression is restricted to germline cells, spermatocytes and spermatids (Qiao, Zeeman et al. 2002), but expression has also been demonstrated in heart, brain, kidney, pancreas, prostate, skeletal

muscle and CD34+ haematopoietic precursors, where expression is lost during *in vitro* differentiation culture (Sharma, Nelson et al. 2001).

1.3.3.2 Mammalian PIWI interacting RNAs - piRNAs

In 2006, a novel class of small RNAs was found to associate with PIWI clade proteins in both flies (Saito, Nishida et al. 2006; Vagin, Sigova et al. 2006) and mammals (Aravin, Gaidatzis et al. 2006; Girard, Sachidanandam et al. 2006; Grivna, Beyret et al. 2006; Lau, Seto et al. 2006; Watanabe, Takeda et al. 2006), the Piwi-interacting RNA (piRNA). Whilst in flies a heavy bias towards sequences derived from repeat elements supported a primary role in control of retrotransposons, mammalian piRNAs showed less specificity for repeat elements, suggesting more diverse mechanisms of action in mammals. Interestingly, the balance of sequences derived from repeat elements versus exonic elements varies during murine embryogenesis, suggesting also a dynamic function for these complexes (Aravin, Sachidanandam et al. 2007; Gan, Lin et al. 2011).

Tens of thousands of piRNA sequences are transcribed from a relatively limited number (50-100) of piRNA clusters (Aravin, Gaidatzis et al. 2006; Girard, Sachidanandam et al. 2006; Grivna, Beyret et al. 2006; Lau, Seto et al. 2006; Watanabe, Takeda et al. 2006). They are believed to be processed from long (25-35 kilobases) single stranded RNA precursors (Betel, Sheridan et al. 2007) by an unidentified endonuclease and 2'-O-methylated at the 3' end by HEN1 (Kirino and Mourelatos 2007). This methylation both stabilises piRNAs and contributes to their PIWI protein binding specificity (Tian, Simanshu et al. 2011).

1.3.4 Non-canonical RNAi pathways – PIWI/piRNAs functions

The original view of Piwi function in Drosophila germline stem cells was that of retrotransposon control. Whilst a novel mechanism is thought to underlie piRNA generation, known as the Ping-Pong cycle, the subsequent effector function is believed to be similar to canonical RNAi mediated by miRNAs, namely post-transcriptional mRNA degradation. As PIWI clade proteins have conserved RNase III domains it seems likely that this remains an important function in mammalian homologues of Piwi, but increasing evidence points towards additional functions of the PIWI/piRNA complex in mammals.

The most work on non-canonical PIWI/piRNA function exists in the field of methylation. Mammalian PIWI proteins are expressed during embryogenesis, and in particular during the period of *de novo* DNA methylation of embryonic germ cells (day 14.5 post coitum – day 2/3 after birth). This is a critical period for establishing methylation of transposons in the male germline, preventing the subsequent expression of these genes. Both *Mili* and *Miwi* knockout mice have reduced testis specific CpG methylation of retrotransposons and an associated increase in retrotransposon expression, a phenotype similar to that of DNA methylation deficient *dnmt3L* knockout animals (Aravin, Sachidanandam et al. 2007; Carmell, Girard et al. 2007; Kuramochi-Miyagawa, Watanabe et al. 2008). Furthermore, a role for piRNAs in the maintenance of paternal imprinting has been described (Watanabe, Tomizawa et al. 2011).

Whilst the vast majority of studies have concentrated on the function of PIWI proteins in the control of retrotransposons in germline specific stem cells, a recent study in *Aplysia*, the sea hare, has demonstrated a quite different function for PIWI/piRNA complexes. Thomas Tuschl's and Eric Kandel's laboratories identified piRNAs in small RNA libraries produced from *Aplysia* neurones (Rajasethupathy, Antonov et al. 2012). PIWI/piRNA complexes modulated the expression of the transcriptional repressor *CREB2* in response to the neuromodulatory agent serotonin, known to be involved in learning and memory. This effect was mediated by a sequence specific methylation of the *CREB2* promoter. This study provides the first example of the effect of

PIWI/piRNA directed CpG methylation on gene function, in addition to which it is an extragonadal effect.

A small number of less comprehensive studies have also suggested possible roles for PIWI proteins in maintaining histone modifications, both in mice (Wang, Han et al. 2011) and humans (Sugimoto, Kage et al. 2007).

1.3.5 PIWI/piRNAs in malignancy

Whilst the study of PIWI/piRNA function in humans has been limited by the apparent gonadal specificity of this system, there have been a number of studies looking at the expression of *PIWI* genes in human malignancy. The recent interest in cancer stem cells, combined with the germline stem cell specificity of PIWI proteins has raised the possibility that PIWI/piRNA complexes may be involved in modulating a stem cell programme in malignancy.

The expression of *PIWIL1* has been demonstrated in seminoma (Qiao, Zeeman et al. 2002), pancreatic adenocarcinoma (Grochola, Greither et al. 2008), and endometrial adenocarcinoma (Liu, Jiang et al. 2010). Furthermore, single studies have found expression of *PIWIL1* to be associated with poorer outcome in gastric carcinoma (Liu, Sun et al. 2006; Wang, Liu et al. 2012), glioma (Sun, Wang et al. 2011), low stage/non-metastatic colorectal carcinoma (Zeng, Qu et al. 2011), hepatocellular carcinoma (Zhao, Zhou et al. 2012), oesophageal carcinoma (He, Wang et al. 2009) and soft-tissue sarcoma (Taubert, Greither et al. 2007) whilst expression has been associated with tumour invasion in squamous cell carcinoma of the cervix (Liu, Jiang et al. 2010).

Whilst these studies have not looked at other members of the PIWI-Like proteins, additional studies have examined the expression of *PIWIL2* in malignancy. Similar to *PIWIL1*, expression of *PIWIL2* has been demonstrated in cervical carcinoma (Lee, Schutte et al. 2006; Feng, Peng et al. 2009; He, Chen et al. 2010), breast carcinoma (Lee, Jung et al. 2010) and most recently papillary thyroid carcinoma (Yin, Li et al. 2011).

These studies, however, have almost exclusively looked at the correlation between malignancy/stage and expression of PIWI-Like genes, often in a non-quantitative fashion. Consequently, little is known about the mechanism of action of PIWI-Like genes in malignancy. A small number of studies have, however, started to examine two potential mechanisms. Firstly is a proposed anti-apoptotic effect of *PIWIL2*. One of the initial papers demonstrating expression of *PIWIL2* in malignancy demonstrated a correlation between

expression of *PIWIL2* and expression of the anti-apoptotic pathway genes STAT3 and Bcl-X_L (Lee, Schutte et al. 2006). This study did not, however, demonstrate the mechanism by which *PIWIL2* might influence expression of these anti-apoptotic genes. A more recent study has demonstrated the association of PIWIL2 with STAT3 and c-src in a complex which appears to phosphorylate STAT3 (Lu, Zhang et al. 2012). This results in increased binding to the promoter of, and subsequent down-regulation of expression of, p53.

The second mechanism which has been investigated is the role of *PIWIL2* in the control of methylation. Siddiqi *et al* demonstrated a directly tumorigenic effect of a *PIWIL2* expression vector in mesenchymal stem cells, resulting in the production of soft tissue sarcomas (Siddiqi, Terry et al. 2012). This study also found that clones expressing *PIWIL2* showed global hypermethylation, whilst *PIWIL2* knockdown resulted in decreased growth and reduced DNA methylation. Interestingly, this study found no evidence of methylation changes specific to promoter CpG islands, although this may relate to the relatively short period of culture following transfection. Instead they postulate that the function of PIWIL2 directed methylation in sarcoma continues to be the suppression of retrotransposon expression.

To date, however, no studies have attempted to identify the expression of piRNAs in malignancy, and none have suggested a plausible link between the sequence specific role of PIWI/piRNA complexes and their function in malignancy.

1.4 Functional screening by shRNA library

1.4.1 Introduction to shRNA screens

High throughput RNAi screens have proved a powerful tool in identifying genes involved in diverse biological processes. The original screen structure tended to involve individual construct transfection of siRNAs in a 96-well format. Such screens have been enormously informative in *Drosophila* and other lower organisms, as well as mammalian systems. More recently, virally delivered, shRNA based libraries have allowed the development of massively parallel functional RNAi screens. This approach has identified existing and novel genes important to embryonic (Hu, Kim et al. 2009; Chia, Chan et al. 2010; Westerman, Braat et al. 2011) and haematopoietic (Ali, Karlsson et al. 2009; Hope, Cellot et al. 2010; Baudet, Karlsson et al. 2012) stem cell maintenance, neuronal synaptic development (Valakh, Naylor et al. 2012) and circadian clock modification (Zhang, Liu et al. 2009).

shRNA based functional screens are most obviously suited to identification of genes involved in cell proliferation and survival, and have therefore become an important tool in cancer research.

The general procedure for lentiviral based screening involves:

1. Transduction of target cells with lentiviral pools to give one construct per cell;
2. Harvesting of sample to provide a baseline for construct prevalence;
3. Period of culture to determine effect of gene knockdown, combined with antibiotic selection to remove untransduced cells;
4. Harvesting second sample for comparison with baseline.

The effect of knockdown of individual genes can be assessed by analysing the variation in construct prevalence between the baseline sample taken prior to knockdown and the second sample taken after a period of culture with knockdown.

1.4.2 Screening approaches – Positive and Negative screening

In broad terms, screens can be conducted to produce either positive selection or negative selection outputs. The former involves an assay in which the output is colonies which have been positively selected following exposure to an environmental pressure (e.g. irradiation, drug treatment, immunophenotypic selection etc) (Westbrook, Martin et al. 2005; Brummelkamp, Fabius et al. 2006; Gazin, Wajapeyee et al. 2007). This approach requires sequencing of individual clones and provides information only on genes deleterious to survival under selective pressure (constructs targeting these genes will be identified from proliferating clones). Negative selection involves analysis of the whole population at the end of the assay and therefore provides information on the prevalence of constructs relative to each other, including those which have become more, and those which have become less, prevalent (Schlabach, Luo et al. 2008; Silva, Marran et al. 2008). Information is therefore gathered on genes which are beneficial to cell growth and survival (in this instance, hairpin construct prevalence will decrease) as well as those which are deleterious.

1.4.3 The development of shRNA libraries

In order to identify the causative construct of a phenotype generated by a pooled/parallel RNAi screen, it is essential that: 1) each cell should receive only a single construct; 2) individual constructs can be identified at the end of the experiment.

The first constraint was addressed by development of microRNA adapted hairpins by the laboratories of Greg Hannon, Cold Spring Harbour Laboratories, and Stephen Elledge, Harvard University (Dickins, Hemann et al. 2005; Silva, Li et al. 2005; Stegmeier, Hu et al. 2005). By cloning the hairpin into the middle of a sequence derived from miRNA-30 (mir30), the construct is believed to follow a more natural pathway from transcription to RISC via processing by DROSHA and DICER. This results in more efficient processing, active export from the nucleus and active loading into RISC, see Figure 1-4. A substantially greater knockdown is achieved, making just a single lentiviral integration effective.

The second of these requirements has been satisfied by sequencing or array based analysis of a specific construct. The two principle approaches available are the analysis of either a barcode, unique to an individual shRNA construct, or analysis of the unique construct itself. Modern iterations of the second option tend to rely on analysis of a “half-hairpin” to avoid probe self-annealing during array analysis (Schlabach, Luo et al. 2008; Silva, Marran et al. 2008).

1.4.4 Limitations of shRNA screens

Genome-wide screens are so complex, that is to say they require accurate transduction of a very large number of constructs with a high degree of coverage, that to perform them in primary material is unfeasible. This presents a difficulty, as hits identified by the *in vitro* analysis may not remain valid in an *in vivo* situation. In a head-to-head comparison of *in vitro* and *in vivo* screening, Meacham *et al* found less than 10% of differentially represented hairpin constructs to be shared between the two situations (Meacham, Ho et al. 2009).

Another difficulty with the complexity of whole-genome screens is that constructs have not been prospectively validated. Instead, hairpin sequences are derived from algorithms which are designed to maximise knockdown whilst minimising off-target effects, especially those derived from the critical seed-region, see 1.3.2. Despite this, it is presumed that a proportion of hairpins will not achieve substantial levels of knockdown, and it has been demonstrated that at least 1% of hits in an siRNA screen is an off-target effect (Schultz, Marenstein et al. 2011). Recently, an analysis tool for predicting off-target effects from the seed-region has been described (Sigoillot, Lyman et al. 2012). Such approaches may help to limit false positive hits generated by off-target effects, although it may still be prudent to confirm that potential therapeutic targets are indeed false hits.

Finally, a small number of studies have demonstrated an innate toxicity to virally delivered shRNAs in the setting of murine models of disease. This effect can be modulated by expressing shRNAs from a miRNA context (McBride, Boudreau et al. 2008). Use of miRNA adapted hairpins is not without complications however. Two studies have shown the potential for shRNAmir constructs, which produce much higher levels of shRNA precursors than do native shRNAs, to saturate the endogenous RNAi machinery, probably at the level of nuclear export by Exportin 5, resulting in disturbed endogenous miRNA function (Grimm, Streetz et al. 2006; Pan, de Ruiter et al. 2011).

1.4.5 shRNA screening in cancer

The unbiased nature of genome-wide functional screening has allowed the identification of novel tumour-suppressor genes (Westbrook, Martin et al. 2005; Zender, Xue et al. 2008; Iorns, Ward et al. 2012) as well as potential therapeutic targets (Cole, Huggins et al. 2011; Zuber, Shi et al. 2011). RNAi screens can also be adapted to more complex assays. Within cancer research this includes synthetic lethality screens used to look for factors potentiating the effect either of an existing therapy (Azorsa, Gonzales et al. 2009), or of a commonly mutated gene such as RAS (Luo, Emanuele et al. 2009) or BRCA2 (Hattori, Skoulidis et al. 2011). The standard cellular survival screen can also be adapted to look at specific cell populations and has been applied to candidate brain tumour stem cell populations to identify genes essential for cancer stem cell survival and maintenance (Wurdak, Zhu et al. 2010; Goidts, Bageritz et al. 2012).

1.5 Hypothesis and objectives of the study

As has been discussed (see 1.2.5 and 1.2.6), lymphoblastic leukaemic blasts with propagating potential are frequent and are not arranged in a hierarchy. This frequency of propagating potential argues in favour of self-renewal being an intrinsic feature of all lymphoblasts which they are able to activate stochastically in response to intrinsic, or more probably extrinsic, cues such as niche availability. However, the genetic determinants of this biological characteristic are unknown.

The **hypothesis** underpinning this study is that:

> Acute lymphoblastic leukaemic blasts derive an intrinsic propagating potential from their lymphoid cell of origin

The **objectives** of this study were:

1) To investigate the role of the candidate stemness gene *PIWIL2* in the propagation of childhood acute lymphoblastic leukaemia
2) To perform an unbiased functional RNAi screen to identify genes involved in the propagation of childhood acute lymphoblastic leukaemia
3) To develop techniques to allow the validation of candidate regulators of lymphoblastic leukaemic propagation using patient-derived leukaemic blasts in an *in vivo* model of leukaemic propagation

Chapter 2

Materials & Methods

2.1 Approvals

2.1.1 Ethical approval for studies using patient-derived leukaemic material

Bone marrow and peripheral blood samples were collected at presentation, as part of the initial diagnostic investigations. Written informed consent was obtained from patients and/or legal guardians according to protocols approved by the corresponding review boards. Patient samples from the UK that were stored before September 2006 were exempt from specific consent for research by the Human Tissue Act. Samples were retrieved from Newcastle Haematological BioBank under the generic BioBank approval given by the Newcastle & North Tyneside Ethics Committee (REC reference number: 07/H0906/109).

Anonymised umbilical cord blood specimens, stored with consent for research, were obtained from the Newcastle Umbilical Cord Blood Bank, Department of Academic Haematology, Newcastle University.

Healthy peripheral blood samples were collected from volunteer donors working within Newcastle University.

2.1.2 Home Office approval for animal research

Animal experiments were performed by trained researchers who had all completed approved Home Office training at Newcastle University and held current Personal Licences under the Animals (Scientific Procedures) Act 1986. Intrafemoral injection and sampling was performed by one of three trained researchers from within the research group. All work was conducted in accordance with the Home Office Project Licence PPL60/3846.

2.2 Materials

2.2.1 Laboratory equipment

ABI Prism 7000 Sequence Detection System	Applied Biosystems
Avanti J-26 XP centrifuge	Beckman Coulter
Chromato-Vue UV trans-illuminator	Ultra-Violet Products Ltd
DMR Fluorescence microscope	Leica
E-410 digital SLR camera	Olympus
EPI2500 Electroporator	RL Fischer, Heidelberg
FACS Calibur	Beckton Dickinson
FACS Canto II	Becton Dickinson
Fluostar Omega plate reader	BMG Labtech
GelDoc Imager	BioRad
GeneAmp PCR System 9700	Applied Biosystems
IMT-2 phase contrast microscope	Olympus
LSM 700 confocal scanning microscope	Carl Zeiss Ltd
Mediphot937 X-ray film developer	Colenta Labortechnik
Model 680 microplate reader	BioRad
Nanodrop 1000 spectrophotometer	Labtech International
Optima L-100 XP ultracentrifuge	Beckman Coulter
RS320 X Ray System	Gulmay Ltd
SPOT RT digital CCD camera	SPOT Imaging
Spectrum IVIS bioluminescent camera	Caliper Life Sciences

2.2.2 Software

FACSDiva	Becton Dickinson
Graphpad Prism 4	Graphpad Software Inc.
Living Image 4.0	Caliper Life Sciences
Modfit LT	Verity Software House
Primer express 2.0	Applied Biosystems
SDS 2.0	Applied Biosystems
SPOT Advanced 4.0	SPOT Imaging Solutions
ZEN 2009	Carl Zeiss Ltd

2.2.3 Chemicals and reagents

2.2.3.1 General chemicals

General chemicals were purchased from Sigma unless otherwise stated.

2.2.3.2 Specific chemicals and reagents

Ampicillin	Sigma
Betaine 5 M solution	Sigma
Bovine pancreatic DNase I	Sigma
Carprofen	Pfizer
ccdB Survival™ 2 T1R competent cells	Invitrogen
Chloramphenicol	Sigma
Coomasie Brilliant Blue G-250	Thermo Scientific
DH5α competent cells	Invitrogen
Dimethylsulphoxide	Sigma
DNA loading buffer	Fermentas
dNTP set 100 mM	Fermentas
Ethidium bromide	Sigma
Ficoll-Paque	GE Life Sciences
Gateway pENTR1A dual selection vector	Invitrogen
GelRed	VWR International
IPTG	Sigma
JM109 competent cells	Promega
Kanamycin	Sigma
Lipofectamine LTX Plus	Invitrogen

Lymphoprep	Axis-Shield
Phosphate buffered saline (PBS)	GIBCO
Polybrene	Sigma
Propidium Iodide	Sigma
Protamine	Sigma
Protein Assay Reagent	BioRad
Puromycin	Sigma
RNase A	QIAGEN
Trypan blue	Sigma
XenoLight™ D-Luciferin	Caliper Life Sciences
X-Gal	Sigma

2.2.3.3 Experimental kits

Decode shRNAmir library	Thermo Scientific
DNA blood mini kit	QIAGEN
Endofree Plasmid Maxi Kit	QIAGEN
Gateway Vector Conversion System	Invitrogen
Gateway LR Clonase II enzyme kit	Invitrogen
miScript Reverse Transcription Kit	QIAGEN
miScript Primer Assay	QIAGEN
miScript SYBR Green PCR Kit	QIAGEN
Nucleospin miRNA kit	Macherey-Nagel
pGEM-T Easy Vector cloning kit	Promega
Platinum® SYBR® Green SuperMix UDG	Applied Biosystems
QIAquick PCR purification kit	QIAGEN
QIAprep Spin Miniprep Kit	QIAGEN
RevertAid™ H Minus cDNA Synthesis Kit	Fermentas
RNeasy Micro kit	QIAGEN
RNeasy Mini Kit	QIAGEN
SuperSignal West Dura	ThermoScientific
Wizard SV Gel and PCR clean-up kit	Promega

2.2.3.4 Enzymes

All restriction enzymes were purchased from Fermentas. Additional enzymes were purchased as follows:

Calf Intestinal Alkaline Phosphatase	Invitrogen
Klenow fragment	Fermentas
KOD Hot Start DNA polymerase	Novagen/Merck
Shrimp Alkaline Phosphatase	Fermentas
Vent polymerase	New England Biolabs

2.2.3.5 Cell culture media and supplements

Fetal calf serum	Sigma
Fetal calf serum	Gibco
Recombinant human IL-3	R & D Systems
Recombinant human IL-7	R & D Systems
Recombinant human stem cell factor	R & D Systems
RPMI1640	Sigma
RPMI1640	PAA

2.2.4 Buffers and media

2.2.4.1 DNA electrophoresis buffers

DNA loading buffer (6x)	0.25% w/v bromophenol blue 0.25% w/v xylene cyanol 20 mM EDTA 30% w/v glycerine
TAE	40 mM Tris 0.114% v/v acetic acid 1 mM EDTA pH 8
TBE	90 mM Tris 90 mM H_3BO_3 2 mM EDTA pH8
TE buffer	10 mM Tris·Cl, pH 8.0 1 mM EDTA

2.2.4.2 Western immunoblotting buffers

Urea buffer	9 M urea 4% w/v CHAPS 1% w/v DTT
Stacking gel buffer (4x)	500 mM Tris 0.4% w/v SDS pH 6.8
Separating gel buffer (4x)	1.5 M Tris 0.4% w/v SDS pH 8.8
Protein loading buffer (6x)	350 mM Tris, pH 6.8

	12 mM EDTA 500 mM DTT 0.012% w/v bromophenol blue 30% w/v glycerol
Electrophoresis buffer	25 mM Tris 190 mM glycine 0.1% w/v SDS
Transfer buffer	25 mM Tris 10% v/v methanol 190 mM glycine
TST	10 mM Tris 150 mM NaCl 0.1% v/v Tween 20
Coomassie blue stain	Coomassie G-250 0.1% w/v Acetic acid 10% v/v Methanol 40% v/v
Ponceau stain	Ponceau S 0.1% w/v Acetic acid 5% v/v

2.2.4.3 Bacterial culture buffers

LB medium	1% w/v tryptone 0.5% w/v yeast extract 1% w/v NaCl pH 7.5
LB agar plates	LB medium 1.5% w/v agar
SOC medium	2% w/v tryptone 0.5% w/v yeast extract 10 mM NaCl 2.5 mM KCl 20 mM Mg^{2+} 20 mM glucose
Calcium Chloride/Magnesium Chloride	80 mM Magnesium chloride 20 mM calcium chloride 0.2 µm filtered
Calcium chloride	0.1 M calcium chloride 0.2 µm filtered
X-Gal-solution	0.4% w/v NN-dimethylformamide

2.2.4.4 Cell culture media

Cells	**Medium**	**Supplements**
HS-5, Kasumi-1, M2-10B4, SEM, 833K	RPMI 1640, HEPES modification	2 mM L-Glutamine 10% v/v fetal calf serum
REH, 697	RPMI 1640	2 mM L-Glutamine 10% v/v fetal calf serum
293T	Dulbecco's Modified Eagle's Medium	4 mM L-Glutamine 1 mM sodium pyruvate 10% v/v fetal calf serum

2.2.4.5 Lentiviral transfection buffers

Calcium chloride 0.5 M	0.5 M CaCl 0.2 µm filtered
HeBS, 2x	0.28 M NaCl 0.05 M HEPES 1.5 mM NaOH pH 7.0 (critical 6.95<pH>7.05) 0.2 µm filtered
2.5 mM HEPES, pH 7.3	0.25% v/v 1 M HEPES pH 7.3 0.2 µm filtered

2.2.4.6 Cell purification and flow cytometry buffers

Citrate buffer	0.25 M Sucrose 40 mM sodium citrate pH 7.6
DNA staining and lysis buffer	20 µg/ml propidium iodide 0.5% v/v NP-40 0.5 mM EDTA PBS
Dilution buffer	2 mM EDTA PBS pH 7.2
Red cell lysis buffer	150 mM ammonium chloride 10 mM potassium bicarbonate 1 mM EDTA
DNase digest buffer	0.5 mM MgCl 1 mM CaCl PBS
MACS buffer	0.5% w/v bovine albumin 2 mM EDTA PBS pH 7.2 degassed in vacuum chamber

2.2.4.7 RNAi buffers

siRNA hybridisation buffer	100 mM NaCl 25 mM Tris pH 7.5

2.2.5 Oligonucleotide sequences

2.2.5.1 Quantitative Reverse Transcriptase PCR primers - mRNA

AML1/ETO	Fw:	5’ - AAT CAC AGT GGA TGG GCC C - 3'
	Re:	5' - TGC GTC TTC ACA TCC ACA GG - 3'
ANGPT1	Fw:	5’ - TCT CTT CCC AGA AAC TTC AAC ATC T - 3’
	Re:	5’ - TCA TGT TTT CCA CAA TGT AAT TCT CA - 3’
BMI1	Fw:	5’ - AAT CCC CAC CTG ATG TGT GT - 3’
	Re:	5’ - GCT GGT CTC CAG GTA ACG AA - 3’
CD34	Fw:	5’ - AAA GCA CCA ATC TGA CCT GAA AA - 3’
	Re:	5’ - CGA GGT GAC CAG TGC AAT CA - 3’
EZH2	Fw:	5’ - GGG CTC CAA AAA GCA TCT ATT G - 3’
	Re:	5’ - TGC ACA GGA TCT TTG ATA AAA ATC C - 3’
GAPDH	Fw:	5’ - GAA GGT GAA GGT CGG AGT C - 3’
	Re:	5’ - GAA GAT GGT GAT GGG ATT TC - 3’
HMGA2	Fw:	5’ - CCC AAA GGC AGC AAA AAC AA - 3’
	Re:	5’ - GCC TCT TGG CCG TTT TTC TC - 3’
MEIS1	Fw:	5’ - GCA TGC AGC CAG GTC CAT - 3’
	Re:	5’ - TAA AGC GTC ATT GAC CGA G - 3´
MLL/AF4	Fw:	5´ - CAG AAG CCC ACG GCT TAT GT - 3’
	Re:	5´ - GCA AAC CAC CCT GGG TGT TA - 3’
PIWIL1	Fw:	5’ - GGA TTT GTT GCC AGC ATC AAT - 3’
	Re:	5’ - AGG CAG ACT TTG AGC CCA TCT - 3’
PIWIL2	Fw:	5’ - AGG CAG AGG CCA TGT ATT TGG - 3’
	Re:	5’ - AAG CAT TTC CCG TTT CAG - 3’
PIWIL3	Fw:	5’ - GGA AAT TTG GAG AGC GCC ATA - 3’
	Re:	5’ - CCA TTC CAC TCT CCG CTC TTT - 3’
PIWIL4	Fw:	5’ - TGC TGG AAG TAC CTT CAT GGA - 3’
	Re:	5’ - GCT GCC AGT CTT GGG GAA AA - 3’

TBP	Fw:	5’ - CCT AAA GAC CAT TGC ACT TCG T - 3’
	Re:	5’ - GTT CGT GGC TCT CTT ATC CTC A - 3’
TERT	Fw:	5’ - GGA GAA CAA GCT GTT TGC GG - 3’
	Re:	5’ - AGG TTT TCG CGT GGG TGA G - 3’

2.2.5.2 Quantitative Reverse Transcriptase PCR primers - shRNA

shAGF1	5’ - TTC TCA GTA CGA TTT CGA GG - 3’
shANGPT1	5’ - TTA GTG CAA AGA TTG ACA AGG T - 3’
shMA6	5’ - TGG AGT AGG TCT GCT TTT CTT TT - 3’

2.2.5.3 siRNA sequences

siAGF1	sense	5’ - CCU CGA AAU CGU ACU GAG AAG - 3'
	antisense	3’ - UUG GAG CUU UAG CAU GAC UCU - 5’
siMA6	sense	5’ - AAG AAA AGC AGA CCU ACU CCA - 3'
	antisense	3' - UUU UCU UUU CGU CUG GAU GAG GU - 5'
siPIWIL2_5	sense	5’ - CUG CUA AUC UGG UAC GCA AUU - 3’
	antisense	*3’ - UUG ACG AUU AGA CCA UGC GUU - 5’*
siPIWIL2_e	sense	5’ - GGA UGA GUG UAC UAA GCU U - 3’
	antisense	3’ - AAG CUU AGU ACA CUC AUC C - 5’

2.2.5.4 Gateway cloning primers

pTRIPZ shRNAmir30 restriction cloning primers

Fw: 5' - GAG AGA GTC GAC CGA CCT CCC TAG CAA ACT GGG - 3'

GTC GAC – *Sal*I restriction site

Re: 5' - GAG AGA GAT ATC GAC TCA CTA TAG GGC CCG CC - 3'

GAT ATC – *EcoR*V restriction site

Entry vector sequencing reverse primer

5' - GTA ACA TCA GAG ATT TTG AGA CAC - 3'

Destination vector sequencing primer 1

5' - CAC ATT ATA CGA GCC GGA AGC AT - 3'

Destination vector sequencing primer 2

5' - CAG TGT GCC GGT CTC CGT TAT CG - 3'

2.2.6 Mammalian cell lines

The following derived mammalian cell lines were used either for cell culture experiments, for production of recombinant lentivirus, to condition medium or as a source of RNA for qRT PCR.

Table 2-1. Mammalian cell lines

Cell line	Origin	Cytogenetics
SEM	Relapsed childhood ALL	t(4;11)(q21;q23) – MLL/AF4
REH	Relapsed childhood ALL	t(12;21)(p13;q22) – TEL/AML1
697	Relapsed childhood ALL	t(1;19)(q23;p13) – E2A/PBX1
Kasumi-1	Relapsed childhood AML	t(8;21)(q22;q22) – RUNX1/RUNX1T1
SKNO-1	Relapsed young adult AML	t(8;21)(q22;q22) – RUNX1/RUNX1T1
833K	Metastatic testicular seminoma	
293T	Human embryonic kidney	Transformed by adenovirus 5 transfection
M2-10B4	Murine bone marrow stroma	Derived from a C57BL/6J X C3H/HeJ F1 mouse
HS-5	Human bone marrow stroma	Transformed by HPV-16 E6/E7 transfection
NALM6	Relapsed young adult ALL	t(5;12)(q33.2;p13.2) – ETV6/PDGFRβ
RS4;11	Relapsed adult ALL	t(4;11)(q21;q23) – MLL/AF4
MV4;11	Diagnosis childhood AML	t(4;11)(q21;q23) – MLL/AF4
THP-1	Relapsed childhood AML	t(9;11)(p22;q23) – MLL/AF9

2.2.7 Bacterial strains

The following bacterial strains were used either for the maintenance of lentiviral plasmids (2.3.7), during the TA cloning of qRT PCR amplicons prior to sequencing (2.3.8) or during the Gateway cloning project (2.3.12).

Table 2-2. Bacterial strains

Strain	Experimental use	Genetics
JM109	Maintenance of lentivectors TA cloning	*end*A1, *rec*A1, *gyr*A96, *thi*, *hsd*R17 (rk⁻, mk+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F′, *tra*D36, *pro*AB, *laq*IqZΔM15]
DH5α	Gateway cloning	F- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (r_k-, m_k+) *pho*A *sup*E44 λ- *thi*-1 *gyr*A96 *rel*A1
ccdB Survival 2T1R	Gateway cloning	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G *fhuA::IS2*

2.2.8 Antibodies

2.2.8.1 Fluorochrome labelled antibodies for flow cytometry

All antibodies were purchased from Becton Dickinson

Specificity	Clone	Fluorochrome
CD3	SK7	Fluorescein isothiocyanate (FITC)
CD14	M5E2	Phycoerythrin (PE)
CD15	MMA	FITC
CD19	SJ25C1	PE
CD19	SJ25C1	Allophycocyanin (APC)
CD20	L27	PerCP Cy5.5
CD34	8G12	APC
CD45 (murine)	30-F11	PE-Cy7
TER-119/Erythroid (murine)	TER-119	PE-Cy7

2.2.8.2 Primary antibodies for Western immunoblotting

Specificity	Species	Dilution	Manufacturer
PIWIL2	Mouse	1:500	Abnova
PIWIL2	Goat	1:100-1:25000	Santa Cruz
PIWIL2	Rabbit	1:200-1:2000	Abcam
GAPDH	Mouse	1:50000	HyTest

2.2.8.3 Secondary antibodies for Western immunoblotting

Specificity	Species	Dilution	Manufacturer
Anti-goat	Donkey	1:5000	Santa Cruz
Anti-mouse	Rabbit	1:2500-1:5000	GE Healthcare
Anti-rabbit	Goat	1:5000	Dako

2.3 Methods

2.3.1 General cell culture methods

2.3.1.1 Thawing viable cells

Viable cells were retrieved from liquid nitrogen and thawed rapidly using a water bath at 37^{o}C. Cells were carefully transferred to a 20 ml universal container. Nine volumes of the appropriate medium at 4^{o}C was added drop wise. Cells were centrifuged 300 x g, 10 minutes and the supernatant aspirated. Cells were resuspended in the appropriate medium, at the desired concentration and transferred to a culture flask or dish.

2.3.1.2 Cell counting using a Haemocytometer

An estimate of the cell concentration was made using an Improved Neubauer Counting Chamber. Cells in suspension were diluted 1:1 with trypan blue, or in the case of homogenised splenic samples, methylene blue to lyse red blood cells, and 8-10 µl placed into the chamber. Viable cells, being those which did not take-up trypan blue, within a quadrant etched on the chamber, were counted using an inverted phase contrast microscope. As the quadrant has a volume of 0.1 mm^3 the number of cells in 2 quadrants (allowing for a 1:1 dilution) multiplied by 10000, is the concentration of cells per millilitre.

2.3.1.3 Routine cell culture – suspension cells

Leukaemic cell lines SEM, REH, pre-B 697 and Kasumi-1 were cultured in upright culture flasks in an incubator set to 37^{o}C, 5% CO_2, fully humidified. Cells were counted every 2-3 days, see 2.3.1.2, and resuspended in fresh medium at a concentration between 5 and 8 x 10^5 cells/ml.

2.3.1.4 Routine cell culture – adherent cells

The adherent cell lines 293T, 833K, HS-5, and M2-10B4 were grown in treated culture dishes or flasks in an incubator set to 37^{o}C, 5% CO_2, fully humidified. Cells were split at confluence or for 293T, whilst still sub-confluent. To split cells, medium was aspirated, the plate washed with a volume of warmed PBS

equal to the initial volume of medium and then cultured at 37^{o}C with 1x trypsin (10x trypsin, Sigma, diluted with PBS) for 1-10 minutes, depending on the cell line. The volume of trypsin used was 2.5 ml for 100mm dishes and 5 ml for larger dishes and flasks. Following incubation, cells were released by gentle tapping and then the trypsin activity neutralised by addition of 7.5 ml or 20 ml culture medium containing 10% v/v fetal calf serum (FCS). The cells were resuspended by pipetting, diluted further with fresh medium and transferred to fresh dishes.

2.3.1.5 Production of conditioned medium

HS-5 cells were grown to near confluence in 100 mm culture dishes. The medium was aspirated and replaced with fresh alpha modified minimum essential medium (αMEM) supplemented with 10% v/v FCS. This medium was harvested after 48 hours and 0.45 µm sterile filtered. The conditioned medium was aliquoted and stored at -20^{o}C.

2.3.1.6 Freezing cells

Cell stocks were made by resuspending cells at 1 x 10^{7} cells/ml in ice-cold FCS with 10% v/v dimethylsulphoxide. Between 250 and 1000 µl of suspension was transferred to a cryovial and placed in an insulated freezing box. Cells were immediately transferred to a -80^{o}C freezer for 48-72 hours before being moved for long-term storage in a liquid nitrogen tank.

2.3.2 Specific cell culture methods

2.3.2.1 Expansion and growth arrest of M2-10B4 feeder cells

The murine bone marrow stromal cell derived cell line M2-10B4 was used for co-culture experiments with SEM cells. This cell line was chosen as the laboratory had experience of growth arresting these cells using irradiation.

M2-10B4 feeder cells were cultured in either 150 mm culture dishes or 175 cm^2 culture flasks in RPMI, 2 mM L-Glutamine and 10% v/v FCS. Cells were grown to confluence, trypsinised and reseeded at a 1:5 dilution, as described in 2.3.1.4.

To growth arrest cells, five dishes or flasks were trypsinised as described in 2.3.1.4. The cell suspension was centrifuged 300 x g, 10 minutes, the cells resuspended in 7.5 ml complete medium and carefully pipetted into a 75 cm^2 flask. This produced a suspension of 1 mm depth when the flask was laid down. Cells were irradiated for 24 minutes at 310 kV, 10mA to give a dose of 80 Gy. Initially, cells were incubated for 2 hours to allow recovery before being re-trypsinised. For later experiments, cells were immediately resuspended with fresh medium, without the need for trypsinisation. Resuspended cells were either seeded out immediately or frozen down for later use, see 2.3.1.6.

Thawed or freshly irradiated feeder cells were seeded into fresh 150 mm dishes or 175 cm^2 flasks at 5 x 10^6 cells per flask. This cell density provided a confluent/sub-confluent cell layer for co-culture.

2.3.2.2 Production of puromycin resistant feeder cells

To produce feeder cells which could be co-cultured in the presence of puromycin, M2-10B4 cells were stably transfected with a plasmid encoding the puromycin N-acetyl-transferase gene (PAC). Transfection was performed using Lipofectamine™ LTX, a cationic lipid preparation which complexes with DNA to form liposomes. Liposomes are endocytosed and the plasmid is subsequently released into the cytoplasm where it can be transcribed.

M2-10B4 cells were seeded at 30% confluence in a six well plate. The following day, Lipofectamine/DNA complexes were prepared as follows. Serum free Optimem medium, 500 µl, was mixed with 2.5 µg PAC plasmid, 2.5 µl Plus Reagent (supplied with Lipofectamine) and incubated at room temperature for 5 minutes. Between 3.5 – 15 µl Lipofectamine was added and the mixture shaken vigorously for 3 minutes and incubated at room temperature for 40 minutes. Each of the wells containing M2-10B4 cells was aspirated of medium and washed with 5 ml Optimem to remove FCS. The Lipofectamine/DNA mixture was then added, 500 µl per well, and the plate incubated for 5 hours at 37^{o}C, 5% CO_2. After 5 hours the Lipofectamine was aspirated off and replaced with 4 ml complete medium.

On day 3 post transfection, 1 µg/ml puromycin was added to the wells. This was increased to 2 µg on day 4 and maintained like this thereafter. A small number of colonies grew and were expanded by trypsinisation, initially into a 12 well plate and subsequently into larger dishes. These cells were maintained in 2 µg/ml puromycin thereafter.

2.3.2.3 Density gradient centrifugation

Density centrifugation was used to separate mononuclear cells from mixed cell populations in peripheral blood or to isolate viable cells from apoptotic cells/debris during the puromycin selection step of the RNAi screen, section 2.3.10.4. This process uses an iso-osmotic medium with a known density (1.077 g/ml) to separate viable mononuclear cells, which have a lower density, from erythrocytes, granulocytes and apoptotic cells which all have a higher density and thus pass through the separation medium during centrifugation. The mononuclear cell layer is then recovered from the interface between centrifugation medium and the original cell suspension medium and washed to provide purified, viable mononuclear cells.

Cells were used either as whole blood diluted 3:1 with dilution buffer (phosphate buffered saline (PBS) supplemented with 2 mM EDTA, pH 7.4), or as cultured cell lines resuspended in PBS. Peripheral blood, 35 ml, was carefully layered onto 15 ml Ficoll-Paque in a 50 ml polypropylene tube. Alternatively, 10 ml cultured cells were layered onto 8 ml Lymphoprep in a 20 ml Universal tube.

The layered cells were carefully transferred to a centrifuge to avoid mixing at the interface and centrifuged at 400 x g for 40 minutes (peripheral blood) or 20 minutes (cultured cells), ambient temperature with no brake applied. Viable cultured cells were then recovered by pipette and washed once in 20 ml PBS, 300 x g, 10 minutes and resuspended in culture medium. Peripheral mononuclear cells were recovered and washed three times in 50 ml dilution buffer: 300 x g, 10 minutes, 200 x g, 15 minutes, 200 x g 15 minutes. This wash step removed both retained Ficoll-Paque and platelets.

Granulocytes were isolated from the fraction which passed through Ficoll-Paque by resuspending the erythrocyte/granulocyte pellet in 2 ml dilution buffer and addition of 50 ml red cell lysis buffer (150 mM ammonium chloride, 10 mM potassium bicarbonate, 1 mM EDTA). After incubation for 10 minutes at room temperature, granulocytes were washed as described above for peripheral blood mononuclear cells.

Peripheral blood mononuclear cells and granulocytes were further purified by MACS immunomagnetic bead sorting as described in 2.3.2.5.

2.3.2.4 Thawing of umbilical cord blood specimens

Stored viable umbilical cord blood specimens in 10% v/v DMSO were kindly donated by the Newcastle Cord Blood Bank, Department of Academic Haematology, Newcastle University. Specimens were retrieved from liquid nitrogen and thawed in a water bath, 37^{o}C, until just a few ice crystals remained. One volume of RPMI, 10% v/v FCS, 1% v/v L-glutamine, 4^{o}C was added to the cord blood over 2 minutes to allow equilibration of DMSO, followed by a further 4 minutes of equilibration. The specimen was then transferred to a sterile 500 ml bottle and made up to 600 ml with medium RPMI medium supplemented as above, 4^{o}C, to dilute DMSO to 1% v/v. The cell suspension was then aliquoted into 50 ml polypropylene tubes and centrifuged at 100 x g, 15 minutes, 4^{o}C.

Cell pellets showed significant degree of clumping due to DNA released from lysed granulocytes which comprise a substantial proportion of a cord blood specimen. The pellets were therefore resuspended in 50 ml PBS, 0.5 mM $MgCl_2$, 1 mM CaCl and 0.5% w/v bovine serum albumin and 5000 units bovine

DNase1 added. Cells were incubated for 45 minutes at room temperature. Residual solid particles were allowed to settle over 2 minutes before the suspension was filtered through a 40 µm cell strainer. The solid particles were then washed once with 50 ml PBS which was also passed through the same cell strainer. Cells were then centrifuged at 200 x g, 15 minutes and resuspended in dilution buffer. If there was heavy red cell contamination, an ammonium chloride red cell lysis was performed as described in 2.3.2.3.

2.3.2.5 MACS microbead purification of cell populations

Immunomagnetic cell sorting using MACS microbeads, Miltenyi Biotec, was used to isolate CD3, CD14, CD15, CD19 and CD34 positive populations from peripheral blood or umbilical cord blood (CD34). CD133 labelling was also used to isolate SEM cells from M2-10B4 feeder co-cultures during the Decode RNAi screening experiments. This technique uses antibody conjugated microbeads to label cells expressing the desired antigen. Cells are then passed down a column containing ferrous spheres, held within a magnet. Microbead labelled cells are held within the column whilst non-labelled cells are washed away. Labelled cells were then recovered and can be used for on-going cell culture or nucleotide extraction.

Cell suspensions were prepared by density gradient centrifugation as described in 2.3.2.3 and 2.3.2.4, or by trypsinisation of SEM-GIPZ/M2-10B4 co-cultures, as described for adherent cells in 2.3.1.4. Cells were resuspended in MACS buffer (PBS, 2 mM EDTA, 0.5% w/v bovine serum albumin, degassed, 4^{o}C) at 10^7 cells in 80 µl buffer (CD3/14/15/19) or 10^8 cells in 300 µl buffer (CD34/133). FcR blocking agent was added to umbilical cord blood and SEM-GIPZ cells, 100 µl per 10^8 cells. Peripheral blood mononuclear cells from each donor were divided into aliquots, one for each of CD3/14/19 labelling. CD15 labelling was performed in cells isolated from the granulocyte population. Microbeads were then added to each aliquot, 20 µl per 10^7 cells (CD3/14/15/19) or 100 µl per 10^8 cells (CD34/133). Cells were incubated at 4^{o}C for 20 (CD3/14/15/19) or 30 minutes (CD34/133), and vortexed after ten minutes. Following incubation, cells were washed twice with 4 or 10 ml MACS buffer and centrifuged: peripheral leukocytes and SEM-GIPZ cells 300 x g, 5 minutes, 4^{o}C; umbilical

cord blood 200 x g, 10 minutes, 4^{o}C. Cells were resuspended in 500-1000 µl MACS buffer.

Selection columns were primed with 500 µl (MS column) or 3 ml (LS column) MACS buffer. Cells suspended in MACS buffer were added to the columns and the flow through collected. In the case of umbilical cord blood specimens, this unlabelled, CD34-, flow-through specimen was collected to provide a control population. The column was washed three times with 500 µl (MS) or 3 ml (LS) MACS buffer to remove unlabelled cells. The column was then removed from the magnet and cells eluted with either 1 ml (MS) or 5 ml (LS) MACS buffer. For CD34 labelled cord blood specimens, cells were passed down a second, primed LS column to further purify the CD34+ population.

Following isolation, cells were counted. In the case of umbilical cord blood an aliquot was labelled with anti-CD34 APC antibody as described in 2.3.9.6 and analysed using the FACS Calibur. Remaining cells were lysed with buffer RLT for extraction of RNA according to 2.3.5.1. In the case of isolated SEM-GlPZ cells, a pellet was frozen at -20^{o}C for subsequent DNA extraction according to 2.3.3.1.

2.3.3 Isolation and purification of DNA

2.3.3.1 Isolation of genomic DNA – Qiagen spin preparations

Genomic DNA was isolated from either cultured cell lines or patient derived leukaemic blasts using the QIAamp DNA Blood Mini kit, QIAGEN. The lysis step is followed by adsorption of DNA onto the silica membrane of the QIAamp spin column and subsequent washing steps to remove protein and other contaminants. Eluted DNA is predominantly 20-30 kilobases in length.

Cells were either taken fresh from *in vitro* culture, harvested fresh from a xenografted mouse or stored as a cell pellet at -20^{o}C. Protease K, 20 µl, and RNase A, 4 µl, were added to a 1.5 ml microfuge tube. Cells were resuspended in 200 µl PBS and added to the tube. Two hundred microliters buffer AL was added to the cell suspension and pulse mixed by vortex for 15 seconds. The sample was then incubated at 56^{o}C for 10 minutes. The microfuge tube was centrifuged and 200 µl ethanol added. The tube was again pulse mixed by vortex for 15 seconds. Having again centrifuged the microfuge tube, the lysate was applied to a QIAamp mini spin column and centrifuged at 6000 x g for 1 minute. The column was transferred to a clean collection tube. The column was washed with 500 µl buffer AW1, 6000 x g 1 minute, followed by 500 µl buffer AW2, 20000 x g, 3 minutes. The column was again transferred to a clean collection tube and centrifuged at 20000 x g, 1 minute. The collection tube was transferred to a 1.5ml microfuge tube and 200 µl of either nuclease free water (Sigma) or buffer AE was pipetted onto the membrane. The column was incubated at room temperature for between 1 and 5 minutes before being centrifuged at 6000 x g, for 1 minute. DNA concentration and purity were assessed by spectrophotometry using a Nanodrop 2000, following which the DNA was stored at -80^{o}C.

2.3.3.2 Purification of PCR products – Qiagen spin preparations

QIAquick PCR purification kit, QIAGEN, was used to purify PCR amplification products, removing residual primers, nucleotides, enzymes and salts. The resultant amplicons were suitable for subsequent cloning.

Five volumes of buffer PB were added to 1 volume of PCR sample, pipette mixed and transferred to the QIAquick spin column. The column was centrifuged 20000 x g, 1 minute and the flow-through discarded. The membrane was washed with 750 µl buffer PE, centrifuged 20000 x g, 1 minute and the flow-through discarded. The column was placed in a fresh collection tube and centrifuged again at 20000 x g, 1 minute. The spin column was transferred to a clean 1.5 ml microfuge tube, 50 µl buffer EB added to the membrane and incubated at room temperature for 1 minute. The column was centrifuged 20000 x g, 1 minute. DNA concentration and purity were assessed by spectrophotometry using a Nanodrop 2000 and stored at -80^{o}C, or the DNA was used directly for TA cloning, see 2.3.8.

2.3.3.3 Purification of PCR/restriction digest products for cloning

During the cloning of the shRNAmir30 constructs into the Gateway Entry vector, restriction digestion of shRNAmir30 amplicons and the pENTR1a dual selection vector was followed by separation of the residual vector by gel electrophoresis, see 2.3.12.3. The desired bands were visualised using a Chromato-Vue UV trans-illuminator, cut out into a weighed 2 ml microfuge tube, reweighed and the DNA isolated using the QIAquick gel extraction kit. For DNA extraction all centrifuge steps were carried out at 18000 x g for 1 minute.

Three volumes of buffer QG were added to 1 volume of gel (volume to weight) and incubated at 50^{o}C until the gel slice was completely dissolved. Isopropanol equivalent to the volume of the initial gel was added and pipette mixed. The sample was then added to the QIAquick column, centrifuged and the flow-through discarded. Five hundred microliters buffer QG was added to the column, centrifuged and the flow-through discarded. The membrane was washed by adding 750 µl buffer PE, centrifuging and discarding the flow-through. The spin column was centrifuged again to remove residual ethanol and the column transferred to a fresh 1.5 ml microfuge tube. DNA was eluted by addition of 30-50 µl buffer EB directly to the membrane, incubation at room temperature for 1 minute and centrifugation. DNA concentration and purity were assessed by spectrophotometry using a Nanodrop 2000 and stored at 4^{o}C prior to use for cloning, see 2.3.12.3.

2.3.3.4 Purification of amplified Decode RNAi screen barcode region

Decode RNAi library barcode regions, amplified by hot start PCR, were purified by gel electrophoresis. PCR product (100 µl) was mixed with 25 µl 5x DNA loading dye (Fermentas) and run in multiple wells of a 1.25% w/v agarose/TBE gel with 1:10000 GelRed DNA stain. The desired band at 250 bp was visualised using a Chromato-Vue UV trans-illuminator, cut out into a weighed 2 ml microfuge tube, reweighed and the DNA isolated using the Wizard SV gel and PCR clean-up kit. For DNA extraction all centrifuge steps were carried out at 16000 x g for 1 minute.

One volume of Membrane Binding Solution was added to 1 volume of gel (volume to weight) and incubated at 56^{o}C until the gel slice was completely dissolved. This process was aided by intermittent vortexing. Seven hundred microliters of dissolved gel were added to an SV Mini column, incubated at room temperature for 1 minute and centrifuged. The flow-through was discarded. Further aliquots of dissolved gel from the same time-point were added to the column up to a total of 7 ml. Having loaded all the dissolved gel onto the column, the membrane was washed with 700 µl Membrane Wash Solution and centrifuged. The column was transferred to a clean collection tube and centrifuged to remove any residual ethanol. The column was transferred a clean 1.5 ml microfuge tube and 50 µl nuclease free water was added directly to the membrane. The column was incubated for 1 minute and centrifuged. DNA concentration and purity were assessed by spectrophotometry using a Nanodrop 2000 and stored at 4^{o}C prior to shipping for microarray analysis.

2.3.4 Transient RNA interference using siRNA

2.3.4.1 Introductory principles

The normal cellular RNA interference (RNAi) machinery comprises of short non-coding RNA sequences, principally microRNAs in mammals, and effector proteins. These effectors proteins include the endoribonucleases, Drosha and Dicer, which process microRNAs and load them into the RNA Induced Silencing Complex, RISC. One of the two strands of a processed non-coding RNA is bound to the RNA binding domain and acts as a guide for RISC to bind target mRNA sequences and either degrade them or prevent translation, depending on the degree of complementarity.

For the current experiments, synthesised siRNAs were delivered by electroporation of suspension cells. This method offers efficient delivery with minimal toxicity to the cell lines studied. siRNAs are taken up by the endogenous RNAi machinery which then degrades complementary mRNA, providing a mechanism for experimental post-transcriptional modulation of target genes.

2.3.4.2 Electroporation

Cells were counted and resuspended at 1×10^7 cells/ml. The desired number of cells was transferred to an electroporation cuvette with a 4 mm electrode gap (Eurogentec or Peqlab). The desired volume of siRNA was added, the cuvette gently shaken and then locked into the electroporator. A square wave electrical impulse was delivered over 10 ms, ranging from 330 – 350 V. The cuvette was incubated at ambient temperature for a minimum of 15 minutes before the cells were resuspended at 0.5×10^6/ml and transferred back to the incubator for on-going culture.

For serial electroporation experiments, cells were counted after 48-72 hours, aliquots taken for RNA isolation or analysis as described in 2.3.4.3 and 2.3.4.4, whilst the required number of cells was resuspended at 1×10^7 cells/ml and re-electroporated as described above.

2.3.4.3 Analysis of cell proliferation

To analyse cell proliferation, cells were counted prior to each electroporation. The total proliferation was then adjusted according to the number of cells carried forward to the subsequent electroporation, using the equation:

$$N(B)corrected = N(B) \times \{N(A)/N(A\ elec)\}$$

Where: N(B) = Number of cells counted at time point B

N(A) = Number of cells counted at time point A (pre first electroporation)

N(A elec) = Number of cells electroporated at time point A

This same equation was applied to each time point to create a corrected running cell count, from which the kinetics of cell proliferation could be calculated.

2.3.4.4 Analysis of cell cycle

Cell cycle analysis was performed using propidium iodide staining of DNA. This technique uses the variation of cellular DNA content during the cell cycle to attribute cells to: G0/G1 – diploid cells; S phase intermediate ploidy; G2 – tetraploid cells. The degree of propidium iodide staining was assessed by flow cytometry and the data analysed using ModFit LT software.

Between 2.5-5 x 10^5 cells were washed with 3 ml PBS, 300 x g, 4 minutes. The supernatant was aspirated and the cells resuspended in 100 µl citrate buffer. To this, 400 µl DNA staining/lysis buffer and 1 µl RNase A were added and incubated at 4^{o}C, in the dark, for a minimum of 30 minutes. Cells were analysed using a FACS Calibur flow cytometer, with 20000 events recorded.

2.3.5 Gene expression analysis using quantitative reverse transcribed polymerase chain reaction

2.3.5.1 Isolation of RNA – RNeasy method

Isolation of mRNA was performed using the RNeasy kit from QIAGEN which binds RNA to a silica membrane. This is followed by washing steps to remove protein and buffer constituents and subsequent elution of purified RNA. For small samples, <5 x 10^5 cells, the RNeasy micro kit was used, see 2.3.5.2, as this allows elution into very small volumes, maximising RNA concentration.

Between 0.5-5 x 10^6 cells were washed in PBS and lysed with 350 µl buffer RLT. The lysate was loaded into a QIAshredder column and centrifuged 20000 x g, 2 minutes. Homogenised lysate was either stored at -20^{o}C or used immediately. One volume (350 µl) of 70% ethanol was added to the lysate, the sample loaded into an RNeasy spin column and centrifuged 10000 x g, 30 seconds. For isolation of cellular protein, the flow-through from this step was retained and made up to 2 ml with ice-cold acetone and stored at -20^{o}C. Protein was isolated according to 2.3.6.1. The membrane was then washed in three steps: 700 µl buffer RW1, centrifugation 10000 x g, 30 seconds; 500 µl buffer RPE, centrifugation 10000 x g, 30 seconds; 500 µl buffer RPE, centrifugation 10000 x g, 2 minutes. The column was then transferred to a fresh collection tube and centrifuged 20000 x g, 1 minute to remove residual ethanol. The column was transferred to a fresh collection tube and the RNA eluted in 25 - 50 µl nuclease free water. RNA purity and concentration were assessed by spectrophotometry using a Nanodrop 2000, following which the RNA was stored at -20^{o}C.

2.3.5.2 Isolation of RNA – RNeasy micro method

Less than 5 x 10^5 cells were resuspended in 350 µl buffer RLT Plus, loaded into a QIAshredder column and centrifuged 20000 x g, 2 minutes. One volume (350 µl) of 70% ethanol was added to the homogenised lysate, the sample loaded into an RNeasy micro column and centrifuged 10000 x g, 30 seconds. The membrane was then washed in three steps: 700 µl buffer RW1, centrifugation 10000 x g, 30 seconds; 500 µl buffer RPE, centrifugation 10000 x g, 30 seconds; 500 µl 80% ethanol, centrifugation 10000 x g, 2 minutes. The column was then transferred to a fresh collection tube and centrifuged with the lid open at 20000 x g, 5 minutes to remove residual ethanol. The column was transferred to a fresh collection tube and the RNA eluted in 14 µl nuclease free water. RNA purity and concentration were assessed by spectrophotometry using a Nanodrop 2000, following which the RNA was stored at -20^{o}C.

2.3.5.3 Isolation of small RNA – NucleoSpin method

For analysis of shRNA expression an alternative method was required to isolate RNA fragments of less than 200 nucleotides. The NucleoSpin kit again uses a similar principle of binding nucleic acids to a silica membrane with an additional on column DNA digest step to remove residual DNA. Depletion of proteins and alteration of buffer conditions then allows binding of short RNA (18-200 nucleotides) to the silica membrane. Whilst the two RNA fractions (<200 and >200 nucleotides) can be isolated using separate columns, for the present study they were bound to the same membrane and eluted as a single fraction.

Between 1-5 x 10^6 cells were washed once in PBS and lysed by addition of 300 µl buffer ML. Lysate was loaded into NucleoSpin Filter column and centrifuged 11000 x g, 1 minute. The lysate was either stored at -20^{o}C until subsequent RNA isolation or used immediately.

Room temperature lysate was mixed with 150 µl ethanol and vortex mixed for 5 seconds. The sample was incubated at room temperature for 5 minutes before being loaded into a NucleoSpin RNA column and centrifuged 14000 x g, 1 minute. The flow-through, containing protein and small RNA, was retained. The RNA column, with bound large RNA, was washed once with 350 µl buffer

MDB, 11000 x g, 1 minute. An on column DNA digest was performed by addition of 150 µl recombinant RNase free DNase (provided with the kit), for 15 minutes at room temperature.

During the DNA digest step, the flow-through containing short RNA was cleared of protein by addition of 300 µl buffer MP and vortexing for 5 seconds. The sample was loaded into a Nucleospin Protein Removal Column and centrifugation 11000 x g, 1 minute. The binding conditions of the flow-through were adjusted by addition of 800 µl buffer MX and vortexing 5 seconds. The sample was added in 600 µl aliquots to the RNA column containing bound long RNA and centrifugation 11000 x g, 30 seconds.

The RNA column with bound long and short fragment RNA was washed in three steps: 600 µl buffer MW1, centrifugation 11000 x g, 30 seconds; 700 µl buffer MW2, 11000 x g, 30 seconds; 250 µl buffer MW2, centrifugation 11000 x g, 2 minutes. The column was transferred to a fresh collection tube and centrifuged 11000 x g, 1 minute. RNA was then eluted into a fresh collection tube in 30-50 µl nuclease free water. RNA purity and concentration were assessed by spectrophotometry using a Nanodrop 2000, following which the RNA was stored at -20^{o}C.

2.3.5.4 Preparation of cDNA from mRNA by reverse transcription

RNA was reverse transcribed to produce cDNA using the RevertAid™H Minus First Strand cDNA Synthesis Kit. Up to 1 µg RNA was added to each reaction along with 1 µl random hexamer primers and DEPC treated water to a total of 12 µl. This initial reaction was incubated at 70^{o}C for 5 minutes to disrupt the secondary structure of the RNA. After cooling to 4^{o}C, 8 µl of master mix was added to each reaction as follows: 5x reaction buffer, 4 µl; 10 mM dNTPs, 2 µl; Ribolock RNase inhibitor, 1 µl; RevertAid H minus reverse transcriptase, 1 µl. The total 20 µl reaction was the reverse transcribed under the following conditions: 25^{o}C, 10 minutes; 42^{o}C, 60 minutes; 70^{o}C, 10 minutes; held at 4^{o}C. cDNA was diluted for use using 30 µl DEPC treated water.

2.3.5.5 Plating of RT PCR for mRNA quantitation

Quantitative reverse transcribed polymerase chain reaction (qRT PCR) was performed in a 384 well format using an ABI Prism 7000 Sequence Detection System. Each well was pipetted with 8 µl of PCR mastermix as follows: 5 µl 2x Platinum® SYBR® Green SuperMix UDG; 0.3 µl 10 µM primer mix; 2.7 µl sterilised deionised water. One exception was the qRT PCR analysis of the MLL/AF4 fusion transcript in SEM cells for which the mastermix contained 0.1 µl 5 µM primer mix; 2.9 µl sterilised deionised water. This lower primer concentration was used to avoid the formation of primer dimers. For each primer pair, 2 µl cDNA was pipetted into triplicate wells. The plate was sealed with an adhesive cover and centrifuged 2000 x g, 30 seconds. The reaction was performed as follows: 50^{o}C, 2 minutes; 95^{o}C, 10 minutes; 60^{o}C, 15 seconds; followed by 40 cycles of 95^{o}C, 15 seconds; 60^{o}C, 1 minute; followed by a single cycle of 95^{o}C, 15 seconds; 60^{o}C 15 seconds; 95^{o}C 15 seconds. The completed plate was either discarded or stored at -20^{o}C prior to analysis of the product by gel electrophoresis or TA cloning, see 2.3.8.

2.3.5.6 Preparation of cDNA from shRNA by reverse transcription

cDNA for analysis of shRNA expression was reverse transcribed using miScript reverse transcription kit. This kit uses a mixture of poly(A) polymerase and a reverse transcriptase so that miRNAs, which are not polyadenylated, are both polyadenylated and reverse transcribed in parallel. Reverse transcription can then by primed by oligo-dT primers which also have a universal tag sequence which allows binding of a universal primer during the PCR step, see 2.3.5.7.

Each reaction contained 4 µl reaction buffer (5x), 1 µl reverse transcriptase mix, 1 µg RNA and DEPC treated water to a total volume of 20 µl. The reaction was incubated at 37^{o}C for 60 minutes followed by 95^{o}C for 5 minutes. The reaction was made up to 50 µl with DEPC treated water and stored at -20^{o}C prior to qRT PCR.

2.3.5.7 Plating of qRT PCR for shRNA quantitation

qRT PCR for analysis of shRNA expression was performed in a 384 well format using an ABI Prism 7000 Sequence Detection System. Each well was pipetted

with 8 µl of PCR mastermix as follows: 5 µl 2x QuantiTect SYBR Green PCR mastermix; 1 µl miScript Universal primer (10x); 0.3 µl specific primer (10 µM), see 2.2.5.2; 1.7 µl sterilised deionised water. cDNA produced according to 2.3.5.6 was diluted further 1:10. For each primer pair 2 µl of cDNA was pipetted into triplicate wells. The plate was sealed with an adhesive cover and centrifuged 2000 x g, 30 seconds. The reaction was performed as follows: 95^{o}C, 15 minutes; followed by 40 cycles of: 94^{o}C, 15 seconds; 55^{o}C, 30 seconds.

2.3.5.8 Analysis of data

qRT PCR data was analysed using SDS 2.0 software to produce the cycle threshold (Ct) value for each replicate. The mean Ct value was calculated for each set of triplicates, with an acceptable standard deviation of 0.3 cycles. A ΔCt value was generated by subtracting the Ct of the reference gene (Ref), commonly GAPDH, from the Ct of the gene of interest (GOI) to provide an internally controlled value for gene expression. This ΔCt value could then be compared between control and experimental samples to provide a comparison of the level of expression of the gene of interest expressed as the number of PCR cycles, the ΔΔCt.

$$\Delta Ct = (Ct^{GOI} - Ct^{Ref})$$

$$\Delta\Delta Ct = \left(\Delta Ct^{Control} - \Delta Ct^{Experiment}\right)$$

A ratio of the expression of the gene of interest in the experimental sample compared to the control sample was derived as follows, assuming that each cycle of the reaction resulted in a doubling of the DNA content of the well:

$$Ratio = 2^{\Delta\Delta Ct}$$

2.3.6 Western immunoblotting

2.3.6.1 Isolation of cellular protein

Protein was precipitated by addition of at least two volumes of acetone at -20^{o}C to the flow through recovered from RNA isolation as detailed in 2.3.5.1 and 2.3.5.2. Precipitate was centrifuged at 16000 x g, 4^{o}C for 15 minutes. Acetone was aspirated and the precipitate centrifuged further for 1 minute to allow removal of residual acetone. The protein pellet was air dried for 5-10 minutes before being resuspended in 30 µl of 9 M urea buffer. Protein suspensions were stored at -20^{o}C for quantification and Western Blot analysis.

2.3.6.2 Polyacrylamide gel electrophoresis

Polyacrylamide gels were cast with a thin stacking gel poured on top of a 8% separating gel, as follows. Separating gel buffer was diluted to 1x with deionised water and acrylamide-bis-acrylamide, final concentration 8% w/v. This was thoroughly mixed. To this were added 0.1% w/v ammonium persulphate and 0.01% w/v TEMED and the gel poured into a 0.75 mm casting chamber (BioRad). Two hundred microliters isopropanol was pipetted onto the surface of the separating gel to give a uniform surface against which to pour the stacking gel.

Once the separating gel was set, the isopropanol was poured off and the 4x stacking gel buffer was diluted with deionised water and acrylamide-bis-acrylamide, final concentration 4% w/v. After mixing, 0.1% w/v ammonium persulphate and 0.01% w/v TEMED were added and the gel poured onto the separating gel.

Gels were loaded into an electrophoresis tank which was filled with electrophoresis buffer. Either marker or a protein sample, 10 µg, was loaded into each well. Electrophoresis was performed at 50 V until the samples had passed into the separating gel. The voltage was then increased to 150 V and the sample run until sufficient separation of the marker bands was achieved.

2.3.6.3 Blotting procedure

Protein samples were transferred from the electrophoresis gel onto a PVDF membrane as follows. The membrane was wetted in methanol and then placed into transfer buffer (10% v/v methanol). The separating gel was placed against the membrane and clamped in a transfer cassette with filter paper and sponges to ensure close apposition. The assembly was performed in transfer buffer to minimise the risk of air bubbles being trapped between gel and membrane.

The transfer cassette was placed in a transfer chamber in transfer buffer with an ice block to minimise heating of the buffer. The transfer was performed at 150 V for 1 hour.

2.3.6.4 Immunolabelling

Following transfer, the membrane was cut along the 70 kDa markers. The membrane with proteins greater than 70 kDa was then cut to allow labelling with different primary antibody concentrations. All sections of the membrane were incubated in 5-10% w/v milk in TST for 1 hour to block non-specific protein binding sites. The membrane sections were then incubated in primary antibodies diluted in 5-10% milk in TST as detailed in 2.2.8.2, 4^{o}C, overnight, with the lower, <70 kDa, membrane being labelled with anti GAPDH antibody.

Following overnight incubation, membranes were washed three times in TST for 5-10 minutes. Washed membranes were then incubated in secondary antibodies, detailed in 2.2.8.3, diluted in 5-10% milk in TST, at room temperature for 1 hour. Membranes were again washed three times in TST for 5-10 minutes.

2.3.6.5 Membrane development

Development of horseradish peroxidase conjugated secondary antibodies was performed using SuperSignal West Dura chemiluminescence substrate kit. This kit contains two components, a peroxidase solution and a luminol solution. These were mixed 1:1 and pipetted onto a smoothed piece of clingfilm. Blots were placed protein side down onto the mixture and incubated at room temperature for 5 minutes, except for GAPDH which was incubated for 1 minute. Excess solution was blotted away at the end of the incubation and the membrane wrapped in clingfilm and taped into a film cassette.

The membranes were then exposed to x-ray film in a dark room for between 1 second and 30 minutes. The x-ray film was processed using a Mediphot937 developer.

2.3.7 General cloning techniques

2.3.7.1 Production of competent bacteria

Competent cells were produced using a magnesium chloride/calcium chloride method(Sambrook and Russel 2001). Non-competent E.coli, strain JM109 (Promega), were subcultured in 4 ml Luria Bertani (LB) medium overnight, 37^{o}C. Four millilitres of the overnight culture was inoculated into 400 ml LB medium, which was incubated with regular measurements made of the optical density using a SmartSpec™ Plus spectrophotometer (BioRad). When the OD_{600} reached 0.4, cells were split into eight, 50 ml polypropylene tubes and cooled on ice for 10 minutes. The cultures were centrifuged at 2700 x g, 10 minutes, 4^{o}C and the supernatant removed. Each cell pellet was resuspended by swirling with 30 ml MgCl-CaCl solution, 4^{o}C, before being centrifuged as above. Each cell pellet was resuspended in 2 ml CaCl solution at 4^{o}C, following which all eight suspensions were pooled. Five hundred and sixty microlitres of dimethyl sulfoxide (DMSO) was added to the 16 ml of suspension, which was incubated on ice for 15 minutes. A further 560 µl of DMSO was added and 100 µl aliquoted into pre chilled polypropylene tubes. Cells were stored at -80^{o}C.

2.3.7.2 Transformation of competent bacteria

Competent E. coli were thawed on ice. Twenty five to fifty microliters of thawed bacteria were transferred to polypropylene tubes containing 1 µl (1-3 µg) of plasmid and incubated on ice for 30 minutes. Cells were then heat shocked in a water bath at 42^{o}C for 45 seconds before being returned to ice for 2 minutes. Each culture received 850 µl SOC medium and was then incubated at 37^{o}C on a horizontal rotator shaker for 1 hour.

Following incubation, cells were plated out at different concentrations onto 1.5% w/v LB/agar culture plates containing the appropriate selection antibiotic. Cell suspensions were allowed to absorb into the agar before the plates were placed in an incubator at 37^{o}C for overnight culture.

Following overnight culture, individual colonies were picked using a pipette tip, and transferred to a sterile universal container containing 5 ml of LB medium

supplemented with selective antibiotic. This culture was again incubated overnight at 37^{o}C in a rotatory shaker at 180 rpm. Plasmid DNA was isolated using the Miniprep technique described in 2.3.7.3 and analysed by restriction digestion.

2.3.7.3 Plasmid purification – Qiagen Miniprep

The QIAprep miniprep kit, Qiagen, was used for the isolation of small quantities of plasmid DNA for analysis or to take forward to further cloning steps. This kit uses a modified alkaline lysis protocol followed by binding of DNA to a silica membrane under low salt and pH conditions. Centrifuge wash steps remove endonucleases (buffer PB) and salts (buffer PE). DNA is then eluted in a small volume of buffer EB.

A culture of 5 ml LB medium with appropriate selection antibiotic was inoculated with a scraping from a glycerol stock or with 100 µl from a previous culture. This pre-culture was incubated overnight in a rotatory shaker at 37^{o}C.

Two millilitres of culture was centrifuged at 8000 x g, 2 minutes and the supernatant discarded. The pellet was resuspended in 250 µl buffer P1 supplemented with 100 µg/ml and Lyse Blue 0.1% v/v. Cell lysis was performed by addition of 250 µl buffer P2 followed by vigorous mixing and incubation at room temperature for 5 minutes. The lysis step was terminated by addition of 350 µl buffer N3 and immediate vigorous mixing. The resultant precipitate was centrifuged out at 18000 x g, 10 minutes. The supernatant was then transferred to a QIAprep spin column and centrifuged at 18000 x g, 1 minute. The flow-through was discarded. The spin column was then washed with 500 µl buffer PB followed 750 µl buffer PE, being centrifuged at 18000 x g, 1 minute for each wash. The spin column was centrifuged at 18000 x g, 1 minute to dry the membrane. DNA was eluted by addition of 30-50 µl buffer EB or nuclease free water to the membrane, incubation for 1 minute and subsequent centrifugation at 18000 x g, 1 minute. DNA concentration and purity were assessed by spectrophotometry using a Nanodrop 2000, following which the plasmid was stored at 4^{o}C.

2.3.7.4 Plasmid purification – Qiagen Maxiprep

The Endofree Plasmid Maxiprep kit, Qiagen, was used for the large scale purification of plasmids. This kit uses a modified alkaline lysis protocol, followed by removal of endotoxin and subsequent binding of supercoiled DNA to Qiagen Anion-Exchange Resin in the Qiagen-tip 500, under low salt and pH conditions. After wash steps, the purified DNA is eluted in a high salt buffer and then precipitated using isopropanol.

A pre-culture of 5 ml LB medium with appropriate selection antibiotic was inoculated with a scraping from a glycerol stock or with 100 µl from a previous culture used for Plasmid Miniprep and analytical restriction digest. This pre-culture was incubated overnight in a rotatory shaker at 37^{o}C. The following day, 4 ml of the pre-culture was inoculated into 400 ml of LB medium with appropriate selection antibiotic and cultured overnight in a rotatory shaker at 37^{o}C.

The bacterial culture was centrifuged at 6000 x g, 15 minutes, 4^{o}C. The bacterial pellet was resuspended in 10 ml buffer P1 supplemented with RNase 100 µg/ml and Lyse Blue 0.1% v/v to demonstrate homogeneous mixing. Cell lysis was performed by addition of 10 ml buffer P2 followed by vigorous mixing and incubation at room temperature for 5 minutes. The lysis step was terminated by addition of 10 ml buffer P3 at 4^{o}C and vigorous mixing. The resultant lysate was immediately poured into the QIAfilter cartridge and incubated at room temperature for 10 minutes to allow the precipitate, containing proteins, genomic DNA and detergent to rise to the top. The lysate was then eluted from the cartridge into a fresh 50 ml container. Endotoxin was removed by addition of 2.5 ml buffer ER, mixing and incubation on ice for 30 minutes.

The Qiagen-tip 500 was equilibrated by applying 10 ml buffer QBT. The filtered lysate was then applied and allowed to pass through by gravity flow. The Qiagen-tip 500 was washed by two additions of 30 ml buffer QC. Plasmid DNA was eluted with 15 ml buffer QN and precipitated by addition of 10.5 ml (0.7 volumes) of room-temperature isopropanol. The precipitated DNA was centrifuged at 15000 x g, 4^{o}C, 30 minutes. The supernatant was carefully

decanted leaving the pellet intact. The pellet was washed with 5 ml 70% ethanol in endotoxin-free water and centrifuged 15000 x g, 4^{o}C, 10 minutes. The supernatant was again carefully decanted leaving the pellet intact. The pellet was air dried and then dissolved in 100-200 µl buffer TE. DNA concentration and purity were assessed by spectrophotometry using a Nanodrop 2000, following which the plasmid was stored at 4^{o}C.

2.3.7.5 Plasmid analysis by restriction enzyme digestion and gel *electrophoresis*

Restriction enzyme digests were performed to validate the identity of isolated plasmids. Plasmid DNA was digested with a combination of enzymes chosen to provide fragments of known length, which could then be resolved by gel electrophoresis. One microgram of plasmid DNA was digested with enzyme and the appropriate buffer (Fermentas) made up to 20 µl with sterilised deionised water. Digestion was performed at 37^{o}C for 1 hour and terminated either by thermal denaturation or addition of 1 µl 0.5 M EDTA. Details of more complex digestions are provided in the individual methods sections.

Digestion products were resolved using TBE or TAE/agarose gels containing GelRed 0.01% v/v (VWR International), or stained using 2 ng/ml ethidium bromide. Gels were imaged using a GelDoc imager (BioRad).

2.3.7.6 Storage of plasmids as glycerol stocks

Bacterial cultures grown for plasmid isolation were frozen to allow initiation of future cultures. Aliquots of 700 µl of bacterial cultures were transferred to sterile 1.7 ml microtubes. Sterile glycerol, 300 µl, was added to each tube and mixed by inversion. These stocks were stored at -80^{o}C.

2.3.8 TA cloning of PCR products

2.3.8.1 Introduction

Sequencing of quantitative RT PCR products was used to confirm their specificity. Amplicons were TA cloned into the pGEM-T Easy vector (Promega). This approach uses the template independent addition of a 3' deoxyadenosine by DNA polymerase as part of the PCR reaction, pairing it with 3' terminal thymidine bases found at each end of the linearised vector.

Using the pGEM-T Easy vector system, colonies can be screened for successful ligation by Blue-White screening. The vector contains the β-Galactosidase gene (*Lac*Z) which metabolises X-Gal to give a blue colour. Successful ligation of PCR products normally disrupts the reading frame of the *Lac*Z, resulting in white colonies. Isopropyl β-D-1-thiogalactopyranoside (IPTG) binds to the Lac repressor allowing transcription of the gene.

2.3.8.2 Ligation into the pGEM-T Easy vector

PCR products from quantitative RT PCR for PIWI-Like transcripts were purified as described in 2.3.3.2. Ligation reactions were pipetted as described in Table 2-3, using 1:1-1:3 insert:vector ratios and incubated at 4^{o}C, overnight.

Ligation reactions (2 µl) were transformed into competent JM109 cells according to 2.3.7.2. Culture plates containing 100 µg/ml ampicillin were prepared for Blue-White screening by spreading 50 µl IPTG, 200mM, on each plate and incubating at room temperature for 15 minutes. Each plate was then spread with 20 µl X-Gal and incubated at 37^{o}C for 2 hours. Transformed bacteria were then spread onto the culture plates and incubated at 37^{o}C overnight.

The following day, white colonies were picked and cultured overnight in 6 ml LB medium containing 100 µg/ml ampicillin. Plasmid DNA was extracted as described in 2.3.7.3. Plasmids were analysed by control restriction digest using EcoRI and a second appropriate enzyme, as described in 2.3.7.5. Clones with appropriate restriction patterns were sent for commercial sequencing, DBS Genomics (Durham University), using the M13 forward primer.

Table 2-3. Reaction mixture for ligation of qRT PCR amplicons and pGEM-T Easy vector

Reagent	Volume
Amplicon DNA	---
Linearised vector (50 ng/µl)	1 µl
T4 DNA ligase (3 Weiss units/µl)	1 µl
T4 ligase buffer (2x)	5 µl
Sterilised deionised water	---
Total volume	10 µl

2.3.9 Lentiviral vectors

2.3.9.1 Introduction

The use of recombinant lentiviral vector particles allows for the transduction of mammalian cells. Unlike other retroviridae, which gain access to the host genome only following nuclear membrane breakdown during nuclear division, lentiviridae express a membrane associated matrix protein which is recognised by mammalian nuclear import proteins and results in active transport of the pre-integration complex into the host nucleus. Once in the host nucleus, reverse transcribed double stranded DNA is integrated into the host genome by the viral protein integrase. These characteristics make lentiviral vectors suitable for transduction of slowly, or non-, dividing cells such as patient derived leukaemia blasts and stem cells.

The lentiviral genome has been separated to place genes necessary for viral packaging (packaging vector, e.g. pCMVΔ8.91), production of envelope proteins (envelope vector, e.g. pMD2.G – coding for vesiculo stomatitis virus G protein, VSV-G) and the gene of interest (lentiviral vector/transfer vector, e.g. pSLIEW) onto three separate plasmids (Appendix B). Separation of these three critical components of the lentiviral genome, combined with deletion of many auxillary but non-essential viral genes such as *vpr*, *vpu*, *nef* and *vif*, provides a level of biosecurity. The transfer vector is the only vector to contain the encapsidation signal, ψ, required for packaging of viral RNA within the viral particle. Consequently, the transfer vector is the only genetic material contained within the virus and subsequently transferred to the target cell. Further security has been afforded by the production of self-inactivating (SIN) viruses, in which there is a deletion of a critical 133 base pair region in the 3' long terminal repeat sequence (LTR). Following reverse transcription, this deletion is transferred to the 5' LTR, removing transcriptional capacity from the resultant provirus.

As a consequence of these engineered security features, viral particles must be produced in a producer cell line. Expression of viral genes from each of the three plasmids results in production of proteins able to package a complete

virus particle. The use of the VSV-G pseudotype results in the ability to infect a broad range of mammalian cells.

Lentiviral plasmids were transformed, isolated and their identity confirmed using the procedures detailed in 2.3.6. Virus production and subsequent transduction of mammalian cells are detailed below.

Lentivirus production was performed using the transformed human embryonic renal cell line 293T. It represents a highly transfectable and rapidly growing cell line which produces large yields of recombinant lentivirus. Co-transfection of lentiviral plasmids was performed by calcium phosphate precipitation of DNA. 293T cells avidly take up calcium phosphate crystals, and therefore the associated plasmid DNA. Transcription of the viral genome results in production of viable but replication incompetent virus, which is released into the culture supernatant, from where it can be harvested.

2.3.9.2 Co-transfection of 293T cells by calcium phosphate precipitation

293T cells were passaged at sub-confluence prior to transduction. On the day prior to transduction, cells were plated at 2-3 x 10^6 cells per 100 mm culture dish to provide a monolayer with 30-40% confluence on the day of transfection. On the day of transfection, 0.5 M CaCl solution, 2x HeBS and 2.5 mM HEPES, pH 7.3 were brought to room temperature. Lentiviral plasmids were mixed in a sterile, endotoxin free microtube (1.7ml): pMD2.G envelope plasmid - 5 µg; pCMVΔ8.91 packaging plasmid - 15 µg; pSLIEW transfer plasmid - 20 µg. The volume of the plasmid mixture was adjusted to 250 µl with 2.5 mM HEPES, pH 7.3. To this, 250 µl 0.5 M CaCl solution was added and mixed. The plasmid/CaCl solution was then added drop wise to the 2x HeBS, whilst bubbling air through the 2x HeBS to ensure even mixing and avoid uneven calcium phosphate precipitation. The mixture was incubated at room temperature for 30-35 minutes before being added drop wise to 293T cells. The culture dish was gently shaken, not swirled, to evenly distribute the calcium phosphate/DNA precipitates.

On the morning following transfection, the medium was aspirated from the culture dish, cells were washed once with PBS, 37^{o}C, and 10 ml fresh medium

replaced. Cells were cultured for a further 48-96 hours prior to harvesting the viral supernatant.

2.3.9.3 Viral harvesting and concentration by ultracentrifugation

On day 3-4 following transfection, the medium was aspirated from the 293T cells and centrifuged at 300 x g, 10 minutes to pellet cellular debris. The supernatant was then filtered using a low protein binding PVDF 0.45 µm filter (Millipore). If a second harvest was required, then 8-10 ml fresh Dulbecco's MEM medium was put onto the cells and harvested, as above, 24 hours later.

Viral particles were concentrated by ultracentrifugation using a Beckman Optima ultracentrifuge. Disposable conical centrifuge inserts were sterilised with 70% ethanol and air dried in a laminar flow hood. Viral supernatant was transferred to the conical inserts. Medium was added to bring the supernatant to within 10 mm of the top of the insert to prevent it collapsing under vacuum and to balance the centrifuge bucket to within 0.2 g of its paired bucket. The ultracentrifugation step was performed 120000 x g, 4^{o}C, 2 hours, with a constant vacuum applied by the centrifuge.

Following ultracentrifugation, supernatant was carefully decanted from the insert. Viral particles were then resuspended in 50 - 1000 µl of the desired medium, aliquoted and stored at -80^{o}C.

2.3.9.4 Titration of lentivirus particles

Frozen lentivirus particle were titrated using the t(4;11)(q21;q23) B precursor cell line SEM and the standard spinfection protocol, see 2.3.9.5. Lentiviral concentrate was added to wells at 0.5 µl – 25 µl. The expression of EGFP was analysed by flow cytometry (FACS Calibur) on day 4 post transduction, as described in 2.3.9.6. Lentiviral concentration was calculated using the two following methods.

For accurate calculation of titre when aiming for low multiplicity of infection, the volume of inoculum which produced 20 – 30% transduction was applied to the following formula:

$$titre\ (TU/ml) = \{F \times (Co/V)\} \times D$$

where: F is the frequency of EGFP positive cells; Co is the total number of target cells infected; V is the volume of the inoculum; D is the dilution factor of the inoculum.

As the rate of transduction was not linearly associated with the volume of the inoculum, an alternative method was used when estimating viral titres required for higher rates of transduction. In this instance the percentage of transduced cells was plotted against volume of the inoculum and a logarithmic curve fitted using Excel 2010 (Microsoft). The volume required to produce the desired percentage of transduction was calculated using the formula:

$$Volume\ of\ virus = e^{\{(y-b)/a\}}$$

where y is the desired proportion of transduced cells and a and b are functions of the equation describing the curve, $y = a \times \ln(x) + b$.

2.3.9.5 Lentiviral transduction of target cells

Transduction of suspension culture cells by lentivirus was performed using a spinfection technique. This involves adding lentivirus to suspended cells and then centrifuging the suspension, frequently in the presence of the cationic polymer polybrene, to bring cells and virus into proximity on the surface of the plate. Whilst the centrifuge speeds used are substantially slower than the ultracentrifugation step detailed in 2.3.9.3, the viral particles retain fragments of cell membrane from the producer cell line, 293T. This results in viral particles being brought down to the surface of the plate at comparatively low centrifuge speeds. Following spinfection, a period of culture allows transduced cells to express the transfer vector before transduction efficiency is assessed using a reporter gene, e.g. green fluorescent protein (GFP), or selective culture conditions are applied, e.g. puromycin selection.

Target cells were resuspended at either 1 x 10^6/ml or 2.5 x 10^6/ml, either in complete medium or Optimem serum free medium (Gibco/Invitrogen). Suspensions of patient-derived blasts were augmented with additional growth factors. Standard spinfection conditions also required augmentation with polybrene 8 µg/ml, although this was reduced to 4 µg/ml in some experiments or replaced with protamine 5 µg/ml.

Cells were plated into a 48 well plate with 0.5 x 10^6 cells per well. Virus concentrate was added to the cells before plates were balanced and wrapped in Parafilm (Pechiney Plastic Packaging Co). Spinfection was initially performed at 1500 x g, 32^{o}C, 2 hours. Later spinfections were performed at 900 x g, 34^{o}C, 50 minutes. A modification was made for spinfection of larger numbers of cells. For transduction of 5 x 10^6 cells, suspended in 2 ml in 6 well plates, an initial spin at 50 x g for 20 minutes settled cells on to the bottom of the plate before increasing the speed for the spinfection.

Following spinfection, Parafilm was removed and cells were incubated either for 6-8 hours or overnight. Following overnight incubation, medium containing polybrene and virus was either replaced with fresh complete medium, or diluted with fresh complete medium in the case of very concentrated spinfections. Cells were transferred to larger wells for on-going incubation which allowed expression of transduced genes prior to analysis or selection.

2.3.9.6 Assessment of transduction efficiency by flow cytometry

Transduction efficiency was assessed by measuring expression of a reporter gene encoded by the transfer vector. In the case of pSLIEW this was enhanced green fluorescent protein (EGFP), whilst for pGIPZ it was turbo GFP (tGFP). A non-transduced population was used to control for auto-fluorescence.

Cells were resuspended and a small aliquot transferred to a FACS tube. Samples were analysed using a FACS Calibur (BD Biosciences) fitted with a 488 nm laser and 530/30 nm filter. Dead cells and debris were gated out using forward and side scatter profiles. A threshold was set using a non-transduced population, so that 99.7% of non-transduced cells were below that threshold. Positive cells were defined as those cells in the transduced population which had a signal greater than the threshold.

2.3.10 Decode library screening methodologies

2.3.10.1 Introduction

To perform an unbiased functional screen for genes involved in the self-renewal of leukaemic cells, a genome-wide short hairpin RNA (shRNA) library, produced commercially by Open Biosystems, was used. The transfer vector used in this library is the pGIPZ construct, which encodes a Turbo GFP (tGFP) reporter, shRNA construct and puromycin N-acetyl-transferase gene, PAC.

2.3.10.2 Determining the puromycin sensitivity of experimental cell lines

Puromycin is an antibiotic which inhibits protein synthesis and is toxic to a broad range of both prokaryotic and eukaryotic cells. As such it can be used to positively select mammalian cells expressing the PAC gene, which is encoded by the pGIPZ vector.

The minimum concentration of puromycin required was determined by WST-1 assay, which relies on the metabolism of the tetrazolium salt of WST-1 to formazan by mitochondrial dehydrogenases. The formazan dye produced by viable cells can be quantitated by colorimetry at 440 nm, providing an estimate of cell viability and therefore the cellular toxicity of puromycin.

Candidate cell lines were seeded into 96 well plates at 5×10^4 cells/well and cultured overnight in the presence of increasing concentrations of puromycin, 0.1-2 µg/ml. After 6 days of culture, 10 µl WST1 reagent was added to each well and the cells incubated at 37^{o}C. Colorimetric measurements were made using a Model 680 microplate reader at 1, 2 and 4 hours during the incubation period.

2.3.10.3 Transduction of cell lines with pGIPZ pools

Multiple test transductions were performed to optimise transduction conditions as well as to determine the relative transduction efficiency of the experimental cell lines. Initially, the functional titre of control virus (provided with the library) in each tested cell line was calculated as described for viral titration in 2.3.9.4. The functional titre was then used to calculate the relative transduction

efficiency (RTE) by comparison to the functional titre of the control virus stock in HEK293T cells (as determined by Open Biosystems), as follows:

$$RTE = \frac{functional\ titre\ in\ cell\ line}{functional\ titre\ in\ HEK293T}$$

The RTE then allows an estimation of the functional viral titre in each of the seven experimental pools in the given cell line.

Knowing the functional viral titre in the cell line to be tested allows the calculation of the volume of virus required to produce the desired rate of transduction in a given number of cells. In order to minimise the number of cells receiving multiple viral integrations, the proportion of cells to be transduced should be 0.3. This is termed the multiplicity of infection (MOI) and, at 0.3, is predicted to result in just 3% of cells receiving more than one integration.

$$Virus\{\mu l\} = \frac{(No.of cells\ to\ be\ transduced \times functional\ titre\{TU/ml\})}{1000}$$

The number of cells to be transduced should be sufficient to provide 100 fold coverage of the pool, that is, each hairpin construct should be integrated into at least 100 cells. This is calculated as follows:

$$N = \frac{\ln(1-p)}{\ln(1-f)} \times \frac{C}{T}$$

N is the required number of transduced cells; p the desired probability of representation - 0.95; and f the fractional proportion of a single shRNA in the library, i.e. the reciprocal of the complexity of 10,000 *per* pool. T represents the fraction of transfected cells and equals approximately the MOI. C is the coverage of each single shRNA, which should be 100-fold for a negative selection. The desired cell number, N, is therefore 1×10^7 cells per spinfection.

Transduction of cell lines using pools of non-silencing control virus particles was carried out according to the standard spinfection protocol detailed in 2.3.9.5, with modifications to the medium designed to optimise transduction efficiency.

2.3.10.4 Puromycin selection and cell sampling

Having assessed the transduction efficiency by flow cytometry on day 4 following transduction, see 2.3.9.6, transduced cells were selected by the addition of 0.4-0.8 µg/ml puromycin to the culture medium. The proportion of cells expressing tGFP was monitored by flow cytometry, see 2.3.9.6.

During the test experiments with the single lentiviral pool, pool 7, low transduction (7.5 – 15%) on day 4 raised concerns about the proportion of transduced cells harvested for the baseline sample. Cells were therefore cultured overnight in 1 µg/ml puromycin, followed by density gradient centrifugation, see 2.3.2.3, and harvesting of cells for DNA extraction.

2.3.10.5 Seeding of transduced cells onto M2-10B4 feeder cells

Following a period of selection in puromycin containing medium, pGIPZ transduced cells were seeded out onto irradiated M2-10B4-PAC cells. To maintain the 5×10^6 cell coverage, ten plates/flasks of feeder cells each received 50 ml RPMI medium, 1% v/v FCS, 0.4 – 0.8 µg/ml puromycin and 5×10^5 cells. This equated to 1×10^4 cells/ml, or approximately 3×10^4 cells/cm^2.

Following a period of co-culture, when pGIPZ transduced cells had expanded sufficiently (aiming for 3 doublings), the medium was aspirated and saved. The plates/flasks were washed once with 10ml PBS and the wash fraction saved. The adherent cells were then trypsinised, see 2.3.1.4. Trypsinised cells were resuspended in complete medium and pooled. All fractions were centrifuged, 300 x g, 10 minutes, the cells resuspended in complete medium, counted and pooled. An aliquot was taken for analysis by flow cytometry and, following the first plating, 5×10^6 cells were reseeded onto fresh irradiated M2-10B4-PAC cells, as above. The remaining cells were then divided: $1\text{-}2 \times 10^7$ cells were taken for MACS microbead sorting, using CD133 to isolate leukaemic cells from feeders, whilst any remaining cells were frozen viable as back-up.

Cells harvested for DNA extraction were positively selected using MACS CD133 microbead kit, as described in 2.3.2.5. Selected cells were stored as pellets of 5×10^6 cells at -20^{o}C for DNA extraction according to 2.3.3.1.

2.3.10.6 Amplification and purification of barcode regions

pGIPZ transduced cells which had been harvested during the test experiments with pool 7 and stored as pellets, had their DNA extracted according to 2.3.3.1. The barcode region, unique to each shRNAmir30 construct, was then amplified by PCR and purified by gel electrophoresis. The resultant amplicons were then deconvoluted by co-hybridisation of baseline and experimental samples to a 2x105K Agilent array provided with the Decode library.

The PCR was performed according to the Decode library manual, as follows. A minimum of 4 reactions was carried out per time-point sample to ensure sufficient coverage. Reactions were pipetted to a total volume of 100 µl, see Table 2-4. A positive control used 1 ng of plasmid DNA, containing a barcode region, supplied with the library. A negative control contained no template DNA. The reactions were performed using a GeneAmp PCR System 9700, programmes according to Table 2-5.

Table 2-4. Reagents for PCR amplification of Decode barcode regions

	Volume (µl)	Final concentration
5 M Betaine	10	0.5M
10x KOD hotstart polymerase buffer	10	1x
dNTP's (2 mM each)	10	0.2 mM (each)
$MgSO_4$ (25 mM)	4	1 mM
Forward primer (5 pmol/µl)	6	0.3 µM
Reverse primer (5 pmol/µl)	6	0.3 µM
KOD hotstart polymerase (1 U/µl)	2	0.02 U/µl
Template DNA – 850 ng	---	---
Nuclease free water	---	---
Total volume	100 µl	

Table 2-5. Reaction conditions for PCR amplification of Decode barcode regions

	Temperature	Time	Cycles
Hot start enzyme activation	94^{o}C	3 minutes	1
Melt	94^{o}C	35 seconds	30
Anneal	62^{o}C	35 seconds	
Extension	72^{o}C	1 minute	

The PCR reactions were then purified by electrophoresis on a 1.25% w/v TBE/agarose gel. The band at approximately 250 bp was visualised using Chromato-Vue UV trans-illuminator, excised and extracted according to 2.3.3.4. Purified DNA was quantitated by spectrophotometry and sent to the Functional Genomics and Proteomics Facility, University of Birmingham, for commercial labelling, hybridisation and scanning, according to the Open Biosystems Decode manual.

2.3.10.7 Analysis of data from genome-wide RNAi screen

Data from the genome-wide screen were analysed commercially through the University of Birmingham using a customised version of "Agi4x44PreProcess" in R (R version 2.15.0). The principle steps involved subtraction of background signal, removal of control probes and summarisation of the multiple probes present for each construct (generation of the median value). Normalisation was performed using “Normalize the columns of a matrix to have the same quantiles, allowing for missing values”. Constructs whose probes had a median fold change of >2.5 were compared between different arrays (representing different time point comparisons in the trial screen).

2.3.11 Transduction of patient-derived material with lentiviral particles

2.3.11.1 Transduction setup and spinfection

Aliquots of patient-derived peripheral blood or bone marrow blasts were retrieved from liquid nitrogen either on the day of, or the day prior to spinfection, thawed and counted with Trypan Blue to assess viability as described in 2.3.1. Cells were centrifuged at 300 x g, 5 mins and resuspended at 1 x 10^6 cells/ml in either αMEM supplemented with 10% v/v FCS, 10% v/v horse serum, 10% v/v HS-5 conditioned medium, see 2.3.1.5, or in serum free expansion medium (SFEM) supplemented with 10% v/v FCS. Five hundred microliters (5 x 10^5 cells) were pipetted into wells of a 48 well culture plate and supplemented with the appropriate polybrene or protamine. For some experiments additional cytokines were added to specific wells to assess their effect on transduction and cell survival. These included stem cell factor, 50 ng/ml, interleukin 7 (IL-7), 10 ng/ml and IL-3, 20 ng/ml. SLIEW/SLMIEW lentivirus was added to the wells, with one well left without virus to provide an untransduced control. The plate was balanced to within 0.2 g and sealed with Parafilm.

Spinfection was performed as detailed in section 2.3.9.5.

Following overnight culture, 350 µl transduction medium was aspirated from the cells and replaced with 850 µl fresh medium. For those specimens transduced in the presence of cytokines, the fresh medium was supplemented with the same cytokine concentrations. Cells were either transplanted into NSG mice, see 2.3.11.4, or analysed my five-colour flow cytometry on day 3 or 4 following transduction, see 2.3.11.2.

2.3.11.2 Assessment of transduction by multicolour flow cytometry

Analysis of transduction was performed by flow cytometry using either single colour assessment of EGFP (FACS Calibur, 2 samples only) or four-colour assessment of EGFP, anti-CD19 PE, anti-CD20 PerCP Cy5.5, anti-CD34 APC. For samples which had been harvested following xeno-transplantation, murine cells were excluded by staining with anti-murine CD45/Ter119 PE Cy7.

Cells were resuspended and washed once in 3 ml PBS in a flow cytometry tube, 450 x g, 4 minutes. Cells were resuspended in 100 µl PBS and stained with anti-CD19, 10 µl, anti-CD20, 10 µl, anti-CD34, 2.5 µl with or without anti-murine CD45, 2.5 µl, anti-Ter119, 2.5 µl, at room temperature, protected from light for 20 minutes. Cells were then washed twice in 2.5 ml PBS, 450 x g, 4 minutes and resuspended in 300 µl PBS ahead of analysis using a FACS Canto II flow cytometer.

2.3.11.3 Washing procedure for removal of residual lentivirus

In order to remove residual, infectious lentivirus, transduced cells were stringently washed each day from day 1 post-transduction onwards. This protocol was developed during the course of this project.

On Day 1 following transduction, 350 µl supernatant was removed from the transduced cells. Cells were resuspended in 850 µl medium followed by the addition of a further 20 ml RPMI 1640, 10% FCS. Cells were centrifuged 300 x g for 10 minutes. The supernatant was aspirated using a plastic pipette (to avoid the risk of sharps related injuries) and the cells resuspended to repeat the above wash procedure. Following the second wash, cells were resuspended in 1 ml SFEM supplemented as for the transduction experiment. The wash procedure was repeated on subsequent days.

2.3.11.4 Xenotransplantation studies using NSG mice

Xenotransplantation studies were conducted using intrafemoral injection into NOD/scid IL-2 receptor gamma common chain knockout (NSG) mice (NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ) mice (Shultz, Lyons et al. 2005) (Jackson Laboratories, supplied by Charles River UK Ltd). These mice, in addition to the severe combined immune deficiency (scid) background, have a further deletion of the IL-2 receptor gamma common chain. The product of this gene is an intrinsic component of a number of cell surface receptors within the adaptive immune system and these mice therefore have profound immune deficiency. Functionally they lack B- and T-cells, NK cells and have disorders of macrophage and complement function. In addition to currently providing the most sensitive model for leukaemic engraftment (McDermott, Eppert et al.

2010), NSG mice also have a longer lifespan than NOD/scid mice and are therefore well suited to transplantation studies which can take many months.

Intrafemoral transplantation was performed by one of three trained members of the laboratory. Transplants were performed in a laminar flow hood with aseptic technique, under general anaesthesia. Mice received subcutaneous Carprofen analgesia, 5 mg/kg, whilst anaesthetised. A suspension of patient derived leukaemic cells or derived cell line cells of between 20 – 30 µl was injected into the femur through a hole made in the distal femoral joint, as described in (le Viseur, Hotfilder et al. 2008).

2.3.11.5 In vivo imaging of transduced, fLuc+ leukaemia cells

In vivo imaging of mice transplanted with leukaemic blasts (patient-derived or cell line) transduced with SLIEW or SLMIEW viruses was performed using an IVIS Spectrum bioluminescent camera. For standard planar imaging, mice received an intraperitoneal injection of 3 µg D-Luciferin. For 3 dimensional imaging, mice received 4.5 µg D-Luciferin. Mice were anaesthetised using 3% v/v isofluorane in an induction chamber so that imaging could begin, with on-going 2% v/v isofluorane, 10 minutes after the D-Luciferin injection.

2.3.11.6 Harvesting mice for engraftment leukaemic material

At the end of an experiment, mice were killed by cervical dislocation. Material was harvested from the spleen and bone marrow, and for one experiment also from the liver and kidney (collected as for the spleen). The spleen was dissected out of the abdomen, weighed and macerated through a 40 µm cell strainer into PBS. Femurs and tibias were dissected out of both legs, the end cut off and the marrow flushed out with PBS using a 25G needle. Bone marrow was also then passed through a 40 µm cell strainer to provide a single cell suspension and remove any fragments of bone.

2.3.11.7 Harvesting organs for immunohistochemistry

For one mouse transplanted with L578-SLIEW, the brain, kidneys, liver and spleen were dissected out and placed into formalin for fixation. Organs were then embedded in paraffin and cut into 4 µm sections and mounted on glass

slides. Haematoxylin and eosin staining was performed by Dr Jennifer Jackson, Northern Institute for Cancer Research. Immunohistochemical staining was performed commercially by the Department of Pathology, Royal Victoria Infirmary using a Ventana Benchmark XT or Ventana Benchmark Ultra and Ventana Ultraview detection system (CD10, CD79a, TdT) and University College London (CD19).

Slides for fluorescent microscopy were mounted using DAPI mount by Dr J Jackson and viewed using a Leica DSM fluorescence microscope and captured with SPOT RT digital CCD camera.

2.3.12 Cloning of a single vector combining *in vitro* analysis, *in vivo* analysis and RNAi – pSLMIEW

2.3.12.1 Introduction

The modification of the existing pSLIEW vector to include a microRNA adapted sort hairpin sequence would allow patient-derived material to receive EGFP (*in vitro* analysis), fLuc (*in vivo* imaging) and an shRNA (gene knockdown) with a single transduction. The approach used required: a) the cloning of the shRNAmir30 sequence from Open Biosystems pTRIPZ vectors into a Gateway pENTR1a vector; b) the cloning of Invitrogen's Gateway destination cassette into the *BamHI* site in pSLIEW. These two vectors would then be recombined to produce the final pSLMIEW vectors.

This work was performed by Simon Bomken (2.3.12.2, 2.3.12.3 and 2.3.12.6) with the assistance of two MRes students, Cara Hernon (2.3.12.2, 2.3.12.4 and 2.3.12.6) and Nana Anim-Addo (2.3.12.5 and 2.3.12.6). The strategy was designed by SB.

2.3.12.2 PCR amplification of shRNAmir30

The shRNAmir30 sequence to be cloned into the pENTR1a vector was amplified by PCR using primers designed to add *SalI* and *EcoRV* restriction sites to 5' and 3' ends respectively, see 2.2.5.4. Additional 5' GAGAGA sequences are required to allow efficient restriction digestion of the amplicon ends. PCR amplification was performed using plasmid DNA of pTRIPZ constructs containing shMA6 (specific for *MLL/AF4* fusion transcript), shAGF1 (specific for *RUNX1/RUNX1T1* fusion transcript) and the Open Biosystems non-targeting control construct which contains an 18bp sequence in place of a hairpin.

The amplification was pipetted on ice according to Table 2-6 and used a high fidelity polymerase, to reduce non-specific amplification and polymerisation errors. The reaction was performed in 2 stages, Table 2-7. The first 3 cycles used an annealing temperature specific for that portion of the primers which were complementary to the plasmid sequence, i.e. assuming the restriction site and 5' GAGAGA sequences to be unbound. The final 20 cycles used a higher

annealing temperature to ensure specific binding of the full primer sequence, now including the restriction sites and terminal GAGAGA bases.

Amplification products were purified as described in 2.3.3.2, assessed by Nanodrop and restriction digested immediately.

Table 2-6. Reaction mixture for PCR amplification of shRNAmir30 sequences

Reagent	Volume	Final concentration
ThermoPol buffer (10x)	5 µl	1x
dNTPs (10 mM)	1 µl	200 µM
Primer mix (10 µM)	1 µl	0.2 µM
Plasmid DNA (20 ng)	---	---
Vent polymerase (1:5 dilution)	1 µl	1 unit
Nuclease free water	---	---
Total volume	50 µl	

Table 2-7. Reaction conditions for PCR amplification of shRNAmir30 sequences

	Temperature	Time	Cycles
	95°C	5 minutes	
Annealing of incompletely complementary primers	95°C	30 seconds	3 cycles
	59°C	30 seconds	
	72°C	30 seconds	
Annealing of complementary primers	95°C	30 seconds	20 cycles
	64°C	30 seconds	
	72°C	30 seconds	
	72°C	5 minutes	

2.3.12.3 Cloning of pENTR-shRNAmir30

Purified shRNAmir30 amplicons from 2.3.12.2 and pENTR1a vector were restriction digested using *Sal*I and *Eco*RV at 37°C for 2 hours to ensure complete digestion. Reactions mixtures were as described in Table 2-8. The

digestion products were again purified according to 2.3.3.2. The pENTR1a digestion products were isolated by gel electrophoresis using a 1.25% w/v agarose/TBE gel, whilst the shRNAmir30 digestion products were separated using a 2% w/v agarose/TBE gel. The appropriate bands excised and purified according to 2.3.3.3.

Table 2-8. Reaction mixture for restriction digest of shRNAmir30 amplicon and pENTR vector

Reagent	Quantity
DNA shRNAmir30 (500 ng) pENTR (1500 ng)	--- ---
*Sal*I (10 units/ µl)	1 µl
*Eco*RV (10 units/ µl)	2 µl
Buffer O (10x)	2 µl
Sterilised deionised water	---
Total volume	20 µl

Purified restriction digest products were ligated together in 1:1, 1:3 and 1:6 ratios of vector to insert, at 16^{o}C, overnight. The reaction mixtures are shown in Table 2-9. Ligated DNA, 2 µl, was transformed into competent DH5α cells as described in 2.3.7.2. Transformations were plated onto culture plates containing 50 µg/ml kanamycin and cultured overnight. Colonies were picked the following day and grown up in 5 ml LB medium supplemented with 50 µg/ml kanamycin. Plasmid DNA was isolated from these pre-cultures as described in 2.3.7.3 and sent for commercial sequencing by DBS Genomics, Durham, using the Entry vector reverse sequencing primer, 2.2.5.4.

Table 2-9. Reaction mixture for ligation of shRNAmir30 and pENTR

Reagent	Volume
DNA	---
T4 ligase (5 units/µl dilution in SDW)	1 µl
T4 ligase buffer (10x)	1 µl
Sterilised deionised water	---
Total volume	10 µl

2.3.12.4 Cloning of ANGPT1 shRNA into pENTR-shRNAmir30

To demonstrate the principle of cloning a new shRNA sequence into the pENTR-shRNAmir30 vector, the shRNA specific for *ANGPT1* was excised from a commercially available pTRIPZ vector and ligated into the digested mir30 sequence of the non-targeting pENTR-shRNAmir30 plasmid cloned in 2.3.12.3.

The pTRIPZ and pENTRshRNAmir30 plasmids were digested using *Xho*I and *Eco*RI according to Table 2-10 . The reaction was performed for 90 minutes at 30^{o}C and terminated by heating to 80^{o}C for 20 minutes. Digestion products were purified by gel electrophoresis as described in 2.3.12.3. Ligations reactions were performed using 1:1 and 1:3 ratios of vector to insert as described in Table 2-9.

Table 2-10. Reaction mixture for restriction digest to clone shANGPT1 hairpin into pENTR-shRNAmir30

Reagent	Quantity
Plasmid DNA (1 µg)	---
*Xho*I (10 units/ µl)	1 µl
*Eco*RI (10 units/ µl)	2 µl
Buffer O (10x)	2 µl
Sterilised deionised water	---
Total volume	20 µl

Ligation products were transformed into competent DH5α cells, as described in 2.3.7.2, cultured on culture plates containing kanamycin 50 µg/ml. Colonies were picked, cultured overnight in LB medium containing kanamycin 50 µg/ml and plasmid DNA isolated, as described in 2.3.7.3. A control digest was performed using *Sal*I and *Eco*RI, as described in Table 2-8, and the resultant products analysed by gel electrophoresis using a 2% w/v agarose/TBE gel. Clones with the appropriate bands were sent for commercial sequencing by DBS Genomics using the Entry vector reverse sequencing primer, 2.2.5.4.

2.3.12.5 Cloning of pSLIEW destination vector

The Gateway conversion cassette, reading frame A, was cloned into the *Bam*HI restriction site 5' to fLuc and 3' to IRES in pSLIEW. pSLIEW was digested at 37^{o}C for 2 hours and heat inactivated at 80^{o}C for 20 minutes. The reaction mixture is described in Table 2-11. The ends of the linearised plasmid were then blunted by a Klenow fill-in reaction, described in Table 2-12. This was performed at 37^{o}C for 10 minutes and heat inactivated at 75^{o}C for 10 minutes.

Table 2-11. Reaction mixture for *Bam*HI digest of pSLIEW

Reagent	Volume
pSLIEW DNA (5 µg)	1.5 µl
*Bam*HI buffer (10x)	2 µl
*Bam*HI	1 µl
Sterilised deionised water	15.5 µl
Total volume	20 µl

Table 2-12. Reaction mixture for Klenow fill-in of linear pSLIEW

Reagent	Volume
Digested pSLIEW	20 µl
dNTP mix (500 µM each)	20 µl
Klenow reaction buffer (10x)	20 µl
Klenow fragment	0.5 µl
Sterilised deionised water	139.5 µl
Final volume	200 µl

The blunt-ended *BamHI* pSLIEW digest was purified as described in 2.3.3.2, and analysed using a Nanodrop. A 5' dephosphorylation reaction was performed using shrimp alkaline phosphatase at 37^{o}C for 30 minutes, as described in Table 2-13. The enzyme was heat inactivated at 65^{o}C for 15 minutes. The DNA was again assessed by Nanodrop and adjusted to 20-50 ng/µl with buffer TE, pH 8.0.

Table 2-13. Reaction mixture for 5' dephosphorylation of linear pSLIEW

Reagent	Volume
DNA (target 1 pmol of termini, actual 0.21 pmol)	17 µl
SAP Reaction buffer (10x)	2 µl
Shrimp Alkaline Phosphatase (1 unit/µl)	1 µl
Total volume	20 µl

The Gateway conversion cassette was then ligated into the dephosphorylated plasmid at 16^{o}C overnight, as described in Table 2-14.

Table 2-14. Reaction mixture for ligation of conversion cassette into linear pSLIEW

Reagent	Volume
Linear plasmid (50 ng/µl)	0.5 µl
50% PEG 4000	2 µl
Gateway conversion cassette (5 ng/µl)	2 µl
T4 DNA ligase buffer (10x)	1 µl
T4 DNA ligase (5 units/µl)	1 µl
Nuclease free water	12.5 µl

The ligated pSLIEW-DEST plasmid was transformed into ccdB Survival 2T1R competent cells at 0.2 µl, 1 µl, 2 µl per 25 µl cells as described in 2.3.7.2. These bacteria have a mutation of DNA gyrase (*gyr*A462), which makes it resistant to interference caused by the CcdB protein product of the *ccdB* gene in the conversion cassette. The chloramphenicol resistance gene on the conversion cassette allows negative selection of un-ligated plasmids, see Figure 2-1. Transformed bacteria were spread on 1.5% w/v agar/LB plates with 30 µg/ml chloramphenicol and incubated over 2 nights at 37^{o}C. Colonies were picked into 5 ml pre-cultures of LB medium with 30 µg/ml chloramphenicol and again grown overnight.

Figure 2-1. Gateway conversion cassette. attR – attachment sites R1 & R2, CmR – chloramphenicol resistance gene, ccdB – ccdB gene for negative selection.

Plasmid DNA was extracted from positive cultures, as described in 2.3.7.3. A control restriction digest was performed using *Pvu*II according to the method in 2.3.7.5. In addition, four clones were sent for sequencing at DBS Genomics, using the Destination vector sequencing primers 1 and 2, see 2.2.5.4.

2.3.12.6 Recombination of vectors to produce pSLMIEW

Each of the four pENTR-shRNAmir plasmids (shMA6, shAGF1, shANGPT1 and the non-targeting sequence) were recombined with pSLIEW-DEST to give completed Gateway expression vectors, known as pSLMIEW. This process uses a proprietary enzyme mix from Invitrogen, LR Clonase II enzyme mix. The reaction was incubated overnight at 16^{o}C. The reaction was stopped by addition of 2 µg Proteinase K and incubation at 37^{o}C for 10 minutes. Each reaction was transformed into competent DH5α cells as described in 2.3.7.2, which were spread onto culture plates containing 100 µg/ml ampicillin and cultured overnight. Colonies were picked and cultured overnight in 5 ml LB medium containing 100 µg/ml ampicillin. Plasmid DNA was extracted according to 2.3.7.3 and analysed by restriction digest using *Pvu*II. Successfully recombined clones were expanded in LB medium supplemented with 100 µg/ml ampicillin. Plasmid DNA was extracted according to 2.3.7.4 and used for lentiviral production according to 2.3.9.

Chapter 3

Expression and function of the candidate stemness gene *PIWIL2* in normal and malignant lymphoid populations

3.1 Introduction

Previous studies in both Drosophila and mice, have demonstrated a critical role for PIWI clade proteins in the maintenance of germline stem cells. Due to the specificity of their small RNA partners, the PIWI interacting RNAs (piRNAs), for retrotransposons, it was presumed that their role is to silence mobile genetic elements by post translational degradation of mRNA, preventing insertional mutagenesis. This is especially important during the period of global demethylation at the end of foetal development during which retrotransposons would otherwise be freely transcribed

More recent studies, however, have suggested a wider role for PIWI/piRNA complexes as they have been shown to be essential for the *de novo* methylation of germline DNA which occurs during foetal development (Kuramochi-Miyagawa, Watanabe et al. 2008). In addition, PIWI/piRNA sequence directed methylation of the *CREB2* promoter modified its expression in Aplysia (Sea Hare) neurones, mediating the role of this transcription factor in response to serotonin (Rajasethupathy, Antonov et al. 2012).

Following the identification of the expression of *PIWIL2* in malignancy (Lee, Schutte et al. 2006), this project aimed to investigate the expression and function of *PIWIL2* in acute lymphoblastic leukaemia. Given the stem cell specific nature of its expression, this project particularly intended to investigate its role in the biology of ALL self-renewal/propagation. Preliminary experiments within the laboratory (conducted by Svetlana Myssina) demonstrated the expression of *PIWIL2* in a small cohort of presentation bone marrow specimens from children with acute lymphoblastic leukaemia (Figure 3-1). Furthermore, paired samples taken at the end of remission induction showed loss of expression in all but one case. This project aimed to further define the expression and function of PIWI-Like genes in childhood ALL.

3.2 Aims of the project

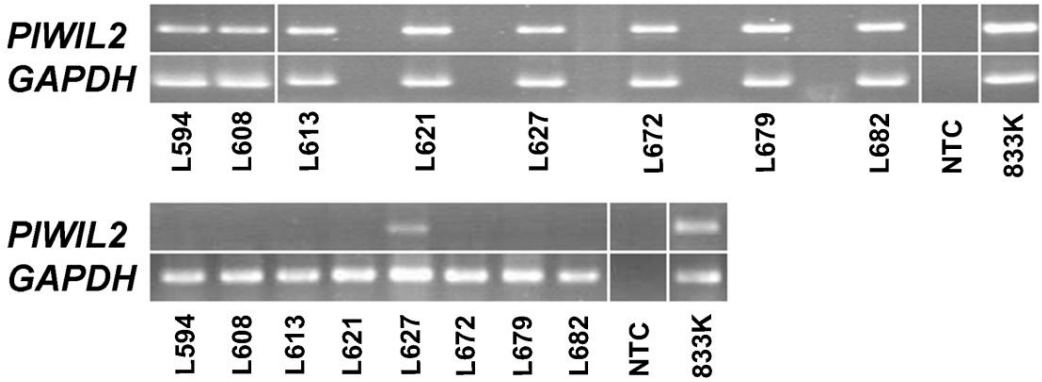
The aims of this project were:

1. To investigate the expression pattern of PIWI-Like genes in childhood acute lymphoblastic leukaemia;
2. To determine the functional significance of PIWI-Like genes in ALL by serial electroporation and siRNA mediated gene knockdown

During the course of this project, peripheral blood lymphocytes (PBLs) from healthy donors were used to provide normal control RNA. Unexpectedly, expression of *PIWIL2* was identified in unfractionated PBLs. This led to a further project aim being developed.

3. To determine the expression of PIWI-Like genes in fractionated peripheral blood lymphocytes

Figure 3-1. Expression of *PIWIL2*, assessed by RT PCR. Top panel – Presentation samples. Bottom panel – Paired remission samples. NTC – Non-target control. 833K – Positive control cell line. Patient details are recorded in Appendix A.



3.3 Results

3.3.1 Optimising the qRT PCR

Primers suitable for quantitative RT PCR were designed to span exon-exon boundaries and therefore be specific for mRNA transcript. A number of primer pairs were designed for each of the PIWI-Like genes, PIWIL1-4. qRT PCR was performed using the testicular seminoma cell line 833K as a positive control. The final primer pairs to be used were chosen to ensure:

1. A uniform melting curve during qRT PCR
2. Primer efficiency close to 2, ensuring accurate quantitation by the ΔΔCt method
3. The appropriate amplicon length, as determined by gel electrophoresis

Using SDS 2.0, the melting curve was assessed. Each of the primer sets was analysed and found to have a single peak, representing a single amplicon without the formation of primer dimers. This analysis was performed following each individual qRT PCR run.

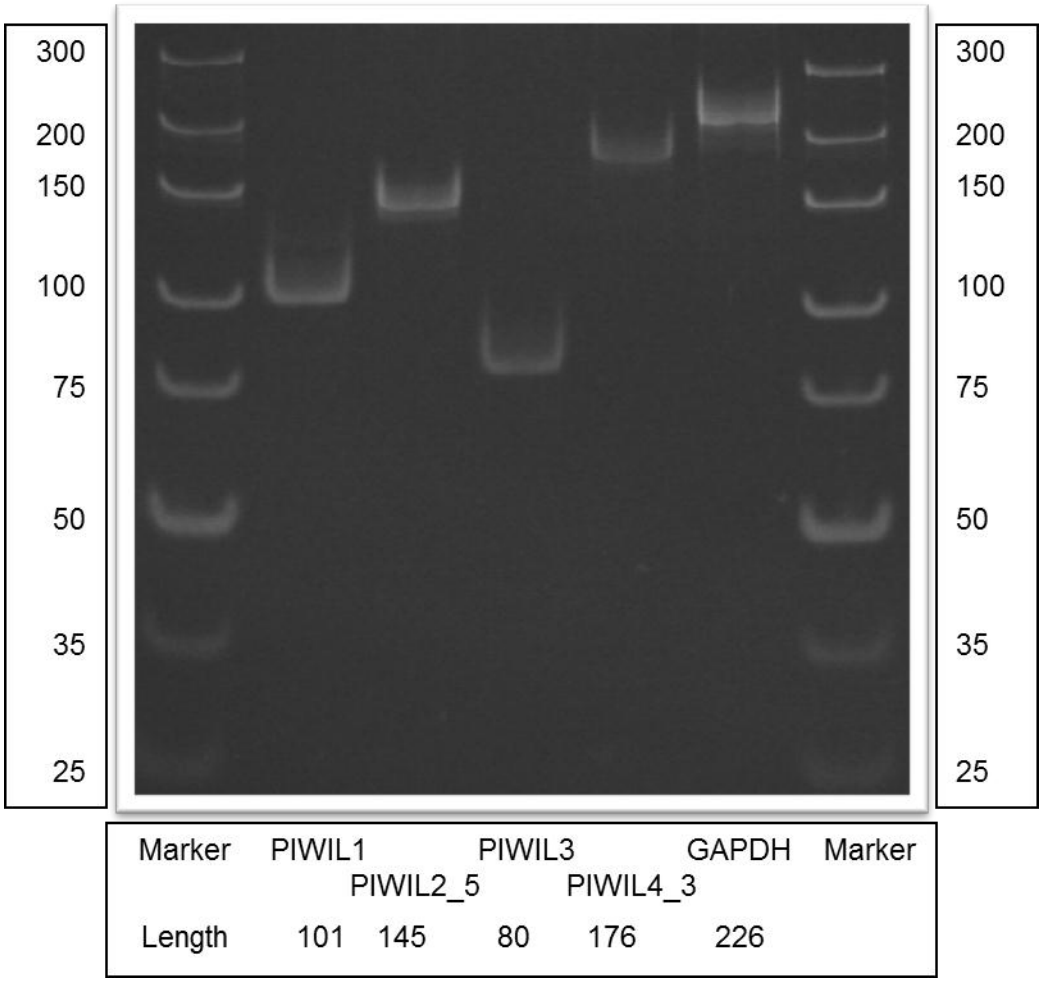
Analysis of the primer efficiency was performed by pipetting serial dilutions of cDNA into the reaction as follows - 1:1, 1:10, 1:100, 1:1000 and 1:10000. A line of best fit was plotted on a graph of ΔCt against cDNA dilution and the efficiency calculated by:

$$E = 10^{(1|b)}$$

where E is the primer efficiency and b describes the gradient of the line of best fit. Acceptable primer efficiencies of 1.9-2.2 were achieved with each of the primers chosen for subsequent experiments.

Amplicon sizes were predicted from the known transcript sequences. Following qRT PCR, amplicons were analysed by electrophoresis using an 8% polyacrylamide gel, stained with ethidium bromide. This analysis demonstrated a single amplicon of the expected size for each of the primer pairs which had been selected following melting curve and efficiency analysis (Figure 3-2).

Figure 3-2. Analysis of PIWI-Like amplicons by gel electrophoresis. The predicted amplicon length in base pairs is displayed beneath the gel. Marker sizes (base pairs) are displayed beside the gel, Fermentas UltraLow ladder.



3.3.2 Confirmation of amplicon specificity by sequencing

Following the unexpected finding of *PIWIL2* expression in normal peripheral blood lymphocytes, see 3.3.4, the specificity of the qRT PCR primers was confirmed by sequencing of the amplicons. This was performed by DBS Genomics, Durham, following the TA cloning of PCR amplicons into the pGEM-T Easy vector. The sequences generated demonstrate >98% sequence homology for the predicted amplicons of *PIWIL1-3* (Figure 3-3, Appendix B). Despite repeated attempts, it was not possible to clone the amplicons of *PIWIL4* and *GAPDH* into the pGEM-T Easy vector and so these amplicons were sequenced in two directions using the qRT PCR primers. By combining the sequences derived in 3' and 5' directions it was possible to determine the complete sequence homology of the amplicons. These sequences also showed >98% homology to the predicted amplicon (Appendix B).

Overall, this sequencing data gave a high degree of confidence in the novel finding that *PIWIL2* and *PIWIL4* were expressed in normal PBLs from healthy control donors.

Figure 3-3. Sequencing of *PIWIL2* amplicon. A) Raw sequencing data generated following TA cloning of qRT PCR amplicons into the pGEM-T Easy vector and sequencing by DBS Genomics.

A

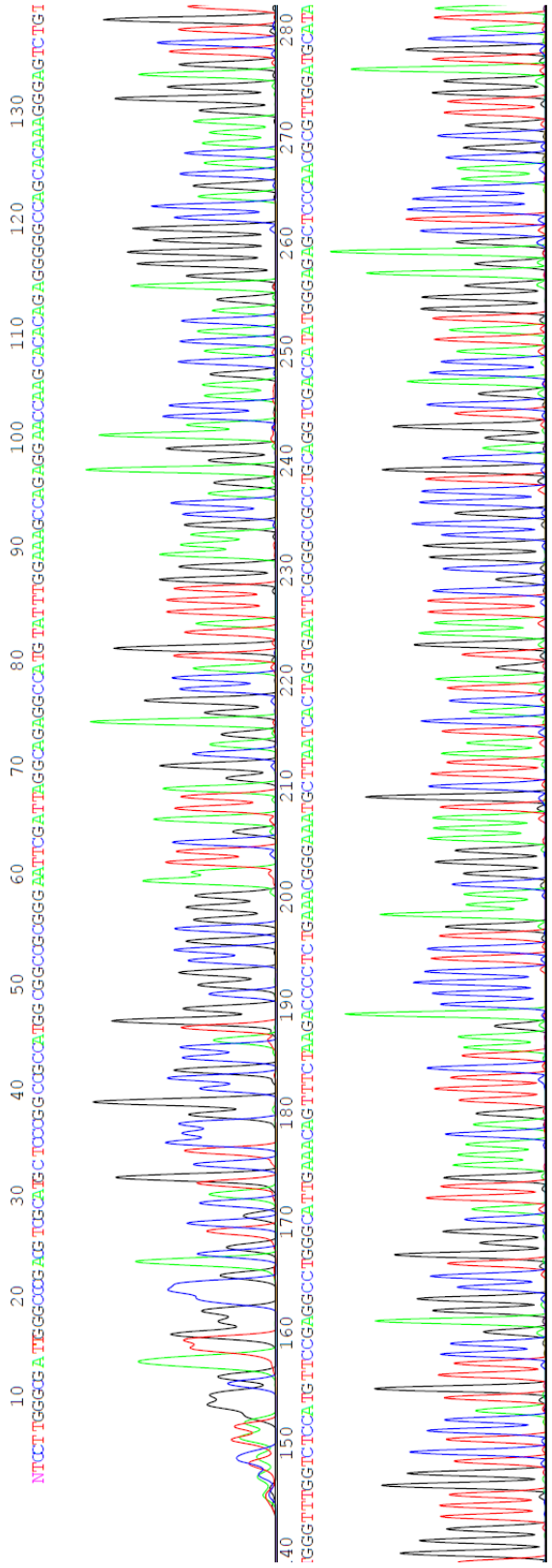


Figure 3-3. B) Sequence alignment of clones 1 and 6 performed using ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) demonstrates homology to the predicted *PIWIL2* sequence

B

```
Clone1          ----------------------------------------------NTCCTTGGGCGATTG 15
Clone6          AGCTATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGC 60
Predicted       ------------------------------------------------------------

Clone1          GGCCCG-ACGTCGCATGCTCCC------GGCCGCCATGGCGGCCGCGG-GAATTC----- 61
Clone6          ATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCGAATTCACTAG 120
Predicted       ------------------------------------------------------------

Clone1          -GATTAGGCAGAGGCCATGTATTTGGAAAGCCAGAGGAACCAAGCACACAGAGGGGGCCA 121
Clone6          TGATTAGGCAGAGGCCATGTATTTGGAAAGCCAGAGGAACCAAGCACACAGAGGGGGCCA 180
Predicted       -----AGGCAGAGGCCATGTATTTGGAAAGCCAGAGGAACCAAGCACACAGAGGGGGCCA 55

Clone1          GCACAAAGGGAGTCTGTGGGTTTGGTCTCCATGTTCCGAGGCCTGGGCATTGAAACAGTT 181
Clone6          GCACAAAGGGAGTCTGTGGGTTTGGTCTCCATGTTCCGAGGCCTGGGCATTGAAACAGTT 240
Predicted       GCACAAAGGGAGTCTGTGGGTTTGGTCTCCATGTTCCGAGGCCTGGGCATTGAAACAGTT 115

Clone1          TCTAAGACCCCTCTGAAACGGGAAATGCTTAATCACTAGTGAATTCGCGGCCGCCTGCAG 241
Clone6          TCTAAGACCCCTCTGAAACGGGAAATGCTTAATC------GAATTCCCG-CGGCCGCCAT 293
Predicted       TCTAAGACCCCTCTGAAACGGGAAATGCTT------------------------------ 145

Clone1          GTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGT 301
Clone6          GGCGGCC------GGGAGCATGCGACGTCGGGCCCAATTCGCCCNNNNATCT-------- 339
Predicted       ------------------------------------------------------------
```

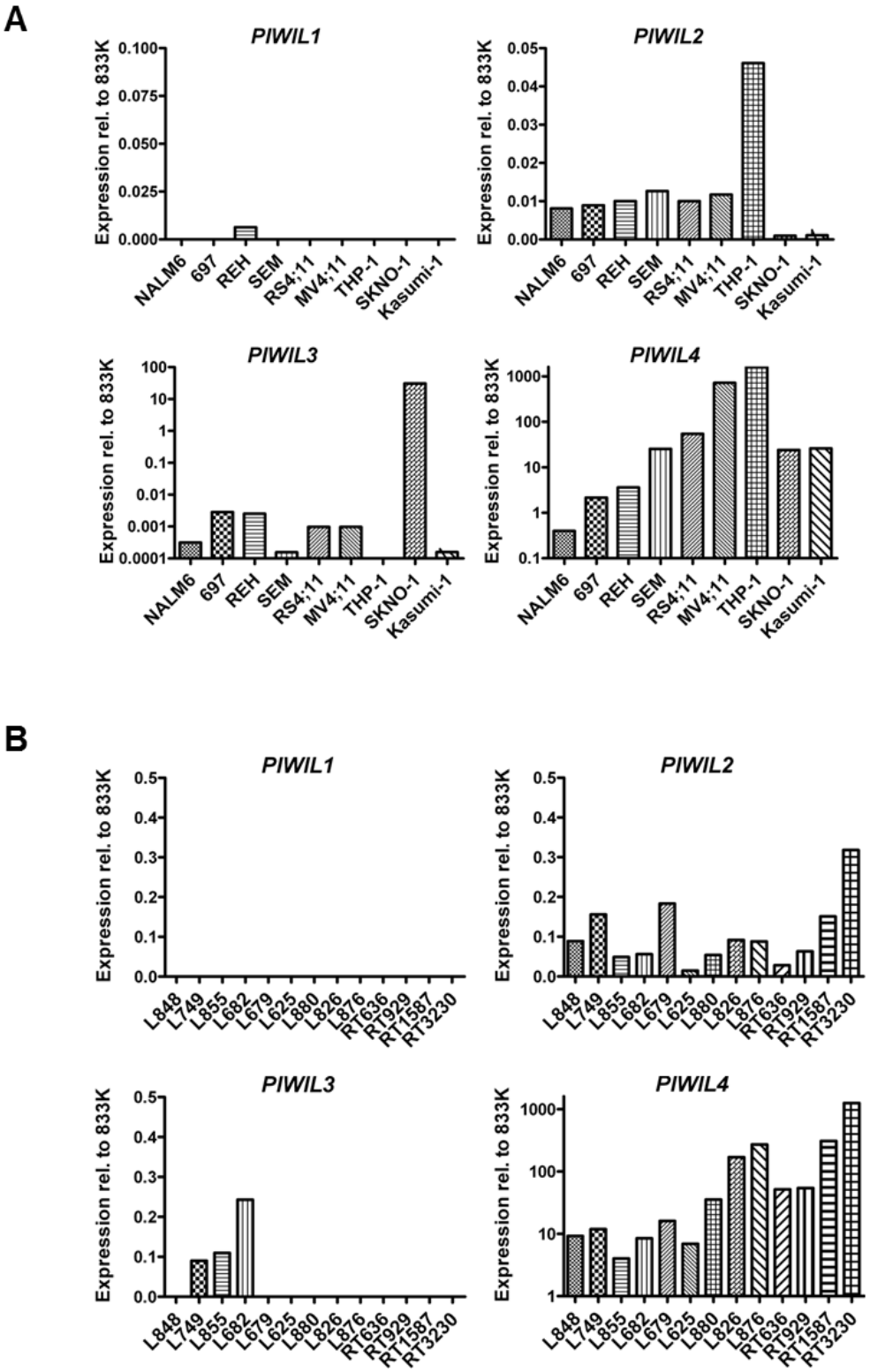
3.3.3 Expression of PIWI-Like genes in acute leukaemia

The expression of PIWI-Like genes was initially analysed in a panel of acute leukaemia cell lines (Figure 3-4A). Cells in stable culture were harvested and RNA extracted. cDNA was prepared and the expression of all 4 PIWI-Like genes analysed relative to their expression in the testicular seminoma cell line 833K. There was substantial variation between cell lines and between PIWI-Like genes. *PIWIL1* was not expressed in either ALL or AML cell lines. *PIWIL2* was expressed in all ALL and *MLL* rearranged AML cell lines, but not in AML cell lines with the *RUNX1-RUNX1T1* translocation. Expression was 1-10% of that seen in 833K. Expression of *PIWIL3* was more variable between cell lines, showing no correlation with ALL versus AML or with *MLL* rearrangement. *PIWIL4* was expressed in all cell lines. In all but NALM-6, *PIWIL4* expression was substantially higher than in 833K cells.

Expression of PIWI-Like genes was then analysed in a panel of patient-derived (details in Appendix A) acute lymphoblastic leukaemias (Figure 3-4B). RNA was extracted from stored viable bone marrow cells. RNA samples RT636, RT929, RT1587 and RT3230 were kindly donated by Ronald Stam, Rotterdam. As with the derived cell line analyses, *PIWIL1* was not expressed. *PIWIL2* was expressed in all patient-derived samples, at levels comparable to those seen in the ALL and *MLL* rearranged AML cell lines, (Figure 3-4A). *PIWIL3* expression was not consistent across samples, whilst *PIWIL4* expression was again found to be 10-1000 times that seen in 833K cells.

These analyses, together with the non-quantitative preliminary data in Figure 3-1, demonstrate the consistent expression of *PIWIL2* in childhood B precursor ALL. Expression of *PIWIL4*, which is believed to be ubiquitous, was both substantially higher than *PIWIL2* expression and higher than expression of *PIWIL4* in 833K cells which were chosen as a positive control cell line due to their robust expression of PIWI-Like genes. *PIWIL1* was not found to be expressed, whilst *PIWIL3* expression showed substantial variation.

Figure 3-4. Expression of PIWI-Like genes in acute leukaemia. A) Derived cell line expression: NALM-6, 697, REH – B precursor ALL; SEM, RS4;11 – *MLL* rearranged B precursor ALL; MV4;11, THP-1 – *MLL* rearranged AML; SKNO-1, Kasumi-1 – *RUNX1/RUNX1T1* translocated AML. B) Patient-derived B-precursor ALL samples – L880, L826, L876, RT636, RT929, RT1587, RT3230 have rearrangement of *MLL*.



3.3.4 Expression of *PIWIL2* in sorted lymphoid populations

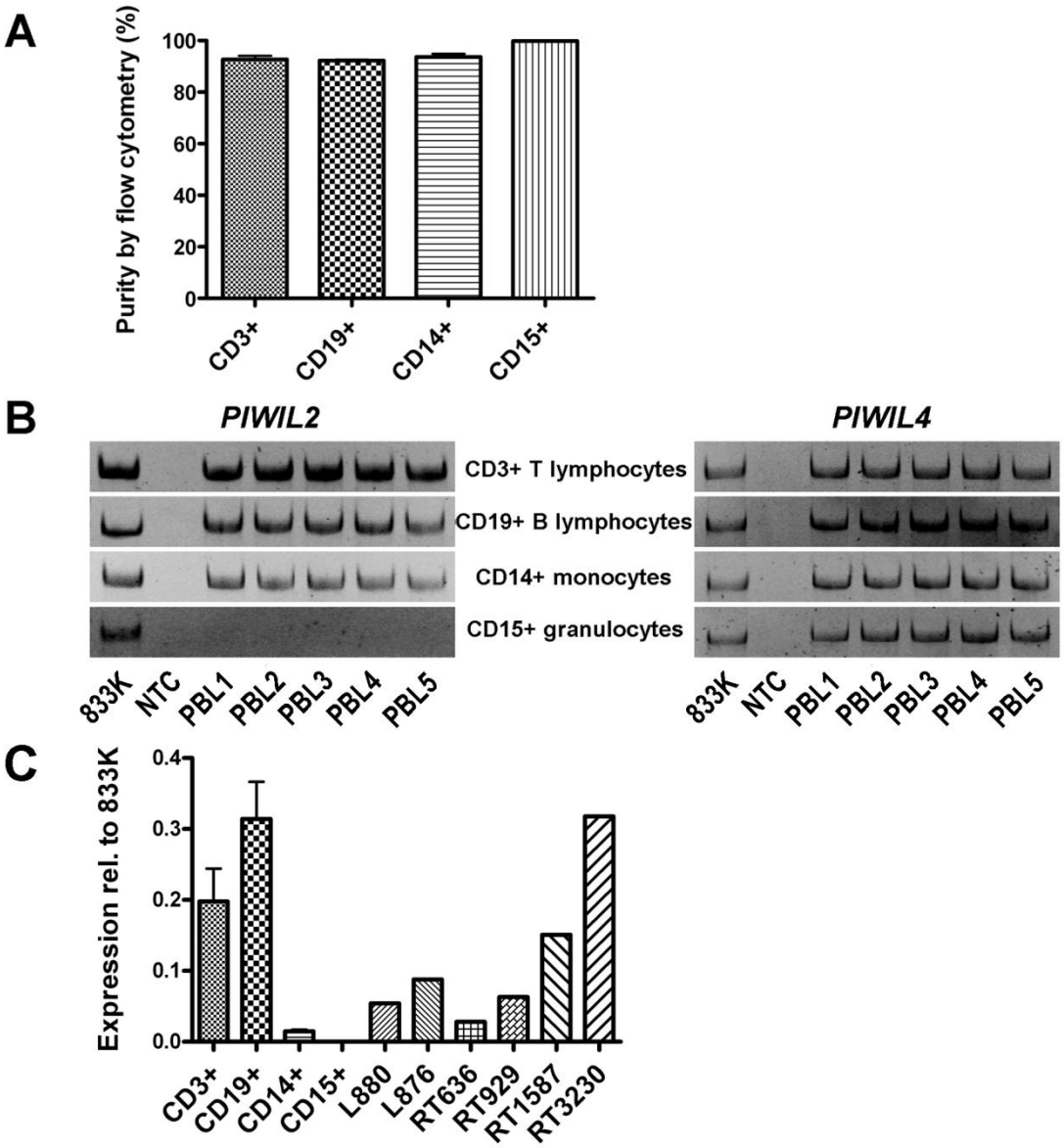
Normal mononuclear cells had been isolated from the peripheral blood of healthy volunteers to provide a negative control for the analysis of PIWI-Like gene expression detailed in 3.3.3. Unexpectedly, repeated qRT PCR on five different volunteer samples consistently demonstrated expression of *PIWIL2* and *PIWIL4*. *PIWIL1* and *PIWIL3* were not expressed.

In order to identify the peripheral blood leukocyte population responsible for the unexpected expression, CD19+ B lymphocyte, CD3+ T lymphocytes, CD14+ monocytes and CD15+ granulocytes were isolated from five further peripheral blood samples donated by healthy volunteers. Isolation was performed by immunomagnetic bead sorting with Macs Microbeads (Miltenyi) directed against CD19, CD3, CD14 or CD15. Purified cells were labelled with the respective fluorochrome labelled antibody and analysed by flow cytometry. This demonstrated a high degree of cell purity (mean; range): CD19+ - 92.2%, 91.3-93.1%; CD3+ - 92.7%, 89.5-97.7%; CD14 – 93.6%, 89.6-97.4%; CD15+ - 99.9%, 99.6-100% (Figure 3-5A).

End point analysis following qRT PCR demonstrated expression of *PIWIL4* in all populations, but expression of *PIWIL2* only in CD19/CD3+ lymphocytes and CD14+ monocytes, Figure 3-5B. Quantitative analyses were compared using one way analysis of variance with Bonferroni correction for multiple testing. This demonstrated significantly lower expression of *PIWIL2* in monocytes compared with both CD19+ B lymphocytes and CD3+ T lymphocytes ($p<0.05$) (Figure 3-5C). As the monocyte fraction is derived from the mononuclear cell layer formed by density gradient centrifugation it is likely that the expression seen in the CD14+ monocyte fraction is consistent with 3-10% lymphocyte contamination present in isolated CD14+ populations. Further purification was not attempted due to the limited cell numbers, but pooling samples and performing serial sorts, as described in 3.3.5, could identify the origin of this low level expression. Expression in B and T cells was comparable with that seen in patient-derived ALL specimens (*MLL* rearranged specimens from Figure 3-4B presented for comparison) (Figure 3-5C).

This analysis shows that, in contrast to previous reports, *PIWIL2* expression in healthy controls is found outside of the germline stem cell compartment. Restriction of expression to lymphocytes is consistent with a role in "self-renewal", as lymphocytes, unlike myeloid lineage cells, retain the ability to clonally expand in response to either successful B cell receptor rearrangement or antigenic stimulation. Importantly, this capacity is not lost with terminal differentiation. As patient-derived leukaemias express similar levels of *PIWIL2* to normal lymphocytes, it is possible that leukaemic blasts may access the clonal expansion programme of normal lymphoid cells to facilitate leukaemic expansion/propagation.

Figure 3-5. Expression of *PIWIL2* and *PIWIL4* in normal peripheral blood leukocytes. A) Purity of isolated leukocyte populations by flow cytometry (mean + standard error of the mean, s.e.m). B) End-point analysis of *PIWIL2* and *PIWIL4* expression by RT PCR. Expression is shown in B and T lymphocytes and monocytes, but *PIWIL2* expression is absent from granulocytes. C) Quantitative analysis of *PIWIL2* expression shows significantly higher expression in B ($p<0.01$) and T ($p<0.001$) lymphocytes compared with either myeloid lineage (mean + s.e.m). Expression is comparable with that seen in leukaemic populations.

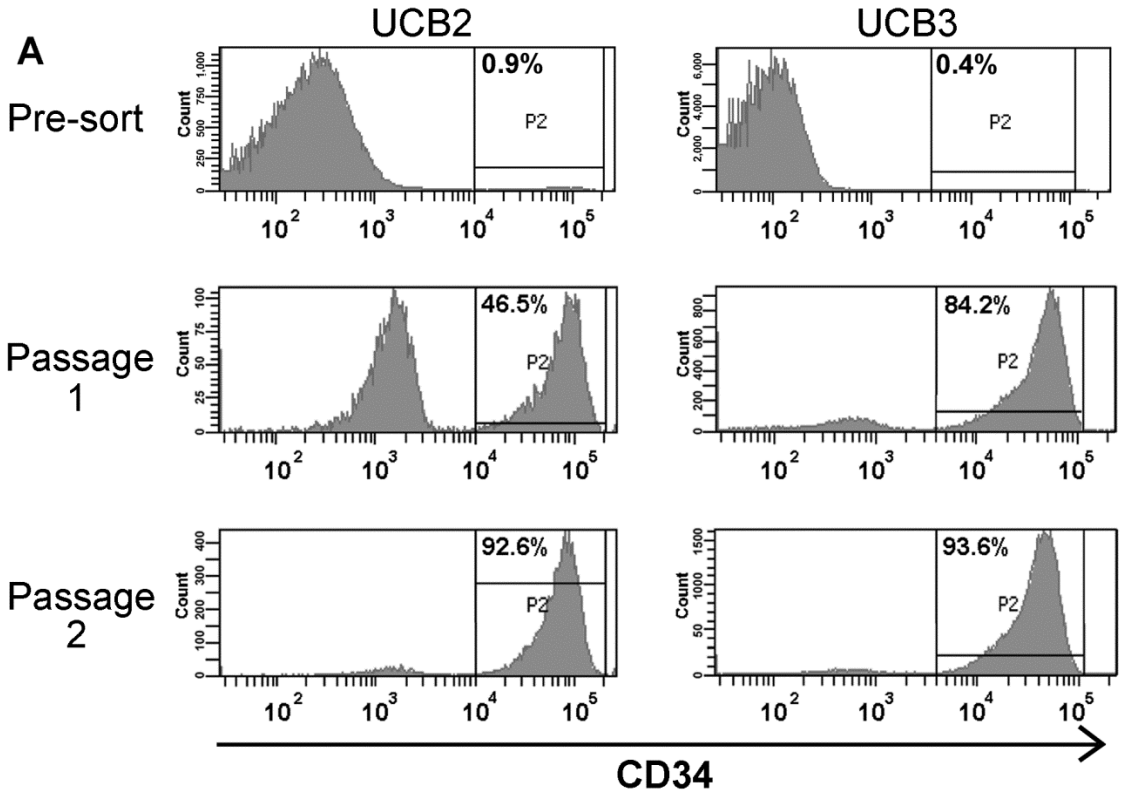


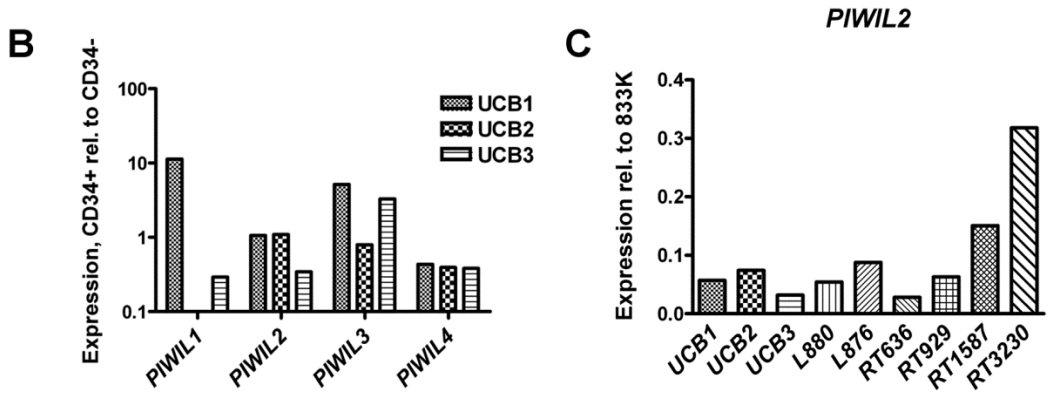
3.3.5 Expression of *PIWIL2* in CD34+ umbilical cord blood cells

Following the unexpected finding of *PIWIL2* expression in mature peripheral lymphocytes, the expression of PIWI-Like genes was assessed in CD34+ haematopoietic progenitor cells to investigate a wider role for these genes during haematopoiesis. Stored umbilical cord blood cells were thawed, labelled with CD34 Macs Microbeads and sequentially passaged down two Macs columns. Aliquots were taken for assessment of CD34+ cell purity prior to isolation and after each column and demonstrated a final purity of greater than 90% compared with <1% in un-purified cord blood (Figure 3-6A). This analysis was not performed for the first specimen as there were insufficient cells.

Expression of the PIWI-Like genes was compared with the flow-through from the Macs isolation columns, representing a source of CD34- cord blood cells. This analysis did not demonstrate enrichment for expression of *PIWIL2*, or any other PIWI-Like gene, in CD34+ umbilical cord blood progenitor cells (Figure 3-6B). This is in contrast to a previous report of a role for *PIWIL1* in human CD34+ bone marrow progenitor cells (Sharma, Nelson et al. 2001). The lack of enrichment in CD34+ cells may result from the presence of a substantial population of lymphocytes in the CD34- population. As the expression of *PIWIL2* was comparable with that seen in patient-derived leukaemias (Figure 3-6C) and normal lymphocytes (Figure 3-5C), greater fractionation of umbilical cord blood, achieved by multicolour FACS sorting, would allow accurate assessment of the differential expression throughout haematopoiesis.

Figure 3-6. Analysis of PIWI-Like gene expression in CD34+ umbilical cord blood progenitor cells. A) Analysis of purity of umbilical cord blood CD34+ cells by flow cytometry. Purity increases sequentially from <1% in both umbilical cord blood (UCB) 2 and 3 to greater than 90%. UCB1 – insufficient cells for analysis. B) Expression of PIWI-Like genes in isolated CD34+ cells compared with CD34- umbilical cord blood cells. C) Expression of *PIWIL2* in purified CD34+ cells compared with patient-derived leukaemic specimens.





3.3.6 Transient transfection with siRNA - confirmation of knockdown

To investigate the functional significance of *PIWIL2* in ALL, transient transfection of the childhood ALL cell line SEM was performed by electroporation in the presence of 500 nM siRNA. Following electroporation, qRT PCR was used to assess the reduction in *PIWIL2* mRNA levels, known as the knockdown. The controls used were siMA6, specific for the *MLL/AF4* breakpoint present in SEM cells, siAGF1, specific for the *RUNX1/RUNX1T1* breakpoint found in Kasumi-1 cells and not present in SEM cells, and Mock in which no siRNA was used but cells were electroporated just as for all other experimental samples.

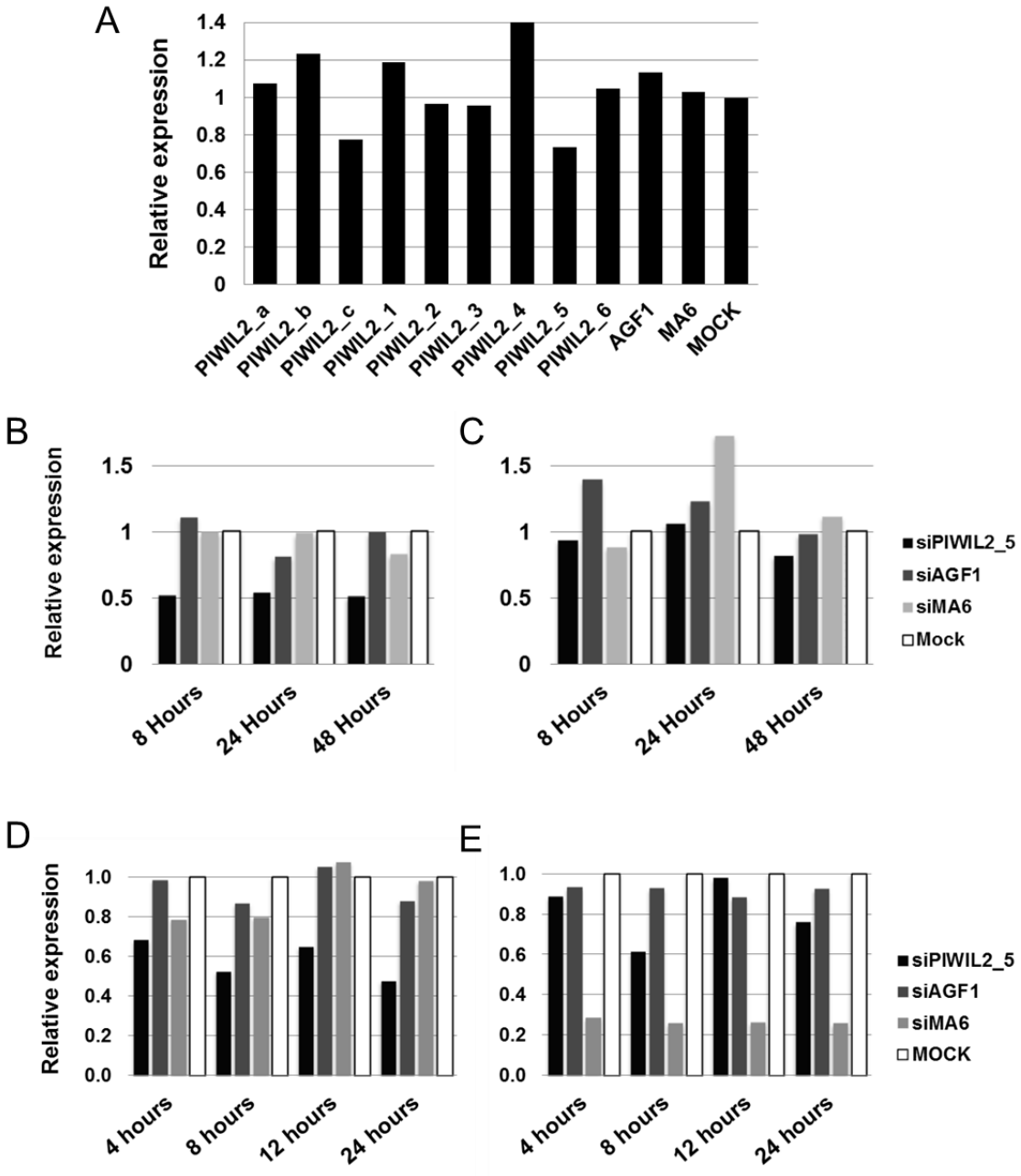
Initially, the efficacy of a number of different siRNAs from different manufacturers was assessed at 24 and 48 hours post electroporation (Figure 3-7A). The two siRNAs found to have the greatest effect, siPIWIL2_c and siPIWIL2_5 (designed by two independent manufacturers) were subsequently identified as having the same sequence.

As only modest knockdown (40-50%) was achieved with the most effective siRNA, siPIWIL2_5, the cellular localisation of *PIWIL2* mRNA was investigated. As experimental knockdown occurs only in the cytoplasm (where siRNAs and RISC co-localise), a nuclear predominance of *PIWIL2* transcript would mask potentially effective cytoplasmic knockdown in whole cell RNA preparations. Fractionation of cytoplasmic and nuclear mRNA was followed by qRT PCR, demonstrating a 2-3 fold higher level of mRNA in the nucleus than in the cytoplasm. However, knockdown in the cytoplasm remained approximately 50% (Figure 3-7B), whilst no knockdown was seen in the nuclear fraction (Figure 3-7C).

A subsequent early time course experiment aimed to identify whether a very early peak knockdown had been missed by sampling at 24 hours. This showed that the peak *PIWIL2* knockdown of approximately 50% was achieved by 8 hours and maintained thereafter (Figure 3-7D). The positive control experiment demonstrates 75-80% knockdown of the *MLL/AF4* fusion transcript by siMA6 as early as 4 hours post electroporation (Figure 3-7E).

As it was not possible to identify an experimental reason for the modest knockdown achieved, a further three commercially designed siRNAs were purchased. The best amongst these, siPIWIL2_e, was found to have a similar efficacy to siPIWIL2_5. These two siRNAs were used for subsequent experiments of the functional effect of PIWIL2 in SEM cells, see 3.3.8, either individually or in combination at 250 nM each (697 cells only).

Figure 3-7. Knockdown of *PIWIL2* electroporation with siRNAs. A) Knockdown 48 hours after transfection. PIWIL2_c and PIWIL2_5 show the greatest knockdown. (PIWIL2_4 result 1.89, not shown in full). Fractionation of cytoplasmic and nuclear mRNA - B) Cytoplasmic knockdown remains approximately 50% whilst C) no nuclear knockdown is demonstrated. D) Time course of *PIWIL2* knockdown. 50% reduction in mRNA is achieved by 8 hours. E) Positive control demonstrates ≥75% knockdown of *MLL-AF4* by 4-8 hours. Relative expression is ΔΔCt relative to Mock.



3.3.7 Analysis of knockdown of protein by Western blot

In order to confirm the effect of mRNA knockdown on PIWIL2 protein levels, protein collected at the same time as mRNA was probed for PIWIL2 by Western immunoblotting. Proteins were separated by gel electrophoresis using an 8% polyacrylamide gel and blotted onto PVDF membrane. A number of commercially available antibodies were used to probe for PIWIL2 protein, as described below.

Initially, a polyclonal antibody from Abnova was used. An exposure for 10 seconds produced a very clear band at 40 kDa (Figure 3-8A). However, full length PIWIL2 protein is approximately 110 KDa. In retrospect, the manufacturer's website specified a 187 amino acid, 40 kDa target. Analysis of the immunogenic sequence showed no homology to PIWIL2. No further experiments were performed with this antibody.

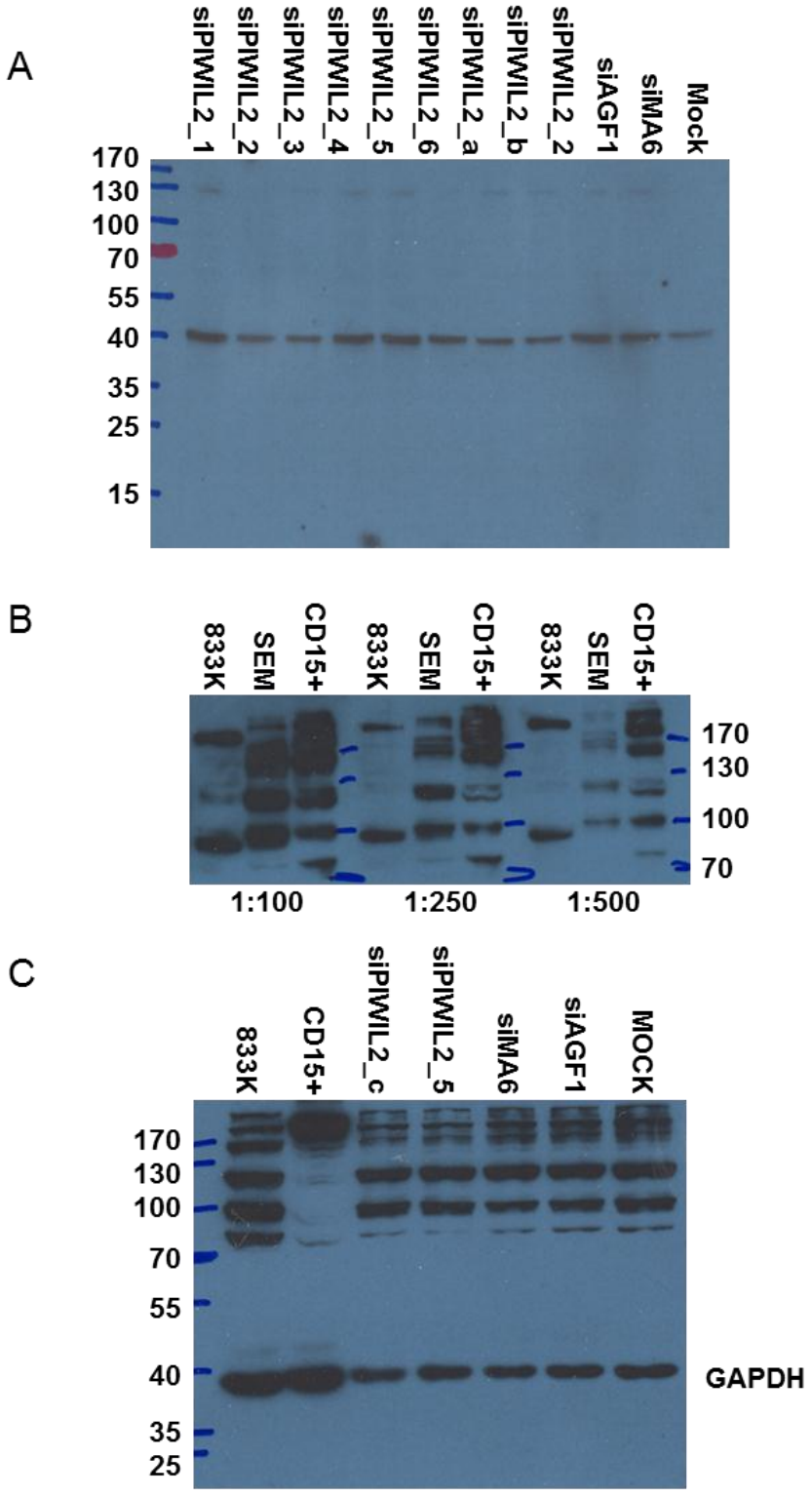
The next antibody to be tested was a polyclonal antibody from Santa Cruz. An initial blot was performed using antibody dilutions between 1:100 and 1:500, as suggested by the manufacturer. This produced multiple bands between 70 and >170 kDa, but with no band present at the correct size (Figure 3-8B). In order to try and reduce non-specific background labelling, a repeat experiment used antibody dilutions of 1:1000-1:25000. Up to 1:5000 dilution gave similar bands to the first trial, but with no clear band at 110 kDa.

A trial blot of samples taken 24 hours following electroporation was performed with the Santa Cruz antibody. There was no observable decrease in the size of any of the bands (Figure 3-8C), despite knockdown of mRNA (Figure 3-7D).

A third antibody from AbCam was trialled in two separate experiments, but gave no signal on either occasion. The loading control analysis of GAPDH gave clear bands of the appropriate size.

Following the failure to demonstrate a band of the correct size, or any band which was modulated by knockdown, using any of these three antibodies, and with no other commercial antibodies available, it was decided to monitor knockdown with qRT PCR alone and investigate whether any functional phenotype could be elicited.

Figure 3-8. Western immunoblot analysis of PIWIL2. A) Abnova antibody demonstrates 40 kDa protein of unknown identity. B) Initial dilutions of Santa Cruz antibody from 1:100-1:500 demonstrated multiple bands, but no clear predominance at 100 kDa. C) Santa Cruz antibody following 48 hours of knockdown with siPIWIL2_c and siPIWIL2_5. No difference is seen in any of the bands present. 833K – Positive control. CD15+ - Granulocyte – Negative control. GAPDH – Loading control. Marker sizes in kDa are recorded alongside each blot.



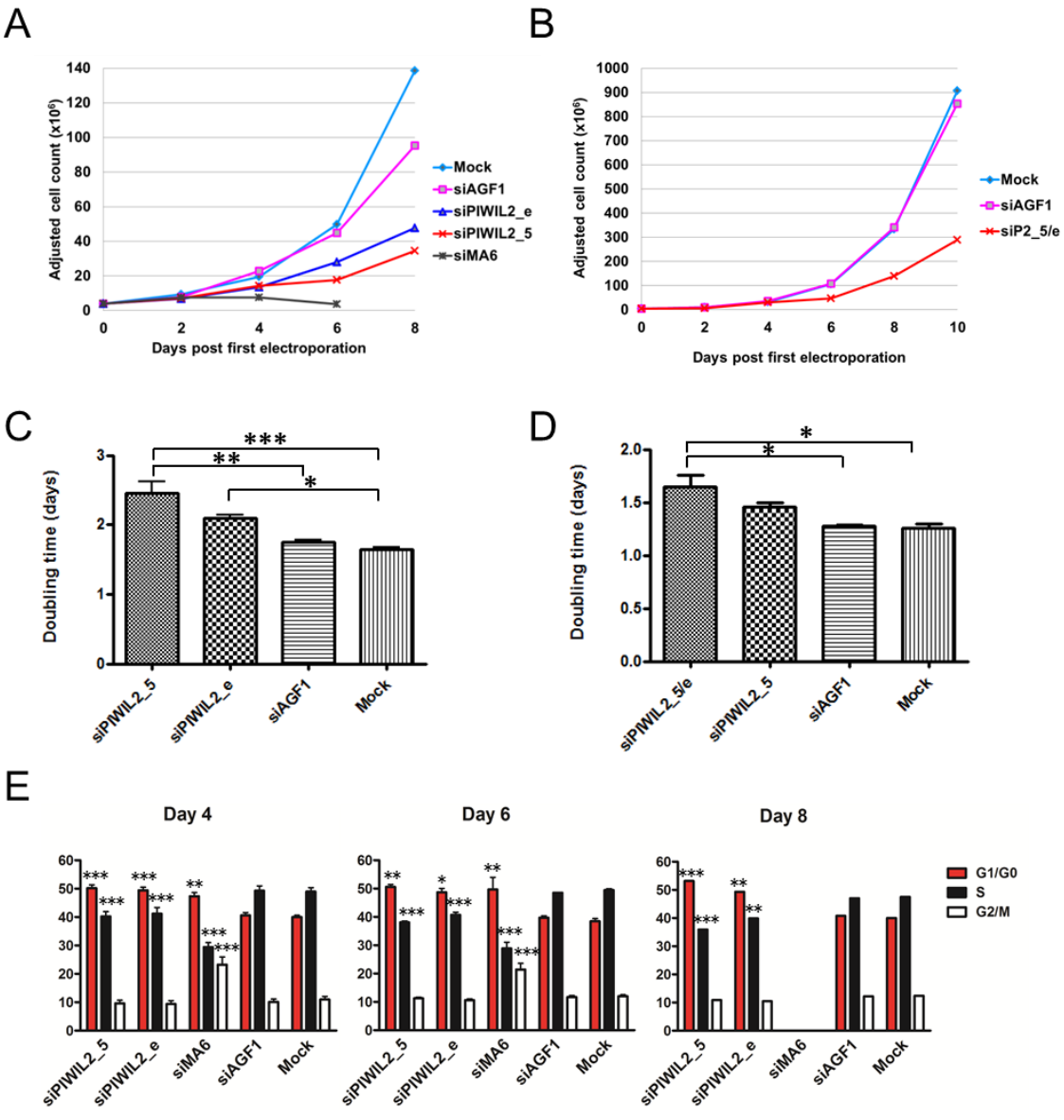
3.3.8 Functional assessment following *PIWIL2* knockdown in SEM cells

Having established a loss of function assay with electroporation and siRNA, an assessment of the functional significance of PIWIL2 was made using serial electroporation of both SEM (performed by SB) and 697 (performed by SB and Hesta McNeill) cells to provide sustained knockdown. Electroporation was performed every 48-72 hours. At each of these time points, cells were counted and an aliquot taken for re-electroporation. Further aliquots were taken for RNA extraction, to confirm knockdown, and for analysis of cell cycle distribution. For SEM cells, three independent experiments were performed over 6 (n=1) or 8 (n=2) days. For 697 cells, three independent experiments were performed over 10 (n=2) or 14 (n=1) days.

Serial cell counts were adjusted as described in 2.3.4.3. These data were then plotted against time and an exponential trendline fitted. From the equation describing the trendline, $y = a \times e^{\mu x}$, doubling times were calculated as $td = \mu / ln^2$. Doubling times were analysed using one way analysis of variance with a Bonferroni correction for multiple testing, with each *PIWIL2* specific siRNA being compared to each negative control.

PIWIL2 knockdown resulted in decreased proliferation in both SEM (Figure 3-9A) and 697 cells (Figure 3-9B). This was quantified as a significant prolongation of doubling times in both SEM and 697 cells (Figure 3-9C and D). Cell cycle analysis of SEM cells demonstrated a significant, although not substantial, G1/G0 arrest, accompanied by a decreased proportion of cells in S phase, following *PIWIL2* knockdown (Figure 3-9E). Cell cycle analysis in 697 cells did not show a consistent pattern of change of the distribution of cells within the cell cycle.

Figure 3-9. *PIWIL2* knockdown affects cell proliferation and cell cycle distribution. A) Proliferation of electroporated SEM cells. B) Proliferation of electroporated 697 cells. Electroporation performed with combination of 250 nM each of siPIWIL2_5 and siPIWIL2_e. Cell counts were performed 2 days after each electroporation. A single representative experiment is shown in each of A and B. Doubling times of electroporated C) SEM or D) 697 cells. Data are mean of three experiments, error bars show standard error of the mean. E) Cell cycle analysis performed 2 days after each serial electroporation of SEM cells. Cell cycle phases are compared against the same phase in Mock treated cells at the same time point. Data are mean values of three experiments, error bars show standard error of the mean. For panels B-D, * p<0.05, ** p<0.01, *** p<0.001.



3.4 Discussion

The importance of *PIWIL2* in mammalian germline stem cells (Kuramochi-Miyagawa, Kimura et al. 2004; Aravin, Hannon et al. 2007; Carmell, Girard et al. 2007), combined with its expression in a wide range of human malignancies (Lee, Schutte et al. 2006; Feng, Peng et al. 2009; He, Chen et al. 2010; Lee, Jung et al. 2010; Yin, Li et al. 2011) suggested that *PIWIL2* may play a critical role in the regulation of self-renewal, or alternative stem cell characteristics, in cancer. This project was designed to investigate the role of *PIWIL2* in childhood acute lymphoblastic leukaemia, particularly in the context of identifying genetic determinants of leukaemia propagating cell biology.

Expression of *PIWIL2* and *PIWIL4* was demonstrated in childhood acute lymphoblastic leukaemia for the first time, using both derived cell lines representing a spectrum of cytogenetic backgrounds and patient derived specimens. During the validation of the qRT PCR analysis of PIWI-Like expression, peripheral blood mononuclear cells were used as a source of negative control RNA. Unexpectedly, expression of *PIWIL2* was also identified in these control cells. By fractionating the peripheral blood leukocytes it was possible to demonstrate that *PIWIL2* was expressed in both B and T lymphoid cells, but not in myeloid lineage cells. Furthermore, expression was not seen in either of two acute myeloid leukaemia cell lines harbouring the *RUNX1/RUNX1T1* translocation. These findings demonstrate that *PIWIL2* expression is a feature common to both normal and malignant lymphoid populations. As the role of *PIWIL2* in normal lymphocytes is not known, and the present study was unable to determine its role in acute leukaemia, it remains to be determined whether *PIWIL2* is critical to cellular self-renewal, as it is in germline stem cells, or whether it has an alternative function in lymphoid malignancy. Intriguingly, *PIWIL2* may provide an example of a normal lymphoid derived genetic determinant of leukaemic biology. It will be important to determine whether or not *PIWIL2* represents one component of a clonal expansion programme which is shared by normal and malignant lymphoid populations.

Expression was also seen in the *MLL* rearranged AML cell lines MV4;11 and THP-1. This may result directly from the powerful transforming potential of this translocation, although knockdown of *MLL/AF4* in SEM cells did not affect *PIWIL2* expression. A broader analysis of AML samples is required to know whether expression is restricted to *MLL* rearranged AML or whether *RUNX1/RUNX1T1* associated AML is unusual in its low expression. As expression of the PIWI-Like genes in human umbilical cord blood stem cells was not clearly defined by the present study, it remains to be determined what impact the cell of origin might have on *PIWIL2* expression.

Modulating *PIWIL2* expression in both SEM and 697 cells proved challenging, with 40-60% knockdown being the best achieved. However, despite achieving only modest knockdown, a consistent phenotype was observed, initially in SEM cells. Reduction in *PIWIL2* expression in SEM cells resulted in a decrease in cellular proliferation which may be accompanied by a G1/G0 cell cycle arrest. Repeat experiments in 697 cells also demonstrated a significant decrease in cellular proliferation, but no consistent effect on cell cycle distribution was seen. This may again result from the modest degree of knockdown achieved, or may indicate an alternative causation for the reduced proliferation other than cell cycle arrest.

To maintain a consistent degree of knockdown and accurate cell numbers over 5-7 serial electroporations of 697 cells proved challenging. Whilst it had been hoped that prolonged knockdown would allow the investigation of changes in promoter methylation in ALL cells, it was not possible to maintain consistent knockdown and rates of proliferation for long enough to be confident of identifying PIWI/piRNA methylation targets. This analysis was therefore not performed.

Following the departure of a collaborator, Professor Karim Nayernia, with substantial experience in PIWIL2 biology and with the difficulties experienced in: 1) achieving strong knockdown; 2) the lack of a protein based assay for demonstrating knockdown; 3) maintaining consistency over a large number of serial electroporations, it was decided to end the *PIWIL2* project. An alternative, unbiased approach to identifying the genetic determinants of ALL

propagating cell biology was adopted. This whole-genome RNAi screen is described in Chapter 5.

To end the PIWIL2 project was a strategic decision and not based on the data produced. Indeed, the finding of *PIWIL2* expression in normal lymphoid populations is consistent with the hypothesis underlying this Fellowship, that "acute lymphoblastic leukaemic blasts derive an intrinsic propagating potential from their lymphoid cell of origin". Development of our understanding of PIWIL2 biology has therefore remained interesting. This interest has grown in the light of recent publications determining a sequence specific role for PIWI/piRNAs outside of the germline in both neuronal cells (Rajasethupathy, Antonov et al. 2012) and especially in the lymphoid lineage derived NK cell (Cichocki, Lenvik et al. 2010). Promoter methylation is the mechanism of action in both of these settings, and provides a feasible mechanism of PIWIL2 mediated control of self-renewal/propagation in lymphoid cells. In addition, the identification of piRNAs in murine lymphoid cells which, in contrast to repeat sequence specific piRNAs found in haematopoietic precursors and myeloid lineage cells, show specificity for 5' UTRs, 3' UTRs, coding and intronic regions (Yan, Hu et al. 2011), suggests a specific role for PIWI/piRNAs in lymphoid cells. With our improved knowledge of PIWI/piRNA biology and with developments both in commercially available technology (improved methylation arrays, new antibodies) and particularly with the routine use of commercially available microRNA adapted shRNA lentivectors in our laboratory, there is now the opportunity to provide more consistent modulation of *PIWIL2* in lymphoid malignancy, allowing assessment of the effect of *PIWIL2* loss on promoter methylation. In parallel, a new collaboration will allow the identification of lymphoid piRNAs, providing an additional source of information on potential PIWIL2 targets. This project is described in greater detail in Chapter 9.

Chapter 4

Expression of stemness related genes in childhood ALL

1. Introduction

In acute myeloid leukaemia, a clear hierarchy has been identified within which leukaemia stem cells with the greatest repopulating capacity sit at the apex and have a CD34+CD38- immunophenotype, mirroring normal haematopoiesis (Lapidot, Sirard et al. 1994; Bonnet and Dick 1997). Leukaemia propagating capacity is now thought to be present in more diverse populations than was once believed, but it still appears that this capacity reduces as cells "differentiate" towards more mature immunophenotypes (Hope, Jin et al. 2004; Goardon, Marchi et al. 2011). Recent identification of stemness signatures has demonstrated real differences between leukaemic populations in expression of genes associated with normal haematopoietic stem cells, suggesting that leukaemia stem cells utilise stemness programmes derived from normal stem cells, either embryonic (Somervaille, Matheny et al. 2009) or haematopoietic (Eppert, Takenaka et al. 2011).

In contrast, leukaemia propagating cells in ALL are more frequent and present in a diverse range of immunophenotypic populations (Kong, Yoshida et al. 2008; le Viseur, Hotfilder et al. 2008; Rehe, Wilson et al. 2012)(Appendix D). Sorting blasts for expression of CD34 has demonstrated a correlation between the loss of expression of CD34 and up-regulation of several B-precursor associated genes, including *IRF4, MS4A1* (CD20), IgH constant region, and IgL κ and λ loci (le Viseur, Hotfilder et al. 2008). Recent re-analysis of this data, comparing the most differentially expressed genes within sorted CD34+ and CD34- leukaemic blasts, has demonstrated the similarity of these populations to their respective normal stem/precursor populations. CD34+ leukaemic blasts cluster with normal CD34+ haematopoietic stem cells. Furthermore, applying the HSC/LSC signature generated by Eppert *et al* (Eppert, Takenaka et al. 2011), CD34+ and CD34- ALL blasts cluster together and not in association with AML CD34+ cells or normal HSCs (Rehe, Wilson et al. 2012)(Appendix D), demonstrating that leukaemia propagating potential in ALL is not identified by this candidate LSC stemness signature.

This project aimed to analyse the expression of a number of candidate stemness related genes in sorted, patient-derived ALL specimens. These genes were derived from published studies of the control of haematopoietic or leukaemic stemness.

2. Aims of the project

The aims of this project were:

1. To sort patient-derived leukaemic blasts according to surface expression of CD34
2. To analyse expression of the candidate stemness genes *BMI1*, *CD34*, *EZH2*, *HMGA2*, *MEIS1* and *TERT* in sorted leukaemic blasts and umbilical cord blood stem cells

3. Results

4.1.1 Purity of sorted populations

Patient-derived leukaemic blasts (see Appendix A for patient details) were retrieved from storage in liquid nitrogen, thawed and labelled with CD19 PE and CD34 FITC antibodies. Primografted samples were also labelled with murine CD45 PECy7 and Ter119 PECy7 to exclude murine cells. Samples were then sorted using a FACS Vantage cell sorter (Hesta McNeill). For comparison, RNA was taken from umbilical cord blood cells sorted for expression of CD34, as described in 3.3.5.

Post sorting purity of leukaemic blasts was assessed using the FACS Vantage. Except for sample 2510, purity was >85% in all populations (Table 4-1). The purity of UCB1 was not assessed by flow cytometry as there were insufficient cells. However, the substantial difference in CD34 expression indicates a substantially enriched population (Figure 4-1A). UCB2 was labelled with CD34 APC and analysed by flow cytometry as described in 3.3.5.

Table 4-1. Purity of sorted patient-derived leukaemic cells.

Specimen	CD34+	CD34-
EMCR1	91.0%	98.5%
L784	87.8%	97.7%
L49101	96.8%	92.6%
2510	88.9%	63.7%
L4951	97.1%	95.6%
L826	94.7%	91.8%
Mean	92.7%	90.0%

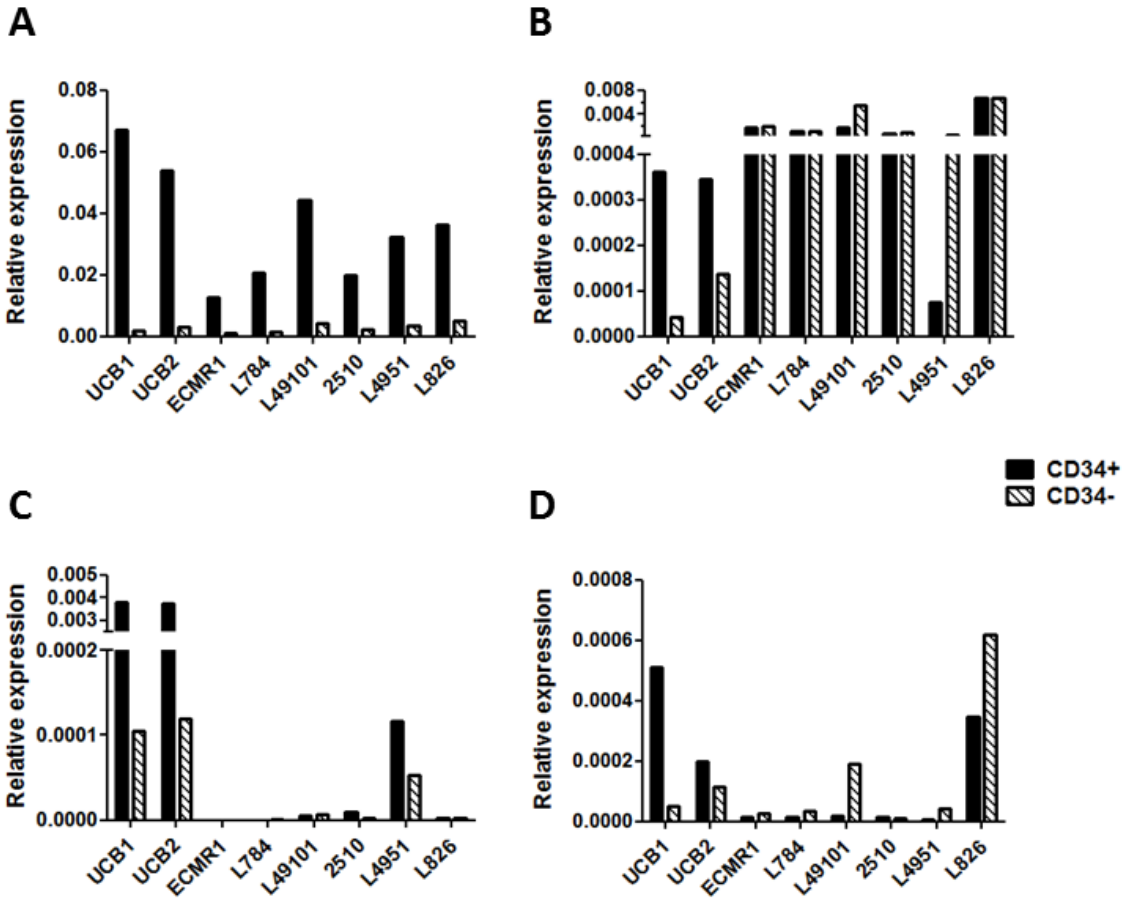
4.1.2 Expression of candidate stemness genes

Sorted populations were washed and RNA was extracted. Quantitative RT PCR was performed with primers specific for the candidate stemness genes *BMI1*, *CD34*, *EZH2*, *HMGA2*, *MEIS1* and *TERT*. Expression was analysed by ΔCt using GAPDH as a reference gene. Using Wilcoxon matched-pairs signed rank analysis, only CD34 expression differed significantly between sorted leukaemic populations, p=0.03, as it did for sorted umbilical cord blood specimens.

Unexpectedly, the polycomb group genes *BMI1* and *EZH2* showed no consistent difference in expression between CD34+ and CD34- populations from either leukaemic or umbilical cord blood specimens. However, expression of *TERT, MEIS1* and *HMGA2* showed substantially higher expression in CD34+ umbilical cord blood stem cells than in the corresponding CD34- population, as would be expected (Figure 4-1B-D).

In contrast to sorted umbilical cord blood HSCs, expression of *TERT* was not different between sorted leukaemic populations and overall was higher than in either of the HSC populations, suggesting an important function across both populations of leukaemic blasts (Figure 4-1B). Expression of *MEIS1* and *HMGA2* was more restricted. Neither of these genes showed higher expression in CD34+ cells, except for the expression of *HMGA2* in the Philadelphia chromosome positive ALL sample, L4951 (Figure 4-1C). However, the two fold higher expression seen in CD34+ leukaemic cells from this sample was substantially lower than the 31-36 fold difference in expression seen in sorted HSC populations. *MEIS1* was characteristically highly expressed in the one *MLL* rearranged sample tested (Figure 4-1D). Together, these findings argue against these candidate stemness genes having an important role in a CD34+ leukaemia stem cell population, consistent with the lack of a stem cell hierarchy in ALL.

Figure 4-1. Expression of candidate stemness genes in CD34+ and CD34- leukaemic blasts. A) CD34 expression is substantially higher in CD34+ cells (black bars) from both umbilical cord blood (UCB) and leukaemic blasts. B) *TERT* expression is substantially higher in CD34+ UCB cells than in CD34- cells (hashed bars). In leukaemic blasts, no difference in *TERT* expression is seen. The same pattern is seen in expression of both C) *HMGA2* and D) *MEIS1*, although expression in leukaemic blasts is less consistent. Relative expression is $2^{-\Delta Ct}$, reference gene GAPDH



4. Discussion

As leukaemia propagating cells in ALL are present in both CD34+ and CD34- sorted populations (Kong, Yoshida et al. 2008; le Viseur, Hotfilder et al. 2008), both populations would be predicted to express genes required for extensive cellular proliferation. This project aimed to assess the expression of a number of candidate stemness genes in sorted CD34+ and CD34- leukaemic blasts using qRT PCR. A normal haematopoietic control was provided by two umbilical cord blood specimens, also sorted for expression of CD34. The differential expression of the CD34 gene in all samples provided a control for the validity of this approach.

The first candidate showing differential expression in umbilical cord blood stem cells is telomerase reverse transcriptase (*TERT*). *TERT* encodes a key component of the telomerase complex, which functions to maintain telomere length and prevent cellular senescence. As such it is an essential component of the self-renewal programme of normal and malignant stem cells. In leukaemia, *TERT* expression is under the control of the leukaemic fusion genes *MLL/AF4* and *RUNX1/RUNX1T1 (Gessner, Thomas et al. 2010)*. Down-regulation resulted in reduced clonogenicity and induction of replicative senescence. Baseline expression is however, too low for *TERT* expression to be reliably represented in expression array analyses and therefore it is not surprising that it does not appear in the stemness signature of Eppert *et al* (Eppert, Takenaka et al. 2011).

MEIS1 is a homeobox protein implicated in the leukaemogenesis of *MLL* rearranged leukaemias and contained in the AML stemness signature defined by Eppert *et al* (Eppert, Takenaka et al. 2011). *MEIS1* is up-regulated by *MLL* fusion proteins, whilst knockdown of *MEIS1* in *MLL* rearranged leukaemic cell lines impairs engraftment and proliferation in murine xenograft models (Kumar, Li et al. 2009; Orlovsky, Kalinkovich et al. 2011). The present project demonstrated higher expression of *MEIS1* in CD34+ cord blood HSCs, compared with CD34- HSCs. Leukaemic expression was low in all samples other than L826, the only *MLL* rearranged leukaemia in the panel tested and

L4951, a Philadelphia chromosome positive ALL. In both of these samples, however, expression was higher in the CD34- population.

HMGA2 encodes the high mobility group protein A2 which is highly expressed during embryogenesis but is then down-regulated (Gattas, Quade et al. 1999). It functions to control growth and development, as evidenced by abnormal growth both in mutated mice (Zhou, Benson et al. 1995; Battista, Fidanza et al. 1999) and in a child with constitutional mutation of this gene (Ligon, Moore et al. 2005). *HMGA2* expression has been identified in haematopoietic stem cells as well as blasts from both acute and chronic myeloid leukaemias (Meyer, Krisponeit et al. 2007). Expression in the present study was limited to four samples, although levels were substantially higher in sample L4951, positive for the Philadelphia chromosome. Overall there was no evidence for higher expression in CD34+ leukaemic cells.

The polycomb group proteins BMI1 and EZH2 were not differentially expressed in either sorted HSCs or leukaemic blasts. This was a somewhat unexpected finding as both of these genes are essential to the maintenance of HSC self-renewal capacity by generating bivalent chromatin domains (Konuma, Oguro et al. 2010) and control the proliferative potential of leukaemia stem cells (Lessard and Sauvageau 2003). The reasons for the lack of differential expression in the present study are not clear. It is possible that the single marker sorting strategy was insufficiently effective in isolating HSCs resulting in analysis of a mixed population.

This project has shown that the candidate leukaemic stemness genes *TERT*, *MEIS1* and *HMGA2* are differentially expressed in sorted CD34+ and CD34- normal haematopoietic stem cells. In contrast, expression in sorted leukaemic blasts showed no bias towards higher expression in CD34+ cells consistent with the lack of a stem cell hierarchy in acute lymphoblastic leukaemia (Rehe, Wilson et al. 2012)(Appendix D).

Chapter 5

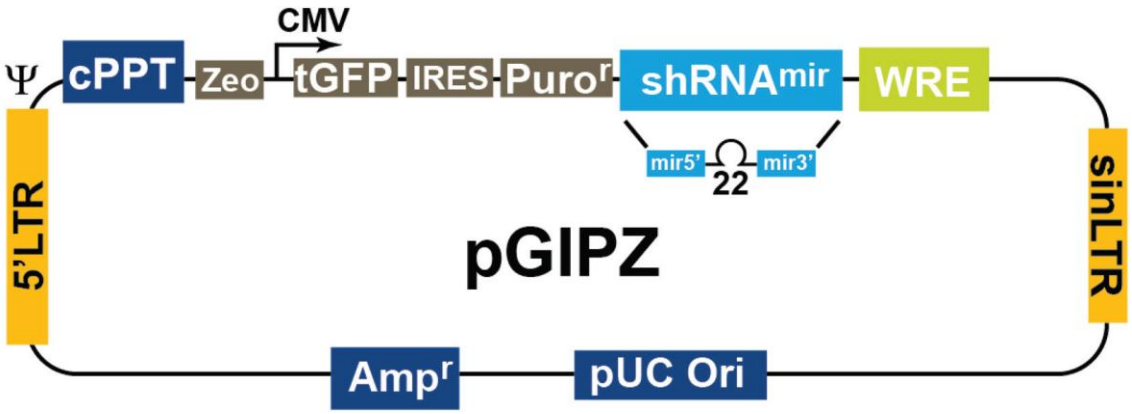
Identifying novel leukaemia propagating genes using a genome-wide RNAi screen

5.1 Introduction

Following the decision to stop the candidate (PIWIL2) self-renewal gene project, the alternative strategy to investigate the cancer stem cell programme in childhood acute lymphoblastic leukaemia was to perform an unbiased, genome-wide functional screen. This project used a commercially available short hairpin RNA (shRNA) library, produced by Open Biosystems. The Decode™ negative selection library contains approximately 70000 different lentiviral constructs, each coding for a unique shRNA. This provides coverage of the entire genome with 2-3 constructs per gene. The library is divided into 7 pools, each of which contains lentiviral particles produced from approximately 10000 individual constructs.

The transfer vector used is the Open Biosystems pGIPZ vector, Figure 5-1. Here, expression of a single transcript is driven by the RNA polymerase II promoter CMV and codes for Turbo GFP (reporter gene), puromycin N-acetyl-transferase (PAC, mammalian selection) and an shRNAmir30 construct (RNAi). Cloning the shRNA within a microRNA (mir30) sequence is believed to result in the transcript being cleaved by the normal microRNA processing pathway, namely Drosha within the nucleus, active transport into the cytoplasm via Exportin 5 and subsequent Dicer cleavage followed by active loading into the RNAi Induced Silencing Complex, RISC, see section 1.3. This results in more effective loading of shRNA into RISC and improved degradation of target transcript – known as knockdown (Dickins, Hemann et al. 2005; Silva, Li et al. 2005). This is a critical feature in a pooled library screen as each target cell must be transduced with only a single viral particle to allow deconvolution of constructs at the end of the experiment. That is, if cells were to contain multiple constructs, it would not be possible to determine which of those constructs were responsible for the altered phenotype of the cell. Immediately 3' to the shRNAmir construct is an approximately 60 nucleotide barcode sequence which is specific for the hairpin and is used to deconvolute the pools by microarray analysis following completion of the screen.

Figure 5-1. The Open Biosystems pGIPZ vector. A single transcript codes for the turbo GFP, puromycin selection marker and short hairpin RNA which is sited within a larger sequence derived from microRNA30. cPPT – Central polypurine tract, Zeo – zeomycin selection marker, CMV – CMV promoter, tGFP – Turbo GFP reporter, IRES – Internal Ribosomal Entry Site, Puro^r – Puromycin-N-Acetyltransferase, shRNAmir – short hairpin construct in mir30 context, WRE – Woodchuck response element, sinLTR – self-inactivating 3' long terminal repeat, pUC Ori – E.coli origin of replication, Amp^r – ampicillin resistance, 5'LTR – 5' long terminal repeat (Image taken from www.openbiosystems.com/rnai/shrnamirlibraries/gipzlentiviralshrnamir).



5.2 Aim of the project

This project sought to develop a genome-wide RNAi screen which would allow the identification of novel candidate genes important to the self-renewal of acute lymphoblastic leukaemia propagating cells. These genes, or the pathways to which they belong, may offer novel therapeutic opportunities. Specifically, the aims of this project were to:

1. Optimise transduction of ALL cell lines using the pGIPZ lentiviral system to allow screening with the Decode shRNA library
2. Investigate the potential for performing the screen using co-culture with bone marrow stromal feeder cells to simulate leukaemic-niche interactions
3. Perform a trial screen using a single lentiviral pool

5.3 Results

5.3.1 Optimisation of Decode screening methodologies

5.3.1.1 Optimisation of lentiviral transduction

Transduction was performed using a spinfection technique. The initial work within this project involved optimising the spinfection to maximise the cell coverage available from each vial of titred virus provided with the library. These experiments also allowed estimation of the transduction efficiency of leukaemic cells lines. Comparison of transduction efficiency relative to HEK293T cells, within which the library had been titred by the manufacturer, allows a calculation of the volume of virus required to achieve a desired rate of transduction in a known number of cells. This is known as the multiplicity of infection (MOI).

The standard transduction protocol used within the laboratory involves suspending 10^6 cells in 1 ml medium supplemented with 8 µg/ml polybrene. Virus is added and 500 µl of suspended cells pipetted per well of a 48 well plate. The plate is then centrifuged at 1500 x g for 2 hours at 32^{o}C. This protocol was initially modified by performing the spinfection in a 48 well plate, rather than 24 well. This modification brought the cells closer together, increasing the likelihood contact with virus. When compared to a previous experiment using the original 24 well protocol in the *MLL* rearranged cell line SEM (performed by Svetlana Myssina, post-doctoral researcher), the reduction in surface area was accompanied by an increase in the relative transduction efficiency from 14% to approximately 50% (Table 5-1).

Additional factors were subsequently examined. These were: 1) reduction in polybrene concentration from 8 µg/ml to 4 µg/ml; 2) use of protamine 5 µg/ml instead of polybrene; 3) use of serum free Optimem medium without polybrene/protamine; 4) reduction in transduction time from 18 hours to 6 hours to limit toxicity of exposure to polybrene.

The relative transduction efficiency was also determined for additional lymphoblastic leukaemic cell lines – 697 and REH, and t(8;21) AML cell lines Kasumi-1 and SKNO-1. The results of the six test transductions performed to assess these variables are presented in Table 5-1.

In parallel to assessments of transduction efficiency, cell proliferation and viability were assessed during transductions 2-5. This demonstrated a proliferative advantage to those cells transduced in the presence of protamine as compared with polybrene (4 or 8 µg/ml), but due to the lower transduction efficiency this did not result in a greater number of viable transduced cells.

Transductions 4-6 also investigated scaling up the transduction to a 6 well-plate format. For this transduction, 5×10^6 cells were resuspended in 2 ml medium and pipetted into a 6 well plate. A modification was made to the spinfection protocol, with a 20 minute, 50 x g spin performed to settle the cells evenly onto the plate surface, followed immediately by increasing the speed to the full rate of 1500 x g. This prevented the cells layering along one edge of the well, as was seen to happen if the centrifuge was set initially to 1500 x g. During these three test transductions, each of which was conducted with SEM cells and polybrene 8 µg/ml, the relative transduction efficiency was 81%, 101% and 54%. The third experiment (and the whole of transduction 6) was conducted with a new stock of SEM cells, bought fresh from DSMZ (Braunschweig, Germany), as these were the cells to be used for the final screen experiment.

The multiple test transductions demonstrated that no alternative method offers an advantage over the standard protocol of polybrene 8 µg/ml, incubated in the presence of virus overnight. Experiments to scale up the transduction produced similar outcomes to the initial transductions, although there was still a substantial variability between transductions of 54-101% relative transduction efficiency. Therefore, this scaled up standard protocol was used for the trial screen described in 5.3.2.

Table 5-1. Relative transduction efficiency determined in each of six test transductions using GIPZ non-target control virus particles. Relative transduction efficiency is calculated relative to efficiency in HEK293T cells (viral titre data provided by manufacturer). Transduction 6 used new SEM cells purchased from DSMZ.

		Maximum relative transduction efficiency (%)			
Transduction	**Cell line**	**Polybrene 8 µg/ml**	**Polybrene 4 µg/ml**	**Protamine 5 µg/ml**	**Serum free**
1	SEM	91			54
2	SEM	55			
	697	120			
	REH	120			
	Kasumi-1	8.2			
3	SEM	44		15.5	6.5
	697	7.3		3.7	2.8
	REH	29		8.5	1.9
	Kasumi-1	7.0		1.5	-
	SKNO-1	9.0		2.8	-
4	SEM	53		21	8.5
	REH	42		14	4.8
5	SEM -6 hrs	105	78	33	
	SEM -18 hrs	133	100	68	
6	SEM	62	45		

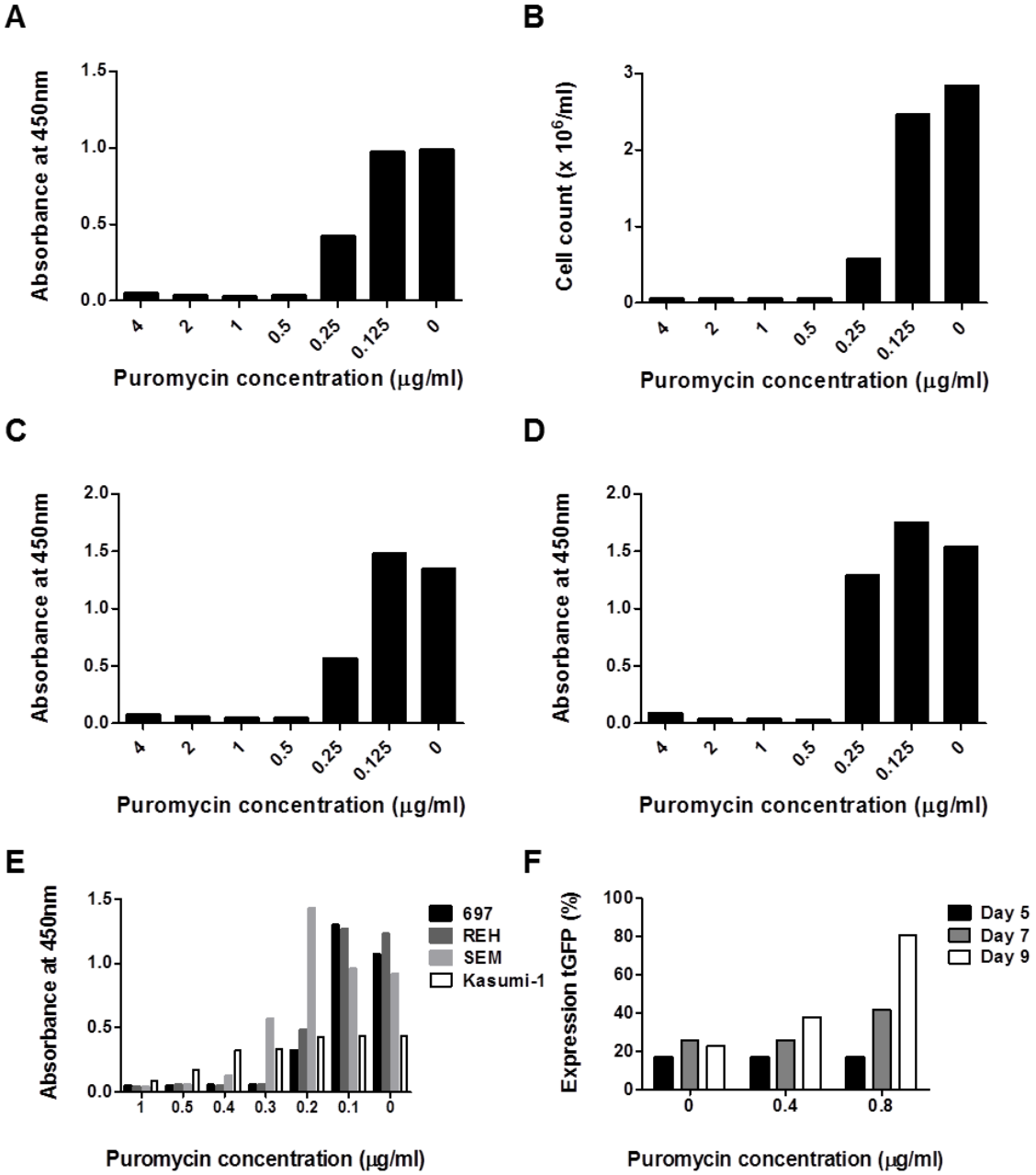
5.3.1.2 Determining puromycin toxicity

Following transduction targeting an MOI of 0.3, transduced cells were to be selected by a period of culture in the presence of puromycin. In order to determine the minimum concentration of puromycin effective for selection, two experiments cultured untransduced cells in the presence of increasing concentrations of puromycin from 0 – 4 µg/ml for up to 14 days. Cells were monitored by performing cell counts and by using a WST-1 assay as a measure of metabolic activity.

In the first experiment, both cell numbers and metabolic activity were affected by day 4, with cells grown in 0.5 µg/ml puromycin showing a substantial drop in numbers and loss of metabolic activity (Figure 5-2A-B). This threshold did not change up to day 10 or 14 (Figure 5-2C-D). As there was a substantial difference between 0.25 µg/ml and 0.5 µg/ml, a second experiment was conducted with further concentrations tested between 0 µg/ml and 0.5 µg/ml. In this second experiment, three other cell lines were also tested, namely REH, 697 and Kasumi-1. This experiment confirmed that the minimum concentration required to negatively select untransduced SEM cells was 0.4 µg/ml (Figure 5-2E). The results varied in other cell lines, with both REH and 697 cells showing a clear threshold of 0.3 µg/ml. Kasumi-1 cells gave a less clear result, with gradually increasing toxicity observed between 0 – 2 µg/ml.

Initially 0.4 µg/ml puromycin was used to positively select transduced SEM-GIPZ cells. However, during the course of the test transductions described in 5.3.1.1, trial selections demonstrated that 0.8 µg/ml puromycin could be used without excessive cell death and produced a more rapid positive selection (Figure 5-2F). This became the standard concentration for positive selection of SEM-GIPZ cells.

Figure 5-2. Assessment of puromycin toxicity. A) Metabolic activity (assessed by WST-1 assay) and B) cell counts in SEM cells at day 4 of puromycin treatment. Metabolic activity in SEM cells at C) day 10 and D) day 14. E) Metabolic activity at day 6 of puromycin treatment in 697, REH, SEM and Kasumi-1 cells. F) Positive selection of SEM-GIPZ cells with different concentrations of puromycin.



5.3.1.3 M2-10B4 feeder cell culture and growth arrest

M2-10B4 murine bone marrow stromal cells were grown to confluence in 175 cm^2 dishes/flasks before being divided 1:5 every 3-4 days. As they grow rapidly, two different approaches were taken to arrest their growth. Initially, mitomycin C was used at concentrations between 2 - 20 µg/ml. Cells were exposed for 2 ½ hours before being washed and fresh medium added. However, growth arrest was not achieved using this protocol. Using an alternative approach, M2-10B4 cells were irradiated with 80 Gy as described in 2.3.2.1 and seeded into fresh culture dishes/flasks at 5 x 10^6 cells/ml. Using this approach, good growth arrest was achieved with very little cell death or formation of proliferative foci. All experiments assessing feeder cell co-culture used these irradiated feeder cells.

5.3.1.4 Leukaemic – feeder cell interactions

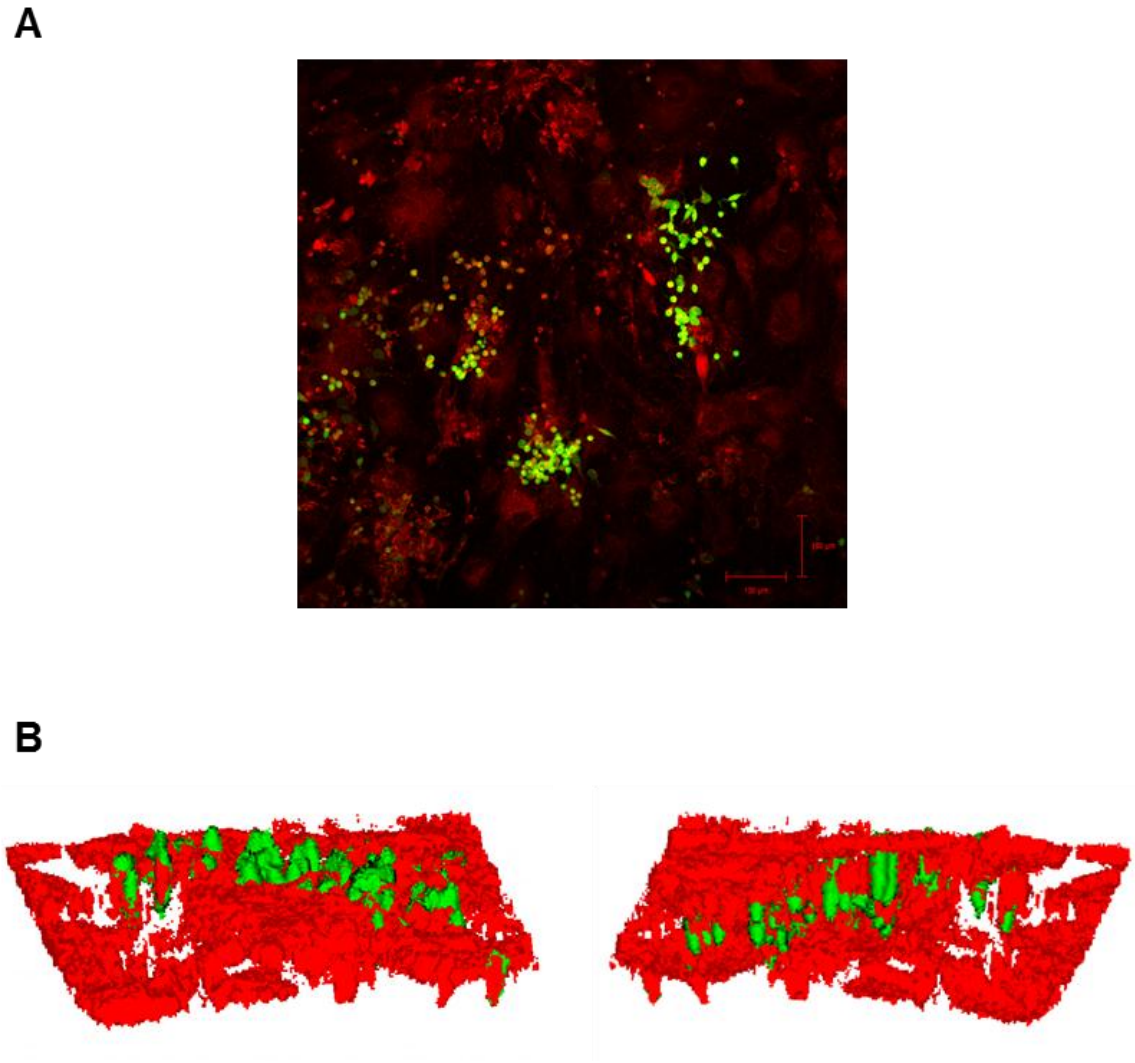
In order to provide the greatest chance of identifying genes involved in leukaemic propagation *in vivo*, the standard suspension culture/outgrowth assay used in genome-wide RNAi screening was modified to involve a re-plating assay. Lymphoblastic leukaemic cell lines do not form colonies in standard colony formation assays, so co-culture on a bone marrow stromal cell derived feeder layer, M2-10B4, was investigated. This offered not only the potential for an assay of colony formation from low cell density, but also provided the potential for leukaemic-niche interaction.

Initially, the relevance of the feeder layer was investigated. It was assumed that SEM cells seeded at a very low density of 10^4 cells/ml (standard density 5 x 10^5 cells/ml), 50 ml per 175 cm^2 dish/flask, would not proliferate. However, when seeded onto a sub-confluent layer of irradiated feeder cells (5 dishes) in RPMI 1640, 10% v/v FCS, GIPZ transduced SEM cells expanded from 2.5 x 10^6 cells to 1.7 x 10^8 cells, approximately 6 doublings, in 7 days. In a second experiment, serial re-plating showed almost 4 doublings in 5 days, followed by more than 4 doublings in 4 days of the second plating and more than 2 doublings in 3 days of the third replating. During these replating steps Turbo GFP (tGFP) expression was also analysed. Having been 88% prior to the first plating, tGFP expression dropped to 78% at the first harvest. After the second

plating, tGFP remained stable at 78%, before reducing further to 66% after a third re-plating step. As expression of shRNA and tGFP are driven by the same promoter, a drop in tGFP expression raises concerns over the expression of shRNA. Maintaining the expression of tGFP by on-going puromycin selection pressure was therefore examined further and is described in 5.3.1.6.

To look at the physical interactions of SEM and M2-10B4 cells, irradiated M2-10B4 cells were seeded into a 7 cm^2 glass culture dish to give a sub-confluent layer. Once adherent, feeder cells were counter stained with Rhodamine B, 20 ng/ml, for 30 minutes. A total of 3 x 10^4 SEM-GIPZ cells were seeded onto the feeder cells and cultured for 6 days before being viewed using a confocal microscope. This analysis demonstrated the growth of SEM-GIPZ cells in discrete colonies, supporting the use of this assay as a modified colony formation assay (Figure 5-3A). Furthermore, 3-dimensional images, reconstructed by stacking multiple planar images taken along the Z-axis, demonstrated that SEM-GIPZ cells were growing down between the feeder cells, not simply growing adjacent to them (Figure 5-3B). This is consistent with presence of physical leukaemic-niche interaction.

Figure 5-3. Confocal imaging of SEM-GIPZ cells growing in colonies on M2-10B4 feeder cells. A) Planar image of SEM-GIPZ cells, expressing tGFP (green), growing in discrete colonies on M2-10B4 feeder cells, counterstained with Rhodamine B (red). Scales show 100 µm. B) Z-stacked 3-dimensional images showing SEM-GIPZ cells growing down into M2-10B4 feeder cells. Left image is surface view. Right image shows same z-stack rotated 180^{o} about the Z axis so that the underside of the feeder monolayer (red) is now seen. This demonstrates SEM-GIPZ cells growing down through the feeder cell layer.

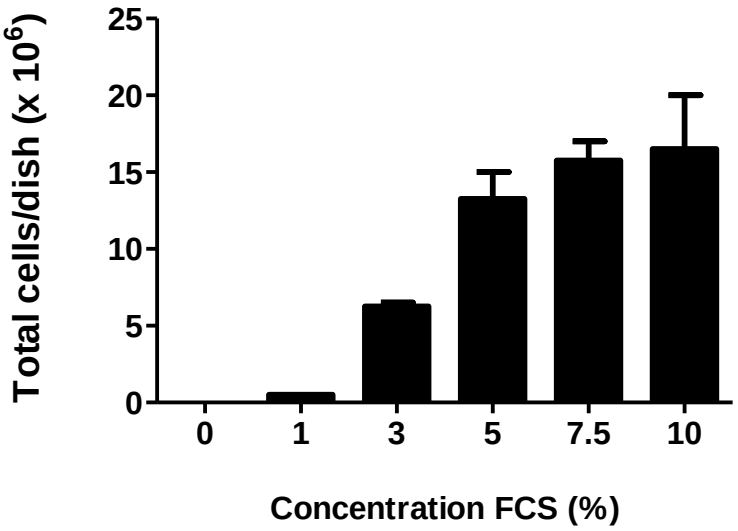


During validation of the assay, an experiment was performed to demonstrate that SEM cells would not grow when seeded at 10^4 cells/ml in the absence of feeder cells. In the first experiment, SEM-GIPZ cells were plated out in standard medium (RPMI 1640, 10% v/v FCS) at 10^4 cells/ml, 50 ml in a 175 cm^2 culture dish. Unexpectedly the cells proliferated well, undergoing more than 5 doublings in 7 days, a rate consistent with standard culture conditions.

It was hypothesised that decreasing the concentration of FCS in the culture medium would make the leukaemic cells reliant on niche support. Two independent experiments were performed to identify the concentration of FCS required to support the growth of SEM cells, at low seeding density (10^4 cells/ml in 50 ml) in the absence of feeder cells. Cells were plated out in 175 cm^2 dishes with concentrations of FCS from 0-10% v/v and cell counts performed 7 days later. These two experiments confirmed that, in the presence of 10% v/v FCS, SEM cells were able to undergo more than 5 doublings in 7 days (Figure 5-4). The growth rate decreased as the concentration of FCS decreased, until no growth was seen with 1% v/v FCS.

One further experiment seeded SEM-GIPZ cells at 10^4 cells/ml, 50 ml per dish in medium with 1% v/v FCS either with or without feeder cells. This again showed no growth in the cells grown in the absence of feeder cells, whilst those cells grown on feeder cells underwent approximately 4 doublings in 7 days (final cell count 7.8 x 10^6 cells in total). It was therefore decided that during the screen, SEM-GIPZ cells would be re-plated with irradiated M2-10B4 feeder cells (5 x 10^6 cells/175cm^2) in RPMI 1640, 1% v/v FCS.

Figure 5-4. Proliferation of SEM cells at low seeding density with decreasing concentrations of fetal calf serum. SEM cells proliferated normally in 7.5-10% FCS but failed to proliferate in 1% or 0%. Graph shows mean of two independent experiments with bars indicating standard error of the mean.



5.3.1.5 Expression of tGFP during culture on feeder cells

Having established a standard culture condition for SEM-GIPZ cells on irradiated M2-10B4 feeder cells (10^4 cells/ml RPMI1640, 1% v/v FCS, 50 ml per 175 cm^2 culture dish/flask), the maintenance of tGFP expression was examined. This followed an early re-plating experiment, conducted with 10% v/v FCS, in which expression of tGFP was seen to reduce from 88% prior to plating to 78%, 78% and then 66% following each of 3 serial re-plating steps, described in 5.3.1.4. As tGFP and the shRNAmir construct are expressed from the same promoter, a reduction in expression of tGFP is likely to be accompanied by a reduction in expression of shRNA. An attempt was therefore made to co-culture SEM-GIPZ cells on M2-10B4 feeder cells in the presence of puromycin.

SEM-GIPZ cells were seeded onto two plates of feeder cells using the standard conditions described above. In addition, puromycin 1 µg/ml was included. Feeder cells were resuspended by trypsinisation on day 5 (one plate) and day 11 (second plate). Cells trypsinised on day 5 were reseeded in standard conditions with the addition of either 1 µg/ml or 0.4 µg/ml puromycin and re-trypsinised 6 days later. For the first plating, analysis at day 5 showed a total of 7.5×10^5 cells (less than one doubling), with expression of tGFP >98%. Analysis of the second plate at day 11 showed 1.4×10^6 cells (less than 2 doublings). Expression of tGFP remained high, >99%. For the SEM-GIPZ cells re-plated after 5 days on feeder cells, analysis on day 6 showed that with 1 µg/ml puromycin no increase in cell numbers had occurred, whilst with 0.4 µg/ml puromycin the total cell count was 1.25×10^6 cells (less than 2 doublings). Expression of tGFP was >98% in both samples.

These experiments demonstrated that tGFP expression can be maintained during feeder cell co-culture by the addition of 0.4 µg/ml puromycin, but even at this concentration, the growth of SEM-GIPZ cells is substantially reduced as a result of feeder cell toxicity. M2-10B4 feeder cells were therefore stably transfected with pPAC, a plasmid encoding puromycin N-acetyltransferase.

5.3.1.6 Puromycin resistant feeder cells

In order to allow the on-going selective pressure of culture with puromycin during the co-culture assay, M2-10B4 feeder cells were stably transfected with pPAC. Initially, transfection was attempted using calcium phosphate precipitation, as described for transfection of 293T cells with lentiviral plasmids. This approach was not successful, with all M2-10B4 cells dying under puromycin selection on two occasions. An alternative approach was taken using Lipofectamine LTX. M2-10B4 cells were grown to sub-confluence in 6 well plates and then incubated with pPAC plasmid and varying concentrations of Lipofectamine LTX. Having allowed 72 hours for expression of the plasmid, puromycin was added at 1 µg/ml and then increased to 2 µg/ml the following day. A small number of surviving colonies were isolated and expanded under continued selection with 2 µg/ml puromycin.

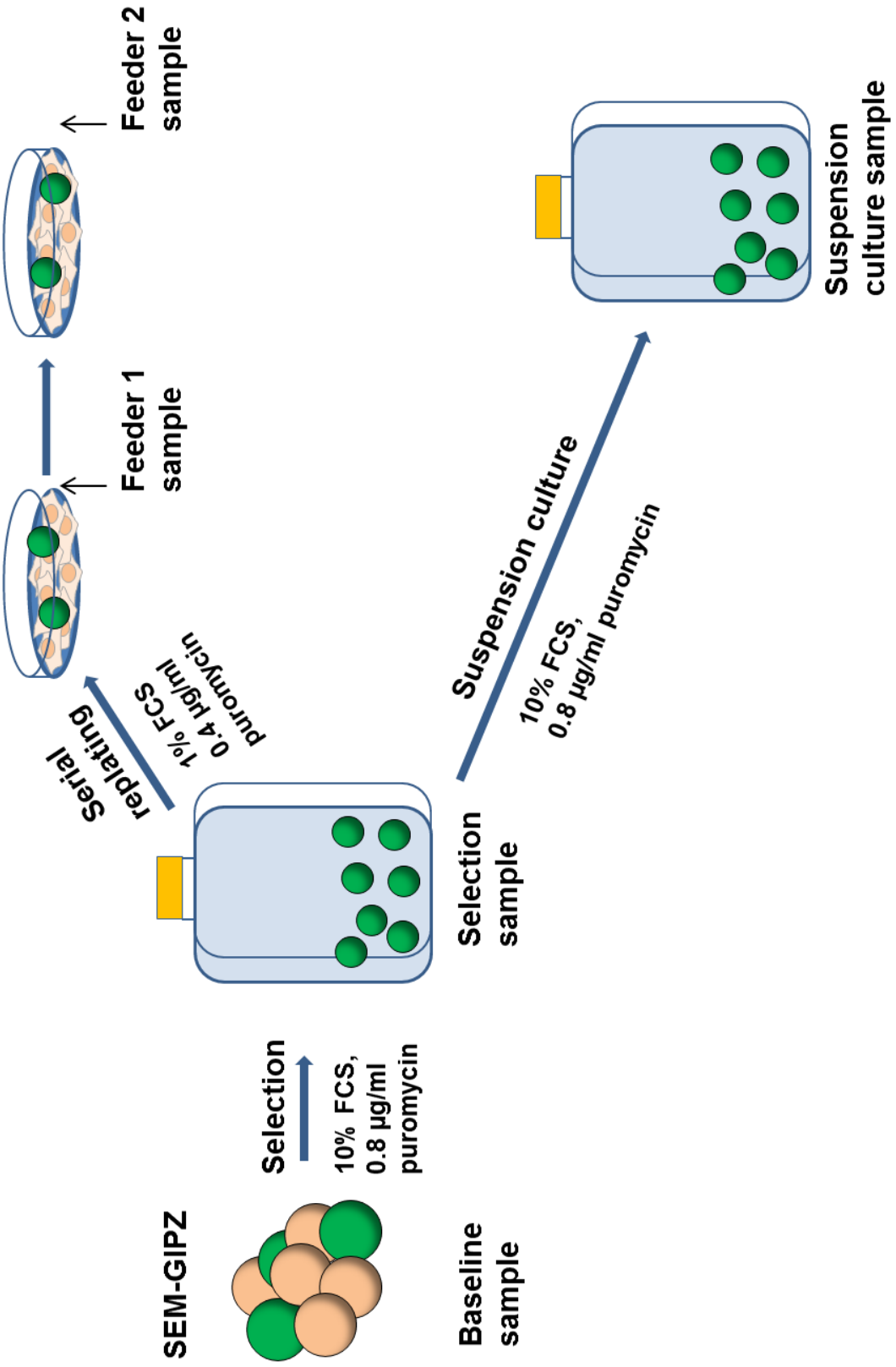
To test whether M2-10B4-PAC cells were able to support the growth of SEM-GIPZ cells with on-going puromycin selection pressure, 5×10^5 SEM-GIPZ cells were seeded onto irradiated M2-10B4-PAC feeder cells in 50 ml RPMI 1640, 1% v/v FCS and 0.4 µg/ml puromycin. After 4 days the feeder cells were trypsinised, counted and analysed by flow cytometry. This confirmed that SEM-GIPZ cells had expanded 3 fold and that expression of tGFP was maintained at 95%.

5.3.2 First trial screen

Through the protocol optimisation described in 5.3.1 an experimental design was developed which would allow the screening of individual pools of lentiviral constructs in leukaemic cell lines. Initial transduction would be followed by a short period of positive selection to achieve >85% expression of tGFP. Transduced leukaemic cells would then be serially re-plated on bone marrow stromal cells to recreate leukaemic propagation in a bone marrow niche micro environment.

This protocol, however, represents a substantially more complex experiment than the standard outgrowth assay used most frequently in RNAi screens. As such, it was important to demonstrate both the relevance and feasibility of this approach by performing a trial screen (Figure 5-5). Having completed this, a replicate trial was performed using the same virus pool to assess the reproducibility of the protocol. In order to determine the number of re-plating steps necessary, trial samples were taken at baseline (required to provide a comparator), following puromycin selection (to achieve >85% selection), following a period of suspension culture equivalent to 11-12 doublings in total and following the first and second re-plating steps. Comparison of the results derived from each of these time points would allow determination of the final experimental protocol to be used for the complete screen. Trial screens were performed using pool 7 from the decode library as this pool had the highest viral titre.

Figure 5-5. Experimental design for trial screens. Transduced cells undergo positive selection with puromycin prior to being split into two parallel arms. The first arm is plated twice at low density onto puromycin resistant bone marrow stromal feeder cells. The second arm is maintained using a "standard" suspension culture approach.



5.3.2.1 Transduction

Transductions were performed in 6 well plates, 5 x 10^6 cells per well in 2 ml medium. The standard transduction medium was used (RPMI 1640, 10% v/v FCS, polybrene 8 µg/ml) and standard spinfection protocol (1500 x g, 2 hours, 32^{o}C). For the first trial experiment, an estimate of the number of cells required was made as follows: the desired coverage for the 10,000 construct pool was 1000 fold (this is equivalent to 95% probability of achieving a 100 fold coverage, see section 2.3.10.3); the desired MOI was 0.3; cell survival following transduction would be 90% (as seen in test transductions using 6 well plate).

$$\frac{1000}{0.3\ x\ 0.9}\ x\ 10{,}000 = 3.7\ x\ 10^7$$

If the relative transduction efficiency was 90% then this transduction would require 12.9 µl virus from pool 7 (9.56 x 10^8 TU/ml). In order to use the entire 25 µl vial of virus, the number of cells was scaled up to 7.5 x 10^7 cells. If the actual relative transduction efficiency was closer to the lower estimate generated in 5.3.1.1, approximately 50%, the total number of transduced cells would be 1.2 x 10^7, still sufficient to achieve a 1000 fold coverage.

The day following transduction, cells were resuspended, pooled into two aliquots (14 ml) and centrifuged 300 x g, 10 minutes. Seven millilitres of supernatant was removed from each aliquot to reduce the polybrene concentration and then cells resuspended with addition of a further 30ml medium to give a final cell concentration of 10^6 cells/ml and pooled. Resuspended cells were divided into one of two 75 cm^2 flasks.

5.3.2.2 Analysis of transduction

Analysis of tGFP expression on day 4 following transduction demonstrated only 7.7% of cells expressing tGFP (Figure 5-6A). From a starting number of 75 x 10^6 cells, this equated to a total number of 5.8 x 10^6 transduced cells, although an increase in the cell number meant that by day 4 there were 190 x 10^6 total cells and 14.6 x 10^6 transduced cells. With such a low proportion of cells containing an integrated construct, the baseline DNA sample would significantly under-represent integrations, failing to maintain the fold coverage into the array

analysis. Transduced cells were therefore enriched by overnight culture in puromycin.

5.3.2.3 Density centrifugation of viable cells

Selection with 1 µg/ml puromycin was performed overnight using one third of the total number of cells in culture – 4.7 x 10^6 transduced cells. The following day, day 5, expression of tGFP was 17% of viable cells (Figure 5-6A). Viable cells were enriched by density gradient centrifugation, from which a total of 28 x 10^6 cells were retrieved to provide the baseline sample. Expression of tGFP in the enriched population was stable at 15%. This equated to 4.2 x 10^6 tGFP positive cells retrieved, 450 fold coverage. On-going culture was performed with 0.8 µg/ml puromycin.

5.3.2.4 Monitoring of positive selection

The expression of tGFP was monitored by flow cytometry throughout the period of initial selection. This showed an increase in expression to >95% in the first 7 days of selection (day 11 post transduction) (Figure 5-6A). However, throughout this period there was no increase in viable cell numbers (Figure 5-6B). Large numbers of apoptotic cells were present. On day 14 following transduction a further enrichment of viable cells was performed using density gradient centrifugation. This recovered a total of 4 x 10^6 cells. Following this enrichment SEM-GIPZ cells again proliferated, doubling approximately 3.7 times between days 14 and 23. This gave sufficient leukaemic cells for seeding onto feeder cells (5 x 10^6 cells), continuing suspension culture (32 x 10^6 cells) and harvesting for DNA (10^7 cells). Cells for DNA were washed with PBS and stored as a cell pellet at -20^{o}C.

Cells maintained in suspension culture (puromycin 0.8 µg/ml) were passaged three times per week until day 35 when they were harvested for DNA. Cells harvested for DNA were washed with PBS and stored as a cell pellet at -20^{o}C.

5.3.2.5 Re-plating steps

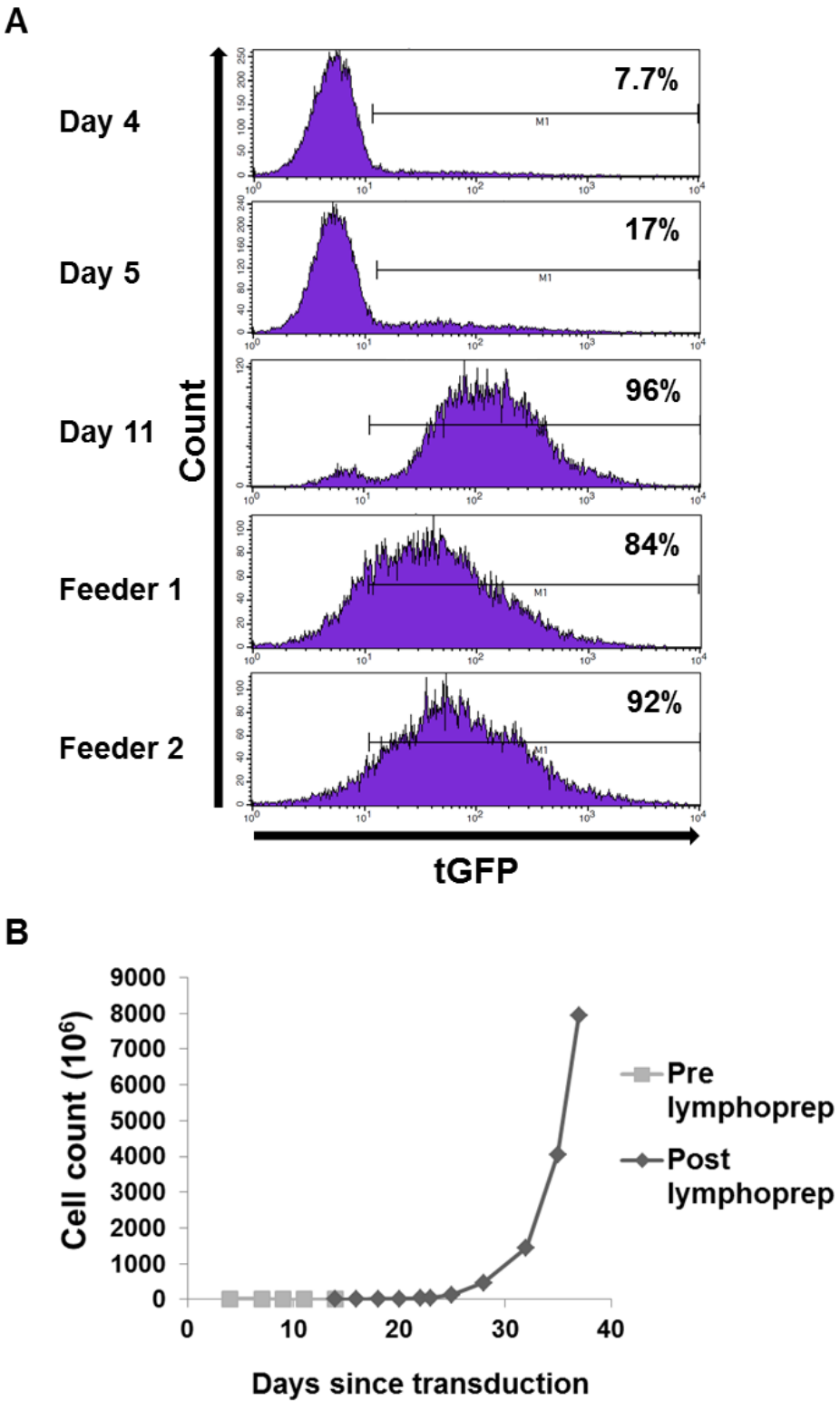
SEM-GIPZ cells were seeded onto 10 plates of M2-10B4 feeder cells at a concentration of 10^4 cells/ml, 50 ml per plate. The medium used was RPMI 1640 with 1% v/v FCS and 0.4 µg/ml puromycin. Cells were harvested by trypsinisation after 7 days. At this time, 5 of the plates, all seeded with the same batch of feeder cells, showed a marked difference to the other 5 plates which had received a different batch of feeder cells. The total cell count in the first 5 plates was 21.8 x 10^6 cells whilst in the second 5 plates the count was just 5 x 10^6 cells. The decision was taken to carry forward only cells from the first 5 plates, which had doubled 3 times since seeding. The expression of tGFP in cells from the selected dishes had dropped slightly to 84% (Figure 5-6A).

SEM-GIPZ cells were either harvested for DNA (5 x 10^6 cells), re-seeded onto a further 11 plates of feeder cells (5.5 x 10^6 cells) or stored as back-up in liquid nitrogen (11.5 x 10^6 cells). The second plating of SEM-GIPZ cells was harvested after 7 days. Cells were recovered both from the supernatant (1.1 x 10^6 cells) and following trypsinisation of adherent SEM-GIPZ/feeder cells (2.2 x 10^6 cells). A total yield of only 3.3 x 10^6 cells demonstrated almost no expansion in cell numbers over the period of co-culture.

The feeder cells used in the second plating had, unlike the first plating, been irradiated and then stored ready for use in liquid nitrogen. When seeded out prior to the plating, a substantial number of these cells had failed to adhere to the dish and the morphology of those which did was unusual, being elongated and spindle-like. A repeat re-plating was therefore performed using feeder cells which were irradiated and then seeded out immediately into culture dishes. SEM-GIPZ cells stored in liquid nitrogen after the first plating were thawed and cultured in standard medium with 0.4 µg/ml puromycin. After 72 hours SEM-GIPZ cells were seeded out onto 8 dishes of fresh feeder cells and cultured for 7 days. Having observed a drop in expression of tGFP over the first plating, the concentration of puromycin was increased to 0.6 µg/ml for the second plating. Four of these dishes developed fungal infection, leaving four dishes to be harvested (total cell number seeded out 2 x 10^6). Harvesting these plates yielded 13.2 x 10^6 cells, representing between 2 and 3 doublings. Of these 6 x

10^6 cells were enriched by CD133 MACS sorting, which in turn yielded 2 x 10^6 cells for immediate isolation of DNA. Analysis of this sorted population showed that with the higher concentration of puromycin, expression of tGFP was maintained at 92% (Figure 5-6A).

Figure 5-6. Selection and growth of SEM-GIPZ cells – first trial RNAi screen. A) Expression of tGFP rises rapidly following overnight puromycin selection and density gradient centrifugation (day 4 to day 5) and then continues to rise with on-going puromycin selection. There is a drop in tGFP expression following the first plating on feeder cells (Feeder 1, puromycin 0.4 µg/ml) but the second plating results in a return to higher expression (Feeder 2, puromycin 0.6 µg/ml). B) Adjusted transduced cell numbers during suspension culture in the presence of puromycin 0.8 µg/ml. No growth is seen prior to eradication of apoptotic cells. Growth afterwards shows an appropriate doubling time of approximately 50 hours.



5.3.3 Second trial screen

The initial trial screen had been subject to a number of difficulties. These included:

1. A low rate of transduction
2. Poor growth of transduced cells during early positive selection
3. Fungal infection of feeder cells
4. Drop in tGFP expression during first plating on feeder cells
5. Poor attachment of feeder cells which had been stored in liquid nitrogen
6. Poor growth on feeder cells

Whilst the second trial was intended to assess the reproducibility of the protocol, a number of modifications were made to address the problems encountered during the first trial.

1. Fewer SEM cells were used to reduce the number of untransduced cells subject to puromycin selection (assumed relative transduction efficiency of 0.6, therefore used 4.7×10^7 cells)
2. Second density gradient centrifugation was performed early (day 8 post transduction) to reduce the period without leukaemic cell growth
3. 175 cm^2 flasks used instead of culture dishes to reduce airborne contamination by fungus
4. Puromycin in plating assays used at 0.8 µg/ml to maintain tGFP expression
5. Feeder cells irradiated and seeded straight into flasks without freezing

5.3.3.1 Summary of results of second trial

The expression of tGFP on day 4 following transduction was 9.8% (Figure 5-7), giving a total of 4.7 x 10^6 transduced cells. Puromycin was again applied at 1 µg/ml overnight and density gradient centrifugation performed to enrich for viable cells. The expression of tGFP following enrichment on day 5 was 26% (Figure 5-7). A total of 5.4 x 10^6 transduced cells were recovered from the enrichment to provide the baseline sample. By day 8 post transduction (day 4 of positive selection), there had been no increase in viable cell numbers. A second density gradient centrifugation enrichment was performed on day 8, returning 6.6 x 10^6 viable cells with 96% expressing tGFP (Figure 5-7).

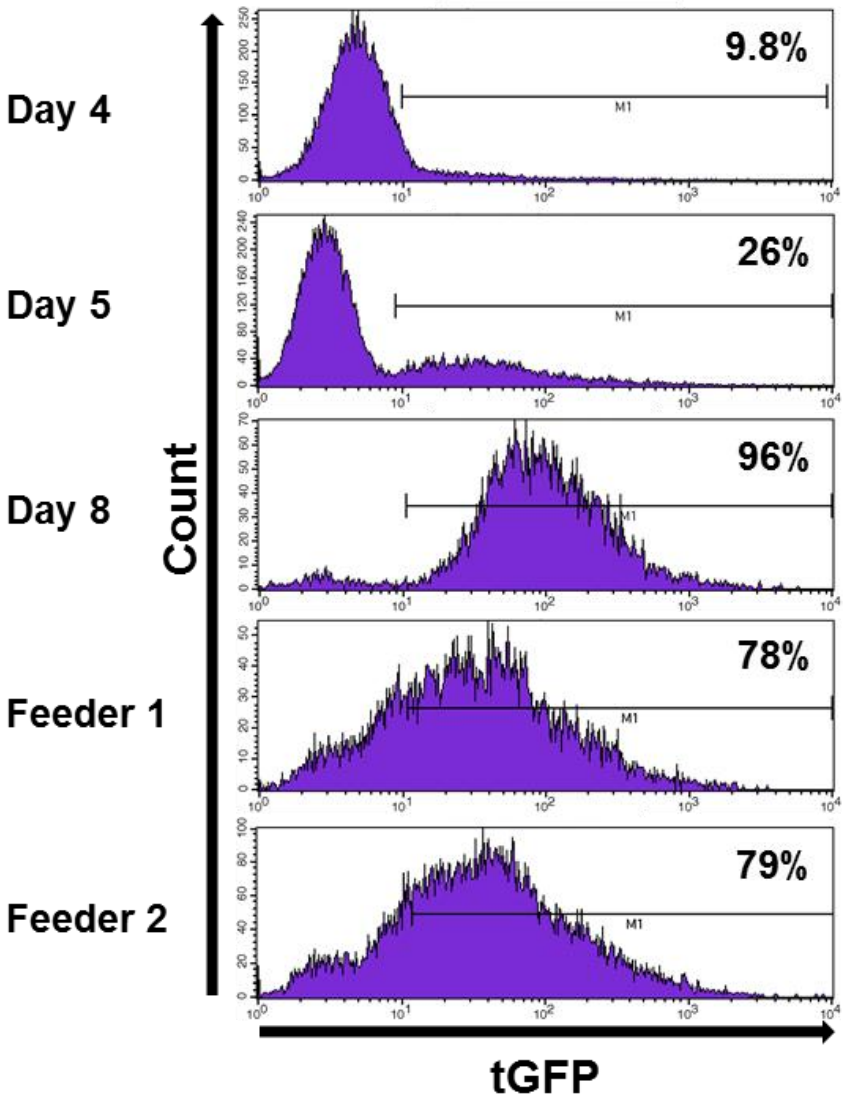
Following the earlier enrichment, cell growth was rapid with 3.7 doublings (equivalent to that seen between days 14 and 23 during trial screen 1) being completed by day 14. Transduced cells were seeded out at this point into 10 flasks of feeder cells (5 x 10^5 cells per flask at a concentration of 10^4 cells/ml, 50 ml per flask). The medium used was RPMI 1640 with 1% v/v FCS and 0.8 µg/ml puromycin. Cells were harvested after 7 days, at which time they had undergone exactly 3 doublings (as during plating 1 of trial screen 1). SEM cells were purified by CD133 MACS sorting. Expression of tGFP in these sorted cells was 78% (Figure 5-7). Samples were taken for DNA and back-up cells frozen and stored in liquid nitrogen.

A second plating was performed into 10 flasks of M2-10B4 feeder cells, as for the first plating. After 7 days the cells did not look as densely grown as after the first plating. Flasks were trypsinised, giving a total of 17 x 10^6 cells, equivalent to just 1.8 doublings. SEM cells were isolated from feeder cells by CD133 MACS sorting. However, only 10^6 cells were recovered from this sort. The second plating was therefore repeated using SEM cells frozen after trypsinisation of the first plating. SEM were plated into 10 flasks of freshly irradiated feeder cells as described for the first plating. After seven days the cell density did not look sufficiently high so half of the medium was aspirated off, the cellular content centrifuged out at 300 x g, 10 minutes and then resuspended in an equal volume of fresh complete medium which was returned to the flasks. Flasks were then harvested after a total of 10 days of culture, at which point a total of 4 x 10^7 SEM were recovered, representing 3 doublings,

similar to that achieved for the second plating in the first trial screen. Expression of tGFP in recovered SEM cells was 79% (Figure 5-7).

The suspension culture, maintained in parallel to the two re-plating steps was harvested on day 24 after 10.7 doublings, equivalent to the 10 doublings achieved in the first trial screen.

Figure 5-7. Selection and growth of SEM-GIPZ cells – second trial RNAi screen.
Expression of tGFP rises rapidly following overnight puromycin selection and density gradient centrifugation (day 4 to day 5) and then continues to rise with on-going puromycin selection. There is a drop in tGFP expression following the first plating on feeder cells but this level of expression is maintained through the second plating (both platings with puromycin at 0.8 µg/ml).



5.3.4 Sample processing and data analysis

5.3.4.1 Amplification and purification of barcodes for pool deconvolution

DNA was extracted from freshly harvested or frozen SEM-GIPZ cells stored as cell pellets at -20^{o}C. Viral barcode sequences were amplified using the negative selection primers contained within the Decode library kit. Between 4 and 22 individual PCR reactions, each containing 850 ng of DNA, were performed for samples taken at each time point. The variation in number of reactions performed results from the varying proportions of transduced cells in the population, that is varying quantities of template, as well as the varying requirements for the arrays - 6 µg of DNA for the baseline sample which provided the comparator for the other time points, each of which required 1.5 µg. The PCR products were analysed by gel electrophoresis to ensure an appropriate 250 bp product was present in each reaction. Reactions were then pooled and purified by gel electrophoresis. DNA concentration was analysed by spectrophotometry before being sent to the Functional Genomics and Proteomics Facility, Birmingham University, who performed the labelling, hybridisation and scanning of the custom microarrays provided with the Decode library kit.

5.3.4.2 Data analysis

Data analysis was performed commercially through the Functional Genomics and Proteomics Facility, Birmingham University. For the first trial screen this analysis demonstrated a low signal intensity and high signal to noise ratio, presumed to result from the use of only one of seven viral pools, and therefore only 1 in 7 array probes being hybridised. The distribution of Cy3 and Cy5 signals was appropriate (Figure 5-8A). For the second screen, the signal to noise ratio was again high. However, in addition three of the four arrays demonstrated a technical problem, with the hybridisation giving an uneven distribution of signal (Figure 5-8B). The second trial screen has therefore not been included in further analysis and the amplification and array analysis will be repeated.

The baseline sample was competitively hybridised with each of the other four time points – day 23, day 35, replating 1 and replating 2, allowing comparison between these four later time points. Using day 23 as an early point during stable culture, comparison was made to day 35 (10 doublings) as well as to feeder cell replating 1 (F1) and feeder cell replating 2 (F2). Hairpins showing greater than a 2.5 fold change in abundance were identified (Figure 5-9). This analysis identified: 24 constructs showing >2.5 fold decrease and 38 showing >2.5 fold increase between day 23 and day 35; 58 constructs showing >2.5 fold decrease and 42 showing a >2.5 fold increase between day 23 and F1; 18 constructs showing >2.5 fold decrease and 9 showing a >2.5 fold increase between day 23 and F2.

Figure 5-8. Example spatial distributions of positive (red) and negative (green) log ratios. A) First trial screen. B) Second trial screen.

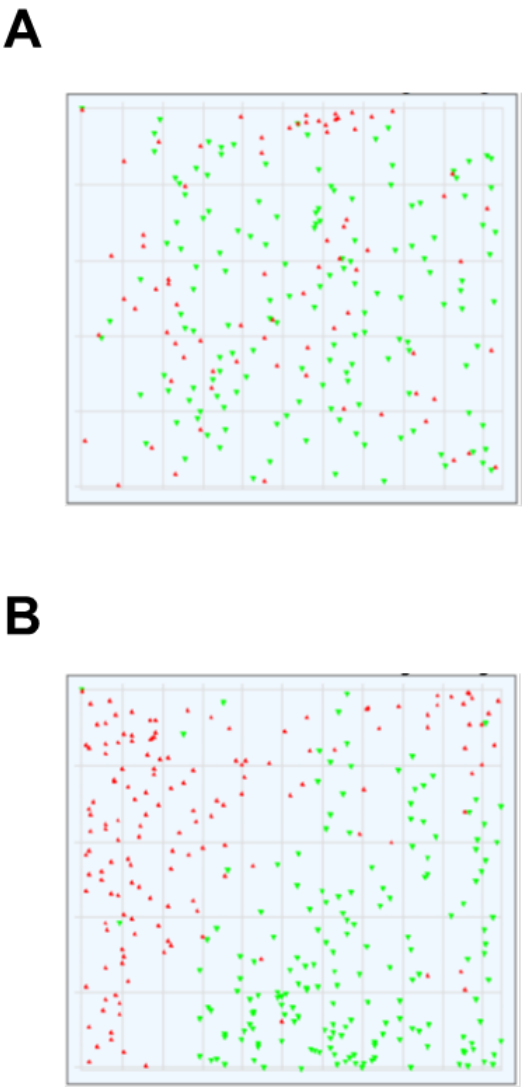


Figure 5-9. Candidate constructs showing >2.5 change in abundance between day 23 and subsequent time points. A) Candidates showing a >2.5 fold reduction in abundance. B) Candidates showing a >2.5 fold increase in abundance.

A

Day 23 v Day 35	
GRM3	-103.7
ANGPT1	**-9.9**
WASL	-7.6
SMYD5	-5.6
ITGAE	-5.5
KCNJ16	-5.4
KLHL17	-4.4
RPLP0	-4.2
AXIN2	-4.0
F2R	-3.6
WNT5A	-3.6
CLDN8	-3.1
SIPA1L2	-2.9
SEC23A	-2.9
UCN	-2.8
LIPG	-2.8
SERPINB	-2.8
SMG1	-2.7
VPS13B	-2.6
MMP3	-2.6
CNN2	-2.6
B4GALT1	-2.6
SLC2A13	-2.5
CRABP1	-2.5

Day 23 v F1	
GRM3	-107.9
CLDN8	-53.0
CD3G	-35.8
NEFH	-29.7
DUSP23	-26.3
ALG8	-16.9
TPST2	-14.9
ANGPT1	**-9.1**
WASL	-8.7
NRXN3	-7.8
THG1L	-6.8
EFCAB10	-6.4
SATB2	-6.3
AXIN2	-6.0
ITGAE	-5.4
SEC23A	-5.0
VPS13B	-4.9
GYPC	-4.8
KCNJ16	-4.7
UCN	-4.5
PDGFD	-4.3
KLHL17	-4.2
RNF5	-3.9
PDE6A	-3.7
SIPA1L2	-3.4
PIGR	-3.4
RPP30	-3.3
APBA3	-3.1
PDE8A	-3.1
AGXT2	-3.1
NEUROD2	-3.0
SLFN5	-3.0
SMYD5	-3.0
SLC2A13	-3.0
PLAC1L	-3.0
MMP3	-3.0
ETFB	-2.9
NEUROD2	-2.8
INHA	-2.8
YBX1	-2.8
RPLP0	-2.8
HSCB	-2.8
ZNF138	-2.7
POLR3A	-2.7
CD53	-2.7
EIF3E	-2.7
LTV1	-2.7
HSPH1	-2.7
GALNT5	-2.6
SOCS4	-2.6
SNRPA	-2.6
RBM12	-2.6
SMG1	-2.6
F2R	-2.5
GNG11	-2.5
GJD2	-2.5
EIF2AK2	-2.5
TUSC2	-2.5

Day 23 v F2	
GRM3	-82.3
NEFH	-43.6
DUSP23	-27.8
ALG8	-14.9
ANGPT1	**-8.9**
ITGAE	-4.8
KLHL17	-3.8
UCN	-3.7
KCNJ16	-3.5
PDGFD	-3.2
THG1L	-3.1
SMYD5	-3.1
PDE6A	-3.1
SMG1	-3.0
WNT5A	-2.8
F2R	-2.7
EIF2AK2	-2.7
SLFN5	-2.6

B

Day 23 v Day 35	
CYB5B	143.5
CA8	19.5
PIP	19.4
SNX27	17.6
SCG3	13.6
CMIP	12.8
COBLL1	11.9
APOBEC3G	9.1
ENPP3	8.6
GYS2	8.1
ANO5	7.5
PMS2L2	6.6
PSME3	6.5
ITSN2	6.3
VTI1B	6.3
ALDH1A1	6.2
PTS	6.2
COL1A2	5.9
LSM1	5.6
RAB3GAP2	5.4
HDAC9	5.2
PLA2G2A	5.1
XRCC5	4.5
PC	4.4
ADAMDEC1	4.3
COL5A3	4.1
SLC38A2	3.8
CAMKK2	3.1
OIT3	2.9
DNASE2B	2.8
BRP44	2.8
LSG1	2.8
ACTC1	2.7
TBRG4	2.7
PPIL6	2.7
TPCN1	2.6
ERBB3	2.6
BRAF	2.5

Day 23 v F1	
CYB5B	131.1
ZMYM5	56.4
PRRT1	34.1
SLC38A2	9.1
ENPP3	8.6
OIT3	8.2
PC	7.7
ATP6V1D	5.7
ACADL	5.0
YBX1	4.9
SHPRH	4.7
RCAN3	4.6
HEXIM1	4.3
ETFDH	4.3
FBXO27	4.2
ADAMTS18	3.6
DCK	3.5
PPP1R3C	3.4
TNIP3	3.3
CHST12	3.1
C3AR1	3.1
GABRG2	3.1
SHROOM4	3.0
EPHB1	3.0
ACAD10	3.0
CIDEC	2.8
TWF1	2.8
ITIH4	2.7
ZBTB10	2.7
MEIS2	2.7
HSF5	2.7
IMPACT	2.7
CCNA1	2.7
EPS15L1	2.7
GOLM1	2.7
SLC35E2	2.7
TLR7	2.7
DNALI1	2.6
ST3GAL3	2.6
P11	2.6
RNF186	2.5
SNX2	2.5

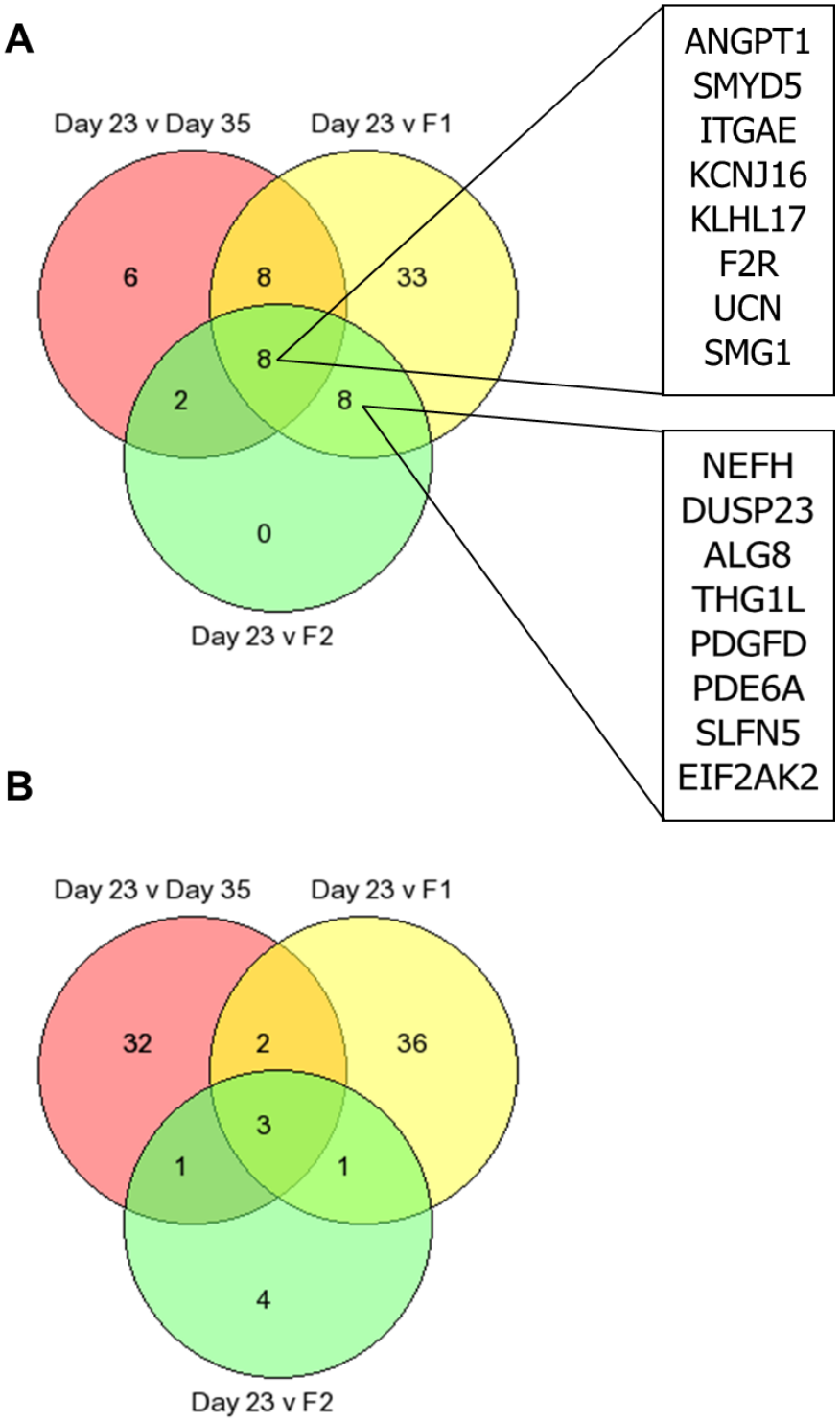
Day 23 v F2	
CYB5B	74.5
SNX27	8.6
ENPP3	8.6
OIT3	4.6
DEFB118	3.8
PIP4K2C	3.3
RCAN3	3.3
GSS	3.0
NUPR1	2.7

A substantial overlap was identified between the constructs demonstrating a change at each of these time points. In total, eight constructs decreased >2.5 fold in all of the three later time points (Figure 5-10A). A further eight constructs showed a >2.5 fold decrease in each of the feeder replating steps, but not the suspension culture (day 35) suggesting that these may represent genes critical for the colony formation assay/niche interaction provided by the feeder cell co-culture (Figure 5-10A).

In contrast, less commonality was seen in constructs demonstrating an increase in abundance. Only 3 constructs were common to all time points, with a single additional construct common only to the two replating steps (Figure 5-10B).

Amongst all of these candidates, one stood out as being of immediate interest, *ANGPT1* (Figure 5-9A). *ANGPT1* encodes the secreted protein Angiopoietin 1 which is the activating ligand for the vascular endothelial and haematopoietic progenitor cell surface receptor, TEK/Tie2. Previous work from our laboratory has identified ANGPT1 as being highly expressed in *MLL* rearranged leukaemias with down-regulation being seen in response to *MLL/AF4* knockdown both in the SEM cell line and patient-derived leukaemic blasts (GarridoCastro, Bomken et al.). This target was therefore carried forward to the *in vivo* analysis project described in Chapter 7 for further investigation.

Figure 5-10. Candidate genes identified by the first trial RNAi screen in SEM cells. A) Overlap of candidates demonstrating >2.5 fold decrease in abundance between day 23 and later time points. Those genes common to all later time points (top right panel) and those common to the two feeder replating steps are detailed (lower right panel). B) Overlap of candidates demonstrating >2.5 fold increase in abundance between day 23 and later time points.



5.4 Discussion

The optimisation of the screening protocol was intended not only to develop understanding of this new methodology within the Northern Institute for Cancer Research, but also to ensure that efficient transduction would be performed. This was important given the limited volume of virus within each library. Balanced against this was the risk of multiple integrations per cell, which was minimised by determining the viral titre required to achieve a multiplicity of infection of 0.3. Despite the optimisation performed, during both trial screens the rate of transduction was substantially lower than had been predicted. The reasons for this remain unclear. During the first trial screen a number of researchers within the NICR had cell lines which were growing very slowly or even dying. This problem was investigated at an Institute level and was felt most likely to be the result of a poor quality batch of fetal calf serum. This was replaced half way through the first trial screen, ahead of the repeat attempt of the second plating step. This would not, however, explain the poor, although slightly improved, rate of transduction in the second trial screen. Loss of viral titre in the pools used is one possibility, although investigating this would be difficult. It is possible that the storage of these vials over the preceding three years (somewhat longer than the suggested shelf life) resulted in a decrease in titre.

The other principle aim of the preparatory work during this project was to investigate the potential for, and relevance of, performing a co-culture assay. The aim of this approach was to provide leukaemic-niche interactions to maximise the chance of identifying targets relevant to the *in vivo* microenvironment. Unexpectedly, under standard culture conditions SEM cells seeded at low density (10^4 cells/ml) were able to grow in the absence of feeder cells. However, when the concentration of fetal calf serum was reduced from 10% to 1%, co-culture with feeder cells was required to maintain proliferation. In keeping with this finding, SEM-GIPZ cells seeded onto M2-10B4 feeder cells grew in distinct colonies with individual cells growing through and underneath the feeder cell layer. Whilst an absolute need for physical interaction between leukaemic and feeder cells has not been demonstrated here, these two pieces of evidence suggest a genuine role for this *in vitro* modelled bone marrow niche

environment in the present project. In addition to the eight constructs whose abundance showed >2.5 fold reduction at all final time points, a further subset of eight targets was identified as being important for cell survival/proliferation/maintenance during feeder cell co-culture in the first trial screen. These candidates will need to be validated by the repeat hybridisation of the second trial screen. The result of this second analysis will inform the decision on whether to complete the full screen using a co-culture assay based approach.

Despite the potential benefits of performing the full screen using re-plating steps on feeder cell layers, this approach does add a substantial degree of complexity to the experiment. Feeder cells need to be expanded and irradiated to be available at the right time for plating steps. Leukaemic cells need to be re-isolated from feeder cells at the end of a plating step. Leukaemic cell proliferation during re-plating has twice been lower than during the initial plating step, a situation complicated by not being able to accurately assess leukaemic cell numbers without trypsinising the feeder/leukaemic culture flasks. Finally the length of time required to perform the screen will vary substantially between the two approaches. In order to be able to respond to changes and difficulties experienced with the more complex re-plating approach, it is reasonable to expect that only a single pool can be assayed at any one time. In contrast, suspension culture could reasonably be performed in sequence, transducing the next pool as soon as the previous pool is in a period of exponential growth. It is therefore important to decide whether the potential benefits offered by the re-plating approach justify the delay in generating the data. Whilst the co-culture end-points provided eight candidate genes not identified by the suspension culture, there was also a subset of genes identified at all three time points. Amongst this subset was the single strongest candidate identified by the trial screen, *ANGPT1*. This gene has previously been identified within our laboratory as being highly expressed in *MLL* rearranged ALL, and is regulated by *MLL* fusion oncogenes (GarridoCastro, Bomken et al.). Whilst this provides some validation for the screen described in this project, it also suggests that key genes contributing to leukaemic development/maintenance can be identified using a standard suspension culture approach.

The final decision on how to conduct the screen will require the information gained from the repeat amplification and hybridisation of DNA obtained from the second screen. With these data it will be possible to make a statistical assessment of the strength of the candidates identified by the suspension culture versus the co-culture assay. If there is reasonable evidence that co-culture is likely to identify candidates important to leukaemic-niche interactions that would not be identified by suspension culture then it will be worthwhile investing the additional resources required to complete this more complex experiment. If not, the relative simplicity of the suspension culture assay will make it preferable.

However the screen is conducted, it will be critical to both identify and validate individual candidate targets. Validation will initially involve generating an shRNA library with a much lower complexity than the genome-wide library described here – perhaps just 1000 candidates. shRNAmir constructs specific for candidates identified by the genome-wide screen will be purchased and pooled. A more limited screen will then be performed either *in vitro* or *in vivo* using additional leukaemic cell lines. Candidates which are again identified during this screen will then be validated using patient-derived leukaemic blasts *in vivo*. To prepare for this final validation step, two further projects conducted within the present Fellowship aimed to develop techniques and technologies required for *in vivo* candidate analysis. Firstly, a protocol was developed for the lentiviral transduction of patient-derived leukaemic blasts. Secondly, a novel vector allowing the *in vitro* and *in vivo* tracking of leukaemia, combined with an shRNAmir construct to allow the modulation of candidate genes, was designed, cloned and validated. These two projects are described in Chapter 6 and Chapter 7.

Chapter 6

Development of a protocol for the lentiviral transduction of patient-derived leukaemic blasts

6.1 Introduction

Candidate elements of the leukaemia stem cell/propagation programme derived from the whole-genome RNAi screen described in Chapter 6 will initially be validated *in vitro*, as described in section 6.4. However, a subsequent validation step will involve modulation of novel candidates using patient-derived leukaemic blasts. Due to the limited viability of patient-derived blasts *in vitro* this approach will require the use of xenograft assays. As substantial differences have been seen between targets identified using *in vitro* and *in vivo* screening approaches (Meacham, Ho et al. 2009), this final validation will also serve to confirm the relevance of targets to the *in vivo*, albeit murine, microenvironment.

The use of lentiviral transduction of patient-derived leukaemic cells will allow for the delivery of a short hairpin construct into cells with a low proliferative capacity. Such an approach is essential in patient-derived blasts which do not divide well *in vitro*. Introduction of shRNAs will allow the modulation of candidate genes throughout prolonged xenograft assays and potentially through serial engraftment, validating their role in leukaemic propagation. However, successful lentiviral transduction of patient-derived leukaemic cells offers a wider potential. Viral integration site analysis can be used for studying clonal complexity. Competitive transplantation of differentially labelled populations will examine fitness in engraftment. Tracking of cells labelled with Luciferase constructs will allow the real-time *in vivo* analysis of the effects of novel agents on disease progression with potential to look at specific sites of disease, such as the central nervous system.

This project used the pSLIEW transfer vector, Appendix B. This vector encodes both firefly Luciferase (fLuc) and enhanced green fluorescent protein (EGFP), separated by an internal ribosomal entry site (IRES) sequence, allowing translation of either gene from a single mRNA transcript. The expression of both of these genes is driven by the spleen focus forming virus (SFFV) promoter, as this promoter is subject to minimal silencing in haematopoietic cells.

The process of optimising the lentiviral transduction protocol involved two, occasionally opposing, objectives. Firstly, as patient-derived material is difficult to transduce it is important to maximise the efficiency of transduction. Secondly, however, maintaining viability of the patient-derived material *in vitro* is important so that viable cells are available to be transplanted at a suitable time point. As a number of the options available for increasing rates of transduction can affect viability, achieving a balance between them was critical.

6.2 Aims of the project

The aims of this project were to:

1. Develop a protocol for the lentiviral transduction of patient-derived leukaemic cells followed by the safe transplantation into NSG mice
2. Demonstrate successful transduction of the leukaemia propagating compartment by serial engraftment of NSG mice
3. Investigate the potential for *in vivo* tracking of leukaemic progression using a firefly Luciferase expressing vector

The approach taken was a step-wise modification to an existing spinfection protocol used for the lentiviral transduction of derived cell lines. Spinfection involves suspending target cells in the presence of infectious virus and then centrifuging the suspension to bring the cells and virus into apposition on the surface of the culture plate. Virus is then taken up by the cell, its RNA genome reverse transcribed into DNA which is then integrated into the target cell's genome by the viral protein Integrase.

6.3 Results

6.3.1 Optimising the transduction protocol

The standard spinfection protocol and the final modified patient-derived material protocol are shown in Figure 6-1. The following sections describe the stepwise modification from the standard to the modified protocol. The details of each of the nine transductions performed during this project are presented in Table 6-1. The details of patients are recorded in Appendix A.

6.3.1.1 Spinfection surface area

This standard protocol involved spinfection of cells in a 24 well plate. The first modification which was made was to switch to using a 48 well plate. The reduction in the surface area of the well from 1.9 cm^2 to 0.95 cm^2 brought the cells closer together, increasing the likelihood of viral particles making contact with a cell. Having blasts in greater proximity may also help improve viability by providing cell surface contact signals. The validity of these assumptions was not tested experimentally however.

Figure 6-1. Spinfection protocol. A) Standard spinfection protocol for transduction of derived cell lines. RPMI 1640 – Roswell Park Memorial Institute medium. FCS – Fetal calf serum. B) Modified spinfection protocol for the lentiviral transduction of patient derived material, elimination of residual infectious virus and intrafemoral transplantation. SFEM – Serum free expansion medium. IL - Interleukin

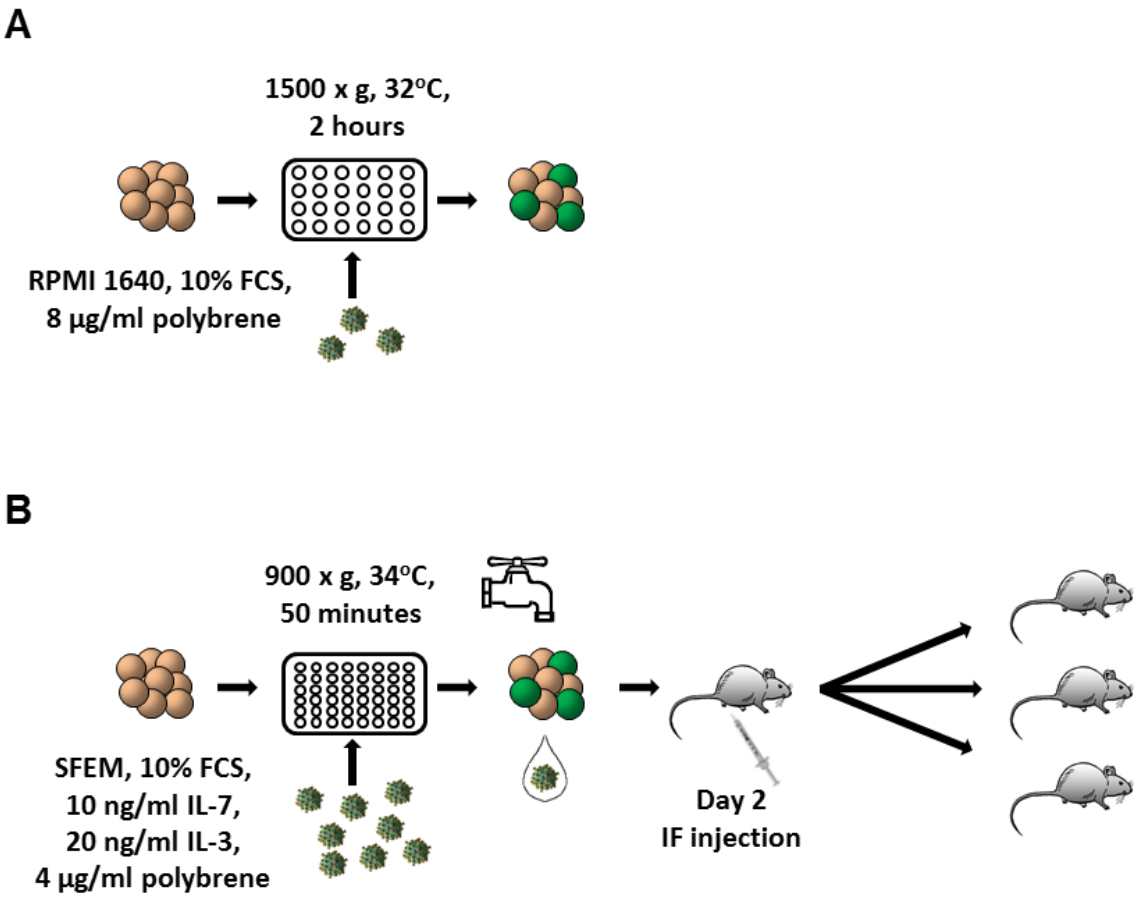


Table 6-1. Details of nine transductions using patient-derived leukaemic blasts. PB – Peripheral blood. BM – Bone marrow. T – Tumour (xenograft). αMEM – Alpha modified minimum essential medium. SFEM – Serum free expansion medium. SCF – Stem cell factor. IL- Interleukin. HS – Horse serum. HS-5 CM – Medium conditioned using HS-5 cells. EGFP – Enhanced green fluorescent protein.

Transduction	Sample ID	PB/BM/T	Cytogenetic subtype	Source	Medium	Adjunct	Growth factors	%EGFP
1	L826	PB	t(4;11)(q23;q21)	Primary	αMEM	Polybrene 8 µg/ml	SCF	19.6
							SCF/IL-7	24.0
							SCF/IL-7/IL-3	23.3
					Optimem	-	SCF/IL-7/IL-3	17.8
2	L826	BM	t(4;11)(q23;q21)	Primary	αMEM	Polybrene 8 µg/ml	SCF/IL-7	16.0
							SCF/IL-7/IL-3	10.4
						Protamine 5 µg/ml	SCF/IL-7	18.8
							SCF/IL-7/IL-3	18.6
					Optimem	-	SCF/IL-7/IL-3	19.0
3	L4951	BM	t(9;22)	Primograft	αMEM	Polybrene 8 µg/ml	SCF/IL-7	40.2
						Protamine 5 µg/ml	SCF/IL-7	50.0
4	L578	T	High Hyperdiploidy	Primograft	SFEM	Polybrene 8 µg/ml	-	15.0
							HS-5 CM/HS	20.4
5	L826	PB	t(4;11)(q23;q21)	Primary	SFEM	Polybrene 8 µg/ml	HS	27.5
						Polybrene 8 µg/ml	SCF/IL-7	35.6
						Polybrene 4 µg/ml	HS	26.2
						Polybrene 4 µg/ml	SCF/IL-7	31.9
						Protamine 5 µg/ml	HS	14.6
						Protamine 5 µg/ml	SCF/IL-7	19.7
6	L826	BM	t(4;11)(q23;q21)	Primograft	SFEM	Polybrene 4 µg/ml	SCF/IL-7	9.1
7	L880		t(4;11)(q23;q21)	Primary	SFEM	Polybrene 4 µg/ml	SCF	3.4
							IL-7	8.2
							SCF/IL-7	6.4
8	L826	BM	t(4;11)(q23;q21)	Primary	SFEM	Polybrene 4 µg/ml	IL-7/IL-3	14.0
9	L839	BM	t(2;10)(q2?1;q2?2)	Primary	SFEM	Polybrene 4 µg/ml	IL-7/IL-3	41.8

6.3.1.2 Transduction medium

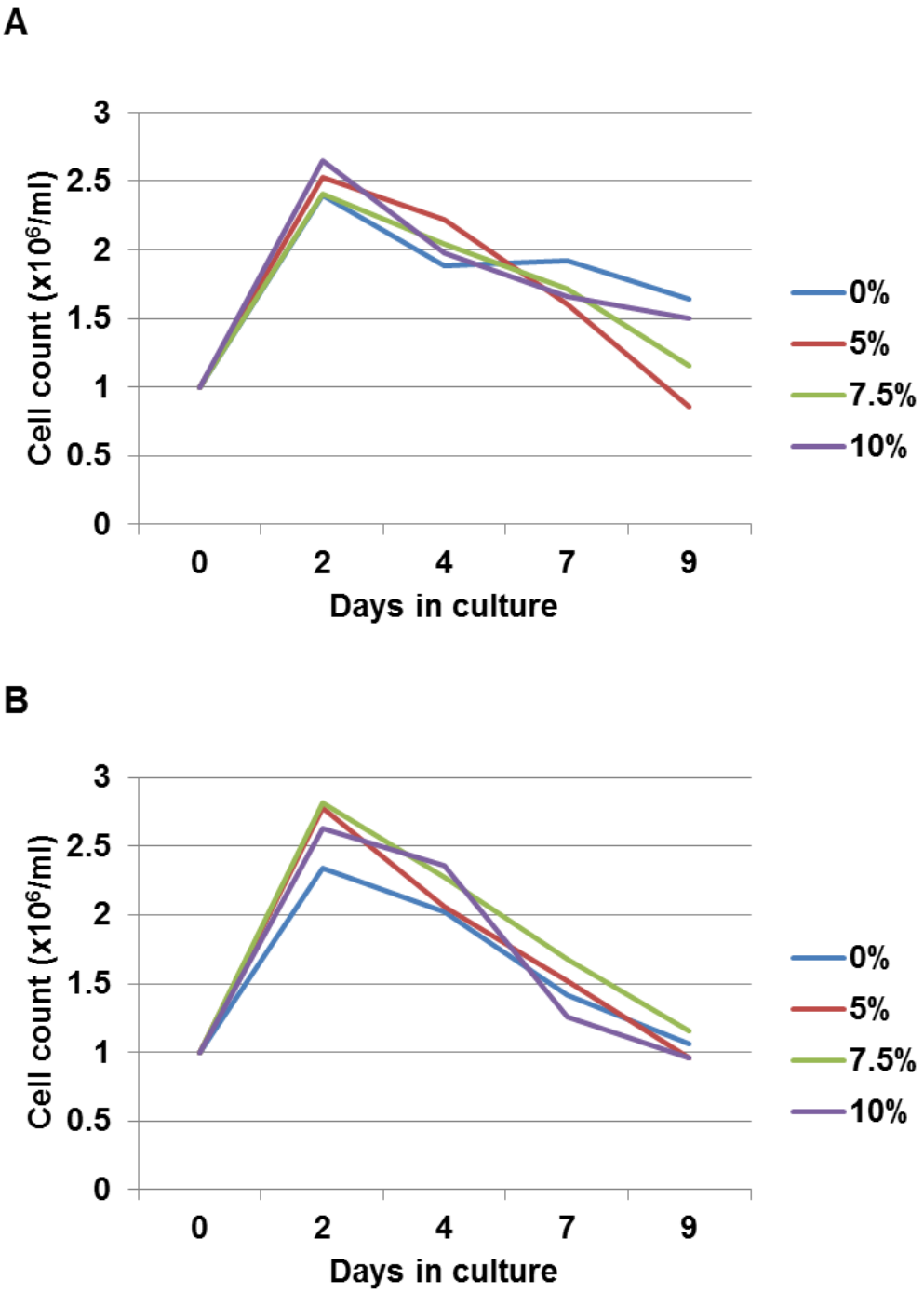
Transduction of derived cell lines used the standard culture medium for those cells, typically RPMI 1640 with 10% fetal calf serum. In order to maximise the viability of patient derived material *in vitro*, a more enriched medium, alpha modified minimal essential medium Eagle (αMEM), was initially used in comparison with a serum free approach using Optimem. The potential benefits of using Optimem were: 1) that in some circumstances the absence of serum may improve rates of transduction; 2) transduction without serum can be performed without the need for polybrene which is toxic to most mammalian cells. αMEM was supplemented with 20% v/v fetal calf serum plus a further 20% v/v horse serum which provides a rich source of growth factors. RPMI 1640 which had been conditioned with HS-5 feeder cells for 72 hours was added to 10% v/v. Finally, cytokines were added as described in 6.3.1.4 and polybrene was used at 8 µg/ml. This medium was compared with a serum free alternative Optimem, supplemented with 20% v/v Optimem conditioned with HS-5 feeder cells as described for αMEM. This comparison was performed in transductions 1 and 2 (Table 6-1).

In the first of these experiments, the rate of transduction with αMEM was 23.3% whilst with Optimem it was 17.8%. In the second experiment the rate of transduction was 10.4% with αMEM but 19.0% with Optimem. Whilst this was inconclusive, the results of the transduction of cells lines were consistently showing a substantial advantage to using complete medium with fetal calf serum and polybrene (Table 5-1). At the same time, the basic medium was switched to be Serum Free Expansion Medium (SFEM), a medium specially formulated for the maintenance of haematopoietic stem and progenitor cells.

A comparison of the effect of different concentrations of horse serum (0%, 5%, 7.5% or 10% v/v) and HS-5 conditioned medium (0 or 10% v/v) on cell viability *in vitro* was made. The material used was primografted L578, harvested from a thigh tumour. The tumour was disaggregated through a 40 µm cell strainer and resuspended in SFEM/10% v/v FCS. Cells were plated into a 24 well plate at 10^6 cells/ml, 1 ml per well and incubated at 37^{o}C, 5% CO_2. Counts of viable and non-viable cells (by trypan blue exclusion) were made each day between days 1 and 9. This showed that freshly harvested cells initially expanded *in*

vitro in all media. When grown in SFEM/10% v/v FCS, there was no difference in viable cell counts between samples incubated without (Figure 6-2A) or with (Figure 6-2B) 10% v/v HS-5 conditioned medium. Neither did the viable cell count vary with different concentrations of horse serum between 0-10% v/v (Figure 6-2). The standard transduction medium used thereafter was therefore SFEM with 10% v/v FCS.

Figure 6-2. Proliferation and survival of primografted L578 *in vitro*. A) Viable cell numbers in SFEM/10% v/v FCS without conditioned medium but with the addition of 0-10% v/v horse serum. B) Viable cell numbers in SFEM/10% v/v FCS with the addition of 10% v/v HS-5 conditioned medium and 0-10% v/v horse serum.

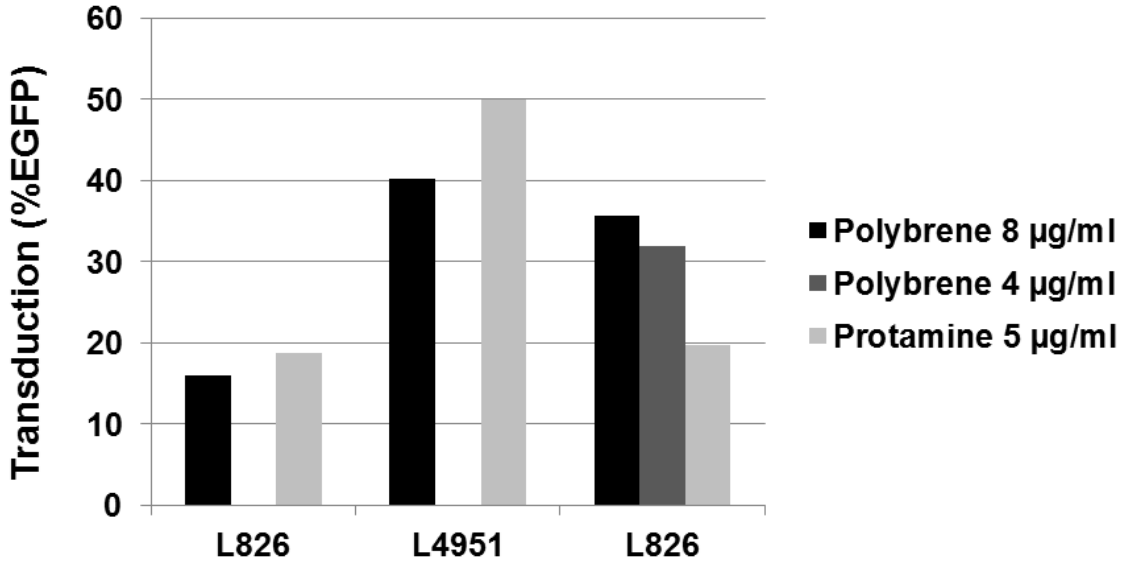


6.3.1.3 Cationic transduction adjuncts – polybrene and protamine

The cationic polymer polybrene is commonly used to facilitate viral transductions. The positive charge binds to, and neutralises, negatively charged cell surface glycoproteins, facilitating binding of negatively charged viral particles. However, polybrene is toxic to eukaryotic cells. A three way comparison was made between polybrene at a standard concentration of 8 µg/ml, polybrene at a reduced concentration of 4 µg/ml and an alternative cation, protamine sulphate, which may offer a lower degree of toxicity. Three transductions (2, 3 and 5, see Table 6-1) were performed comparing polybrene (8 µg/ml) with protamine (5 µg/ml). Of these, transduction 5 also compared these adjuncts with polybrene 4 µg/ml.

The relative efficacy of polybrene 8 µg/ml compared to protamine 5 µg/ml was inconsistent between transductions 2 and 3, performed in parallel, and transduction 5 (Figure 6-3). However, as with use of Optimem described in 6.3.1.2, repeated transductions with cell lines demonstrated the superiority of polybrene over protamine (Table 5-1). This was confirmed in transduction 5 (polybrene 8 µg/ml 35.6%, protamine 5 µg/ml 19.7%, Figure 6-3). As the relative efficacy of polybrene 8 µg/ml and polybrene 4 µg/ml were very similar (35.6% compared to 31.9%), it was decided to use polybrene 4 µg/ml to limit the associated toxicity.

Figure 6-3. Transduction efficiency using different cationic adjuncts. The efficiency of lentiviral transduction, determined by flow cytometry of %EGFP expression. Three experiments are shown, L826 (transduction 2, left bars), L4951 (transduction 3, middle bars) and L826 (transduction 5, right bars).



6.3.1.4 Recombinant cytokine adjuncts

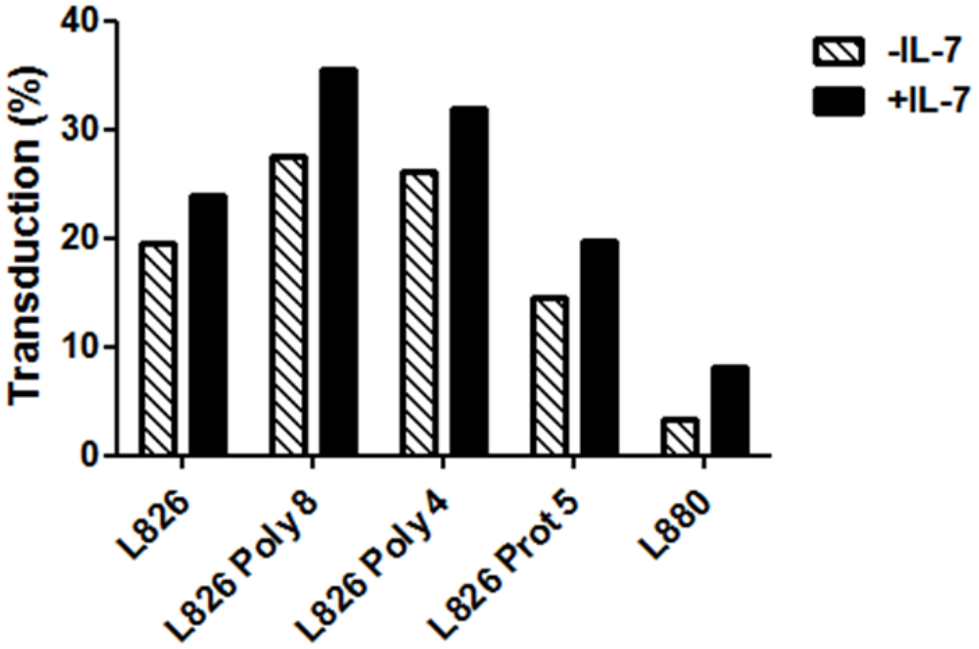
The recombinant human cytokines stem cell factor (SCF), interleukin (IL) 7 and IL-3 were used to support the viability of patient-derived leukaemic blasts *in vitro*. This combination had been successfully used to maintain the growth of leukaemic blasts *in vitro* (Cox, Evely et al. 2004). In order to identify which of the three cytokines were most important to leukaemic survival and efficiency of lentiviral transduction, comparisons were made of the effect of individual cytokines versus combinations, where the number of available blasts and quantity of virus allowed.

Three individual transductions (1, 2 polybrene and 2 protamine) compared SCF/IL-7 versus SCF/IL-7/IL-3 and three transductions (1 αMEM, 5 and 7) compared SCF versus SCF/IL-7. The last of these, transduction 7, also compared SCF alone versus IL-7 alone versus the combination of both.

Three comparisons of SCF/IL-7 versus SCF/IL-7/IL-3 demonstrated no improvement in transduction efficiency with the addition of IL-3. This cytokine was therefore withdrawn from subsequent transductions. It was, however, reintroduced for transductions 8 and 9 due to the potential offered to support leukaemic survival and growth *in vitro* (Uckun, Gesner et al. 1989; Eder, Ottmann et al. 1990; Brown, Nosaka et al. 1999). No experiments were performed to validate this.

In the three transductions comparing IL-7 with medium not containing IL-7 (transductions 1, 5 and 7), there was a significantly higher rate of transduction in IL-7 containing experiments, range 4.4-8.1% ($p=0.03$, Wilcoxon matched-pairs signed rank test) (Figure 6-4). Transduction 7 confirmed not only that IL-7 produces a higher rate of transduction, but combination of SCF and IL-7 did not give a higher rate of transduction than IL-7 alone, 6.4% versus 8.3% respectively (Table 6-1). SCF was therefore withdrawn from the final protocol.

Figure 6-4. Effect of IL-7 on rates of transduction. Three protocols (performed as part of transductions 1, 5 and 7) examining IL-7 demonstrate a significantly higher rate of transduction than those without IL-7, p=0.03.

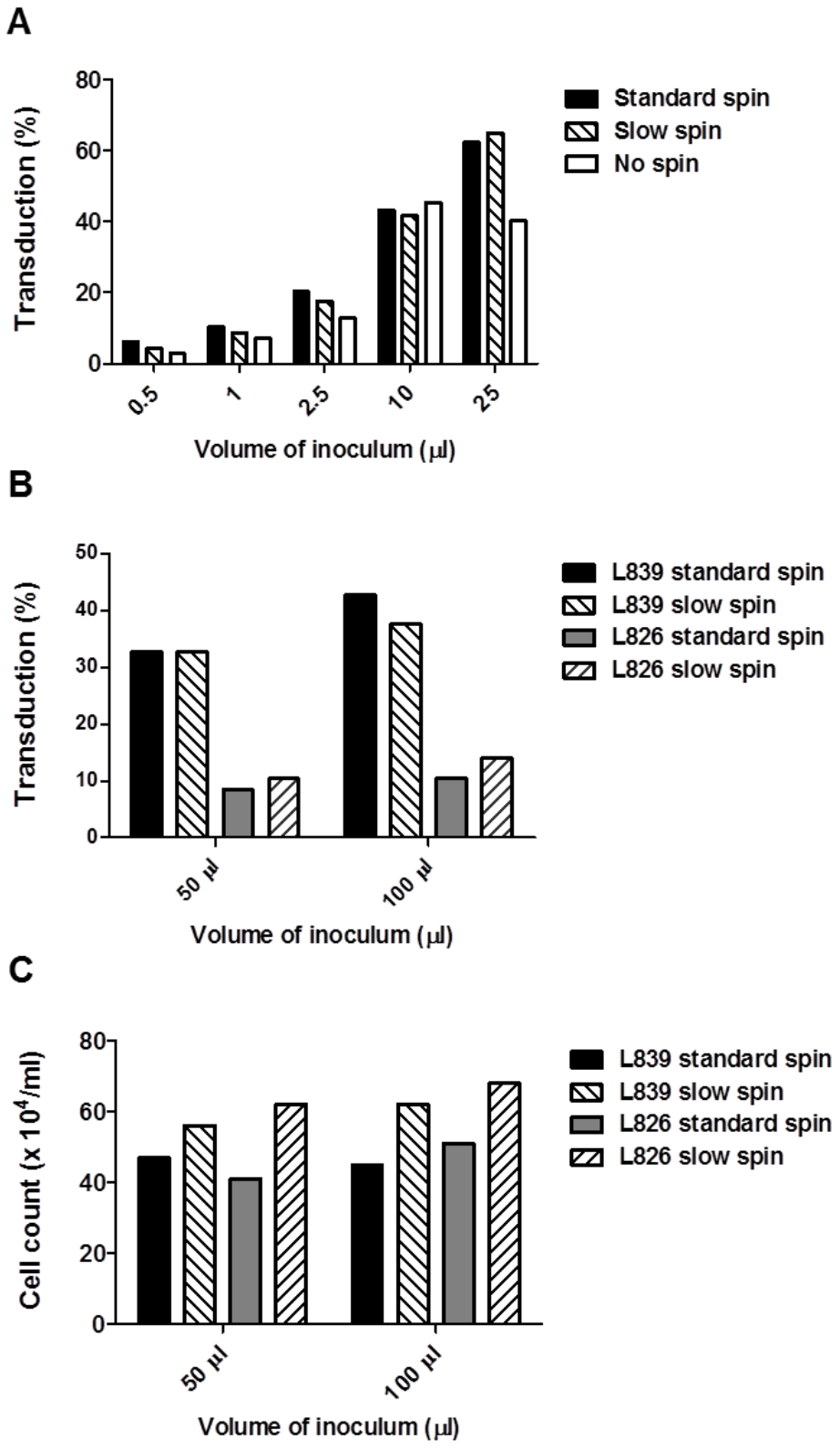


6.3.1.5 Spinfection protocol

As the standard spinfection protocol involves centrifugation at high speed (1500 x g) for a prolonged period (2 hours) a modification was made to reduce these parameters to limit the resultant cell damage. A new spinfection protocol of 900 x g for 50 minutes was initially trialled in SEM cells in direct comparison to the standard protocol and a third plate which was not centrifuged but immediately placed into a cell culture incubator. In this setting, there was a 1.8-2.6% lower rate of transduction at the slower speed, and a 3.3-21.9% lower rate of transduction with no centrifuge step (Figure 6-5A).

The slower spinfection protocol was trialled in two patient-derived specimens (transductions 8 and 9). Here, both transduction efficiency and cell viability were assessed. This confirmed that there was no loss of transduction efficiency (Figure 6-5B). However, all populations in the slow spin plate showed improved cell viability on day 4 post transduction, range 19-51% p=0.048, compared with the standard spin plate (Figure 6-5C).

Figure 6-5. Effect of modified spinfection protocol. The modified, slow spinfection protocol does not substantially reduce the rate of transduction in either A) SEM or B) 2 patient derived specimens. C) The modified protocol does result in a significantly higher cell viability on Day 4 post transduction, p=0.048.



6.3.1.6 Relative transduction efficiency

For the two final transductions, 8 and 9, the complete modified protocol was used along with viral suspensions which had been titred in SEM cells using the standard spinfection protocol. This allowed estimation of the relative transduction efficiency of these two patient-derived samples compared with SEM cells. The titration in SEM cells had been repeated on two occasions, giving titres of 5.6×10^7 TU/ml and 1.2×10^8 TU/ml respectively. The functional titre in sample L826 was 7.0×10^5-1.6×10^6 TU/ml, resulting in a relative transduction efficiency compared to SEM cells of 0.6-3.4%. For sample L839 the functional titre was 1.9-5.3×10^6 TU/ml, giving a relative transduction efficiency of 1.6-11%.

This result gives a broad idea of the titre of virus required to transduce patient-derived cells. However, it also shows that different patient specimens will need different viral titres applied, and that for accurate rates of transduction the relative transduction efficiency will need to be determined in advance for each specimen.

6.3.2 Serial transplantation and *in vivo* monitoring of transduced patient derived material

In order to assess the functional potential of transducing patient-derived leukaemic cells, three samples were transplanted intrafemorally into highly immunocompromised NSG mice. The first two samples came from transductions 2 and 3 and were transplanted by Klaus Rehe, who also performed the bone marrow punctures. The third sample was transduced and transplanted by Lars Buechler, but monitored, harvested and the second generation experiments conducted by SB. The second generation transplants of this third sample, L578, were performed by Frida Ponthan.

6.3.2.1 Transplantation of transduced L826 and L4951 cells

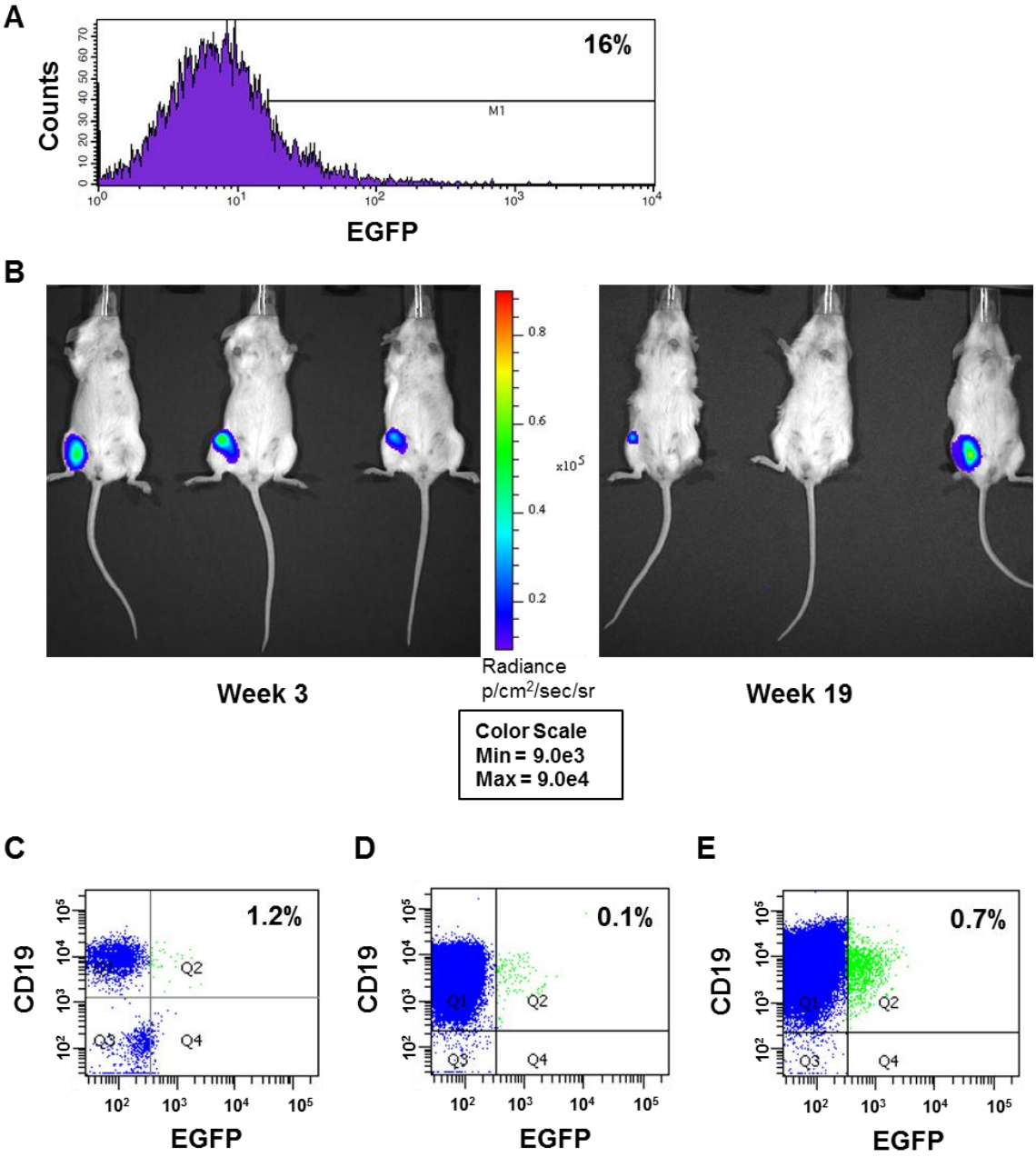
On day 1 following transduction, L826 and L4951 cells were washed twice with complete αMEM and resuspended at approximately 8×10^4 cells (L826) or 1×10^5 cells (L4951) per 30 µl transplant aliquot. This aliquot was injected into the right femur of each of three female NSG mice per sample: L826 – mice 1-3; L4951 – mice 4-6. Cells from the initial transduction which had not been used for transplantation were analysed by flow cytometry on day 4 post transduction to provide a baseline for the rate of transduction. This analysis demonstrated that sample L826 was transplanted with 16% expression of EGFP (Figure 6-6A) and L4951 transplanted with 40% expression of EGFP.

Engraftment was monitored using a Spectrum IVIS bioluminescent camera from week 2 onwards, every 1-3 weeks. Bone marrow punctures were performed at weeks 4, 8 and 13. At harvest, cells were recovered from femurs/tibias and spleens. Expression of EGFP was assessed by multicolour flow cytometry and cells were either stored viable in liquid nitrogen or re-transplanted fresh into a second generation of mice (L4951 only, bone marrow and spleen).

For sample L826, bioluminescent imaging demonstrated a moderate strength signal from the right femur of all three mice, present from week 3 (Figure 6-6B). The signal remained at the same intensity throughout the experiment. Furthermore, there was no evidence of disease dissemination over this period. Bone marrow punctures demonstrated engraftment of CD19+ cells, but <2% of

these expressed EGFP. At harvest, analysis of bone marrow from mouse 1 showed only 1.2% of CD19+ cells expressed EGFP, despite a persistent right femoral signal (Figure 6-6C). Mouse 2, which has lost the right femoral luciferase signal showed good engraftment but <0.1% of CD19+ cells expressed EGFP (Figure 6-6D). Mouse 3, also demonstrating a persistent right femoral signal, showed good engraftment but with only 0.7% CD19+ cells expressing EGFP. These low rates of EGFP expression in the presence of persistent femoral Luciferase signal, especially in mouse 1, raised concerns that murine tissue may have been transduced by SLIEW virus which had not been removed by standard cell washing procedures. No cells from this sample were re-transplanted.

Figure 6-6. Transplantation of L826-SLIEW. A) L826 was transplanted on day 1 following transduction with 16% of cells expressing EGFP. B) Serial bioluminescent monitoring demonstrated a signal from the right femoral injection site at week 3 (left panel) but showed no evidence of dissemination or local increase in signal intensity. Mouse 2 (centre mouse, right panel) lost the femoral signal by week 19. Bone marrow analysed at harvest showed very low expression of EGFP: C) Mouse 1 showed only poor engraftment (1 x 10^6 cells analysed); D) Mouse 2 showed strong engraftment but <0.1% of CD19+ cells expressed EGFP (1.5 x 10^5 cells analysed); E) Mouse 3 showed good engraftment and 0.7% of CD19+ cells expressed EGFP (2.5 x 10^5 cells analysed).

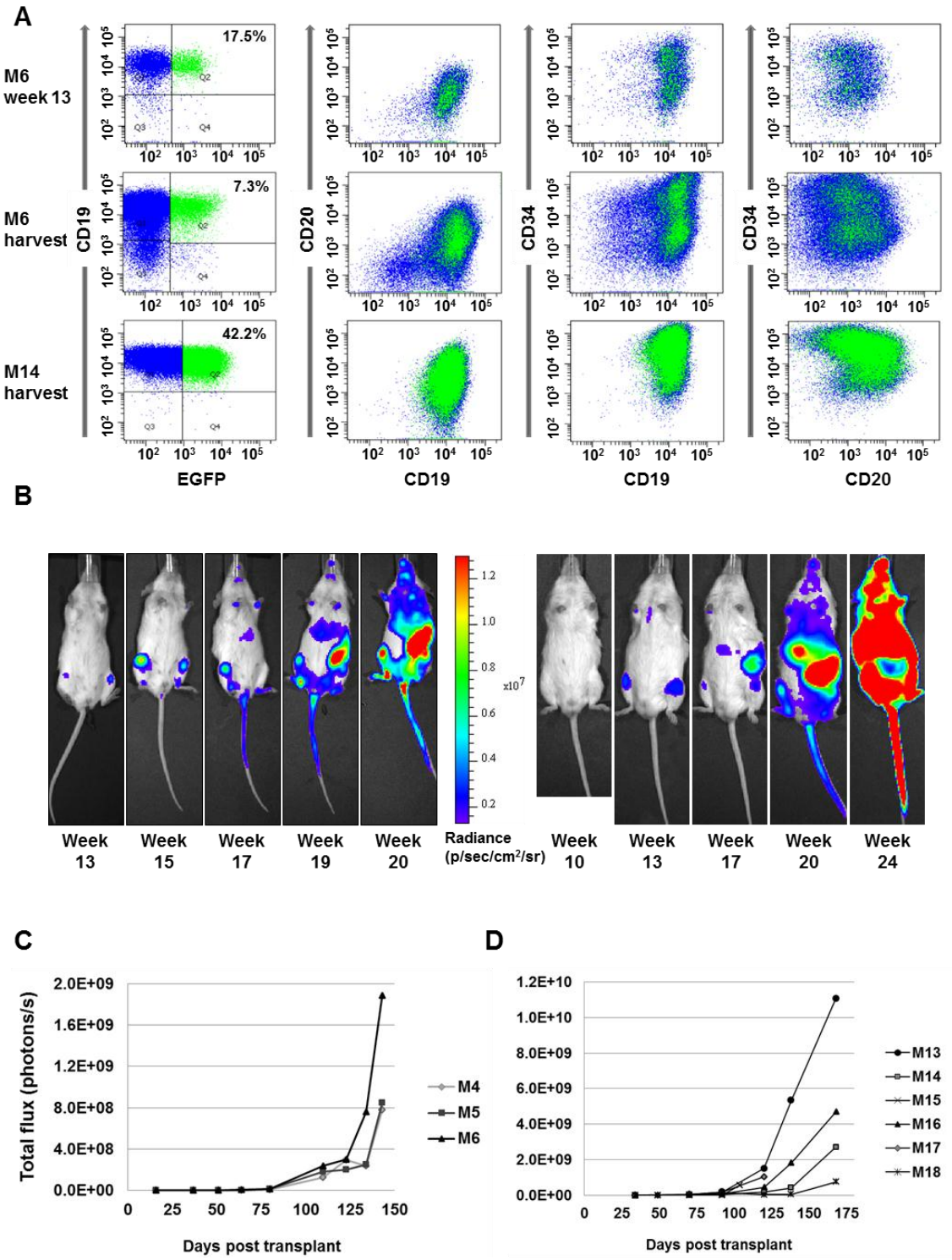


For sample L4951, both serial bone marrow punctures (Figure 6-7A) and bioluminescent imaging (Figure 6-7B) demonstrated the on-going expression of EGFP and imaging demonstrated the dissemination of transduced leukaemic cells to the contralateral femur, spleen, vertebrae and skull/meninges/brain. At the third bone marrow puncture, EGFP expression was found to have fallen from 40% at transplantation, to 8-17% in the three mice. At harvest this had fallen further in both spleen and bone marrow, being 5.3-10%. However, transduction was still seen equally within each of the immunophenotypic populations present (Figure 6-7A). Signal intensity increased exponentially as engraftment progressed (Figure 6-7C).

Twelve secondary recipient mice each received a total of 1×10^5 cells from either the bone marrow (n=3) or spleen (n=3) from either mouse 4 or mouse 6. Eleven out of twelve mice showed engraftment by bioluminescent imaging. For the recipients of cells harvested from mouse 4, 2 of 3 bone marrow and 3 of 3 spleen recipients engrafted. The pattern of dissemination did not vary between cells of different origin. At harvest, the expression of EGFP had dropped further to 0.2-1.8%.

All six recipients of cells harvested from mouse 6 engrafted. Again, the pattern of dissemination of cells from bone marrow and spleen did not differ. Surprisingly, however, expression of EGFP in CD19+ cells at harvest had increased in all mice from 7.3% (bone marrow) or 8.0% (spleen) to between 17 and 42% of CD19+ cells. Again, there was no bias in the populations expressing EGFP (Figure 6-7A).

Figure 6-7. Serial monitoring of engrafted L4951-SLIEW. A) Multicolour flow cytometry following bone marrow puncture of (top panels) and harvest (middle panels) of mouse 6 and harvest of mouse 14 (bottom panels). B) Serial monitoring of primary recipient mouse 6 (left panels) and secondary recipient mouse 14 (right panels). Mouse 14 received 1 x 10^5 bone marrow cells from mouse 6. Images show progressive dissemination to contralateral femur, spleen, liver and vertebrae. Analysis of bioluminescence from each C) primary or D) secondary recipient showing exponential increase in signal intensity during the period of engraftment.

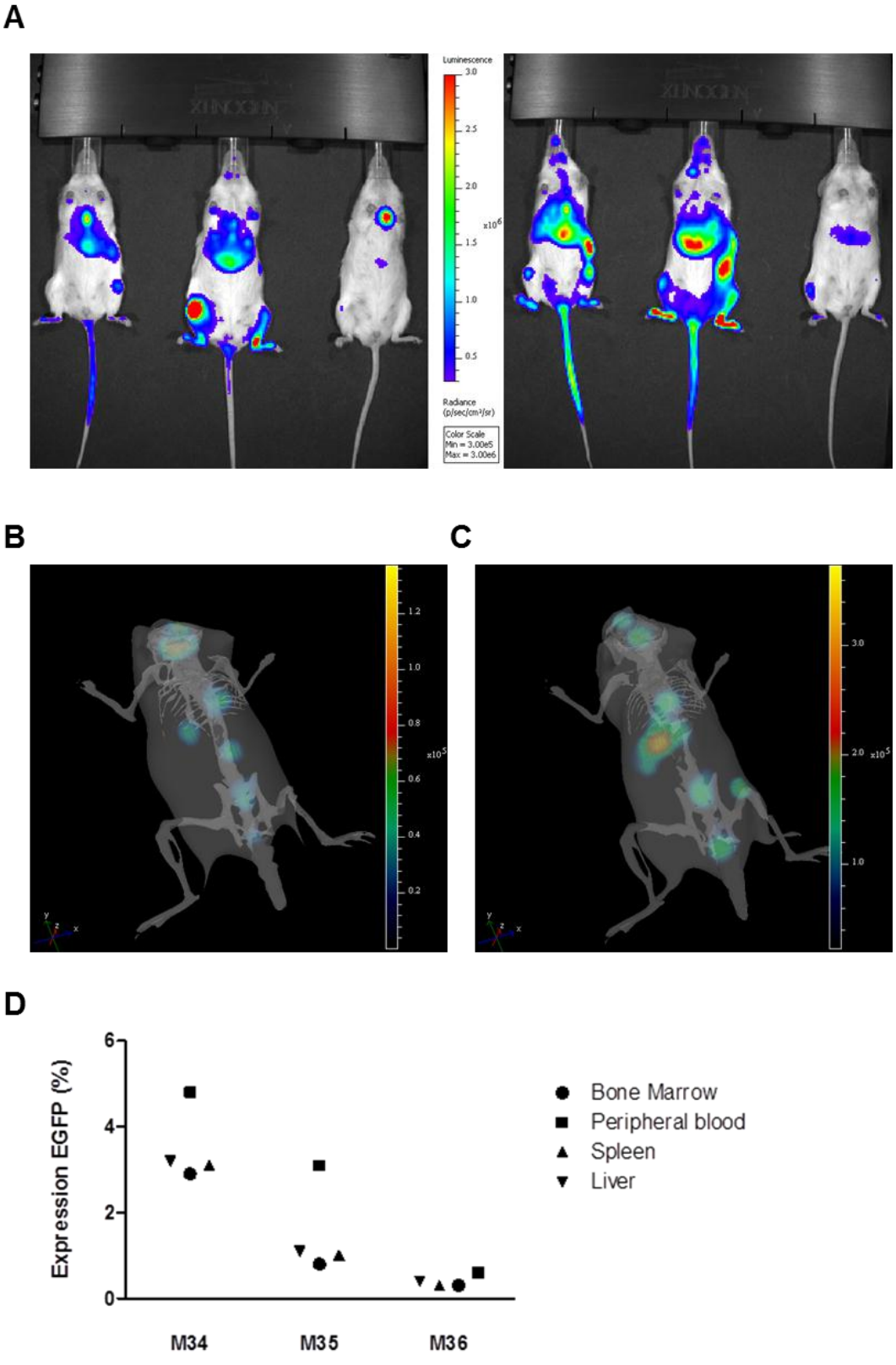


6.3.2.2 Monitoring and re-transplantation of L578

Leukaemic sample L578 had been transduced in a separate experiment. Expression of EGFP was substantially lower in this sample, being 2%. A total of 1 x 10^6 cells was transplanted on day 4 post transduction. As the initial experiment was discontinued, this mouse was brought into the present project and imaging continued by SB. This mouse showed widespread leukaemic dissemination by bioluminescent imaging. At harvest, expression of EGFP remained relatively stable at 1.6% (bone marrow) and 1.9% (spleen).

Bone marrow was re-transplanted into three recipients at a total dose of 1 x 10^5 cells per mouse. Serial bioluminescent imaging demonstrated disease dissemination (Figure 6-8A), which could be localised further by 3-dimensional reconstructions. These show that in this sample with a low rate of transduction, early engraftment is seen in discrete areas of skeleton and spleen (Figure 6-8B-C). These three mice were harvested at week 17, at which point the expression of EGFP in bone marrow was 0.3-2.9%. Flow cytometric analysis of cells isolated from peripheral blood, bone marrow, spleen, liver and kidney showed no difference in expression of EGFP between these tissues (Figure 6-8D).

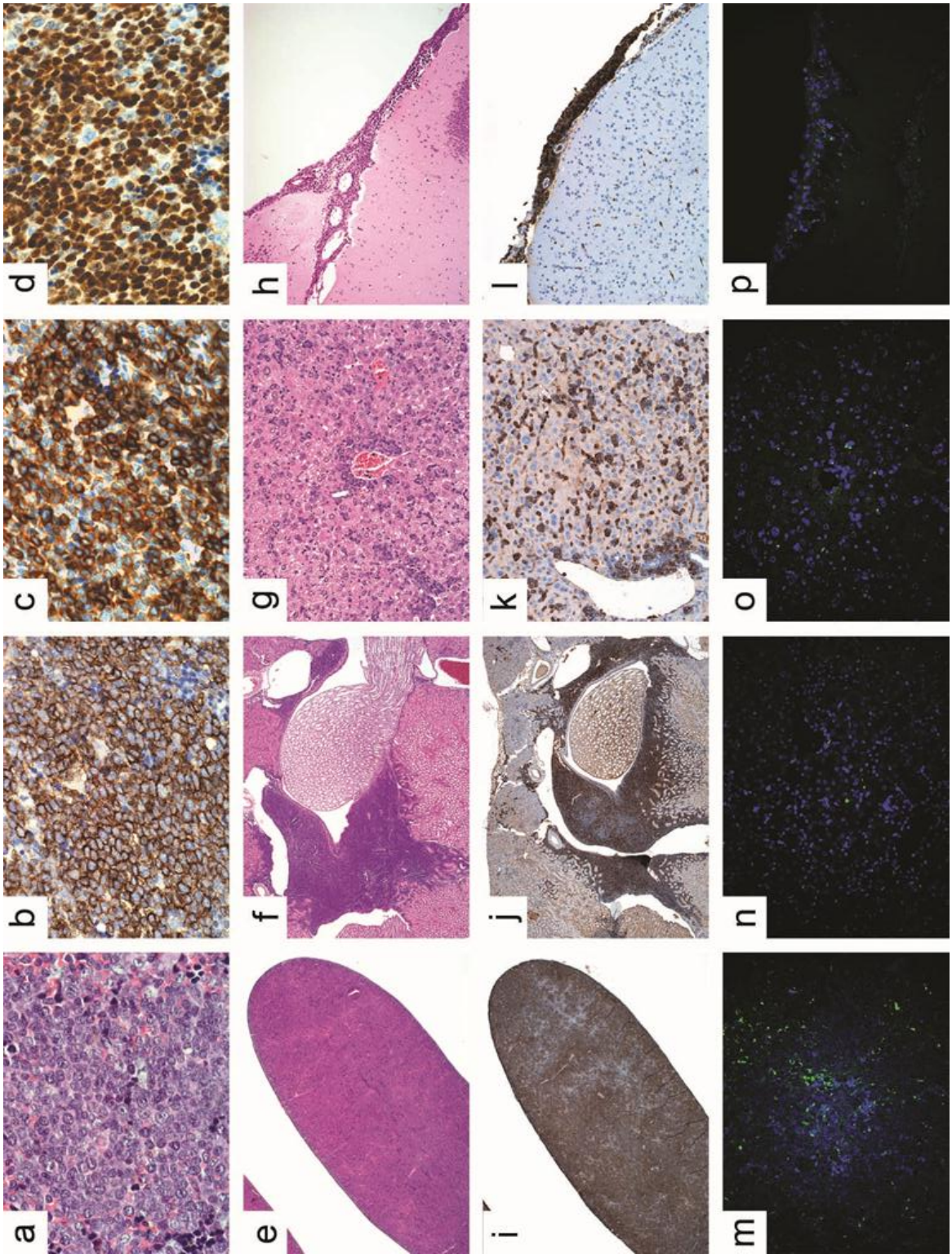
Figure 6-8. Bioluminescent monitoring of L578-SLIEW. A) *In vivo* monitoring demonstrating disease progression from week 14 (left panel) to week 17 (right panel). 3-dimensional reconstructions of mouse 35 (centre mouse in A) at B) week 12 and C) week 14, showing focal engraftment of vertebrae, increase in splenic signal and appearance of right femoral signal. Scales are photons/second. The skeleton is projected for orientation and not derived from imaging of this animal. D) Organ specific engraftment with EGFP+ blasts.



To further validate the model, brain, spleen, liver and kidneys harvested from the primary recipient of L578-SLIEW were stained with haematoxylin and eosin and immunolabelled with antibodies specific for CD19, CD10, CD79a and TdT (Figure 6-9). This showed a diffuse infiltrate of lymphoblasts expressing these characteristic B-precursor proteins. Whilst the bioluminescent imaging had been unable to differentiate between infiltration of the skull vault and the brain/meninges, histology confirmed meningeal infiltration (Figure 6-9H-L).

The presence of EGFP positive cells amongst the engrafting blasts was demonstrated by fluorescence microscopy, even in this sample with low rates of transduction (Figure 6-9M-O). This analysis confirmed the presence of EGFP positive cells within the meninges (Figure 6-9P).

Figure 6-9. Engraftment of primary recipient of L578-SLIEW. Engrafted L578 leukaemic cells showed typical lymphoblastic morphology (a) and expressed characteristic proteins CD19 (b), CD10 (c) and TdT (d). There was widespread dissemination to multiple sites including spleen (e, i, m), kidney (f, j, n), liver (g, k, o) and meninges (h, l, p) as illustrated by staining with H&E (e-h) and for CD79a (i-l). EGFP positive cells are seen sparsely throughout the organs (m-p), consistent with the low levels of transduction in this experiment.

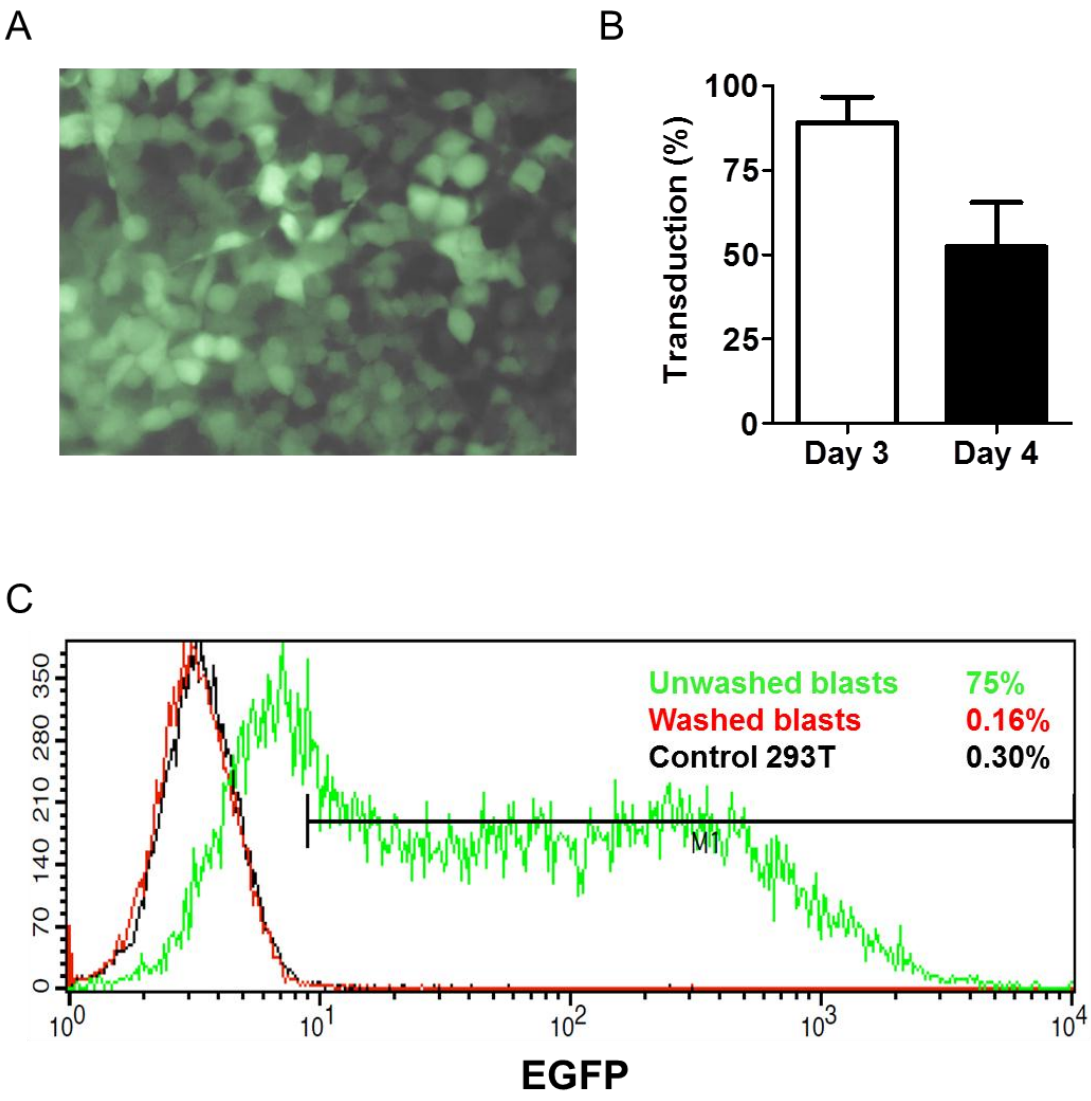


6.3.3 Ensuring safety prior to xenotransplantation

During the primary transplantations with specimens L826 and L4951, medium was taken from the cells prior to transplantation and put onto 293T cells growing in 24 well plates. These cells were incubated overnight before replacing the supernatant with standard complete DMEM. These cells were viewed with a fluorescent microscope at day 4 following transduction and seen to express EGFP strongly. This unexpected finding, along with the persistent luciferase signal from the site of injection in mice 1-3, suggested the persistence of infectious viral particles not only 24 hours post transduction but also following two washes of the cells prior to transplantation. An alternative protocol was therefore developed to ensure the removal of infectious virus prior to transplantation. This is essential to prevent transduction of murine cells (and false positive imaging results) as well as to protect the transplanter who uses sharp needles to perform the transplant.

A modified washing procedure was devised, as detailed in section 2.3.11.3. In three separate experiments, supernatant was taken from cells which had been stringently washed daily since day one post transduction and control cells which had not. Supernatant was again incubated with 293T cells overnight and then washed off the following morning. Supernatant from the unwashed cells was shown to have a high titre of infectious virus as long as four days post transduction (Figure 6-10A-B). However, virus was effectively removed by just two stringent washes, making it safe to transplant on day 2 post transduction (Figure 6-10C).

Figure 6-10. Persistence of infectious SLIEW virus following transduction of patient-derived material. 293T cells were assessed for transduction by expression of EGFP both by A) fluorescence microscopy - supernatant taken day 3 post transduction, and B) flow cytometry - supernatant was taken at either day 3 (n=3) or day 4 (n=2) (mean + standard deviation) from transduced but unwashed leukaemic blasts. Both demonstrate high residual transduction capacity. C) Effect of washing protocol - by day 2, washing transduced leukaemic blasts has effectively removed residual transduction capacity from the supernatant, whilst the supernatant from unwashed blasts retains a high viral titre.



6.4 Discussion

This project has developed a protocol for the transduction of patient-derived leukaemic blasts using a lentiviral construct expressing EGFP and firefly luciferase. This has allowed the *in vivo* tracking of disease progression using an NSG xenotransplantation model. Successful engraftment of two generations of mice with low doses of transduced cells (as low as 2×10^4 followed by 1.6×10^3) has confirmed the transduction of the leukaemia propagating cell compartment. However, the reasons for the substantial changes in EGFP expression remain unclear. Whilst in some samples (L826 primary transplant and L4951 mouse 4 secondary transplants) expression of EGFP was either lost or substantially reduced, in others it was stable (L578) or even increased (L4951 mouse 6 secondary transplant). Work is currently on-going within the laboratory to elucidate the changes seen in clonal complexity with serial transplantation, using ligation mediated PCR analysis of lentiviral integration sites. This work should define how many cells contribute to disease engraftment and may explain the variation in EGFP expression seen in the present project. However, initial results have demonstrated that oligoclonal outgrowth does not explain the increase in EGFP expression seen in secondary recipients of mouse 6 bone marrow, mice 13 and 14 (Bomken, Buechler et al. 2012), Appendix D , as persistence of stable polyclonal engraftment has been demonstrated.

Further improvements to the transduction protocol should allow both the reduction in the titre of virus required as well as improvements in cell viability. The use of additional or alternative cytokines may support cell growth, survival and consequently transduction. Co-culture with a feeder cell layer to support leukaemic viability was considered but not investigated here. The concern with using feeder cells is the almost certain transduction and subsequent transplantation of the feeder cell line. Whilst it may be possible to sort blasts from feeders by FACS, preliminary trials of sorting patient-derived blasts, post-transduction, have demonstrated extremely poor viability post-sorting. Immunomagnetic bead sorting may provide an alternative, but would not allow concurrent sorting using EGFP expression.

Whilst the xeno-transplantation studies described here have only sought to track disease progression *in vivo*, many more potential uses exist for this protocol. Examples include:

- using shRNA containing constructs to modulate the expression of genes of interest – the development of such a vector is described in Chapter 7;
- monitoring disease progression during testing of novel therapies/drug combinations – an early experiment using a novel therapy is currently in progress using FACS sorted tertiary transplants of L4951 harvested from the spleen of mouse 14;
- transplantation with blasts transduced with vectors containing different fluorescent/luminescent markers to examine fitness to engraft in a competitive situation.

The on-going development of this protocol will allow for these experiments to be developed both in leukaemic cells and, potentially, in other malignancies as well.

Chapter 7
Development of a single lentiviral vector for the transduction of primary material

7.1 Introduction

This project was designed to modify the existing pSLIEW lentivector to include a shRNA sequence in addition to the existing firefly luciferase and EGFP genes (Appendix B). The purpose behind the cloning of a single vector capable of *in vitro* monitoring/cell sorting, *in vivo* bioluminescent monitoring and RNAi was to create a tool for the transduction and modulation of patient-derived leukaemic blasts for xenotransplantation assays. Existing commercial vectors would only allow two of the three functions. Therefore, to achieve all three functions would require multiple transductions of primary blasts. The limited viability of ALL blasts *in vitro* makes it highly unlikely that this could be performed successfully.

The requirements for the strategy were:

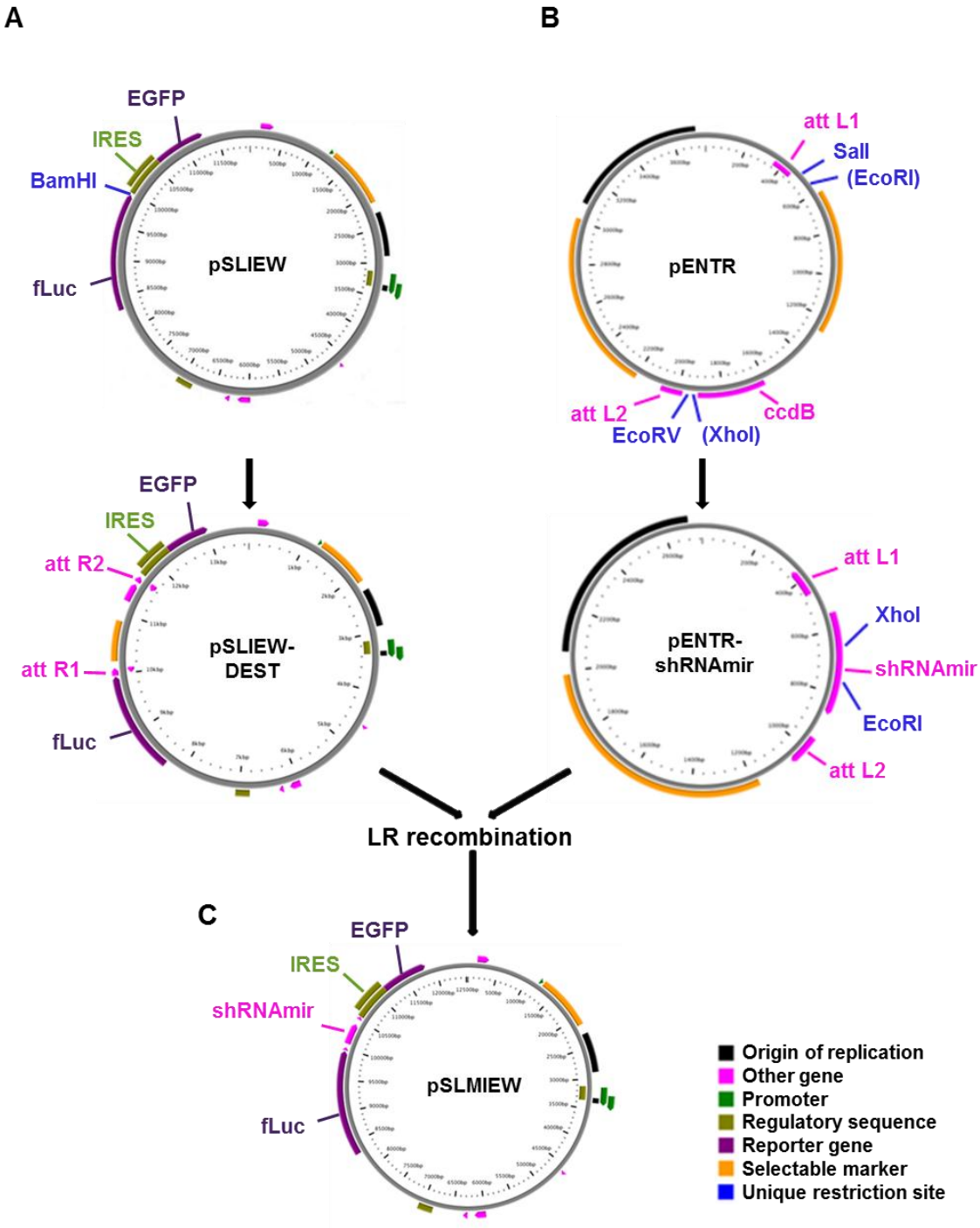
1) To clone into the single remaining unique restriction site within the transcribed region of the provirus - BamHI, 3' of fLuc, 5' of the IRES
2) To allow easy, repeated cloning of different shRNA constructs

The approach taken was to develop two vectors derived from the Gateway recombination system from Invitrogen. Conventional restriction cloning was used to produce a Destination vector and an Entry vector. Employing a system developed from lambda bacteriophage recombination, these two vectors are then recombined at complementary attachment (*att*) sites, switching the sequence contained between *att* sites on one vector with that contained between the *att* sites of the other vector. For the current project, that meant cloning the Gateway Destination cassette into the BamHI site of pSLIEW (to give pSLIEW-DEST) (Figure 7-1A) and the shRNA sequence into a commercial Gateway Entry vector (to give pENTR-shRNAmir) (Figure 7-1B). A recombination reaction would then clone the shRNA sequence between fLuc and the IRES site of the resultant pSLMIEW (Figure 7-1C).

In order to maximise the efficiency of knockdown in hard to transduce patient derived cells, a microRNA adapted shRNA sequence was used, as described in Chapter 5. This sequence has unique XhoI and EcoRI restriction sites within

the microRNA backbone, allowing new shRNAs to be inserted into the backbone by conventional restriction cloning. The Entry vector cloning strategy was designed to exploit this potential by excising XhoI and EcoRI sites from the multiple cloning site of pENTR1a (Figure 7-1B).

Figure 7-1. Cloning strategy to create pSLMIEW. A) The Gateway Destination cassette was to be cloned into the BamHI site between the fLuc and IRES of pSLIEW to produce pSLIEW-DEST. B) The microRNA adapted short hairpin sequences produced by PCR amplification from existing pTRIPZ/pGIPZ vectors were cloned into pENTR using SalI and EcoRV restriction sites, thereby excising EcoRI and XhoI restriction sites from the residual multiple cloning site of pENTR-shRNAmir. The remaining XhoI and EcoRI sites within the shRNAmir sequence were therefore unique, allowing repeated cloning of new shRNA sequences into the microRNA backbone of pENTR-shRNAmir. C) Lambda recombination of complementary *att*L and *att*R sites would then insert the shRNAmir30 sequence into the pSLIEW vector to create a new pSLMIEW transfer vector.



7.2 Aims of the project

The aims of this project were to:

1. Clone the Gateway Destination cassette into the BamHI site in the open reading frame of pSLIEW
2. Clone an shRNAmir30 construct into the multiple cloning site of pENTR
3. Clone a novel shRNA construct into the microRNA backbone sited in pENTR (proof of principle)
4. Recombine the new pSLIEW-Destination vector and pENTR-shRNAmir vector to produce a single vector expressing fLuc, shRNAmir30 and EGFP – pSLMIEW
5. Functionally validate this vector by:
    a. Producing infectious virus
    b. Analysing shRNA production, target knockdown and phenotype
    c. Validating residual fLuc expression
6. Perform a preliminary *in vivo* assessment of the vector following xenotransplantation

The cloning component of this project was completed as part of the supervision of two BSc students, Nana Anim-Addo who cloned the pSLIEW-DEST vector (7.3.1) and performed the initial recombination reactions (7.3.4) and Cara Hernon who cloned the shANGPT1 construct into the existing pENTR-shRNAmir vector (7.3.3) and performed the respective recombination reaction (7.3.4). The cloning strategy was designed by SB. All lentiviral work was completed by SB.

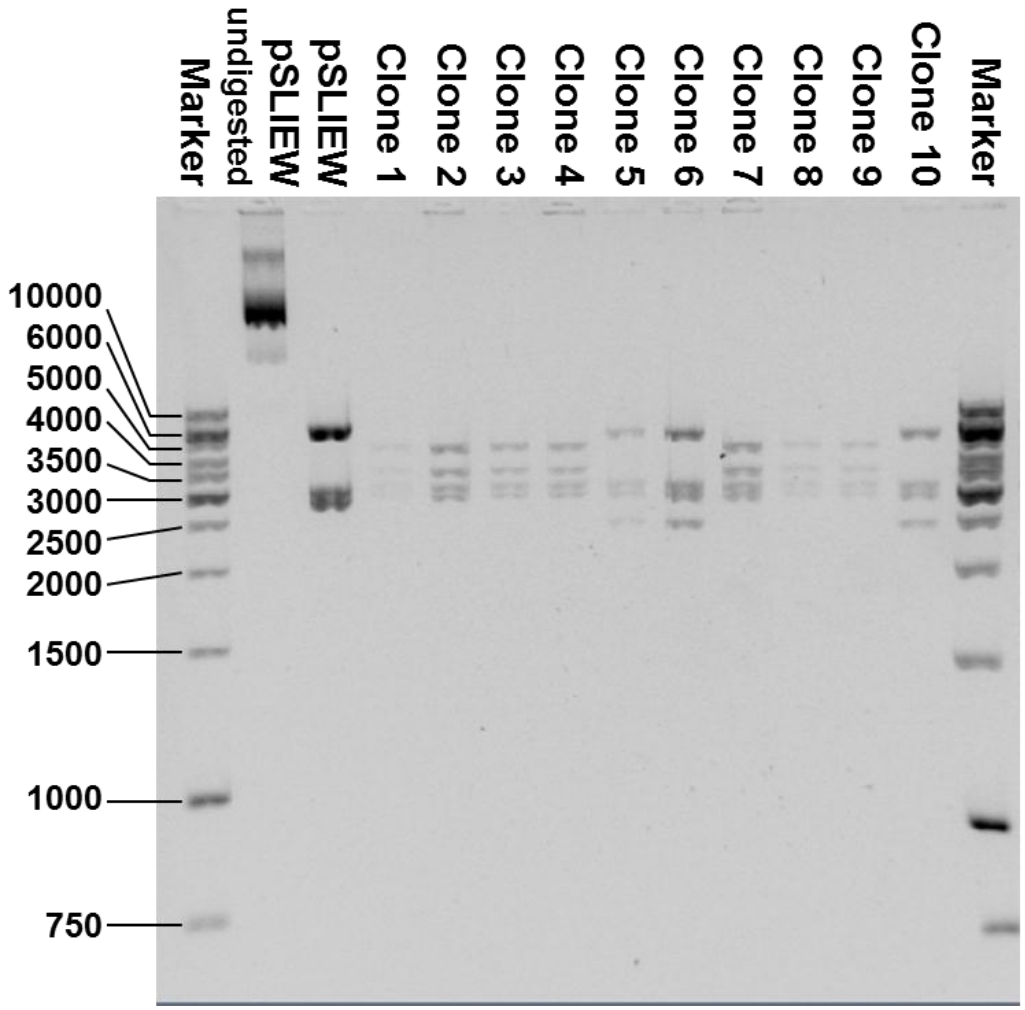
7.3 Results

7.3.1 Cloning of pSLIEW-Destination vector

pSLIEW was digested using BamHI and the resultant 5' overhang filled in with a Klenow reaction. In order to stop the linearised vector from re-ligating, 5' phosphate groups were removed with calf intestinal alkaline phosphatase. The Gateway destination cassette, reading frame A, was then ligated in using T4 DNA ligase. The initial trial of this protocol did not produce any colonies from transformed bacteria, so the protocol was modified to use Shrimp alkaline phosphatase which is more efficiently heat inactivated. With this modification, 24 colonies were produced from transformed ccdB Survival 2T1R competent cells. Ten colonies were grown overnight and plasmid DNA extracted. A control digest was performed using PvuII which was able to determine the orientation of the destination cassette. In 7/10 clones this digest produced bands at 2850bp, 3066bp, 3364bp and 4296bp as would be expected for cassette insertion in the correct orientation. In the remaining 3/10 clones it demonstrated bands at 2355bp, 2850bp, 3066bp and 5305bp as would be expected for reverse insertion of the cassette (Figure 7-2).

Clones 2, 4, 6 and 10 were sequenced commercially using the Destination vector sequencing primers 1 and 2 (DBS Genomics). This confirmed the insertion of the correct sequence at the BamHI site with fLuc 5', and IRES 3', to the cassette. Clone 2 was expanded and plasmid DNA extracted for use in recombination experiments (sequence and plasmid map are detailed in Appendix B).

Figure 7-2. Restriction digest of pSLIEW-Destination vector. PvuII digest showing the bands expected of a correctly orientated Gateway destination cassette (2850bp, 3066bp, 3364bp and 4296bp) in clones 1-4, 7-9). The bands expected of a reverse orientated cassette (2355bp, 2850bp, 3066bp and 5305bp) are seen in clones 5, 6 and 10). Marker – Fermentas 1kb DNA marker.



7.3.2 Cloning of pENTR-shRNAmir constructs

Three shRNAmir30 sequences were amplified by PCR from existing Open Biosystems vectors. These vectors contained shRNAs specific for the MLL/AF4 fusion transcript (variant found in SEM cells), the RUNX1/RUNX1T1 fusion transcript (found in Kasumi-1 and SKNO-1 cells) and a non-targeting sequence provided by Open Biosystems as a negative control. Primers were designed to add SalI and EcoRV restriction sites at the 5' and 3' ends respectively. The PCR products were restriction digested using SalI and EcoRV, as was the pENTR vector. Gel purification of the digested pENTR vector prevented reincorporation of the excised multiple cloning site.

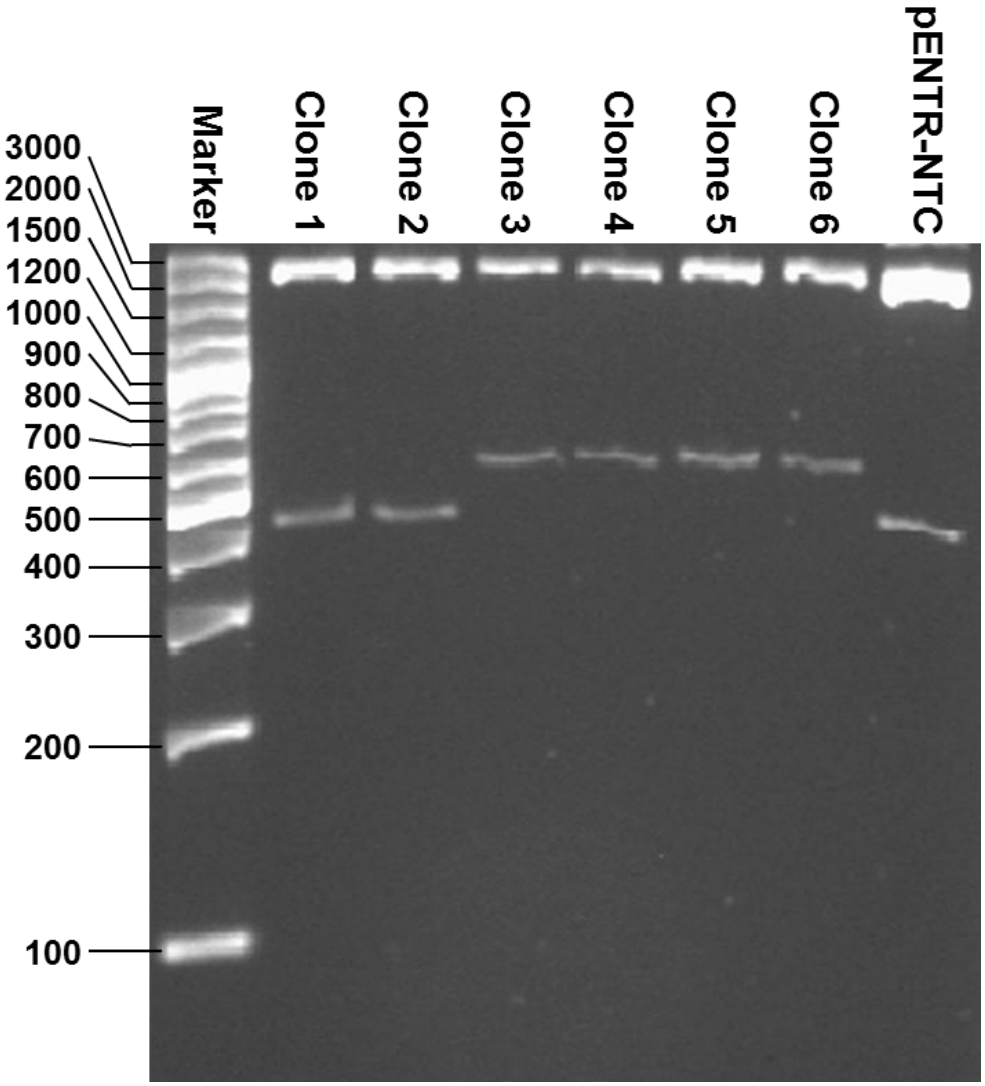
Digested shRNAmir30 sequences and pENTR were ligated to give the three pENTR-shRNAmir constructs. The ligated plasmids were transformed into DH5α cells, cultured overnight and colonies picked and expanded over a second night. Plasmid DNA was extracted from successful cultures. In view of the PCR step, all clones were sequenced commercially (DBS Genomics) using the pENTR reverse sequencing primer to demonstrate the accuracy of the shRNAmir sequence. One correctly sequenced clone for each construct was expanded and plasmid DNA isolated for recombination experiments (sequences and plasmid map are detailed in Appendix B).

7.3.3 Cloning of the novel shANGPT1 hairpin into pENTR-shRNAmir

In order to demonstrate the principle of cloning a new shRNA sequence into the newly developed pSLMIEW/pENTR-shRNAmir30 system, an shRNA sequence specific for *ANGPT1* was excised from a commercially available pTRIPZ vector using XhoI and EcoRI. The existing non-targeting sequence was excised from the control pENTR-shRNAmir construct described in 7.3.2. After gel purification of the digested pENTR-shRNAmir vector, the shANGPT1 sequence was ligated in using T4 DNA ligase and transformed into DH5α cells. Colonies were picked, expanded overnight and plasmid DNA extracted from successful cultures.

Double restriction digests using EcoRI and SalI failed to give the predicted bands of 382bp for the non-targeting control vector and 468bp for the successfully clones shANGPT1 vector (Figure 7-3). However, as clones 3-6 gave a band approximately 80-90bp longer than the non-targeting vector (included as a negative control), it was presumed that these clones now had the alternative shANGPT1 construct. All six clones were sequenced using the pENTR reverse sequencing primer (DBS Genomics). This demonstrated that the 4 clones giving the larger bands (clones 3-6) had the correct sequence for the shANGPT1 construct (sequences and plasmid map are detailed in Appendix B).

Figure 7-3. Restriction digest of pENTR-shANGPT1. EcoRI/SalI double restriction digest to excise the shRNAmir30 sequence including the new shANGPT1 sequence. The control digest of pENTR-non-targeting control (pENTR-NTC, far right lane) shows a band larger than the predicted 382bp, but equal to the bands from clones 1 and 2. Clones 3-6 have a larger band (predicted to be 86bp larger). Marker is Fermentas 100-3000bp marker.

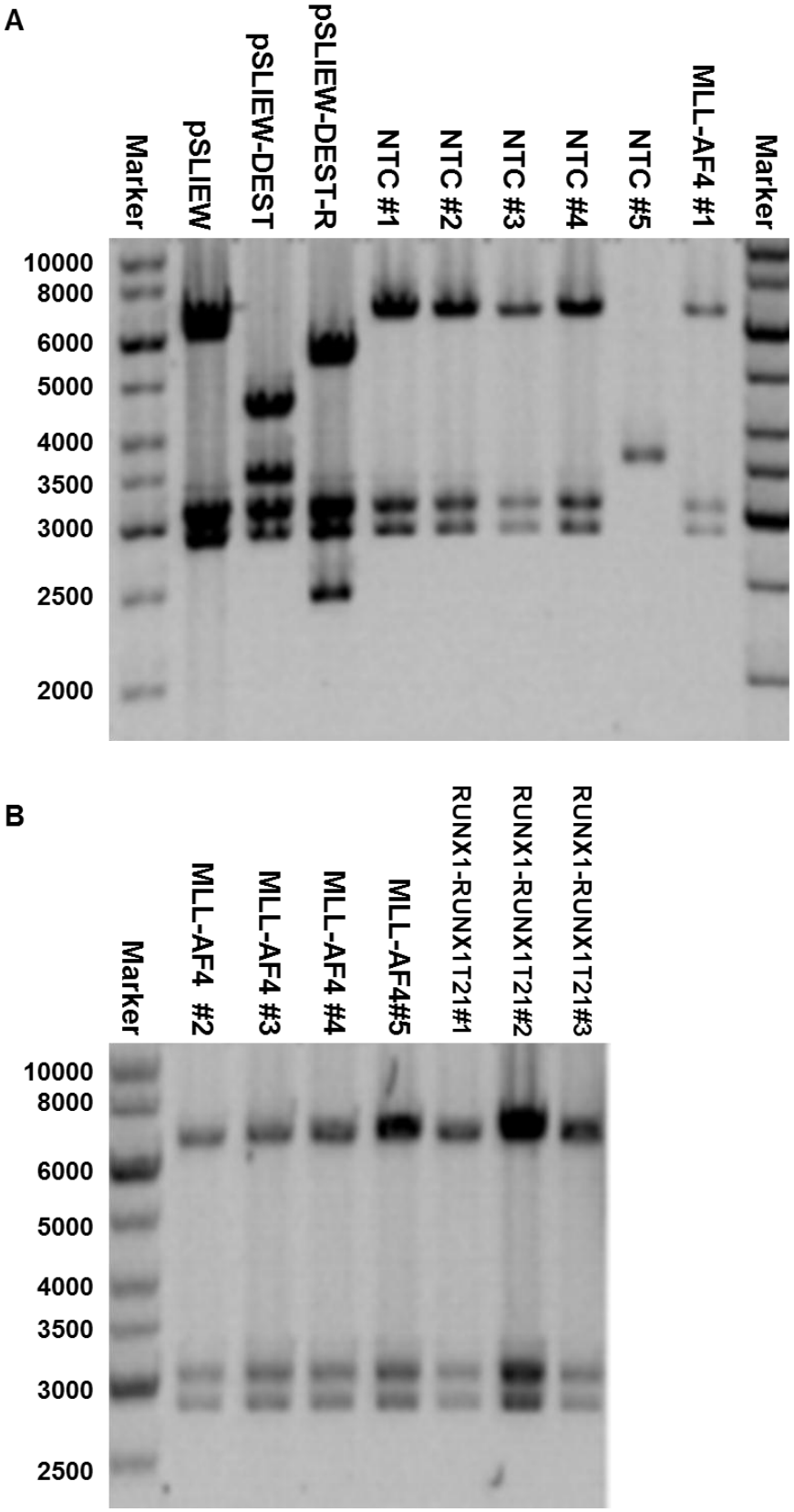


7.3.4 Recombination of pSLIEW-Destination and pENTR-shRNAmir

Successfully sequenced vectors were recombined with pSLIEW-DEST cloned in 7.3.1. Recombination products were transformed into DH5α cells and incubated overnight. Colonies were picked and grown further overnight. Plasmid DNA was extracted from overnight cultures and digested using PvuII. Successful recombination removed the PvuII restriction site within the Gateway destination cassette so that the digest pattern returns to approximately that seen in the native pSLIEW vector (Figure 7-4).

Successfully recombined clones were expanded and used for production of lentivirus for subsequent validation experiments.

Figure 7-4. Restriction digests of recombined pSLMIEW vectors. PvuII digests showing the same pattern of bands in pSLIEW and successfully recombined clones (expected bands at 2849bp, 3064bp and 6532bp) of: A) Non-target control (NTC) clones 1-4 and shMLL/AF4 clone 1. NTC clone 5 shows an incorrect digest pattern; B) shMLL/AF4 clones 2-5 and RUNX1/RUNX1T1 clones 1-3.



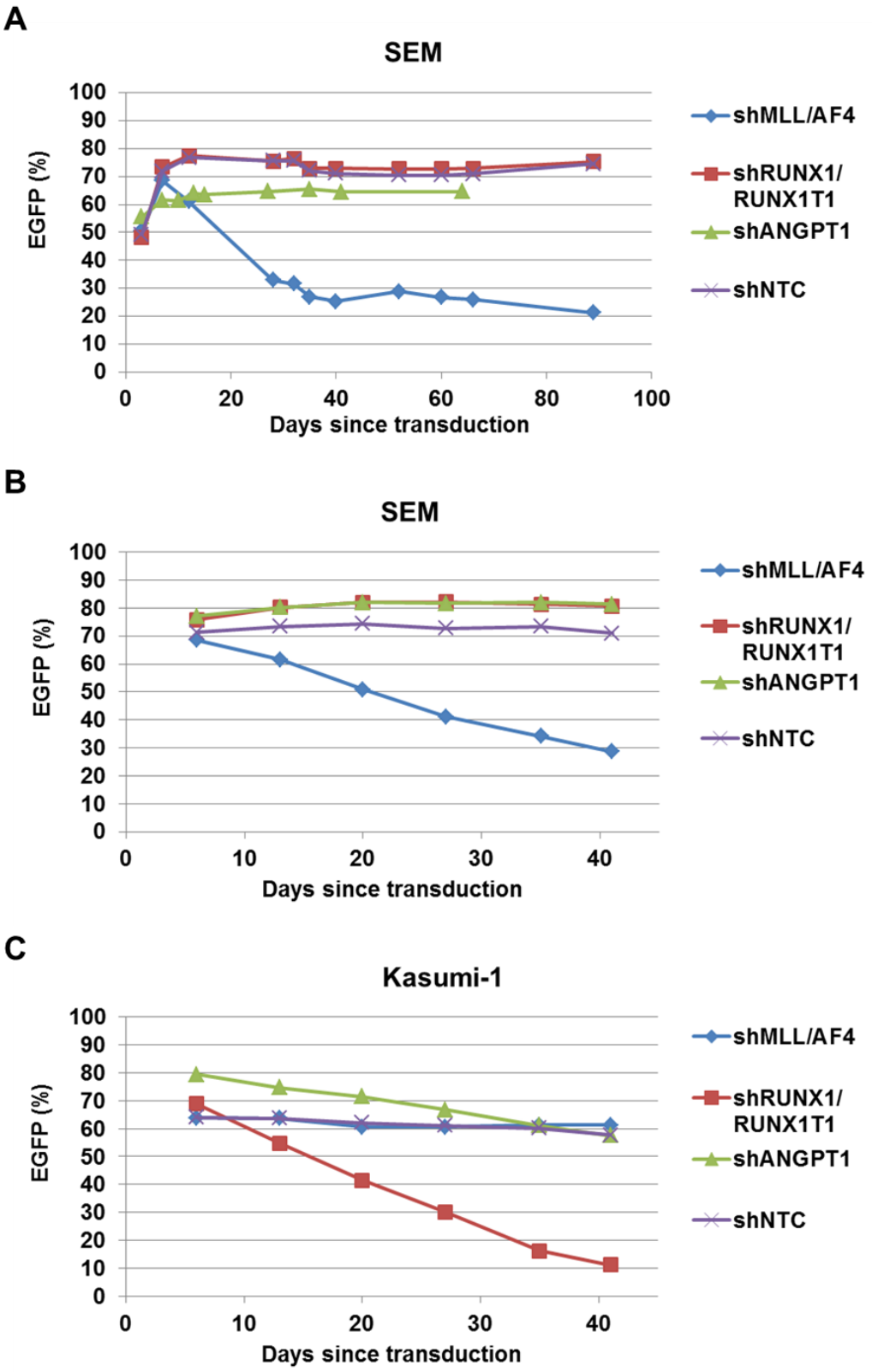
7.3.5 Validation of pSLMIEW function *in vitro*

Virus was produced using the four pSLMIEW vectors described above. SEM and Kasumi-1 cells were transduced with each of the four viruses, resulting in expression of EGFP. The functional effect of knockdown was assessed *in vitro* using a competitive proliferation assay.

SEM (n=2) and Kasumi-1 (n=1) cells were transduced with each of the constructs at approximately equal rates. Cells were then maintained in suspension culture without sorting. The proportion of cells expressing EGFP was monitored by serial flow cytometric measurements. Knockdown of the relevant fusion transcript (*MLL/AF4* in SEM cells, *RUNX1/RUNX1T1* in Kasumi-1 cells) resulted in a reduction in the expression of EGFP (Figure 7-5), demonstrating a proliferative disadvantage for transduced cells and confirming the on-going importance of the fusion oncogene to leukaemic maintenance *in vitro*. In the first experiment, the expression of EGFP appeared to stabilise in SEM-shMLL/AF4 cells between 20-30% after approximately 40 days (Figure 7-5A). The same plateau was not seen in Kasumi-1-shRUNX1/RUNX1T1 cells (Figure 7-5C).

Kasumi-1 cells transduced with shANGPT1 also showed a proliferative disadvantage compared with control cells, but this was not as marked as in the Kasumi-1 cells transduced with shRUNX1/RUNX1T1 (Figure 7-5C). Surprisingly, given that *ANGPT1* is up-regulated by the *MLL/AF4* fusion product (Figure 7-6C), SEM cells transduced with shANGPT1 did not show a proliferative disadvantage (Figure 7-5A-B). The reason for this remains unclear but is the subject of on-going experiments within the laboratory.

Figure 7-5. Functional assessment of SLMIEW constructs *in vitro*. A-B) Two independent transductions of SEM cells demonstrate a proliferative disadvantage to shMLL/AF4 transduced cells. C) Transduction of Kasumi-1 cells shows proliferative disadvantage to shRUNX1/RUNX1T1 and shANGPT1 transduced cells.



Production of shRNA and target gene knockdown over time were confirmed by qRT PCR. Expression of the hairpin constructs was directly assessed by qRT PCR 13 days following transduction. Strong expression of both shRUNX1/RUNX1T1 and shANGPT1 was seen in SEM and Kasumi-1 cells (Figure 7-6A). In contrast, shMLL/AF4 was expressed at a substantially lower, although easily detectable, level (Figure 7-6A).

Despite differential expression of shRNA, effective knockdown was achieved for each of the three target transcripts. *MLL/AF4* knockdown, assessed in SEM cells, showed a maximal knockdown on day 6 (Figure 7-6B). At this time point a knockdown of 48% was seen in a mixed population with 68% EGFP+ cells. The knockdown, adjusted for the rate of transduction, was therefore 71%. As the proportion of transduced cells decreases throughout the experiment, so the measured knockdown is seen to reduce (Figure 7-6B).

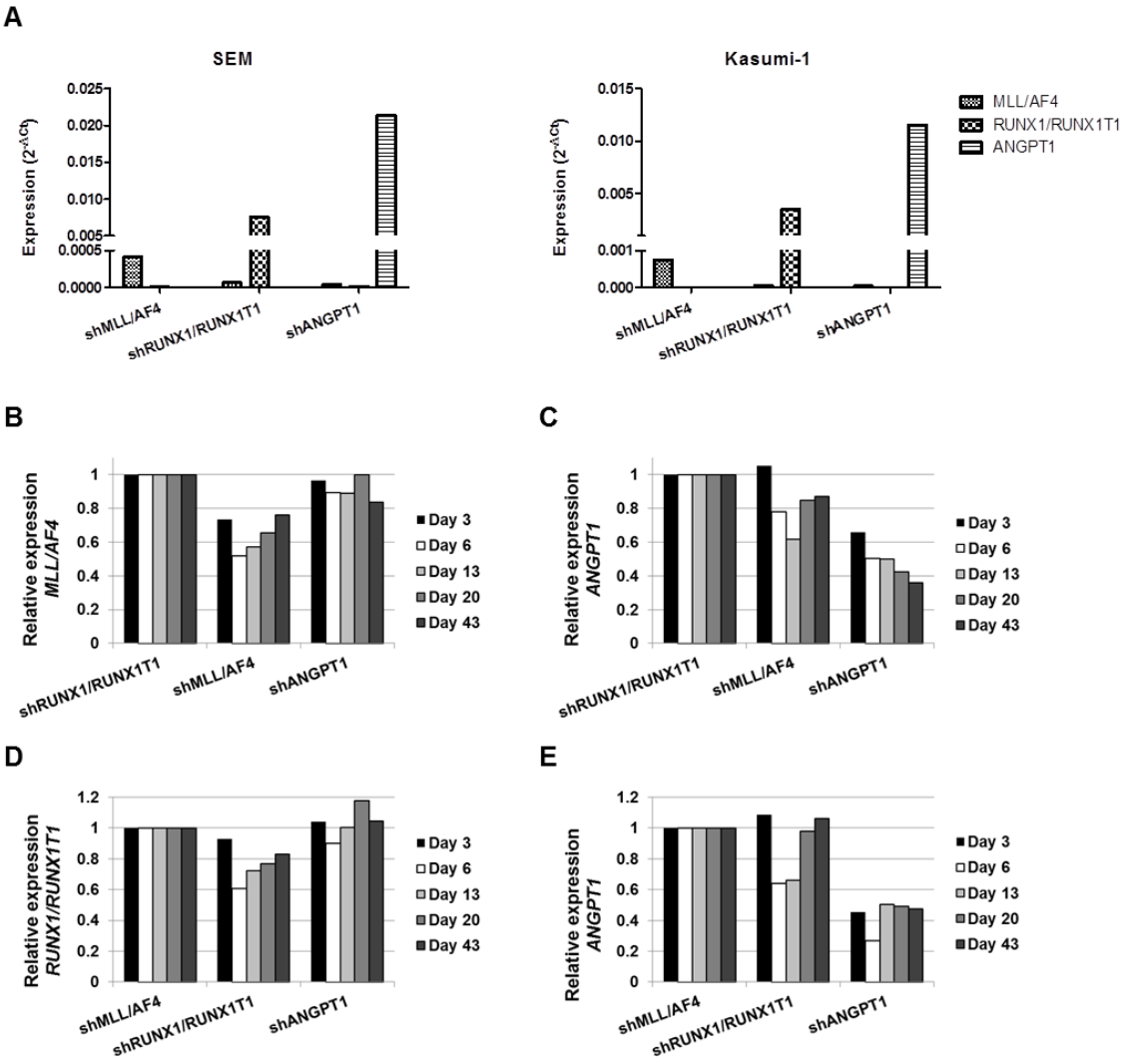
shRUNX1/RUNX1T1 produced a slightly weaker knockdown, as assessed in Kasumi-1 cells (Figure 7-6D). Here, a maximal knockdown of 40% was again seen at day 6. As only 69% of the population expressed EGFP on this day, the knockdown was adjusted to 58% for the rate of transduction. As described for SEM cells, transduction of Kasumi-1 cells with the relevant fusion specific shRNA construct (shRUNX1/RUNX1T1) resulted in a proliferative disadvantage and so the knockdown is seen to decrease as the proportion of transduced cells decreases (Figure 7-6D).

Finally, *ANGPT1* knockdown was assessed in both SEM (Figure 7-6C) and Kasumi-1 (Figure 7-6E) cells and peaked on day 6 in Kasumi-1, consistent with the proliferative disadvantage in this population. As SEM cells showed no proliferative disadvantage, knockdown continued to increase through to the end of the experiment on day 43. In 79% (Kasumi-1) and 81% (SEM) transduced populations, knockdown was 74% and 64% respectively, adjusting to 94% knockdown in Kasumi-1 cells and 79% in SEM cells.

Analysis of *ANGPT1* also identified the reduction of expression of this transcript associated with knockdown of the respective fusion oncogene. In both SEM cells transduced with shMLL/AF4 (Figure 7-6C) and Kasumi-1 cells transduced with shRUNX1RUNX1T1 (Figure 7-6E), expression of *ANGPT1* reduces initially

(Day 6-13) before rising again as the proportion of transduced cells in each of these populations decreases. This is in keeping with other studies within our laboratory which have demonstrated the control of *ANGPT1* expression by the *MLL/AF4* and *RUNX1/RUNX1T1* fusion oncogenes.

Figure 7-6. *In vitro* expression and knockdown by SLMIEW constructs. A) Expression of shRNA by SEM cells (left panel) and Kasumi-1 cells (right panel). Expression is presented as $2^{-\Delta Ct}$, relative to the control small nuclear RNA, RNU6-2. Groups along the X-axis represent the transduced SLMIEW construct. Minor columns adjacent to the column of shRNA expression expected from an individual construct represent non-specific background signal. B) Knockdown of *MLL/AF4* in SEM cells. C) Knockdown of *ANGPT1* in SEM cells. D) Knockdown of *RUNX1/RUNX1T1* in Kasumi-1 cells. E) Knockdown of *ANGPT1* in Kasumi-1 cells. Relative expression is ΔΔCt using *GAPDH* as the reference transcript and normalising against expression in shRUNX1/RUNX1T1 transduced cells (B-C) or shMLL/AF4 transduced cells (D-E).



The final element of *in vitro* validation was to analyse the luciferase activity from each of the constructs. As the shRNA construct is sited immediately 3' to the Luciferase gene, processing of the single transcript by endogenous RNAi pathway proteins (Drosha/Dicer) would reduce the proportion of transcript available for translation into Luciferase protein. As some shRNA constructs may be more efficiently processed than others, it was important to investigate how this affected Luciferase activity.

Transduced SEM and Kasumi-1 cells were adjusted for the percentage expression of EGFP and 10^5 EGFP positive cells pipetted into triplicate wells. The total volume was made up to 175 µl. D-Luciferin, 0.075 µg in 25 µl, was added to each well and the plate analysed using a Fluostar Omega fluorescence plate reader, 1 second per well, at different time intervals.

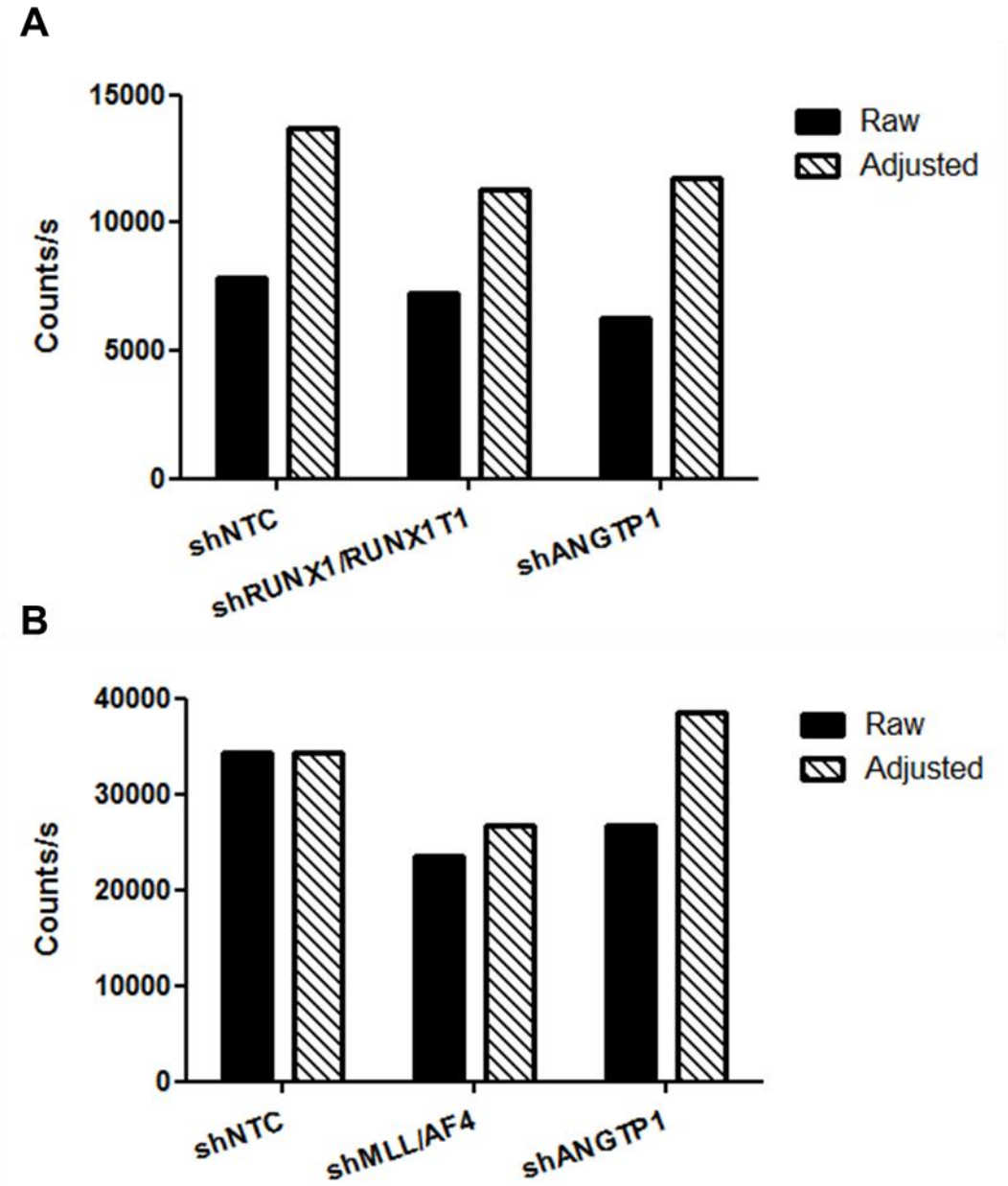
Within each of the two cells lines tested, luciferase activity demonstrated a moderate variation. This did not change with the interval time to reading, three minutes giving the highest counts. In SEM cells, the non-target control construct had 25% higher luciferase activity than the lowest activity construct, shANGPT1 (Figure 7-7A). Adjusting the luciferase activity for the geometric mean of EGFP expression, used as a measure of expression intensity, resulted in a small reduction in this difference to 16% (Figure 7-7A).

In Kasumi-1 cells, the greatest luciferase activity was again in the non-target control construct (Figure 7-7B). There was a 46% difference between the non-target construct and the shMLL/AF4 construct which showed the lowest luciferase activity. In Kasumi-1 cells, these differences showed a more marked correction with adjustment for the geometric mean of EGFP expression.

Most strikingly however, was the difference between SEM cells and Kasumi-1 cells. In Kasumi-1 cells the luciferase activity of the non-target construct was over 4 times that in SEM cells. Even having adjusted for the EGFP geometric mean, the activity in Kasumi-1 cells was still 2.5 fold that in SEM cells.

These preliminary results suggest that the luciferase activity will need to be assessed for any new shRNA constructs created and, potentially, adjustments made accordingly to data derived from *in vivo* imaging studies.

Figure 7-7. Luciferase activity of SLMIEW constructs *in vivo*. The raw and adjusted (normalising data by comparing the geometric mean of individual cells/constructs against the geometric mean of Kasumi-1 –shNTC) luciferase activity for A) SEM and B) Kasumi-1 cells.



7.3.6 Xenotransplantation of SEM-SLMIEW constructs

In order to test the effectiveness of the single vector in a xenotransplantation experiment, SEM cells were transduced with SLMIEW virus containing either shMLL/AF4, shRUNX1/RUNX1T1 (negative control) or shANGTP1 constructs. The shANGPT1 construct was chosen as data from the laboratory have shown high expression of *ANGPT1* in *MLL* rearranged leukaemias as well as control of the expression by the *MLL/AF4* fusion product (Figure 7-6C).

An estimation of the volume of inoculum required to give an MOI of 0.8 was made by putting a line of best fit onto the standard curves produced from a viral titration experiment. SEM cells were transduced using the standard transduction protocol and washed on subsequent days using the stringent washing procedure described in 2.3.11.3. On day 4 following transduction SEM-SLMIEW cells were FACS sorted at the Newcastle University Flow Cytometry Core Facility, Centre for Life. Cell sorts achieved a 94% efficiency. Cells were incubated overnight in RPMI1640, 20% v/v FCS and 1% v/v penicillin/streptomycin mixture. The following day, cells were counted and resuspended. This showed that a substantial proportion of cells were no longer viable. However, 3×10^4 viable cells were injected into each of six mice (3 female, 3 male) per construct. Engraftment was monitored by serial bioluminescent imaging.

Unfortunately, the mouse colony suffered from a problem (presumed to be an undiagnosed infection) during this experiment, with a large number of both transplanted and stock mice developing distended abdomens and dying. In total this problem killed 4 shRUNX1/RUNX1T1 negative control mice and 2 shANGPT1 mice. None of the shMLL/AF4 mice were affected.

In the shMLL/AF4 arm, 3/6 mice engrafted (mice 51, 52 and 53). Engraftment of transduced cells showed a delayed dissemination away from the right femoral site of injection and a weaker overall engraftment, even at week 8 (Figure 7-8A, Figure 7-9A), one week after most other mice were killed. The two surviving shRUNX1/RUNX1T1 negative control mice showed different degrees of engraftment, with mouse 62 showing a stronger luciferase signal

than mouse 60. However, both of these mice showed widespread dissemination to the contralateral femur, vertebrae, spleen and skull/meninges (Figure 7-8B, Figure 7-9B). Mice transplanted with SEM-SLMIEW cells containing the shANGPT1 construct also showed incomplete engraftment. Of the 4 surviving mice, engraftment was seen in 3 (mice 64, 65 and 66). Mouse 63 also engrafted but then died during week 5. The 3 surviving engrafted mice also showed widespread dissemination, similar to the negative control group (Figure 7-8C, Figure 7-9C).

Figure 7-8. *In vivo* bioluminescent imaging of NSG mice transplanted with SLMIEW transduced SEM cells. A) shMLL/AF4 mice (week 8) show reduced dissemination and lower engraftment of transduced cells compared to B) negative control shRUNX1/RUNX1T1 (week 7) mice or C) shANGPT1 mice (week 7). Signal scales are in photons/s/cm^2/sr.

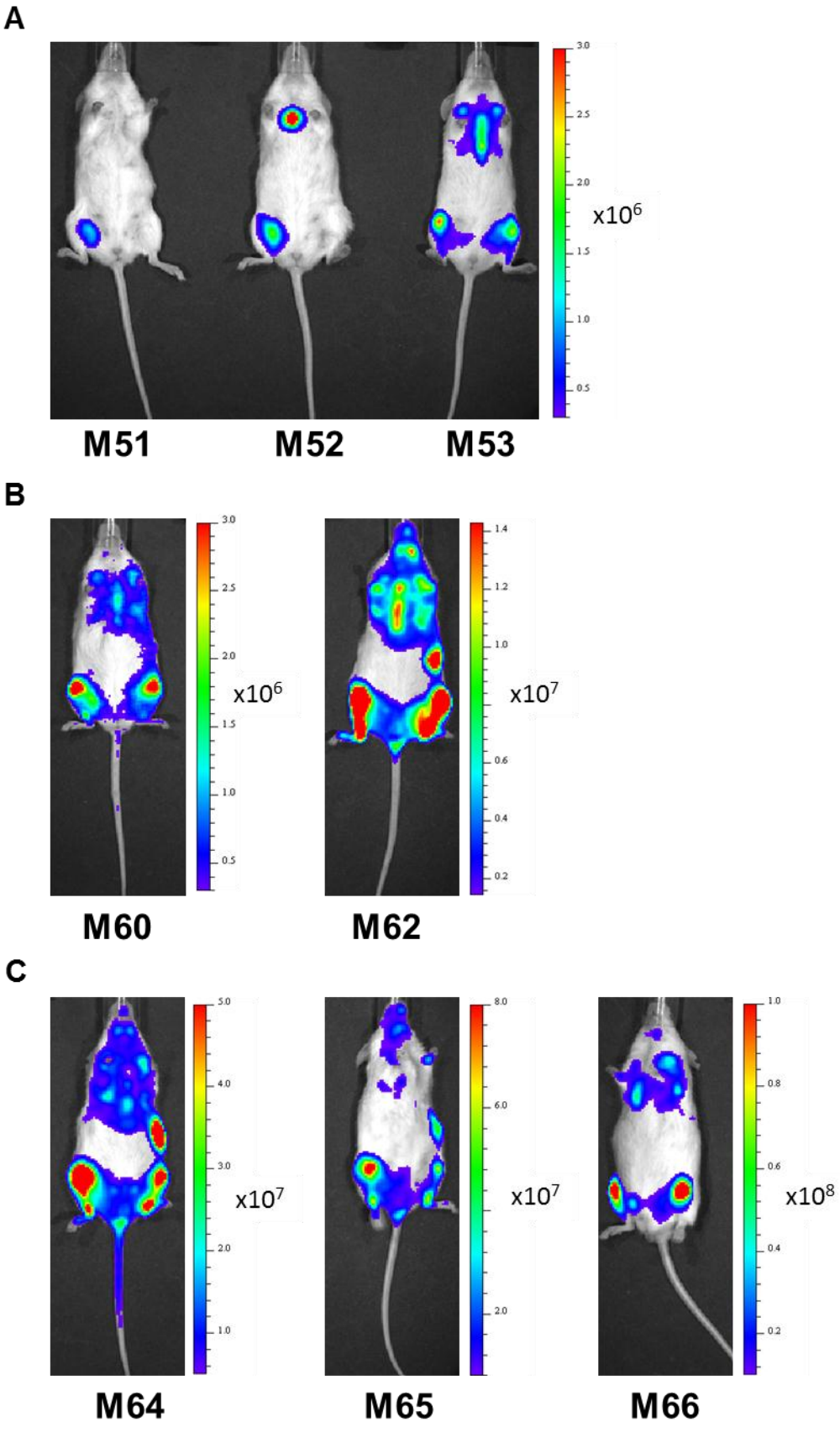
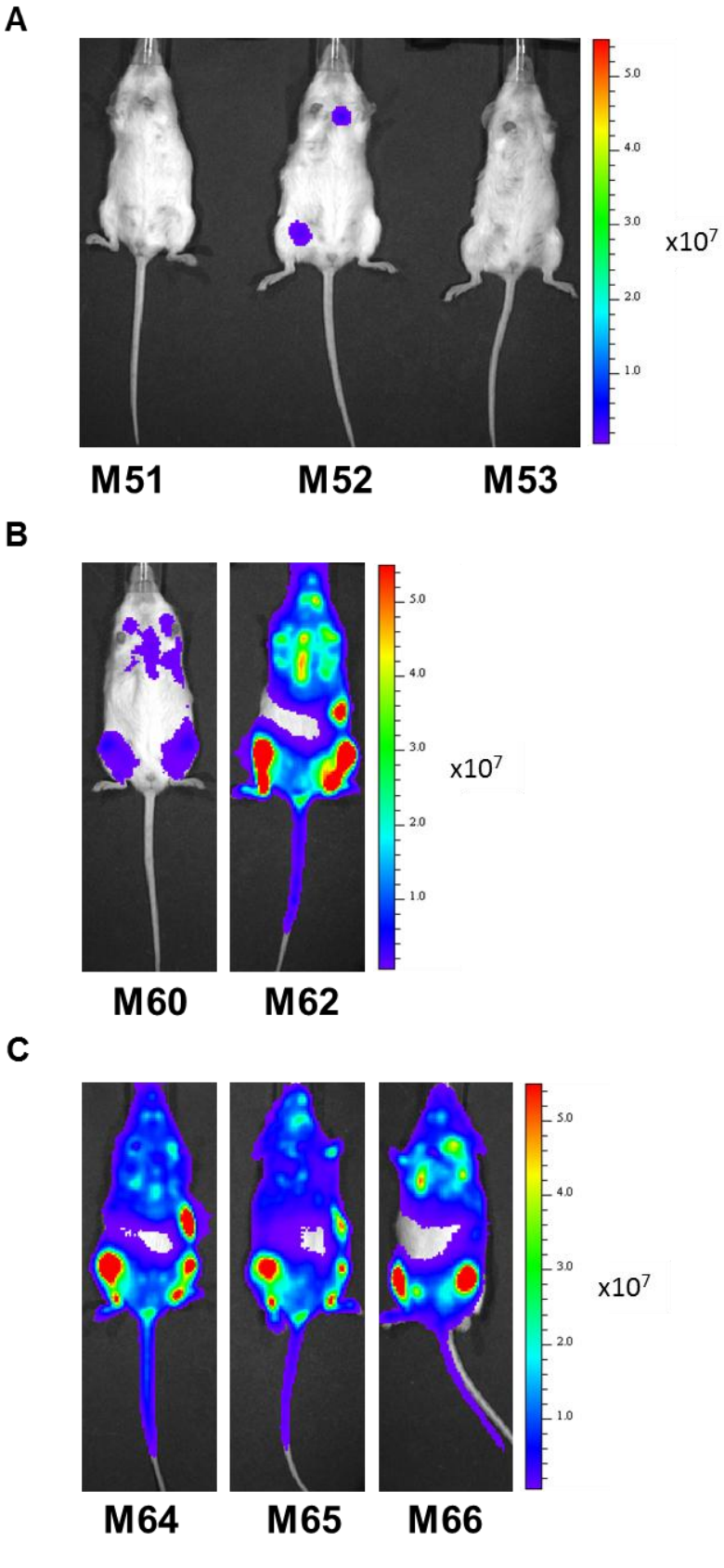


Figure 7-9. *In vivo* bioluminescent imaging of NSG mice transplanted with SLMIEW transduced SEM cells. Mice are demonstrated in the same order as in **Figure 7-8** but with all images taken at week 7 and identical scales provided for comparison. A) shMLL/AF4 mice. B) negative control shRUNX1/RUNX1T1 mice or C) shANGPT1 mice. Signal scales are in photons/s/cm^2/sr.



Analysis of the total luminescence of each mouse demonstrated a reduced signal in the shMLL/AF4 group compared with either the negative control or shANGPT1 group (Figure 7-10A), consistent with the images seen in Figure 7-8. There was no difference between the shANGTP1 group and the negative control shRUNX1/RUNX1T1 group (Figure 7-8, Figure 7-9).

Mice were killed and harvested on week 7 (shRUNX1/RUNX1T2 mouse 62 and shANGPT1 mice 64, 65, 66 and 68) or week 8 (shMLL/AF4 mice 51-56 and shRUNX1/RUNX1T1 mouse 60). The details of engrafted mice are contained in Table 7-1. Bone marrow and splenic material was harvested from all engrafted mice and analysed by multicolour flow cytometry. There was a close correlation between the expression of EGFP in the bone marrow and that in the spleen (Figure 7-10B). SEM cells were isolated from bone marrow by CD133 MacsBead sorting and analysed for expression of *MLL/AF4* and *ANGPT1*.

Two of three mice engrafted with shMLL/AF4 transduced SEM cells had the lowest expression of EGFP, demonstrating a proliferative disadvantage to these cells. The third mouse in this group, mouse 52, showed higher proportions of EGFP expression (bone marrow 86%, spleen 67%). Overall, the expression of EGFP in harvested samples correlates well with the intensity of the bioluminescent signal produced with *in vivo* monitoring (Figure 7-10C).

The proliferative disadvantage of shMLL/AF4 transduced SEM cells is corroborated by the absence of *MLL/AF4* knockdown in re-isolated SEM cells (Figure 7-11A). In contrast, knockdown of *ANGPT1* persisted in re-isolated shANGPT1 transduced cells, albeit with some variation (Figure 7-11B). shRNAmir expression (Figure 7-11C) is similar to that seen *in vitro* (Figure 7-6A).

Figure 7-10. Analysis of luminescence of engrafted mice. A) Total luminescence from mice surviving in each of the three groups. Negative control (shRUNX1/RUNX1T1) and experimental (shANGPT1) groups show more rapid and greater engraftment of transduced cells than the positive control group (shMLL/AF4). B) Bone marrow and splenic expression of EGFP correlate closely. C) End-point total luminescence correlates with expression of EGFP in the bone marrow at harvest. Arrow shows samples from mouse 52

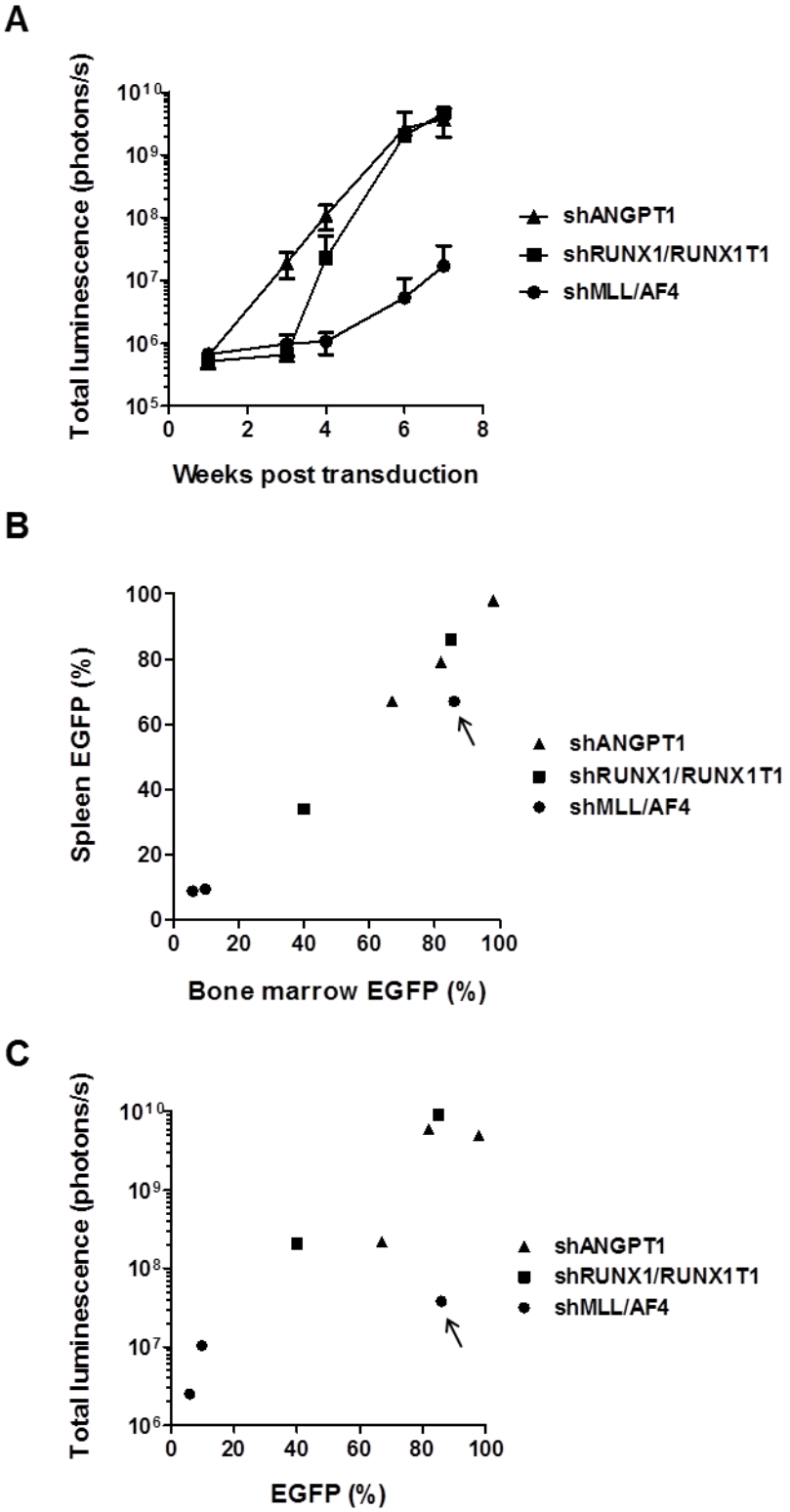
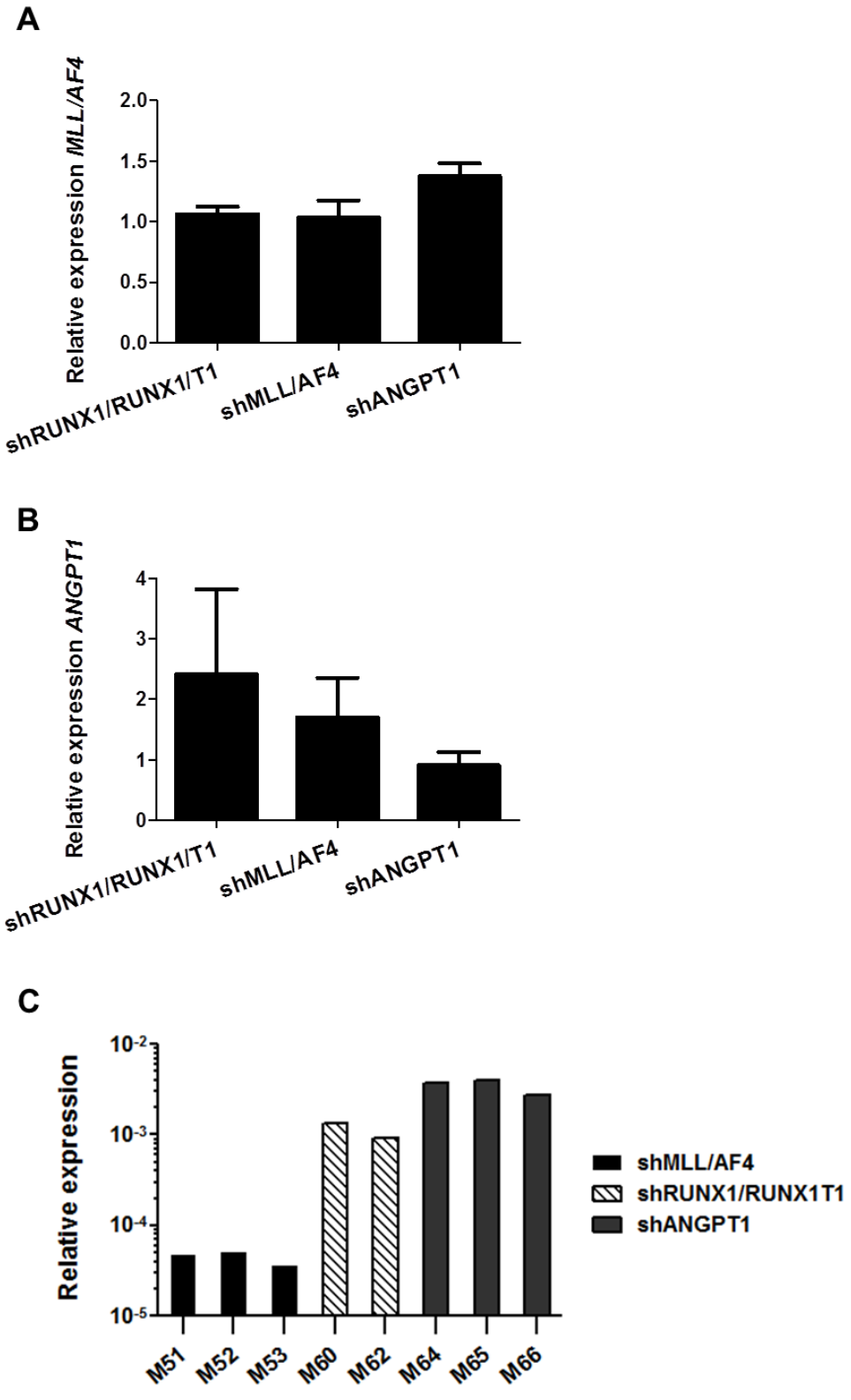


Table 7-1. Details of mice transplanted with SLMIEW transduced SEM cells. BM – bone marrow. EGFP – Enhanced green fluorescent protein. %EGFP expression is percentage of CD19+ cells expressing EGFP. Mice marked – in the engraftment column died before demonstrating engraftment. Mouse 63 engrafted but was found dead on week 5.

	Mouse	Engrafted	BM EGFP	Spleen (g)	Spleen EGFP
shMLL/AF4	51	Y	5.8%	0.06	8.8%
	52	Y	86%	0.05	67%
	53	Y	9.7%	0.08	9.4%
	54	N	-	0.04	-
	55	N	-	0.04	-
	56	N	-	0.05	-
shRUNX1/RUNX1T1	57	-	-	-	-
	58	-	-	-	-
	59	-	-	-	-
	60	Y	40%	0.16	34%
	61	-	-	-	-
	62	Y	85%	0.13	86%
shANGPT1	63	Y	-	-	-
	64	Y	82%	0.29	79%
	65	Y	67%	0.5	67%
	66	Y	98%	0.05	98%
	67	-	-	-	-
	68	N		-	

Figure 7-11. Knockdown and expression of shRNAmir30 constructs in SEM cells re-isolated from murine bone marrow. A) Expression of *MLL/AF4*. B) Expression of *ANGPT1*. Columns represent mean and standard error of the mean of mice transplanted with SEM cells transduced with the each of the three shRNAmir constructs. Relative expression is ΔΔCt compared to mouse 62 (shRUNX1/RUNX1T1) using GAPDH as the reference transcript. C) Expression of shRNAmir constructs in re-isolated SEM cells. Expression is $2^{-\Delta Ct}$ using RNU6-2 as a reference transcript.



7.4 Discussion

The aim of this project was to construct a lentiviral transfer vector which would allow *in vitro* analysis/sorting, *in vivo* monitoring and modulation of target genes in a single transduction. Such a vector would allow for functional analysis of candidate genes using patient-derived material in xenotransplantation assays. The cloning strategy was designed to allow new short hairpin constructs to be easily cloned into a microRNA backbone which could then be inserted into a transfer vector. This concept has been proven to work using an shRNA construct specific for *ANGTP1*.

Both of the original shRNAmir constructs derived from pGIPZ/pTRIPZ vectors (shMLL/AF4, shRUNX1/RUNX1T1) and the newly cloned shANGPT1 construct have been shown to express shRNA and produce effective knockdown. There was some variation in the knockdown efficacy, but this would be expected in any experimental RNAi context and represents a sequence specific phenomenon. Due to the restrictions placed on the design of fusion transcript specific hairpins by the very limited sequence availability, it is not surprising that these two constructs gave the weakest knockdown, whilst the shANGPT1 construct achieved the best knockdown at 79-94%. The variation seen in knockdown was not as great as the variation seen in shRNA expression, suggesting that perhaps the shRNA qRT PCR was subject to sequence specific differences, giving a falsely low analysis of shMLL/AF4 expression.

Having successfully demonstrated the knockdown achieved by the new constructs, the continued functioning of the original reporter genes was assessed, as it was possible that processing of the miRNA transcripts might differentially affect translation of either the fLuc or EGFP transcripts. As the basic measure used to quantitate the efficiency of transduction, EGFP expression has been preserved. Further validation of the efficiency of EGFP expression is difficult, as any analysis would be dependent on the rate of transduction. Luciferase activity, however, can easily be compared between populations with known rates of transduction, assuming constant EGFP expression. Analysis of luciferase activity demonstrated variation between constructs *in vitro.* However, there was a good correlation between the

proportion of EGFP positive cells and the luciferase signal obtained by *in vivo* bioluminescent imaging of mice engrafted with SEM cells transduced with the three SLMIEW constructs. Whilst new constructs will need to be analysed to determine Luciferase activity, in principle the *in vitro* validation of the new vector system was successful.

In an initial experiment to demonstrate the value of the single vector in xenograft assays, the cell line SEM was transduced with three SLMIEW constructs – shMLL/AF4, shANGPT1 and shRUNX1/RUNX1T1 (negative control). Unfortunately, this experiment was significantly disrupted by a presumed infection within the NSG colony and insufficient time was available to repeat it within this Fellowship. Despite this, SEM cells transduced with the shMLL/AF4 construct demonstrated substantially slower and less extensive engraftment than control cells (shRUNX1/RUNX1T1). This, combined with the substantial drop in EGFP expression in 2/3 engrafted populations and the loss of *MLL/AF4* knockdown in re-isolated SEM cells demonstrates for the first time the role *MLL/AF4* expression in leukaemic maintenance *in vivo*. This result is consistent with the importance of on-going expression of *MLL/AF4* in leukaemic cells *in vitro* demonstrated both in the present study and in previous work from our laboratory *(Thomas, Gessner et al. 2005)*. Furthermore, this result not only demonstrates the effectiveness of bioluminescent monitoring in this assay, but also supports the value of tracking proportional EGFP expression as an experimental endpoint.

The rapid engraftment of the two surviving mice in the shRUNX1/RUNX1T1 negative control group again demonstrates the effectiveness of this approach. Optimising the consistency within this negative control group will be important. It is hoped that this can be achieved by improving the viability of cells at transplantation.

Despite the up-regulation of *ANGPT1* by the *MLL/AF4* fusion product present in SEM cells, the shANGTP1 construct did not produce a consistent phenotype in this xenotransplantation experiment. Expression of the shANGPT1 construct remained strong following re-isolation. Whilst the *ANGPT1* knockdown data showed some variation, when combined with the strong shANGPT1 expression it seems unlikely that poor knockdown was the cause of this result. One

alternative was that the experimental design itself may be the cause of the negative finding. It is possible that whatever the function of *ANGPT1* in primary leukaemias, it is not required in the context of a derived cell line or within a murine microenvironment. Repeating the experiment using patient-derived material would investigate the first possibility but addressing the second would require a substantially different approach. An alternative explanation is that any advantage offered by *ANGPT1* may only be uncovered in a competitive situation. This possibility would be addressed by moving to a lower multiplicity of infection and transplanting without FACS sorting. Finally, it is possible, especially given the lack of effect from *ANGPT1* knockdown *in vitro* that this is a true negative finding.

In addition to demonstrating the importance of *MLL/AF4* in leukaemic engraftment/maintenance *in vivo*, this experiment also provided a substantial amount of information on the functioning of this vector and potential improvements to methodological design. Sorting the transduced cells resulted in substantial cell death which presumably reduced the engrafting cell dose below that which was transplanted. Any difference in cell viability between populations would provide a source of bias to the timing or likelihood of engraftment. A preferable approach would be to transduce all populations to a known/predicted rate, based on lentiviral titration, and then monitor the change in EGFP expression as one outcome. This has the added advantage of providing a competitive situation, which may demonstrate more subtle differences in engraftment potential than engraftment of a sorted population.

In summary, this project has successfully produced a novel vector offering three functions from a single transduction. In combination with the modified transduction protocol described in Chapter 6, this will offer the opportunity to perform *in vivo* tracking of gene modulation studies using patient-derived material. When considering the difficulties experienced during the xenotransplantation experiment described in 7.3.6, in combination with published concerns regarding the toxicity associated with high levels of exogenous shRNA (Grimm, Streetz et al. 2006; Pan, de Ruiter et al. 2011), moving towards an approach with a lower multiplicity of infection (e.g. 0.3-0.4) would offer a number of advantages. Firstly, it would reduce potential toxicity

from high levels of shRNA. Secondly, transplanting unsorted populations will improve cell viability and consistency between populations. Finally, introducing a competitive element to the experiment may identify more subtle advantages than are seen in a straightforward engraftment situation. These changes will be introduced with the next experiments using this vector system.

Chapter 8
General Discussion

The identification of leukaemia propagating cells in ALL has been complicated by the influence of xenograft assay variables (mouse strain, route of injection, antibodies used for cell sorting and mouse gender) on the engraftment of patient-derived blasts (Vormoor 2009; McDermott, Eppert et al. 2010; Taussig, Vargaftig et al. 2010). Recent studies using the most permissive models have shown engraftment with low numbers of primary or primografted lymphoblasts (Morisot, Wayne et al. 2010; Schmitz, Breithaupt et al. 2011), without enrichment being identified in any specific immunophenotypic population (Kong, Yoshida et al. 2008; le Viseur, Hotfilder et al. 2008). Remaining questions in the determination of LPC identity in lymphoblastic leukaemia have recently been addressed by our own laboratory, in the most comprehensive study to date, transplanting sorted primary blasts at limiting dilution (Rehe, Wilson et al. 2012)(Appendix D). Together, these data demonstrate that leukaemia propagation in lymphoblastic leukaemia cannot be enriched for by immunophenotypic sorting of B-cell precursor markers and is best described by a stochastic model.

Whilst cancer stem cells have been enriched from a number of malignancies, examples exist where tumour propagating capacity is common and cannot be enriched, notably melanoma (Quintana, Shackleton et al. 2008; Quintana, Shackleton et al. 2010) and lymphoma (Kelly, Dakic et al. 2007). Melanoma excepted, in the majority of malignancies investigated, a primitive looking population has been identified which contains the bulk of the tumour propagating capacity. The unusually diverse nature of propagating cells in B precursor ALL (Kong, Yoshida et al. 2008; le Viseur, Hotfilder et al. 2008) led us to consider what might underlie this feature of lymphoid malignancy.

Lymphoid cells are, perhaps, the only lineage of cells in which maturation is not associated with a loss of proliferative capacity. Indeed, it is a central feature of lymphoid development that both B cell precursors and germinal centre B cells are required to clonally expand. For B cell precursors a round of clonal expansion occurs following successful rearrangement of the B cell receptor heavy chain, producing a clone able to rearrange the light chain genes in the hope of creating a functional B cell receptor. In the germinal centre, somatic hypermutation and consequent affinity based selection are accompanied by

several rounds of clonal expansion to produce both high affinity antibody producing plasma cells and memory B cells (Victora and Nussenzweig 2012). Subsequent re-exposure to antigen is accompanied by further clonal expansion of memory B cells and further rounds of affinity maturation. It therefore seems plausible that the unusually broad capacity for the propagation of both leukaemias (Kong, Yoshida et al. 2008; le Viseur, Hotfilder et al. 2008; Morisot, Wayne et al. 2010; Schmitz, Breithaupt et al. 2011) and lymphomas (Kelly, Dakic et al. 2007) may reflect a property inherent to the originating cell lineage.

The aim of this project was to investigate the genetic programmes underlying the leukaemia propagating capacity of childhood ALL. The hypothesis was that "acute lymphoblastic leukaemic blasts derive an intrinsic propagating potential from their lymphoid cell of origin". However, little is known about the genetics driving lymphoid clonal expansion. The best studied element is the transcriptional repressor *BCL6* which is characteristically expressed in germinal centre B cells (Victora and Nussenzweig 2012). *BCL6* is also expressed following successful V_H-DJ_H rearrangement, in response to preB receptor signalling, where the repression of Arf and p53 prevents apoptosis in response to Ig light chain gene single strand breaks (Duy, Yu et al. 2010). In ALL, *BCL6* has been seen to be highly expressed in Down syndrome associated ALL (Hertzberg, Vendramini et al. 2010) and is regulated by the BCR/ABL1 fusion product where experimental *BCL6* inhibition results in re-sensitisation of previously resistant leukaemia propagating cells to tyrosine kinase inhibitors (Duy, Hurtz et al. 2011).

Given the relatively limited understanding of the control of lymphoid clonal expansion, this PhD took two new approaches to identifying genes relevant to leukaemic propagation. Firstly, a candidate approach was designed to investigate the role of the germline stem cell maintenance associated gene, *PIWIL2,* which had previously been identified in a number of malignancies including leukaemias. Secondly, an unbiased approach was developed using a genome-wide functional RNAi screen. The third aim was to develop techniques and technologies required for the *in vivo* validation of candidates identified in the earlier components of this PhD.

The first project within this PhD was able to confirm the expression of *PIWIL2* in childhood acute lymphoblastic leukaemia. This was accompanied by the unexpected finding of *PIWIL2* expression in lymphoid populations from healthy donor peripheral blood. This project was further able to demonstrate the functional relevance of *PIWIL2* in two ALL cell lines, SEM (*MLL/AF4*) and 697 (*E2A/PBX1*). However, it was not possible to demonstrate PIWIL2 by Western blot. Serial electroporation, an approach which had been highly successful over a limited number of repetitions (Gessner, Thomas et al. 2010), proved extremely challenging given the duration of *PIWIL2* depletion required to demonstrate a phenotype. Finally, the departure of the project collaborator resulted in the strategic decision to stop this project and shift the focus of the PhD towards the functional RNAi screen. Since this decision, further publications have demonstrated a potential role for this novel pathway in normal and malignant lymphoid cells. Firstly, the demonstration of PIWI/piRNA control of DNA methylation (Aravin, Sachidanandam et al. 2008; Watanabe, Tomizawa et al. 2011; Rajasethupathy, Antonov et al. 2012), a mechanism of action which might plausibly function in the control of lymphoid clonal expansion. Secondly, identification of a bias of piRNA sequence specificities in lymphoid cells away from intergenic/repeat sequences and towards coding regions, introns and untranslated regions (UTRs) (Yan, Hu et al. 2011), makes a role in the control of gene expression feasible. This project will therefore be restarted, employing techniques recently available both within our laboratory (microRNA adapted shRNA knockdown) and commercially (improved antibodies, improved methylation arrays) and will look to identify DNA methylation events targeted in a sequence specific manner by PIWI/piRNA complexes (Chapter 9).

Having stopped the candidate based approach, the focus of this PhD became the unbiased screen for pathways involved in leukaemic propagation. The project was aimed at developing this new technique within the Northern Institute for Cancer Research, as well as investigating the potential for performing a genome-wide screen using a feeder cell co-culture assay. Early work both optimised the transduction procedure and determined the functional relevance of feeder cells to the t(4;11) acute lymphoblastic leukaemic cell line SEM. A trial screen was then devised to further validate the use of feeder cell co-culture

by comparing the results of negative selection screening with and without co-culture replating steps.

Despite relatively consistent results during test transductions, the rate of transduction achieved during the trial screens was approximately half that predicted. This led to the need for an additional early puromycin selection step prior to taking the baseline sample and subsequently for an additional density gradient centrifugation step as the presence of substantial amounts of apoptotic cells resulted in minimal expansion of the remaining viable cells. The reason for the poor transduction is not clear, but comparison of the same protocol using a newly purchased aliquot of pooled viral constructs may help determine whether there has been a loss of titre in the original library virus

Analysis of the first trial screen identified eight candidate genes common to both suspension culture and feeder cell co-culture approaches. Within this "common" group of candidates was *ANGPT1*. This gene has previously been identified within our laboratory, not only as being highly expressed in *MLL* rearranged lymphoblastic leukaemias, but also as being down-regulated following *MLL/AF4* knockdown. There was, however, a group of 41 up-regulated constructs only identified using the feeder cell co-culture approach. Re-analysis of the second trial screen, aimed at achieving successful array hybridisation, will allow analysis of the strength of each of these candidates. This information will then inform the decision on whether the complete screen is conducted using suspension culture or feeder cell co-culture. If there is not strong evidence of a likely benefit to using co-culture then the pragmatic approach, given the experimental difficulties experienced during the trial screens, may be to simplify the screen. This approach may involve use of the more limited "Annotated gene" library containing only 3 pools of 10000 constructs and/or use of suspension culture only, rather than co-culture. The feeder cell co-culture approach could be re-applied during a later, more limited, validation screen.

The final aim of this PhD was to develop techniques necessary for the *in vivo* validation of candidate genes identified by the *in vitro* RNAi screen. This aim had two focuses: firstly, developing a protocol for the lentiviral transduction of patient-derived leukaemic blasts; secondly, cloning of a single vector capable of

in vitro analysis/ cell sorting (EGFP), *in vivo* tracking (Luciferase) and candidate gene modulation (shRNA). Developing the transduction protocol adopted a pragmatic approach. This was necessary both because of the limited availability of patient-derived material and the high titres of virus required to achieve good rates of transduction. The final protocol represents a functioning, practical approach which balances rates of transduction against maintenance of cell viability. Unexpectedly, the persistence of infectious virus up to four days after transduction required a further protocol for the removal of residual virus following transduction in order to both prevent direct transduction of murine tissues and to remove the risk posed by needle stick injury to an experimental operator. This was achieved using a stringent washing procedure. The utility of the transduction protocol has already been demonstrated in preliminary experiments investigating the clonal complexity of engrafting populations (Bomken, Buechler et al. 2012). On-going work on this is being performed within the laboratory by Dr Alex Elder. In addition, one leukaemic sample transduced during this project (L4951) has been transplanted in a further experiment which is using bioluminescent monitoring to determine the efficacy of a novel therapeutic agent. Despite these early applications, further improvements to the protocol will be possible. These may include the use of Retronectin coated plates to reduce spinfection speeds/times further or the trial of additional/alternative supportive cytokines to improve cell viability.

The design and successful cloning of a novel vector allowing *in vitro* analysis, *in vivo* tracking and RNAi from a single transduction will allow the use of patient-derived material for this critical *in vivo* validation step. The design of the vector allows repeated cloning of novel hairpin sequences into the “companion” pENTR-shRNAmir vector, which can then be recombined into the pSLIEW-Destination construct to create a new pSLMIEW vector. This system allows the simple screening of multiple hairpins against a candidate, as well as the validation of multiple candidate genes. Whilst the pSLMIEW vector is equally suitable for derived cell-line transductions, the constitutive expression of the miRNA adapted shRNA will mean that for some experiments it will be preferable to dual transduce – once with a vector allowing sorting/tracking and once with a vector allowing sorting/inducible RNAi.

The aim of this PhD was to determine genetic programmes responsible for the propagation of acute lymphoblastic leukaemia. Whilst this aim has not been fully realised, many of the techniques required to achieve it have now been developed within our laboratory. Furthermore, through the PIWIL2 project, a potential novel component of the clonal expansion programme of normal lymphoid cells has been identified as being functionally relevant in malignant lymphoid cells also. Having developed substantial experience with lentiviral approaches to RNAi and with the publication of several lines of evidence which suggest a lymphoid specific role for PIWI/piRNAs, the PIWIL2 project will now recommence. The on-going validation and final screening using the Decode RNAi library will provide additional leukaemia propagation candidates to investigate further.

In summary, this PhD has functionally validated a novel component of malignant lymphoid cells, *PIWIL2*, which in addition was identified in normal lymphoid cells for the first time. Further candidate elements of the leukaemic propagation programme are yet to be identified by a genome-wide functional RNAi screen. Candidates which are identified by this screen, as well as further elements identified by on-going work into the pathways affected by PIWI/piRNAs, can be validated *in vivo*, with patient-derived material, using the transduction protocol and lentiviral vector described here. Successfully validated candidates can then be assessed for their suitability as novel therapeutic targets. The hope is that drugs specific for targets critical to leukaemic propagation will offer improved rates of cure for resistant or relapsed disease whilst reducing the side-effects associated with current, less well targeted therapies.

Chapter 9

Future work

9.1 Understanding the mechanism of action of *PIWIL2* in lymphoblastic leukaemia propagating cells

Following the strategic decision to stop the *PIWIL2* project detailed in Chapter 3, further published work has developed our understanding of potential mechanisms of action for PIWI-Like genes in leukaemia. As a result of these publications I have developed a succession project to investigate further the role of *PIWIL2* in ALL. In particular this concentrates on the mechanism of action of *PIWIL2* and piRNAs in ALL. I will be Principle Investigator for this project and have led on both finance applications and appointing a research assistant who is due to start the study on 1st January 2013.

In the medium term, a successful outcome to this project will prompt a parallel study in B cell non-Hodgkin lymphoma, my intended future research and clinical sub-specialty. This transition is based on the hypothesis that the clonal expansion programme which is expressed by ALL blasts is common to all B lineage cells able to clonally expand, including peripheral/germinal centre B cells which are believed to be the cell of origin in B cell non-Hodgkin lymphoma.

9.1.1 Novel data relating to PIWI/piRNA function in mammals

The first study which further developed our understanding of the function of PIWI/piRNAs outside of the genome was a study by Tom Tuschl and Eric Kandel's labs (Rajasethupathy, Antonov et al. 2012). They incidentally found piRNAs in small RNA libraries prepared from *Aplysia* (Sea Hare) neurones. They demonstrated that PIWI/piRNA complexes directed CpG methylation, in a sequence specific manner, to the promoter region of the transcriptional regulator *CREB2*, in response to the neuronal modulator serotonin. This was the first demonstration of PIWI/piRNA directed DNA methylation outside of germline cells. This potential mechanism of action was particularly interesting as DNA methylation is important in both B-cell precursor development (Challen, Sun et al. 2012) and in germinal centre B cells (Shaknovich, Cerchietti et al. 2011).

As the project developed, three more studies provided evidence to support further investigation of the mechanism of *PIWIL2* action in ALL/lymphocytes. Firstly, the demonstration of piRNA expression in murine lymphocytes (Kuchen, Resch et al. 2010) and secondly a re-analysis of the same data to show that, unlike other cell types, lymphoid piRNAs showed a high degree of sequence specificity for 5'UTRs, coding regions, introns and 3'UTRs (Yan, Hu et al. 2011). This is in dramatic contrast to piRNAs from murine myeloid lineage cells, which are largely specific for unannotated regions of the genome and piRNAs from *Drosophila*, which show almost exclusive homology to repeat sequences. This change in the balance of piRNA sequences is suggestive of a specific role in lymphoid lineage cells, and therefore potentially in lymphoid malignancies.

Finally, a fourth study demonstrated the production of an antisense transcript from the proximal promoter of KIR3DLI, which codes for one of the killer Ig-like receptors (KIR) expressed on the surface of NK cells, a cell closely related to the lymphoid lineage (Cichocki, Lenvik et al. 2010). This antisense transcript was processed to produce a single piRNA-like small RNA which was responsible for directing CpG methylation of the promoter and suppressing expression of the sense transcript.

Together these studies have: 1) demonstrated a role for PIWI-Like genes and piRNAs in methylation of promoters outside the germline; 2) demonstrated a bias, specifically in lymphoid expressed piRNAs, for sequence homology for promoter and coding regions, rather than repeat sequences; 3) provided an example of a lymphoid related cell using a single piRNA to regulate the methylation and subsequent expression of a key gene.

The project will use a lentiviral approach to knockdown both *PIWIL2* and *PIWIL4* (a ubiquitously expressed PIWI family member capable of histone modification (Sugimoto, Kage et al. 2007)) expression in childhood ALL cell lines, combined with array based analysis of transcriptome and DNA methylation changes. Alongside this, a collaboration with Tom Tuschl, Rockefeller University (New York, USA), will generate piRNA libraries which will be sequenced to identify potential promoter homologies. The intersect of these approaches will provide candidate genes for validation and investigation as potential therapeutic targets.

This project is supported by an MRC Centenary Early Career Award (£20,000). Funding has also been received from the North of England Children's Cancer Research Fund (NECCR, £10,000) and an extension has been granted to use remaining funds from the present MRC Clinical Research Training Fellowship (£12,000). A further award has been made by the Newcastle Healthcare Charity (£14,898).

9.2 Completion of a genome-wide functional RNAi screen in ALL

Following the unsuccessful hybridisation of the second trial screen, barcode sequences will be re-amplified and a repeat hybridisation performed. If successful, this will provide the opportunity to analyse the reproducibility of the screening approaches detailed in Chapter 5, as well as assess the probability that any candidates identified are genuine. This information will assist in the critical decision making around how the final screen is performed. If reproducibility and a substantial benefit (in terms of quality of candidate identified) can be demonstrated using the co-culture approach, then this will be used for the final screen. In this case it may be necessary to reduce the screen to use only pools targeting annotated genes to reduce the complexity of the experimental process (total of 3 pools rather than 7). If sufficient benefit is not demonstrated then use of a suspension culture approach will greatly simplify the screen and the genome-wide library can be used.

Candidates derived from the final screen will be validated using a custom library containing 3-5 constructs targeting each candidate. As approximately 1000 constructs are expected to be returned, the complexity of this validation screen requires that it will be conducted *in vitro*. In order to identify candidates with global importance in childhood ALL, the cell lines REH/t(12;21) and 697/t(1;19) will be used in place of SEM/t(4;11). If the actual number of candidates returned by the initial screen is substantially lower (e.g. 300), this additional *in vitro* step may be omitted and validation proceed immediately to the *in vivo* stage. *In vivo* validation will consist of ALL cell line clones which have been transduced with the pSLIEW vector being further transduced with the relevant inducible pTRIPZ constructs and engrafted in an immunocompromised mouse assay. This will allow *in vivo* disease tracking alongside candidate gene knockdown. In order to select therapeutic candidates which are not critical to normal haematopoiesis, retained hits from the sequential screening process will be analysed for function using haematopoietic precursor cells derived from umbilical cord blood. Candidates which are found to have relevance across the childhood ALL subtypes described, but found not to have a role in normal haematopoiesis will be considered as potential targets for novel drug development. Additional targets may be derived based on known function, e.g.

involvement in B cell receptor signalling pathway. Each of these candidates will be further investigated using patient-derived material transduced with the pSLIEW vector (see Chapter 6) allowing *in vivo* tracking during trials of novel therapeutic agents, combined with a complementary shRNA approach using the pSLMIEW vector (see Chapter 7). This represents a substantial amount of work and will be conducted as part of a wider programme of ALL propagating cell research within the Leukaemia Stem Cell laboratory.

9.3 Summary

The two pieces of on-going work described here will generate new information about the genetic control of malignant lymphoid propagation. Both projects are capable of identifying targets and pathways critical to this process, and potentially identifying targets for novel therapeutic agents designed specifically to reduce the clonal expansion of malignant B lineage cells. These targets may be specific to B-precursor ALL, to mature B cell lymphomas or have therapeutic benefit in both. Ultimately this work is intended to result in improved treatment, with fewer side effects, for children and young people affected by leukaemia or non-Hodgkin lymphoma.

References

Ahlbom, A., N. Day, et al. (2000). "A pooled analysis of magnetic fields and childhood leukaemia." Br J Cancer **83**(5): 692-698.

Al-Hajj, M., M. S. Wicha, et al. (2003). "Prospective identification of tumorigenic breast cancer cells." Proc Natl Acad Sci U S A **100**(7): 3983-3988.

Ali, N., C. Karlsson, et al. (2009). "Forward RNAi screens in primary human hematopoietic stem/progenitor cells." Blood **113**(16): 3690-3695.

Alvero, A. B., R. Chen, et al. (2009). "Molecular phenotyping of human ovarian cancer stem cells unravels the mechanisms for repair and chemoresistance." Cell Cycle **8**(1): 158-166.

Anderson, K., C. Lutz, et al. (2011). "Genetic variegation of clonal architecture and propagating cells in leukaemia." Nature **469**(7330): 356-361.

Aravin, A., D. Gaidatzis, et al. (2006). "A novel class of small RNAs bind to MILI protein in mouse testes." Nature **442**(7099): 203-207.

Aravin, A. A., G. J. Hannon, et al. (2007). "The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race." Science **318**(5851): 761-764.

Aravin, A. A., R. Sachidanandam, et al. (2008). "A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice." Mol Cell **31**(6): 785-799.

Aravin, A. A., R. Sachidanandam, et al. (2007). "Developmentally regulated piRNA clusters implicate MILI in transposon control." Science **316**(5825): 744-747.

Armstrong, G. T., Q. Liu, et al. (2009). "Late mortality among 5-year survivors of childhood cancer: a summary from the Childhood Cancer Survivor Study." J Clin Oncol **27**(14): 2328-2338.

Azorsa, D. O., I. M. Gonzales, et al. (2009). "Synthetic lethal RNAi screening identifies sensitizing targets for gemcitabine therapy in pancreatic cancer." J Transl Med **7**: 43.

Basyuk, E., F. Suavet, et al. (2003). "Human let-7 stem-loop precursors harbor features of RNase III cleavage products." Nucleic Acids Res **31**(22): 6593-6597.

Battista, S., V. Fidanza, et al. (1999). "The expression of a truncated HMGI-C gene induces gigantism associated with lipomatosis." Cancer Res **59**(19): 4793-4797.

Baudet, A., C. Karlsson, et al. (2012). "RNAi screen identifies MAPK14 as a druggable suppressor of human hematopoietic stem cell expansion." Blood **119**(26): 6255-6258.

Betel, D., R. Sheridan, et al. (2007). "Computational analysis of mouse piRNA sequence and biogenesis." PLoS Comput Biol **3**(11): e222.

Blank, U. and S. Karlsson (2011). "The role of Smad signaling in hematopoiesis and translational hematology." Leukemia **25**(9): 1379-1388.

Bohnstedt, C., M. Taskinen, et al. (2009). "Poor treatment compliance in children with down syndrome and acute lymphoblastic leukemia." J Pediatr Hematol Oncol **31**(1): 79-80.

Boiko, A. D., O. V. Razorenova, et al. (2010). "Human melanoma-initiating cells express neural crest nerve growth factor receptor CD271." Nature **466**(7302): 133-137.

Bomken, S., L. Buechler, et al. (2012). "Lentiviral marking of patient-derived acute lymphoblastic leukaemic cells allows in vivo tracking of disease progression." Leukemia.

Bonnet, D. and J. E. Dick (1997). "Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell." Nat Med **3**(7): 730-737.

Brennecke, J., A. Stark, et al. (2005). "Principles of microRNA-target recognition." PLoS Biol **3**(3): e85.

Brown, M. P., T. Nosaka, et al. (1999). "Reconstitution of early lymphoid proliferation and immune function in Jak3-deficient mice by interleukin-3." Blood **94**(6): 1906-1914.

Brummelkamp, T. R., A. W. Fabius, et al. (2006). "An shRNA barcode screen provides insight into cancer cell vulnerability to MDM2 inhibitors." Nat Chem Biol **2**(4): 202-206.

Carmell, M. A., A. Girard, et al. (2007). "MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline." Dev Cell **12**(4): 503-514.

Case, M., E. Matheson, et al. (2008). "Mutation of genes affecting the RAS pathway is common in childhood acute lymphoblastic leukemia." Cancer Res **68**(16): 6803-6809.

Castor, A., L. Nilsson, et al. (2005). "Distinct patterns of hematopoietic stem cell involvement in acute lymphoblastic leukemia." Nat Med **11**(6): 630-637.

CCRG. (2010). "National Registry of Childhood Tumours." from http://www.ccrg.ox.ac.uk/.

Challen, G. A., D. Sun, et al. (2012). "Dnmt3a is essential for hematopoietic stem cell differentiation." Nat Genet **44**(1): 23-31.

Chen, J., Y. Li, et al. (2012). "A restricted cell population propagates glioblastoma growth after chemotherapy." Nature **488**(7412): 522-526.

Chia, N. Y., Y. S. Chan, et al. (2010). "A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity." Nature **468**(7321): 316-320.

Cichocki, F., T. Lenvik, et al. (2010). "Cutting edge: KIR antisense transcripts are processed into a 28-base PIWI-like RNA in human NK cells." J Immunol **185**(4): 2009-2012.

Clarke, M. F., J. E. Dick, et al. (2006). "Cancer stem cells--perspectives on current status and future directions: AACR Workshop on cancer stem cells." Cancer Res **66**(19): 9339-9344.

Cobaleda, C., N. Gutierrez-Cianca, et al. (2000). "A primitive hematopoietic cell is the target for the leukemic transformation in human philadelphia-positive acute lymphoblastic leukemia." Blood **95**(3): 1007-1013.

Coghill, R. W., J. Steward, et al. (1996). "Extra low frequency electric and magnetic fields in the bedplace of children diagnosed with leukaemia: a case-control study." Eur J Cancer Prev **5**(3): 153-158.

Cole, K. A., J. Huggins, et al. (2011). "RNAi screen of the protein kinome identifies checkpoint kinase 1 (CHK1) as a therapeutic target in neuroblastoma." Proc Natl Acad Sci U S A **108**(8): 3336-3341.

Collins, A. T., P. A. Berry, et al. (2005). "Prospective identification of tumorigenic prostate cancer stem cells." Cancer Res **65**(23): 10946-10951.

Cox, C. V., R. S. Evely, et al. (2004). "Characterization of acute lymphoblastic leukemia progenitor cells." Blood **104**(9): 2919-2925.

Cox, D. N., A. Chao, et al. (1998). "A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal." Genes Dev **12**(23): 3715-3727.

CRUK. (2011). "Childhood Cancer Statistics." from http://info.cancerresearchuk.org/cancerstats/childhoodcancer/incidence/.

Curley, M. D., V. A. Therrien, et al. (2009). "CD133 expression defines a tumor initiating cell population in primary human ovarian cancer." Stem Cells **27**(12): 2875-2883.

Dalerba, P., S. J. Dylla, et al. (2007). "Phenotypic characterization of human colorectal cancer stem cells." Proc Natl Acad Sci U S A **104**(24): 10158-10163.

Deng, W. and H. Lin (2002). "miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis." Dev Cell **2**(6): 819-830.

Dickins, R. A., M. T. Hemann, et al. (2005). "Probing tumor phenotypes using stable and regulated synthetic microRNA precursors." Nat Genet **37**(11): 1289-1295.

Doll, R. and R. Wakeford (1997). "Risk of childhood cancer from fetal irradiation." Br J Radiol **70**: 130-139.

Doulatov, S., F. Notta, et al. (2012). "Hematopoiesis: a human perspective." Cell Stem Cell **10**(2): 120-136.

Driessens, G., B. Beck, et al. (2012). "Defining the mode of tumour growth by clonal analysis." Nature **488**(7412): 527-530.

Duy, C., C. Hurtz, et al. (2011). "BCL6 enables Ph+ acute lymphoblastic leukaemia cells to survive BCR-ABL1 kinase inhibition." Nature **473**(7347): 384-388.

Duy, C., J. J. Yu, et al. (2010). "BCL6 is critical for the development of a diverse primary B cell repertoire." J Exp Med **207**(6): 1209-1221.

Eder, M., O. G. Ottmann, et al. (1990). "Effects of recombinant human IL-7 on blast cell proliferation in acute lymphoblastic leukemia." Leukemia **4**(8): 533-540.

Elbashir, S. M., W. Lendeckel, et al. (2001). "RNA interference is mediated by 21- and 22-nucleotide RNAs." Genes Dev **15**(2): 188-200.

Eppert, K., K. Takenaka, et al. (2011). "Stem cell gene expression programs influence clinical outcome in human leukemia." Nat Med **17**(9): 1086-1093.

Eramo, A., F. Lotti, et al. (2008). "Identification and expansion of the tumorigenic lung cancer stem cell population." Cell Death Differ **15**(3): 504-514.

Farazi, T. A., S. A. Juranek, et al. (2008). "The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members." Development **135**(7): 1201-1214.

Fear, N. T., E. Roman, et al. (2003). "Vitamin K and childhood cancer: a report from the United Kingdom Childhood Cancer Study." Br J Cancer **89**(7): 1228-1231.

Feng, D., C. Peng, et al. (2009). "Identification and characterization of cancer stem-like cells from primary carcinoma of the cervix uteri." Oncol Rep **22**(5): 1129-1134.

Ford, A. M., C. A. Bennett, et al. (1998). "Fetal origins of the TEL-AML1 fusion gene in identical twins with leukemia." Proc Natl Acad Sci U S A **95**(8): 4584-4588.

Forestier, E., S. Izraeli, et al. (2008). "Cytogenetic features of acute lymphoblastic and myeloid leukemias in pediatric patients with Down syndrome: an iBFM-SG study." Blood **111**(3): 1575-1583.

Fukuda, K., Y. Saikawa, et al. (2009). "Tumor initiating potential of side population cells in human gastric cancer." Int J Oncol **34**(5): 1201-1207.

Gan, H., X. Lin, et al. (2011). "piRNA profiling during specific stages of mouse spermatogenesis." RNA **17**(7): 1191-1203.

GarridoCastro, P., S. Bomken, et al. "ANGIOPOIETIN-1, a novel growth factor implicated in MLL-rearranged ALL." Manuscript in preparation.

Gattas, G. J., B. J. Quade, et al. (1999). "HMGIC expression in human adult and fetal tissues and in uterine leiomyomata." Genes Chromosomes Cancer **25**(4): 316-322.

Gazin, C., N. Wajapeyee, et al. (2007). "An elaborate pathway required for Ras-mediated epigenetic silencing." Nature **449**(7165): 1073-1077.

Gessner, A., M. Thomas, et al. (2010). "Leukemic fusion genes MLL/AF4 and AML1/MTG8 support leukemic self-renewal by controlling expression of the telomerase subunit TERT." Leukemia **24**(10): 1751-1759.

Gilham, C., J. Peto, et al. (2005). "Day care in infancy and risk of childhood acute lymphoblastic leukaemia: findings from UK case-control study." BMJ **330**(7503): 1294.

Girard, A., R. Sachidanandam, et al. (2006). "A germline-specific class of small RNAs binds mammalian Piwi proteins." Nature **442**(7099): 199-202.

Goardon, N., E. Marchi, et al. (2011). "Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia." Cancer Cell **19**(1): 138-152.

Goidts, V., J. Bageritz, et al. (2012). "RNAi screening in glioma stem-like cells identifies PFKFB4 as a key molecule important for cancer cell survival." Oncogene **31**(27): 3235-3243.

Golub, T. R., G. F. Barker, et al. (1995). "Fusion of the TEL gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia." Proc Natl Acad Sci U S A **92**(11): 4917-4921.

Goto, H., T. Inukai, et al. (2011). "Acute lymphoblastic leukemia and Down syndrome: the collaborative study of the Tokyo Children's Cancer Study Group and the Kyushu Yamaguchi Children's Cancer Study Group." Int J Hematol **93**(2): 192-198.

Gregory, R. I., K. P. Yan, et al. (2004). "The Microprocessor complex mediates the genesis of microRNAs." Nature **432**(7014): 235-240.

Grimm, D., K. L. Streetz, et al. (2006). "Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways." Nature **441**(7092): 537-541.

Grivna, S. T., E. Beyret, et al. (2006). "A novel class of small RNAs in mouse spermatogenic cells." Genes Dev **20**(13): 1709-1714.

Grochola, L. F., T. Greither, et al. (2008). "The stem cell-associated Hiwi gene in human adenocarcinoma of the pancreas: expression and risk of tumour-related death." Br J Cancer **99**(7): 1083-1088.

Gupta, P. B., C. L. Chaffer, et al. (2009). "Cancer stem cells: mirage or reality?" Nat Med **15**(9): 1010-1012.

Hammond, S. M., E. Bernstein, et al. (2000). "An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells." Nature **404**(6775): 293-296.

Harewood, L., H. Robinson, et al. (2003). "Amplification of AML1 on a duplicated chromosome 21 in acute lymphoblastic leukemia: a study of 20 cases." Leukemia **17**(3): 547-553.

Hasle, H., I. H. Clemmensen, et al. (2000). "Risks of leukaemia and solid tumours in individuals with Down's syndrome." Lancet **355**(9199): 165-169.

Hattori, H., F. Skoulidis, et al. (2011). "Context dependence of checkpoint kinase 1 as a therapeutic target for pancreatic cancers deficient in the BRCA2 tumor suppressor." Mol Cancer Ther **10**(4): 670-678.

He, G., L. Chen, et al. (2010). "Piwil2 expressed in various stages of cervical neoplasia is a potential complementary marker for p16." Am J Transl Res **2**(2): 156-169.

He, W., Z. Wang, et al. (2009). "Expression of HIWI in human esophageal squamous cell carcinoma is significantly associated with poorer prognosis." BMC Cancer **9**: 426.

Heerema, N. A., E. A. Raetz, et al. (2011). iAMP21 is associated with inferior outcomes in children with acute lymphoblastic leukaemia (ALL) on contemporary Children's Oncology Group (COG) Studies. American Society of Hematology Annual Meeting. San Diego, Blood. **118**.

Hemmati, H. D., I. Nakano, et al. (2003). "Cancerous stem cells can arise from pediatric brain tumors." Proc Natl Acad Sci U S A **100**(25): 15178-15183.

Hermann, P. C., S. L. Huber, et al. (2007). "Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer." Cell Stem Cell **1**(3): 313-323.

Hertzberg, L., E. Vendramini, et al. (2010). "Down syndrome acute lymphoblastic leukemia, a highly heterogeneous disease in which aberrant expression of CRLF2 is associated with mutated JAK2: a report from the International BFM Study Group." Blood **115**(5): 1006-1017.

Ho, M. M., A. V. Ng, et al. (2007). "Side population in human lung cancer cell lines and tumors is enriched with stem-like cancer cells." Cancer Res **67**(10): 4827-4833.

Hong, D., R. Gupta, et al. (2008). "Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia." Science **319**(5861): 336-339.

Hope, K. J., S. Cellot, et al. (2010). "An RNAi screen identifies Msi2 and Prox1 as having opposite roles in the regulation of hematopoietic stem cell activity." Cell Stem Cell **7**(1): 101-113.

Hope, K. J., L. Jin, et al. (2004). "Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity." Nat Immunol **5**(7): 738-743.

Hu, G., J. Kim, et al. (2009). "A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal." Genes Dev **23**(7): 837-848.

Hutvagner, G., J. McLachlan, et al. (2001). "A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA." Science **293**(5531): 834-838.

Iorns, E., T. M. Ward, et al. (2012). "Whole genome in vivo RNAi screening identifies the leukemia inhibitory factor receptor as a novel breast tumor suppressor." Breast Cancer Res Treat.

Joo, K. M., S. Y. Kim, et al. (2008). "Clinical and biological implications of CD133-positive and CD133-negative cells in glioblastomas." Lab Invest **88**(8): 808-815.

Kelly, P. N., A. Dakic, et al. (2007). "Tumor growth need not be driven by rare cancer stem cells." Science **317**(5836): 337.

Kirino, Y. and Z. Mourelatos (2007). "2'-O-methyl modification in mouse piRNAs and its methylase." Nucleic Acids Symp Ser (Oxf)(51): 417-418.

Kong, Y., S. Yoshida, et al. (2008). "CD34+CD38+CD19+ as well as CD34+CD38-CD19+ cells are leukemia-initiating cells with self-renewal capacity in human B-precursor ALL." Leukemia **22**(6): 1207-1213.

Konuma, T., H. Oguro, et al. (2010). "Role of the polycomb group proteins in hematopoietic stem cells." Dev Growth Differ **52**(6): 505-516.

Krivtsov, A. V., D. Twomey, et al. (2006). "Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9." Nature **442**(7104): 818-822.

Kuchen, S., W. Resch, et al. (2010). "Regulation of microRNA expression and abundance during lymphopoiesis." Immunity **32**(6): 828-839.

Kumar, A. R., Q. Li, et al. (2009). "A role for MEIS1 in MLL-fusion gene leukemia." Blood **113**(8): 1756-1758.

Kuramochi-Miyagawa, S., T. Kimura, et al. (2004). "Mili, a mammalian member of piwi family gene, is essential for spermatogenesis." Development **131**(4): 839-849.

Kuramochi-Miyagawa, S., T. Kimura, et al. (2001). "Two mouse piwi-related genes: miwi and mili." Mech Dev **108**(1-2): 121-133.

Kuramochi-Miyagawa, S., T. Watanabe, et al. (2008). "DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes." Genes Dev **22**(7): 908-917.

Landgraf, P., M. Rusu, et al. (2007). "A mammalian microRNA expression atlas based on small RNA library sequencing." Cell **129**(7): 1401-1414.

Lapidot, T., C. Sirard, et al. (1994). "A cell initiating human acute myeloid leukaemia after transplantation into SCID mice." Nature **367**(6464): 645-648.

Lau, N. C., A. G. Seto, et al. (2006). "Characterization of the piRNA complex from rat testes." Science **313**(5785): 363-367.

le Viseur, C., M. Hotfilder, et al. (2008). "In childhood acute lymphoblastic leukemia, blasts at different stages of immunophenotypic maturation have stem cell properties." Cancer Cell **14**(1): 47-58.

Lee, J. H., C. Jung, et al. (2010). "Pathways of proliferation and antiapoptosis driven in breast cancer stem cells by stem cell protein piwil2." Cancer Res **70**(11): 4569-4579.

Lee, J. H., D. Schutte, et al. (2006). "Stem-cell protein Piwil2 is widely expressed in tumors and inhibits apoptosis through activation of Stat3/Bcl-XL pathway." Hum Mol Genet **15**(2): 201-211.

Lee, Y., C. Ahn, et al. (2003). "The nuclear RNase III Drosha initiates microRNA processing." Nature **425**(6956): 415-419.

Lee, Y., K. Jeon, et al. (2002). "MicroRNA maturation: stepwise processing and subcellular localization." EMBO J **21**(17): 4663-4670.

Lessard, J. and G. Sauvageau (2003). "Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells." Nature **423**(6937): 255-260.

Lewis, B. P., C. B. Burge, et al. (2005). "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." Cell **120**(1): 15-20.

Li, C., D. G. Heidt, et al. (2007). "Identification of pancreatic cancer stem cells." Cancer Res **67**(3): 1030-1037.

Li, H., J. M. Lahti, et al. (1996). "Molecular cloning and chromosomal localization of the human cyclin C (CCNC) and cyclin E (CCNE) genes: deletion of the CCNC gene in human tumors." Genomics **32**(2): 253-259.

Ligon, A. H., S. D. Moore, et al. (2005). "Constitutional rearrangement of the architectural factor HMGA2: a novel human phenotype including overgrowth and lipomas." Am J Hum Genet **76**(2): 340-348.

Lim, L. P., N. C. Lau, et al. (2005). "Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs." Nature **433**(7027): 769-773.

Lin, H. and A. C. Spradling (1997). "A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary." Development **124**(12): 2463-2476.

Lingel, A., B. Simon, et al. (2004). "Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain." Nat Struct Mol Biol **11**(6): 576-577.

Liu, W. K., X. Y. Jiang, et al. (2010). "Expression of PSCA, PIWIL1 and TBX2 and its correlation with HPV16 infection in formalin-fixed, paraffin-embedded cervical squamous cell carcinoma specimens." Arch Virol **155**(5): 657-663.

Liu, W. K., X. Y. Jiang, et al. (2010). "Expression of PSCA, PIWIL1, and TBX2 in endometrial adenocarcinoma." Onkologie **33**(5): 241-245.

Liu, X., Y. Sun, et al. (2006). "Expression of hiwi gene in human gastric cancer was associated with proliferation of cancer cells." Int J Cancer **118**(8): 1922-1929.

Lu, Y., K. Zhang, et al. (2012). "Piwil2 suppresses p53 by inducing phosphorylation of signal transducer and activator of transcription 3 in tumor cells." PLoS One **7**(1): e30999.

Luis, T. C., M. Ichii, et al. (2012). "Wnt signaling strength regulates normal hematopoiesis and its deregulation is involved in leukemia development." Leukemia **26**(3): 414-421.

Luo, J., M. J. Emanuele, et al. (2009). "A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene." Cell **137**(5): 835-848.

Ma, J. B., K. Ye, et al. (2004). "Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain." Nature **429**(6989): 318-322.

Ma, S., K. W. Chan, et al. (2007). "Identification and characterization of tumorigenic liver cancer stem/progenitor cells." Gastroenterology **132**(7): 2542-2556.

Ma, X., P. A. Buffler, et al. (2002). "Daycare attendance and risk of childhood acute lymphoblastic leukaemia." Br J Cancer **86**(9): 1419-1424.

Magee, J. A., E. Piskounova, et al. (2012). "Cancer stem cells: impact, heterogeneity, and uncertainty." Cancer Cell **21**(3): 283-296.

Mahmoud, H. H., S. A. Ridge, et al. (1995). "Intrauterine monoclonal origin of neonatal concordant acute lymphoblastic leukemia in monozygotic twins." Med Pediatr Oncol **24**(2): 77-81.

Maia, A. T., V. H. van der Velden, et al. (2003). "Prenatal origin of hyperdiploid acute lymphoblastic leukemia in identical twins." Leukemia **17**(11): 2202-2206.

Maloney, K. W. (2011). "Acute lymphoblastic leukaemia in children with Down syndrome: an updated review." Br J Haematol **155**(4): 420-425.

Maloney, K. W., W. L. Carroll, et al. (2010). "Down syndrome childhood acute lymphoblastic leukemia has a unique spectrum of sentinel cytogenetic lesions that influences treatment outcome: a report from the Children's Oncology Group." Blood **116**(7): 1045-1050.

Mar, B. G., D. Amakye, et al. (2011). "The controversial role of the Hedgehog pathway in normal and malignant hematopoiesis." Leukemia **25**(11): 1665-1673.

McBride, J. L., R. L. Boudreau, et al. (2008). "Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi." Proc Natl Acad Sci U S A **105**(15): 5868-5873.

McDermott, S. P., K. Eppert, et al. (2010). "Comparison of human cord blood engraftment between immunocompromised mouse strains." Blood **116**(2): 193-200.

Meacham, C. E., E. E. Ho, et al. (2009). "In vivo RNAi screening identifies regulators of actin dynamics as key determinants of lymphoma progression." Nat Genet **41**(10): 1133-1137.

Meyer, B., D. Krisponeit, et al. (2007). "Quantitative expression analysis in peripheral blood of patients with chronic myeloid leukaemia: correlation between HMGA2 expression and white blood cell count." Leuk Lymphoma **48**(10): 2008-2013.

Moorman, A. V. (2012). "The clinical relevance of chromosomal and genomic abnormalities in B-cell precursor acute lymphoblastic leukaemia." Blood Rev **26**(3): 123-135.

Moorman, A. V., H. M. Ensor, et al. (2010). "Prognostic effect of chromosomal abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: results from the UK Medical Research Council ALL97/99 randomised trial." Lancet Oncol **11**(5): 429-438.

Moorman, A. V., S. M. Richards, et al. (2007). "Prognosis of children with acute lymphoblastic leukemia (ALL) and intrachromosomal amplification of chromosome 21 (iAMP21)." Blood **109**(6): 2327-2330.

Mori, H., S. M. Colman, et al. (2002). "Chromosome translocations and covert leukemic clones are generated during normal fetal development." Proc Natl Acad Sci U S A **99**(12): 8242-8247.

Morisot, S., A. S. Wayne, et al. (2010). "High frequencies of leukemia stem cells in poor-outcome childhood precursor-B acute lymphoblastic leukemias." Leukemia **24**(11): 1859-1866.

Mullighan, C. G., J. R. Collins-Underwood, et al. (2009). "Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia." Nat Genet **41**(11): 1243-1246.

Mullighan, C. G., S. Goorha, et al. (2007). "Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia." Nature **446**(7137): 758-764.

Nicholson, L., T. Knight, et al. (2012). "Casitas B lymphoma mutations in childhood acute lymphoblastic leukemia." Genes Chromosomes Cancer **51**(3): 250-256.

Notta, F., C. G. Mullighan, et al. (2011). "Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells." Nature **469**(7330): 362-367.

Novershtern, N., A. Subramanian, et al. (2011). "Densely interconnected transcriptional circuits control cell states in human hematopoiesis." Cell **144**(2): 296-309.

Nowell, P. C. (1976). "The clonal evolution of tumor cell populations." Science **194**(4260): 23-28.

O'Brien, C. A., A. Pollett, et al. (2007). "A human colon cancer cell capable of initiating tumour growth in immunodeficient mice." Nature **445**(7123): 106-110.

O'Neill, K. A., K. J. Bunch, et al. (2012). "Immunophenotype and cytogenetic characteristics in the relationship between birth weight and childhood leukemia." Pediatr Blood Cancer **58**(1): 7-11.

Ogden, A. T., A. E. Waziri, et al. (2008). "Identification of A2B5+CD133- tumor-initiating cells in adult human gliomas." Neurosurgery **62**(2): 505-514; discussion 514-505.

Okuda, T., S. A. Shurtleff, et al. (1995). "Frequent deletion of p16INK4a/MTS1 and p15INK4b/MTS2 in pediatric acute lymphoblastic leukemia." Blood **85**(9): 2321-2330.

Orlovsky, K., A. Kalinkovich, et al. (2011). "Down-regulation of homeobox genes MEIS1 and HOXA in MLL-rearranged acute leukemia impairs engraftment and reduces proliferation." Proc Natl Acad Sci U S A **108**(19): 7956-7961.

Pan, Q., P. E. de Ruiter, et al. (2011). "Disturbance of the microRNA pathway by commonly used lentiviral shRNA libraries limits the application for screening host factors involved in hepatitis C virus infection." FEBS Lett **585**(7): 1025-1030.

Parker, R. and U. Sheth (2007). "P bodies and the control of mRNA translation and degradation." Mol Cell **25**(5): 635-646.

Pearce, M. S., J. A. Salotti, et al. (2012). "Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study." Lancet **380**(9840): 499-505.

Preston, D. L., S. Kusumi, et al. (1994). "Cancer incidence in atomic bomb survivors. Part III. Leukemia, lymphoma and multiple myeloma, 1950-1987." Radiat Res **137**(2 Suppl): S68-97.

Pui, C. H., S. C. Raimondi, et al. (1993). "Immunophenotypes and karyotypes of leukemic cells in children with Down syndrome and acute lymphoblastic leukemia." J Clin Oncol **11**(7): 1361-1367.

Pui, C. H., M. V. Relling, et al. (2004). "Acute lymphoblastic leukemia." N Engl J Med **350**(15): 1535-1548.

Qiao, D., A. M. Zeeman, et al. (2002). "Molecular characterization of hiwi, a human member of the piwi gene family whose overexpression is correlated to seminomas." Oncogene **21**(25): 3988-3999.

Quintana, E., M. Shackleton, et al. (2010). "Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized." Cancer Cell **18**(5): 510-523.

Quintana, E., M. Shackleton, et al. (2008). "Efficient tumour formation by single human melanoma cells." Nature **456**(7222): 593-598.

Raimondi, S. C., J. L. Frestedt, et al. (1995). "Acute lymphoblastic leukemias with deletion of 11q23 or a novel inversion (11)(p13q23) lack MLL gene rearrangements and have favorable clinical features." Blood **86**(5): 1881-1886.

Rajasethupathy, P., I. Antonov, et al. (2012). "A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity." Cell **149**(3): 693-707.

Raynaud, S., H. Cave, et al. (1996). "The 12;21 translocation involving TEL and deletion of the other TEL allele: two frequently associated alterations found in childhood acute lymphoblastic leukemia." Blood **87**(7): 2891-2899.

Rehe, K., K. Wilson, et al. (2012). "Acute B lymphoblastic leukaemia propagating cells are present at high frequency in diverse lymphoblast populations." EMBO Molecular Medicine, accepted for publication.

Ricci-Vitiani, L., D. G. Lombardi, et al. (2007). "Identification and expansion of human colon-cancer-initiating cells." Nature **445**(7123): 111-115.

Rives, S., J. Estella, et al. (2011). "Intermediate dose of imatinib in combination with chemotherapy followed by allogeneic stem cell transplantation improves early outcome in paediatric Philadelphia chromosome-positive acute lymphoblastic leukaemia (ALL): results of the Spanish Cooperative Group SHOP studies ALL-94, ALL-99 and ALL-2005." Br J Haematol **154**(5): 600-611.

Russell, L. J., M. Capasso, et al. (2009). "Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia." Blood **114**(13): 2688-2698.

Saito, K., K. M. Nishida, et al. (2006). "Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome." Genes Dev **20**(16): 2214-2222.

Sambrook, J. and D. Russel (2001). Preparation and transformation of competent E. coli using calcium chloride. Molecular cloning: A laboratory manual. Cold Spring Harbor, Cold Spring Harbor Laboratory Press. **1:** 1.116-111.118.

Sasaki, T., A. Shiohama, et al. (2003). "Identification of eight members of the Argonaute family in the human genome small star, filled." Genomics **82**(3): 323-330.

Schatton, T., G. F. Murphy, et al. (2008). "Identification of cells initiating human melanomas." Nature **451**(7176): 345-349.

Schepers, A. G., H. J. Snippert, et al. (2012). "Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas." Science **337**(6095): 730-735.

Schlabach, M. R., J. Luo, et al. (2008). "Cancer proliferation gene discovery through functional genomics." Science **319**(5863): 620-624.

Schmitz, M., P. Breithaupt, et al. (2011). "Xenografts of highly resistant leukemia recapitulate the clonal composition of the leukemogenic compartment." Blood **118**(7): 1854-1864.

Schultz, K. R., W. P. Bowman, et al. (2009). "Improved early event-free survival with imatinib in Philadelphia chromosome-positive acute lymphoblastic leukemia: a children's oncology group study." J Clin Oncol **27**(31): 5175-5181.

Schultz, N., D. R. Marenstein, et al. (2011). "Off-target effects dominate a large-scale RNAi screen for modulators of the TGF-beta pathway and reveal microRNA regulation of TGFBR2." Silence **2**: 3.

Shaknovich, R., L. Cerchietti, et al. (2011). "DNA methyltransferase 1 and DNA methylation patterning contribute to germinal center B-cell differentiation." Blood **118**(13): 3559-3569.

Sharma, A. K., M. C. Nelson, et al. (2001). "Human CD34(+) stem cells express the hiwi gene, a human homologue of the Drosophila gene piwi." Blood **97**(2): 426-434.

Shultz, L. D., B. L. Lyons, et al. (2005). "Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells." J Immunol **174**(10): 6477-6489.

Siddiqi, S., M. Terry, et al. (2012). "Hiwi mediated tumorigenesis is associated with DNA hypermethylation." PLoS One **7**(3): e33711.

Sigoillot, F. D., S. Lyman, et al. (2012). "A bioinformatics method identifies prominent off-targeted transcripts in RNAi screens." Nat Methods **9**(4): 363-366.

Silva, J. M., M. Z. Li, et al. (2005). "Second-generation shRNA libraries covering the mouse and human genomes." Nat Genet **37**(11): 1281-1288.

Silva, J. M., K. Marran, et al. (2008). "Profiling essential genes in human mammary cells by multiplex RNAi screening." Science **319**(5863): 617-620.

Singh, S. K., I. D. Clarke, et al. (2003). "Identification of a cancer stem cell in human brain tumors." Cancer Res **63**(18): 5821-5828.

Somervaille, T. C., C. J. Matheny, et al. (2009). "Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells." Cell Stem Cell **4**(2): 129-140.

Song, J. J., S. K. Smith, et al. (2004). "Crystal structure of Argonaute and its implications for RISC slicer activity." Science **305**(5689): 1434-1437.

Stegmeier, F., G. Hu, et al. (2005). "A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells." Proc Natl Acad Sci U S A **102**(37): 13212-13217.

Strefford, J. C., F. W. van Delft, et al. (2006). "Complex genomic alterations and gene expression in acute lymphoblastic leukemia with intrachromosomal amplification of chromosome 21." Proc Natl Acad Sci U S A **103**(21): 8167-8172.

Sugimoto, K., H. Kage, et al. (2007). "The induction of H3K9 methylation by PIWIL4 at the p16Ink4a locus." Biochem Biophys Res Commun **359**(3): 497-502.

Sun, G., Y. Wang, et al. (2011). "Clinical significance of Hiwi gene expression in gliomas." Brain Res **1373**: 183-188.

Takaishi, S., T. Okumura, et al. (2009). "Identification of gastric cancer stem cells using the cell surface marker CD44." Stem Cells **27**(5): 1006-1020.

Taubert, H., T. Greither, et al. (2007). "Expression of the stem cell self-renewal gene Hiwi and risk of tumour-related death in patients with soft-tissue sarcoma." Oncogene **26**(7): 1098-1100.

Taussig, D. C., J. Vargaftig, et al. (2010). "Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34(-) fraction." Blood **115**(10): 1976-1984.

Thomas, M., A. Gessner, et al. (2005). "Targeting MLL-AF4 with short interfering RNAs inhibits clonogenicity and engraftment of t(4;11)-positive human leukemic cells." Blood **106**(10): 3559-3566.

Tian, Y., D. K. Simanshu, et al. (2011). "Structural basis for piRNA 2'-O-methylated 3'-end recognition by Piwi PAZ (Piwi/Argonaute/Zwille) domains." Proc Natl Acad Sci U S A **108**(3): 903-910.

Uckun, F. M., T. G. Gesner, et al. (1989). "Leukemic B-cell precursors express functional receptors for human interleukin-3." Blood **73**(2): 533-542.

UKCCSInvestigators (1999). "Exposure to power-frequency magnetic fields and the risk of childhood cancer. UK Childhood Cancer Study Investigators." Lancet **354**(9194): 1925-1931.

UKCCSInvestigators (2002). "The United Kingdom Childhood Cancer Study of exposure to domestic sources of ionising radiation: 1: radon gas." Br J Cancer **86**(11): 1721-1726.

Vagin, V. V., A. Sigova, et al. (2006). "A distinct small RNA pathway silences selfish genetic elements in the germline." Science **313**(5785): 320-324.

Valakh, V., S. A. Naylor, et al. (2012). "A large-scale RNAi screen identifies functional classes of genes shaping synaptic development and maintenance." Dev Biol **366**(2): 163-171.

Valent, P., D. Bonnet, et al. (2012). "Cancer stem cell definitions and terminology: the devil is in the details." Nat Rev Cancer **12**(11): 767-775.

van Delft, F. W., S. Horsley, et al. (2011). "Clonal origins of relapse in ETV6-RUNX1 acute lymphoblastic leukemia." Blood **117**(23): 6247-6254.

Victora, G. D. and M. C. Nussenzweig (2012). "Germinal centers." Annu Rev Immunol **30**: 429-457.

Vormoor, H. J. (2009). "Malignant stem cells in childhood acute lymphoblastic leukemia: the stem cell concept revisited." Cell Cycle **8**(7): 996-999.

Wang, J., P. O. Sakariassen, et al. (2008). "CD133 negative glioma cells form tumors in nude rats and give rise to CD133 positive cells." Int J Cancer **122**(4): 761-768.

Wang, Q. E., C. Han, et al. (2011). "Stem cell protein Piwil2 modulates chromatin modifications upon cisplatin treatment." Mutat Res **708**(1-2): 59-68.

Wang, Y., Y. Liu, et al. (2012). "The PIWI protein acts as a predictive marker for human gastric cancer." Int J Clin Exp Pathol **5**(4): 315-325.

Watanabe, T., A. Takeda, et al. (2006). "Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes." Genes Dev **20**(13): 1732-1743.

Watanabe, T., S. Tomizawa, et al. (2011). "Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse Rasgrf1 locus." Science **332**(6031): 848-852.

Westbrook, T. F., E. S. Martin, et al. (2005). "A genetic screen for candidate tumor suppressors identifies REST." Cell **121**(6): 837-848.

Westerman, B. A., A. K. Braat, et al. (2011). "A genome-wide RNAi screen in mouse embryonic stem cells identifies Mp1 as a key mediator of differentiation." J Exp Med **208**(13): 2675-2689.

Williams, R. T., W. den Besten, et al. (2007). "Cytokine-dependent imatinib resistance in mouse BCR-ABL+, Arf-null lymphoblastic leukemia." Genes Dev **21**(18): 2283-2287.

Wurdak, H., S. Zhu, et al. (2010). "An RNAi screen identifies TRRAP as a regulator of brain tumor-initiating cell differentiation." Cell Stem Cell **6**(1): 37-47.

Yan, Z., H. Y. Hu, et al. (2011). "Widespread expression of piRNA-like molecules in somatic tissues." Nucleic Acids Res **39**(15): 6596-6607.

Yi, R., Y. Qin, et al. (2003). "Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs." Genes Dev **17**(24): 3011-3016.

Yin, D. T., H. Q. Li, et al. (2011). "Expression of Piwil2 and its relationship with tumor invasion and metastasis in papillary thyroid carcinoma." Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi **46**(3): 237-239.

Zender, L., W. Xue, et al. (2008). "An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer." Cell **135**(5): 852-864.

Zeng, Y., L. K. Qu, et al. (2011). "HIWI expression profile in cancer cells and its prognostic value for patients with colorectal cancer." Chin Med J (Engl) **124**(14): 2144-2149.

Zhang, E. E., A. C. Liu, et al. (2009). "A genome-wide RNAi screen for modifiers of the circadian clock in human cells." Cell **139**(1): 199-210.

Zhang, S., C. Balch, et al. (2008). "Identification and characterization of ovarian cancer-initiating cells from primary human tumors." Cancer Res **68**(11): 4311-4320.

Zhao, Y. M., J. M. Zhou, et al. (2012). "HIWI is associated with prognosis in patients with hepatocellular carcinoma after curative resection." Cancer **118**(10): 2708-2717.

Zhou, M., L. Gu, et al. (1995). "Homozygous deletions of the CDKN2 (MTS1/p16ink4) gene in cell lines established from children with acute lymphoblastic leukemia." Leukemia **9**(7): 1159-1161.

Zhou, X., K. F. Benson, et al. (1995). "Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C." Nature **376**(6543): 771-774.

Zuber, J., J. Shi, et al. (2011). "RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia." Nature **478**(7370): 524-528.