Involvement of *ID4* (Inhibitor of DNA Binding 4) in Haematopoietic Malignancies.

Submission for the degree of Doctor of Philosophy

2011

Palminder Kaur Dusanjh, BSc (Hons), MSc.

Involvement of *ID4* (Inhibitor of DNA Binding 4) in Haematopoietic Malignancies.

Palminder Kaur Dusanjh

Abstract

Chromosomal translocations involving the immunoglobulin heavy chain (*IGH*) locus have been described in B-cell precursor acute lymphoblastic leukaemia (BCP-ALL). One of these, t (6;14)(p22;q32), involved the Inhibitor of DNA Binding 4 *(ID4)* gene, resulting in deregulated *ID4* expression. Members of the ID family are helix-loop-helix proteins that lack DNA binding domains. They are thought to regulate cellular differentiation and proliferation by binding and sequestering E proteins from their DNA targets. The aim of this thesis was to examine the expression and oncogenic potential of *ID4* in B-cells.

ID4 expression was detected in a subset of chronic lymphocytic leukaemia (CLL) and BCP-ALL lacking the *ID4/IGH* chromosomal translocation. *ID4* expression defined a subtype of BCP-ALL with a distinctive gene signature.

ID4, unlike other members of the *ID* family is not expressed in normal B-cells. Unexpectedly, ectopic ID4 protein expression in most but not all derived B-cell lines reduced proliferation and induced cell cycle arrest. B-cells in ID4-induced cell cycle arrest showed a partial resistance to apoptosis. Proteomic analysis of cell lines expressing ectopic ID4 showed that ID4 interacted with E proteins, although these data could not be confirmed with endogenous ID4 for technical reasons. Primary normal mouse B lymphocyte progenitors showed impaired B-cell development in an *in vitro* assay following expression of *ID4*. In contrast, *ID4* cooperated with *myc* over-expression in inducing proliferation of murine B lymphocyte progenitors.

Collectively, these data suggest that *ID4* may be a novel genetic “driver” in different subtypes of B-cell malignancy. *ID4* may contribute to leukemogenesis of B-cell progenitors by inducing inappropriate gene expression and affecting mitosis. This collectively promotes genetic instability and generates a preleukemic stage.

Acknowledgements

I would like to take this opportunity to thank my supervisor Professor Dyer for the chance to do this PhD and for his supervision. I would like to thank Dr Loraine Karran for her supervision, but also for her support and patience throughout. I am grateful to Dr Aneela Majid for the CLL prognostic data and tissue samples, and also for her advice and assistance with the project. I would also like to thank Dr Takashi Akasaka for the B-cell transfected clones.

I am grateful to Mrs Rebecca Jukes-Jones, Dr Robert Boyd, Mr David Brown and Dr Kelvin Kain for their advice with the immunoprecipitations and for performing the mass spectrometry. I would also like to thank Dr Alison Michie and Dr Karen Blyth (Paul O'Gorman Leukaemia Research Centre and the Beatson Institute, University of Glasgow) for the collaboration with the *cdkn2a-/-* and *Emu-myc* transgenic mice.

I would like to acknowledge all people within the MRC Toxicology Unit for their advice over the years with various aspects of this project and for making my time there more enjoyable. I am grateful for the coffee/pub group of colleagues and friends for their support and welcomed distractions.

I would also like to thank my parents and in-laws for their support and encouragement. I am grateful to my siblings and extended family for providing the welcomed distractions and 'lending an ear' when needed. Last, but not least, I would like to thank my husband Manjinder Singh for his patience and support, without which, none of this would have been possible.

The contributions towards this research are as follows:

Dr Takashi Akasaka (Professor Dyer laboratory) for providing all the cell lines over-expressing ID4 and the constructs used to create them.

Dr Alison Michie (Paul O'Gorman Leukaemia Research centre, University of Glasgow) and Dr Karen Blyth (Beatson Institute, University of Glasgow) for the *Emu-myc* and *cdkn2a -/-* mice, for the genotyping of the embryos and for allowing the work to be carried out in their laboratories.

Dr Kelvin Cain, Dr Robert Boyd, Mr David Brown and Mrs Rebecca Jukes-Jones (MRC Toxicology Unit) for their advice regarding the immunoprecipitation experiments. Dr Robert Boyd and Mrs Rebecca Jukes-Jones for processing the gels and generating the Mascot data for the mass spectrometry analysis.

Dr Aneela Majid (Professor Dyer Laboratory) for access to the CLL patient prognostic data.

To Dr Lüder Meyer and Dr Manon Queudeville (University of Ulm) for the xenograft RNAs.

Table of contents

List of figures and tables

Chapter 1: Introduction.

Chapter 2: Materials and Methods.

Chapter 3: Expression of ID4 in normal and malignant tissues.

Chapter 4: Effects of ID4 over-expression in B-cell lines.

Chapter 5: Effects of ID4 expression in primary B-cell progenitors.

Chapter 6: Identification of proteins which interact with ID4.

Chapter 7: Discussion.

Abbreviations

ABF-1	Activated B cell factor.
AD1	Activation domain 1.
AD2	Activation domain 2.
AID	Activation induced cytidine deaminase.
AKT	Serine/threonine protein kinase Akt .
ALL	Acute Lymphoblastic Leukaemia.
Alt	Alternative.
AML	Acute myeloid leukaemia.
APC/C	Anaphase promoting complex cyclosome.
ARF	ADP -ribosylation factors.
ATM	Ataxia telangcectasia-mutated gene .
BCP-ALL	B cell precursor ALL.
BCR	B cell receptor.
BHLH	Basic HLH.
BMP	Bone morphogenic protein.
B-NHL	B-cell non Hodgkins lymphoma.
BP	Base pairs.
Can	Canonical.
Cdc20	Cell division cycle protein 20.
Cdh1	Cdc20 homologue 1.
cDNA	Complementary DNA
CDR	Complementary determining regions.
CEBP	CCAAT/enhancer binding protein .
CLL	Chronic lymphocytic leukaemia.
CLP	Common lymphoid progenitor.
CML	Chronic myeloid leukaemia.
Co SMAD	Common partner SMAD.
CoA	Coactivators.
CRLF2	Cytokine receptor-like factor 2.
CSN	COP9 signalosome.
CSR	Class switch recombination.
D_H	Diversity region of the *IGH*
DLBCL	Diffuse large B-cell lymphoma.
E	Glutamic acid.
E2A	E2A immunoglobulin enhancer binding factors E12/E47.
EBF	Early B cell factor.
ECL	Enhanced chemilluminescence.
EDTA	Ethylenediaminetetraacetic acid.
ELP	Early lymphoid progenitor.
EST	Expressed sequence tag.
EV	Empty vector.
FGFR3	Fibroblast Growth Factor Receptor.

FISH	Fluorescent *in situ G*enomic Hybridisation.
GC	Germinal Centres.
GFP	Green fluorescent protein.
GMP	Granulocyte macrophage progenitors.
GPI	Glycosylphosphatidylinositol.
HCL	Hydrochloric acid.
HEEBO	Human exonic evidence based oligonucleotide.
HLH	Helix loop helix.
HRP	Horse radish peroxidise.
HRS	Hodgkin Reed Sternberg cells.
HSC	Haematopoietic Stem Cell.
I SMAD	Inhibitory SMAD.
ID	Inhibitor of DNA binding.
ID1	Inhibitor of DNA binding 1.
ID2	Inhibitor of DNA binding 2.
ID3	Inhibitor of DNA binding 3.
ID4	Inhibitor of DNA binding 4.
IG	Immunoglobulin.
IGF	Insulin like growth factor.
IGH	Immunoglobulin heavy chain.
IKZF1	Ikaros1.
IP	Immunoprecipitation.
J_H	Joining region of the *IGH*
L	Leucine.
l7	Interleukin 7.
LDI PCR	Long distance inverse PCR.
LMPP	Lymphoid primed MPP.
LT-HSC	Long term HSC.
M	Mutated
MAPK	Mitogen activated protein kinase.
MEEBO	Mouse exonic evidence based oligonucleotide.
MEP	Megakaryocyte erythrocyte progenitors.
MMSET	Multiple myeloma SET domain .
MPP	Multipotent progenitors.
MRD	Minimum residual disease.
NES	Nuclear Export Signal.
NHL	Non-Hodgkin Lymphoma.
NK	Natural killer.
NP40	Tergitol-type NP-40.
ORF	Open reading frame.
PAGE	Polyacrylamide gel electophoresis
PCNA	Proliferating cell nuclear antigen.
PCR	Polymerase chain reaction.
pDc	Plasmacytoid dentdritic cells.

PEL	Primary effusion lymphoma.
PI	Propidium idodide.
Pi	Protease inhibitor.
PI3K	Phosphoinositide 3 kinase (PI3K).
Q	Glutamine.
R	Arginine.
R SMAD	Receptor regulated SMAD.
Rb	Retinoblastoma.
RGLS	Restriction Landmark Genomic Scanning.
RSU_1	Ras supressor protein 1.
RT-PCR	Reverse transcriptase PCR.
SCLC	Small Cell Lung Carcinoma.
SD	Standard deviation.
SDS	Sodium Dodecyl Sulphate.
SEM	Standard error of the mean.
SHM	Somatic hypermutation.
SOSC7	Suppressor of cytokine signalling 7.
ST-HSC	Short term HSC.
T-ALL	T cell ALL.
TBS	Tris buffered saline.
TBS T	TBS tween.
TCF	Transcription factor.
TdT	Deoxynucleotidyl transferase .
TEMED	N,N,N',N'-Tetramethylethylenediamine.
TGFβ	Transforming growth factor pathway beta .
TPM	Transcripts per million.
Tris	Tris(hydroxymethyl)aminomethane.
TSLP	Thymic stromal lymphopoietin.
U	Unmutated
UPS	Ubiquitin proteosome system.
UTR	Un-translated region.
V_H	Variable region of the *IGH*
ZNF289	Zinc finger protein 289.

Chapter 1: Introduction

1.1 Cancer and genetic instability.

The development of cancer requires the disruption of regulatory circuits which maintain cellular homeostasis. The hallmarks of cancer include self sufficiency in growth signals, insensitivity to anti-growth signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis, tissue invasion and metastasis (Hanahan and Weinberg, 2000).

These changes arise as a consequence of genetic events resulting in gene deregulation. Genes can be inactivated as a result of deletions, mutations or epigenetic mechanisms. Genes may also be deregulated by amplifications or chromosomal translocations (Mitelman *et al.*, 2007; Rabbitts, 2009).

The influences of these key genetic events in certain cancers is well characterised, particularly in haematological malignancies where chromosomal translocations result in deregulated gene expression or gene fusions (Mitelman *et al.*, 2007; Rabbitts, 2009). For example, the identification of the Philadelphia chromosomal translocation in chronic myeloid leukaemia (CML) involving a fusion protein between the *BCR* gene (on chromosome 22) and the *ABL* gene (on chromosome 9). This was the first example of a recurring oncogenic fusion protein being involved in a specific disease (Nowell, 2007) and led to the development of a tumour specific therapy (Deininger *et al.*, 2005).

However, there is often more than one genetic aberration present in many cancers. Cancer cells acquire somatic mutations as part of normal development which differ from any germline mutations. Those mutations which give proliferative and/or survival responses to cells are acquired as part of an on-going process during development. These somatic mutations are categorised as ‘drivers’ or ‘passengers’. ‘Drivers’ are defined as those mutations which give the cells a survival advantage. ‘Passenger’ mutations are those mutations which allow the ancestry of the cell to be mapped but are not beneficial for the cells survival (Stratton *et al.*, 2009). ‘Driver’ somatic mutations offer new selective targets for therapy and are identified by whole genome sequencing (Stratton *et al.*, 2009).

This was recently demonstrated in an acute myeloid leukaemia (AML) patient where somatic mutations of the malignant tissue were examined and compared with the normal tissue. This study identified 12 somatic mutations which targeted the coding sequences of genes and 52 somatic point mutations present in conserved or regulatory regions of genes. The 12 mutations targeting coding sequences possibly contributed to the malignancy because of the disruption of specific proteins. One 'new' driver mutation identified in this study targeted the *IDH1* gene, which was also detected in 15 out of 187 other AML cases (Mardis *et al.*, 2009). The other 52 mutations in non-coding regions were thought to be random events that were present in the haematopoietic cell that had been transformed (Mardis *et al.*, 2009).

The whole cancer genome sequencing approach has also been demonstrated in other cancers, such as gliomas (Parsons *et al.*, 2008) and lung cancer (Paez *et al.*, 2004) where new somatic mutations in genes were identified, offering new targets for therapy.

B-cell malignancies develop as leukaemias or lymphomas from specific stages of B-cell development (Figure 1.1) and can display chromosomal translocations involving the immunoglobulin heavy chain (*IGH*) locus on chromosome 14 (Willis and Dyer, 2000). This thesis will examine the consequences of aberrant expression of a gene, inhibitor of DNA binding 4 (*ID4*), which was initially shown to be targeted by an *IGH* translocation in cases of B-cell precursor acute lymphoblastic leukaemia (BCP-ALL), but may have oncogenic function in other malignancies.

1.2 Normal B-cell development.

B lymphocyte development begins within the foetal liver, where, in mice, two types of B-cells are produced, the B1 and B2 cells. The B1 cells are a CD5+ sub-population of B-cells, which differ from the B2 cells by their antibody production, anatomical localisation and phenotype (LeBien and Tedder, 2008). The B2 cells are part of the adaptive immune responses, involving the specific immune responses (Dorshkind and Montecino-Rodriguez, 2007). This thesis will focus on the development of the B2 cells.

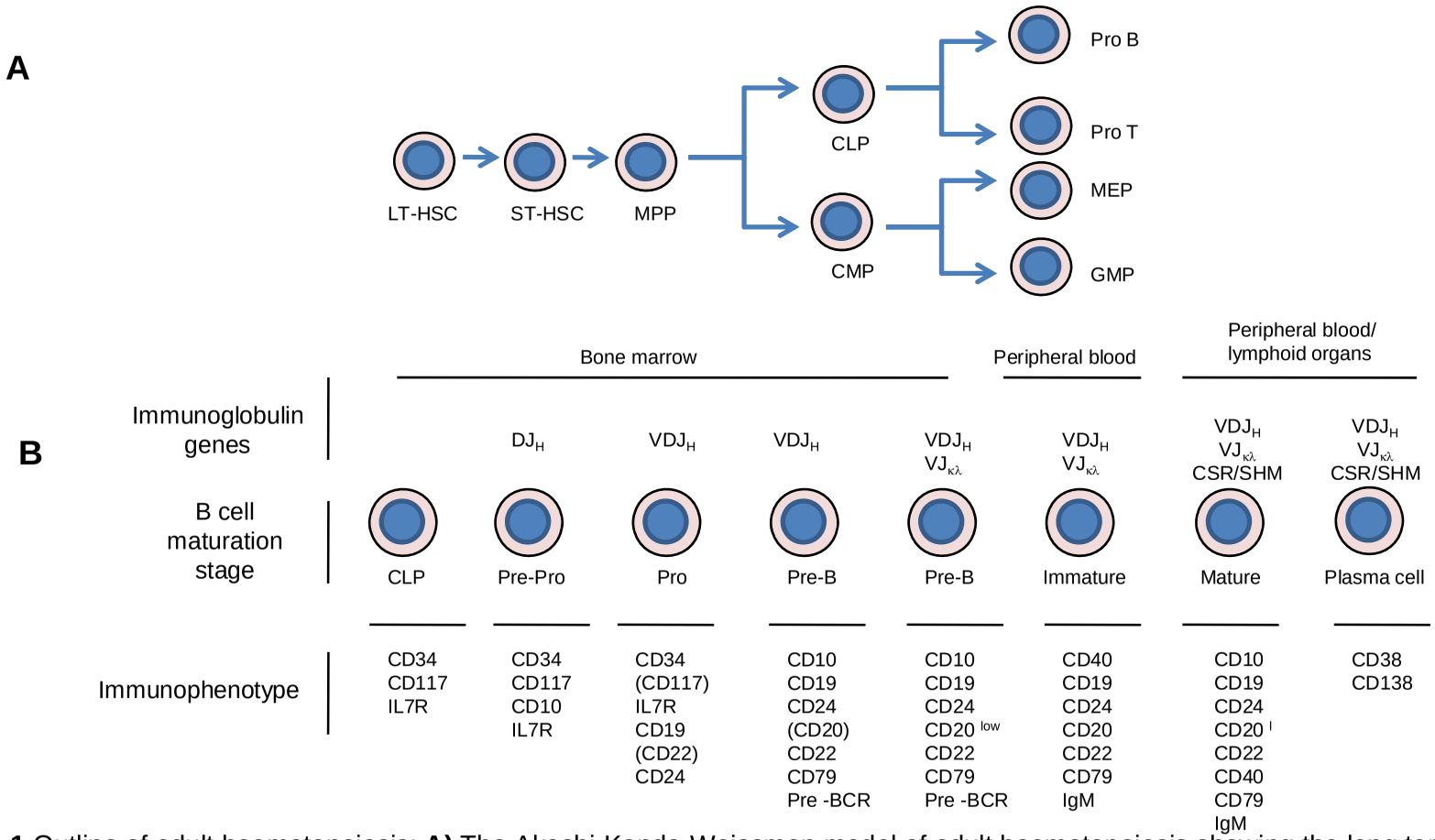


Figure 1.1 Outline of adult haematopoiesis: **A)** The Akashi-Kondo-Weissman model of adult haematopoiesis showing the long term and short term haematopoietic stem cells (LT-HSC, ST-HSC); the multipotent progenitor (MPP), the common lymphoid progenitor (CLP), common myeloid progenitor (CMP), megakaryocyte erythrocyte progenitors (MEP) and granulocyte macrophage progenitors (GMP) (Laiosa *et al.*, 2006) **B)** The expression of cell surface markers at different stages of B-cell development and the progress of immunoglobulin rearrangements (Cobaleda and Sanchez-Garcia, 2009).

B-cell development begins from the pluripotent haematopoietic stem cell (HSC) in the foetal liver and then later in the postnatal bone marrow (Hardy and Hayakawa, 2001). There are many models of haematopoiesis; one is the Akasi-Kondo-Weismann scheme (Figure 1.1 panel A).

A coordinated stepwise process drives the development of the HSC, from long term HSC (LT-HSC) to short term HSC (ST-HSC) to the multipotent progenitors (MPP) which have multi-lineage potential. These then branch off into the lymphoid primed MPPs (LMPP) which become the early lymphoid progenitors (ELP). The ELPs are the precursors for the common lymphoid progenitors which give rise to the B-cells, natural killer cells (NK), dendritic cells and T-cells. The ELP then leads to the earliest B-cell progenitor, the PrePro B-cells (Nutt and Kee, 2007) proceeding to further B-cell commitment stages (Figure 1.1 Panel B).

B-cell lineage determination from the HSC is regulated by the coordinated expression of different transcription factors. These can be seen in Table 1.1 and Figure 1.2.

The development of the CLP from the LMPP is marked by the expression of the interleukin 7 receptor (IL7R) and the FLT3 receptor (FLT3R) which further regulate development. *Il7r* and *flt3r* expression is regulated by the transcription factors PU.1 and Ikaros (Nutt and Kee, 2007). *Pu.1, ikzf1, bcl11a* and *e2a* are required for the lineage determination of B-cells, whereas genes such as *ebf* and *pax5* are required to commit to the B-cell lineage (Ramirez *et al.*, 2010).

Collectively, these transcription factors allow the expression of genes important for B-cell precursor development which include *mb-1, b29, λ5* and *vpreb* (Figure 1.2 panel B) (Nutt and Kee, 2007).

Table 1.1 Transcription factors required for B-cell development. (adapted from (Nutt and Kee, 2007)).

Transcription factor	**Function**	**Ectopic expression phenotype.**	**Knockout phenotype in mice**
PU.1	ETS family required for *il7r* expression and regulates commitment at LMPP stage between myeloid and B-cell.	In *pu.1 -/-* foetal liver cells graded amounts of *pu.1* rescue B- cell and macrophage development	Lack foetal B-cells and CLPs. Conditional knockout shows normal amount of B2 and B1 cell expansion.
Ikaros	Transcriptional activator and repressor. Involved in *flt3r* expression in LMPP.		Lack all stages of B-cell differentiation.
E2A	Required for B-cell lineage specification.	Induces cell-cycle arrest and apoptosis in T or B-cell lines. In the 70Z/3 macrophage line induces B-cell lineage conversion.	Block prior to prepro B-cells; Conditional knockout reduced survival in pre-B-cell lines. Ectopic E protein antagonist: Pro B-cell-growth arrest, decreased pro B-cell gene expression.
Early B-cell factor (EBF)	Regulates B-cell transcription	Induces B-cell differentiation in multipotent progenitors, rescues B lymphopoiesis in knockouts of *pu.1, e2a* or *Il-7r* progenitors.	Arrest at CLP to pre pro B-cell transition, no cells with *igh* gene rearrangement.
PAX5	Important for commitment to the B-cell lineage.	Impairs T-cell development and promotes T-cell lymphoma formation. Variably affects myeloid and erythroid differentiation.	Foetal liver lacks B lineage cells. Adult bone marrow block at pro-B-cell stage. Have *d-jh* but only a few proximal *v-dj*$_h$ rearrangements; Conditional knockout shows that *pax5* is required for the maintenance of B-cell fate and repression of plasma-cell differentiation.
BCL11A	Involved in early B-cell lymphopoiesis		Knockout results in no B lineage cells.

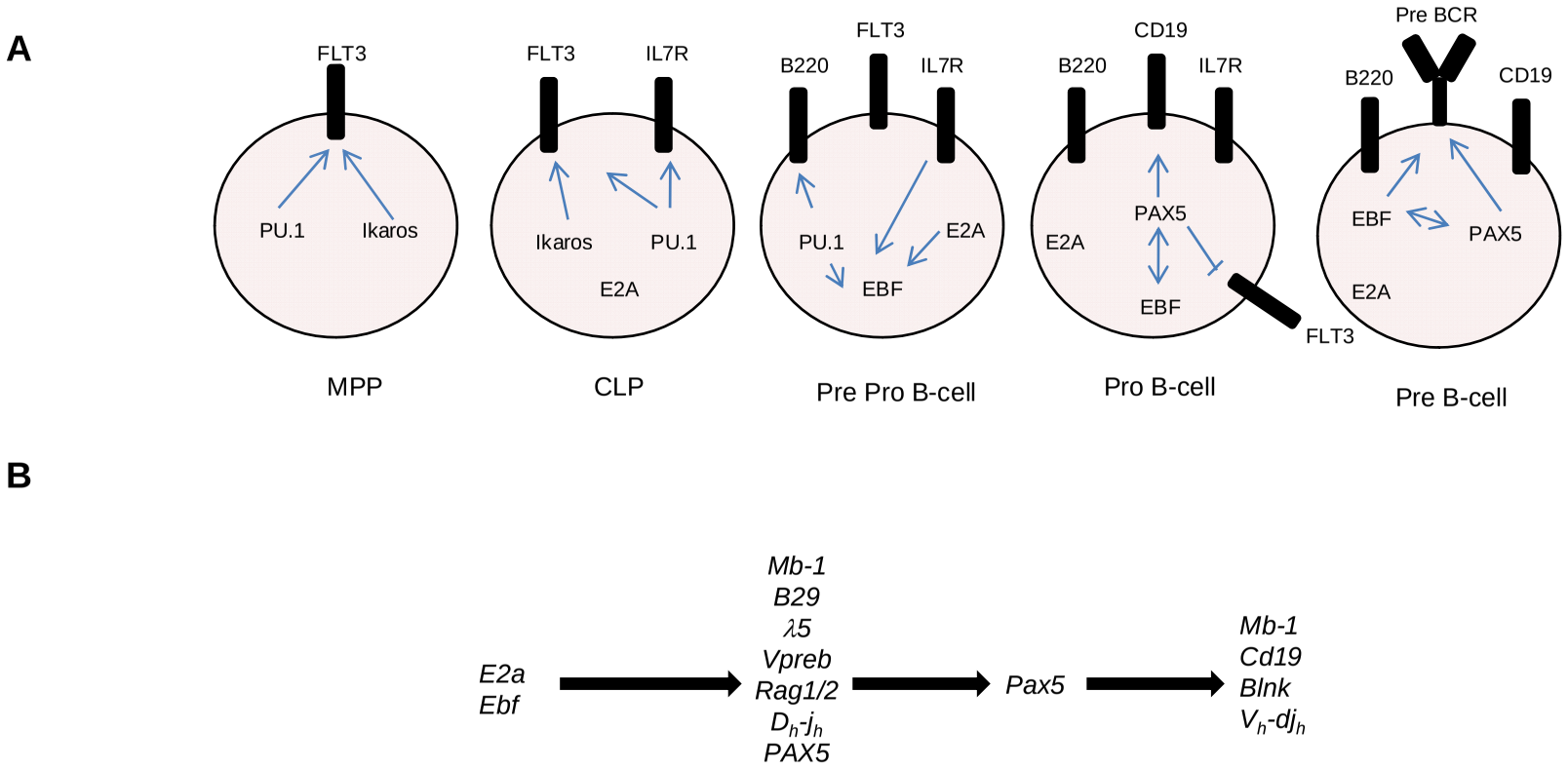


Figure 1.2 Gene regulatory networks that govern B-cell development from the HSC. **A)** The appropriate transcription factors required for early B-cell development beginning from the MPP to the Pre B-cell stage. The arrows represent transcriptional regulation of the target. The transcriptional expression hierarchy in early B-cells begins with IL7 and E2A-dependent activation of *EBF1α*. This induces the expression of further *EBF1α*. *EBF1α* induces the expression of *PAX5* which with PU.1 and Ets1 drives the expression of *EBF 1β* (Nutt and Kee, 2007; Roessler *et al.*, 2007; Thal *et al.*, 2009). **B)** Gene expression hierarchy required for B-cell development. E2A regulates the expression of *EBF* and both *E2A* and *EBF* expression is required for regulation of *PAX5* expression which seals commitment to a B-cell phenotype (adapted from Nutt and Kee, 2007; Singh *et al.*, 2005).

Cell lineage determination is regulated by a mixture of transcription factors of which the timed expression, correct levels and cross antagonism regulate lineage determination (Rosenbauer *et al.*, 2005). An example of this is the *GATA1* and *PU.1* regulation of monocytic and erythroid cells. *GATA1* expression mediates myelomonocytic cell fate by inhibiting *PU.1* expression. Conversely, *PU.1* expression in erythroid precursors represses GATA1 mediated gene expression (Graf and Enver, 2009).

Within the bone marrow, B-cell development from the common lymphoid progenitor (CLP) relies on a series of transcription and growth factors that coordinate gene expression allowing rearrangements of the immunoglobulin (*IG*) loci and resulting in the generation of a B-cell receptor (BCR) from dispersed gene segments by somatic recombination (Hardy and Hayakawa, 2001) (Figure 1.3).

The somatic rearrangement of the *IGH* locus is of particular interest. The Rag 1/2 complexes induce DNA double strand breaks within the heavy chain segments, specifically between two coding segments, at their flanking recombination signal sequences. The recombination of DNA initially produces D_H-J_H recombination followed by V_H to D_HJ_H recombination by a process of non-homologous end-joining in pro B-cells (Jung *et al.*, 2006).

This rearrangement produces a mature *IGH* gene and by a process of splicing with the constant region produces an IgM protein. This pairs with a surrogate light chain in B-cells comprising of the proteins λ5 and VpreB, forming the pre-BCR. At this point the pre B-cells (immature B-cells) undergo a pre-BCR checkpoint to test for defective *IGH* rearrangement (Hardy and Hayakawa, 2001). Correct rearrangement allows a further clonal expansion and progression of the immature B-cells to mature B-cells. At this stage there is the rearrangement of the light chains to form the mature BCR.

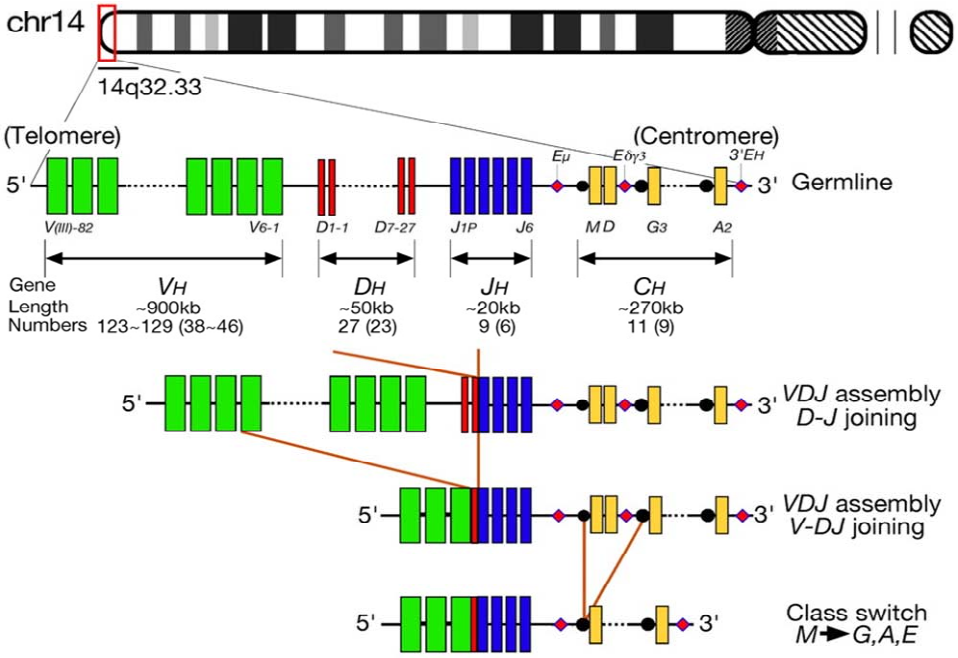


Figure 1.3 Schematic of the *IGH* locus and representations of *VDJ* recombination and CSR in the *IGH* locus**.** The human *IGH* locus is located adjacent to the telomere of chromosome 14q. The locus is composed of dispersed variable (V_H, denoted by green bars), diversity (D_H, red), and joining (J_H, blue) region segments that undergo somatic recombination in B-cell precursors to produce a functional *VDJ* gene. These are expressed with constant region gene segments (C_H, yellow) with transcription being driven by enhancers (denoted by pink diamonds) (Dyer *et al.*, 2010).

Cells that have successful light chain rearrangements and are not self reactive can mature further into transient transition T1 and T2 cells. These mature cells then undergo further testing to ensure correct rearrangement of the BCR. Many different combinations of V_H, D_H and J_H gene segments and the action of the enzyme terminal deoxynucleotidyl transferase (TdT), which adds non-templated nucleotides at the V_H-D_H and D_H-J_H junctions, contribute to the large antibody repertoire (Hardy and Hayakawa, 2001).

T1 cells have the ability to re-circulate and reside within the bone marrow and the spleen and do not express IgD. The T2 cells are B-cells that have acquired the expression of IgD. These cells can reside within the spleen and lymph-nodes and then mature to generate the follicular and marginal zones (Pillai and Cariappa, 2009).

Within the spleen and lymph nodes, mature B-cells are activated and differentiate, producing germinal centres (GC).The next round of genetic maturation, somatic hypermutation (SHM), occurs here. This ensures the production of high affinity antibodies, by targeting the complementary determining regions (CDR) on the V_H sequence for mutation. It is after SHM that B-cells can mature into memory or plasma B-cells. Further B-cell maturation involves the DNA of the *VDJ* segment rearranging to another constant region in order to produce an antibody of another class; the process is referred to as class switch recombination (CSR) (LeBien and Tedder, 2008) (Figure 1.3).

The expression of the proteins Rag1 and Rag2 are required for the *IGH* rearrangements. In mature B-cells the enzyme, activation induced cytidine deaminase (AID), is required for SHM and CSR (LeBien and Tedder, 2008).

1.3 B-cell malignancies and *IGH* chromosomal translocations.

Subversion of the normal mechanisms of *IGH* rearrangements can result in illegitimate chromosomal recombination which involve chromosomes, other than chromosome 14, transposing into the *IGH* locus. The *IG* loci contain strong transcriptional enhancers that drive deregulated expression of the juxtaposed

genes that are now translocated out of their normal environment (Dyer *et al.*, 2010).

Cytogenetic techniques, such as interphase fluorescent *in situ* genomic hybridisation (FISH), have allowed specific *IGH* translocations to be associated with subtypes of B-cell malignancies. Examples include the *CyclinD1*/ *BCL1* locus on chromosome 11 into the *IGH* locus resulting in t(11;14)(q13;q32) typical of mantle cell lymphoma (Raynaud *et al.*, 1993), t(14;18)(q32;q21) involving the *IGH* locus and *BCL2* on chromosome 18 resulting in follicular lymphoma (Akasaka *et al.*, 1998) and t(8;14)(q24;q32) involving the *MYC* gene on chromosome 8 in Burkitts lymphomas (Magrath, 1990). These target genes are over-expressed as a consequence of the translocation and have been shown to be oncogenes contributing to the pathogenesis of malignancy (Willis and Dyer, 2000). Genes may also be deregulated by other chromosomal aberrations, such as deletions and duplications and epigenetic phenomena such as promoter methylation.

1.4 B-cell precursor acute lymphoblastic leukaemia (BCP-ALL).

Acute lymphoblastic leukaemia (ALL) is a malignancy of lymphoid progenitor cells and is the most common form of childhood leukaemia, but also presents in adults. Paediatric B-cell precursor ALL (BCP-ALL) has a B-cell precursor phenotype, presents between the ages of 2 to 5 years and has a favourable outcome with an 80% cure rate. However, some cases remain resistant to therapy or relapse following therapy. In adults, this manifests at between 15-25 years and over 75 years of age. Adult BCP-ALL presents with 78-93% cure rate, but has an increased relapse rate compared to paediatric BCP-ALL (Fielding *et al.*, 2007). The cause of ALL remains unknown, but there are a number of epidemiological and genetic predisposing factors (Greaves, 2006; Papaemmanuil *et al.*, 2009; Pui *et al.*, 2008)

Approximately less than 5% of ALLs are associated with inherited disorders, including Downs Syndrome, Blooms syndrome, Ataxia-telangiectasia and Nijmegen breakage syndrome (Pui *et al.*, 2008). These are syndromes associated with chromosomal aberrations and genetic instability partly due to

defective DNA repair mechanisms. High birth weight (Hjalgrim *et al.*, 2003), residential magnetic fields (Ahlbom *et al.*, 2000), parental occupation, maternal reproductive history, parental alcohol or tobacco use, maternal diet and exposure to pesticides and solvents and exposure to ionising radiation or chemotherapeutic drugs have all been suggested to increase the risk to ALL(Ahlbom *et al.*, 2000; Biondi *et al.*, 2000; Greaves, 2006; Hjalgrim *et al.*, 2003; Pui *et al.*, 2008; Spector *et al.*, 2005).

This increased susceptibility has been associated with mutations, instability and defective repair mechanisms in DNA which are induced by environmental or man-made causes. Unborn children can be affected by the damage to parental DNA and can be exposed *in utero* to environmental damage.

The ability of children to mount an effective immune response may affect the susceptibility of an individual to leukaemia (Greaves, 2006). An infection-based hypothesis for ALL has been suggested by both Kinlen and Greaves (Greaves, 2006; Kinlen, 2004). In affluent societies, where there may be an absence of common infections during childhood, an exposure to common pathogens at a later date may result in an aberrant response from the immune system, predisposing the individual to ALL (Greaves, 2006).

Kinlen proposed the ‘population mixing’ (infective) hypothesis, based on the finding that viruses cause several animal leukaemias. In this model susceptible individuals who show rare responses to infection are generally exposed to common infections by population mixing with carriers, such as that seen with the urbanisation of rural areas in the last 60 years (Kinlen, 2004).

Greaves proposed a ‘delayed infection’ hypothesis. This relies on a minimal two hit model whereby an individual has a genetic susceptibility to developing ALL i.e. a pre-leukemic clone which is transformed by a secondary event.

The hypotheses by both Greaves and Kinlen suggest a susceptibility which may be in the form of a pre-leukemic clone which develops into leukaemia after acquiring further ‘driver’ mutations.

This has been demonstrated with the *ETV6/RUNX1* (also known as *TEL/AML1)* translocation as a pre-leukemic initiating event in paediatric ALL, *in utero*. This is then proceeded by further mutations in genes involved in B-cell development or the cell cycle (Bateman *et al.*; 2010; Hong *et al.*, 2008) generating leukaemia.

Twins harbouring this pre-leukemic clone, show a common single cell origin which in one twin can manifest as an overt leukaemia and leave the other twin unaffected. This clone has the ability to self-renew as well as differentiate (Bateman *et al.*; 2010; Hong *et al.*, 2008).

1.4.1 Common genetic aberrations in BCP-ALL.

BCP-ALL is a clonal malignant disease defined by an accumulation of blasts at a particular stage of B-cell differentiation (Cobaleda and Sanchez-Garcia, 2009). BCP-ALL theoretically originates from the cell where the first genetic lesion took place, as was demonstrated by the *ETV6/RUNX1* study described above. The common genetic events in adult and paediatric BCP-ALL are outlined in Figure 1.4. Defining the cell of origin in BCP-ALL is difficult because of the accruing of further somatic mutations which can arise *in utero* or later in adult life (Cobaleda and Sanchez-Garcia, 2009).

1.4.2 BCP-ALL and cooperating mutations.

Whole genome wide analysis studies in BCP-ALL have identified recurring mutations and deletions in genes which are required for B-cell development (Mullighan *et al.*, 2007). These mutations cooperate with the initial genetic event to form leukaemia. This includes genes such as *PAX5*, *EBF*, *TCF3, IKZF1,IKZF3, BLNK, CEBPE* and *ARID5B* (Mullighan *et al.*, 2007; Papaemmanuil *et al.*, 2009).

The disruption of early B-cell development by the targeting of specific genes is found in approximately 40% of BCP-ALLs (Mullighan *et al.*, 2007). *PAX5* is required for commitment to the B-cell lineage and has been found to be targeted for somatic deletion in 37.1% of BCP-ALL cases.

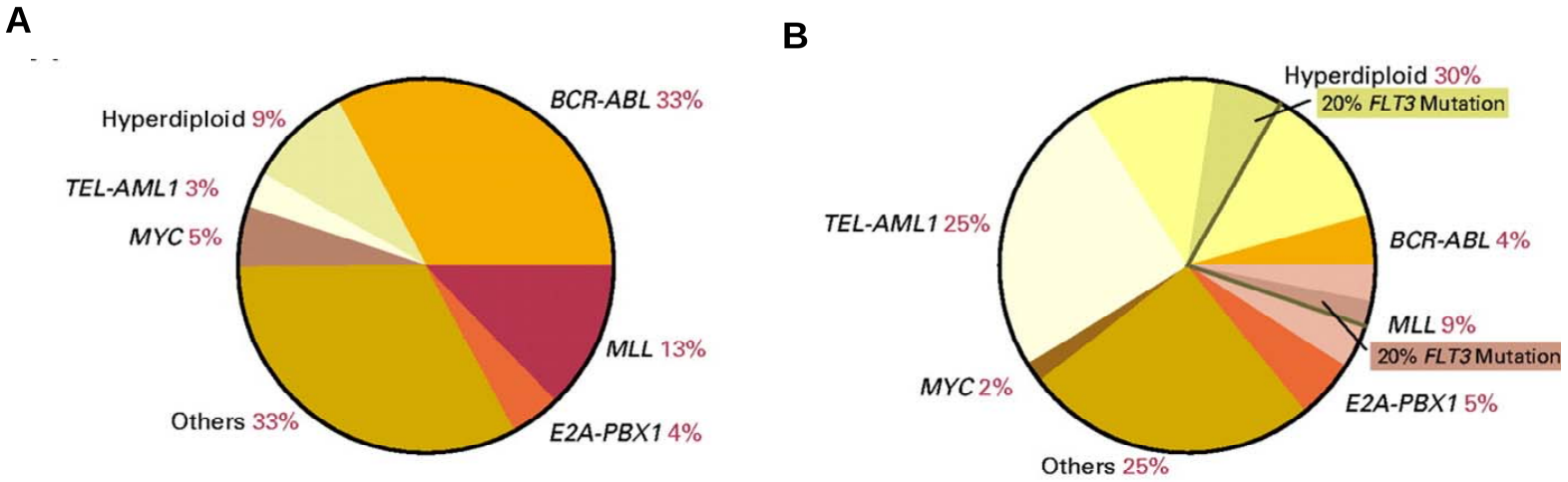


Figure 1.4 Molecular genetics of BCP-ALL in **A)** adult **B)** paediatric BCP-ALL (adapted from Armstrong and Look, 2005).

TEL-AML also referred to as *ETV6/RUNX1*

Four types of deletions have been found involving the *PAX5* locus, including deletions of *PAX5* (both homozygous and heterozygous) and flanking genes and large chromosome 9 or 9p deletions or the deletion of all of chromosome 9. These *PAX5* deletions were present in the majority of the cases in the dominant leukemic clone (Mullighan *et al.*, 2007).

PAX5 has also been found to be targeted by translocations or to have specific point mutations. The point mutations result in the loss of function of the DNA binding and/or the trans-activation domains of PAX5*. PAX5-ETV6, PAX5-FOXP1* and *PAX5-ZNF251* translocations have also been found (Mullighan *et al.*, 2007).

EBF, another gene important for the regulation of B-cell development (see Figure 1.3) has been found to be deleted in approximately 3% of BCP-ALLs and presents with both mono-allelic deletions and bi-allelic deletions (Mullighan *et al.*, 2007).

Recurrent mono and bi-allelic deletions in the *CDKN2A* and *CDKN2B* loci have also been implicated in B and T-cell ALLs as secondary events (Bateman *et al.*; 2010; Sulong *et al.*, 2009), especially with the *ETV6/RUNX1* translocation cases (Bateman *et al.*; 2010; Mullighan *et al.*, 2007; Sherborne *et al.*). The deletions of chromosome 9q21 in many leukaemias are thought to arise because of the aberrant Rag protein activity in these cells (Kohno and Yokota, 2006; Raschke *et al.*, 2005). The *CDKN2A* locus which encodes the proteins p16 and p14 Arf, is important in controlling the G1/S phase of the cell cycle.

1.4.3 IGH translocations in BCP-ALL.

Approximately one third of individuals with paediatric BCP-ALL lack recurrent chromosomal translocations and the pathogenesis of this disease subtype is not well understood (Mullighan *et al.*, 2009). This implicates other genetic aberrations. Genome wide profiling is an invaluable tool for the identification recurring mutations, however this technique does not detect aberrant *IGH* translocations.

Studies in the Dyer laboratory (Dyer *et al.*, 2010) have identified several *IGH* translocations representing approximately 3% of total BCP-ALLs screened. The results of this screen can be seen in Table 1.2.

IGH translocations are uncommon in BCP-ALL, but some of the genes identified are also deregulated by other mechanisms.

For example, the cytokine receptor-like factor 2 gene (*CRLF2*) encodes a receptor for the thymic stromal lymphopoietin (TSLP) cytokine. This cytokine is involved in B-cell lymphopoiesis and proliferation (Astrakhan *et al.*, 2007). *CRLF2*, as well as being targeted by *IGH* translocation (Russell *et al.*, 2009), is also deregulated in BCP-ALL by other mechanisms. Mullighan *et al.* have demonstrated the *P2RY8-CRFL2* translocation in 7% of BCP-ALLs and in 58% of Down syndrome BCP-ALLs (Mullighan *et al.*, 2009). *CRLF2* genetic aberrations have been found in several studies which confirm an association with high risk BCP-ALL. *CRLF2* aberrations correlate with mutations in *JAK1*, *JAK2* and with *IKZF1* deletions (Cario *et al.*, 2010; Harvey *et al.*, 2010; Mullighan *et al.*, 2009; Yoda *et al.*, 2009).

1.4.4 t(6;14)(p22;q32).

The t(6;14)(p22;q32) translocation involves the gene, *ID4,* with the *IGH* locus (Table 1.2). ID4 is a member of the inhibitor of DNA binding (ID) family of proteins which has three other members (ID1, ID2 and ID3). These proteins are characterised by their helix loop helix domain (HLH), which is required for hetero-dimerisation with transcription factors containing the same domain and thereby antagonising them. ID1, ID2 and ID3 are required in normal developing tissues and have been implicated as oncogenes in many malignancies (Perk *et al.*, 2005). These genes will be discussed in more detail in section 1.7.

Table 1.2 Summary of *IGH* translocations identified in BCP-ALL(adapted from Dyer *et al.*, 2010).

Target Gene	Chromosomal location
CRLF2	Xp22 Yp22
EPOR	19p13
ID4	6p21
CEBPD	8q11
CEBPA	19q13
CEBPE	14q11
CEBPB	20q13
CEBPG	19q13
IL3	5q31
IGF2BP1	17q21.3
MiR125b	11q24.1
BCL9	1q21
hTERT	5p15.3
LHX4	1q21
DUX4	10q24
BCL2	18q21

The t(6;14)(p22;q32) translocation was initially identified by Bellido *et al.* using FISH and inverse PCR in one case of adult BCP-ALL and showed *ID4* RNA expression. This study also identified the over-expression of ID4 protein by immunohistochemical staining in four other cases of BCP-ALL which harboured other genetic abnormalities such as a t(8;14) translocation, a *P53* mutation, 46XY12p- and hyperdiploidy. These cases all showed a nuclear staining pattern for ID4. There were also two cases of T-cell prolymphocytic leukaemia which showed *ID4* RNA expression. The sub-cellular localisation of ID4 protein in these patients was nuclear (Bellido *et al.*, 2003).

The involvement of *ID4* in an *IGH* translocation was identified by Russell *et al.* in 13 patients with BCP-ALL (Russell *et al.*, 2008). The t(6;14)(p22;q32) was cloned from 3 cases using long distance inverse PCR (LDI-PCR) and showed the translocation of *IGHJ* into the centromeric 3' portion of the *ID4* gene resulting in the over expression of *ID4* RNA (Figure1.5). These cases were from adult or adolescent BCP-ALL, but overall, presented with a good prognosis (Russell *et al.*, 2008).

Of these cases, 12 patients also presented with an abnormality on the short arm of chromosome 9 with the monoallelic or biallelic deletion of *CDKN2A* and the mono allelic deletion of *PAX5*. The reduced levels of *PAX5* RNA were confirmed by real time PCR (Russell *et al.*, 2008). All 10 patient samples examined, harboured deletions of *CDKN2A* and *PAX5,* implicating these two genes in the transformation pathway together with *ID4*.

In an independent study with collaborators in Kiel (Siebert, personal communication), approximately 11.5% (n=297) of BCP-ALL cases harbouring *ID4* RNA expression have been identified. On total gene expression analysis, these cases all cluster together showing that *ID4* RNA is expressed in a subtype of BCP-ALL. These cases are both with and without an *IGH* translocation (Table 1.2).

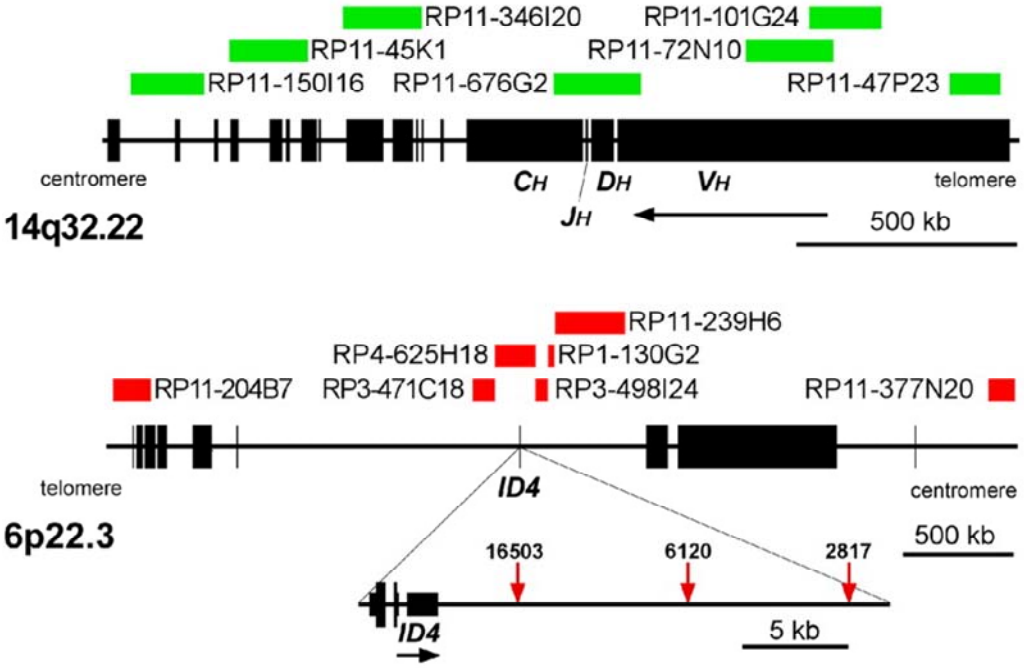


Figure 1.5 Involvement of *ID4* in the translocation, t(6;14)(p22;q32). Diagram showing the location of clones used as dual-colour, dual-fusion FISH probes (red and green) and breakpoints (red arrows) cloned by LDI-PCR from *IGHJ6* segments (adapted from Russell *et al.*, 2008).

As can be seen in Table 1.2 the *CEBP* family have all been implicated in BCP-ALL as targets of *IGH* translocations. However, *ID4* is the only member of the ID family to be involved with this disease indicating that a feature unique to *ID4* is important for the pathogenesis of BCP-ALL. Therefore, by analogy with the other *IGH* chromosomal translocations, *ID4* appears to be a dominant oncogene in a subgroup of BCP-ALL.

The current literature and preliminary gene expression experiments therefore suggest that *ID4* may be a tumour suppressor or an oncogene in a context dependent manner. The aim of this project was to investigate whether *ID4* is an oncogene in a B-cell context.

1.5 Chronic lymphocytic Leukaemia (CLL).

ID4 expression in normal and malignant haematological tissues is limited. Yu *et al.* saw no *ID4* RNA expression in primary CLL cases (Yu *et al.*, 2005). *ID4* RNA expression in primary CLL cases requires further investigation and this is addressed in this thesis.

CLL is a disease of mature B-cells and is defined by several prognostic markers. One of these includes whether or not the *IGH* variable regions have undergone SHM and hence cases are defined as mutated or unmutated. The latter have a poorer prognosis. Recurrent cytogenetic abnormalities also define prognostic subgroups of CLL and include deletions of 13q, 11q and 17p deletions as well as trisomy of chromosome 12 (Zenz *et al.*, 2009).

Deletions of 13q are associated with a favourable prognosis and this deletion targets *mir-15a* and *mir-16-1*. The 11q deletions target the Ataxia telangcectasia-mutated gene *(ATM)*, a tumour suppressor gene involved with the repair of double strand DNA breaks. Deletions of 17p target the tumour suppressor, *P53*, and are associated with a poor prognosis. Trisomy 12 is also associated with a poor prognosis, however the specific genes affected are unknown (Zenz *et al.*, 2009).

1.6 Helix loop helix (HLH) proteins.

The ID proteins belong to a class of HLH proteins involved in the development of many cell types. There are four members of this family, ID1, ID2, ID3 and ID4. The literature on the ID proteins mainly focuses on ID1, ID2 and ID3.

The ID protein's key motif is the HLH domain which allows them to dimerise with other HLH domain containing proteins and regulate their function (Massari and Murre, 2000). There are seven classes of HLH proteins categorised by their tissue distribution, dimerisation and tissue binding abilities and these are summarised in Table 1.3 below.

The class I HLH proteins, also known as basic HLH (bHLH) proteins and E proteins, contain a motif of positively charged amino acids adjacent to the HLH domain and bind to the E box sequence CANNTG. These proteins have the ability to homodimerise or heterodimerise and their binding specificity is limited to E boxes. The class II proteins are also capable of forming homodimers and heterodimerise with the class I proteins. The class IV proteins interact with the class III proteins. The class V proteins (including ID4) lack a DNA binding domain and ID1, ID2 and ID3 have been shown to negatively regulate class I and II proteins (Massari and Murre, 2000) and therefore affect differentiation of many tissues (Perk *et al.*, 2005)

Table 1.3 The classes of the HLH proteins with their specific domains and tissue distribution (adapted from (Massari and Murre, 2000).

Class	Examples of Proteins	Specific domains	Tissue distribution
I	E2A (E12/E47), TCF12, TCF4	Basic DNA binding domain	Ubiquitously expressed
II	Myo D, Neuro D, ABF-1	Basic DNA binding domain	More tissue specific expression
III	MYC family of transcription factors	Basic DNA binding domain, Leucine zipper frame	Ubiquitously expressed
IV	Mad, Max Mixi, c-MYC	Basic DNA binding domain	Ubiquitously expressed
V	ID1, ID2, ID3 and ID4	No basic DNA binding domain	Varied tissue specific expression
VI	Hairy (Drosophila)	Basic DNA binding domain and a proline in the DNA binding domain.	
VII	Aromatic hydrocarbon receptor	Basic DNA binding domain with a PAS domain.	

1.7 E proteins.

The E proteins consist of E12 and E47 (collectively known as E2A, but also known as TCF3; ITF1; VDIR; bHLHb21; MGC129647; MGC129648), TCF12 (also known as HEB; bHLHb20; HsT17266; HTF4) and TCF4 (also known as E2-2;PTHS; SEF2; SEF2-1; SEF2-1A; SEF2-1B; bHLHb19; MGC149723; MGC149724; ITF2). These proteins are important for haematopoiesis and are regulated by ID1, ID2 and ID3 proteins. The role of these E proteins in general haematopoiesis is outlined below. The protein structures of E proteins and ID proteins is shown in Figure 1.6.

1.7.1 E2A.

The E2A proteins E12 and E47 are both transcribed from the gene *TCFE2A* and are generated as the result of alternative splicing. They also differ in their binding affinities to E box sequences.

In B-cell development *e2a* expression becomes active at the CLP stage (Semerad *et al.*, 2009), progressing to the highest levels within the Pre Pro B-cell stage, but it is reduced again once the cells mature through the pre-BCR checkpoint (Quong *et al.*, 2004). Levels are raised again at Pre B-cell stage. The conditional knockout of *e2a* results in reduced numbers of cells in the GC which proliferate normally with normal levels of *aid*, SHM and CSR (Kwon *et al.*, 2008).

E2A is required for *Igh* recombination by regulating the expression of *rag1* and *rag2* by binding to the Erag enhancer present in the *rag2* promoter (Hsu *et al.*, 2003). E2A also regulates CSR and SHM by regulating *aid* expression (Kwon *et al.*, 2008). E2A protein binds the HS1 enhancer in the 5' region of the v_h gene segment; this region is also bound by PAX5 and PU.1 and can regulate v_h gene usage in T-cells (Bain *et al.*, 1999).

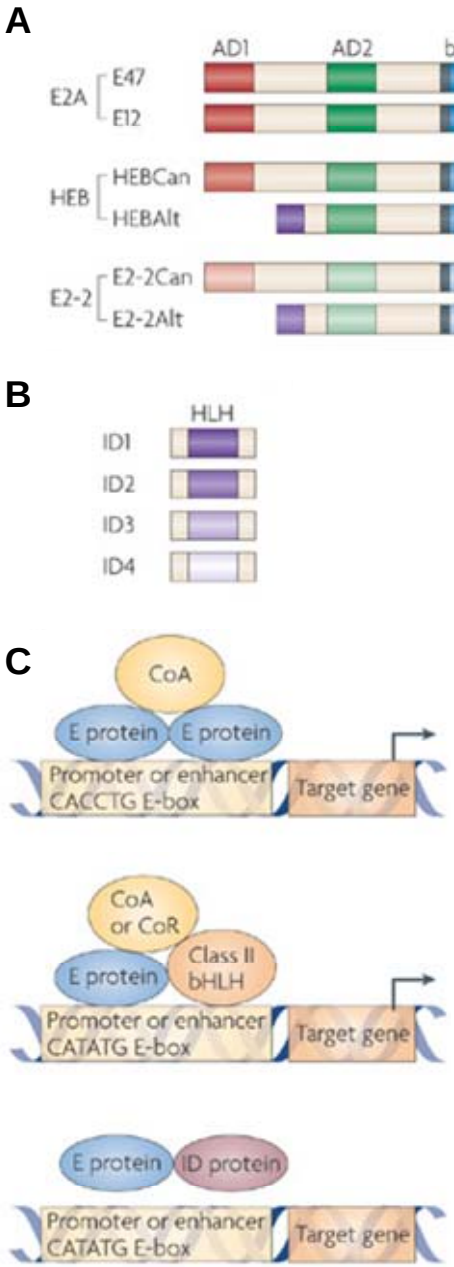


Figure 1.6 Schematic representation of the HLH proteins and their transcriptional activities. **A)** E proteins showing the key protein domains activation domain 1 (AD1) in red, activation domain 2 (AD2) in green and bHLH domain in blue. Alternative transcription start sites give rise to canonical (Can) and alternative (Alt) forms of TCF12 (HEB) and TCF4 (E2-2) proteins. **B)** The four ID proteins with their HLH domains in purple. **C)** The transcriptional activities of the E proteins are determined by the AD1 and AD2 domains. AD1 can bind transcriptional activators through the recruitment of co-activators (CoA) such as p300. E proteins interact with tissue specific class II HLH proteins in target genes promoters and function as transcriptional activators or repressors depending on the proteins that are recruited (co-activators or co-repressors (CoR, such as ETO or mSin3A) (middle). E protein–ID protein heterodimers fail to bind DNA and do not activate gene transcription (bottom) (Kee, 2009; Murre, 2005).

E2A protein also regulates *igl* κ chains by cooperatively binding with the protein interferon regulatory factor 4 at the κ*e3'* enhancer regions. A disruption of these regions impairs κ recombination in pre B-cells (Lazorchak *et al.*, 2006).

From the Pro B-cell stage E2A regulates the expression of *ebf*. *Ebf* gene expression is regulated by 2 promoters ebf 1 α and ebf 1β. The ebf1α promoter is active in Pro B-cells and is auto-regulated and also regulated by E2A. The *ebf1*β promoter is active in mature B lymphocytes and is regulated by Ets1, PU.1 and PAX5 (Roessler *et al.*, 2007). Ectopic expression of *pax5* in *e2a -/-* progenitors is capable of promoting the development of pro B-cells, possible because of the ability of PAX5 protein to regulate *ebf* expression (Kwon *et al.*, 2008).

E2A is therefore required in early and late B-cell development to regulate the expression of B-cell appropriate genes required for the B-cell lineage development and for the development of the BCR.

The ectopic expression of *ebf* is capable of inducing B-cell lineage gene expression in *e47-/-* progenitors, through interaction with the other E proteins, TCF12 and TCF4. This demonstrates the ability of the E proteins TCF12 and TCF4 to partially compensate for the absence of E2A. However in *e2a-/-* mice , ectopic *ebf* expression in pro B-cells results in a failure to proliferate in response to IL7 because of the requirement for E2A to induce *n-myc* via IL7 R signalling (Seet *et al.*, 2004).

Ebf knockout experiments demonstrate that *ebf* is crucial (Kee, 2009) for B-cell lineage commitment because it can repress myeloid gene expression such as that of *pu.1* and *cebpa*. This block in differentiation can be rescued by the ectopic expression of *ebf* or *pax5*, but not *e2a* (Pongubala *et al.*, 2008) .

E2a -/- mice also show a reduced population of CLP and as a consequence are deficient in mature B-cells, specifically a reduced B220+ CD43 progenitor population (Bain *et al.*, 1994; Kee, 2009; Massari and Murre, 2000; Seet *et al.*, 2004). The resulting reduced LT-HSC compartment affects the MPP's and the CLP. There is an increase in the expression of *gata1* that induces a

megakaryocyte phenotype and therefore there is an increase in this population. As well as a reduced HSC compartment, the HSC's show an increased cell cycle rate which can be attributed to a lack of p21, which E2A can up-regulate (Semerad *et al.*, 2009). *E2a +/-* mice present reduced levels of B-cell development, demonstrating a dose dependent response to E2A proteins in B-cell development.

E2A proteins have a regulatory function in early B-cell development and in GC development, but a portion of this transcriptional activity can be attributed to EBF which can regulate the expression of *pax5* and therefore, collectively, expression of these two proteins commiT-cells to the B-cell lineage. E2A and EBF proteins are important for the expression of B-cell lineage specific genes and therefore determine B-cell differentiation.

PAX5 is crucial for the commitment to the B-cell lineage (Nutt *et al.*, 1999). In *pax5-/-* adult mice, B-cell development proceeds up to the Pro B-cell stage, but B-cells maintain the ability to differentiate into other lineage if exposed to inappropriate cytokines and gene expression (Nutt *et al.*, 1999). Therefore, *pax5* is crucial for maintaining the committed stage in B lymphopoiesis (Cobaleda *et al.*, 2007). Impaired *PAX5* expression and function has been implicated in many BCP-ALL cases (Mullighan *et al.*, 2007).

E2A also affects T-cell development and works in concert with TCF12 to promote T-cell development. The regulation of *e2a* expression in T and B-cell development is governed by Notch signalling (Kee, 2009). High Notch expression correlates with T-cell development. Notch1 signalling determines B-cell fate by degrading the JAK kinases (involved in the IL7 signalling pathway) and E2A proteins (by MAPK signalling which is highly expressed in B but not T-cells) (Nie *et al.*, 2008). However, the regulation of *e2a* expression by Notch1 in determining cell fate may also require other processes (Kee, 2009; Nie *et al.*, 2008).

1.7.2 TCF12.

TCF12 (also known as HEB; bHLHb20; HsT17266; HTF4) has two alternative isoforms, a TCF12 canonical (TCF12 can) and TCF12 alternative (TCF12 alt) (Figure 1.6). TCF12 alt lacks the first AD1 domain and has a unique N terminus (Figure 1.6) (Wang *et al.*, 1996).

The role of *tcf12* in haematopoietic development has been examined in knockout mice by Zhuang *et al*. (Zhuang *et al.*, 1996). Each of the E proteins was knocked out and the impact of haematopoiesis examined. *Tcf12 -/-* mice die two weeks after birth, whereas *tcf12+/-* mice appear normal. When all three E proteins (Table 1.3) were knocked-out, the foetal liver examinations revealed a decline in Pro B-cells, with *e2a -/-* showing the most profound defect. Heterozygote knockouts of *e2a* and *tcf12* appear to affect CD4 and CD5 expression on developing thymocytes (Zhuang *et al.*, 1996).

TCF12 and E2A heterodimerise and bind to CD4 promoters involved in T-cell development. Bardant *et al.* designed a construct of *tcf12* with a disrupted DNA binding domain which ablates both TCF12 and E2A function. This study showed mice expressing this construct demonstrated a block in T-cell development at the double negative to immature single positive stage and in the double positive to single positive stage. TCF12 in cooperation with E2A is required for T-cell differentiation rather than commitment. *Tcf12* mutation also mildly affected the numbers of pro, pre and immature B-cells and a reduction in NK cells was also been observed (Barndt *et al.*, 2000).

The role of *tcf12* appears to be in the further differentiation of T-cells, but *e2a* is required for the commitment at the DN1 stage of development, by regulation of $v\beta$ germ-line transcription (Jia *et al.*, 2008).

1.7.3 TCF4.

E2A appears to be the key E protein involved in regulating lymphopoiesis, however the other E proteins, TCF12 and TCF4, play a compensatory role in the absence of E2A (Seet *et al.*, 2004; Zhuang *et al.*, 1996).

TCF4 (also known as E2-2;PTHS; SEF2; SEF2-1; SEF2-1A; SEF2-1B; bHLHb19; MGC149723; MGC149724; ITF2) has two isoforms (Figure 1.6) and is important for plasmacytoid dentdritic cells (pDc) development. Whilst pDc cells express all three E proteins, TCF4 is most abundant. Knockout studies have revealed that TCF4 regulates gene expression for pDc development. This includes genes such as *dntt, mgl1, ldhb, ccr9* and most importantly s*piB,* an ETS factor crucial for pDC signalling. Haploinsufficiency of *tcf4* has been detected in the autoimmune disease Pitt-Hopkins syndrome, resulting in aberrant interferon responses in pDC's (Cisse *et al.*, 2008; Nagasawa *et al.*, 2008).

1.8 The ID proteins.

The cloning of the 3 *Id* family members *id1* (Benezra *et al.*, 1990), *id2* (Sun, 1994) and *id3* (Christy *et al.*, 1991) identified these proteins as inhibitors of differentiation and stimulators of proliferation in many tissue types (Perk *et al.*, 2005). *Id4* was cloned by screening cDNA libraries from mouse bone marrow (Riechmann *et al.*, 1994). There is extensive literature on ID1, ID2 and ID3 proteins, but limited information on ID4. An overview of the, RNA and protein data of the *id* genes can be seen in Table 1.4.

The expression of all the *id* genes has been shown during embryonic development and they have overlapping expression (*id4* expression is more restricted) (Jen *et al.*, 1996). The roles of the *id* genes have been defined in mouse knockout studies and their overlapping function is evident, with only the double and triple knockouts exerting lethal effects on development (Fraidenraich *et al.*, 2004; Lyden *et al.*, 1999). The double knockouts of *id1* with *id3* present with embryonic lethality and have cranial haemorrhages, small brain size due to fewer proliferating cells and abnormal vascularisation (Lyden *et al.*, 1999).

These features are due to abnormal angiogenesis, an increase in neurological class II HLH factors (*math1,math3,mash-1, ngn1* and *ngn2*) and increases in p16 and p27 expression in these cells. This resulted in premature differentiation and the withdrawal of the neuroblasts from the cell cycle (Lyden *et al.*, 1999).

Table 1.4 Overview of the genomic, RNA and protein features of the *ID* family.

ID protein	Genomic location	NCBI accession numbers	Coding sequence (base pairs)	Predicted molecular weight (kDa)	Other names
ID1	20q11	ID1a=NM_002165.2 ID1b=NM_181353.1	467 449	ID1 a 17 ID1 b 16.4	ID; bHLHb24
ID2	2p25	NM_002166.4	404	14.7	GIG8; ID2A; ID2H; bHLHb26; MGC26389
ID3	1p36.13-p36.12	NM_002167.3	359	13	HEIR-1; bHLHb25
ID4	6p22-p21	NM_001546.2	485	17.7	IDB4; bHLHb27

Triple knockouts of *id1, id2* and *id3* resulted in embryonic lethality (Fraidenraich *et al.*, 2004).

1.8.1 *ID4.*

The human *ID4* has two coding exons and produces a protein predicted to be approximately 17kDa. In mice, *id4* RNA is expressed predominately in developing and mature neuronal tissue and differs from the expression patterns of *id1*, *id2* and *id3* which overlap in many different tissues (Jen *et al.*, 1996). *Id4* RNA expression is normally found in the normal brain, the bladder and the bone marrow and there is limited expression in other haematopoietic tissues.

Id4 -/- mice show a 20% mortality rate. The adults present with smaller brains due to impaired proliferation and increased apoptosis of neuronal precursor cells and increased differentiation of the astrocytes (Bedford *et al.*, 2005; Riechmann *et al.*, 1994; Yun *et al.*, 2004). In neural stem cells *id4* expression is required for the G1/S phase transition and the knockout mice showed an impaired cell cycle (Yun *et al.*, 2004).

1.9 The ID proteins in normal haematopoiesis.

While there is no embryonic lethality associated with individual *id* knockout phenotypes, *id1, id2* and *id3* genes appear to play an important role in general development (including angiogenesis). For the purposes of this thesis the roles of the ID proteins in haematopoietic development and in determining B-cell lineage commitment will be discussed.

1.9.1 Id1 in normal haematopoiesis.

Id1 RNA is expressed at the LT-HSC stage, but is not at the LMPP stage. GMP's express *id1* RNA, but there are low to no levels of *id1* RNA expression in CLP's; MEP's and CMP's (Cochrane *et al.*, 2009; Jankovic *et al.*, 2007; Leeanansaksiri *et al.*, 2005).

The ectopic expression of *id1* in HSC's promotes the formation of the myeloid lineage extending to macrophages, but not granulocytes. Myeloid promoting

cytokines such as IL3 and GM-CSF when used to treat multipotent progenitors induce *id1* expression (Cochrane *et al.*, 2009; Leeanansaksiri *et al.*, 2005).

Id1 over-expression represses lymphoid development and promotes myeloid development. There are no major effects on haematopoiesis when *id1* expression is perturbed, but when all the ID proteins and other bHLH proteins such as SCL/TAL1 are inhibited with the chimaeric ET2 protein (an inhibitory protein which has the ability to inhibit ID proteins and SCL/TAL1 proteins from antagonising E proteins) there is a block in the myeloid lineage. This effect is only apparent at the LMPP stage and not at further committed stages of development (Cochrane *et al.*, 2009).

There are two conflicting roles of *id1* in haematopoietic development suggested here. *Id1* over-expression favours myeloid lineage specification over lymphoid, but in the absence of *id1* there is a premature commitment to myeloid differentiation. A recent study by Suh *et al.* has demonstrated that *id1 -/-* mice also have defects in the bone marrow micro-environment. When cells from the bone marrows of wild type mice are injected into these *id1 -/-* mice, the effects on haematopoiesis are the same as those exhibited in *id1 -/-* mice (Suh *et al.*, 2009). Therefore *id1* has an important effect not only on the myeloid specification of these cells, but on the environment and growth factors required to nurture them.

Id1 -/- mice show increased numbers of granulocytes, monocytes, erythrocytes and platelets but the bone marrow mice show overall reduced cellularity. The number of HSC remains the same but there is evidence of a reduced self-renewal capacity and an increase in numbers of cells in the S phase. There is no overall change observed in the number of LT-HSC and ST-HSC and the other compartments appear unaffected but there is an increased population of MEP's with reduced GMP's. There is an overall premature commitment to myeloid differentiation in the absence of *id1* in haematopoiesis (Jankovic *et al.*, 2007; Suh *et al.*, 2009).

The *id1 -/-* phenotype partially improves the mortality of *e2a-/-* mice, but cannot reverse the block in B-cell development and the generation of T-cell tumours

(Yan *et al.*, 1997). The *id1* -/- phenotype also induces skin cancers which are attributed to the 50% reduction in cutaneous γ/δ T-cells. This reduction is due to a defect in the homing mechanism of these T-cells possibly due to a down-regulation of the chemokine receptor, CXCR4 (Sikder *et al.*, 2003).

1.9.2 ID2 in normal haematopoiesis.

Id2 -/- mice demonstrate defects in lymph node formation, Peyers patches formation and the development of langerhans cells and splenic dendritic cells. (Hacker *et al.*, 2003; Ji *et al.*, 2008; Perk *et al.*, 2005). Other consequences of *id2* knock out are hydronephrosis, enteric neoplastic lesions and defects in mammary proliferation (Perk *et al.*, 2005).

Ectopic *id2* expression blocks B-cell development at the pro B-cell stage recapitulating the *e2a* -/- phenotype. Endogenous *id2* RNA levels have been detected in HSC and CMP's, GMP's express lower levels of *id2* RNA and MEP's express higher levels. *Id2* ectopic expression in the correct conditions induces erythroid lineage differentiation in CMP's and MEP's. In this instance ID2 protein does not work by antagonising E2A, but by inhibiting PU.1 which allows the transcription factor GATA1 to drive erythroid transcriptional activity (Ji *et al.*, 2008).Other functions of ID2 in the haematopoietic system have been demonstrated in mature B-cells in the spleen (Figure 1.7).

1.9.3 ID3 in normal haematopoiesis.

Id3 and *id2* expression is down-regulated in Pre Pro B-cells by EBF (Thal *et al.*, 2009). *Id3* expression is up-regulated when E47 protein levels decline upon the expression of the pre-BCR (Schebesta *et al.*, 2002). *Id3* knockout studies have shown no general defect in development in the haematopoietic system or in general development, but does disrupt humoral activity (Pan *et al.*, 1999).

The ectopic expression of *id3* has been shown to cooperate in TGFβ signal transduction pathways and induce apoptosis of B lymphocyte progenitors; this apoptosis was rescued by the ectopic expression of *e47* (Kee, 2005).

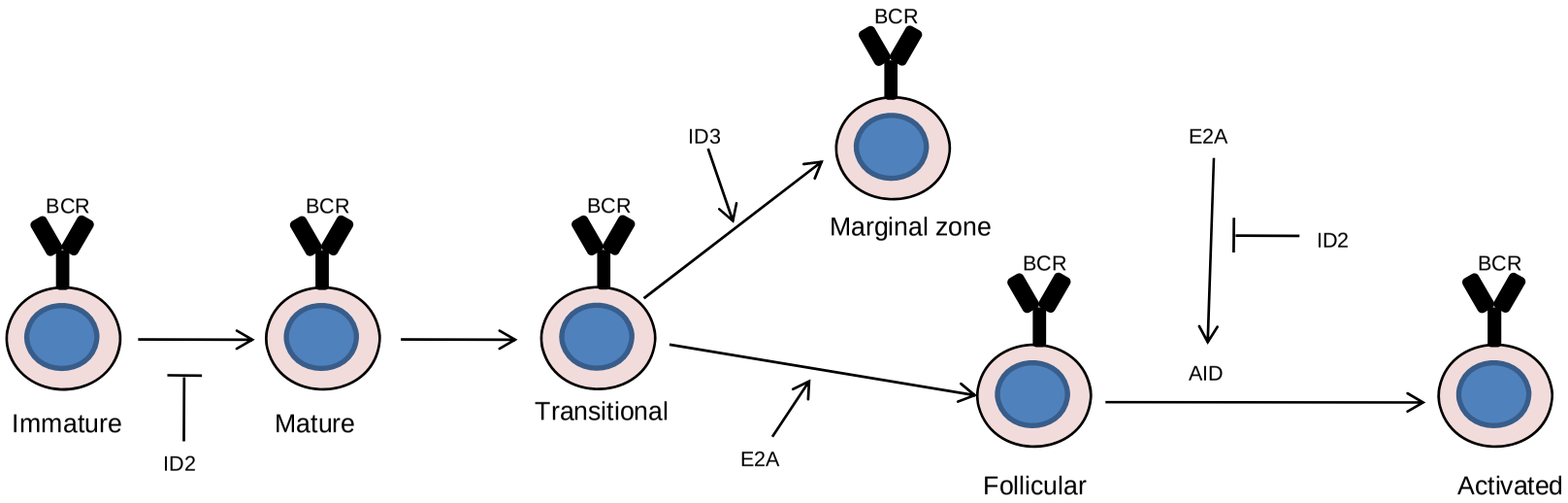


Figure 1.7 E2A and ID protein synergy in B-cell maturation. ID2 inhibits the differentiation of immature (IgM^+IgD^-) to mature (IgM^+IgD^+) B-cells by antagonising E2A. During normal differentiation of these cells *id2* expression is up-regulated in immature cells and then down-regulated in mature B-cells. In the absence of *id2* expression, these cells differentiate into mature B-cells. In *id2* knockout mice the marginal zone in the spleen is absent (Becker-Herman *et al.*, 2002). ID2 can also regulate CSR in B-cells by inhibiting E2A activity, for example, in *id2* knockout mice, high levels of IgE were detected in the serum of mice. This was found to be a result of E2A not being antagonised by ID2 and therefore free to bind the ε promoter along with cooperative signalling with IL4, driving the expression of *IgE* (Sugai *et al.*, 2003).There is also evidence that high levels of E2A can drive *id2* expression and that EBF can inhibit the expression of *id2* and *id3* by binding their promoters (Thal *et al.*, 2009). ID1, ID2 and ID3 proteins all bind PAX proteins and inhibit them, forming complexes with DNA (Roberts *et al.*, 2001). *Id2* expression inhibits *aid* expression by antagonising both E2A and PAX5 by directly binding to them. *ID2* expression is associated with the expression of a splice variant of *PAX5* in AID negative CLL cases (Oppezzo *et al.*, 2005). Abundant E2A favours a follicular cell fate and abundant ID3 favours marginal zone development (Quong *et al.*, 2004).

Ectopic *id3* expression also antagonises E2A, affecting later stages of B-cell development, specifically, at the T1 and T2 stages where further receptor editing of the BCR is required. ID3 protein has the ability to antagonise E2A function, mimicking the *e47-/+* system and result in a reduced T2 population and therefore an increased marginal zone within the spleen (Quong *et al.*, 2004).

Id3 -/- mice demonstrate a reduced proliferative capacity in response to BCR signalling and it appears that ID3 protein has a negative role during isotype switching (Pan *et al.*, 1999). The role of ID3 protein in mature B-cell development is shown in Figure 1.7.

1.9.4 ID4 in normal haematopoiesis.

There is little information on *ID4* expression in the haematopoietic system. The lack of *ID4* expression in normal haematopoietic tissues (human and mouse) has been described (Cooper *et al.*, 1997; Kersten *et al.*, 2006; Nogueira *et al.*, 2000; Riechmann *et al.*, 1994). Furthermore, the knockdown of *id4* in mice does not affect B-cell development (personal communication Professor Stablitzky and Dr M Capasso).

However, Yu *et al.* describe *id4* RNA expression in the spleen, thymus and the bone marrow (Yu *et al.*, 2005) of mice. Nogueria *et al.* also found low levels of *id4* RNA (with all the other ID transcripts) in *in vitro* embryonic stem cell line cultures, but this disappeared as the cells differentiated into blast colonies (Nogueira *et al.*, 2000). *Id4* was also originally cloned from a mouse bone marrow cDNA library (Riechmann *et al.*, 1994). Therefore the expression of ID4 in the haematopoietic system requires further examination.

1.10 The ID proteins in tumourigenesis.

ID1, ID2 and ID3 proteins do not appear to be true oncogenes and the literature focuses on the ID proteins cooperating with various oncogenic pathways. The *ID* RNA's (*ID1, ID2* and *ID3*) have been shown to be elevated in numerous cancers. Elevated levels have been reported in carcinomas of the prostate, breast, ovary, colon, rectum, pancreas, liver, endometrium, cervix and thyroid,

in squamous cell carcinomas of the nasopharynx, oesophagus and oral cavity, in neural tumours, melanoma, Ewing's sarcoma, seminoma, as well as gastric adenocarcinoma and leukaemia (Perk *et al.*, 2005).

The limited literature on ID4 makes its possible roles in oncogenesis difficult to determine. To understand how *ID4* could be functioning as an oncogene, the following sections focus on the involvement of ID1, ID2 and ID3 proteins in haematological malignancies and any reports of ID4 involvement in any malignancy.

1.10.1 The expression of ID proteins in haematological malignancies.

ID1 expression is prevalent in numerous types of malignancies. In haematopoietic malignancies, *ID1* has been demonstrated to be up-regulated as a consequence of oncogenic tyrosine kinases in myeloproliferative cases (Tam *et al.*, 2008). *ID1* expression has been found in a subset of multiple myelomas which harbour the t(4;14)(16.3;q32) translocation, which produces an oncogenic fusion protein involving the Multiple myeloma SET domain (*MMSET*) gene and the Fibroblast growth factor receptor (*FGFR3)* gene (Hudlebusch *et al.*, 2005).

ID2 is up-regulated by MYC signalling in neurones (Lasorella *et al.*, 2000). In Burkitts lymphomas, *ID2* displays high levels of expression, but is dispensable for MYC induced transformation in these cells (Nilsson *et al.*, 2004). In Hodgkin Reed Sternberg cells (HRS), ID2 protein in collaboration with the activated B-cell factor (ABF-1) antagonise E2A and therefore deregulate the B-cell transcriptional program. *ABF-1* and *ID2* expression in HRS cells results in the down-regulation of B-cell associated genes and the expression of some T-cell and macrophage associated markers (Mathas *et al.*, 2006; Renne *et al.*, 2006). This loss of B-cell lineage identity has also been associated with *ID2* expression in Primary Effusion Lymphoma (PEL) (Lietz *et al.*, 2007).

In HRS cells, high *ID2* expression has been found to correlate with low levels of *PAX5* and *E2A* in these cells (Renne *et al.*, 2006). *ID2* expression is also up-

regulated as a consequence of *LMP1* expression in GC B-cells (Vockerodt *et al.*, 2008).

There is no literature to suggest the involvement of ID3 in haematological malignancies.

1.10.2 ID4 expression in malignancies.

The sub-cellular localisation of the ID proteins is considered nuclear, but differs with cell type, differentiation stage and post translational modification (Deed *et al.*, 1997; Trausch-Azar *et al.*, 2004). In normal rat breast tissue ID4 sub-cellular localisation changes with the differentiation stage of the epithelia, with more nuclear staining in proliferating epithelia of mid-stage pregnant rats. At all other stages ID4 is either predominately cytoplasmic or very weakly expressed (Shan *et al.*, 2003). In chemically induced breast carcinomas in rats, ID4 protein is up-regulated, but only 10% of these show nuclear localisation of ID4. The nuclear localisation of ID4 in these cases is associated with increased proliferation of these cells (Shan *et al.*, 2003)

ID4 expression has been linked to breast tumourigenesis and ID4 protein down-regulates *BRCA1* expression (Beger *et al.*, 2001). In basal like breast cancers *ID4* expression positively correlates with cytokeratin 5 and 6 expression and reduces *BRCA1* expression (Turner *et al.*, 2007). *ID4* expression is inversely correlated with oestrogen receptor expression in breast cancer malignancies (de Candia *et al.*, 2006). *ID4* promoter methylation is also evident in breast cancer and correlates with nodal metastasis (Umetani *et al.*, 2005).

ID4 expression in breast cancer demonstrates its role as a potential tumour suppressor and oncogene in a context dependent manner.

Enforced *ID4* expression in mammary epithelial cells induces their proliferation and increases colony formation, however, *ID4* expression inhibits the differentiation of these cells (Shan *et al.*, 2003). *ID4* expression is also regulated by ‘gain of function’ p53-R175H mutations, which with E2F1 bind to the *ID4* promoter to drive transcription. Increased levels of ID4 protein can then bind to

and stabilise the cytokines IL-8 and GRO-α to promote angiogenesis (Fontemaggi *et al.*, 2009).

The expression of *ID4* in the brain and its development also relates to *ID4* over-expression in glioblastomas. A study by Jeon *et al.* demonstrated the cooperation of *id4* expression in *cdkn2a* -/- mouse astrocytes. *Id4* expression inhibited differentiation with increased Jagged 1 protein expression, activated Notch1 and increased the expression of cyclin E protein resulting in increased proliferation (Jeon *et al.*, 2008).

ID4 expression in prostate cancer is associated with an increased risk of metastasis (Yuen *et al.*, 2006). A more recent study by Carey *et al*. found *ID4* promoter methylation in prostate cancer and ectopic expression of *ID4* results in cell cycle arrest (Carey *et al.*, 2009). *ID4* promoter methylation and loss of expression has also been documented in colorectal adenocarcinomas (Gomez Del Pulgar *et al.*, 2008; Yu *et al.*, 2005).

ID4, therefore, has been implicated in malignancies other than BCP-ALL and may function as both a tumour suppressor and/or an oncogene in a context dependent manner. The role of *ID4* as a driver in these malignancies remains to be established.

1.11 The ID proteins in other cellular functions.

The ability of ID1, ID2 and ID3 proteins to affect haematopoiesis and drive transformation has been discussed in the earlier sections to the extent that the ID proteins modify transcription in cells. The ID proteins also interact with other cellular processes both transcriptionally and post transcriptionally. The next section examines the cellular dynamics of the ID proteins.

1.11.1 Cellular localisation.

The cellular location of the ID proteins is dependent on post translational modification, interaction with the E proteins and the cell type. The ID proteins, because of their small size, can diffuse between the nucleus and the cytoplasm freely. ID2 also contains two nuclear export signals (NES) to facilitate this

process. NES1 is located in the HLH domain, whereas NES2 is in the C terminus, it is NES2 which is required for ID2 export out of the nucleus in a chromosome region maintenance protein 1 dependent manner (Trausch-Azar *et al.*, 2004).

The nuclear location of ID2 provides a greater repression of the E proteins (Matsumura *et al.*, 2002). Whilst ID2 contains no nuclear localisation signal, some basic domains in the HLH domain are sufficient for nuclear localisation in ID2 (Kurooka and Yokota, 2005). ID1 also contains a domain which allows nuclear localisation between amino acids 68 and 84 (Trausch-Azar *et al.*, 2004).

The localisation of the ID proteins is generally considered to be nuclear (Trausch-Azar *et al.*, 2004; Tu *et al.*, 2003). In the haematopoietic progenitor cells 32D, ID2 was found to be both nuclear and cytoplasmic, but once the cells were induced to differentiate, the ID2 localised predominately to the cytoplasm (Tu *et al.*, 2003). In smooth muscle cells the location of ID2 is within in the nucleus, however, ablating the phosphorylation of ID2 alters the cellular localisation to the cytoplasm (Matsumura *et al.*, 2002).

Interaction with the E proteins also affects the localisation of the ID proteins. In oligodendrocytes, ID2 and ID4 sequester the class II HLH factors OLIG1 and OLIG2 into the cytoplasm inhibiting their interaction with the E proteins and thereby affecting the differentiation of the cells (Samanta and Kessler, 2004). ID1 is sequestered into the nuclei of Hela cells by the classII HLH protein, MyoD, a feature which is also controlled by the E2A proteins (Lingbeck *et al.*, 2005; Trausch-Azar *et al.*, 2004) and ID3 is also chaperoned into the nucleus by interactions with E47 in Cos cells (Deed *et al.*, 1997). The nuclear localisation of the ID proteins also extends their half life (Deed *et al.*, 1997; Lingbeck *et al.*, 2005; Trausch-Azar *et al.*, 2004).

1.11.2 Post translational modification, protein stability and degradation.

ID2 and ID3 contain an N terminal region targeted by Cyclin A and Cyclin E cdk2 complexes for phosphorylation. This region is encoded by an SPVR site of which a serine residue is phosphorylated; the SPVR region is also present on

ID4 but absent in ID1 (Hara *et al.*, 1997). Phosphorylation of either ID2 or ID3 affects their binding interactions, cell cycle control and cellular location.

Phosphorylation of ID2 inhibits the binding to E proteins, whereas the phosphorylation of ID3 changes the specificity of the binding to either E12 homodimers or E12-MyoD heterodimers in B-cells (Deed *et al.*, 1997; Hara *et al.*, 1997). The inhibition of the phosphorylation of either ID2 or ID3 also affects their ability to promote S phase entry and therefore affects the proliferative potential of the ID2 and ID3 (Deed *et al.*, 1997; Matsumura *et al.*, 2002).

The half life of the ID proteins varies between 15 minutes to 1hour depending on cell type and cellular localisation. As was demonstrated by the interaction of ID2 with APC/C complex in neurones, the disruption of the degradation of the ID proteins can induce transformation (Lasorella *et al.*, 2006).

The degradation of the ID proteins has been linked to the ubiquitin proteasome system (UPS). This system relies on an ubiquitin activating enzyme, E1, an ubiquitin carrier protein, E2, and an ubiquitin ligase, E3. The addition of multiple molecules of ubiquitin then targets the protein for degradation via the proteasome (Bounpheng *et al.*, 1999; Fajerman *et al.*, 2004). This process of degradation is apparent with ID1, ID2 and ID3, but not ID4. However, ID4 is still ubiquitinated (Bounpheng *et al.*, 1999). ID1 also interacts with a component of the 26S proteasome, which disrupts its interaction with MyoD.

The UPS system relies on N terminal ubiquitination on an internal lysine group within the protein and this has been demonstrated with ID2 in Cos cells (Fajerman *et al.*, 2004). ID1 and ID3, but not ID2 or ID4, interact with components of the COP9 signalosome (CSN). The CSN cooperates with the UPS system to phosphorylate proteins and target them for degradation via the proteasome. Over-expression of CSN2 (a component of CSN) helps to stabilise ID3 potentially by increasing the half life of E47 in HELA cells (Berse *et al.*, 2004). Both ID1 and ID2 have also been shown to be targeted for degradation by the E3 APC/C (see section 1.11.4).

1.11.3 Cell cycle.

ID1 and ID2 proteins are required for cell cycle progression through the G1 phase and levels of these proteins decline as cells enter senescence in fibroblasts (Hara *et al.*, 1994). The retinoblastoma (Rb) protein represents a crucial protein in cell cycle regulation and the disruption of Rb is one hallmark of cancer (Giacinti and Giordano, 2006). An overview of the function of the cell cycle can be seen in Figure 1.8.

ID2 RNA levels, after cells are released from serum deprivation, peak at the G0/G1 transition and again at the late G1 phase. In U2OS cells, which have functional Rb and p53, *ID2* over expression results in increased proliferation. This is not the case in SAOS-2 cells, which lack functional Rb, indicating that ID2 protein exerts its effect on the cell cycle via Rb (Iavarone *et al.*, 1994). Furthermore, the loss of *id2* expression rescues *rb-/-* embryos and delays tumourgenesis of pituitary tumours in *rb+/-* mice (Lasorella *et al.*, 2000; Lasorella *et al.*, 2005). ID2 is therefore an effector of cell cycle regulation and proliferation by the regulation of Rb. An overview on the role of ID1, ID2 and ID3 in the cell cycle is shown in Figure 1.9.

The effects of ID4 protein expression on the cell cycle have been described by Yun *et al.* to be required for the G1to S phase transition in neural precursors (Yun *et al.*, 2004). A report on prostate cancer cell lines shows that ID4 protein has a different effect on the cell cycle to the other ID proteins. Carey *et al.* demonstrated that *ID4*, when ectopically over-expressed in prostate cancer cell lines, induced S phase arrest and apoptosis. This was attributed to ID4 up-regulating the expression of *E2A* which in turn increased the levels of *p21* and *p27* expression (Carey *et al.*, 2009). ID4 has also been demonstrated to increase *cyclin e* levels in *cdkn2a -/-* astrocytes, which drives their proliferation (Jeon *et al.*, 2008).

The ID proteins in cells are regulated by mitogenic factors, which allow their up-regulation and driving of the cell cycle. Furthermore, the ID proteins are also postranslationally modified by cell cycle proteins (see section 1.10.2)

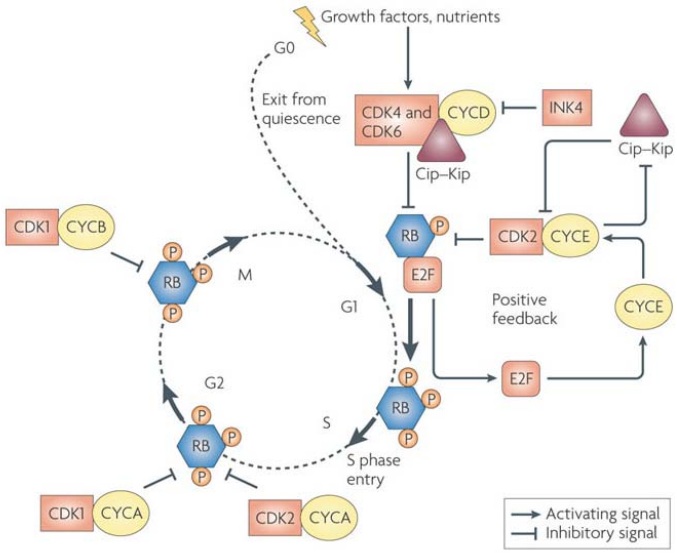


Figure 1.8 The cell cycle. The different stages of the cell cycle are regulated by a series of cyclins and cyclin dependent kinases (cdk). Mitogenic stimuli such as growth factors release CDK4 and CDK6 from the inhibitory *CDKN2A* family members (p16, p15, p18 and p19) to form active complexes with the D type cyclins which phosphorylate the retinoblastoma protein (Rb). The Rb (p105) protein family consists of two other members, p130 and p107, which when unphosphorylated, bind and antagonise cellular factors such as the E2F family of transcription factors and induce cell cycle arrest at the G1 phase. Phosphorylated Rb releases E2F thereby permitting the expression of proteins such as cyclin E required for G1 to S phase progression. In the late G1 phase CDK2 and cyclin E heterodimers reinforce the phosphorylation of Rb which allows progression through a restriction point to the S phase. From this stage onwards, Rb is maintained in a hyperphosphorylated state by the activities of cyclin A – CDK2, cyclin A-CDK1 and cyclin B-CDK1 complexes. CDK2 is inhibited by the Cip-Kip family members, p27, p21 and p57. The cyclin D CDK4/6 complexes can bind the Cip-Kip family members targeting them for degradation (Lapenna and Giordano 2009). During late S phase/ early G2 phase, the centrosomes duplicate and separate to allow the beginning of mitosis. This process is regulated by CDK1 and CDK2 serine threonine kinases protein kinases including aurora A and polo like kinase 1. CDK1 and cyclin A complex in the G2 M phase to allow chromosomal condensation and segregation and microtubule formation. Cyclin A is then degraded and cyclin B CDK1 complexes progress the cell cycle through the M phase. For complete cell division, the CDK1 is inactivated by the degradation of cyclin B through the anaphase promoting complex or cyclosome (APC/C) (Lapenna and Giordano, 2009). Mitotic checkpoints such as the spindle assembly checkpoint are activated once the cells enter mitosis. This prevents the inappropriate degradation of regulators of mitosis such as cyclin B by the APC/C, resulting in mitotic arrest. The appropriate separation of the microtubules signals the degradation of SAC allowing APC/C to degrade components such as cyclin B allowing mitosis (Lapenna and Giordano, 2009).

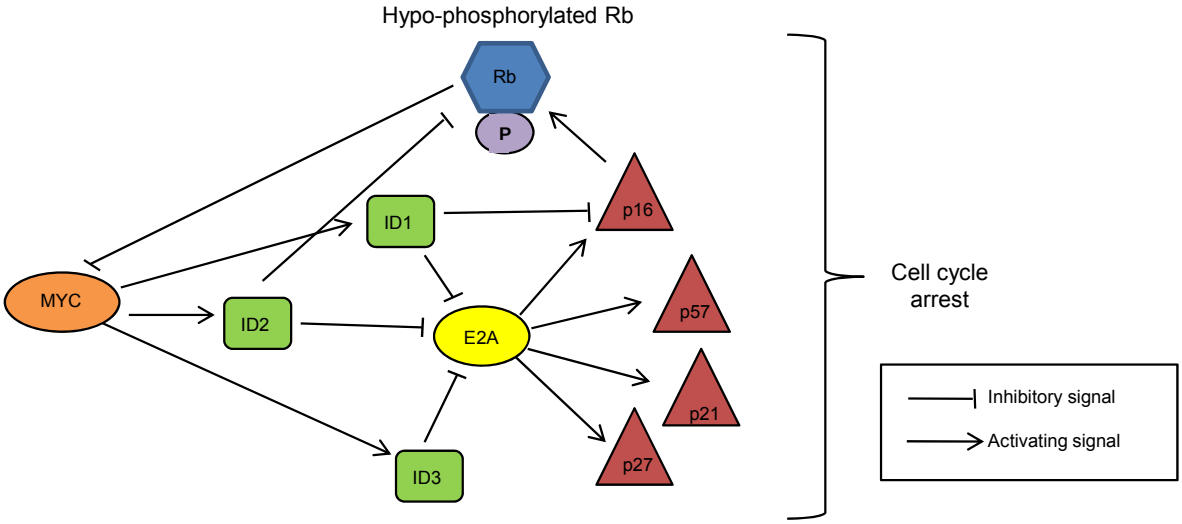


Figure 1.9 The effects of ID1, ID2 and ID3 on the cell cycle. *ID1, ID2* and *ID3* are up-regulated in MYC induced hyperplastic epidermis as a consequence of MYC expression (Murphy *et al.*, 2004). *ID1* has also been found to be up-regulated in breast epithelial cells in correlation with c-MYC. In this instance, the up-regulation of either protein resulted in the proliferation of MCF7 cells in low serum conditions and reduced Rb phosphorylation as a consequence of reduced cyclin D and cyclin E expression (Swarbrick *et al.*, 2005). ID2 protein binds un-phosphorylated Rb via its HLH domains, thereby inactivating it; ID1 and ID3 do not show these interactions (Iavarone *et al.*, 1994; Lasorella *et al.*, 1996). By inactivating Rb function ID2 also abolishes the function of p16, which acts upstream of Rb (Lasorella *et al.*, 1996). In neuroblastoma cell lines n-MYC and c-MYC both directly up-regulate *ID2* by binding to its promoter, Rb function down-regulates MYC function. MYC signalling therefore requires ID2 to bypass Rb inhibition (Lasorella *et al.*, 2000). ID1, ID2 and ID3 have been shown to antagonise the effects of E2A proteins on cell cycle (Lasorella *et al.*, 2006; Rothschild *et al.*, 2006; Trabosh *et al.*, 2009). Over-expressed *e2a* induce G1 arrest in numerous cell types and *e2a* ectopic expression can rescue *e2a* -/- thymic lymphomas by inducing apoptosis (Engel and Murre, 1999; Peverali *et al.*, 1994; Prabhu *et al.*, 1997). E47 can also transcriptionally up-regulate *p16, p21* and *p57* expression (Ouyang *et al.*, 2002; Prabhu *et al.*, 1997; Rothschild *et al.*, 2006; Semerad *et al.*, 2009). The regulation of *p16* expression by ID1 is mediated by ID1 interacting with ETS2, which is regulated by the Ras-Raf-Mek signalling pathway. ID1 directly interacts with ETS2 in 293T-cells and inhibits the ability of ETS2 to up-regulate *p16* (Ohtani *et al.*, 2001). In NIH3T3 cells, ID3 antagonises *p27* expression by inhibiting ELK1 which binds to the *CDKN1B* promoter to regulate its activity and as with the effects of ID1 on *p16*, the consequential cell cycle proliferation results in an increase in the phosphorylation of Rb (Chassot *et al.*, 2007). Depletion of *ID3* correlates with an increase in *P27* in human dermal fibroblasts (Chassot *et al.*, 2007) and the knockdown of *id3* in neural crest progenitor cells results in a decrease in proliferating cells as a consequence of *p27* up-regulation (Kee, 2005). In human dermal fibroblasts, ID3 was found to antagonise the *p27* expression in a non-E2A dependent manner.

1.11.4 ID proteins and genomic instability.

The ID proteins have been shown to de-regulate the G1 to S phase of the cell cycle, and in doing so affect proliferation. The ID proteins can also affect mitosis which can result in genomic instability.

The APC/C complex is a multi subunit E3 ubiquitin ligase required for the degradation of components of mitosis, thereby allowing efficient chromosomal segregation. The APC/C component is regulated by its sub-units, cell division cycle protein 20 (cdc20) and cdc20 homologue 1 (cdh1), which confer substrate specificity and activation of APC/C (Pesin and Orr-Weaver, 2008).

The APC^{cdh1} complex targets a canonical D box motif present on ID1, ID2 and ID4, but not ID3. However the N terminus of ID1 is also important in this degradation pathway (Lasorella *et al.*, 2006; Wang *et al.*, 2008). Both ID1 and ID2 have also been shown to be targeted for degradation by the E3 APC/C with the co-activators cdh1 and cdc20 via the D box motif (residues 100-107) and independently of N terminal ubquitination. This targeted degradation allows for coordinated differentiation programs in neuronal development (Lasorella *et al.*, 2006).

ID1 interacts with both cdc20 and cdh1, thereby disrupting the normal function of the APC/C complex. The interaction of ID1 and cdc20 results in early mitosis where as the interaction with cdh1 results in the failure of cytokinesis. Both interactions were via the D box motif. ID1 itself was targeted for degradation by the APC/C complex (Wang *et al.*, 2008). This effect in prostate cell lines resulted in ploidy and genomic instability.

ID1 can also inhibit the degradation Aurora A, a centrosome localised mitotic kinase protein involved in the assembly and the stability of the mitotic spindle. This is by ID1 binding cdh1 and disrupting APC activity (Man *et al.*, 2008).
ID2 is targeted for degradation by the APC^{cdh1} complex in primary neurones. This targeted degradation of ID2 allows coordinated neuronal differentiation by allowing bHLH protein function in neurones. The disruption of this process allows elevated ID2 levels which overcome cell cycle arrest, demonstrating a potential pathway for transformation (Lasorella *et al.*, 2006).

1.11.5 TGFβ pathway.

The transforming growth factor beta (TGFβ) family signalling pathway has been implicated in B-cell development and in the regulation of the ID proteins (Perk *et al.*, 2005).

TGFβ signalling proteins are structurally related cytokines and include TGFβ and bone morphogenitc proteins (BMP). These proteins play an important role in normal development and are involved in haematopoiesis. The signalling pathway is mutated in many cancers. These cytokines work by binding type II serine threokine kinase receptors which heterodimerise with type II receptors, thereby phosphorylating them. This then instigates signal transduction via SMAD proteins. There are three types of SMAD proteins, Common partner SMADS (CoSMAD), receptor regulated SMAD (R-SMAD) and inhibitory SMAD (I-SMAD). On signal transduction, R-SMAD proteins are phosphorylated and complex with the CoSMAD, SMAD 4. This complex then translocates to the nucleus to regulate transcription (Waite and Eng 2003).

BMP6 expression up-regulates expression of *ID1, ID2* and *ID3* in progenitor B-cells (Kersten, Sivertsen et al. 2005). TGFβ expression in differenT-cell types exhibits different responses. In epithelial cells the *id* genes are down-regulated, and in B-cells *id2* and *id3* have been shown to be up-regulated (Spender and Inman, 2009). Treating B-cell precursors with TGFβ blocks their proliferation and thereby inhibits B-cell development by inducing G1 arrest in these cells. This effect has been attributed to the ability of TGFβ to up-regulate *id3*, which is required for the initial stages of G1 arrest. Low levels of ID3 protein expression can mimic a TGFβ response in B lymphocyte progenitors (Kee, Rivera et al. 2001).

TGFβ induces cell cycle arrest in cells and this has been attributed to the down-regulation of c-MYC which allows expression of the CDKi p21, p15 and p27 (Siegel *et al.*, 2003; Spender and Inman, 2009). In Burkitts lymphoma cell lines, where there is over-expression of *MYC* due to the t(8;14) translocation, TGFβ treatment of these cells still induces cell cycle arrest and this has been

attributed to the down regulation of *E2F-1*, a factor required for the progression of G1 to S phase entry of the cell cycle (Spender and Inman 2009).

The TGF β family control proliferation, differentiation, migration and apoptosis in many cell types. The *ETV6/RUNX1* fusion was shown to overcome TGF β mediated signalling in B-lymphocyte progenitors by acting within and reducing a cells response to the TGFβ signalling pathway (Ford *et al.*, 2009).

Overall, ID1, ID2 and ID3 proteins affect the cell cycle and differentiation, disruption of which are hallmarks of neoplasia. These proteins work in a cellular context-dependent manner and their function is governed by their degradation, cellular location and the proteins they interact with. There is a role for the ID proteins in haematopoiesis, and the over-expression of ID2 is associated with Burkitts and Hodgkin lymphoma. The *ID* genes have been described as oncogenes, but often require cooperating mutations in order for malignancies to arise (Perk *et al.*, 2005). The potential of *ID4* to be an oncogene in a B-cell environment has to be investigated taking these factors into consideration.

1.12 Aim and objectives.

The aim of this thesis is to discover whether *ID4* functions as an oncogene in a B-cell context.

The objectives of this thesis are:

- To examine the expression of *ID4* RNA in normal and malignant tissues using bioinformatics, RT-PCR and western blotting of B-cell lines, BCP-ALL and xenograft material.
- To study the possible effects of *ID4* expression in B-cell lines and primary B-cell progenitors, by examining sub-cellular localisation of the ID4 protein and features such as proliferation and cooperation with other oncogenes.
- To identify pathways by which ID4 protein could induce oncogenesis by identifying proteins which ID4 interacts with and identifying mechanisms by which ID4 is deregulated.

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Tissue culture.

2.1.1.1 Cell lines.

Table 2.1 Lists of the cell lines used and their pathologies

Cell line	Pathology	Other comment (see section 2.2.1)
NALM-6	Human BCP-ALL	Complete RPMI medium
LK63	Human BCP-ALL	Complete RPMI medium
SEM	Human BCP-ALL	Complete RPMI medium
RCH-ACV	Human BCP-ALL	Complete RPMI medium
Nalm27	Human BCP-ALL	Complete RPMI medium
380	Human BCP-ALL	Complete RPMI medium
MUTZ5	Human BCP-ALL	Complete RPMI medium
Cemo1	Human BCP-ALL	Complete RPMI medium
Per 365	Human BCP-ALL	Complete RPMI medium
Per 377	Human BCP-ALL	Complete RPMI medium
SUDHL4	Human BCP-ALL	Complete RPMI medium
LILA	Human BCP-ALL	Complete RPMI medium
RS4;11	Human BCP-ALL	Complete RPMI medium
HB	Human BCP-ALL	Complete RPMI medium
KOPN8	Human BCP-ALL	Complete RPMI medium
697	Human BCP-ALL	Complete RPMI medium
WEHI-3B	Mouse myelomonocytic leukaemia	Cultured in D-MEM Invitrogen, 31965-025) supplemented with 10% FCS
Ba/F3	Mouse progenitor B-cells	RPMI medium supplemented with 10% IL3
LP1	Myeloma	Complete RPMI medium
OCILY10	DLBCL	Complete RPMI medium
OCIYL3	DLBCL	Complete RPMI medium
JURKAT	T cell ALL	Complete RPMI medium
SSK41	Marginal zone lymphoma	Complete RPMI medium
RPMI 8226	Myeloma	Complete RPMI medium
HUT78	T cell lymphoma	Complete RPMI medium
RIVA	DLBCL	Complete RPMI medium
785 BG	Normal	Complete RPMI medium
HBL2	Mantle cell lymphoma	Complete RPMI medium
ELIJAH	Burkitts lymphoma	Complete RPMI medium
RAJI 2.2.5	Burkitts lymphoma	Complete RPMI medium
KHM2B	Burkitts lymphoma	Complete RPMI medium
RAMOS	Burkitts lymphoma	Complete RPMI medium
BJAB	Burkitts lymphoma	Complete RPMI medium
MUTU III	Burkitts lymphoma	Complete RPMI medium
HDMYZ	Hodgkin lymphoma	Complete RPMI medium
DAUDI	Burkitts lymphoma	Complete RPMI medium
MHH-PREB-1	Burkitts lymphoma	Complete RPMI medium
DG75	Burkitts lymphoma	Complete RPMI medium
SP53	Mantle cell lymphoma	Complete RPMI medium
OCILY 19	DLBCL	Complete RPMI medium
IRM2	Mantle cell lymphoma	Complete RPMI medium
REC-1	Mantle cell lymphoma	Complete RPMI medium
MOLT4	T cell ALL	Complete RPMI medium
SCL1	DLBCL	Complete RPMI medium
JVM-2	Mantle cell lymphoma	Complete RPMI medium
UPN1	Mantle cell lymphoma	Complete RPMI medium
VALLOIS	DLBCL	Complete RPMI medium
SUDHL-9	DLBCL	Complete RPMI medium
FL18	DLBCL	Complete RPMI medium
JEKO1	Mantle cell lymphoma	Complete RPMI medium
SP1	Mantle cell lymphoma	Complete RPMI medium
K422	DLBCL	Complete RPMI medium
OP9	Mouse bone marrow stromal layer	Complete IMDM
Phoenix Alpha	Viral Packaging cell line	Complete IMDM with 2% Hepes

DLBCL- Diffuse large B-cell lymphoma
ALL-Acute lymphoblastic leukaemia
BCP-ALL- B-cell precursor ALL

2.1.1.2 Transfected cell lines.

Table 2.2 Clones from cell lines transfected with the relevant constructs (see section 2.2.3 for vector maps).

Parental cell Line	Clone	Vectors and Inserts	Selection in medium	Comments
*Ba/F3†	ID4 8	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin (Invitrogen ant-zn-5p) 500µg/ml	Complete RPMI medium with 10% conditioned IL3
	ID4 11	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin 500µg/ml	
	ID4 16	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin 500µg/ml	
	EV14	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin 500µg/ml	
	EV15	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin 500µg/ml	
	EV16	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin 500µg/ml	
	MUT 4	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin 500µg/ml	Complete RPMI medium with 10% conditioned IL3 ID4 mutated in amino acids R139G and L143V
	MUT 9	pcDNA6/TR pcDNA4/TOID4myc-his	Zeocin 500µg/ml	Complete RPMI medium with 10% conditioned IL3 ID4 mutated in domains R139G and L143V
*NALM-6	TR	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin 50µg/ml	Complete RPMI medium (Tetracycline free FCS used)
*RCH-ACV	TR 1	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin 50µg/ml	Complete RPMI medium (Tetracycline free FCS used)
	TR 2	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin 50µg/ml	
	TR 3	pcDNA6/TR pcDNA4/TOID4myc-His	Zeocin 50µg/ml	

*All parent cell lines were transfected using the T-REx™ (see section 2.2.2 and Figure 2.1) to allow the inducible expression of ID4.
† All BA/F3 clones constitutively expressed ID4. Tetracycline was not used to induce ID4 expression in these clones.

2.1.1.3 Tissue culture reagents

10x Trypsin EDTA	0.25% Trypsin, 5.3mM EDTA - 5.0g/l trypsin, 2.0g/l EDTA-4Na and 8.5g/l NaCl. Working solution is diluted to 1 x in PBS
Phosphate Buffered Saline (PBS)	PBS tablets (Oxoid BR0014G) in H_20.

2.1.2 Western Blot solutions

Tris buffered saline (TBS) 1x:	Tris Base (20mM) (Fisher Scientific, BPE152-1) NaCl (137mM) (Fisher Scientific: S/3160153) and HCL (to pH 7.6) (Fisher Scientific H/1000/PB17).
TBS tween (TBST) 1X	TBS 1 x, 0.01%Tween 20 (Sigma Ultra P7949)
Electrode Buffer 1x:	Tris Base (25mM), Gylcine (250mM) (Fisher Scientific, BPE381) and 0.1% SDS (Fisher Scientific BNO:16416) in pH 8.3.
Transfer Buffer1x:	Tris Base (25mM), Glycine (192mM) and 20% Methanol (Fisher Scientific M/3950/PB17).
1x SDS Sample Buffer:	Tris HCl (100mM pH6.8), 20% Glycerol (Sigma Aldrich,G5516), 2% SDS, 5% 2-Mercaptoethanol (Fisher Scientific12547-0010), 0.1% Bromophenol Blue (Sigma B8026).
Blocking buffer	5% Marvel in TBST
Ammonium Persulphate:	10% stock solution (Sigma Aldrich, BPE179-125)
1.5M Tris (pH8.8):	(Biorad, 1610798)
0.5M Tris (pH6.8):	(Biorad, 1610799)
TEMED	(Sigma Aldrich, T-724).
Acrylamide/ Bis Acrylamide (37.5:1)	(GeneFlow, EC-89).

2.1.3 Antibodies.

Table 2.3 Primary antibodies, dilutions and incubation times.

Antibody	Western Blotting		Immunofluorescence	Immunoprecipitation		
	Concentration	Incubation	Concentration (all antibodies were incubated in 10 % goat serum (Sigma Aldrich G9023) at 4°C overnight)	Concentration of antibody/ 20µl beads	Beads used	Incubation
Anti ID4 H70 Santa Cruz (SC-13047)	1µg/ml TBST no milk used in the secondary antibody	Overnight	1µg/ml	1µg	A type	Over night
Anti E47 N-649 SantaCruz (Sc-763)	1µg/ml TBST milk	Over night	2µg/ml	1µg	A type	Overnight
Anti HTF4 A-20 Santa Cruz (Sc-357)	2µg/ml TBST	Overnight	2µg/ml	1µg	A type	Overnight
ZNF283 MO3 Abnova (H00084364-M03)	2µg/ml TBST milk	Overnight	10µg/ml	Not applicable	Not applicable	Not applicable
RSU_1 M01 Abnova (H00006251-M01)	2.5µg/ml TBST	Overnight	10µg/ml	Not applicable	Not applicable	Not applicable
Rb 4H1 Cell Signalling Technology (9309)	1:2000 TBST in milk	One hour at room temperature	1: 100	Not applicable	Not applicable	Not applicable
Anti β Actin AC-15 Sigma (A5441)	1:10000 TBST	One hour at room temperature	Not applicable	Not applicable	Not applicable	Not applicable
Myc Tag 9B11 Cell Signalling Technology (2276)	1:2000 TBST in milk	One hour at room temperature	1:1000	2.5µg	A type	2 hours
PCNA BD Transduction Laboratories (610665)	1/10000 TBST	One hour at room temperature	Not applicable	Not applicable	Not applicable	Not applicable
Anti α Tubulin DMA1 Calbiochem (CP06)	25ng/ml TBST	One hour at room temperature	Not applicable	Not applicable	Not applicable	Not applicable
Anti E2ABD Pharamigen (554199)	Not applicable	Overnight	10µg/ml	Not applicable	Not applicable	Not applicable
Anti Laminin Serotec (MCA 1444)	1:100 TBST 3%BSA	Overnight	Not applicable	Not applicable	Not applicable	Not applicable
Anti Mono and Poly-Ubiquitin FK2 (Enzo BML-PW8810)	1/1000 TBST	Overnight	Not applicable	Not applicable	Not applicable	Not applicable

Table 2.4 Secondary antibodies, dilutions and incubation times

Antibody	Concentration	Incubation	Use
Donkey anti rabbit horse radish peroxidise (HRP) (GE Healthcare NA934V)	1:10000	Blocking buffer One hour at room temperature	Western blotting
Sheep anti mouse HRP (GE Healthcare NA931V)	1:10000	Blocking buffer One hour at room temperature	Western blotting
Goat anti mouse Alexa Fluor 568 (Invitrogen A11077)	2µg/ml	10% Goat serum One hour at room temperature	Immunofluorescence
Goat anti rabbit Alexa Fluor 488 (Invitrogen A11034)	2µg/ml	10% Goat serum One hour at room temperature	Immunofluorescence

2.1.4 Immunoprecipitations (IP) and mass spectrometry.

Ripa lysis buffer	Tris (pH7.6) 50mM, NaCl 150mM, Triton X 0.5%,
Homogenisation Buffer	250mM Sucrose (Sigma S7903), 1mM $MgCl_2$ (Fisher M/0550/53) in 10mM Hepes (pH 7.2) (Sigma H3375),
Hypotonic Buffer	10mM Hepes (pH 7.1), 10mM $MgCl_2$, 1mM EDTA,
NP40 buffer	15mMTris pH8, 10% Glycerol, 15mM NaCl, 0.2% NP40
Buffer 1	25mM Ammonium bicarbonate/5% acetonitrile
Buffer 2	25mM Ammonium bicarbonate/50% acetonitrile
Buffer 3	Trypsin buffer (PromegaV5111), 1.8ml 25mM ammonium bicarbonate

2.1.5 Flow cytometry

Annexin Buffer	10mM Hepes (Sigma H3375)/NaOH (pH7.4)(Fisher BP359), 150mM NaCl, 5mM KCl (Fisher P/4240/53), 1mM $MgCl_2$ (Fluka 68475), 1.8mM $CaCl_2$ (Fisher C/1240/53),

2.1.6 Polymerase chain reaction (PCR) and agarose gel electrophoresis

5xTris Borate EDTA (TBE)	Tris 450mM, Boric Acid (Fisher B/3800/53) 445mM, 0.001M EDTA.

2.1.7 Calcium phosphate transfections

Calcium chloride	250mM calcium chloride (Sigma 21101)
2xHBS	Hepes (pH7.1) 50mM, $NaCl_2$ 0.3M and Na_2HPO_4 (SigmaS7907)

2.2 Methods.

2.2.1 Cell culture.

The cell lines used are in Table 2.1. All cell lines, unless otherwise stated, were cultured in RPMI (Invitrogen 31870-025) supplemented with 10% foetal bovine serum and 2mM GlutaMAX (Gibco). All cell lines were cultured in an 37°C incubator in a humidified atmosphere of 5% CO_2 and 95% air. Adherent cells required trypsinisation and so were washed in PBS before an incubation in 1x trypsin EDTA at 37°C for five minutes. The cells were then resuspended in the relevant medium and spun down at 200 G for 3 minutes before resuspending and replating at the desired confluency.

2.2.1.1 Generation of IL3 conditioned medium.

The Ba/F3 cell line is a mouse progenitor B-cell line which requires the growth factor IL3 for survival. WEHI-3B cells were seeded into 50ml tissue culture flasks in medium and supplemented with 500μM 2-Mercaptoethanol. The medium was harvested after the cells had reached a confluency of $1x10^6$ /ml. The medium was centrifuged at 200 G for 3 minutes at room temperature after which it was passed through a 0.2μM filter to remove any cell debris. This was then stored at -20°C.

2.2.2 Generation and culture cell lines over-expressing ID4.

All cell lines used and their culture conditions can be seen in Table 2.2. These parental cell lines were transfected using the T-REx™ (Invitrogen K1020-01) system (see Figure 2.1) to provide the inducible expression of ectopic *ID4*. All clones were generated by electroporation using the Amaxa nucleofection system (Lonza). Once transfected these cells were single cell cloned. All clones were kindly generated by Dr T. Akasaka (Professor Dyer laboratory).

2.2.2.1 Electroporation.

$1x10^6$ cells were electroporated using 1μg of plasmid DNA (Table 2.2) in the Amaxa nucleofection V buffer (Lonza VCA-1003) using program T20 according to the manufacturers' instructions. Cells were then resuspended as single cell cultures under zeocin (500μg/ml) selection.

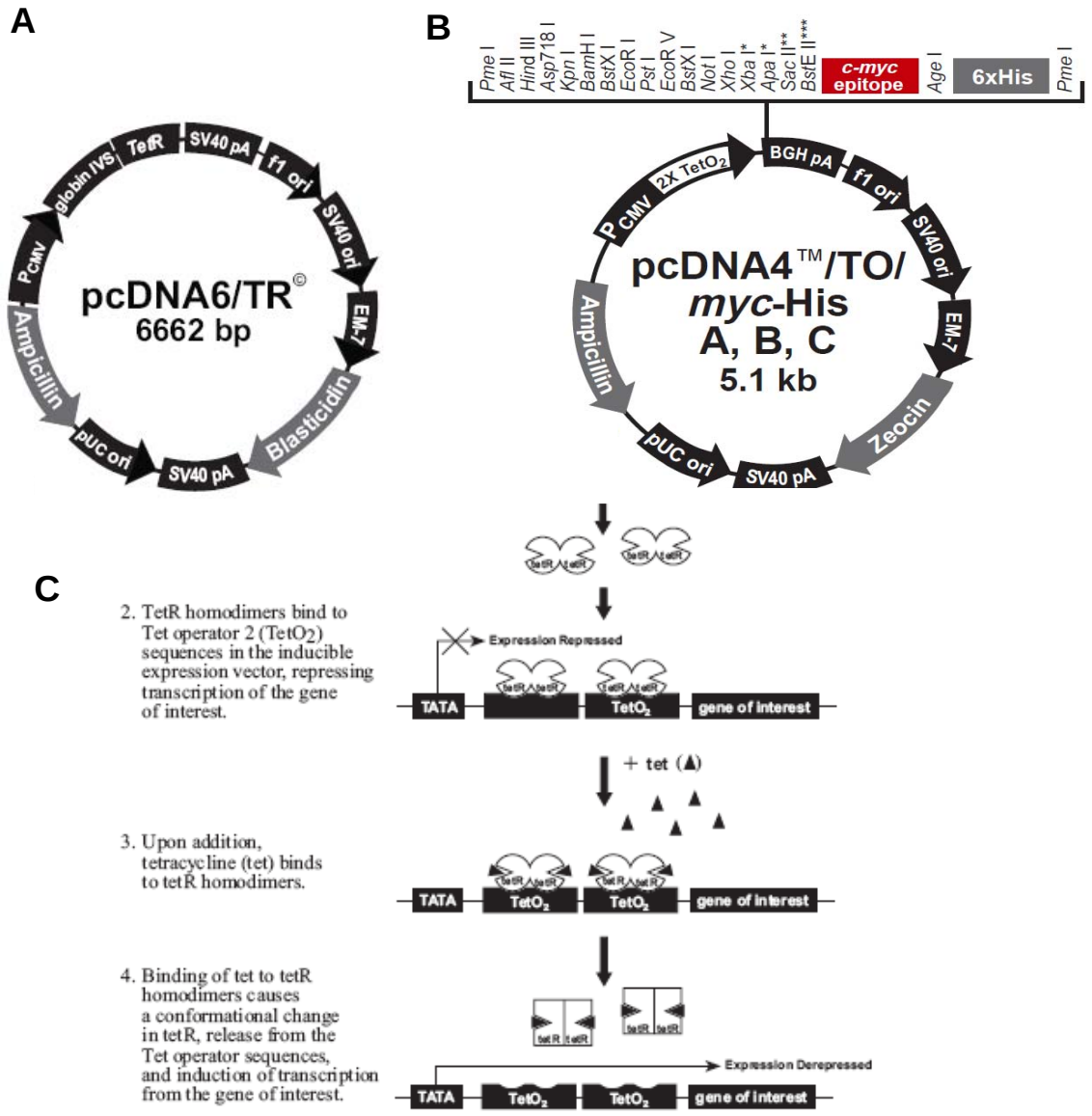


Figure 2.1 Overview of the T-REx™ system. The T-REx™ system allows the expression of a gene of interest in the presence of tetracycline. This system uses the *E.coli* Tet repressor protein (Tet R) expressed by pcDNA6/TR (**A**) under the control of a CMV promoter. TetR can bind to the Tet operator 2 ($TetO_2$) sequences. Two $TetO_2$ sequences are present in the CMV promoter of the pcDNA4/TO plasmid (**B**). In the absence of tetracycline, Tet R binds with high affinity to the $TetO_2$, thereby repressing transcription. Tetracycline can bind to the TetR and induce a conformational change which no longer allows it to bind to the $TetO_2$ molecules, thereby allowing transcription to occur (**C**). Source: Invitrogen T-REx™ manual.

A final concentration of 2µg/ml of Tetracycline Hydrochloride (Calbiochem 583411) in the complete medium was used to drive expression of the pcDNA4/TO ID4 myc-His ™ expressing clones (see Figure 2.1).

2.2.3 Vectors

All constructs used were generated by Dr T Akasaka (Professor Dyers laboratory).The human cDNA sequence for human *ID4* (clone MGC: 20126) was cloned from a pOTB7 construct (Geneservice). The coding sequence, preceded by a Kozak consensus sequence, was subcloned into a pcDNA4/TO myc-His vector (Invitrogen) using the EcoRI and XbaI restriction sites (Figure 2.1 Panel B). On treatment with tetracycline (see section 2.2.2), cells expressing the *ID4* pcDNA4/TO plasmid expressed ID4 protein with a myc-His tag.

All mutated clones were generated by Dr T Akasaka by site directed mutagenesis.

ID4 was also cloned into a green fluorescent protein (GFP) bicistronic retroviral vector, *Mig*RI (Pear *et al.*, 1998) (Figure 2.2 Panel A), using Bgl II and XhoI restriction sites. The VSV-G vector (Panel B), encoding viral envelope protein was co-transfected with the *Mig*RI vector.

2.2.4 Preparation of cell lysates for western blotting.

Approximately $2x10^7$ cells were collected by centrifugation at 200 G for 5 minutes at room temperature. The cells were washed with PBS and pelleted again. The pellet was resuspended in 200µl of 1 x sample buffer and protease inhibitors (Pi) (according to the manufacturers' instructions, Sigma P2714-1BTL), followed by sonication at an amplitude of 10 microns for 10 seconds on ice. The lysates were boiled for 5 minutes, clarified by centrifugation at 16060 G and loaded on protein gels immediately. Lysates were stored for short term at -20°C and long term at -80°C.

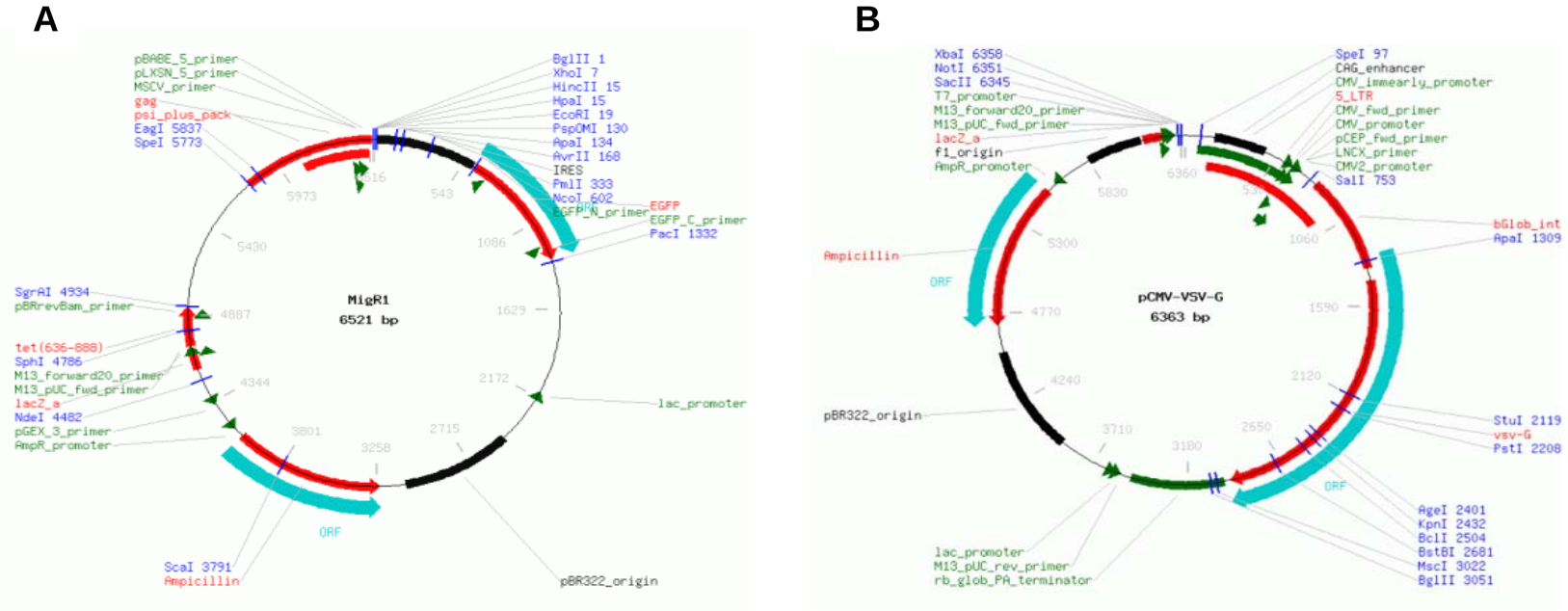


Figure 2.2 Vector maps. **A**) *Mig*RI retrovirus **B**) VSV-G viral envelope plasmid.

2.2.5 Bradford assay.

BSA (Sigma Aldrich, A3803) was diluted in 1ml of diluted Protein Assay Dye Reagent Concentrate (Biorad, 500-0006) at concentrations of, 1µg/ml, 2µg/ml, 3µg/ml, 4µg/ml, 6µg/ml and 8µg/ml. The ODs were read on a spectrophotometer at λ 595 in order to draw a standard curve. Protein samples of interest were then diluted appropriately with the diluted Protein Assay Dye Reagent Concentrate and their OD used to calculate the concentrations from the standard curve.

2.2.6 Western Blotting.

Protein gel electrophoresis was performed using SDS polyacrylamide gels, running at approximately 80 Vh-1 in electrode buffer. The formulation of the various percentage gels can be seen in Tables 2.5 and 2.6. Precision Plus Protein Dual Colour Standard (according to the manufacturers' instructions, Biorad 161-0374) was used as a size marker for the proteins.

The proteins were transferred over night (20 Vh-1 at room temperature) or for three hours (60 Vh-1 at 4°C) onto a nitrocellulose membrane (Hybond C Extra Nitrocellulose, Amersham Bioscience, RPN303E) using transfer buffer.

The blots were blocked overnight in blocking buffer after which they were washed twice for 5 minutes in TBST before the primary antibody was added (see Table 2.3). Following incubation with the primary antibody, the blots were washed twice for 5 minutes in blocking buffer and twice for 5 minutes in TBST after which the secondary was applied (see Table 2.4) followed by the same series of washes. The bands binding the primary and secondary antibodies on the blot were detected using enhanced chemilluminescence (ECL) according to the manufacturers' guidelines (ECL Western Blotting Substrate, Pierce, 32106). To visualise residual protein after transfer, gels were stained with Gelcode Blue Stain Reagent (Pierce, 24590).

Table 2.5 Resolving gel for SDS PAGE

Gel	10%	12%	15%
H_2O	12.0ml	10.0ml	7.0ml
1.5M Tris (pH8.8)	7.5ml	7.5ml	7.5ml
10%SDS	300μL	300μL	300μL
Acrylamide (30%)	10ml	12ml	15ml
10% Ammonium Persulphate	150μL	150μL	150μL
TEMED	15μL	15μL	15μL

Table 2.6 5% stacking gel for SDS PAGE.

H_2O	4.02ml
0.5M Tris (pH6.8)	1.67ml
10%SDS	66.7μL
Acrylamide (30%)	0.87ml
10% Ammonium Persulphate	33μL
TEMED	6.7μL

2.2.7 Immunofluorescent staining.

Suspension cells were washed in PBS and resuspended in 4% PFA. Approximately $5x10^4$ cells in suspension were placed on a single spot outlined with a wax pen on a poly-L- lysine coated slide (Fisher MNJ800010F).

Adherent cells were collected and a cell suspension of $2x10^5$ was added to a 6 well tissue culture plate containing a sterilised coverslip (Fisher MNJ 380-020H). After 24 hours the coverslips were washed and left to fix in 4% PFA.

The cells were left to fix for 30 minutes at room temperature or at 4°C overnight. All incubations were performed in a dark and humidified chamber. All solutions were at 4°C.

Permeabilisation was performed for 10 to 15 minutes at room temperature using either 0.1% (v/v) Triton buffer for adherent cells and 0.5% (v/v) Triton or 0.5% (w/v) SDS for the suspension cells. The cells were then washed three times for 10 minutes in PBS. 100µl of primary antibody resuspended in 10% (v/v) goat serum was applied (see Table 2.3) and incubated overnight at 4°C.

The cells were gently washed three times for 10 minutes in PBS and incubated with the secondary antibodies for 1 hour at room temperature. One 10 minute wash in PBS was followed by a 10 minute incubation in 0.5µg/ml of Hoechst 33342 in PBS (Molecular Probes H-3570). After a final wash in PBS, the cells were mounted on to a slide using 50µl of Vectashield (Vector Labs H-1200).

Images were collected on a ZeissLSM510 with Axiovert 200 microscope and analyzed with laser scanning microscopy and ImageJ software.

2.2.8 IP.

Dynal beads were used as the capture support. Optimisations were performed to find the correct bead type for the antibody (see Table 2.3). A type beads (Invitrogen 100-01D) were prepared by washing three times in PBS with 0.05% Tween pH8.1, G type beads (Invitrogen 100-03D) were washed in PBS with

0.05% Tween pH5.1. Antibodies were incubated with the beads for 1 hour at room temperature in the relevant PBS solution on a roller. The antibody bead complexes were washed three times in the relevant PBS solution. Any excess PBS was removed and the protein lysate applied. The original volume of the beads was not exceeded and any extra volume of protein lysate required was made up using the relevant lysis buffer. This mixture was then incubated for the required time (see Table 2.3) on a roller at 4°C.

2.2.9 Generation of protein lysates for IP.

All procedures were performed using Pi and at 4°C. Protein was concentrated by using protein concentrator columns with a 10000kDa cut off (Fisher VS0101 or VS001).

2.2.9.1 Ripa lysates (whole cell lysates).

Approximately $1x10^6$ cells were washed with PBS and then resuspended in 100μl of the Ripa lysis buffer. The lysate was placed at 4°C for 10 minutes, after which it was resuspended again and left at 4°C for a further 10 minutes. The lysate was clarified at 16060 G at 4°C for 5 minutes.

2.2.9.2 Cytosol preparation.

Approximately $1x10^8$ cells were washed in PBS and resuspended in 5ml of homogenisation buffer. The cells were then passed through a pre-chilled ball bearing homogeniser (8.002mm ball) several times until approximately 90% of the cells membranes had been ruptured. Cell viability was monitored using trypan blue. Nuclei were left intact.

The lysate was clarified by spinning at 690 G for 5 minutes to remove the nuclei. The cytoplasmic fraction was then centrifuged at 100,000 G (using Beckman tubes P80223) for 1 hour. The cytosol fraction was the supernatant leaving the remaining membrane fraction in the bottom of the tube.

The remaining nuclear and membrane fractions were solubilised using Ripa buffer.

2.2.9.3 Hypotonic buffer cell lysis.

Approximately $1x10^8$ cells were washed in PBS and resuspended in 5ml of hypotonic buffer and left at 4°C for 10 minutes. The same procedure as the cytosol purification was used except a pestle was used to break open the cells. The lysate was clarified twice by spinning at 690 G for 5 minutes to remove the nuclei. The remaining nuclei were washed in hypotonic buffer before being solubilised in RIPA buffer.

2.2.9.4 IP pre-optimisation conditions.

Optimised conditions used in these experiments involved varying the lysis buffers used, the method of elution of the IPs and the clean up steps. The optimal cell solubilisation conditions for IP had been determined in preliminary work (data not shown) and for the experiments described in this thesis, IPs were performed using the lysis conditions described above.

Immunoprecipitated protein often appeared as smearing on western blotting. This smearing was reduced in some instances by eluting the IPs in 2xSDS buffer with agitation only and without β mercaptoethanol.

The concentration of protein lysates results in the increased concentration of salts or detergents from the lysis buffer, which can disrupt the IP and the western blotting. For this reason, the lysates/ IPs were processed with an UPPA kit (GenoTech 786-121) according to the manufacturers' instructions.

2.2.10 Staining and processing IP gels for mass spectrometry.

Gels were coomassie stained by using the Brilliant Blue dye kit (Sigma 379204) according to the manufacturers' instructions. Gels were silver stained using the SilverQuest Silver Staining Kit (Invitrogen LC6070) according to the manufacturers' instructions.

Gel slices of approximately 1-1.5mm were cut from the gel lanes and processed using the Montage In-GelDigest$_{zp}$ kit (Millipore LSKG DZP 96). The slices were placed in PCR plates and incubated in 100µl of buffer 1 for 30 minutes at room temperature after which the buffer was removed. This was then followed by

further 30 minutes incubation with 100 µl of buffer 2 (performed twice). After the removal of buffer 2, the gel slices were dehydrated with 100µl of 100% acetonitrile for 10 minutes. This was then removed and replaced with 20µl of buffer3; the plate was then sealed and incubated overnight at 30°C.

The next day, the samples were extracted into 100 µl of 0.2% trifluoroacetic acid for 1 hour at room temperature. The trifluoroacetic acid extracts were then removed and dried down in a speed vac at room temperature. These samples were then resuspended in 5% formic acid and submitted for mass spectrometrical analysis.

2.2.11 Mass spectrometry analysis.

An essential tool in proteomics is tandem mass spectrometry. A process of shotgun proteomics was used in these experiments to identify proteins which interact with ID4. An overview of this process is outlined in Figure 2.3.The raw data from the mass spectrometer is then processed using the ProteinLynx Global SERVER (Waters) to generate ‘peak list’ files, which can then be applied to the Mascot program (Matrix Science), which applies algorithms. This then applies probabilities for the amino acid sequences identified belonging to particular proteins.

The Scaffold 2 Proteome Software program was then used to apply our specific analysis criteria. In these experiments the data has been analysed by the unique number of peptides recognised and this is how this data is listed.

These peptides were then further sorted using a ‘minimum peptide filter’, which sets the unique number of peptides; the ‘minimum peptide probability filter’, which sets the minimum probability of the peptide identity from a least one spectrum and the ‘minimum protein’, which defines the probability that the protein identification is correct. This is supported further by analysing the spectra for the peptides identified.

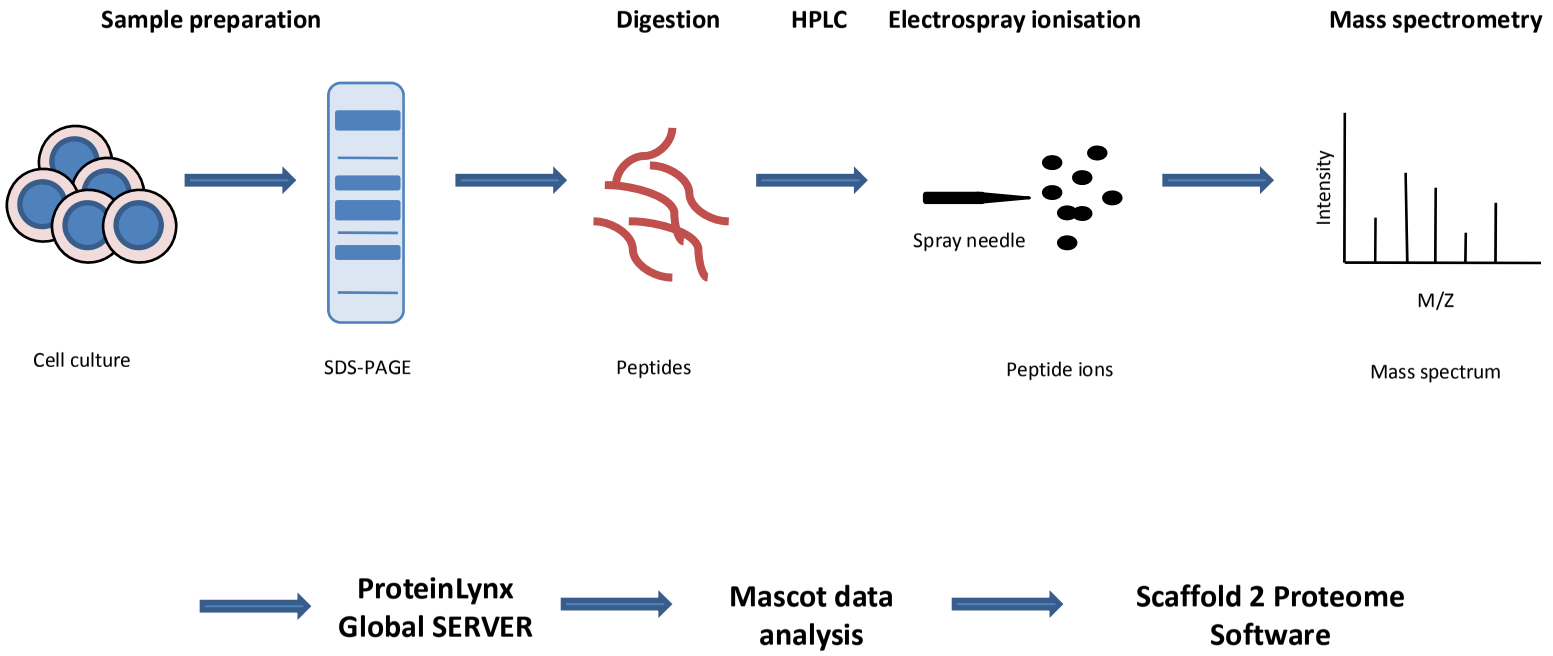


Figure 2.3 Shotgun proteomics methodology outline. Sample preparation in the appropriate lysis buffer and purification by SDS-PAGE. Visualised bands on the gel are then excised, purified and digested by trypsinoslysis to produce peptides. Capillary liquid chromatography passes the peptides through a gradient (organic to aqueous phase) separating the peptides based on their hydrophobicity. The peptides are ionised by a process called electrospray ionisation; these ions then pass through a vacuum into the mass spectrometer where the ions are fragmented by colliding with argon molecules and the mass to charge ratio (m/z) is calculated. Search algorithm tools such as Mascot analyse data bases. An amino acid sequence is then determined from the data (Steen and Mann, 2004) and analysed using Scaffold Proteome Software.

The spectra are produced as described in Figure 2.3, the B ions produced represent the peptide fragment ions which are charged on their amino terminal and the Y ions on the carboxy terminal (Steen and Mann, 2004).

2.2.12 Annexin V Propidium iodide (PI) staining.

Approximately 1x 10^6 cells were washed in PBS and resuspended in 1ml Annexin Buffer. 1.5µl of FITC conjugated Annexin V (Bender MedSytems BMS306 FI) was added and the sample incubated at room temperature for 10 minutes. 250ng of PI (Sigma P4864) was added and the samples analysed immediately on a Becton Dickenson FacScan flow cytometer.

2.2.13 IL3 withdrawal.

BA/F3 transfectants and parental cells were taken whilst in log phase and washed twice in PBS by pelleting at 200G for 3 minutes at room temperature. The cells were then seeded at a confluency of 2x10^5/ml in standard RPMI without IL3. Cells were then analysed for apoptosis using Annexin PI staining (see section 2.2.12) at the relevant time points.

2.2.14 Induction of apoptosis in G1 arrested cells.

NALM-6 TR cells were induced to apoptose with and without tetracycline treatment (see section 2.2.2.1) after 24 hours. Apoptosis was induced using TRAIL ligand (kind gift provided by Dr Marion MacFarlane) at 500ng/µl or 1mM Staurosporine (Sigma S5921) for 4 hours. Cells were analysed for apoptosis using Annexin PI staining (see section 2.2.12).

2.2.15 Cell cycle analysis.

For cell cycle analysis 1x 10^6 cells were washed in PBS and resuspended in 2ml of ice cold 70% ethanol in PBS and left for 30 minutes at 4°C. The cells were then collected, 10µg of PI and 100µg of RNase (Sigma R6513) were added and the final volume made up to 1ml with PBS. This was followed by an incubation at 37°C for 30 minutes after which the cells were kept at 4°C overnight before analysis on a Becton Dickenson FacScan flow cytometer

2.2.16 Purification of lymphocytes from peripheral blood.

Peripheral blood was received in Li-Heparin tubes. The blood was diluted 1:2 with PBS and then gently layered onto a Ficoll cushion (H8889 Sigma) in a 2:1 ratio. The gradient was then centrifuged at 400G for 30 minutes at room temperature. The cells at the lymphocyte interface were collected and washed by diluting the pellet in 3 times the volume in PBS and pelleting at 250G for 10 minutes at room temperature. The cells were resuspended in complete RPMI medium and kept at 4°C from this point. Cells were counted on a Sharfe Coulter cell counter. Excess cells were stored in liquid nitrogen.

2.2.17 Transfections.

Transfections with Lipofectamine (Invitrogen, 11668-019) were carried out according to the manufacturers' instructions, except that half the recommended volume of reagent was used. Four hours post transfection, the medium was changed.

Transfections with calcium phosphate were performed by mixing the DNA in 1ml of 250mM calcium calcium chloride (equivalent for a 10cm dish) and water up to 1ml. This was then added drop wise to 1ml of 2x HBS under gentle agitation; this solution was left at room temperature for 20 minutes after which it was added drop wise to the cells.

2.2.18 Foetal liver assay.

2.2.18.1 Generation of virus.

Approximately $3x10^5$ Phoenix Alpha cells were seeded onto a 10 cm dish 24 hours before transfection. These cells were then transfected with 20µg of the *Mig*R1 (IRES GFP) with *ID4* or *EV* and 1.2µg of the VSV-G vector (Figure 2.2 Panels A and B respectively) using the calcium phosphate method of transfection (see section 2.2.17). The medium was replaced 12 hours post infection and then collected 12 hours, 24 hours and 36 hours after this point. Retroviral medium was stored in aliquots at -80°C.

2.2.18.2 Infection and culture of foetal livers.

Foetuses were obtained from *cdkn2a* knock out (*cdkn2a* -/-) (Serrano *et al.*, 1996) and *Eμ-myc* mice (Adams *et al.*, 1985) which were kind gifts from Dr Michie and Dr Blyth (Paul O'Gorman Leukaemia Research Centre and The Beatson Institute, Glasgow). Foetal livers were removed from E13.5 stage embryos.

The livers from the *cdkn2a* -/- embryos were pooled (these were all null for *cdkn2a*), whereas the livers from the *Eμ-myc* mice were handled individually since embryos from the same litter could be either positive or negative for the *Eμ-myc* transgene. In these experiments, the embryos negative for the *Eμ-myc* transgene were used as wildtype controls. The *Eμ-myc* embryos were genotyped by PCR for the transgene (courtesy of Dr Michie and Dr Blyth).

The foetal livers were sieved through a 70μm mesh and ficolled by layering on top of lympholyte M (CL5030) (see section 2.2.16).

Viral supernatants were thawed and supplemented with stem cell factor (10ng/ml), FLT3 ligand (10ng/ml), IL7 (10ng/ml) and polybrene 4μg/ml. Cells from each foetal liver were resuspended in 500μl of supplemented viral medium and spinoculated at 900G at room temperature for 90 minutes; the cells were then resuspended in 700μl of fresh viral supernatant and 1ml DMEM (all supplemented) and left overnight in an incubator. The next day the cells were collected, pelleted and resuspended in 500μl DMEM and 500μl virus (all supplemented) and spinoculated as before.

After spinoculation, the cells were then sorted on lineage negative (CD11b, GR-1, CD3, CD4, CD8, Ter119) and Sca-1^+ and $CD117^+$, thereby identifying the stem cells/ progenitor B-cell populations. From this population, cells were also sorted as GFP positive and GFP negative populations. All sorting was performed using the Aria II Bectson Dickenson cell sorter (courtesy of Dr A. Michie).

A mixture of GFP positive and GFP negative cells were finally resuspended in 1ml of alphaMEM and layered into wells containing 20-25% confluent OP9

stromal cells. The cells were sampled on days 2, 5, 8 and 14 for GFP positive populations and immunophenotyping by flow cytometry staining (courtesy of Dr Michie).

2.2.19 RNA extractions.

All RNA extractions for standard reverse transcriptase PCR (RT-PCR) were extracted according to the manufacturers' instructions using the RNeasy Midi kit (Invitrogen 75144).

All RNA extractions used for expression microarrays were extracted using QIAzol Lysis Reagent (QIAGEN 79306) according to the manufacturers' instructions.

2.2.20 DNA extractions.

All DNA extractions were performed using the DNeasy Blood and Tissue Kit (Qiagen 69504) according to the manufacturers' instructions.

2.2.21 Complementary DNA (cDNA) synthesis.

cDNA synthesis was performed using the reverse transcriptase SuperScript III (Invitrogen 18080-051) according to the manufacturers' instructions. For the cDNA synthesis, both Oligo dT and random hexamers were used to prime 500ng of RNA

2.2.22 Xenograft cDNA synthesis

BCP-ALL xenograft RNA samples were provided by the Dr Meyer and Dr Queudeville (University Ulm). These samples had been retrieved from the spleens of non obese diabetic/ severe combined immunodeficiency (NOD/SCID) mice which had been transplanted with unsorted primary human BCP-ALL bone marrow via intravenous injection. cDNA synthesis was performed as described in 2.2.21.

2.2.23 Polymerase chain reaction (PCR) and agarose gel electrophoresis

2.2.23.1 Primer Design

PCR primers were designed using the free online service Primer 3 v0.4.0. http://fokker.wi.mit.edu/primer3/input.htm. The individual primers are listed in Appendix Table A1.1 Primers were purchased from Invitrogen.

2.2.23.2 PCR.

PCR was performed using the enzymes Taq Polymerase (Promega M830B). The individual PCR conditions used for each pair of primers can be seen in Appendix Tables A1.2 to A1.4. All PCR reactions were run on a thermocycler.

2.2.23.3 Agarose gel electrophoresis.

Agarose gels of various concentrations were made by dissolving agarose powder (Invitrogen 16500500) in 1xTBE solution and heating till fully dissolved. The solution was allowed to cool before a final concentration of 0.5µg/ml ethidium bromide (Sigma E2512) was added and the gel set in the appropriate cast.

Approximately 5µl of PCR product was loaded onto the gels and ran at 80Vh-1 in 1XTBE buffer. 5µl of DNA ladder (Abgene SLL-100/LD) was run as a control (Appendix Figure A1.1). The PCR products were visualised on a UV transilluminator.

Chapter 3: Expression of ID4 in normal and malignant tissues

3.1 Introduction.

The existence of four ID proteins, which share a common HLH motif, suggests that each protein, whilst very similar, has a specific and distinctive role. One aim of this chapter was to examine the ID4 protein for any domains which allow it to function differently to the other ID proteins.

The lack of *ID4* RNA expression in normal haematopoietic tissues has been described (Cooper *et al.*, 1997; Kersten *et al.*, 2006; Nogueira *et al.*, 2000; Riechmann *et al.*, 1994), with *ID4* being defined as a tumour suppressor in haematopoietic malignancies (Yu *et al.*, 2005). Furthermore, the knockout of *id4* in mice does not affect B-cell development (personal communication Professor Sablitzky and Dr Capasso).

However, the involvement of *ID4* in the t(6;14)(p22;q32) translocation in several cases of BCP-ALL(Russell *et al.*, 2008) and *ID4* RNA expression in several cases of BCP-ALL independent of the translocation (personal communication Professor R Siebert) imply that *ID4* is a driver of BCP-ALL.

In contrast, in primary CLL cases, the *ID4* promoter was found to be methylated and no *ID4* RNA expressed as a consequence (Yu *et al.*, 2005). This implies that *ID4* may function as an oncogene in a context dependent manner. Therefore, this chapter examined the expression of *ID4* in normal and malignant states in more detail, using bioinformatics, in-house cDNA screening of cell lines, BCP-ALL xenograft material and primary CLL cases.

3.2 The ID proteins contain multiple domains.

Amino acid alignments of the different human ID proteins were performed to identify any domains unique to ID4 and comparisons made between the ID4 sequences of different species, to identify the conservation of these domains across evolution (Figure 3.1).

Box 1, 3 and 4; the HLH domain is located in box 2. ID4 shares approximately 41%, 43% and 45% amino acid homology with ID1, ID2 and ID3 respectively.

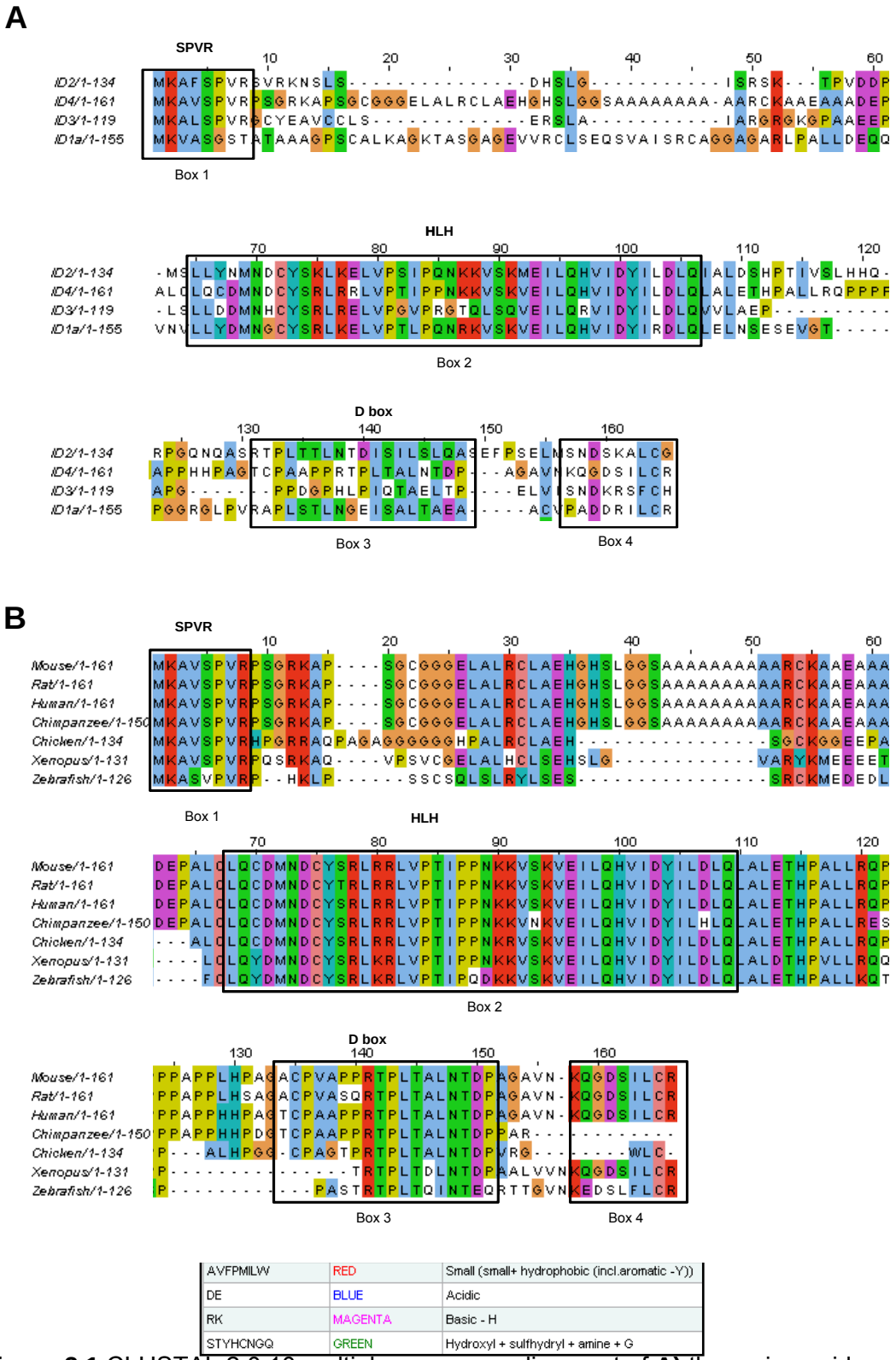


Figure 3.1 CLUSTAL 2.0.10 multiple sequence alignment of **A)** the amino acid sequences of the different ID proteins (for accession numbers see Table 1.4) **B)** The ID4 protein amino acid sequences across the human, rat, mouse, chimpanzee, chicken, zebra fish and *Xenopus* (accession numbers NP_001537.1, NP_783172.1, NP_00112443.1, XP_001170946.1, NP_001035079.1, NP_001080704.1). Legend to amino acid composition can be seen in the box below.

In panel A the region of highest homology is within the HLH region, although there are other small areas of homology across the proteins and are labelled.

Within box 1, the SPVR site is conserved between ID2, ID3 and ID4. The SPVR site was described by Hara *et al.* as a site targeted by cdk2 Cyclin A or cyclin E complexes for phosphorylation (Hara *et al.*, 1997). Box 2 encompasses the HLH domain, which is highly conserved in all the ID proteins. With respect to human ID4 in panel A, there are two significant amino acid changes.

Amino acid 66 is a glutamine (Q), in the position where the other ID proteins have a leucine (L), the former is a polar neutral amino acid and the latter is non-polar neutral amino acid. Amino acid 78 is an arginine (R) in ID4 and a glutamic acid (E) in the other ID proteins. This changes the amino acid 78 from a polar acidic amino acid to a polar basic one. The difference at amino acid 78 involves a change in charge, which may affect interactions in the HLH domain.

The amino acid composition of ID4 is conserved across species (Figure 3.1 panel B), ID4 shares 98% and 95% amino acid homology with the mouse and the rat; 66%, 65% and 56% homology with the chicken, *Xenopus* and the zebrafish and 96% homology with the chimpanzee amino acid sequence. The zebrafish deviates slightly from the other species specifically at the N and C terminuses of the protein, including the SPVR phosphorylation site. Both the zebrafish and the chimpanzee have 2 alternative amino acid changes in the HLH domain.

The D box motif has a core recognition sequence for the APC/C complex which is RXXLXXXN and was defined by examining cell cycle mediated proteolysis of cyclins (Glotzer *et al.*, 1991). The D box motif is conserved in ID1, ID2 and ID4 (Lasorella *et al.*, 2006). The sequence for ID1 is RAPLSTLN, ID2 is RTPLTTLN and ID4 is RTPLTALN (Lasorella *et al.*, 2006). The D box motif is conserved across all species (Figure 3.1 panel B) implicating the importance of its role in the degradation of the ID4 protein.

Two nuclear export signal (NES) regions were identified in ID2 by Kurooka *et al.* A NES1 domain is located in the HLH domain (box 2, see Figure 3.1 panel

A) and is encoded by LQHVIDYILDL (residues 95-105). This domain is conserved across all four ID proteins, but ID2 and ID4 exhibit the highest levels of homology. This domain is also conserved across species in ID4, but the chimpanzee has one amino acid substitution (Figure 3.1 panel B).

NES2 is located in residues 134-150 (LTTLNTDISISLLQASE); this region is only present in ID2. There is a region of high positive charge density present in ID1 with a sequence of RLKELVPTLPQNRKVSK and this region was found to be important for nuclear localisation (Trausch-Azar *et al.*, 2004) and was found in both ID2 and ID4.

These differences in human ID4 compared to the other ID proteins potentially dictate the binding specificities and affinities for other HLH proteins. These differences in amino acid are maintained across the species is shown in Figure 3.1 panel B.

Overall this analysis has shown that ID4 protein domains are conserved across species and that these domains are shared by ID1, ID2 and ID3. All the ID proteins differ either side of the HLH domain, which may allow different binding partners and define different properties.

3.3 The HLH domain is the only conserved domain across the *ID* genes.

3.3.1 The ID mRNA untranslated regions (UTRs) differ.

A comparison was made of the four *ID* family members at the nucleotide level to identify any unique features of *ID4*. This alignment can be seen in the Appendix Figure A 1.2.

Gene expression is regulated by several mechanisms such as transcription factors and promoter methylation. Gene expression is also regulated post-transcriptionally at the mRNA level, where motifs within the 5’ and 3’ UTRs regulate the translation and stability of mRNA.

The open reading frame (ORF) encodes the protein (Appendix Figure A 1.2). The region with the highest homology between the different *ID* RNAs is in the HLH region within the ORF. The 5’ and 3’ UTRs all differ in length between the

different *ID* RNAs. *ID3* and *ID4* have the longest 5'UTRs (405bp and 369bp respectively) and *ID2* and *ID4* have the longest 3'UTRs (814bp and 1534bp respectively) compared to *ID1* and *ID3*. *ID4* shares approximately 74%, 77% and 71% overall nucleotide homology with *ID1, ID2* and *ID3*, respectively.

3.4 Expression of *ID1, ID2, ID3* and *ID4* in normal and malignant adult tissues.

Searching gene expression databases can help identify tissues in which a gene is highly expressed and, in the case of malignancies, highlights patterns of aberrant gene expression. The differential gene expression of the four *IDs* in normal tissues is shown in Table 3.1. These data were derived from the NCBI databases http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi.

The tissues expressing the highest levels of *ID4* mRNA are the embryonic tissues, thyroid, ovary, lung, vascular tissue and nerve tissue. There is no *ID4* expression in any of the haematopoietic tissues except for the bone marrow which has a transcript per million (TPM) of 40. All the *ID* RNAs share expression in several tissues, such as the brain, nerve, ovary, mammary gland and lung, which implies each ID protein has a unique role in these tissues and that they may be transcriptionally regulated by a similar mechanism.

In the haematopoietic tissues *ID1*, *ID2* and *ID3* are all expressed in the blood, the lymph node and thymus. There is only expression of *ID2* and *ID3* in the bone marrow; *ID1* and *ID3* are expressed in the lymph. Only *ID3* RNA is expressed in the spleen and the tonsil.

Therefore in the haematopoietic system, *ID4* RNA is only expressed at low levels in the bone marrow of the haematopoietic tissues. There are no other tissues where *ID4* is expressed without the other *ID* RNAs, but there are instances where *ID4* is not expressed and the other *ID* RNAs are. This shows that *ID4* expression regulates a distinct process which is not required in all tissues.

Table 3.1 Gene expression of *ID1, ID2, ID3* and *ID4* in human tissues. Gene expression patterns as inferred from expressed sequence tag (EST) counts as transcripts per million (TPM) of *ID1, ID2, ID3* and *ID4* in human tissues. Source: UniGene EST Profile Viewer http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Hs.519601. The haematopoietic tissues are highlighted in yellow.

Tissue	*TPM			
	ID1	***ID2***	***ID3***	***ID4***
Adipose tissue	0	76	76	0
Adrenal gland	0	209	0	29
Ascites	649	49	124	0
Bladder	298	66	165	33
Blood	32	80	48	0
Bone	389	153	222	97
Bone marrow	0	61	40	40
Brain	75	99	76	80
Cervix	206	0	226	0
Connective tissue	73	40	106	13
Ear	122	244	1101	0
Embryonic tissue	250	143	227	106
Oesophagus	49	49	0	0
Eye	108	75	127	75
Heart	110	254	11	33
Intestine	182	110	127	16
Kidney	98	84	84	47
Larynx	0	81	81	0
Liver	153	235	47	23
Lung	212	136	88	206
Lymph	90	0	135	0
Lymph node	54	97	10	0
Mammary gland	25	71	129	19
Mouth	14	0	29	0
Muscle	9	0	147	18
Nerve	189	189	316	126
Ovary	126	107	185	282
Pancreas	199	97	55	65
Parathyroid	96	2567	0	0
Pharynx	24	0	24	0
Pituitary gland	0	59	59	0
Placenta	95	45	144	28
Prostate	110	20	73	52
Salivary gland	345	0	295	0
Skin	99	85	330	37
Spleen	0	0	55	0
Stomach	740	164	257	20
Testis	15	24	9	51
Thymus	49	24	12	0
Thyroid	83	41	20	375
Tonsil	0	0	234	0
Trachea	19	19	0	0
Umbilical cord	0	0	0	0
Uterus	158	115	265	38
Vascular	77	57	19	173

* The results are represented as the EST profiles of approximate gene expression patterns as inferred by EST counts as TPM in cDNA pools.

ID4 RNA is expressed in various malignancies (Table 3.2). The levels of *ID4* expression are highest in the ovarian and germ cell tumours and gliomas; *ID1*, *ID2* and *ID3* are also expressed in these malignancies. There is no *ID4* expressed in leukaemia and lymphomas. The other *ID* RNAs are expressed in a number of malignancies; they are expressed in leukaemias and lymphomas, with *ID3* showing the highest expression levels. Therefore *ID4* RNA is expressed in different diseases, but not in haematological malignancies.

3.5 *ID4* is expressed in a subset of derived malignant cell lines.

Examining gene expression across cell lines can also identify malignancies in which the gene is aberrantly expressed. One online database available with this information is Oncomine (www.Oncomine.org). Using this tool, the RNA expression of *ID4* across the 'Wooster' panel of cell lines was examined; this study was chosen because of its comprehensive screen of cell lines derived from various malignancies, totalling 318 samples (Wooster, unpublished data).These data are summarised in Appendix Figure A1.3.

High level of expression of *ID4* was seen in several types of cell lines. The small cell lung carcinoma (SCLC) panel expressed the highest levels of *ID4* (Appendix Figure A1.3 and Table A1.5). The other lung cancer cell lines also show *ID4* expression along with several other cell lines belonging to ovarian adenocarcinomas, breast cancers, endometrial carcinomas, melanomas, other epithelial cancers and neurological cell lines (Appendix Figure A1.3 and Table A1.5).

Using the cutaneous melanoma cell lines as an example, there are cell lines from one malignancy type with and without *ID4* expression, indicating that *ID4* expression may represent an individual subset of that disease; this is also true of the lung adenocarcinoma cell lines and non-SCLC. 96% of SCLC cell line panel express high levels of *ID4* implicating *ID4* in the pathogenesis of this disease (Figure 3.2 panel A and Appendix Table A1.5).

Table 3.2 Gene expression of *ID1, ID2, ID3* and *ID4* in malignant human tissues. Gene expression patterns as inferred from EST counts as transcripts per million (TPM) of *ID1, ID2, ID3* and *ID4* in human tissues. Source: UniGene EST Profile Viewer http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Hs.519601. The haematopoietic tissues are highlighted in yellow.

Tissue	***TPM**			
	ID1	***ID2***	***ID3***	***ID4***
Adrenal tumour	0	233	0	0
Bladder carcinoma	281	0	168	56
Breast (mammary gland) tumour	31	21	190	21
Cervical tumour	289	0	289	0
Chondrosarcoma	410	84	301	24
Colorectal tumour	52	34	60	8
Oesophageal tumour	57	57	0	0
Gastrointestinal tumour	509	100	200	25
Germ cell tumour	41	18	33	15
Glioma	530	102	232	288
Head and neck tumour	80	21	87	101
Kidney tumour	129	57	144	0
Leukaemia	41	72	72	0
Liver tumour	144	196	72	20
Lung tumour	231	19	96	0
Lymphoma	69	83	111	0
Non-neoplasia	184	41	71	41
Ovarian tumour	155	116	129	349
Pancreatic tumour	190	19	76	9
Primitive neuroectodermal tumour.	47	110	102	63
Prostate cancer	77	0	67	0
Retinoblastoma	42	0	21	0
Skin tumour	127	15	461	15
Soft tissue/muscle tissue tumour	158	63	373	31
Uterine tumour	132	110	154	44

* The results are represented as the EST profiles of approximate gene expression patterns as inferred by EST counts as transcripts per million (TPM) in cDNA pools.

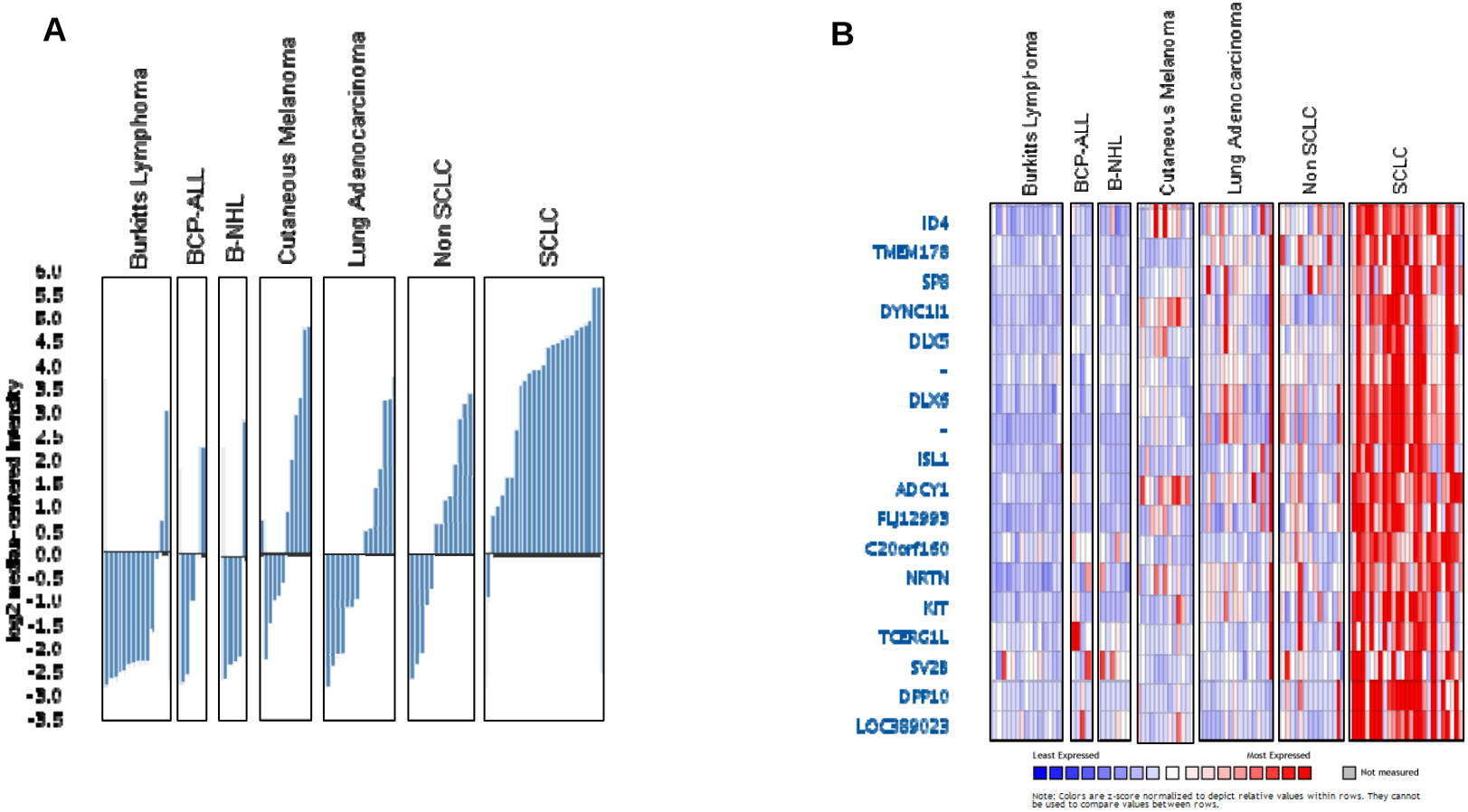


Figure 3.2 Cell lines which express *ID4* in the Wooster cell line collection. Data retrieved from www.oncomine.org using the cancer verses cancer analysis, p=0.0001. Analysed using the Affymetrix Human Genome U133 Plus 2.0 Array. **A)** Example of malignancies with cell lines both positive and negative for *ID4* expression. The list of cell lines expressing *ID4* from these panels can be seen in Appendix Table A1.5 **B)** Gene expression profiles of these malignancies.

Most haematological derived cell lines do not show any expression of *ID4* RNA. However, there are some cell lines that do (Appendix Figure A1.3 and Figure 3.2 panel A). The BCP-ALL cell lines examined in the Wooster Cell Line panel include TANOUE, SEM and RCH-ACV which all have no *ID4* RNA expression; NALM-6 in contrast shows *ID4* expression. The cell lines DG-75 and CA46 (in the Burkitts Lymphoma panel) also show elevated levels of *ID4* as does the B-cell non-Hodgkin's lymphoma (NHL) cell line MHH-PREB-1 (Figure 3.2 panel A).

The relative cytogenetics of each of these cell lines can be seen in Appendix Table A1.6; MHH-PREB-1, DG75 and CA46 all harbour an *MYC/IGH* translocation, otherwise there was no common rearrangement. The expression of *ID4* in some B-cell lines suggests that *ID4* expression may be representative of a subset of disease.

The data was further analysed by examining the gene expression profiles of diseases in which there were cell lines with and without *ID4* expression (Figure 3.2 panel B and Appendix Figure A1.4). There was co-ordinate gene expression of many other genes in the *ID4* positive SCLC cases; this was also true of the lung adenocarcinomas, non-SCLC and cutaneous melanoma cell lines, but not of the haematopoietic cell lines. There were no similar gene expression profiles between different malignancy groups, implying that *ID4* expression exerts a context-dependent effect.

Overall, *ID4* RNA was expressed in a number of different cell lines from different origins, but was restricted within the haematopoietic malignancies. *ID4* expression was found in 4 B-cell lines, which do not share a common origin or gene expression profile.

3.6 *ID4* is expressed in a subset of primary BCP-ALL cases with a unique gene expression profile.

Further analysis using the Oncomine web tool was performed, focusing on studies performed using primary BCP-ALL cases. These studies are summarised in Table 3.3.

Table 3.3 The frequency of *ID4* expressing BCP-ALL cases and associated cytogenetics. Primary data retrieved from www.oncomine.org. Cases are analysed at the 95% outlier with a p=0.0001 significance.

Study	*ID4* positive cases of total BCP-ALL.	Sampling group	Other correlations of *ID4* positive samples	Platform
Cario (Cario *et al.*, 2005)	8/52	Minimum residual disease (MRD) patients	No *ID4* expressing cases associated with translocations. No specific correlation with minimum residual disease risk.	In house
Bhojwani (Bhojwani *et al.*, 2006)	8/103	Patients in relapse	2 *ID4* positive cases harbour *ETV6/RUNX1* fusion.	Human Genome U133 Plus 2.0 Array
Bhojwani Study 2 (Bhojwani *et al.*, 2008)	13/99	High risk ALL.	4 *ID4* expressing cases harbour the *TCF3/PBX1* translocation.	Human Genome U133 Plus 2.0 Array
Andersson (Andersson *et al.*, 2007)	12/53	Patients from Sweden.	*3 ID4 expressing cases harbour the ETV6/RUNX1 translocation.* *2 cases harbour the MYC/IGH* translocation. 2 cases harbour the *TCF3/PBX1 translocation.*	In house
Ross (Ross *et al.*, 2003)	13/132	Samples chosen based on cytogenetics.	2 *ID4* positive cases harbour *TCF3/PBX1 translocation*	Human Genome U133A Array,
Sorich (Sorich *et al.*, 2008)	10/161	Newly diagnosed patients.		Human Genome U133A Array
Holleman (Holleman *et al.*, 2004)	12/146	Newly diagnosed patients.		Human Genome U133A Array

ID4 RNA was consistently expressed in approximately 10% of paediatric BCP-ALL in all the cases studied. The cytogenetic analysis showed no correlation with any particular chromosomal translocation. However there were several studies where some *ID4* expressing cases also harboured the *TCF3/PBX1* translocation.

The Andersson study showed 2 BCP-ALL cases harbouring the *MYC/IGH* translocation, which also expressed *ID4*; this is similar to what has been seen with the cell lines DG75, MHH-PREB-1 and CA46 (Appendix Table A1.6), which also have the *MYC/IGH* translocation. The cases with increased levels of *ID4* did not appear to have a uniformly poor prognosis; however this requires further analysis from prospective studies.

The *ID4* expressing BCP-ALL cases were then further analysed for coordinate gene expression as was seen with the cell lines. Analysis of the primary BCP-ALL cases showed the patients over-expressing *ID4* also cluster together (Figure 3.3). A list of some the genes found to be co-ordinately altered across the 7 studies analysed in Table 3.3 are listed in Table 3.4.

There were several genes which were up-regulated across more than one of the studies in correlation with *ID4* over-expression. These genes include *PGBD5, TOX3, PDGFRA, PCDH17, SATB2, PTPRM, CHST2 and MCAM*. This clustering therefore defines a subset of BCP-ALL which express *ID4* and also have a unique genetic signature.

The gene *PDGFRA* is expressed in the Andersson, Bhojwani 2, Sorich and Holleman studies and encodes for the protein platelet-derived growth factor receptor, alpha polypeptide, which is a tyrosine kinase. This gene is often the target of aberrant translocations in myeloproliferative disorders, such as that of the *FIP1L1-PDGFRA* in Hypereosinophilic syndrome and other malignancies such as glioblastomas (Andrae *et al.*, 2008).

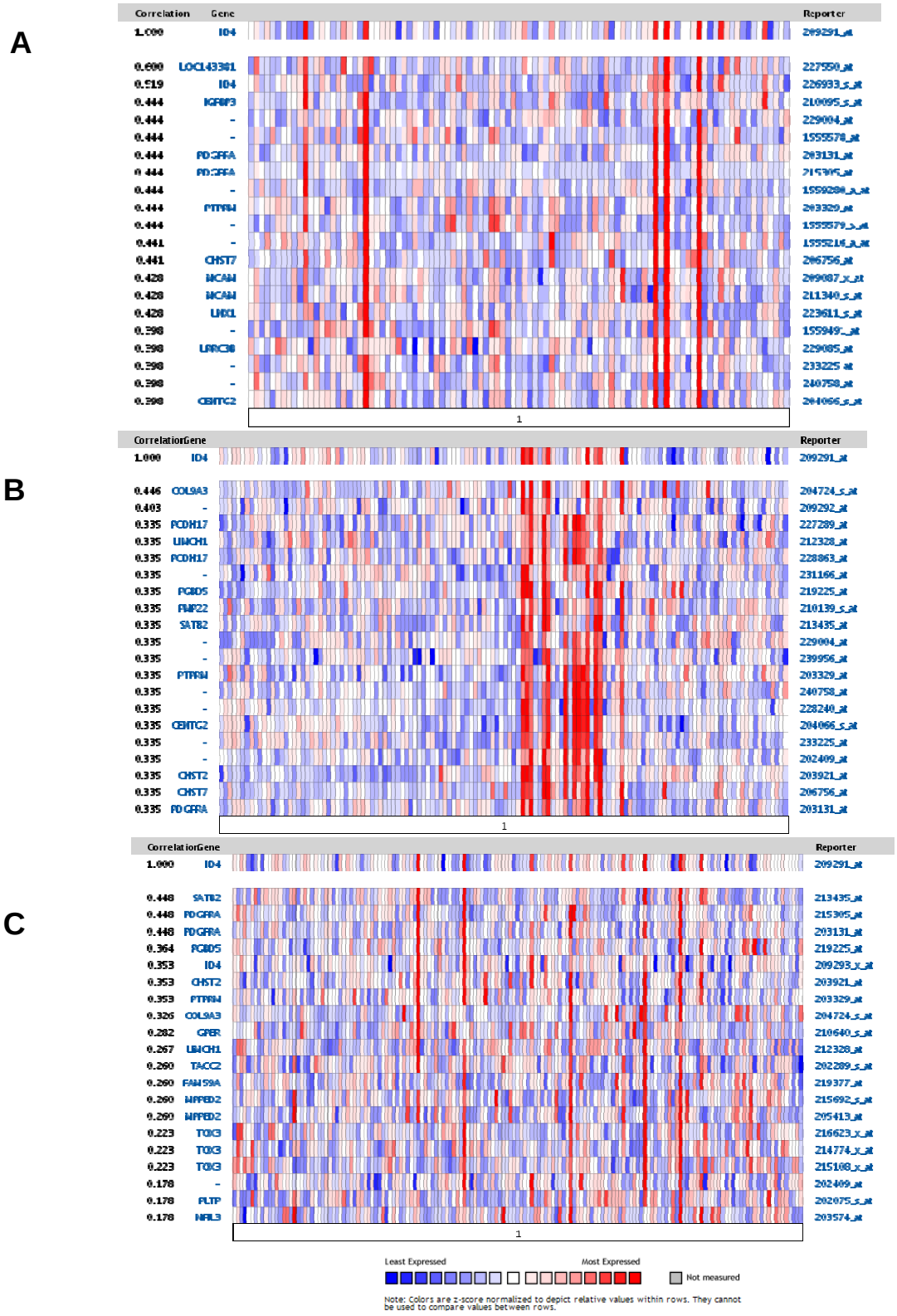


Figure 3.3 Gene expression profiles for 3 different BCP-ALL studies in which *ID4* RNA expression was found. **A)** The Bhojwani 2 study (Bhojwani *et al.*, 2008). **B)** The Ross study (Ross *et al.*, 2003). **C)** The Sorich study (Sorich *et al.*, 2008). Data retrieved from www.oncomine.org using the cancer verses cancer analysis. p=0.0001. Analysed using the Human Genome U133 Plus 2.0Array.

Table 3.4 List of deregulated genes observed in a gene expression analysis in 7 different studies where *ID4* is over-expressed in BCP-ALL cases. Source: www.Oncomine.org. Recurring genes are highlighted in yellow.

Cario	Ross	Sorich	Holleman	Bhojwani 2	Andersson	Bhojwani
TOX3	*TOX3*	*TOX3*	*RHBDL2*	*LNX1*	*SOX9*	*CYP2F1*
PGBD5	*PGBD5*	*PGBD5*	*PGBD5*	*IGFBP3*	*SMOC2*	*PODXL2*
SCN2A	*FLJ13137*	*PDGFRA*	*PDGFRA*	*PDGFRA*	*PDGFRA*	*SMARCA2*
C20orf174	*PMP22*	*TACC2*	*CENTG2*	*CENTG2*	*RGS6*	*MAGEA1*
RGMB	*LIMCH1*	*PPFIA*	*MCAM*	*MCAM*	*CA13*	*HNRNPL*
LRRC3B	*PTPRM*	*PTPRM*	*PTPRM*	*PTPRM*	*IGF2*	*PHLPP*
CLTC	*SATB2*	*SATB2*	*D DN*		*LIMCH1*	*GCNT3*
TSPAN8	*CHST2*	*CHST2*	*CHST2*		*ADAMTS9*	*HOXA3*
ITGB8	*PCDH17*	*GPER*	*PLTP*		*PCDH17*	*ST8SIA3*
SNTB1	*CHST7*	*COL9A3*	*TPBG*		*NXN*	*TACR3*
LRG5		*KCNS3*	*WT1*		*SUSD5*	
CNR1		*NFIL3*	*CAPN3*		*RGS9*	
		PTPRG			*GLDC*	
		FAM59A				
		MPPED2				

The gene *PBDG5* encodes the protein piggyBac transposable element derived 5 and could be implicated in genomic instability. The gene *PCDH17* encodes the protein protocadherin 17 and is involved in cell-cell specific interactions in the brain. There is no reference to either of these genes in relevance to malignancies.

The *TOX3* gene encodes the protein TOX high mobility group box family member 3. This gene belongs to the High mobility group box family of genes, one of which, *TOX*, is involved in T cell development (Wilkinson *et al.*, 2002). TOX3 is a calcium dependent transactivator in neurones (Yuan *et al.*, 2009).

CENTG2 expression is a member of the GTPase family of proteins for ADP-ribosylation factors (Arfs). Arfs are important for the assembly of coated carrier vesicles for transport between the golgi apparatus and the endoplasmic reticulum (Saitoh *et al.*, 2009). *SATB2* is a homologue of *SATB1*, which has been associated with a poor prognosis in breast cancer. *SATB2*, is expressed in the brain, kidney and in Pre B-cells where it augments *IGHμ* gene expression (Dobreva *et al.*, 2003).

CHST2 encodes an enzyme which belongs to the sulfotransferases family. *MCAM* encodes the melanoma cell adhesion molecule, which involved in angiogenesis responses and metastasis and is implicated in cancers such as melanomas and prostate cancer (Ouhtit *et al.*, 2009).

Of all the genes listed in Table 3.4, *PDGFRA* is the most interesting because of its involvement in haematological malignancies of myeloid origin. If *ID4* is involved in an oncogenic pathway with PDGFRA protein, PDGFRA provides a target for therapy.

3.7 *ID4* mRNA expression in haematopoietic cell lines.

Whilst *ID4* is described as a tumour suppressor in haematopoietic malignancies, this analysis of published data has shown *ID4* RNA expression (independent of the *IGH* translocation) in several BCP-ALL cases from different studies. Furthermore *ID4* RNA was seen in 4 B-cell lines. Therefore, *ID4* may not always function as a tumour suppressor in haematopoietic cells. The next

section expanded the screen for *ID1-ID4* RNA expression by RT-PCR in a larger panel of B-cell lines.

Section 3.5 identified the expression of *ID4* RNA in 4 B-cell lines. This was then confirmed by RT-PCR for *ID4* RNA in *ID4* positive cell lines NALM-6, MHH-PREB-1 and DG75 (the cell line CA46 was not available) and the *ID4* negative cell line RCH-ACV (Figure 3.4 panel A).
These results confirmed the expression of *ID4* RNA in NALM-6, MHH-PREB-1 and DG75 and the lack of *ID4* RNA expression in RCH-ACV.

The expression of *ID1-ID4* RNA was examined in a panel of haematopoietic cell lines. The RT-PCR results are summarised in Table 3.5 and Figure 3.4 Panel B.

The BCP-ALL cell lines (Figure 3.4 Panel B) all express *ID1* mRNA, this is also the case with *ID2* with the exception of HB. *ID3* is expressed in 33% of the BCP-ALL cell lines. *ID4* is only expressed in NALM-6. The mRNAs of *ID1, ID2* and *ID3* are all expressed in most of the cell lines screened, but there does not appear to be a pattern associated with any malignancy and *ID1* has a more restricted expression pattern compared to *ID2* and *ID3* in this screen.
Overall *ID4* mRNA expression is present in NALM-6, MHH-PREB-1 and DG75 but not in RCH-ACV which confirms the data from the Wooster cell line panel (Figure 3.2). Also, *ID1*, *ID2* and *ID3* are all expressed in NALM-6 and MHH-PREB-1.Further work in this section would involve examining levels of *ID4* expression in these cells by using real time PCR to identify low levels of expression that might be present in certain cell lines.

Table3.5 RT-PCR screen of *ID* mRNAs in malignant haematopoietic cell lines. Red signifies positive samples and green signifies negative samples.

Pathology	Cell Line	*ID1*	*ID2*	*ID3*	*ID4*
BCP-ALL	NALM-6	+	+	+	+
	LK6	+	+	+	-
	SEM	+	+	+	-
	RCH-ACV	+	+	+	-
	NALM27	+	+	+	-
	380	+	+	+	-
	MUTZ- 5	+	+	+	-
	CEMO-1	+	+	+	-
	PER-377	+	+	+	-
	SUDHL-4	+	+	-	-
	LiLa	+	+	-	-
	RS4;11	+	+	-	-
	HB	+	-	-	-
	KOPN 8	+	+	-	-
	697	+	+	+	-
Burkitts Lymphoma	Elijah	+	+	+	-
	RAJI 2.2.5	-	+	+	-
	KHMB2	+	+	+	-
	RAMOS	+	+	+	-
	BJAB	+	+	+	-
	MUTU III	+	+	+	-
	Daudi	+	+	+	-
	DG75	+	-	+	+
	*MHH-PREB-1	+	+	+	+
Mantle cell lymphoma	SP53	+	+	+	-
	IRM-2	-	+	+	-
	REC-1	+	-	+	-
	JVM-2	-	+	+	-
	Jeko-1	+	+	+	-
	SP1	+	+	+	-
	UPM1	+	+	+	-
DLBCL (Germinal centre)	VALLIOS	+	+	+	-
	SUDHL-9	+	+	+	-
	OCILY 19	-	+	-	-
	SCL1	+	+	+	-
	K422	-	+	+	-
	FL-18	+	+	-	-
Activated DLBCL	OCILY10	-	+	+	-
	OCILY3	+	+	+	-
	RIVA	-	+	+	-
	HBL 2	+	+	+	-
Hodgkin Lymphoma	HDMYZ	-	-	-	-
Marginal Zone	SSK41	+	+	+	-
Myeloma	LP-1	+	+	+	-
	RPMI 8226	-	+	+	-
T cell Lymphoma	Hut 78	-	+	+	-
	JURKAT	+	+	+	-
	Molt-4	+	+	+	-
Normal	785BG	-	+	+	-

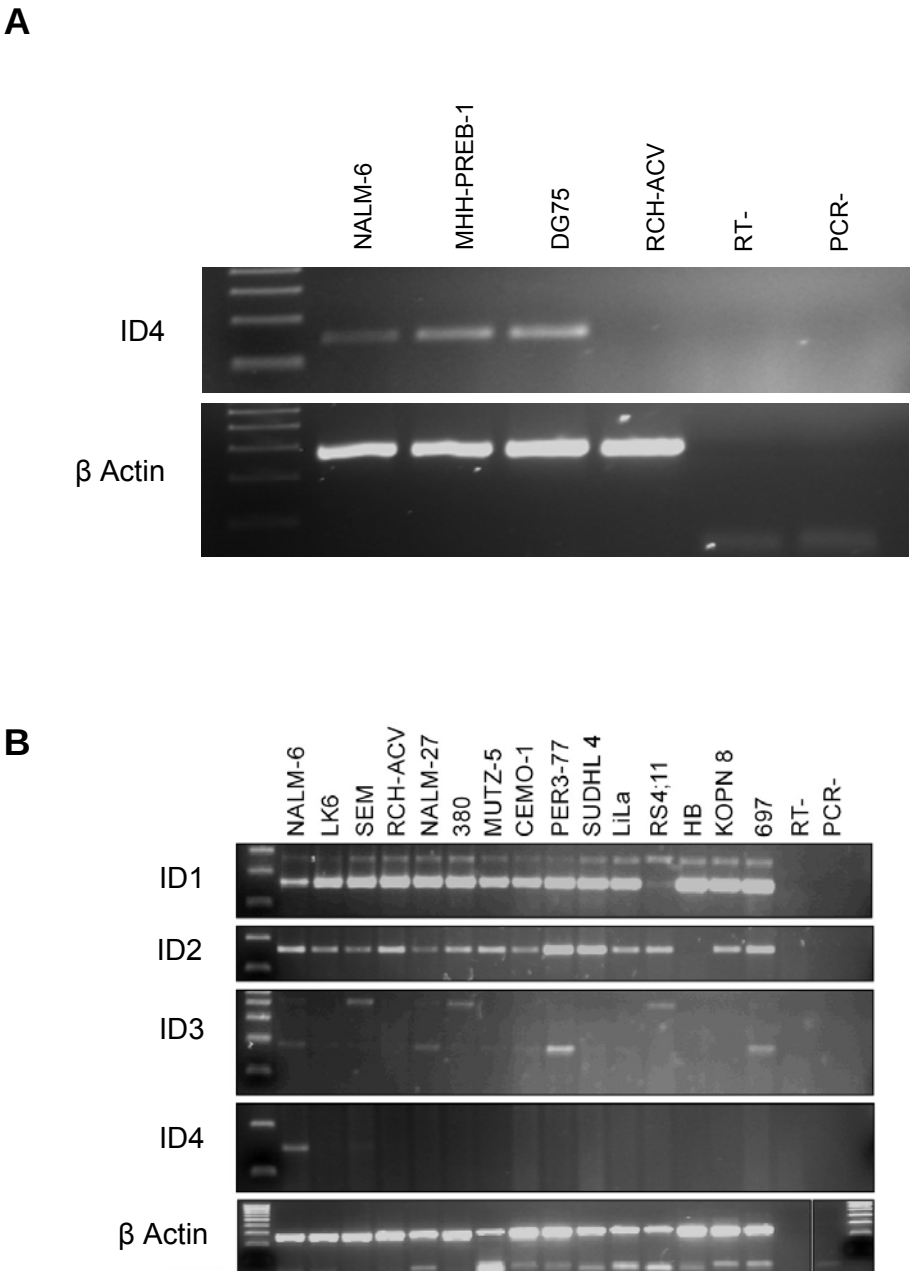


Figure 3.4 RT-PCR screen **A)** PCR products showing *ID4* expression in the B-cell lines positive for *ID4* expression on the Oncomine database **B)** BCP-ALL cell lines for *ID1, ID2, ID3* and *ID4* RNAs. Refer to Appendix Table A1.1 for PCR product sizes and primers used.

3.8 *ID4* mRNA expression in primary BCP-ALL xenograft material.

The expression of *ID4* RNA in BCP-ALL cases without the *IGH* translocation did not correlate with any prognostic group or translocation in the cases examined (Table 3.3). This prompted a further screen of primary BCP-ALL cases harbouring specific cytogenic translocations. Our lack of primary material from cases of BCP-ALL was overcome by obtaining mouse BCP-ALL xenograft RNA from collaborators at the University of Ulm (Dr Meyer and Dr Queudeville).

Mouse xenograft BCP-ALL material is generated by the intravenous injection of primary human BCP-ALL material (patient bone marrow) into immunocompromised NOD/SCID mice. The patient material then engrafts in the host mice allowing expansion of the malignant clone, which can then be further passaged into other mice. In this instance, the RNA was extracted by our collaborators (Dr Meyer and Dr Queudeville, University of Ulm) from the spleens of these mice. This material is therefore representative of the primary malignancy.

23 BCP-ALL xenograft cases with a variety of genetic aberrations were screened for *ID1, ID2, ID3* and *ID4* expression by RT-PCR, using MHH-PREB-1 mRNA as a control. These results can be seen in Table 3.6 and Figure 3.5. There was no *ID4* detected, but there was expression of *ID1, ID2* and *ID3* in 8 %, 60% and 100% of the cases respectively. There was a more restricted expression pattern for *ID1*. Overall there was no correlation of the expression of *ID1, ID2, ID3* and *ID4* with BCP-ALL specific translocations. It may be necessary to screen many more cases to identify such correlations.

Table 3.6 *ID1, ID2, ID3* and *ID4* screen by RT-PCR in BCP-ALL xenograft samples. Red signifies positive samples and green signifies negative samples.

Sample number	Cytogenetics *	Mouse Passage	*ID1*	*ID2*	*ID3*	*ID4*
120		1	+	+	+	-
029	*MLL/ENL*	6	+	-	+	-
064		2	-	+	+	-
112		0	-	+	+	-
002		6	-	+	+	-
026		6	-	-	+	-
006		5	-	-	+	-
052	*MLL/ENL*	0	-	+	+	-
036	*TEL/AML1*	2	-	+	+	-
037	*BCR/ABL*	1	-	-	+	-
022	*TEL/AML1*	3	-	+	+	-
024		3	-	+	+	-
076		3	-	-	+	-
016	*MLL/ENL*	5	-	+	+	-
048	*TEL/AML1*	2	-	+	+	-
039		4	-	-	+	-
089		2	-	+	+	-
118		0	-	-	+	-
093		1	-	+	+	-
007	*TEL/AML1*	4	-	-	+	-
116		0	-	+	+	-
125		0	-	+	+	-
032		0	-	+	+	-

* The cytogenetics represents detected abnormalities. Samples negative in this column may have cytogenetic abnormalities not known.

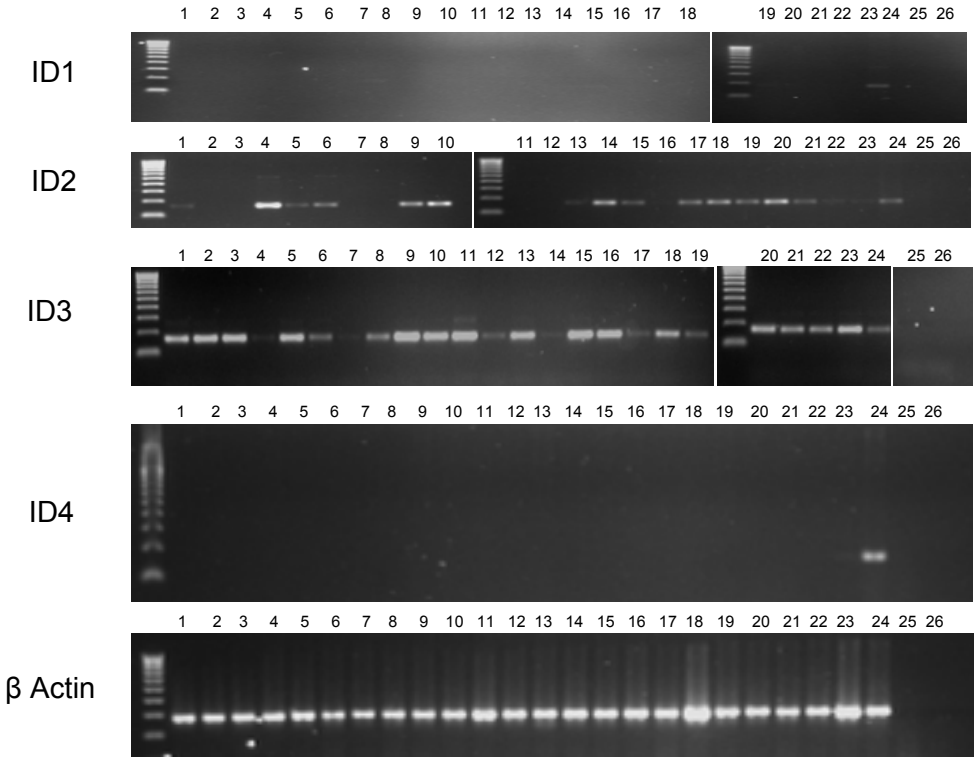


Figure 3.5 RT-PCR screen of xenograft samples for *ID1, ID2, ID3* and *ID4* RNA. For product sizes refer to Appendix Table A1.1.

3.9 *ID4* RNA expression in CLL.

The observation of *ID4* RNA expression in haematopoietic malignancies prompted the re-examination of *ID4* expression in CLL. Although Yu *et al.* stated a lack of *ID4* RNA expression because of promoter methylation, *ID4* RNA expression was never shown in this paper; also different degrees of promoter methylation do not always completely block RNA expression. Therefore, CLL patient samples from an in house tissue bank were chosen at random and screened for *ID4* mRNA and protein expression by RT-PCR and by western blotting respectively (this work was done in collaboration with Dr A Majid).

3.9.1 ID4 RNA is expressed in some cases of CLL.

ID4 RNA expression was found in 34% of the samples screened (21 *ID4* positive cases and 62 negative cases). The results obtained were then cross-referenced against a database containing genetic and molecular screens and general clinical information on the patients. The results of the RT-PCR screen of the CLL patients can be seen in Table 3.7.

These data were further statistically analysed (using the Fisher exact test) to see whether *ID4* expression in these CLL cases was more frequent with any prognostic marker.

ID4 was not preferentially expressed in a subset of CLL associated with an *IGHV* mutated or unmutated phenotype. *ID4* expression did not correlate with any specific cytogenetic abnormality compared to the cases which do not express *ID4*.

ID4 protein expression was also examined in the *ID4* RNA positive CLL cases, but no protein was detected.

Therefore *ID4* RNA was expressed in some CLL cases, which was an unexpected finding.

Table 3.7 RT-PCR screen of CLL samples for *ID4* RNA expression and related clinical data.

Sample number	*ID4*	SEX	*IGHV* Mutated (M)/ Unmutated (U)	11q deletion	13q deletion	17p deletion	Trisomy 12	Other abnormality
597	+	M	U	0	yes	0	0	0
709	+	F	M	0	yes	0	0	t(14;18)(q32; q21)
984	+	M	M					
1045	+	F	M	0	yes	0	0	0
1129	+	M		0	yes	0	0	0
1146	+	M	U	0	yes	0	0	0
1150	+	F	M	0	0	0	0	0
1151	+	M	M	yes	0	0	yes	0
1168	+	M		0	yes	0	0	0
1172	+	M	M	0	yes	0	0	0
1196	+	F	U	0	0	0	yes	14q del
1223	+	M	U	0	0	yes	0	0
1228	+	M	M	0	0	0	0	0
1229	+	M	M	0	yes	0	0	0
1239	+	M	M	0	0	0	0	0
1246	+	M	M	0	0	0	0	0
1289	+	M	U	0	0	0	yes	t(7;15), t(9;22)
1333	+	F	U	yes	0	0	yes	loss x and gain 5
1357	+	M	M	0	0	0	yes	del(2)(p12p22)
1154	+	M	U	0	0	0	0	t(?11;17)

3.10 ID4 protein expression in haematopoietic cell lines.

The expression of ID4 protein was examined in haematopoietic cell lines by western blot (Figure 3.6). Panel A shows the western blot for ID4 expression in NALM-6, MHH-PREB-1 and DG75. ID4 protein runs at approximately 17kDa and appears as a doublet. MHH-PREB-1 expressed the highest levels of ID4 followed by DG75 and then NALM-6. The non-SCLC cell line H1299 was also western blotted to examine the expression of ID4 in a non-SCLC cell line; it was negative.

Panel B shows the expression of ID4 in HEK 293 cells which express endogenous *ID4* RNA as shown by gene expression data (http://www.ebi.ac.uk/gxa/gene/ENSG00000172201). The protein levels of ID4 in HEK 293 and NALM-6 are comparable. Panel C shows the BCP-ALL cell line panel screen. The blots show non-specific bands present at 20kDa and 15 kDa, which are evident when the blots are over exposed. Therefore the majority of these cell lines are negative which agrees with the RNA expression data seen in Table 3.5.

ID4 RNA and protein expression is limited within the haematopoietic cell lines; there are more cell lines which express *ID4* RNA than those that express ID4 protein, but this finding may be due to the low detection capabilities of the H70 antibody.

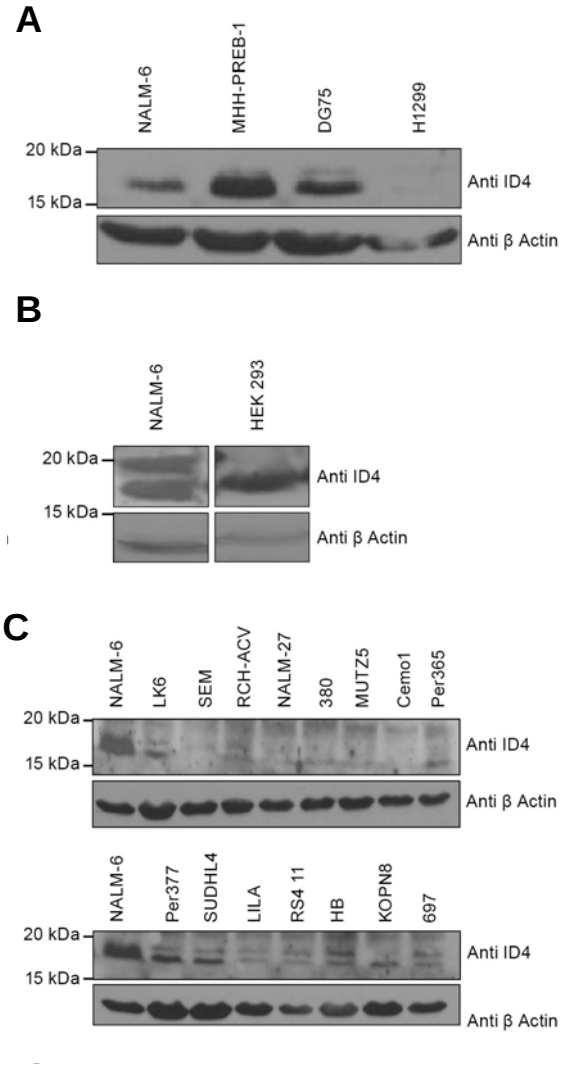


Figure 3.6 Western blot screening of B-cell lines for ID4 protein. Approximately $1.5x10^6$ cells were lysed in 2x SDS loading buffer and loaded on to a 15% SDS PAGE gel. The blots were then probed with anti ID4 anti β Actin sequentially **A)** Panel of B-cell lines that screened positive for *ID4* RNA expression NALM-6, MHH-PREB-1 and DG75 with the non SCLC cell line H1299. **B)** NALM-6 and HEK 293 cells. **C)** BCP-ALL cell line panel.

3.11 Conclusions.

The aim of this chapter was to examine any novel features of ID4 protein and gene which allow it to function differently to ID1, ID2 and ID3 and to examine the expression patterns of *ID4* in normal and disease states.

Structurally the ID4 protein shares similar motifs to the other ID proteins, such as the HLH domain, the SPVR site and the D box domain. Many of the unique features of ID4, such as the amino acids outside the HLH domain and the amino acid substitutions in the HLH domain are conserved across species and therefore are important for its function.

ID4 RNA is expressed within thyroid, ovary, lung, vascular tissue, nervous tissue and embryonic tissue; the bone marrow is the only haematopoietic tissue with *ID4* expression. Where *ID4* is expressed in the bone marrow, requires further investigation since it is not expressed in the haematopoietic system (Cooper *et al.*, 1997; Kersten *et al.*, 2006; Riechmann *et al.*, 1994). ID4 may be expressed in early HSC for a short time (Nogueira *et al.*, 2000).The expression pattern of *ID4* RNA was more restricted than *ID1, ID2* and *ID3* in normal tissues. Overall, *ID4* was never expressed without the other *ID* RNAs in normal or malignant tissues. This implicates a common mechanism of transcriptional regulation of the ID genes, but also that ID4 protein has a different function from the other IDs.

ID4 expression in BCP-ALL and subsets of other diseases shows a clustered gene expression profile. This implies that *ID4* expression in these diseases associated with a new specific subgroup of each disease and has a specific effect within each cell type. This chapter has also shown that ID4 expression does not function as ID1, ID2 and ID3 do in B-cells, by altering B-cell specific transcription. The possible clinical relevance of this observation remains to be determined.

Primary CLL cases identified the expression of *ID4* RNA in 34% of cases. This does not agree with the data from Yu *et al.* (Yu *et al.*, 2005) who identified *ID4* promoter methylation in CLL, which would imply that no *ID4* is being

transcribed, but never showed real time data for *ID4* RNA. Our result may be a consequence of contaminating T cells in the samples and therefore a B-cell purification needs to be done. This may also be a result of sensitivity on detection and we still need to screen more samples and perform real time to measure the levels of *ID4* RNA expression. ID4 protein expression also needs to be confirmed in these cases.

Overall this chapter has shown that ID4 is not expressed within the normal haematopoietic system but is expressed in some haematopoietic malignancies and other diseases where it defines a subtype of these diseases. This therefore defines *ID4* as a context dependent oncogene in haematopoietic malignancies.

Chapter 4: Effects of ID4 over-expression in B-cell lines

4.1 Introduction

The previous chapter demonstrated that *ID4* RNA expression was detected in cases of primary BCP-ALL and CLL, as well as B-cell lines without an *IGH* translocation. These data in conjunction with the t (6;14)(p22;q32) translocation in BCP-ALL suggest, that *ID4* might be a dominant oncogene in B-cell malignancies.

Whether the expression of *ID4* alone is sufficient to cause cellular transformation of B-cells is unclear. To examine this possibility, B-cell lines constitutively or transiently expressing ID4 protein were assessed for ID4 sub-cellular localisation, proliferative and apoptotic responses and effects on the cell cycle.

The B-cell lines used in this chapter include NALM-6, RCH-ACV and Ba/F3 cells. The cell line NALM-6 originates from the peripheral blood of a 19 year old man with relapsed BCP-ALL (Hurwitz *et al.*, 1979); this cell line has no defining chromosomal translocation and has a near diploid karyotype (www.DSMZ.com). NALM-6 cells express endogenous ID4 protein (Figure 3.6).

RCH-ACV was established from the bone marrow of a relapsed BCP-ALL 8 year old patient (Jack *et al.*, 1986); this cell line harbours the *E2A/PBX1* translocation. RCH-ACV cells were shown to be negative for *ID4* RNA expression (Figure 3.4 Panel A).

Stable clones of NALM-6 and RCH-ACV expressing the tetracycline inducible *ID4* pcDNA4/TO myc-His construct (NALM-6 TR and RCH-ACV TR) were provided by Dr Takashi Akasaka (Professor Dyer Laboratory) and generated by electroporation and single cell cloning under antibiotic selection (see sections 2.2.2 and 2.2.3). Tetracycline treatment of either NALM-6 TR or RCH-ACV TR cells, induces the expression of ID4 protein co-expressing the myc-His tag (see Materials and Methods, section 2.2.2 and Figure 2.1). There was one NALM-6 TR clone and 3 RCH-ACV TR clones used in these experiments.

Ba/F3 is an IL3 dependent mouse progenitor B-cell line established from the BALB/c mouse (Palacios *et al.*, 1984). This cell line is dependent on IL3 for survival. In the absence of IL3, these cells apoptose and therefore serve as a good model for growth factor withdrawal, a hallmark of transformation (Hanahan and Weinberg, 2000).

Stable *ID4* pcDNA4/TO myc-His expressing Ba/F3 cells were also generated as described above, but these cells constitutively expressed ID4 protein with a myc-His tag without tetracycline treatment. There were 3 ID4 expressing Ba/F3 clones and 3 EV expressing clones used in these experiments (see Table 2.3 Materials and Methods).

4.2 ID4 protein is localised to the cytoplasm and nucleus of B-cell lines.

The sub-cellular localisation of proteins ID1, ID2 and ID3 has been studied in normal and malignant states. ID1, ID2 and ID3 are both nuclear and cytoplasmic (Deed *et al.*, 1996; Lingbeck *et al.*, 2005; Trausch-Azar *et al.*, 2004). However the sub-cellular localisation of these proteins can change depending on the differentiation stage or the severity of the malignancy.

For example, in the murine hematopoietic progenitor cells 32D, ID2 changes it cellular localisation from nuclear and cytoplasmic to predominately nuclear once the cells are stimulated to differentiate (Tu *et al.*, 2003). In oligodendrocytes, both ID2 and ID4 sequester bHLH proteins into the cytoplasm thereby inhibiting the differentiation of these cells (Samanta and Kessler, 2004). The nuclear localisation of ID2 and ID4 in prostate cancer cells is associated with an increased risk of metastasis (Yuen *et al.*, 2006).

Therefore, the sub-cellular localisation of endogenous ID4 protein was examined in the cell line NALM-6 and the localisation of ectopic ID4 expression was examined in NALM-6 TR, RCH-ACV TR and the Ba/F3 clones. The cellular localisation of ID4 was studied by sub-cellular fractionation and confocal microscopy.

4.2.1 Endogenous ID4 is expressed in the cytoplasm of NALM-6 cells.

Sub-cellular fractions of NALM-6 were produced by using a hypotonic buffer to obtain the cytoplasmic fraction. The nuclear fractions were solubilised with either a Triton-based buffer (Ripa) or an NP40 based buffer. The remaining un-solubilised material was then solubilised in an SDS loading buffer. On western blotting of NALM-6 sub-cellular fractions (Figure 4.1), the majority of the ID4 was in the cytoplasm fraction and ran at the predicted size of 17kDa as a doublet. There was some ID4 in the nuclear fractions with both solubilisation methods; all the ID4 solubilised as there was no remaining ID4 in the nuclear pellet.

Overall, the sub-cellular fractionation of NALM-6 shows ID4 protein to be located within the cytoplasm. Immunofluorescent staining with the anti ID4 (H70) antibody on B-cell lines for endogenous ID4 protein was not done because of the non-specific binding of the antibody which is shown in Figure 3.6.

4.2.2 Ectopic ID4 is located in the cytoplasm of NALM-6 TR and RCH-ACV TR cells.

The sub-cellular localisation of ectopic ID4 was then examined in NALM-6 TR, RCH-ACV-TR (clone 2) cells before and after ID4 induction with tetracycline (Figure 4.2). This was carried out using immunofluorescent staining with the anti Myc tag antibody.

In Panel A, the NALM-6 TR cells show predominately cytoplasmic localisation of induced ID4 protein, which agrees with the fractionation data for the parental NALM-6 cells. There was low level background staining with the anti Myc tag antibody, both in the un-induced NALM-6 TR cells and the parental NALM-6 cells. This could be attributed to the endogenous Myc protein.

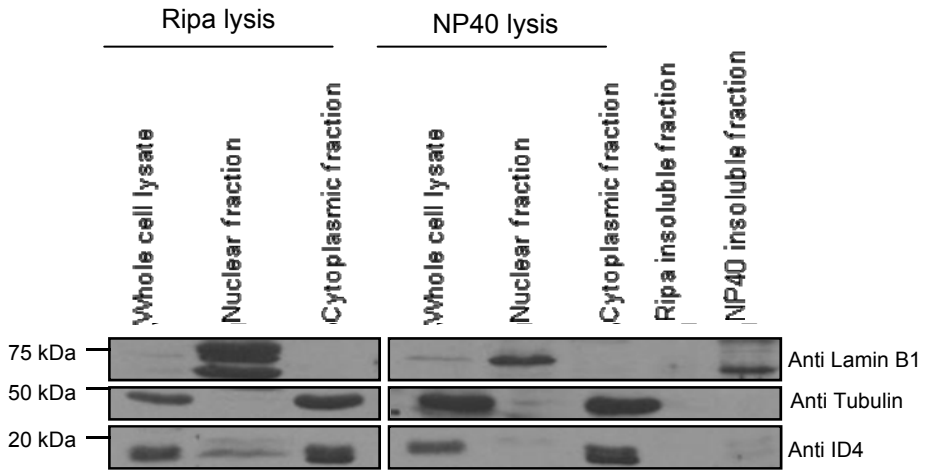


Figure 4.1 ID4 sub-cellular fractionation in NALM-6 cells. Nuclear and cytoplasmic fractionation of NALM-6 was attempted; the nuclear fraction was solubilised with the Ripa lysis buffer or the NP40 buffer. Any un-solubilised nuclear fraction was then solubilised in 2xSDS loading buffer. Approximately 160µg of each fraction was run on a 10% PAGE-SDS gel. A NALM-6 whole cell lysate was run as a control (approximately $1.5x10^6$ cells). The blot was sequentially probed with anti Lamin B1, anti α tubulin and anti ID4 antibodies.

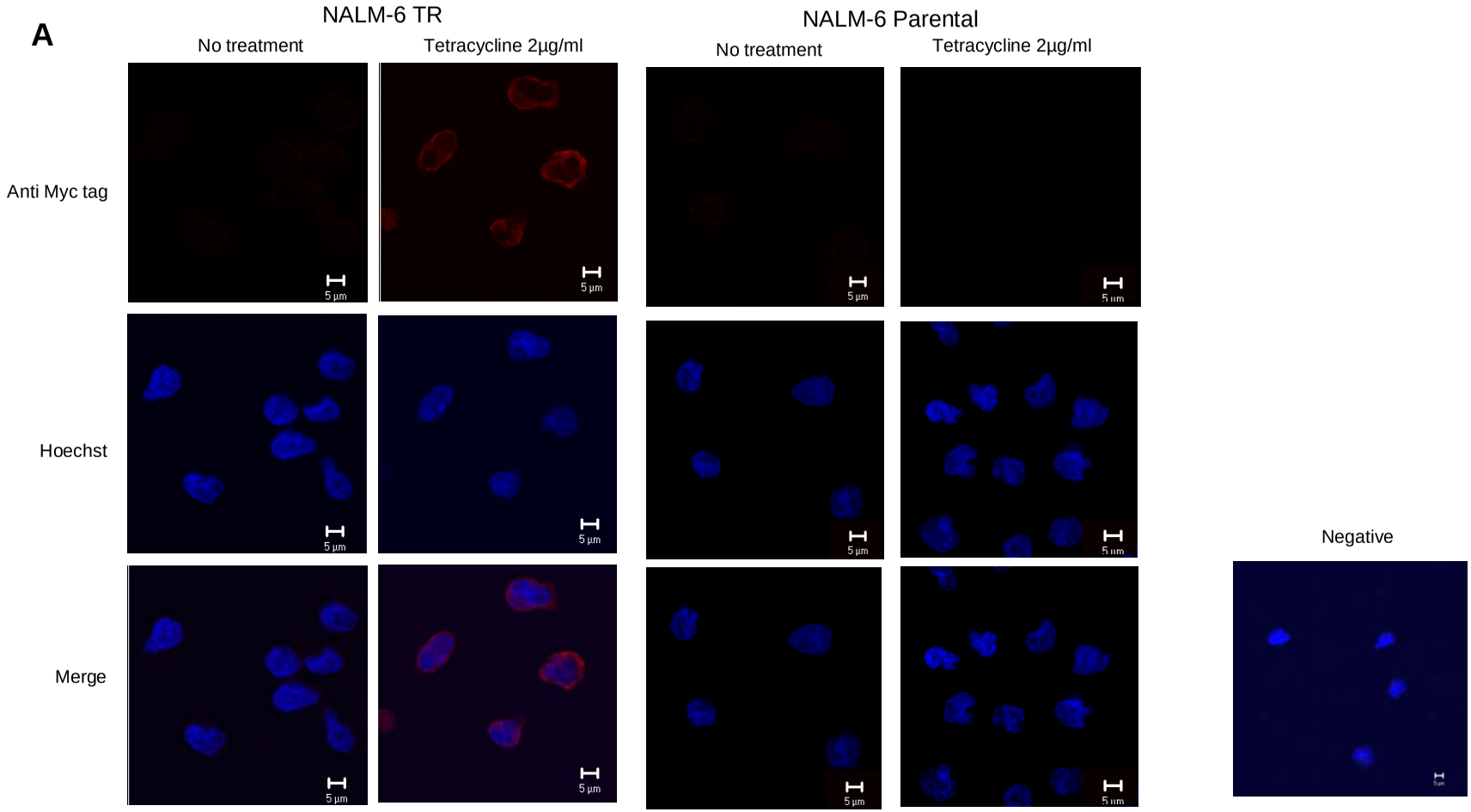
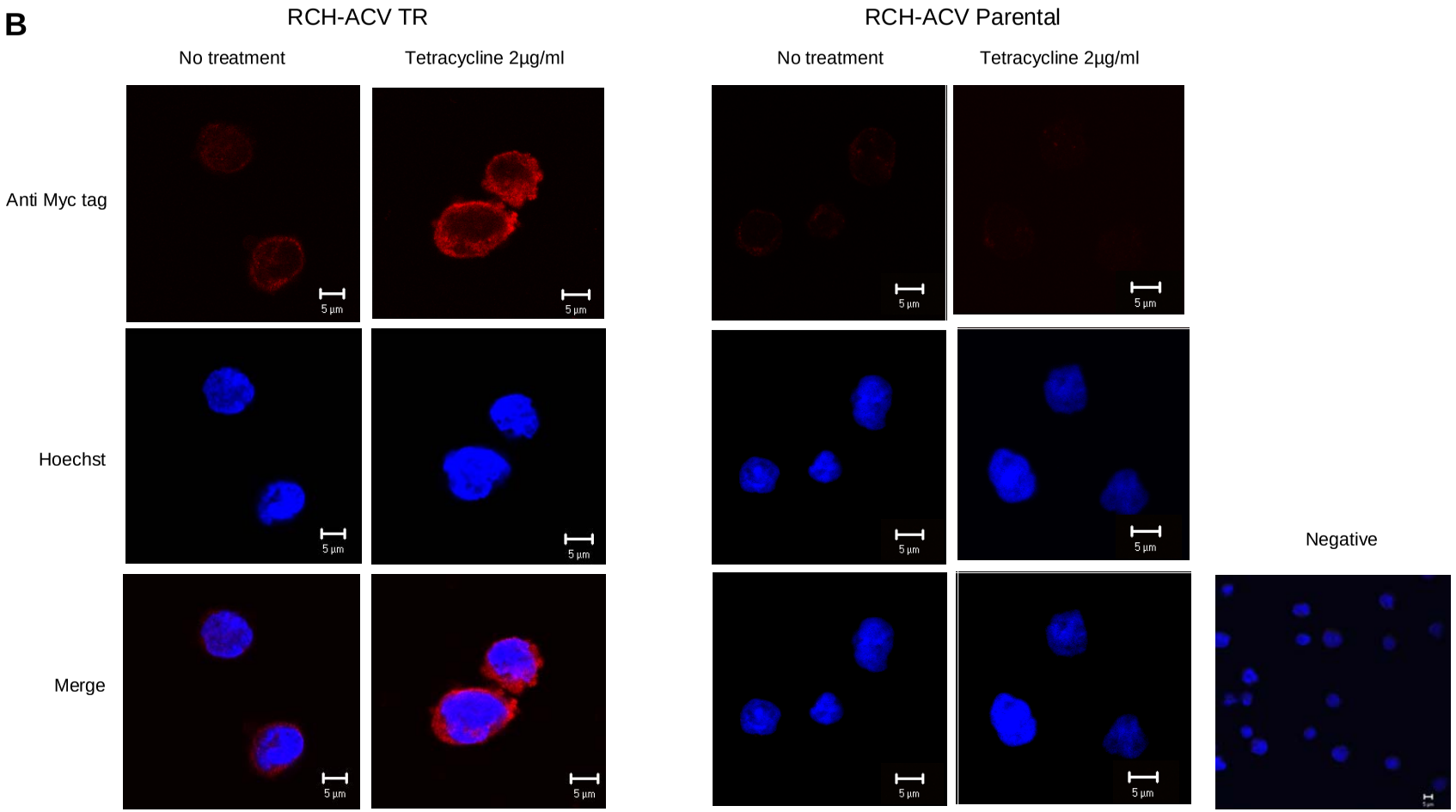
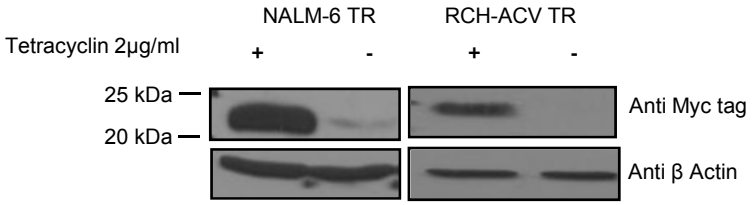


Figure 4.2 Sub-cellular localisation of myc tagged ID4 protein in the clones NALM-6 TR and RCH-ACV TR (clone 2) cells, before after treatment with 2µg/ml tetracycline for 24 hours. Immunofluorescent staining of **A)** NALM-6 TR and **B)** RCH-ACV TR stained with the anti Myc tag antibody (red). The negative control with each TR cell line represents the staining procedure without a primary antibody. **C)** Corresponding western blots to panels A and B, approximately $1x10^6$ cells were lysed in 2XSDS buffer and run on a 15% PAGE-SDS gel and probed with anti Myc tag and anti β Actin antibodies sequentially.


B
RCH-ACV TR
RCH-ACV Parental
No treatment
Tetracycline 2µg/ml
No treatment
Tetracycline 2µg/ml
Anti Myc tag
Hoechst
Merge
Negative
5 µm

C


NALM-6 TR
RCH-ACV TR
Tetracyclin 2µg/ml
+
-
+
-
25 kDa
20 kDa
Anti Myc tag
Anti β Actin

In Panel B, RCH-ACV TR cells also showed a predominately cytoplasmic staining pattern for induced ectopic ID4 protein. There was again a low level background staining with the anti Myc tag antibody in both the un-induced and parental cells, this could also be attributed to endogenous Myc protein in these cells.

The corresponding western blot to panels A and B can be seen in panel C, the Myc tagged ID4 protein runs at approximately 22 kDa. This blot shows with tetracycline treatment there is the marked expression of ID4 protein in NALM-6 TR and RCH-ACV TR cells.

Overall, the induced over-expression of ID4 protein in NALM-6 TR and RCH-ACV TR cells resulted in a predominantly cytoplasmic sub-cellular localisation of ID4 protein after 24 hours. Fractionation had been attempted in these cells, but was not successful.

4.2.3 Constitutive ectopic ID4 expression is both nuclear and cytoplasmic in the Ba/F3 cells.

Immunofluorescent staining was then performed on the Ba/F3 clones expressing ID4 myc-His protein constitutively. Figure 4.3 panel A shows the immunofluorescent staining of Ba/F3 clones 16, 8 and 11, stained with the anti Myc tag antibody. In all three clones, the sub-cellular localisation of ID4 protein was within the nucleus and the cytoplasm. In the Ba/F3 empty vector (EV) clones (panel B) there was low level non-specific background staining with the anti Myc tag antibody, which could be attributed to endogenous Myc protein in these cells.

Western blotting of the 3 Ba/F3 clones expressing ID4 myc-His protein showed a band at approximately 22 kDa, endogenous ID4 protein in NALM-6 and MHH-PREB-1 cells ran at 17kDa (panel C).

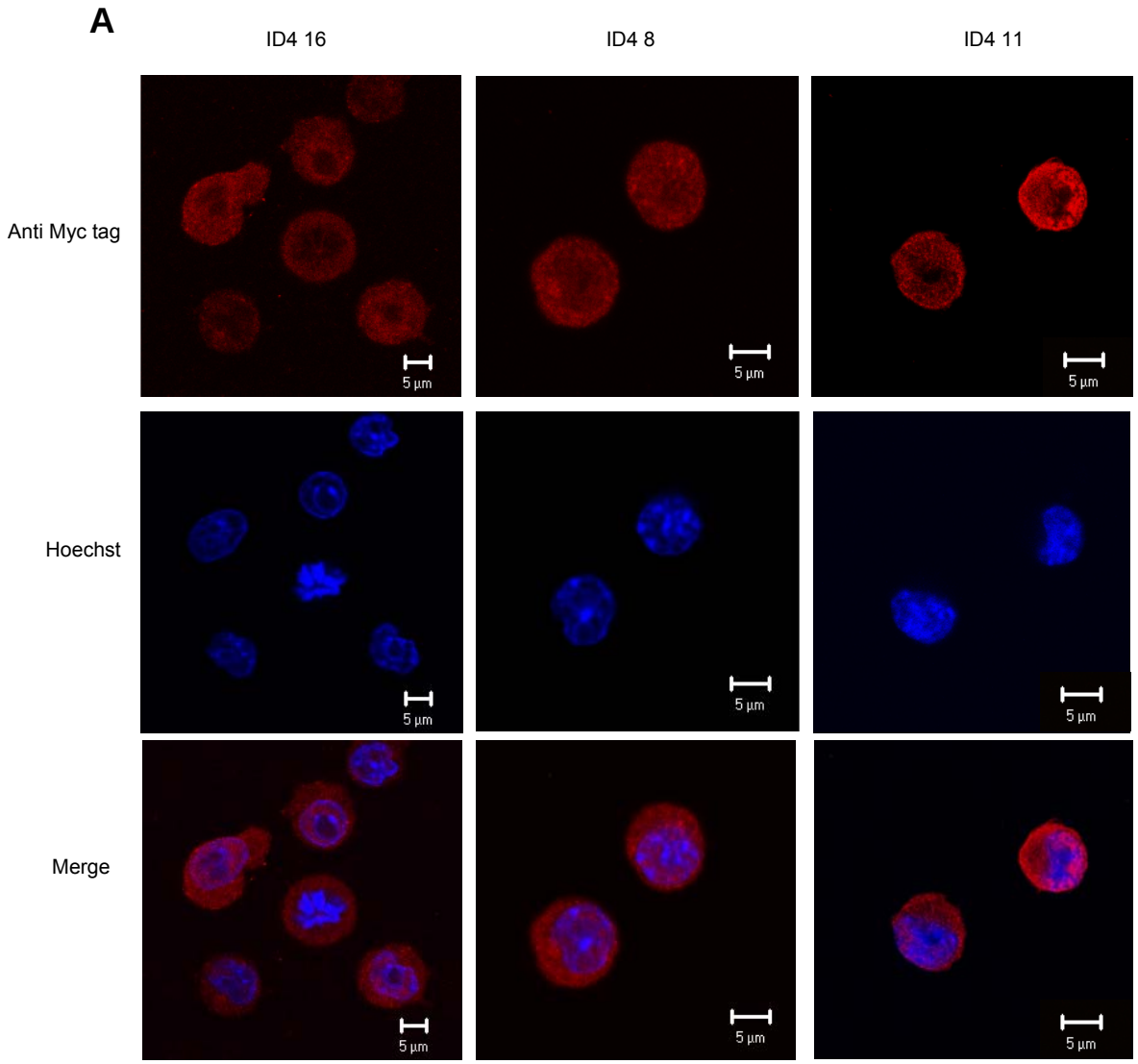
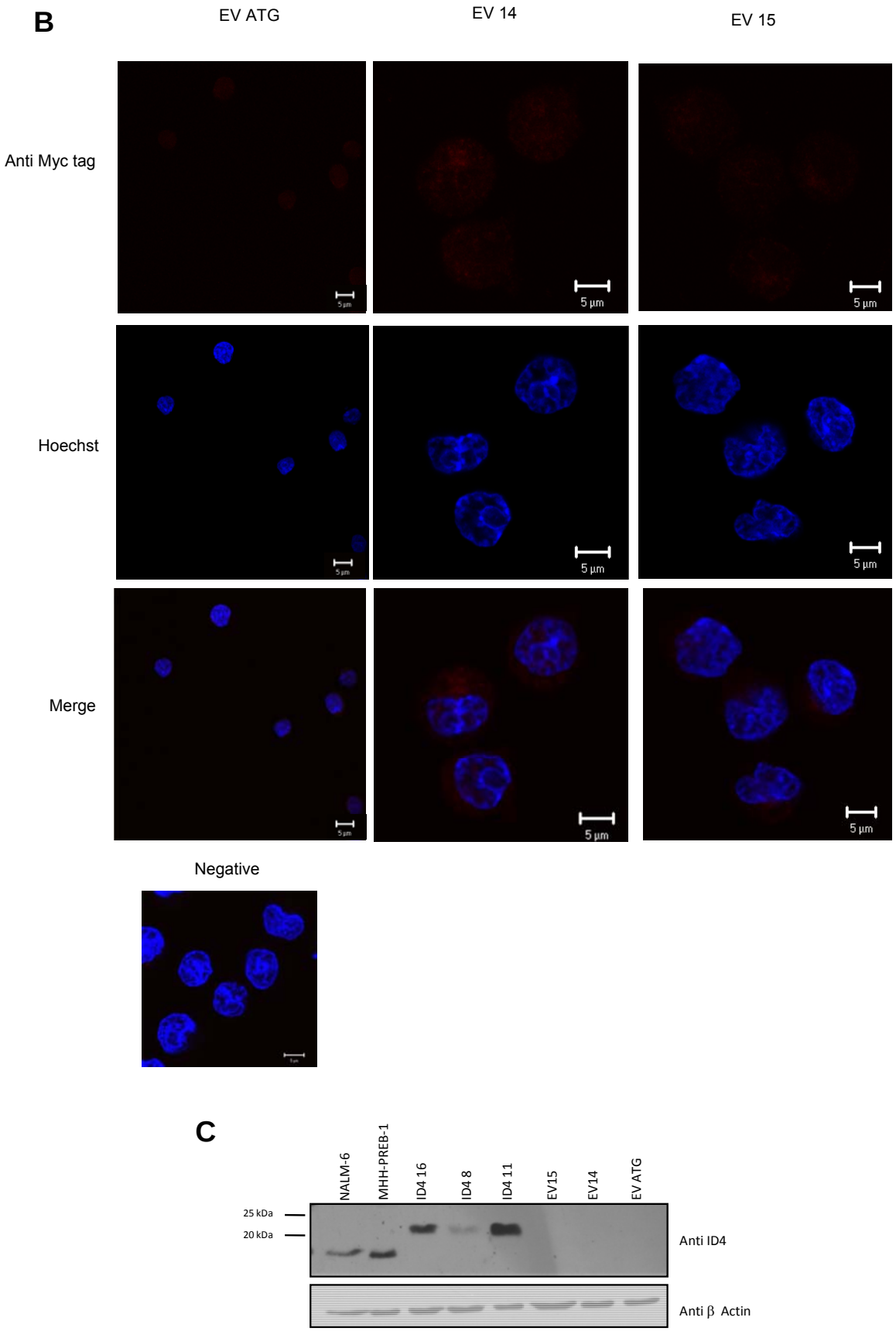


Figure 4.3 Expression and localisation of ectopic ID4 in Ba/F3 clones. Immunofluorescent staining of the Ba/F3 **A)** ID4 expressing clones **B)** EV clones, using the anti Myc tag antibody (red). The negative control represents the staining procedure without a primary antibody. **C)** Ripa lysates were made from each of Ba/F3 clones to show the expression of different levels of ID4 protein. 100µg of lysate was loaded onto a 15% PAGE-SDS gel. NALM-6 and MHH-PREB-1 lysates (generated in the same way as the Ba/F3 clones) were run as controls. The blot was the probed with anti ID4 and anti β Actin sequentially


B
EV ATG
EV 14
EV 15
Anti Myc tag
Hoechst
Merge
5 µm
Negative
C
NALM-6
MHH-PREB-1
ID4 16
ID4 8
ID4 11
EV15
EV14
EV ATG
25 kDa
20 kDa
Anti ID4
Anti β Actin

The Ba/F3 clone ID4 11 had the highest expression of ID4, which was higher than observed in MHH-PREB-1 and NALM-6. The Ba/F3 clone 16 expressed equal levels of ID4 protein to MHH-PREB-1; Ba/F3 clone ID4 8 expressed the lowest levels of ID4 protein. All the Ba/F3 clones used did not express endogenous ID4 protein, however the expression of endogenous *ID4* RNA had not been examined in these cells.

The sub-cellular localisation of ID4 was overall cytoplasmic whether ectopic or endogenous in the human B-cell lines, but both nuclear and cytoplasmic in the mouse Ba/F3 clones.

4.3 Ectopic ID4 over-expression inhibits proliferation of human BCP-ALL cell lines.

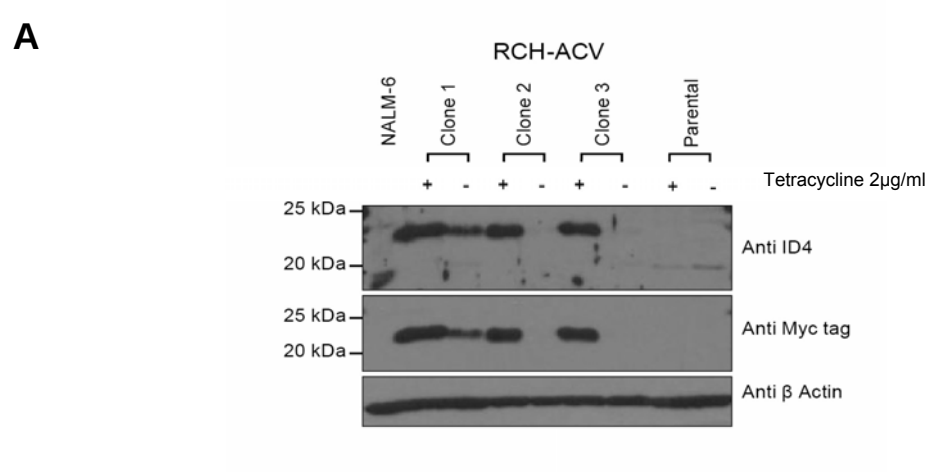
ID1, ID2 and ID3 have been shown to drive proliferation and inhibit differentiation in many cell types (Perk *et al.*, 2005). The effects of ectopic ID4 expression on cell proliferation were examined in the cell lines RCH-ACV TR, NALM-6 TR and in the Ba/F3 clones.

4.3.1 Induced ID4 expression in RCH-ACV TR reduces cellular growth.

The effects of induced ID4 expression in the 3 RCH-ACV TR clones compared to RCH-ACV parental cells was examined (Figure 4.4 panel A). After 24 hours, the RCH-ACV TR clones 2 and 3 expressed ID4 when treated with tetracycline.

Clone 1 constitutively expressed ID4 without tetracycline treatment and was therefore considered leaky, but nevertheless increased levels of ID4 protein were seen after tetracycline induction. All 3 clones when treated with tetracycline expressed similar levels of ID4 protein.

Figure 4.4 panel B examined the growth of the RCH-ACV TR clones and parental cells with and without tetracycline treatment, the cell counts can be seen in Table 4.1. An ideal negative control at this stage would have been RCH-ACV cells expressing an empty vector to assess the effects of tetracycline and zeocin selection on proliferation in these cells, but this was not available.



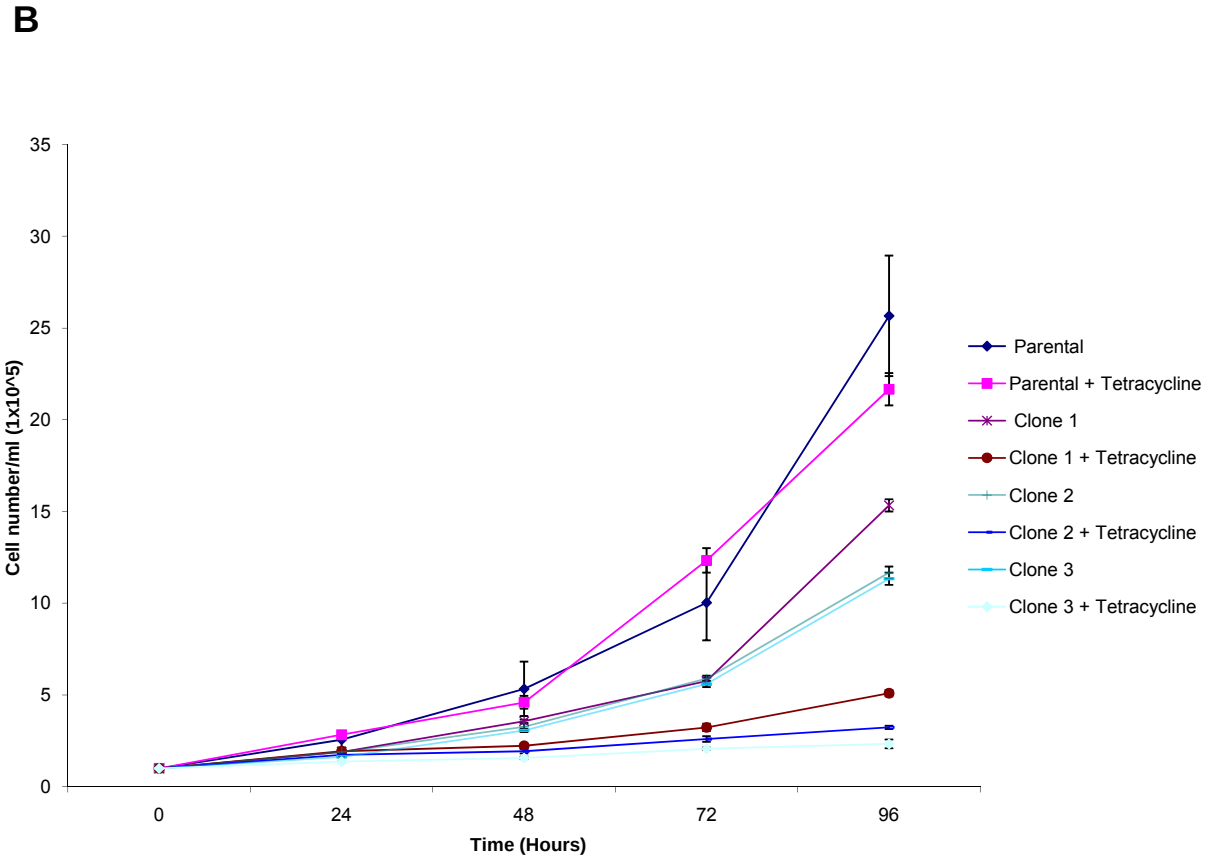


Figure 4.4 Effects of ID4 induced expression on RCH-ACV TR clones 1, 2 and 3. **A)** Western blot of all RCH-ACV TR clones and RCHACV parental cells with and without treatment with 2µg/ml tetracycline for 24 hours. Approximately 100µg of Ripa buffer cell lysate was run on a 15% SDS PAGE gel and western blotted. The blot was then probed with anti ID4, anti Myc tag and anti β Actin antibodies sequentially. B**)** Cells were seeded at a concentration of $1x10^5$/ml (in triplicate) and cell growth analysed over 96 hours at 24 hour intervals after the addition of tetracycline. The results shown are the mean of each clone (n=3) ± Standard deviation (SD).

Table 4.1 Average cell counts x10^5/ml of RCH-ACV parental and RCH-ACV TR clones with or without tetracycline treatment (2µg/ml) over 96 hours. Cells were seeded at a confluency of 1x10^5 cells/ml (in triplicate) at time 0. Cells were counted every 24 hours. The results shown are averages (n=3) at each time point.

	No tetracycline				*With tetracycline*			
	24 hours	*48 hours*	*72hours*	*96 hours*	*24 hours*	*48 hours*	*72 hours*	*96 hours*
Parental	*2.7*	*5.3*	*10*	*25.7*	*2.8*	*4.6*	*12.3*	*21.2*
Clone 1	*1.9*	*3.7*	*5.8*	*15.3*	*1.9*	*2.2*	*3.2*	*5.1*
Clone 2	*1.9*	*3.3*	*5.9*	*11.7*	*1.7*	*1.9*	*2.6*	*3.2*
Clone 3	*1.6*	*3.1*	*5.6*	*11.3*	*1.4*	*1.6*	*2.1*	*2.3*

There were three patterns of growth. The parental cells grew at the fastest rate and tetracycline treatment had no effect on cell proliferation. The RCH-ACV TR clones 1, 2 and 3 with no ID4 induction grew at similar pace, until 96hours, where clone 1 had more cells than clones 2 and 3. Overall these cells grew slower than the parental cells. This is due to the cells being under zeocin antibiotic selection. The slowest growing cells were the clones where ID4 protein had been induced.

At 96 hours clones 2 and 3 had 73% and 79% (respectively) fewer cells/ml when ID4 expression was induced compared to the un-induced cells. Clone 1, with ID4 induction showed a 67% decrease in cell number at 96 hours. Therefore the effects of ID4 induction were less pronounced in clone 1.

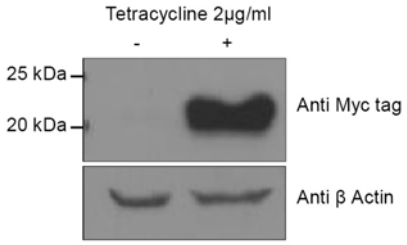
From these data, it can be seen that in RCH-ACV TR clones 2 and 3, induction of ID4 protein in a cell line negative for ID4 expression, reduces proliferation. Clone 1 expressed significant ID4 protein in the absence of tetracycline. These cells proliferated marginally faster than clones 2 and 3 with or without ID4 induction. The different proliferation rates observed in these clones may be due to different levels of ID4 protein being induced.

4.3.2 Induced ID4 expression in NALM-6 TR cells reduces cellular growth.

The effect of inducing ID4 expression in a cell line which already expresses endogenous ID4 protein was examined. The western blot showed that ectopic ID4 protein was still being expressed 72 hours after tetracycline treatment (Figure 4.5 panel A).

The proliferation of NALM-6 TR cells with and without ID4 induction can be seen in Table 4.2 and Figure 4.5 Panel B. These results show that after ID4 induction at 72 hours, there were 47% fewer cells than those without induced ID4 expression. The effect of induced ID4 expression in this cell line was not as pronounced as in RCH-ACV TR cells.

A



B

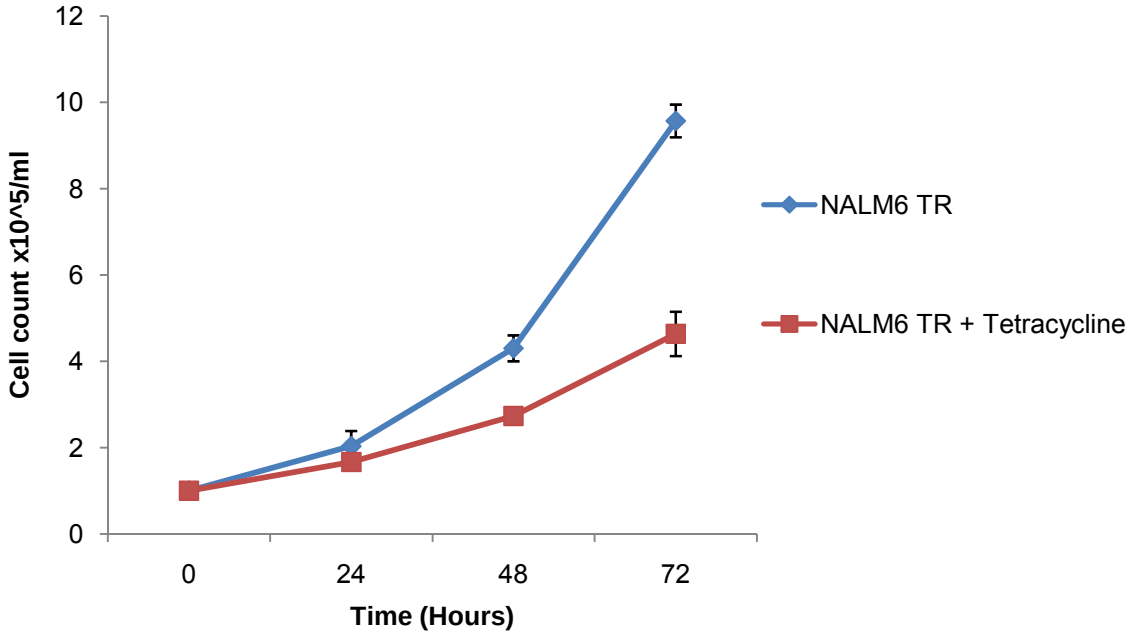


Figure 4.5 Cell proliferation in NALM-6 TR cells following ID4 induction with tetracycline. **A)** Western blot the expression of ID4 after 72 hours with and without treatment of 2µg/ml Tetracycline . Approximately 50µg of Ripa buffer cell lysate was run on a 12% SDS PAGE gel and western blotted. The blot was then probed with anti Myc tag and anti β Actin antibodies sequentially. **B)** Cells were seeded at a concentration of $1x10^5$/ml (in triplicate) and cell growth analysed over 72 hours at 24 hour intervals after the addition of 2µg/ml tetracycline. The results shown are the mean (n=3) ± SD.

Table 4.2 Average cell counts x10^5/ml of NALM-6 TR cells with or without tetracycline treatment (2µg/ml) over 96 hours. Cells were seeded at a confluency of 1x10^5cells/ml (in triplicate) at time 0. Cells were counted every 24 hours. The results shown are averages (n=3) at each time point.

	No tetracycline			With tetracycline		
	24 hours	48 hours	72 hours	24 hours	48 hours	72 hours
NALM-6 TR	2	4.3	9.6	1.7	2.7	4.6

Tetracycline is a broad spectrum antibiotic that inhibits translation in bacteria, but can be toxic to mammalian cells. The recommended concentration of tetracycline used in these experiments is not toxic to cells (www.Invitrogen.com). However, in combination with another antibiotic Zeocin (which works in a different manner to tetracycline), this concentration of tetracycline may affect cellular proliferation. Therefore, an ideal control lacking in these experiments were parental NALM-6 cells and EV expressing clones to assess the effects of tetracycline alone and with zeocin selection on the cells proliferation.

4.3.3 Constitutive ID4 expression in Ba/F3s has no affect on cellular growth.

Figure 4.6 shows the proliferation rates for the Ba/F3 clones. Three clones constitutively expressing ID4 and three EV clones (for controls) were selected for analysis. The constitutive expression of ID4 in the Ba/F3 clones did not affect proliferation when compared to the EV expressing clones.

Overall, the induced over-expression of ID4 in B-cell lines reduces proliferation from 24 hours onwards. This was not apparent in the constitutively expressing Ba/F3 cells.

4.4 ID4 induces cell cycle arrest when expressed in RCH-ACV TR and NALM-6 TR cells.

ID1, ID2 and ID3 are all involved in cell cycle regulation involving the G1 phase, inducing proliferation; and mitosis, which results in increased genomic instability. ID4 has been shown to induce S phase arrest in prostate cancer cell lines in a p53 independent manner (Carey *et al.*, 2009), but has also been shown to be required for G1/S progression in neurones (Yun *et al.*, 2004). These results have shown that ectopic ID4 expression in human cells does not show a normal cell doubling time. Therefore the effects of ectopic ID4 expression on the cell cycle in B-cells were examined.

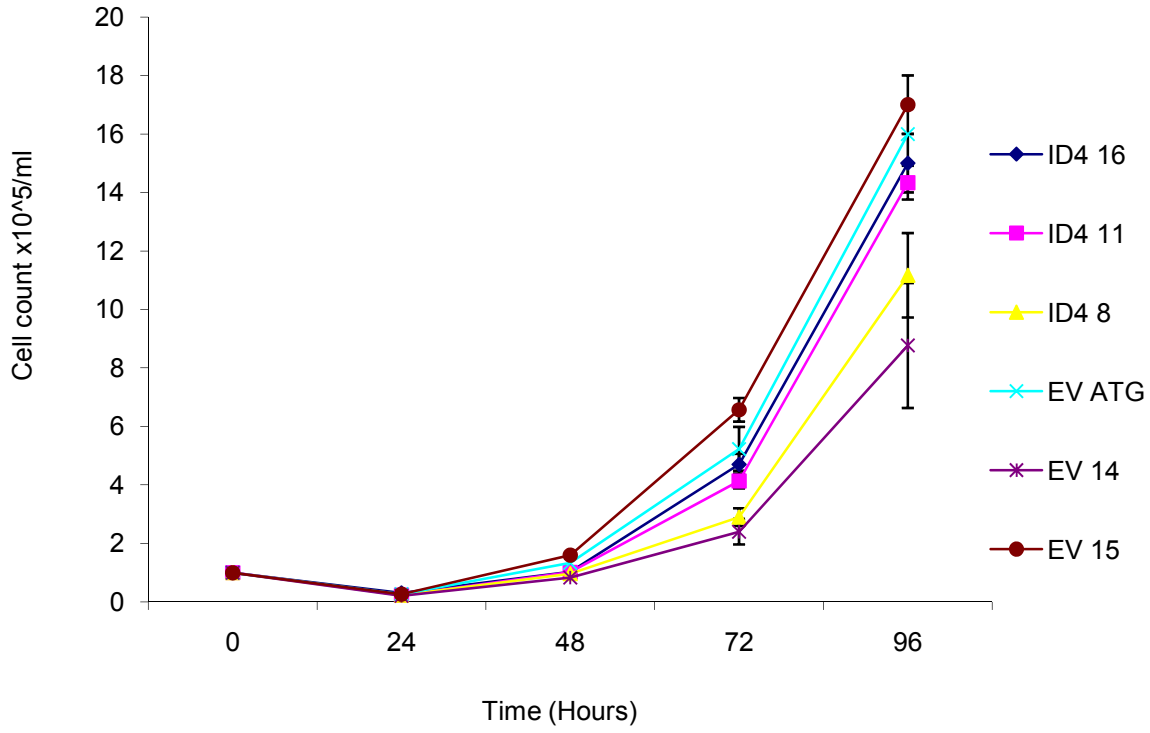


Figure 4.6 Cell proliferation in Ba/F3 clones. Cell proliferation of the Ba/F3 EV and ID4 clones over 96 hours. Cells were seeded at a concentration of $1x10^4$/ml and cell proliferation analysed over 96 hours at 24 hour intervals. The results shown are the mean (n=3) ± SD.

4.4.1 Induced ID4 expression in RCH-ACV TR clone 2 cells induces cell cycle arrest.

ID4 expression was maximally induced in RCH-ACV TR cells after 8 hours (Figure 4.7 panel A). After 24 hours, the levels of ID4 declined, after which time it remained at a steady level. The cell cycle profiles in panel B showed at 12 hours post ID4 induction, the S phase of the cell cycle declined, and by 24 hours the cells were in G1 arrest. At 36 and 48 hours, the G1 block had been partially released but did not fully recover.

4.4.2 Induced ID4 expression in NALM-6 TR cells initially induces cell cycle arrest.

In NALM6-TR, ID4 expression was maximally induced 4 hours after induction; these levels then reduced significantly after the 8 hour time point (Figure 4.8 panel A). The cell cycle profiles in panel B showed that the S phase began to decline from 8 hours onwards; at 36 hours the S phase began to recover and at 48 hours the cells were cycling again, but not as at time 0.

4.4.3 Constitutive ID4 expression in Ba/F3 cells does not affect the cell cycle.

Examination of the cell cycle of the Ba/F3 clones was performed, but there was no difference in the cell cycle distribution between ID4 expressing and EV clones (data not shown).

Therefore the induction ID4 protein in both RCH-ACV TR and NALM-6 TR cells resulted in G1 arrest. In the case of RCH-ACV TR, this had not recovered at 48 hours. However, in NALM-6 TR cells, loss of ID4 induced potential was observed 24 hours onwards resulting in the recovery of the cell cycle. These experiments lacked the relevant controls to examine the effects of tetracycline treatment on the cell cycle in parental or EV expressing cells and will need to be addressed in future work.

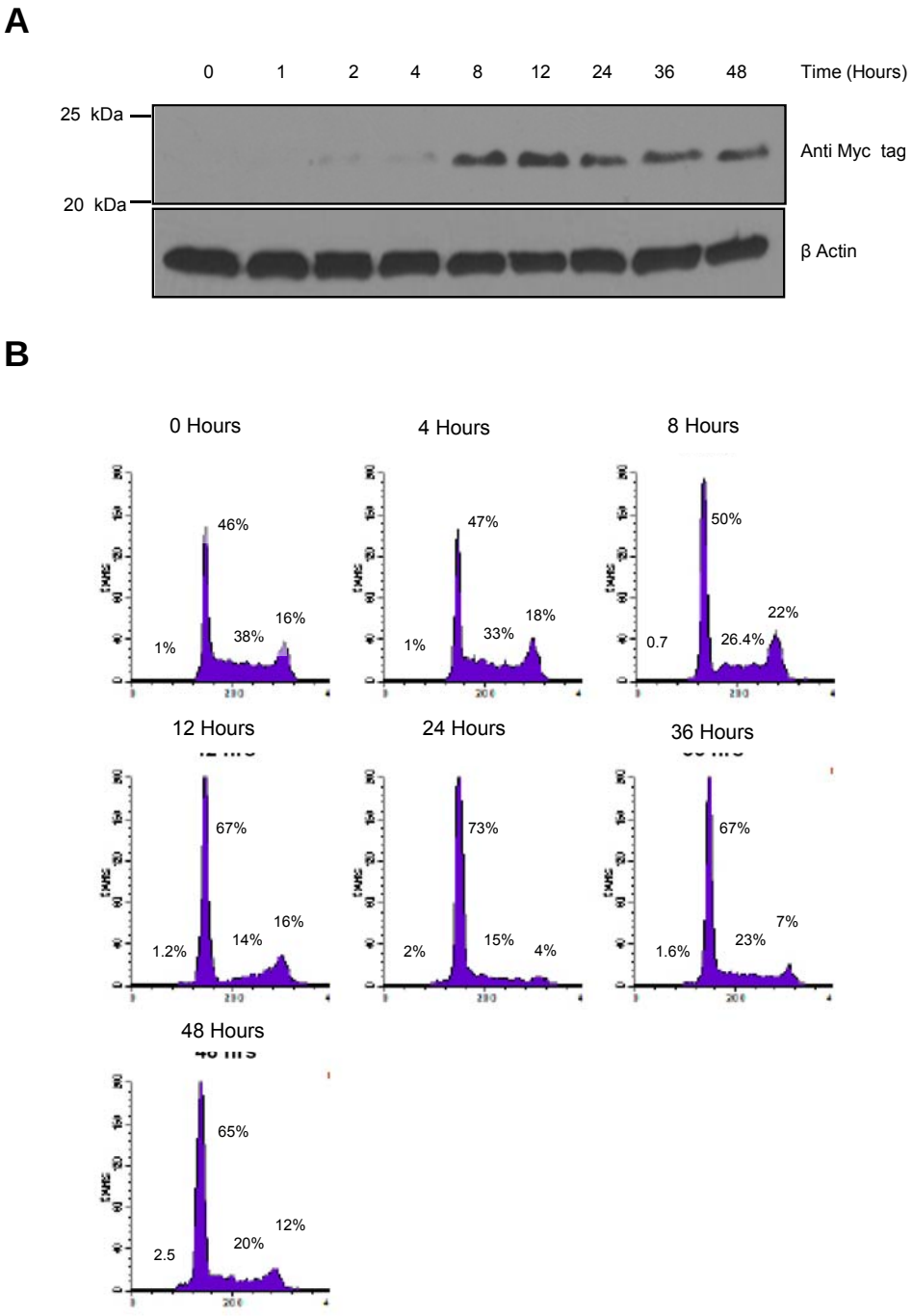


Figure 4.7 Cell cycle time course on RCH-ACV TR clone 2 cells. **A)** Western blot showing a time course examining induction of ID4. Cells were lysed in Ripa lysis buffer and 25μg of lysate was run on a 12% SDS PAGE gel and western blotted. The blot was then probed with anti Myc tag and anti β Actin antibodies sequentially. **B)** Histograms showing the distribution of the cell cycle across the time course for the induction of ID4.

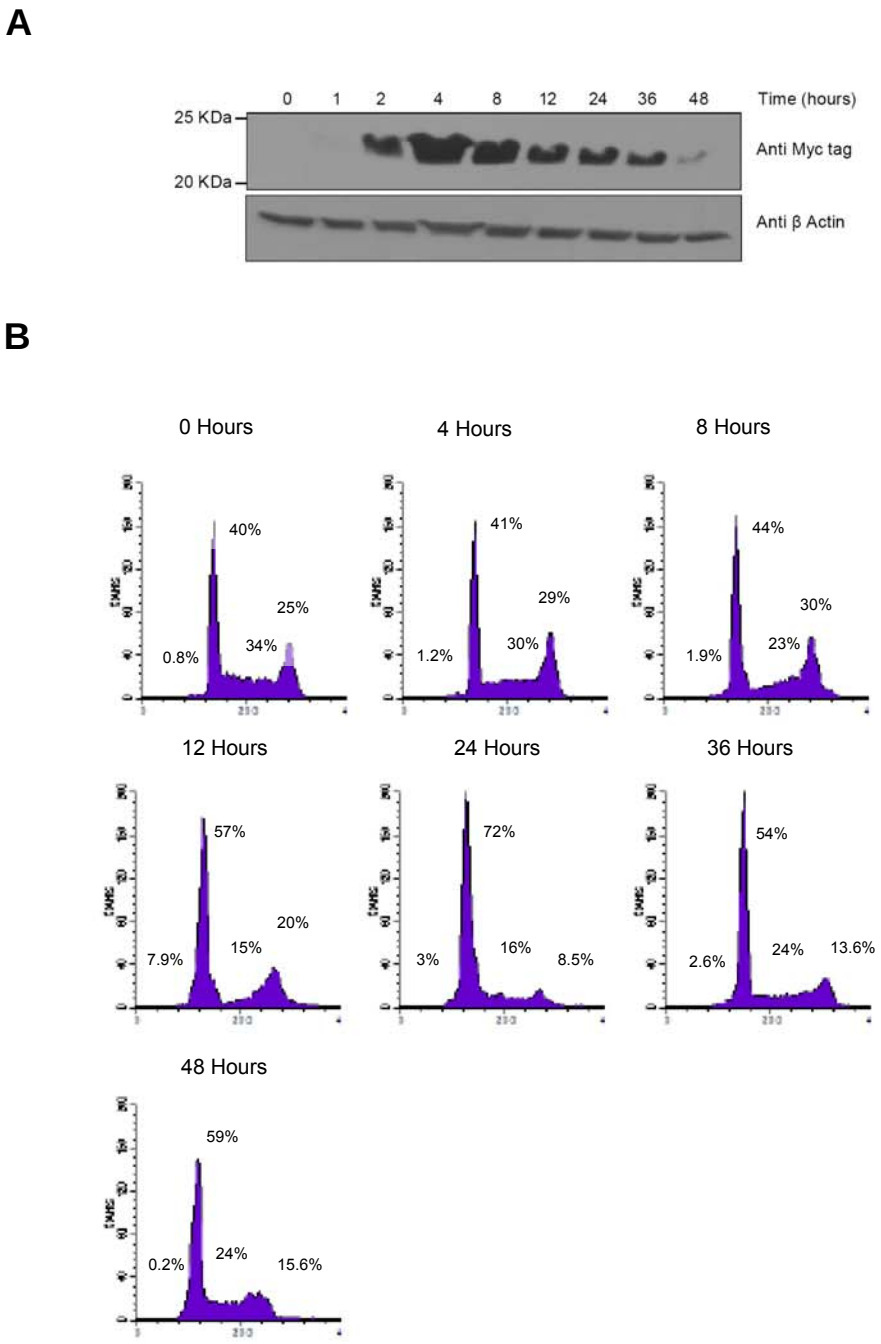


Figure 4.8 Cell cycle time course on NALM-6 TR cells. **A)** Western blot showing a time course induction of ID4. Cells were lysed in Ripa lysis buffer and 25μg of lysate was run on a 12% SDS PAGE gel and western blotted. The blot was then probed with anti Myc tag and anti β Actin antibodies sequentially. **B)** Histograms showing the distribution of the cell cycle phases across the time course for the induction of ID4.

4.5 Examination of cell cycle proteins with ID4 induced expression.

Since ID4 expression in NALM-6 TR and RCH-ACV TR resulted in G1arrest, changes in cell cycle proteins were therefore examined in both. The cell cycle proteins Rb, Cyclin D3, cdk6, p15, p16, p21 and p27 were examined in NALM-6 TR following ID4 induction (Figure 4.9 panel A). These proteins were selected because they are involved in the G1 to S phase transition.

ID4 expression was expressed at its highest levels between 4 to 8 hours after induction with tetracycline in NALM-6 TR. Approximately 24 hours after induction, ID4 protein levels declined. These cells were in G1 arrest after 24 hours of induction, which in NALM-6 TR cells recovered after 36 hours induction.

Cyclin D3 decreased between 4, 8 and 12 hours, but increased again at 24 hours onwards. p15 decreased between 4, 8 and 12 hours, but increased from 24 hours onwards. Cdk6 increased within this time frame and dropped back to base level after 24 hours. This coincided with cell cycle arrest beginning (see panel A).

Cdk4, p16 and cyclin D1 were not detected. p27 overall appeared to increase at 8 hours and there were indications of ubiquitination as is apparent by the smearing above the band of interest. p21 increased over time from 12 hours onwards. There were no changes observed in Rb protein levels until 72 and 96 hours, where it declined. Further work will require then examination of the phosphorylation status of Rb.

Cyclin D3 levels decreased when ID4 was highly expressed and G1 arrest occurred. After this, once the cell cycle began to recover, the levels of cyclin D3 returned to basal level. Within this time frame p15 decreased, possibly being unable to inhibit cyclin D3 and cdks from phosphorlyating Rb and in response there is an increase in p16.

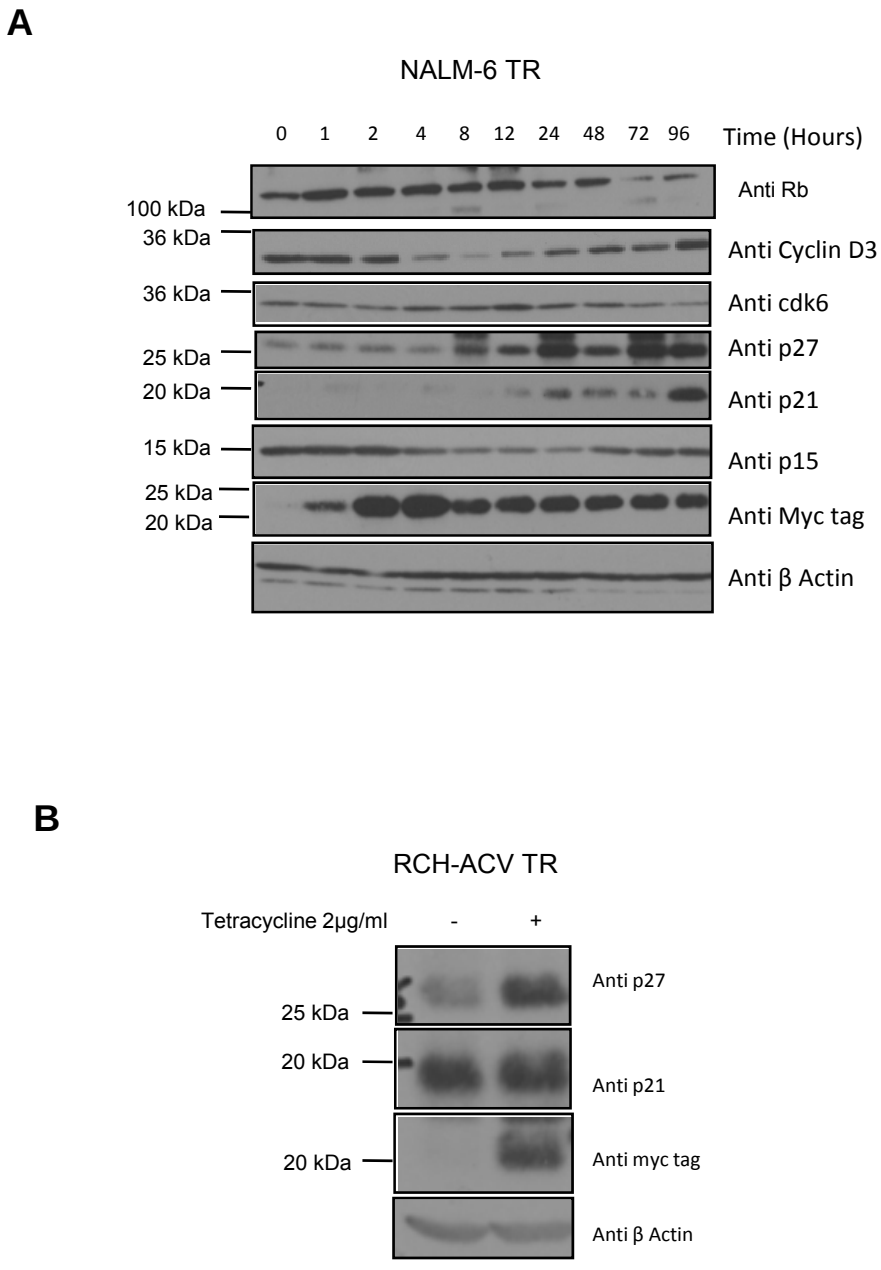


Figure 4.9 Cell cycle protein analyses with ID4 induction with tetracycline treatment (2µg/ml). **A)** NALM-6 TR cells had ID4 expression induced over 96 hours; protein lysates were prepared using Ripa buffer and approximately 50µg run on a 15% SDS PAGE gel. The blot was then probed with anti p27, p21, p15, cdk6, cyclin D3, anti Myc tag and anti β actin antibodies. **B)** RCH-ACV TR representative protein samples at 24 hours were western blotted and probed with anti p21, p27, Myc tag and β Actin sequentially.

Levels of p21 and p27 were again examined in both RCH-ACV TR 24 hours after ID4 induction (Figure 4.9 panel B). In this instance there were subtle changes the levels of p21 with ID4 expression after 24hours, but there were more substantial changes in p27. Further examination of changes in levels of p21 and 27 RNA is required to confirm this.

Overall these results show that ID4 expression in NALM-6 TR and in RCH-ACV TR is inducing G1 arrest by the up-regulation of p21 and p27. Further work would require examination cdk2 Cyclin A and Cyclin E as well as proteins involved in other stages of the cell cycle. Further controls would involve examining the effects of tetracycline treatment on the cell cycle proteins in parental or EV expressing cells

4.6 Induced ID4 expression alters apoptosis responses in B-cells.

Ectopic ID4 expression in the RCH-ACV TR and the NALM-6 TR cells reduced cell numbers which could be attributed to the transient cell cycle arrest. The effects of induced ID4 expression on apoptosis in the cell lines RCH-ACV TR and NALM-6 TR was also examined to rule out its involvement in the reduced numbers of cells where ID4 expression had been induced.

4.6.1 Induced ID4 expression causes apoptosis in NALM-6 TR and RCH-ACV TR cells.

Apoptosis (assessed by annexin PI staining) was examined in RCH-ACV TR and NALM-6 TR after the induction of ID4 (Figure 4.10). In RCH-ACV TR (panel A), there was 26% apoptosis (above background apoptosis) after ID4 is induced after 96 hours respectively. In NALM-6 TR cells (panel B) after the induction of ID4 there was approximately 26% apoptosis (above background levels) after 96 hours. There was no apoptosis in either RCH-ACV TR or NALM-6 TR 24 hours after the induction of ID4.

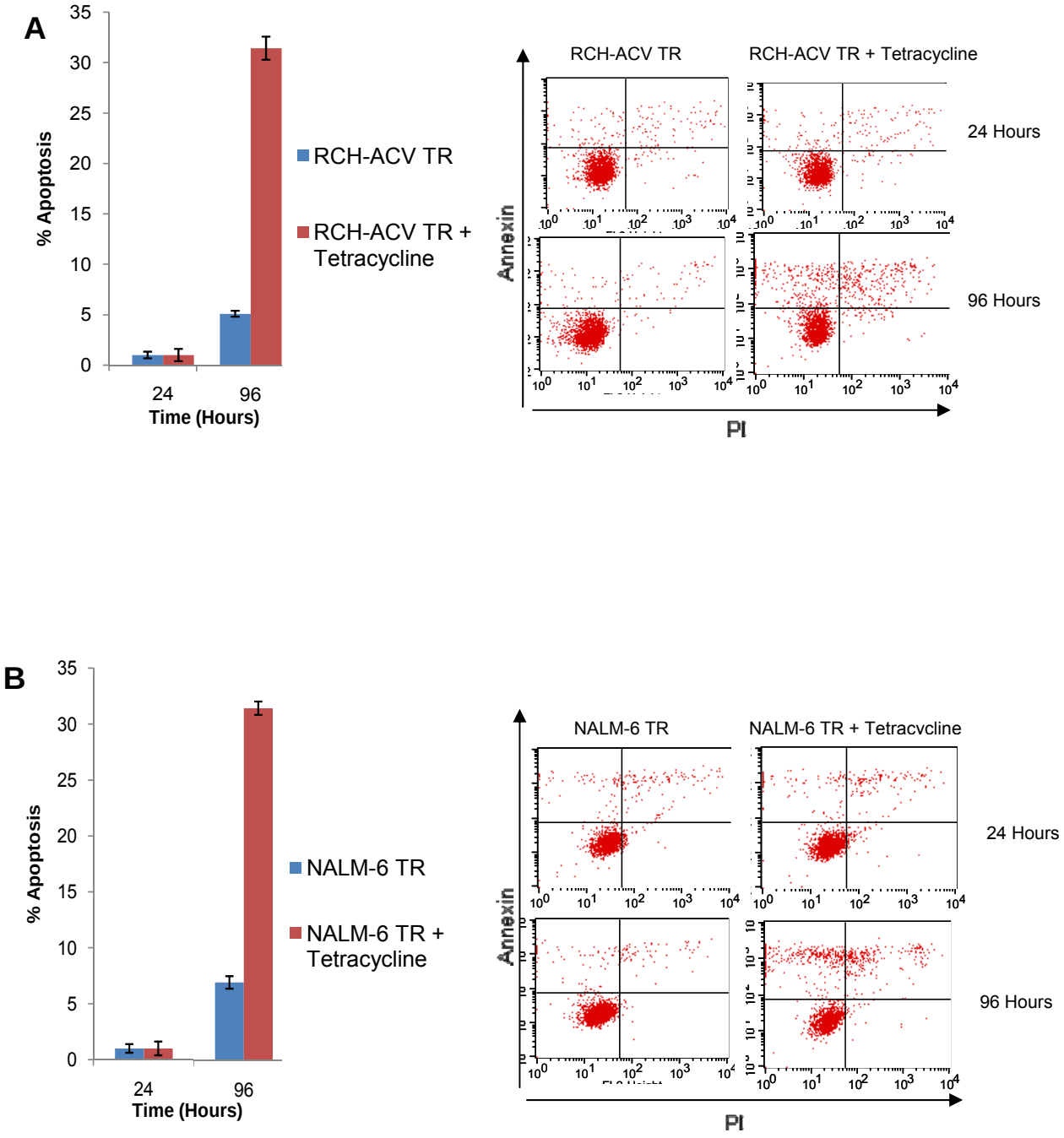


Figure 4.10 Apoptosis in the cell lines RCH-ACV TR and NALM-6 TR with or without ID4 induction after treatment with 2µg/ml tetracycline. Cells were seeded at a concentration of $1x10^4$/ml (in triplicate for each treatment). Apoptosis was examined in the cell lines **A)** RCH-ACV TR (clone 2) and **B)** NALM-6 TR with or without ID4 induction between 24 and 96 hours and their relative (annexin PI staining) flow cytometry profiles. The results shown are the mean (n=3) ± SD.

Therefore, the reduced numbers of cells observed with induced ID4 expression in RCH-ACV TR and NALM-6 TR cannot be attributed to apoptosis. There were also a lack of parental and EV expressing clones to examine the effects of long term culture with tetracycline on cell viability were required in these experiments.

4.6.2 ID4 expression does not confer resistance to apoptosis on IL3 withdrawal in Ba/F3 cells.

Ba/F3 cells are reliant on the IL3 cytokine for survival (Palacios *et al.*, 1984). The ability of ectopic ID4 to confer IL3 independence in these cells was examined. The Constitutive ID4 expression in Ba/F3 cells did not affect proliferation or induce cell cycle arrest when compared to the EV expressing clones.

The effects of IL3 withdrawal were examined in the clones expressing ID4 compared to the EV expressing clones. Apoptosis was assessed by annexin/PI analysis with IL3 withdrawal.

Constitutive ID4 expression did not have a consistent effect on cytokine withdrawal, but the two clones 8 and 16 do exhibit more apoptosis compared to clones 11 which exhibits similar rates of the EV clones. By 24 hours all the cells had died (data not shown). Therefore, there was no consistent effect observed with IL3 withdrawal of Ba/F3 clones expressing ID4 compared to those expressing EV.

4.6.3 ID4 induced G1 arrest confers resistance to apoptosis in NALM-6 TR cells.

Short term expression of ID4 protein induced G1 arrest (from which there is partial recovery) and overall reduced cell numbers. Apoptotic responses in B-cells were examined, whilst in ID4 induced G1 arrest to assess if the cells were in resistant to apoptosis when arrested.

Apoptosis was induced with reagents which do not target DNA damage responses. Staurosporine is a protein kinase inhibitor which can induce cell

cycle arrest and induce apoptosis in a caspase 3 dependent manner; however it has a broad spectrum of activity (Chae *et al.*, 2000). TRAIL is a member of the tumour necrosis family of ligands which induce apoptosis via caspase 8 (MacFarlane, 2003).

NALM-6 TR cells showed sensitivity to TRAIL ligand and Staurosporine, but not to Fas ligand induced apoptosis. RCH-ACV TR cells showed no sensitivity to TRAIL ligand induced apoptosis and therefore was not used.

The NALM-6 TR cells were treated with tetracycline for 24 hours before being treated with TRAIL or Staurosporine for 4 hours (Figure 4.11 and Appendix Figure A1.5). Staurosporine (1mM) and TRAIL ligand (500ng/ml) induced approximately 90% apoptosis in NALM-6 TR cells without ID4 induction.

The NALM-6 TR cells with ID4 induction showed approximately 75% apoptosis, this is 15% less than the cells without ID4 expression. Therefore, G1 arrest induced by ID4 expression in cells showed a reduced response to apoptotic stimuli. This experiment requires further controls, to assess the effects of tetracycline on the sensitivity of NALM-6 TR cells to apoptosis.

4.7 Conclusions.

The previous chapter showed *ID4* RNA expression in B-cell malignancies independent of the *IGH* translocation, implicating ID4 as a potential oncogene in these cells. This chapter therefore examined the transforming potential of ID4 over-expression in B-cell lines by looking at ID4 cellular localisation, the effects of ID4 expression on cell proliferation, cell cycle and apoptosis.

The cellular localisation of ID4 protein (endogenous or ectopically expressed) was predominately cytoplasmic in the human BCP-ALL cell lines. The over-expression of ID4 did not change its cellular localisation. The Ba/F3 cells showed both nuclear and cytoplasmic staining for ectopic ID4. The differences in ID4 localisation in different cell lines may be attributed to the different stages of differentiation these cells. Ba/F3 cells are of a more immature stage than NALM-6 and RCH-ACV which are both at the Pre B-cell stage.

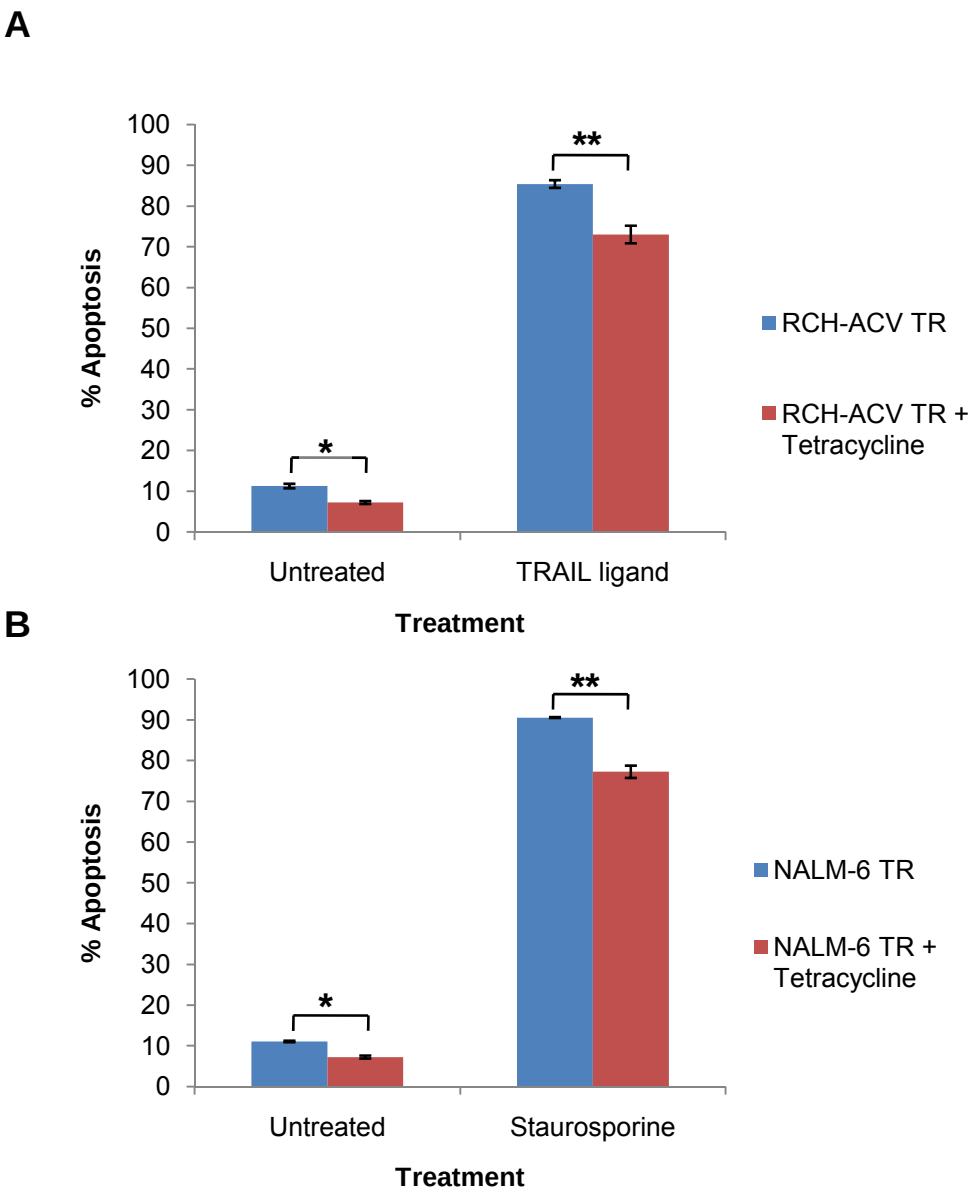


Figure 4.11 Apoptosis induction in NALM6 TR cells with and without ID4 induction (post 24 hours treatment with 2µg/,ml tetracycline). Cells were seeded at a concentration of $1x10^4$/ml (in triplicate for each treatment). **A)** Cells were treated with 500ng/ml of TRAIL ligand for 4 hours. *p=0.0005, **p=0.0008 **B)** Cells were treated with 1mM Stauorsporine for four hours.*p0.00005, **p=0.0001. The results shown are the mean (n=3) ± SD. P values were calculated using an unpaired t test. This experiment has been repeated three times with consistent results.

Ba/F3 cells are of mouse origin, therefore expressing human ID4 protein in these cells may not exhibit the same response as would be seen with expressing mouse ID4 protein.

These experiments were limited by the lack of fractionations for ectopically expressed ID4 and immunofluorescent staining for endogenous ID4 because of the non-specific binding of the ID4 antibody.

The myc-His tagged ID4 protein ran at approximately 22kDa on western blotting. The native ID4 protein has a molecular weight of 17kDa and the myc His tag adds approximately 2kDa. Therefore the extra 2 kDa observed could be due to posttranslational modification; this is supported by the fractionation data for NALM-6 (Figure 4.1) where the cytoplasmic fraction of ID4 protein ran as a doublet. This would require further examination.

The induced over-expression of ID4 reduced cell numbers, which appears to be a consequence of the transient cell cycle arrest. There was some apoptosis observed at 96 hours, but there were no controls to assess the effects of tetracycline alone on these cells. Also, no sub G1 peak was observed when doing the cell cycle analysis (another measure for apoptosis), but this may be the result of the sample processing. Further work would require also the use of the proliferation marker ki67 as a more sensitive means to measure the effects of ID4 expression on cellular proliferation.

The RCH-ACV TR clone1 was leaky for ID4 expression and when treated with tetracycline, the cells did express more ID4 protein. The effects of ID4 induction in this clone was not as pronounced as with clones 2 and 3, which were not leaky for ID4 expression. Also, the cells only showed a transient G1 arrest with induced ID4 expression. This all implies that the effects of ID4 are ‘tolerated’ over time and this is possibly why no effects of constitutive ID4 expression were observed in the Ba/F3 cells.

The sensitivity of ID4 over-expressing cells to apoptotic stimuli was also assessed and ID4 induced G1 arrest provided a significant resistance to

apoptotic stimuli. Further work would require examining the cells' responses after 24 hours ID4 induction, where their cell cycle partially recovers.

Due to time constraints, these experiments were sometimes limited by the lack of relevant controls to examine the effects of tetracycline on cellular proliferation, cell cycle and responses to apoptosis. This should be addressed in subsequent work to clarify the role of ID4 in the responses identified in this chapter.

Chapter 5: Effects of *ID4* expression in primary B-cell progenitors.

5.1 Introduction.

Examining cellular responses to *ID4* over-expression in malignant B-cell lines is hindered by the possibility of co-existing mutations masking these effects. The next approach was to examine the effects of expressing human *ID4* in B-cell progenitors obtained from the livers of E13.5 embryos of normal mice, mice with the *cdkn2a* locus deleted (*cdkn2a-/-)* or expressing the *Eμ-myc* transgene.

The *CDKN2A* locus encodes for the cdk inhibitors p16 and p14 ARF (p19 ARF in mouse). Both proteins are transcribed by a process of alternative splicing from 2 exons (see Figure 5.1). This locus is frequently targeted in many cancers such as BCP-ALL (Mullighan *et al.*, 2007).

Id4 expression was shown to drive proliferation in *cdkn2a-/-* mouse astrocytes (lacking p16 and p19 ARF expression) (Jeon *et al.*, 2008). Furthermore, the t(6;14)(p22;q32) translocation BCP-ALL cases also presented with deletions of the *CDKN2A* locus (Russell *et al.*, 2008). This defines *ID4* as a dominant oncogene in this background.

To examine this further, c*dkn2a -/-* E13.5 foetal livers were obtained from *cdkn2a -/-* mice (Serrano *et al.*, 1996) and infected with the *ID4 Mig*R1 retrovirus (see sections 2.2.3 and 2.2.18 and Figure 2.2). In this model, exons 1α and 2 of the *cdkn2a* locus have been deleted and as a consequence no p16 is expressed and a truncated p19ARF protein is expressed in these mice (Serrano *et al.*, 1996).

MYC, when over expressed is a potent oncogene which has been implicated in B-cell lymphomas (Magrath, 1990). Transgenic mice expressing the oncogene *c-myc* under the control of the *IGH* enhancer element (*Eμ-myc)* limit *c-myc* expression to B lymphocytes, thereby allowing the effects of *c-myc* over expression to be examined in this context.

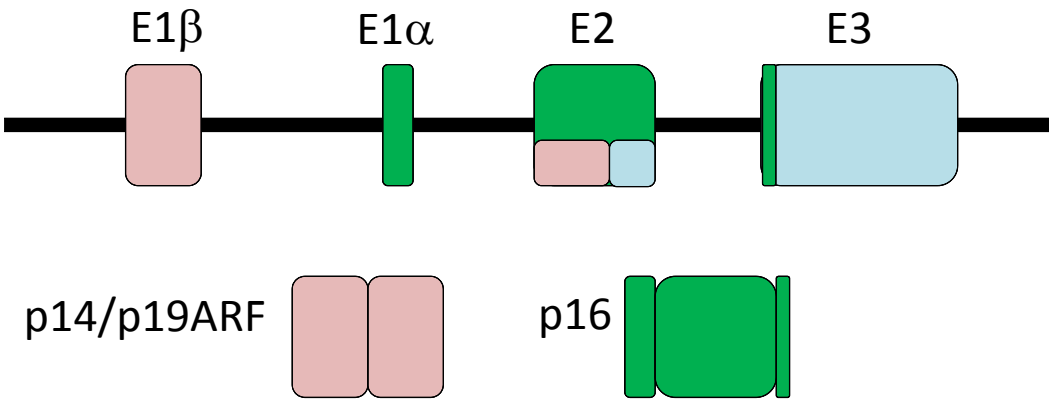


Figure 5.1 The *CDKN2A* locus. The *CDKN2A* locus is located on chromosome 9 in humans and chromosome 5 in mice. This gene has 2 exons which, by a process of alternative splicing can produce 2 proteins, p16 and p14 ARF(p19 ARF in mice). These proteins regulate key processes, such as senescence and apoptosis by inhibiting cdk proteins (p16) and stabilising p53. Adapted from Gil and Peters, 2006 .

These mice were generated by Adams *et al.* and have been shown to generate B-cell lymphomas a few months from birth (Adams *et al.*, 1985). *ID4* was therefore over expressed in this system to examine cooperation with or inhibition of *myc* induced lymphogenesis.

The *Eμ-myc* transgenic mice produced a mixture of embryos which positive and negative for the transgene in one litter. The embryos which were negative for the transgene were used as wild type controls in these experiments. These were identified using PCR (courtesy of Dr A Michie and Dr K Blythe-see below).

All embryos were provided by and PCRs performed in the laboratories of Dr A Michie and Dr K Blythe (University of Glasgow and the Beatson Institute). The retroviral infections and analysis was undertaken by myself in the laboratory of Dr A Michie.

E13.5 foetal liver cells were infected with *ID4 Mig*R1 or empty vector (*EV) Mig*R1 retrovirus. This virus expresses GFP under the control of a separate IRES enhancer. The presence of a GFP signal was used as confirmation of expression of either *ID4* or *EV* in the infected cells. The expression of *ID4* using this retrovirus was verified by Dr T Akasaka (Professor Dyer laboratory).

Cells were infected, sorted and grown on an OP9 stromal layer with the appropriate B-cell cytokines. This process is summarised in Figure 5.2. The percentage of GFP positive cells and stages of B-cell differentiation were examined over 14 days (see Materials and Methods section 2.2.17).

5.2 *ID4* over-expression in B-cell progenitors does not induce proliferation.

Figure 5.3 shows the percentage of GFP positive cells from the *cdkn2a-/-* and *Eμ-myc* foetal livers. *ID4* expression did not affect proliferation as assessed by the percentage of GFP positive cells in the *cdkn2a -/-* foetal livers (panel A). This experiment is not statistically significant and requires repeats before a conclusion can be made about the effects of *ID4* over expression in *cdkn2a -/-* foetal livers.

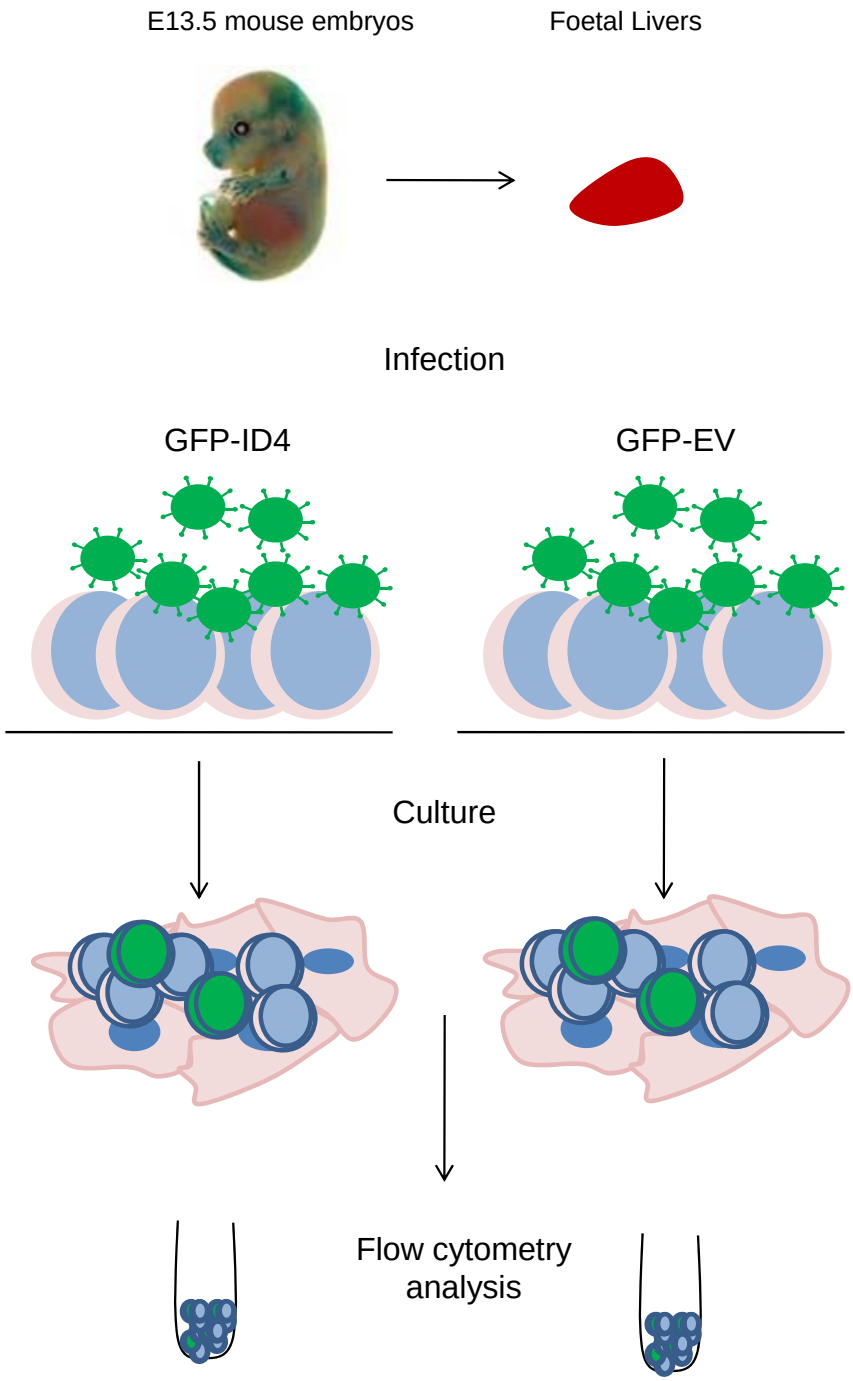


Figure 5.2 Summary of the foetal liver assay. Foetal livers were obtained from E13.5 *cdkn2a -/-* or *Eμ-myc* mice. Cells from each liver were then infected with either *ID4* or *EV* retroviruses after which they were sorted (see Materials and Methods section 2.2.18.2) and cultured on OP9 stromal cells with the relevant cytokines. At Days 2, 5, 8 and 14, the cells were collected for flow cytometry analysis to determine GFP positive populations and their immunophenotype.

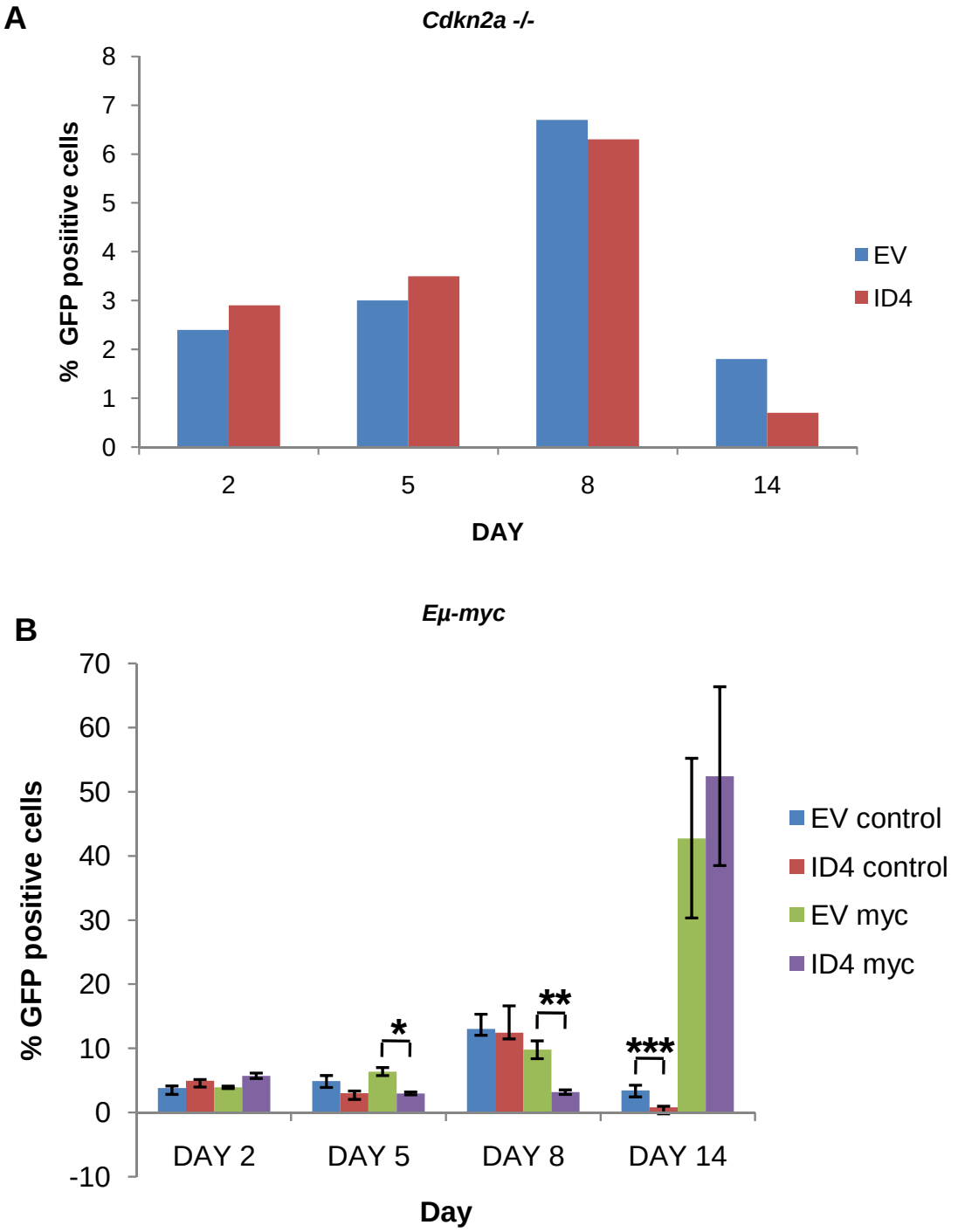


Figure 5.3 B-cell progenitor proliferation with *EV* or *ID4* over-expression over 14 days**. A)** *cdkn2a-/-* foetal liver cells (n=1). Several foetal livers were prepared and pooled before infection. These cells were then spilt to be infected with either *ID4 Mig*R1or *EV Mig*R1retrovirus. This graph shows the result of one experiment. **B)** *Eμ-myc* foetal liver cells. Each foetal liver was infected individually (with either *ID4 Mig*R1 or *EV Mig*R1*)* and treated as one experiment. The graph shows the mean of control foetal liver cells (n=6 foetal livers) and *Eμ-myc* foetal liver cells (n=5 foetal livers) ± standard error of the mean (SEM). P values were calculated using an un-paired t test. *p=0.0022, ** p=0.0037 and ***p=0.0208.

Figure 5.3 panel B shows the percentage of GFP positive cells from the *Eμ-myc* foetal livers. The effects of *c-myc* on proliferation were visible at day 14, where there were approximately 40% more GFP positive cells compared to the *EV* control. Any effects of *ID4* expression in these *Eμ-myc* B-cell progenitors was only evident at day 8, at which time there were fewer cells, but the numbers recovered at day 14.

In the control cells (negative for the *μu-myc* transgene), the effects of *ID4* expression was only evident at day 14, where there were significantly fewer cells compared to the *EV*.

In the *EV Eμ-myc* cells, at day 14, there were approximately 30% more B-cells when compared to day 8. In the *ID4 Eμ-myc* cells, there were approximately 50% more cells. Therefore *ID4* expression in *Eμ-myc* cells promotes a proliferative advantage larger than that seen with the EV expressing cells between days 8 and 14.

Therefore, the expression of *ID4* in B-cell progenitors did not induce proliferation. *ID4* expression did not cooperate with the *cdkn2a-/-* phenotype to induce proliferation, but did cooperate with *Eμ-myc* between days 8 and 14 to drive proliferation of these cells. Further work in the *cdkn2a-/-* foetal livers is required to confirm these data in order to perform statistical analysis. A further control would be to confirm the expression of ID4 protein in the GFP positive cells.

5.3 *ID4* expression promotes pre B-cell development, but this is reduced as the B-cells mature.

The cells were also immunophenotyped to examine if *ID4* expression was affecting B-cell differentiation using FACS analysis. Progenitor populations were gated as GFP^+, $CD11B^-$, $NK1.1^-$, $CD117^+$, $CD19^-$ since this would include the CLP and the pre-pro B-cells. Further populations and the overall CD19+ population were analysed (as described in Figure 5.4). All the B-cell populations were assessed within the GFP positive population only.

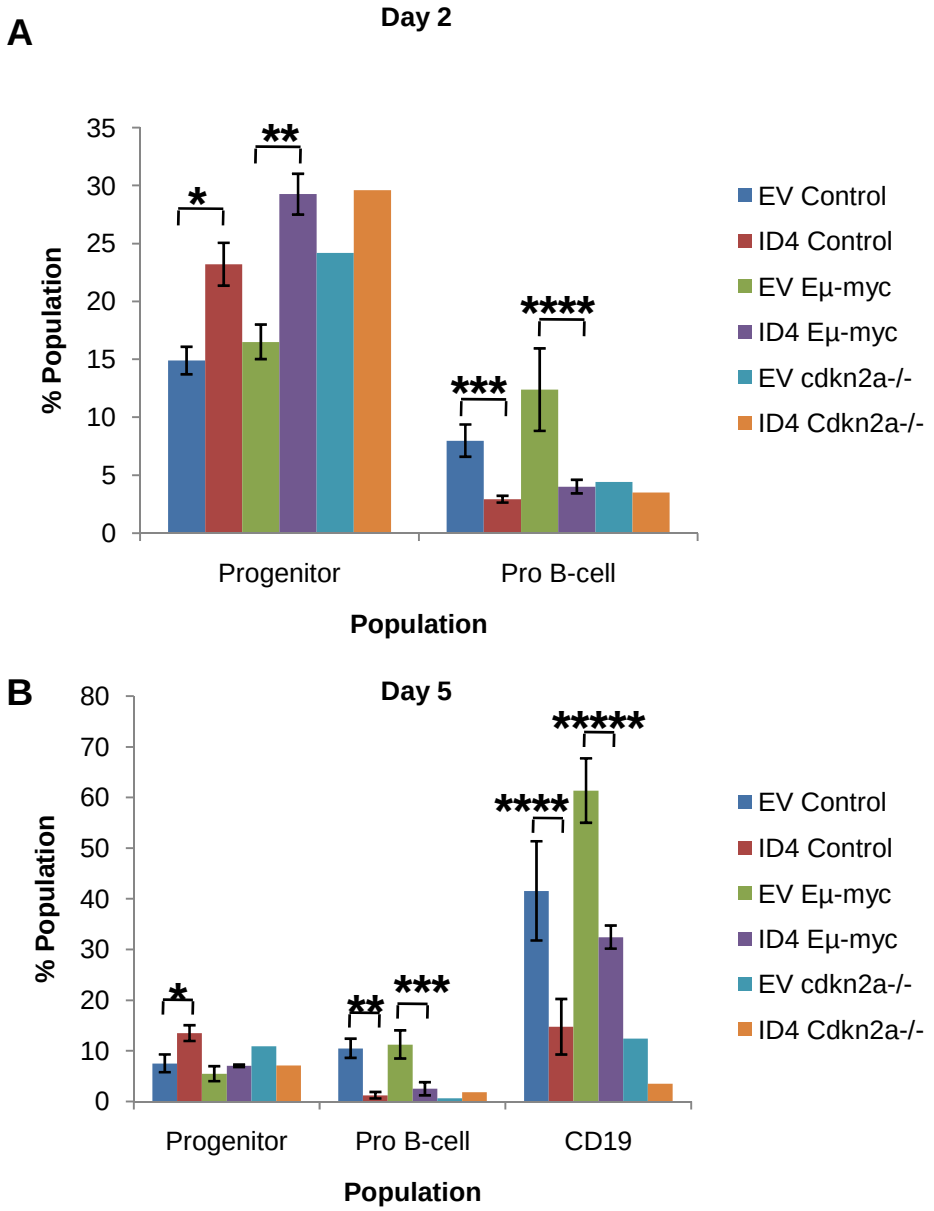
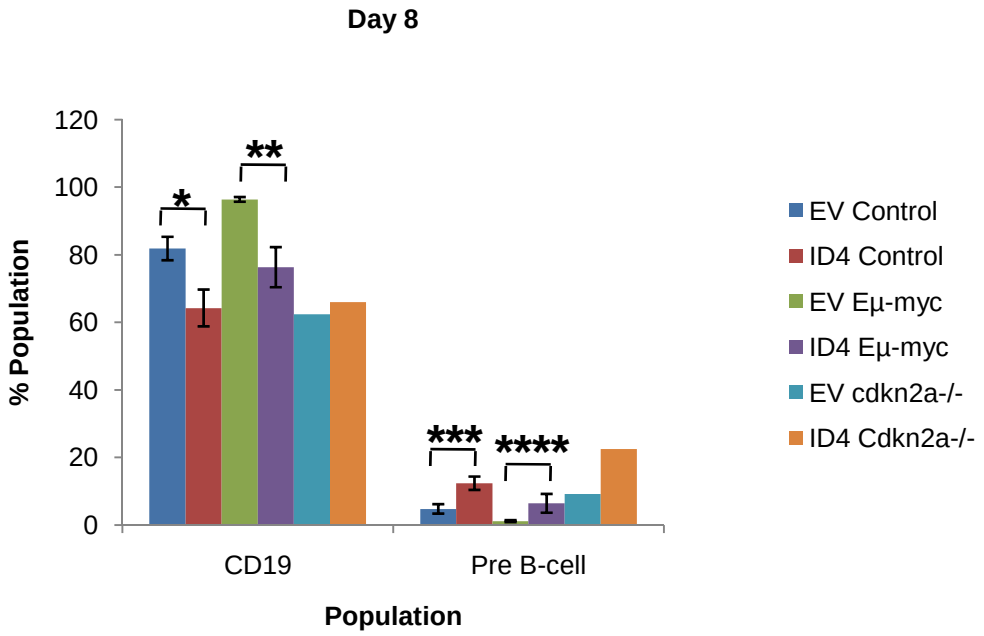
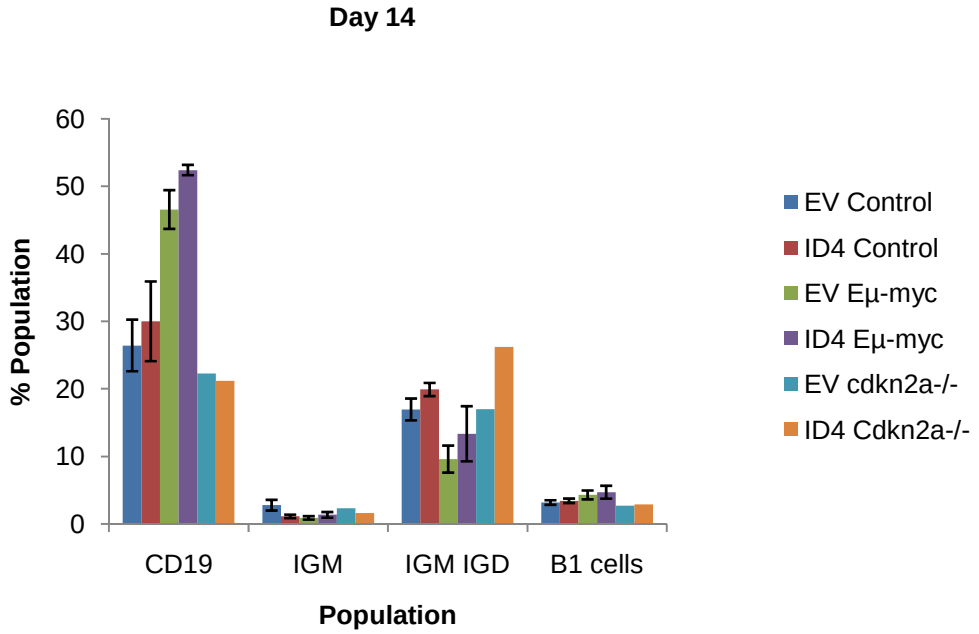


Figure 5.4 B-cell development over 14 days with *ID4* or *EV* expression in control cells, *Eμ-myc* cells and *cdkn2a-/-* cells. As with Figure 5.3, *cdkn2a-/-* foetal livers were pooled and spilt to be infected with either *ID4 Mig*R1 or *EV Mig*R1. The embryos from the *Eμ-myc* transgenic mouse, each foetal liver was infected (with either *ID4 Mig*R1 or *EV Mig*R1*)* and treated as one experiment. Immunophenotyping was set up with gating for progenitors (GFP^+, $CD11B^-$, $NK1.1^-$, $CD117^+$, $CD19^-$), Pro B-cells (GFP^+, $CD11B^-$, $NK1.1^-$, $CD117^+$, $CD19^+$), overall $CD19^+$ (GFP^+ $CD19^+$), Pre B-cells (GFP^+,$CD19^+$, IGM^-, $CD25^+$), IGM (GFP^+,$CD19^+$, IGM^+, IGD^-) IGM^+ IGD^+ (GFP^+,$CD19^+$,IGM^+ IGD^+), and B1 cells (GFP^+,$CD19^+$, IGM^+ $CD5^+$) populations. The graphs show the mean of the control foetal livers (n=6 foetal livers) and *Eμ-myc* expressing foetal livers (n=5 foetal livers) ± SEM and one set of the *cdkn2a-/-* cells. **A)** Day 2 *P=0.0026, **P=0.0015, ***P=0.004, ****P=0.0594. **B)** Day5. *P=0.0344, **P=0.0038, ***P=0.0206, ****P=0.0405, *****P=0.0137 **C)** Day 8. *P=0.0215, **P<0.0001, ***P=0.0107, ****P=0.0648 **D)** Day 14. P values were calculated using an un-paired t test.

C


Day 8
120
100
80
60
40
20
0
% Population
*
**
***
****
CD19
Pre B-cell
Population
EV Control
ID4 Control
EV Eμ-myc
ID4 Eμ-myc
EV cdkn2a-/-
ID4 Cdkn2a-/-

D


Day 14
60
50
40
30
20
10
0
% Population
CD19
IGM
IGM IGD
B1 cells
Population
EV Control
ID4 Control
EV Eμ-myc
ID4 Eμ-myc
EV cdkn2a-/-
ID4 Cdkn2a-/-

The day 2 results show there were more progenitors in all the populations expressing *ID4* (panel A). The Pro B-cell populations (defined as GFP^+, $CD11B^+$, $NK1.1^-$, $CD117^+$, $CD19^+$) were reduced where *ID4* was expressed in the control cells; the differences in the *Eμ-myc* populations were not significant and there was no difference seen in the *cdkn2a-/-* foetal livers.

At day 5 (panel B), in the control cells, the *ID4* expressing cells had more progenitors than the EV expressing cells; this effect is not visible in the *Eμ-myc* and the *cdkn2a-/-* populations. The Pro B-cells were reduced where *ID4* was expressed in the control and the *Eμ-myc* populations; the *cdkn2a-/-* cells overall had a small Pro population. The overall CD19 population is reduced by 50% across all the foetal livers expressing *ID4*.

At day 8 (panel C), the overall CD19 populations expressing *ID4* were reduced by 20% compared to the control and the *Eμ-myc* foetal livers, there was no difference with the *cdkn2a-/-* foetal livers. The control cells expressing *ID4* had a significantly increased population of Pre B-cells. This was also observed with the *cdkn2a-/-* and in the *Eμ-myc* foetal livers, but, in the *Eμ-myc* cells these results were not significant.

The day 14 results (panel D) showed the overall CD19 population in the *myc* expressing cells (*EV* and *ID4* infected) had a 45% CD19 population compared to the corresponding *EV* infected *myc* expressing cells, the expression of *ID4* at this stage had no effect. The other populations were all comparable irrespective of background or ID4 expression.

Overall, these results have shown that ectopic *ID4* expression impaired B-cell development. This was evidenced by the increased progenitor populations and reduced pro B-cells in the first week and the reduced CD19 levels across all the populations. This effect recovered by day 14.

5.4 Conclusions.

This chapter aimed to examine any cooperation of *ID4* with an oncogene, c-*myc* and the loss of a tumour suppressor, *cdkn2a* -/- in B-cell progenitors.

ID4 expression gave no proliferative advantage in normal B-cell progenitors. Further work would require proliferation to be examined in more detail with markers such as ki67. *ID4* expression did not cooperate with the *cdkn2a-/-* genotype, but further work would require more repeats to establish statistical significance. *ID4* expression, did not overcome the proliferative effects of *myc* in this system, but between days 8 and 14 provided a larger proliferative advantage than in the *EV* cells. This implies that *ID4* can cooperate with the *myc* oncogene in B-cells. GFP positive cells were assumed to be expressing *ID4*. A further control of a Western or a PCR was required to confirm this since GFP was expressed on an IRES.

The immunophenotyping showed that *ID4* inhibited the differentiation of B-cells, irrespective of the genetic background. At day 2, *ID4* expression was impairing development from the progenitors to the pre pro B-cells. The most profound effects included the reduced CD19 population at day 8 and the increased Pre B-cell population.

This implies that *ID4* may be antagonising bHLH proteins which reduces the CD19 expression; however the accumulation of the Pre B-cells implicated *ID4* at this stage of B-cell development, because by day 14 the cells recovered. These findings imply that at the Pre B-cell stage *ID4* may have the potential with the correct cooperating mutation to generate leukaemia.

Chapter 6: Identification of proteins which interact with ID4.

6.1 Introduction.

In the previous chapters, ID4 expression in primary B-cell progenitors was shown to disrupt normal B-cell development *in* vitro, this represents one hallmark of transformation (Hanahan and Weinberg, 2000). However the gene expression analysis of primary BCP-ALL cases expressing *ID4* RNA (using Oncomine) showed no changes in genes associated with B-cell development. This suggests that *ID4* expression exerts a context dependent effect at different stages of B-cell development.

In B-cells, both ID1 and ID2 have been shown to bind E2A (comprising of the proteins E47 and E12) and in doing so disrupt B-cell development (Cochrane *et al.*, 2009; Ji *et al.*, 2008; Mathas *et al.*, 2006). However, ID2 interactions with the APC/C complex in neurones disrupts normal differentiation and proliferation (Lasorella *et al.*, 2006). This suggests that ID4 may function predominantly by a post-transcriptional mode of action. Therefore, the aim of this chapter was to identify proteins which interact with ID4 in B-cells.

The experimental approach adopted for the studies involved identifying interacting proteins by immunoprecipitation (IP). Confirmation of these interactions was done (where possible) by reverse IP's. Further to this, IPs were analysed using shot gun mass spectrometry to confirm the interactions and identify new interacting proteins. The processing of the protein gels for mass spectrometry and Mascot data generation (see Figure 2.3) was performed by Mrs Rebecca Jukes-Jones and Dr Robert Boyd (MRC Toxicology Unit, Leicester).

One limitation with these experiments was the anti ID4 antibody. Previous chapters have shown that the anti ID4 antibody had limited sensitivity and gives a lot of background with western blotting and therefore hindered progress with these studies.

The IPs in these experiments was performed using optimised techniques outlined in the Materials and Methods chapter (section 2.2.8). IPs were performed using a whole cell lysis (Ripa buffer) and a cytosolic preparation

(without detergent). Therefore detergent based and non-detergent based methods were used.

Cell lines used in these experiments included NALM-6 cells and ectopic ID4 expressing cell lines, Ba/F3 clone ID4 11 and NALM-6 TR cells (outlined in chapter 4, section 4.1). The Ba/F3 cells were used to IP from because of their stable constitutive expression of myc-His tagged ID4 protein (see Figure 4.3). The NALM-6 TR cells provided a system where myc-His tagged ID4 protein expression could be induced.

6.2 The interaction of ID4 and E2A.

ID4 expression in primary B-cell progenitors disrupted normal B-cell development, possibly by ID4 interacting with and antagonising the transcription factor E2A.Therefore the interactions between ID4 and E2A were examined in NALM-6 cells, the Ba/F3 clone ID4 11 and NALM-6 TR cells.

6.2.1 IP of endogenous ID4.

An attempt was made to IP endogenous ID4 using the cell line NALM-6 (Figure 6.1). An IP performed from the Ba/F3 clone ID4 11 using the anti ID4 antibody showed that very little of the ID4 (Myc tagged) immunoprecipitated (panel A).

An IP was attempted in the NALM-6 cells (panel B), but unsurprisingly no endogenous ID4 was immunoprecipitated (IP lane).

This was also attempted using a whole cell preparation of both NALM-6 and MHH-PREB-1, but proved unsuccessful (data not shown). The difficulty of immunoprecipitating endogenous ID4 meant that the IPs had to be performed in the ectopic ID4 expressing systems, using an anti Myc tag antibody.

6.2.2 Ectopic ID4 interacts with E47 in B-cell lines.

The IPs were performed using both cytosolic and whole cell lysates. With each condition, the IPs were performed using the anti Myc tag antibody and the anti E47 antibody (Figure 6.2). The E47 protein is approximately 72 kDa. The blots show a lower band than this which may be a cross reacting product.

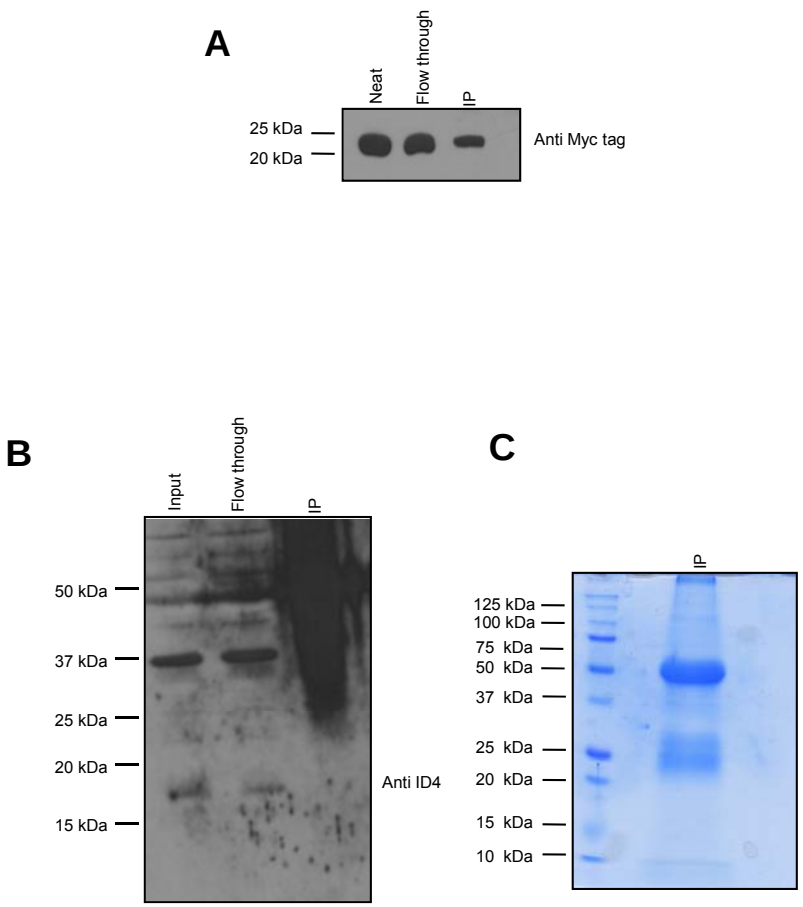


Figure 6.1 IP of endogenous ID4 with anti ID4 antibody. **A)** 708µg of cytosolic lysate from Ba/F3 clone ID4 11 was immunoprecipitated for 2 hours at 4°C, with the ID4 antibody (40 µl of beads). 5µl of each reaction was run on a 12% SDS PAGE gel and probed with the anti Myc tag antibody. **B)** NALM-6 IP using anti ID4 antibody. Cytosol was extracted from $1.4x10^9$ cells. The IP was performed using 3000µg of protein with 300µl of beads for two hours. Approximately 45µg of each fraction were run on a 12% SDS PAGE gel, which was western blotted and probed with anti ID4 antibody. **C)** As with B, the remaining captured fraction was purified using the UPPA kit (see Materials and Methods, section 2.2.9.4) and run on a 12% SDS PAGE gel and coomassie stained.

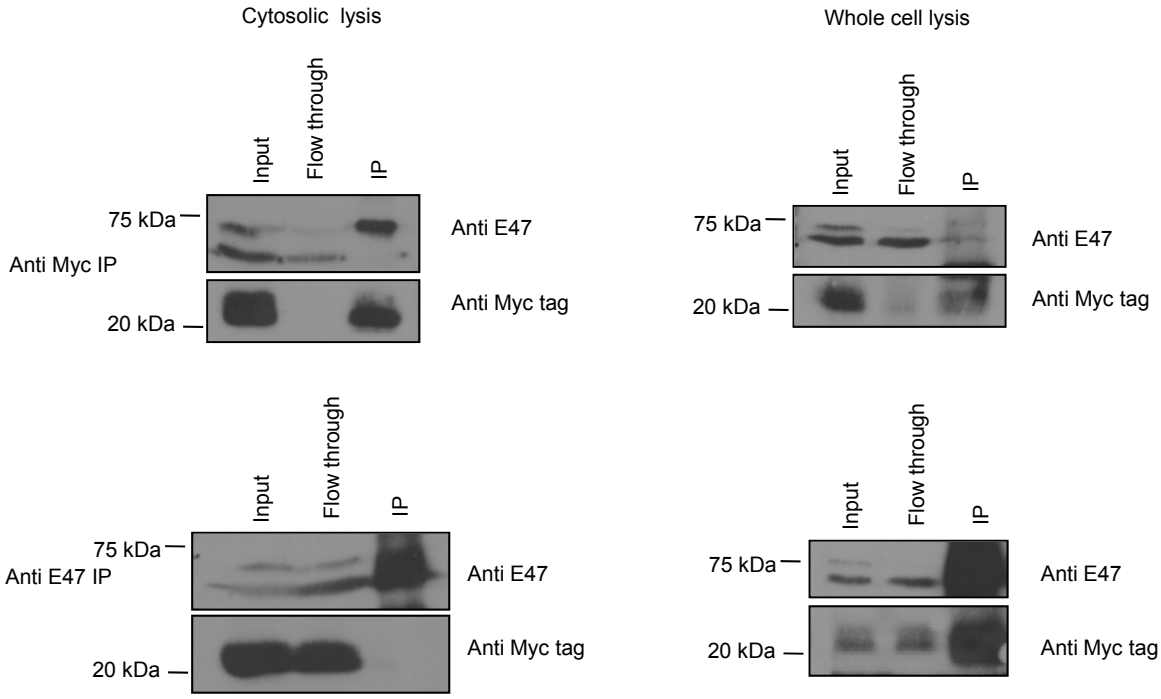


Figure 6.2 Interaction between ID4 and E47 in Ba/F3 cells. IP of both cytosolic and whole cell preparations using the anti Myc tag antibody (20µl of beads) and the E47 antibody (40µl of the beads). For the cytosolic lysis preparation, 120µg of lysate was used and for the whole cell preparation, 84µg of lysate was used. The samples were run on a 12% SDS PAGE gel and western blotted. The blot was then probed with anti E47 and anti Myc tag antibodies sequentially.

In both the whole cell and the cytosolic preparations, the anti Myc tag IP showed an interaction between ID4 and E47.

The anti E47 IP was not successful (in either lysis condition), firstly because of the lack of E47 that was immunoprecipitated and secondly because of the resulting heavy smearing on the western blotting making subtle bands more difficult to detect.

In the Ba/F3 cells constitutively expressing ectopic ID4, ID4 interacted with endogenous E47 as demonstrated by the anti Myc tag IPs. An IP was attempted with the E47 antibody in both NALM-6 and MHH-PREB-1, but no interaction between E47 and ID4 was seen (data not shown).

The interaction between ectopic ID4 expression and endogenous E47 was then examined in the human cell line NALM-6 TR using IP. ID4 expression was induced in these cell lines and after 4 hours (when there is maximal expression) and IPs performed using the anti Myc tag antibody (Figure 6.3).

Figure 6.3 shows that ectopically expressed ID4 interacted with E47 in NALM-6 TR cells, but not all of the E47 had immunoprecipitated with ID4. The gels without induced ID4 expression (no treatment) showed a weak signal for E47 in both preparations in the IP lanes, this may be a result of a non-specific binding but the IPs performed from the cells with induced ID4 expression showed a much stronger interaction. The ectopic ID4 protein interacted with endogenous E47 protein.

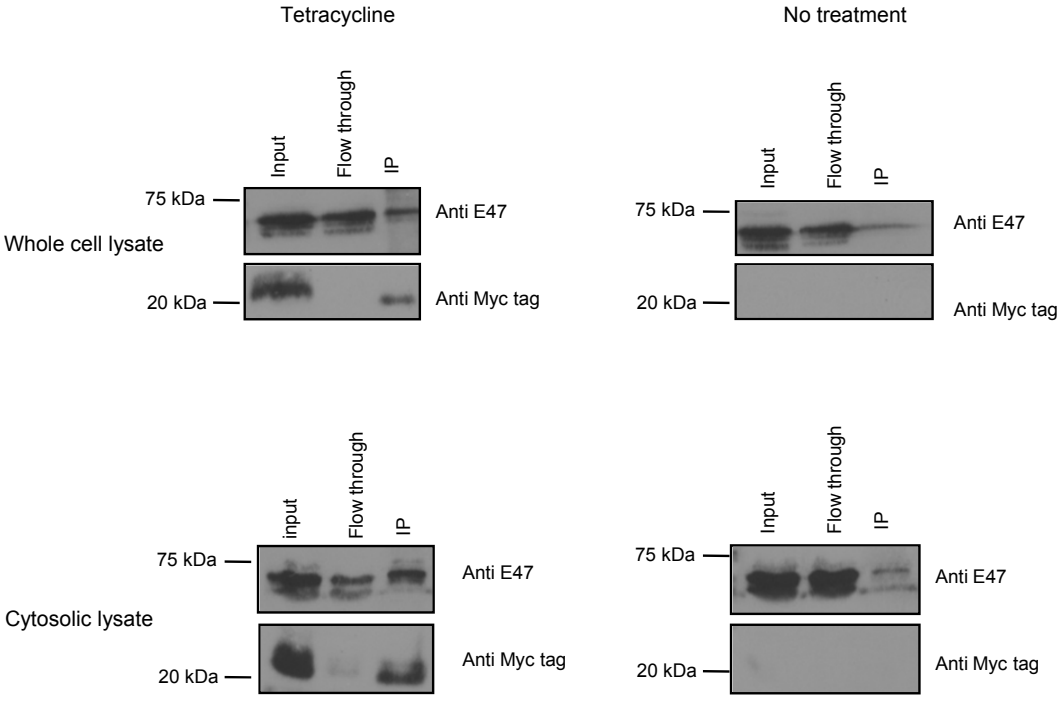


Figure 6.3 ID4 interacts with E47 in NALM-6 TR cells. IP of ID4 using the anti Myc tag antibody from NALM-6 TR cells with and without treatment with 2μg/ml tetracycline for 4 hours. Whole cell lysates were produced from $1x10^8$ cells and the IP performed using 500μg of protein and 40μl of beads. The cytosol preparations were produced from $1.8x10^8$ cells and the IP performed with 670μg of protein. 20 μl of immunoprecipitated fraction and the other relevant fractions were run on a 12% SDS PAGE gel and western blotted. The blots were then probed with anti E47 and anti Myc tag antibodies sequentially.

6.3 ID4 interacts with other proteins.

Ectopically expressed ID4 and not endogenously expressed ID4 interacted with endogenous E2A in cell lines. Proteins other than E2A which may interact with the ID proteins include the ID1 and ETS interaction (Ohtani *et al.*, 2001), ID2 and Rb (Iavarone *et al.*, 1994; Lasorella *et al.*, 2000; Lasorella *et al.*, 2005), ID2 and PU.1 (Ji *et al.*, 2008), the ID proteins interactions with PAX5 (Gonda *et al.*, 2003; Roberts *et al.*, 2001) and the interaction of ID2 with components of the APC/C complex (Lasorella *et al.*, 2006).

This raised the question of which other proteins ID4 may interact with in B-cells. To address this, IPs for Myc tagged ID4 were performed in the cell lines Ba/F3 ID4 clone 11 and NALM-6 TR and mass spectral analysis performed on the immunoprecipitated proteins (detected on coomasie/ silver stained gels).

The approach to mass spectrometry involved running the immunoprecipitated protein on SDS PAGE gels which were coomassie or silver stained. Visible bands were processed for mass spectrometry.

Analysis of the mass spectrometry data was performed using the Scaffold 2 Proteome software program. Using this program the analysis criterion for detected proteins was set by a number of criteria including the unique number of peptides recognised (see Materials and Methods, section 2.2.11).

6.3.1 ID4 interacting proteins in Ba/F3 ID4 clone 11 cells.

An IP was performed using the anti Myc tag antibody and cytosolic fractions from Ba/F3 clones ID4 11 and Ba/F3 EV14. A cytosolic lysis was performed because it reduced the non-specific proteins that immunoprecipitated with Myc tagged ID4 in these cells. The Ba/F3 EV clone was used as a control. The immunoprecipitated fractions were run on an SDS PAGE and the gel was initially coomasie stained, but no band corresponding to ID4 could be seen. This gel was then silver stained (Figure 6.4).

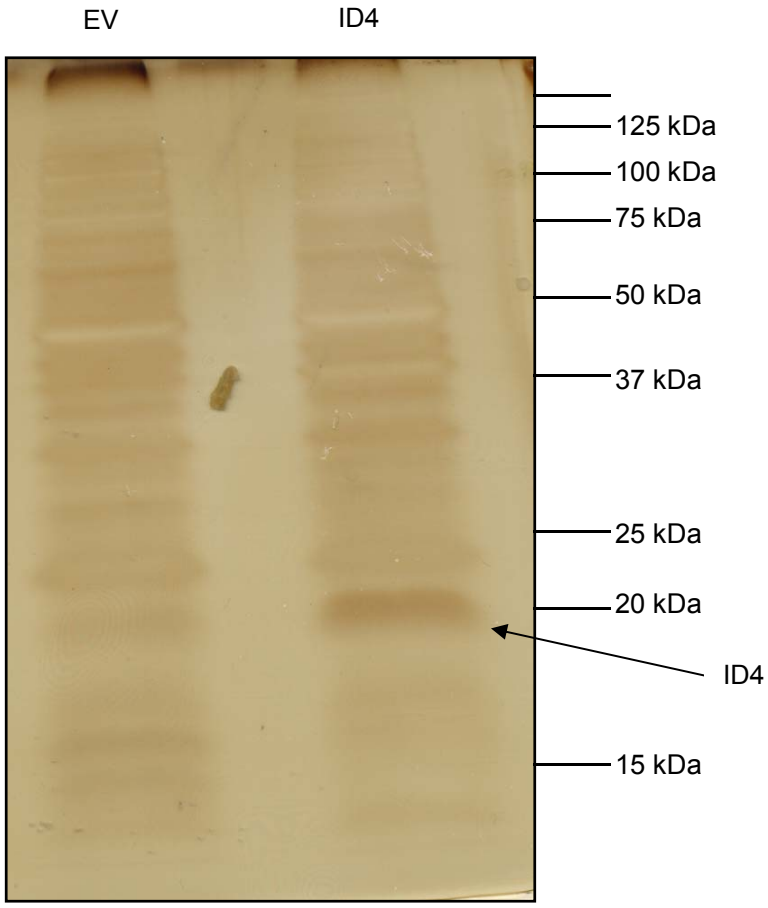


Figure 6.4 IP using Ba/F3 clones ID4 11 and EV14. $5.3x10^8$ cells from clone ID4 11 and $3.4 x10^8$ cells from clone EV14 were used to prepare cytosolic fractions. 700µg of concentrated cytosol was then incubated with 40µl of beads and anti Myc tag antibody for 2 hours at 4°C. The IPs were then run on a 12% SDS PAGE gel, which was silver stained.

There was a strong band visible between 20 and 25 kDa in the 'ID4' lane. This band was excised and processed for mass spectrometry, which confirmed that it was ID4 (mass spectra data are found in Appendix Figure A1.6 and Table A1.7). The gel also showed numerous bands present in both lanes. In the 'ID4' lane alone there was one outstanding band that ran at just under 50 kDa.

The remainder of the silver stained gel was then processed for mass spectrometry (for both the ID4 and EV tracks) and peptide lists generated for each track. Any peptides recognised from the EV track were removed from the ID4 peptides list, therefore removing any non-specific proteins (Appendix Table A1.8).

The list in Appendix Table A1.8 was re-analysed by increasing the criteria stringency settings. These were the maximum criteria at which ID4 protein could be detected. These peptides were then analysed (those protein which gave more than 1 peptide hit had an adequate mascot score and a clean spectra) to generate the list of proteins in Table 6.1.

The protein TCF12 was present on the list in Appendix Table A1.8. This is a bHLH protein belonging to the same family as E47. The peptides in Appendix Table A1.8 were then re-analysed setting the criteria thresholds to detect TCF12 (those protein which gave more than 1 peptide hit had an adequate mascot score and a clean spectra) to generate the list of proteins in Table 6.2 below. This data was analysed with lower thresholds and with more emphasis on the mascot scores and spectral layout.

From the proteins co-immunoprecipitated with ID4 in the Ba/F3 ID4 clone (Tables 6.1 and 6.2), TCF12 was examined further.

TCF12 is a bHLH protein, which like E2A is involved in haematopoietic development. TCF12 is expressed in developing T lymphocytes and has been shown to be important for the development of Pro B-cells (Zhuang *et al.*, 1996).The mass spectral data for this protein is in Appendix Figure A1.7 and Table A1.9.

Table 6.1 List of proteins with 'strong' interactions with ID4 based on spectral analysis and functional significance. Analysis was done using peptide thresholds at 90% minimum, protein thresholds at 90% minimum and 1 peptide minimum.

Identified Proteins	Molecular Weight	Unique peptides
DNA-binding protein inhibitor ID-4 (Inhibitor of DNA binding 4) - Mus musculus (Mouse)	17 kDa	2
Septin-11 - Mus musculus (Mouse)	50 kDa	2
Proteasome activator complex subunit 2 - Mus musculus (Mouse)	27 kDa	2
Ras suppressor protein 1 (Rsu-1) - Mus musculus (Mouse)	32 kDa	2
GTPase-activating protein ZNF289 - Mus musculus (Mouse)	57 kDa	2

Table 6.2 List of proteins with 'weak' interactions with ID4 based on spectral analysis and functional significance. Analysis was done using peptide thresholds at 20% minimum, protein thresholds at 50% minimum and 1 peptide minimum.

Identity	Molecular weight	Unique Peptides
Septin-2 - Mus musculus (Mouse)	42 kDa	4
COP9 signalosome complex subunit 7a - Mus musculus (Mouse)	30 kDa	1
Glucose-6-phosphate isomerase (EC 5.3.1.9) - Mus musculus (Mouse)	63 kDa	1
Calponin-2 Mus musculus (Mouse)	33 kDa	1
ADP/ATP translocase 1 - Mus musculus (Mouse)	33 kDa	2
Transcription factor 12 - Mus musculus (Mouse)	76 kDa	1
G1 to S phase transition protein 1 homolog - Mus musculus (Mouse)	56 kDa	1

IPs were then attempted to confirm the interaction between ID4 and TCF12 (Figure 6.5). TCF12 has 5 isoforms, 3 of which give 3 different molecular weight proteins of 56kDa, 75 kDa and 78 kDa. All 3 isoforms were seen on the western blots in panel A, but only the highest molecular weight proteins interacted with ID4 using both lysis methods. The flow through fraction in the cytosol lysis showed that not all the TCF12 interacted with ID4.

Based on these results, only TCF12 was examined further by performing IPs with the anti TCF12 antibody. As with the IPs using the anti E2A antibody the TCF12 antibody gave a heavy smearing, which made it difficult to detect any protein in the IP fraction.

This section has shown that the ID4 IP in the Ba/F3 clone 11 identified a number of interacting proteins. The interaction with the bHLH protein TCF12 was confirmed in this cell line. No E2A protein was detected in this IP.

6.3.2 ID4 interacting proteins in NALM-6 TR cells.

A similar approach to detecting interacting proteins with ID4 in Ba/F3 cells was adopted in NALM-6 TR cells with and without induced ID4 expression. IPs were performed using the anti Myc tag antibody and a cytosolic preparation (because of the reduced background) run on an SDS PAGE gel which was coomassie stained and processed for mass spectrometry (data not shown). As with the Ba/F3 cells, 2 lists of immunoprecipitated proteins were produced and the proteins that appeared in the sample where ID4 had not been induced were subtracted from the list with ID4 induced expression. These results can be seen in Appendix Table A1.10.

Analysis of the proteins in Appendix Table A1.10 showed ID4 and two other bHLH proteins E2A and TCF4; the mass spectra data to these proteins are in Appendix Figure A 1.8 and Appendix Table 1.11.

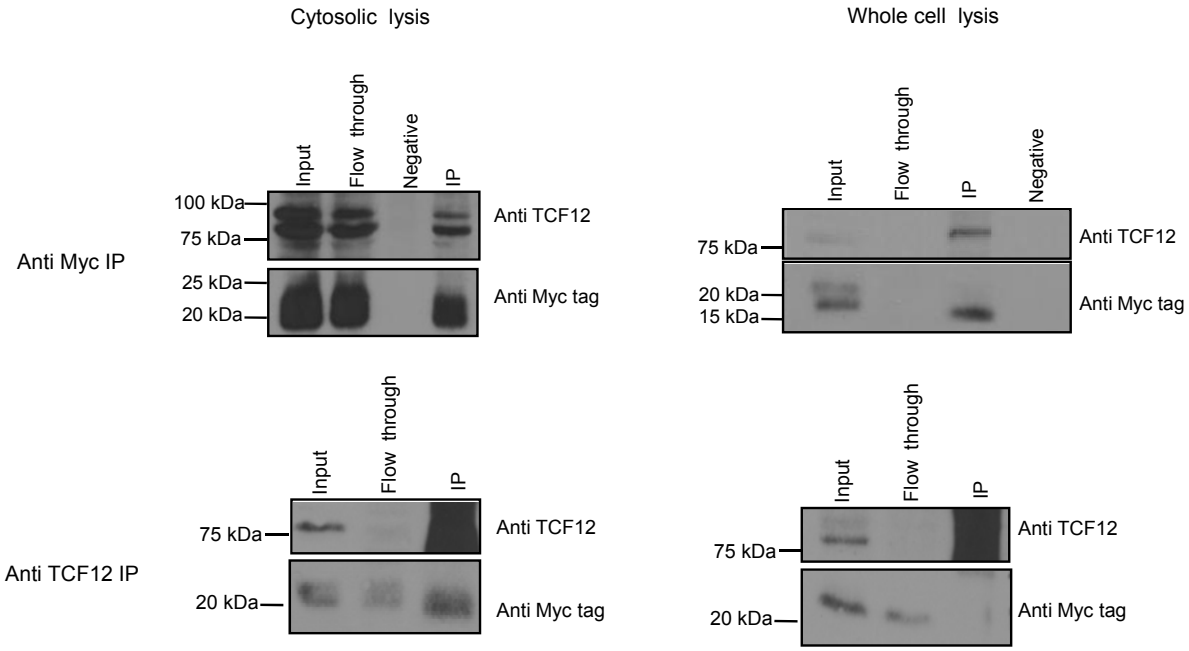


Figure 6.5 Validation of TCF12 interaction in Ba/F3 ID4 clone 11. **A)** IP using both cytosolic and whole cell preparations with the anti Myc tag and anti TCF12 antibodies using 40μl of beads. For the cytosolic IPs, 120μg of lysate was used, and for the whole cell lysates, 84μg of protein was used. Negative controls were used in some instances to identify any non-specific binding and involved performing the IP without the anti Myc tag antibody. The samples were run on a 12% SDS PAGE gel and western blotted. The blot was then probed with anti TCF12 and anti Myc tag antibody sequentially.

Two peptides for E2A were identified with acceptable mass ion scores and spectra. TCF4 only had one peptide that was identified and a mascot score which was just under the threshold, but has adequate spectra. GSTP1 was the only protein which also was among the proteins immunoprecipitated from the Ba/F3 clone ID4 11 (see Table 6.2).

ID4 was found to interact with E2A and TCF4 in NALM-6 TR cells, but no TCF12 was detected. The interaction with E2A and ID4 in NALM-6 TR cells had been confirmed (Figure 6.3). The interaction between ID4 and TCF12 was confirmed by western blotting of immunoprecipitated fractions (Figure 6.6). There was no available antibody to immunoblot for TCF4.

Figure 6.6 showed that the largest isoforms of TCF12 interacted with ID4 in NALM-6 TR cells. This interaction was similar to what was observed with the Ba/F3 clone ID4 11 cells where not all the TCF12 in the cells immunoprecipitated with ID4. This interaction was more pronounced in the whole cell lysates. Overall, ID4 in NALM-6 TR cell interacts with the 3 bHLH proteins (E2A, TCF12 and TCF4) important for lymphocyte development.

6.3.3 Confirmation of other ID4 interactions.

ID4 interactions with the bHLH protein E2A and TCF12 would potentially antagonise them and disrupt normal B-cell development (see chapter 5). The gene expression data in the primary BCP-ALL cases showed no changes in B-cell specific gene expression (chapter 3). Therefore, other interactions listed in Tables 6.1 and 6.2 were examined further. Based on functional significance, the proteins RSU_1 and ZNF289 were chosen for further analysis.

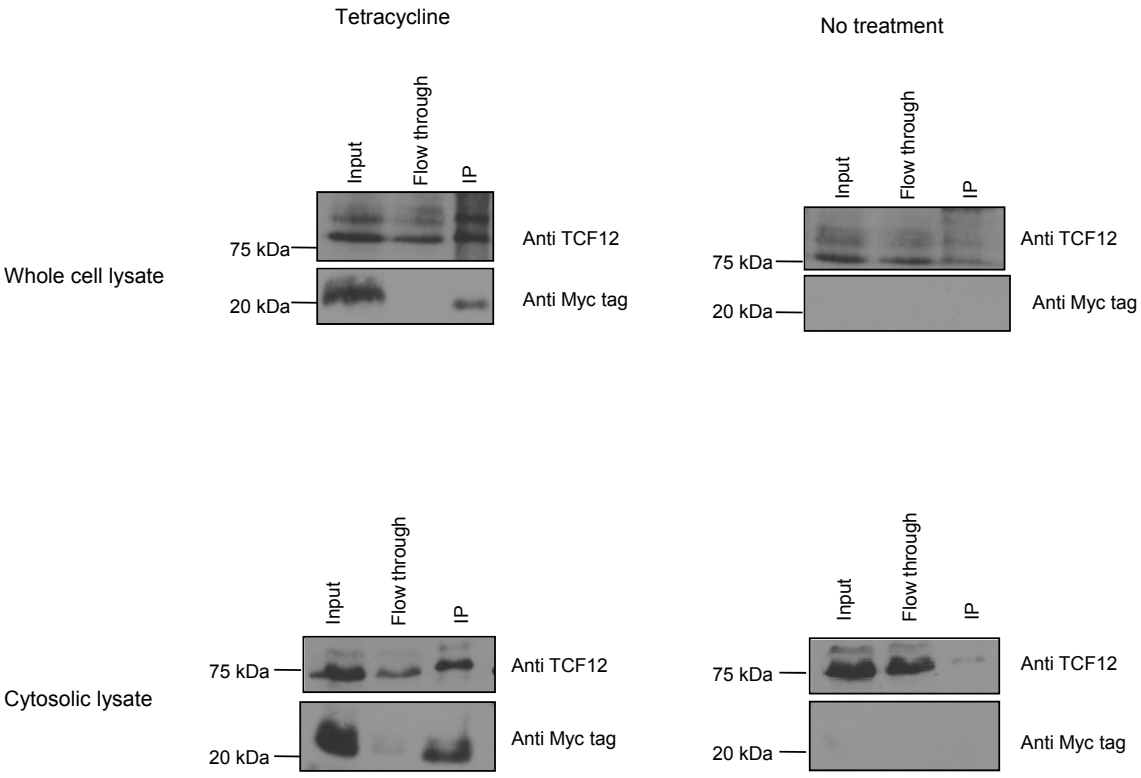


Figure 6.6 ID4 and TCF12 interaction in NALM-6 TR cells. IP using the Myc tag antibody of ID4 from NALM-6 TR ID4 cells treated with 2μg/ml tetracycline for 4 hours. Whole cell lysates were produced from $1x10^8$ cells and the IP performed using 500μg of protein and 40μl of beads. Cytosol preparations were produced from $1.8x10^8$ cells and the IP performed with 670μg. 20 μl of immunoprecipitated fraction and the other relevant fractions were run on a 12% SDS PAGE gel and western blotted. The blots were then probed with anti TCF12 antibody and anti Myc tag antibody sequentially.

Ras suppressor protein1 (RSU_1) protein is ubiquitously expressed and a known tumour suppressor because of its ability to inhibit transformation by the Ras oncogene. RSU_1 is deleted in high grade gliomas, but alterantive splice forms of this protein are expressed in 30% of high grade gliomas and correlate with active Ras expression. Loss of RSU_1 results in increased migration properties of cells (Dougherty *et al.*, 2008). Therefore ID4 could interact with and affect the function of RSU_1, thereby assisting in transformation.

The transcription factor zinc finger 289 (ZNF289) expression is a member of the GTPase family of proteins for ADP-ribosylation factors (Arfs). Arfs are important for the assembly of coated carrier vesicles for transport between the golgi apparatus and the endoplasmic reticulum.

ZNF289 expression is regulated by ID1 in a breast cancer cell line (Saitoh *et al.*, 2009; Singh *et al.*, 2005).

As with TCF12, IPs were performed with the anti Myc tag antibody to confirm the interaction between ID4 and RSU_1 and ZNF289 (Figure 6.7). The relative mass spectral data are in Appendix Figure A 1.9 and Table A 1.12.

The protein ZNF289 has a predicted molecular weight of 56 kDa, and on the blots in panel A, ZNF289 ran at approximately this size. ZNF289 immunoprecipitated with the Myc tag pull in the cytosolic preparation, but there was also a band in the negative lane. In the whole cell preparation, ZNF289 did not IP with the Myc tag IP. Therefore ZNF289 and ID4 do not co-IP in Ba/F3 ID4 clone 11 cells.

RSU_1 has 2 isoforms, which have predicted molecular weights of 25 kDa and 30.5 kDa. The blots in panel A show only one protein band at approximately 30 kDa. RSU_1 did not interact with ID4 in Ba/F3 ID4 clone 11 cells.

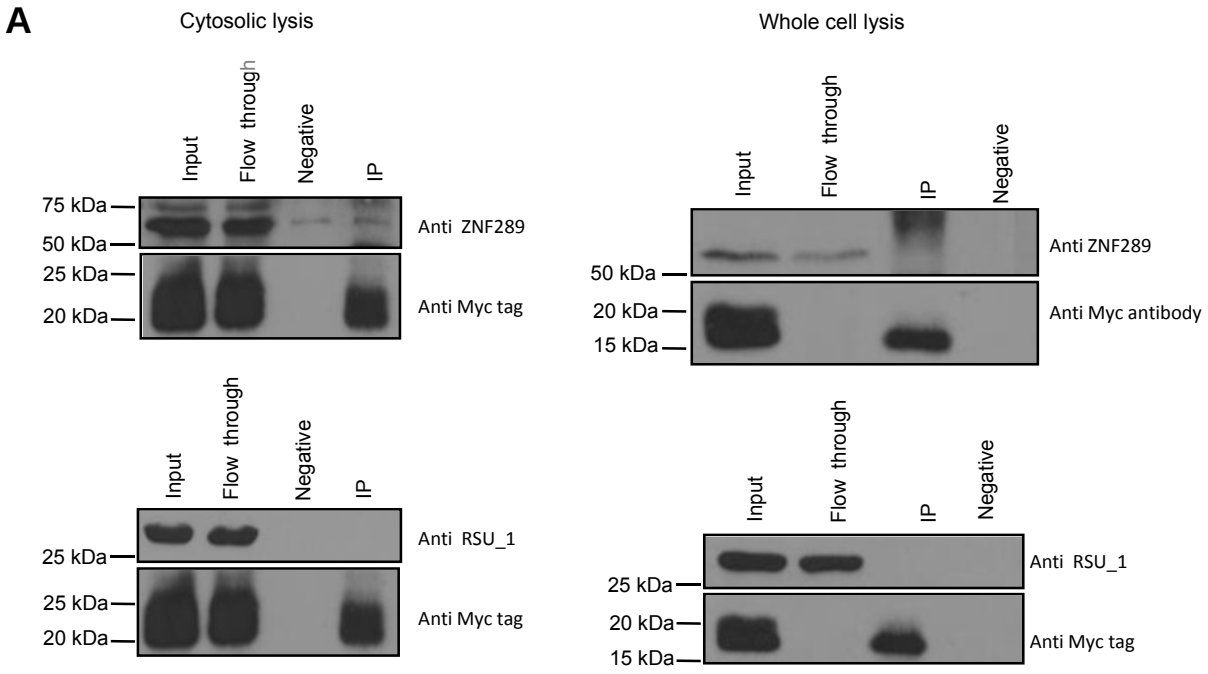


Figure 6.7 Validation of ZNF289 and RSU_1 interactions in Ba/F3 ID4 clone 11 cells. IP using both cytosolic and whole cell preparations with the anti Myc tag antibody and using 40μl of beads.

These results confirm the ID4 interaction with RSU_1 or ZNF289 were not reproducible. Further work requires the analysis of the other proteins listed in Table 6.1 and 6.2.

6.4 ID4 is targeted for degradation via the proteasome.

The identification of proteins which interact with ID4 required a large amount of protein and although ID4 immunoprecipitated, it was difficult to improve the spectra of the immunoprecipitated protein. This was most evident in the Ba/F3 cells which constitutively express ID4. The ID proteins have a short half life and are rapidly degraded. The degradation of the ID proteins has been linked to the ubiqitin proteasome system (Bounpheng *et al.*, 1999; Fajerman *et al.*, 2004). The degradation of ID4 by the UPS was then examined.

6.4.1 Ectopically expressed ID4 is targeted for proteasomal degradation.

IPs attempted in Ba/F3 clone 16 had proved difficult because over time, this clone down-regulated the expression of ID4 protein (Figure 6.8 panel A). To examine if ID4 protein is regulated at the level of protein stability the Ba/F3 clones ID4 11, EV clone 14 and NALM-6 cells were treated with MG132, a proteasomal inhibitor which reduces the degradation of ubiquitin conjugated proteins by the 26S proteasome without affecting the ATPase or isopeptidase activity (www.merck-chemicals.com).

Blocking the proteasome resulted in an accumulation of ubiquitinated proteins as is evident by a heavy smear detected by staining with the anti-ubiquitin antibody (Figure 6.8). However, these experiments do not clearly show that there is an accumulation of ubiquitinated ID4 protein; this would have to be clarified by IP and should be addressed in future work. Proteasomal inhibition increased the amount of ID4 protein within the Ba/F3 clone ID4 11 cells. This effect was not evident with the NALM-6 cells.

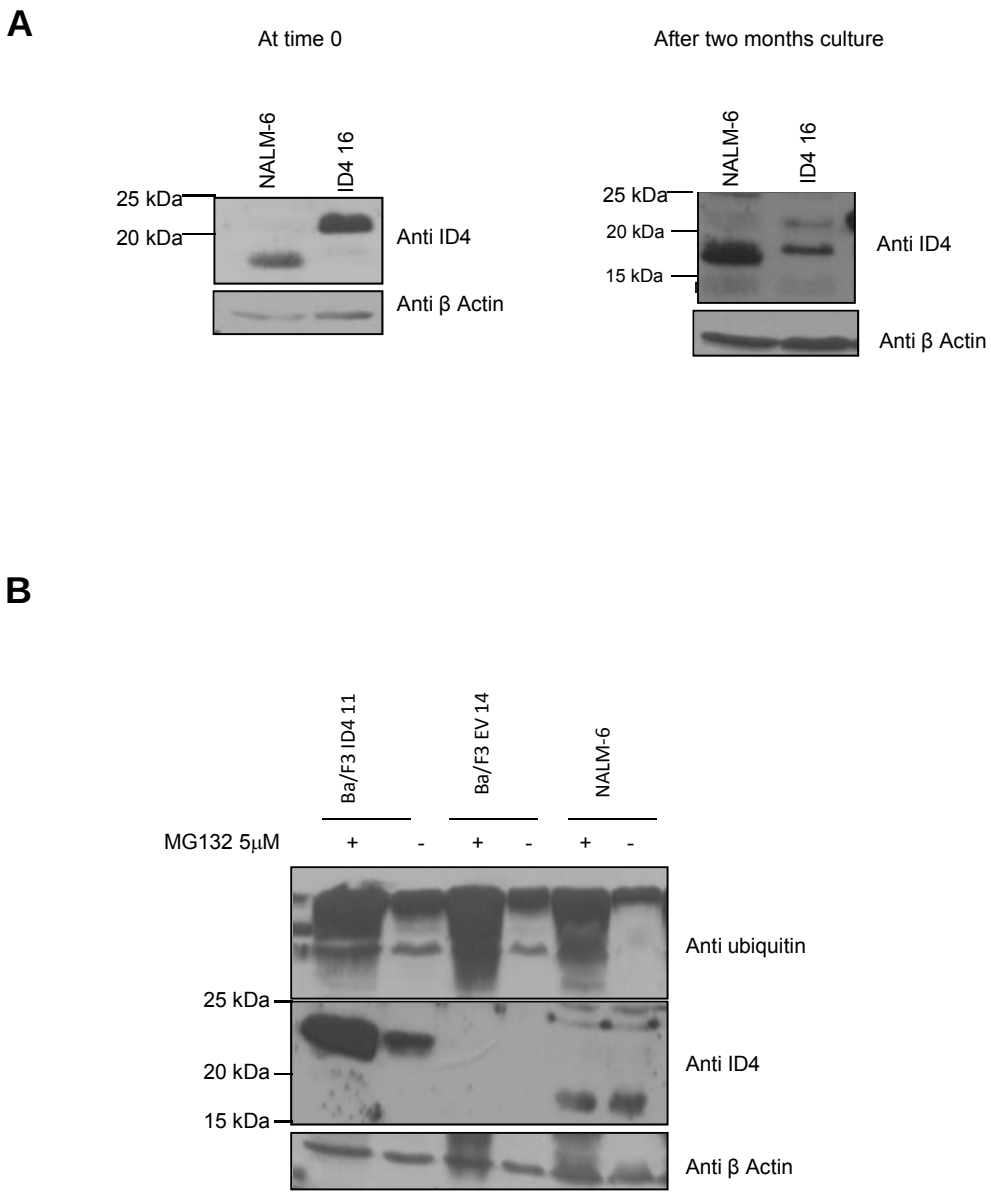


Figure 6.8 ID4 is targeted for degradation via the proteasome. **A)** ID4 expression was lost in Ba/F3 ID4 clone16 (2 months apart) $1.5x10^6$ cells were lysed in 2xSDS loading buffer and run on a 15% SDS PAGE gel. The blots were then probed with anti ID4 antibody and anti β Actin sequentially. **B)** Ba/F3 ID4 clone 11 and EV clone 14 and NALM-6 cells were treated with 5μM MG132 for 5 hours. $2x10^6$ cells were lysed in 2xSDS buffer and run on a 15% SDS PAGE gel. The blots were then probed with anti ubiquitin antibody, anti ID4 and anti β Actin antibodies sequentially.

6.4.2 The D box motif is not required for the proteasomal mediated degradation of ID4.

The D box motif in the ID4 protein, targets it for proteasomal mediated degradation (Lasorella *et al.*, 2006). Ba/F3 clones expressing ID4 (mut 4 and 9) with a mutated D box (domains R139G and L143V) were used to examine if this motif is essential for ID4 degradation via the proteasome.

Ba/F3 clone Mut 4 and Mut 9 were treated to 5μM MG132 for 5 hours and the accumulation of ID4 protein examined (Figure 6.9). Inhibition of the proteasome with MG132 resulted in the partial accumulation of ID4 protein, but not to the same extent as observed with the Ba/F3 ID4 clone 11 cells (see Figure 6.8). However, this control was not used in this experiment and would need to be considered for future experiments.

Therefore the D box motif in ID4 is required for its proteasomal mediated degradation; however it is not the only mechanism by which ID4 is targeted for degradation.

These experiments have shown that ectopic, but not endogenous ID4 protein is regulated at the level of protein stability in B-cells in a mechanism that is not dependent on the D box motif.

6.5 Conclusions.

The function of the ID proteins is defined by their interacting partners. This chapter aimed to identify the interacting partners of ID4 in a B-cell environment.

This was approached by confirming any interactions with the main bHLH protein involved in B-cell development, E2A (composed of E12 and E47). It was only when ID4 was ectopically over-expressed that it showed an interaction with E2A, confirmed with IPs. This interaction did not involve all the E2A in the cells.

Whilst these results support the data in chapter 5, where disrupted B-cell development was observed, possibly by ID4 expression antagonising E2A, the BCP-ALL cases expressing *ID4* RNA did not show changes in genes involved in B-cell development. Therefore other ID4 partners need to be identified.

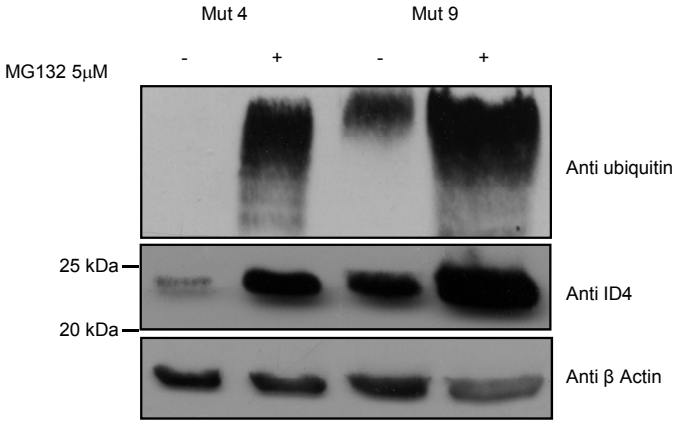


Figure 6.9 ID4 D box mutants are not resistant to proteasomal mediated degradation. Ba/F3 ID4 Mut clones 4 and 9, were treated with 5μM MG132 for 5 hours. $2x10^6$ cells were lysed in 2xSDS buffer and run on a 15% SDS PAGE gel. The blots were then probed with anti ubiquitin antibody, anti ID4 antibody and anti β Actin sequentially.

To address this, entire IPs were processed by mass spectrometry. This provided a large list of proteins including the bHLH proteins E2A, TCF12 and TCF4. Other interacting proteins which had functional significance were examined further, but failed to show any interactions. Therefore further work is required to re-analyse the mass spectrometry data and identify other interacting proteins (see discussion).

Difficulties in these experiments involved the limitations of the ID4 (H70) antibody in immunoprecipitating endogenous ID4. Also, whilst over-expressed ID4 protein is regulated at the level of protein stability independent of its D-box motif, this was tested using MG132. MG132 also targets lysosomal proteases as well as the proteasome. Therefore other proteases and proteolytic components of the cell cannot be ruled out. Other inhibitor s such as epoxomicin or lactacystin that are more specific for the proteasome would need to be tested. Also, an IP is required to confirm that there is an increase in ubiquitinated ID4 protein if ID4 is really regulated at the level of protein stability. Further work is required to address these issues.

Chapter 7: Discussion

ID4 was found to be the target of an *IGH* translocation in 13 cases of BCP-ALL (Russell *et al.*, 2008). Further studies have shown that *ID4* is up-regulated in a distinct subset of BCP-ALL cases without the translocation and is not associated with any other known subgroup of BCP-ALL (this thesis and personal communication Dr Gunnar Cario and Professor Reiner Siebert, Kiel, Germany). These data suggest that *ID4* may act as an oncogene in BCP-ALL. The aim of this thesis was to examine the possible functions of ID4 in cells of the B-cell lineage, with the eventual aim of determining whether ID4 might be a therapeutic target. The objectives of this thesis have been addressed by examining *ID4* expression, cellular localisation, effects of ID4 over-expression in B-cells and the interacting partners of ID4.

7.1 *ID4* expression in BCP-ALL.

Bioinformatics identified studies in which primary BCP-ALL cases were examined. Overall, *ID4* RNA expression was identified in 10% (n=746) of these cases. Where the data were available, *ID4* RNA expression did not correlate with any cytogenetic abnormality. The primary BCP-ALL cases with *ID4* expression shared a similar and very distinctive gene expression profile across different studies. This implied that ID4 was either “driving” a major transcriptional effect or that it was itself expressed as a consequence of another oncogenic pathway in these cells.

ID4 is not normally expressed in B-cells and therefore the precise functions of this gene in B-cell malignancies remain unknown. In contrast, the over-expression of ID2 was detected in HRS cells of classical Hodgkin lymphoma, where it was found to antagonise E2A function, resulting in the down-regulation of B-cell specific genes such as *CD19, CD79a, RAG1, RAG2 and EBF*. ID2 expression also induced inappropriate gene expression, such as *GATA3, TCF7, MAF* and *CSF1R* (Mathas *et al.*, 2006). ID3 ectopic expression in B-cells results in a similar phenotype (Kee *et al.*, 2001).

BCP-ALL expressing *ID4* RNA did not show significant down-regulation of B-cell specific genes in the bioinformatic studies. Therefore, it seems likely that ID4 does not function like ID2 and ID3 in B-cell progenitors and that ID4 has

specific functions in B-cells distinctive from those of other family members. Consistent with this, screening of BCP-ALL xenografts showed expression of both *ID2* and *ID3* RNA in the majority of the cases, suggesting that the effects of *ID4* expression in BCP-ALL cases would have to work in competition with *ID2* and *ID3* or by a different mechanism. It is also of interest that only *ID4* appears to be involved in chromosomal translocations in BCP-ALL. In contrast, all members of the *CEBP* gene family are involved in *IGH* translocations (Dyer *et al.*, 2010). Collectively, these data suggest a major and specific role for *ID4* expression in BCP-ALL.

In *ID4* expressing BCP-ALL, it appears that ID4 does not target B-cell development, but instead may drive inappropriate gene expression. This gene expression profile (Table 3.4) involves genes important for transcriptional regulation and neurological development. It may also be the case that *ID4* is itself, induced by these pathways. Both *ID4* and some of the associated up-regulated genes such as *PDGFRA* could be targeted therapeutically (Andrae *et al.*, 2008).

7.2 *ID4* expression in CLL.

ID4 RNA expression was identified in 34% of the CLL cases screened. *ID4* RNA expression proved more reliable than examining ID4 protein expression in these cases, since the antibody available can generate false positives. The data from Yu *et al*. showed that the *ID4* promoter was methylated in CLL cases, but *ID4* RNA expression was not examined in this paper (Yu *et al.*, 2005). It may be that promoter methylation does not always completely block transcription and therefore allows low levels of RNA expression.

The samples screened in our study comprised of B and T lymphocytes. A further B-cell purification was performed on one CLL case (courtesy of Dr Majid) and confirmed that *ID4* RNA was expressed in B-cells. ID4 protein expression also needs to be confirmed in *ID4* positive CLL cases.

ID4 expressing cases did not appear to associate with any specific prognostic factor in CLL. The CLL cases expressing *ID4* overall presented with more

cytogenetic abnormalities compared to the *ID4* negative CLL cases. *ID4* expression has been shown to be induced by p53 gain of function mutants in breast cancer, thereby promoting angiogenesis (Fontemaggi *et al.*, 2009). *P53* mutational status is an important prognostic factor in CLL and further work is required to examine the correlation between *ID4* expression and p53 status in CLL.

Previous studies have shown ID2 expression in CLL; ID2 expression in AID deficient CLL cases was associated with the expression of alternative *PAX5* transcripts (Oppezzo *et al.*, 2005). It would have been beneficial to examine *PAX5* expression and overall gene expression in these cases and should be addressed in further work.

7.3 Expression of *ID4* in other malignancies.

ID4 RNA was also expressed in several malignancies, where it was co-expressed with the other *ID* RNAs. Within each of these disease types, only a sub-group expressed *ID4* RNA. Interestingly, as with the BCP-ALL, these *ID4* positive subgroups displayed a different and distinctive gene expression profile to the cell lines not expressing *ID4*. Furthermore, there was no correlation of genes between different malignancies. These data support the hypothesis that *ID4* expression significantly alters the gene expression profiles in a cell-context dependent manner. High *ID4* expression in most of the SCLC cell lines, implicates ID4 in the pathogenesis of this disease and this would require further examination.

There were 4 B-cell lines identified as expressing *ID4* RNA, of which 3 were examined further and showed ID4 protein expression. These cell lines showed no common gene expression profile, but all but one harboured an *MYC/IGH* translocation.

ID4 expression in normal tissues is more restricted than the other ID family members. It is unlikely that *ID4* expression in many different malignancies results from a common aberrant pathway in all these diseases, because of the lack of overlap between the gene expression profiles. Instead, ID4 may act as a

driver in these cases, exerting a different transcriptional effect with each disease.

7.4 Effects of ID4 expression in B-cells.

7.4.1 ID4 induces G1 cell cycle arrest.

Induced ID4 expression in NALM-6 TR and RCH-ACV TR resulted in cell cycle arrest at 24 hours. The NALM-6 TR cells did begin to proliferate again from 36 hours onwards. The RCH-ACV TR cells did not recover to the same extent. ID4 protein levels declined after 24 hours to a steady state level in both cell lines. What modulates this change in protein level requires further investigation. All ID proteins have a very short half-life and are subject to proteasomal degradation (Bounpheng *et al.*, 1999; Fajerman *et al.*, 2004; Lasorella *et al.*, 2006).

The expression of both ID1 and ID2 is normally within the G1 stage of the cell cycle and their absence will result in G1 arrest in human embryonic lung fibroblasts (Hara *et al.*, 1994). ID1, ID2 and ID3 promote the cell cycle. ID1 antagonises the expression of p16 (Ohtani *et al.*, 2001), ID2 expression can interact with Rb and promote G1 to S phase transition (Iavarone *et al.*, 1994). ID1, ID2 and ID3 all decrease the expression of p27 and or p21 in either an E2A dependent or independent manner (Engel and Murre, 1999; Peverali *et al.*, 1994; Prabhu *et al.*, 1997). Similarly, ID4 was found to be required for G1 to S phase transition in neurones (Yun *et al.*, 2004). In B-cell progenitors, ID3 was shown to induce cell cycle arrest and induce apoptosis (Kee, 2005).

However, in contrast, over-expression of *ID4* in a murine tumour lymphoid cell line Yac1 induced cell cycle arrest and apoptosis (Yu *et al.*, 2005). The over-expression of *ID4* in a prostate cell line induced an S phase arrest (Carey *et al.*, 2009). It is only in astrocytes that the over-expression of *ID4* has been associated with an increased proliferation following cyclin E up-regulation on a *cdkn2a-/-* background (Jeon *et al.*, 2008). In the human B-cell lines examined in this thesis, ID4 induced the expression of p27 in both RCH-ACV TR and NALM-6 TR; this alone may be inducing the G1 arrest.

Carey *et al.* also found ectopic *ID4* expression induced p21 and p27 in prostate cancer cell lines and this was attributed to ID4 protein inducing the expression of *E47* (Carey *et al.*, 2009). RCH-ACV cells harbour the *E2A/PBX1* translocation and therefore show haplo-insufficiency of *E2A* (Curry *et al.*, 2001) and would most likely show a reduced responsiveness to *ID4* expression, which was the case with p21 expression. p21 and p27 are also up-regulated by other mechanisms, such as the TGFβ signalling pathway and the p53 pathway. The potential involvement of ID4 in these pathways needs to be examined.

The ability of ID4 to reduce proliferation and induce cell cycle arrest in cells suggests that it might act as a tumour suppressor. The *ETV6-RUNX1* translocation (a prominent translocation in BCP-ALL) was shown in Ba/F3 cells to reduce proliferation and induce cell cycle arrest; however, these cells were resistant to the anti-proliferative effects of TGFβ (Ford *et al.*, 2009). Therefore, the effects of ID4 expression in B-cells requires further examination.

The cells in ID4 induced G1 arrest also showed partial resistance to apoptosis induced by TRAIL and staurosporine. This experiment requires a further control to measure the normal rates of apoptosis of cells in G1 arrest. These results are significant, but would benefit from a repeat with a western blot to observe any effects on the caspase signalling pathways.

The effects of tetracycline on these cells (parental or expressing EV) was not assessed in these experiments. This should be addressed in future work to confirm that the observed effects are due to *ID4* expression and not tetracycline treatment.

No effects of ID4 expression were observed in the Ba/F3 ID4 clones 11, 16 and 8 in terms of proliferation or growth factor withdrawal. One reason for this appears to be the constitutive expression of ID4 protein in these cells which appears to be regulated at the level of protein stability (see section 6.4). This cell line is also a mouse progenitor B-cell line at a more immature stage compared to NALM-6 and RCH-ACV cells and therefore may elicit a different response to ID4 over expression.

7.4.2 ID4 expression in normal mouse foetal liver B lymphocyte progenitors does not induce proliferation, but impairs B-cell differentiation.

ID4 expression in normal mouse B lymphoid progenitors impaired B-cell development, but this recovered as the cells matured (Figure 5.4) Similarly, ID1 and ID2 inhibit B-cell development in favour of other lineages (Cochrane *et al.*, 2009; Ji *et al.*, 2008). By analogy with the ID1 and ID2 data, *ID4* expression may be antagonising E2A, but other transcription factors, such as EBF (which can be induced by other transcription factors) can subsequently take over, allowing B-cell development to progress, but at a slower pace (Roessler *et al.*, 2007).

This is a similar profile that has been observed in *e47* +/- mice and the recovery in B-cell development at later stages is due to different stages of B-cells not always requiring *e2a* expression (Quong *et al.*, 2004). However, this model does not support the increased pre B-cell population with *ID4* expression, when it should be reduced with the lack of E2A. Therefore ID4 expression may be exerting some specific effect at the pre-B-cell stage of development.

The effects observed in the B-cells with *ID4* expression, suggest that ID4 is affecting B-cell development and proliferation at specific stages of differentiation. Further work would require examination of the gene expression profiles of the B lymphocyte progenitors and also examining the effects of *ID4* expression *in vivo* in an *ID4* transgenic mouse model.

7.4.3 ID4 expression in Eμ-myc mouse foetal liver B lymphocyte progenitors induces proliferation.

The expression of *ID4* in B-cell progenitors with an *cdkn2a-/-* background allowed the recapitulation of the BCP-ALL cases with the t(6;14)(p22;q32) translocation, most of which also harboured deletions of the *CDKN2A* locus. Furthermore, *id4* expression in gliomas cooperated with the *cdkn2a-/-* genotype (Jeon *et al.*, 2008). However, *ID4* did not cooperate with this phenotype in the *cdkn2a-/-* B lymphocyte progenitors, but this data set was not significant and

further repeats are required. This further supports the concept that *ID4* may induce transformation in a context dependent manner

In contrast, *ID4* expression in *Eμ-myc* infected livers did not inhibit or cooperate with the response to *myc* until day 14 where there was an increased proliferation with *ID4 myc* expressing cells compared to the *EV myc* expressing cells. The response to *myc* expression was not seen till day 14, when the B-cells had reached maturity.

ID2 is up-regulated by MYC signalling in neurones (Lasorella *et al.*, 2000). In Burkitts lymphoma *ID2* displays high levels of expression, but is dispensable for *MYC* induced transformation in these cells (Nilsson *et al.*, 2004). *ID4* therefore may cooperate with *MYC* to induce transformation, but since not all malignancies with *ID4* expression also show deregulated *MYC* expression, this effect may be context dependent. This requires further examination.

7.5 Sub-cellular localisation of ID4

The sub-cellular localisation of the ID proteins can obviously influence their functions. For example, ID3 and ID1 are imported into the nuclei because of their interaction with bHLH proteins (Deed *et al.*, 1997; Trausch-Azar *et al.*, 2004). Conversely, ID2 and ID4 also sequester the bHLH proteins OLIG1 and OLIG2 into the cytoplasm of neurones, preventing the differentiation of these cells (Samanta and Kessler, 2004). The nuclear localisation of ID4 and ID2 in prostate cancer epithelia is associated with a poor prognosis and increased risk of metastasis (Yuen *et al.*, 2006). Thus subcellular localisation is cell-context dependent.

In human malignant B-cells, ID4 was predominately cytoplasmic, irrespective of being endogenously or ectopically expressed, however in the mouse progenitor Ba/F3 cells, ectopic ID4 protein was both nuclear and cytoplasmic.

ID2 cellular localisation changes from nuclear to cytoplasmic as cells differentiate (Matsumura *et al.*, 2002). The Ba/F3 cells are a more immature than the BCP-ALL cell lines NALM-6 and RCH-ACV cells and this could explain the differences in ID4 protein localisation in these cells. Further work to address

this would involve examining the localisation of ID4 protein in the cells of the foetal liver assay during different stages of B-cell development.

Another factor that could also influence ID4 cellular localisation is its size, since proteins under 50 kDa can diffuse freely between the nucleus and the cytoplasm (Trausch-Azar *et al.*, 2004).

ID4 may also be regulated in a cell cycle dependent manner. The phosphorylation of ID2 on the amino-terminal SPVR motif (on the serine residue by the Cyclin A and Cyclin E cdk2 complexes) sequesters it into the cytoplasm (Matsumura *et al.*, 2002) and the nuclear localisation of ID proteins increases their half life (Deed *et al.*, 1997; Lingbeck *et al.*, 2005; Trausch-Azar *et al.*, 2004). ID4 contains the same SPVR domain as ID2 and this may locate it to the cytoplasm when the cells are arrested. ID4 did not show any interaction with Rb as was found with ID2 (Iavarone *et al.*, 1994; Lasorella *et al.*, 2000; Lasorella *et al.*, 2005).

7.6 ID4 protein interactions.

The effects of ID4 expression in B-cell lines and B lymphocyte progenitors suggested that ID4, like ID1 and ID2 (Cochrane *et al.*, 2009; Ji *et al.*, 2008) might be antagonising E2A function and other bHLH transcription factors, which would affect the cell cycle and B-cell development. The proteomic data for immunoprecipitated ID4 interacting proteins showed that ectopically expressed ID4 interacted with the E proteins E2A, TCF12 and TCF4 in B-cell lines However, ID4 did not interact with all the E2A and TCF12.

ID4 interacting with the E proteins, but not down-regulating the E protein genes implies that ID4 may function by antagonising the E proteins differently to the other ID proteins. The AML1-ETO fusion protein interacts with E2A and TCF12 via their AD1 (activation domain1) domains and in doing so induces a co-factor exchange in these cells (refer to Figure 1.6). This alters the gene expression profile which predisposes the cells to further leukemicogenic events (Zhang *et al.*, 2004). This was the consequence of expressing *ETO* in B-cells, a polypeptide not normally expressed in the haematopoietic system (Zhang *et al.*,

2004); this may also be the case with *ID4* expression, since it is not normally expressed in B-cells.

ID4 therefore may interact with the E proteins via its HLH domain, but instead of antagonising the transcriptional activity of these proteins, ID4 may change their transcriptional function; this may rely on domains other than the HLH in ID4. The lack of correlation in gene expression across different malignancies may be because of the presence of other tissue specific transcription factors, which also influence the E proteins. Also, supporting this theory, is the lack of tissue specific class II bHLH proteins in the IPs with ID4.

Other ID4 interacting proteins which may influence its transforming ability also require further examination. Some of these interactions include septins 2 and 11 which were found interacting with ID4 in the Ba/F3 clone 11 cells but not in NALM-6 TR cells. Septins are members of the GTP binding family of proteins required for cell division, cytokinesis (Spiliotis *et al.*, 2005), re-organisation of the actin cytoskeleton and maintenance of the suppressor of cytokine signalling 7(SOCS7) NCK pathway (Kremer *et al.*, 2007; Spiliotis *et al.*, 2005). Disruption of the septins results in binucleate cells as a result of the disruption of chromosome alignment and segregation (Spiliotis *et al.*, 2005). ID4 in these cells could disrupt septin function resulting in chromosomal instability and ploidy.

The septins sequester NCK proteins into the cytoplasm and during an apoptosis response, NCK proteins translocate to the nucleus, this process is dependent on the SOSC7 protein. The NCK protein, once in the nucleus, is important for the activation of down-stream signalling of the DNA damage cascade and results in activation of the chk2/ p53 pathway. The loss of septins results in the nuclear accumulation of NCK, but not the activation of the chk2/p53 pathway without ATM/ATR activation (Kremer *et al.*, 2007). SOCS7 was also identified in an IP of Ba/F3 ID4 clone 11 cells. Therefore, ID4 expression could bind septins releasing the SOCS7/NCK protein into the nucleus allowing cell cycle arrest to occur. This implicates ID4 in mitosis.

Based on this data, ID4, appears to be involved in transcriptional regulation (by the interaction with E proteins), but also in the regulation of cell division.

7.7 ID4 degradation.

Treatment of the Ba/F3 clones with the proteasomal inhibitor MG132, resulted in the accumulation of the ID4 protein, suggesting that ID4 is degraded in a proteasomal manner. If this degradation pathway is disrupted, high levels of ID4 protein could drive transformation. MG132 does not disrupt all components of the proteasome and further work would require examining ID4 protein stability with different compounds which could target other degradation pathways. However, what this data showed was that ID4 is regulated at the level of protein stability.

ID4 degradation is dependent on the UPS, which can target domains such as the D box (Lasorella *et al.*, 2006). The D box domain is targeted by the APC/C or by the CSN complex, which is also cell cycle dependent. A component of the COP9 complex, CSN7a has also been identified in an IP with ID4 in Ba/F3 cells (Tables 6.1 and 6.2). A component of the CSN is known to prevent the ubquitination of ID3 and therefore increase its half life (Berse *et al.*, 2004). It is possible that there is a feedback network involved with the UPS system, ID4, COP9 and possibly even the APC/C in a cell cycle-dependent manner.

The D box mutants of ID4 were not resistant to proteasomal mediated degradation and a publication by Wang *et al.* demonstrated that mutation of the D box in ID1also did not protect it from proteasomal mediated degradation. This study concluded that the N terminus in ID1 is also important for its ubiquitination (Wang *et al.*, 2008) and therefore this region may also be targeted in ID4.

The UPS mediated degradation of *ID4* in NALM-6 was not evident and this may be a result of *ID4* being mutated in this cell line. This would have to be confirmed by sequencing and checking whether the same thing happened in other cells expressing endogenous ID4.

The RCH-ACV TR clone 1 showed that *ID4* expression was tolerated from a leaky tetracycline expression system, but induced over-expression still had an effect on the cells. Levels of ID4 protein therefore may influence transformation which may be based on competition with other ID proteins.

The effects of *ID4* over-expression in B-cells differed depending on how ID4 was expressed (constitutive and induced) and between the different stages of B-cell development (as was observed with the B lymphocyte progenitors). What remains to be addressed is how the cell lines expressing endogenous ID4 can continue to proliferate based on the data with ectopic expression.

7.8 Final hypothesis.

ID4 is not expressed in normal haematopoietic cells. The expression of *ID4* in malignant B-cells is therefore ectopic. *ID4* expression in B-cell malignancies may directly “drive” transformation. Alternatively, cooperation with other oncogenes may be necessary for the malignant phenotype to emerge. Further work is needed to establish this point. For example, RNA knockdown in cell lines and primary B-cells expressing endogenous ID4 could be investigated.

ID4 may induce transformation through its ability to alter the transcriptional profile. This transcriptional profile change may depend on ID4 interacting with the E proteins, but instead of antagonising their function and inhibiting a B-cell transcriptional profile, ID4 may alter the cofactors that these E proteins interact with, thereby altering their transcriptional targets. This would allow inappropriate gene expression in differenT-cells. The types of genes expressed would be dependent on the types of tissue specific cofactors and other bHLH proteins available in the cells.

ID4 may also influence other events post transcriptionally, such as the cell cycle and promote genetic instability by cooperating with or influencing the deregulation of mitosis and/or components of the APC/C complex. The different interactions of ID4 may be reliant on its protein domains outside of the HLH domain, which appear to be conserved across species.

Collectively, these events are summarised in Figure 7.1.

Further work required to assess this includes:

- Further define the interacting partners of ID4
- Examine changes in gene expression in B lymphocyte progenitors
- Examine any association of ID4 with p53 mutation or dysfunction.
- Further analyse the degradation pathway for ID4.
- Examination on what drives *ID4* expression in malignant B-cells.
- Knock-down *ID4* RNA in B-cell lines to define its role in these malignancies.

7.9 Conclusion.

The data in this thesis identify the expression of *ID4* in different malignancies. This expression of *ID4* cannot be attributed to random somatic mutation, because of its recurrent involvement in BCP-ALL cases (with and without *IGH* involvement) and because of its expression in different malignancies. *ID4* expressing malignancies exhibited altered transcriptional profiles. ID4 may function via its interaction with E proteins to alter transcription and/or may function post- transcriptionally, predisposing cells to chromosomal instability and alternative gene expression.

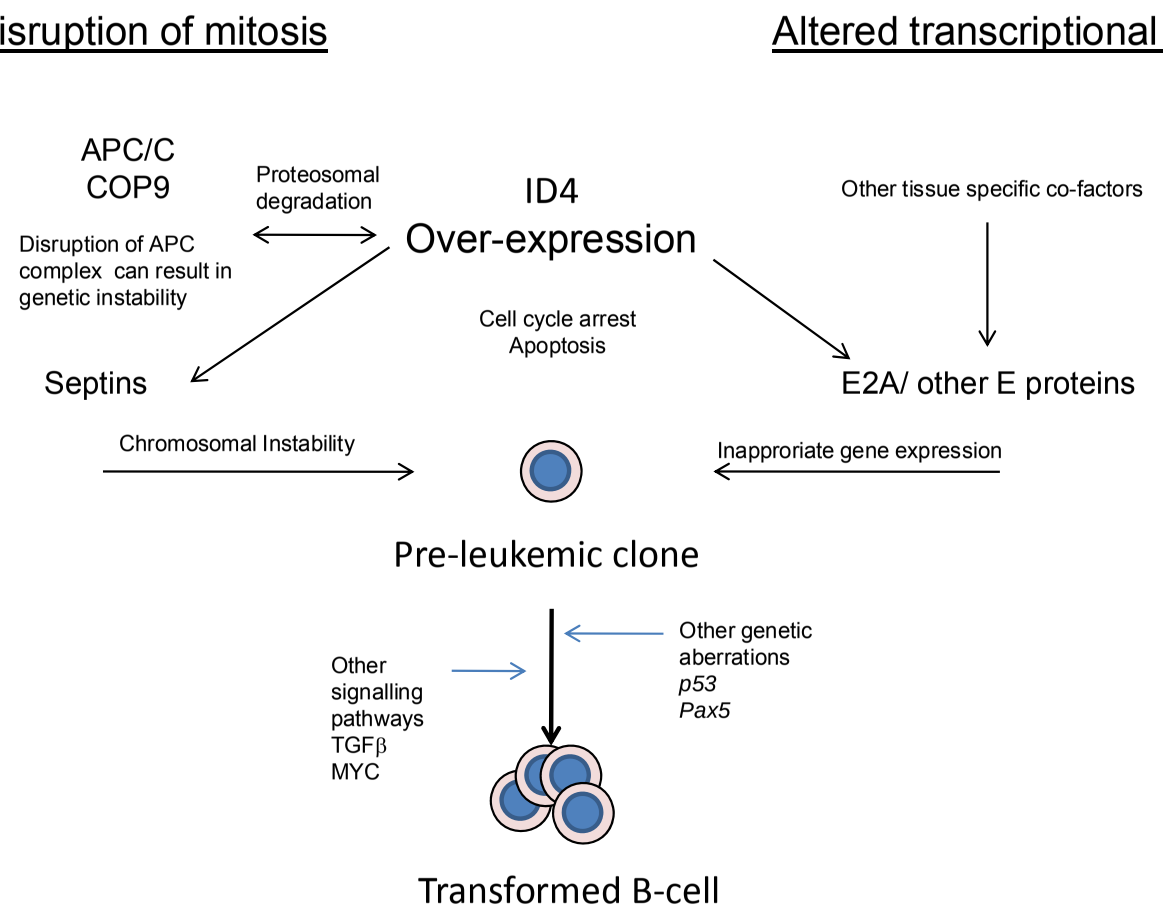


Figure 7.1 Proposed pathway by which ID4 induces transformation. ID4 expression in B-cells can target chromosomal instability and induce inappropriate gene expression profiles. ID4 is involved in the mitotic pathway via its interaction with components of the APC/C complex and septins. ID4 can also interact with E proteins and alter their transcriptional activity. All this generates a pre-leukemic clone which is primed for further leukemogenic events.

Appendix

Table A 1.1 PCR primers sequences and relative details.

Primer	Sequence(5'to 3')	Expected product size (bp)	If genomic DNA present product size	Tm (°C)	Designed
Beta Actin F *	tcatgaagtgtgacgttgacatccgt	285	511	55	Promega (catalogue G5740)
Beta Actin R*	cttagaagcatttgcggtgcacgatg	285	511	55	Promega (catalogue G5740)
ID4 F*	tccgaagggagtgactagga	152	N/A	49	Dr A Majid
ID4 R*	cccgagcccaacaatgac	152	N/A	47	Dr A Majid
ID1 F*	cctcaacggcgagatcag	190 (1B 420)	408	60.5	P.Dusanjh
ID1 R*	atcggtcttgttctccctca	190 (1B 420)	408	59.7	P.Dusanjh
ID2 F*	cctcaacacggatatcagca	198	1302	60	P.Dusanjh
ID2R*	cctccttgtgaaatggttgaa	198	1302	59.6	P.Dusanjh
ID3F*	actcagcttagccaggtgga	167	268	60	P. Dusanjh
ID3R*	aagctccttttgtcgttgga	167	268	59.9	P.Dusanjh

Table A1.2 A β Actin PCR with Taq polymerase.

Step/ cycles		Temperature	Time
Initial denaturation		94°C	5 minutes
40 cycles	Denaturation	94°C	30 seconds
	Annealing	65°C	1 minute
	Extension	68°C	2 minutes
Final extension		68°C	7 minutes

Table A1.3 ID1, ID2 and ID3 PCR with Taq polymerase.

Step/ cycles		Temperature	Time
Initial denaturation		94°C	5 minutes
30cycles	Denaturation	94°C	30 seconds
	Annealing	59°C	30 seconds
	Extension	72°C	45 seconds
Final extension		72°C	5 minutes

Table A 1.4 ID4 PCR with Taq polymerase.

Step/ cycles		**Temperature**	**Time**
Initial denaturation		94°C	5 minutes
30cycles	Denaturation	94°C	30 seconds
	Annealing	54°C	30 seconds
	Extension	72°C	30 seconds
Final extension		72°C	5 minutes

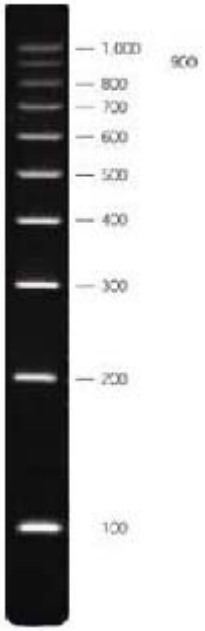
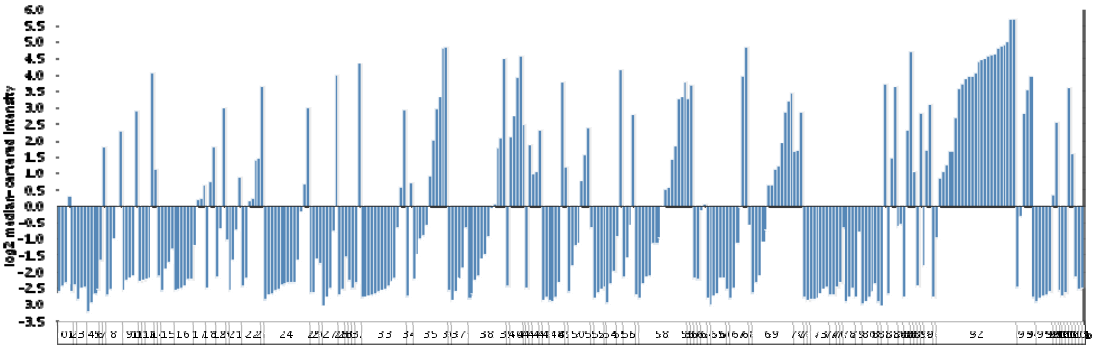


Figure A 1.1 DNA Super Ladder 1000bp (AbGene SLL-100L)

Figure A 1.2. CLUSTAL 2.0.10 multiple sequence alignment of mRNA transcripts for *ID1, ID2, ID3* and *ID4* (see Table 1.4 for accession numbers). The ORF is highlighted in blue and the HLH domain is highlighted in yellow.*Denotes sequence homology.

```
ID1a            CACCCTCAACGGCGA----GATCAGCGCCCTGACGGCCGAGG----------------- 526
ID1b            CACCCTCAACGGCGA----GATCAGCGCCCTGACGGCCGAGGTGAGATCCAGATCCGACC 544
ID4             CCGCGCCGCCGCGGACCCCGCTCACTGCGCTCAACACCGACCCGGCCGGCGCGGTGAACA 826
ID3             C----TCACTCCGGA----ACTTGTCATCTCCAACGACAAAAGGAGCT------TTTGCC 760
ID2             CACCCTCAACACGGA----TATCAGCATCCTGTCCTTGCAGGCTTCTGAATTCCCTTCTG 550
                *    *      **      *                  *

ID1a            ------------------------------------------------------------
ID1b            ACTAGATCATCCTTATACCGACGGGGAAACGGAGGCCAGAGAG-GGCGTGGGCGCTTGCA 603
ID4             AGCAGGGCGACAGCATTCTG---------TGCCGCTGAGCCGC-GCTGTCCAGGTGTGCG 876
ID3             ACTGACTCGGCCGTGTCCTG-------------ACACCTCCAG-AACGCAGGTGCTGGCG 806
ID2             AGTTAATGTCAAATGACAGCAAAGCACTGTGTGGCTGAATAAGCGGTGTTCATGATTTCT 610

ID1a            ------------------------------------------------------------
ID1b            CCACTTCCGTCCCATCCTTGCGGGTACCTGGCTATGCGGGGGTGCCTAAGGAGCCTGGAA 663
ID4             GCCGCCTGAGCCCGAGCCAGGAGCACTAGAGAGGGAGGGGGAAGAGCAGAAGTTAGAGAA 936
ID3             CCCGTTCTGCCTGGGACCCCGGGAACC------------------------------- 833
ID2             TTTATTCTTTGCACAACAACAACAACAACAAATTCACGGAATCTTTTAAGTGCTGAACTT 670

ID1a            ------------------------------------------------------------
ID1b            AAAGCGCTCCCCCGTCG--TGCTTCCTGGGGAAGGGGGCGTTCGCTGCGCTCGGAGCGGC 721
ID4             AAAAAGCCACCGGAGGA--AAGGAAAAAACATCGGCCAACCTAGAAACGTTTTCATTCGT 994
ID3             ------------------------------------------------------------
ID2             ATTTTTCAACCATTTCACAAGGAGGACAAGTTGAATGGACCTTTTTAAAAAGAAAAAAAA 730

ID1a            ---------------------------------------------CGGCATGCGTTCCTGC 542
ID1b            GTCCCTTCCAACCCGCCGGTCTCATTTCTTCTCGTTTTCACAGGCGGCATGCGTTCCTGC 781
ID4             CATTCCAAGAGAGAGAGGAAAGAAAAATACAACTTTCATTCTTTCTTTGCACGTTCAT 1054
ID3             ---------------------------------------------TCTCCTGCCGGAAGCC 849
ID2             AATGGAAGGAAAACTAAGAATGATCATCTTCCCAGGGTGTTCTCTTACTTGGACTGTGAT 790
                                                              **

ID1a            GGACGATC-GCATCTTGTGTCGCTGAAGCGCCTCCCCCAGGGACCGGCGGACCCCAGCCA 601
ID1b            GGACGATC-GCATCTTGTGTCGCTGAAGCGCCTCCCCCAGGGACCGGCGGACCCCAGCCA 840
ID4             AAACATTCTACATACGTATTCTCTTTTGTCTCTTCATTTATAACTGCTGTGAATTGTACA 1114
ID3             GGACGGCAGGGATGGGCCCCAACTTCGCCCTGCCCACTTGACTTCACCAAATCC----CT 905
ID2             ATTCGTTATTTATGAAAAAGACTTTTAAATGCCCTTTCTGCAGTTGGAAGGTTTTCTTTA 850
                   *       **          *

ID1a            TCCAGGGGGCAAGAGGAATTACGT-GCTCTGTGGGTCTCCCCCAACGC-GCCTCGCCGGA 659
ID1b            TCCAGGGGGCAAGAGGAATTACGT-GCTCTGTGGGTCTCCCCCAACGC-GCCTCGCCGGA 898
ID4             TTTCTGTGTTTTTTGGAGGTGCA--GTTAAACTTTTAAGCTTAAGTGT-GACAGGACTGA 1171
ID3             TCCTGGAGACTAAACCTGGTGCTCAGGAGCGAAGGACTGTGAACTTGT-GGCCTGAAGAG 964
ID2             TATACTATTCCCACCATGGGGAGCGAAAACGTTAAAATCACAAGGAATTGCCCAATCTAA 910
                *                                              * *

ID1a            TCTGAGGGAGAACAAGA------CCGATC---GGCGGCCAC-----TGCGCCCTTAACTG 705
ID1b            TCTGAGGGAGAACAAGA------CCGATC---GGCGGCCAC-----TGCGCCCTTAACTG 944
ID4             TAAATAGAAGATCAAGAGTAGATCCGACTTTAGAAGCCTAC-----TTTGTGACCAAGGA 1226
ID3             CCAGAGCTAGCTCTGGC----CACCAGCTGGGCGACGTCAC-----CCTGCTCCCACCCC 1015
ID2             GCAGACTTTGCCTTTTTTCAAAGGTGGAGCGTGAATACCAGAAGGATCCAGTATTCAGTC 970
                        *                           *

ID1a            CATCCAGCCTGGGGCTGAGGCT-----GAGGCACTGGCGAGGAGAGGGCGCTCCTCTCTG 760
ID1b            CATCCAGCCTGGGGCTGAGGCT-----GAGGCACTGGCGAGGAGAGGGCGCTCCTCTCTG 999
ID4             GCTCAATTTTTGTTTTGAAGCTTTACTAATCTACCAGAGCATTGTAGATATTTTTTTTT- 1285
ID3             ACCC---CCAAGTTCTAAGGTCTCTTCAGAGCGTGGAGGTGTGGAAGGAGTGGCTGCTCT 1072
ID2             ACTTAAATGAAGTCTTTTGGTC---AGAAATTACCTTTTTGACACAAGCCTACTGAATGC 1027
                           *   *   *

ID1a            CACACCTACTAGTCACCAGAGAC---TTTAGGGGGTGGGA------------TTCCACTC 805
ID1b            CACACCTACTAGTCACCAGAGAC---TTTAGGGGGTGGGA------------TTCCACTC 1044
ID4             TACATCTATTGTTTAAAATAGATGATTATAACGGGGCAGAGAACTTTCTTTTCTCTGCAA 1345
ID3             CCAAACTATGCCAAGGCGGCGGCAGAGCTGGTCTTCTGGT-----------CTCCTTGGA 1121
ID2             TGTGTATATATTTATATATAAATATATCTATTTGAGTGAAACCTTGTGAACTCTTTAATT 1087
                      **                    *

ID1a            GTGTGTTTC-TATTTTTTGA-A-AAGCA---GACATTTTAAAAAATGGTCACGTTTGGTG 859
ID1b            GTGTGTTTC-TATTTTTTGA-A-AAGCA---GACATTTTAAAAAATGGTCACGTTTGGTG 1098
ID4             GAATGTTACATATTGTATAG-ATAAATGAGTGACATTTCATACCATGTATATATAGAGAT 1404
ID3             GAAAGGTTC-TGTTGCCCTG-ATTTATGAACTCTATAATAGAGTATATAGGTTTTGTACC 1179
ID2             AGAGTTTTCTTGTATAGTGGCAGAGATG---TCTATTTCTGCATTCAAAAGTGTAATGAT 1144
                      * * * *         *           **              *

ID1a            CTTCTCAGATTTCTGAGGAAAT-TG---CTTTGTA-TTGTATA-TTACAATGATCACCGA 913
ID1b            CTTCTCAGATTTCTGAGGAAAT-TG---CTTTGTA-TTGTATA-TTACAATGATCACCGA 1152
ID4             GTTCT-ATAAGTGTGAGAAAGTATA---TGCTTTA-ATAGATA-CTGTAATTATAAGATA 1458
ID3             TTTTTTACAGGAAGGTGACTTTCTGTAACAATGCG-ATGTATA-TTAAACTTTTTATAAA 1237
ID2             GTACTTATTCATGCTAAACTTTTTATAAAAGTTTAGTTGTAAACTTAACCCTTTTATACA 1204
                 *  * *              * *       *     *  * *  *       * *   *

ID1a            CTGA-----AAATATTGTTTTACAATAGTTCTGTGGGGCTGTTTTTTTGTTATTAAACAA 968
ID1b            CTGA-----AAATATTGTTTTACAATAGTTCTGTGGGGCTGTTTTTTTGTTATTAAACAA 1207
ID4             TTTTTAATTAAATATTTTTTTGTAAATATTATGTGTGTGTTTTTTTTTAAT-CTATGGGA 1517
ID3             AGTTA----ACATTTTGCATAATAAACGATTTTTAAACACTTGTGTATATGATGA----- 1288
ID2             AAATAAATCAAGTGT-GTTTATTGAATGGTGATTGCCTGCTTTATTTCAGAGGACCAGTG 1263
                        *  * *    *    *    *   *       *   *

ID1a            ATAATTTAGATGGTGAAAAAAAAAA---------------------------------- 993
ID1b            ATAATTTAGATGGTGAAAAAAAAAAAA------------------------------- 1234
ID4             ATATTTCTTTTGGAAAATCATTTTTCAGCTCAATTACAGAGCTCTTGATATCTTGAATGT 1577
ID3             ------------------------------------------------------------
ID2             CTTTGATTTTTATTATGCTATGTTATAACTGAACCCAAATAAATACAAGTTCAAATTTAT 1323
```

```
ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             CTTTTCTGTTTGGCCTGGCTCTTAATTTGCTTTTGTTTTGCCCAGTATAGACTCGGAAGT 1637
ID3             ------------------------------------------------------------
ID2             GTAGACTGTATAAGATTATAATAAAACATGTCTGAAGTCAAAAAAAAAAAAAAAAAAAAAA 1383


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             AACAGTTATAGCTAGTGGTCTTGCATGATTGCATGAGATGTTTAATCACAAATTAAACTT 1697
ID3             ------------------------------------------------------------
ID2             AAAAAAAAAAAAAAAAAA------------------------------------------ 1402


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             GTTCTGAGTCCATTCAAATGTGTTTTTTTAAATGTAGATTGAAATCTTTGTATTTGAAGC 1757
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             ATACATGTTGAAAATACACCTTATCAGTTTTTAAGTACAGGGTTTTATAGTGTAATATAT 1817
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             ACAGAGTAAGTGTTTGTTTTTGTTTTTCAACTGAGGTCAAAATGGATTCTGAATGATTTT 1877
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             GCATATGGGATGAGGAAATGCTTGGATCCTTAAGGAGTTTACGAAATCTGCTGTTTTATC 1937
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             AAAGTGAAAAAAAAATTGCTTATTACTCTTCATTTTACACTAAAGCTTAATGTCACTAAGT 1997
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             TTCATGTCTGTACAGATTATTTAAATCATGGAAATGAAAAAAATGTTCTCTGCTTGCTAC 2057
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             CAAAGGACAAACTCTTGGAAATGAACACTTTCTGCTTTCCTTCCTCCAAAGAATTAATAG 2117
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             GCAACAGTGGGAGAAAAAAAAGGCATAATGGCAAATCCTTCAAGCAGGGATAAAAGTCGA 2177
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             TCTTCAAACATTAACTTAAGCAGACCAAAAATTCTGATGACCGCATCTAGATTATTTTTT 2237
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             TATAAAAATGATTTTCACTATAGCTATGTTACGCTAAGCTACTGTCCCATCTCTTGTGAT 2297
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            ------------------------------------------------------------
ID1b            ------------------------------------------------------------
ID4             GTGTAACTTTTACATGTGAATATTAAAGTAGATTTCTCTGTCTTGTAAAAAAAAAAAAAA 2357
ID3             ------------------------------------------------------------
ID2             ------------------------------------------------------------


ID1a            --------------------------------
ID1b            --------------------------------
ID4             AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2389
ID3             --------------------------------
ID2             --------------------------------
```



0. No value (4)	26. Cecum Carcinoma (1)	53. Hypopharyngeal Squamous Cell Carcinoma (1)	80. Primary Effusion Lymphoma (5)
1. Acute Biphenotypic Leukaemia (1)	27. Cervical Cancer (5)	54. Immunoblastic Lymphoma (3)	81. Prostate Adenocarcinoma (1)
2. Acute Monocytic Leukaemia (1)	28. Cervical Squamous Cell Carcinoma (2)	55. Invasive Ductal Breast Carcinoma (2)	82. Prostate Carcinoma (2)
3. Acute Myeloid Leukaemia (3)	29. Chondrosarcoma (1)	56. Large Cell Lung Carcinoma (4)	83. Rectal Adenocarcinoma (3)
4. Acute Myelomonocytic Leukaemia (3)	30. Chronic Lymphocytic Leukaemia (1)	57. Liposarcoma (1)	84. Renal Carcinoma (2)
5. Acute Promyelocytic Leukaemia (1)	31. Chronic Myelogenous Leukaemia (1)	58. Lung Adenocarcinoma (15)	85. Renal Cell Carcinoma (2)
6. Adrenal Cortex Carcinoma (1)	32. Clear Cell Renal Cell Carcinoma (2)	59. Lung Atypical Carcinoid Tumour (1)	86. Renal Leiomyoma (1)
7. Amelanotic Skin Melanoma (1)	33. Colon Adenocarcinoma (14)	60. Lung Carcinoid Tumour (1)	87. Renal Wilms Tumour (1)
8. B-Cell Acute Lymphoblastic Leukaemia (5)	34. Colon Carcinoma (2)	61. Lung Cancer (1)	88. Retinoblastoma (1)
9. B-Cell Non-Hodgkin's Lymphoma (5)	35. Cutaneous Melanoma (11)	62. Mantle Cell Lymphoma (1)	89. Rhabdoid Tumour of the Kidney (1)
10. B-Cell Prolymphocytic Leukaemia (1)	36. Cutaneous T-Cell Non-Hodgkin's Lymphoma (1)	63. Medulloblastoma (2)	90. Rhabdomyosarcoma (3)
11. Barrett's Adenocarcinoma (1)	37. Diffuse Large B-Cell Lymphoma (5)	64. Mesothelioma (1)	91. Sezary Syndrome (1)
12. Bladder Cancer (3)	38. Ductal Breast Carcinoma (12)	65. Multiple Myeloma (5)	92. Small Cell Lung Carcinoma (25)
13. Bladder Papillary Urothelial Carcinoma (1)	39. Embryonal Rhabdomyosarcoma (1)	66. Mycosis Fungoides (1)	93. Squamous Cell Lung Carcinoma (5)
14. Bladder Squamous Cell Carcinoma (1)	40. Endometrial Adenocarcinoma (4)	67. Neuroblastoma (6)	94. Synovial Sarcoma (1)
15. Bladder Urothelial Carcinoma (4)	41. Endometrial Carcinoma (1)	68. Neuroglial Tumour, NOS (1)	95. T-Cell Acute Lymphoblastic Leukaemia (5)
16. Blast Phase Chronic Myelogenous Leukaemia (6)	42. Erythroleukemia (1)	69. Non-Small Cell Lung Carcinoma (13)	96. Thyroid Gland Follicular Carcinoma (1)
17. Bone Osteosarcoma (4)	43. Esophageal Adenocarcinoma (1)	70. Ovarian Adenocarcinoma (3)	97. Thyroid Gland Medullary Carcinoma (1)
18. Brain Astrocytoma (3)	44. Esophageal Squamous Cell Carcinoma (2)	71. Ovarian Clear Cell Adenocarcinoma (1)	98. Thyroid Gland Papillary Carcinoma (1)
19. Brain Glioblastoma (3)	45. Ewing's Sarcoma of Bone (1)	72. Ovarian Serous Surface Papillary Carcinoma (1)	99. Thyroid Gland Undifferentiated (Anaplastic) Carcinoma (1)
20. Breast Adenocarcinoma (1)	46. Fibrosarcoma (2)	73. Pancreatic Adenocarcinoma (6)	100. Ureter Urothelial Carcinoma (1)
21. Breast Carcinoma (4)	47. Gastric Adenocarcinoma (1)	74. Pancreatic Carcinoma (2)	101. Uterine Corpus Leiomyosarcoma (1)
22. Bronchioloalveolar Carcinoma (6)	48. Gastric Cancer (4)	75. Pancreatic Ductal Adenocarcinoma (1)	102. Uterine Corpus Sarcoma (1)
23. Bronchogenic Carcinoma (1)	49. Glioblastoma (1)	76. Papillary Lung Adenocarcinoma (1)	103. Vulvar Carcinoma (1)
24. Burkitt's Lymphoma (14)	50. Hepatocellular Carcinoma (7)	77. Pharyngeal Carcinoma (1)	104. Vulvar Leiomyosarcoma (1)
25. Cecum Adenocarcinoma (3)	51. Histiocytoma (1)	78. Placental Choriocarcinoma (3)	105. Vulvar Squamous Cell Carcinoma (1)
	52. Hodgkin's Lymphoma (3)	79. Pleomorphic Hepatocellular Carcinoma (2)	

Figure A1.3. Cell lines which over-express ID4 in the Wooster Cell Line 2 collection. Data retrieved from www.oncomine.org using the cancer verses cancer analysis. p=0.0001. Using the Human Genome U133 Plus 2.0Array.

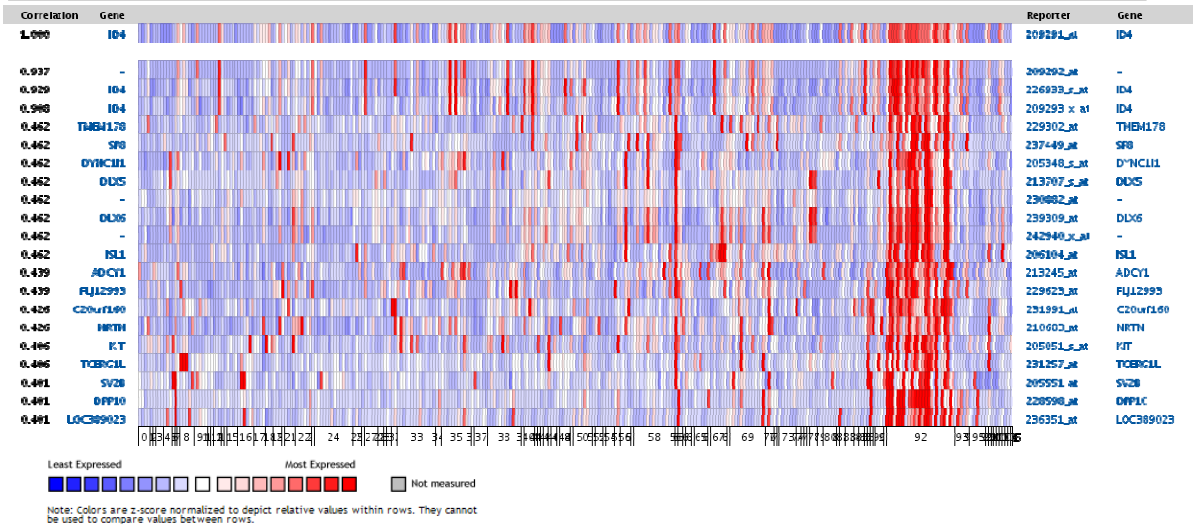
Table A1.5 Cell lines which express ID4 in the Wooster Cell Line collection. Data retrieved from www.oncomine.org using the cancer verses cancer analysis. p=0.0001.

Cell line	Malignancy	Normalised expression units
WI-38	Foetal lung	0.1552
C32TG	Amelanotic skin melanoma	0.933
MHH-PREB-1	B-cell NHL	1.495
HT-1376	Bladder cancer	1.992
MG-63	Bone oseteosarcoma	0.099
HOS		0.113
SJSA-1		0.313
CCF-STTG1	Brain Astrocytoma	0.399
SW 1088		0.901
SNB-19	Brain Glioblastoma	1.498
DU-4475	Breast carcinoma	0.420
NCI-H1666	Bronchioalveolar Carcinoma	0.072
SW 1573		0.117
NCI-H1659		0.703
NCI-H358		0.727
ChaGo-K-1	Bronchiogenic Carcinoma	1.847
CA46	Burkitts Lymphoma	0.351
DG-75		0.155
C-33 A	Cervical Cancer	2.0.14
LS 174T	Colon Adenocarcinoma	0.29
SW1417		1.476
T84	Colon Carcinoma	0.362
Caki-2	Clear Cell Renal Cell Carcinoma	2.201
COLO 829	Cutaneous Melanoma	0.460
A7		1.023
SH-4		1.521
SK-MEL-1		1.696
HMCB		2.425
CHL-1		2.522
HCC1954	Ductal Breast Carcinoma	0.892
HCC1143		1.037
HCC38		2.277

HEC-1-A	Endometrial Adenocarcinoma	1.101
HEC-1-B		1.410
AN3 CA		2.089
KLE		2.414
RL95-2		1.317
OE19	Oesophageal Adenocarcinoma	0.979
OE21	Oesophageal Squamous Cell Carcinoma	0.479
KYSE30		0.536
RD-ES	Ewings Sarcoma of Bone	1.146
CGTH-W-1	Thyroid Gland Follicular Carcinoma	0.183
SNU-1	Gastric Carcinoma	1.995
DK-MG	Glioblastoma	0.669
SNU-183	Hepatocellular Carcinoma	0.402
Hep 3B2.1-7		0.783
SNU-398		1.436
HCC 1599	Invasive Ductal Breast Carcinoma	2.065
NCI-H661	Large Cell Lung Carcinoma	1.513
NCI-H2228	Lung Adenocarcinoma	0.263
NCI-H2009		0.275
NCI-H1623		0.721
NCI-H1395		0.969
NCI-H1573		1.669
NCI-H1355		1.696
NCI-H522		1.995
NCI-H720	Atypical Lung Carcinoid Tumour	1.599
UMC-11	Lung Carcinoid Tumour	1.895
TT	Thyroid Glad Medullary Carcinoma	1.258
SK-N-F1	Neuroblastoma	2.017
MC-1XC		2.426
NCI-H1155	Non-Small Cell Lung Carcinoma	0.313
NCI-H1651		0.330
NCI-H1770		0.583
NCI-H1793		0.647
NCI-H1975		0.951
NCI-H810		1.421

NCI-H212		1.557
NCI-H2030		1.691
Caov-3	Ovarian Adenocarcinoma	0.810
COLO-704		0.872
SK-OV-3		1.426
BFTC-905	Bladder Papillary Urothelial Carcinoma	0.556
NALM-6	BCP-ALL	1.183
22Rv1	Prostate Carcinoma	1.986
NCI-H630	Rectal Adenocarcinoma	0.759
SW837		1.789
769-P	Renal cell Carcinoma	1.153
G-402	Renal Leiomyoma	2.360
SK-NEP-1	Renal Wilms Tumour	0.554
G-401	Rhabdoid Tumour of the Kidney	1.406
A-204	Rhabdomysarcoma	0.867
A-673		1.676
NCI-H2107	Small cell lung carcinoma	0.506
COLO 668		0.513
NCI-H82		0.622
NCI-H524		0.841
DMS 53		0.869
NCI-H446		1.346
NCI-H2081		1.790
NCI-H694		1.81
NCI-H1930		1.971
DMS 273		2.0
NCI-H187		2.083
NCI-H69		2.098
NCI-H1048		2.184
NCI-H1436		2.202
COR-L279		2.249
COR-L88		2.282
DMS 153		2.337
NCI-H1092		2.413

NCI-H1618		2.440
DMS 79		2.442
NCI-H2171		2.461
DMS 114		2.516
NCI-H2195		2.801
NCI-H748		2.804
NCI-H2170	Squamous Cell Lung Carcinoma	1.462
NCI-H226		1.818
SW 900		1.950
SK-UT-1	Uterine Corpus Leiomyosarcoma	1.813
MES-SA	Uterine Corpus Sarcoma	0.833



0. No value (4)	26. Cecum Carcinoma (1)	53. Hypopharyngeal Squamous Cell Carcinoma (1)	80. Primary Effusion Lymphoma (5)
1. Acute Biphenotypic Leukaemia (1)	27. Cervical Cancer (5)	54. Immunoblastic Lymphoma (3)	81. Prostate Adenocarcinoma (1)
2. Acute Monocytic Leukaemia (1)	28. Cervical Squamous Cell Carcinoma (2)	55. Invasive Ductal Breast Carcinoma (2)	82. Prostate Carcinoma (2)
3. Acute Myeloid Leukaemia (3)	29. Chondrosarcoma (1)	56. Large Cell Lung Carcinoma (4)	83. Rectal Adenocarcinoma (3)
4. Acute Myelomonocytic Leukaemia (3)	30. Chronic Lymphocytic Leukaemia (1)	57. Liposarcoma (1)	84. Renal Carcinoma (2)
5. Acute Promyelocytic Leukaemia (1)	31. Chronic Myelogenous Leukaemia (1)	58. Lung Adenocarcinoma (15)	85. Renal Cell Carcinoma (2)
6. Adrenal Cortex Carcinoma (1)	32. Clear Cell Renal Cell Carcinoma (2)	59. Lung Atypical Carcinoid Tumour (1)	86. Renal Leiomyoma (1)
7. Amelanotic Skin Melanoma (1)	33. Colon Adenocarcinoma (14)	60. Lung Carcinoid Tumour (1)	87. Renal Wilms Tumour (1)
8. B-Cell Acute Lymphoblastic Leukaemia (5)	34. Colon Carcinoma (2)	61. Lung Cancer (1)	88. Retinoblastoma (1)
9. B-Cell Non-Hodgkin's Lymphoma (5)	35. Cutaneous Melanoma (11)	62. Mantle Cell Lymphoma (1)	89. Rhabdoid Tumour of the Kidney (1)
10. B-Cell Prolymphocytic Leukaemia (1)	36. Cutaneous T-Cell Non-Hodgkin's Lymphoma (1)	63. Medulloblastoma (2)	90. Rhabdomyosarcoma (3)
11. Barrett's Adenocarcinoma (1)	37. Diffuse Large B-Cell Lymphoma (5)	64. Mesothelioma (1)	91. Sezary Syndrome (1)
12. Bladder Cancer (3)	38. Ductal Breast Carcinoma (12)	65. Multiple Myeloma (5)	92. Small Cell Lung Carcinoma (25)
13. Bladder Papillary Urothelial Carcinoma (1)	39. Embryonal Rhabdomyosarcoma (1)	66. Mycosis Fungoides (1)	93. Squamous Cell Lung Carcinoma (5)
14. Bladder Squamous Cell Carcinoma (1)	40. Endometrial Adenocarcinoma (4)	67. Neuroblastoma (6)	94. Synovial Sarcoma (1)
15. Bladder Urothelial Carcinoma (4)	41. Endometrial Carcinoma (1)	68. Neuroglial Tumour, NOS (1)	95. T-Cell Acute Lymphoblastic Leukaemia (5)
16. Blast Phase Chronic Myelogenous Leukaemia (6)	42. Erythroleukemia (1)	69. Non-Small Cell Lung Carcinoma (13)	96. Thyroid Gland Follicular Carcinoma (1)
17. Bone Osteosarcoma (4)	43. Esophageal Adenocarcinoma (1)	70. Ovarian Adenocarcinoma (3)	97. Thyroid Gland Medullary Carcinoma (1)
18. Brain Astrocytoma (3)	44. Esophageal Squamous Cell Carcinoma (2)	71. Ovarian Clear Cell Adenocarcinoma (1)	98. Thyroid Gland Papillary Carcinoma (1)
19. Brain Glioblastoma (3)	45. Ewing's Sarcoma of Bone (1)	72. Ovarian Serous Surface Papillary Carcinoma (1)	99. Thyroid Gland Undifferentiated (Anaplastic) Carcinoma (1)
20. Breast Adenocarcinoma (1)	46. Fibrosarcoma (2)	73. Pancreatic Adenocarcinoma (6)	100. Ureter Urothelial Carcinoma (1)
21. Breast Carcinoma (4)	47. Gastric Adenocarcinoma (1)	74. Pancreatic Carcinoma (2)	101. Uterine Corpus Leiomyosarcoma (1)
22. Bronchioloalveolar Carcinoma (6)	48. Gastric Cancer (4)	75. Pancreatic Ductal Adenocarcinoma (1)	102. Uterine Corpus Sarcoma (1)
23. Bronchogenic Carcinoma (1)	49. Glioblastoma (1)	76. Papillary Lung Adenocarcinoma (1)	103. Vulvar Carcinoma (1)
24. Burkitt's Lymphoma (14)	50. Hepatocellular Carcinoma (7)	77. Pharyngeal Carcinoma (1)	104. Vulvar Leiomyosarcoma (1)
25. Cecum Adenocarcinoma (3)	51. Histiocytoma (1)	78. Placental Choriocarcinoma (3)	105. Vulvar Squamous Cell Carcinoma (1)
	52. Hodgkin's Lymphoma (3)	79. Pleomorphic Hepatocellular Carcinoma (2)	

Figure A1.4. Gene expression profiles with ID4 in the Wooster cell line panel. Data retrieved from www.oncomine.org using the cancer verses cancer analysis. p=0.0001. Using the Human Genome U133 Plus 2.0Array.

Table A 1.6 The cytogenetic data for the 4 haematological cell lines identified from the Wooster Cell Line panel. This data was obtained using the DSMZ, the German resource centre for biological material (http://www.dsmz.de/).

Cell line	Cytogenetics	Common cytogenetic feature.
NALM-6	Human near diploid karyotype - 46(43-47)<2n>XY, t(5;12)(q33.2;p13.2)	*IL3/ETV6*
MHH-PREB-1	Human hyperdiploid karyotype with 12% polyploidy - 48<2n>XYY, +21, ?ins(2;2)(q21;p15p22), t(8;14)(q24;q32)	*MYC/IGH*
CA46	Human near diploid karyotype with about 3% triploidy - 46(45-48)<2n>X/XY, dup(1)(q21q32), dup(7)(q12q22), t(8;14)(q24;q32) - additional rearrangements were present in a sideline, viz. t(6;13)(p21;q32) and del(11)(p11) - a supernumerary dmin was present in about 25% metaphases	*MYC/IGH* *Cyclin D3/IGH*
DG-75	Human near diploid karyotype with 6% polyploidy - 46<2n>XY, t(8;14)(q24;q32)	*MYC/IGH*
RCH-ACV	Human flat-moded hyperdiploid karyotype with 2% polyploidy - 43-50<2n>XX, +8, t(1;19)(q23;p13.3) - sideline with idem, +16, +der(19)t(1;19) - carries t(1;19) effecting E2A-PBX1 fusion - matches published karyotype	*E2A/PBX1*
Ba/F3	Murine near diploid karyotype with 33% polyploidy - 40(36-42)<2n>	

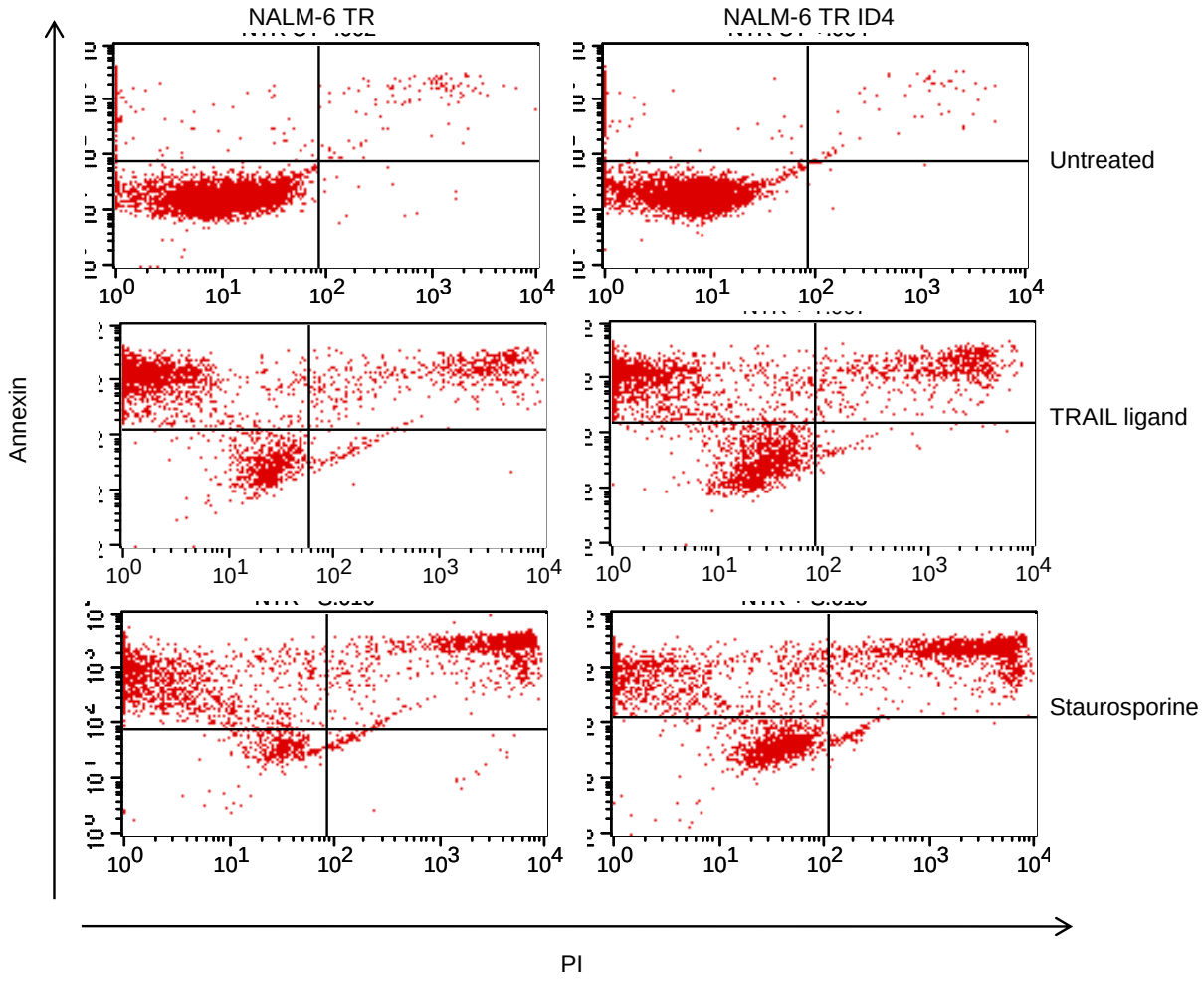
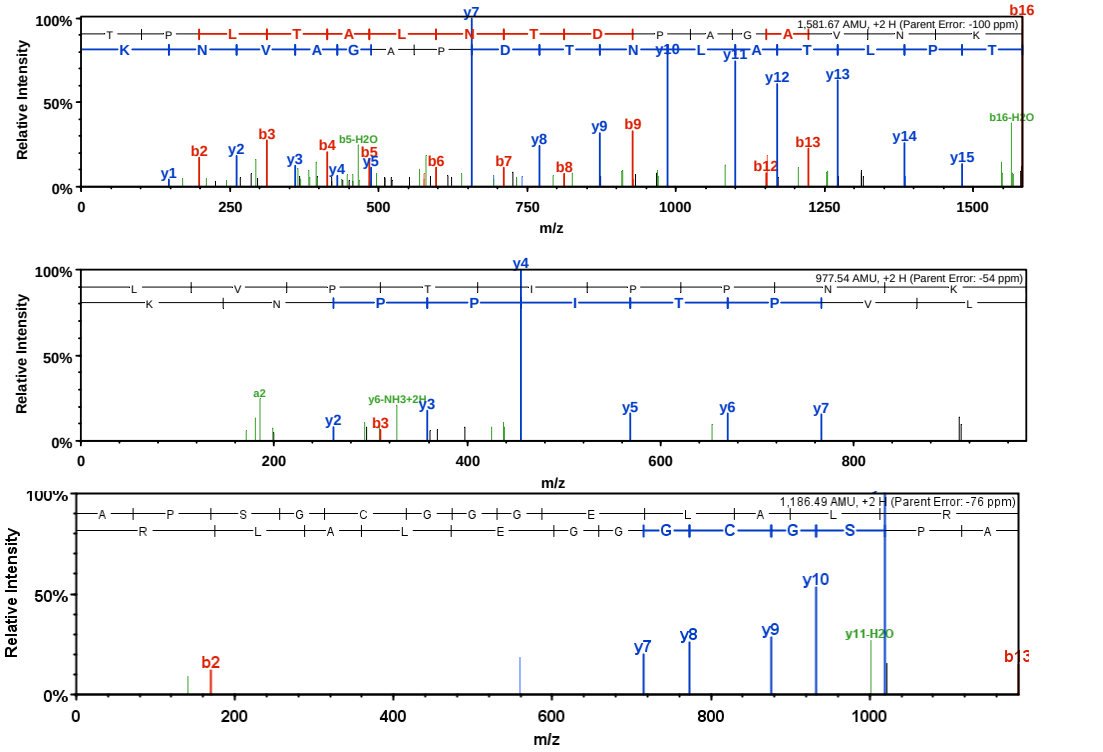


Figure A1.5 Representative flow cytometry scatter plots for the annexin PI analysis of apoptosis induction in NALM6 TR cells with and without ID4 induction (post 24 hours). Cells were treated with 500ng/ml of TRAIL ligand, or 1mM Stauorsporine for four hours.



ID4_MOUSE (100%), 16,596.5 Da
DNA-binding protein inhibitor ID-4 (Inhibitor of DNA binding 4) - Mus musculus (Mouse)
3 unique peptides, 3 unique spectra, 3 total spectra, 38/161 amino acids (24% coverage)

MKAVSPVRPS GRKAPSGCGG GELALRCLAE HGHSLGGSAA AAAAAAAARC
KAAEAAADEP ALCLQCDMND CYSRLRRLVP TIPPNKKVSK VEILQHVIDY
ILDLQLALET HPALLRQPPP PAPPLHPAGA CPVAPPRTPL TALNTDPAGA
VNKQGDSILC R

Figure A 1.6 Mass spectral analysis of captured ID4. 3 peptides were identified with the relevant spectral read out. Analysis was done using peptide thresholds at 90% minimum, protein thresholds at 90% minimum and 1 peptide minimum, using the Scaffold 2 proteomics software. The peptides were then searched on http://blast.ncbi.nlm.nih.gov/Blast.cgi confirming that the recognised peptides belong to ID4.

Table A1.7 Mass spectral analysis of the peptides recognised for ID4 in Figure A1.6

	Molecular weight	Peptides	Peptide	Mascot score	M/z	Charge
DNA-binding protein inhibitor ID-4	17 kDa	3	(R)LVPTIPPNK(K)	41.95	489.8	2
			(R)TPLTALNTDPAGAVNK(Q)	111	791.8	2
			(K)APSGCGGGELALR(C)	34.3	594.25	2

Table A1.8. Total list of proteins from an IP performed in Ba/F3 clones ID4 11 and EV 14 Using the anti myc tag antibody 9B11.

Identified Proteins	Accession Number	Molecular Weight	Taxonomy	Peptides
Coatomer subunit delta (Delta-coat protein) (Delta-COP) (Archain) - Mus musculus (Mouse)	COPD_MOUSE	57 kDa	Mus musculus	9
Coronin-1A (Coronin-like protein p57) (Coronin-like protein A) (Clipin-A) (Tryptophan aspartate-containing coat protein) (TACO) - Mus musculus (Mouse)	COR1A_MOUSE	51 kDa	Mus musculus	8
Plastin-2 (L-plastin) (Lymphocyte cytosolic protein 1) (LCP-1) (65 kDa macrophage protein) (pp65) - Mus musculus (Mouse)	PLSL_MOUSE	70 kDa	Mus musculus	8
Annexin A2 (Annexin II) (Lipocortin II) (Calpactin I heavy chain) (Chromobindin-8) (p36) (Protein I) (Placental anticoagulant protein IV) (PAP-IV) - Mus musculus (Mouse)	ANXA2_MOUSE	39 kDa	Mus musculus	7
Importin beta-3 (Karyopherin beta-3) (Ran-binding protein 5) (RanBP5) - Mus musculus (Mouse)	IMB3_MOUSE	124 kDa	Mus musculus	6
14-3-3 protein gamma - Mus musculus (Mouse)	1433G_MOUSE	28 kDa	Mus musculus	5
ADP/ATP translocase 2 (Adenine nucleotide translocator 2) (ANT 2) (ADP,ATP carrier protein 2) (Solute carrier family 25 member 5) - Mus musculus (Mouse)	ADT2_MOUSE	33 kDa	Mus musculus	5
Glutathione transferase omega-1 (EC 2.5.1.18) (GSTO 1-1) (p28) - Mus musculus (Mouse)	GSTO1_MOUSE	27 kDa	Mus musculus	4
Actin-related protein 2/3 complex subunit 2 (ARP2/3 complex 34 kDa subunit) (p34-ARC) - Mus musculus (Mouse)	ARPC2_MOUSE	34 kDa	Mus musculus	4
Septin-2 (Protein NEDD5) (Neural precursor cell expressed developmentally down-regulated protein 5) - Mus musculus (Mouse)	SEPT2_MOUSE	42 kDa		4
Putative ATP-dependent RNA helicase Pl10 (EC 3.6.1.-) - Mus musculus (Mouse)	PL10_MOUSE	73 kDa	Mus musculus	4
Vesicle-fusing ATPase (EC 3.6.4.6) (Vesicular-fusion protein NSF) (N-ethylmaleimide sensitive fusion protein) (NEM-sensitive fusion protein) (Suppressor of K(+) transport growth defect 2) (Protein SKD2) - Mus musculus (Mouse)	NSF_MOUSE	83 kDa	Mus musculus	4
DNA ligase 3 (EC 6.5.1.1) (DNA ligase III) (Polydeoxyribonucleotide synthase [ATP] 3) - Mus musculus (Mouse)	DNL3_MOUSE	113 kDa	Mus musculus	3
Myosin-binding protein C, cardiac-type (Cardiac MyBP-C) (C-protein, cardiac muscle isoform) - Mus musculus (Mouse)	MYPC3_MOUSE	141 kDa	Mus musculus	3
DNA-binding protein inhibitor ID-4 (Inhibitor of DNA binding 4) - Mus musculus (Mouse)	ID4_MOUSE	17 kDa	Mus musculus	3
General transcription factor 3C polypeptide 1 (Transcription factor IIIC-subunit alpha) (TF3C-alpha) (TFIIIC 220 kDa subunit) (TFIIIC220) (TFIIIC box B-binding subunit) - Mus musculus (Mouse)	TF3C1_MOUSE	237 kDa	Mus musculus	3
Tumour protein D54 (Tumour protein D52-like 2) - Mus musculus (Mouse)	TPD54_MOUSE	24 kDa	Mus musculus	3
Elongation factor 1-beta (EF-1-beta) - Mus musculus (Mouse)	EF1B_MOUSE	25 kDa	Mus musculus	3
Proteasome activator complex subunit 2 (Proteasome activator 28-subunit beta) (PA28beta) (PA28b) (Activator of multicatalytic protease subunit 2) (11S regulator complex subunit beta) (REG-beta) - Mus musculus (Mouse)	PSME2_MOUSE	27 kDa	Mus musculus	3

Chloride intracellular channel protein 1 (Nuclear chloride ion channel 27) (NCC27) - Mus musculus (Mouse)	CLIC1_MOUSE	27 kDa	Mus musculus	3
14-3-3 protein eta - Mus musculus (Mouse)	1433F_MOUSE	28 kDa	Mus musculus	3
14-3-3 protein theta (14-3-3 protein tau) - Mus musculus (Mouse)	1433T_MOUSE	28 kDa	Mus musculus	3
Ras suppressor protein 1 (Rsu-1) (RSP-1) - Mus musculus (Mouse)	RSU1_MOUSE	32 kDa	Mus musculus	3
Macrophage-capping protein (Actin-regulatory protein CAP-G) (Myc basic motif homolog 1) (Actin-capping protein GCAP39) - Mus musculus (Mouse)	CAPG_MOUSE	39 kDa	Mus musculus	3
Septin-1 (Differentiation protein 6) (Diff6 protein) - Mus musculus (Mouse)	SEPT1_MOUSE	42 kDa	Mus musculus	3
Actin-like protein 2 (Actin-related protein 2) - Mus musculus (Mouse)	ARP2_MOUSE	45 kDa	Mus musculus	3
Septin-11 - Mus musculus (Mouse)	SEP11_MOUSE	50 kDa	Mus musculus	3
Heterogeneous nuclear ribonucleoprotein K - Mus musculus (Mouse)	HNRPK_MOUSE	51 kDa	Mus musculus	3
T-complex protein 1 subunit gamma (TCP-1-gamma) (CCT-gamma) (Matricin) (mTRiC-P5) - Mus musculus (Mouse)	TCPG_MOUSE	61 kDa	Mus musculus	3
Septin-9 (SL3-3 integration site 1 protein) - Mus musculus (Mouse)	SEPT9_MOUSE	66 kDa	Mus musculus	3
78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kDa protein 5) (Immunoglobulin heavy chain-binding protein) (BiP) - Mus musculus (Mouse)	GRP78_MOUSE	72 kDa	Mus musculus	3
Calpain-1 catalytic subunit (EC 3.4.22.52) (Calpain-1 large subunit) (Calcium-activated neutral proteinase 1) (CANP 1) (Calpain mu-type) (muCANP) (Micromolar-calpain) - Mus musculus (Mouse)	CAN1_MOUSE	82 kDa	Mus musculus	3
Islet amyloid polypeptide precursor (Amylin) (Diabetes-associated peptide) (DAP) - Mus musculus (Mouse)	IAPP_MOUSE	10 kDa	Mus musculus	2
Rho guanine nucleotide exchange factor 1 (Lymphoid blast crisis-like 2) (Lbc's second cousin) - Mus musculus (Mouse)	ARHG1_MOUSE	103 kDa	Mus musculus	2
ADAMTS-15 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 15) (ADAM-TS 15) (ADAM-TS15) - Mus musculus (Mouse)	ATS15_MOUSE	104 kDa	Mus musculus	2
Catenin alpha-2 (Alpha-catenin-related protein) (Alpha N-catenin) - Mus musculus (Mouse)	CTNA2_MOUSE	105 kDa	Mus musculus	2
Polycystin-2 (Polycystic kidney disease 2 protein homolog) - Mus musculus (Mouse)	PKD2_MOUSE	109 kDa	Mus musculus	2
N-acetyl-beta-glucosaminyl-glycoprotein 4-beta-N-acetylgalactosaminyltransferase 1 (EC 2.4.1.244) (NGalNAc-T1) (Beta-1,4-N-acetylgalactosaminyltransferase IV) (Beta4GalNAc-T4) (Beta4GalNAcT4) - Mus musculus (Mouse)	B4GN4_MOUSE	117 kDa	Mus musculus	2
JmjC domain-containing histone demethylation protein 3C (EC 1.14.11.-) (Jumonji domain-containing protein 2C) - Mus musculus (Mouse)	JHD3C_MOUSE	120 kDa	Mus musculus	2
Kinesin-like protein KIF1C - Mus musculus (Mouse)	KIF1C_MOUSE	122 kDa	Mus musculus	2
AF4/FMR2 family member 1 (Protein AF-4) (Proto-oncogene AF4) - Mus musculus (Mouse)	AFF1_MOUSE	132 kDa	Mus musculus	2
Treacle protein (Treacher Collins syndrome protein homolog) - Mus musculus (Mouse)	TCOF_MOUSE	135 kDa	Mus musculus	2

Misshapen-like kinase 1 (EC 2.7.11.1) (Mitogen-activated protein kinase kinase kinase kinase 6) (MAPK/ERK kinase kinase kinase 6) (MEK kinase kinase 6) (MEKKK 6) (Misshapen/NIK-related kinase) (GCK family kinase MiNK) - Mus musculus (Mouse)	MINK1_MOUSE	147 kDa	Mus musculus	2
RUN and SH3 domain-containing protein 2 - Mus musculus (Mouse)	RUSC2_MOUSE	161 kDa	Mus musculus	2
Bromodomain adjacent to zinc finger domain protein 1B (Williams-Beuren syndrome chromosome region 9 protein homolog) (WBRS9) - Mus musculus (Mouse)	BAZ1B_MOUSE	171 kDa	Mus musculus	2
Ankyrin repeat domain-containing protein 26 - Mus musculus (Mouse)	ANR26_MOUSE	181 kDa	Mus musculus	2
Clathrin heavy chain - Mus musculus (Mouse)	CLH_MOUSE	192 kDa	Mus musculus	2
Nebulin-related-anchoring protein (N-RAP) - Mus musculus (Mouse)	NRAP_MOUSE	196 kDa	Mus musculus	2
Dedicator of cytokinesis protein 2 (Protein Hch) - Mus musculus (Mouse)	DOCK2_MOUSE	212 kDa	Mus musculus	2
Ras-related protein Rab-10 - Mus musculus (Mouse)	RAB10_MOUSE (+6)	23 kDa	Mus musculus	2
Ras-related protein Rab-7a - Mus musculus (Mouse)	RAB7A_MOUSE	23 kDa	Mus musculus	2
Myoferlin (Fer-1-like protein 3) - Mus musculus (Mouse)	MYOF_MOUSE	233 kDa	Mus musculus	2
Brefeldin A-inhibited guanine nucleotide-exchange protein 3 - Mus musculus (Mouse)	BIG3_MOUSE	240 kDa	Mus musculus	2
Polycystic kidney disease and receptor for egg jelly-related protein precursor (PKD and REJ homolog) - Mus musculus (Mouse)	PKDRE_MOUSE	241 kDa	Mus musculus	2
Platelet-activating factor acetylhydrolase IB subunit beta (EC 3.1.1.47) (PAF acetylhydrolase 30 kDa subunit) (PAF-AH 30 kDa subunit) (PAF-AH subunit beta) (PAFAH subunit beta) - Mus musculus (Mouse)	PA1B2_MOUSE	26 kDa	Mus musculus	2
Eukaryotic translation initiation factor 4H (eIF-4H) (Williams-Beuren syndrome chromosome region 1 protein homolog) - Mus musculus (Mouse)	IF4H_MOUSE	27 kDa	Mus musculus	2
14-3-3 protein beta/alpha (Protein kinase C inhibitor protein 1) (KCIP-1) - Mus musculus (Mouse)	1433B_MOUSE	28 kDa	Mus musculus	2
ADP/ATP translocase 1 (Adenine nucleotide translocator 1) (ANT 1) (ADP,ATP carrier protein 1) (Solute carrier family 25 member 4) (ADP,ATP carrier protein, heart/skeletal muscle isoform T1) (mANC1) - Mus musculus (Mouse)	ADT1_MOUSE	33 kDa	Mus musculus	2
Coiled-coil domain-containing protein 68 - Mus musculus (Mouse)	CCD68_MOUSE	38 kDa	Mus musculus	2
Guanine nucleotide-binding protein G(i), alpha-2 subunit (Adenylate cyclase-inhibiting G alpha protein) - Mus musculus (Mouse)	GNAI2_MOUSE (+1)	40 kDa	Mus musculus	2
Endonuclease VIII-like 1 (EC 3.2.2.-) (EC 4.2.99.18) (Nei-like 1) (DNA glycosylase/AP lyase Neil1) (DNA-(apurinic or apyrimidinic site) lyase Neil1) (NEH1) - Mus musculus (Mouse)	NEIL1_MOUSE	44 kDa	Mus musculus	2
Nucleosome assembly protein 1-like 1 (NAP-1-related protein) (Brain protein DN38) - Mus musculus (Mouse)	NP1L1_MOUSE	45 kDa	Mus musculus	2
Docking protein 1 (Downstream of tyrosine kinase 1) (p62(dok)) - Mus musculus (Mouse)	DOK1_MOUSE	52 kDa	Mus musculus	2
Dolichyl-P-Man:Man(7)GlcNAc(2)-PP-dolichyl-alpha-1,6-mannosyltransferase (EC 2.4.1.-) (Mannosyltransferase ALG12 homolog) - Mus musculus (Mouse)	ALG12_MOUSE	54 kDa	Mus musculus	2

Ciliary dynein heavy chain 8 (Axonemal beta dynein heavy chain 8) - Mus musculus (Mouse)	DYH8_MOUSE	541 kDa	Mus musculus	2
Inosine-5'-monophosphate dehydrogenase 2 (EC 1.1.1.205) (IMP dehydrogenase 2) (IMPDH-II) (IMPD 2) - Mus musculus (Mouse)	IMDH2_MOUSE	56 kDa	Mus musculus	2
GTPase-activating protein ZNF289 - Mus musculus (Mouse)	ZN289_MOUSE	57 kDa	Mus musculus	2
Intercellular adhesion molecule 1 precursor (ICAM-1) (MALA-2) - Mus musculus (Mouse)	ICAM1_MOUSE	59 kDa	Mus musculus	2
T-complex protein 1 subunit epsilon (TCP-1-epsilon) (CCT-epsilon) - Mus musculus (Mouse)	TCPE_MOUSE	60 kDa	Mus musculus	2
Stress-induced-phosphoprotein 1 (STI1) (Hsc70/Hsp90-organizing protein) (Hop) (mSTI1) - Mus musculus (Mouse)	STIP1_MOUSE	63 kDa	Mus musculus	2
Uncharacterized protein KIAA1949 homolog - Mus musculus (Mouse)	K1949_MOUSE	66 kDa	Mus musculus	2
Echinoderm microtubule-associated protein-like 2 (EMAP-2) - Mus musculus (Mouse)	EMAL2_MOUSE	71 kDa	Mus musculus	2
Inactive serine protease RAMP precursor (Regeneration-associated muscle protease homolog) - Mus musculus (Mouse)	RAMP_MOUSE	80 kDa	Mus musculus	2
IQ and AAA domain-containing protein - Mus musculus (Mouse)	IQCA_MOUSE	100 kDa	Mus musculus	1
Afadin (Protein Af-6) (Fragment) - Mus musculus (Mouse)	AFAD_MOUSE	101 kDa	Mus musculus	1
WD repeat protein 47 - Mus musculus (Mouse)	WDR47_MOUSE	102 kDa	Mus musculus	1
Gamma-tubulin complex component 2 (GCP-2) - Mus musculus (Mouse)	GCP2_MOUSE	103 kDa	Mus musculus	1
Trifunctional purine biosynthetic protein adenosine-3 [Includes: Phosphoribosylamine--glycine ligase (EC 6.3.4.13) (GARS) (Glycinamide ribonucleotide synthetase) (Phosphoribosylglycinamide synthetase); Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1) (AIRS) (Phosphoribosyl-aminoimidazole synthetase) (AIR synthase); Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2) (GART) (GAR transformylase) (5'-phosphoribosylglycinamide transformylase)] - Mus musculus (Mouse)	PUR2_MOUSE (+1)	107 kDa	Mus musculus	1
Serine/threonine-protein kinase Nek9 (EC 2.7.11.1) (NimA-related protein kinase 9) (Never in mitosis A-related kinase 9) - Mus musculus (Mouse)	NEK9_MOUSE	107 kDa	Mus musculus	1
E3 ubiquitin-protein ligase Arkadia (EC 6.3.2.-) (RING finger protein 111) - Mus musculus (Mouse)	RN111_MOUSE	108 kDa	Mus musculus	1
Resistin-like beta precursor (RELMbeta) (Cysteine-rich secreted protein FIZZ2) (Cysteine-rich secreted protein A12-beta) - Mus musculus (Mouse)	RETNB_MOUSE	11 kDa	Mus musculus	1
Uncharacterized protein C1orf130 homolog - Mus musculus (Mouse)	CA130_MOUSE	11 kDa	Mus musculus	1
Ephrin type-A receptor 4 precursor (EC 2.7.10.1) (Tyrosine-protein kinase receptor SEK) (MPK-3) - Mus musculus (Mouse)	EPHA4_MOUSE	110 kDa	Mus musculus	1
Ellis-van Creveld syndrome protein homolog - Mus musculus (Mouse)	EVC_MOUSE	113 kDa	Mus musculus	1
Ankyrin repeat domain-containing protein 28 - Mus musculus (Mouse)	ANR28_MOUSE	113 kDa	Mus musculus	1
Integrin alpha-4 precursor (Integrin alpha-IV) (VLA-4) (Lymphocyte Peyer patch adhesion molecules subunit alpha) (LPAM subunit alpha) (CD49d antigen) - Mus musculus (Mouse)	ITA4_MOUSE	116 kDa	Mus musculus	1

Probable ATP-dependent RNA helicase DDX46 (EC 3.6.1.-) (DEAD box protein 46) - Mus musculus (Mouse)	DDX46_MOUSE	117 kDa	Mus musculus	1
Calmin - Mus musculus (Mouse)	CLMN_MOUSE	117 kDa	Mus musculus	1
Proline-, glutamic acid- and leucine-rich protein 1 (Modulator of nongenomic activity of estrogen receptor) - Mus musculus (Mouse)	PELP1_MOUSE	118 kDa	Mus musculus	1
60S ribosomal protein L36 - Mus musculus (Mouse)	RL36_MOUSE	12 kDa	Mus musculus	1
Zinc finger Ran-binding domain-containing protein 3 (EC 3.6.1.-) - Mus musculus (Mouse)	ZRAB3_MOUSE	121 kDa	Mus musculus	1
Ubiquitin-protein ligase E3C (EC 6.3.2.-) - Mus musculus (Mouse)	UBE3C_MOUSE	124 kDa	Mus musculus	1
Ubiquitin carboxyl-terminal hydrolase 43 (EC 3.1.2.15) (Ubiquitin thioesterase 43) (Ubiquitin-specific-processing protease 43) (Deubiquitinating enzyme 43) - Mus musculus (Mouse)	UBP43_MOUSE	124 kDa	Mus musculus	1
Pleckstrin homology domain-containing family A member 7 - Mus musculus (Mouse)	PKHA7_MOUSE	127 kDa	Mus musculus	1
Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (Leukocyte adhesion receptor MO1) (CD11b antigen) - Mus musculus (Mouse)	ITAM_MOUSE	127 kDa	Mus musculus	1
Structural maintenance of chromosomes protein 5 (mSMC5) - Mus musculus (Mouse)	SMC5_MOUSE	129 kDa	Mus musculus	1
Thrombospondin-2 precursor - Mus musculus (Mouse)	TSP2_MOUSE	130 kDa	Mus musculus	1
Zinc finger protein 687 - Mus musculus (Mouse)	ZN687_MOUSE	130 kDa	Mus musculus	1
Partitioning-defective 3 homolog B (PAR3-beta) (Partitioning-defective 3-like protein) (PAR3-L protein) (Amyotrophic lateral sclerosis 2 chromosome region candidate gene 19 protein homolog) - Mus musculus (Mouse)	PAR3L_MOUSE	133 kDa	Mus musculus	1
Zinc finger E-box-binding homeobox 2 (Zinc finger homeobox protein 1b) (Smad-interacting protein 1) - Mus musculus (Mouse)	ZEB2_MOUSE	136 kDa	Mus musculus	1
Tax1-binding protein 3 (Tax interaction protein 1) (TIP-1) - Mus musculus (Mouse)	TX1B3_MOUSE	14 kDa	Mus musculus	1
LIM domain-containing protein 2 - Mus musculus (Mouse)	LIMD2_MOUSE	14 kDa	Mus musculus	1
TC10/CDC42 GTPase-activating protein (Sorting nexin-26) - Mus musculus (Mouse)	TCGAP_MOUSE	140 kDa	Mus musculus	1
Pleckstrin homology-like domain family B member 2 (Protein LL5-beta) - Mus musculus (Mouse)	PHLB2_MOUSE	141 kDa	Mus musculus	1
Dynactin subunit 1 (150 kDa dynein-associated polypeptide) (DP-150) (DAP-150) (p150-glued) - Mus musculus (Mouse)	DCTN1_MOUSE (+2)	142 kDa	Mus musculus	1
Liprin-alpha-2 (Protein tyrosine phosphatase receptor type f polypeptide-interacting protein alpha-2) (PTPRF-interacting protein alpha-2) - Mus musculus (Mouse)	LIPA2_MOUSE	143 kDa	Mus musculus	1
Ninein-like protein - Mus musculus (Mouse)	NINLP_MOUSE	144 kDa	Mus musculus	1
Astrotactin-1 precursor (Neuronal migration protein GC14) - Mus musculus (Mouse)	ASTN1_MOUSE (+2)	145 kDa	Mus musculus	1
Splicing factor 3B subunit 1 (Spliceosome-associated protein 155) (SAP 155) (SF3b155) (Pre-mRNA-splicing factor SF3b 155 kDa subunit) - Mus musculus (Mouse)	SF3B1_MOUSE (+2)	146 kDa	Mus musculus	1

Paired amphipathic helix protein Sin3a (Transcriptional corepressor Sin3a) (Histone deacetylase complex subunit Sin3a) - Mus musculus (Mouse)	SIN3A_MOUSE (+1)	146 kDa	Mus musculus	1
60S ribosomal protein L23 - Mus musculus (Mouse)	RL23_MOUSE	15 kDa	Mus musculus	1
Protein FAM19A2 precursor (Chemokine-like protein TAFA-2) - Mus musculus (Mouse)	F19A2_MOUSE	15 kDa	Mus musculus	1
Putative ATP-dependent RNA helicase DHX29 (EC 3.6.1.-) (DEAH box protein 29) - Mus musculus (Mouse)	DHX29_MOUSE (+2)	154 kDa	Mus musculus	1
Protein phosphatase Slingshot homolog 2 (EC 3.1.3.48) (EC 3.1.3.16) (SSH-2L) (mSSH-2L) - Mus musculus (Mouse)	SSH2_MOUSE	158 kDa	Mus musculus	1
Rho-associated protein kinase 1 (EC 2.7.11.1) (Rho-associated, coiled-coil-containing protein kinase 1) (p160 ROCK-1) (p160ROCK) - Mus musculus (Mouse)	ROCK1_MOUSE (+1)	158 kDa	Mus musculus	1
Proteasome maturation protein (Proteassemblin) (Protein UMP1 homolog) (mUMP1) - Mus musculus (Mouse)	POMP_MOUSE	16 kDa	Mus musculus	1
UPF0279 protein C14orf129 homolog - Mus musculus (Mouse)	CN129_MOUSE	16 kDa	Mus musculus	1
Nitric-oxide synthase, brain (EC 1.14.13.39) (NOS type I) (Neuronal NOS) (N-NOS) (nNOS) (Constitutive NOS) (NC-NOS) (bNOS) - Mus musculus (Mouse)	NOS1_MOUSE	160 kDa	Mus musculus	1
Collagen alpha-3(IV) chain precursor [Contains: Tumstatin] - Mus musculus (Mouse)	CO4A3_MOUSE	162 kDa	Mus musculus	1
Protein capicua homolog - Mus musculus (Mouse)	CIC_MOUSE	164 kDa	Mus musculus	1
Alpha-2-macroglobulin precursor (Pregnancy zone protein) (Alpha-2-M) [Contains: Alpha-2-macroglobulin 165 kDa subunit; Alpha-2-macroglobulin 35 kDa subunit] - Mus musculus (Mouse)	A2M_MOUSE (+5)	166 kDa	Mus musculus	1
Eukaryotic translation initiation factor 5A-1 (eIF-5A-1) (eIF-5A1) (Eukaryotic initiation factor 5A isoform 1) (eIF-5A) (eIF-4D) - Mus musculus (Mouse)	IF5A1_MOUSE (+1)	17 kDa	Mus musculus	1
Prostaglandin E synthase (EC 5.3.99.3) (mPGES-1) - Mus musculus (Mouse)	PTGES_MOUSE	17 kDa	Mus musculus	1
Brain-specific angiogenesis inhibitor 1 precursor - Mus musculus (Mouse)	BAI1_MOUSE	173 kDa	Mus musculus	1
E3 ubiquitin-protein ligase HECW1 (EC 6.3.2.-) (HECT, C2 and WW domain-containing protein 1) (NEDD4-like ubiquitin-protein ligase 1) (mNEDL1) - Mus musculus (Mouse)	HECW1_MOUSE	179 kDa	Mus musculus	1
DEP domain-containing protein 2 (Phosphatidylinositol 3,4,5-trisphosphate-dependent Rac exchanger 2) (PtdIns(3,4,5)-dependent Rac exchanger 2) (P-Rex2) - Mus musculus (Mouse)	DEPD2_MOUSE (+2)	182 kDa	Mus musculus	1
BCL-6 corepressor (BCoR) - Mus musculus (Mouse)	BCOR_MOUSE	193 kDa	Mus musculus	1
Lens fiber membrane intrinsic protein (MP17) (MP18) (MP19) (MP20) - Mus musculus (Mouse)	LMIP_MOUSE	20 kDa	Mus musculus	1
Coatomer subunit zeta-1 (Zeta-1 coat protein) (Zeta-1 COP) - Mus musculus (Mouse)	COPZ1_MOUSE	20 kDa	Mus musculus	1
Proteasome-associated protein ECM29 homolog (Ecm29) - Mus musculus (Mouse)	ECM29_MOUSE	204 kDa	Mus musculus	1

Zinc finger protein 106 (Zfp-106) (Zinc finger protein 474) (H3a minor histocompatibility antigen) (Son of insulin receptor mutant) - Mus musculus (Mouse)	ZF106_MOUSE (+15)	209 kDa	Mus musculus	1
Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-) (Proprotein convertase PC5) (Subtilisin/kexin-like protease PC5) (PC6) (Subtilisin-like proprotein convertase 6) (SPC6) - Mus musculus (Mouse)	PCSK5_MOUSE (+3)	209 kDa	Mus musculus	1
Histone H1.0 (Histone H1(0)) (Histone H1') - Mus musculus (Mouse)	H10_MOUSE	21 kDa	Mus musculus	1
60S ribosomal protein L18a - Mus musculus (Mouse)	RL18A_MOUSE	21 kDa	Mus musculus	1
Major urinary protein 4 precursor (MUP 4) - Mus musculus (Mouse)	MUP4_MOUSE	21 kDa	Mus musculus	1
Low-density lipoprotein receptor-related protein 4 precursor (LDLR dan) - Mus musculus (Mouse)	LRP4_MOUSE (+1)	212 kDa	Mus musculus	1
Voltage-dependent L-type calcium channel subunit alpha-1S (Voltage-gated calcium channel subunit alpha Cav1.1) (Calcium channel, L type, alpha-1 polypeptide, isoform 3, skeletal muscle) - Mus musculus (Mouse)	CAC1S_MOUSE	213 kDa	Mus musculus	1
Zinc finger protein 638 (Nuclear protein 220) (Zinc-finger matrin-like protein) - Mus musculus (Mouse)	ZN638_MOUSE	218 kDa	Mus musculus	1
40S ribosomal protein S7 - Mus musculus (Mouse)	RS7_MOUSE	22 kDa	Mus musculus	1
Ras-related protein Rab-22A (Rab-22) (Rab-14) - Mus musculus (Mouse)	RB22A_MOUSE	22 kDa	Mus musculus	1
60S ribosomal protein L18 - Mus musculus (Mouse)	RL18_MOUSE	22 kDa	Mus musculus	1
Nascent polypeptide-associated complex subunit alpha, muscle-specific form (Alpha-NAC, muscle-specific form) - Mus musculus (Mouse)	NACAM_MOUSE	221 kDa	Mus musculus	1
Myosin-4 (Myosin heavy chain 4) (Myosin heavy chain 2b) (MyHC-2b) - Mus musculus (Mouse)	MYH4_MOUSE	223 kDa	Mus musculus	1
Myosin-6 (Myosin heavy chain 6) (Myosin heavy chain, cardiac muscle alpha isoform) (MyHC-alpha) - Mus musculus (Mouse)	MYH6_MOUSE (+6)	224 kDa	Mus musculus	1
Rootletin (Ciliary rootlet coiled-coil protein) - Mus musculus (Mouse)	CROCC_MOUSE	227 kDa	Mus musculus	1
Synaptosomal-associated protein 23 (SNAP-23) (Vesicle-membrane fusion protein SNAP-23) (Syndet) - Mus musculus (Mouse)	SNP23_MOUSE	23 kDa	Mus musculus	1
Homeobox protein Rhox5 (Reproductive homeobox on chromosome X 5) (Homeobox protein Pem) (Placenta and embryonic expression protein) - Mus musculus (Mouse)	RHOX5_MOUSE	23 kDa	Mus musculus	1
Rho GDP-dissociation inhibitor 1 (Rho GDI 1) (Rho-GDI alpha) (GDI-1) - Mus musculus (Mouse)	GDIR_MOUSE	23 kDa	Mus musculus	1
Oxygen-regulated protein 1 (Retinitis pigmentosa RP1 protein homolog) - Mus musculus (Mouse)	RP1_MOUSE (+1)	234 kDa	Mus musculus	1
Dedicator of cytokinesis protein 9 (Cdc42 guanine nucleotide exchange factor zizimin-1) - Mus musculus (Mouse)	DOCK9_MOUSE (+5)	235 kDa	Mus musculus	1
Uncharacterized protein C1orf49 homolog - Mus musculus (Mouse)	CA049_MOUSE	24 kDa	Mus musculus	1
Ras-related protein Rab-11A (Rab-11) - Mus musculus (Mouse)	RB11A_MOUSE (+1)	24 kDa	Mus musculus	1
Cardiotrophin-like cytokine factor 1 precursor (B cell-stimulating factor 3) (BSF-3) (Novel neurotrophin-1) (NNT-1) - Mus musculus (Mouse)	CLCF1_MOUSE	25 kDa	Mus musculus	1

Lipoma HMGIC fusion partner-like 3 protein - Mus musculus (Mouse)	LHPL3_MOUSE	25 kDa	Mus musculus	1
UPF0193 protein EVG1 homolog - Mus musculus (Mouse)	EVG1_MOUSE	25 kDa	Mus musculus	1
Claudin-17 - Mus musculus (Mouse)	CLD17_MOUSE	25 kDa	Mus musculus	1
Protein-L-isoaspartate(D-aspartate) O-methyltransferase (EC 2.1.1.77) (Protein-beta-aspartate methyltransferase) (PIMT) (Protein L-isoaspartyl/D-aspartyl methyltransferase) (L-isoaspartyl protein carboxyl methyltransferase) - Mus musculus (Mouse)	PIMT_MOUSE	25 kDa	Mus musculus	1
Host cell factor 2 (HCF-2) (C2 factor) - Mus musculus (Mouse)	HCFC2_MOUSE	26 kDa	Mus musculus	1
High affinity immunoglobulin epsilon receptor beta-subunit (FcERI) (IgE Fc receptor subunit beta) (Fc epsilon receptor I beta-chain) (Membrane-spanning 4-domains subfamily A member 2) - Mus musculus (Mouse)	FCERB_MOUSE	26 kDa	Mus musculus	1
Dihydropteridine reductase (EC 1.5.1.34) (HDHPR) (Quinoid dihydropteridine reductase) - Mus musculus (Mouse)	DHPR_MOUSE	26 kDa	Mus musculus	1
Proteasome subunit alpha type 5 (EC 3.4.25.1) (Proteasome zeta chain) (Macropain zeta chain) (Multicatalytic endopeptidase complex zeta chain) - Mus musculus (Mouse)	PSA5_MOUSE	26 kDa	Mus musculus	1
EF-hand domain-containing protein 2 (Swiprosin-1) - Mus musculus (Mouse)	EFHD2_MOUSE	27 kDa	Mus musculus	1
Calcyclin-binding protein (CacyBP) (Siah-interacting protein) - Mus musculus (Mouse)	CYBP_MOUSE	27 kDa	Mus musculus	1
Receptor expression-enhancing protein 2 - Mus musculus (Mouse)	REEP2_MOUSE	28 kDa	Mus musculus	1
Antolefinin (Antolefinine) - Mus musculus (Mouse)	ANTOL_MOUSE	28 kDa	Mus musculus	1
Vesicle-associated membrane protein-associated protein A (VAMP-associated protein A) (VAMP-A) (VAP-A) (33 kDa Vamp-associated protein) (VAP-33) - Mus musculus (Mouse)	VAPA_MOUSE (+1)	28 kDa	Mus musculus	1
FKBP12-rapamycin complex-associated protein (FK506-binding protein 12-rapamycin complex-associated protein 1) (Rapamycin target protein) (RAPT1) (Mammalian target of rapamycin) (mTOR) - Mus musculus (Mouse)	FRAP_MOUSE (+6)	289 kDa	Mus musculus	1
Origin recognition complex subunit 6 - Mus musculus (Mouse)	ORC6_MOUSE	29 kDa	Mus musculus	1
Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonic anhydrase II) (Carbonate dehydratase II) (CA-II) - Mus musculus (Mouse)	CAH2_MOUSE	29 kDa	Mus musculus	1
COP9 signalosome complex subunit 7a (Signalosome subunit 7a) (SGN7a) (JAB1-containing signalosome subunit 7a) - Mus musculus (Mouse)	CSN7A_MOUSE	30 kDa	Mus musculus	1
Protein FAM79A - Mus musculus (Mouse)	FA79A_MOUSE	30 kDa	Mus musculus	1
Prohibitin (B-cell receptor-associated protein 32) (BAP 32) - Mus musculus (Mouse)	PHB_MOUSE	30 kDa	Mus musculus	1
Caspase-3 precursor (EC 3.4.22.56) (CASP-3) (Apopain) (Cysteine protease CPP32) (Yama protein) (CPP-32) (SREBP cleavage activity 1) (SCA-1) (LICE) [Contains: Caspase-3 p17 subunit; Caspase-3 p12 subunit] - Mus musculus (Mouse)	CASP3_MOUSE	31 kDa	Mus musculus	1
Receptor-transporting protein 1 - Mus musculus (Mouse)	RTP1_MOUSE	31 kDa	Mus musculus	1
Calretinin (CR) - Mus musculus (Mouse)	CALB2_MOUSE	31 kDa	Mus musculus	1
Splicing factor, arginine/serine-rich 7 - Mus musculus (Mouse)	SFRS7_MOUSE	31 kDa	Mus musculus	1

Phosphatidylinositol transfer protein alpha isoform (PtdIns transfer protein alpha) (PtdInsTP) (PI-TP-alpha) - Mus musculus (Mouse)	PIPNA_MOUSE	32 kDa	Mus musculus	1
Voltage-dependent anion-selective channel protein 2 (VDAC-2) (mVDAC2) (mVDAC6) (Outer mitochondrial membrane protein porin 2) - Mus musculus (Mouse)	VDAC2_MOUSE	32 kDa	Mus musculus	1
Voltage-dependent anion-selective channel protein 1 (VDAC-1) (mVDAC1) (mVDAC5) (Outer mitochondrial membrane protein porin 1) (Plasmalemmal porin) - Mus musculus (Mouse)	VDAC1_MOUSE (+1)	32 kDa	Mus musculus	1
Mortality factor 4-like protein 2 (MORF-related gene X protein) (Transcription factor-like protein MRGX) (Sid 393) - Mus musculus (Mouse)	MO4L2_MOUSE	32 kDa	Mus musculus	1
Calponin-2 (Calponin H2, smooth muscle) (Neutral calponin) - Mus musculus (Mouse)	CNN2_MOUSE	33 kDa	Mus musculus	1
Interferon-stimulated gene 20 kDa protein (EC 3.1.13.1) (Promyelocytic leukemia nuclear body-associated protein ISG20) (DnaQL protein) - Mus musculus (Mouse)	ISG20_MOUSE	33 kDa	Mus musculus	1
ATP synthase gamma chain, mitochondrial precursor (EC 3.6.3.14) - Mus musculus (Mouse)	ATPG_MOUSE	33 kDa	Mus musculus	1
Max-like protein X (Max-like bHLHZip protein) (BigMax protein) (Protein Mlx) (Transcription factor-like protein 4) - Mus musculus (Mouse)	MLX_MOUSE	33 kDa	Mus musculus	1
Coiled-coil domain-containing protein C1orf110 homolog - Mus musculus (Mouse)	CA110_MOUSE	33 kDa	Mus musculus	1
Ankyrin repeat family A protein 2 (RFXANK-like 2) - Mus musculus (Mouse)	ANRA2_MOUSE	34 kDa	Mus musculus	1
Elongation of very long chain fatty acids protein 2 - Mus musculus (Mouse)	ELOV2_MOUSE	34 kDa	Mus musculus	1
Mitochondrial 2-oxoglutarate/malate carrier protein (OGCP) (Solute carrier family 25 member 11) - Mus musculus (Mouse)	M2OM_MOUSE	34 kDa	Mus musculus	1
Pituitary homeobox 1 (Homeobox protein P-OTX) (Pituitary OTX-related factor) (Hindlimb expressed homeobox protein backfoot) (Ptx1) - Mus musculus (Mouse)	PITX1_MOUSE	34 kDa	Mus musculus	1
Uncharacterized protein C13orf26 homolog - Mus musculus (Mouse)	CM026_MOUSE	35 kDa	Mus musculus	1
Probable rRNA-processing protein EBP2 - Mus musculus (Mouse)	EBP2_MOUSE	35 kDa	Mus musculus	1
Vomeronasal type-1 receptor A7 (Vomeronasal receptor 3) (Pheromone receptor VN3) - Mus musculus (Mouse)	VN1A7_MOUSE (+1)	35 kDa	Mus musculus	1
Dual specificity protein phosphatase 2 (EC 3.1.3.48) (EC 3.1.3.16) (Dual specificity protein phosphatase PAC-1) - Mus musculus (Mouse)	DUS2_MOUSE	35 kDa	Mus musculus	1
Striated muscle-specific serine/threonine protein kinase (EC 2.7.11.1) (Aortic preferentially expressed protein 1) (APEG-1) - Mus musculus (Mouse)	SPEG_MOUSE	354 kDa	Mus musculus	1
Annexin A4 (Annexin IV) - Mus musculus (Mouse)	ANXA4_MOUSE	36 kDa	Mus musculus	1
Autophagy-related protein 3 (APG3-like) - Mus musculus (Mouse)	ATG3_MOUSE	36 kDa	Mus musculus	1
Transcription factor MafB (V-maf musculoaponeurotic fibrosarcoma oncogene homolog B) (Transcription factor MAF1) (Segmentation protein KR) (Kreisler) - Mus musculus (Mouse)	MAFB_MOUSE (+5)	36 kDa	Mus musculus	1

Homeobox protein Hox-A1 (Hox-1.6) (Homeotic protein ERA-1-993) (Early retinoic acid 1) (Homeoboxless protein ERA-1-399) - Mus musculus (Mouse)	HXA1_MOUSE	36 kDa	Mus musculus	1
Transaldolase (EC 2.2.1.2) - Mus musculus (Mouse)	TALDO_MOUSE	37 kDa	Mus musculus	1
Serine/threonine-protein phosphatase PP1-alpha catalytic subunit (EC 3.1.3.16) (PP-1A) - Mus musculus (Mouse)	PP1A_MOUSE (+2)	38 kDa	Mus musculus	1
Suppressor of G2 allele of SKP1 homolog - Mus musculus (Mouse)	SUGT1_MOUSE	38 kDa	Mus musculus	1
Leukotriene B4 receptor 1 (LTB4-R 1) - Mus musculus (Mouse)	LT4R1_MOUSE	38 kDa	Mus musculus	1
PDZ and LIM domain protein 2 (PDZ-LIM protein mystique) - Mus musculus (Mouse)	PDLI2_MOUSE	38 kDa	Mus musculus	1
Thioredoxin-like protein 2 (PKC-interacting cousin of thioredoxin) (PKC-theta-interacting protein) (PKCq-interacting protein) - Mus musculus (Mouse)	TXNL2_MOUSE	38 kDa	Mus musculus	1
Annexin A1 (Annexin I) (Lipocortin I) (Calpactin II) (Chromobindin-9) (p35) (Phospholipase A2 inhibitory protein) - Mus musculus (Mouse)	ANXA1_MOUSE	39 kDa	Mus musculus	1
Nuclear inhibitor of protein phosphatase 1 (NIPP-1) (Protein phosphatase 1 regulatory inhibitor subunit 8) - Mus musculus (Mouse)	PP1R8_MOUSE	39 kDa	Mus musculus	1
Twinfilin-1 (Protein A6) - Mus musculus (Mouse)	TWF1_MOUSE	40 kDa	Mus musculus	1
Ubiquitin-like protein 7 - Mus musculus (Mouse)	UBL7_MOUSE	40 kDa	Mus musculus	1
Laminin subunit alpha-5 precursor - Mus musculus (Mouse)	LAMA5_MOUSE	404 kDa	Mus musculus	1
5-hydroxytryptamine 5B receptor (5-HT-5B) (Serotonin receptor 5B) - Mus musculus (Mouse)	5HT5B_MOUSE	41 kDa	Mus musculus	1
Membrane protein MLC1 - Mus musculus (Mouse)	MLC1_MOUSE	42 kDa	Mus musculus	1
3 beta-hydroxysteroid dehydrogenase type 4 (3 beta-hydroxysteroid dehydrogenase type IV) (3-beta-HSD IV) (NADPH-dependent 3-beta-hydroxy-Delta(5)-steroid dehydrogenase) (EC 1.1.1.-) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase) - Mus musculus (Mouse)	3BHS4_MOUSE (+1)	42 kDa	Mus musculus	1
Testis-specific serine/threonine-protein kinase 5 (EC 2.7.11.1) (TSSK-5) (Testis-specific kinase 5) (TSK-5) - Mus musculus (Mouse)	TSSK5_MOUSE	42 kDa	Mus musculus	1
P-selectin glycoprotein ligand 1 precursor (PSGL-1) (Selectin P ligand) - Mus musculus (Mouse)	SELPL_MOUSE	42 kDa	Mus musculus	1
Selenoprotein P precursor (SeP) (Plasma selenoprotein P) - Mus musculus (Mouse)	SEPP1_MOUSE	42 kDa	Mus musculus	1
Dystrophin - Mus musculus (Mouse)	DMD_MOUSE (+1)	426 kDa	Mus musculus	1
Uncharacterized protein C9orf114 homolog - Mus musculus (Mouse)	CI114_MOUSE (+1)	43 kDa	Mus musculus	1
Prostaglandin E synthase 2 (EC 5.3.99.3) (Microsomal prostaglandin E synthase 2) (mPGES-2) (GATE-binding factor 1) (GBF-1) [Contains: Prostaglandin E synthase 2 truncated form] - Mus musculus (Mouse)	PGES2_MOUSE	43 kDa	Mus musculus	1
Methionyl-tRNA formyltransferase, mitochondrial precursor (EC 2.1.2.9) (MtFMT) - Mus musculus (Mouse)	FMT_MOUSE	43 kDa	Mus musculus	1
Progestin and adipoQ receptor family member 9 (Progestin and adipoQ receptor family member IX) - Mus musculus (Mouse)	PAQR9_MOUSE	43 kDa	Mus musculus	1
Prostaglandin E2 receptor, EP1 subtype (Prostanoid EP1 receptor) (PGE receptor, EP1 subtype) - Mus musculus (Mouse)	PE2R1_MOUSE	43 kDa	Mus musculus	1

RNA-binding motif, single-stranded-interacting protein 1 (Single-stranded DNA-binding protein MSSP-1) - Mus musculus (Mouse)	RBMS1_MOUSE (+1)	44 kDa	Mus musculus	1
Alpha-(1,3)-fucosyltransferase (EC 2.4.1.-) (Galactoside 3-L-fucosyltransferase) (Fucosyltransferase 7) (FUCT-VII) - Mus musculus (Mouse)	FUT7_MOUSE	44 kDa	Mus musculus	1
Natural killer cell receptor 2B4 precursor (NKR2B4) (NK cell type I receptor protein 2B4) (CD244 antigen) (Non MHC restricted killing associated) - Mus musculus (Mouse)	CD244_MOUSE	45 kDa	Mus musculus	1
DnaJ homolog subfamily A member 1 (Heat shock 40 kDa protein 4) (DnaJ protein homolog 2) (HSJ-2) - Mus musculus (Mouse)	DNJA1_MOUSE	45 kDa	Mus musculus	1
Adherens junction-associated protein 1 - Mus musculus (Mouse)	AJAP1_MOUSE	45 kDa	Mus musculus	1
EMI domain-containing protein 1 precursor (Protein Emu1) (Emilin and multimerin domain-containing protein 1) - Mus musculus (Mouse)	EMID1_MOUSE	46 kDa	Mus musculus	1
Protein pellino homolog 1 (Pellino-1) - Mus musculus (Mouse)	PELI1_MOUSE	46 kDa	Mus musculus	1
Keratin, type I cuticular Ha1 (Hair keratin, type I Ha1) (Keratin-31) (HKA-1) - Mus musculus (Mouse)	K1H1_MOUSE	47 kDa	Mus musculus	1
Argininosuccinate synthase (EC 6.3.4.5) (Citrulline--aspartate ligase) - Mus musculus (Mouse)	ASSY_MOUSE	47 kDa	Mus musculus	1
SET and MYND domain-containing protein 5 (Protein NN8-4AG) - Mus musculus (Mouse)	SMYD5_MOUSE	47 kDa	Mus musculus	1
Squalene synthetase (EC 2.5.1.21) (SQS) (SS) (Farnesyl-diphosphate farnesyltransferase) (FPP:FPP farnesyltransferase) - Mus musculus (Mouse)	FDFT_MOUSE (+2)	48 kDa	Mus musculus	1
Leucine-rich repeat-containing protein 42 - Mus musculus (Mouse)	LRC42_MOUSE (+5)	48 kDa	Mus musculus	1
Intraflagellar transport 52 homolog (Protein NGD5) - Mus musculus (Mouse)	IFT52_MOUSE	48 kDa	Mus musculus	1
Inward rectifier potassium channel 16 (Potassium channel, inwardly rectifying subfamily J member 16) (Inward rectifier K(+) channel Kir5.1) - Mus musculus (Mouse)	IRK16_MOUSE	48 kDa	Mus musculus	1
Nucleoporin 50 kDa (Nuclear pore-associated protein 60 kDa-like) - Mus musculus (Mouse)	NUP50_MOUSE	49 kDa	Mus musculus	1
Septin-6 - Mus musculus (Mouse)	SEPT6_MOUSE	50 kDa	Mus musculus	1
Iroquois-class homeodomain protein IRX-1 (Iroquois homeobox protein 1) (Homeodomain protein IRXA1) - Mus musculus (Mouse)	IRX1_MOUSE (+1)	50 kDa	Mus musculus	1
Malcavernin (Cerebral cavernous malformations protein 2 homolog) (Osmosensing scaffold for MEKK3) - Mus musculus (Mouse)	CCM2_MOUSE (+4)	50 kDa	Mus musculus	1
Protein FAM113A - Mus musculus (Mouse)	F113A_MOUSE	51 kDa	Mus musculus	1
Inositol-trisphosphate 3-kinase A (EC 2.7.1.127) (Inositol 1,4,5-trisphosphate 3-kinase A) (IP3K A) (IP3 3-kinase A) - Mus musculus (Mouse)	IP3KA_MOUSE	51 kDa	Mus musculus	1
L-2-hydroxyglutarate dehydrogenase, mitochondrial precursor (EC 1.1.99.2) (Duranin) - Mus musculus (Mouse)	L2HDH_MOUSE	51 kDa	Mus musculus	1
Zinc finger protein 622 - Mus musculus (Mouse)	ZN622_MOUSE	53 kDa	Mus musculus	1
6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44) - Mus musculus (Mouse)	6PGD_MOUSE	53 kDa	Mus musculus	1
FAD-dependent oxidoreductase domain-containing protein 1 - Mus musculus (Mouse)	FXRD1_MOUSE	54 kDa	Mus musculus	1

Serine/threonine-protein kinase Chk1 (EC 2.7.11.1) - Mus musculus (Mouse)	CHK1_MOUSE	54 kDa	Mus musculus	1
Target of Myb protein 1 - Mus musculus (Mouse)	TOM1_MOUSE	54 kDa	Mus musculus	1
RUN and FYVE domain-containing protein 4 - Mus musculus (Mouse)	RUFY4_MOUSE	54 kDa	Mus musculus	1
Beta-Ala-His dipeptidase (EC 3.4.13.20) (Carnosine dipeptidase 1) (CNDP dipeptidase 1) - Mus musculus (Mouse)	CNDP1_MOUSE (+5)	55 kDa	Mus musculus	1
Archaemetzincin-1 (EC 3.-.-.-) (Archeobacterial metalloproteinase-like protein 1) - Mus musculus (Mouse)	AMZ1_MOUSE	55 kDa	Mus musculus	1
Myotilin (Titin immunoglobulin domain protein) (Myofibrillar titin-like Ig domains protein) - Mus musculus (Mouse)	MYOTI_MOUSE	55 kDa	Mus musculus	1
Transcriptional regulator ERG - Mus musculus (Mouse)	ERG_MOUSE	55 kDa	Mus musculus	1
Ubiquitin carboxyl-terminal hydrolase 14 (EC 3.1.2.15) (Ubiquitin thioesterase 14) (Ubiquitin-specific-processing protease 14) (Deubiquitinating enzyme 14) - Mus musculus (Mouse)	UBP14_MOUSE	56 kDa	Mus musculus	1
TBC1 domain family member 10A (EBP50-PDX interactor of 64 kDa) (EPI64 protein) - Mus musculus (Mouse)	TB10A_MOUSE	56 kDa	Mus musculus	1
G1 to S phase transition protein 1 homolog - Mus musculus (Mouse)	GSPT1_MOUSE	56 kDa	Mus musculus	1
Solute carrier family 2, facilitated glucose transporter member 10 (Glucose transporter type 10) - Mus musculus (Mouse)	GTR10_MOUSE	57 kDa	Mus musculus	1
Serine/threonine-protein kinase Sgk3 (EC 2.7.11.1) (Serum/glucocorticoid-regulated kinase 3) (Serum/glucocorticoid-regulated kinase-like) (Cytokine-independent survival kinase) - Mus musculus (Mouse)	SGK3_MOUSE (+2)	57 kDa	Mus musculus	1
T-complex protein 1 subunit beta (TCP-1-beta) (CCT-beta) - Mus musculus (Mouse)	TCPB_MOUSE	57 kDa	Mus musculus	1
Serine/threonine-protein kinase Nek3 (EC 2.7.11.1) (NimA-related protein kinase 3) - Mus musculus (Mouse)	NEK3_MOUSE	57 kDa	Mus musculus	1
Protein aurora borealis - Mus musculus (Mouse)	BORA_MOUSE	57 kDa	Mus musculus	1
Cytochrome P450 7B1 (EC 1.14.13.100) (25-hydroxycholesterol 7-alpha-hydroxylase) (Oxysterol 7-alpha-hydroxylase) (HCT-1) - Mus musculus (Mouse)	CP7B1_MOUSE (+11)	58 kDa	Mus musculus	1
T-complex protein 1 subunit zeta (TCP-1-zeta) (CCT-zeta) (CCT-zeta-1) - Mus musculus (Mouse)	TCPZ_MOUSE	58 kDa	Mus musculus	1
Cytochrome P450 3A16 (EC 1.14.14.1) (CYPIIIA16) - Mus musculus (Mouse)	CP3AG_MOUSE	58 kDa	Mus musculus	1
Immunity-related GTPase family Q protein - Mus musculus (Mouse)	IRGQ_MOUSE	59 kDa	Mus musculus	1
Zinc finger protein 14 (Zfp-14) (Krox-9 protein) - Mus musculus (Mouse)	ZFP14_MOUSE	59 kDa	Mus musculus	1
RING finger protein 37 (Ubiquitin-conjugating enzyme 7-interacting protein 5) (UbcM4-interacting protein 5) (U-box domain-containing protein 5) - Mus musculus (Mouse)	RNF37_MOUSE	59 kDa	Mus musculus	1
Gamma-taxilin - Mus musculus (Mouse)	TXLNG_MOUSE	60 kDa	Mus musculus	1
T-complex protein 1 subunit alpha A (TCP-1-alpha) (CCT-alpha) (Tailless complex polypeptide 1A) (TCP-1-A) - Mus musculus (Mouse)	TCPA1_MOUSE (+1)	60 kDa	Mus musculus	1
Synaptic vesicle 2-related protein (SV2-related protein) - Mus musculus (Mouse)	SVOP_MOUSE (+1)	61 kDa	Mus musculus	1

Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit gamma isoform (PP2A, B subunit, B' gamma isoform) (PP2A, B subunit, B56 gamma isoform) (PP2A, B subunit, PR61 gamma isoform) (PP2A, B subunit, R5 gamma isoform) (PP2A, B subunit, B'alpha3 isoform) - Mus musculus (Mouse)	2A5G_MOUSE (+6)	61 kDa	Mus musculus	1
Mitogen-activated protein kinase 15 (EC 2.7.11.24) (Extracellular signal-regulated kinase 7) - Mus musculus (Mouse)	MK15_MOUSE (+2)	61 kDa	Mus musculus	1
Glucose-6-phosphate isomerase (EC 5.3.1.9) (GPI) (Phosphoglucose isomerase) (PGI) (Phosphohexose isomerase) (PHI) (Neuroleukin) (NLK) - Mus musculus (Mouse)	G6PI_MOUSE	63 kDa	Mus musculus	1
Suppressor of cytokine signaling 7 (SOCS-7) - Mus musculus (Mouse)	SOCS7_MOUSE	63 kDa	Mus musculus	1
Ubiquitin carboxyl-terminal hydrolase 21 (EC 3.1.2.15) (Ubiquitin thioesterase 21) (Ubiquitin-specific-processing protease 21) (Deubiquitinating enzyme 21) - Mus musculus (Mouse)	UBP21_MOUSE	63 kDa	Mus musculus	1
Ran GTPase-activating protein 1 - Mus musculus (Mouse)	RGP1_MOUSE	64 kDa	Mus musculus	1
GTP-binding protein ARD-1 (ADP-ribosylation factor domain protein 1) (Tripartite motif-containing protein 23) - Mus musculus (Mouse)	ARD1_MOUSE (+1)	64 kDa	Mus musculus	1
Bifunctional purine biosynthesis protein PURH [Includes: Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3) (5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase) (AICAR transformylase); IMP cyclohydrolase (EC 3.5.4.10) (Inosinicase) (IMP synthetase) (ATIC)] - Mus musculus (Mouse)	PUR9_MOUSE	64 kDa	Mus musculus	1
Collagen alpha-2(IX) chain precursor - Mus musculus (Mouse)	CO9A2_MOUSE (+1)	65 kDa	Mus musculus	1
T-box transcription factor TBX18 (T-box protein 18) - Mus musculus (Mouse)	TBX18_MOUSE	65 kDa	Mus musculus	1
Transcription factor COE4 (Early B-cell factor 4) (EBF-4) (Olf-1/EBF-like 4) (OE-4) (O/E-4) - Mus musculus (Mouse)	COE4_MOUSE	65 kDa	Mus musculus	1
Atlastin-2 (ADP-ribosylation factor-like protein 6-interacting protein 2) - Mus musculus (Mouse)	ATLA2_MOUSE	66 kDa	Mus musculus	1
Zinc finger protein 354A (Transcription factor 17) (Kidney, ischemia, and developmentally-regulated protein 1) (Renal transcription factor Kid-1) - Mus musculus (Mouse)	Z354A_MOUSE	66 kDa	Mus musculus	1
Cytidine monophosphate-N-acetylneuraminic acid hydroxylase precursor (EC 1.14.18.2) (CMP-N-acetylneuraminate monooxygenase) (CMP-N-acetylneuraminic acid hydroxylase) (CMP-NeuAc hydroxylase) (CMP-Neu5Ac hydroxylase) - Mus musculus (Mouse)	CMAH_MOUSE	67 kDa	Mus musculus	1
SET domain-containing protein 3 (Endothelial differentiation inhibitory protein D10) - Mus musculus (Mouse)	SETD3_MOUSE (+1)	67 kDa	Mus musculus	1
Interleukin-1 receptor type I precursor (IL-1R-1) (p80) (CD121a antigen) - Mus musculus (Mouse)	IL1R1_MOUSE	67 kDa	Mus musculus	1
Lysyl-tRNA synthetase (EC 6.1.1.6) (Lysine--tRNA ligase) (LysRS) - Mus musculus (Mouse)	SYK_MOUSE	68 kDa	Mus musculus	1
Ezrin (p81) (Cytovillin) (Villin-2) - Mus musculus (Mouse)	EZRI_MOUSE (+2)	69 kDa	Mus musculus	1
Afamin precursor (Alpha-albumin) (Alpha-Alb) - Mus musculus (Mouse)	AFAM_MOUSE	69 kDa	Mus musculus	1

Zinc finger protein 667 (Myocardial ischemic preconditioning up-regulated protein 1) - Mus musculus (Mouse)	ZN667_MOUSE	70 kDa	Mus musculus	1
Poly(A)-specific ribonuclease PARN (EC 3.1.13.4) (Polyadenylate-specific ribonuclease) - Mus musculus (Mouse)	PARN_MOUSE	72 kDa	Mus musculus	1
Zinc finger protein 235 (Zinc finger protein 93) (Zfp-93) - Mus musculus (Mouse)	ZN235_MOUSE	73 kDa	Mus musculus	1
Protein KIAA0020 - Mus musculus (Mouse)	K0020_MOUSE	73 kDa	Mus musculus	1
F-box/LRR-repeat protein 19 (F-box and leucine-rich repeat protein 19) - Mus musculus (Mouse)	FXL19_MOUSE	74 kDa	Mus musculus	1
Zinc finger protein 710 - Mus musculus (Mouse)	ZN710_MOUSE	75 kDa	Mus musculus	1
Transcription factor 12 (Transcription factor HTF-4) (E-box-binding protein) (DNA-binding protein HTF4/TCF12) (Class A helix-loop-helix transcription factor ME1) - Mus musculus (Mouse)	HTF4_MOUSE	76 kDa	Mus musculus	1
Cirhin (Testis-expressed gene 292 protein) - Mus musculus (Mouse)	CIR1A_MOUSE (+4)	77 kDa	Mus musculus	1
Protein CASP - Mus musculus (Mouse)	CASP_MOUSE (+1)	77 kDa	Mus musculus	1
SH3 adapter protein SPIN90 (NCK-interacting protein with SH3 domain) (SH3 protein interacting with Nck, 90 kDa) (VacA-interacting protein, 54 kDa) (VIP54) (N-WASP-interacting protein, 90kDa) (Wiskott-Aldrich syndrome protein-binding protein WISH) (N-WASP-binding protein) - Mus musculus (Mouse)	SPN90_MOUSE	79 kDa	Mus musculus	1
Guanine nucleotide-binding protein G(I)/G(S)/G(O) gamma-4 subunit precursor - Mus musculus (Mouse)	GBG4_MOUSE	8 kDa	Mus musculus	1
RING finger protein C13orf7 homolog - Mus musculus (Mouse)	CM007_MOUSE	80 kDa	Mus musculus	1
G protein-regulated inducer of neurite outgrowth 3 (GRIN3) - Mus musculus (Mouse)	GRIN3_MOUSE (+1)	80 kDa	Mus musculus	1
Myeloperoxidase precursor (EC 1.11.1.7) (MPO) [Contains: Myeloperoxidase light chain; Myeloperoxidase heavy chain] - Mus musculus (Mouse)	PERM_MOUSE	81 kDa	Mus musculus	1
Poly(A) polymerase alpha (EC 2.7.7.19) (PAP) (Polynucleotide adenylyltransferase) - Mus musculus (Mouse)	PAPOA_MOUSE	82 kDa	Mus musculus	1
Tripartite motif-containing protein 42 - Mus musculus (Mouse)	TRI42_MOUSE	83 kDa	Mus musculus	1
ATP-dependent RNA helicase DDX1 (EC 3.6.1.-) (DEAD box protein 1) - Mus musculus (Mouse)	DDX1_MOUSE	83 kDa	Mus musculus	1
Tripartite motif-containing protein 46 (Tripartite, fibronectin type-III and C-terminal SPRY motif protein) - Mus musculus (Mouse)	TRI46_MOUSE	83 kDa	Mus musculus	1
Catenin beta-1 (Beta-catenin) - Mus musculus (Mouse)	CTNB1_MOUSE (+2)	85 kDa	Mus musculus	1
Leucine-rich repeat-containing protein 41 (Protein Muf1) - Mus musculus (Mouse)	LRC41_MOUSE (+2)	86 kDa	Mus musculus	1
Myb-related protein A (A-Myb) - Mus musculus (Mouse)	MYBA_MOUSE (+2)	86 kDa	Mus musculus	1
Ribosomal protein S6 kinase alpha-6 (EC 2.7.11.1) (S6K-alpha 6) - Mus musculus (Mouse)	KS6A6_MOUSE	87 kDa	Mus musculus	1
Cullin-4A (CUL-4A) - Mus musculus (Mouse)	CUL4A_MOUSE (+1)	88 kDa	Mus musculus	1
RelA-associated inhibitor (Inhibitor of ASPP protein) (Protein iASPP) (PPP1R13B-like protein) (NFkB-interacting protein 1) - Mus musculus (Mouse)	IASPP_MOUSE	89 kDa	Mus musculus	1

Transmembrane and tetratricopeptide repeat containing 2 - Mus musculus (Mouse)	TMTC2_MOUSE	94 kDa	Mus musculus	1
Lymphocyte-specific helicase (EC 3.6.1.-) (Proliferation associated SNF2-like protein) - Mus musculus (Mouse)	HELLS_MOUSE	95 kDa	Mus musculus	1
ATP-binding cassette sub-family F member 1 - Mus musculus (Mouse)	ABCF1_MOUSE	95 kDa	Mus musculus	1
Programmed cell death 6-interacting protein (ALG-2-interacting protein X) (ALG-2-interacting protein 1) (E2F1-inducible protein) (Eig2) - Mus musculus (Mouse)	PDC6I_MOUSE	96 kDa	Mus musculus	1
Glycogen phosphorylase, muscle form (EC 2.4.1.1) (Myophosphorylase) - Mus musculus (Mouse)	PYGM_MOUSE	97 kDa	Mus musculus	1
Kinase suppressor of ras-1 (Kinase suppressor of ras) (mKSR1) (Hb protein) - Mus musculus (Mouse)	KSR1_MOUSE	97 kDa	Mus musculus	1
Lethal(3)malignant brain tumor-like 3 protein (L(3)mbt-like 3 protein) - Mus musculus (Mouse)	LMBL3_MOUSE	99 kDa	Mus musculus	1

Analysed with a protein threshold at 20% and a minimum peptide threshold of 20% and 1 peptide.

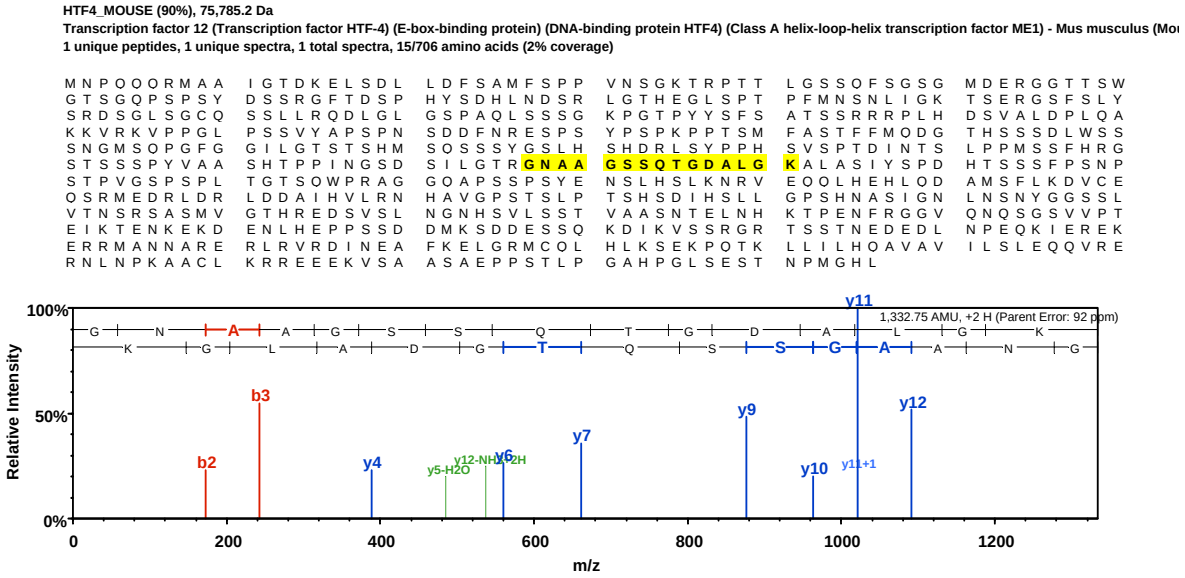


Figure A 1.7 Mass spectra data for the immunoprecipitated protein TCF12 (HTF4) with ID4 from the Ba/F3 ID4 clone 11.

Table A1.9: Mass spectral analysis of the TCF12 peptide recognised for ID4 in Figure A1.7.

	Molecular weight	Peptides	Peptide	Mascot score	M/z	Charge
Transcription factor 12 (TCF12)	76 kDa	1	R)GNAAGSSQTGDALGK(A)	65.88	667.4	2

Table A1.10 List of proteins from an IP performed in NALM6 TR cells using the anti myc tag antibody and a cytosolic preparation.

Identified Proteins (72)	Accession Number	Molecular Weight	Taxonomy	Number of unique peptides
Tubulin beta chain - Homo sapiens (Human)	TBB5_HUMAN	50 kDa	Homo sapiens	6
Vimentin - Homo sapiens (Human)	VIME_HUMAN	54 kDa	Homo sapiens	11
Tubulin alpha-1A chain - Homo sapiens (Human)	TBA1A_HUMAN (+2)	50 kDa	Homo sapiens	4
Nucleophosmin - Homo sapiens (Human)	NPM_HUMAN	33 kDa	Homo sapiens	3
Transcription factor E2-alpha - Homo sapiens (Human)	TFE2_HUMAN	68 kDa	unknown	4
Cardiotrophin-like cytokine factor 1 precursor - Homo sapiens (Human)	CLCF1_HUMAN	25 kDa	Homo sapiens	1
Brain-specific angiogenesis inhibitor 1 precursor - Homo sapiens (Human)	BAI1_HUMAN	174 kDa	Homo sapiens	1
Heat shock cognate 71 kDa protein - Homo sapiens (Human)	HSP7C_HUMAN	71 kDa	Homo sapiens	2
Peroxiredoxin-1 - Homo sapiens (Human)	PRDX1_HUMAN	22 kDa	Homo sapiens	3
Myosin light chain 2, lymphocyte-specific - Homo sapiens (Human)	MYLPL_HUMAN	25 kDa	Homo sapiens	1
Amyotrophic lateral sclerosis 2 chromosomal region candidate gene 13 protein - Homo sapiens (Human)	AL2SC_HUMAN	62 kDa	unknown	1
Glyceraldehyde-3-phosphate dehydrogenase - Homo sapiens (Human)	G3P_HUMAN	36 kDa	Homo sapiens	3
Trypsin-2 precursor - Homo sapiens (Human)	TRY2_HUMAN (+2)	26 kDa	Homo sapiens	1
Elongation factor 1-alpha 1 - Homo sapiens (Human)	EF1A1_HUMAN (+2)	50 kDa	Homo sapiens	1
Pyruvate kinase isozymes M1/M2 - Homo sapiens (Human)	KPYM_HUMAN (+1)	58 kDa	Homo sapiens	1
T-cell activation Rho GTPase-activating protein - Homo sapiens (Human)	TAGAP_HUMAN	81 kDa	Homo sapiens	1
40S ribosomal protein SA - Homo sapiens (Human)	RSSA_HUMAN	33 kDa	Homo sapiens	1
60S ribosomal protein L4 - Homo sapiens (Human)	RL4_HUMAN	48 kDa	unknown	2
Plasminogen activator inhibitor 1 RNA-binding protein - Homo sapiens (Human)	PAIRB_HUMAN	45 kDa	Homo sapiens	2
RING finger protein C1orf164 - Homo sapiens (Human)	CA164_HUMAN	63 kDa	unknown	2
Nuclease-sensitive element-binding protein 1 - Homo sapiens (Human)	YBOX1_HUMAN	36 kDa	Homo sapiens	2
Keratin, type II cytoskeletal 1 - Homo sapiens (Human)	K2C1_HUMAN (+1)	66 kDa	Homo sapiens	1
Keratin, type I cytoskeletal 10 - Homo sapiens (Human)	K1C10_HUMAN (+6)	60 kDa	Homo sapiens	1
Sulfotransferase 1C2 - Homo sapiens (Human)	ST1C2_HUMAN	35 kDa	unknown	1
Heterogeneous nuclear ribonucleoprotein K - Homo sapiens (Human)	HNRPK_HUMAN	51 kDa	Homo sapiens	1
60S ribosomal protein L13 - Homo sapiens (Human)	RL13_HUMAN	24 kDa	unknown	1
Alkyldihydroxyacetonephosphate synthase, peroxisomal precursor - Homo sapiens (Human)	ADAS_HUMAN	73 kDa	Homo sapiens	1
Transcription factor 4 - Homo sapiens (Human)	ITF2_HUMAN	71 kDa	Homo sapiens	1
Carboxypeptidase O precursor - Homo sapiens (Human)	CBPO_HUMAN	43 kDa	unknown	1
Lymphoid-restricted membrane protein - Homo sapiens (Human)	LRMP_HUMAN	62 kDa	unknown	1
Inositol monophosphatase 2 - Homo sapiens (Human)	IMPA2_HUMAN	31 kDa	unknown	1
Cystinosin - Homo sapiens (Human)	CTNS_HUMAN (+3)	42 kDa	unknown	1
Heterogeneous nuclear ribonucleoprotein G - Homo sapiens (Human)	HNRPG_HUMAN (+2)	42 kDa	Homo sapiens	1

Regulator of G-protein signaling 9 - Homo sapiens (Human)	RGS9_HUMAN	77 kDa	unknown	1
Ras GTPase-activating protein-binding protein 1 - Homo sapiens (Human)	G3BP1_HUMAN	52 kDa	Homo sapiens	1
Proliferation-associated protein 2G4 - Homo sapiens (Human)	PA2G4_HUMAN	44 kDa	Homo sapiens	1
Oligoribonuclease, mitochondrial precursor - Homo sapiens (Human)	ORN_HUMAN	27 kDa	unknown	1
Brain-specific angiogenesis inhibitor 1-associated protein 2-like protein 2 - Homo sapiens (Human)	BI2L2_HUMAN (+3)	59 kDa	Homo sapiens	1
Transmembrane protein 97 - Homo sapiens (Human)	TMM97_HUMAN	21 kDa	Homo sapiens	1
Histone H1x - Homo sapiens (Human)	H1X_HUMAN (+1)	22 kDa	Homo sapiens	1
Dedicator of cytokinesis protein 4 - Homo sapiens (Human)	DOCK4_HUMAN	225 kDa	unknown	1
SECIS-binding protein 2 - Homo sapiens (Human)	SEBP2_HUMAN	95 kDa	unknown	1
Mediator of RNA polymerase II transcription subunit 12 - Homo sapiens (Human)	MED12_HUMAN	243 kDa	unknown	1
Pleckstrin homology-like domain family A member 2 - Homo sapiens (Human)	PHLA2_HUMAN	17 kDa	Homo sapiens	1
Protein asteroid homolog 1 - Homo sapiens (Human)	ASTE1_HUMAN	77 kDa	Homo sapiens	1
Autophagy-related protein 9B - Homo sapiens (Human)	ATG9B_HUMAN (+4)	101 kDa	unknown	1
Interleukin-25 precursor - Homo sapiens (Human)	IL25_HUMAN	20 kDa	Homo sapiens	1
Pre-mRNA-processing factor 17 - Homo sapiens (Human)	PRP17_HUMAN	66 kDa	Homo sapiens	1
Phosphoinositide 3-kinase regulatory subunit 5 - Homo sapiens (Human)	PI3R5_HUMAN	97 kDa	Homo sapiens	1
Ferrochelatase, mitochondrial precursor - Homo sapiens (Human)	HEMH_HUMAN	48 kDa	Homo sapiens	1
Protocadherin-17 precursor - Homo sapiens (Human)	PCD17_HUMAN	126 kDa	Homo sapiens	1
Splicing factor, arginine/serine-rich 10 - Homo sapiens (Human)	TRA2B_HUMAN	34 kDa	Homo sapiens	1
CCR4-NOT transcription complex subunit 1 - Homo sapiens (Human)	CNOT1_HUMAN	267 kDa	Homo sapiens	1
Neurobeachin - Homo sapiens (Human)	NBEA_HUMAN	328 kDa	Homo sapiens	1
Phosphofurin acidic cluster sorting protein 2 - Homo sapiens (Human)	PACS2_HUMAN (+2)	98 kDa	Homo sapiens	1
Claudin-12 - Homo sapiens (Human)	CLD12_HUMAN	27 kDa	unknown	1
RING finger and SPRY domain-containing protein 1 precursor - Homo sapiens (Human)	RSPRY_HUMAN	64 kDa	unknown	1
Ankyrin repeat domain-containing protein 17 - Homo sapiens (Human)	ANR17_HUMAN	274 kDa	unknown	1
Transmembrane and TPR repeat-containing protein 3 - Homo sapiens (Human)	TMTC3_HUMAN (+4)	104 kDa	unknown	1
Cell division protein kinase 4 - Homo sapiens (Human)	CDK4_HUMAN	34 kDa	Homo sapiens	1
Kynurenine--oxoglutarate transaminase 3 - Homo sapiens (Human)	KAT3_HUMAN	51 kDa	unknown	1
Insulin receptor substrate 4 - Homo sapiens (Human)	IRS4_HUMAN	134 kDa	unknown	1
Proteasome activator complex subunit 1 - Homo sapiens (Human)	PSME1_HUMAN	29 kDa	Homo sapiens	1
Schlafen family member 11 - Homo sapiens (Human)	SLN11_HUMAN	103 kDa	Homo sapiens	1
Zinc finger protein 148 - Homo sapiens (Human)	ZN148_HUMAN	89 kDa	unknown	1
Histone acetyltransferase PCAF - Homo sapiens (Human)	PCAF_HUMAN	93 kDa	Homo sapiens	1
Ras-related protein Rab-3C - Homo sapiens (Human)	RAB3C_HUMAN	26 kDa	unknown	1
Sperm-associated antigen 1 - Homo sapiens (Human)	SPAG1_HUMAN	104 kDa	Homo sapiens	1

E3 ubiquitin-protein ligase DZIP3 - Homo sapiens (Human)	DZIP3_HUMAN (+1)	139 kDa	unknown	1
Plakophilin-2 - Homo sapiens (Human)	PKP2_HUMAN	97 kDa	Homo sapiens	1
Fatty aldehyde dehydrogenase - Homo sapiens (Human)	AL3A2_HUMAN (+6)	55 kDa	unknown	1
Myelodysplasia syndrome 1 protein - Homo sapiens (Human)	MDS1_HUMAN	19 kDa	unknown	1

Analysed with a protein threshold at 20% and a minimum peptide threshold of 20% and 1 peptide.

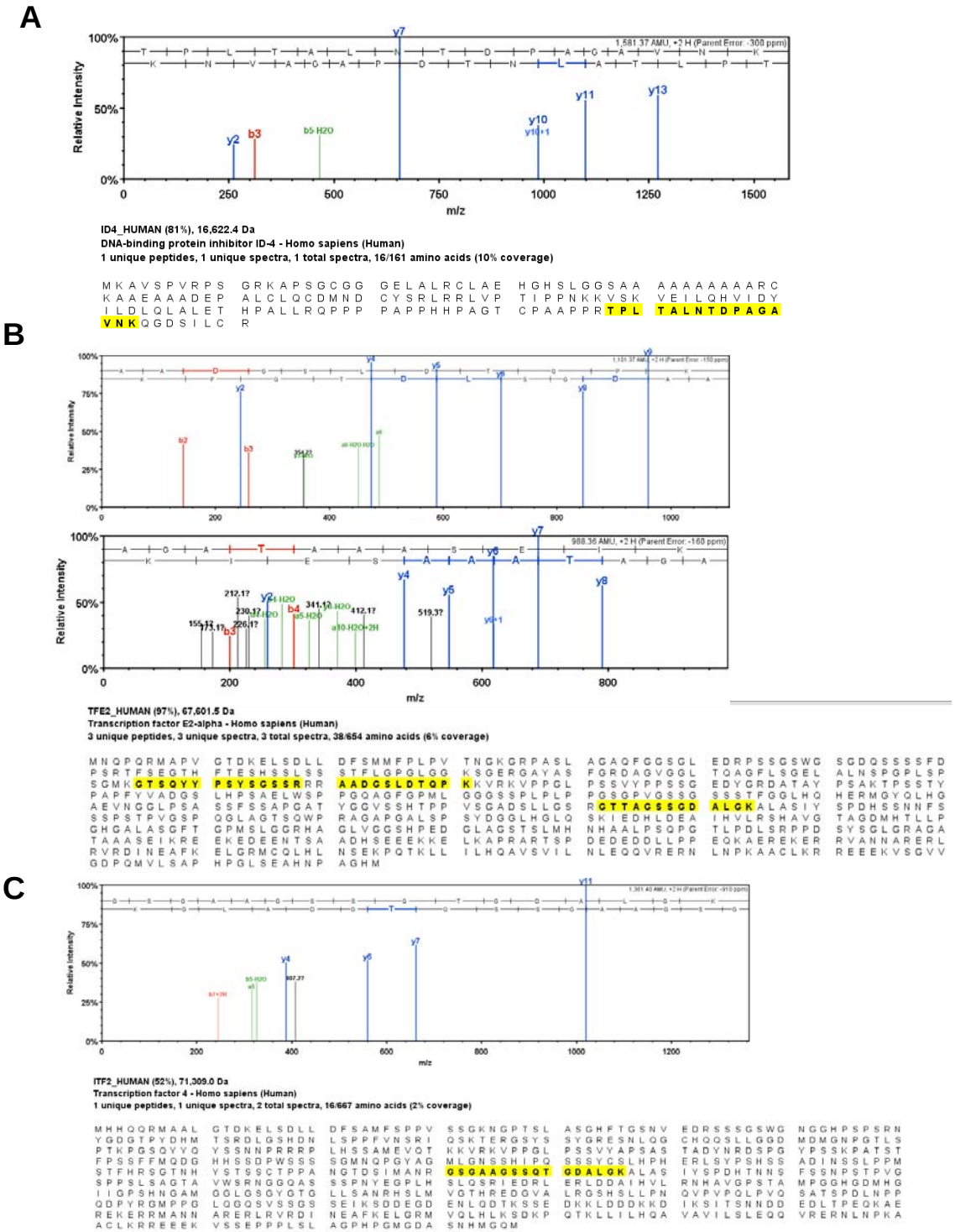


Figure A1.8 Mass spectra data for the immunoprecipitated proteins **A)** ID4 **B)** E2A and **C)** TCF4 (ITF2), with ID4 from the NALM-6 TR cells with induced ID4 expression.

Table A1.11 Mass spectral analysis of the ID4, E2A and TCF12 peptides recognised for in Figure A1.8.

	Molecular weight	Peptides	Peptide	Mascot score	M/z	Charge
DNA-binding protein inhibitor ID-4	17 kDa	1	(R)TPLTALNTDPAGAVNK(Q)	37.8	791.69	2
Transcription factor E2-alpha	68 kDa	2	(R)AGATAAASEIK(R)	45.5	495.19	2
			(R)AADGSLDTQPK(K)	54	51.69	2
Transcription factor 4	71 kDa	1	R)GSGAAGSSQTGDALGK(A)	29.7	681.71	2

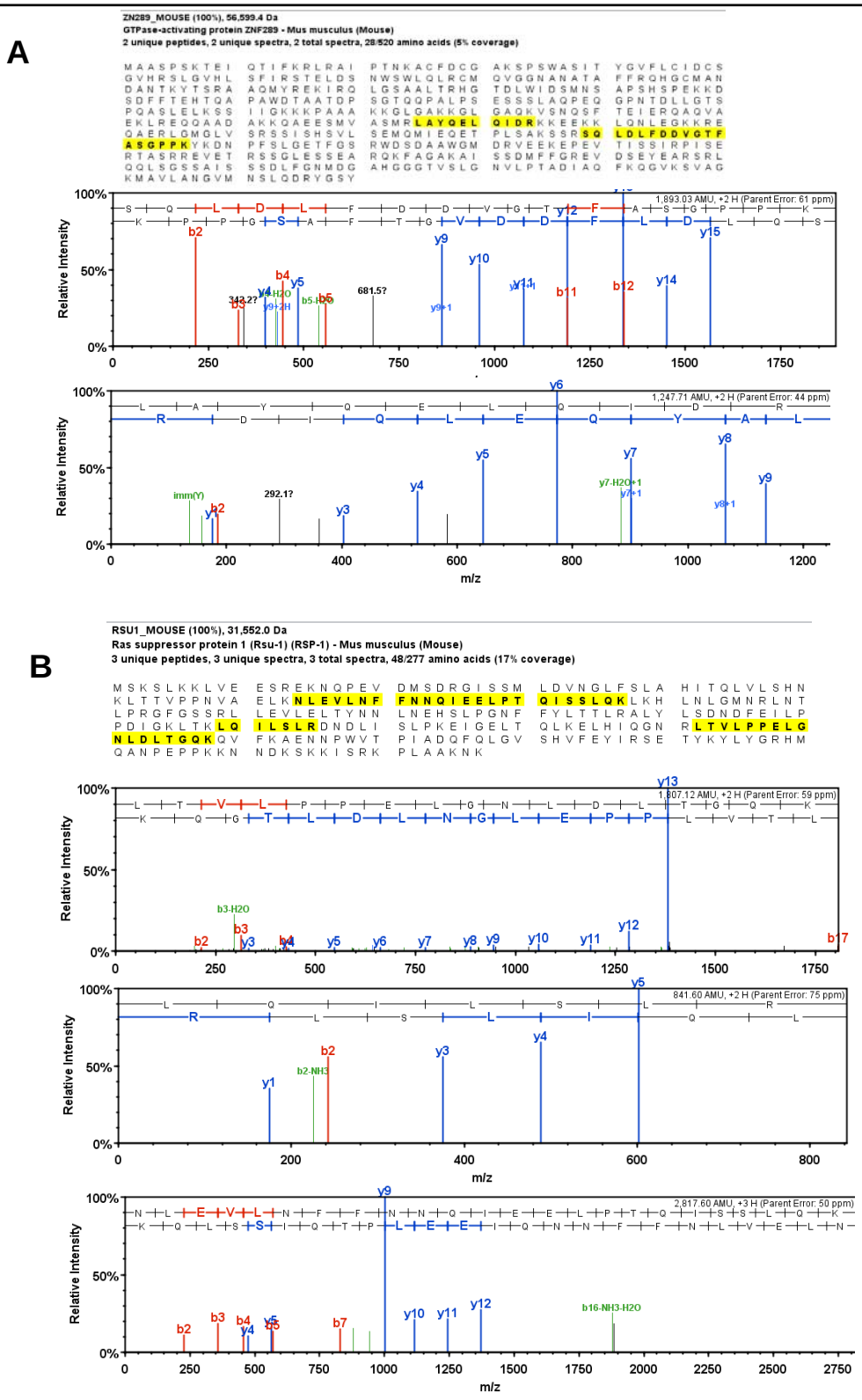


Figure A 1.9 Mass spectra data for the immunoprecipitated proteins **A)** ZNF289 and **B)** RSU_1with ID4 from the Ba/F3 ID4 clone 11.

Table A1.12: Mass spectral analysis of the peptides recognised for ID4 in Figure A1.9.

	Molecular weight	Peptides	Peptide	Mascot score	M/z	Charge
GTPase-activating protein ZNF289 - Mus musculus (Mouse)	57 kDa	2	R)LAYQELQIDR(K)	61.47	624.9	2
			(R)SQLDLFDDVGTFASGPPK(Y)	80.88	947.5	2
Ras suppressor protein 1 (Rsu-1) (RSP-1) - Mus musculus (Mouse)	32 kDa	3	K)LQILSLR(D)	33.09	421.8	2
			R)LTVLPPELGNLDLTGQK(Q)	71.29	904.6	2
			(K)NLEVLNFFNNQIEELPTQISSLQK(L)	52.44	940.2	2

References

Adams JM, Harris AW, Pinkert CA, Corcoran LM, Alexander WS, Cory S *et al* (1985). The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature* **318:** 533-8.

Ahlbom A, Day N, Feychting M, Roman E, Skinner J, Dockerty J *et al* (2000). A pooled analysis of magnetic fields and childhood leukaemia. *Br J Cancer* **83:** 692-8.

Akasaka T, Akasaka H, Yonetani N, Ohno H, Yamabe H, Fukuhara S *et al* (1998). Refinement of the BCL2/immunoglobulin heavy chain fusion gene in t(14;18)(q32;q21) by polymerase chain reaction amplification for long targets. *Genes Chromosomes Cancer* **21:** 17-29.

Andersson A, Ritz C, Lindgren D, Eden P, Lassen C, Heldrup J *et al* (2007). Microarray-based classification of a consecutive series of 121 childhood acute leukemias: prediction of leukemic and genetic subtype as well as of minimal residual disease status. *Leukemia* **21:** 1198-203.

Andrae J, Gallini R, Betsholtz C (2008). Role of platelet-derived growth factors in physiology and medicine. *Genes Dev* **22:** 1276-312.

Armstrong SA, Look AT (2005). Molecular genetics of acute lymphoblastic leukemia. *J Clin Oncol* **23:** 6306-15.

Ashida S, Furihata M, Katagiri T, Tamura K, Anazawa Y, Yoshioka H *et al* (2006). Expression of novel molecules, MICAL2-PV (MICAL2 prostate cancer variants), increases with high Gleason score and prostate cancer progression. *Clin Cancer Res* **12:** 2767-73.

Astrakhan A, Omori M, Nguyen T, Becker-Herman S, Iseki M, Aye T *et al* (2007). Local increase in thymic stromal lymphopoietin induces systemic alterations in B cell development. *Nat Immunol* **8:** 522-31.

Bain G, Maandag EC, Izon DJ, Amsen D, Kruisbeek AM, Weintraub BC *et al* (1994). E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. *Cell* **79:** 885-92.

Bain G, Romanow WJ, Albers K, Havran WL, Murre C (1999). Positive and negative regulation of V(D)J recombination by the E2A proteins. *J Exp Med* **189:** 289-300.

Barbaric D, Byth K, Dalla-Pozza L, Byrne JA (2006). Expression of tumor protein D52-like genes in childhood leukemia at diagnosis: clinical and sample considerations. *Leuk Res* **30:** 1355-63.

Barndt RJ, Dai M, Zhuang Y (2000). Functions of E2A-HEB heterodimers in T-cell development revealed by a dominant negative mutation of HEB. *Mol Cell Biol* **20:** 6677-85.

Barnhart BC, Lee JC, Alappat EC, Peter ME (2003). The death effector domain protein family. *Oncogene* **22:** 8634-44.

Bateman CM, Colman SM, Chaplin T, Young BD, Eden TO, Bhakta M *et al* (2010) Acquisition of genome-wide copy number alterations in monozygotic twins with acute lymphoblastic leukemia. *Blood* **115:** 3553-8.

Becker-Herman S, Lantner F, Shachar I (2002). Id2 negatively regulates B cell differentiation in the spleen. *J Immunol* **168:** 5507-13.

Bedford L, Walker R, Kondo T, van Cruchten I, King ER, Sablitzky F (2005). Id4 is required for the correct timing of neural differentiation. *Dev Biol* **280:** 386-95.

Beger C, Pierce LN, Kruger M, Marcusson EG, Robbins JM, Welcsh P *et al* (2001). Identification of Id4 as a regulator of BRCA1 expression by using a ribozyme-library-based inverse genomics approach. *Proc Natl Acad Sci U S A* **98:** 130-5.

Bellido M, Aventin A, Lasa A, Estivill C, Carnicer MJ, Pons C *et al* (2003). Id4 is deregulated by a t(6;14)(p22;q32) chromosomal translocation in a B-cell lineage acute lymphoblastic leukemia. *Haematologica* **88:** 994-1001.

Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61:** 49-59.

Berse M, Bounpheng M, Huang X, Christy B, Pollmann C, Dubiel W (2004). Ubiquitin-dependent degradation of Id1 and Id3 is mediated by the COP9 signalosome. *J Mol Biol* **343:** 361-70.

Bhojwani D, Kang H, Menezes RX, Yang W, Sather H, Moskowitz NP *et al* (2008). Gene expression signatures predictive of early response and outcome in high-risk childhood acute lymphoblastic leukemia: A Children's Oncology Group Study [corrected]. *J Clin Oncol* **26:** 4376-84.

Bhojwani D, Kang H, Moskowitz NP, Min DJ, Lee H, Potter JW *et al* (2006). Biologic pathways associated with relapse in childhood acute lymphoblastic leukemia: a Children's Oncology Group study. *Blood* **108:** 711-7.

Biondi A, Valsecchi MG, Seriu T, D'Aniello E, Willemse MJ, Fasching K *et al* (2000). Molecular detection of minimal residual disease is a strong predictive factor of relapse in childhood B-lineage acute lymphoblastic leukemia with medium risk features. A case control study of the International BFM study group. *Leukemia* **14:** 1939-43.

Bounpheng MA, Dimas JJ, Dodds SG, Christy BA (1999). Degradation of Id proteins by the ubiquitin-proteasome pathway. *FASEB J* **13:** 2257-64.

Boutros R, Byrne JA (2005). D53 (TPD52L1) is a cell cycle-regulated protein maximally expressed at the G2-M transition in breast cancer cells. *Exp Cell Res* **310:** 152-65.

Brunner C, Laumen H, Nielsen PJ, Kraut N, Wirth T (2003). Expression of the aldehyde dehydrogenase 2-like gene is controlled by BOB.1/OBF.1 in B lymphocytes. *J Biol Chem* **278:** 45231-9.

Bueno MJ, Gomez de Cedron M, Laresgoiti U, Fernandez-Piqueras J, Zubiaga A, Malumbres M (2010). Multiple E2F-induced microRNAs prevent replicative stress in response to mitogenic signalling. *Mol Cell Biol*.

Carey JP, Asirvatham AJ, Galm O, Ghogomu TA, Chaudhary J (2009). Inhibitor of differentiation 4 (Id4) is a potential tumor suppressor in prostate cancer. *BMC Cancer* **9:** 173.

Cario G, Stanulla M, Fine BM, Teuffel O, Neuhoff NV, Schrauder A *et al* (2005). Distinct gene expression profiles determine molecular treatment response in childhood acute lymphoblastic leukemia. *Blood* **105:** 821-6.

Cario G, Zimmermann M, Romey R, Gesk S, Vater I, Harbott J *et al* (2010). Presence of the P2RY8-CRLF2 rearrangement is associated with a poor prognosis in non-high-risk precursor B-cell acute lymphoblastic leukemia in children treated according to the ALL-BFM 2000 protocol. *Blood*.

Chae HJ, Kang JS, Byun JO, Han KS, Kim DU, Oh SM *et al* (2000). Molecular mechanism of staurosporine-induced apoptosis in osteoblasts. *Pharmacol Res* **42:** 373-81.

Chassot AA, Turchi L, Virolle T, Fitsialos G, Batoz M, Deckert M *et al* (2007). Id3 is a novel regulator of p27kip1 mRNA in early G1 phase and is required for cell-cycle progression. *Oncogene* **26:** 5772-83.

Christy BA, Sanders LK, Lau LF, Copeland NG, Jenkins NA, Nathans D (1991). An Id-related helix-loop-helix protein encoded by a growth factor-inducible gene. *Proc Natl Acad Sci U S A* **88:** 1815-9.

Cisse B, Caton ML, Lehner M, Maeda T, Scheu S, Locksley R *et al* (2008). Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development. *Cell* **135:** 37-48.

Cobaleda C, Sanchez-Garcia I (2009). B-cell acute lymphoblastic leukaemia: towards understanding its cellular origin. *Bioessays* **31:** 600-9.

Cobaleda C, Schebesta A, Delogu A, Busslinger M (2007). Pax5: the guardian of B cell identity and function. *Nat Immunol* **8:** 463-70.

Cochrane SW, Zhao Y, Welner RS, Sun XH (2009). Balance between Id and E proteins regulates myeloid-versus-lymphoid lineage decisions. *Blood* **113:** 1016-26.

Cooper CL, Brady G, Bilia F, Iscove NN, Quesenberry PJ (1997). Expression of the Id family helix-loop-helix regulators during growth and development in the hematopoietic system. *Blood* **89:** 3155-65.

Curry JD, Glaser MC, Smith MT (2001). Real-time reverse transcription polymerase chain reaction detection and quantification of t(1;19) (E2A-PBX1) fusion genes associated with leukaemia. *Br J Haematol* **115:** 826-30.

de Candia P, Akram M, Benezra R, Brogi E (2006). Id4 messenger RNA and estrogen receptor expression: inverse correlation in human normal breast epithelium and carcinoma. *Hum Pathol* **37:** 1032-41.

Deed RW, Armitage S, Norton JD (1996). Nuclear localization and regulation of Id protein through an E protein-mediated chaperone mechanism. *J Biol Chem* **271:** 23603-6.

Deed RW, Hara E, Atherton GT, Peters G, Norton JD (1997). Regulation of Id3 cell cycle function by Cdk-2-dependent phosphorylation. *Mol Cell Biol* **17:** 6815-21.

Deininger M, Buchdunger E, Druker BJ (2005). The development of imatinib as a therapeutic agent for chronic myeloid leukemia. *Blood* **105:** 2640-53.

Dobreva G, Dambacher J, Grosschedl R (2003). SUMO modification of a novel MAR-binding protein, SATB2, modulates immunoglobulin mu gene expression. *Genes Dev* **17:** 3048-61.

Dorshkind K, Montecino-Rodriguez E (2007). Fetal B-cell lymphopoiesis and the emergence of B-1-cell potential. *Nat Rev Immunol* **7:** 213-9.

Dougherty GW, Jose C, Gimona M, Cutler ML (2008). The Rsu-1-PINCH1-ILK complex is regulated by Ras activation in tumor cells. *Eur J Cell Biol* **87:** 721-34.

Dyer MJ, Akasaka T, Capasso M, Dusanjh P, Lee YF, Karran EL *et al* (2010). Immunoglobulin heavy chain locus chromosomal translocations in B-cell precursor acute lymphoblastic leukemia: rare clinical curios or potent genetic drivers? *Blood* **115:** 1490-9.

Engel I, Murre C (1999). Ectopic expression of E47 or E12 promotes the death of E2A-deficient lymphomas. *Proc Natl Acad Sci U S A* **96:** 996-1001.

Fajerman I, Schwartz AL, Ciechanover A (2004). Degradation of the Id2 developmental regulator: targeting via N-terminal ubiquitination. *Biochem Biophys Res Commun* **314:** 505-12.

Fielding AK, Richards SM, Chopra R, Lazarus HM, Litzow MR, Buck G *et al* (2007). Outcome of 609 adults after relapse of acute lymphoblastic leukemia (ALL); an MRC UKALL12/ECOG 2993 study. *Blood* **109:** 944-50.

Fontemaggi G, Dell'Orso S, Trisciuoglio D, Shay T, Melucci E, Fazi F *et al* (2009). The execution of the transcriptional axis mutant p53, E2F1 and ID4 promotes tumor neo-angiogenesis. *Nat Struct Mol Biol* **16:** 1086-93.

Ford AM, Palmi C, Bueno C, Hong D, Cardus P, Knight D *et al* (2009). The TEL-AML1 leukemia fusion gene dysregulates the TGF-beta pathway in early B lineage progenitor cells. *J Clin Invest* **119:** 826-36.

Fraidenraich D, Stillwell E, Romero E, Wilkes D, Manova K, Basson CT *et al* (2004). Rescue of cardiac defects in id knockout embryos by injection of embryonic stem cells. *Science* **306:** 247-52.

Giacinti C, Giordano A (2006). RB and cell cycle progression. *Oncogene* **25:** 5220-7.

Gil J, Peters G (2006). Regulation of the INK4b-ARF-INK4a tumour suppressor locus: all for one or one for all. *Nat Rev Mol Cell Biol* **7:** 667-77.

Glotzer M, Murray AW, Kirschner MW (1991). Cyclin is degraded by the ubiquitin pathway. *Nature* **349:** 132-8.

Gomez Del Pulgar T, Valdes-Mora F, Bandres E, Perez-Palacios R, Espina C, Cejas P *et al* (2008). Cdc42 is highly expressed in colorectal adenocarcinoma and downregulates ID4 through an epigenetic mechanism. *Int J Oncol* **33:** 185-93.

Gonda H, Sugai M, Nambu Y, Katakai T, Agata Y, Mori KJ *et al* (2003). The balance between Pax5 and Id2 activities is the key to AID gene expression. *J Exp Med* **198:** 1427-37.

Graf T, Enver T (2009). Forcing cells to change lineages. *Nature* **462:** 587-94.

Greaves M (2006). Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer* **6:** 193-203.

Hacker C, Kirsch RD, Ju XS, Hieronymus T, Gust TC, Kuhl C *et al* (2003). Transcriptional profiling identifies Id2 function in dendritic cell development. *Nat Immunol* **4:** 380-6.

Hanahan D, Weinberg RA (2000). The hallmarks of cancer. *Cell* **100:** 57-70.

Hara E, Hall M, Peters G (1997). Cdk2-dependent phosphorylation of Id2 modulates activity of E2A-related transcription factors. *EMBO J* **16:** 332-42.

Hara E, Yamaguchi T, Nojima H, Ide T, Campisi J, Okayama H *et al* (1994). Id-related genes encoding helix-loop-helix proteins are required for G1 progression and are repressed in senescent human fibroblasts. *J Biol Chem* **269:** 2139-45.

Hardy RR, Hayakawa K (2001). B cell development pathways. *Annu Rev Immunol* **19:** 595-621.

Harvey RC, Mullighan CG, Chen IM, Wharton W, Mikhail FM, Carroll AJ *et al* (2010). Rearrangement of CRLF2 is associated with mutation of JAK kinases, alteration of IKZF1, Hispanic/Latino ethnicity, and a poor outcome in pediatric B-progenitor acute lymphoblastic leukemia. *Blood*.

Hjalgrim LL, Westergaard T, Rostgaard K, Schmiegelow K, Melbye M, Hjalgrim H *et al* (2003). Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies. *Am J Epidemiol* **158:** 724-35.

Holleman A, Cheok MH, den Boer ML, Yang W, Veerman AJ, Kazemier KM *et al* (2004). Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. *N Engl J Med* **351:** 533-42.

Hong D, Gupta R, Ancliff P, Atzberger A, Brown J, Soneji S *et al* (2008). Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. *Science* **319:** 336-9.

Hsu LY, Lauring J, Liang HE, Greenbaum S, Cado D, Zhuang Y *et al* (2003). A conserved transcriptional enhancer regulates RAG gene expression in developing B cells. *Immunity* **19:** 105-17.

Hudlebusch HR, Theilgaard-Monch K, Lodahl M, Johnsen HE, Rasmussen T (2005). Identification of ID-1 as a potential target gene of MMSET in multiple myeloma. *Br J Haematol* **130:** 700-8.

Hurwitz R, Hozier J, LeBien T, Minowada J, Gajl-Peczalska K, Kubonishi I *et al* (1979). Characterization of a leukemic cell line of the pre-B phenotype. *Int J Cancer* **23:** 174-80.

Iavarone A, Garg P, Lasorella A, Hsu J, Israel MA (1994). The helix-loop-helix protein Id-2 enhances cell proliferation and binds to the retinoblastoma protein. *Genes Dev* **8:** 1270-84.

Jack I, Seshadri R, Garson M, Michael P, Callen D, Zola H *et al* (1986). RCH-ACV: a lymphoblastic leukemia cell line with chromosome translocation 1;19 and trisomy 8. *Cancer Genet Cytogenet* **19:** 261-9.

Jankovic V, Ciarrocchi A, Boccuni P, DeBlasio T, Benezra R, Nimer SD (2007). Id1 restrains myeloid commitment, maintaining the self-renewal capacity of hematopoietic stem cells. *Proc Natl Acad Sci U S A* **104:** 1260-5.

Jen Y, Manova K, Benezra R (1996). Expression patterns of Id1, Id2, and Id3 are highly related but distinct from that of Id4 during mouse embryogenesis. *Dev Dyn* **207:** 235-52.

Jeon HM, Jin X, Lee JS, Oh SY, Sohn YW, Park HJ *et al* (2008). Inhibitor of differentiation 4 drives brain tumor-initiating cell genesis through cyclin E and notch signaling. *Genes Dev* **22:** 2028-33.

Ji M, Li H, Suh HC, Klarmann KD, Yokota Y, Keller JR (2008). Id2 intrinsically regulates lymphoid and erythroid development via interaction with different target proteins. *Blood* **112:** 1068-77.

Jia J, Dai M, Zhuang Y (2008). E proteins are required to activate germline transcription of the TCR Vbeta8.2 gene. *Eur J Immunol* **38:** 2806-20.

Jung D, Giallourakis C, Mostoslavsky R, Alt FW (2006). Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol* **24:** 541-70.

Kee BL (2005). Id3 induces growth arrest and caspase-2-dependent apoptosis in B lymphocyte progenitors. *J Immunol* **175:** 4518-27.

Kee BL (2009). E and ID proteins branch out. *Nat Rev Immunol* **9:** 175-84.

Kee BL, Rivera RR, Murre C (2001). Id3 inhibits B lymphocyte progenitor growth and survival in response to TGF-beta. *Nat Immunol* **2:** 242-7.

Kersten C, Dosen G, Myklebust JH, Sivertsen EA, Hystad ME, Smeland EB *et al* (2006). BMP-6 inhibits human bone marrow B lymphopoiesis--upregulation of Id1 and Id3. *Exp Hematol* **34:** 72-81.

Kersten C, Sivertsen EA, Hystad ME, Forfang L, Smeland EB, Myklebust JH (2005). BMP-6 inhibits growth of mature human B cells; induction of Smad phosphorylation and upregulation of Id1. *BMC Immunol* **6:** 9.

Kinlen L (2004). Infections and immune factors in cancer: the role of epidemiology. *Oncogene* **23:** 6341-8.

Kohno T, Yokota J (2006). Molecular processes of chromosome 9p21 deletions causing inactivation of the p16 tumor suppressor gene in human cancer:

deduction from structural analysis of breakpoints for deletions. *DNA Repair (Amst)* **5:** 1273-81.

Komor M, Guller S, Baldus CD, de Vos S, Hoelzer D, Ottmann OG *et al* (2005). Transcriptional profiling of human hematopoiesis during in vitro lineage-specific differentiation. *Stem Cells* **23:** 1154-69.

Kremer BE, Adang LA, Macara IG (2007). Septins regulate actin organization and cell-cycle arrest through nuclear accumulation of NCK mediated by SOCS7. *Cell* **130:** 837-50.

Kurooka H, Yokota Y (2005). Nucleo-cytoplasmic shuttling of Id2, a negative regulator of basic helix-loop-helix transcription factors. *J Biol Chem* **280:** 4313-20.

Kwon K, Hutter C, Sun Q, Bilic I, Cobaleda C, Malin S *et al* (2008). Instructive role of the transcription factor E2A in early B lymphopoiesis and germinal center B cell development. *Immunity* **28:** 751-62.

Laiosa CV, Stadtfeld M, Graf T (2006). Determinants of lymphoid-myeloid lineage diversification. *Annu Rev Immunol* **24:** 705-38.

Lapenna S, Giordano A (2009). Cell cycle kinases as therapeutic targets for cancer. *Nat Rev Drug Discov* **8:** 547-66.

Lasorella A, Iavarone A, Israel MA (1996). Id2 specifically alters regulation of the cell cycle by tumor suppressor proteins. *Mol Cell Biol* **16:** 2570-8.

Lasorella A, Noseda M, Beyna M, Yokota Y, Iavarone A (2000). Id2 is a retinoblastoma protein target and mediates signalling by Myc oncoproteins. *Nature* **407:** 592-8.

Lasorella A, Rothschild G, Yokota Y, Russell RG, Iavarone A (2005). Id2 mediates tumor initiation, proliferation, and angiogenesis in Rb mutant mice. *Mol Cell Biol* **25:** 3563-74.

Lasorella A, Stegmuller J, Guardavaccaro D, Liu G, Carro MS, Rothschild G *et al* (2006). Degradation of Id2 by the anaphase-promoting complex couples cell cycle exit and axonal growth. *Nature* **442:** 471-4.

Lazorchak AS, Schlissel MS, Zhuang Y (2006). E2A and IRF-4/Pip promote chromatin modification and transcription of the immunoglobulin kappa locus in pre-B cells. *Mol Cell Biol* **26:** 810-21.

LeBien TW, Tedder TF (2008). B lymphocytes: how they develop and function. *Blood* **112:** 1570-80.

Leeanansaksiri W, Wang H, Gooya JM, Renn K, Abshari M, Tsai S *et al* (2005). IL-3 induces inhibitor of DNA-binding protein-1 in hemopoietic progenitor cells and promotes myeloid cell development. *J Immunol* **174:** 7014-21.

Lietz A, Janz M, Sigvardsson M, Jundt F, Dorken B, Mathas S (2007). Loss of bHLH transcription factor E2A activity in primary effusion lymphoma confers resistance to apoptosis. *Br J Haematol* **137:** 342-8.

Lingbeck JM, Trausch-Azar JS, Ciechanover A, Schwartz AL (2005). E12 and E47 modulate cellular localization and proteasome-mediated degradation of MyoD and Id1. *Oncogene* **24:** 6376-84.

Lyden D, Young AZ, Zagzag D, Yan W, Gerald W, O'Reilly R *et al* (1999). Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts. *Nature* **401:** 670-7.

MacFarlane M (2003). TRAIL-induced signalling and apoptosis. *Toxicol Lett* **139:** 89-97.

Magrath I (1990). The pathogenesis of Burkitt's lymphoma. *Adv Cancer Res* **55:** 133-270.

Man C, Rosa J, Yip YL, Cheung AL, Kwong YL, Doxsey SJ *et al* (2008). Id1 overexpression induces tetraploidization and multiple abnormal mitotic phenotypes by modulating aurora A. *Mol Biol Cell* **19:** 2389-401.

Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, Chen K *et al* (2009). Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med* **361:** 1058-66.

Massari ME, Murre C (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol* **20:** 429-40.

Mathas S, Janz M, Hummel F, Hummel M, Wollert-Wulf B, Lusatis S *et al* (2006). Intrinsic inhibition of transcription factor E2A by HLH proteins ABF-1 and Id2 mediates reprogramming of neoplastic B cells in Hodgkin lymphoma. *Nat Immunol* **7:** 207-15.

Matsumura ME, Lobe DR, McNamara CA (2002). Contribution of the helix-loop-helix factor Id2 to regulation of vascular smooth muscle cell proliferation. *J Biol Chem* **277:** 7293-7.

Mitelman F, Johansson B, Mertens F (2007). The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer* **7:** 233-45.

Mullighan CG, Collins-Underwood JR, Phillips LA, Loudin MG, Liu W, Zhang J *et al* (2009). Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. *Nat Genet* **41:** 1243-6.

Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD *et al* (2007). Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446:** 758-64.

Murphy DJ, Swigart LB, Israel MA, Evan GI (2004). Id2 is dispensable for Myc-induced epidermal neoplasia. *Mol Cell Biol* **24:** 2083-90.

Nagasawa M, Schmidlin H, Hazekamp MG, Schotte R, Blom B (2008). Development of human plasmacytoid dendritic cells depends on the combined action of the basic helix-loop-helix factor E2-2 and the Ets factor Spi-B. *Eur J Immunol* **38:** 2389-400.

Nie L, Perry SS, Zhao Y, Huang J, Kincade PW, Farrar MA *et al* (2008). Regulation of lymphocyte development by cell-type-specific interpretation of Notch signals. *Mol Cell Biol* **28:** 2078-90.

Nilsson JA, Nilsson LM, Keller U, Yokota Y, Boyd K, Cleveland JL (2004). Id2 is dispensable for myc-induced lymphomagenesis. *Cancer Res* **64:** 7296-301.

Nogueira MM, Mitjavila-Garcia MT, Le Pesteur F, Filippi MD, Vainchenker W, Dubart Kupperschmitt A *et al* (2000). Regulation of Id gene expression during embryonic stem cell-derived hematopoietic differentiation. *Biochem Biophys Res Commun* **276:** 803-12.

Nowell PC (2007). Discovery of the Philadelphia chromosome: a personal perspective. *J Clin Invest* **117:** 2033-5.

Nutt SL, Heavey B, Rolink AG, Busslinger M (1999). Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401:** 556-62.

Nutt SL, Kee BL (2007). The transcriptional regulation of B cell lineage commitment. *Immunity* **26:** 715-25.

Ohtani N, Zebedee Z, Huot TJ, Stinson JA, Sugimoto M, Ohashi Y *et al* (2001). Opposing effects of Ets and Id proteins on p16INK4a expression during cellular senescence. *Nature* **409:** 1067-70.

Oppezzo P, Dumas G, Lalanne AI, Payelle-Brogard B, Magnac C, Pritsch O *et al* (2005). Different isoforms of BSAP regulate expression of AID in normal and chronic lymphocytic leukemia B cells. *Blood* **105:** 2495-503.

Ouhtit A, Gaur RL, Abd Elmageed ZY, Fernando A, Thouta R, Trappey AK *et al* (2009). Towards understanding the mode of action of the multifaceted cell adhesion receptor CD146. *Biochim Biophys Acta* **1795:** 130-6.

Ouyang XS, Wang X, Ling MT, Wong HL, Tsao SW, Wong YC (2002). Id-1 stimulates serum independent prostate cancer cell proliferation through inactivation of p16(INK4a)/pRB pathway. *Carcinogenesis* **23:** 721-5.

Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S *et al* (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304:** 1497-500.

Palacios R, Henson G, Steinmetz M, McKearn JP (1984). Interleukin-3 supports growth of mouse pre-B-cell clones in vitro. *Nature* **309:** 126-31.

Pan L, Sato S, Frederick JP, Sun XH, Zhuang Y (1999). Impaired immune responses and B-cell proliferation in mice lacking the Id3 gene. *Mol Cell Biol* **19:** 5969-80.

Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E *et al* (2009). Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nat Genet* **41:** 1006-10.

Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, Angenendt P *et al* (2008). An integrated genomic analysis of human glioblastoma multiforme. *Science* **321:** 1807-12.

Pear WS, Miller JP, Xu L, Pui JC, Soffer B, Quackenbush RC *et al* (1998). Efficient and rapid induction of a chronic myelogenous leukemia-like myeloproliferative disease in mice receiving P210 bcr/abl-transduced bone marrow. *Blood* **92:** 3780-92.

Perk J, Iavarone A, Benezra R (2005). Id family of helix-loop-helix proteins in cancer. *Nat Rev Cancer* **5:** 603-14.

Pesin JA, Orr-Weaver TL (2008). Regulation of APC/C activators in mitosis and meiosis. *Annu Rev Cell Dev Biol* **24:** 475-99.

Peverali FA, Ramqvist T, Saffrich R, Pepperkok R, Barone MV, Philipson L (1994). Regulation of G1 progression by E2A and Id helix-loop-helix proteins. *EMBO J* **13:** 4291-301.

Pillai S, Cariappa A (2009). The follicular versus marginal zone B lymphocyte cell fate decision. *Nat Rev Immunol* **9:** 767-77.

Pongubala JM, Northrup DL, Lancki DW, Medina KL, Treiber T, Bertolino E *et al* (2008). Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. *Nat Immunol* **9:** 203-15.

Prabhu S, Ignatova A, Park ST, Sun XH (1997). Regulation of the expression of cyclin-dependent kinase inhibitor p21 by E2A and Id proteins. *Mol Cell Biol* **17:** 5888-96.

Pui CH, Robison LL, Look AT (2008). Acute lymphoblastic leukaemia. *Lancet* **371:** 1030-43.

Quong MW, Martensson A, Langerak AW, Rivera RR, Nemazee D, Murre C (2004). Receptor editing and marginal zone B cell development are regulated by the helix-loop-helix protein, E2A. *J Exp Med* **199:** 1101-12.

Rabbitts TH (2009). Commonality but diversity in cancer gene fusions. *Cell* **137:** 391-5.

Ramirez J, Lukin K, Hagman J (2010). From hematopoietic progenitors to B cells: mechanisms of lineage restriction and commitment. *Curr Opin Immunol* **22:** 177-84.

Raschke S, Balz V, Efferth T, Schulz WA, Florl AR (2005). Homozygous deletions of CDKN2A caused by alternative mechanisms in various human cancer cell lines. *Genes Chromosomes Cancer* **42:** 58-67.

Raynaud SD, Bekri S, Leroux D, Grosgeorge J, Klein B, Bastard C *et al* (1993). Expanded range of 11q13 breakpoints with differing patterns of cyclin D1 expression in B-cell malignancies. *Genes Chromosomes Cancer* **8:** 80-7.

Renne C, Martin-Subero JI, Eickernjager M, Hansmann ML, Kuppers R, Siebert R *et al* (2006). Aberrant expression of ID2, a suppressor of B-cell-specific gene expression, in Hodgkin's lymphoma. *Am J Pathol* **169:** 655-64.

Riechmann V, van Cruchten I, Sablitzky F (1994). The expression pattern of Id4, a novel dominant negative helix-loop-helix protein, is distinct from Id1, Id2 and Id3. *Nucleic Acids Res* **22:** 749-55.

Roberts EC, Deed RW, Inoue T, Norton JD, Sharrocks AD (2001). Id helix-loop-helix proteins antagonize pax transcription factor activity by inhibiting DNA binding. *Mol Cell Biol* **21:** 524-33.

Roessler S, Gyory I, Imhof S, Spivakov M, Williams RR, Busslinger M *et al* (2007). Distinct promoters mediate the regulation of Ebf1 gene expression by interleukin-7 and Pax5. *Mol Cell Biol* **27:** 579-94.

Romanow WJ, Langerak AW, Goebel P, Wolvers-Tettero IL, van Dongen JJ, Feeney AJ *et al* (2000). E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells. *Mol Cell* **5:** 343-53.

Rosenbauer F, Koschmieder S, Steidl U, Tenen DG (2005). Effect of transcription-factor concentrations on leukemic stem cells. *Blood* **106:** 1519-24.

Ross ME, Zhou X, Song G, Shurtleff SA, Girtman K, Williams WK *et al* (2003). Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. *Blood* **102:** 2951-9.

Rothschild G, Zhao X, Iavarone A, Lasorella A (2006). E Proteins and Id2 converge on p57Kip2 to regulate cell cycle in neural cells. *Mol Cell Biol* **26:** 4351-61.

Russell LJ, Akasaka T, Majid A, Sugimoto KJ, Loraine Karran E, Nagel I *et al* (2008). t(6;14)(p22;q32): a new recurrent IGH@ translocation involving ID4 in B-cell precursor acute lymphoblastic leukemia (BCP-ALL). *Blood* **111:** 387-91.

Russell LJ, Capasso M, Vater I, Akasaka T, Bernard OA, Calasanz MJ *et al* (2009). Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. *Blood* **114:** 2688-98.

Saitoh A, Shin HW, Yamada A, Waguri S, Nakayama K (2009). Three homologous ArfGAPs participate in coat protein I-mediated transport. *J Biol Chem* **284:** 13948-57.

Samanta J, Kessler JA (2004). Interactions between ID and OLIG proteins mediate the inhibitory effects of BMP4 on oligodendroglial differentiation. *Development* **131:** 4131-42.

Schebesta M, Pfeffer PL, Busslinger M (2002). Control of pre-BCR signaling by Pax5-dependent activation of the BLNK gene. *Immunity* **17:** 473-85.

Seet CS, Brumbaugh RL, Kee BL (2004). Early B cell factor promotes B lymphopoiesis with reduced interleukin 7 responsiveness in the absence of E2A. *J Exp Med* **199:** 1689-700.

Semerad CL, Mercer EM, Inlay MA, Weissman IL, Murre C (2009). E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors. *Proc Natl Acad Sci U S A* **106:** 1930-5.

Serrano M, Lee H, Chin L, Cordon-Cardo C, Beach D, DePinho RA (1996). Role of the INK4a locus in tumor suppression and cell mortality. *Cell* **85:** 27-37.

Shan L, Yu M, Qiu C, Snyderwine EG (2003). Id4 regulates mammary epithelial cell growth and differentiation and is overexpressed in rat mammary gland carcinomas. *Am J Pathol* **163:** 2495-502.

Sherborne AL, Hosking FJ, Prasad RB, Kumar R, Koehler R, Vijayakrishnan J *et al* Variation in CDKN2A at 9p21.3 influences childhood acute lymphoblastic leukemia risk. *Nat Genet*.

Siegel PM, Shu W, Massague J (2003). Mad upregulation and Id2 repression accompany transforming growth factor (TGF)-beta-mediated epithelial cell growth suppression. *J Biol Chem* **278:** 35444-50.

Sikder H, Huso DL, Zhang H, Wang B, Ryu B, Hwang ST *et al* (2003). Disruption of Id1 reveals major differences in angiogenesis between transplanted and autochthonous tumors. *Cancer Cell* **4:** 291-9.

Singh H, Medina KL, Pongubala JM (2005). Contingent gene regulatory networks and B cell fate specification. *Proc Natl Acad Sci U S A* **102:** 4949-53.

Singh J, Itahana Y, Parrinello S, Murata K, Desprez PY (2001). Molecular cloning and characterization of a zinc finger protein involved in Id-1-stimulated mammary epithelial cell growth. *J Biol Chem* **276:** 11852-8

Sorich MJ, Pottier N, Pei D, Yang W, Kager L, Stocco G *et al* (2008). In vivo response to methotrexate forecasts outcome of acute lymphoblastic leukemia and has a distinct gene expression profile. *PLoS Med* **5:** e83.
Spector LG, Xie Y, Robison LL, Heerema NA, Hilden JM, Lange B *et al* (2005). Maternal diet and infant leukemia: the DNA topoisomerase II inhibitor hypothesis: a report from the children's oncology group. *Cancer Epidemiol Biomarkers Prev* **14:** 651-5.

Spender LC, Inman GJ (2009). TGF-beta induces growth arrest in Burkitt lymphoma cells via transcriptional repression of E2F-1. *J Biol Chem* **284:** 1435-42.

Spiliotis ET, Kinoshita M, Nelson WJ (2005). A mitotic septin scaffold required for Mammalian chromosome congression and segregation. *Science* **307:** 1781-5.

Steen H, Mann M (2004). The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* **5:** 699-711.

Stratton MR, Campbell PJ, Futreal PA (2009). The cancer genome. *Nature* **458:** 719-24.

Sugai M, Gonda H, Kusunoki T, Katakai T, Yokota Y, Shimizu A (2003). Essential role of Id2 in negative regulation of IgE class switching. *Nat Immunol* **4:** 25-30.

Suh HC, Ji M, Gooya J, Lee M, Klarmann KD, Keller JR (2009). Cell non-autonomous function of Id1 in the hematopoietic progenitor cell niche. *Blood*.

Sulong S, Moorman AV, Irving JA, Strefford JC, Konn ZJ, Case MC *et al* (2009). A comprehensive analysis of the CDKN2A gene in childhood acute lymphoblastic leukemia reveals genomic deletion, copy number neutral loss of heterozygosity, and association with specific cytogenetic subgroups. *Blood* **113:** 100-7.

Sun XH (1994). Constitutive expression of the Id1 gene impairs mouse B cell development. *Cell* **79:** 893-900.

Swarbrick A, Akerfeldt MC, Lee CS, Sergio CM, Caldon CE, Hunter LJ *et al* (2005). Regulation of cyclin expression and cell cycle progression in breast epithelial cells by the helix-loop-helix protein Id1. *Oncogene* **24:** 381-9.

Tam WF, Gu TL, Chen J, Lee BH, Bullinger L, Frohling S *et al* (2008). Id1 is a common downstream target of oncogenic tyrosine kinases in leukemic cells. *Blood* **112:** 1981-92.

Thal MA, Carvalho TL, He T, Kim HG, Gao H, Hagman J *et al* (2009). Ebf1-mediated down-regulation of Id2 and Id3 is essential for specification of the B cell lineage. *Proc Natl Acad Sci U S A* **106:** 552-7.

Trabosh VA, Divito KA, B DA, Simbulan-Rosenthal CM, Rosenthal DS (2009). Sequestration of E12/E47 and suppression of p27KIP1 play a role in Id2-induced proliferation and tumorigenesis. *Carcinogenesis* **30:** 1252-9.

Trausch-Azar JS, Lingbeck J, Ciechanover A, Schwartz AL (2004). Ubiquitin-Proteasome-mediated degradation of Id1 is modulated by MyoD. *J Biol Chem* **279:** 32614-9.

Tu X, Baffa R, Luke S, Prisco M, Baserga R (2003). Intracellular redistribution of nuclear and nucleolar proteins during differentiation of 32D murine hemopoietic cells. *Exp Cell Res* **288:** 119-30.

Turner NC, Reis-Filho JS, Russell AM, Springall RJ, Ryder K, Steele D *et al* (2007). BRCA1 dysfunction in sporadic basal-like breast cancer. *Oncogene* **26:** 2126-32.

Umetani N, Mori T, Koyanagi K, Shinozaki M, Kim J, Giuliano AE *et al* (2005). Aberrant hypermethylation of ID4 gene promoter region increases risk of lymph node metastasis in T1 breast cancer. *Oncogene* **24:** 4721-7.

Vockerodt M, Morgan SL, Kuo M, Wei W, Chukwuma MB, Arrand JR *et al* (2008). The Epstein-Barr virus oncoprotein, latent membrane protein-1, reprograms germinal centre B cells towards a Hodgkin's Reed-Sternberg-like phenotype. *J Pathol* **216:** 83-92.

Waite KA, Eng C (2003). From developmental disorder to heritable cancer: it's all in the BMP/TGF-beta family. *Nat Rev Genet* **4:** 763-73.

Wang J, Huang Q, Tang W, Nadal-Ginard B (1996). E2F1 inhibition of transcription activation by myogenic basic helix-loop-helix regulators. *J Cell Biochem* **62:** 405-10.

Wang X, Di K, Zhang X, Han HY, Wong YC, Leung SC *et al* (2008). Id-1 promotes chromosomal instability through modification of APC/C activity during mitosis in response to microtubule disruption. *Oncogene* **27:** 4456-66.

Wilkinson B, Chen JY, Han P, Rufner KM, Goularte OD, Kaye J (2002). TOX: an HMG box protein implicated in the regulation of thymocyte selection. *Nat Immunol* **3:** 272-80.

Willis TG, Dyer MJ (2000). The role of immunoglobulin translocations in the pathogenesis of B-cell malignancies. *Blood* **96:** 808-22.

Yan W, Young AZ, Soares VC, Kelley R, Benezra R, Zhuang Y (1997). High incidence of T-cell tumors in E2A-null mice and E2A/Id1 double-knockout mice. *Mol Cell Biol* **17:** 7317-27.

Yoda A, Yoda Y, Chiaretti S, Bar-Natan M, Mani K, Rodig SJ *et al* (2009). Functional screening identifies CRLF2 in precursor B-cell acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A* **107:** 252-7.

Yu L, Liu C, Vandeusen J, Becknell B, Dai Z, Wu YZ *et al* (2005). Global assessment of promoter methylation in a mouse model of cancer identifies ID4 as a putative tumor-suppressor gene in human leukemia. *Nat Genet* **37:** 265-74.

Yuan SH, Qiu Z, Ghosh A (2009). TOX3 regulates calcium-dependent transcription in neurons. *Proc Natl Acad Sci U S A* **106:** 2909-14.

Yuen HF, Chua CW, Chan YP, Wong YC, Wang X, Chan KW (2006). Id proteins expression in prostate cancer: high-level expression of Id-4 in primary prostate cancer is associated with development of metastases. *Mod Pathol* **19:** 931-41.

Yun K, Mantani A, Garel S, Rubenstein J, Israel MA (2004). Id4 regulates neural progenitor proliferation and differentiation in vivo. *Development* **131:** 5441-8.

Zenz T, Mertens D, Kuppers R, Dohner H, Stilgenbauer S (2009). From pathogenesis to treatment of chronic lymphocytic leukaemia. *Nat Rev Cancer* **10:** 37-50.

Zhang J, Kalkum M, Yamamura S, Chait BT, Roeder RG (2004). E protein silencing by the leukemogenic AML1-ETO fusion protein. *Science* **305:** 1286-9.

Zhuang Y, Cheng P, Weintraub H (1996). B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. *Mol Cell Biol* **16:** 2898-905.