Factors Regulating the Nuclear Localization of β-catenin in Acute Myeloid Leukaemia

by

Jenna Ridsdale

2018

A thesis presented to Cardiff University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

Department of Haematology, Division of Cancer and Genetics, School of Medicine, Cardiff University, CF14 4XN, Cardiff, United Kingdom

Supervisors: Professor Richard Darley and Dr. Alex Tonks

Acknowledgements

I would like to thank my supervisors Professor Richard Darley and Dr. Alex Tonks for their support, patience, expertise and guidance throughout this study.

I would also like to thank the following individuals from Cardiff University:

- Sandra Kreuser, Sara Davies and Sarah Baker for their technical assistance and training.
- Dr. Paul Hole and Dr. Marie Gilmore for their support, advice and training in the laboratory.
- Members of the Cardiff ECMC group for allowing me to use their instrumentation and communal reagents.
- Dr. Andrew Tee for his advice about immunoprecipitation and co-immunoprecipitation techniques.

In addition, I would like to acknowledge the following individuals:

- Prof. Tony Whetton and Dr. Andrew Pierce (The University of Manchester, UK) for their advice and expertise on mass spectrometric techniques and analysis.
- Dr. Alex Greenhough (University of Bristol, UK) for his advice on β-catenin immunoprecipitation and provision of an effective immunoprecipitation protocol.
- Dr. Kate Heesom (University of Bristol, UK) for answering my questions about sample preparation for mass spectrometry and mass spectrometric data analysis.
- Dr Chinmay Munje (University of Glasgow, UK) for his advice, support and training in the laboratory.

I would also like to thank Cancer Research Wales (CRW) for funding this project and allowing me to conduct this study.

Foreword

Supplementary information is provided in the Appendices section at the end of this document.

Raw mass spectrometry data is provided in Appendix 4 on the attached CD disk. Two Microsoft Excel documents are included; one for data generated using a 1% false discovery rate (FDR), and one for data generated using a 5% FDR.

Abstract

Acute myeloid leukaemia (AML) is a highly heterogenous haematological malignancy that accounts for approximately 34% of newly diagnosed leukaemia cases in the United Kingdom. It is a disease that primarily affects the elderly, and the prognosis for AML is generally poor, with a 5-year survival rate of 25% in adults. Identifying factors that contribute to the pathogenesis of AML is important in developing new targeted treatments and improving patient outcome. The Wnt signalling pathway is one of the most commonly dysregulated signalling pathways in AML, and overexpression of its principal effector, β-catenin, is associated with a poor prognosis. β-catenin is a transcription factor that regulates key cellular processes, including proliferation and cell survival, by binding in complex with T-cell factor/lymphoid enhancer-binding factor (TCF/LEF) proteins to Wnt target gene promoters. Nuclear localization of β-catenin is fundamental for its role as a transcription factor. Despite this, little is known about the mechanisms regulating this process in AML; previous work has shown AML blasts and cell lines are highly variable in their ability to translocate β-catenin from the cytosol to the nucleus. This study used a mass spectrometric approach to identify candidate β-catenin nuclear localization factors in AML. This approach was based on comparisons of β-catenin binding partners in the nucleus and cytoplasm of leukaemia cell lines that i) freely translocate β-catenin to the nucleus (K562 and HEL), or ii) resist the nuclear localization of β-catenin (THP-1, U937 and NOMO-1). Cytoplasmic/nuclear fractionation and immunoprecipitation of β-catenin were optimized in K562 cells, before being used to prepare samples of each cell line for mass spectrometry. Following mass spectrometric analysis, eight candidate β-catenin nuclear localization factors were identified, of which two proteins, RUNX1 and LEF-1, were verified by western blotting. Since LEF1 has been previously implicated as a β-catenin nuclear import factor in other contexts and because of its known role in AML, LEF-1 was chosen for further analysis. To determine the relevance of LEF-1 mediated β-catenin nuclear localization in AML pathogenesis, knockdown and overexpression studies were conducted. Knockdown of LEF-1 was successful in leukaemia lines. In cells stimulated with the Wnt agonists (Wnt3a and BIO), LEF-1 knockdown decreased the nuclear translocation of β-catenin and impacted β-catenin mediated transcription. Additionally, knockdown of LEF-1 appeared to reduce proliferation of leukaemia cells, but did not impact their migration or survival. Ectopic nuclear overexpression of LEF1 proved difficult to achieve in myeloid cells due to protein instability; making reciprocal demonstration of the role of LEF1 difficult to demonstrate. Overall this study has identified candidate β-catenin nuclear localization factors and has validated the role of one of these proteins (LEF1). This work provides insight into the potential mechanisms governing β-catenin nuclear localization in AML, and in the longer term, may lead to novel approaches for the treatment of AML.

Abbreviations

Ab	Antibody
ABL	Abelson murine leukemia viral oncogene homolog 1
ADH	Alcohol dehydrogenase
ADP	Adenosine diphosphate
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
AML1	Acute myeloid leukemia 1 protein
APC	Adenomatous polyposis coli
ASXL	Additional sex combs-like
ATP	Adenosine triphosphate
β-cat	β-catenin
BCL	B-cell lymphoma
BCR	Breakpoint cluster region protein
BIO	6-bromoindirubin-3′-oxime
BSA	Bovine serum albumin
CBF	Core binding factor
CBP	CREB binding protein
CD	Cluster of differentiation
cDNA	Complementary DNA
CEB	Cell extraction buffer
C/EBP	CCAAT-enhancer-binding protein
CHAPS	(3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate)
CK	Casein kinase
CLP	Common lymphoid progenitor
CML	Chronic myeloid leukaemia
CMP	Common myeloid progenitor
C/N	Cytoplasmic/nuclear
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRM1	Chromosomal maintenance 1
CtBP	C-terminal binding protein
CXCL	C-X-X motif chemokine
CXCR	C-X-C chemokine receptor
DAAM	Dishevelled associated activator of morphogenesis
dH_2O	Distilled H_2O
DKK	Dickkopf-related protein
DMEM	Dulbecco’s Modified Eagles Medium
DMP	Dimethyl phthalate
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DNA-PK	DNA dependent protein kinase
DNMT	DNA methyltransferase
dNTP	Deoxyribose nucleoside triphosphate
DTT	Dithiothreitol
Dvl	Dishevelled
EGFP	Enhanced green fluorescent protein
EPO	Erythropoietin
ER	Estrogen receptor
ETO	Eight twenty one
FAB	French-American-British
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FDG	Fluorodeoxyglucose

FDR	False discovery rate
FLT3	FMS-like tyrosine kinase 3
FSC	Forward scatter
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GATA	GATA-binding protein
G-CSF	Granulocyte colony stimulating factor
GFI1	Growth factor independent 1 transcriptional repressor
GFP	Green fluorescent protein
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GSK	Glycogen synthase kinase
HAT	Histone acetyltransferase
HBSS	Hanks balanced salt solution
HDAC	Histone deacetylase
HEPES	4-(2hydroxyethyl)-1-piperazine-ethanesulfonic acid
HOX	Homeobox protein
HPLC	High-performance liquid chromatography
HRP	Horseradish peroxidase
HSC	Haematopoietic stem cell
ICAT	β-catenin interacting protein
IDH	Isocitrate dehydrogenase
Ig	Immunoglobulin
IL	Interleukin
IMDM	Iscove's Modified Dulbecco's Medium
IP	Immunoprecipitation
IRES	Internal ribosome entry site
IRS-1	Insulin receptor substrate 1
ITD	Internal tandem duplication
ITRAQ	Isobaric tags for relative and absolute quantitation
JAG1	Jagged 1
JNK	c-Jun N-terminal kinase
KD	Knockdown
KDM	Lysine demethylase
KI	Knock-in
LB	Luria Bertani
LDS	Lithium-dodecyl sulphate
LEF	Lymphoid enhancer-binding factor
LIC	Leukaemia initiating cell
LRP	Low density lipoprotein receptor-related protein
LSC	Leukemic stem cell
LSK	$lin^{-}Sca1^{+}kit^{+}$
LT-HSC	Long term haematopoietic stem cell
LTR	Long terminal repeat
MACS	Magnetic-antibody cell sorting
miRNA	microRNA
MLL	Mixed lineage leukemia
MRC	Medical Research Council
MS	Mass spectrometry/mass spectrometric
MSC	Mesenchymal stem cell
MUC-1	Mucin-1
Myc	Myelocytomatosis oncogene
NES	Nuclear export signal
NF-E2	Nuclear factor erythroid derived-2
NLS	Nuclear localization signal
NPM1	Nucleophosmin
NSD	Nuclear Receptor Binding SET Domain Protein

NUP	Nucleoporin
OPN	Osteopontin
PAGE	Polyacrylamide gel electrophoresis
pBARV	β-catenin activated Venus reporter
PBS	Phosphate-buffered saline
PBST	PBS Tween
PCP	Planar cell polarity
PHF	PHD finger protein
PIC	Protease inhibitor cocktail
Pin1	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1
PKC	Protein kinase C
PML	Promyelocytic leukemia protein
PORCN	Protein-serine O-palmitoleoyltransferase
PP	Protein phosphatase
PTD	Partial tandem duplication
PTK	Protein tyrosine kinase
PTP	Protein tyrosine phosphatase
Pygo	Pygopus
Rac	Ras-related C3 botulinum toxin substrate
RAG1	Recombination activating gene 1
RanBP	Ran binding protein
RAR	Retinoic acid receptor
RNA	Ribonucleic acid
ROR	Receptor tyrosine kinase-like orphan receptor
ROS	Reactive oxygen species
RPMI	Roswell Park Memorial Institute
RUNX	Runt-related transcription factor
Ryk	Related to receptor tyrosine kinase
SCF	Stem cell factor
SDF	Stromal cell-derived factor
SDS	Sodium dodecyl sulphate
SFRP	Secreted frizzled-related protein
shRNA	Short hairpin RNA
SILAC	Stable isotope labelling with amino acids in cell culture
SN	Supernatant
SSC	Side scatter
ST-HSC	Short term haematopoietic stem cell
SUMO	Small ubiquitin-like modifier
TAE	Tris-Acetate-EDTA
TBS	Tris-buffered saline
TBST	TBS Tween
TCF	T-cell factor
TCF7L2	Transcription factor 7-like 2
TCR	T-cell receptor
TEAB	Triethylamonium bicarbonate
TET	Ten-eleven translocation
TGF	Transforming growth factor
TIP	TAT-interactive protein
TKD	Tyrosine kinase domain
TPO	Thrombopoietin
TrCP	Transducing repeat containing-protein
TRRAP	Transformation/transcription domain associated protein
UC	Universal container
UV	Ultra violet
VCAM-1	Vascular cell adhesion molecule-1

WHO	World Health Organization
WIF	Wnt inhibitory factor
Wnt	Wingless-type MMTV integration site family
WT1	Wilms tumor protein

Table of Contents

Summary of Figures

Summary of Tables

1 Introduction

1.1 Haematopoiesis

1.1.1 *Overview of Haematopoiesis*

Haematopoiesis is the process by which blood cells are made. It involves the self-renewal and differentiation of the haematopoietic stem cell (HSC) into a variety of cell types, which are classified into two groups; the myeloid and lymphoid lineages. Classically haematopoietic cells are organised in a hierarchy, in which there is an ordered progression from HSCs into haematopoietic progenitor cells (HPCs) and then finally into terminally differentiated haematopoietic cells which have distinct roles (**Figure 1.1**). Briefly, the cells of the lymphoid lineage mediate the immune response, whereas the cells of the myeloid lineage are involved in a wider range of processes, including blood clotting (platelets) and oxygen transport (erythrocytes). This section will focus on HSCs and cells of the myeloid lineage only, which is the focus of this study.

HSCs are a rare population of cells that retain the ability to differentiate into all other blood cell types (Spangrude *et al*, 1988). They are split into two groups; long-term haematopoietic stem cells (LT-HSCs) and short-term haematopoietic stem cells (ST-HSCs). ST-HSCs have a reduced self-renewal capacity compared to LT-HSCs, however, this view may be oversimplified, with HSCs representing a more fluid and heterogenous population of cells (Liu *et al*, 2012). A delicate balance between self-renewal and differentiation is important in maintaining the pool of HSCs, ensuring that blood cells can be replenished throughout an individual's lifetime.

There are two main models of HSC division; asymmetric and symmetric. According to the asymmetric theory, a single stem cell always give rise to one daughter stem cell and one differentiated daughter cell. In contrast, in the symmetric theory (or "stochastic" model) a single stem cell can either give rise to two stem daughter cells or two differentiated daughter cells. The symmetric division of stem cells may play an important role, for example, when responding

to increasing demand for differentiated blood cells following infection or injury (Morrison & Kimble, 2006).

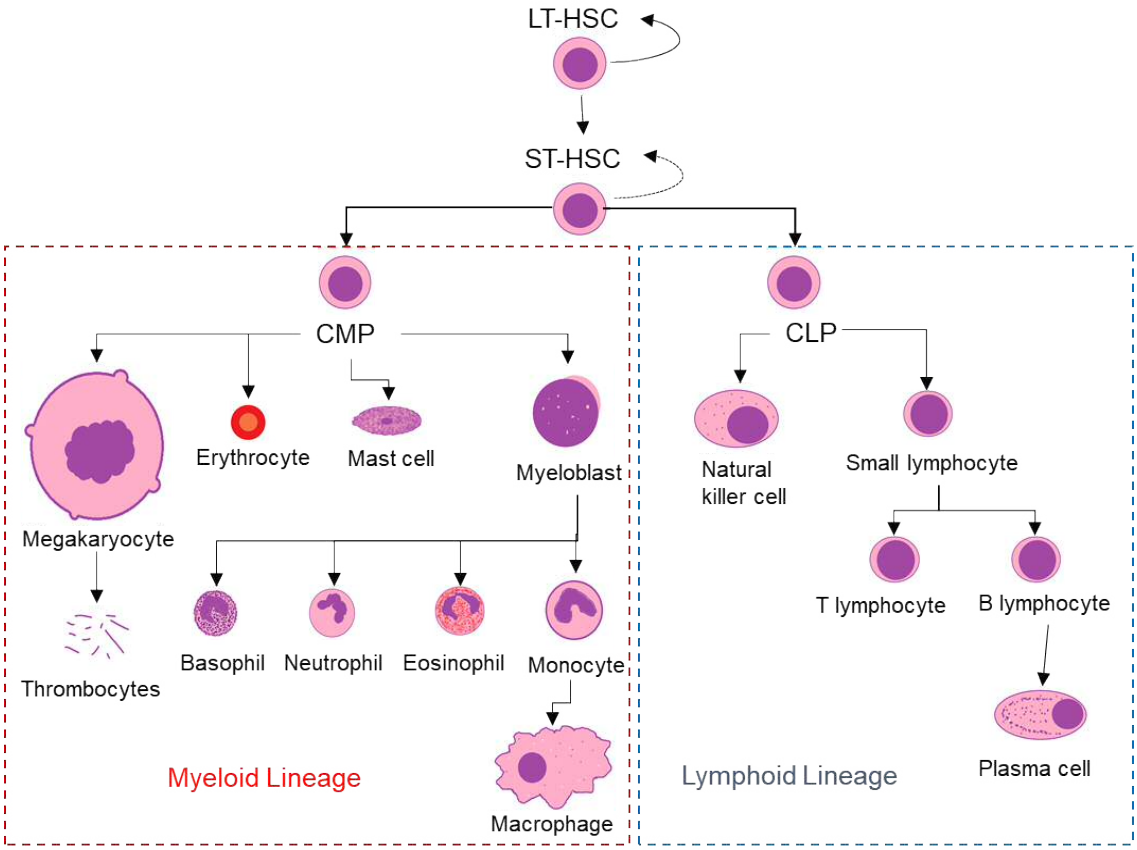


Figure 1.1 A diagram of the haematopoietic hierarchy.

The cells of the haematopoietic system are split into two lineages; the myeloid lineage (left) and the lymphoid lineage (right). The main cells in each lineage are outlined above, but other progenitor cells exist that have not been included in this diagram, for example, the MEP (megakaryocyte erythroid progenitor), which gives rise to erythroid and megakaryocyte cells. Terminally differentiated blood cells have distinct roles. Megakaryocytes produce thrombocytes (platelets), which are involved in blood clotting. Erythrocytes transport oxygen to the cells of the body. Mast cells are involved in wound healing and the immune response. Basophils are involved in the inflammatory response. Neutrophils are phagocytes that are involved in the innate immune response. Eosinophils have varied roles, including in the inflammatory response. Monocytes are involved in phagocytosis, cytokine production and antigen presentation. Macrophages are phagocytes that perform many roles, including in the immune response and inflammation. Abbreviations; CMP (common myeloid progenitor) and CLP (common lymphoid progenitor).

In the classical model of haematopoiesis, as HSCs differentiate, each branch (or bifurcation) in lineage progression is associated with a decreased capacity of the cell type to respond to external factors governing the development to the other lineage, or differentiation path. This is a gradual process, that involves the differentiation of early progenitors into pools of different progenitor cell types, such as the megakaryocyte-erythroid progenitor cell (MEP), that subsequently differentiate into the different blood cell types.

More recent data, from studies looking at HSCs and early haematopoiesis using single cell analysis, is incompatible with this classical model. It suggests that HSCs do not need to go through the defined stages of differentiation. Instead, lineage commitment can occur much earlier than expected and bypass the stages of early progenitor differentiation outlined in the classical model (Velten *et al*, 2017).

1.1.2 ***Regulation of haematopoiesis by extrinsic and intrinsic factors***

Haematopoiesis is regulated by a combination of extrinsic and intrinsic factors. Extrinsic factors include cytokines, growth factors and other environmental cues that can trigger changes in haematopoietic cells by activating cell signalling pathways. Many of these factors are expressed on the cell surface of (or secreted by) cells of the haematopoietic stem cell niche (**Figure 1.2**). Intrinsic factors include signalling pathway components and transcription factors that drive key haematopoietic processes, including self-renewal, differentiation, survival and cell growth.

1.1.2.1 *Cytokines, growth factors and transcription factors*

Cytokines and growth factors control haematopoiesis by regulating downstream signalling pathway components. These pathways are transmitted from membrane receptors at the cell surface and activate haematopoietic transcription factors. Examples of cytokines and growth factors that regulate haematopoiesis are outlined in **Table 1.1**. Knockdown studies suggest that there is a high degree of redundancy between these factors, and that the loss of one factor generally does not lead to a failure in haematopoiesis. One exception is the cytokine erythropoietin (EPO), which is essential to produce erythrocytes in response to

haematopoietic stress. Following haematopoietic injury, EPO is produced in the kidney and its increased presence in the bone marrow acts to i) bias the differentiation of HSCs into erythroid progenitors and ii) rapidly induce erythroid progenitors to proliferate and differentiate into erythroid cells (de Haan *et al*, 1996; De Maria *et al*, 1999).

Cytokines and downstream effector proteins converge on haematopoietic transcription factors, such as members of the C/EBP family and PU.1 (**Table 1.2**). These transcription factors regulate target gene expression to drive changes in progenitor cells that regulate their differentiation down the haematopoietic hierarchy.

Many of the cytokines, growth factors, signalling components and transcription factors that regulate normal haematopoiesis are dysregulated in leukaemia (**1.2**). The role of Wnt signalling pathway components in the regulation of haematopoiesis and leukaemia are discussed in **1.3**.

Table 1.1. Cytokines and growth factors that regulate myeloid development.

Cytokines that regulate HSCs and cells of the myeloid lineage. Many of these factors are expressed by cells in the bone marrow niche, and work in combination to regulate myeloid development by activating downstream signalling pathways and transcription factors in HSCs and myeloid progenitor cells.

Cytokine/growth factor	Example of role
Interleukin-1 (IL-1)	Emergency response following injury. Regulates proliferation of HSCs/early progenitors and promotes commitment to a myeloid fate (Pietras *et al*, 2016).
Interleukin-6 (IL-6)	Works in synergy with other factors such as SCF, to promote survival and self-renewal of HSCs and early progenitors. Promotes commitment of early progenitors to a myeloid fate (Bernad *et al*, 1994; Schürch *et al*, 2014).
Granulocyte-macrophage colony-stimulating factor (GM-CSF)	Stimulates the differentiation of progenitors into granulocytes and macrophages (Ushach & Zlotnik, 2016).
Stem cell factor (SCF)	Works in synergy with other growth factors to regulate the self-renewal and differentiation of HSCs and early progenitors (Broudy, 1997).
Erythropoietin (EPO)	Promotes the survival, proliferation and differentiation of erythroid progenitors (de Haan *et al*, 1996; De Maria *et al*, 1999).
Thrombopoietin (TPO)	Stimulates the differentiation of progenitors into megakaryocytes and regulates HSC quiescence (Kaushansky *et al*, 1995; Qian *et al*, 2007)
Interleukin-5 (IL-5)	Promotes the maturation, differentiation and survival of eosinophils (Sanderson, 1992; Yamaguchi *et al*, 1988)
Granulocyte colony stimulating factor (G-CSF)	Stimulates the differentiation of progenitors into neutrophils (Demetri & Griffin, 1991).
Jagged-1 (JAG1)	Supports the self-renewal and expansion of HSCs (Poulos *et al,* 2013; Weber & Calvi, 2010).
Wnt ligands	Promote the homing of HSCs to the bone marrow niche and self-renewal of HSCs. This is discussed further in **1.3.2.**
Angiopoentin-1	Regulates HSC quiescence (Arai *et al*, 2004).
C-X-C motif chemokine 12 (CXCL12)	Involved in maintaining the HSC pool (Sugiyama *et al*, 2006).
Transforming growth factor beta (TGFβ)	Promotes HSC quiescence (Vaidya & Kale, 2015).
Osteopontin	Promotes HSC quiescence (Nilsson *et al*, 2005).

Table 1.2. Haematopoietic transcription factors

Haematopoietic transcription factors that regulate HSCs and cells of the myeloid lineage.

Transcription factor	Example of role
PU.1	Many; for example, HSC commitment to both lymphoid and myeloid lineages (regulates HSC differentiation into GMP and GLP), biases monocytic commitment of the GMP etc (Iwasaki *et al,* 2005).
CCAAT-enhancer-binding protein alpha (C/EBPα)	Promotes granulocyte differentiation. Activates other haematopoietic transcription factors, e.g. PU.1 (Friedman *et al*, 2003).
Growth factor independent 1 transcriptional repressor (GFI1)	Granulocyte differentiation (de la Luz Sierra *et al*, 2010).
GATA-1/GATA-2	Erythropoiesis (Suzuki *et al*, 2013; Weiss *et al*, 1997).
Runt-related transcription factor 1 (RUNX1)	Homeostasis of HSC/early haematopoietic progenitor cell numbers (Burns *et al*, 2005; Ichikawa *et al,* 2013).
JunB	Represses proliferation and differentiation of HSCs (Passegué *et al*, 2004; Santaguida *et al*, 2009).
Myc	Regulates the balance between self-renewal, differentiation and survival of HSCs (Laurenti *et al*, 2008; Wilson *et al,* 2004).
Homeobox protein B4 (Hox-B4)	HSC proliferation/expansion (Antonchuk *et al*, 2002; Björnsson *et al*, 2003).
Nuclear factor erythroid derived-2 (NF-E2)	Formation of megakaryocytes/platelets (Lecine *et al,* 1998; Shivdasani *et al*, 1995).

1.1.2.2 *Regulation of HSCs in the bone marrow niche*

The bone marrow is the primary site of haematopoiesis in adults, and classically it is thought to contain two niches in which HSCs reside; the osteoblastic niche and vascular niche. These niches are composed of different cell types, including osteoclasts, osteoblasts, stromal cells, extracellular matrix components and a vascular network (**Figure 1.2**). The niche cells play a vital role in haematopoiesis, by regulating key processes in HSCs and progenitor cells, including adhesion, self-renewal, survival and differentiation. This is achieved through direct cell-cell contact with HSCs/progenitors and niche cells, as well as secretion of regulatory factors into the microenvironment.

The importance of the osteoblastic niche in regulating HSCs remains unclear. *In vivo* studies that increase osteoblast numbers lead to an increase in HSCs (Calvi *et al*, 2003; Zhang *et al,* 2003) and in the reverse experiment, involving the conditional removal of osteoblasts from the niche, decreased HSC numbers and abnormal haematopoiesis are observed (Visnjic *et al,* 2004). Although this suggests that the osteoblastic niche is important in regulating HSCs, in the study by Visnjic *et al*, the effect of depleted osteoblasts on HSCs was only observed after a decrease in differentiated blood cells. This suggests that osteoblasts regulate other cell types and that the observed decrease in HSC number may be a result of changes in other cell types. Additionally, more recent *in vivo* studies, using HSC markers to visualize HSC localization in murine bone marrow, found that most HSCs do not reside in the osteoblastic niche and are primarily localized around sinusoids (Acar *et al,* 2015; Morrison & Scadden, 2014). Furthermore, deletion of the key niche cytokines Scf and Cxcl12 from murine osteoblasts did not affect HSC number or function *in vivo* (Ding & Morrison, 2013; Ding *et al*, 2012b). In contrast, conditional knockout of Scf and Cxcl12 in perivascular/vascular niche stromal cells led to HSC depletion. Together, these studies suggest that the perivascular/vascular niche surrounding sinusoids may play a more prominent role in regulating HSCs than the osteoblastic niche.

Given the importance of the regulation of HSCs by niche cells, it is not surprising that factors controlling adhesion of HSCs and niche components are also important in regulating haematopoiesis. These include vascular cell adhesion

molecule-1 (VCAM-1), integrin, N-cadherin(N-cad), β-catenin (**1.3.5**) and osteopontin (OPN) (Yin & Li, 2006).

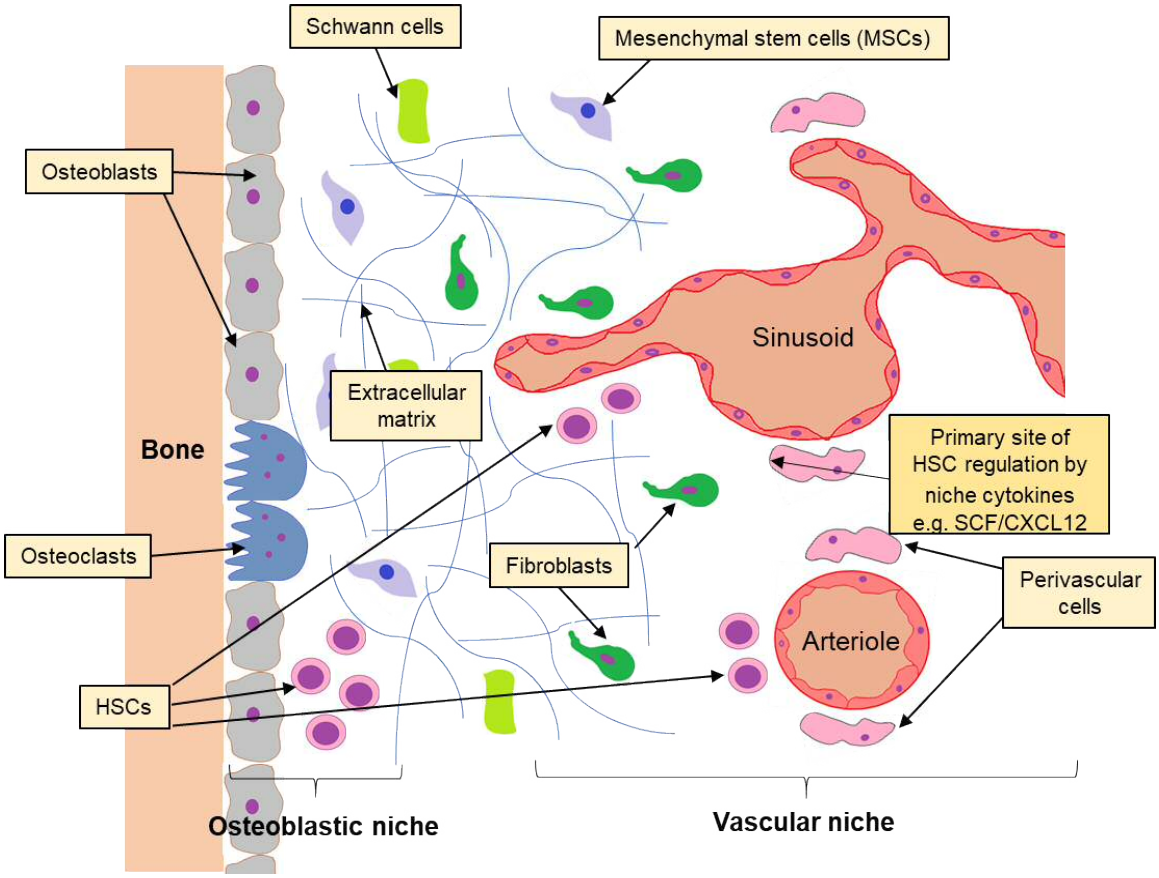


Figure 1.2. The bone marrow niche.

The bone marrow contains two niches where HSCs reside (osteoblastic and vascular niches). Some of the cell types found in these niches are highlighted.

In addition to niche cells, HSCs are also regulated by other factors in the bone marrow, such as oxygen tension. The bone marrow niche has low levels of oxygen (is hypoxic), which promotes HSC quiescence, maintaining them in an undifferentiated state. This implicates the osteoblastic niche as the site for maintenance of quiescent HSCs, and the vascular niche (next to blood vessels that provide higher oxygen levels) as the site of HSC proliferation and differentiation. The low level of oxygen in the bone marrow niche may provide protection from harmful reactive oxygen species (ROS) that can cause mutations and impact the ability of HSCs to promote normal haematopoiesis. (Ludin *et al*, 2014; Mohyeldin *et al*, 2010; Zhou *et al*, 2013).

1.2 Acute Myeloid Leukaemia

1.2.1 ***Overview***

Leukaemia is a term used to describe haematological malignancies that usually develop in the bone marrow. It is generally associated with a high level of abnormal white blood cells and reduced levels of other blood cell types such as erythrocytes and platelets. This leads to a multitude of symptoms including anaemia, fatigue, bruising and immunodeficiency.

Acute myeloid leukaemia (AML) is a highly heterogenous disease and is characterized by a block in normal myeloid development and accumulation of abnormal immature blast cells (**Figure 1.3**). In the UK, 3.4 people per 100,000 are diagnosed with AML per annum and it is most common in the elderly, with a median age of diagnosis of approximately 72 years old[1]. The prognosis for AML is generally poor and varies depending on different prognostic risk factors (1.2.5.1). Current treatment for AML usually involves the use of combinations of chemotherapeutic agents, and to a lesser extent growth factor therapy and radiotherapy (1.2.5.2). Although around 40-80% of patients go into remission following these treatments (Tallman, 2005), the rate of relapse is high, particularly in the elderly, with ~85% of AML patients over the age of 60 relapsing within two years (Burnett, 2012). In addition, the toxicity of these treatments is a major

[1] https://www.hmrn.org (accessed on 25/03/2018)

disadvantage. As such, developing more effective treatments for AML is important for improving patient outcome and quality of life.

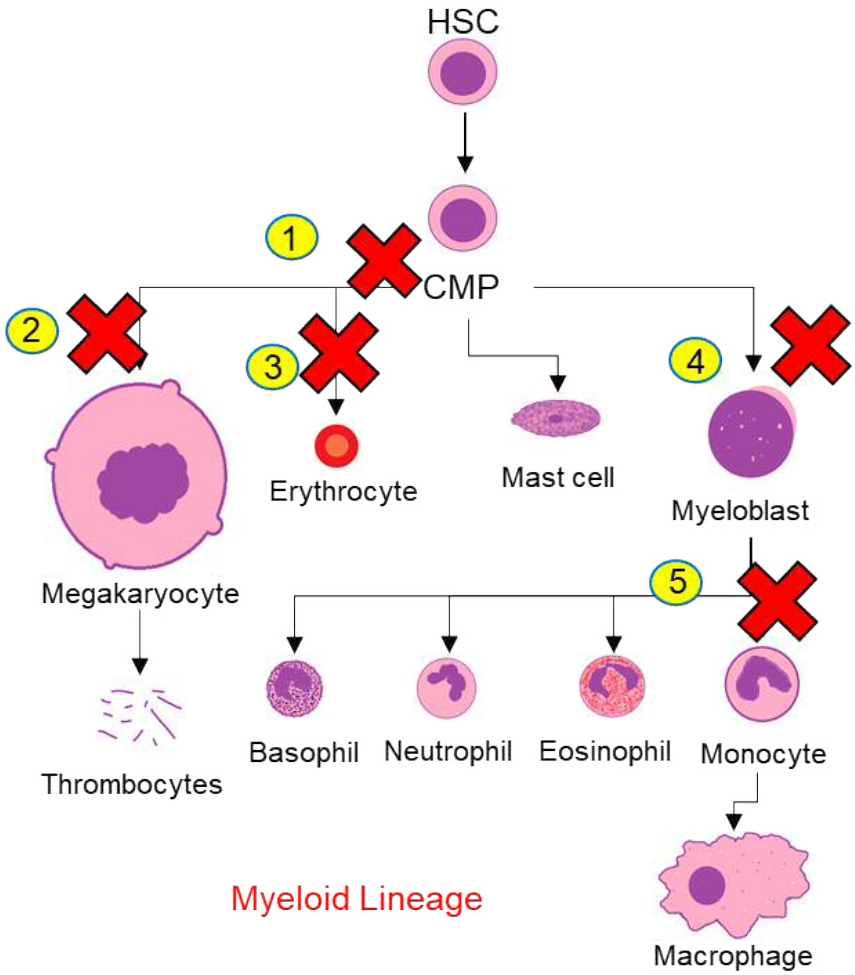


Figure 1.3. Differentiation bock in AML.

The mutations leading to a differentiation block and accumulation of blasts in AML can occur in cells at different stages of the haematopoietic hierarchy (some examples are marked by an X above). This leads to individuals with AML blasts of different differentiation stages, which can be used to classify AML (Table 1.4). Abbreviations; CMP = common myeloid progenitor, HSC= haematopoietic stem cell.

1.2.2 **Molecular *abnormalities in AML***

Mutations involved in AML have historically been split into two different classes based on the 'two hit hypothesis' of AML progression: class I involved mutations that increase proliferation, whereas class II inhibited differentiation and apoptosis (Kelly & Gilliland, 2002). More recently a third mutation class has been added to this classification system, class III, which involves epigenetic regulators (Dombret, 2011; Kao *et al*, 2014). Studies have highlighted a wide array of mutation types in AML that do not fit into the class system described above and the combination of mutations in each patient are thought to collectively regulate increased proliferation and impaired differentiation and apoptosis required for AML pathogenesis.

A recent study using whole genome sequencing, exome sequencing, RNA/miRNA sequencing and DNA methylation analysis of AML patients provides an overview of the common molecular abnormalities in AML (**Table 1.3**). Briefly, the two most commonly mutated genes in AML are the cytoplasmic-nuclear shuttling protein nucleophosmin (NPM1) and the tyrosine kinase FLT3. NPM1 mutations are observed in 27% of AML patients, the FLT3-internal tandem duplication (FLT3-ITD) mutation occurs in 24% of AML patients, and mutations of the FLT3 tyrosine kinase domain (FLT3-TKD) are observed in 4.8% of AML patients (Bacher *et al*, 2008; Lee *et al*, 2007). These mutations have different prognostic values; *NPM1* mutations are associated with a favourable prognosis and FLT3-ITD is associated with a poor prognosis (De Kouchkovsky & Abdul-Hay, 2016). There are also common chromosomal abnormalities in AML, including the t(8;21) translocation and the inversion inv(16) which result in the abnormal proteins RUNX1-ETO and MHY11-CBFβ respectively. MLL-fusions are also commonly observed in AML (**Table 1.3**) (most commonly MLL-AF9) (Meyer *et al*, 2013). Signalling genes and transcription factors also have a high mutation rate in AML, for example, *RUNX1* and *CEBPA* (Ley *et al*, 2013). *RUNX1* mutations are observed in ~ 33% of patients with cytogenetically normal AML, and mutations in *CEBPA* are present in ~10% of AML patients (Fasan *et al*, 2014; Schnittger *et al*, 2011).

Table 1.3. Common molecular abnormalities in AML

Information in this table was taken from (Ley et al, 2013).

Type of mutation	% AML patients	Example(s)
Activated signalling	59%	FLT3, KIT, KRAS, NRAS, PTPs, Ser/Thr kinases and other Tyr kinases
DNA methylation	46%	TET1, TET2, IDH1, IDH2, DNMT3B, DNMT1, DNMT3A.
Chromatin modifiers	30.5%	MLL fusions, MLL PTD, NUP-98-NSD1, ASXL1, EZ112, KDM6A.
NPM1	27%	
Myeloid transcription factors	22%	RUNX1, CEBPA and others.
Transcription factor fusions	18%	PML-RARA, MYH11-CBFB, RUNX1-RUNX1T1, PICALM-MLLT10.
Tumour suppressors	16.5%	TP53, WT1, PHF6
Spliceosome	13.5%	
Cohesin complex	13%	

It is unlikely that novel high frequency abnormalities will be found in AML because extensive genomic classification/whole genome sequencing has been conducted on large AML patient cohorts (Ilyas *et al*, 2015; Ley *et al*, 2013; Papaemmanuil *et al*, 2016) with similar high frequency mutations being characterised across studies, but interestingly AML occurs in patients without DNA mutations. In one study, gene profiling of 1540 AML patients found that 4% of the cohort carried no mutations in any of the currently recognised driver genes associated with the pathogenesis of AML (Papaemmanuil *et al,* 2016). This indicates that other factors can cause disease pathogenesis. These factors include abnormal protein expression levels and epigenetic factors such as methylation of genes and non-coding RNA (e.g. miRNA). Abnormal expression levels of the Wnt pathway component β-catenin will be discussed in **1.3.3**.

1.2.3 *AML diagnosis and classification*

In the past, individuals were diagnosed with AML if their bone marrow presented a blast count of 30% or more, and following diagnosis, French-American-British (FAB) typing was used to classify the AML (**Table 1.4**), based on their morphology and cytochemistry (Bennett *et al*, 1976). Although this system could account for some of the heterogeneity observed in AML, it had limitations for distinguishing AML based on the wide array of genetic and clinical features it can present (Vardiman *et al*, 2002) and was generally poor at predicting outcome.

As a result, FAB typing has been replaced with a newer classification system, outlined by the World Health Organization (WHO) and updated in 2016 (Arber *et al*, 2016). This system was developed to more accurately reflect the heterogeneity observed in AML. It considers the morphological, cytochemistical, immunophenotypic, genotypic and clinical features of the disease. AML is now diagnosed if a patient presents with a bone marrow blast count of 20% and the blasts have a myeloid origin or contain at least one of the genetic abnormalities outlined by the WHO classification system (**Table 1.5**). This classification of AML into subtypes is important in assigning prognosis and best treatment options for patients. In academic research it also enables comparison of different subtypes of AML to establish the relevance of any findings to a clinical setting e.g. if a

protein or drug being studied will benefit all AML patients, or just a specific subtype.

Table 1.4. The French-American-British classification system for AML

Information in this table is from (Bennett et al, 1976).

Subtype	Feature
M0	Undifferentiated
M1	Myeloblastic without maturation
M2	Myeloblastic with maturation
M3	Promyelocytic
M4	Myelomonocytic
M4 EO	Myelomonocytic with bone marrow eosinophilia
M5	Monocytic
M6	Erythroleukemic
M7	Megakaryocytic

Table 1.5. The World Health Organisation classification system for AML

Information in this table is taken from (Arber et al, 2016).

Category	Features
Acute myeloid leukaemia with recurrent genetic abnormalities	AML with t(8;21) (q22;q22), (*AML1/ETO*) AML with inv(16) (p13q22) or t(16;16) (p13;q22), (*CBFβ/MYH11*) APL with t(15;17) (q22;q12), (*PML/RARα*) and variants AML with t(9;11) (p21.3;q23.3);MLLT3-KMT2A AML with t(9;11) (p22;q23); MLLT3-MLL AML with t(6;9) (p23;q34.1);DEK-NUP214 AML with inv(3) (q21.3q26.2) or t(3;3) (q21.3;q26.2); GATA2, MECOM AML (megakaryoblastic) with t(1;22) (p13.3;q13.3);RBM15-MKL1 AML with mutated NPM1 AML with biallelic mutations of CEBPA
AML with myelodysplasia-related changes	
Therapy-related myeloid neoplasms	
AML, not otherwise specified	AML with minimal differentiation AML without maturation AML with maturation Acute myelomonocytic leukaemia Acute monoblastic/monocytic leukaemia Pure erythroid leukaemia Acute megakaryoblastic leukaemia Acute basophilic leukaemia Acute panmyelosis with myelofibrosis

1.2.4 ***Theories of AML propagation***

There are different theories of AML propagation and this section will briefly summarise each.

1.2.4.1 *The clonal evolution theory*

The clonal evolution theory is based on the observation that in AML there is a heterogenous population of leukaemia cells that contain different molecular abnormalities (**1.2.2**). These are subject to selection pressures, favouring subtypes of cells that have an advantage over others (**Figure 1.4A**), for example, increased survival or proliferation. This results in the 'evolution' of AML blasts and a population of cells that is different than the one present at the beginning of disease pathogenesis (Ding *et al*, 2012a).

1.2.4.2 *The hierarchical model*

The hierarchical model (leukaemia stem cell model) theorises that the heterogenous clones in AML are organised in a structure like the normal haematopoietic hierarchy (**Figure 1.1**). According to the model, all AML blasts arise from leukemic stem cells (LSCs) which can establish and maintain disease (**Figure 1.4B**). Early xenograft mouse models suggested that LSCs were derived from normal HSCs or early HPCs. This was based on the observation that only $CD34^+$ $CD38^-$ AML cells engrafted into mice and led to the establishment of leukaemia blasts (Bonnet & Dick, 1997; Lapidot *et al*, 1994). Subsequently it was discovered that antibody labelling of $CD38^+$ cells resulted in immune clearing in these initial experiments and later experiments have shown that $CD38^+$ cells are in fact able to engraft and cause disease, though at a lower frequency (Taussig *et al*, 2008). Studies focusing on the immunophenotype of AML LSCs have identified cell surface markers that are more strongly expressed on LSCs. These include CD96 (Hosen *et al*, 2007), CD123 (Du *et al*, 2011; Jordan *et al*, 2000) and CD47 (Majeti *et al*, 2009b).

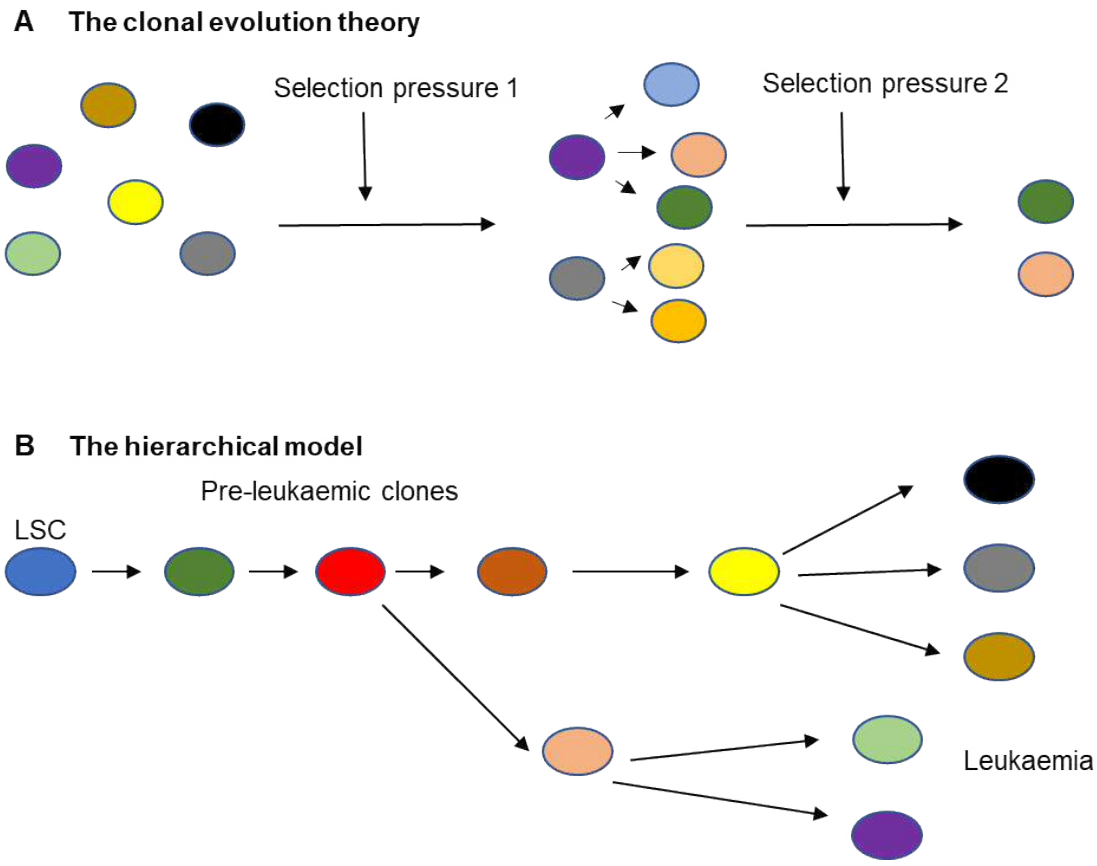


Figure 1.4. Theories of AML leukaemogenesis.

(**A**) The clonal evolution theory. In this model selection pressures act on AML clones, which causes clones with certain characteristics to survive. Over time this changes the population of the AML blasts, so the disease is constantly changing due to environmental factors. This can be used to explain relapse in patients following treatment, as the treatment acts as a selection pressure. (**B**) The hierarchical model. In this model all of the AML blasts arise from an LSC.

1.2.5 *AML prognosis and treatment*

1.2.5.1 *Prognosis*

There are different factors that influence the prognosis of AML patients, including age, initial treatment, white blood cell count, cytogenetic status and secondary disease (Estey & Döhner, 2006). Age is an important predictor of prognosis in patients with AML. Survival of patients enrolled in MRC clinical trials between 1970 and 2009, highlight the significant progress that has been made in the treatment of patients below the age of 60, with a five-year survival rate of >50%. This decreases to 25% survival in patients aged 60-69 and 10% in patients 70+ years old (Burnett, 2012).

Cytogenetics is also a reliable predictor of AML prognosis, and recurring abnormalities are classified as favourable, intermediate or adverse according to their influence on prognosis (**Table 1.6**). For example, FLT3-ITD is associated with an adverse patient outcome prognosis, whereas NPM1 is associated with a favourable patient outcome (De Kouchkovsky & Abdul-Hay, 2016). Karyotypic abnormalities are common in AML, occurring in around 60% of AML patients before treatment (Mrózek *et al*, 2001) and these abnormalities have a profound clinical relevance. For example, patients with deletion of chromosomes 5 or 7 have a poor prognosis and are unlikely to respond to treatment, whereas patients with CBF abnormalities such as inv(16) or t(8;21) have a favourable prognosis and are more likely to respond to treatment (Grimwade *et al*, 1998).

In AML, the largest cytogenetic group is a normal karyotype, which confers an intermediate risk. In these patients, the presence of other molecular abnormalities (**1.2.2**) are more useful for predicting prognosis. For example, NPM1 mutations predict a favourable outcome (Döhner *et al*, 2005) with such accuracy that they have been included in the updated WHO classification system of AML (**Table 1.5)**

The presence of multiple mutations complicates the prognosis of AML. For example, although NPM1 mutations predict a favourable outcome in AML patients, this is not the case when they are found in combination with FLT3-ITD mutations (Döhner *et al*, 2005; Thiede *et al*, 2006).

Table 1.6. Risk groups based on cytogenetics and other molecular abnormalities

The information in this table was taken from (De Kouchkovsky & Abdul-Hay, 2016).

Risk Group	Abnormality
Favourable	t(8;21) without c-KIT mutation t(15;17) inv(16) NPM1 without FLT3-ITD mutation (normal karyotype) Biallelic CEBPA mutation (normal karyotype)
Intermediate	Normal karyotype (unless included in the other risk groups) t(8;21) with c-KIT mutation t(9;11) Other cytogenetic abnormalities not included in the other risk groups
Adverse	TP53 mutation FLT3-ITD (normal karyotype) DNMT3A (normal karyotype) KMT2A-PTD (normal karyotype) inv (3) t(6;9) 11q abnormalities (except t(9;11)) -5 del(5q) -7 Complex karyotype (greater than or equal to three or four chromosomal abnormalities in the absence of any recurrent genetic abnormalities outlined by the WHO classification system (Mrózek, 2008; Stölzel *et al*, 2016)).

1.2.5.2 *Treatment*

Currently, the treatment of AML involves the use of chemotherapy to reduce the bulk blast population and induce long term complete remission (<5% bone marrow blasts and differentiation/maturation of other lineages) (Gale *et al*, 2005). Chemotherapy drugs are generally non-specific and target processes that are important for both AML blasts and normal cells (e.g. DNA replication). As such, there are a multitude of side effects including nausea, infertility, hair loss, immune-suppression and vomiting.

The therapeutic approach used for most patients is the 3+7 approach, involving treatment of patients with daunorubicin for 3 days and cytarabine (AraC) for 7 days (Appelbaum *et al*, 2001). These agents work by targeting DNA replication and cell metabolism, and this approach induces complete remission in 65-75% of adults (18-60 years old) (Tallman, 2005). If patients do not achieve complete remission or shortly relapse after treatment, bone marrow transplants are considered as a next step in treatment. Owing to common complications, high rates of relapse and toxicity of bone marrow transplants, however, they are only suitable for the fittest of patients.

Owing to the poor prognosis and lack of increased survival in elderly AML patients, new therapeutic options are necessary. This is likely to involve agents that target specific molecular abnormalities that are present in AML (**1.2.2**). This paradigm has been used successfully in the treatment of other haematological malignancies, for example chronic myeloid leukaemia (CML), in which targeting the Bcr-Abl fusion protein using tyrosine kinase inhibitors, such as Imatinib, has increased eight-year survival rate of CML patients from ~65% to ~87% (Kantarjian *et al*, 2012). Components of the Wnt signalling pathway could provide a target for such therapeutics in AML patients, as members of the Wnt pathway are frequently dysregulated in AML (**1.3.3**).

1.3 Wnt signalling

1.3.1 *Overview of Wnt Signalling*

The Wnt pathway is an evolutionarily conserved signalling pathway that mediates embryonic development and tissue homeostasis. Wnt proteins are encoded by 19 genes in humans and belong to a family of secreted glycoproteins (Clevers, 2006; Nusse & Clevers, 2017) that bind to specific cell membrane receptors. Three known mechanisms of Wnt signalling have been characterized which fit into two groups; canonical Wnt signalling and non-canonical Wnt signalling. The canonical and non-canonical pathways are regulated by different ligands, receptors and downstream components (**Table 1.7**). Non-canonical Wnt signalling is transduced (1) by small G proteins, for example Rho (in the planar cell polarity (PCP) pathway)) or (2) via regulation of intracellular calcium levels (Kim & Kahn, 2014). Canonical Wnt signalling is transduced through the transcription factor β-catenin (**Figure 1.5**). This section will focus on the canonical Wnt pathway because its central mediator, β-catenin, is the focus of this study.

1.3.1.1 *Inactive canonical Wnt signalling*

In the absence of a Wnt ligand, β-catenin is constitutively degraded by a destruction complex comprising glycogen synthase kinase 3β (GSK3β), casein kinase 1 (CK-1), Axin-1/2, adenomatous polyposis coli (APC) and the E3 ligase, β- transducing repeat containing-protein (β-TrCP). Phosphorylation of β-catenin is important in regulating its degradation by this complex; CK-1 phosphorylates β-catenin on Serine 45 (Ser45) and GSK3β phosphorylates β-catenin on Ser33, Ser 37 and Threonine 41 (Thr41). These phosphorylation events generate recognition sites for β-TrCP which ubiquitinates β-catenin, targeting it for proteasomal degradation (Aberle *et al*, 1997; Orford *et al*, 1997; Polakis *et al*, 1999; Salomon *et al*, 1997). In the nucleus, β-catenin target genes (1.3.1.2) are repressed by a complex comprising TCF, Groucho, histone deacetylases (HDACs) and C-terminal binding protein (CtBP) (Billin *et al*, 2000; Roose *et al,* 1998; Willert & Jones, 2006).

Table 1.7. Components of the canonical and non-canonical (PCP) pathways.

Information in this table is taken from (Komiya & Habas, 2008; Siar et al, 2012). Abbreviations; Wnt= wingless, LRP= low-density lipoprotein receptor-related protein, Ryk= receptor-like tyrosine kinase, ROR= receptor tyrosine kinase-like orphan receptor, PTK= protein tyrosine kinase, Dvl= dishevelled, CK= casein kinase, GSK= glycogen synthase kinase, TCF= T-cell factor, LEF= lymphoid enhancer-binding factor, BCL= B cell/CLL lymphoma, DAAM= dishevelled associated activator of morphogenesis, Rac= Ras-related C3 botulinum toxin substrate, JNK= c-Jun N-terminal kinase.

	Canonical	Non-canonical (PCP) pathway
Ligands	Wnt 1, Wnt2, Wnt3, Wnt8a, Wnt8, Wnt10a and Wnt10b	Wnt4, Wnt5a, Wnt5b, Wnt6, Wnt7a, Wnt7b and Wnt 11
Receptors	Frizzled receptors + LRP5/6 co-receptors	Frizzled receptors + proposed co-receptors (e.g. Ryk, ROR2, and PTK7)
Downstream signalling molecules	Dvl, CK1γ and GSK3β, β-catenin, TCF/LEF, pygopus, BCL-9	Dvl, DAAM1, Rho, Rac, JNK,Jun.

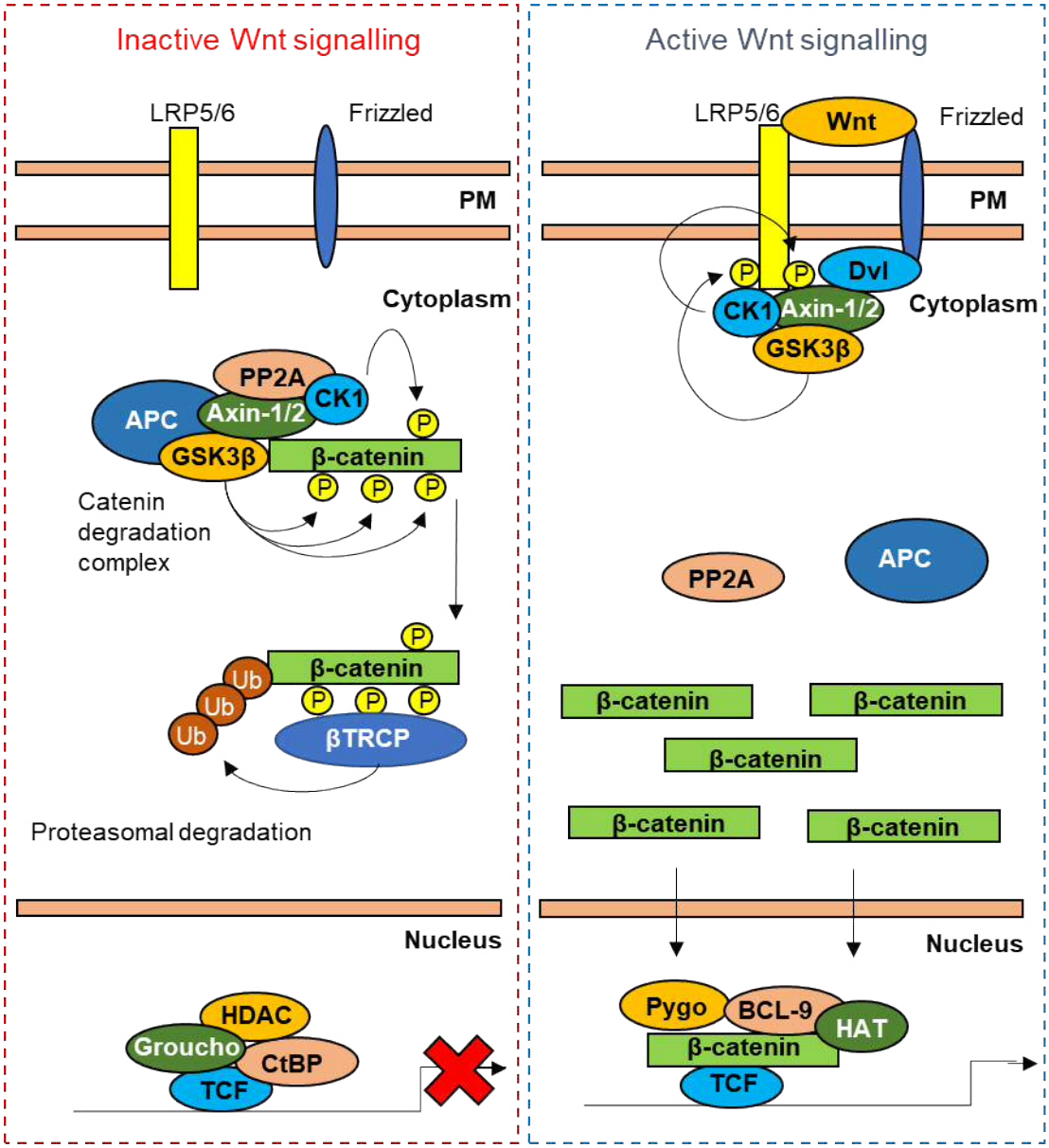


Figure 1.5. Canonical Wnt signalling.

Left; In inactive Wnt signalling, Wnt target gene expression is repressed (X). Right; In active Wnt signalling Wnt target genes are expressed. Abbreviations; PM= plasma membrane, P= phosphorylation, Ub= ubiquitin. Adapted from (Shi *et al*, 2016).

1.3.1.2 *Active canonical Wnt signalling*

Canonical Wnt signalling is activated by binding of Wnt ligands (e.g. Wnt3a) to Frizzled receptors and their co-receptors, low-density lipoprotein receptor-related proteins 5/6 (LRP5/6). Dishevelled (Dvl) is recruited to the receptor complex and assists CK-1 and GSK3β mediated phosphorylation of LRP5/6 (Zeng *et al*, 2008; Zeng *et al*, 2005). The phosphorylated sites on LRP5/6 act as docking sites for Axin-1/2, which detaches from the catenin destruction complex and relocates to the plasma membrane. This leads to the dissociation of the destruction complex, and CK-1 and GSKβ can no longer phosphorylate β-catenin. Without these phosphorylation events, the E3 ubiquitin ligase, β-TrCP, no longer targets β-catenin for proteasomal degradation and β-catenin accumulates in the cytoplasm before translocating to the nucleus (**1.3.6**). In the nucleus, β-catenin displaces Groucho, leading to the dissociation of the transcriptional repressor complex. β-catenin binds the TCF/LEF family of transcription factors in a transcriptional activation complex with pygopus, B-cell CLL/lymphoma-9 protein (BCL-9) and histone acetyl transferases (e.g. CBP/p300) (Evans *et al*, 2010; Kramps *et al*, 2002; Sun *et al*, 2000; Townsley *et al*, 2004). Well characterized β-catenin/TCF target genes include *cyclinD1*, *c-myc*, *TCF-1*, *survivin* and *CD44* (He *et al*, 1998; Roose *et al*, 1999; Shtutman *et al*, 1999; Wielenga *et al*, 1999; Zhang *et al*, 2001). These genes regulate processes such as proliferation and cell survival.

1.3.1.3 *Canonical Wnt pathway antagonists*

The canonical Wnt pathway is negatively regulated at different levels of the pathway. For example, at the ligand-receptor level, proteins can act as Wnt decoys by binding to the Wnt receptor complex without transducing the signal to intracellular Wnt pathway components. Characterized Wnt decoy proteins include Wnt inhibitory factor (WIF), Dickkopf-related protein 1 (DKK1) and soluble frizzled related proteins (sFRPs). In the nucleus, β-catenin can be sequestered by the β-catenin-interacting protein (ICAT), reducing the levels of TCF/LEF bound β-catenin available for activation of Wnt target genes (Daniels & Weis, 2002). Alternative splicing of TCF proteins (including TCF-1 and LEF-1) can also negatively regulate β-catenin mediated transcription. Whilst the longer TCF isoforms can bind to β-catenin, the shorter isoforms lack the β-catenin binding

site and instead bind to target gene promoters and act as transcriptional repressors (Molenaar *et al*, 1996; Roose *et al,* 1999)

1.3.2 *The role of Wnt signalling in haematopoiesis*

There is growing evidence to suggest that canonical and non-canonical Wnt signalling is active in normal haematopoiesis, however, this remains a contentious issue due to conflicting experimental results. Early studies observed the haematopoietic expression of various Wnt pathway components in murine models, including Wnt5A, Wnt10B and Frizzled receptors. *In vitro,* culturing HSCs in conditioned media containing Wnt1, Wnt3a and Wnt10B led to HSC expansion, suggesting that the Wnt pathway promotes the survival and proliferation of HSCs (Austin *et al,* 1997). Other studies characterizing the expression of Wnt genes in human haematopoiesis found that Wnt2B, Wnt5A and Wnt10B were expressed in bone marrow stromal cells, and to varying degrees in haematopoietic cells of the lymphoid and myeloid lineages. In the $Lin^{-}CD34^{+}$ cell population, only Wnt5a was expressed and this cell population expanded in response to Wnt secreting stromal cells (Van Den Berg *et al*, 1998).

Since these early observations, other studies have focused on the role that Wnt signalling plays in the regulation of HSCs and haematopoiesis. *In vivo* studies, in which the Wnt antagonist DKK-1 was overexpressed in the bone marrow niche, led to increased HSC proliferation and a reduction in post-transplant regeneration capacity, suggesting that Wnt signalling regulates the quiescence and reconstitution of HSCs (Fleming *et al*, 2008). In contrast, depleting Wnt3a led to a decrease in HSC numbers, reduced their long-term repopulation capacity in secondary recipient lethally irradiated mice, and reduced the number of myeloid progenitors (Luis *et al,* 2009). This suggested that Wnt signalling regulates the proliferation and long-term survival of HSCs. This was supported by another *in vivo* study, in which inhibition of the catenin destruction complex component GSK3β led to increased β-catenin levels and an increase in self-renewal and reconstitution of HSCs (Trowbridge *et al*, 2006). Additionally, repression of the β-catenin co-activator LEF-1 in $CD34^{+}$ progenitor cells inhibited proliferation and led their apoptosis (Skokowa *et al,* 2006). Wnt signalling also has a proposed role in the homing of HSCs to the stem cell niche. Inhibition of Wnt signalling using DKK-

1 interfered with homing of HSCs to the osteoblastic niche *in vivo* (Lane *et al*, 2011).

In vitro studies overexpressing ICAT (Wnt inhibitor) in murine haematopoietic progenitor cells, interfered with T-cell development (Pongracz *et al*, 2006) suggesting that Wnt signalling is important for normal T-cell development. Conversely, *in vitro* studies overexpressing Wnt3a in human bone marrow progenitor cells had an anti-proliferative effect on B cells, suggesting that Wnt signalling negatively regulates the expansion of this cell type (Døsen *et al*, 2006).

The role of the primary canonical Wnt pathway effector β-catenin in haematopoiesis remains a contentious issue. Some *in vitro* and *in vivo* studies using β-catenin overexpression and knockout mouse models suggest that β-catenin plays an essential role in HSC self-renewal and survival (Nemeth *et al*, 2009; Reya *et al*, 2003; Zhao *et al*, 2007), whereas other studies suggest that activation of β-catenin leads to exhaustion of the HSC pool and a block in multilineage differentiation (Kirstetter *et al,* 2006; Scheller *et al,* 2006). Studies deleting β-catenin and/or γ-catenin (a close homologue of β-catenin) in HSCs did not inhibit normal haematopoiesis upon transplantation of these cells into irradiated mice, however, in these cells TCF/LEF1 expression was still active, suggesting that there are other redundant factors that can regulate their expression in the absence of β-catenin and γ-catenin (Cobas *et al*, 2004; Jeannet *et al,* 2008). In another study, overexpression of β-catenin in bone marrow stromal cells led to proliferation of HSCs, but when β-catenin was overexpressed in HSCs this led to exhaustion of the HSC pool (Kim *et al*, 2009). Stromal cells overexpressing β-catenin also exhibited an enrichment of Notch ligands and downstream Notch signalling. Notch signalling can regulate the self-renewal of HSCs and repress their differentiation (Weber & Calvi, 2010). This suggests that crosstalk between the Wnt and Notch pathways may be important in regulating HSC self-renewal and maintaining them in an undifferentiated state. Studies overexpressing β-catenin in myeloid and lymphoid progenitor cells led to their increased self-renewal and a differentiation block. This suggests that β-catenin's role in haematopoiesis could primarily be to regulate cell ''stemness'' (Baba *et al*, 2005). β-catenin may also play a role in regulating the adhesion of HSCs to stem cell niche cells (1.1.2.2). In one study, *in vitro* imaging of HSCs and MSCs

identified colocalization of N-Cadherin and β-catenin at the point of contact between HSCs and MSCs, suggesting that these components could be involved in regulating their adherence (Wein *et al*, 2010).

The contrasting observations regarding the role Wnt/β-catenin in haematopoiesis in these studies could be due to the different experimental approaches used. An *in vivo* study examining the effect of different APC mutations (leading to a gradient of Wnt signalling), suggested that Wnt signalling regulates haematopoiesis in a dosage dependent manner, which could explain the contrasting observations in previous studies (Luis *et al*, 2011). Whether β-catenin is essential for haematopoiesis or not, tight regulation of Wnt signalling seems to be important in haematopoiesis, and Wnt pathway components including β-catenin, are dysregulated in leukaemia (**1.3.3**).

1.3.3 *Dysregulation of Wnt signalling in AML*

Gene expression analysis of HSCs and LSCs and microarray analysis of AML blasts suggest that the Wnt pathway is one of the most commonly dysregulated pathways in AML (Daud, 2014; Majeti *et al*, 2009a).

In AML, canonical and non-canonical Wnt pathway components are dysregulated at different levels of the signalling pathway. Starting at the cell membrane, aberrant expression of Wnt ligands and receptors have been identified in AML. For example, the Wnt ligands Wnt1 and Wnt2B are expressed in AML blasts but are undetectable in $CD34^+$ progenitor cells (Simon *et al,* 2005). Aberrant expression of Wnt receptors is also observed in AML. For example, overexpression of Frizzled-4 protein is seen in AML blasts and when these cells are induced with Wnt3a, this leads to stabilisation of β-catenin (Tickenbrock *et al,* 2008). Components of the catenin destruction complex are also dysregulated in AML. For example, overexpression of GSK3β is associated with resistance of AML to chemotherapy and missplicing of GSK3β can lead to elevated levels of β-catenin, contributing to leukaemogenesis (Abrahamsson *et al,* 2009; De Toni *et al,* 2006)

Different studies have focused on the central canonical Wnt pathway effector β-catenin and its role in AML. Studies using Western blotting and

immunohistochemistry have shown that β-catenin is expressed in AML blasts. Aberrant *β-catenin* mRNA expression is observed in approximately 20% of AML patients and β-catenin overexpression is associated with poor prognosis and enhanced proliferation of colony forming units *in vitro.* (Chen *et al*, 2009; Simon *et al*, 2005; Xu *et al*, 2008; Ysebaert *et al*, 2006). In the study by *Xu et. al* the level of β-catenin nuclear localization was investigated. The presence of unphosphorylated (active β-catenin) was detected by immunohistochemistry in the nucleus of almost 50% of the AML samples. In the study by *Simon et. al*, β-catenin mediated transcription was evaluated in AML blasts using a TOPFLASH reporter (which measures intracellular TCF/LEF activity). They found that TOPFLASH reporter expression was higher in AML blasts than in normal progenitors. However, they did not correlate this to increased levels of β-catenin. Knockdown of β-catenin in AML cell lines and AML blasts decreases their proliferation *in vitro* (Siapati *et al*, 2011) and impacts the engraftment of AML cell lines to the bone marrow niche *in vivo (*following xenotransplantation in mice*)* (Gandillet *et al*, 2011). Other *in vivo* studies suggest that β-catenin is required for self-renewal of mouse leukaemia initiating cells (LICs) co-expressing the oncogenes (homeobox A9 (HoxA9) or Meis homeobox 1a (Meis1a) (Wang *et al*, 2010). *In vivo* studies looking at preleukaemic stem cell (pre-LSC) and LSC enriched cell populations in AML with MLL fusion proteins have suggested that β-catenin is critical for development of MLL LSCs. Knockdown of β-catenin in MLL LSCs led to delayed disease latency and reduced *in vitro* cloning capacity. Knockout of β-catenin in pre-LSCs led to a failure of these cells to induce leukaemia. This suggests that β-catenin plays a vital role in the establishment of LSCs in MLL AML (Yeung *et al*, 2010). Recently, *in vivo* studies of murine $lin^{-}Sca1^{+}kit^{+}$ (LSK) derived MLL cancer stem cells have demonstrated that HoxA9 suppression sensitises the cells to β-catenin inhibition which leads to the abolishment of the cancer stem cell transcriptional signature and interferes with their transformation capacity (Siriboonpiputtana *et al*, 2017). This suggests that this could be a potential axis for treatment of these aggressive MLL LSCs.

Focusing on another Wnt transcription factor, γ-catenin, studies have observed *γ-catenin* mRNA and protein overexpression in the presence of AML fusion proteins e.g. RUNX1/ETO. This led to downstream activation of Wnt target gene promoters including *c-myc* and *cyclin D1* (Müller-Tidow *et al*, 2004; Zheng *et al*,

2004). Overexpression and aberrant nuclear localization of γ-catenin occurs frequently in AML, and γ-catenin overexpression is associated with the stabilization and nuclear localization of β-catenin (Morgan *et al*, 2013).

In the nucleus, β-catenin binding partners that regulate Wnt target genes have also been implicated in AML. For example, aberrant LEF-1 expression is observed in AML and transplantation of bone marrow cells overexpressing LEF-1 into mice leads to the development of AML and B lymphoblastic leukaemia (Petropoulos *et al*, 2008). LEF-1 was also identified as one of the Wnt pathway components regulating AML by AML associated fusion proteins (Müller-Tidow *et al*, 2004). Microarray analysis of AML patient blasts and normal progenitor cells identified another TCF family member, *TCF7L2* (*TCF-4*) as aberrantly expressed in 78% of patients. In the same study, knockdown of TCF7L2 in leukaemia lines led to a decrease in β-catenin/TCF mediated transcription and an inability of these cells to respond to the Wnt agonists Wnt3a and 6-bromoindirubin-3′-oxime (BIO) (Daud, 2014).

Wnt pathway components are regulated by common AML molecular abnormalities, including FLT3-ITD and fusion proteins (**1.2.2**). For example, the t(8;21) fusion product can bind the promoter of the Wnt antagonist SFRP and repress its function (Cheng *et al*, 2011). In a study examining the promoter methylation status of Wnt antagonist promoters in 269 AML patients, including WIF-1, SFRPs and DKK-1, 62% of patients had methylation of at least one of these factors. WIF-1 hypermethylation was associated with *CEBPA* mutations and the t(15;17) translocation product, and *SFRP-1/SFRP-2* hypermethylation was associated with the t(8;21) translocation product (Hou *et al*, 2011). FLT3-ITD has also been associated with the nuclear localization of the Wnt pathway component β-catenin into the nucleus of AML cells (Kajiguchi *et al*, 2007).

Despite this common dysregulation of Wnt signalling in AML, there is no published evidence of recurrent mutations of canonical Wnt pathway genes in AML.

1.3.4 *Targeting canonical Wnt signalling in the treatment of AML*

Owing to the high relapse rate and poor prognosis in AML (particularly in the elderly) (**1.2.5**), targeted therapies are becoming a major focus for new treatments. Since Wnt signalling is one of the most commonly dysregulated pathways in AML (Daud, 2014; Majeti *et al*, 2009a), it presents a potential target for future therapies. Current research is focusing on the use of novel drugs to target different Wnt pathway components (**Table 1.8**).

Table 1.8. Targeting Wnt signalling in cancer.

Small molecules, antibodies and natural products that can target the Wnt signalling pathway and may be useful for treatment of cancer. Adapted from (Shang et al, 2017).

Agent	Details
LGK974 I	A small molecule that inhibits lipid modification of Wnt by targeting protein-serine O-palmitoleoyltransferase porcupine (PORCN).
OMP-54F28	An antibody that inhibits Wnt ligands.
OMP-18R5	An antibody that inhibits multiple Frizzled receptors.
OTSA101	An antibody that inhibits Frizzled 10.
Celecoxib	A small molecule that promotes GSK3β activation and subsequent targeting of β-catenin for degradation.
DIF1/3	A natural product that activates GSK3β leading to subsequent targeting of β-catenin for degradation.
Genistein	A natural product that activates GSK3β leading to subsequent targeting of β-catenin for degradation.
G007-LK	A small molecule that stabilizes Axin-2 and promotes subsequent targeting of β-catenin for degradation.
XAV939	A small molecule that stabilizes Axin-2 and promotes subsequent targeting of β-catenin for degradation.
JW55	A small molecule that stabilizes Axin-2 and promotes subsequent targeting of β-catenin for degradation.

Agent	Details
WGA	A small molecule that supresses the nuclear pore complex and inhibits β-catenin nuclear translocation.
PRI 724	A small molecule that interrupts β-catenin/CBP interactions. A phase I clinical trial has been completed in AML, but the results are yet to be published[2].
PKF115-584	A small molecule that interrupts β-catenin/TCF interactions. *In vitro* studies treating AML cell lines and AML blasts suggest this molecule could potentially be used in the treatment of AML (Minke *et al*, 2009).
CGP9049090	A small molecule that interrupts β-catenin/TCF interactions. *In vitro* studies treating AML cell lines and AML blasts suggest this molecule could potentially be used in the treatment of AML (Minke *et al*, 2009).
Vitamin D	A natural product that competes with TCF/LEF-1 for β-catenin interaction.
Retinoid acid	A natural product that competes with TCF/LEF-1 for β-catenin interaction.
SAH-BCL-9	A peptide that interrupts β-catenin BCL9 interaction.
CWP232291	Identified as a β-catenin inhibitor. A phase I clinical trial has been completed in AML, but the results are yet to be published[3].

[2] https://clinicaltrials.gov/ct2/show/NCT01606579 (accessed 04/01/2018)
[3] https://clinicaltrials.gov/ct2/show/study/NCT01398462 (accessed 04/01/2018)

1.3.5 *Overview of β-catenin structure and function*

β-catenin is a member of the Armadillo repeat family of proteins and a product of the CTNNB1 gene located at 3p22[4]. This gene gives rise to 15 different transcripts which are mostly protein coding[5]. Four of these transcripts encode full length β-catenin (781 amino acids). Although this project focuses on the role of β-catenin as a transcription factor and component of the Wnt pathway, it is important to note that β-catenin performs other roles, including its regulation of adhesion as a component of adherens junctions. There is some evidence to suggest that β-catenin is involved in regulating the adhesion of HSCs to niche cells (Wein *et al*, 2010)(1.3.2).

The structure of β-catenin is key to its functionality; it features a stretch of 12 armadillo repeats (R1-R12) in addition to distinct C and N terminal domains (**Figure 1.6**). Each armadillo repeat is approximately 42 amino acids long and contains three α-helices (H1, H2 and H3) (Huber *et al*, 1997). These armadillo repeats are the most evolutionarily conserved area of the protein, which is unsurprising given that most of β-catenin's characterized binding partners interact with them (**Table 1.9**). These interactions are vital in regulating β-catenin stability, nuclear localization and β-catenin mediated transcription.

[4] https://www.ncbi.nlm.nih.gov/gene/1499 (accessed on 04/01/2018)

[5] http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000168036;r=3:41194837-41260096 (accessed on 04/01/2018

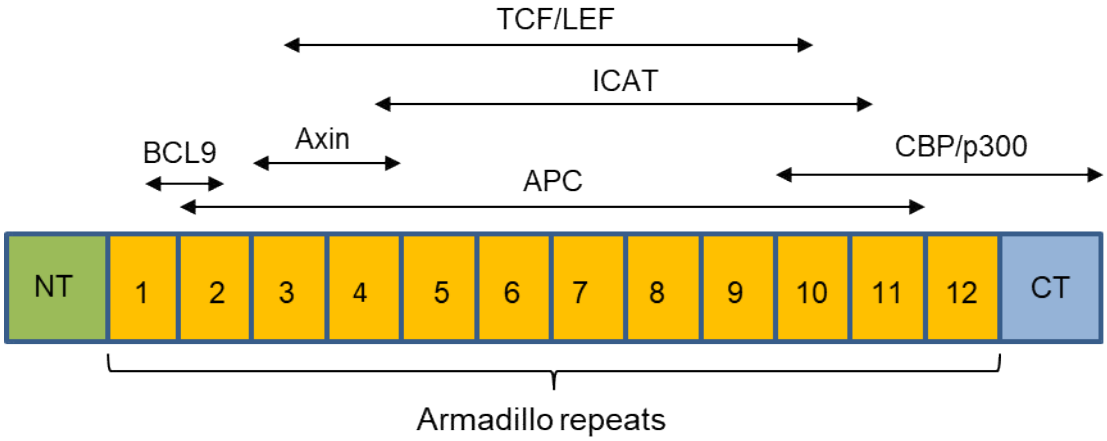


Figure 1.6. Structure of β-catenin and its binding to Wnt pathway components.

The arrows represent areas in the β-catenin structure to which interaction partners bind. More specific information about the most important residues for binding of some of these partners is outlined in **Table 1.9**. The numbers 1-12 represent the 12 armadillo repeats. Abbreviations; NT= N-terminus, CT= C-terminus.

Table 1.9. Binding of interaction partners to β-catenin.

β-catenin binding partner	Binding details
Axin-1	The β-catenin armadillo repeats are vital for its binding to Axin-1. Mutation of histidine^{260} and lysine^{292} completely blocks its binding to Axin-1 (Nakamura *et al*, 1998; von Kries *et al*, 2000).
APC	Armadillo repeats 5 & 6 of β-catenin seem to be the most important in regulating its interaction with APC. There is a strong reduction in binding when arginine^{386}, lysine^{345} and tryptophan^{383} are mutated (von Kries *et al*, 2000).
GSK3β	GSK3β is thought to interact indirectly with β-catenin via its binding to Axin-1 (Wu & Pan, 2010).
PP2A	PP2A is thought to interact indirectly with β-catenin via its binding to other catenin destruction complex components, e.g. Axin-1 (Stamos & Weis, 2013).
CK-1	CK-1 is thought to interact indirectly with β-catenin via its binding to Axin-1 (Stamos & Weis, 2013)
LEF-1	β-catenin armadillo repeats 8 & 9 are important for its binding to LEF-1. Studies mutating residues in this area (specifically residues that flank a hydrophobic area near leucine^{427}) abolished LEF-1 binding. Mutations in other areas did not significantly affect its binding to LEF-1 (von Kries *et al*, 2000).
TCF-4	Residues of β-catenin that are thought to regulate its interaction with TCF-4 are; asparagine^{426}, lysine^{435}, arginine^{469}, histidine^{470}, lysine^{508} and lysine^{312} (Fasolini *et al*, 2003).
BCL9	β-catenin binds to BCL9 via armadillo repeat 1. Leucine^{156}, leucine^{159}, and leucine^{178} form part of the domain that interacts with BCL9 (Sampietro *et al*, 2006).

1.3.6 *Regulation of β-catenin nuclear localization*

Localization of β-catenin to the nucleus is necessary for its role as a transcription factor as part of the canonical Wnt pathway. Studies correlating the cytoplasmic and nuclear levels of β-catenin protein expression have highlighted that stabilization of cytoplasmic β-catenin does not necessarily lead to nuclear localization of β-catenin in AML. Some AML blasts present with a low cytoplasmic level but high nuclear level of β-catenin, and vice versa (Morgan *et al*, 2014). AML cell lines also exhibit variable levels of nuclear β-catenin. Previous analysis of AML cell lines in our laboratory has identified some cell lines that freely translocate β-catenin to the nucleus (termed translocators), and others that resist the nuclear localization of β-catenin (termed non-translocators) (**Figure 1.7**) (Morgan *et al*, 2014). In agreement with the observations made in AML blasts, the level of nuclear β-catenin in these cell lines does not correlate with cytoplasmic levels, suggesting that these cell lines could present a good model system to study the regulation of β-catenin nuclear localization in AML. Identifying factors that regulate β-catenin nuclear localization in AML is important in understanding the mechanism of Wnt signalling in leukaemogenesis and could provide novel therapeutic targets for AML.

Translocators and non-translocators

"Translocation" is a global term used to describe the movement of proteins from one cellular location or compartment to another, and the term "nuclear translocation" is commonly used to describe the movement of β-catenin from the cytoplasm to the nucleus of cells (Chen *et al*, 2005; Griffin *et al*, 2018; Herencia *et al*, 2012; Jian *et al*, 2006). The derivative terms "translocator" and "non-translocator" are not commonly used in the literature to describe cells, but in this study they provide a way of classifying the cell lines according to whether they freely translocate β-catenin to the nucleus or resist nuclear accumulation of β-catenin. The reason for characterizing these cells was so they could be used as a model system to identify potential candidate nuclear localization factors (**1.4**) and so the results could be easily analysed and discussed. This characterization, however, does not consider all the complexity observed in these cell lines, with some "non-translocator" cell lines in fact translocating β-catenin to the nucleus, albeit at a much lower level than "translocator" cell lines (Appendix 1).

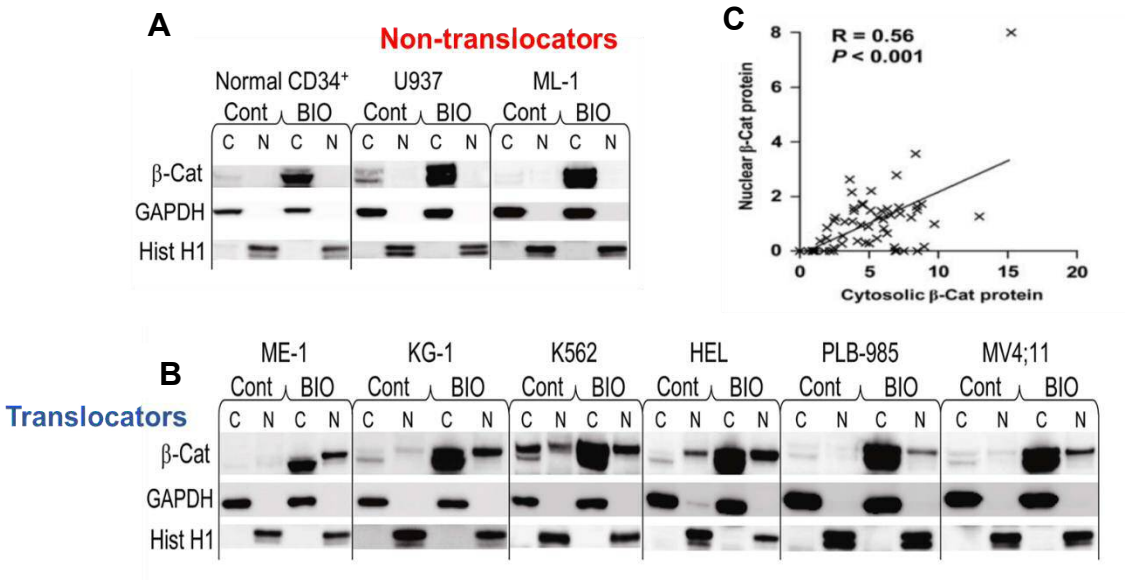


Figure 1.7 Translocator and non-translocator cell lines.

Classification of myeloid cell lines into cell lines that (**A**) resist the nuclear localization of β-catenin (non-translocators) and (**B**) translocate β-catenin to the nucleus (translocators). (**C**) Comparison of cytoplasmic and nuclear levels of β-catenin in AML blasts. This data was generated by Dr. Rhys Morgan (Morgan *et al*, 2014).

β-catenin nuclear import has been studied in various contexts, and different mechanisms governing its localization to the nucleus have been implicated. Some of these mechanisms were reviewed in our paper (Morgan *et al*, 2014). Briefly, β-catenin nuclear localization can be mediated by direct interaction of β-catenin with nuclear membrane components (such as nucleoporins and emerin), CRM1 dependent import, nuclear retention, CRM1 dependent export, cytoplasmic/membrane retention and CRM1 independent export. Different factors have been implicated in the regulation of β-catenin nuclear localization in different contexts (**Table 1.10**). Most studies looking at the regulation of β-catenin localization in cancer have been conducted in solid tumours, but some of these factors could also be relevant to haematological malignancies. Interestingly, some of the factors that mediate β-catenin stability (e.g. Axin and APC) and β-catenin mediated transcription (e.g. LEF-1 and TCF-4) also appear to regulate β-catenin localization to the cytoplasm or nucleus.

Table 1.10. β-catenin nuclear localization factors.

Name of factor(s)	Mode of action	Relevance to cancer	Reference
Nucleoporins	β-catenin binds directly to nucleoporins via its central armadillo repeats (R10-12 and R3-8) and enters the nucleus.	β-catenin is aberrantly expressed and localized in different cancers e.g. leukaemia, colon cancer, breast cancer and ovarian cancer.	(Fagotto *et al*, 1998; Funayama *et al*, 1995; Kau *et al*, 2004; Mikesch *et al*, 2007; Wiechens & Fagotto, 2001; Wiechens *et al*, 2004; Yokoya *et al*, 1999)
Emerin	Binds directly to β-catenin. Limits β-catenin nuclear entry.	Unknown	(Markiewicz *et al*, 2006)
FoxM1	Can directly bind to β-catenin and shuttle it into the nucleus.	FoxM1 overexpression is observed in solid tumours including gliomas and gastric cancer. FoxM1 mediated β-catenin nuclear translocation has been observed in glioma cells.	(Pilarsky *et al*, 2004; Zhang *et al*, 2011)
IRS-1	IRS-1 can shuttle β-catenin into the nucleus.	Overexpression of IRS-1 has been observed in solid tumours and been correlated to β-catenin signalling in ovarian cancer.	(Chen *et al*, 2005)
MUC-1	Binds to β-catenin and promotes nuclear translocation.	MUC-1 has been implicated in cancer and MUC-1 CT expression correlates with β-catenin expression in colorectal and gastric cancers.	(Baldus *et al*, 2004; Kufe, 2009; Li *et al*, 2011; Udhayakumar *et al*, 2007)
BCL-9	Can promote β-catenin nuclear localization and retention.	BCL-9 has been implicated in the pathogenesis of solid tumours and haematological malignancies.	(Townsley *et al*, 2004)
Androgen receptor	Promotes β-catenin nuclear translocation.	Increased androgen receptor signalling is a linked to human	(Mulholland *et al*, 2002)

Name of factor(s)	Mode of action	Relevance to cancer	Reference
		cancer e.g. breast cancer.	
γ-catenin	Promotes β-catenin stabilization and nuclear translocation.	γ-catenin and β-catenin are frequently dysregulated in cancer. γ-catenin has been shown to promote β-catenin nuclear localization in AML.	(Morgan *et al*, 2013)
LEF-1	LEF-1 can bind to β-catenin and promotes nuclear β-catenin localization and retention.	LEF-1 is dysregulated in colon cancer and leukaemia.	(Behrens *et al*, 1996; Fu *et al*, 2014; Huber *et al*, 1996; Jamieson *et al*, 2011; Petropoulos *et al*, 2008; Prieve & Waterman, 1999; Tandon *et al*, 2011)
TCF	XTCF-3 *(Xenopus* TCF-1 homologue) binds directly to β-catenin and promotes its nuclear translocation. TCF-4 overexpression shifts β-catenin localization to the nucleus. In addition, TCF-4 may act by retaining β-catenin in the nucleus.	TCF-4 can regulate nuclear retention of β-catenin in leukaemia patients and may contribute to oncogenesis in colorectal cancer.	(Korinek *et al*, 1997; Krieghoff *et al*, 2006; Molenaar *et al*, 1996; Morin *et al*, 1997)
Axin	Axin can shuttle β-catenin out of the nucleus.	Mutations in Axin have been identified in colorectal cancer.	(Cong & Varmus, 2004; Krieghoff *et al*, 2006; Wiechens *et al*, 2004)
α-Catenin	α-Catenin can export β-catenin from the nucleus and can retain β-catenin at the cell membrane.	A reduction in α-catenin expression has been implicated in solid tumour progression and metastasis. This may also be the case in myeloid leukaemia.	(Giannini *et al*, 2004; Harris & Peifer, 2005; Nelson & Nusse, 2004)
Ran BP-3	Shuttles active β-catenin from the nucleus to the cytoplasm in an APC and	Unknown	(Hendriksen *et al,* 2005)

Name of factor(s)	Mode of action	Relevance to cancer	Reference
	CRM1 independent manner.		
Pin1	Mediates nuclear β-catenin nuclear localization following Wnt3a induction.	Pin1 mediates numerous cancer driving signalling pathways.	(Shin *et al*, 2016; Zhou & Lu, 2016)

1.4 Study aims

The overall aim of this study is to identify factors that regulate β-catenin nuclear localization in AML. To do so, a mass spectrometric (MS) approach (2.6) will be used to identify candidate β-catenin interaction partners that regulate nuclear localization of β-catenin in AML cell lines, by comparing β-catenin interaction partners in the cytoplasm and nucleus of cell lines that i) freely translocate β-catenin to the nucleus (translocator cell lines) and ii) resist the nuclear localization of β-catenin (non-translocator cell lines) (**Figure 1.7**). The reason for using this approach, rather than focusing on previously identified β-catenin nuclear localization factors (**Table 1.10**) was to ensure that important factors regulating the nuclear localization of β-catenin in AML were not overlooked. Analysing all the known β-catenin nuclear localization factors would be time and labour intensive and would not necessarily lead to the identification of β-catenin nuclear localization factors in AML. This is because proteins are expressed and regulated differently depending on their context, so proteins that regulate β-catenin nuclear localization in other contexts will not necessarily regulate this process in AML. An MS approach allows many candidate proteins to be identified, increasing the likelihood that key nuclear localization factors in AML will be discovered. Using an MS approach, the identification of candidate nuclear localization factors in AML will be achieved through the subsequent aims:

1) Optimisation of the nuclear/cytoplasmic fractionation of leukaemia lines and co-immunoprecipitation of β-catenin and its binding partners.

Before MS analysis of β-catenin interaction partners in the cytoplasm and nucleus of AML cell lines can be made, the cell lines need to be fractionated and a co-immunoprecipitation approach needs to be optimized to capture β-catenin interaction complexes.

2) MS analysis of co-immunoprecipitated lysates to identify candidate β-catenin nuclear localization factors.

Identification of interaction partners from β-catenin co-immunoprecipitations will be done using MS. Comparison of nuclear and cytoplasmic interaction partners

will be conducted to identify candidate β-catenin nuclear localization factors. Western blotting will then be used to validate the MS findings.

3) Functional analysis of putative factors regulating β-catenin localization.

Knockdown and overexpression studies will be used to confirm the role of candidate factors in mediating β-catenin nuclear localization and to establish the functional consequences of modulating their expression in AML cell lines in terms of: β-catenin-mediated transcription, cell proliferation, cell migration and cell survival.

2 Materials and Methods

2.1 Cell culture

2.1.1 *General cell culture conditions*

Cell culture was conducted in a Microflow Class II biological safety cabinet (Bioquell, Andover, UK), and prior to cell culture work all surfaces were sterilised with 70% ethanol. Cells were incubated at 37°C in 5% CO_2/air in a Hera Cell humidified incubator (DJB Labcare).

Roswell Park Memorial Institute - 1640 (RPMI-1640), Dulbecco's Modified Eagles Medium (DMEM) and Iscove's Modified Dulbecco's Medium (IMDM) were purchased from Sigma-Aldrich and supplemented with L-Glutamine (Thermo Fisher), FBS (Biowest, Nuaillé, France) and gentamycin (University Hospital of Wales Pharmacy) (**Table 2.1**). All plasticware was purchased pre-sterilised.

Contaminated waste was soaked in Haz-Tabs (Guest Medical, Kent, UK) at 2500 parts per million (ppm) overnight prior to disposal, and if appropriate, discarded into autoclave bins. If retrovirus or lentivirus was used in the cell culture work, the strength of the Haz-Tab solution was doubled.

The cell lines used in this project are summarized in **Table 2.1**. All cultures were maintained at a cell density between $1x10^5$ and $8x10^5$/mL, by passage every 48-72 hours.

Table 2.1. Cell lines used in this project

Cell line	Derivation	AML subgroup/ associations	Culture conditions	Source
K562	Erythroleukemia cell type derived from a 53 year old female CML patient[6].	FAB M1 BCR/ABL1 - t(9;22)(q34;q11) (Andersson *et al*, 2005) c-KIT+ (Hu *et al*, 1994) TP53 mutation[7] HOXA9 mutation[7] ASXL1 mutation[7] KDM3B mutation[7]	RPMI-1640, 10% FBS, 2mM L-Glutamine, 20µg/mL Gentamicin	European Collection of Cell Cultures (ECACC)
HEL	Erythroleukemia cell type derived from a 30 year old male erythroleukemia patient[6].	FAB M6 JAK2 mutation c-KIT+ (Heo *et al*, 2017) TP53 mutation[7]	RPMI-1640, 10% FBS, 2mM L-Glutamine, 20µg/mL Gentamicin	European Collection of Cell Cultures (ECACC)
THP-1	Monocytic cell type derived from a 1 year old male AML patient[6].	FAB M5 MLL translocation: t(9;11)(p22;q23)(Andersson *et al*, 2005) N-RAS mutation[7] TP53 mutation (Sugimoto *et al*, 1992)	RPMI-1640, 10% FBS, 2mM L-Glutamine, 20µg/mL Gentamicin	European Collection of Cell Cultures (ECACC)
U937	Monocytic cell type derived from a 37 year old male histiocytic	FAB M5 MLL translocation : t(10;11)(p12;q14) (Andersson *et al*, 2005)	RPMI-1640, 10% FBS, 2mM L-Glutamine,	American Type Culture Collection (ATCC)

[6] https://www.lgcstandards-atcc.org (accessed on 26/03/2018)
[7] https://cancer.sanger.ac.uk/cosmic (accessed on 13/05/2018)

Cell line	Derivation	AML subgroup/ associations	Culture conditions	Source
	lymphoma patient[6].	TP53 mutation (Sugimoto *et al*, 1992)	20µg/mL Gentamicin	
NOMO-1	Derived from a 31 year old female AML patient[8].	FAB M5a MLL translocation: t(9;11)(p22;q23)[8] K-RAS mutation (Weisberg *et al*, 2014) ASXL1 mutation (Abdel-Wahab *et al*, 2012)	RPMI-1640, 10% FBS, 2mM L-Glutamine, 20µg/mL Gentamicin	DSMZ (Braunschweig, Germany)
HEK293T	Embryonic kidney cells (derived from a fetus)[6].		DMEM, 10% FBS, 2mM L-Glutamine, 20µg/mL Gentamicin	European Collection of Cell Cultures (ECACC)
LS174T	Colon cancer cell line derived from a 58 year old female colorectal adenocarcinoma patient[6].		DMEM, 10% FBS, 2mM L-Glutamine, 20µg/mL Gentamicin	These were a kind gift from Dr. Kenneth Ewan (Cardiff University, UK).
Phoenix	Embryonic kidney cell type[6].		DMEM, 10% FBS, 2mM L-Glutamine, 20µg/mL Gentamicin	These were a kind gift from Prof. Garry Nolan (Stanford University, California, USA).

[8] https://www.dsmz.de (accessed on 26/03/2018)

2.1.2 *Cryopreservation of cells for long term storage and culture*

For cryopreservation, $1x10^6$-$1x10^8$ cells were centrifuged at 280 x g for 5 minutes, resuspended in the recommended growth medium (**Table 2.1**) and added drop-wise to freezing medium (*v/v* 50% IMDM, 30% FBS, 20%, dimethylsulphoxide (DMSO)) of equal volume. Cells were transferred to 1.8mL cryopreservation vials, placed into a Mr. Frosty cryopreservation chamber (Thermo Fisher) with 100% isopropanol, and stored at -80°C overnight. The tubes were then transferred to liquid nitrogen for long-term storage.

When recovering cryopreserved cells, vials were removed from liquid nitrogen storage and thawed rapidly in a 37°C water bath. Freezing medium was diluted slowly (drop-wise) by adding 5mL of growth medium, to all cells to adapt to the changing osmolarity of the solution. The cells were centrifuged at 280 x g for 10 minutes, resuspended in 5mL of the recommended growth medium (**Table 2.1**), seeded in F25 culture flasks and incubated overnight for recovery.

2.1.3 *Cell enumeration*

Cells were counted using a haemocytometer (Hawksley, Brighton, UK) and an Eclipse TS100 light transmission microscope (Nikon, Surrey, UK). 8µL cell aliquots were aseptically removed from culture and pipetted under the haematocytometer cover slip. Cell density per mL was calculated by counting the cells in each of the four counting square sections, calculating an average and multiplying by $1x10^4$.

2.1.4 *Purification of $CD34^+$ haematopoietic progenitor cells*

To isolate mononuclear cells, heparinised human cord blood (obtained with ethical permission from University Hospital of Wales) was diluted 1:1 in Hanks balanced salt solution (HBSS) containing 100µg/mL gentamicin, 25mM 4-(2hydroxyethyl)-1-piperazine-ethanesulfonic acid (HEPES) and 1U/mL heparin. This mixture was carefully layered on top of Ficoll ®-Paque (GE Healthcare, Hemel Hempstead, UK) in a 50mL falcon tube. This was done at a ratio of 8:5 (blood:Ficoll). The tube was then centrifuged for 40 min at 400 x g with the brake disengaged and slow acceleration. Whilst the cells were being centrifuged, 15 mL

of wash medium (RPMI-1640, 10% FBS, 1U/mL heparin and 20µg/mL Gentamicin) was added to a UC. After centrifugation, the mononuclear cells formed a visible layer between the Ficoll and plasma and were aspirated into the UC containing wash medium. The UC was centrifuged for 10 min at 200 x g and the cells were washed again as described above, until there was no longer any visible platelet contamination in the supernatant. The mononuclear cells were counted as described in 2.1.3 and resuspended in RPMI-1640 +10% FBS. The cells were cryopreserved (2.1.2).

$CD34^+$ cells were purified from these mononuclear cells using a magnetic-activated cell separation (miniMACS™) kit (Miltenyi Biotec, UK). Mononuclear cells were recovered from liquid nitrogen storage (2.1.2) and $1x10^8$ cells were resuspended in 150µL magnetic-antibody cell sorting (MACS) buffer (*v/v* 1 x phosphate-buffered saline (PBS), 0.5% BSA, 5mM $MgCl_2$). The cells were centrifuged for 15 min at 4° C with 50µL hapten-conjugated monoclonal CD34 antibody (clone QBEND/10) in the presence of FcR blocking agent. To stop the reaction, 5 mL of MACS buffer was added and the cells were centrifuged for 5 min at 200 x g. The cell pellet was washed and resuspended in MACS buffer before incubation at 4° C for 15 min in the presence of 50µL anti-hapten microbeads. The cells were washed again, resuspended in MACS buffer and applied to a magnetised column. The column was removed from the magnet and 1 mL MACS buffer was used to elute the magnetically labelled $CD34^+$ cells. To maximise purity, this fraction was passed through a second column.

To assess the purity of the eluted $CD34^+$ cells, $1x10^4$ cells were resuspended in staining buffer (1 x PBS, 0.5% BSA, 0.02% Sodium Azide) and 2.5µg/mL R-PE-conjugated anti-human monoclonal CD34 antibody (Clone 8G12, BD) was added. The cells were incubated at 4° C for 30 min and analysed by flow cytometry (2.4).

2.1.5 *Stabilization of β-catenin using BIO and Wnt3a*

To stabilize β-catenin, (2'*Z*,3'*E*)-6-Bromoindirubin-3'-oxime (BIO) (Tocris Bioscience) and Wnt3a (R&D systems) were used. BIO stabilizes β-catenin by competing with ATP binding to the catenin destruction complex component GSK3β. This inhibits the kinase activity of GSK3β and subsequent phosphorylation of β-catenin and β-catenin degradation. To induce cells, 2.5μM BIO or 1μg/mL Wnt3a was added to $2x10^5$/mL cells in 500μl, and the cells were incubated overnight.

2.2 DNA techniques

2.2.1 *Subcloning*

Reaction buffers, enzymes, DNA ladders, dNTP mix and BSA used for subcloning were purchased from New England Biolabs (Hitchin, UK) and molecular grade water was purchased from Sigma-Aldrich. Digestions and ligation conditions used in this study are outlined in **Table 2.2**.

Table 2.2. List of enzymes used in this study and their reaction conditions

Enzyme	Buffer	Incubation
*Kpn*1	NEB1 buffer + BSA	37°C for 1 hour
*Xba*1	NEB2 buffer + BSA	37°C for 1 hour
*Hpa*1	NEB2 buffer + BSA	37°C for 1 hour
*BamH*1	NEB4 buffer + BSA	37°C for 1 hour
Klenow	NEB1 buffer + dNTP mix	25°C for 15 min
T4 ligase	Ligase buffer	16°C for at least 4 hours

2.2.2 *Agarose gel electrophoresis and DNA purification*

Restriction digested DNA was electrophoresed on a 1% SeaKem GTG agarose gel, prepared in Tris-Acetate-EDTA (TAE) buffer. The DNA was diluted in loading buffer (*v/v* 30% Glycerol, 0.25%, 0.25% Xylene Cyanol, Bromophenol Blue) and compared to a 1 kilobase (kb) New England Biolabs DNA ladder. The gel was submerged in TAE buffer and electrophoresed at 40V for 1-2 hours. The gel was then stained in PeqGREEN (VWR, UK) diluted 1:10 in TAE buffer for 15 mins, followed by destaining for 15 mins in water. To visualise and isolate DNA fragments for purification, the gel was placed on saran wrap and a long UV wavelength box was used on high power. A clean scalpel was used to excise the required DNA fragment(s) in low melting point agarose.

A QIAquick™ Gel Extraction kit (Qiagen, Sussex, UK) was used for purification of DNA, according to the manufacturer's instructions.

2.2.3 *Transformation of competent cells and plasmid purification*

To make growth agar and broth for competent *E.coli* culture, Luria Bertani (LB) agar or LB broth (Sigma-Aldrich) was dissolved in 1L of sterile water and autoclaved to ensure sterility (using standard laboratory practice). Before use, agar was melted in a microwave, cooled to approximately 45-50^{o}C and ampicillin (University Hospital of Wales Pharmacy) was added to the concentration of 100µg/mL. 30 mL of agar was poured into each 90 mm Petri dish plate and the plates were cooled before being stored at 4°C overnight.

To amplify vector DNA, competent *E.coli* cells were transformed. Retroviral plasmids contain components required for plasmid replication by bacterial cells, including a component for ampicillin resistance. This allows ampicillin to be used to select for colonies containing the retroviral plasmid, and vector of interest. Vectors for amplification and One Shot® TOP10 or *Stbl*3 Chemically Competent *E.coli* (Invitrogen) were thawed on ice. For each vector, 20ng of DNA was aseptically transferred into a vial of competent cells. The vials were gently mixed by tapping and incubated on ice for 30 min. The cells were then placed in a 42˚C

water bath for 30-45 seconds and placed on ice for 2 min. 250µL of pre-warmed S.O.C medium (Invitrogen) was added to each vial and the vials were shaken at 37ºC for 1 hour at 225 revolutions per minute (rpm) in a shaking incubator. Bacteria from each vial was spread on a separate pre-warmed selective LB-agar plate and incubated overnight at 37° C. Colonies were selected for test digests, grown in LB broth overnight at 37° C in a shaking incubator and plasmid DNA was purified using QIAprep® Mini and Maxi-prep kits (Qiagen, Sussex, UK) and the manufacturer's protocol.

DNA quantification was performed using a NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies, Delaware, USA) and stored at -20° C. DNA purity was assessed using the 260/280 absorbance ratio, with a value of greater than or equal to 1.8 indicating a 'pure' sample. The 260/230 absorbance ratio was measured as a secondary assessment of DNA purity, with a value of ~2 or above indicating a 'pure' sample.

If required, samples were sent for sequencing using Eurofins Genomic's SmartSeq kit sequencing service. 50-100ng/µL of DNA was diluted in 15µL of DNA elution buffer (supplied with Qiagen's QIAprep® Mini and Maxi-prep, or Gel Extraction kits). The relevant primer was added at a concentration of 10pmol/µL (10 µM). The premixed samples were sent for sequencing to Eurofins Genomics, Germany.

2.3 Transduction of cell lines

2.3.1 *Retroviral and lentiviral vectors*

Retroviruses are RNA viruses that belong to the Retroviridae viral family. They are duplicated in host cells and use the enzyme, reverse transcriptase, to produce DNA. The DNA incorporates into the host's genome using an integrase enzyme and then replicates with the host's DNA. Lentiviruses are the only retroviral family member that can integrate in non-dividing cells, making them one of the most efficient gene delivery systems. The constructs used in the project are outlined in **Table 2.3** and **Figure 2.1**. Each of the viruses used in this study are replication defective.

Table 2.3. Vectors/constructs used to transduce cell lines

Vector	Description	Source
pcDNA3-β-catenin-FLAG	Nonviral vector. Contains a neomycin selectable marker.	Addgene
pHIV-EGFP/pHIV-EGFP-β-catenin-FLAG	Lentiviral vector. Contains an EGFP selectable marker (expressed from an IRES in the same message as the insert).	pHIV-EGFP was purchased from Addgene pHIV-EGFP-β-catenin-FLAG was created using subcloning as outlined in 3.3.1.1.
polyPOZ/ β-catenin-polyPOZ	Retroviral vector. Contains a LacZ selectable marker.	These were a kind gift from Prof. Trevor Dale (Cardiff University, UK).
pcDNA3-PKC-FLAG	Nonviral vector. Contains a neomycin selectable marker.	Addgene
pLV-EGFP:T2A:Puro-EF1A>hLEF1	Lentiviral LEF-1 overexpression vector. Contains puromycin and EGFP selectable markers.	VectorBuilder (Cyagen)
PLKO.1-puro shRNA control	Lentiviral knockdown control vector. Contains a puromycin selectable marker.	Sigma
LEF-1 shRNATRCN0000020163 PLKO.1-puro	Lentiviral LEF-1 knockdown vector. Contains a puromycin selectable marker.	Sigma
LEF-1 shRNATRCN0000413476 PLKO.1-puro	Lentiviral LEF-1 knockdown vector. Contains a puromycin selectable marker.	Sigma

Vector	Description	Source
LEF-1 shRNATRCN0000418104 PLKO.1-puro	Lentiviral LEF-1 knockdown vector. Contains a puromycin selectable marker.	Sigma
LEF-1 shRNATRCN0000428178 PLKO.1-puro	Lentiviral LEF-1 knockdown vector. Contains a puromycin selectable marker.	Sigma
LEF-1 shRNATRCN0000428355 PLKO.1-puro	Lentiviral LEF-1 knockdown vector. Contains a puromycin selectable marker.	Sigma
pBARVUbR	Lentiviral TCF/LEF mediated transcription reporter construct. Contains a Venus (variant of EYFP) reporter and a DsRed selectable marker (**Figure 2.1**).	This was a kind gift from Prof. Randall Moon (University of Washington, Seattle, WA, USA)

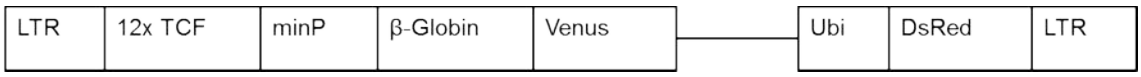


Figure 2.1. Structure and components of the pBARV construct.

The construct contains a concatemer of 12 TCF response elements that have been separated by unique linkers that minimize recombination which would lead to loss of TCF binding sites. The functional promoter driving transcription of Venus (variant of EYFP) is Prol11cga's minP minimal promoter. The Ubi (Ubiquitin) promoter drives DsRed as a selectable marker and the reporters are inserted between the long terminal repeats (LTRs) of a lentivirus transducing plasmid.

2.3.2 *Production of lentivirus and retrovirus by transfection*

In this study, Phoenix amphotropic or HEK293T packaging cell lines (designed to yield high titre retrovirus or lentivirus respectively) were seeded at 7.0×10^6 in an F75 tissue culture flask in 10mL appropriate medium (**Table 2.1**). The next day, the medium was replaced with 15mL of fresh medium and transfection of plasmid DNA was performed. To do this, 20µg of the relevant plasmid DNA was added to 125mM Calcium Chloride ($CaCl_2$) and made up to 450µL in sterile water. This was then added to 450µL HEPES-buffered saline dropwise using gentle 'bubbling'. The mixture was vortexed briefly and incubated at room temperature for 20 min to allow a precipitate of calcium-phosphate-DNA complexes to develop. 5 minutes before the end of the incubation, 25µM chloroquine was added to the cultures to be transformed to promote the uptake of DNA. At the end of the 20min incubation, the 900µL of precipitated DNA was pipetted onto the cells, and the flask was gassed with 5% CO_2 and incubated at 37˚C overnight. The next morning, the medium was replaced with 8mL of fresh medium and the flask was gassed with 5% CO_2 and incubated again at 37˚C overnight. The next day harvesting of the virus was conducted, by centrifuging the medium at 200 x g for 10 minutes. The supernatant containing the retrovirus was then aliquoted into 1mL cryovials, snap frozen in liquid nitrogen and stored at -80˚C.

2.3.3 *Transduction of cell lines*

Cell lines were transduced using a Retronectin based spin infection protocol. Retronectin is a recombinant human fibronectin peptide fragment that binds cell surface proteins and viruses. It significantly enhances retroviral/lentiviral transduction of mammalian cells when coated on the surface of cell culture plates (Tonks *et al*, 2005). Cells were transduced in a sterile untreated 24 multi-well plate using a spin infection procedure. Wells were coated with 250µL of Retronectin (Takara Bio, Europe) and incubated for 2 hours at room temperature, or overnight at 4°C. The Retronectin was then replaced with 200µL of PBS (+1% BSA) and incubated for 30 min at room temperature. A few minutes before the end of the incubation, the relevant viruses were rapidly thawed at 37˚C and placed in the culture cabinet. After blocking with 1% BSA was complete, it was

aspirated from each well and immediately replaced with 1mL of retrovirus. The plate was then centrifuged at 4°C for 120 min at 930 x g.

During the centrifugation, cells were prepared at a density of $1x10^5$/mL in the relevant culture medium (**Table 2.1**). Once the centrifugation was complete, the supernatant in the wells was replaced with 1mL of cells and the plate was incubated at 37˚C overnight. The next day, a second spin infection was carried out. Cells that had been transduced with vectors containing puromycin resistance markers were selected for. To do this, 2 x 10^5 cells in fresh medium were placed in a new 24-well plate incubated with 1 µg/mL puromycin at 37° C overnight in 5% CO_2. Selection was continued in the presence of puromycin until all control uninfected cells were killed.

2.4 Flow cytometry

2.4.1 *Flow cytometry overview*

Flow cytometry is a technique that can be used to measure the light scattering properties and fluorescence emission of single cells in a heterogenous mixture. During flow cytometry, single cell suspension samples are taken into the instrument and surrounded by sheath fluid which acts as a buffer. Hydrodynamic focusing organises the cell suspension into a stream of single cells that travel to an interrogation point. Here a laser light beam illuminates the cell, and some of the light scatters and is measured by the flow cytometer as forward (FSC) or side (SSC) scatter. This can be used to determine the relative cell size and give an indication of its cellular structure, respectively. The laser also excites fluorophores in the cell, which produce a fluorescence emission. These emissions pass through a set of filters and dichromic mirrors that isolate light of specific wavelengths. This light is then converted into a digital signal that is displayed on the flow cytometer's computer software[9]. When using a flow cytometer to measure the presence of fluorophores, a knowledge of its excitation/emission peaks is necessary so that the correct laser is used for excitation and the correct

[9] http://www.thermofisher.com/uk/en/home/life-science/cell-analysis/cell-analysis-learning-center/molecular-probes-school-of-fluorescence/flow-cytometry-basics/flow-cytometry-fundamentals/how-flow-cytometer-works.html (accessed on 04/01/2018)

filters are used to collect the emitted light. If more than one fluorophore is used a knowledge of their excitation and emission points also allows for correct "compensation". **Table 2.4** outlines the filter settings that were used for the fluorophores in this study, and unless otherwise stated, an Accuri C6 benchtop Flow Cytometer™ (Accuri Cytometers, Ann Arbor, MI, USA) was used for analysis. 15,000 events were measured per sample, and backflushing was performed between each acquisition to reduce cell transfer between samples. Flow cytometric data analysis was conducted using FCS Express 6 (De Novo Software, California, USA). Forward and side scatter characteristics were used to exclude debris from the analysis and 7-AAD was added to the samples to allow gating out of dead cells (**Table 2.4**).

Table 2.4. Fluorophores used in flow cytometric analysis

Fluorophore	Excitation/emission wavelength (nm)	Filter set (Accuri C6)
7-AAD	546/647	FL2
To-PRO3	624/661	FL3
Venus	515/528	FL1
Fluorescein	490/525	FL1
EGFP/GFP	489/509	FL1

2.4.2 *FDG sorting*

Fluorodeoxyglucose (FDG) can be used for cell sorting. In the presence of β-galactosidase (the LacZ gene product) it is converted into fluorescein (Plovins *et al*, 1994) which can be measured by flow cytometry (**Table 2.4**). For FDG sorting, cells were counted (2.1.3) and 500K cells were diluted in 250µL of PBS+1% BSA in 5mL BD FACS tubes. Hypotonic shock was conducted by warming the tubes for 2 min in a 37° C water bath. 250µL of prewarmed FDG was added to the cells and mixed by pipetting up and down twice. The tubes were returned to the water bath and incubated for 60s. 2500µL of ice cold stop solution (PBS+1% BSA + 1µg/mL 7-AAD) was added, and the tubes were mixed and immediately placed on ice. The tubes were incubated on ice for 2 hours. The cells were transferred to a UC and centrifuged at 280 x g for 5 min. The cells were resuspended in 500µL of stop solution and passed through a 70µM cell strainer (Miltenyi Biotec) into clean 5mL FACS tubes. These samples were sent for fluorescence activated cell sorting (FACS) at the Flow Cytometry Facility, Cardiff University. A small sample of the cells were analysed on the Accuri flow cytometer prior to FACS to determine the percentage of cells emitting fluorescein. Dead cells were removed by gating out cells with a high uptake of 7-AAD. This was repeated after the sort to determine how successful the sorting process had been.

2.4.3 *Proliferation and survival assay*

To measure proliferation and survival of cells a proliferation/viability assay was conducted. Cells in log phase growth (4-8 x10^5 cells/mL) were pelleted in a universal container (UC) and washed in 10 mL of serum free medium. The medium was discarded, and the pellet was again washed in 10mL of fresh serum free medium. 90µL of cells were distributed into each required well of a 96 well dish using a multichannel pipette. 10µL serum dilutions were added to the cells (to yield 0%, 1%, 3% and 10% *v/v* serum), the lid was put on the plate and it was vortexed to mix. The plate was boxed up and incubated overnight (2.1.1). Three replicate plates were set up for analysis over the next three days. The residual cells (in the UC) were recounted using the flow cytometer to give starting densities for the assay: cell suspension (99µL) was added to a 1mL tube

containing 1µL of 100µg/mL 7-AAD, viable cell density was calculated from a calibrated uptake of 10µL of cell suspension.

The next day 1µL of 100µg/mL 7-AAD was added to a 1mL tube for each sample. The cells were mixed using a multichannel pipette and 99µL of cells were transferred to each minitube. The Accuri flow cytometer was used to count the cells (as above) and the % viability was calculated for each sample. This was repeated for the plate replicates over the next two days.

2.4.4 *Migration assay*

To measure the migration of cells, 24-well chemotaxis chambers (pore size 5.0uM) (Corning Costar) were used. The inserts were placed into the wells and rinsed with PBS to remove any debris. 600µL of chemotaxis medium (IMDM + 1% BSA) was added to wells in a receiver plate. 100µL of chemotaxis medium was added to the inserts which were placed in medium containing-wells. These were incubated (2.1.1) whilst the cells were prepared. Cells were counted (2.1.3) and 100K cells were aspirated into a UC which was topped up with 10mL of prewarmed chemotaxis medium. The cells were centrifuged for 10min at 350 x g and the supernatant was aspirated and discarded. The cells were resuspended in 590µL prewarmed chemotaxis medium, giving a density of 1.7 $x10^5$ cells/mL. The inserts were transferred to empty wells and their contents were aspirated and replaced with 100µL of cells. 5µL SDF-1 (PBS+1% BSA) was added to the wells giving final concentrations of 100ng/mL, 30ng/mL, 10ng/mL and 3ng/mL. The inserts were returned to the wells and incubated (2.1.1) for four hours. After the incubation, the inserts were lifted from the wells and the contents of the inserts were transferred to 1mL tubes. The inserts were rinsed with 100µL of chemotaxis medium which was combined in the 1mL tubes. This process was repeated for each well and 1µg/mL 7-AAD was added to each tube. Cells were counted using the Accuri flow cytometer (2.4.1), which was set to acquire 50µL of each sample. The number of events in this volume was measured for each sample.

2.4.5 *Statistics*

Statistical analyses of the proliferation and migration assay results were conducted using a two sample T test and evaluating p values. Values of $p < 0.01$ and < 0.05 were considered significant differences.

2.5 Western Blotting

2.5.1 *Cytoplasmic/nuclear fractionation for use in western blotting*

For cytoplasmic/nuclear fractionation prior to western blotting, a modified version of the protocol provided with the BioVision fractionation (Biovision, California, USA) kit was used. Cells ($2x10^6$) were pelleted and washed twice with 20mL Tris-buffered saline (TBS), followed by centrifugation for 10 min at 280 x g. Next, cells were resuspended in 200µL cytosol extraction buffer A (CEB-A) containing 1x protease inhibitor cocktail (PIC)(Sigma) and 1mM dithiothreitol (DTT), vortexed for 15 seconds and incubated on ice for 10 min. Eleven µL of cytosol extraction buffer B (CEB-B) was then added to the cells followed by a quick vortex and incubation on ice for 1 min. Cells were vortexed quickly before being centrifuged at 3,000 x g for 8 minutes in a 4˚C microcentrifuge. The supernatant (cytoplasmic extract) was removed to a clean mini Eppendorf tube and stored at -80°C. The pellet was washed with 500µL of pellet wash buffer (1 x PBS + 5mM $MgCl_2$) and centrifuged at 3000 x g for 3 minutes. The supernatant was removed, and the nuclear pellet was snap frozen in liquid nitrogen. The pellets were then freeze thawed three times, and 1µL of Benzonase (Sigma) was added directly to the pellet. The tube was tapped to mix and incubated on ice for 30-60 min. During this incubation, the tube was tapped regularly. Fifty µL of TEAB buffer (0.5M TEAB, 0.05% SDS, PIC) was added and the tube was incubated on ice for 30 min, with vortexing every 10 minutes. The tubes were centrifuged at 10,000 x g for 10 min and the supernatant (nuclear extract) was transferred to a clean tube and stored at -80° C.

2.5.2 *Protein quantification*

The protein concentration of cytoplasmic/nuclear fractions was calculated using the Bradford's protein assay (Bradford, 1976). Protein standards (0, 10, 40, 70 and 100 g/mL BSA) were diluted in fractionation buffer. Ten µL of protein standards were loaded onto wells of a 96-well plate, using 2 duplicates of each. Next, 10µL of each cell fraction was loaded into wells in duplicate. 190µL of Bradford's reagent was added to the samples and incubated for 10 min. Absorbances for each sample were read at 590nm using an ASYS Hitech Expert plus spectrophotometer (Biochrom, Cambridge, UK). Protein concentrations were calculated by interpolation from the standard curve.

2.5.3 *LDS-PAGE gel electrophoresis and electroblotting*

Lithium-dodecyl sulphate polyacrylamide gel electrophoresis (LDS-PAGE) separates proteins in a gel matrix with an applied electric field. Addition of LDS to protein samples results in them having a negative charge. During electrophoresis the proteins travel through the gel towards the cathode, and are separated based on their size, with smaller proteins migrating more quickly through the gel matrix. Following gel electrophoresis, the proteins are transferred to a membrane for western blotting.

Gel electrophoresis and western blotting was conducted using the NuPAGE® PreCast Gel System (Invitrogen), and unless otherwise stated all reagents were purchased from Invitrogen (Thermo Fisher).

Protein lysates were thawed on ice, and equivalent amounts of total protein were prepared for each sample. To do this, samples were diluted in NuPAGE® 1xLDS, 50nM NuPAGE® reducing agent and dH_20, well mixed and incubated at 70˚C for 10 minutes to denature the proteins. The samples were then placed back on ice.

Pre-cast NuPAGE® Novex 4-12% bis-Tris gels were washed with water and the wells were rinsed with running buffer (1 x NuPAGE® MOPS-sodium dodecyl sulphate (MOPS-SDS)) twice. The gels were placed into an XCell SureLock™ Mini Cell gel electrophoresis tank and the inner chamber was filled with 200mL of running buffer (+ 500µL NuPAGE® anti-oxidant). The samples were loaded into the wells, and a MagicMark™ XP Western Protein Standard ladder was

included. The outer chamber was filled with running buffer, and gel electrophoresis was conducted at 200V for 50 minutes.

The gel was removed from the tank and one surface of the gel was dampened with transfer buffer (1 X NuPAGE® Transfer Buffer, 10% MeOH, dH_20, 1mL antioxidant). A piece of pre-soaked Whatman 3M filter paper was layered on the gel and the other side of the gel was moistened with transfer buffer. A pre-soaked nitrocellulose membrane was placed on the gel surface and another piece of pre-soaked filter paper was layered on top. The gel and membrane were placed between pre-soaked blotting pads (**Figure 2.2**).

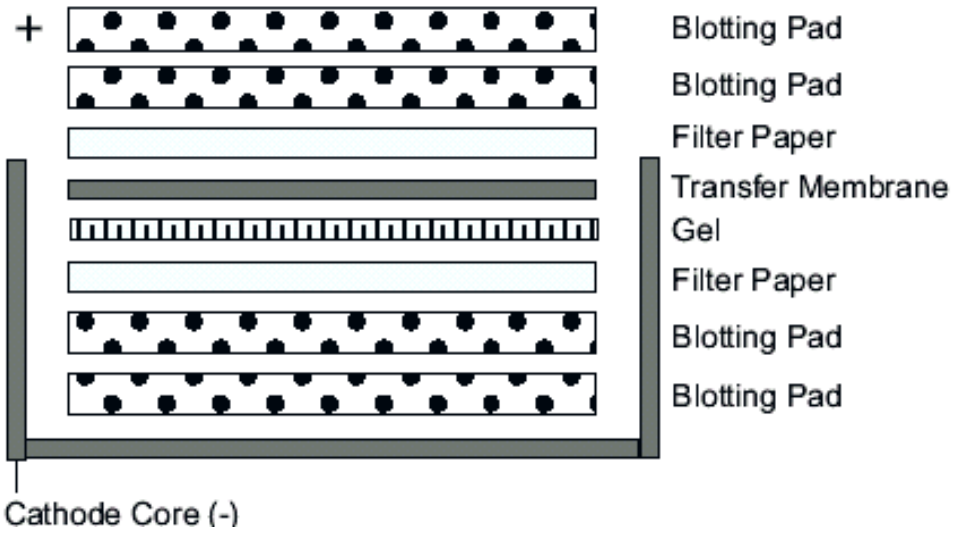


Figure 2.2. Diagram of the blot module set up.

This was placed into an XCell II™ Blot Module, which was filled with transfer buffer. The tank was filled with distilled water and the transfer was conducted at 30V for 1 hour.

Once the electroblotting was complete, the gel and membrane were removed from the blot module and the membrane was rinsed twice with ultra-pure water. The membrane was then washed twice for 5 min with ultra-pure water on a rotary shaker. The membranes were incubated in 30mL of Ponceau S solution for 30 seconds with agitation, to check for efficient protein transfer. Two more washes of the membrane for 5 minutes in ultra-pure water were conducted.

2.5.4 *Protein detection using chemiluminescence*

Proteins on the membrane were visualized using the Amersham ECL™ Prime Western Blotting Detection Kit (GE Healthcare). The membrane was placed in 10mL of blocking solution (2.5% marvel milk in TBST (TBS +1% Tween20)) and incubated at room temperature for 1h on a shaker. The membrane was washed in TBST once for 15 mins and three times for 5 mins on a shaker. Primary antibody was prepared to the appropriate concentration (**Table 2.5**) in 10mL of blocking solution. The membrane was incubated in this primary antibody solution overnight at 4˚C with gentle rotation. The next day, the membrane was washed as outlined above and HRP-conjugated anti-rabbit or anti-mouse secondary antibodies (GE Healthcare) were prepared in 1% marvel milk solution. The secondary antibody solution was added to the membrane and the membrane was incubated at room temperature for 1h with gentle shaking. The membrane was washed as outlined above.

Detection was performed according to the manufacturer's protocol. Briefly, equal volumes of solution A and solution B (5 mL in total per membrane) were combined at room temperature and protected from light. The membrane surface was blotted using filter paper to dry it and 4mL of the chemiluminescent substrate (combined solution A and B) was added. The membrane was incubated with the substrate for 5 min in the dark and excess substrate was blotted away. A piece of clean acetate was placed on top of the membrane and a digital image was captured using a LAS3000 digital scanner (FUJIFILM UK Ltd, Bedfordshire, UK). Exposures between 30 sec-30 min were included.

Table 2.5. Antibodies used for western blotting

Antibody	Western blotting dilution	Supplier (catalogue number)
β-catenin (14/Beta-Catenin) mouse monoclonal IgG1	1/1K	BD Biosciences
LEF-1 (C12A5) rabbit monoclonal	1/5K	Cell Signalling Technology (#2230)

Antibody	Western blotting dilution	Supplier (catalogue number)
Flag (m2) mouse monoclonal	1/2K	Sigma (F1804)
GAPDH (D16H11) rabbit monoclonal IgG	1/10K	Cell Signalling Technology (#5174)
Histone H1 (AE-4) mouse monoclonal	1/5K	BioRad (4974-7808)
RUNX1 rabbit polyclonal	1/1K	Sigma (HPA004176)
TCF7 (C46C7) rabbit monoclonal	1/1K	Cell Signalling Technology (#2206)
HRP-conjugated anti-rabbit (NA934)	1/5K	Amersham (GE Health Care)
HRP-conjugated anti-mouse (NXA931)	1/5K	Amersham (GE Health Care)

2.5.5 *Densitometric analysis of proteins*

The digital images were analysed using Advanced Image Data Analyzer (AIDA) software v4.2.0 (Raytek Scientific, UK) and a pixel-based densitometry approach, in which the intensity of pixels were used to quantitate the protein on the blot. A region of interest was generated around the band(s), which produced a peak intensity histogram. This was used to measure background pixel intensity (calculated using the area surrounding the protein band). The area under the curve was used to calculate an intensity value for the band(s), which allowed estimates of percentage overexpression or underexpression/knockdown to be calculated.

2.6 Mass spectrometry

2.6.1 *Experimental approach*

In this project a mass spectrometric (MS) approach was used to identify candidate nuclear localization factors (**Figure 2.3**). In this approach, AML cell lines were fractionated to isolate/enrich cytoplasmic and nuclear proteins, and β-catenin immunoprecipitation was conducted prior to MS analysis of β-catenin interaction partners. β-catenin interaction partners present in the cytoplasm and nucleus of AML cell lines that i) freely translocate β-catenin to the nucleus and ii) resist the nuclear accumulation of β-catenin were compared. Based on the localization of β-catenin interaction partners in these cell lines, candidate factors regulating the nuclear localization of β-catenin were identified (Chapter 4).

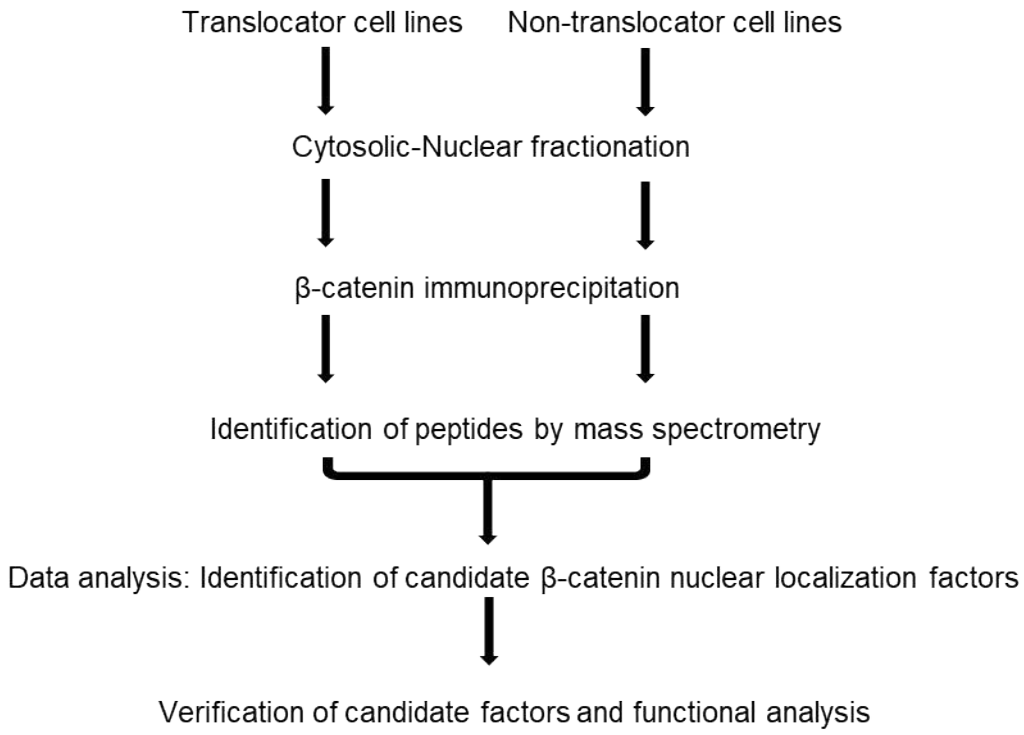


Figure 2.3. Diagram of the MS approach used in this project

2.6.2 *Cytoplasmic/nuclear fractionation using CHAPS buffer*

CHAPS buffer (0.8% CHAPS, 150mM NaCl, 30mM Tris-HCl, 1mM DTT, PIC) was used to fractionate AML cell lines prior to β-catenin immunoprecipitation. To isolate cytoplasmic fractions, $1x10^7$ cells were pelleted and washed twice with 20mL TBS, followed by centrifugation for 10 min at 280 x g. Next, cells were resuspended in 1mL CHAPS buffer, vortexed for 15 seconds and incubated on ice for 20 min followed by a quick vortex and centrifugation at 3000 x g for 8 minutes in a 4˚C microcentrifuge. The supernatant (cytoplasmic extract) was removed to a clean mini Eppendorf tube and stored at -80°C. To isolate nuclear fractions, a nuclear pellet was generated following the steps outlined in 2.5.1, but scaled up for $1x10^7$ cells and the pellet was washed with 500µL of pellet wash buffer (1 x PBS + 5mM $MgCl_2$) followed by centrifugation at 3000 x g for 3 minutes. The supernatant was removed, and the nuclear pellet was snap frozen in liquid nitrogen. The pellet was then freeze thawed three times, and 1µL of Benzonase was added directly to the pellet. The tube was tapped to mix and incubated on ice for 30-60 min. During this incubation, the tube was tapped regularly. 250µL of CHAPS buffer was added and the tube was incubated on ice for 30 min, with vortexing every 10 minutes. The tube was centrifuged at 10,000xg for 10 mins and the supernatant (nuclear extract) was transferred to a clean tube and stored at -80°C.

2.6.3 *Immunoprecipitation using protein G agarose beads*

Immunoprecipitations were first attempted using protein G agarose beads (Santa Cruz). First a preclearing step was conducted by combining 300µg of protein, 10µL of agarose G beads and 0.5µg of mouse IgG1 antibody (**Table 2.6**) in 2mL of CHAPS buffer and incubating on ice for 30 min. The sample was centrifuged at 580 x g for 5 min at 4°C and the supernatant (precleared lysate) was transferred into two clean 1.5mL protein LoBind Eppendorf tubes (Sigma). 2µg of β-catenin antibody (14/Beta-Catenin) or mouse (G3A1) control antibody (**Table 2.6**) was added and the tubes were incubated on ice for 1h. 20µL of agarose beads was added and the tubes were incubated overnight at 4°C with inversion/rotation. The next day the tubes were centrifuged at 580 x g for 5 min at 4°C and the supernatant was transferred to a new Eppendorf tube and stored

at -80°C. The beads (with immunoprecipitated β-catenin) were washed four times with 1mL CHAPS buffer and centrifuged at 580 x g for 5 min at 4°C. 40µL of LDS buffer (NuPAGE® 1xLDS, 50nM NuPAGE® reducing agent) was added to the beads, well mixed and the tubes were incubated at 70˚C for 10 minutes to elute the proteins from the beads.

2.6.4 *Immunoprecipitation using protein G Dynabeads*

Following unsuccessful immunoprecipitation using protein agarose G beads, a different protocol using protein G Dynabeads was used prior to MS. All steps unless otherwise specified were performed on ice. Firstly, β-catenin or IgG1control antibodies were crosslinked to the beads. To do this, the protein G Dynabeads were vortexed to mix them and 165µL of beads (enough for 5 immunoprecipitations) were placed into a 1.5mL protein LoBind Eppendorf tube, a magnet was applied and the supernatant (storage buffer) was removed. The beads were then resuspended in 165µL of PBST (*v/v* PBS + 0.02% Tween-20), 40µg of antibody was added and then the mixture was made up to 1mL in PBST. The tube was incubated at room temperature for 2h with rotation. Following incubation, a magnet was applied to the tube and the supernatant was discarded. The antibody-bound beads were washed three times in PBST with inverting. The beads were then washed three times in coupling buffer (PBST + 0.2M triethanolamine pH 9) before incubation with 1mL freshly prepared dimethyl phthalate (DMP) solution (coupling buffer + 2mM DMP) for 30 min at room temperature with rotation. A magnet was applied, and the supernatant was discarded before repeating the incubation with fresh DMP solution. The DMP solution was discarded and the beads were incubated for 30 min at room temperature with 1mL of quenching buffer (PBST + 50mM ethanolamine). The beads were washed three times in elution buffer (0.2M glycine pH2.5, 0.01% Tween-20 pH.2.5 dh_20) and then resuspended in 165µL of TBST and stored at 4° C.

Following crosslinking of antibodies to beads, immunoprecipitation of β-catenin and its binding partners was conducted. First a preclearing step was conducted. 33µL of IgG1 control antibody-bound beads were placed in an Eppendorf tube and the supernatant was removed and replaced with 33µL of CHAPS buffer. One

thousand µg of protein lysate was added and made up to 500µL in CHAPS buffer. The tube was incubated for 6 hours in a cold room with rotation. A magnet was applied and the supernatant (precleared lysate) was transferred to a clean tube. Thirty-three µL of antibody-bound beads was added to the lysate and the tube was incubated overnight in the cold room with rotation. The next day, a magnet was applied to the tube and the supernatant was transferred to a clean tube and stored at -80° C for use in western blotting analysis of immunoprecipitation efficiency. The beads were washed 6 times with 500µL of CHAPS buffer with inverting and incubation on ice for 5 min. To elute β-catenin and its binding partners from the beads, they were boiled in 2x LDS buffer (NuPAGE® 2xLDS, 50nM NuPAGE® reducing agent and dH20) at 95°C for 5 min. The eluted protein was stored at -20° C.

Table 2.6. Antibodies used in immunoprecipitation

Antibody (clone)	Supplier (catalogue number)
β-catenin (14/Beta-Catenin) mouse monoclonal IgG1. Binds β-catenin amino acids 571-781.	BD Biosciences (610154)
β-catenin (L87A12) mouse monoclonal IgG1. Binds residues surrounding Asp56 of β-catenin.	Cell Signalling Technology (#2698)
β-catenin rabbit polyclonal. Binds residues surrounding Asp56 of β-catenin.	Cell signalling technology (#9581)
β-catenin (6B3) rabbit monoclonal IgG. Binds C-terminus of β-catenin.	Cell Signalling technology (#9582)
β-catenin (H-102) rabbit polyclonal. Binds β-catenin amino acids 680-781.	Santa Cruz (sc-7199)
Mouse monoclonal IgG1 (G3A1)	Cell Signalling Technology (#5415)
Rabbit monoclonal IgG (DA1E)	Cell Signalling Technology (#3900)

2.6.5 *Overview of MS*

Mass spectrometers measure the mass-to-charge ratio (m/z) of ionized molecules, and can be used to identify proteins in a complex mixture. For protein identification, the proteins are cleaved into smaller peptides, their absolute mass is measured by a mass spectrometer and compared to a protein sequence database for peptide/protein identification (Henzel *et al*, 1993; James *et al*, 1993; Mann *et al*, 1993; Pappin, 2003). In this project, a tandem MS (MS/MS) approach, outlined in **Figure 2.4** was used to identify peptides. Briefly, the AML cell line nuclear and cytoplasmic fractions were run on an SDS-PAGE gel, each gel lane was cut into slices and in-gel tryptic digestion was performed. The peptides were then fractionated, and ionized prior to mass spectra acquisition by tandem MS. These spectra were processed and quantified before database searching.

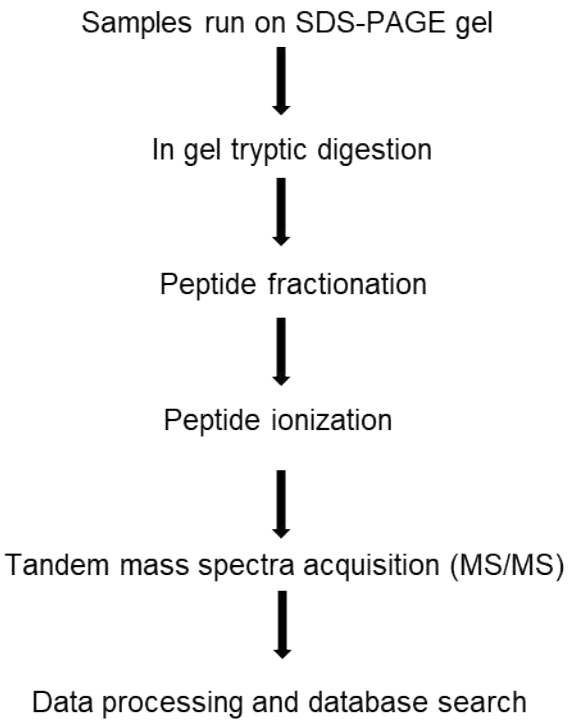


Figure 2.4. Overview of MS procedure.

2.6.6 *MS procedure*

The samples were sent to Bristol Proteomics Facility (Bristol University, UK) for MS to be performed. There, the samples were run on a 10% SDS-PAGE gel until the dye front moved ~3 cm into the separating gel. Each lane was cut into 3 equal slices, followed by in-gel tryptic digestion of each slice using a DigestPro automated digestion unit (Intavis Ltd.) The peptides were fractionated using an Ultimate 3000 nanoHPLC system in conjunction with an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific). Briefly, peptides in 1% formic acid were injected onto an Acclaim PepMap C18 nano-trap column (Thermo Scientific). The peptides were washed with 0.5% acetonitrile and 0.1% formic acid peptides were resolved on a 250 mm × 75 μM Acclaim PepMap C18 reverse phase analytical column (Thermo Scientific). This was done over a 150min organic gradient, using 7 gradient segments of aqueous 80% acetonitrile in 0.1% formic acid (1-6%, 6-15%,15-32%, 32-40%, 40-90%, 90% and then reduced to 1%) with a flow rate of 300 nL min^{-1}.

Peptides were ionized at 2.1 kV using a stainless-steel emitter with an internal diameter of 30 μM (Thermo Scientific) and a capillary temperature of 250°C. Tandem mass spectra were measured using an LTQ- Orbitrap Velos mass spectrometer and Xcalibur 2.1 software (Thermo Scientific) in data-dependent acquisition mode. The Orbitrap was configured to analyse survey scans at 60,000 resolution (at m/z 400) in the mass range m/z 300 to 2000. The top twenty multiply charged ions in each duty cycle were selected for MS/MS in the LTQ linear ion trap. Charge state filtering (where unassigned precursor ions were not selected for fragmentation) and dynamic exclusion (repeat count, 1; repeat duration, 30s; exclusion list size, 500) were used. The LTQ fragmentation conditions were set to normalized collision energy, 40%; activation q, 0.25; activation time 10ms; and minimum ion selection intensity, 500 counts.

The raw data files were processed/quantified using Proteome Discoverer software v1.4 (Thermo Scientific) and the UniProt Human database (131351 entries) was searched using the SEQUEST algorithm. MS/MS tolerance was set to 0.8Da and peptide precursor mass tolerance was set at 10ppm. MS/MS. searches were performed with full tryptic digestion and a maximum of 1 missed

cleavage was allowed. The reverse database search option was enabled, and all peptide data was filtered to satisfy false discovery rate (FDR) of 1% or 5%.

The output of this analysis was 2 Excel Spreadsheets, one for FDR 1% and one for FDR 5%. These contained a list of proteins that were identified by MS, which samples they were present in, and metadata about each of the peptides, including a peptide score, molecular weight and number of amino acids. The spreadsheets containing the raw data are included in Appendix 4 (see the attached CD disk).

3 Development of fractionation and immunoprecipitation techniques for analysis of leukaemia cell lines by MS

3.1 Introduction

β-catenin is dysregulated in AML and its overexpression is associated with a poor prognosis (Chen *et al,* 2009; Simon *et al*, 2005; Xu *et al*, 2008; Ysebaert *et al*, 2006). Key to its role as a transcription factor is its localization to the nucleus, and nuclear unphosphorylated β-catenin is observed in ~50% of AML blasts (Xu *et al*, 2008). In AML blasts and cell lines, nuclear β-catenin levels do not correlate with cytoplasmic β-catenin levels (Morgan *et al*, 2014), suggesting that cytoplasmic accumulation of β-catenin does not necessarily lead to its nuclear localization and there are additional factors regulating this process. Factors that regulate β-catenin nuclear localization in other contexts have been identified (1.3.6), but little is known about the nuclear localization of β-catenin in AML. In this project, an MS approach was used to identify factors that regulate the nuclear localization of β-catenin in AML (**Figure 3.1**). In this approach, cytoplasmic and nuclear fractionation of cell lines was conducted, followed by β-catenin immunoprecipitation and MS identification of β-catenin binding partners. Comparisons of binding partners in translocator and non-translocator cell lines were made to identify candidate β-catenin nuclear localization factors, and functional analysis of identified candidates was conducted. This approach was taken rather than focusing on previously identified β-catenin nuclear localization factors (**Table 1.10**) to ensure that important factors regulating the nuclear localization of β-catenin in AML were not overlooked. Analysing known β-catenin nuclear localization factors would be time and labour intensive and may not lead to the identification of β-catenin nuclear localization factors in AML. This is because proteins are expressed and regulated differently depending on their context, so proteins that regulate β-catenin nuclear localization in other contexts will not necessarily regulate it in AML. An MS approach allows multiple

candidates to be identified, increasing the likelihood of finding key β-catenin nuclear localization factors in AML. This chapter outlines the optimization steps taken to efficiently extract β-catenin in leukaemia cell line fractions and immunoprecipitate β-catenin ready for MS analysis. This involved i) expression of epitope tagged β-catenin for use in pull down experiments (to decrease the likelihood of background binding which would interfere with identification of candidate localization factors by MS analysis) or expression of wild type β-catenin (to increase the amount of β-catenin available for analysis) ii) cytoplasmic/nuclear fractionation with CHAPS buffer (to maintain protein-protein interactions that are important in the identification of interaction partners by MS) and iii) immunoprecipitation of β-catenin and its binding partners.

3.2 Aims

- To express FLAG-β-catenin in leukaemia cells lines for use in immunoprecipitation.
- Develop a fractionation technique for leukaemia cell lines using a detergent that maintains protein-protein interactions.
- To optimize immunoprecipitation of β-catenin.

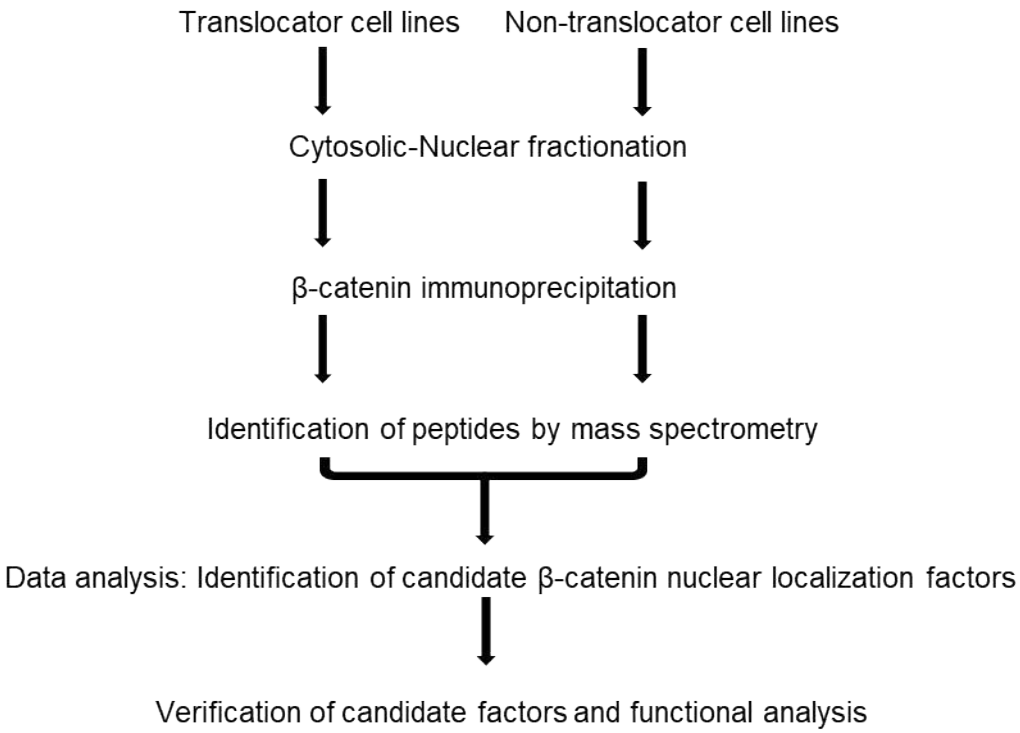


Figure 3.1. Schematic of the experimental approach to identify candidate β-catenin nuclear localization factors in leukaemia lines.

This diagram summarizes the steps taken to identify candidate β-catenin nuclear localization factors in leukaemia lines.

3.3 Results

3.3.1 *Overexpression of β-catenin constructs in K562 cells*

To isolate β-catenin binding partners in leukaemia cell lines, epitope tagged β-catenin was chosen for use in pulldown experiments, because antibodies raised against epitope tags have a low reactivity with other cellular proteins compared to those raised against endogenous proteins. This was an important consideration because using samples with the lowest possible background binding was key to the success of my MS approach, outlined in (**Figure 3.1**).

Initially I adopted a FLAG-tagged β-catenin overexpression system. FLAG is an artificial antigen (DYKDDDDK) to which high affinity monoclonal antibodies are available (this study employed clone M2 (**Table 2.5**). The FLAG tag is relatively small at 1kDa, decreasing the risk of it interfering with β-catenin protein function. Out of the leukaemia cell lines used in the experimental approach, K562 cells were chosen to test the methodology and reagents before transduction of other lines. This was because they are a rapidly growing cell line and are easy to transduce.

3.3.1.1 *Creation of a FLAG-β-catenin lentiviral construct*

A FLAG-β-catenin pcDNA3 construct was purchased from Addgene. In this construct the FLAG tag is attached to the C-terminal end of β-catenin. This was subcloned into the lentiviral vector (pHIV-EGFP) (**Table 2.3**). This was done because the lentiviral vector is more efficient at stably transducing cells than the original pcDNA3 vector.

Briefly, the original vector (pcDNA3-β-catenin-FLAG) was digested with restriction enzymes to provide some evidence that the construct was correct. To do this, PDRAW32 (AcaClone) was used to identify restriction sites on the vector that could be used as a strategy to help validate the construct. **Figure 3.2A** shows a diagrammatic representation of the vector and the two enzymes that

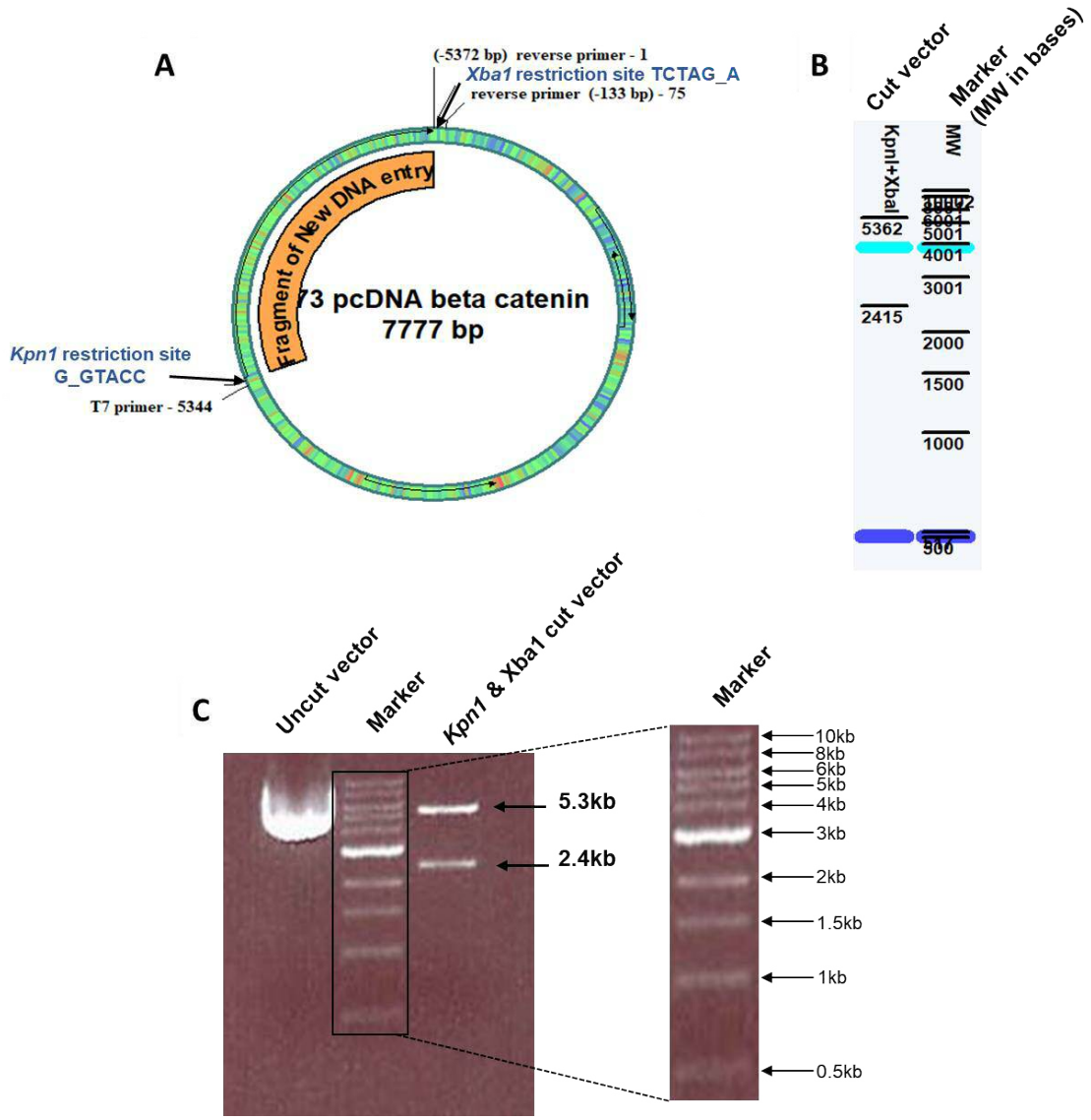


Figure 3.2. Test digest of FLAG-β-catenin pcDNA3 vector (Addgene).

(A) A diagram of the FLAG-β-catenin pcDNA3 vector purchased from Addgene. Restriction sites for *Kpn1* and *XbaI* are highlighted. This diagram was generated using PDRAW32 (Acaclone). **(B)** Expected banding pattern following digestion of the vector with *KpnI* and *XbaI* enzymes. **(C)** Banding pattern observed following digestion of the vector with these enzymes. Abbreviations; MW= molecular weight, kb= kilobases. Marker=NEB 1kb DNA ladder.

were chosen for use in the digestion *(KpnI* and *XbaI*). Each of these enzymes was a 'single cut' vector, with restriction sites in the multiple cloning sites, in-between which the FLAG-β-catenin construct was located. **Figure 3.2B** shows the predicted banding pattern observed when agarose gel electrophoresis (2.2.2) was used to analyse the digest. This included a band at 5.3kb corresponding to the digested vector backbone and a band at 2.4kb corresponding to the excised FLAG-β-catenin. **Figure 3.2C** shows the result of the digestion of the pcDNA3 vector with *KpnI* and *XbaI*. As predicted in **Figure 3.2B**, bands were visible at ~5.3kb and ~2.4kb. This verified that the construct was likely correct and could be used in sub-cloning to transfer FLAG-β-catenin to the lentiviral pHIV-EGFP vector.

For subcloning, *KpnI* was used to digest the FLAG- β-catenin-pcDNA3 vector, and Klenow fragment was used to selectively blunt this end. Next, *XbaI* was used to excise the FLAG-β-catenin from the pcDNA3 vector, for use in the subcloning procedure. Correspondingly, the destination vector, pHIV-EGFP, was prepared by digestion at the multiple cloning site with *XbaI* and the blunt cutter, *HpaI,* allowing directional subcloning of the FLAG-β-catenin into the vector under the control of the EF1A promoter. The pHIV-EGFP vector and FLAG-β-catenin insert were ligated (2.2.1) and Stbl-3 cells were transformed with the ligated vector to amplify the DNA. Three transformed colonies were selected, and the DNA was purified (2.2.3).

Test digests were performed on the DNA from three different colonies to check that the pHIV-EGFP-β-catenin-FLAG vector had been successfully created. **Figure 3.3A** shows a diagrammatic representation of the EGFP-β-catenin-FLAG vector and the restriction for *BamHI* (the enzyme chosen to be used in the test digest). The restriction sites (G'GATC) for this enzyme occurred twice in the constructed vector, at the multiple cloning site of the lentiviral vector and in the 5' sequence of the insert. **Figure 3.3B** shows the banding pattern expected when the samples were electrophoresed on an agarose gel if the subcloning was successful. This included a band at 7.6kb corresponding to the cut vector backbone and a band at 2.4kb corresponding to the excised FLAG-β-catenin. **Figure 3.3C** shows the results of the test digest, following digestion of the vector

with the *BamHI* enzyme. The vector isolated from colony 1 had visible bands at ~7.7kb and ~2.4kb, as predicted in **Figure 3.3B.**

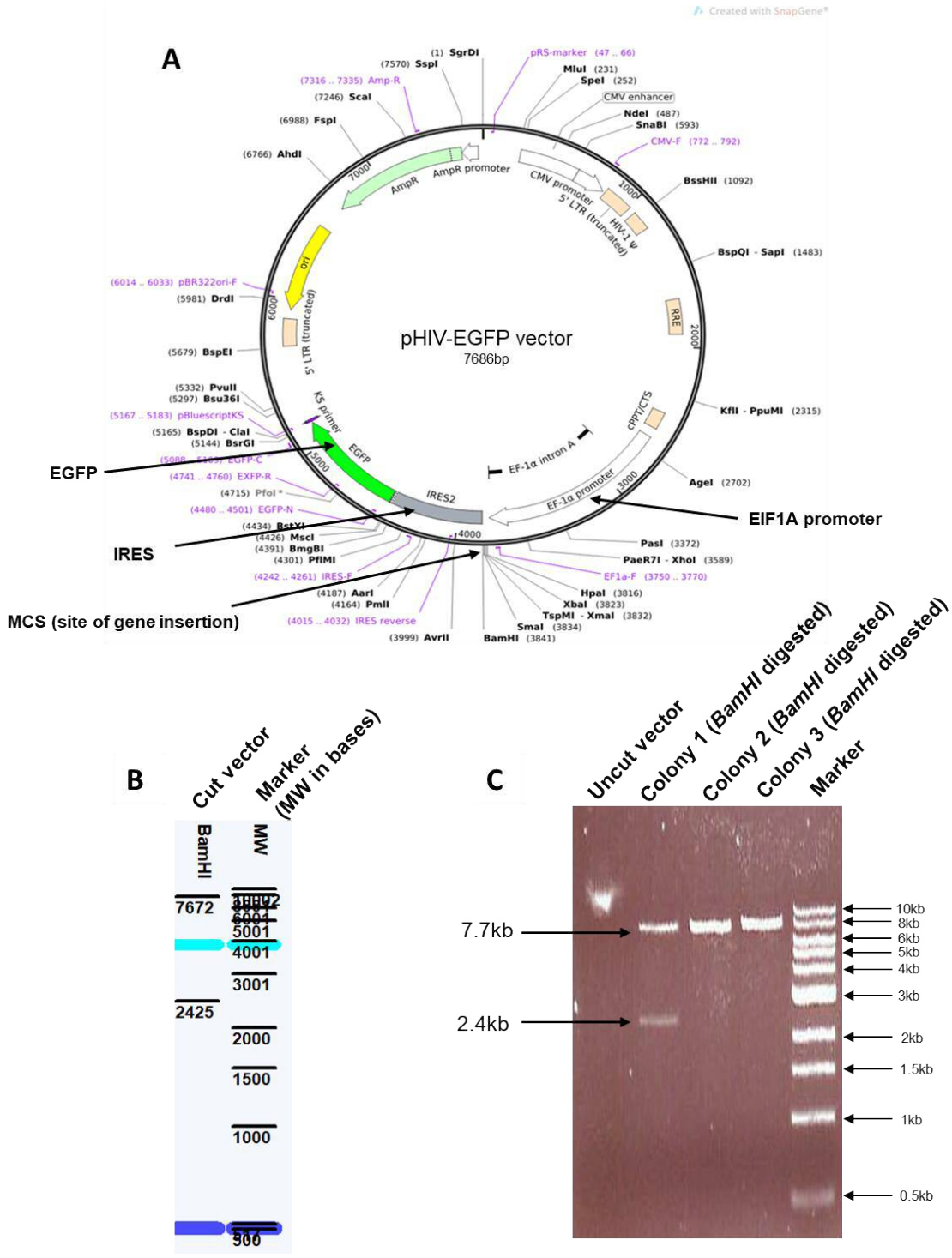


Figure 3.3. Test digest of FLAG-β-catenin pHIV-EGFP vector.

(A) A diagram of the β-catenin-FLAG pHIV-EGFP vector. Expression of the inserted gene is driven by the EF1A promoter, which generates an mRNA which also contains an internal ribosome entry site (IRES) and an EGFP sequence. **(B)** The predicted banding pattern following digestion of the vector with the *BamHI* enzyme. **(C)** The banding pattern observed following digestion of the vector from transformed *E.coli* colonies (Stbl3) with *BamHI*. Abbreviations; MCS= multiple cloning site, MW= molecular weight, kb= kilobases. Marker= NEB 1kb DNA ladder.

The other colonies analysed did not contain insert. This suggested that the construct in colony 1 was the successfully ligated pHIV-EGFP-β-catenin-FLAG that could be used in subsequent experiments to transduce leukaemia cell lines with FLAG-β-catenin. To verify this was the case, samples of the construct were sent for sequencing (Eurofins). The results (Appendix 5) verified that the construct was correct.

3.3.1.2 *Transduction of K562 cells with FLAG-β-catenin*

After successful generation of the FLAG-β-catenin construct, K562 cells were lentivirally transduced with the pHIV-EGFP-β-catenin-FLAG or pHIV-EGFP control vector (2.3). Briefly, K562 cells were transduced, BIO induced (2.1.5), fractionated into cytoplasmic and nuclear lysates and western blotted using the FLAG M2 antibody (2.5). K562 cells transduced with the empty pHIV-EGFP vector were included as a control. Cells were also blotted with β-catenin antibody (BD) in addition to the FLAG M2 antibody.

Figure 3.4A shows the results of this test transduction in the K562 cells. When cells were immunoblotted with β-catenin antibody (BD) there was only a faint signal in the cytoplasmic or nuclear fractions of uninduced K562 cells. In contrast, in induced cells (+BIO), bands were visible in the cytoplasmic and nuclear fractions at ~90kDa, corresponding to full length β-catenin; however, there was no difference in overall β-catenin expression between the control and the FLAG-β-catenin transduced cells.

When cells were blotted with FLAG antibody (Sigma) there was no visible signal in the cytoplasmic or nuclear fractions of uninduced K562 cells. This was the same in both control and FLAG-β-catenin transduced cells. In addition, no signal was detected in induced cells. This suggested that the FLAG-β-catenin construct was either not being expressed, or the FLAG antibody was not successfully binding to FLAG.

To verify that the K562 cells had been successfully transduced with the FLAG-β-catenin construct, flow cytometry was used (2.4) to measure GFP expression (expressed from an IRES in the same mRNA as FLAG-β-catenin). **Figure 3.4B&C** shows the results of this flow cytometric analysis.

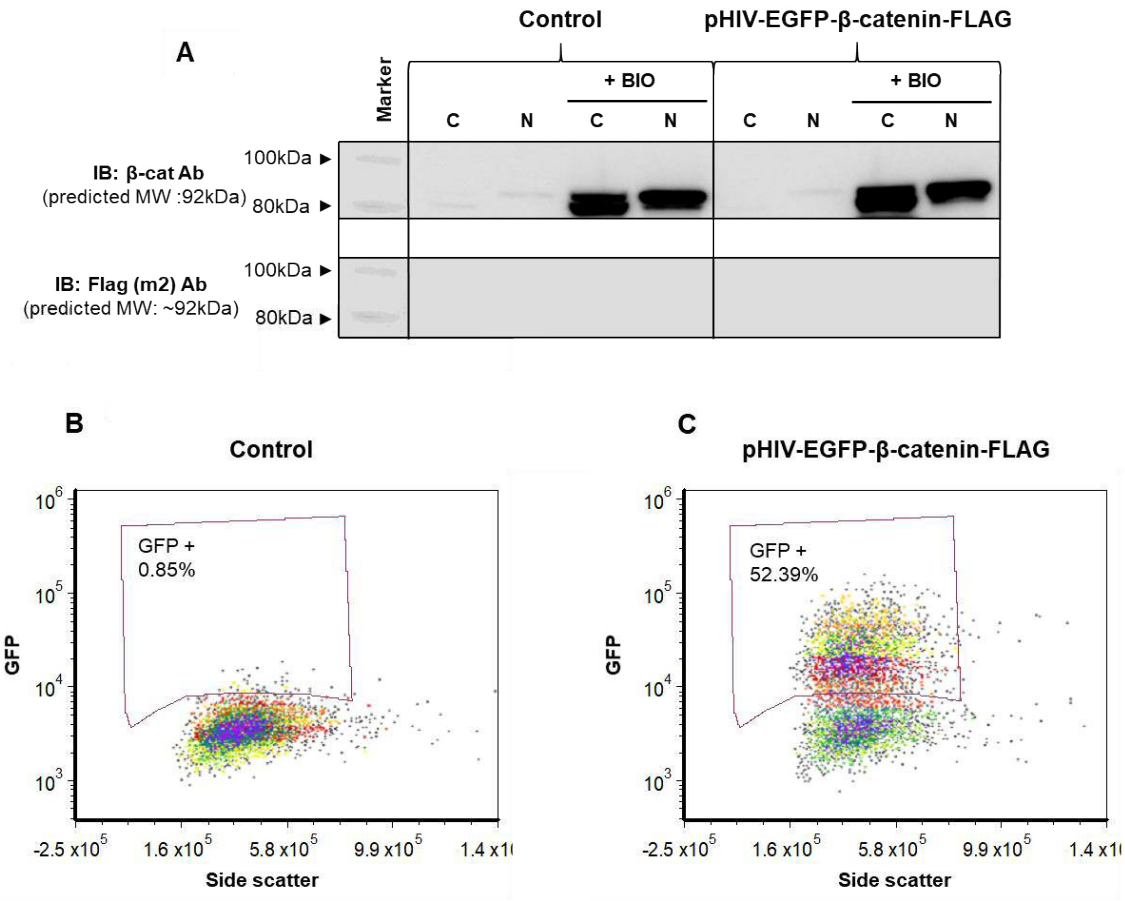


Figure 3.4. Transduction of K562 cells with flag-tagged β-catenin.

(A) Western blotting of K562 cytoplasmic and nuclear fractions following transduction with control and FLAG-β-catenin vectors. BIO was used to stabilize β-catenin. Abbreviations; C= cytoplasmic fraction, N= nuclear fraction, IB= immunoblot, β-cat= β-catenin, Ab= antibody, MW= molecular weight, kDa= kilodalton. Marker = MagicMark XP (Thermo Fisher). **(B)** Measurement of auto fluorescence in control transduced K562 cells by flow cytometry. Region defines threshold (c1%) for GFP expression in C. **(C)** Measurement of GFP expression in K562 cells transduced with pHIV-EGFP-β-catenin-FLAG by flow cytometry. GFP= green fluorescent protein. To gate out dead cells, a viable gate was applied which excluded cells permeable to the DNA-intercalating fluorescent compound, 7-AAD (data not shown).

Background auto fluorescence was established using untransduced K562 cells. In FLAG-β-catenin transduced cells, the FL-1 (GFP emission detection channel) signal was 52%. This suggested that the FLAG-β-catenin construct was transduced into ~50% of the cells and that mRNA for FLAG-*β-catenin* is being expressed in these cells suggesting there must be another reason for the lack of protein expression in K562 cells observed in **Figure 3.4A**. These and later data (discussed below) suggested that there may be an issue of protein stability with the FLAG-β-catenin construct. To rule out the possibility that the FLAG epitope was responsible for this, I assessed whether overexpression could be achieved in K562 cells using wildtype β-catenin. Despite the robust increase in β-catenin following BIO induction, I wanted to also overexpress β-catenin to further increase the amount available for immunoprecipitation. This was an important consideration because the more β-catenin that was immunoprecipitated, the more likely β-catenin interaction partners would be identified by mass spectrometry, particularly low abundance or transiently binding partners.

3.3.1.3 *Transduction of K562 cells with β-catenin polyPOZ*

To assess whether wildtype β-catenin could be overexpressed, K562 cells were transduced with a β-catenin polyPOZ retroviral vector which co-expresses lacZ as a selectable marker (**Table 2.3**). Transduced K562 cells were incubated with the lacZ fluorogenic substrate, FDG, and FACS sorted (based on the presence of lacZ in the vector) (2.4.2). **Figure 3.5** shows the effectiveness of this FDG sorting. Prior to FDG sorting only 21% of cells were lacZ positive **(Figure 3.5A**), suggesting that the β-catenin polyPOZ vector was only expressed in ~20% of cells. After FDG sorting the percentage of cells that were lacZ positive increased to 74% (**Figure 3.5B**), suggesting that the β-catenin polyPOZ vector was enriched to ~70% of cells. These cells were fractionated and immunoblotted with β-catenin antibody (2.5). **Figure 3.5C** shows the results of this western blot. Despite the successful transduction of cells with the β-catenin polyPOZ vector, the bands observed at ~90kDa (corresponding to full length β-catenin) were at a similar intensity for both control and β-catenin transduced cells. This suggested that there was no overexpression of β-catenin in the transduced cells. Owing to the similar results observed for the FLAG-β-catenin construct, this suggested that

a mechanism may be present in these cells to prevent the stable overexpression of β-catenin (3.4.1).

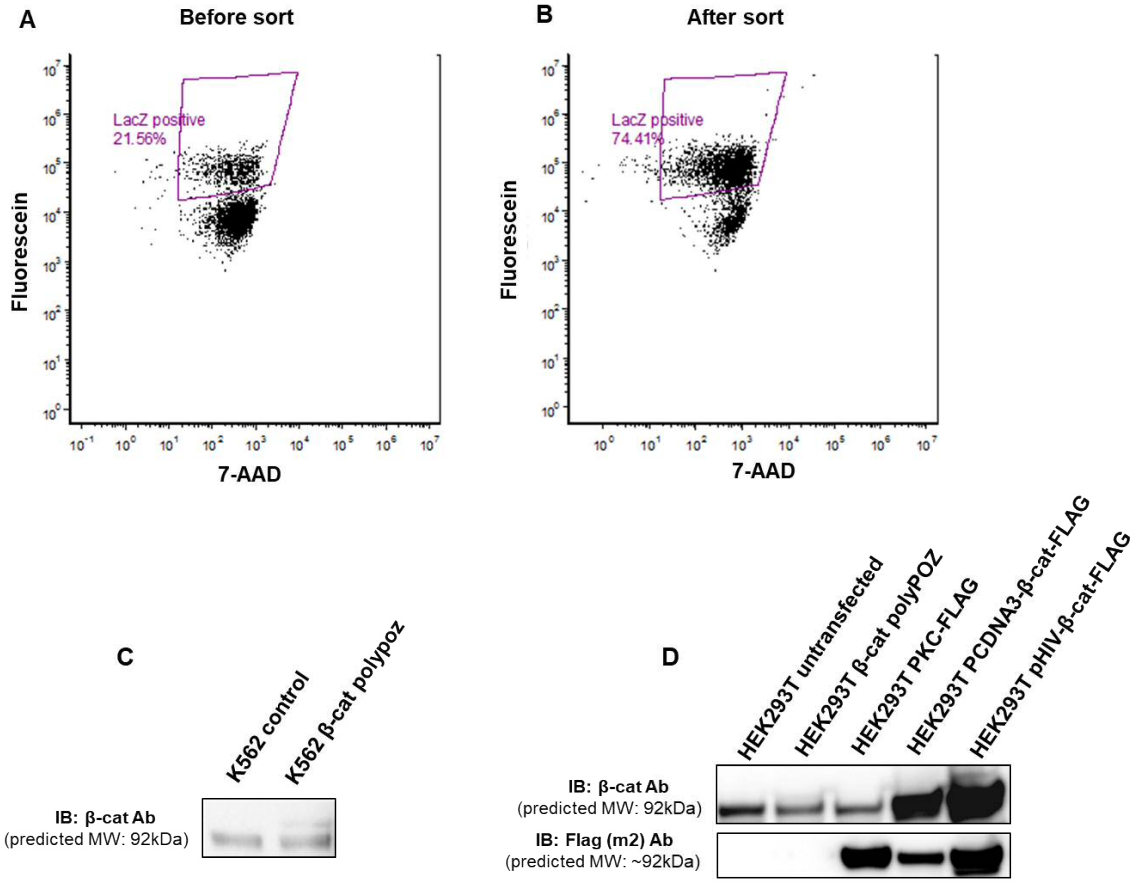


Figure 3.5. Transduction of cells with β-catenin polyPOZ and flag-tagged β-catenin.

(**A**) Percentage of K562 cells with fluorescein following transduction with β-catenin polyPOZ. (**B**) Percentage of K562 cells with fluorescein following sorting. 7-AAD was included to check the viability of the cells and debris was excluded using forward vs. side scatter (data not shown). (**C**) Western blotting of K562 whole cell lysates from sorted cultures of K562 cells transduced with control and β-catenin polyPOZ vectors. (**D**) Western blotting of HEK293T whole cell lysates following transfection with β-catenin vectors. The western blot marker was not visible at the relevant MW, but the relative migration was consistent with the signal being due to β-catenin. Abbreviations; IB= immunoblot, Ab = antibody, kDa= kilodalton, MW= molecular weight.

3.3.1.4 *Transfection of HEK293T cells with β-catenin and FLAG-β-catenin*

To determine if the lack of β-catenin construct overexpression was specific to K562 cells, the β-catenin and FLAG-β-catenin constructs were overexpressed in a different context (HEK293T cells). HEK293T cells were transfected with both β-catenin vectors, lysed and western blotted with β-catenin and FLAG M2 antibodies (Chapter 0). A FLAG-PKC construct was included as a positive control to verify the FLAG antibody. The FLAG-β-catenin pcDNA3 vector was also included as another comparison, to determine if the sub-cloning of FLAG-β-catenin into the pHIV-EGFP vector (3.3.1.1) had led to issues with overexpressing the protein.

Figure 3.5D shows the results of these western blots. Upon blotting with β-catenin antibody, a band was visible at ~90kDa in all lysates (corresponding to full length β-catenin). The intensity of this band did not increase upon transfection of HEK293T cells with β-catenin polyPOZ. This suggested that the β-catenin polyPOZ vector was not overexpressed in this line and could indicate a problem with this vector. In contrast, the intensity of the banding pattern increased in cells transfected with the FLAG-β-catenin vectors. This suggested that FLAG-β-catenin was expressed in the transfected HEK293T cells.

Upon blotting with FLAG M2 antibody, as expected no bands were visible in the untransfected cells nor in cells transfected with β-catenin polyPOZ; however, a band was visible at ~90kDa in the lysates of cells transfected with the positive control PKC-FLAG (which has a similar molecular weight to β-catenin) and both vectors containing FLAG-β-catenin.

To conclude, the β-catenin polyPOZ vector was not expressed in HEK293T cells, suggesting that there could be a problem with this construct. In contrast, FLAG-β-catenin vectors were expressed in HEK293T cells. This suggested that the lack of expression in K562 cells observed in 3.3.1.2 was context specific and did not arise from problems with vector design or performance.

Owing to the general issues of overexpression of β-catenin and the time constraints of the project, the decision was made to rely on endogenous β-catenin in subsequent experiments, and use cell lines treated with BIO to increase the

levels of β-catenin in both the cytosolic and nuclear compartments sufficiently for immunoprecipitation and MS analysis.

3.3.2 *Fractionation of K562 cells using CHAPS buffer*

The next step in optimizing the methodology for the MS approach was to optimize a different cytoplasmic-nuclear fractionation method to the one used so far (2.6.2). The reason for optimizing a new technique was because of the proprietary formulation of the Biovision C/N fractionation kit used in our standard laboratory method (incorporating buffer CEB-A) and the fact that common lysis buffer components are incompatible with MS. The new cytoplasmic-nuclear fractionation method was optimized in K562 cells prior to its use in other cell lines.

First, the extraction of β-catenin from the cytosol and nucleus of K562 cells using the detergent 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) was tested. CHAPs was chosen as a potential detergent for the fractionation buffer, because it had been shown to maintain protein-protein interactions in other contexts (Kim *et al*, 2002; Rosner & Hengstschläger, 2012). Maintaining protein-protein interactions was important in my experimental approach, which relied on identifying β-catenin binding partners. A range of CHAPS concentrations (0.5%, 0.6%, 0.8%, 1%) were tested and compared to our standard protocol buffers (CEB-A and TEAB extraction buffer (2.5.1)) to identify the most effective conditions for β-catenin extraction. In our standard protocol, CEB-A is used to isolate cytoplasmic protein and TEAB extraction buffer is used to isolate nuclear protein. Samples were immunoblotted and probed with β-catenin antibody (2.5)

Figure 3.6A shows the relative extraction of β-catenin following fractionation with buffers containing different CHAPS concentrations. In the cytoplasmic fractions, bands were visible at ~90kDa in cells fractionated with CEB-A buffer and cells fractionated with the range of CHAPS concentrations. This corresponded to the molecular weight of full length β-catenin, suggesting that β-catenin was successfully extracted from the cytoplasmic fraction of K562 cells using all the buffers. The intensity of the band was higher in the cytoplasmic extracts fractionated with CHAPS, compared to using our standard CEB-A buffer. This

suggested that using the CHAPS buffer would be a more efficient way of extracting β-catenin from the cytoplasm of leukaemia lines. The percent increase in cytoplasmic β-catenin extraction for each CHAPS concentration compared to extraction with CEB-A was calculated. These were 837% for 0.5% CHAPS, 837% for 0.6% CHAPS, 679% for 0.8% CHAPS and 815% for 1% CHAPS.

The results for the nuclear fractions were similar. Bands were visible at ~90kDa in cells fractionated with TEAB buffer and cells fractionated with the range of CHAPS concentrations. This corresponded to the molecular weight of full length β-catenin, suggesting that β-catenin was successfully extracted from the nuclear fraction of K562 cells, with each buffer. The intensity of the band was higher in the nuclear extracts fractionated with CHAPS, compared to using our standard TEAB buffer. This suggested that using the CHAPS buffer would be a more efficient way of extracting β-catenin from the nucleus of leukaemia cells.

For nuclear extracts, the percent difference in β-catenin nuclear extraction for each CHAPS concentration was compared to extraction with TEAB was calculated. These were a 42% increase for 0.5% CHAPS, a 65% increase for 0.6% CHAPS, a 79% increase for 0.8% CHAPS and a 42% increase for 1% CHAPS. 0.8% CHAPS was chosen for use in subsequent fractionations because it gave a higher amount of β-catenin isolation than using TEAB buffer. It was chosen in preference over 1% CHAPS because lower detergent levels were less likely to interfere with downstream applications (immunoprecipitation and MS).

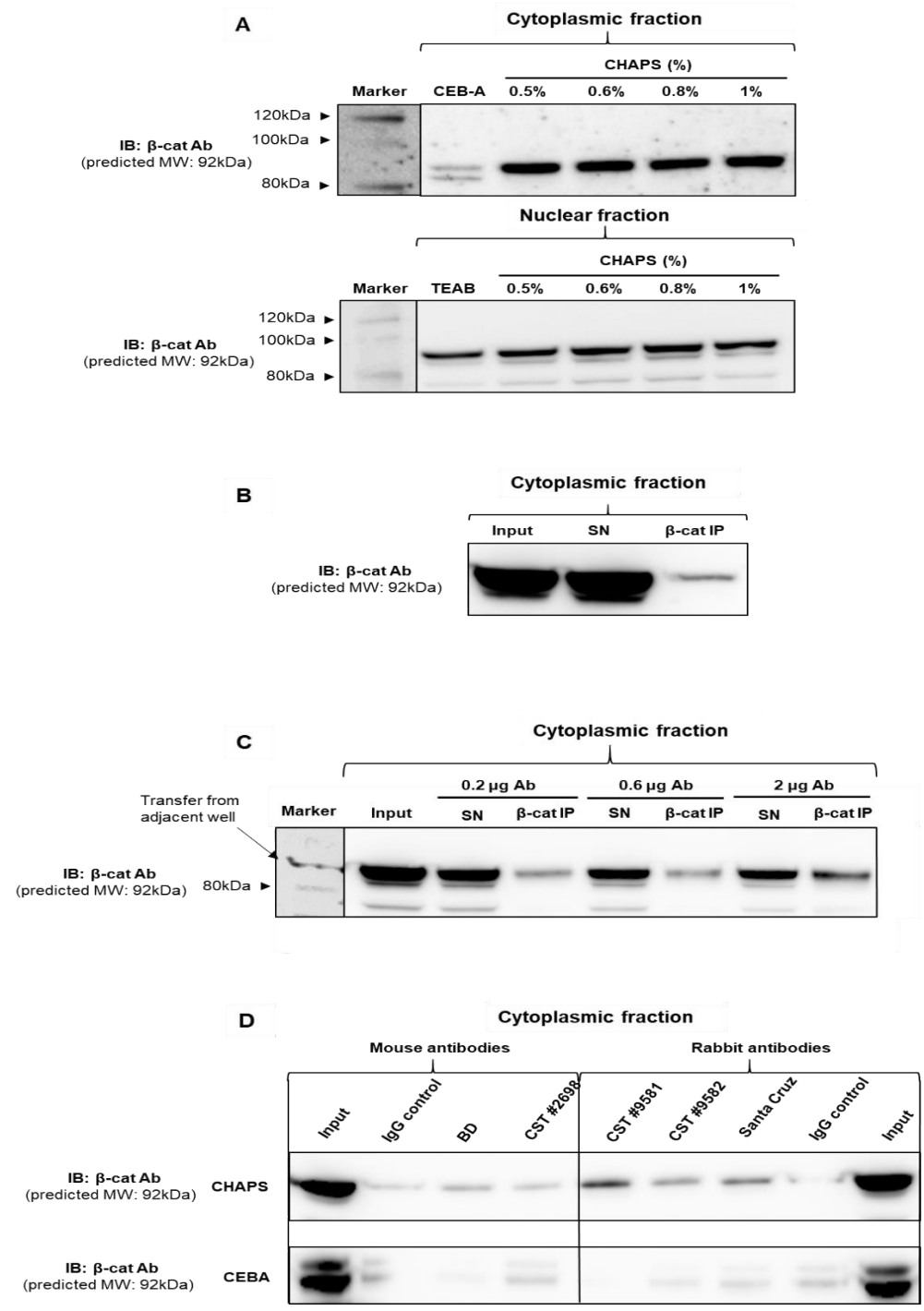


Figure 3.6. Fractionation of K562 cells with CHAPS and immunoprecipitation of β-catenin.

(**A**) Western blotting of K562 fractions following fractionation with buffers containing different levels of CHAPS. CEB-A and TEAB were included as controls. (**B**) β-catenin immunoprecipitation in K562 cytoplasmic fractions using agarose G beads. The marker was not visible at the relevant MW, but the relative migration was consistent with the signal being due to β-catenin. (**C**) β-catenin immunoprecipitation in K562 cytoplasmic fractions using different antibody concentrations. (**D**) β-catenin immunoprecipitation of K562 cytoplasmic fractions using different antibodies. The marker was not visible at the relevant MW, but the relative migration was consistent with the signal being due to β-catenin. See **Table 2.6** for information about the antibodies used. Abbreviations; IB= immunoblot, MW= molecular weight, Ab= antibody, kDa = kilodalton, SN= supernatant (immunodepleted lysate following the IP reaction), IP= Immunoprecipitation. Marker = MagicMark XP (Thermo Fisher).

3.3.3 *Immunoprecipitation of β-catenin using Protein G agarose beads*

Next, immunoprecipitation of β-catenin and its binding partners was optimized in K562 cell fractions. Protein G agarose beads were used in these immunoprecipitations due to their wide usage and low cost. The immunoprecipitation protocol provided with the beads (from Santa Cruz[10]) was used. K562 cells were BIO induced and fractionated (using CHAPS). The cytoplasmic fraction was then immunoprecipitated using β-catenin antibody (BD) and western blotted (Chapter 0).

Figure 3.6B shows the results following blotting with β-catenin antibody. A band was visible in the input, supernatant and immunoprecipitation at ~90kDa, corresponding to the expected molecular weight of full length β-catenin. The intensity of this banding was similar in the input and supernatant lanes, suggesting that only a very small amount of β-catenin had left the supernatant by binding to the beads. The very low intensity of the band observed in the immunoprecipitation verified that only a small amount of β-catenin had bound to the beads. This showed that the immunoprecipitation efficiency was very low, and improving this efficiency became the aim of subsequent optimization steps.

3.3.3.1 *Optimization of immunoprecipitation by changing antibody conditions*

One factor that could have caused the low immunoprecipitation efficiency was an insufficient concentration of antibody in the reaction. This is because if the antibody concentration is too low, not enough antibody binds to the agarose beads and there are less antibody-bead complexes for β-catenin to bind to. A range of different antibody concentrations (0.2µg, 0.6µg and 2µg) were tested in K562 cells, following the same procedure as described previously (3.3.3.). **Figure 3.6C** shows the results following western blotting of the immunoprecipitations. A band was visible at ~90kDa in the inputs, supernatant and immunoprecipitations at each antibody concentration. This corresponded to the expected molecular weight of full length β-catenin. Though the yield of immunoprecipitated protein

[10] http://datasheets.scbt.com/sc-2002.pdf

improved with increasing antibody the efficiency of immunoprecipitation remained well below 50%.

Another factor that could have led to the low immunoprecipitation efficiency was the antibody used. This is because different antibodies have different binding affinities and the low efficiency of β-catenin binding could be due to a low binding efficiency under the conditions of the extraction. Additionally, β-catenin binding partners could be masking the antibody epitope and preventing the binding of the antibody to β-catenin. As such, different antibodies, which bound to different β-catenin domains (**Table 2.6**), were tested using both CHAPS and CEB-A buffers. **Figure 3.6D** shows the results of blotting of these immunoprecipitations. Banding was observed at ~90kDa in each of the immunoprecipitations. The intensity of the banding remained low in all the immunoprecipitations when comparing them to the input, for both CHAPS and CEB-A buffers.

In addition, **Figure 3.7A** shows the results of combining two antibodies (with C and N terminal epitopes) into one immunoprecipitation. The banding observed was also very low. These results suggest that the antibody being used in the immunoprecipitation is not responsible for the low immunoprecipitation efficiency. As a result, the original β-catenin antibody (BD) was used for subsequent immunoprecipitations.

3.3.3.2 *Immunoprecipitation efficiency in buffers containing different CHAPS concentrations*

Another factor that may have led to a low immunoprecipitation efficiency was the concentration of CHAPS in the buffer. Detergents can alter the structure and function of proteins and therefore interfere with immunoprecipitation by reducing antibody-target binding and/or antibody-bead binding. To determine if the level of CHAPS in the buffers was interfering with the immunoprecipitation, K562 cells were BIO induced, fractionated with 0.4% or 0.8% CHAPS, immunoprecipitated with β-catenin antibody and western blotted (Chapter 0). **Figure 3.7A** shows the results of these immunoprecipitations following blotting of K562 cytoplasmic fractions with β-catenin antibody. Banding was observed at ~90kDa, which corresponded to the expected molecular weight of full length β-catenin. The

intensity of the banding was the same (very low) for the immunoprecipitations with 0.4% and 0.8% CHAPS. In addition, the intensity of the banding in the supernatants of both immunoprecipitations was similar to the input lane. This suggested that very little β-catenin had left the supernatant and bound to the agarose beads and that the level of CHAPS in the buffer was not responsible for the low immunoprecipitation efficiency.

3.3.3.3 *Immunoprecipitation efficiencies using different amounts of agarose beads.*

Another factor that could have led to the low immunoprecipitation efficiency was the volume of beads in the reaction. It was theorized that increasing the volume of beads in the immunoprecipitation reaction would increase the surface area for β-catenin-antibody complex binding, thereby increasing immunoprecipitation efficiency. To test this, K562 fractions were immunoprecipitated as before, but using a range of bead volumes (20-80μL). **Figure 3.7B** shows the results of western blotting of these immunoprecipitations with β-catenin antibody. Bands were visible at ~90kDa in the input and the immunoprecipitations for each bead volume (20μL, 40μL and 80μL). This corresponded to the expected banding pattern for full length β-catenin. The intensity of the band was similar for all the bead volumes, and very low compared to the input. This suggested that the immunoprecipitation efficiency was low, regardless of the volume of beads in the reaction.

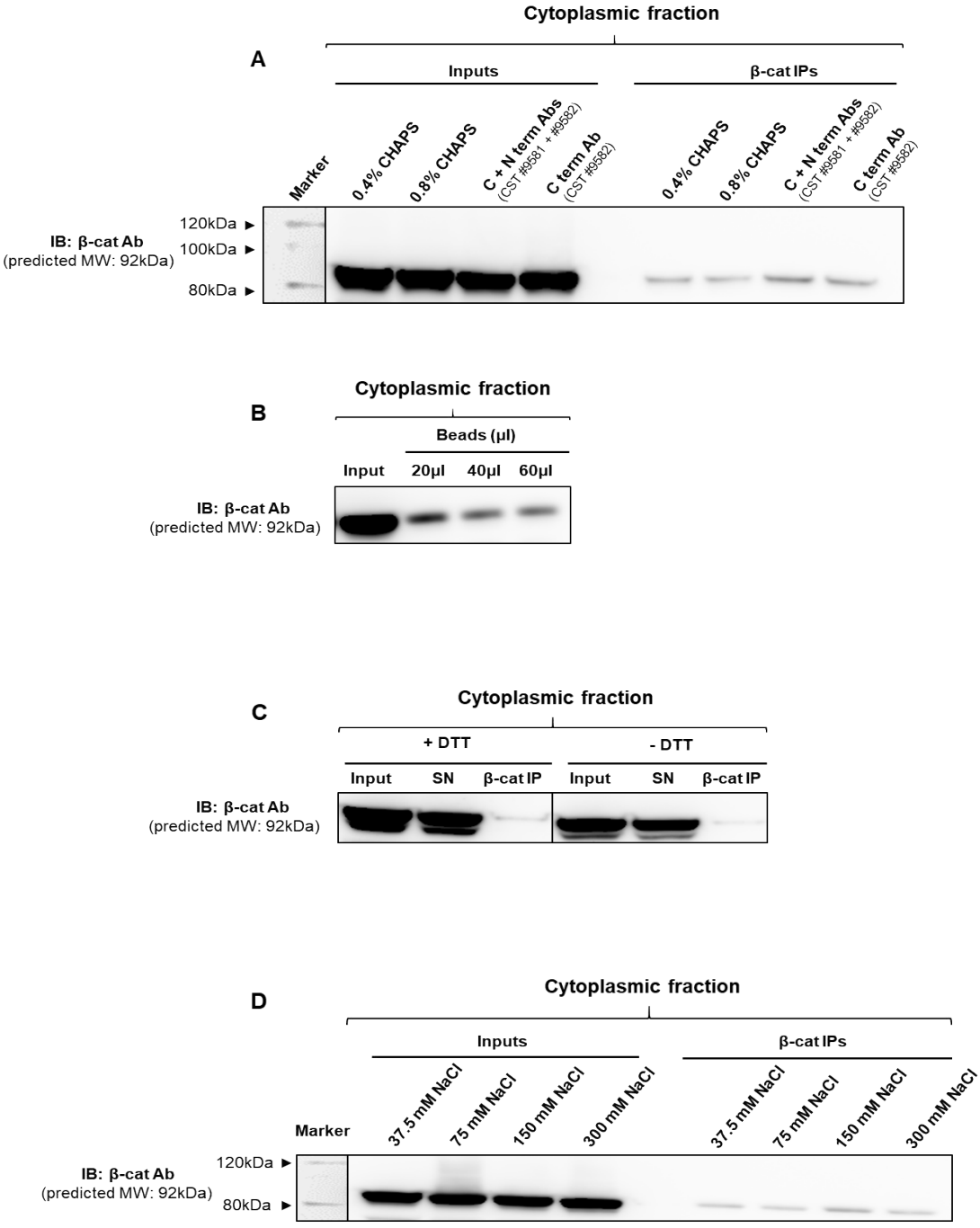


Figure 3.7 β-catenin immunoprecipitation efficiency under different buffer conditions.

(**A**) β-catenin immunoprecipitation in K562 cytoplasmic fractions using buffers containing 0.4% CHAPS, 0.8% CHAPS and containing 2 different β-catenin antibodies with N and C terminal epitopes (CST #9581 and CST #9582) (**Table 2.6**). (**B**) β-catenin immunoprecipitation in K562 cytoplasmic fractions using different amounts of beads. The marker was not visible at the relevant MW, but the relative migration was consistent with the signal being due to β-catenin. (**C**) β-catenin immunoprecipitation in K562 cytoplasmic fractions in the presence of DTT (+DTT) and in the absence of DTT (-DTT). The marker was not visible at the relevant MW, but the relative migration was consistent with the signal being due to β-catenin. (**D**) β-catenin immunoprecipitation in K562 cytoplasmic fractions in buffers containing different salt concentrations. Abbreviations; IB= immunoblot, Ab= antibody, MW= molecular weight, kDa= kilodalton, IP= Immunoprecipitation, SN= supernatant (immunodepleted lysate following the IP reaction). Marker = MagicMark XP (Thermo Fisher).

3.3.3.4 *Immunoprecipitation efficiencies when altering DTT and salt levels in the buffer*

Other components in the buffer could have caused the low efficiency of immunoprecipitation. Firstly, DTT (Dithiothreitol) can damage the structure of proteins, therefore reducing the amount of target-antibody-bead complexes formed. To determine if DTT was causing the low immunoprecipitation efficiency observed in my experiments, K562 cells were BIO induced and fractionated with and without DTT. Cytoplasmic fractions were then immunoprecipitated and western blotted with β-catenin antibody. **Figure 3.7C** shows the results of this western blot. Banding was observed at ~90kDa in both inputs and supernatants (-/+ DTT). In contrast, only very faint banding was observed in the immunoprecipitation in the presence of DTT, and no band could be seen in the immunoprecipitation without DTT in the buffer. This suggested that only a very small amount of β-catenin was immunoprecipitated both in the presence and absence of DTT, suggesting that DTT was not responsible for the low β-catenin immunoprecipitation efficiency.

Another component in the buffer that could be leading to low immunoprecipitation efficiency was the salt level. Too much salt can interfere with protein binding and thereby affect binding of antibody to beads or target protein. To test if the salt (NaCl) concentration in the CHAPS buffer used for fractionation was impacting the efficiency of immunoprecipitation, a range of salt concentrations were compared. K562 cells were BIO induced and fractionated with buffers containing a range of NaCl concentrations (37.5mM, 75mM, 150mM and 300mM). Cytoplasmic fractions were immunoprecipitated and western blotted with β-catenin antibody. **Figure 3.7D** shows the results of this western blot. Banding was observed at ~90kDa in all the supernatants and immunoprecipitations. This corresponded to full length β-catenin. The intensity of the banding was similar for all the immunoprecipitations (very faint). This suggested that the immunoprecipitation efficiencies were not impacted by the salt level in the buffer.

Taken together, the results of this section suggest that individual components of the fractionation buffer were not responsible for the low β-catenin immunoprecipitation efficiency and that there was a major limiting factor in the

protocol, either because of the reagents or the method used. Subsequently, a collaboration was established with Dr. Alex Greenhough (Bristol, UK). In their laboratory, they had successfully immunoprecipitated β-catenin and its binding partners in colon cancer cell lines.

3.3.3.5 *Optimization of immunoprecipitation by cross-linking the antibodies and beads*

Upon comparing my protocol with the one used by Dr. Greenhough's laboratory, a few key differences were observed which could be leading to the low β-catenin immunoprecipitation efficiency in my experiments. Firstly, in Dr. Greenhough's protocol, there was a step which cross-linked the antibody and beads prior to immunoprecipitation. This is an important difference because cross-linking strengthens the binding of antibodies to the beads and reduces the likelihood that antibody/protein complexes are lost during the washing steps, which could lead to a low immunoprecipitation efficiency. They were also using different buffers and different beads. Each of these factors was tested to determine if I could increase immunoprecipitation efficiency.

Firstly, to test if cross-linking would increase the immunoprecipitation efficiency, cells were BIO induced and cytoplasmic fractions were immunoprecipitated as before (2.6.3). However, β-catenin antibody was cross-linked to the agarose beads prior to immunoprecipitation. K562 and the colon adenocarcinoma cell line LS174T (positive control) cells were included in the experiment and the immunoprecipitations were western blotted using β-catenin antibody. **Figure 3.8A** shows the results of this western blot. A band was visible at ~90kDa in the input and supernatant of both K562 and LS174T cells. This corresponded to the expected weight of full length β-catenin. A band was not visible in the immunoprecipitations without cross-linking. This suggested that in both K562 and LS174T cells, only a very small amount (if any) of β-catenin was immunoprecipitated. A very faint band was observed in the immunoprecipitations with cross-linking. This suggested that there was an increase in immunoprecipitation efficiency in both K562 and LS174T cells after cross-linking of β-catenin to the agarose beads. This increase in efficiency was only modest,

however, suggesting that the lack of cross-linking was not responsible for the low immunoprecipitation efficiencies observed so far.

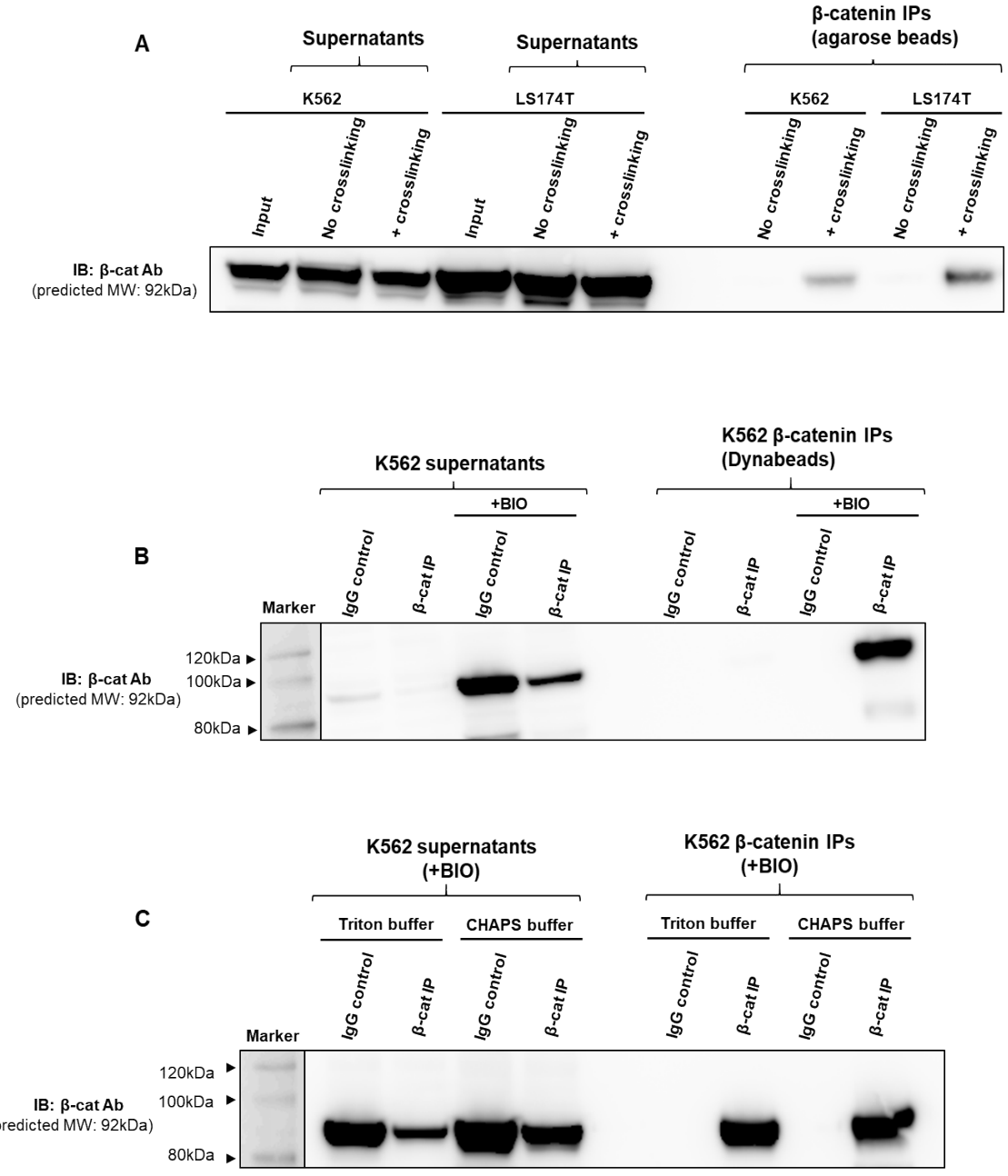


Figure 3.8 Immunoprecipitation of β-catenin using two methods in K562 and LS174T cell lines.

(**A**) Immunoprecipitation of β-catenin in K562 and LS174T cells using 0.8% CHAPS buffer and protein G agarose beads with and without cross-linking. The western blot marker was not visible at the relevant MW, but the relative migration was consistent with the signal being due to β-catenin. (**B**) β-catenin immunoprecipitation in K562 whole cell lysates (-/+ BIO) using Triton buffer (Appendix 3). The higher than expected banding pattern observed in the K562 Dynabeads β-catenin IP is due to slower sample migration on one side of the gel. (**C**) Repeat of the immunoprecipitation in K562 lysates prepared using Triton and CHAPS buffers. Abbreviations; IB= immunoblot, Ab= antibody, MW= molecular weight, SN=supernatant

(immunodepleted lysate following the IP reaction), IP= immunoprecipitated protein sample. Marker = MagicMark XP (Thermo Fisher).

3.3.4 *Immunoprecipitation of β-catenin using protein G Dynabeads*

Due to the failure to increase β-catenin immunoprecipitation using agarose beads, the decision was made to abandon this method in favour of the method from Dr.Greenhough's laboratory (2.6.4). To test this method in my cell line, K562 cells were BIO induced, lysed with Triton buffer and immunoprecipitated using protein G Dynabeads. Immunoprecipitations were then western blotted using β-catenin antibody. **Figure 3.8B** shows the results of this western blot. A band was visible at ~90kDa in the supernatant of K562 cells following induction with BIO, corresponding to full length β-catenin. The intensity of this band was lower in the immunoprecipitation supernatant, compared to IgG control. There was also intense banding visible in the immunoprecipitation of BIO induced K562 cells. Taken with the reduction of banding in the supernatant this suggested that β-catenin was efficiently immunoprecipitated using this new method in K562 whole cell lysates. This indicated that β-catenin could be immunoprecipitated in K562 cells and that BIO did not interfere with the reaction. Next, this method needed testing in K562 fractions (using CHAPS buffer).

K562 cells were BIO induced, fractionated using 0.8% CHAPS and cytoplasmic fractions were immunoprecipitated using β-catenin antibody. K562 whole cell lysate (lysed with Triton buffer) was also included as a positive control. **Figure 3.8C** shows the results of these immunoprecipitations following blotting with β-catenin antibody. A band was visible at ~90kDa in the supernatants of the K562 cytoplasmic fraction and whole cell lysate. This corresponded to full length β-catenin. The intensity of the band was lower in the immunoprecipitation supernatants compared to IgG control. An intense band was also visible in the immunoprecipitations from the cytoplasmic fraction and whole cell lysate. This band was absent in the IgG controls. This suggested that β-catenin was successfully immunoprecipitated at high efficiency using both buffers.

To conclude, attempts to overexpress β-catenin constructs in leukaemia cell lines were unsuccessful. C/N fractionation was successfully achieved using CHAPS

buffer in K562 cells, and a 0.8% CHAPS buffer was chosen for use in subsequent fractionations. Immunoprecipitation of β-catenin was unsuccessful using protein G agarose beads, despite attempts to increase immunoprecipitation efficiency by using different antibodies, different antibody concentrations, different CHAPs levels and different bead amounts, as well as altering the levels of DTT and salt in the buffer and cross-linking antibodies to the beads prior to immunoprecipitation. Immunoprecipitation was successful using an alternative method and protein G Dynabeads. This method will be used for subsequent immunoprecipitations (Chapter 4).

3.4 Discussion

The aim of the experiments presented in this chapter was to optimize cytoplasmic/nuclear fractionation and immunoprecipitation techniques for use in the MS analysis approach summarized in (**Figure 3.1**). This section discusses the approach taken to optimize these techniques as well as some of the issues that were encountered throughout this process.

3.4.1 *β-catenin overexpression was unsuccessful in leukaemia cell lines*

To isolate β-catenin binding partners in leukaemia cell lines, I attempted to overexpress β-catenin, this was done to increase the total amount of β-catenin available for immunoprecipitation. A FLAG-tagged β-catenin construct was chosen for use in pulldown experiments. The reason for this is that, firstly, antibodies raised against epitope tags have a low reactivity with other cellular proteins compared to those raised against endogenous proteins, and can be eluted from IP beads gently using short peptides rather than harsh methods such as boiling in SDS. This should result in less background binding proteins in MS analysis (Gingras *et al*, 2005). Additionally, an anti-FLAG antibody was readily available (clone M2, Sigma), and the small size of the FLAG protein decreased the risk of it interfering with β-catenin protein function. Test digests were performed to verify that the construct bought from Addgene (PCDNA3-β-catenin-FLAG) contained the correct construct. Sequencing could have been used to verify that the base level sequence was correct, however, this was done for the

subcloned pHIV-EGFP-β-catenin-FLAG vector, which verified that the β-catenin-FLAG construct was the correct sequence. When overexpression of this construct was unsuccessful in K562 cells, attempts to overexpress wild type β-catenin were made instead in case the addition of FLAG was destabilising the expression of β-catenin. Again, this was unsuccessful in K562 cells.

Despite lack of protein expression, *β-catenin-FLAG* mRNA appeared to be successfully expressed, as highlighted by the detection of GFP in transduced cells (which was expressed from an internal ribosome entry site (IRES) in the same message as *β-catenin-FLAG* mRNA). An IRES initiates ribosome binding and mRNA translation in a cap independent manner. This allows two proteins to be translated from the same mRNA sequence, in this case *β-catenin-FLAG* and GFP. The *β-catenin-FLAG* sequence is transcribed before the GFP sequence, and therefore if incomplete transcription of *β-catenin-FLAG* occurs, no GFP transcription should occur, and subsequently no mRNA should be translated into GFP protein. This allows GFP expression to be used as an indicator of *β-catenin-FLAG* mRNA expression.

The expression of *β-catenin-FLAG* mRNA, but not *β-catenin-FLAG* protein was not entirely unexpected given that β-catenin is a protein whose expression is chiefly regulated at the post-transcriptional level by the catenin destruction complex (1.3.1.1) as evidenced by the strong increase in protein expression following treatment with Bio.

Further expression of β-catenin-FLAG was successful in HEK293T cells (**Figure 3.5D**), this suggests that the FLAG-β-catenin protein can be expressed in other contexts. Given that HEK293T cells showed relatively good endogenous expression of β-catenin, it could be the case that overexpression was observed in this case because the destruction complex is less active in these cells.

3.4.2 *Using CHAPS for cytoplasmic/nuclear fractionation*

CHAPs was chosen for use in cytoplasmic/nuclear fractionation because it had been shown to maintain protein-protein interactions in other contexts, for example, maintenance of the raptor-mTOR complex in HEK293T cells and fibroblasts, when other detergents such as Triton-X-100 and NP-40 led to dissociation of this complex (Kim *et al*, 2002; Rosner & Hengstschläger, 2012). This was an important consideration for my experimental approach, whose success relied on identification of β-catenin interaction partners.

Being a mild buffer, CHAPS is usually used to isolate cytoplasmic fractions, or at higher concentrations, whole cell lysates (Mancini *et al*, 2017; Rosner & Hengstschläger, 2008; Shaiken & Opekun, 2014). Since different detergents can influence the binding partners that are identified following immunoprecipitation and MS (Arachea *et al,* 2012), I wanted to use CHAPS in the nuclear fractionation buffer as well. To isolate nuclear fractions, freeze thawing of isolated nuclear pellets was used prior to lysis with CHAPS buffer. This technique has been used previously by another group (Rosner & Hengstschläger, 2008). Overall, this approach was designed to ensure that protein-protein interactions were maintained for the MS approach, however, it would have been more thorough to compare other detergents as well to determine if using them would have increased the efficiency of β-catenin immunoprecipitation. If this approach had been taken it would also be important to compare the levels of known binding partners following β-catenin immunoprecipitation using different detergent buffers, to ensure that protein-protein interactions were still maintained.

3.4.3 *Optimization of β-catenin immunoprecipitation.*

The first decision made for the starting immunoprecipitation protocol was the type of beads to use; agarose, sepharose or magnetic. Each of these has their own advantages and disadvantages, including ease of use, surface area and cost. One important consideration was the amount of non-specific binding for each bead type, which can differ in cytoplasmic and nuclear cell fractions. In one study comparing non-specific protein binding to sepharose, agarose and magnetic beads in HeLa cytoplasmic and nuclear fractions, non-specific binding of nuclear

proteins was shown to be lowest using magnetic beads, and a similar amount of nuclear proteins bound to agarose and sepharose beads. In contrast, for cytoplasmic fractions the lowest amount of non-specific binding was observed using agarose beads, followed by sepharose and then magnetic beads (Trinkle-Mulcahy *et al*, 2008). Since in this project cytoplasmic and nuclear proteins were being studied, the decision about which beads to use was based on other factors.

Agarose beads were ultimately chosen because they are widely used and inexpensive compared to other beads. In this study, β-catenin immunoprecipitation using agarose beads was ineffective which led to different conditions being optimized to successfully immunoprecipitate β-catenin.

3.4.3.1 *Antibody conditions*

One factor that could have caused the low immunoprecipitation efficiency was an insufficient concentration of antibody in the reaction. This is because if the antibody concentration is too low, not enough antibody binds to the agarose beads and there are less antibody-bead complexes for β-catenin to bind to. As such, the effect of using different antibody amounts in the immunoprecipitation was assessed. No significant differences in immunoprecipitation efficiency were observed by increasing the antibody concentration.

Another factor that can lead to the low immunoprecipitation efficiency is the antibody used. This is because different antibodies have different binding efficiencies and the low efficiency of β-catenin binding could be due to a low binding efficiency. Additionally, β-catenin binding partners could be masking the antibody epitope and preventing the binding of the antibody to β-catenin. The antibody used for initial immunoprecipitations (from BD) was chosen because it was a monoclonal antibody, which have less variability batch-to-batch compared with polyclonal antibodies and are less likely to bind non-specifically to other proteins because they only bind to one specific epitope, however, this is not always the case and the specificity of antibodies varies greatly from antibody to antibody and from supplier to supplier (Gilda *et al*, 2015; Vanli *et al*, 2017). This was an important consideration for the consistency and reliability of the immunoprecipitation results because having the lowest possible amount of

background binding was important for the MS approach. One disadvantage of using monoclonal antibodies for immunoprecipitation, is that they tend to have a much lower overall binding affinity for their targets than polyclonal antibodies because they only bind to one epitope rather than multiple epitopes[11], however, this is not always the case. In one study looking at the immunoprecipitation efficiencies of 11 antibodies for the estrogen receptor β (ERβ) in MCF-7 cells, a high degree of variability in immunoprecipitation efficiencies was observed, and this was not correlated with the antibodies being monoclonal or polyclonal (Weitsman *et al*, 2006).

As such, a range of monoclonal and polyclonal antibodies were selected for optimization in this project. When selecting these antibodies, their IgG isotype was taken into consideration because different antibody isotypes have different binding affinities for protein G, which coated the agarose beads in my experiment[12]. Only IgG isotypes that bind to protein G with medium or high affinity were selected (mouse IgG1 and rabbit IgG) (**Table 2.6**).

The antibody epitopes were also considered. Whilst many of the proteins that interact with β-catenin do so by binding to the armadillo repeats in the central part of the protein (**Figure 1.6**), other established binding partners interact with β-catenin's C terminus (e.g. the histone acetyltransferases CBP/p300 and TRRAP p400 and TIP60) and N terminus (e.g. the E3 ligase β-TrCP) (Valenta *et al*, 2012; Xu & Kimelman, 2007). As a result, I wanted to make sure that the ineffective immunoprecipitation was not due to binding partners masking the epitope for the BD antibody. When selecting antibodies for optimization, a selection of N terminal and C terminal binding antibodies were included. Comparisons of antibodies suggested that epitope screening was not a factor in IP efficiency of β-catenin in this study.

[11] http://www.abcam.com/protocols/a-comparison-between-polyclonal-and-monoclonal (accessed on 04/01/2018)

[12] https://www.neb.com/tools-and-resources/selection-charts/affinity-of-protein-ag-for-igg-types-from-different-species (accessed on 04/01/2018)

3.4.3.2 *Detergent concentrations*

The concentration of CHAPS in the buffer was another factor that was optimized to try and successfully immunoprecipitate β-catenin. Detergents can have different effects on proteins and protein interactions at different concentrations. For example, an experiment studying the effect of the NP-40 detergent on the recovery of alcohol dehydrogenase (ADH) in immunoprecipitation experiments determined that low levels of this detergent increased the efficiency of the immunoprecipitation compared to immunoprecipitations conducted with no detergent. This suggests that the detergent may change the conformation of the antibody and/or target protein in such a way that more effective target-antibody binding is observed. In contrast, higher levels of NP-40 decreased immunoprecipitation efficiency (Yang *et al*, 2009). The effect of detergent on antibody-antigen interactions is different for each detergent, with detergents like SDS destroying their interaction at relatively low levels (0.01%) and detergents like Triton-X-100 not having a profound effect on antibody-antigen binding up to concentrations of about 5% (Qualtiere *et al*, 1977)

Given the differences in antibody-antigen binding affinities for each protein and antibody, the effect of detergents on the efficiency of immunoprecipitation is not always predictable. To determine if the concentration of CHAPS was influencing the efficiency of my immunoprecipitation, different concentrations of CHAPS (0.4% and 0.8%) were compared.

In hindsight, this may not have been the best approach to take, as the effect of detergents on the efficiency of immunoprecipitations is not linear. In the study by Yang *et al*. 2009, the efficiency of immunoprecipitation decreased between 0.02% and 0.1% NP-40, and then the efficiency remained relatively constant at higher concentrations. As a result, it would have been a better approach to use a wider range of CHAPS concentrations to determine if it was influencing the efficiency of β-catenin immunoprecipitation in my experiments.

3.4.3.3 *DTT and salt concentration*

Another component in the buffer that can lead to low immunoprecipitation efficiency is the salt level. Salts can impact protein stability and impact the efficiency of immunoprecipitation by disrupting antibody-target binding and/or antibody bead binding. For example, the salt ammonium sulfate was found to decrease protein stability of the proteins catalase and a-lactalbumin at ~1M concentration, ribonuclease A at ~1.5M and bovine serum albumin at ~ 2M. In contrast, the stability of myoglobin decreased only slightly at 2.5M (Dumetz *et al*, 2007). This shows that different proteins are affected by the same salt, in this case ammonium sulfate, at different concentrations and determining the affect of salts on specific proteins is an important consideration when using a proteomics approach.

Different studies have suggested that NaCl concentration does not have a significant effect on protein stability (Dumetz *et al*, 2007; Lindman *et al*, 2006). In one study, NaCl concentration only marginally impacted the stability of protein G. In this study, the Tm (temperature at the denaturation midpoint) was used to determine protein stability and only a very small decrease in the Tm was observed when comparing low salt concentration to 150mM. This suggests that there was a small decrease in protein stability, but this was not significant (Lindman *et al*, 2006). Similarly, in the study by Dumtez *et al*. 2007 the effect of a range of NaCl (0-4M) concentrations on the second osmotic virial coefficient (b2) values of seven proteins was measured using self-interaction chromatography. b2 values are a measure of non-ideal solution behaviour resulting from interactions between two solutes (in this case proteins). A positive b2 value indicates repulsion between the two solutes, whereas a negative b2 value indicates attraction between the two solutes (Alford *et al*, 2008). For each of the proteins, the b2 values remained constant at all NaCl concentrations, suggesting that NaCl does not impact protein stability in this context. In addition, 150mM NaCl is a standard component of lysis buffers and immunoprecipitation buffers, and β-catenin and its binding partners have been co-immunoprecipitated successfully using 150mM containing buffer (Chen *et al*, 2012; Li *et al*, 2012; Palka-Hamblin *et al*, 2010).

Although this evidence suggests that NaCl does not impact the stability and interaction of proteins at the concentrations used in this study, the complexities of individual protein interactions in different contexts meant that I still wanted to rule this out as a factor leading to my low immunoprecipitation efficiency. In agreement with the literature outlined above, NaCl concentration did not affect β-catenin immunoprecipitation in this project.

Another factor that was considered was the effect of DTT on the efficiency of my immunoprecipitations. DTT targets disulphide bonds, which are important in maintaining protein structure and has been shown to impact the titration of IgM antibodies, however, DTT did not impact the titration of IgG antibodies under the same conditions (Okuno & Kondelis, 1978; Pirofsky & Rosner, 1974). Since I was using IgG antibodies in my study, this suggested that DTT may not be impacting the efficiency of my immunoprecipitations.

A different study found that DTT does cleave IgG, but only at a concentration of 2mM (Crivianu-Gaita *et al*, 2015). This is twice as high as the 1mM DTT used in my study. Despite this, DTT could not be ruled out as a factor that could be impacting the immunoprecipitation efficiency and was included in this analysis since it could potentially impact the structure of β-catenin, changing its confirmation and interfering with the binding of the antibody to its epitope.

To summarize, conditions for β-catenin extraction were optimised using buffers known to favour protein-protein interactions. Issues with IP efficiency led to the investigation of different variables affecting IP. An alternative protocol was adopted which used different beads, and this overcame the problems with poor IP efficiency. It was concluded that a defective batch of protein G beads was probably responsible for the poor immunoprecipitation efficiency.

4 Identification of candidate factors regulating the nuclear localization of β-catenin by MS

4.1 Introduction

β-catenin is dysregulated in AML and its overexpression is associated with a poor prognosis (Chen *et al,* 2009; Simon *et al*, 2005; Xu *et al*, 2008; Ysebaert *et al*, 2006). Key to its role as a transcription factor is its localization to the nucleus. Although β-catenin nuclear localization factors have been identified in other contexts (Morgan *et al*, 2014), little is known about the factors regulating β-catenin nuclear localization in AML. In this project an MS approach was conducted to identify candidate β-catenin nuclear localization factors in AML, using AML cell lines as a model system. In the previous chapter, β-catenin co-immunoprecipitation was optimized in leukaemia lines. In this chapter, this protocol was subsequently used to co-immunoprecipitate β-catenin and its binding partners in nuclear and cytoplasmic fractions of leukaemia cell lines. These samples were sent for MS analysis at Bristol Proteomics Facility and analysed to produce a candidate list of potential β-catenin nuclear localization factors. The MS approach taken to identify candidate β-catenin nuclear localization factors is introduced in 2.6 and summarized in **Figure 4.1**. Briefly, the presence/absence of β-catenin binding partners was compared in the nuclear and cytoplasmic fractions of translocator and non-translocator lines. Candidate β-catenin nuclear localization factors were identified based on the presence or absence of factors in each compartment, as shown in **Table 4.1**. These proteins were filtered and their potential role as regulators of β-catenin nuclear localization was validated.

4.2 Aims

- To use the optimized co-immunoprecipitation protocol to prepare protein samples for MS analysis;
- To identify candidate β-catenin nuclear import and export factors in leukaemia cell lines.

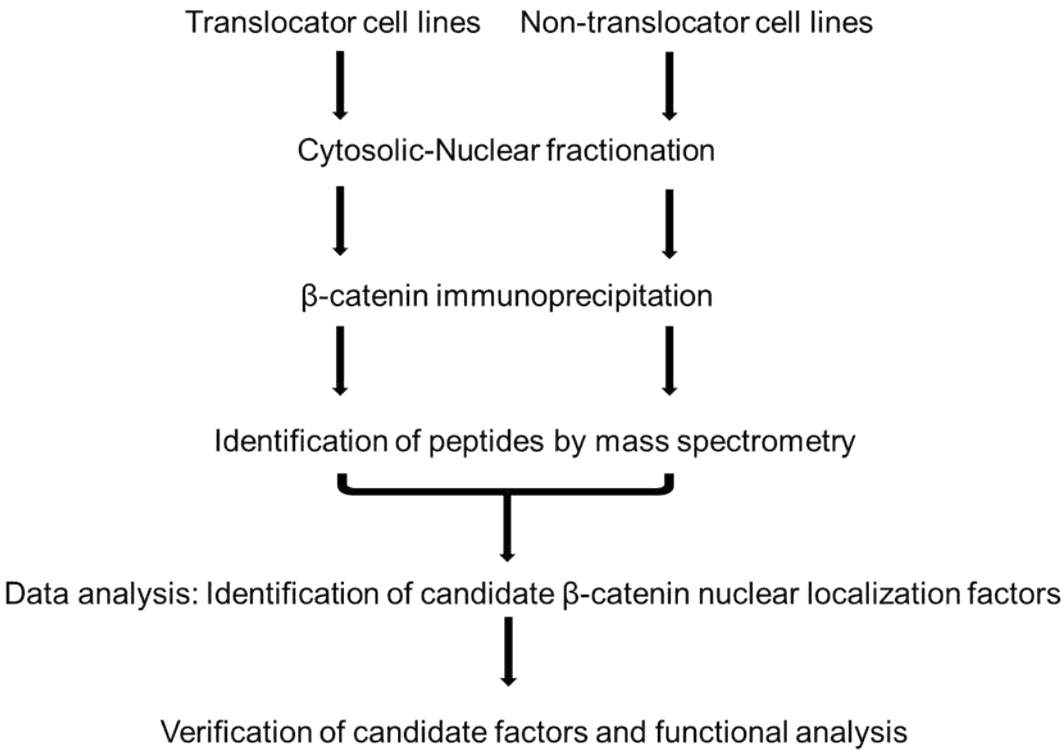


Figure 4.1. Schematic of the experimental approach used to identify candidate β-catenin nuclear localization factors in leukaemia cell lines.

This diagram is reproduced from 2.6 for convenience. It summarizes the steps taken to identify candidate β-catenin nuclear localization factors in leukaemia cell lines.

Table 4.1. Expected localization of β-catenin binding partners involved in its nuclear localization.

The mechanism of β-catenin nuclear localization and expected localization of proteins involved in the process are highlighted below. Dark green (checked pattern) = presence, light green= possible presence and yellow (diagonal lined pattern) = absence. C= Cytoplasm and N= Nucleus. Note; data for the nuclear fractions of non-translocator cell lines is not available.

Method	Translocator		Non-translocator	
	C	N	C	N
Nuclear retention				
Cytoplasmic-nuclear shuttling				
Nuclear-cytoplasmic shuttling				
Cytoplasmic retention				

4.3 Results

4.3.1 *Co-immunoprecipitation of β-catenin and its binding partners for MS analysis*

4.3.1.1 *Assessment of the purity of leukaemia cell line cytoplasmic and nuclear fractions*

To identify candidate β-catenin nuclear localization factors, leukaemia cell lines were fractionated, followed by co-immunoprecipitation of β-catenin and its binding partners (2.6). Prior to β-catenin co-immunoprecipitation, cytosolic and nuclear fractions were immunoblotted with GAPDH and Histone H1 antibodies (2.5) to verify the purity/enrichment of the fractions. This was important because the identification of candidate β-catenin nuclear localization factors was dependent on the presence and absence of proteins in the cell compartments (**Figure 4.1**) and contamination between the fractions could confound the conclusions drawn from the MS analysis.

Figure 4.2A outlines the purity of the cell fractions. The blot was probed with both GAPDH and Histone H1 antibodies, with the expectation that the different molecular weights of these proteins would clearly discriminate the signal from each antibody. In the cytoplasmic fractions of all the lines, a band was visible at approximately 37kDa, corresponding to the expected weight of GADPH. Additionally, there was a cluster of weak bands in the nuclear fractions at the same molecular weight. This was likely due to unexpected signal from the histone H1 antibody at a high molecular weight rather than GAPDH signal.

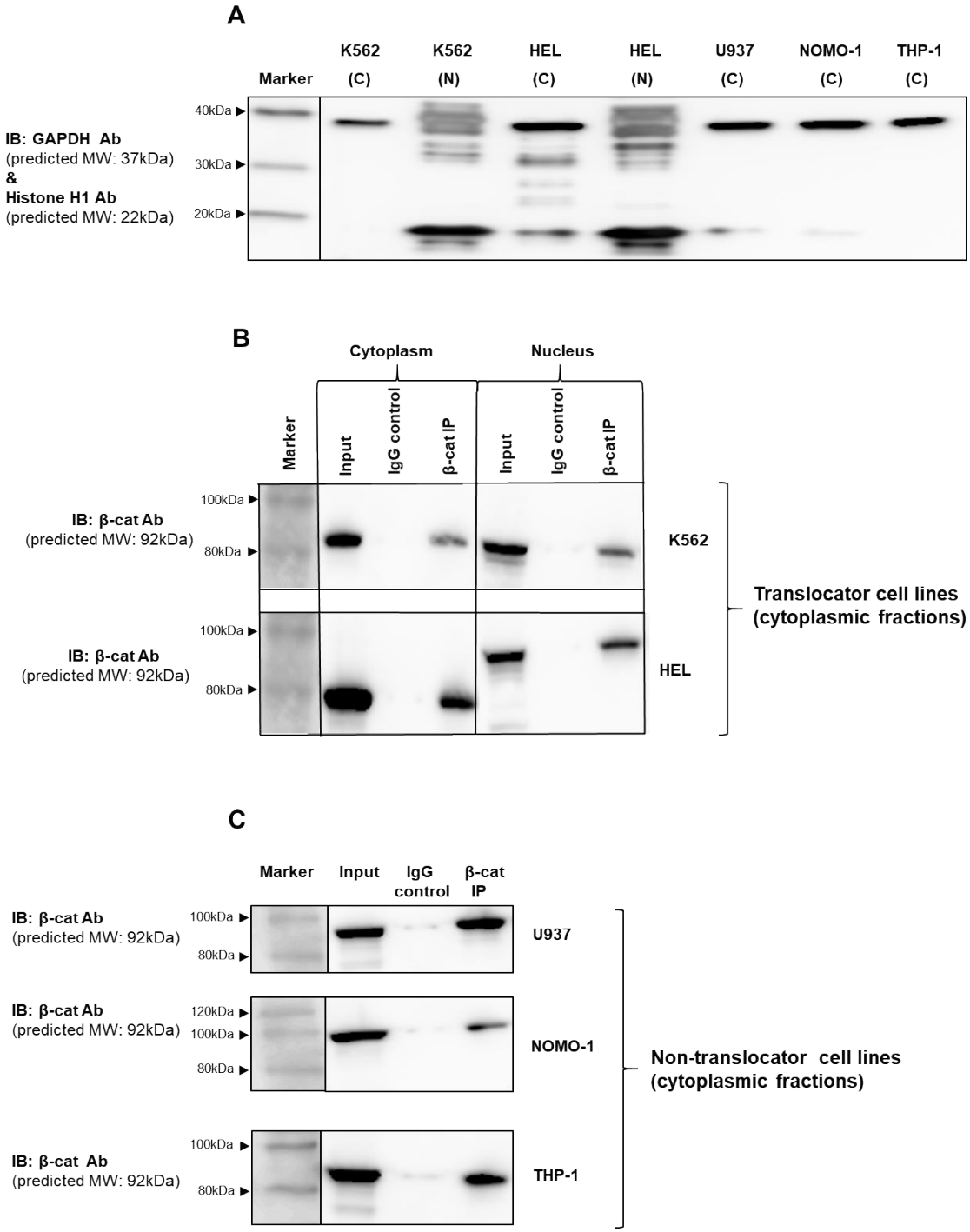


Figure 4.2. Co-immunoprecipitation of β-catenin in leukaemia cell lines for MS analysis.

(**A**) Inputs for each cell line and fraction were immunoblotted with GAPDH and Histone H1 to determine the purity of the fractions prior to co-immunoprecipitation (2.5). (**B**) Western blotting of samples following co-immunoprecipitation with β-catenin antibody in translocator lines. The higher than expected banding pattern observed in the HEL nuclear fractions is due to slower sample migration on one side of the gel. (**C**) Western blotting of samples following co-immunoprecipitation with β-catenin antibody in non-translocator lines. An IgG control co-immunoprecipitation was also analysed for each cell line to allow removal of background binding proteins (non-specific binding). Abbreviations; β-cat=β-catenin, C= cytosolic fraction, N= nuclear fraction, IB= immunoblot, Ab= antibody, MW= molecular weight, kDa= kilodalton. Marker= MagicMark XP (Thermo Fisher).

In the nuclear fractions, two bands were visible just below the 20kDa marker, corresponding to Histone H1. As expected it was absent in the cytoplasmic fractions, except for HEL cells, where there was a small amount of nuclear contamination. Since the level of β-catenin in the nucleus of these cell lines is minimal compared to the amount in the cytoplasm (**Figure 1.7**), this was unlikely to impact the MS results. Therefore, these samples were considered 'pure' enough to be used for subsequent co-immunoprecipitation and MS analysis.

4.3.1.2 *Determination of non-specific IgG-β-catenin binding and successful β-catenin co-immunoprecipitation*

Co-immunoprecipitation of β-catenin and its binding partners was conducted following the method optimized in Chapter 3 (2.6). Prior to sending the samples for MS, non-specific binding of β-catenin to IgG was assessed. IgG control and β-catenin co-immunoprecipitations were immunoblotted with β-catenin antibody (**Figure 4.2B&C**). There was only very faint β-catenin signal in the IgG lanes of some of the samples. This suggested that non-specific binding of β-catenin to IgG was unlikely to be an issue and I could confidently remove background binding partners from the analysis of MS data without removing β-catenin binding partners.

In contrast, β-catenin signal was detected in all the β-catenin co-immunoprecipitation samples. This confirmed that β-catenin co-immunoprecipitation was successful, and the samples could be analysed using MS at Bristol Proteomics Facility. In Bristol, the samples were analysed by tandem MS using an LTQ- Orbitrap Velos mass spectrometer (2.6.6).

4.3.2 ***Filtering of raw data to remove background proteins***

Details of the methodology used for MS are outlined in (2.6.6). Briefly, samples were electrophoresed on an SDS-PAGE gel and each lane was cut into three equal slices. Each slice was subject to in gel tryptic digestion and fractionated prior to MS. The raw data files were processed using Proteome Discoverer software (Thermo Scientific) and false discovery rates (FDR) of 1% and 5% were

used to produce two MsExcel files of identified peptides (2.6.6) (Appendix 4 - see the attached CD disk).

These data were filtered to remove proteins that non-specifically bound to IgG (**Figure 4.3A**). Peptides present in all IgG lanes were filtered out. This removed proteins from the list that were only non-specific IgG binding proteins.

The next step was to take each cell line at a time and remove background binding proteins specific to that cell line. To do this Excel filtering was used on the Area columns for IgG and β-catenin co-immunoprecipitation. The IgG column was filtered based on “Equals 0” and the co-immunoprecipitation column was filtered based on “Does Not Equal 0”. Next, proteins with a peptide score of less than 3 were removed by filtering the Peptide column based on “Greater than 2”.

The remaining proteins were used to identify candidate nuclear β-catenin localization factors.

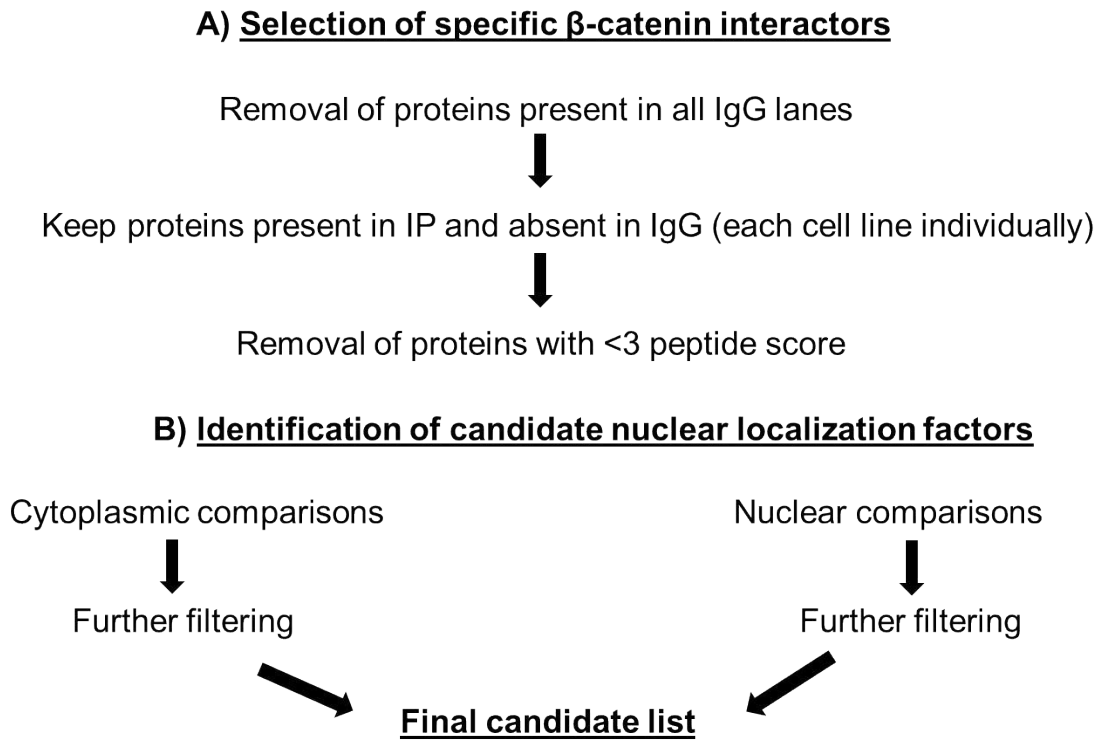


Figure 4.3. Schematic of the data analysis approach used to identify candidate β-catenin nuclear localization factors.

This is an overview of the analysis strategy discussed in the text. Data received from Bristol Proteomics Facility was analysed following the steps above. (**A**) Removal of non-specific binding partners (or background proteins). (**B**) Identification of candidate factors regulating the nuclear localization of β-catenin.

4.3.3 *Identification of candidate nuclear β-catenin exporters*

To identify candidate β-catenin exporters, β-catenin binding partners from translocator and non-translocator cytoplasmic fractions were compared. This approach was based on the expected localization of nuclear export proteins in the cytoplasm and nucleus of non-translocator lines, outlined in **Table 4.1.**

Briefly, proteins present in the cytoplasmic fraction of non-translocator cell lines and absent in the cytoplasmic fraction of translocator cell lines were identified as potential β-catenin exporters. This approach was based on the presumption that the presence of β-catenin exporters in leukaemia cell lines is sufficient for nuclear β-catenin export. Comparisons of peptides between different cell lines and cell compartments were conducted using an online Venn diagram creator[13]. This produced Venn diagrams comparing peptide presence based on their accession numbers.

Firstly, peptides present in the cytoplasmic fraction of all non-translocator cell lines were identified (**Figure 4.4A**). 52 proteins from NOMO-1 cells, 22 proteins from U937 cells and 15 proteins from THP-1 cells were compared. Out of these, 3 proteins were common to all non-translocator lines; α-catenin 1, α-catenin 2, and APC. This represents 6% of NOMO-1 cytoplasmic proteins, 14% of U937 cytoplasmic proteins and 20% of THP-1 cytoplasmic proteins.

Next, peptides present in the cytoplasmic fraction of both translocator lines were identified (**Figure 4.4B**). 45 peptides from K562 cells and 142 peptides from HEL cells were compared. Out of these, 24 peptides were present in both translocator cell lines. This represents 53% of K562 cytoplasmic proteins and 17% of HEL cytoplasmic proteins.

Finally, the three non-translocator cytoplasmic proteins and the 24 translocator cytoplasmic proteins were compared (**Figure 4.4C**). No peptides were present in the cytoplasm of non-translocator lines and absent in the cytoplasm of translocator lines. This means that no candidate β-catenin nuclear export factors were identified in this project.

[13] http://bioinformatics.psb.ugent.be/webtools/Venn/ (accessed on 04/01/2018)

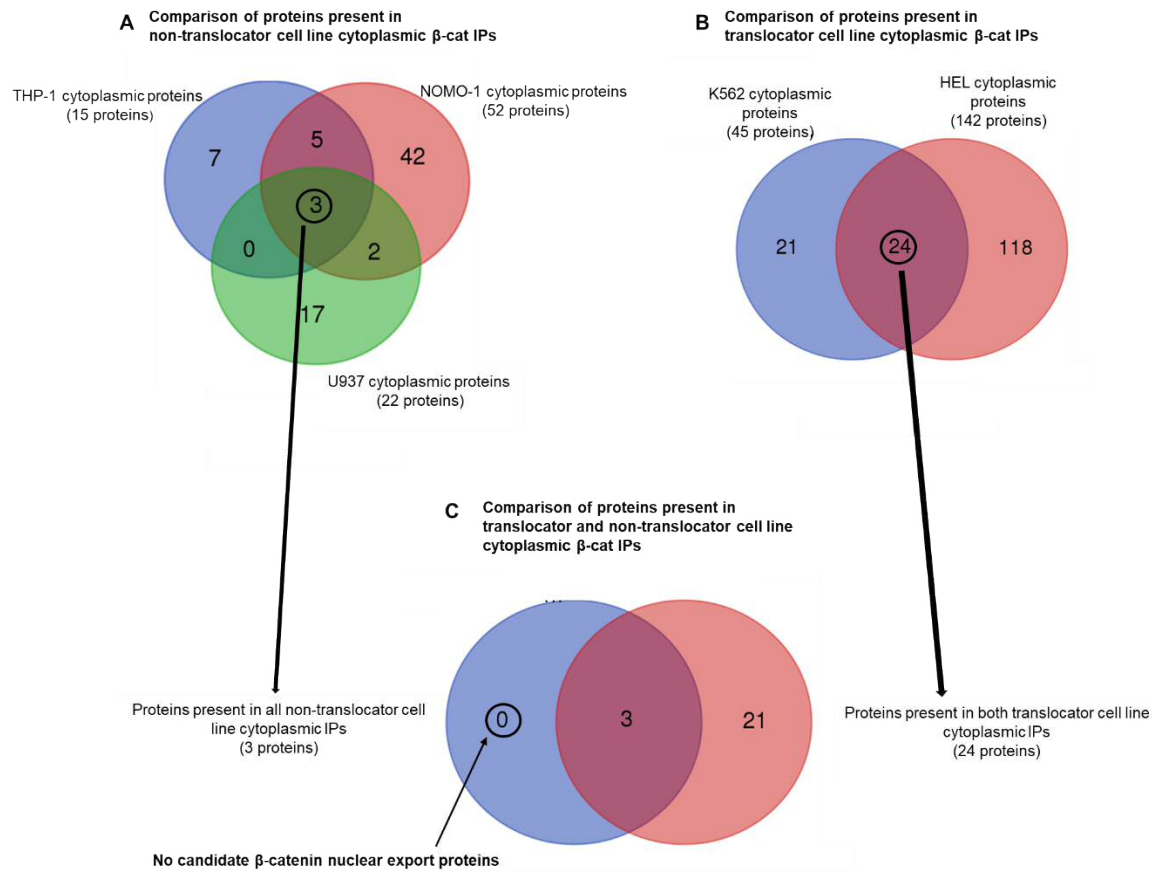


Figure 4.4. Identification of candidate nuclear β-catenin exporter proteins following removal of background binding partners.

(**A**) Venn diagrams were created to compare proteins found in the cytoplasm of non-translocator cell lines. (**B**) Venn diagrams were created to compare proteins found in the cytoplasm of translocator cell lines. (**C**) The 3 proteins identified in the cytoplasm of all non-translocator cell lines and the 24 proteins identified in the cytoplasm of all translocator cell lines were compared. Abbreviation; IP= immunoprecipitation

4.3.4 *Identification of candidate nuclear β-catenin import and retention factors*

To identify candidate β-catenin importers and nuclear retention factors, nuclear β-catenin binding partners were compared. This approach was based on the expected localization of nuclear importers to the nucleus of translocator cell lines only (**Table 4.1**). Owing to the absence of β-catenin in the nuclear fraction of non-translocator lines, peptides present in the nucleus of translocator and non-translocator cell lines could not be compared. Instead, peptides present in the β-catenin co-immunoprecipitations of both translocator nuclear fractions were viewed as potential candidates. To investigate if the candidates were potential cytoplasmic-nuclear shuttling proteins (rather than retention factors) the localization of these peptides in the cytoplasm of translocator lines was compared. It would be expected that candidate cytoplasmic-nuclear shuttling proteins would be predominantly present in the cytoplasmic co-immunoprecipitations, whereas nuclear retention factors would be predominately in the nucleus (**Table 4.1**). Comparisons of peptides between different cell lines and cell compartments were conducted using an online Venn diagram creator[14]. This produced Venn diagrams comparing peptide presence based on their accession numbers.

Firstly, peptides present in the nuclear fraction of both translocator cell lines were identified (**Figure 4.5A**). 107 proteins from HEL (**Table 4.2**) cells and 53 proteins from K562 cells (**Table 4.3**) were compared. Out of these, 27 proteins were present in both translocator cell lines (**Table 4.4**) and were identified as potential candidate β-catenin nuclear localization factors.

Next, to identify which of these factors were potential cytoplasmic-nuclear shuttling proteins, these 27 proteins were compared to those present in the cytoplasmic fractions of translocator cell lines (**Figure 4.5B**). 7 peptides were present in the cytoplasmic and nuclear fractions of both cell lines and were potential cytoplasmic-nuclear shuttling factors. 11 peptides were present in the nucleus of both translocator cell lines and the cytoplasm of HEL cells only. Due

[14] http://bioinformatics.psb.ugent.be/webtools/Venn/ (accessed on 04/01/2018)

to the presence of nuclear contamination in the cytoplasmic fraction of HEL (**Figure 4.2**), for further analysis these proteins were taken as candidate nuclear retention factors, along with the 8 peptides present in the nucleus of both translocator cell lines only.

In summary, 27 candidate β-catenin nuclear localization factors were identified, of which 7 were highlighted as potential cytoplasmic-nuclear shuttling proteins and 19 factors were identified as potential nuclear retention factors.

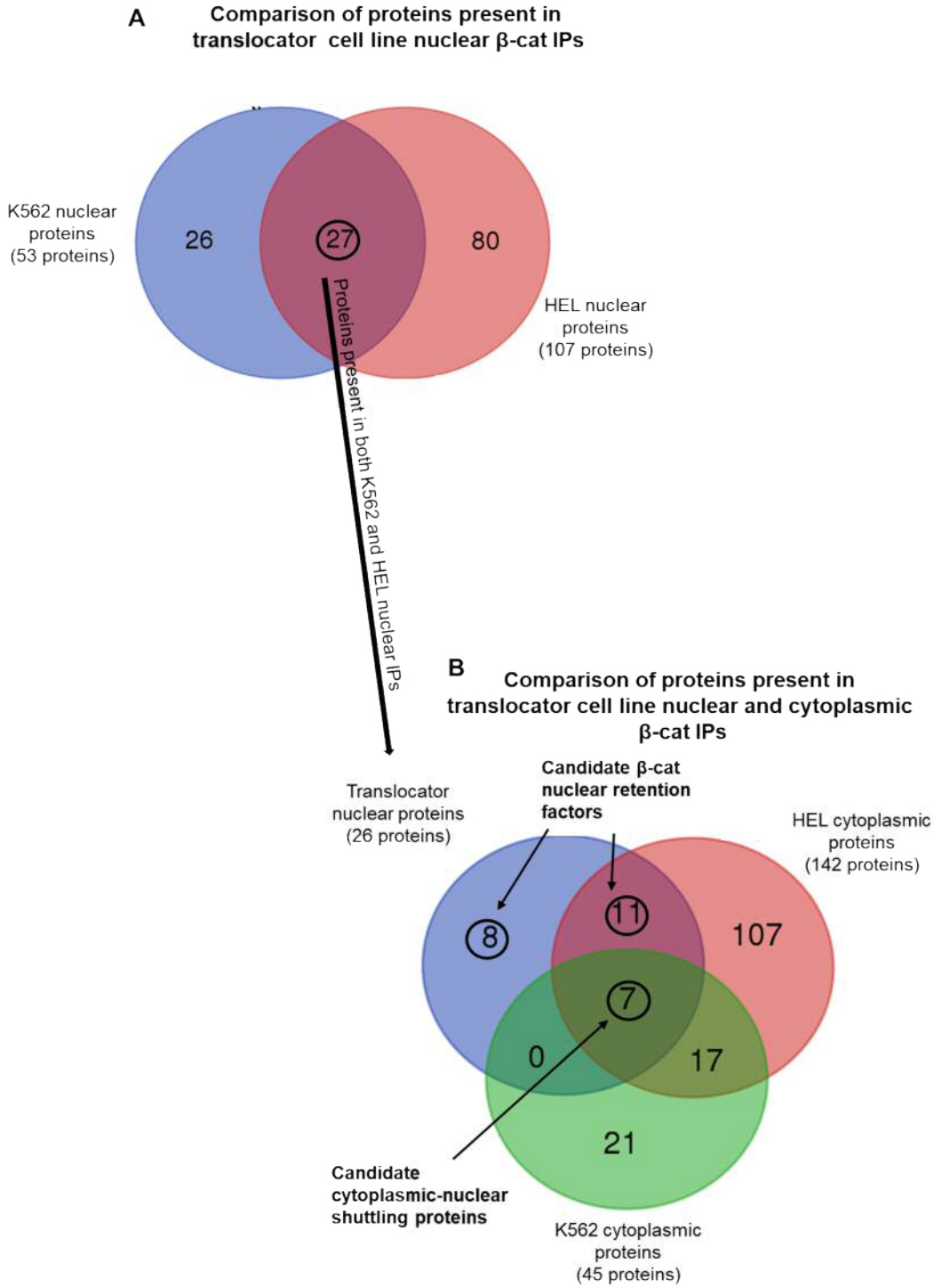


Figure 4.5. Identification of candidate nuclear β-catenin importers and retention factors, following removal of background binding partners.

(**A**) Venn diagram comparing proteins found in the nucleus of translocator lines. (**B**) The 27 proteins identified in (A) were compared to proteins present in the cytoplasm of translocator lines. The discrepancies in the total number of translocator nuclear proteins (27 in (A) and 26 in (B) are due to the removal of the α-catenin 2 protein from the list prior to analysis. This was done due to its detection in the cytoplasm of non-translocator cell lines in earlier analysis. Abbreviation; IP= immunoprecipitation..

Table 4.2. β-catenin binding partners in the HEL nuclear fraction prior analysis.

A total of 107 proteins were present in the nuclear β-catenin IP and absent from the nuclear IgG control sample in HEL cells. These proteins are listed below and were used for the comparisons made in **Figure 4.5**. Proteins with HUGO gene names of N/A could not be found in the HUGO database based on their accession number or protein name.

Accession number	HUGO gene name (symbol)[15]	Protein name
P07437	Tubulin beta class I (TUBB)	Tubulin beta chain
P68371	Tubulin beta 4B class IVb (TUBB4B)	Tubulin beta-4B chain
P84095	Ras homolog family member G (RHOG)	Rho-related GTP-binding protein RhoG
Q59H57	N/A	Fusion (Involved in t(12;16) in malignant liposarcoma) isoform a variant (Fragment)
Q9BTQ7	N/A	Similar to ribosomal protein L23 (Fragment)
B2R6L0	N/A	cDNA, FLJ93005, highly similar to Homo sapiens tubulin, beta polypeptide (TUBB), Mrna
B3KQ75	N/A	cDNA FLJ33018 fis, clone THYMU1000459, highly similar to Homo sapiens transcription factor 7 (T-cell specific, HMG-box) (TCF7), transcript variant 5, mRNA
H0Y2W2	ATPase family, AAA domain containing 3A (ATAD3A)	ATPase family AAA domain-containing protein 3A (Fragment)
Q9UJU2	Lymphoid enhancer binding factor 1 (LEF1)	Lymphoid enhancer-binding factor 1
Q2TAM6	Runt related transcription factor 1 (RUNX1)	RUNX1 protein (Fragment)
P42167	Thymopoietin (TMPO)	Lamina-associated polypeptide 2, isoforms beta/gamma OS
F5GXX5	Defender against cell death 1 (DAD1)	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit DAD1
Q96T67	ATPase family, AAA domain containing 3B (ATAD3B)	TOB3
E2GH18	Transcription factor 7 like 2 (TCF7L2)	T-cell factor-4 variant C
P05141	Solute carrier family 25 member 5 (SLC25A5)	ADP/ATP translocase 2
P12236	Solute carrier family 25 member 6 (SLC25A6)	ADP/ATP translocase 3

[15] https://www.genenames.org/ (accessed on 26/03/2018)

Accession number	HUGO gene name (symbol)[15]	Protein name
O75955	Flotillin 1 (FLOT1)	Flotillin-1
P36402	Transcription factor 7 (TCF7)	Transcription factor 7
P25705	ATP synthase F1 subunit alpha (ATP5F1A)	ATP synthase subunit alpha, mitochondrial
O95167	NADH:ubiquinone oxidoreductase subunit A3 (NDUFA3)	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 3
B7Z597	N/A	cDNA FLJ54373, highly similar to 60 kDa heat shock protein, mitochondrial
O00161	Synaptosome associated protein 23 (SNAP23)	Synaptosomal-associated protein 23
B4E2V5	N/A	cDNA FLJ52062, highly similar to Erythrocyte band 7 integral membrane protein
P42166	Thymopoietin (TMPO)	Lamina-associated polypeptide 2, isoform alpha
C6ZRK5	Transcription factor 7 like 2 (TCF7L2)	TCF7L2 isoform pFC8A_TCF7L2_ex1-11-13-14
A0A0D9SGH8	Transcription factor 7 like 2 (TCF7L2)	Transcription factor 7-like 2
O95298	NADH:ubiquinone oxidoreductase subunit C2 (NDUFC2)	NADH dehydrogenase [ubiquinone] 1 subunit C2
Q8N3E6	N/A	Putative uncharacterized protein DKFZp761L1023 (Fragment)
Q9P0J0	NADH:ubiquinone oxidoreductase subunit A13 (NDUFA13)	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 13
B4DLA6	N/A	cDNA FLJ54365, highly similar to DNA replication licensing factor MCM4
B4DQY2	N/A	cDNA FLJ59388, highly similar to Mitochondrial inner membrane protein
Q96PK6	RNA binding motif protein 14 (RBM14)	RNA-binding protein 14
A0A0C4DGS1	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase non-catalytic subunit (DDOST)	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit
B4DKE6	N/A	cDNA FLJ60629, highly similar to Replication factor C subunit 3

Accession number	HUGO gene name (symbol)[15]	Protein name
Q5T2N8	ATPase family, AAA domain containing 3C (ATAD3C)	ATPase family AAA domain-containing protein 3C
P17483	Homeobox B4 (HOXB4)	Homeobox protein Hox-B4
B4DDX2	N/A	cDNA FLJ54590, highly similar to KH domain-containing, RNA-binding, signaltransduction-associated protein 1
A0A0D9SFB3	DEAD-box helicase 3, X-linked (DDX3X)	ATP-dependent RNA helicase DDX3X
P62140	Protein phosphatase 1 catalytic subunit beta (PPP1CB)	Serine/threonine-protein phosphatase PP1-beta catalytic subunit
Q14145	Kelch like ECH associated protein 1 (KEAP1)	Kelch-like ECH-associated protein 1
P17482	Homeobox B9 (HoxB9)	Homeobox protein Hox-B9
P20700	Lamin B1 (LMNB1)	Lamin-B1
P35249	Replication factor C subunit 4 (RFC4)	Replication factor C subunit 4
B9A062	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase (MTHFD2)	Bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase, mitochondrial
P22695	Ubiquinol-cytochrome c reductase core protein 2 (UQCRC2)	Cytochrome b-c1 complex subunit 2, mitochondrial
P35453	Homeobox D13 (HOXD13)	Homeobox protein Hox-D13
P78527	Protein kinase, DNA-activated, catalytic polypeptide (PRKDC)	DNA-dependent protein kinase catalytic subunit
E9PQ56	Poly(U) binding splicing factor 60 (PUF60)	Poly(U)-binding-splicing factor PUF60 (Fragment)
P40939	Hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex subunit alpha (HADHA)	Trifunctional enzyme subunit alpha, mitochondrial
H3BQZ7	N/A	HCG2044799
A5PLK7	Regulator of chromosome condensation 2 (RCC2)	RCC2 protein (Fragment)
B4DU18	N/A	cDNA FLJ51093, highly similar to Cadherin-5

Accession number	HUGO gene name (symbol)[15]	Protein name
B2R8R5	N/A	cDNA, FLJ94025, highly similar to Homo sapiens tripartite motif-containing 28 (TRIM28), mRNA
Q9Y285	Phenylalanyl-tRNA synthetase alpha subunit (FARSA)	Phenylalanine--tRNA ligase alpha subunit
Q86YV0	Ras protein activator like 3 (RASAL3)	RAS protein activator like-3
B7Z8L8	N/A	cDNA FLJ50339, highly similar to A-kinase anchor protein 8
B4DP70	N/A	cDNA FLJ51017, highly similar to ATP-dependent RNA helicase DDX1 (EC 3.6.1.-) (Fragment)
O43707	Actinin alpha 4 (ACTN4)	Alpha-actinin-4
P24468	Nuclear receptor subfamily 2 group F member 2 (NR2F2)	COUP transcription factor 2
B7WNQ9	GATA binding protein 1 (GATA1)	Erythroid transcription factor
O15169	Axin 1 (AXIN1)	Axin-1
P43246	MutS homolog 2 (MSH2)	DNA mismatch repair protein Msh2
B3KUJ5	N/A	Histone deacetylase
B7Z9V0	N/A	cDNA, FLJ78964, highly similar to Coronin-1C
Q13422	IKAROS family zinc finger 1 (IKZF1)	DNA-binding protein Ikaros
Q7Z518	N/A	NADH dehydrogenase
B4DGI9	Transcription factor 12 (TCF12)	Transcription factor 12 (Fragment)
B4E0S6	N/A	cDNA FLJ55635, highly similar to pre-mRNA-splicing factorATP-dependent RNA helicase DHX15 (EC 3.6.1.-)
B0I1T2	Myosin IG (MYO1G)	Unconventional myosin-Ig
Q13547	Histone deacetylase 1 (HDAC1)	Histone deacetylase 1
B4DUD5	N/A	cDNA FLJ58787, highly similar to Cleavage stimulation factor 64 kDa subunit
Q08211	DExH-box helicase 9 (DHX9)	ATP-dependent RNA helicase A

Accession number	HUGO gene name (symbol)[15]	Protein name
B4DJ81	NADH:ubiquinone oxidoreductase core subunit S1 (NDUFS1)	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial
Q5JTC6	APC membrane recruitment protein 1 (AMER1)	APC membrane recruitment protein 1
H3BT57	Promyelocytic leukemia (PML)	Protein PML
V9P4H4	N/A	Magi1d
D3DV75	Adenosine deaminase, RNA-specific (ADAR)	Adenosine deaminase, RNA-specific, isoform CRA_b
B2RBM8	N/A	cDNA, FLJ95596, highly similar to Homo sapiens activity-dependent neuroprotector (ADNP), mRNA
B0AZS5	N/A	Kinesin-like protein
O75475	PC4 and SFRS1 interacting protein 1 (PSIP1)	PC4 and SFRS1-interacting protein
B3KM90	N/A	cDNA FLJ10529 fis, clone NT2RP2000965, highly similar to Targeting protein for Xklp2
B7Z2Z1	N/A	cDNA FLJ59523, highly similar to Scaffold attachment factor B
Q92878	RAD50 double strand break repair protein (RAD50)	DNA repair protein RAD50
Q92621	Nucleoporin 205 (NUP205)	Nuclear pore complex protein Nup205
Q4FD37	Zinc finger protein 148 (ZNF148)	ZBP-89 delta-Nter isoform
B0AZQ4	N/A	Structural maintenance of chromosomes protein
O75326	Semaphorin 7A (John Milton Hagen blood group) (SEMA7A)	Semaphorin-7A
Q9Y2X9	Zinc finger protein 281 (ZNF281)	Zinc finger protein 281
P98175	RNA binding motif protein 10 (RBM10)	RNA-binding protein 10
Q14151	Scaffold attachment factor B2 (SAFB2)	Scaffold attachment factor B2
O00308	WW domain containing E3 ubiquitin protein ligase 2 (WWP2)	NEDD4-like E3 ubiquitin-protein ligase WWP2

Accession number	HUGO gene name (symbol)[15]	Protein name
A0A0G2JMS7	Scribbled planar cell polarity protein (SCRIB)	Protein scribble homolog
B4DF22	N/A	cDNA FLJ55615, highly similar to SWI/SNF-related matrix-associatedactin-dependent regulator of chromatin subfamily C member 2
P35580	Myosin heavy chain 10 (MYH10)	Myosin-10
Q9UBW7	Zinc finger MYM-type containing 2 (ZMYM2)	Zinc finger MYM-type protein 2
O14776	Transcription elongation regulator 1 (TCERG1)	Transcription elongation regulator 1
Q15149	Plectin (PLEC)	Plectin
A0A024QZW7	Nucleoporin 153 (NUP153)	Nucleoporin 153kDa, isoform CRA_a
O75533	Splicing factor 3b subunit 1 (SF3B1)	Splicing factor 3B subunit 1
X5D2J9	General transcription factor Iii (GTF2I)	General transcription factor IIi isoform D (Fragment)
Q9C0D5	Tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 1 (TANC1)	Protein TANC1
G8JLG1	Structural maintenance of chromosomes 1A (SMC1A)	Structural maintenance of chromosomes protein
D3DR32	Kinesin family member 20B (KIF20B)	Kinesin-like protein
P49792	RAN binding protein 2 (RANBP2)	E3 SUMO-protein ligase RanBP2

Table 4.3. β-catenin binding partners in the K562 nuclear fraction prior to analysis.

A total of 53 proteins were present in the nuclear β-catenin IP and absent from the nuclear IgG control sample in K562 cells. These proteins are listed below and were used for the comparisons made in **Figure 4.5**. Proteins with HUGO gene names of N/A could not be found in the HUGO database based on their accession number or protein name.

Accession number	HUGO gene name (symbol)[16]	Protein name
P62249	Ribosomal protein S16 (RPS16)	40S ribosomal protein S16
P39019	Ribosomal protein S19 (RPS19)	40S ribosomal protein S19
P19474	Tripartite motif containing 21 (TRIM21)	E3 ubiquitin-protein ligase TRIM21
B7Z4V2	N/A	cDNA FLJ51907, highly similar to Stress-70 protein, mitochondrial
Q59H57	N/A	Fusion (Involved in t(12;16) in malignant liposarcoma) isoform a variant (Fragment)
P35232	Prohibitin (PHB)	Prohibitin
P50402	Emerin (EMD)	Emerin
F5GY37	Prohibitin 2 (PHB2)	Prohibitin-2
Q01196	Runt related transcription factor 1 (RUNX1)	Runt-related transcription factor 1
B3KQ75	N/A	cDNA FLJ33018 fis, clone THYMU1000459, highly similar to Homo sapiens transcription factor 7 (T-cell specific, HMG-box) (TCF7), transcript variant 5, mRNA
H0Y2W2	ATPase family, AAA domain containing 3A (ATAD3A)	ATPase family AAA domain-containing protein 3A (Fragment)
Q9UJU2	Lymphoid enhancer binding factor 1 (LEF1)	Lymphoid enhancer-binding factor 1
Q2TAM6	Runt related transcription factor 1 (RUNX1)	RUNX1 protein (Fragment)
Q96T67	ATPase family, AAA domain containing 3B (ATAD3B)	TOB3
E2GH18	Transcription factor 7 like 2 (TCF7L2)	T-cell factor-4 variant C
P12236	Solute carrier family 25 member 6 (SLC25A6)	ADP/ATP translocase 3
P36402	Transcription factor 7 (TCF7)	Transcription factor 7

[16] https://www.genenames.org/ (accessed on 26/03/2018)

Accession number	HUGO gene name (symbol)[16]	Protein name
Q96AG4	Leucine rich repeat containing 59 (LRRC59)	Leucine-rich repeat-containing protein 59
B7Z597	N/A	cDNA FLJ54373, highly similar to 60 kDa heat shock protein, mitochondrial
Q92841	DEAD-box helicase 17 (DDX17)	Probable ATP-dependent RNA helicase DDX17
Q9UPN1	Protein phosphatase 1 catalytic subunit gamma (PPP1CC)	Serine/threonine-protein phosphatase (Fragment)
Q4ZG57	Minichromosome maintenance complex component 6 (MCM6)	Putative uncharacterized protein MCM6 (Fragment)
C6ZRK5	Transcription factor 7 like 2 (TCF7L2)	TCF7L2 isoform pFC8A_TCF7L2_ex1-11-13-14
A0A0D9SGH8	Transcription factor 7 like 2 (TCF7L2)	Transcription factor 7-like 2
M0QXS5	Heterogeneous nuclear ribonucleoprotein L (HNRNPL)	Heterogeneous nuclear ribonucleoprotein L (Fragment)
B4DN41	N/A	cDNA FLJ53366, highly similar to Probable ATP-dependent RNA helicase DDX5 (EC 3.6.1.-)
Q96BA7	N/A	HNRPU protein
Q8N3E6	N/A	Putative uncharacterized protein DKFZp761L1023 (Fragment)
P55084	Hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex subunit beta (HADHB)	Trifunctional enzyme subunit beta, mitochondrial
B2R659	N/A	cDNA, FLJ92803, highly similar to Homo sapiens hydroxysteroid (17-beta) dehydrogenase 4 (HSD17B4), mRNA
B4DLA6	N/A	cDNA FLJ54365, highly similar to DNA replication licensing factor MCM4
B4DQY2	N/A	cDNA FLJ59388, highly similar to Mitochondrial inner membrane protein
M0QZM1	Heterogeneous nuclear ribonucleoprotein M (HNRNPM)	Heterogeneous nuclear ribonucleoprotein M (Fragment)
Q96PK6	RNA binding motif protein 14 (RBM14)	RNA-binding protein 14

Accession number	HUGO gene name (symbol)[16]	Protein name
A0A024QZP7	Cyclin dependent kinase 1 (CDK1)	Cell division cycle 2, G1 to S and G2 to M, isoform CRA_a
Q5T2N8	ATPase family, AAA domain containing 3C (ATAD3C)	ATPase family AAA domain-containing protein 3C
P42357	Histidine ammonia-lyase (HAL)	Histidine ammonia-lyase
B4DDX2	N/A	cDNA FLJ54590, highly similar to KH domain-containing, RNA-binding, signaltransduction-associated protein 1
P62140	Protein phosphatase 1 catalytic subunit beta (PPP1CB)	Serine/threonine-protein phosphatase PP1-beta catalytic subunit
A0A087WTP3	KH-type splicing regulatory protein (KHSRP)	Far upstream element-binding protein 2
P35249	Replication factor C subunit 4 (RFC4)	Replication factor C subunit 4
B4E0E1	N/A	cDNA FLJ53442, highly similar to Poly (ADP-ribose) polymerase 1 (EC 2.4.2.30)
P78527	Protein kinase, DNA-activated, catalytic polypeptide (PRKDC)	DNA-dependent protein kinase catalytic subunit
Q8NBX0	Saccharopine dehydrogenase (putative) (SCCPDH)	Saccharopine dehydrogenase-like oxidoreductase
Q6NUK7	LYN proto-oncogene, Src family tyrosine kinase (LYN)	Non-specific protein-tyrosine kinase (Fragment)
O15169	Axin 1 (AXIN1)	Axin-1
A0A024R5M9	Nuclear mitotic apparatus protein 1 (NUMA1)	Nuclear mitotic apparatus protein 1, isoform CRA_a
P23490	Loricrin (LOR)	Loricrin
Q92621	Nucleoporin 205 (NUP205)	Nuclear pore complex protein Nup205
O00308	WW domain containing E3 ubiquitin protein ligase 2 (WWP2)	NEDD4-like E3 ubiquitin-protein ligase WWP2

4.3.5 *Filtering of candidate nuclear β-catenin import factors*

Following the identification of potential candidates, these data were further refined by removing proteins based on certain criteria, as outlined in **Table 4.4**.

Two proteins were removed as they were suspected to be present because of mitochondrial contamination, rather than nuclear proteins. These were the well characterized mitochondrial proteins ATPase family AAA domain-containing protein 3C and ADP/ATP translocase 3. Six candidates were removed because they were only evident at the transcript level (cDNA) or had only been experimentally verified at a transcript level[17] and three proteins were removed because they were present in non-translocator cytoplasmic fractions.

Next, proteins present in more than two IgG co-immunoprecipitates were removed from the candidate list. Their presence in the IgG samples suggests they bind to IgG non-specifically (rather than being a specific β-catenin binding partner). Proteins present in more than two IgG samples were chosen as the cut-off point resulting in the removal of three candidates.

The remaining proteins were organised into a final candidate list (**Table 4.5**) consisting of 12 proteins. It is important to point out that some of these candidate factors (DNA-dependent protein kinase, TCF7 and TCF7L2 (TCF-4)) were detected in the cytoplasmic fraction of one of the non-translocator cell lines with a peptide score of >2, however they were included in the candidate list for specific reasons, as outlined in **Table 4.6**. Briefly, DNA-dependent protein kinase was included in the list because in the cell line in which it was detected, it was also detected in the IgG lane. TCF7L2 and TCF7 were included in the list because of their identification as candidate nuclear localization factors in other contexts (4.4.2).

Candidate proteins with commercially available antibodies and of functional relevance (had already been identified as potential β-catenin nuclear localization factors; **Table 1.10**) were chosen for validation. Due to time constraints, not all the proteins could be validated.

[17] http://www.uniprot.org (accessed on 04/01/2018)

Table 4.4. Filtering of candidate β-catenin nuclear localization factors.

Proteins were filtered out based on the following properties: green; present in non-translocator (NT) cytoplasmic IPs, grey; probable mitochondrial contamination, peach; only evident at transcript level, blue; present in >2 IgG samples. The reason that only 24 proteins are included in this list (rather than the 27 identified in **Figure 4.5**) is because α-catenin 1, α-catenin 2 and APC were removed, owing to their presence in the cytoplasm of non-translocator cell fractions as well. Proteins with HUGO gene names of N/A could not be found in the HUGO database based on their accession number or protein name.

Accession number	HUGO gene name (symbol)[18]	Candidate protein name	Potential Mechanism
O00308	WW domain containing E3 ubiquitin protein ligase 2 (WWP2)	NEDD4-like E3 ubiquitin-protein ligase WWP2	Nuclear retention
B4DDX2	N/A	cDNA FLJ54590, highly similar to KH domain-containing, RNA-binding, signaltransduction-associated protein 1	Nuclear retention
P35249	Replication factor C subunit 4 (RFC4)	Replication factor C subunit 4	Nuclear retention
O15169	Axin1 (AXIN1)	Axin-1	Nuclear retention
P36402	Transcription factor 7 (TCF7)	Transcription factor 7	Nuclear retention
A0A0D9SGH8	Transcription factor 7 like 2 (TCF7L2)	Transcription factor 7-like 2	Nuclear retention
Q92621	Nucleoporin 205 (NUP205)	Nuclear pore complex protein Nup205	Nuclear retention
B3KQ75	N/A	cDNA FLJ33018 fis, clone THYMU1000459, highly similar to Homo sapiens transcription factor 7 (T-cell specific, HMG-box) (TCF7), transcript variant 5, mRNA	Nuclear retention
Q96T67	ATPase family, AAA domain containing 3B (ATAD3B)	TOB3	Cytoplasmic/nuclear shuttling

[18] https://www.genenames.org/ (accessed on 26/03/2018)

Accession number	HUGO gene name (symbol)[18]	Candidate protein name	Potential Mechanism
H0Y2W2	ATPase family, AAA domain containing 3A (ATAD3A)	ATPase family AAA domain-containing protein 3A	Cytoplasmic/nuclear shuttling
Q59H57	N/A	Fusion (Involved in t(12;16) in malignant liposarcoma) isoform a variant	Cytoplasmic/nuclear shuttling
Q9UJU2	Lymphoid enhancer binding factor 1 (LEF1)	Lymphoid enhancer-binding factor 1	Cytoplasmic/nuclear shuttling
Q2TAM6	Runt related transcription factor 1 (RUNX1)	RUNX1 protein	Cytoplasmic/nuclear shuttling
P78527	Protein kinase, DNA-activated, catalytic polypeptide (PRKDC)	DNA-dependent protein kinase catalytic subunit	Cytoplasmic/nuclear shuttling
Q96PK6	RNA binding motif protein 14 (RBM14)	RNA-binding protein 14	Cytoplasmic/nuclear shuttling
Q5T2N8	ATPase family, AAA domain containing 3C (ATAD3C)	ATPase family AAA domain-containing protein 3C	Nuclear retention
Q8N3E6	N/A	Putative uncharacterized protein DKFZp761L1023	Nuclear retention
B4DQY2	N/A	cDNA FLJ59388, highly similar to Mitochondrial inner membrane protein	Nuclear retention
P62140	Protein phosphatase 1 catalytic subunit beta (PPP1CB)	Serine/threonine-protein phosphatase PP1-beta catalytic subunit	Nuclear retention
B7Z597	N/A	cDNA FLJ54373, highly similar to 60 kDa heat shock protein, mitochondrial	Nuclear retention
B4DLA6	N/A	cDNA FLJ54365, highly similar to DNA replication licensing factor MCM4	Nuclear retention
E2GH18	Transcription factor 7 like 2 (TCF7L2)	T-cell factor-4 variant C	Nuclear retention

Accession number	HUGO gene name (symbol)[18]	Candidate protein name	Potential Mechanism
C6ZRK5	Transcription factor 7 like 2 (TCF7L2)	TCF7L2 isoform pFC8A_TCF7L2_ex1-11-13-14	Nuclear retention
P12236	Solute carrier family 25 member 6 (SLC25A6)	ADP/ATP translocase 3	Nuclear retention

Table 4.5. Final candidate list.

Following filtering of candidate nuclear localization factors (4.3.5) a final candidate list was produced.

Accession number	HUGO gene name (symbol)[19]	Candidate protein name	Role	Potential mechanism
O15169	Axin1 (AXIN1)	Axin-1	Part of the Canonical Wnt Pathway catenin degradation complex.	Nuclear retention
P36402	Transcription factor 7 (TCF7)	Transcription factor 7	Canonical Wnt Pathway component that is a co-activator of Wnt target genes.	Nuclear retention
A0A0D9SGH8	Transcription factor 7 like 2 (TCF7L2)	Transcription factor 7-like 2 (TCF-4)		Nuclear retention
P78527	Protein kinase, DNA-activated, catalytic polypeptide (PRKDC)	DNA-dependent protein kinase catalytic subunit	Ser/thr kinase involved in sensing DNA damage	Nuclear retention
Q96PK6	RNA binding motif protein 14 (RBM14)	RNA-binding protein 14	Ribonucleoprotein. General nuclear co-activator.	Nuclear retention
P62140	Protein phosphatase 1 catalytic subunit beta (PPP1CB)	Serine/threonine-protein phosphatase PP1-beta catalytic subunit		Nuclear retention
Q2TAM6	Runt related transcription factor 1 (RUNX1)	RUNX1 protein	Transcription factor involved in regulating haematopoiesis. Aberrations in RUNX1 are involved in leukaemia.	Nuclear retention
Q9UJU2	Lymphoid enhancer binding factor 1 (LEF1)	Lymphoid enhancer-binding factor 1	Canonical Wnt Pathway component that is a co-activator of Wnt target genes.	Cytoplasmic-nuclear shuttling

[19] https://www.genenames.org/ (accessed on 26/03/2018)

Table 4.6 Reasons for keeping specific proteins in the candidate list

Candidate name	Reason
DNA-dependent protein kinase	The non-translocator cell line in which this protein was detected also contained it in the IgG control lane, so this could be due to non-specific binding in this cell line meaning that it could be falsely excluded by this criterion. Therefore, this protein was kept in the candidate list
TCF7	This has relevance to β-catenin mediated transcription and possibly to β-catenin nuclear localization in other contexts (4.4.2.2). Although this protein was detected in the cytoplasm of one of the non-translocator cell lines, the lack of data for the nuclear fractions of non-translocator cell lines meant that this factor could not be completely ruled out as a candidate β-catenin nuclear localization factor in this context.
TCF7L2 (TCF-4)	This was kept in the candidate list for the same reasons as TCF7.

4.3.6 *Validation of candidate β-catenin nuclear localization factors*

Following the establishment of a final candidate list, their expression and subcellular localization needed validating to verify their potential role as β-catenin nuclear localization factors. To do this, leukaemia cell lines were fractionated and western blotting was used to verify that the localization of candidate proteins was consistent with the MS results (Chapter 0). Axin-1, TCF7, LEF-1 and RUNX1 were all chosen for analysis based on the criteria above (4.3.5). The reasons for selecting these factors and not others for western blotting is discussed further in 4.4.

Axin-1 could not be verified as no signal was detected following western blotting (data not shown). **Figure 4.6** shows the results of the western blots for the other candidate factors. Firstly, translocator and non-translocator cell lines were immunoblotted with TCF7 antibody (**Table 2.5**). Three bands were visible in the nuclear fraction of K562 cells at approximately 43-51kDa, corresponding to the expected band for full length TCF7. This was not visible in HEL cells or non-translocator cell lines. This contrasted with the MS results, in which TCF7 was distributed in the nucleus of both translocator cell lines. This may be because the peptide count for HEL cells was lower than K562 cells and/or because of differences in sensitivity between the two techniques. This suggests that TCF7 could be a potential candidate regulating the nuclear localization of β-catenin (based on the MS results) but the levels on TCF7 in HEL cells were undetectable by western blot and hence TCF7 could not be validated.

Next, leukaemia cell lines were immunoblotted with LEF-1 and RUNX1 antibodies. In both translocator cell line nuclear fractions, bands were visible at ~50kDa upon blotting with both antibodies, corresponding to the expected bands for LEF-1 and RUNX1. A weak signal was observed in the nuclear fractions of non-translocator cell lines when blotted with RUNX1 antibody. This may correlate with β-catenin localization in these cells, as although they are classed as non-translocator cell lines, a low level of β-catenin does localize to the nucleus in these cells (Appendix 1). No LEF-1 signal was detected in the nucleus of non-translocator cell lines. The localization of LEF-1 and RUNX1 in these cell lines

corresponds to the expected localization of a potential β-catenin nuclear import factor (**Table 4.1**), suggesting LEF-1 and RUNX1 could be potential candidates.

To summarize, Axin-1 could not be verified as a potential β-catenin nuclear localization factor. The localization of TCF7 was inconsistent with the results of the MS analysis and could not be fully verified as a potential β-catenin nuclear localization factor. The localization of LEF-1 and RUNX1 was consistent with the MS results (absent in the cytoplasm of all cell lines, present in the nucleus of translocator cell lines and absent in the nucleus of non-translocator cell lines) and corresponded to that of potential nuclear β-catenin localization factors.

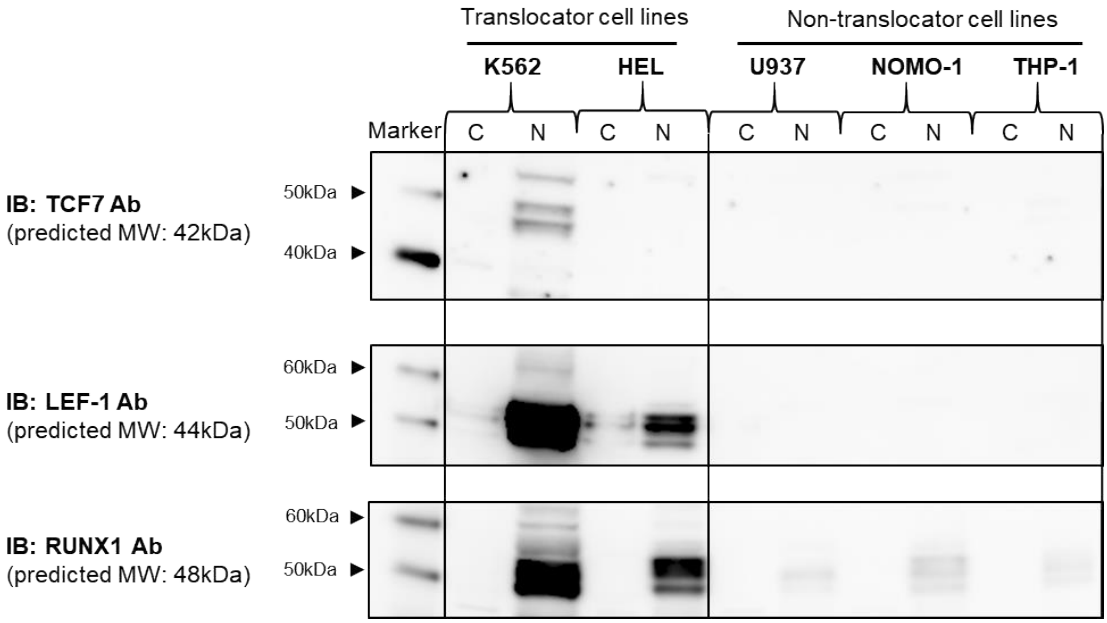


Figure 4.6. Validation of candidate factors by Western blotting in leukaemia cell lines.

Western blots of translocator (left) and non-translocator (right) cells following cytoplasmic/nuclear fractionation. Antibodies for the candidate β-catenin nuclear localization factors TCF7, LEF-1 and RUNX1 were used (**Table 2.5**). Abbreviations; C= cytoplasmic fraction N= nuclear fraction, IB= immunoblot, Ab= antibody, MW= molecular weight, kDa= kilodalton. Marker= MagicMark XP (Thermo Fisher).

4.4 Discussion

Nuclear localization of β-catenin is fundamental for its role as a transcription factor. Despite this, little is known about the mechanisms regulating this process in AML; previous work has shown AML blasts and cell lines are highly variable in their ability to translocate β-catenin from the cytosol to the nucleus (Morgan *et al*, 2014). The aim of the experiments presented in this chapter was to analyse the data generated from the MS approach outlined in (**Figure 4.1**) to identify candidate β-catenin nuclear localization factors in AML. This section discusses aspects of the experimental techniques used, as well as the factors that were identified as candidate β-catenin nuclear localization factors in AML.

4.4.1 ***Experimental techniques***

4.4.1.1 *Sample quality*

One concern regarding the MS experiment was the quality of the samples, particularly with regards to the presence of proteins in the IgG control samples. Each of the samples had a high number of proteins in the IgG control lane. To calculate the extent of the problem, the number of proteins detected in the IgG lane for each sample was calculated as a percentage of the total number of proteins in each sample. The percentage of proteins detected in the IgG lanes was 72.5% in the K562 cytoplasmic fraction, 53.7% in the HEL cytoplasmic fraction, 76.2% in the THP-1 cytoplasmic fraction, 71.9% in the NOMO-1 cytoplasmic fraction, 85.5% in the U937 cytoplasmic fraction, 79.5% in the K562 nuclear fraction and 73.8% in the HEL nuclear fraction. Unfortunately, this high degree of IgG protein binding could not be detected prior to running the MS analysis, although steps were taken throughout the sample preparation to reduce the amount of background binding (Chapter 3). Briefly, a monoclonal antibody was used in the immunoprecipitation, a preclearing step was used to remove proteins that bound to IgG control antibody/beads from the cell fractions, the bead complex was washed thoroughly, and a fresh tube was used for elution (Moser *et al*, 2009). Benzonase was included in the nuclear buffer to break down DNA and prevent the sticking of DNA and DNA bound proteins to the beads in nuclear

fractions (Benedik & Strych, 1998; Liao *et al,* 2007). In addition, gloves were always worn and the workspace was kept as clean as possible whilst the samples were prepared to limit the amount of keratin contamination (Hodge *et al*, 2013). Finally, Sypro staining was used to detect bands present in the IgG control and immunoprecipitation samples (data not shown) (Chevalier, 2010). The Sypro stain did not detect the large number of bands that were observed in the MS of IgG samples.

Perhaps the detergent used in this experiment was the reason so many background binding proteins were present. To maintain protein-protein interactions between β-catenin and its binding partners, a mild detergent, CHAPS, was used for cytoplasmic/nuclear fractionation and in the immunoprecipitation buffer. As a trade-off for maintaining β-catenin interaction partner binding, it could also have made the immunoprecipitation washes less effective, with less disruption of non-specific weakly bound proteins (Huang & Kim, 2013).

4.4.1.2 *β-catenin was detected in the IgG controls*

Following on from the previous section, another troubling issue was the presence of β-catenin in the IgG control lanes of each sample, with peptide scores in the range of 1-5. This may be due to some of the reasons outlined in the section above, and was problematic for data analysis. It was necessary to remove proteins present in the IgG lanes as part of the data analysis approach (**Figure 4.3**). As a result, additional candidate factors may have been removed as part of the filtering process, if they were present in the IgG samples because of their co-localization with β-catenin. This could not be avoided, however, because it was necessary to remove the proteins in the IgG samples to remove background binding proteins. It is likely that the results of this experiment have been particularly biased towards proteins that bind strongly to β-catenin and that are expressed at higher levels, with proteins that bind more weakly and/or have a lower expression being missed.

4.4.2 *Identification of candidate β-catenin nuclear localization factors*

Eight candidate β-catenin nuclear localization factors were identified (**Table 4.5**) in this project using β-catenin immunoprecipitation and MS analysis of β-catenin interaction partners. Each of these candidates are discussed in more detail below.

4.4.2.1 *Axin-1*

Axin-1 is a component of the canonical Wnt pathway catenin degradation complex that has a well-established role in regulating the stability of cytoplasmic β-catenin (1.3.1.1) and mutations in Axin-1 have been observed in solid tumours (Mazzoni & Fearon, 2014).

In this project, Axin-1 was identified as a candidate β-catenin nuclear retention factor. This contrasts with other studies in which Axin-1 has been identified as a β-catenin nuclear export factor. In HEK293 cells, overexpression of Axin-1 promoted a shift in β-catenin from the nucleus to the cytoplasm (Cong & Varmus, 2004). In this study, nuclear import and nuclear export signals (NLS and NES) were identified in Axin-1, and the results of this study suggested that Axin-1 functions as a nuclear-cytoplasmic β-catenin shuttling protein. In another study, Axin-1 expression in HEK293T cells promoted the cytoplasmic localization of β-catenin, however, the proposed mechanism was by cytoplasmic retention. This was because expressing Axin-1 did not increase the speed of β-catenin transport across the nuclear membrane (Krieghoff *et al*, 2006).

Since Axin-1 has been identified as a β-catenin nuclear export protein, the observation that Axin-1 may be a β-catenin nuclear localization factor in leukaemia cell lines was unexpected, but given the lack of data for the nuclear fractions of non-translocator cells and the fact that Axin-1 localization could not be validated by western blotting, the most that can be stated is that Axin-1 is a nuclear β-catenin binding partner in the translocator cell lines studied. Further studies would be necessary to determine if it acts as a nuclear β-catenin localization factor in this context.

The reasons for the ineffective detection of Axin-1 by western blots could be because of its size and/or low protein expression levels. Axin-1 is a large protein with a molecular weight of 110kDa. Larger proteins can sometimes be difficult to detect by western blotting because they transfer less efficiently from the SDS-PAGE gel to the membrane. If Axin-1 protein levels are also low in these cells, this could explain why it was difficult to detect it by western blotting. *Axin-1* mRNA levels in these cell lines are not low[20], however, it could be that protein levels do not correlate with mRNA levels in these cells. In other studies, the success of Axin-1 detection by western blotting is variable depending on the context. While Axin-1 can be detected easily in some contexts, for example, platelets and breast cancer cells (MDA-MB-231 and MCF-7 cells), in others e.g. HEK293T cells, although detection was successful, only a very faint band was observed using the same antibody (Arnold *et al*, 2009; Bao *et al*, 2012; Steele *et al*, 2009). If more time had been available, I would have attempted to detect Axin-1 by immunoprecipitation in my cells.

4.4.2.2 *TCF7*

TCF7 (also known as TCF-1) is a member of the TCF family of transcription factors with well-established roles in regulating β-catenin mediated transcription (1.3), and it has recently been shown to promote the self-renewal and survival of LSCs, maintaining the LSC in the bone marrow and propagating disease (Yu *et al,* 2016).

In this study, TCF7 was identified by MS as a candidate β-catenin nuclear retention protein, which agrees with the observations made by Molenaar, 1996. Injection of the Xenopus TCF7 homologue (XTcf-3) into Xenopus embryos led to translocation of β-catenin from the cytoplasm and plasma membrane to the nucleus (Molenaar *et al*, 1996).

The results of western blotting of translocator and non-translocator cell line fractions with TCF7 antibody (**Figure 4.6**) contrasted with the results of the MS analysis. In the western blots, TCF7 was only detected in the nucleus of K562 cells, whereas using MS it was detected in the nucleus of both translocator cells.

[20] https://www.proteinatlas.org/ENSG00000103126-AXIN1/cell (accessed on 04/01/2018)

As mentioned previously, this could be due to the differences in sensitivity of the two techniques, but could also be due to differences in the level of TCF7 in these cell lines. The levels of *TCF7* mRNA expression are low in K562 cells, but not detected in HEL cells[21]. This could indicate that protein levels might also be very low in HEL cells, which would explain the results of the western blot. Due to the inconsistencies between the MS and western blotting, no further analysis of this protein was conducted.

4.4.2.3 *TCF7L2*

TCF7L2 (also known as TCF-4) is a member of the TCF family of transcription factors with well-established roles in regulating β-catenin mediated transcription (1.3). TCF7L2 is also dysregulated in AML; microarray analysis of AML blasts from 223 patients identified TCF7L2 as the most significantly dysregulated Wnt factor in this patient cohort, with overexpression of *TCF7L2* mRNA being observed in 78% of samples (Daud, 2014).

In this project, TCF7L2 was identified as a potential β-catenin nuclear retention factor in leukaemia cell lines. This agrees with other studies that have identified TCF7L2 as a β-catenin nuclear localization factor. Expression of TCF7L2 in MCF-7 (breast cancer) cells with low levels of TCF proteins led to the nuclear localization of β-catenin (Jamieson *et al*, 2016). In another study, overexpression of TCF7L2 in HEK293T cells localized β-catenin to the nucleus, and the authors proposed that it was acting as a nuclear retention factor. This was based on the observation that TCF7L2 overexpression slowed down β-catenin transport across the nuclear membrane, rather than speeding it up (Krieghoff *et al*, 2006).

Another study looking at the correlation of nuclear β-catenin and nuclear TCF7L2 in lung cancer brain metastases, suggests that TCF7L2 does not regulate the nuclear localization of β-catenin (Bleckmann *et al*, 2013). In this study, TCF7L2 was detected in the nucleus of 100% of the samples used, whereas nuclear β-catenin was only detected in 36% of samples. It is difficult to draw conclusions from this observation because no functional analysis was conducted (e.g. knockdown or overexpression studies). It could be that specific post translational

[21] https://www.proteinatlas.org/ENSG00000081059-TCF7/cell (accessed on 04/01/2018)

modifications and/or interaction partners are necessary for TCF7L2 mediated nuclear localization of β-catenin, and the factors responsible for this are not present in this context. Alternatively, TCF7L2 may regulate β-catenin nuclear localization, but another factor may be present that more efficiently exports β-catenin out of the nucleus in the cells in which nuclear β-catenin was not detected.

Although TCF7L2 was identified as a candidate nuclear β-catenin localization factor by my MS approach, it was not validated by western blotting because previous studies in our laboratory have found that the localization of TCF7L2 and β-catenin does not correlate in these lines, suggesting that it might not regulate β-catenin nuclear localization in this context (Appendix 2).

4.4.2.4 *DNA dependent protein kinase*

DNA dependent protein kinase (DNA-PK) is a serine/threonine protein kinase that plays a role in DNA repair, mitosis, chromatin remodelling and telomere maintenance (Smith & Jackson, 1999). In this project, DNA-PK was identified as a candidate β-catenin nuclear retention factor. Interestingly, a previous study using LEF-1 immunoprecipitation and MS analysis in adenocarcinoma cell lines (SW480 and HT29) identified DNA-PK as a candidate LEF-1 interaction partner. Co-immunoprecipitation and immunohistochemistry were used to validate the binding of LEF-1 and DNA-PK and colocalization of LEF-1 and DNA-PK in the nucleus of these cells (Shimomura *et al*, 2013). It is possible that although we have proposed DNA-PK as a β-catenin interaction partner in this project, it may in fact bind indirectly via LEF-1. If this is the case, it could be that DNA-PKs localization in the nucleus of translocator lines does not mediate the nuclear localization of β-catenin, but is due to its co-localization with LEF-1 (see below).

An interesting hypothesis is that DNA-PK could regulate LEF-1 mediated β-catenin nuclear localization. Since LEF-1 has also been identified as a candidate nuclear β-catenin localization factor and DNA-PK has been implicated in the regulation of nuclear import of other proteins by phosphorylation of residues close to their NLS (Xiao *et al*, 1997), perhaps DNA-PK could regulate the nuclear localization of β-catenin by interacting with LEF-1, phosphorylating its NLS and promoting LEF-1 mediated β-catenin nuclear localization. This might explain the

inefficient nuclear localization of LEF-1 in non-translocator cell lines in which LEF-1 was overexpressed (Chapter 5).

This protein was not chosen as a factor for western blotting to validate its localization, since most of the other factors identified were established β-catenin binding partners and had already been implicated in the nuclear localization of β-catenin in other contexts and/or had been implicated in AML. As such they were favoured for further analysis and due to the time constraints of the project not all the factors could be analysed.

Given the lack of data for the nuclear fractions of non-translocator cells and the fact that DNA-PK localization was not validated by western blotting, the most that can be stated is that DNA-PK is a nuclear β-catenin binding partner in the translocator cell lines studied. Further studies would be necessary to determine if it acts as a nuclear β-catenin localization factor in this context.

4.4.2.5 *RNA binding protein 14*

RNA-binding protein 14, also known as co-activator activator (CoAA) is a general nuclear transcriptional regulator. In this study it was identified as a candidate β-catenin nuclear retention factor. To my knowledge, it has not been identified previously as a β-catenin interaction partner. Interestingly, RNA-binding protein 14 has been demonstrated to bind to the runt domain of RUNX proteins (4.4.2.7) (Li *et al*, 2009), which could suggest that it binds indirectly to β-catenin via its interaction with RUNX1.

This protein was not chosen as a factor for western blotting to validate its localization in leukaemia cell lines, since most of the other factors identified were established β-catenin binding partners and had already been implicated in the nuclear localization of β-catenin in other contexts and/or had been implicated in AML. As such they were favoured for further analysis and due to the time constraints of the project not all the factors could be analysed.

Given the lack of data for the nuclear fractions of non-translocator cells and the fact that RNA-binding protein 14 localization was not validated by western blotting, the most that can be stated is that RNA-binding protein 14 interacts in a complex with β-catenin in the nucleus of the translocator cell lines studied.

4.4.2.6 *PP1-β*

Protein phosphatase 1-beta (PP1-β*)* is a global serine/threonine phosphatase that performs varied roles, for example, in glycogen metabolism, cell-cycle progression and protein synthesis (Korrodi-Gregório *et al*, 2014; Rebelo *et al*, 2015). PP1 is involved in the regulation of canonical Wnt signalling by its interaction with and dephosphorylation of Axin, leading to less effective destruction of β-catenin by the catenin destruction complex, and subsequent stabilization of β-catenin (Kim *et al*, 2013; Luo *et al*, 2007).

In this project PP1-β was identified as a candidate β-catenin nuclear retention factor. To my knowledge, PP1 has not been identified as a direct binding partner of β-catenin, so it could be that PP1 binds indirectly via its association with Axin-1, which was also identified as a candidate β-catenin nuclear retention factor. Similarly, another protein phosphatase family member, PP2A, also binds β-catenin indirectly via other catenin destruction complex components e.g. Axin (Hsu *et al*, 1999; Stamos & Weis, 2013). Again, to my knowledge PP1 has not been implicated as a direct regulator of β-catenin mediated localization, except because of stabilization of β-catenin following dephosphorylation of Axin (Kim *et al*, 2013; Luo *et al*, 2007). This does not apply in a haematopoietic context, in which stabilization of β-catenin does not normally lead to its nuclear localization (Morgan *et al*, 2014).

Owing to the uncertainty about the regulation of β-catenin nuclear localization by PP1, it was not chosen for validation by western blotting.

4.4.2.7 *RUNX1*

Runt-related transcription factor 1 (RUNX1) is an important haematopoietic transcription factor and is commonly dysregulated in AML (1.2.2); either fused with other genes through chromosomal translocations or mutated. In this project, RUNX1 was identified as a candidate β-catenin nuclear retention factor in leukaemia lines. To my knowledge, RUNX1 has not been identified as a β-catenin interaction partner, but it has been identified as a LEF-1 interaction partner (Kahler & Westendorf, 2003), suggesting that RUNX1 may indirectly interact with β-catenin via LEF-1 in the nucleus of leukaemia cell lines.

RUNX1 was one of the two proteins whose localization was verified by western blotting (**Figure 4.6**), being present in the nucleus of translocator cell lines and absent in non-translocator cell . This corresponded to the localization of LEF-1, and together these two proteins were the only two factors identified by the MS approach that were validated as potential nuclear β-catenin localization factors. Due to time constraints, RUNX1 was not included in further functional analysis, as LEF-1 was favoured based on its more established role as a candidate β-catenin nuclear localization factor in other contexts.

4.4.2.8 *LEF-1*

LEF-1 is a member of the TCF family of transcription factors with well-established roles in regulating β-catenin mediated transcription (1.3). It is frequently dysregulated in AML and its overexpression has clinical significance, with LEF-1 overexpression being associated with a favourable prognosis (Fu *et al*, 2014; Metzeler *et al*, 2012). In this project, LEF-1 was identified as a candidate nuclear localization factor. LEF-1 has been shown to regulate the nuclear localization of β-catenin in other contexts, including, SW480, NIH-3T3 and MDCK cell lines and mouse embryos (Henderson *et al*, 2002; Huber *et al*, 1996; Jamieson *et al*, 2016; Jamieson *et al*, 2011; Simcha *et al*, 1998). In SW480 colon cancer cells, immunohistochemistry was used to detect the localization of β-catenin in cells i) transfected with APC and ii) co-transfected with APC and LEF-1. Whilst transfection of cells with APC reduced nuclear β-catenin levels, co-transfection of cells with APC and LEF-1 prevented APC mediated export of β-catenin (Henderson *et al*, 2002), suggesting that LEF-1 regulates the nuclear localization of β-catenin in these cells. LEF-1 has also been implicated as a β-catenin nuclear localization factor following GSK3β inhibition in NIH-3T3 and SW480 cells (Jamieson *et al*, 2011). In another study, nuclear injection of LEF-1 cDNA into blastomeres of murine two-cell stage embryos, led to co-localization of β-catenin and LEF-1 in the nucleus. In blastomeres without injection of LEF-1 cDNA, this localization was not observed (Huber *et al*, 1996). Another study, in which immunohistochemistry was used to detect the localization of β-catenin in LEF-1 overexpressing canine MDCK cells, also suggests that LEF-1 acts as a β-catenin nuclear localization factor (Simcha *et al*, 1998).

In summary, MS analysis identified eight potential candidate nuclear localization factors. Out of these, factors were primarily chosen for further analysis due to their identification as candidate β-catenin nuclear localization factors in other contexts. This contrasts with the project aim outlined in **1.4**, to identify novel candidate β-catenin nuclear localization factors, however, given the tight time constraints towards the end of this project, this represented the approach most likely to identify β-catenin nuclear localization factors in AML. Since these factors had not yet been identified as candidates in this context, this approach still had potential to discover novel findings in AML. Out of the eight candidates, two of the chosen factors, were successfully validated by western blotting (LEF-1 and RUNX1). Again, due to the time constraints of the project, only one of these factors could be examined further, and LEF-1 was chosen for further analysis, based on its well-established role in canonical Wnt signalling, and its identification as a candidate β-catenin nuclear localization factor in other contexts.

5 The role of LEF-1 in regulating the nuclear localization of β-catenin and its functional relevance in leukaemia

5.1 Introduction

Chapter 4 identified candidate factors involved in regulating the nuclear localization of β-catenin. Further analysis of these proteins is necessary to validate their role in this process and their functional relevance in leukaemia. Due to time constraints, analysis of every candidate is outside the scope of this project. Instead, LEF-1 was chosen for further analysis due to its established role in regulating the Wnt pathway. Upon Wnt activation, LEF-1 (a TCF family member) binds in a complex with active β-catenin and regulates transcription of β-catenin target genes (1.3.1.2). Additionally, LEF-1 has been implicated in the nuclear localization of β-catenin in different contexts by cytosolic-nuclear shuttling and/or nuclear retention (Behrens *et al*, 1996; Henderson *et al*, 2002; Huber *et al*, 1996; Jamieson *et al*, 2016; Jamieson *et al*, 2011). LEF-1 has also been implicated in AML pathogenesis. Transplantation of bone marrow cells overexpressing LEF-1 into mice leads to the development of AML, and LEF-1 was also identified as one of the Wnt pathway components regulating AML by AML associated fusion proteins (Müller-Tidow *et al*, 2004; Petropoulos *et al*, 2008). LEF-1 also has clinical relevance in AML, with LEF-1 overexpression being associated with a favourable prognosis (Fu *et al*, 2014; Metzeler *et al*, 2012). Together, these factors make LEF-1 a good candidate for further analysis.

5.2 Aims

- To establish whether LEF-1 regulates the nuclear localization of β-catenin in leukaemia cell lines.
- To examine the effects of modulating LEF-1 on β-catenin mediated transcription.
- To determine whether LEF-1 mediated β-catenin nuclear localization is important in regulating key cancer processes in leukaemia cell lines including; proliferation, viability, migration, survival and differentiation.

5.3 Results

5.3.1 *LEF-1 regulates the nuclear localization of β-catenin in leukaemia cell lines*

LEF-1 has been implicated in the regulation of β-catenin nuclear localization in different contexts via cytoplasmic-nuclear shuttling or retention of β-catenin in the nucleus. To examine whether LEF-1 regulates the nuclear localization of β-catenin in leukaemia, lentiviral constructs were used to knockdown and overexpress full length LEF-1 in pBARV reporter lines (2.3). Cell lines were induced with the GSK3β inhibitor BIO (Meijer *et al*, 2003) to stabilize β-catenin, fractionated and western blotted (Chapter 0).

5.3.1.1 *Knockdown of LEF-1 in translocator cell lines*

To study the role of LEF-1 in translocator cell lines (K562 and HEL) an shRNA LEF-1 construct was used to knockdown LEF-1 in these cell lines. To choose a suitable shRNA construct, that effectively knocked down LEF-1 and did not have adverse effects on the cells, the knockdown efficiencies of five different LEF-1 shRNA constructs (**Table 2.3**) were tested in K562 cells. Cells were grown, transduced with LEF-1 shRNA and fractionated into nuclear and cytoplasmic fractions. A control shRNA vector (harbouring a sequence to a non-human gene) was included for comparison (**Table 2.3**). Due to the exclusive localization of LEF-1 in the nucleus of K562 cells, only nuclear fractions were analysed by western

blotting. Each of the LEF-1 shRNA constructs reduced expression of LEF-1 (**Figure 5.1A**). None of the constructs had adverse consequences for the cells with all of them coming through puro selection equivalently to shRNA control lines. The percentage of LEF-1 knockdown for each of the LEF-1 shRNA constructs was calculated. These were; ~82% for shRNATRCN0000020163, ~62% for shRNATRCN0000413476, ~90% for shRNATRCN0000418104, ~83% for shRNATRCN0000428178 and ~95% for shRNATRCN0000428355. The most effective LEF-1 shRNA construct (shRNATRCN0000428355) was chosen for use in further analysis in K562 and HEL cells. A control shRNA vector was included in all experiments (**Figure 5.1B**). Knockdown efficiency using the chosen LEF-1 shRNA was 78% in K562 cells and 76% in HEL cells (**Figure 5.2A**).

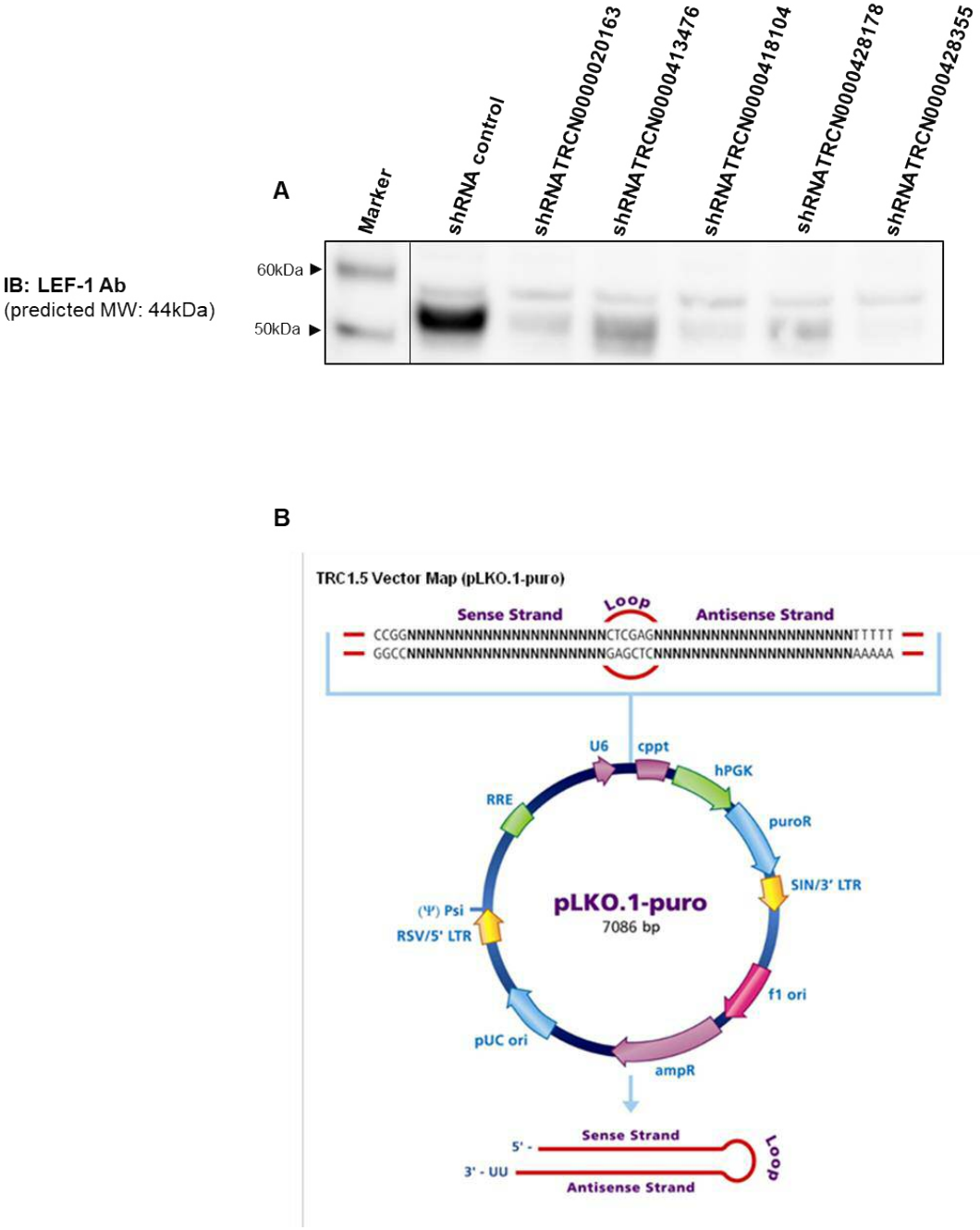


Figure 5.1. LEF-1 shRNA vectors.

(**A**) Western blot of the five LEF-1 shRNA vectors tested in K562 cells and immunoblotted with LEF-1 antibody (**Table 2.5**). A control shRNA vector was included for comparison. (**B**) shRNA control vector map. Abbreviations; IB= immunoblot, Ab= antibody, MW= molecular weight. Marker= MagicMark XP (Thermo Fisher).

5.3.1.2 *Overexpression of LEF-1 in non-translocator cell lines*

To study the role of LEF-1 in non-translocator cell lines (THP-1 and U937), cells were lentivirally transduced with a LEF-1 overexpression vector, pLV-EGFP:T2A:Puro-EF1A>hLEF1. The cells were grown, transduced and fractionated into cytosolic and nuclear fractions (Chapter 0). A control vector was included in all experiments (pLV-EGFP-T2A-Puro-EF1A). To determine if the overexpression of LEF-1 in non-translocator cell lines was successful, control and LEF-1 transduced cells nuclear and cytosolic fractions were immunoblotted for LEF-1 (2.5). BIO was used to stabilize β-catenin prior to fractionation.

In U937 cells a band was seen in the cytoplasmic fraction at ~45kDa, corresponding to full length LEF-1. There were additional bands at lower molecular weights, suggesting LEF-1 degradation in the cytoplasmic fraction of U937 cells. LEF-1 was not observed in the nucleus of U937 cells (**Figure 5.2B**). The lack of LEF-1 banding in the nuclear fraction could be due to a strong regulatory mechanism preventing its translocation to the nucleus. Additionally, incubation with BIO led to stabilization of LEF-1 in the cytoplasmic fraction of U937 cells.

In THP-1 cells, no bands were observed in the cytoplasmic or nuclear fractions following incubation with LEF-1 antibody. In contrast, in THP-1 cells induced with BIO, bands were observed in the nuclear fraction. (**Figure 5.2B**). These bands were lower than those expected for full length LEF-1 suggesting that following BIO induction, LEF-1 translocates to the nucleus of THP-1 cells and is subsequently degraded. This reaffirms the observations made in U937 cells and suggests that LEF-1 could be regulated by degradation in non-translocator cell lines. The banding observed in THP-1 cells is suggestive of a mechanism by which LEF-1 enters the nucleus but is subsequently degraded.

To summarize, LEF-1 overexpression was observed in both non-translocator lines following induction with BIO. This suggests that regulation of LEF-1 protein levels in non-translocator cell lines is complex, involving transcriptional regulation, post transcriptional regulation by BIO and regulation by subcellular localization, as observed in U937 cells.

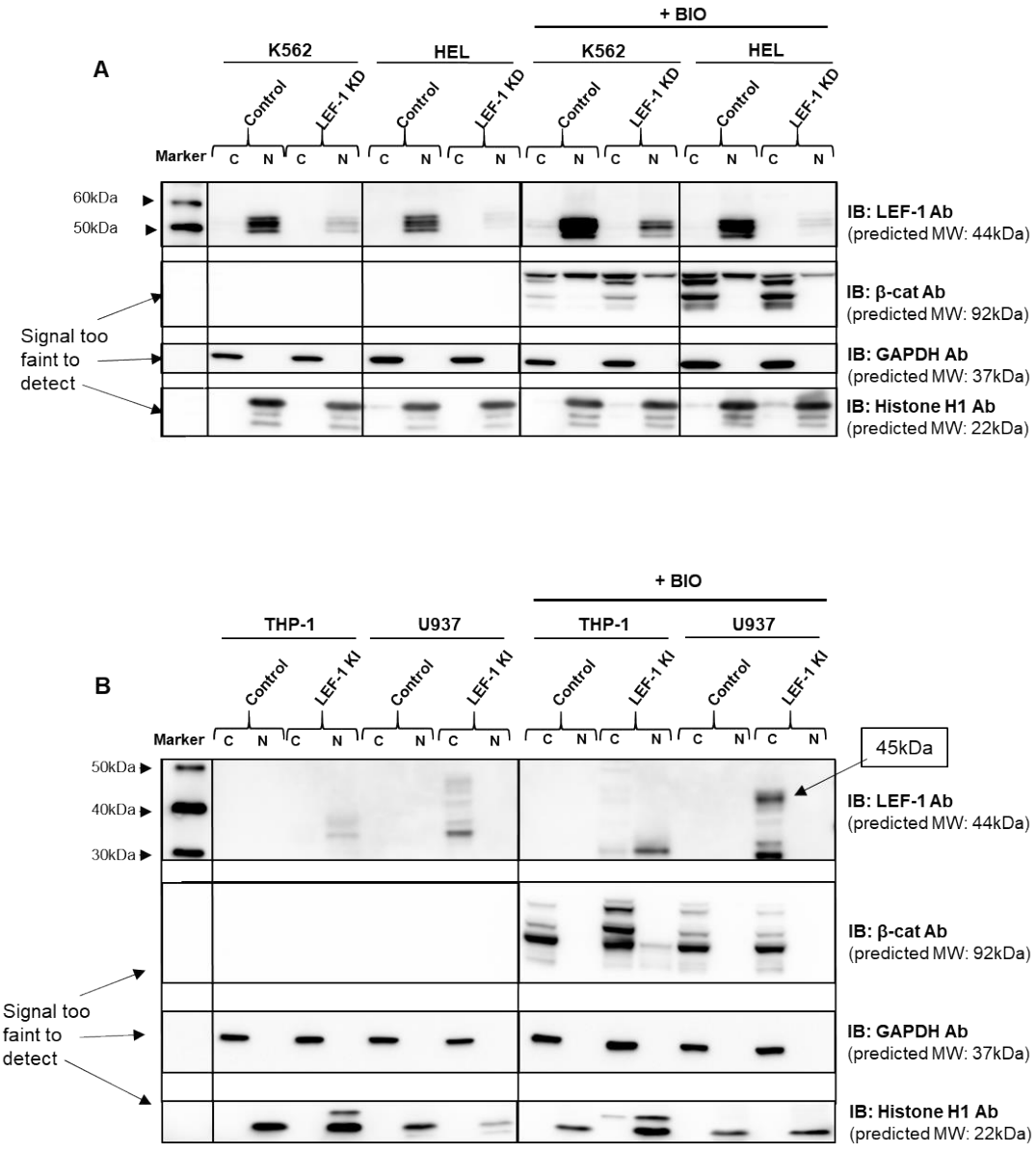


Figure 5.2. The effect of LEF-1 on the nuclear localization β-catenin in leukaemia cell lines.

Western blotting of cytoplasmic/nuclear fractions in (**A**) translocator cell lines lentivirally transduced with LEF-1 shRNA or corresponding shRNA control and (**B**) non-translocator cell lines lentivirally transduced with a LEF-1 overexpression vector and corresponding control (**Table 2.3**). In the lanes without BIO induction, the β-catenin lanes appear empty due to the relative exposure (BIO gives a very strong signal). MW markers were not visible on re-probed blots, but the relative migration of detected bands was consistent with the indicated proteins. Abbreviations; C= cytoplasmic fraction, N= nuclear fraction, KD= knockdown, KI= knock in, IB= immunoblot, MW= molecular weight, Ab= antibody, kDa= kilodalton. Marker= MagicMark XP (Thermo Fisher). BIO was added to stabilize β-catenin. GAPDH and Histone H1 antibodies were used as loading and fractionation efficiency controls.

5.3.1.3 *β-catenin nuclear localization is regulated by LEF-1*

To examine the role of LEF-1 in regulating nuclear localization of β-catenin, LEF-1 knockdown and overexpression cell lines were fractionated and western blotted with β-catenin antibody (2.5). shRNA and overexpression control lines were included and BIO was added to stabilize β-catenin, thereby increasing the amount of β-catenin available for nuclear translocation.

In translocator cell lines (K562 and HEL) induced with BIO, knockdown of LEF-1 reduced the level of nuclear β-catenin (**Figure 5.2A**) by 52% in K562 cells and 60% in HEL cells. This suggests that LEF-1 promotes the nuclear localization of β-catenin in these cell lines, and reducing the amount of LEF-1 lowers the amount of nuclear β-catenin.

In non-translocator cell lines (U937 and THP-1) the overexpression of LEF-1 and associated changes in β-catenin localization were more complicated (**Figure 5.2B**). In THP-1 cells, overexpression of LEF-1 resulted in a weak β-catenin signal in the nucleus of induced cells (+BIO) not observed in the induced control line. This is consistent with the weak nuclear expression of LEF-1 in this line and suggests that LEF-1 may play a role in regulating the nuclear localization of β-catenin in THP-1 cells. In contrast, in the U937 line, no β-catenin signal was observed in the nuclear fraction of LEF-1 overexpressing cells. Given the absence of LEF-1 in the nucleus of LEF-1 overexpressing U937 cells this was expected.

In conclusion, LEF-1 knockdown and overexpression studies suggest LEF-1 plays a role in regulating the nuclear localization of β-catenin, however, the contrasting results in THP-1 and U937 cells suggest the presence of multiple mechanisms to exclude LEF-1 (and correspondingly, β-catenin) from the nucleus.

5.3.2 *LEF-1 regulates β-catenin mediated transcription*

Following the confirmation that LEF-1 regulates the nuclear localization of β-catenin in leukaemia cell lines the next step was to examine whether subsequent changes in β-catenin mediated transcription occurred. Flow cytometry was used to measure β-catenin mediated transcription (visualised by Venus GFP reporter (2.4).

5.3.2.1 *LEF-1 positively regulates β-catenin mediated transcription in BIO induced leukaemia cell lines*

To study β-catenin mediated transcription, BIO was used to stabilize β-catenin and induce β-catenin mediated nuclear translocation and subsequent transcription in leukaemia cell lines.

In translocator cell lines induction of Wnt signalling by BIO resulted in an increase in reporter activity representing an increase in β-catenin mediated transcription (**Figure 5.3A**). In K562 control cells, adding BIO resulted in ~40% of cells expressing a level of reporter activity above the defined threshold value. Knockdown of LEF-1 reduced this to only 10% of cells. The threshold value is defined by the background reporter readout (from uninduced cells) hence the percentage induction represents the proportion of cells exhibiting higher reporter activity than that observed in uninduced cells. In HEL control cells, adding BIO resulted in ~52% of cells expressing a level of reporter activity above the defined threshold value. Knockdown of LEF-1 reduced this to only 2% of cells. These decreases in reporter activity in translocator cell lines are indicative of decreased β-catenin mediated transcription upon knockdown of LEF-1.

In LEF-1 overexpressing cell lines the study of β-catenin mediated transcription was complicated by the presence of GFP in the LEF-1 overexpression vector which interfered with the readout of the Venus reporter. The effect of GFP in the overexpression vector on the reporter readout can be observed in **Figure 5.4A**. This is based on the transduction of cell lines with the control LEF-1 overexpression vector (expressing the background GFP). Any additional reporter activity was taken as a readout of β-catenin mediated transcription. In non-translocator cell lines treatment with BIO did not result in reporter readout above

the defined threshold value (**Figure 5.4B&C**). This is because β-catenin does not translocate to the nucleus in these lines and is not able to regulate the reporter gene. Small increases observed in these parental lines are likely due to the fluorescence of BIO itself.

In THP-1 cells LEF-1 overexpression resulted in an increase in reporter readout following BIO induction (**Figure 5.4B**). In THP-1 control cells, adding BIO resulted in ~5% of cells expressing a level of reporter activity above the defined threshold value. Overexpression of LEF-1 increased this to 42% of cells. This suggests that there may be an increase in β-catenin mediated transcription resulting from overexpression of LEF-1 in this cell line. Together with the western blot data (**Figure 5.2B**) this supports the hypothesis that LEF-1 promotes nuclear localization of β-catenin and subsequent transcriptional activation in THP-1 cells.

In U937 cells, LEF-1 overexpression and BIO induction only resulted in 5% of cells expressing reporter activity above the defined threshold value (**Figure 5.4B**). This is supported by the western blot data (**Figure 5.2B**) which suggests LEF-1 (and β-catenin) does not translocate to the nucleus of U937 cells and subsequently cannot regulate β-catenin mediated transcription.

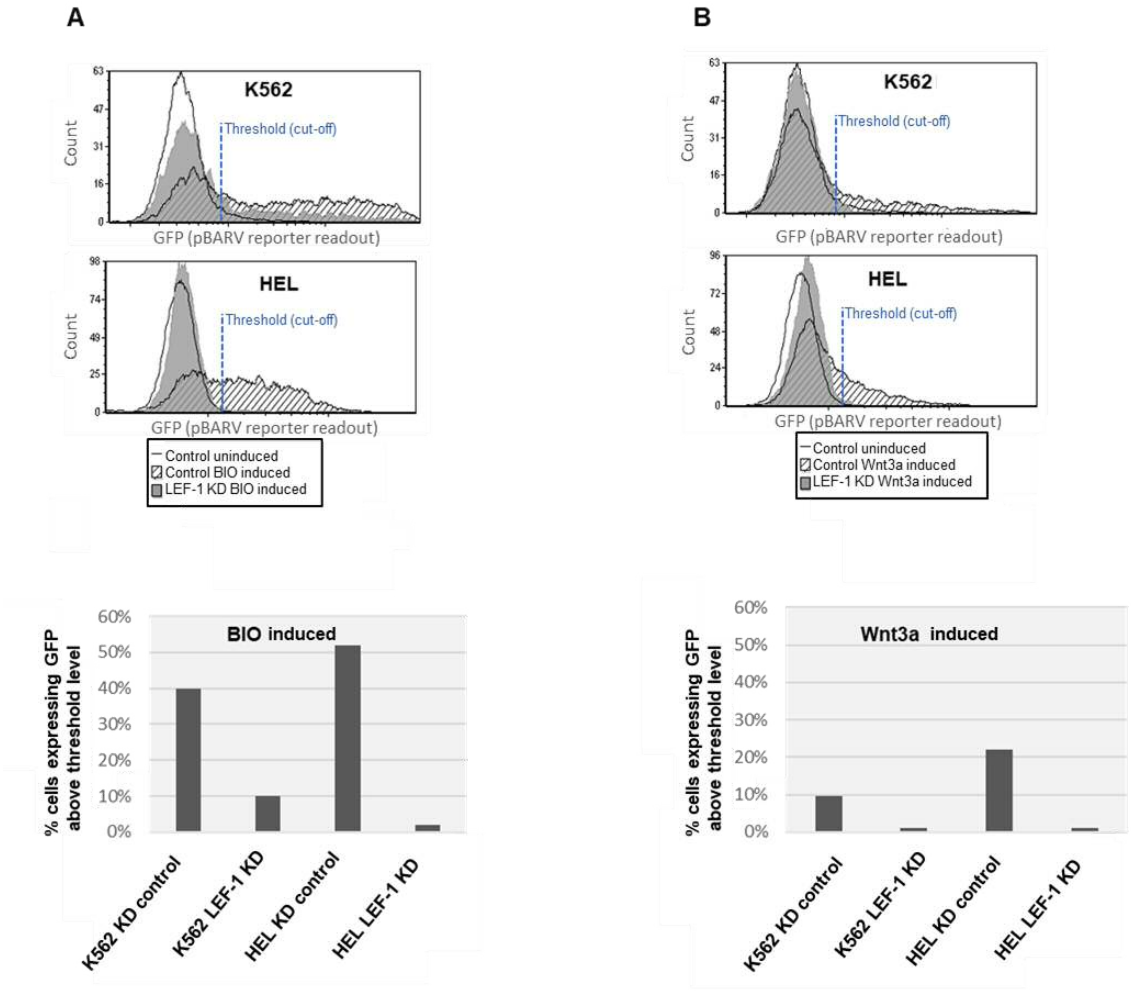


Figure 5.3. The effect of LEF-1 on β-catenin mediated transcription in translocator cell lines.

Top: Flow cytometric analysis of translocator lines expressing the pBARV reporter to measure β-catenin mediated transcription in cells transduced with LEF-1 shRNA or shRNA control (**Table 2.3**). Wnt signalling was induced with (**A**) BIO and (**B**) Wnt3a. Bottom: Bar charts representing the percent of cells expressing reporter activity above the defined threshold value (relative to uninduced) with addition of (**A**) BIO and (**B**) Wnt3a.

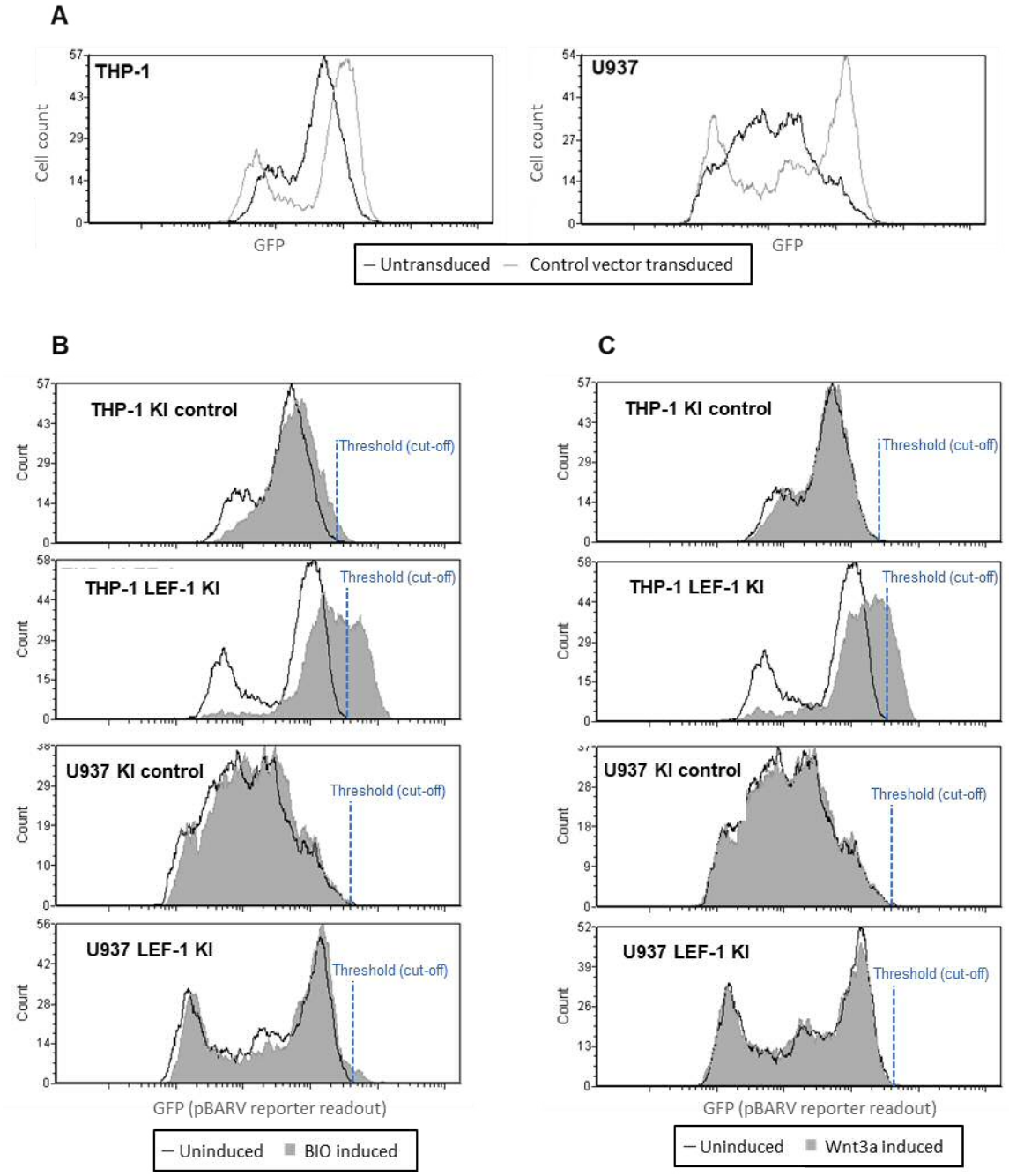


Figure 5.4. β-catenin mediated transcription in non-translocator cell lines in response to LEF-1 overexpression.

(**A**) Flow cytometric analysis of GFP expression in parental and control infected cells. (**B**) Flow Analysis of β-catenin mediated transcription in pBARV lines in non-translocator cell lines -/+ BIO and Wnt3a. Abbreviation; KI= knock-in (LEF-1 overexpression).

5.3.2.2 *LEF-1 positively regulates β-catenin mediated transcription in Wnt3a induced leukaemia cell lines*

In addition to using BIO, Wnt3a was included as another agonist to stimulate Wnt signalling and study β-catenin mediated transcription. Wnt3a is a canonical Wnt pathway component and was included in the analysis because it represents a more biologically relevant system than the artificial GSK3β inhibitor BIO. Ideally Wnt3a would have been used to study the effect of LEF-1 on β-catenin mediated localization via western blotting (5.3.1), however, the amount of Wnt3a required for western blot analysis made this prohibitively expensive.

In translocator cell lines (K562 and HEL) induced with Wnt3a, knockdown of LEF-1 resulted in a decrease in reporter activity. In K562 control cells, adding Wnt3a resulted in ~9.6% of cells expressing a level of reporter activity above the defined threshold value. Knockdown of LEF-1 decreased this to 1% of cells. In HEL control cells, adding Wnt3a resulted in 22% of cells expressing a level of reporter activity above the defined threshold value. Knockdown of LEF-1 decreased this to 1% of cells. These reporter readouts suggest there was a reduction in β-catenin mediated transcription following LEF-1 knockdown (**Figure 5.3B**). This supports the data observed in BIO induced cells that LEF-1 regulates β-catenin mediated transcription (5.3.2.1).

In non-translocator cell lines (THP-1 and U937) induced with Wnt3a, β-catenin mediated transcription reporter readout followed the same pattern as with BIO induction (5.3.2.1). In THP-1 LEF-1 overexpressing cells, adding Wnt3a resulted in ~19% of cells expressing a level of reporter activity above the defined threshold value (**Figure 5.4C**). In U937 LEF-1 overexpressing cells no change in reporter readout was observed with Wnt3a induction (**Figure 5.4C**). Again, this can be explained by the absence of overexpressed LEF-1 in the nucleus of U937 cells (**Figure 5.2B**).

In conclusion, knockdown of LEF-1 in translocator cell lines decreased β-catenin mediated transcription in cells induced with BIO and Wnt3a. In addition, induction of LEF-1 overexpressing THP-1 cells resulted in an increase in reporter activity corresponding to an increase in β-catenin mediated transcription. In U937 cells this was not the case, owing to the absence of LEF-1 in the nucleus. Together

with the western blot results (**Figure 5.2A & B**) this data suggests that LEF-1 regulates both the nuclear localization of β-catenin and subsequent β-catenin-mediated transcription.

Ideally repeating this experiment with another LEF-1 vector (without GFP) would be necessary to negate possible confusion with GFP present in the overexpression vector and strengthen these observations.

5.3.3 *LEF-1 knockdown may decrease proliferation of translocator cell lines*

Wnt signalling regulates the cell cycle and proliferation in many contexts. To establish whether LEF-1 regulates proliferation of leukaemia cell lines, a proliferation assay was used to study cell lines with LEF-1 knockdown and overexpression (2.4.3). Briefly, cells were grown in different concentrations of serum to identify any differences in growth rate and serum dependence and counted over a 3-day period using flow cytometry (**Figure 5.5, Figure 5.6** & **Figure 5.7**). Proliferation was measured as a fold change from seeding density.

5.3.3.1 *The effect of LEF-1 knockdown on proliferation of K562 cells*

In K562 cells (**Figure 5.5A**) on days 1-2 the overall trend was that LEF-1 knockdown led to a decreased proliferation rate, particularly at lower serum concentrations. On day 1 there was a significant difference in the proliferation rate between control (~1-fold increase) and LEF-1 knockdown (no change) at 0.3% serum. Similarly, at 3% serum there was a significant difference between control (~1.5-fold increase) and LEF-1 knockdown (~1-fold increase). On day 2 at 0.3% serum there was a significant difference between control (~4-fold increase) and LEF-1 knockdown (~2-fold increase). At 3% serum, there was a significant difference between control (~4.5-fold increase) and LEF-1 knockdown (~3-fold increase).

On day 3 the differences in proliferation rate did not follow the trend. At serum concentrations 0.1%, 0.3% and 3% there was no difference in proliferation rate between control and LEF-1 cells.

5.3.3.2 *The effect of LEF-1 knockdown on the proliferation of HEL cells*

In HEL cells (**Figure 5.5B**) on day 1 there was no difference in the proliferation of control and LEF-1 knockdown cell lines at all serum concentrations. On days 2-3 the general trend was decreased proliferation in LEF-1 knockdown cells compared to control. On day 2 at 0.3% serum there was a significant difference between control (~3-fold increase) and LEF-1 knockdown (~2-fold increase). At 3% serum, there was a significant difference between control (~5-fold increase) and LEF-1 knockdown (~2.5-fold increase).

On day 3 there was a significant difference in proliferation between control (~5.5-fold increase) and LEF-1 knockdown (~2-fold increase) at 0.3% serum. At 3% serum, there was a significant difference between control (~6.5-fold increase) and LEF-1 knockdown (~3.5-fold increase) cells. At 10% serum, there was a significant difference between control (~6.5-fold increase) and LEF-1 knockdown (~4-fold increase) cells. This experiment suggests that LEF-1 may regulate proliferation in HEL cells as knocking down LEF-1 resulted in reduced proliferation.

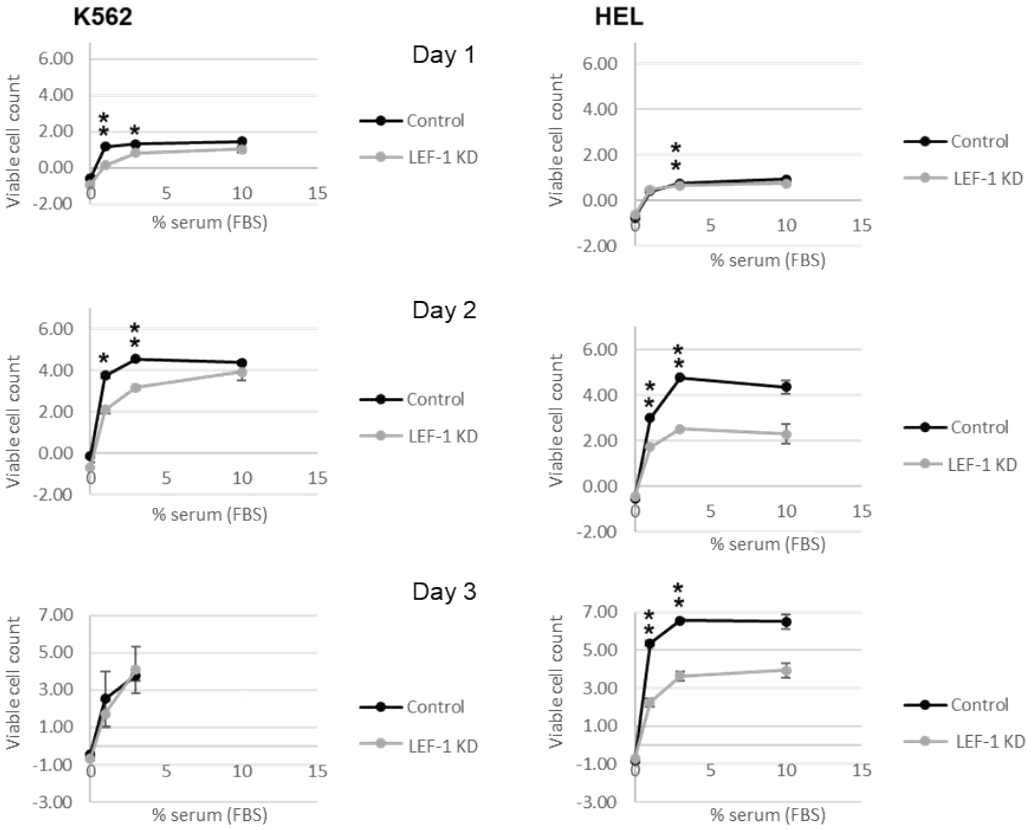


Figure 5.5. Effect of LEF-1 knockdown on proliferation of translocator cells.

Flow cytometric analysis of cell proliferation in control shRNA and LEF-1 shRNA K562 and HEL cells, on day 1 (top), day 2 (middle) and day 3 (bottom). Each data point corresponds to the average cell count from two replicates. The assay was repeated twice, but the second experiment failed. * p=<0.05, ** p=<0.01 (two sample T test. n=2). The starting density of cells was $0.8x10^5$/ml for K562 control, $1x10^5$/ml for K562 LEF-1 KD, $0.8x10^5$/ml for HEL control and $1x10^5$/ml for HEL LEF-1 KD.

5.3.4 *LEF-1 overexpression does not impact the proliferation of non-translocator cell lines*

5.3.4.1 *The effect of LEF-1 overexpression on the proliferation of THP-1 cells*

In both repeat experiments LEF-1 overexpression did not impact the proliferation of THP-1 cells. As with the translocator cell lines, there was a general trend for increased proliferation with increasing serum concentration up to 10% (normal growth conditions). In the first experiment (**Figure 5.6A**), on days 1-2 there were some small significant differences between control and LEF-1 transduced cells at certain serum concentrations (Day 1; 3 % serum. Day 2; 0.1% serum and 3% serum). This data does not provide enough evidence to support the hypothesis that LEF-1 regulates proliferation in THP-1 cells due to the lack of any trend and the very small differences observed.

In the second experiment (**Figure 5.6B**) there was no difference in proliferation between control and LEF-1 overexpressing cells on any days or serum concentrations. Given the issues observed in the second set of repeats for HEL and K562 cells (data not shown), there may be issues with this set of results for THP-1 cells as well.

To conclude, both repeat experiments suggest that LEF-1 does not play a role in regulating the proliferation of THP-1 cells as there were no differences in fold change between control and LEF-1 overexpressing cells.

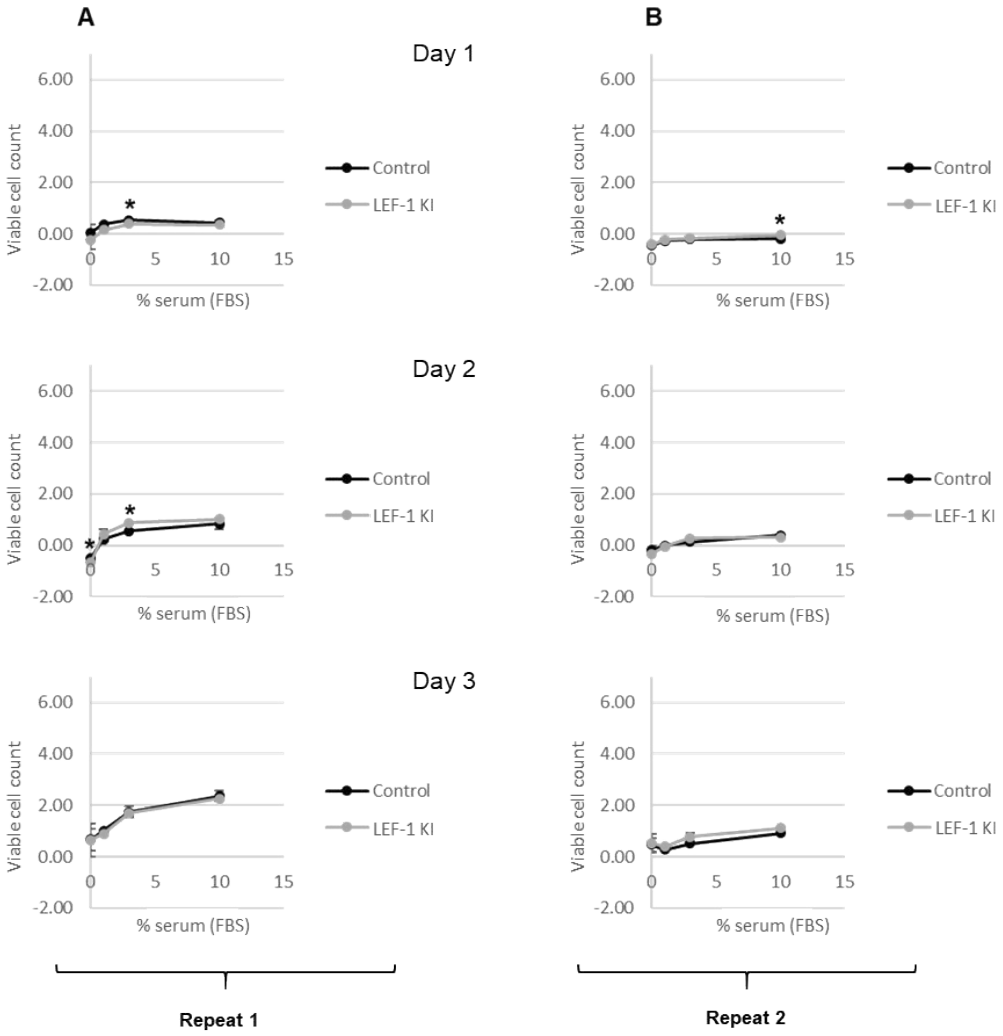


Figure 5.6. Effect of overexpression on proliferation of THP-1 cells.

Flow cytometric analysis of cell proliferation in control and LEF-1 overexpressing THP-1 cells on day 1 (top), day 2 (middle) and day 3 (bottom). Each data point corresponds to the average cell count from two replicates. The assay was repeated twice(A&B). * p=<0.05 (Two sample T test. n=2). The starting densities of cells in (A) were $2.7x10^5$/ml for THP-1 control and $2.9x10^5$/ml for THP-1 LEF-1 KI. The starting densities of cells in (B) were $2.5x10^5$/ml for THP-1 control and THP-1 LEF-1 KI.

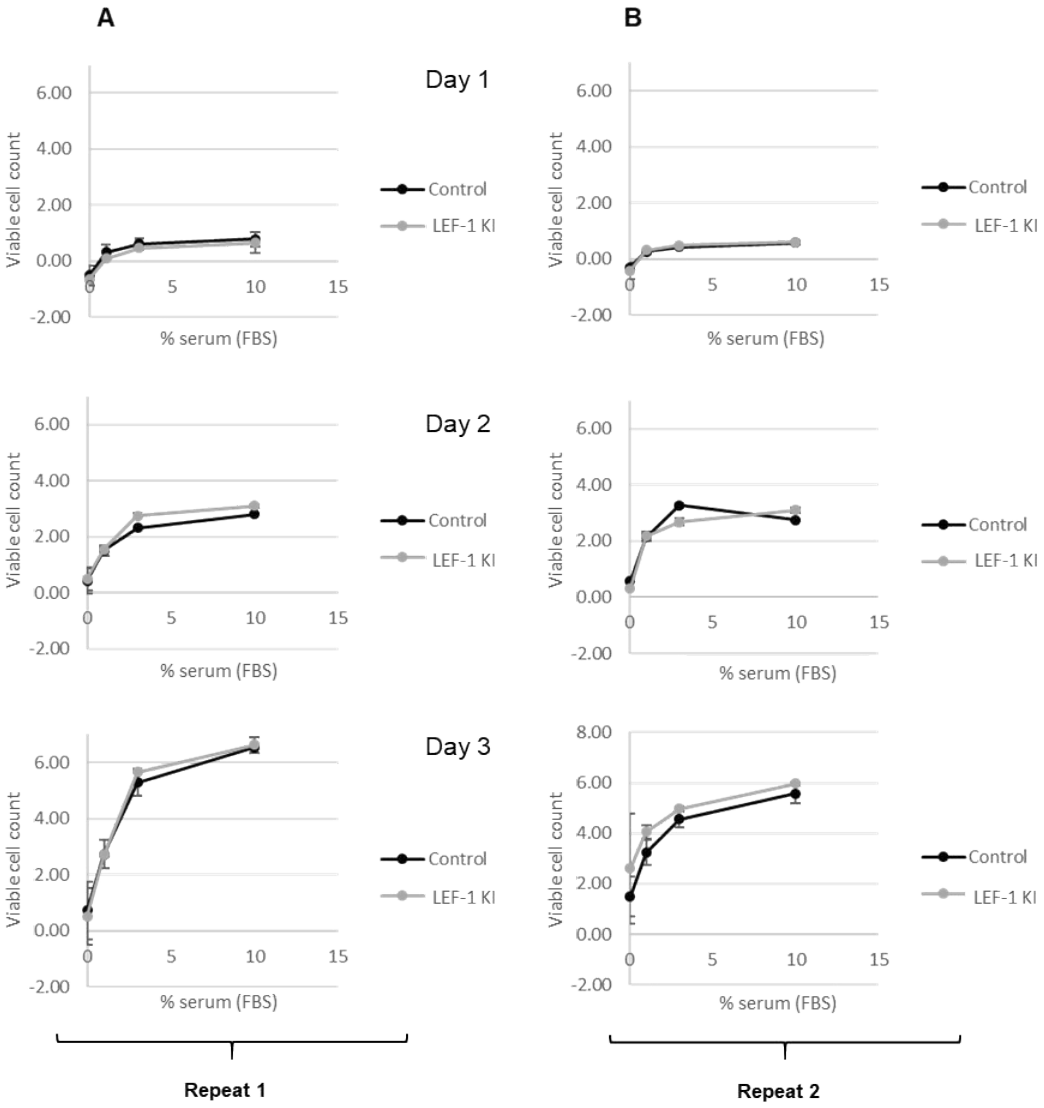


Figure 5.7. Effect of overexpression on proliferation of U937 cells.

Flow cytometric analysis of cell proliferation in control and LEF-1 overexpressing U937 cells. Each data point corresponds to the average cell count from two replicates. The assay was repeated twice(**A**&**B**). The starting densities of cells in (**A**) were $1.4x10^5$/ml for U937 control and $1.3x10^5$/ml for U937 LEF-1 KI. The starting densities of cells in (**B**) were $1.1x10^5$/ml for U937 control and U937 LEF-1 KI.

5.3.4.2 *The effect of LEF-1 overexpression on the proliferation of U937 cells*

In U937 cells the effect of serum concentration was the same as the other cell lines with increasing proliferation up to serum concentration of 10%. In both repeat experiments LEF-1 overexpression did not impact the proliferation of U937 cells (**Figure 5.7A &B**). There are no significant differences between control and LEF-1 overexpressing cells on any day or at any serum concentration.

To conclude, LEF-1 does not impact the proliferation of U937 cells. Taken together with the results observed in THP-1 cells this suggests that LEF-1 does not regulate the proliferation of non-translocator cells.

5.3.5 ***Overexpression or knockdown of LEF-1 does not impact migration of leukaemia cell lines in vitro.***

Migration plays an important role in cancer progression and β-catenin has been implicated in the regulation of migration in different contexts including epithelial cells, solid tumours, and macrophages (Amini-Nik *et al*, 2014; Müller *et al*, 2002; Yang *et al*, 2017). To determine if overexpression or knockdown of LEF-1 impacts the migration capabilities of leukaemia cell lines a migration assay was used (2.4.4). Briefly, a trans-well assay was set up in which cells were loaded into the upper chamber and a chemoattractant (SDF-1) was added to the lower chamber. Flow cytometry was used to measure the number of cells that migrated from the upper to lower chamber. **Figure 5.8** shows the percentage of cells that migrated through to the bottom chamber for each cell line and condition. To validate the assay $CD34^+$ cells from human cord blood were included as a positive control. The migration of each cell line was also measured without the addition of SDF-1 to provide an indication of background migration. If LEF-1 regulates the migration of leukaemia cell lines, we would expect to see a reduction in the migration of cells in translocator cell lines with LEF-1 KD and an increase in migration of cells in non-translocator cell lines with overexpression of LEF-1.

In both translocator cell lines (HEL and K562) LEF-1 knockdown had no effect on the migration of cells from the upper chamber to the lower chamber in this trans-well assay. The lack of response to SDF-1 in these cells prior to LEF-1

knockdown means that the effect of LEF-1 knockdown on migration could not be scored.

Both non-translocator cell lines responded to SDF-1 in this assay, with migration of cells into the lower chamber, despite undetectable LEF-1. This suggests that LEF-1 expression is not required for the migratory response. Overexpression of LEF-1 in non-translocator cell lines gave conflicting results. In the THP-1 cell line a reduction in migration was observed with overexpression of LEF-1 whereas in the U937 cells the opposite was observed. The contrasting results could be due to experimental error or due to mechanistic differences in each of the cell lines. The conflicting results do suggest the absence of a global mechanism regulating LEF-1 mediated changes in migration, with cell line specific mediators likely to be involved.

Given the lack of consistent effects of LEF-1 overexpression on migration in non-translocator cell lines and the lack of response in translocator cell lines, further studies on migration were not pursued and the experiment was not repeated.

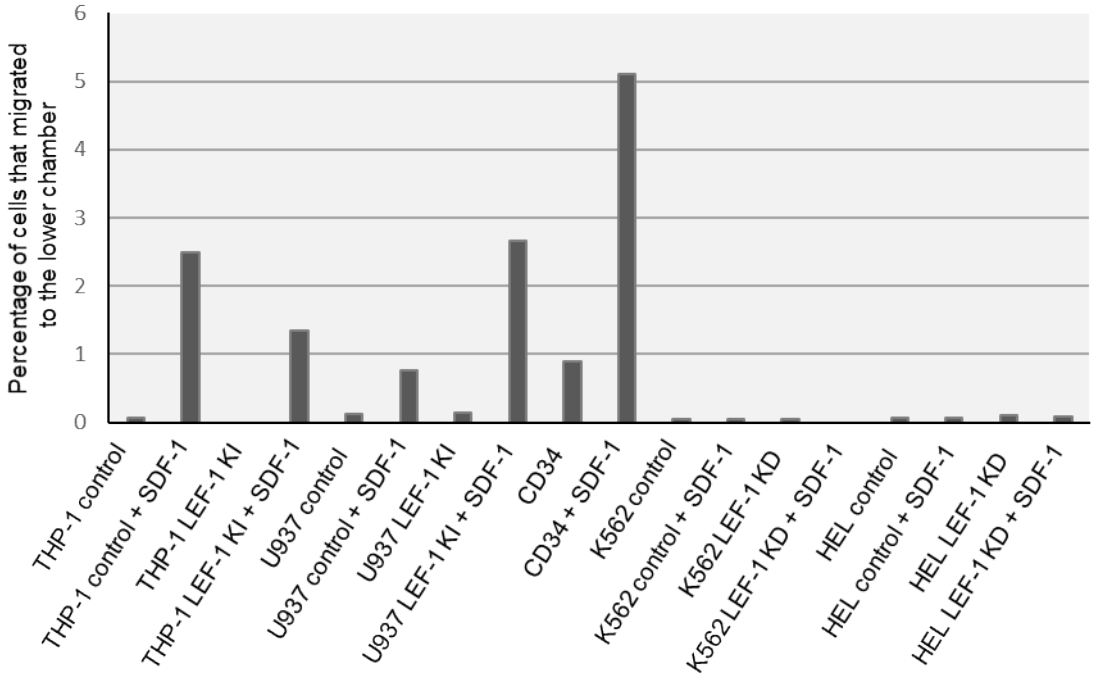


Figure 5.8. Effect of LEF-1 on the migration of leukaemia cell lines.

Migration data for leukaemia cell lines transduced with control, LEF-1 shRNA (KD) or LEF-1 overexpression vectors (KI) (**Table 2.3**). SDF-1 was used as a chemoattractant in the lower well to stimulate migration across the cell permeable membrane. Data is presented as a percentage of total cells that migrated through the membrane. CD34 cells were included as a positive control. Cells were counted using flow cytometry (2.4).

5.3.6 *LEF-1 levels do not impact cell survival in cell leukaemia lines*

Another hallmark of leukaemia is the ability of cancer cells to resist apoptosis. To establish whether changes in LEF-1 levels have an impact on cell survival in leukaemia cell lines a viability assay was used (2.4.3). Control and LEF-1 overexpression/knockdown cell lines were stained with 7-AAD and cells were analysed by flow cytometry to measure viability (2.4). 7-AAD is a cell viability stain that works by binding to the DNA of non-viable cells. In viable cells the stain cannot readily cross the intact cell membrane. As with the proliferation assays (5.3.3 & 5.3.4) the viability assay was conducted under a range of serum concentrations (0.1%, 0.3%, 3% and 10%) to determine if LEF-1 regulates apoptosis under conditions of stress (low serum concentration).

5.3.6.1 *Survival of translocator cell lines following knockdown of LEF-1*

In K562 cells (**Figure 5.9**) there was a general trend of decreased survival with decreasing serum concentrations on all 3 days. Knockdown of LEF-1 did not impact cell survival in K562 cells. The only significant difference in cell viability between control and LEF-1 knockdown lines was on day 1 at 3% serum. Cell viability was ~63% in control cells and ~70% in LEF-1 knockdown cells. Given the lack of significant differences between control and LEF-1 knockdown cells at any other data point, this is likely to be an anomaly. This data suggests that LEF-1 does not regulate cell survival in K562 cells, however, ideally this data would be repeated with a Wnt agonist such as Wnt3a.

In HEL cells (**Figure 5.9**) there was a trend of decreased survival with decreasing serum concentrations. There appeared to be an additional small negative effect on survival with knockdown of LEF-1. There were only significant differences between control and knockdown cells at the lower serum concentrations on days 1 and 3. On Day 1 there was a significant difference in cell viability of ~7% at 0.3% serum. On Day 2 there were significant differences in cell viability of ~7% at 0.1% serum and ~4% at 0.3% serum. This data suggests that LEF-1 could play a role in regulating the survival of HEL cells at low serum concentrations, however, owing to the very small differences observed, repeat experiments would be necessary to strengthen this hypothesis.

Taken together, the results of the viability assays in translocator cell lines (K562 and HEL) provide evidence that LEF-1 may regulate cell survival to some extent in HEL cells but that this is not the case in K562 cells. This suggests that if LEF-1 does regulate the survival of HEL cells, this is not a global mechanism in leukaemia cell lines.

5.3.6.2 *Effect of LEF-1 overexpression on the survival of non-translocator cell lines*

In THP-1 cells, cell viability decreased slightly with decreasing serum concentrations (**Figure 5.10**). Cell viability remained constant between 60-75% for all 3 days in both control and LEF-1 overexpressing cells. There were no significant differences in cell viability between the control and LEF-1 knockdown cells.

Similarly, in U937 cells, cell viability decreased with decreasing serum concentrations (**Figure 5.10**). Viability remained constant for the 3 days and there were no significant differences in cell viability between control and LEF-1 overexpressing cells.

The results of the viability assays in non-translocator cells suggest that overexpressing LEF-1 does not impact cell survival in these cell lines, however, as discussed in 5.4.5.1, overexpression of LEF-1 was not effective in these cell lines making it difficult to draw conclusions from this data.

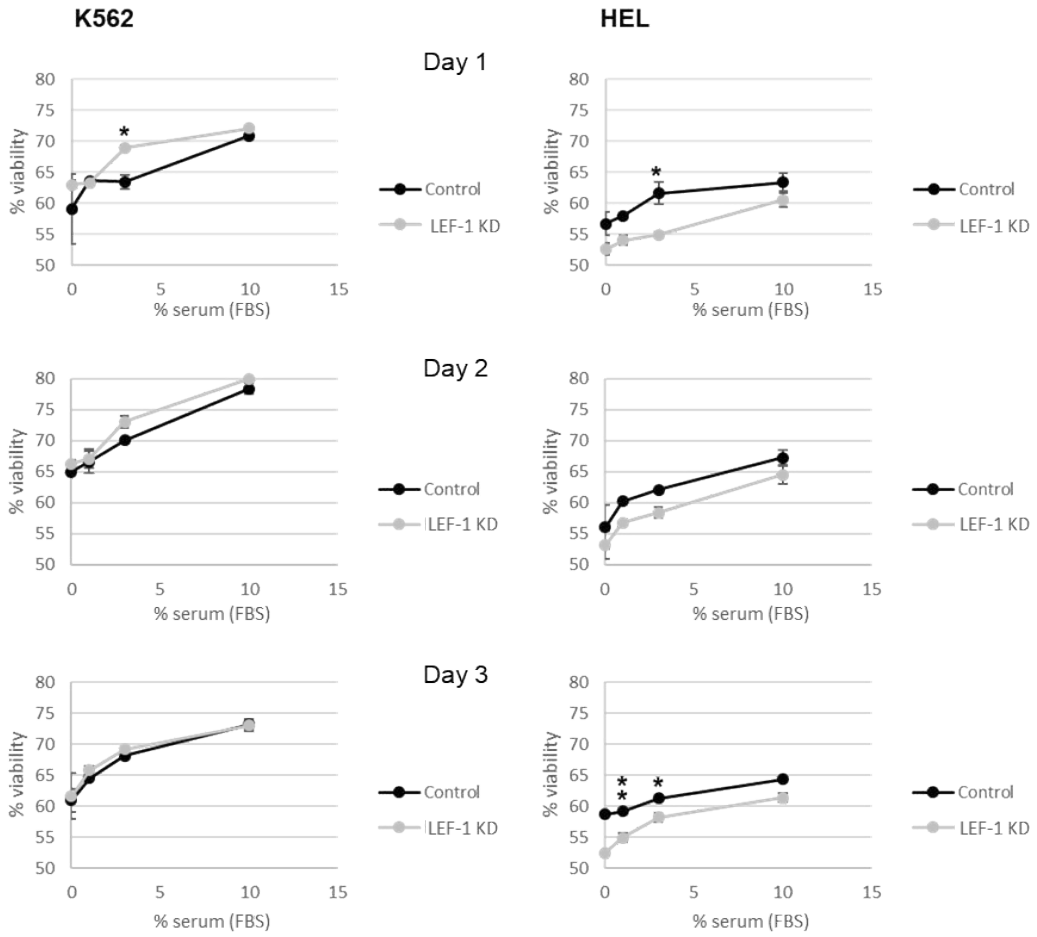


Figure 5.9. Effects of LEF-1 knockdown on survival in translocator lines.

Flow cytometric analysis of cell viability in translocator cell lines. Viability was calculated as a percentage of total cells. Each data point represents the average of two replicate experiments. * p=<0.05, ** p=<0.01 (Two sample T test. n=2). The starting cell densities were $1.1x10^5$/ml for K562 control, $1.2x10^5$/ml for K562 LEF-1 KD, $1.2x10^5$/ml for HEL control and $1.2x10^5$/ml for HEL LEF-1 KD.

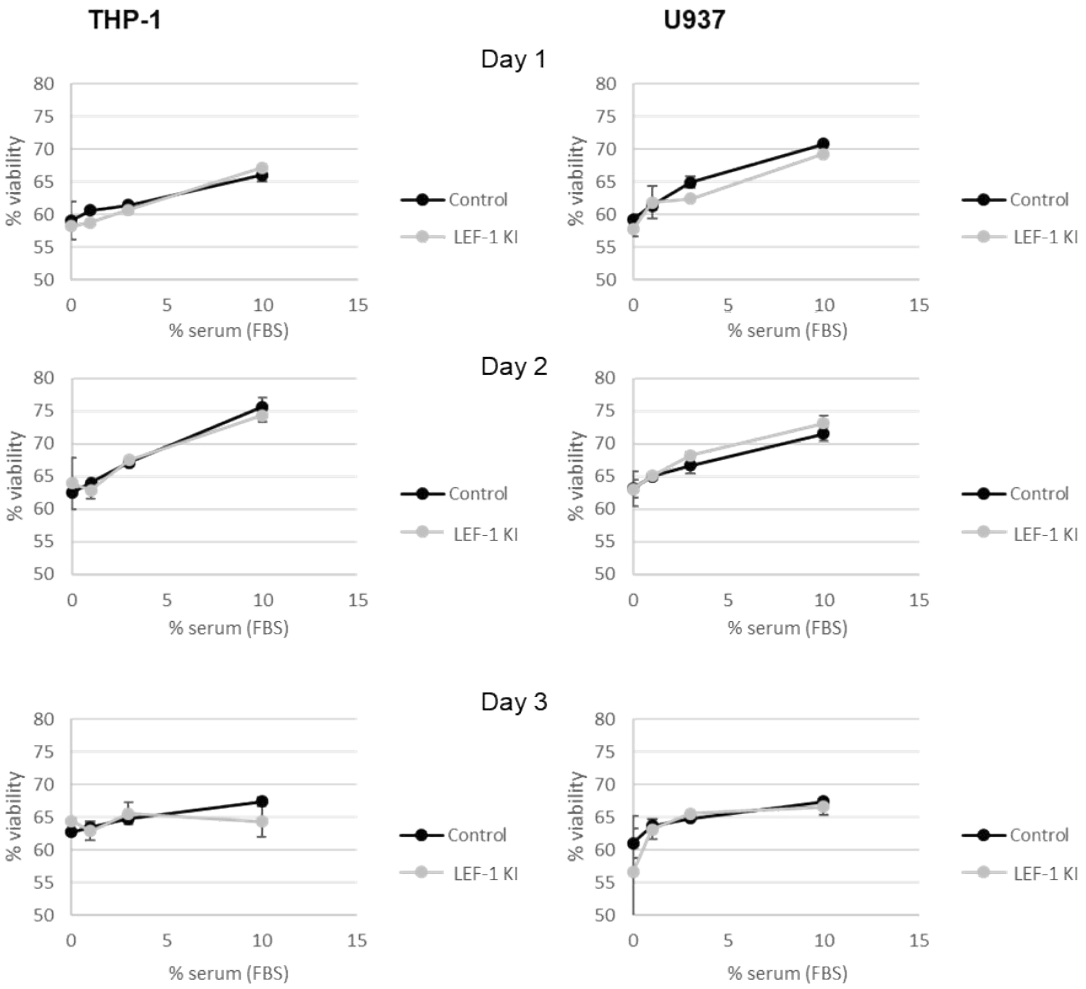


Figure 5.10. Effects of LEF-1 overexpression on survival in non-translocator lines.

Flow cytometric analysis of cell viability in non- translocator cell lines. Viability was calculated as a percentage of total cells. Each data point represents the average of two replicate experiments. The starting cell densities for THP-1 control and THP-1 LEF-1 KI were $2.5x10^5$/ml. The starting cell densities for U937 control and U937 LEF-1 KI were $1.1x10^5$/ml.

5.3.7 *LEF-1 overexpression changes the light scattering properties of THP-1 cells*

Wnt signalling is involved in regulating the differentiation of cells in different contexts, including dendritic cells, embryonic stem cells and HSCs (Davidson *et al*, 2012; Famili *et al*, 2016; Xu *et al*, 2016). One of the characteristics of leukaemia is a block in normal differentiation and an accumulation of immature blasts. To examine the potential role of LEF-1 in changing the morphology and/or differentiation of leukaemia cell lines, flow cytometry was used to analyse the scatter profiles of control and LEF-1 overexpressing/knockdown lines (2.4). In three of the cell lines (K562, HEL and U937) there was no difference in the scatter profiles of control or LEF-1 knockdown/overexpression cell lines (**Figure 5.11A & B**). In the non-translocator cell line THP-1 there was a difference in the scatter profile of LEF-1 overexpressing cells compared to control cells **(Figure 5.11A**). In LEF-1 overexpressing cells there was an increase in forward scatter in 11.4% of cells. This increased to 18.2% and 32.5% in Wnt3a and BIO induced cells. This shift in forward scatter suggests these cells are larger than the control cells. The side scatter profiles of the THP-1 control and LEF-1 overexpressing cells were the same. This indicates that LEF-1 does not impact the granularity of the cells. This data suggests that LEF-1 impacts the morphology of THP-1 cells but due to time constraints this was only repeated once and further experiments would be necessary to determine if this is a significant change and whether it is indicative of differences in differentiation.

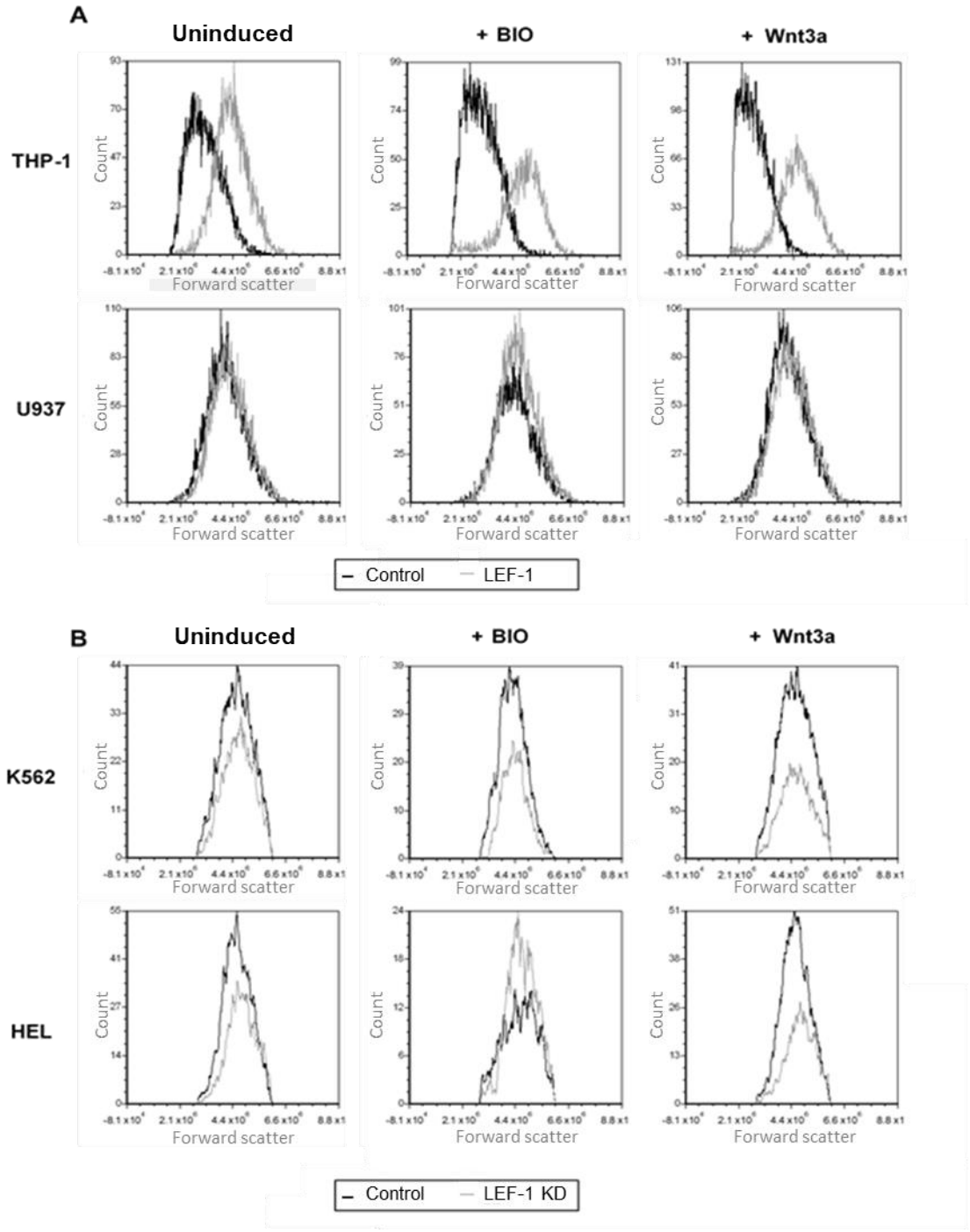


Figure 5.11. Forward scatter analysis of leukaemia cell lines following knockdown and overexpression of LEF-1.

Flow cytometric measurement of cell size using forward scatter profiles in (**A**) non-translocator cells lines transduced with control or LEF-1 overexpression vectors and (**B**) translocator cell lines transduced with control or LEF-1 shRNA vectors. BIO and Wnt3a were used to induce Wnt signalling.

5.4 Discussion

LEF-1 is frequently dysregulated in AML and its overexpression has clinical significance, with LEF-1 overexpression being associated with a favourable prognosis (Fu *et al*, 2014; Metzeler *et al*, 2012). It has been shown to regulate the nuclear localization of β-catenin in different contexts, including, SW480, NIH-3T3, MDCK cell lines and mouse embryos (Henderson *et al*, 2002; Huber *et al*, 1996; Jamieson *et al*, 2016; Jamieson *et al*, 2011; Simcha *et al*, 1998), and has a well-established role in regulating β-catenin mediated transcription as a member of the TCF family of proteins (1.3.1.2). The aim of the experiments presented in this chapter was to establish whether LEF-1 regulates β-catenin nuclear localization in leukaemia cell lines, and to assess the functional relevance of any findings. In this chapter, LEF-1 was knocked down in translocator cell lines and overexpressed in non-translocator cell lines. This was an attempt to "convert" translocator cells into non-translocator cells and vice versa. One thing that was not considered when conducting these experiments was overexpressing LEF-1 in translocator cell lines. This may have identified LEF-1 as a factor that regulates the nuclear localization of β-catenin in translocator cell lines without addition of BIO. This would be more functionally relevant than the artificial induction of β-catenin nuclear localization with BIO and is something that would be interesting to look at in future work.

This section discusses aspects of the experimental techniques used in this chapter, as well as the conclusions made about the role of LEF-1 in mediating β-catenin nuclear localization, β-catenin mediated transcription, proliferation, cell survival and cell migration.

5.4.1 *LEF-1 regulates the nuclear localization of β-catenin in leukaemia cell lines*

Little is known about the factors that regulate β-catenin nuclear localization in AML, and this study aimed to identify factors that regulate this process. In this chapter, the candidate β-catenin nuclear localization factor LEF-1 was studied. LEF-1 was identified as a candidate nuclear localization factor by MS in Chapter 4 and western blotting results (**Figure 5.2**) support this data, suggesting that LEF-

1 regulates the nuclear localization of β-catenin in leukaemia cell lines. This agrees with observations in other contexts, as discussed in 4.4.2.8.

Although LEF-1 has been shown to regulate β-catenin nuclear levels, it does not appear to be essential for β-catenin nuclear localization. In one study, binding of LEF-1 to β-catenin was disrupted in Cos-1 cells by mutating β-catenin. In these cells, nuclear localization of β-catenin was still observed, suggesting that, at least in this context, LEF-1 is not essential for its nuclear localization (Prieve & Waterman, 1999). It is difficult to say whether my results agree with this observation, because although knockdown of LEF-1 was 77-78% efficient in translocator lines, some LEF-1 was still detectable in the nuclei of these cells and so could have been sufficient to localize β-catenin to the nucleus. Recent data shows a strong correlation between nuclear LEF-1 and nuclear β-catenin levels in AML. In 23 AML patients (mainly paediatric) the relative percent nuclear translocation of LEF-1 and β-catenin strongly correlates (and is highly significant). This supports LEF1 as a key translocation partner in AML (Morgan unpublished; 2017).

In summary, LEF-1 has been shown to regulate the nuclear localization of β-catenin in different contexts. Most research into LEF-1 mediated β-catenin nuclear localization in cancer has focused on solid tumours (**Table 1.10**) and little is known about the factors that regulate β-catenin nuclear localization in leukaemia. This study has identified LEF-1 as a regulator of β-catenin nuclear localization in leukaemia cells. It would be interesting to determine if LEF-1 knockdown also impacts the levels of nuclear β-catenin in AML blasts with abnormal nuclear β-catenin localization, which occurs in ~50% of AML patients (Xu *et al*, 2008). This would be particularly interesting given the recent data in support of LEF-1 as a key regulator of β-catenin nuclear localization in AML patients (Morgan unpublished; 2017).

5.4.2 *LEF-1 regulates β-catenin mediated transcription in leukaemia cell lines*

As a central Wnt signalling pathway component, LEF-1 has a well-established role in regulating β-catenin (TCF/LEF) mediated transcription (1.3.1.2). In a haematopoietic context, normal $CD34^+$ cells overexpressing LEF-1 have increased activation of Wnt target gene expression, including *cyclin D1*, *c-myc* and *survivin* (Skokowa *et al*, 2006).

LEF-1 is not essential for Wnt target gene transcription. Wnt target genes are regulated by other redundant factors, such as its close family member TCF-1. In one study, LEF-1, TCF-1 and LEF-1/TCF-1 knock outs were conducted in mice and the effects of these knockouts on the differentiation of thymocytes was observed. Only the double knock out of LEF-1 and TCF-1 affected the transcription of the target genes RAG-1, TCRα and TCRβ, leading to a complete block in T cell differentiation (Okamura *et al*, 1998). This suggests that knockdown of LEF-1 alone is not sufficient to impact target gene expression in this context. Additionally, β-catenin redundant factors, such as γ-catenin, can regulate TCF/LEF mediated transcription. In a study by our laboratory, overexpression of γ-catenin led to increased TCF/LEF mediated transcription. This increase was dampened by β-catenin knockdown, but was not reversed (Morgan *et al*, 2013).

Due to the existence of these redundant factors, it was important to verify that the reduction in nuclear β-catenin levels observed in LEF-1 knockdown cells led to a decrease in TCF/LEF mediated transcription. The results in **Figure 5.3** show that knockdown of LEF-1 in translocator cell lines was accompanied by a reduction in TCF/LEF mediated transcription. This suggests that LEF-1 regulates β-catenin mediated transcription in myeloid leukaemia cell lines, which is consistent with the results observed in $CD34^+$ cells (Skokowa *et al*, 2006). Since β-catenin mediated transcription was not completely abolished by LEF-1 knockdown in K562 and HEL cells, this suggests that in agreement with the findings made in thymocytes (Okamura *et al,* 1998), LEF-1 redundant factors are likely to regulate β-catenin mediated transcription in myeloid leukaemia cells.

5.4.3 *LEF-1 knockdown impacts the proliferation of leukaemia cell lines*

One of the hallmarks of AML is uncontrolled proliferation leading to the accumulation of blast cells. The Wnt pathway effector β-catenin and TCF factors have well established roles in regulating proliferation by upregulating gene expression of target genes such as *cyclinD1* and *c-myc* (He *et al*, 1998; Shtutman *et al*, 1999). In normal haematopoiesis LEF-1 has been shown to regulate proliferation. Knockdown of LEF-1 in murine pre-B cells led to reduced proliferation (Reya *et al*, 2000), although the effects of LEF-1 knockdown on β-catenin mediated transcription were not established, so LEF-1 may have been acting in a β-catenin independent manner in this study. Knockdown of LEF-1 in normal $CD34^+$ cells led to a 50% reduction in proliferation, measured using BrDU staining (Skokowa *et al*, 2006). Interestingly, in the same assay knockdown of β-catenin did not impact the proliferation of $CD34^+$ cells, suggesting that in this context LEF-1 was acting in a β-catenin independent manner.

Dysregulation of LEF-1/β-catenin mediated proliferation is associated with cancer. In one study, elevated levels of β-catenin in colon cancer cells and its binding to LEF-1 led to activation of the *cyclinD1* gene and uncontrolled entry into the cell cycle (Shtutman *et al*, 1999). Similarly, in acute lymphoblastic leukaemia (ALL), LEF-1 is associated with increased *cyclinD1* and *c-myc* mRNA levels and increased proliferation of ALL cells (Guo *et al*, 2015).

This project used a proliferation assay (2.4.3) to determine if LEF-1 mediated nuclear localization of β-catenin and β-catenin mediated transcription affected proliferation in LEF-1 knockdown translocator cell lines. The results in **Figure 5.5** suggest that knockdown of LEF-1, and subsequent reduction in β-catenin nuclear localization and β-catenin mediated transcription, may lead to a reduction in proliferation. This agrees with another study in which the proliferation of leukaemia cell lines was measured following expression of dominant negative β-catenin (Chung *et al*, 2002). In this study, proliferation of a subset of leukaemia cell lines was regulated by β-catenin (Jurkat, HUT-102 and K562), but proliferation of HL-60 cells was not. In another study, knockdown of β-catenin in HL-60 cells and AML blasts decreased their proliferation *in vitro* (Siapati *et al*, 2011).

To conclude, my results suggest that LEF-1/β-catenin may regulate the proliferation of leukaemia cell lines and may play a role in regulating proliferation in AML, and this agrees with other studies looking at the role of β-catenin in regulating proliferation in haematopoietic and leukemic contexts. Owing to the failure of a repeat assay, however, this conclusion is only based on one experiment in two cell lines. It would be interesting to repeat this experiment on other translocator cell lines and on β-catenin translocating AML blasts to determine if knockdown of LEF-1 impacts the proliferation of these cells.

5.4.4 *LEF-1 does not impact the survival or migration of leukaemia lines*

Another hallmark of leukaemogenesis is evasion of apoptosis, which contributes to the accumulation of blast cells. The canonical Wnt pathway has a well-established role in regulating survival by upregulating gene expression of target genes such as *survivin* (Zhang *et al*, 2001). In a haematopoietic context, normal $CD34^+$ progenitor cells overexpressing β-catenin have reduced apoptosis, suggesting that β-catenin regulates cell survival in these cells (Simon *et al*, 2005). In addition, knockdown of LEF-1 has a pro apoptotic effect on $CD34^+$ cells (Skokowa *et al*, 2006).

The results in **Figure 5.9** show that knockdown of LEF-1 did not significantly impact the survival of translocator cell lines, although there was a small reduction in survival of HEL cells with LEF-1 knockdown, which concurs with the observation that in HEL cells there was a greater reduction in proliferation. This contrasts with experiments in chronic lymphocytic leukaemia in which knockdown of LEF-1 by siRNA and inhibition of LEF-1/β-catenin binding by small molecule inhibitors led to increased apoptosis of these blasts (Gandhirajan *et al*, 2010; Gutierrez *et al*, 2010), suggesting that canonical Wnt signalling regulates cell survival in these cells. This suggests that there are differences in the regulation of LEF-1 mediated Wnt signalling in myeloid and lymphoid leukaemia. It could also be the case that the low level of β-catenin/LEF-1 that was detectable in the nucleus of knockdown cells (**Figure 5.2**) in this project was enough to maintain cell survival, or that redundant factors such as γ-catenin and TCF-1 are sufficient to regulate Wnt target genes that mediate cell survival.

Another process that contributes to leukaemogenesis is cell migration. β-catenin mediated signalling has been shown to regulate cell migration in different contexts including in epithelial cells and solid tumours (Müller *et al*, 2002; Yang *et al*, 2017). Macrophage specific deletion of β-catenin leads to impaired migration and cell adhesion (Amini-Nik *et al*, 2014).

Despite the regulation of migration by Wnt signalling in other contexts, the results in **Figure 5.8** do not show any differences in cell migration following knockdown of LEF-1 and the subsequent reduction in nuclear β-catenin and β-catenin mediated transcription in translocator cell lines. This may suggest that LEF-1/β-catenin does not regulate cell migration in AML, or alternatively, the lack of migration observed in my assay could be because these cell lines do not respond to SDF-1. The lack of response to SDF-1 in the translocator lines without LEF-1 knockdown meant that the migration endpoint for LEF-1 knockdown cells could not be measured. This meant that only the data for LEF-1 overexpression in non-translocator cell lines could be used. The lack of response in translocator cell lines could be due to the fact they lack the receptor for SDF-1 (the chemoattractant used in the migration assay). The receptor for SDF-1, CXCR4, is expressed in low levels in HEL cells and is almost undetectable in K562 cells[22]. Both THP-1 and U937 cells have a comparatively higher expression of CXCR4 which could explain why more of these cells migrated in my assay. It would be interesting to express CXCR4 in K562/HEL cells with i) endogenous LEF-1 and ii) knockdown of LEF-1 and repeat this assay to determine if this is the case.

5.4.5 ***Experimental limitations***

5.4.5.1 *Overexpression of LEF-1 was not achieved in non-translocator cell lines*

One factor that limited the conclusions that could be drawn from the analysis of LEF-1 in this project was that overexpression of LEF-1 was ineffective in non-translocator cell lines. In other studies, successful overexpression of LEF-1 was achieved in murine HSCs and normal human $CD34^+$ cells (Petropoulos *et al*, 2008; Skokowa *et al*, 2006). This suggests that despite the unsuccessful

[22] https://www.proteinatlas.org/ENSG00000121966-CXCR4/cell (accessed on 04/01/2018)

overexpression of LEF-1 in this project, it can be overexpressed in a haematopoietic context. Perhaps LEF-1 overexpression is tolerated in HSCs and early progenitor cells but not in myeloid cells. Indeed, in normal mouse myeloid cells, the expression of *LEF-1* mRNA is low[23], suggesting that LEF-1 levels are tightly controlled in these cells. In one study, RT-PCR was used to measure *LEF-1* mRNA levels in different murine bone marrow fractions (Sakhinia *et al*, 2006). In this study, *LEF-1* expression was highest in Lin^- cells, suggesting that *LEF-1* expression is associated with HSCs and early progenitors in normal haematopoiesis. *LEF-1* was not detected in the Lin^+ cell fraction. Next, they analysed single cell BM precursors to determine the levels of *LEF-1* mRNA expression in more differentiated haematopoietic precursors. *LEF-1* was undetectable in these more differentiated cells, except for mast cell and B cell precursors. This suggests that LEF-1 is primarily expressed in HSCs and early progenitors, whereas in more differentiated haematopoietic cells LEF-1 expression is maintained at low levels. This could explain why other studies have managed to overexpress LEF-1 in HSCs and $CD34^+$ cells yet I was unsuccessful in overexpressing LEF-1 in the non-translocator cells used in this study. It could be that the non-translocator cells have retained mechanisms that maintain low LEF-1 levels in normal myeloid cells. The presence of bands in western blots of non-translocator cells at lower molecular weights than would be expected for LEF-1 (**Figure 5.2B**), suggests that this mechanism could be degradation. Although LEF-1 overexpression was unsuccessful in this study, it is observed in AML, is associated with leukaemogenesis (Petropoulos *et al*, 2008), and has clinical significance for patient prognosis and survival (Fu *et al*, 2014; Metzeler *et al*, 2012).

Due to the ineffective overexpression of LEF-1 in this study, the conclusions based on LEF-1 overexpression are not discussed further.

[23] http://servers.binf.ku.dk/bloodspot/?gene=Lef1&dataset=mouse_nl_rna_seq (accessed on 04/01/2018)

5.4.5.2 *Western blotting detected bands at higher molecular weights than the predicted weight of full length LEF-1*

The predicted weight of full length LEF-1 is 45kDa, yet in translocator cell lines, western blotting detected a cluster of bands at ~55kDa, suggesting that endogenous LEF-1 is larger in these cells than expected. It is unlikely that these bands are due to non-specific binding of the antibody to another factor, because bands between 50-60kDa are regularly observed in other contexts using different LEF-1 antibodies (Jamieson *et al*, 2016; Lambertini *et al*, 2009; Warsito *et al*, 2012; Wu *et al*, 2014)

It could be that post-translational modifications are contributing to the 10kDa larger size of LEF-1 than is predicted based purely on its amino acid sequence. One potential candidate for modification of LEF-1, that could account for a 10kDa increase in molecular weight is the SUMO protein (Hilgarth & Sarge, 2005). SUMOylation of LEF-1 has been observed in other contexts, and has been implicated in regulating its localization to nuclear bodies, repressing β-catenin mediated transcriptional activity (Sachdev *et al*, 2001). Other post translational modifications may also be responsible for the detection of bands at larger molecular weights. It would be interesting to determine which post translational modifications of LEF-1, including SUMOylation, occur in translocator cells.

In summary, LEF-1 knockdown studies suggested that LEF-1 regulates the nuclear localization of β-catenin and β-catenin nuclear localization in myeloid leukaemia cell lines, however, due to the experimental limitations outlined above, it was difficult to draw conclusions as to the effects of this on proliferation, survival and migration of these cells.

6 General Discussion

6.1 Summary and conclusions

β-catenin is the principle canonical Wnt pathway effector that regulates the transcription of genes involved in key cellular processes, including proliferation and cell survival. Canonical Wnt signalling has been identified as one of the most commonly dysregulated signalling pathways in AML and β-catenin overexpression is associated with a poor prognosis (Daud, 2014; Majeti *et al*, 2009a). Nuclear localization of β-catenin is key to its role as a transcription factor, and although multiple β-catenin nuclear localization factors have been identified in other contexts, this process is not well understood in AML (Morgan *et al*, 2014). In this study, cytoplasmic/nuclear fractionation and immunoprecipitation of β-catenin was optimized in leukaemia cell lines, and candidate β-catenin nuclear localization factors were identified by immunoprecipitation and MS analysis.

Initially, in Chapter 3, cytoplasmic/nuclear fractionation and β-catenin immunoprecipitation was optimized in K562 cells. Immunoprecipitation with agarose G beads was unsuccessful despite testing of various factors that could be impacting the immunoprecipitation efficiency. Alternatively, a different immunoprecipitation approach was tested, using protein G Dynabeads. This was successfully used to immunoprecipitate β-catenin in the leukaemia cell lines tested and to generate ample material for MS analysis. Although this was sufficient for my analysis, it might be useful to improve the immunoprecipitation efficiency further if this was to be repeated on patient material, which would be more limiting than the cell lines used in this study. Additionally, when optimizing the fractionation of cell lines and immunoprecipitation of β-catenin, it may have been better to focus on different detergent conditions and analyse their effects on co-immunoprecipitation efficiency prior to MS analysis, for example by determining their effects on interactions between β-catenin and known binding partners. Comparing more than one immunoprecipitation protocol earlier in the optimization steps could also have saved more time trying to optimize immunoprecipitation using agarose beads. This may have allowed more time for thorough functional analysis of LEF-1 in Chapter 5.

In Chapter 4, data analysis of MS results from β-catenin immunoprecipitations was conducted. Peptides present in the cytoplasmic and nuclear fractions of translocator and non-translocator cell lines were compared. This identified eight β-catenin interaction partners as potential candidate β-catenin nuclear localization factors. Western blotting of these factors validated RUNX1 and LEF-1 as candidates for further analysis. Due to time constraints and technical difficulties, not all the candidates could be validated by western blotting (4.4) and of the validated factors, LEF-1 was favoured for further analysis over RUNX1, owing to its well established role in canonical Wnt signalling (1.3.1.2) and its identification as a candidate β-catenin nuclear localization factor in other contexts (Henderson *et al*, 2002; Huber *et al*, 1996; Jamieson *et al*, 2016; Jamieson *et al*, 2011; Simcha *et al*, 1998). There is some evidence to suggest that RUNX1 is involved in the regulation of canonical Wnt signalling in other contexts. For example, in estrogen receptor positive breast cancer cells, RUNX1 binds to Axin1 and prevents estrogen receptor mediated Axin1 suppression (Chimge *et al*, 2016). RUNX1 loss in these cells leads to stabilization of β-catenin, suggesting that RUNX1 is a negative regulator of canonical Wnt signalling in this context. RUNX1 also plays a role in regulating Wnt signalling in AML as part of AML associated fusion proteins. In one study, expression of RUNX1-ETO in K562 cells led to a 12 fold increase in catenin/TCF mediated transcription, measured using a TOPFLASH reporter (Yeh *et al*, 2009). In a separate study, expression of RUNX1-ETO in U937 cells led to increased γ-catenin mRNA expression, increased γ-catenin and β-catenin protein levels and association of γ-catenin with the promoter region of the Wnt target gene *c-myc* (Müller-Tidow *et al*, 2004). Taking these data into account, it would be interesting to further establish the role of RUNX1 in regulating β-catenin nuclear localization.

In Chapter 5, functional analysis of LEF-1 mediated β-catenin nuclear localization was conducted. Knockdown and overexpression studies showed that LEF-1 can regulate the nuclear localization of β-catenin and subsequent β-catenin mediated transcription in leukaemia cell lines. This did not have significant effects on proliferation, survival or migration of leukaemia cells, though, this analysis was in part hindered by ineffective LEF-1 overexpression in non-translocator cell lines (5.4.5.1).

6.2 Experimental Approach

The details of the techniques used in this project are discussed in 3.4, 4.4 and 5.4. This section discusses the general approach taken to identify candidate β-catenin nuclear localization factors and its limitations.

One of the main limitations of the MS approach used in this project is that it was based on identifying proteins that were either present or absent in each cell fraction, rather than the relative levels of proteins. This does not consider the possibility that the level of certain β-catenin interaction partners may determine their effects on β-catenin nuclear localization. Initially, a quantitative proteomics approach was considered for analysis of β-catenin interaction partners in this study, but given the number of samples required this was not feasible. In the original experimental design two translocator lines and two non-translocator lines were going to be used, which would result in a total of six samples (cytoplasmic and nuclear translocator cell line fractions and cytoplasmic non-translocator cell line fractions). In addition to this, IgG samples for each would need to be included, bringing the total number of samples to 12. This was too many samples for use in quantitative SILAC or ITRAQ based MS (the two most commonly used quantitative approaches).

One way that a quantitative approach could still have been used for analysis was if I reduced the number of samples used, for example, by only comparing one translocator cell line and one non-translocator cell line. The decision was made to include more cell lines and use a non-quantitative approach rather than use a quantitative approach with only two cell lines. This was because it made it easier to identify candidate β-catenin nuclear localization factors that play a global role in leukaemia, rather than cell line specific factors which could have less relevance for AML.

Another experimental technique that was not included in this project was co-immunoprecipitation of candidate factors and β-catenin. Ideally, following identification of candidate factors, rather than just western blotting of the samples sent for MS, a repeat β-catenin immunoprecipitation for each cell line/compartment would have been conducted and then western blotting using candidate factor antibodies. In addition, immunoprecipitation of candidate factors

would have been conducted followed by western blotting with β-catenin antibody. This would give more confidence in the identification of candidate factors as β-catenin interaction partners, particularly for factors that have not been identified as candidate β-catenin interaction partners previously (4.4). The reason this was not done was because this was not considered a priority in this case, as LEF-1 was already a known β-catenin interacting protein.

6.3 Further work

In this project I found that LEF-1 mediates the nuclear localization of β-catenin in leukaemia cell lines. LEF-1 has previously been identified as a β-catenin nuclear localization factor in other contexts and its expression has been associated with AML (Fu *et al*, 2014; Metzeler *et al*, 2012; Müller-Tidow *et al*, 2004; Petropoulos *et al*, 2008). It would be interesting to determine if LEF-1 expression correlates with nuclear β-catenin in AML blasts, which would indicate that LEF-1 may be regulating nuclear localization of β-catenin in AML patients. In addition, if there was a correlation between LEF-1 expression and β-catenin nuclear localization, it would be interesting to determine if this is also associated with specific molecular abnormalities in AML. It may be particularly interesting to look at MLL abnormalities in cell lines and AML blasts, given the proposed role that β-catenin has in the establishment of LSCs in MLL AML (Yeung *et al*, 2010). Many of the cell lines used in the study of AML have MLL abnormalities, which seem to confer immortalisation, making them useful for *in vitro* culture (Drexler *et al*, 2004). It would be interesting to determine if there are certain MLL abnormalities that correlate with LEF-1 expression and β-catenin nuclear localization in these lines. There is evidence to suggest that β-catenin nuclear localization and LEF-1 expression do not always correlate in AML. For example, certain molecular abnormalities such as FLT3-ITD have been shown to promote the nuclear localization of β-catenin and increase β-catenin mediated transcription. In contrast, FLT3-ITD has been associated with low LEF-1 expression, suggesting that this factor does not impact LEF-1 mediated β-catenin nuclear localization (Kajiguchi *et al*, 2007; Metzeler *et al*, 2012; Tickenbrock *et al*, 2005). This suggests that there are additional mechanisms regulating the nuclear localization of β-catenin in AML.

It would be interesting to get a more complete picture of the factors regulating the nuclear localization of β-catenin by i) validating more of the identified candidate factors (**Table 4.5**) by western blotting, ii) perform functional analysis of these factors to determine their role in regulating β-catenin nuclear localization and β-catenin mediated transcription as well as any consequence for proliferation, cell survival and migration in leukaemia cell lines and iii) determine if these factors also colocalize in primary AML. RUNX1 would be an interesting one to start with given its role as a haematopoietic transcription factor and its frequent dysregulation in AML (1.2.2).

This project focused on *in vitro* studies of leukaemia cell lines, which is limited as a model system. As discussed in 1.1.2, external factors in the niche are important in regulating cell signalling. *In vitro* studies are biased, with only selected agonists being used. Although this can make it easier to study a very specific mechanism, it oversimplifies complex signalling networks and can result in conclusions that are not functionally relevant *in vivo*. It would be important to study the role of LEF-1 and other candidate factors *in vivo* to determine their effects on β-catenin nuclear localization, β-catenin mediated transcription and leukemogenesis.

It would also be interesting to repeat the immunoprecipitation and MS analysis using different fractionation/immunoprecipitation buffers to potentially identify more candidate β-catenin nuclear localization factors. The buffers used to isolate proteins and their binding partners can have a big impact on the interaction partners that are identified. Some buffers can interrupt the binding of certain proteins, and based on some of the issues encountered in this project (3.4 and 4.4), repeating this experiment with different buffers could identify other candidate β-catenin binding partners that were missing from the analysis using CHAPS buffer.

It would also be useful to determine if LEF-1 and other candidate nuclear β-catenin localization factors regulate the nuclear localization of γ-catenin. γ-catenin was not identified as a β-catenin binding factor in these cells, but it has been previously identified as a factor regulating the nuclear localization of β-catenin (Morgan *et al*, 2013). It would be interesting to see if there is a common mechanism regulating the nuclear localization of both catenins in leukaemia.

Finally, one of the main limiting factors in performing functional analysis of LEF-1 and β-catenin in this project was that these are not independent variables. It may be worthwhile to focus further studies on establishing the role of β-catenin itself outside its role in MLL AML. It's role in non MLL AML is still controversial (1.3.3), and knockdown or the use of β-catenin inhibitors may not be sufficient to establish this if only small amounts of β-catenin are required for its function, as has been suggested by (Luis *et al*, 2011). The role of β-catenin in AML could be reinvestigated using new technology such as CRISPR gene editing (Czarnek & Bereta, 2016) to see if AML blasts can survive a double knockout of β-catenin.

References

Abdel-Wahab, O., Adli, M., LaFave, L. M., Gao, J., Hricik, T., Shih, A. H., Pandey, S., Patel, J. P., Chung, Y. R., Koche, R., Perna, F., Zhao, X., Taylor, J. E., Park, C. Y., Carroll, M., Melnick, A., Nimer, S. D., Jaffe, J. D., Aifantis, I., Bernstein, B. E. & Levine, R. L. (2012) ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell*, 22(2), 180-93.

Aberle, H., Bauer, A., Stappert, J., Kispert, A. & Kemler, R. (1997) beta-catenin is a target for the ubiquitin-proteasome pathway. *EMBO J*, 16(13), 3797-804.

Abrahamsson, A. E., Geron, I., Gotlib, J., Dao, K. H., Barroga, C. F., Newton, I. G., Giles, F. J., Durocher, J., Creusot, R. S., Karimi, M., Jones, C., Zehnder, J. L., Keating, A., Negrin, R. S., Weissman, I. L. & Jamieson, C. H. (2009) Glycogen synthase kinase 3beta missplicing contributes to leukemia stem cell generation. *Proc Natl Acad Sci U S A*, 106(10), 3925-9.

Acar, M., Kocherlakota, K. S., Murphy, M. M., Peyer, J. G., Oguro, H., Inra, C. N., Jaiyeola, C., Zhao, Z., Luby-Phelps, K. & Morrison, S. J. (2015) Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. *Nature*, 526(7571), 126-30.

Alford, J. R., Kendrick, B. S., Carpenter, J. F. & Randolph, T. W. (2008) Measurement of the second osmotic virial coefficient for protein solutions exhibiting monomer-dimer equilibrium. *Anal Biochem*, 377(2), 128-33.

Amini-Nik, S., Cambridge, E., Yu, W., Guo, A., Whetstone, H., Nadesan, P., Poon, R., Hinz, B. & Alman, B. A. (2014) β-Catenin-regulated myeloid cell adhesion and migration determine wound healing. *J Clin Invest*, 124(6), 2599-610.

Andersson, A., Edén, P., Lindgren, D., Nilsson, J., Lassen, C., Heldrup, J., Fontes, M., Borg, A., Mitelman, F., Johansson, B., Höglund, M. & Fioretos, T. (2005) Gene expression profiling of leukemic cell lines reveals conserved molecular signatures among subtypes with specific genetic aberrations. *Leukemia*, 19(6), 1042-50.

Antonchuk, J., Sauvageau, G. & Humphries, R. K. (2002) HOXB4-induced expansion of adult hematopoietic stem cells ex vivo. *Cell*, 109(1), 39-45.

Appelbaum, F. R., Rowe, J. M., Radich, J. & Dick, J. E. (2001) Acute myeloid leukemia. *Hematology Am Soc Hematol Educ Program*, 62-86.

Arachea, B. T., Sun, Z., Potente, N., Malik, R., Isailovic, D. & Viola, R. E. (2012) Detergent selection for enhanced extraction of membrane proteins. *Protein Expr Purif*, 86(1), 12-20.

Arai, F., Hirao, A., Ohmura, M., Sato, H., Matsuoka, S., Takubo, K., Ito, K., Koh, G. Y. & Suda, T. (2004) Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche. *Cell*, 118(2), 149-61.

Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., Bloomfield, C. D., Cazzola, M. & Vardiman, J. W. (2016) The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*, 127(20), 2391-405.

Arnold, H. K., Zhang, X., Daniel, C. J., Tibbitts, D., Escamilla-Powers, J., Farrell, A., Tokarz, S., Morgan, C. & Sears, R. C. (2009) The Axin1 scaffold protein promotes formation of a degradation complex for c-Myc. *EMBO J*, 28(5), 500-12.

Austin, T. W., Solar, G. P., Ziegler, F. C., Liem, L. & Matthews, W. (1997) A role for the Wnt gene family in hematopoiesis: expansion of multilineage progenitor cells. *Blood*, 89(10), 3624-35.

Baba, Y., Garrett, K. P. & Kincade, P. W. (2005) Constitutively active beta-catenin confers multilineage differentiation potential on lymphoid and myeloid progenitors. *Immunity*, 23(6), 599-609.
Bacher, U., Haferlach, C., Kern, W., Haferlach, T. & Schnittger, S. (2008) Prognostic relevance of FLT3-TKD mutations in AML: the combination matters--an analysis of 3082 patients. *Blood*, 111(5), 2527-37.
Baldus, S. E., Mönig, S. P., Huxel, S., Landsberg, S., Hanisch, F. G., Engelmann, K., Schneider, P. M., Thiele, J., Hölscher, A. H. & Dienes, H. P. (2004) MUC1 and nuclear beta-catenin are coexpressed at the invasion front of colorectal carcinomas and are both correlated with tumor prognosis. *Clin Cancer Res*, 10(8), 2790-6.
Bao, R., Christova, T., Song, S., Angers, S., Yan, X. & Attisano, L. (2012) Inhibition of tankyrases induces Axin stabilization and blocks Wnt signalling in breast cancer cells. *PLoS One*, 7(11), e48670.
Behrens, J., von Kries, J. P., Kühl, M., Bruhn, L., Wedlich, D., Grosschedl, R. & Birchmeier, W. (1996) Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature*, 382(6592), 638-42.
Benedik, M. J. & Strych, U. (1998) Serratia marcescens and its extracellular nuclease. *FEMS Microbiol Lett*, 165(1), 1-13.
Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A., Gralnick, H. R. & Sultan, C. (1976) Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol*, 33(4), 451-8.
Bernad, A., Kopf, M., Kulbacki, R., Weich, N., Koehler, G. & Gutierrez-Ramos, J. C. (1994) Interleukin-6 is required in vivo for the regulation of stem cells and committed progenitors of the hematopoietic system. *Immunity*, 1(9), 725-31.
Billin, A. N., Thirlwell, H. & Ayer, D. E. (2000) Beta-catenin-histone deacetylase interactions regulate the transition of LEF1 from a transcriptional repressor to an activator. *Mol Cell Biol*, 20(18), 6882-90.
Björnsson, J. M., Larsson, N., Brun, A. C., Magnusson, M., Andersson, E., Lundström, P., Larsson, J., Repetowska, E., Ehinger, M., Humphries, R. K. & Karlsson, S. (2003) Reduced proliferative capacity of hematopoietic stem cells deficient in Hoxb3 and Hoxb4. *Mol Cell Biol*, 23(11), 3872-83.
Bleckmann, A., Siam, L., Klemm, F., Rietkötter, E., Wegner, C., Kramer, F., Beissbarth, T., Binder, C., Stadelmann, C. & Pukrop, T. (2013) Nuclear LEF1/TCF4 correlate with poor prognosis but not with nuclear β-catenin in cerebral metastasis of lung adenocarcinomas. *Clin Exp Metastasis*, 30(4), 471-82.
Bonnet, D. & Dick, J. E. (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med*, 3(7), 730-7.
Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, 72, 248-54.
Broudy, V. C. (1997) Stem cell factor and hematopoiesis. *Blood*, 90(4), 1345-64.
Burnett, A. K. (2012) Treatment of acute myeloid leukemia: are we making progress? *Hematology Am Soc Hematol Educ Program*, 2012, 1-6.
Burns, C. E., Traver, D., Mayhall, E., Shepard, J. L. & Zon, L. I. (2005) Hematopoietic stem cell fate is established by the Notch-Runx pathway. *Genes Dev*, 19(19), 2331-42.
Calvi, L. M., Adams, G. B., Weibrecht, K. W., Weber, J. M., Olson, D. P., Knight, M. C., Martin, R. P., Schipani, E., Divieti, P., Bringhurst, F. R., Milner, L. A.,

Kronenberg, H. M. & Scadden, D. T. (2003) Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature*, 425(6960), 841-6.
Chen, C. C., Gau, J. P., You, J. Y., Lee, K. D., Yu, Y. B., Lu, C. H., Lin, J. T., Lan, C., Lo, W. H., Liu, J. M. & Yang, C. F. (2009) Prognostic significance of beta-catenin and topoisomerase IIalpha in de novo acute myeloid leukemia. *Am J Hematol*, 84(2), 87-92.
Chen, H. J., Hsu, L. S., Shia, Y. T., Lin, M. W. & Lin, C. M. (2012) The β-catenin/TCF complex as a novel target of resveratrol in the Wnt/β-catenin signaling pathway. *Biochem Pharmacol*, 84(9), 1143-53.
Chen, J., Wu, A., Sun, H., Drakas, R., Garofalo, C., Cascio, S., Surmacz, E. & Baserga, R. (2005) Functional significance of type 1 insulin-like growth factor-mediated nuclear translocation of the insulin receptor substrate-1 and beta-catenin. *J Biol Chem*, 280(33), 29912-20.
Cheng, C. K., Li, L., Cheng, S. H., Ng, K., Chan, N. P., Ip, R. K., Wong, R. S., Shing, M. M., Li, C. K. & Ng, M. H. (2011) Secreted-frizzled related protein 1 is a transcriptional repression target of the t(8;21) fusion protein in acute myeloid leukemia. *Blood*, 118(25), 6638-48.
Chevalier, F. (2010) Standard Dyes for Total Protein Staining in Gel-Based Proteomic Analysis. *Materials (Basel)*, 3(10), 4784-4792.
Chimge, N. O., Little, G. H., Baniwal, S. K., Adisetiyo, H., Xie, Y., Zhang, T., O'Laughlin, A., Liu, Z. Y., Ulrich, P., Martin, A., Mhawech-Fauceglia, P., Ellis, M. J., Tripathy, D., Groshen, S., Liang, C., Li, Z., Schones, D. E. & Frenkel, B. (2016) RUNX1 prevents oestrogen-mediated AXIN1 suppression and β-catenin activation in ER-positive breast cancer. *Nat Commun*, 7, 10751.
Chung, E. J., Hwang, S. G., Nguyen, P., Lee, S., Kim, J. S., Kim, J. W., Henkart, P. A., Bottaro, D. P., Soon, L., Bonvini, P., Lee, S. J., Karp, J. E., Oh, H. J., Rubin, J. S. & Trepel, J. B. (2002) Regulation of leukemic cell adhesion, proliferation, and survival by beta-catenin. *Blood*, 100(3), 982-90.
Clevers, H. (2006) Wnt/beta-catenin signaling in development and disease. *Cell*, 127(3), 469-80.
Cobas, M., Wilson, A., Ernst, B., Mancini, S. J., MacDonald, H. R., Kemler, R. & Radtke, F. (2004) Beta-catenin is dispensable for hematopoiesis and lymphopoiesis. *J Exp Med*, 199(2), 221-9.
Cong, F. & Varmus, H. (2004) Nuclear-cytoplasmic shuttling of Axin regulates subcellular localization of beta-catenin. *Proc Natl Acad Sci U S A*, 101(9), 2882-7.
Crivianu-Gaita, V., Romaschin, A. & Thompson, M. (2015) High efficiency reduction capability for the formation of Fab′ antibody fragments from F(ab)2 units. *Biochem Biophys Rep*, 2, 23-28.
Czarnek, M. & Bereta, J. (2016) The CRISPR-Cas system - from bacterial immunity to genome engineering. *Postepy Hig Med Dosw (Online)*, 70(0), 901-16.
Daniels, D. L. & Weis, W. I. (2002) ICAT inhibits beta-catenin binding to Tcf/Lef-family transcription factors and the general coactivator p300 using independent structural modules. *Mol Cell*, 10(3), 573-84.
Daud, S. S. (2014) *The Role of WNT Transcription Factor TCF7L2 in Acute Myeloid Leukaemia*Cardiff University. Available online: https://orca.cf.ac.uk/57554/1/2014DAUDSSphd.pdf [Accessed 04/01/2018].
Davidson, K. C., Adams, A. M., Goodson, J. M., McDonald, C. E., Potter, J. C., Berndt, J. D., Biechele, T. L., Taylor, R. J. & Moon, R. T. (2012) Wnt/β-catenin

signaling promotes differentiation, not self-renewal, of human embryonic stem cells and is repressed by Oct4. *Proc Natl Acad Sci U S A*, 109(12), 4485-90.
de Haan, G., Dontje, B. & Nijhof, W. (1996) Concepts of hemopoietic cell amplification. Synergy, redundancy and pleiotropy of cytokines affecting the regulation of erythropoiesis. *Leuk Lymphoma*, 22(5-6), 385-94.
De Kouchkovsky, I. & Abdul-Hay, M. (2016) 'Acute myeloid leukemia: a comprehensive review and 2016 update'. *Blood Cancer J*, 6(7), e441.
de la Luz Sierra, M., Sakakibara, S., Gasperini, P., Salvucci, O., Jiang, K., McCormick, P. J., Segarra, M., Stone, J., Maric, D., Zhu, J., Qian, X., Lowy, D. R. & Tosato, G. (2010) The transcription factor Gfi1 regulates G-CSF signaling and neutrophil development through the Ras activator RasGRP1. *Blood*, 115(19), 3970-9.
De Maria, R., Testa, U., Luchetti, L., Zeuner, A., Stassi, G., Pelosi, E., Riccioni, R., Felli, N., Samoggia, P. & Peschle, C. (1999) Apoptotic role of Fas/Fas ligand system in the regulation of erythropoiesis. *Blood*, 93(3), 796-803.
De Toni, F., Racaud-Sultan, C., Chicanne, G., Mas, V. M., Cariven, C., Mesange, F., Salles, J. P., Demur, C., Allouche, M., Payrastre, B., Manenti, S. & Ysebaert, L. (2006) A crosstalk between the Wnt and the adhesion-dependent signaling pathways governs the chemosensitivity of acute myeloid leukemia. *Oncogene*, 25(22), 3113-22.
Demetri, G. D. & Griffin, J. D. (1991) Granulocyte colony-stimulating factor and its receptor. *Blood*, 78(11), 2791-808.
Ding, L., Ley, T. J., Larson, D. E., Miller, C. A., Koboldt, D. C., Welch, J. S., Ritchey, J. K., Young, M. A., Lamprecht, T., McLellan, M. D., McMichael, J. F., Wallis, J. W., Lu, C., Shen, D., Harris, C. C., Dooling, D. J., Fulton, R. S., Fulton, L. L., Chen, K., Schmidt, H., Kalicki-Veizer, J., Magrini, V. J., Cook, L., McGrath, S. D., Vickery, T. L., Wendl, M. C., Heath, S., Watson, M. A., Link, D. C., Tomasson, M. H., Shannon, W. D., Payton, J. E., Kulkarni, S., Westervelt, P., Walter, M. J., Graubert, T. A., Mardis, E. R., Wilson, R. K. & DiPersio, J. F. (2012a) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*, 481(7382), 506-10.
Ding, L. & Morrison, S. J. (2013) Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature*, 495(7440), 231-5.
Ding, L., Saunders, T. L., Enikolopov, G. & Morrison, S. J. (2012b) Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature*, 481(7382), 457-62.
Dombret, H. (2011) Gene mutation and AML pathogenesis.: Blood.
Drexler, H. G., Quentmeier, H. & MacLeod, R. A. (2004) Malignant hematopoietic cell lines: in vitro models for the study of MLL gene alterations. *Leukemia*, 18(2), 227-32.
Du, W., Li, X. E., Sipple, J. & Pang, Q. (2011) Overexpression of IL-3Rα on CD34+CD38- stem cells defines leukemia-initiating cells in Fanconi anemia AML. *Blood*, 117(16), 4243-52.
Dumetz, A. C., Snellinger-O'brien, A. M., Kaler, E. W. & Lenhoff, A. M. (2007) Patterns of protein protein interactions in salt solutions and implications for protein crystallization. *Protein Sci*, 16(9), 1867-77.
Döhner, K., Schlenk, R. F., Habdank, M., Scholl, C., Rücker, F. G., Corbacioglu, A., Bullinger, L., Fröhling, S. & Döhner, H. (2005) Mutant nucleophosmin (NPM1) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics: interaction with other gene mutations. *Blood*, 106(12), 3740-6.

Døsen, G., Tenstad, E., Nygren, M. K., Stubberud, H., Funderud, S. & Rian, E. (2006) Wnt expression and canonical Wnt signaling in human bone marrow B lymphopoiesis. *BMC Immunol*, 7, 13.
Estey, E. & Döhner, H. (2006) Acute myeloid leukaemia. *Lancet*, 368(9550), 1894-907.
Evans, P. M., Chen, X., Zhang, W. & Liu, C. (2010) KLF4 interacts with beta-catenin/TCF4 and blocks p300/CBP recruitment by beta-catenin. *Mol Cell Biol*, 30(2), 372-81.
Fagotto, F., Glück, U. & Gumbiner, B. M. (1998) Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of beta-catenin. *Curr Biol*, 8(4), 181-90.
Famili, F., Brugman, M. H., Taskesen, E., Naber, B. E. A., Fodde, R. & Staal, F. J. T. (2016) High Levels of Canonical Wnt Signaling Lead to Loss of Stemness and Increased Differentiation in Hematopoietic Stem Cells. *Stem Cell Reports*, 6(5), 652-659.
Fasan, A., Haferlach, C., Alpermann, T., Jeromin, S., Grossmann, V., Eder, C., Weissmann, S., Dicker, F., Kohlmann, A., Schindela, S., Kern, W., Haferlach, T. & Schnittger, S. (2014) The role of different genetic subtypes of CEBPA mutated AML. *Leukemia*, 28(4), 794-803.
Fasolini, M., Wu, X., Flocco, M., Trosset, J. Y., Oppermann, U. & Knapp, S. (2003) Hot spots in Tcf4 for the interaction with beta-catenin. *J Biol Chem*, 278(23), 21092-8.
Fleming, H. E., Janzen, V., Lo Celso, C., Guo, J., Leahy, K. M., Kronenberg, H. M. & Scadden, D. T. (2008) Wnt signaling in the niche enforces hematopoietic stem cell quiescence and is necessary to preserve self-renewal in vivo. *Cell Stem Cell*, 2(3), 274-83.
Friedman, A. D., Keefer, J. R., Kummalue, T., Liu, H., Wang, Q. F. & Cleaves, R. (2003) Regulation of granulocyte and monocyte differentiation by CCAAT/enhancer binding protein alpha. *Blood Cells Mol Dis*, 31(3), 338-41.
Fu, Y., Zhu, H., Wu, W., Xu, J., Chen, T., Xu, B., Qian, S., Li, J. & Liu, P. (2014) Clinical significance of lymphoid enhancer-binding factor 1 expression in acute myeloid leukemia. *Leuk Lymphoma*, 55(2), 371-7.
Funayama, N., Fagotto, F., McCrea, P. & Gumbiner, B. M. (1995) Embryonic axis induction by the armadillo repeat domain of beta-catenin: evidence for intracellular signaling. *J Cell Biol*, 128(5), 959-68.
Gale, R. E., Hills, R., Kottaridis, P. D., Srirangan, S., Wheatley, K., Burnett, A. K. & Linch, D. C. (2005) No evidence that FLT3 status should be considered as an indicator for transplantation in acute myeloid leukemia (AML): an analysis of 1135 patients, excluding acute promyelocytic leukemia, from the UK MRC AML10 and 12 trials. *Blood*, 106(10), 3658-65.
Gandhirajan, R. K., Staib, P. A., Minke, K., Gehrke, I., Plickert, G., Schlösser, A., Schmitt, E. K., Hallek, M. & Kreuzer, K. A. (2010) Small molecule inhibitors of Wnt/beta-catenin/lef-1 signaling induces apoptosis in chronic lymphocytic leukemia cells in vitro and in vivo. *Neoplasia*, 12(4), 326-35.
Gandillet, A., Park, S., Lassailly, F., Griessinger, E., Vargaftig, J., Filby, A., Lister, T. A. & Bonnet, D. (2011) Heterogeneous sensitivity of human acute myeloid leukemia to β-catenin down-modulation. *Leukemia*, 25(5), 770-780.
Giannini, A., Mazor, M., Orme, M., Vivanco, M., Waxman, J. & Kypta, R. (2004) Nuclear export of alpha-catenin: overlap between nuclear export signal sequences and the beta-catenin binding site. *Exp Cell Res*, 295(1), 150-60.

Gilda, J. E., Ghosh, R., Cheah, J. X., West, T. M., Bodine, S. C. & Gomes, A. V. (2015) Western Blotting Inaccuracies with Unverified Antibodies: Need for a Western Blotting Minimal Reporting Standard (WBMRS). *PLoS One*, 10(8), e0135392.
Gingras, A. C., Aebersold, R. & Raught, B. (2005) Advances in protein complex analysis using mass spectrometry. *J Physiol*, 563(Pt 1), 11-21.
Griffin, J. N., Del Viso, F., Duncan, A. R., Robson, A., Hwang, W., Kulkarni, S., Liu, K. J. & Khokha, M. K. (2018) RAPGEF5 Regulates Nuclear Translocation of β-Catenin. *Dev Cell*, 44(2), 248-260.e4.
Grimwade, D., Walker, H., Oliver, F., Wheatley, K., Harrison, C., Harrison, G., Rees, J., Hann, I., Stevens, R., Burnett, A. & Goldstone, A. (1998) The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood*, 92(7), 2322-33.
Guo, X., Zhang, R., Liu, J., Li, M., Song, C., Dovat, S., Li, J. & Ge, Z. (2015) Characterization of LEF1 High Expression and Novel Mutations in Adult Acute Lymphoblastic Leukemia. *PLoS One*, 10(5), e0125429.
Gutierrez, A., Tschumper, R. C., Wu, X., Shanafelt, T. D., Eckel-Passow, J., Huddleston, P. M., Slager, S. L., Kay, N. E. & Jelinek, D. F. (2010) LEF-1 is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B-cell lymphocytosis. *Blood*, 116(16), 2975-83.
Harris, T. J. & Peifer, M. (2005) Decisions, decisions: beta-catenin chooses between adhesion and transcription. *Trends Cell Biol*, 15(5), 234-7.
He, T. C., Sparks, A. B., Rago, C., Hermeking, H., Zawel, L., da Costa, L. T., Morin, P. J., Vogelstein, B. & Kinzler, K. W. (1998) Identification of c-MYC as a target of the APC pathway. *Science*, 281(5382), 1509-12.
Henderson, B. R., Galea, M., Schuechner, S. & Leung, L. (2002) Lymphoid enhancer factor-1 blocks adenomatous polyposis coli-mediated nuclear export and degradation of beta-catenin. Regulation by histone deacetylase 1. *J Biol Chem*, 277(27), 24258-64.
Hendriksen, J., Fagotto, F., van der Velde, H., van Schie, M., Noordermeer, J. & Fornerod, M. (2005) RanBP3 enhances nuclear export of active (beta)-catenin independently of CRM1. *J Cell Biol*, 171(5), 785-97.
Henzel, W. J., Billeci, T. M., Stults, J. T., Wong, S. C., Grimley, C. & Watanabe, C. (1993) Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci U S A*, 90(11), 5011-5.
Heo, S. K., Noh, E. K., Kim, J. Y., Jeong, Y. K., Jo, J. C., Choi, Y., Koh, S., Baek, J. H., Min, Y. J. & Kim, H. (2017) Targeting c-KIT (CD117) by dasatinib and radotinib promotes acute myeloid leukemia cell death. *Sci Rep*, 7(1), 15278.
Herencia, C., Martínez-Moreno, J. M., Herrera, C., Corrales, F., Santiago-Mora, R., Espejo, I., Barco, M., Almadén, Y., de la Mata, M., Rodríguez-Ariza, A. & Muñoz-Castañeda, J. R. (2012) Nuclear translocation of β-catenin during mesenchymal stem cells differentiation into hepatocytes is associated with a tumoral phenotype. *PLoS One*, 7(4), e34656.
Hilgarth, R. S. & Sarge, K. D. (2005) Detection of sumoylated proteins. *Methods Mol Biol*, 301, 329-38.
Hodge, K., Have, S. T., Hutton, L. & Lamond, A. I. (2013) Cleaning up the masses: exclusion lists to reduce contamination with HPLC-MS/MS. *J Proteomics*, 88, 92-103.

Hosen, N., Park, C. Y., Tatsumi, N., Oji, Y., Sugiyama, H., Gramatzki, M., Krensky, A. M. & Weissman, I. L. (2007) CD96 is a leukemic stem cell-specific marker in human acute myeloid leukemia. *Proc Natl Acad Sci U S A*, 104(26), 11008-13.
Hou, H. A., Kuo, Y. Y., Liu, C. Y., Lee, M. C., Tang, J. L., Chen, C. Y., Chou, W. C., Huang, C. F., Lee, F. Y., Liu, M. C., Yao, M. & Tien, H. F. (2011) Distinct association between aberrant methylation of Wnt inhibitors and genetic alterations in acute myeloid leukaemia. *Br J Cancer*, 105(12), 1927-33.
Hsu, W., Zeng, L. & Costantini, F. (1999) Identification of a domain of Axin that binds to the serine/threonine protein phosphatase 2A and a self-binding domain. *J Biol Chem*, 274(6), 3439-45.
Hu, Z. B., Ma, W., Uphoff, C. C., Quentmeier, H. & Drexler, H. G. (1994) c-kit expression in human megakaryoblastic leukemia cell lines. *Blood*, 83(8), 2133-44.
Huang, B. X. & Kim, H. Y. (2013) Effective identification of Akt interacting proteins by two-step chemical crosslinking, co-immunoprecipitation and mass spectrometry. *PLoS One*, 8(4), e61430.
Huber, A. H., Nelson, W. J. & Weis, W. I. (1997) Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell*, 90(5), 871-82.
Huber, O., Korn, R., McLaughlin, J., Ohsugi, M., Herrmann, B. G. & Kemler, R. (1996) Nuclear localization of beta-catenin by interaction with transcription factor LEF-1. *Mech Dev*, 59(1), 3-10.
Ichikawa, M., Yoshimi, A., Nakagawa, M., Nishimoto, N., Watanabe-Okochi, N. & Kurokawa, M. (2013) A role for RUNX1 in hematopoiesis and myeloid leukemia. *Int J Hematol*, 97(6), 726-34.
Ilyas, A. M., Ahmad, S., Faheem, M., Naseer, M. I., Kumosani, T. A., Al-Qahtani, M. H., Gari, M. & Ahmed, F. (2015) Next generation sequencing of acute myeloid leukemia: influencing prognosis. *BMC Genomics*, 16 Suppl 1, S5.
Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E. A., Iwasaki-Arai, J., Mizuno, S., Arinobu, Y., Geary, K., Zhang, P., Dayaram, T., Fenyus, M. L., Elf, S., Chan, S., Kastner, P., Huettner, C. S., Murray, R., Tenen, D. G. & Akashi, K. (2005) Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood*, 106(5), 1590-600.
James, P., Quadroni, M., Carafoli, E. & Gonnet, G. (1993) Protein identification by mass profile fingerprinting. *Biochem Biophys Res Commun*, 195(1), 58-64.
Jamieson, C., Mills, K. M., Lui, C., Semaan, C., Molloy, M. P., Sharma, M., Forwood, J. K. & Henderson, B. R. (2016) Characterization of a beta-catenin nuclear localization defect in MCF-7 breast cancer cells. *Exp Cell Res*, 341(2), 196-206.
Jamieson, C., Sharma, M. & Henderson, B. R. (2011) Regulation of β-catenin nuclear dynamics by GSK-3β involves a LEF-1 positive feedback loop. *Traffic*, 12(8), 983-99.
Jeannet, G., Scheller, M., Scarpellino, L., Duboux, S., Gardiol, N., Back, J., Kuttler, F., Malanchi, I., Birchmeier, W., Leutz, A., Huelsken, J. & Held, W. (2008) Long-term, multilineage hematopoiesis occurs in the combined absence of beta-catenin and gamma-catenin. *Blood*, 111(1), 142-9.
Jian, H., Shen, X., Liu, I., Semenov, M., He, X. & Wang, X. F. (2006) Smad3-dependent nuclear translocation of beta-catenin is required for TGF-beta1-induced proliferation of bone marrow-derived adult human mesenchymal stem cells. *Genes Dev*, 20(6), 666-74.

Jordan, C. T., Upchurch, D., Szilvassy, S. J., Guzman, M. L., Howard, D. S., Pettigrew, A. L., Meyerrose, T., Rossi, R., Grimes, B., Rizzieri, D. A., Luger, S. M. & Phillips, G. L. (2000) The interleukin-3 receptor alpha chain is a unique marker for human acute myelogenous leukemia stem cells. *Leukemia*, 14(10), 1777-84.
Kahler, R. A. & Westendorf, J. J. (2003) Lymphoid enhancer factor-1 and beta-catenin inhibit Runx2-dependent transcriptional activation of the osteocalcin promoter. *J Biol Chem*, 278(14), 11937-44.
Kajiguchi, T., Chung, E. J., Lee, S., Stine, A., Kiyoi, H., Naoe, T., Levis, M. J., Neckers, L. & Trepel, J. B. (2007) FLT3 regulates beta-catenin tyrosine phosphorylation, nuclear localization, and transcriptional activity in acute myeloid leukemia cells. *Leukemia*, 21(12), 2476-84.
Kantarjian, H., O'Brien, S., Jabbour, E., Garcia-Manero, G., Quintas-Cardama, A., Shan, J., Rios, M. B., Ravandi, F., Faderl, S., Kadia, T., Borthakur, G., Huang, X., Champlin, R., Talpaz, M. & Cortes, J. (2012) Improved survival in chronic myeloid leukemia since the introduction of imatinib therapy: a single-institution historical experience. *Blood*, 119(9), 1981-7.
Kao, H. W., Liang, D. C., Wu, J. H., Kuo, M. C., Wang, P. N., Yang, C. P., Shih, Y. S., Lin, T. H., Huang, Y. H. & Shih, L. Y. (2014) Gene mutation patterns in patients with minimally differentiated acute myeloid leukemia. *Neoplasia*, 16(6), 481-8.
Kau, T. R., Way, J. C. & Silver, P. A. (2004) Nuclear transport and cancer: from mechanism to intervention. *Nat Rev Cancer*, 4(2), 106-17.
Kaushansky, K., Broudy, V. C., Lin, N., Jorgensen, M. J., McCarty, J., Fox, N., Zucker-Franklin, D. & Lofton-Day, C. (1995) Thrombopoietin, the Mp1 ligand, is essential for full megakaryocyte development. *Proc Natl Acad Sci U S A*, 92(8), 3234-8.
Kelly, L. M. & Gilliland, D. G. (2002) Genetics of myeloid leukemias. *Annu Rev Genomics Hum Genet*, 3, 179-98.
Kim, D. H., Sarbassov, D. D., Ali, S. M., King, J. E., Latek, R. R., Erdjument-Bromage, H., Tempst, P. & Sabatini, D. M. (2002) mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell*, 110(2), 163-75.
Kim, J. A., Kang, Y. J., Park, G., Kim, M., Park, Y. O., Kim, H., Leem, S. H., Chu, I. S., Lee, J. S., Jho, E. H. & Oh, I. H. (2009) Identification of a stroma-mediated Wnt/beta-catenin signal promoting self-renewal of hematopoietic stem cells in the stem cell niche. *Stem Cells*, 27(6), 1318-29.
Kim, S. E., Huang, H., Zhao, M., Zhang, X., Zhang, A., Semonov, M. V., MacDonald, B. T., Garcia Abreu, J., Peng, L. & He, X. (2013) Wnt stabilization of β-catenin reveals principles for morphogen receptor-scaffold assemblies. *Science*, 340(6134), 867-70.
Kim, Y. M. & Kahn, M. (2014) The role of the Wnt signaling pathway in cancer stem cells: prospects for drug development. *Res Rep Biochem*, 4, 1-12.
Kirstetter, P., Anderson, K., Porse, B. T., Jacobsen, S. E. & Nerlov, C. (2006) Activation of the canonical Wnt pathway leads to loss of hematopoietic stem cell repopulation and multilineage differentiation block. *Nat Immunol*, 7(10), 1048-56.
Komiya, Y. & Habas, R. (2008) Wnt signal transduction pathways. *Organogenesis*, 4(2), 68-75.
Korinek, V., Barker, N., Morin, P. J., van Wichen, D., de Weger, R., Kinzler, K. W., Vogelstein, B. & Clevers, H. (1997) Constitutive transcriptional activation by

a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science*, 275(5307), 1784-7.
Korrodi-Gregório, L., Esteves, S. L. & Fardilha, M. (2014) Protein phosphatase 1 catalytic isoforms: specificity toward interacting proteins. *Transl Res*, 164(5), 366-91.
Kramps, T., Peter, O., Brunner, E., Nellen, D., Froesch, B., Chatterjee, S., Murone, M., Züllig, S. & Basler, K. (2002) Wnt/wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear beta-catenin-TCF complex. *Cell*, 109(1), 47-60.
Krieghoff, E., Behrens, J. & Mayr, B. (2006) Nucleo-cytoplasmic distribution of beta-catenin is regulated by retention. *J Cell Sci*, 119(Pt 7), 1453-63.
Kufe, D. W. (2009) Mucins in cancer: function, prognosis and therapy. *Nat Rev Cancer*, 9(12), 874-85.
Lambertini, E., Lisignoli, G., Torreggiani, E., Manferdini, C., Gabusi, E., Franceschetti, T., Penolazzi, L., Gambari, R., Facchini, A. & Piva, R. (2009) Slug gene expression supports human osteoblast maturation. *Cell Mol Life Sci*, 66(22), 3641-53.
Lane, S. W., Wang, Y. J., Lo Celso, C., Ragu, C., Bullinger, L., Sykes, S. M., Ferraro, F., Shterental, S., Lin, C. P., Gilliland, D. G., Scadden, D. T., Armstrong, S. A. & Williams, D. A. (2011) Differential niche and Wnt requirements during acute myeloid leukemia progression. *Blood*, 118(10), 2849-56.
Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M. A. & Dick, J. E. (1994) A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature*, 367(6464), 645-8.
Laurenti, E., Varnum-Finney, B., Wilson, A., Ferrero, I., Blanco-Bose, W. E., Ehninger, A., Knoepfler, P. S., Cheng, P. F., MacDonald, H. R., Eisenman, R. N., Bernstein, I. D. & Trumpp, A. (2008) Hematopoietic stem cell function and survival depend on c-Myc and N-Myc activity. *Cell Stem Cell*, 3(6), 611-24.
Lecine, P., Villeval, J. L., Vyas, P., Swencki, B., Xu, Y. & Shivdasani, R. A. (1998) Mice lacking transcription factor NF-E2 provide in vivo validation of the proplatelet model of thrombocytopoiesis and show a platelet production defect that is intrinsic to megakaryocytes. *Blood*, 92(5), 1608-16.
Lee, B. H., Tothova, Z., Levine, R. L., Anderson, K., Buza-Vidas, N., Cullen, D. E., McDowell, E. P., Adelsperger, J., Fröhling, S., Huntly, B. J., Beran, M., Jacobsen, S. E. & Gilliland, D. G. (2007) FLT3 mutations confer enhanced proliferation and survival properties to multipotent progenitors in a murine model of chronic myelomonocytic leukemia. *Cancer Cell*, 12(4), 367-80.
Ley, T. J., Miller, C., Ding, L., Raphael, B. J., Mungall, A. J., Robertson, A., Hoadley, K., Triche, T. J., Laird, P. W., Baty, J. D., Fulton, L. L., Fulton, R., Heath, S. E., Kalicki-Veizer, J., Kandoth, C., Klco, J. M., Koboldt, D. C., Kanchi, K. L., Kulkarni, S., Lamprecht, T. L., Larson, D. E., Lin, L., Lu, C., McLellan, M. D., McMichael, J. F., Payton, J., Schmidt, H., Spencer, D. H., Tomasson, M. H., Wallis, J. W., Wartman, L. D., Watson, M. A., Welch, J., Wendl, M. C., Ally, A., Balasundaram, M., Birol, I., Butterfield, Y., Chiu, R., Chu, A., Chuah, E., Chun, H. J., Corbett, R., Dhalla, N., Guin, R., He, A., Hirst, C., Hirst, M., Holt, R. A., Jones, S., Karsan, A., Lee, D., Li, H. I., Marra, M. A., Mayo, M., Moore, R. A., Mungall, K., Parker, J., Pleasance, E., Plettner, P., Schein, J., Stoll, D., Swanson, L., Tam, A., Thiessen, N., Varhol, R., Wye, N., Zhao, Y., Gabriel, S., Getz, G., Sougnez, C., Zou, L., Leiserson, M. D., Vandin, F., Wu, H. T., Applebaum, F., Baylin, S. B., Akbani, R., Broom, B. M., Chen, K., Motter, T. C., Nguyen, K.,

Weinstein, J. N., Zhang, N., Ferguson, M. L., Adams, C., Black, A., Bowen, J., Gastier-Foster, J., Grossman, T., Lichtenberg, T., Wise, L., Davidsen, T., Demchok, J. A., Shaw, K. R., Sheth, M., Sofia, H. J., Yang, L., Downing, J. R., Eley, G., *et al* (2013) Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*, 368(22), 2059-74.
Li, V. S., Ng, S. S., Boersema, P. J., Low, T. Y., Karthaus, W. R., Gerlach, J. P., Mohammed, S., Heck, A. J., Maurice, M. M., Mahmoudi, T. & Clevers, H. (2012) Wnt signaling through inhibition of β-catenin degradation in an intact Axin1 complex. *Cell*, 149(6), 1245-56.
Li, X., Hoeppner, L. H., Jensen, E. D., Gopalakrishnan, R. & Westendorf, J. J. (2009) Co-activator activator (CoAA) prevents the transcriptional activity of Runt domain transcription factors. *J Cell Biochem*, 108(2), 378-87.
Li, Y., Yi, H., Yao, Y., Liao, X., Xie, Y., Yang, J., Yan, Z., Wang, L., Lu, S., Kuang, Y., Gu, M., Fei, J., Wang, Z. & Huang, L. (2011) The cytoplasmic domain of MUC1 induces hyperplasia in the mammary gland and correlates with nuclear accumulation of β-catenin. *PLoS One*, 6(4), e19102.
Liao, Q., Chiu, N. H., Shen, C., Chen, Y. & Vouros, P. (2007) Investigation of enzymatic behavior of benzonase/alkaline phosphatase in the digestion of oligonucleotides and DNA by ESI-LC/MS. *Anal Chem*, 79(5), 1907-17.
Lindman, S., Xue, W. F., Szczepankiewicz, O., Bauer, M. C., Nilsson, H. & Linse, S. (2006) Salting the charged surface: pH and salt dependence of protein G B1 stability. *Biophys J*, 90(8), 2911-21.
Liu, L., Papa, E. F., Dooner, M. S., Machan, J. T., Johnson, K. W., Goldberg, L. R., Quesenberry, P. J. & Colvin, G. A. (2012) Homing and long-term engraftment of long- and short-term renewal hematopoietic stem cells. *PLoS One*, 7(2), e31300.
Ludin, A., Gur-Cohen, S., Golan, K., Kaufmann, K. B., Itkin, T., Medaglia, C., Lu, X. J., Ledergor, G., Kollet, O. & Lapidot, T. (2014) Reactive oxygen species regulate hematopoietic stem cell self-renewal, migration and development, as well as their bone marrow microenvironment. *Antioxid Redox Signal*, 21(11), 1605-19.
Luis, T. C., Naber, B. A., Roozen, P. P., Brugman, M. H., de Haas, E. F., Ghazvini, M., Fibbe, W. E., van Dongen, J. J., Fodde, R. & Staal, F. J. (2011) Canonical wnt signaling regulates hematopoiesis in a dosage-dependent fashion. *Cell Stem Cell*, 9(4), 345-56.
Luis, T. C., Weerkamp, F., Naber, B. A., Baert, M. R., de Haas, E. F., Nikolic, T., Heuvelmans, S., De Krijger, R. R., van Dongen, J. J. & Staal, F. J. (2009) Wnt3a deficiency irreversibly impairs hematopoietic stem cell self-renewal and leads to defects in progenitor cell differentiation. *Blood*, 113(3), 546-54.
Luo, W., Peterson, A., Garcia, B. A., Coombs, G., Kofahl, B., Heinrich, R., Shabanowitz, J., Hunt, D. F., Yost, H. J. & Virshup, D. M. (2007) Protein phosphatase 1 regulates assembly and function of the beta-catenin degradation complex. *EMBO J*, 26(6), 1511-21.
Majeti, R., Becker, M. W., Tian, Q., Lee, T. L., Yan, X., Liu, R., Chiang, J. H., Hood, L., Clarke, M. F. & Weissman, I. L. (2009a) Dysregulated gene expression networks in human acute myelogenous leukemia stem cells. *Proc Natl Acad Sci U S A*, 106(9), 3396-401.
Majeti, R., Chao, M. P., Alizadeh, A. A., Pang, W. W., Jaiswal, S., Gibbs, K. D., van Rooijen, N. & Weissman, I. L. (2009b) CD47 is an adverse prognostic factor and therapeutic antibody target on human acute myeloid leukemia stem cells. *Cell*, 138(2), 286-99.

Mancini, F., Teveroni, E., Di Conza, G., Monteleone, V., Arisi, I., Pellegrino, M., Buttarelli, M., Pieroni, L., D'Onofrio, M., Urbani, A., Pontecorvi, A., Mazzone, M. & Moretti, F. (2017) MDM4 actively restrains cytoplasmic mTORC1 by sensing nutrient availability. *Mol Cancer*, 16(1), 55.
Mann, M., Højrup, P. & Roepstorff, P. (1993) Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol Mass Spectrom*, 22(6), 338-45.
Markiewicz, E., Tilgner, K., Barker, N., van de Wetering, M., Clevers, H., Dorobek, M., Hausmanowa-Petrusewicz, I., Ramaekers, F. C., Broers, J. L., Blankesteijn, W. M., Salpingidou, G., Wilson, R. G., Ellis, J. A. & Hutchison, C. J. (2006) The inner nuclear membrane protein emerin regulates beta-catenin activity by restricting its accumulation in the nucleus. *EMBO J*, 25(14), 3275-85.
Mazzoni, S. M. & Fearon, E. R. (2014) AXIN1 and AXIN2 variants in gastrointestinal cancers. *Cancer Lett*, 355(1), 1-8.
Meijer, L., Skaltsounis, A. L., Magiatis, P., Polychronopoulos, P., Knockaert, M., Leost, M., Ryan, X. P., Vonica, C. A., Brivanlou, A., Dajani, R., Crovace, C., Tarricone, C., Musacchio, A., Roe, S. M., Pearl, L. & Greengard, P. (2003) GSK-3-selective inhibitors derived from Tyrian purple indirubins. *Chem Biol*, 10(12), 1255-66.
Metzeler, K. H., Heilmeier, B., Edmaier, K. E., Rawat, V. P., Dufour, A., Döhner, K., Feuring-Buske, M., Braess, J., Spiekermann, K., Büchner, T., Sauerland, M. C., Döhner, H., Hiddemann, W., Bohlander, S. K., Schlenk, R. F., Bullinger, L. & Buske, C. (2012) High expression of lymphoid enhancer-binding factor-1 (LEF1) is a novel favorable prognostic factor in cytogenetically normal acute myeloid leukemia. *Blood*, 120(10), 2118-26.
Meyer, C., Hofmann, J., Burmeister, T., Gröger, D., Park, T. S., Emerenciano, M., Pombo de Oliveira, M., Renneville, A., Villarese, P., Macintyre, E., Cavé, H., Clappier, E., Mass-Malo, K., Zuna, J., Trka, J., De Braekeleer, E., De Braekeleer, M., Oh, S. H., Tsaur, G., Fechina, L., van der Velden, V. H., van Dongen, J. J., Delabesse, E., Binato, R., Silva, M. L., Kustanovich, A., Aleinikova, O., Harris, M. H., Lund-Aho, T., Juvonen, V., Heidenreich, O., Vormoor, J., Choi, W. W., Jarosova, M., Kolenova, A., Bueno, C., Menendez, P., Wehner, S., Eckert, C., Talmant, P., Tondeur, S., Lippert, E., Launay, E., Henry, C., Ballerini, P., Lapillone, H., Callanan, M. B., Cayuela, J. M., Herbaux, C., Cazzaniga, G., Kakadiya, P. M., Bohlander, S., Ahlmann, M., Choi, J. R., Gameiro, P., Lee, D. S., Krauter, J., Cornillet-Lefebvre, P., Te Kronnie, G., Schäfer, B. W., Kubetzko, S., Alonso, C. N., zur Stadt, U., Sutton, R., Venn, N. C., Izraeli, S., Trakhtenbrot, L., Madsen, H. O., Archer, P., Hancock, J., Cerveira, N., Teixeira, M. R., Lo Nigro, L., Möricke, A., Stanulla, M., Schrappe, M., Sedék, L., Szczepański, T., Zwaan, C. M., Coenen, E. A., van den Heuvel-Eibrink, M. M., Strehl, S., Dworzak, M., Panzer-Grümayer, R., Dingermann, T., Klingebiel, T. & Marschalek, R. (2013) The MLL recombinome of acute leukemias in 2013. *Leukemia*, 27(11), 2165-76.
Mikesch, J. H., Steffen, B., Berdel, W. E., Serve, H. & Müller-Tidow, C. (2007) The emerging role of Wnt signaling in the pathogenesis of acute myeloid leukemia. *Leukemia*, 21(8), 1638-47.
Minke, K. S., Staib, P., Puetter, A., Gehrke, I., Gandhirajan, R. K., Schlösser, A., Schmitt, E. K., Hallek, M. & Kreuzer, K. A. (2009) Small molecule inhibitors of WNT signaling effectively induce apoptosis in acute myeloid leukemia cells. *Eur J Haematol*, 82(3), 165-75.

Mohyeldin, A., Garzón-Muvdi, T. & Quiñones-Hinojosa, A. (2010) Oxygen in stem cell biology: a critical component of the stem cell niche. *Cell Stem Cell*, 7(2), 150-61.
Molenaar, M., van de Wetering, M., Oosterwegel, M., Peterson-Maduro, J., Godsave, S., Korinek, V., Roose, J., Destrée, O. & Clevers, H. (1996) XTcf-3 transcription factor mediates beta-catenin-induced axis formation in Xenopus embryos. *Cell*, 86(3), 391-9.
Morgan, R. G., Pearn, L., Liddiard, K., Pumford, S. L., Burnett, A. K., Tonks, A. & Darley, R. L. (2013) γ-Catenin is overexpressed in acute myeloid leukemia and promotes the stabilization and nuclear localization of β-catenin. *Leukemia*, 27(2), 336-43.
Morgan, R. G., Ridsdale, J., Tonks, A. & Darley, R. L. (2014) Factors affecting the nuclear localization of β-catenin in normal and malignant tissue. *J Cell Biochem*, 115(8), 1351-61.
Morin, P. J., Sparks, A. B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B. & Kinzler, K. W. (1997) Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science*, 275(5307), 1787-90.
Morrison, S. J. & Kimble, J. (2006) Asymmetric and symmetric stem-cell divisions in development and cancer. *Nature*, 441(7097), 1068-74.
Morrison, S. J. & Scadden, D. T. (2014) The bone marrow niche for haematopoietic stem cells. *Nature*, 505(7483), 327-34.
Moser, J. J., Chan, E. K. & Fritzler, M. J. (2009) Optimization of immunoprecipitation-western blot analysis in detecting GW182-associated components of GW/P bodies. *Nat Protoc*, 4(5), 674-85.
Mrózek, K. (2008) Cytogenetic, molecular genetic, and clinical characteristics of acute myeloid leukemia with a complex karyotype. *Semin Oncol*, 35(4), 365-77.
Mrózek, K., Heinonen, K. & Bloomfield, C. D. (2001) Clinical importance of cytogenetics in acute myeloid leukaemia. *Best Pract Res Clin Haematol*, 14(1), 19-47.
Mulholland, D. J., Cheng, H., Reid, K., Rennie, P. S. & Nelson, C. C. (2002) The androgen receptor can promote beta-catenin nuclear translocation independently of adenomatous polyposis coli. *J Biol Chem*, 277(20), 17933-43.
Müller, T., Bain, G., Wang, X. & Papkoff, J. (2002) Regulation of epithelial cell migration and tumor formation by beta-catenin signaling. *Exp Cell Res*, 280(1), 119-33.
Müller-Tidow, C., Steffen, B., Cauvet, T., Tickenbrock, L., Ji, P., Diederichs, S., Sargin, B., Köhler, G., Stelljes, M., Puccetti, E., Ruthardt, M., deVos, S., Hiebert, S. W., Koeffler, H. P., Berdel, W. E. & Serve, H. (2004) Translocation products in acute myeloid leukemia activate the Wnt signaling pathway in hematopoietic cells. *Mol Cell Biol*, 24(7), 2890-904.
Nakamura, T., Hamada, F., Ishidate, T., Anai, K., Kawahara, K., Toyoshima, K. & Akiyama, T. (1998) Axin, an inhibitor of the Wnt signalling pathway, interacts with beta-catenin, GSK-3beta and APC and reduces the beta-catenin level. *Genes Cells*, 3(6), 395-403.
Nelson, W. J. & Nusse, R. (2004) Convergence of Wnt, beta-catenin, and cadherin pathways. *Science*, 303(5663), 1483-7.
Nemeth, M. J., Mak, K. K., Yang, Y. & Bodine, D. M. (2009) beta-Catenin expression in the bone marrow microenvironment is required for long-term maintenance of primitive hematopoietic cells. *Stem Cells*, 27(5), 1109-19.
Nilsson, S. K., Johnston, H. M., Whitty, G. A., Williams, B., Webb, R. J., Denhardt, D. T., Bertoncello, I., Bendall, L. J., Simmons, P. J. & Haylock, D. N. (2005)

Osteopontin, a key component of the hematopoietic stem cell niche and regulator of primitive hematopoietic progenitor cells. *Blood*, 106(4), 1232-9.
Nusse, R. & Clevers, H. (2017) Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell*, 169(6), 985-999.
Okamura, R. M., Sigvardsson, M., Galceran, J., Verbeek, S., Clevers, H. & Grosschedl, R. (1998) Redundant regulation of T cell differentiation and TCRalpha gene expression by the transcription factors LEF-1 and TCF-1. *Immunity*, 8(1), 11-20.
Okuno, T. & Kondelis, N. (1978) Evaluation of dithiothreitol (DTT) for inactivation of IgM antibodies. *J Clin Pathol*, 31(12), 1152-5.
Orford, K., Crockett, C., Jensen, J. P., Weissman, A. M. & Byers, S. W. (1997) Serine phosphorylation-regulated ubiquitination and degradation of beta-catenin. *J Biol Chem*, 272(40), 24735-8.
Palka-Hamblin, H. L., Gierut, J. J., Bie, W., Brauer, P. M., Zheng, Y., Asara, J. M. & Tyner, A. L. (2010) Identification of beta-catenin as a target of the intracellular tyrosine kinase PTK6. *J Cell Sci*, 123(Pt 2), 236-45.
Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., Potter, N. E., Heuser, M., Thol, F., Bolli, N., Gundem, G., Van Loo, P., Martincorena, I., Ganly, P., Mudie, L., McLaren, S., O'Meara, S., Raine, K., Jones, D. R., Teague, J. W., Butler, A. P., Greaves, M. F., Ganser, A., Döhner, K., Schlenk, R. F., Döhner, H. & Campbell, P. J. (2016) Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med*, 374(23), 2209-2221.
Pappin, D. J. (2003) Peptide mass fingerprinting using MALDI-TOF mass spectrometry. *Methods Mol Biol*, 211, 211-9.
Passegué, E., Wagner, E. F. & Weissman, I. L. (2004) JunB deficiency leads to a myeloproliferative disorder arising from hematopoietic stem cells. *Cell*, 119(3), 431-43.
Petropoulos, K., Arseni, N., Schessl, C., Stadler, C. R., Rawat, V. P., Deshpande, A. J., Heilmeier, B., Hiddemann, W., Quintanilla-Martinez, L., Bohlander, S. K., Feuring-Buske, M. & Buske, C. (2008) A novel role for Lef-1, a central transcription mediator of Wnt signaling, in leukemogenesis. *J Exp Med*, 205(3), 515-22.
Pietras, E. M., Mirantes-Barbeito, C., Fong, S., Loeffler, D., Kovtonyuk, L. V., Zhang, S., Lakshminarasimhan, R., Chin, C. P., Techner, J. M., Will, B., Nerlov, C., Steidl, U., Manz, M. G., Schroeder, T. & Passegué, E. (2016) Chronic interleukin-1 exposure drives haematopoietic stem cells towards precocious myeloid differentiation at the expense of self-renewal. *Nat Cell Biol*, 18(6), 607-18.
Pilarsky, C., Wenzig, M., Specht, T., Saeger, H. D. & Grützmann, R. (2004) Identification and validation of commonly overexpressed genes in solid tumors by comparison of microarray data. *Neoplasia*, 6(6), 744-50.
Pirofsky, B. & Rosner, E. R. (1974) DTT test: a new method to differentiate IgM and IgG erythrocyte antibodies. *Vox Sang*, 27(5), 480-8.
Plovins, A., Alvarez, A. M., Ibañez, M., Molina, M. & Nombela, C. (1994) Use of fluorescein-di-beta-D-galactopyranoside (FDG) and C12-FDG as substrates for beta-galactosidase detection by flow cytometry in animal, bacterial, and yeast cells. *Appl Environ Microbiol*, 60(12), 4638-41.
Polakis, P., Hart, M. & Rubinfeld, B. (1999) Defects in the regulation of beta-catenin in colorectal cancer. *Adv Exp Med Biol*, 470, 23-32.

Pongracz, J. E., Parnell, S. M., Jones, T., Anderson, G. & Jenkinson, E. J. (2006) Overexpression of ICAT highlights a role for catenin-mediated canonical Wnt signalling in early T cell development. *Eur J Immunol*, 36(9), 2376-83.
Poulos, M. G., Guo, P., Kofler, N. M., Pinho, S., Gutkin, M. C., Tikhonova, A., Aifantis, I., Frenette, P. S., Kitajewski, J., Rafii, S. & Butler, J. M. (2013) Endothelial Jagged-1 is necessary for homeostatic and regenerative hematopoiesis. *Cell Rep*, 4(5), 1022-34.
Prieve, M. G. & Waterman, M. L. (1999) Nuclear localization and formation of beta-catenin-lymphoid enhancer factor 1 complexes are not sufficient for activation of gene expression. *Mol Cell Biol*, 19(6), 4503-15.
Qian, H., Buza-Vidas, N., Hyland, C. D., Jensen, C. T., Antonchuk, J., Månsson, R., Thoren, L. A., Ekblom, M., Alexander, W. S. & Jacobsen, S. E. (2007) Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. *Cell Stem Cell*, 1(6), 671-84.
Qualtiere, L. F., Anderson, A. G. & Meyers, P. (1977) Effects of ionic and nonionic detergents on antigen-antibody reactions. *J Immunol*, 119(5), 1645-51.
Rebelo, S., Santos, M., Martins, F., da Cruz e Silva, E. F. & da Cruz e Silva, O. A. (2015) Protein phosphatase 1 is a key player in nuclear events. *Cell Signal*, 27(12), 2589-98.
Reya, T., Duncan, A. W., Ailles, L., Domen, J., Scherer, D. C., Willert, K., Hintz, L., Nusse, R. & Weissman, I. L. (2003) A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature*, 423(6938), 409-14.
Reya, T., O'Riordan, M., Okamura, R., Devaney, E., Willert, K., Nusse, R. & Grosschedl, R. (2000) Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. *Immunity*, 13(1), 15-24.
Roose, J., Huls, G., van Beest, M., Moerer, P., van der Horn, K., Goldschmeding, R., Logtenberg, T. & Clevers, H. (1999) Synergy between tumor suppressor APC and the beta-catenin-Tcf4 target Tcf1. *Science*, 285(5435), 1923-6.
Roose, J., Molenaar, M., Peterson, J., Hurenkamp, J., Brantjes, H., Moerer, P., van de Wetering, M., Destrée, O. & Clevers, H. (1998) The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. *Nature*, 395(6702), 608-12.
Rosner, M. & Hengstschläger, M. (2008) Cytoplasmic and nuclear distribution of the protein complexes mTORC1 and mTORC2: rapamycin triggers dephosphorylation and delocalization of the mTORC2 components rictor and sin1. *Hum Mol Genet*, 17(19), 2934-48.
Rosner, M. & Hengstschläger, M. (2012) Detection of cytoplasmic and nuclear functions of mTOR by fractionation. *Methods Mol Biol*, 821, 105-24.
Sachdev, S., Bruhn, L., Sieber, H., Pichler, A., Melchior, F. & Grosschedl, R. (2001) PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. *Genes Dev*, 15(23), 3088-103.
Sakhinia, E., Byers, R., Bashein, A., Hoyland, J., Buckle, A. M. & Brady, G. (2006) Gene expression analysis of myeloid and lymphoid lineage markers during mouse haematopoiesis. *Br J Haematol*, 135(1), 105-16.
Salomon, D., Sacco, P. A., Roy, S. G., Simcha, I., Johnson, K. R., Wheelock, M. J. & Ben-Ze'ev, A. (1997) Regulation of beta-catenin levels and localization by overexpression of plakoglobin and inhibition of the ubiquitin-proteasome system. *J Cell Biol*, 139(5), 1325-35.
Sampietro, J., Dahlberg, C. L., Cho, U. S., Hinds, T. R., Kimelman, D. & Xu, W. (2006) Crystal structure of a beta-catenin/BCL9/Tcf4 complex. *Mol Cell*, 24(2), 293-300.

Sanderson, C. J. (1992) Interleukin-5, eosinophils, and disease. *Blood*, 79(12), 3101-9.
Santaguida, M., Schepers, K., King, B., Sabnis, A. J., Forsberg, E. C., Attema, J. L., Braun, B. S. & Passegué, E. (2009) JunB protects against myeloid malignancies by limiting hematopoietic stem cell proliferation and differentiation without affecting self-renewal. *Cancer Cell*, 15(4), 341-52.
Scheller, M., Huelsken, J., Rosenbauer, F., Taketo, M. M., Birchmeier, W., Tenen, D. G. & Leutz, A. (2006) Hematopoietic stem cell and multilineage defects generated by constitutive beta-catenin activation. *Nat Immunol*, 7(10), 1037-47.
Schnittger, S., Dicker, F., Kern, W., Wendland, N., Sundermann, J., Alpermann, T., Haferlach, C. & Haferlach, T. (2011) RUNX1 mutations are frequent in de novo AML with noncomplex karyotype and confer an unfavorable prognosis. *Blood*, 117(8), 2348-57.
Schürch, C. M., Riether, C. & Ochsenbein, A. F. (2014) Cytotoxic CD8+ T cells stimulate hematopoietic progenitors by promoting cytokine release from bone marrow mesenchymal stromal cells. *Cell Stem Cell*, 14(4), 460-72.
Shaiken, T. E. & Opekun, A. R. (2014) Dissecting the cell to nucleus, perinucleus and cytosol. *Sci Rep*, 4, 4923.
Shang, S., Hua, F. & Hu, Z. W. (2017) The regulation of β-catenin activity and function in cancer: therapeutic opportunities. *Oncotarget*, 8(20), 33972-33989.
Shi, J., Chi, S., Xue, J., Yang, J., Li, F. & Liu, X. (2016) Emerging Role and Therapeutic Implication of Wnt Signaling Pathways in Autoimmune Diseases. *J Immunol Res*, 2016, 9392132.
Shimomura, A., Takasaki, A., Nomura, R., Hayashi, N. & Senda, T. (2013) Identification of DNA-dependent protein kinase catalytic subunit as a novel interaction partner of lymphocyte enhancer factor 1. *Med Mol Morphol*, 46(1), 14-9.
Shin, H. R., Islam, R., Yoon, W. J., Lee, T., Cho, Y. D., Bae, H. S., Kim, B. S., Woo, K. M., Baek, J. H. & Ryoo, H. M. (2016) Pin1-mediated Modification Prolongs the Nuclear Retention of β-Catenin in Wnt3a-induced Osteoblast Differentiation. *J Biol Chem*, 291(11), 5555-65.
Shivdasani, R. A., Rosenblatt, M. F., Zucker-Franklin, D., Jackson, C. W., Hunt, P., Saris, C. J. & Orkin, S. H. (1995) Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoietin/MGDF in megakaryocyte development. *Cell*, 81(5), 695-704.
Shtutman, M., Zhurinsky, J., Simcha, I., Albanese, C., D'Amico, M., Pestell, R. & Ben-Ze'ev, A. (1999) The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway. *Proc Natl Acad Sci U S A*, 96(10), 5522-7.
Siapati, E. K., Papadaki, M., Kozaou, Z., Rouka, E., Michali, E., Savvidou, I., Gogos, D., Kyriakou, D., Anagnostopoulos, N. I. & Vassilopoulos, G. (2011) Proliferation and bone marrow engraftment of AML blasts is dependent on β-catenin signalling. *Br J Haematol*, 152(2), 164-74.
Siar, C. H., Nagatsuka, H., Han, P. P., Buery, R. R., Tsujigiwa, H., Nakano, K., Ng, K. H. & Kawakami, T. (2012) Differential expression of canonical and non-canonical Wnt ligands in ameloblastoma. *J Oral Pathol Med*, 41(4), 332-9.
Simcha, I., Shtutman, M., Salomon, D., Zhurinsky, J., Sadot, E., Geiger, B. & Ben-Ze'ev, A. (1998) Differential nuclear translocation and transactivation potential of beta-catenin and plakoglobin. *J Cell Biol*, 141(6), 1433-48.
Simon, M., Grandage, V. L., Linch, D. C. & Khwaja, A. (2005) Constitutive activation of the Wnt/beta-catenin signalling pathway in acute myeloid leukaemia. *Oncogene*, 24(14), 2410-20.

Siriboonpiputtana, T., Zeisig, B. B., Zarowiecki, M., Fung, T. K., Mallardo, M., Tsai, C. T., Lau, P. N. I., Hoang, Q. C., Veiga, P., Barnes, J., Lynn, C., Wilson, A., Lenhard, B. & So, C. W. E. (2017) Transcriptional memory of cells of origin overrides β-catenin requirement of MLL cancer stem cells. *EMBO J*, 36(21), 3139-3155.
Skokowa, J., Cario, G., Uenalan, M., Schambach, A., Germeshausen, M., Battmer, K., Zeidler, C., Lehmann, U., Eder, M., Baum, C., Grosschedl, R., Stanulla, M., Scherr, M. & Welte, K. (2006) LEF-1 is crucial for neutrophil granulocytopoiesis and its expression is severely reduced in congenital neutropenia. *Nat Med*, 12(10), 1191-7.
Smith, G. C. & Jackson, S. P. (1999) The DNA-dependent protein kinase. *Genes Dev*, 13(8), 916-34.
Spangrude, G. J., Heimfeld, S. & Weissman, I. L. (1988) Purification and characterization of mouse hematopoietic stem cells. *Science*, 241(4861), 58-62.
Stamos, J. L. & Weis, W. I. (2013) The β-catenin destruction complex. *Cold Spring Harb Perspect Biol*, 5(1), a007898.
Steele, B. M., Harper, M. T., Macaulay, I. C., Morrell, C. N., Perez-Tamayo, A., Foy, M., Habas, R., Poole, A. W., Fitzgerald, D. J. & Maguire, P. B. (2009) Canonical Wnt signaling negatively regulates platelet function. *Proc Natl Acad Sci U S A*, 106(47), 19836-41.
Stölzel, F., Mohr, B., Kramer, M., Oelschlägel, U., Bochtler, T., Berdel, W. E., Kaufmann, M., Baldus, C. D., Schäfer-Eckart, K., Stuhlmann, R., Einsele, H., Krause, S. W., Serve, H., Hänel, M., Herbst, R., Neubauer, A., Sohlbach, K., Mayer, J., Middeke, J. M., Platzbecker, U., Schaich, M., Krämer, A., Röllig, C., Schetelig, J., Bornhäuser, M. & Ehninger, G. (2016) Karyotype complexity and prognosis in acute myeloid leukemia. *Blood Cancer J*, 6, e386.
Sugimoto, K., Toyoshima, H., Sakai, R., Miyagawa, K., Hagiwara, K., Ishikawa, F., Takaku, F., Yazaki, Y. & Hirai, H. (1992) Frequent mutations in the p53 gene in human myeloid leukemia cell lines. *Blood*, 79(9), 2378-83.
Sugiyama, T., Kohara, H., Noda, M. & Nagasawa, T. (2006) Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity*, 25(6), 977-88.
Sun, Y., Kolligs, F. T., Hottiger, M. O., Mosavin, R., Fearon, E. R. & Nabel, G. J. (2000) Regulation of beta -catenin transformation by the p300 transcriptional coactivator. *Proc Natl Acad Sci U S A*, 97(23), 12613-8.
Suzuki, M., Kobayashi-Osaki, M., Tsutsumi, S., Pan, X., Ohmori, S., Takai, J., Moriguchi, T., Ohneda, O., Ohneda, K., Shimizu, R., Kanki, Y., Kodama, T., Aburatani, H. & Yamamoto, M. (2013) GATA factor switching from GATA2 to GATA1 contributes to erythroid differentiation. *Genes Cells*, 18(11), 921-33.
Tallman, M. S. (2005) New strategies for the treatment of acute myeloid leukemia including antibodies and other novel agents. *Hematology Am Soc Hematol Educ Program*, 143-50.
Tandon, B., Peterson, L., Gao, J., Nelson, B., Ma, S., Rosen, S. & Chen, Y. H. (2011) Nuclear overexpression of lymphoid-enhancer-binding factor 1 identifies chronic lymphocytic leukemia/small lymphocytic lymphoma in small B-cell lymphomas. *Mod Pathol*, 24(11), 1433-43.
Taussig, D. C., Miraki-Moud, F., Anjos-Afonso, F., Pearce, D. J., Allen, K., Ridler, C., Lillington, D., Oakervee, H., Cavenagh, J., Agrawal, S. G., Lister, T. A., Gribben, J. G. & Bonnet, D. (2008) Anti-CD38 antibody-mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. *Blood*, 112(3), 568-75.

Thiede, C., Koch, S., Creutzig, E., Steudel, C., Illmer, T., Schaich, M. & Ehninger, G. (2006) Prevalence and prognostic impact of NPM1 mutations in 1485 adult patients with acute myeloid leukemia (AML). *Blood*, 107(10), 4011-20.
Tickenbrock, L., Hehn, S., Sargin, B., Choudhary, C., Bäumer, N., Buerger, H., Schulte, B., Müller, O., Berdel, W. E., Müller-Tidow, C. & Serve, H. (2008) Activation of Wnt signalling in acute myeloid leukemia by induction of Frizzled-4. *Int J Oncol*, 33(6), 1215-21.
Tickenbrock, L., Schwäble, J., Wiedehage, M., Steffen, B., Sargin, B., Choudhary, C., Brandts, C., Berdel, W. E., Müller-Tidow, C. & Serve, H. (2005) Flt3 tandem duplication mutations cooperate with Wnt signaling in leukemic signal transduction. *Blood*, 105(9), 3699-706.
Tonks, A., Tonks, A. J., Pearn, L., Mohamad, Z., Burnett, A. K. & Darley, R. L. (2005) Optimized retroviral transduction protocol which preserves the primitive subpopulation of human hematopoietic cells. *Biotechnol Prog*, 21(3), 953-8.
Townsley, F. M., Thompson, B. & Bienz, M. (2004) Pygopus residues required for its binding to Legless are critical for transcription and development. *J Biol Chem*, 279(7), 5177-83.
Trinkle-Mulcahy, L., Boulon, S., Lam, Y. W., Urcia, R., Boisvert, F. M., Vandermoere, F., Morrice, N. A., Swift, S., Rothbauer, U., Leonhardt, H. & Lamond, A. (2008) Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J Cell Biol*, 183(2), 223-39.
Trowbridge, J. J., Xenocostas, A., Moon, R. T. & Bhatia, M. (2006) Glycogen synthase kinase-3 is an in vivo regulator of hematopoietic stem cell repopulation. *Nat Med*, 12(1), 89-98.
Udhayakumar, G., Jayanthi, V., Devaraj, N. & Devaraj, H. (2007) Interaction of MUC1 with beta-catenin modulates the Wnt target gene cyclinD1 in H. pylori-induced gastric cancer. *Mol Carcinog*, 46(9), 807-17.
Ushach, I. & Zlotnik, A. (2016) Biological role of granulocyte macrophage colony-stimulating factor (GM-CSF) and macrophage colony-stimulating factor (M-CSF) on cells of the myeloid lineage. *J Leukoc Biol*, 100(3), 481-9.
Vaidya, A. & Kale, V. P. (2015) TGF-β signaling and its role in the regulation of hematopoietic stem cells. *Syst Synth Biol*, 9(1-2), 1-10.
Valenta, T., Hausmann, G. & Basler, K. (2012) The many faces and functions of β-catenin. *EMBO J*, 31(12), 2714-36.
Van Den Berg, D. J., Sharma, A. K., Bruno, E. & Hoffman, R. (1998) Role of members of the Wnt gene family in human hematopoiesis. *Blood*, 92(9), 3189-202.
Vanli, G., Cuesta-Marban, A. & Widmann, C. (2017) Evaluation and validation of commercial antibodies for the detection of Shb. *PLoS One*, 12(12), e0188311.
Vardiman, J. W., Harris, N. L. & Brunning, R. D. (2002) The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood*, 100(7), 2292-302.
Velten, L., Haas, S. F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B. P., Hirche, C., Lutz, C., Buss, E. C., Nowak, D., Boch, T., Hofmann, W. K., Ho, A. D., Huber, W., Trumpp, A., Essers, M. A. & Steinmetz, L. M. (2017) Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol*, 19(4), 271-281.
Visnjic, D., Kalajzic, Z., Rowe, D. W., Katavic, V., Lorenzo, J. & Aguila, H. L. (2004) Hematopoiesis is severely altered in mice with an induced osteoblast deficiency. *Blood*, 103(9), 3258-64.

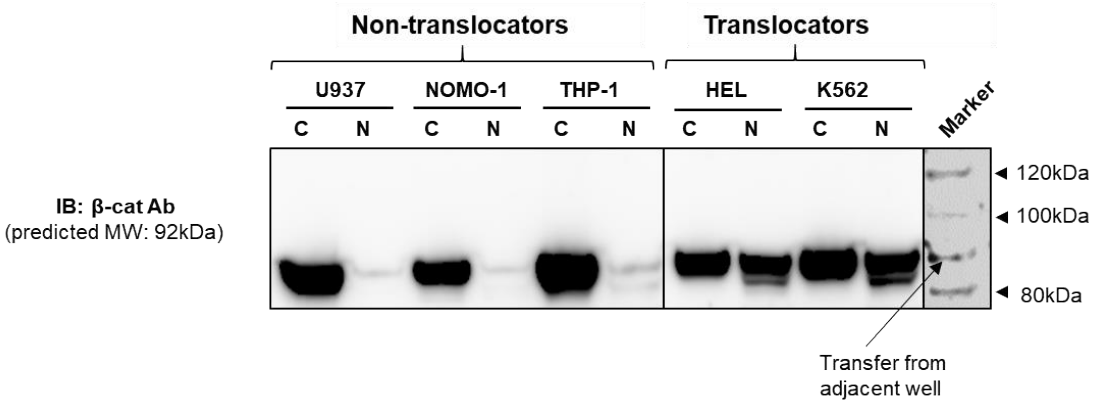
von Kries, J. P., Winbeck, G., Asbrand, C., Schwarz-Romond, T., Sochnikova, N., Dell'Oro, A., Behrens, J. & Birchmeier, W. (2000) Hot spots in beta-catenin for interactions with LEF-1, conductin and APC. *Nat Struct Biol*, 7(9), 800-7.
Wang, Y., Krivtsov, A. V., Sinha, A. U., North, T. E., Goessling, W., Feng, Z., Zon, L. I. & Armstrong, S. A. (2010) The Wnt/beta-catenin pathway is required for the development of leukemia stem cells in AML. *Science*, 327(5973), 1650-3.
Warsito, D., Sjöström, S., Andersson, S., Larsson, O. & Sehat, B. (2012) Nuclear IGF1R is a transcriptional co-activator of LEF1/TCF. *EMBO Rep*, 13(3), 244-50.
Weber, J. M. & Calvi, L. M. (2010) Notch signaling and the bone marrow hematopoietic stem cell niche. *Bone*, 46(2), 281-5.
Wein, F., Pietsch, L., Saffrich, R., Wuchter, P., Walenda, T., Bork, S., Horn, P., Diehlmann, A., Eckstein, V., Ho, A. D. & Wagner, W. (2010) N-cadherin is expressed on human hematopoietic progenitor cells and mediates interaction with human mesenchymal stromal cells. *Stem Cell Res*, 4(2), 129-39.
Weisberg, E., Nonami, A., Chen, Z., Nelson, E., Chen, Y., Liu, F., Cho, H., Zhang, J., Sattler, M., Mitsiades, C., Wong, K. K., Liu, Q., Gray, N. S. & Griffin, J. D. (2014) Upregulation of IGF1R by mutant RAS in leukemia and potentiation of RAS signaling inhibitors by small-molecule inhibition of IGF1R. *Clin Cancer Res*, 20(21), 5483-95.
Weiss, M. J., Yu, C. & Orkin, S. H. (1997) Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. *Mol Cell Biol*, 17(3), 1642-51.
Weitsman, G. E., Skliris, G., Ung, K., Peng, B., Younes, M., Watson, P. H. & Murphy, L. C. (2006) Assessment of multiple different estrogen receptor-beta antibodies for their ability to immunoprecipitate under chromatin immunoprecipitation conditions. *Breast Cancer Res Treat*, 100(1), 23-31.
Wiechens, N. & Fagotto, F. (2001) CRM1- and Ran-independent nuclear export of beta-catenin. *Curr Biol*, 11(1), 18-27.
Wiechens, N., Heinle, K., Englmeier, L., Schohl, A. & Fagotto, F. (2004) Nucleo-cytoplasmic shuttling of Axin, a negative regulator of the Wnt-beta-catenin Pathway. *J Biol Chem*, 279(7), 5263-7.
Wielenga, V. J., Smits, R., Korinek, V., Smit, L., Kielman, M., Fodde, R., Clevers, H. & Pals, S. T. (1999) Expression of CD44 in Apc and Tcf mutant mice implies regulation by the WNT pathway. *Am J Pathol*, 154(2), 515-23.
Willert, K. & Jones, K. A. (2006) Wnt signaling: is the party in the nucleus? *Genes Dev*, 20(11), 1394-404.
Wilson, A., Murphy, M. J., Oskarsson, T., Kaloulis, K., Bettess, M. D., Oser, G. M., Pasche, A. C., Knabenhans, C., Macdonald, H. R. & Trumpp, A. (2004) c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. *Genes Dev*, 18(22), 2747-63.
Wu, D. & Pan, W. (2010) GSK3: a multifaceted kinase in Wnt signaling. *Trends Biochem Sci*, 35(3), 161-8.
Wu, P., Shi, K. J., An, J. J., Ci, Y. L., Li, F., Hui, K. Y., Yang, Y. & Xu, C. M. (2014) The LEF1/CYLD axis and cIAPs regulate RIP1 deubiquitination and trigger apoptosis in selenite-treated colorectal cancer cells. *Cell Death Dis*, 5, e1085.
Xiao, C. Y., Hübner, S. & Jans, D. A. (1997) SV40 large tumor antigen nuclear import is regulated by the double-stranded DNA-dependent protein kinase site (serine 120) flanking the nuclear localization sequence. *J Biol Chem*, 272(35), 22191-8.
Xu, J., Suzuki, M., Niwa, Y., Hiraga, J., Nagasaka, T., Ito, M., Nakamura, S., Tomita, A., Abe, A., Kiyoi, H., Kinoshita, T. & Naoe, T. (2008) Clinical significance

of nuclear non-phosphorylated beta-catenin in acute myeloid leukaemia and myelodysplastic syndrome. *Br J Haematol*, 140(4), 394-401.
Xu, W. & Kimelman, D. (2007) Mechanistic insights from structural studies of beta-catenin and its binding partners. *J Cell Sci*, 120(Pt 19), 3337-44.
Xu, W. D., Wang, J., Yuan, T. L., Li, Y. H., Yang, H., Liu, Y., Zhao, Y. & Herrmann, M. (2016) Interactions between canonical Wnt signaling pathway and MAPK pathway regulate differentiation, maturation and function of dendritic cells. *Cell Immunol*, 310, 170-177.
Yamaguchi, Y., Suda, T., Suda, J., Eguchi, M., Miura, Y., Harada, N., Tominaga, A. & Takatsu, K. (1988) Purified interleukin 5 supports the terminal differentiation and proliferation of murine eosinophilic precursors. *J Exp Med*, 167(1), 43-56.
Yang, C. M., Ji, S., Li, Y., Fu, L. Y., Jiang, T. & Meng, F. D. (2017) β-Catenin promotes cell proliferation, migration, and invasion but induces apoptosis in renal cell carcinoma. *Onco Targets Ther*, 10, 711-724.
Yang, L., Zhang, H. & Bruce, J. E. (2009) Optimizing the detergent concentration conditions for immunoprecipitation (IP) coupled with LC-MS/MS identification of interacting proteins. *Analyst*, 134(4), 755-62.
Yeh, J. R., Munson, K. M., Elagib, K. E., Goldfarb, A. N., Sweetser, D. A. & Peterson, R. T. (2009) Discovering chemical modifiers of oncogene-regulated hematopoietic differentiation. *Nat Chem Biol*, 5(4), 236-43.
Yeung, J., Esposito, M. T., Gandillet, A., Zeisig, B. B., Griessinger, E., Bonnet, D. & So, C. W. (2010) β-Catenin mediates the establishment and drug resistance of MLL leukemic stem cells. *Cancer Cell*, 18(6), 606-18.
Yin, T. & Li, L. (2006) The stem cell niches in bone. *J Clin Invest*, 116(5), 1195-201.
Yokoya, F., Imamoto, N., Tachibana, T. & Yoneda, Y. (1999) beta-catenin can be transported into the nucleus in a Ran-unassisted manner. *Mol Biol Cell*, 10(4), 1119-31.
Ysebaert, L., Chicanne, G., Demur, C., De Toni, F., Prade-Houdellier, N., Ruidavets, J. B., Mansat-De Mas, V., Rigal-Huguet, F., Laurent, G., Payrastre, B., Manenti, S. & Racaud-Sultan, C. (2006) Expression of beta-catenin by acute myeloid leukemia cells predicts enhanced clonogenic capacities and poor prognosis. *Leukemia*, 20(7), 1211-6.
Yu, S., Li, F., Xing, S., Zhao, T., Peng, W. & Xue, H. H. (2016) Hematopoietic and Leukemic Stem Cells Have Distinct Dependence on Tcf1 and Lef1 Transcription Factors. *J Biol Chem*, 291(21), 11148-60.
Zeng, X., Huang, H., Tamai, K., Zhang, X., Harada, Y., Yokota, C., Almeida, K., Wang, J., Doble, B., Woodgett, J., Wynshaw-Boris, A., Hsieh, J. C. & He, X. (2008) Initiation of Wnt signaling: control of Wnt coreceptor Lrp6 phosphorylation/activation via frizzled, dishevelled and axin functions. *Development*, 135(2), 367-75.
Zeng, X., Tamai, K., Doble, B., Li, S., Huang, H., Habas, R., Okamura, H., Woodgett, J. & He, X. (2005) A dual-kinase mechanism for Wnt co-receptor phosphorylation and activation. *Nature*, 438(7069), 873-7.
Zhang, J., Niu, C., Ye, L., Huang, H., He, X., Tong, W. G., Ross, J., Haug, J., Johnson, T., Feng, J. Q., Harris, S., Wiedemann, L. M., Mishina, Y. & Li, L. (2003) Identification of the haematopoietic stem cell niche and control of the niche size. *Nature*, 425(6960), 836-41.
Zhang, N., Wei, P., Gong, A., Chiu, W. T., Lee, H. T., Colman, H., Huang, H., Xue, J., Liu, M., Wang, Y., Sawaya, R., Xie, K., Yung, W. K., Medema, R. H., He, X. & Huang, S. (2011) FoxM1 promotes β-catenin nuclear localization and

controls Wnt target-gene expression and glioma tumorigenesis. *Cancer Cell*, 20(4), 427-42.
Zhang, T., Otevrel, T., Gao, Z., Ehrlich, S. M., Fields, J. Z. & Boman, B. M. (2001) Evidence that APC regulates survivin expression: a possible mechanism contributing to the stem cell origin of colon cancer. *Cancer Res*, 61(24), 8664-7.
Zhao, C., Blum, J., Chen, A., Kwon, H. Y., Jung, S. H., Cook, J. M., Lagoo, A. & Reya, T. (2007) Loss of beta-catenin impairs the renewal of normal and CML stem cells in vivo. *Cancer Cell*, 12(6), 528-41.
Zheng, X., Beissert, T., Kukoc-Zivojnov, N., Puccetti, E., Altschmied, J., Strolz, C., Boehrer, S., Gul, H., Schneider, O., Ottmann, O. G., Hoelzer, D., Henschler, R. & Ruthardt, M. (2004) Gamma-catenin contributes to leukemogenesis induced by AML-associated translocation products by increasing the self-renewal of very primitive progenitor cells. *Blood*, 103(9), 3535-43.
Zhou, F., Shen, Q. & Claret, F. X. (2013) Novel roles of reactive oxygen species in the pathogenesis of acute myeloid leukemia. *J Leukoc Biol*, 94(3), 423-9.
Zhou, X. Z. & Lu, K. P. (2016) The isomerase PIN1 controls numerous cancer-driving pathways and is a unique drug target. *Nat Rev Cancer*, 16(7), 463-78.

Appendices

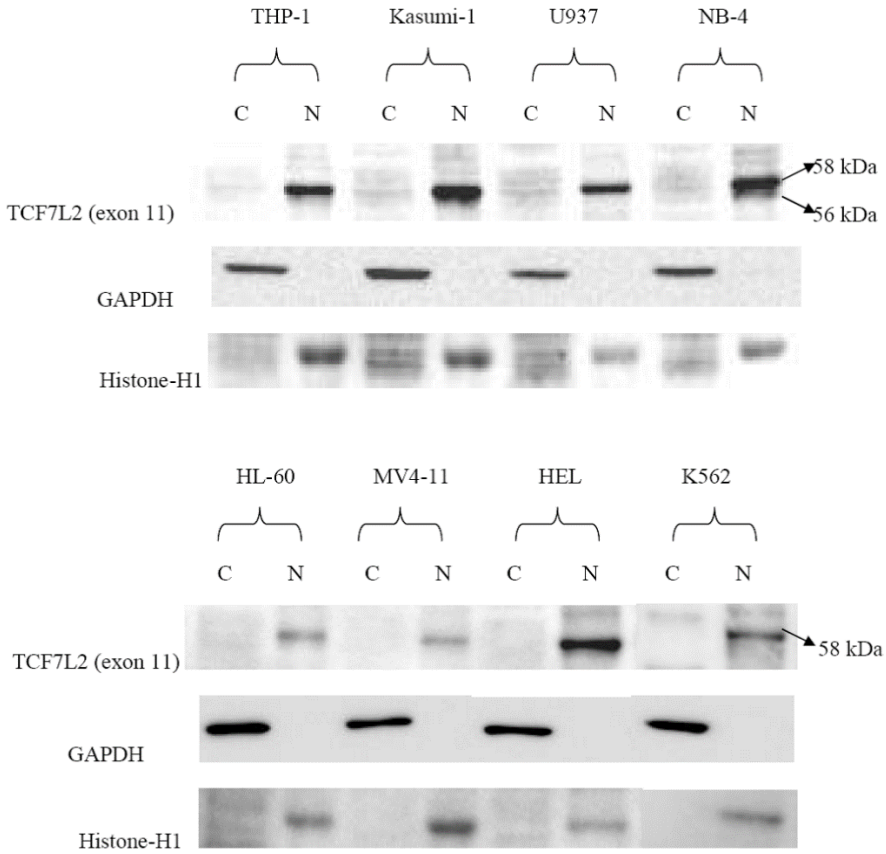
Appendix 1



β-catenin translocation status of cell lines used in this study.

Western blotting of cell lines used in this study. Cells were fractionated using CHAPS buffer (2.6.2) and western blotted using β-catenin antibody (BD) (2.5). Abbreviations; IB= immunoblot, Ab= antibody, MW= molecular weight, kDa= kilodaltons, C= cytoplasm, N= nuclear. Marker= MagicMark XP (Thermo Fisher).

Appendix 2



Distribution of TCF7L2 in myeloid cell lines.

TCF7L2 is present in both translocator and non-translocator cell lines. This figure was created by Dr. Sara Daud.

Appendix 3

Method: Whole cell lysis using Triton buffer

To extract whole cell lysates, cell pellets were thawed on ice in the presence of 1µL DNase (1mg/mL) for 5 min, tapping the tube regularly. 50 µL of Triton buffer (0.25µM sucrose, 10mM HEPES-KOH (pH 7.2), 1mM magnesium acetate, 0.5mM EDTA, 0.5mM EGTA, 10mM BME, 1mM sodium orthovanadate, PIC, *v/v* 1% Triton X-100) was added to the cells and the tube was incubated on ice for 30 min with occasional vortexing. The tubes were centrifuged at 3000 x g for 5 min at 4 °C. The supernatant (cell lysate) was aspirated into a clean 1.5mL Eppendorf tube, snap frozen in liquid nitrogen and stored at -80 °C.

Appendix 4

To view the raw mass spectrometry data generated in this project, please see the attached CD disk.

Appendix 5

The following sequencing data is for the pHIV-EGFP FLAG-β-catenin vector.

Primers used:

- pHIVFor - this is the forward primer for the pHIV-EGFP vector. This was bought from Sigma.
- pHIVRev - this is the reverse primer for the pHIV-EGFP vector. This was bought from Sigma.
- b-catP1(TTCCAGACACGCTATCATGC) - A β-catenin primer designed using Primer3.
- b-catP2 (TGCAGTTCGCCTTCACTATG) - A β-catenin primer designed using Primer3.

These primers combined gave full coverage of the β-catenin-FLAG insert.

The sequencing data was compared to the predicted sequence of the pHIV-EGFP-β-catenin-FLAG vector (predicted-vector-full).

During the first sequencing run using the b-catP1 primer, an extra T was detected. In a repeat sequencing run the extra T was not detected, suggesting that there may have been a sequencing error in the first run.

```
predicted-vector-full      CGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCG 60
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      GAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGAT 120
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      CGCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATA 180
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACAT 240
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      CAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAG 300
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      AACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGA 360
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCA 420
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0
```

```
predicted-vector-full      CCGCACAGCAAGCGGCCGGCCGCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGGGAC 480
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      AATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCA 540
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      CCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCT 600
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATGACGCTG 660
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      ACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGG 720
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      GCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG 780
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGT 840
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAA 900
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0
```

```
predicted-vector-full      TCTCTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAAT 960
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAA 1020
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      CAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAAT 1080
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATA 1140
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      GTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTT 1200
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      CAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGT 1260
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      GGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCGGCACTGCGTGCGCCA 1320
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      ATTCTGCAGACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGG 1380
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      GTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATT 1440
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0
```

```
predicted-vector-full      ACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAGATCC 1500
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      AGTTTGGTTAGTACCGGGCCCGCTCTAGCCGTGAGGCTCCGGTGCCCGTCAGTGGGCAGA 1560
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      GCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGC 1620
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      CTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTT 1680
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCG 1740
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      CAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCT 1800
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      CTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGAT 1860
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTCGAGGCCTTGCGCTTAAGG 1920
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      AGCCCCTTCGCCTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCGTGCG 1980
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0
```

```
predicted-vector-full      AATCTGGTGGCACCTTCGCGCCTGTCTCGCTGCTTTCGATAAGTCTCTAGCCATTTAAAA 2040
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TTTTTGATGACCTGCTGCGACGCTTTTTTTCTGGCAAGATAGTCTTGTAAATGCGGGCCA 2100
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      AGATCTGCACACTGGTATTTCGGTTTTTGGGGCCGCGGGCGGCGACGGGGCCCGTGCGTC 2160
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      CCAGCGCACATGTTCGGCGAGGCGGGGCCTGCGAGCGCGGCCACCGAGAATCGGACGGGG 2220
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      GTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGCCGTGTATCGCCCC 2280
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      GCCCTGGGCGGCAAGGCTGGCCCGGTCGGCACCAGTTGCGTGAGCGGAAAGATGGCCGCT 2340
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGG 2400
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0

predicted-vector-full      TGAGTCACCCACACAAAGGAAAAGGGCCTTTCCGTCCTCAGCCGTCGCTTCATGTGACTC 2460
pHIVFor                    ------------------------------------------------------------ 0
insert-predicted           ------------------------------------------------------------ 0
b-catP1                    ------------------------------------------------------------ 0
b-catP1(repeat)            ------------------------------------------------------------ 0
b-catP2                    ------------------------------------------------------------ 0
pHIVRev                    ------------------------------------------------------------ 0
```

```
predicted-vector-full   CACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGTC 2520
pHIVFor                 ------------------------------------------------------------ 0
insert-predicted        ------------------------------------------------------------ 0
b-catP1                 ------------------------------------------------------------ 0
b-catP1(repeat)         ------------------------------------------------------------ 0
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   GTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGA 2580
pHIVFor                 ------------------------------------------------------------ 0
insert-predicted        ------------------------------------------------------------ 0
b-catP1                 ------------------------------------------------------------ 0
b-catP1(repeat)         ------------------------------------------------------------ 0
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   GACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAATTTGCCCTTTTTGA 2640
pHIVFor                 ------------------------------------------------------------ 0
insert-predicted        ------------------------------------------------------------ 0
b-catP1                 ------------------------------------------------------------ 0
b-catP1(repeat)         ------------------------------------------------------------ 0
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   GTTTGGATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTCTTCCATT 2700
pHIVFor                 ----------------------------AGACAGTGGTTCAAAGTTTTTTTCTTCCATT 31
insert-predicted        ------------------------------------------------------------ 0
b-catP1                 ------------------------------------------------------------ 0
b-catP1(repeat)         ------------------------------------------------------------ 0
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   TCAGGTGTCGTGAGCGGCCGCTGAGTTCGAGCTCGGATCCCCAGCGTGGACAATGGCTAC 2760
pHIVFor                 TCAGGTGTCGTGAGCGGCCGCTGAGTTCGAGCTCGGATCCCCAGCGTGGACAATGGCTAC 91
insert-predicted        ---------------------------CGAGCTCGGATCCCCAGCGTGGACAATGGCTAC 33
b-catP1                 ------------------------------------------------------------ 0
b-catP1(repeat)         ------------------------------------------------------------ 0
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0
                                                  START codon
predicted-vector-full   TCAAGCTGATTTGATGGAGTTGGACATGGCCATGGAACCAGACAGAAAAGCGGCTGTTAG 2820
pHIVFor                 TCAAGCTGATTTGATGGAGTTGGACATGGCCATGGAACCAGACAGAAAAGCGGCTGTTAG 151
insert-predicted        TCAAGCTGATTTGATGGAGTTGGACATGGCCATGGAACCAGACAGAAAAGCGGCTGTTAG 93
b-catP1                 ------------------------------------------------------------ 0
b-catP1(repeat)         ------------------------------------------------------------ 0
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   TCACTGGCAGCAACAGTCTTACCTGGACTCTGGAATCCATTCTGGTGCCACTACCACAGC 2880
pHIVFor                 TCACTGGCAGCAACAGTCTTACCTGGACTCTGGAATCCATTCTGGTGCCACTACCACAGC 211
insert-predicted        TCACTGGCAGCAACAGTCTTACCTGGACTCTGGAATCCATTCTGGTGCCACTACCACAGC 153
b-catP1                 ------------------------------------------------------------ 0
b-catP1(repeat)         ------------------------------------------------------------ 0
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0
```

```
predicted-vector-full     TCCTTCTCTGAGTGGTAAAGGCAATCCTGAGGAAGAGGATGTGGATACCTCCCAAGTCCT 2940
pHIVFor                   TCCTTCTCTGAGTGGTAAAGGCAATCCTGAGGAAGAGGATGTGGATACCTCCCAAGTCCT 271
insert-predicted          TCCTTCTCTGAGTGGTAAAGGCAATCCTGAGGAAGAGGATGTGGATACCTCCCAAGTCCT 213
b-catP1                   ------------------------------------------------------------ 0
b-catP1(repeat)           ------------------------------------------------------------ 0
b-catP2                   ------------------------------------------------------------ 0
pHIVRev                   ------------------------------------------------------------ 0

predicted-vector-full     GTATGAGTGGGAACAGGGATTTTCTCAGTCCTTCACTCAAGAACAAGTAGCTGATATTGA 3000
pHIVFor                   GTATGAGTGGGAACAGGGATTTTCTCAGTCCTTCACTCAAGAACAAGTAGCTGATATTGA 331
insert-predicted          GTATGAGTGGGAACAGGGATTTTCTCAGTCCTTCACTCAAGAACAAGTAGCTGATATTGA 273
b-catP1                   ------------------------------------------------------------ 0
b-catP1(repeat)           ------------------------------------------------------------ 0
b-catP2                   ------------------------------------------------------------ 0
pHIVRev                   ------------------------------------------------------------ 0

predicted-vector-full     TGGACAGTATGCAATGACTCGAGCTCAGAGGGTACGAGCTGCTATGTTCCCTGAGACATT 3060
pHIVFor                   TGGACAGTATGCAATGACTCGAGCTCAGAGGGTACGAGCTGCTATGTTCCCTGAGACATT 391
insert-predicted          TGGACAGTATGCAATGACTCGAGCTCAGAGGGTACGAGCTGCTATGTTCCCTGAGACATT 333
b-catP1                   ------------------------------------------------------------ 0
b-catP1(repeat)           ------------------------------------------------------------ 0
b-catP2                   ------------------------------------------------------------ 0
pHIVRev                   ------------------------------------------------------------ 0

predicted-vector-full     AGATGAGGGCATGCAGATCCCATCTACACAGTTTGATGCTGCTCATCCCACTAATGTCCA 3120
pHIVFor                   AGATGAGGGCATGCAGATCCCATCTACACAGTTTGATGCTGCTCATCCCACTAATGTCCA 451
insert-predicted          AGATGAGGGCATGCAGATCCCATCTACACAGTTTGATGCTGCTCATCCCACTAATGTCCA 393
b-catP1                   ------------------------------------------------------------ 0
b-catP1(repeat)           ------------------------------------------------------------ 0
b-catP2                   ------------------------------------------------------------ 0
pHIVRev                   ------------------------------------------------------------ 0

predicted-vector-full     GCGTTTGGCTGAACCATCACAGATGCTGAAACATGCAGTTGTAAACTTGATTAACTATCA 3180
pHIVFor                   GCGTTTGGCTGAACCATCACAGATGCTGAAACATGCAGTTGTAAACTTGATTAACTATCA 511
insert-predicted          GCGTTTGGCTGAACCATCACAGATGCTGAAACATGCAGTTGTAAACTTGATTAACTATCA 453
b-catP1                   ------------------------------------------------------------ 0
b-catP1(repeat)           ------------------------------------------------------------ 0
b-catP2                   ------------------------------------------------------------ 0
pHIVRev                   ------------------------------------------------------------ 0

predicted-vector-full     AGATGATGCAGAACTTGCCACACGTGCAATCCCTGAACTGACAAAACTGCTAAATGACGA 3240
pHIVFor                   AGATGATGCAGAACTTGCCACACGTGCAATCCCTGAACTGACAAAACTGCTAAATGACGA 571
insert-predicted          AGATGATGCAGAACTTGCCACACGTGCAATCCCTGAACTGACAAAACTGCTAAATGACGA 513
b-catP1                   ------------------------------------------------------------ 0
b-catP1(repeat)           ------------------------------------------------------------ 0
b-catP2                   ------------------------------------------------------------ 0
pHIVRev                   ------------------------------------------------------------ 0

predicted-vector-full     GGACCAGGTGGTGGTTAATAAGGCTGCAGTTATGGTCCATCAGCTTTCTAAAAAGGAAGC 3300
pHIVFor                   GGACCAGGTGGTGGTTAATAAGGCTGCAGTTATGGTCCATCAGCTTTCTAAAAAGGAAGC 631
insert-predicted          GGACCAGGTGGTGGTTAATAAGGCTGCAGTTATGGTCCATCAGCTTTCTAAAAAGGAAGC 573
b-catP1                   ------------------------------------------------------------ 0
b-catP1(repeat)           ------------------------------------------------------------ 0
b-catP2                   ------------------------------------------------------------ 0
pHIVRev                   ------------------------------------------------------------ 0

predicted-vector-full     TTCCAGACACGCTATCATGCGTTCTCCTCAGATGGTGTCTGCTATTGTACGTACCATGCA 3360
pHIVFor                   TTCCAGACACGCTATCATGCGTTCTCCTCAGATGGTGTCTGCTATTGTACGTACCATGCA 691
insert-predicted          TTCCAGACACGCTATCATGCGTTCTCCTCAGATGGTGTCTGCTATTGTACGTACCATGCA 633
b-catP1                   ------------------------------------TCTGCTATTGTACGTACCATGCA 23
b-catP1(repeat)           ------------------------------------TCTGCTATTGTACGTACCATGCA 23
b-catP2                   ------------------------------------------------------------ 0
pHIVRev                   ------------------------------------------------------------ 0
```

```
predicted-vector-full   GAATACAAATGATGTAGAAACAGCTCGTTGTACCGCTGGGACCTTGCATAACCTTTCCCA 3420
pHIVFor                 GAATACAAATGATGTAGAAACAGCTCGTT------------------------------ 720
insert-predicted        GAATACAAATGATGTAGAAACAGCTCGTTGTACCGCTGGGACCTTGCATAACCTTTCCCA 693
b-catP1                 GAATACAAATGATGTAGAAACAGCTCGTTGTACCGCTGGGACCTTGCATAACCTTTCCCA 83
b-catP1(repeat)         GAATACAAATGATGTAGAAACAGCTCGTTGTACCGCTGGGACCTTGCATAACCTTTCCCA 83
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   TCATCGTGAGGGCTTACTGGCCATCTTTAAGTCTGGAGGCATTCCTGCCCTGGTGAAAAT 3480
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TCATCGTGAGGGCTTACTGGCCATCTTTAAGTCTGGAGGCATTCCTGCCCTGGTGAAAAT 753
b-catP1                 TCATCGTGAGGGCTTACTGGCCATCTTTAAGTCTGGAGGCATTCCTGCCCTGGTGAAAAT 143
b-catP1(repeat)         TCATCGTGAGGGCTTACTGGCCATCTTTAAGTCTGGAGGCATTCCTGCCCTGGTGAAAAT 143
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   GCTTGGTTCACCAGTGGATTCTGTGTTGTTTTATGCCATTACAACTCTCCACAACCTTTT 3540
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        GCTTGGTTCACCAGTGGATTCTGTGTTGTTTTATGCCATTACAACTCTCCACAACCTTTT 813
b-catP1                 GCTTGGTTCACCAGTGGATTCTGTGTTGTTTTATGCCATTACAACTCTCCACAACCTTTT 203
b-catP1(repeat)         GCTTGGTTCACCAGTGGATTCTGTGTTGTTTTATGCCATTACAACTCTCCACAACCTTTT 203
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   ATTACATCAAGAAGGAGCTAAAATGGCAGTGCGTTTAGCTGGTGGGCTGCAGAAAATGGT 3600
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ATTACATCAAGAAGGAGCTAAAATGGCAGTGCGTTTAGCTGGTGGGCTGCAGAAAATGGT 873
b-catP1                 ATTACATCAAGAAGGAGCTAAAATGGCAGTGCGTTTAGCTGGTGGGCTGCAGAAAATGGT 263
b-catP1(repeat)         ATTACATCAAGAAGGAGCTAAAATGGCAGTGCGTTTAGCTGGTGGGCTGCAGAAAATGGT 263
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   TGCCTTGCTCAACAAAACAAATGTTAAATTCTTGGCTATTACGACAGACTGCCTTCAAAT 3660
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TGCCTTGCTCAACAAAACAAATGTTAAATTCTTGGCTATTACGACAGACTGCCTTCAAAT 933
b-catP1                 TGCCTTGCTCAACAAAACAAATGTTAAATTCTTGGCTATTACGACAGACTGCCTTCAAAT 323
b-catP1(repeat)         TGCCTTGCTCAACAAAACAAATGTTAAATTCTTGGCTATTACGACAGACTGCCTTCAAAT 323
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   TTTAGCTTATGGCAACCAAGAAAGCAAGCTCATCATACTGGCTAGTGGTGGACCCCAAGC 3720
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TTTAGCTTATGGCAACCAAGAAAGCAAGCTCATCATACTGGCTAGTGGTGGACCCCAAGC 993
b-catP1                 TTTAGCTTATGGCAACCAAGAAAGCAAGCTCATCATACTGGCTAGTGGTGGACCCCAAGC 383
b-catP1(repeat)         TTTAGCTTATGGCAACCAAGAAAGCAAGCTCATCATACTGGCTAGTGGTGGACCCCAAGC 383
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   TTTAGTAAATATAATGAGGACCTATACTTACGAAAAACTACTGTGGACCACAAGCAGAGT 3780
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TTTAGTAAATATAATGAGGACCTATACTTACGAAAAACTACTGTGGACCACAAGCAGAGT 1053
b-catP1                 TTTAGTAAATATAATGAGGACCTATACTTACGAAAAACTACTGTGGACCACAAGCAGAGT 443
b-catP1(repeat)         TTTAGTAAATATAATGAGGACCTATACTTACGAAAAACTACTGTGGACCACAAGCAGAGT 443
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   GCTGAAGGTGCTATCTGTCTGCTCTAGTAATAAGCCGGCTATTGTAGAAGCTGGTGGAAT 3840
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        GCTGAAGGTGCTATCTGTCTGCTCTAGTAATAAGCCGGCTATTGTAGAAGCTGGTGGAAT 1113
b-catP1                 GCTGAAGGTGCTATCTGTCTGCTCTAGTAATAAGCCGGCTATTGTAGAAGCTGGTGGAAT 503
b-catP1(repeat)         GCTGAAGGTGCTATCTGTCTGCTCTAGTAATAAGCCGGCTATTGTAGAAGCTGGTGGAAT 503
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0
```

```
predicted-vector-full   GCAAGCTTTAGGACTTCACCTGACAGATCCAAGTCAACGTCTTGTTCAGAACTGTCTTTG 3900
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        GCAAGCTTTAGGACTTCACCTGACAGATCCAAGTCAACGTCTTGTTCAGAACTGTCTTTG 1173
b-catP1                 GCAAGCTTTAGGACTTCACCTGACAGATCCAAGTCAACGTCTTGTTCAGAACTGTCTTTG 563
b-catP1(repeat)         GCAAGCTTTAGGACTTCACCTGACAGATCCAAGTCAACGTCTTGTTCAGAACTGTCTTTG 563
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   GACTCTCAGGAATCTTTCAGATGCTGCAACTAAACAGGAAGGGATGGAAGGTCTCCTTGG 3960
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        GACTCTCAGGAATCTTTCAGATGCTGCAACTAAACAGGAAGGGATGGAAGGTCTCCTTGG 1233
b-catP1                 GACTCTCAGGAATCTTTCAGATGCTGCAACTAAACAGGAAGGGATGGAAGGTCTCCTTGG 623
b-catP1(repeat)         GACTCTCAGGAATCTTTCAGATGCTGCAACTAAACAGGAAGGGATGGAAGGTCTCCTTGG 623
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   GACTCTTGTTCAGCTTCTGGGTTCAGATGATATAAATGTGGTCACCTGTGCAGCTGGAAT 4020
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        GACTCTTGTTCAGCTTCTGGGTTCAGATGATATAAATGTGGTCACCTGTGCAGCTGGAAT 1293
b-catP1                 GACTCTTGTTCAGCTTCTGGGTTCAGATGATATAAATGTGGTCACCTGTGCAGCTGGAAT 683
b-catP1(repeat)         GACTCTTGTTCAGCTTCTGGGTTCAGATGATATAAATGTGGTCACCTGTGCAGCTGGAAT 683
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   TCTTTCTAACCTCACTTGCAATAATTATAAGAACAAGATGATGGTCTGCCAAGTGGGTGG 4080
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TCTTTCTAACCTCACTTGCAATAATTATAAGAACAAGATGATGGTCTGCCAAGTGGGTGG 1353
b-catP1                 TCTTTCTAACCTCACTTGCAATAATTATAAGAACAAGATGATGGTCTGCCAAGTGGGTGG 743
b-catP1(repeat)         TCTTTCTAACCTCACTTGCAATAATTATAAGAACAAGATGATGGTCTGCCAAGTGGGTGG 743
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   TATAGAGGCTCTTGTGCGTACTGTCCTTCGGGCTGGTGACAGGGAAGACATCACTGAGCC 4140
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TATAGAGGCTCTTGTGCGTACTGTCCTTCGGGCTGGTGACAGGGAAGACATCACTGAGCC 1413
b-catP1                 TATAGAGGCTCTTGTGCGTACTGTCCTTCGGGCTGGTGACAGGGAAGACATCACTGAGCC 803
b-catP1(repeat)         TATAGAGGCTCTTGTGCGTACTGTCCTTCGGGCTGGTGACAGGGAAGACATCACTGAGCC 803
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

predicted-vector-full   TGCCATCTGTGCTCTTCGTCATCTGACCAGCCGACACCAAGAAGCAGAGATGGCCCAGAA 4200
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TGCCATCTGTGCTCTTCGTCATCTGACCAGCCGACACCAAGAAGCAGAGATGGCCCAGAA 1473
b-catP1                 TGCCATCTGTGCTCTTCGTCATCTGACCAGCCGACACCAAGAAGCAGAGATGGCCCAGAA 863
b-catP1(repeat)         TGCCATCTGTGCTCTTCGTCATCTGACCAGCCGACACCAAGAAGCAGAGATGGCCCAGAA 863
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

                                                                        Extra T

predicted-vector-full   TGCAGTTCGCCTTCACTATGGACTACCAGTTGTGGTTAAGCT-CTTACACCCACCATCCC 4259
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TGCAGTTCGCCTTCACTATGGACTACCAGTTGTGGTTAAGCT-CTTACACCCACCATCCC 1532
b-catP1                 TGCAGTTCGCCTTCACTATGGACTACCAGTTGTGGTTAAGCTTCTTACACCCACCATCCC 923
b-catP1(repeat)         TGCAGTTCGCCTTCACTATGGACTACCAGTTGTGGTTAAGCT-CTTACACCCACCATCCC 922
b-catP2                 ------------------------------------------------------------ 0
pHIVRev                 ------------------------------------------------------------ 0

                                                  No Extra T in repeat run

predicted-vector-full   ACTGGCCTCTGATAAAGGCTACTGTTGGATTGATTCGAAATCTTGCCCTTTGTCCCGCAA 4319
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ACTGGCCTCTGATAAAGGCTACTGTTGGATTGATTCGAAATCTTGCCCTTTGTCCCGCAA 1592
b-catP1                 ACTGGCCTCTGATAAAGGCTACTGTTGGATTGATTCGAAATCTTGCCCTTTGTCCCGCAA 983
b-catP1(repeat)         ACTGGCCTCTGATAAAGGCTACTGTTGGATTGATTCGAAATCTTGCCCTTTGTCCCGCAA 982
b-catP2                 ACTGGCCTCTGATAAAGGCTACTGTTGGATTGATTCGAAATCTTGCCCTTTGTCCCGCAA 60
pHIVRev                 ------------------------------------------------------------ 0
```

```
predicted-vector-full   ATCATGCACCTTTGCGTGAGCAGGGTGCCATTCCACGACTAGTTCAGTTGCTTGTTCGTG 4379
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ATCATGCACCTTTGCGTGAGCAGGGTGCCATTCCACGACTAGTTCAGTTGCTTGTTCGTG 1652
b-catP1                 ATCATGCACCTTTGCGTGAGCAGGGTGCCATTCCACGACTAGT---------------- 1026
b-catP1(repeat)         ATCATGCACCTTTGCGTGAGC-------------------------------------- 1003
b-catP2                 ATCATGCACCTTTGCGTGAGCAGGGTGCCATTCCACGACTAGTTCAGTTGCTTGTTCGTG 120
pHIVRev                 ------------------------------------------------------------ 0


predicted-vector-full   CACATCAGGATACCCAGCGCCGTACGTCCATGGGTGGGACACAGCAGCAATTTGTGGAGG 4439
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        CACATCAGGATACCCAGCGCCGTACGTCCATGGGTGGGACACAGCAGCAATTTGTGGAGG 1712
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 CACATCAGGATACCCAGCGCCGTACGTCCATGGGTGGGACACAGCAGCAATTTGTGGAGG 180
pHIVRev                 ------------------------------------------------------------ 0


predicted-vector-full   GGGTCCGCATGGAAGAAATAGTTGAAGGTTGTACCGGAGCCCTTCACATCCTAGCTCGGG 4499
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        GGGTCCGCATGGAAGAAATAGTTGAAGGTTGTACCGGAGCCCTTCACATCCTAGCTCGGG 1772
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 GGGTCCGCATGGAAGAAATAGTTGAAGGTTGTACCGGAGCCCTTCACATCCTAGCTCGGG 240
pHIVRev                 ------------------------------------------------------------ 0


predicted-vector-full   ATGTTCACAACCGAATTGTTATCAGAGGACTAAATACCATTCCATTGTTTGTGCAGCTGC 4559
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ATGTTCACAACCGAATTGTTATCAGAGGACTAAATACCATTCCATTGTTTGTGCAGCTGC 1832
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ATGTTCACAACCGAATTGTTATCAGAGGACTAAATACCATTCCATTGTTTGTGCAGCTGC 300
pHIVRev                 ------------------------------------------------------------ 0


predicted-vector-full   TTTATTCTCCCATTGAAAACATCCAAAGAGTAGCTGCAGGGGTCCTCTGTGAACTTGCTC 4619
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TTTATTCTCCCATTGAAAACATCCAAAGAGTAGCTGCAGGGGTCCTCTGTGAACTTGCTC 1892
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 TTTATTCTCCCATTGAAAACATCCAAAGAGTAGCTGCAGGGGTCCTCTGTGAACTTGCTC 360
pHIVRev                 ------------------------------------------------------------ 0


predicted-vector-full   AGGACAAGGAAGCTGCAGAAGCTATTGAAGCTGAGGGAGCCACAGCTCCTCTGACAGAGT 4679
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        AGGACAAGGAAGCTGCAGAAGCTATTGAAGCTGAGGGAGCCACAGCTCCTCTGACAGAGT 1952
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 AGGACAAGGAAGCTGCAGAAGCTATTGAAGCTGAGGGAGCCACAGCTCCTCTGACAGAGT 420
pHIVRev                 ----------------AGAAGCTATTGAAGCTGAGGGAGCCACAGCTCCTCTGACAGAGT 44


predicted-vector-full   TACTTCACTCTAGGAATGAAGGTGTGGCGACATATGCAGCTGCTGTTTTGTTCCGAATGT 4739
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TACTTCACTCTAGGAATGAAGGTGTGGCGACATATGCAGCTGCTGTTTTGTTCCGAATGT 2012
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 TACTTCACTCTAGGAATGAAGGTGTGGCGACATATGCAGCTGCTGTTTTGTTCCGAATGT 480
pHIVRev                 TACTTCACTCTAGGAATGAAGGTGTGGCGACATATGCAGCTGCTGTTTTGTTCCGAATGT 104


predicted-vector-full   CTGAGGACAAGCCACAAGATTACAAGAAACGGCTTTCAGTTGAGCTGACCAGCTCTCTCT 4799
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        CTGAGGACAAGCCACAAGATTACAAGAAACGGCTTTCAGTTGAGCTGACCAGCTCTCTCT 2072
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 CTGAGGACAAGCCACAAGATTACAAGAAACGGCTTTCAGTTGAGCTGACCAGCTCTCTCT 540
pHIVRev                 CTGAGGACAAGCCACAAGATTACAAGAAACGGCTTTCAGTTGAGCTGACCAGCTCTCTCT 164
```

```
predicted-vector-full   TCAGAACAGAGCCAATGGCTTGGAATGAGACTGCTGATCTTGGACTTGATATTGGTGCCC 4859
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TCAGAACAGAGCCAATGGCTTGGAATGAGACTGCTGATCTTGGACTTGATATTGGTGCCC 2132
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 TCAGAACAGAGCCAATGGCTTGGAATGAGACTGCTGATCTTGGACTTGATATTGGTGCCC 600
pHIVRev                 TCAGAACAGAGCCAATGGCTTGGAATGAGACTGCTGATCTTGGACTTGATATTGGTGCCC 224

predicted-vector-full   AGGGAGAACCCCTTGGATATCGCCAGGATGATCCTAGCTATCGTTCTTTTCACTCTGGTG 4919
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        AGGGAGAACCCCTTGGATATCGCCAGGATGATCCTAGCTATCGTTCTTTTCACTCTGGTG 2192
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 AGGGAGAACCCCTTGGATATCGCCAGGATGATCCTAGCTATCGTTCTTTTCACTCTGGTG 660
pHIVRev                 AGGGAGAACCCCTTGGATATCGCCAGGATGATCCTAGCTATCGTTCTTTTCACTCTGGTG 284

predicted-vector-full   GATATGGCCAGGATGCCTTGGGTATGGACCCCATGATGGAACATGAGATGGGTGGCCACC 4979
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        GATATGGCCAGGATGCCTTGGGTATGGACCCCATGATGGAACATGAGATGGGTGGCCACC 2252
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 GATATGGCCAGGATGCCTTGGGTATGGACCCCATGATGGAACATGAGATGGGTGGCCACC 720
pHIVRev                 GATATGGCCAGGATGCCTTGGGTATGGACCCCATGATGGAACATGAGATGGGTGGCCACC 344

predicted-vector-full   ACCCTGGTGCTGACTATCCAGTTGATGGGCTGCCAGATCTGGGGCATGCCCAGGACCTCA 5039
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ACCCTGGTGCTGACTATCCAGTTGATGGGCTGCCAGATCTGGGGCATGCCCAGGACCTCA 2312
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ACCCTGGTGCTGACTATCCAGTTGATGGGCTGCCAGATCTG---------------------- 761
pHIVRev                 ACCCTGGTGCTGACTATCCAGTTGATGGGCTGCCAGATCTGGGGCATGCCCAGGACCTCA 404

predicted-vector-full   TGGATGGGCTGCCTCCAGGTGACAGCAATCAGCTGGCCTGGTTTGATACTGACCTGGACT 5099
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        TGGATGGGCTGCCTCCAGGTGACAGCAATCAGCTGGCCTGGTTTGATACTGACCTGGACT 2372
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 TGGATGGGCTGCCTCCAGGTGACAGCAATCAGCTGGCCTGGTTTGATACTGACCTGGACT 464

                                FLAG tag

predicted-vector-full   ACAAGGACGACGATGACAAGTAAGCGGCCGCTCGAGCATGCATCTAGAGTACCCGGGCTA 5159
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ACAAGGACGACGATGACAAGTAAGCGGCCGCTCGAGCATGCAT------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ACAAGGACGACGATGACAAGTAAGCGGCCGCTCGAGCATGCATCTAGAGTACCCGGGCTA 524

                STOP codon

predicted-vector-full   GGATCCGCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAG 5219
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 GGATCCGCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAG 584

predicted-vector-full   GCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGA 5279
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 GCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGA 644
```

```
predicted-vector-full   GGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCG 5339
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 GGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTT---------- 694

predicted-vector-full   CCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTT 5399
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   GAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACA 5459
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   GGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCC 5519
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   AGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTAT 5579
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   TCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATCTGATCTGGGGC 5639
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   CTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAACGTCTAGGCCCCCCGA 5699
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   ACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAATATGGCCACAACCATGGTG 5759
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   AGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGAC 5819
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694
```

```
predicted-vector-full   GTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAG 5879
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   CTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTG 5939
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   ACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAC 5999
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   GACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAG 6059
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   GACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAAC 6119
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   CGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTG 6179
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   GAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATC 6239
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   AAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCAC 6299
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694

predicted-vector-full   TACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTG 6359
pHIVFor                 ------------------------------------------------------------ 720
insert-predicted        ------------------------------------------------------------ 2415
b-catP1                 ------------------------------------------------------------ 1026
b-catP1(repeat)         ------------------------------------------------------------ 1003
b-catP2                 ------------------------------------------------------------ 761
pHIVRev                 ------------------------------------------------------------ 694
```

```
predicted-vector-full      AGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTG 6419
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGC 6479
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CGCATCGATACCGTCGACCTCGATCGAGACCTAGAAAAACATGGAGCAATCACAAGTAGC 6539
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AATACAGCAGCTACCAATGCTGATTGTGCCTGGCTAGAAGCACAAGAGGAGGAGGAGGTG 6599
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGAT 6659
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAACGAAGA 6719
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAAC 6779
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      TACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTA 6839
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CCAGTTGAGCAAGAGAAGGTAGAAGAAGCCAATGAAGGAGAGAACACCCGCTTGTTACAC 6899
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694
```

```
predicted-vector-full      CCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTATTAGAGTGGAGGTTTGAC 6959
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AGCCGCCTAGCATTTCATCACATGGCCCGAGAGCTGCATCCGGACTGTACTGGGTCTCTC 7019
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      TGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAG 7079
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCT 7139
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGCATGTGA 7199
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCAT 7259
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC 7319
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCT 7379
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCG 7439
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694
```

```
predicted-vector-full      CTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTG 7499
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGT 7559
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG 7619
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTAC 7679
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 7739
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTT 7799
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTT 7859
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      TCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGA 7919
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      TTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATC 7979
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694
```

```
predicted-vector-full      TAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCT 8039
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      ATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATA 8099
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      ACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCA 8159
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGA 8219
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGA 8279
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTG 8339
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGA 8399
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTT 8459
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCT 8519
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694
```

```
predicted-vector-full      CTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCA 8579
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      TTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAAT 8639
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      ACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA 8699
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCC 8759
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGG 8819
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTC 8879
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTT 8939
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCA 8999
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CCTGACGTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCT 9059
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694
```

```
predicted-vector-full      GCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTG 9119
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGA 9179
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGC 9239
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATA 9299
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGC 9359
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG 9419
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC 9479
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG 9539
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACG 9599
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694
```

```
predicted-vector-full      TATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGAT 9659
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT 9719
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGC 9779
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGCGCGTTTTGCCTGTACT 9839
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCA 9899
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      CTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTG 9959
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      TGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGC 10019
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      AGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCTCTCTCGACGCA 10079
pHIVFor                    ------------------------------------------------------------ 720
insert-predicted           ------------------------------------------------------------ 2415
b-catP1                    ------------------------------------------------------------ 1026
b-catP1(repeat)            ------------------------------------------------------------ 1003
b-catP2                    ------------------------------------------------------------ 761
pHIVRev                    ------------------------------------------------------------ 694

predicted-vector-full      GGACTCGGCTTGCTGAAG   10097
pHIVFor                    ------------------   720
insert-predicted           ------------------   2415
b-catP1                    ------------------   1026
b-catP1(repeat)            ------------------   1003
b-catP2                    ------------------   761
pHIVRev                    ------------------   694
```