Collaboration of Ezh2 and Runx1 Inactivating Mutations in Malignant Haematopoiesis

Christopher Booth

Haematopoietic Stem Cell Biology

Weatherall Institute of Molecular Medicine

Lincoln College

D.Phil Medical Sciences

Trinity Term 2017

Collaboration of Ezh2 and Runx1 Inactivating Mutations in Malignant Haematopoiesis

Christopher Booth
Lincoln College
D.Phil Medical Sciences
Trinity Term 2017
Words: 44,049

Extensive efforts have shed light on the identity and biology of cancer stem cells, required and sufficient for the propagation of hematological malignancies and solid tumours. Much less is understood about the closely related issue as to the identity and properties of the normal stem and progenitor cells targeted by oncogenic lesions, and how the nature of the targeted cell might impact on the biology and clinical picture of the resulting cancer. To address this, we developed a mouse model allowing targeted inactivation of *Ezh2* and *Runx1* to different haematopoietic compartments. Inactivating mutations of *EZH2* and *RUNX1* frequently co-occur in haematological malignancies with markedly different phenotypes including myelodysplastic syndrome (MDS) and early thymic progenitor (ETP) leukaemia.

Inactivation of *Ezh2* and *Runx1* in adult haematopoietic stem cells (HSCs) resulted in perturbed haematopoiesis leading to development of an MDS-like disease. Unexpectedly, this MDS phenotype could be fully reproduced when *Ezh2* and *Runx1* inactivation was targeted to multipotent progenitors (MPPs) using Flt3-Cre. Furthermore, the disease was transplantable by MPPs, but not more committed progenitor populations, demonstrating that MDS tumour propagating potential is not exclusive to intrinsically self-renewing HSCs.

Targeting *Ezh2* and *Runx1* inactivation to early lympho-myeloid progenitors did not result in an MDS phenotype. These mice showed a marked expansion of ETPs within the thymus, combined with a block in thymocyte differentiation. These expanded ETPs displayed transcriptional features characteristic of ETP leukaemia, a treatment-resistant acute leukaemia subtype hypothesised to originate from ETPs. Combination of inactivation of *Ezh2* and *Runx1* in ETPs with the constitutively activating *Flt3-ITD* signalling mutation resulted in an aggressive lympho-myeloid acute leukaemia, which could be propagated by the expanded ETP population. These findings demonstrate the potential of lympho-myeloid progenitors such as ETPs to become leukaemia stem cells which propagate a disease retaining lympho-myeloid features.

We used this novel ETP leukaemia model to explore therapeutic targeting of *Ezh2*-inactivated ETP leukaemias using inhibitors of the bromodomain and extra terminal (BET) proteins. Aberrant transcription resulting from epigenetic changes induced by *Ezh2* loss could be reversed by BET inhibitors, and these compounds showed therapeutic efficacy against both mouse and human ETP leukaemias *in vitro* and *in vivo*.

Acknowledgements

I thank my supervisors Adam Mead and Sten Eirik Jacobsen for their mentorship and support over the last five years, and for giving me the opportunity to perform this research. My DPhil was funded by the MRC WIMM Prize Studentship.

I thank all members of the Mead and Jacobsen labs, past and present, for their help, advice and friendship. I am particularly grateful to Debbie Atkinson, who introduced me to many of the techniques when I first started in the lab, Wen Hao Neo, who has contributed to several of the studies presented in this thesis, and Nikolas Barkas, who analysed a lot of the gene expression data. I also thank Hanane Boukarabila for her input in the early stages of the project.

I would also like to thank the WIMM FACS Facility team, and all of the BMS JR staff, especially for their patience with the JQ1 *in vivo* experiments.

I also thank all the collaborators who have contributed to this work. Elizabeth Soilleux performed histological analysis and immunohistochemical staining, Noushin Farnoud and Elli Papaemmanuil analysed exome sequencing data, George Giotopoulos and Brian Huntly performed I-BET151 *in vivo* experiments, and David Teachey provided patient-derived xenograft samples.

Finally, I thank everyone at OUCKC and FRCC for the fun and adventures.

Table of Contents

Abbreviations

ALL	Acute lymphoblastic leukaemia
allo-SCT	Allogeneic stem cell transplantation
AML	Acute myeloid leukaemia
ATRA	All-trans retinoic acid
aUPD	Acquired uniparental disomy
B-ALL	B-cell acute lymphoblastic leukaemia
BET	Bromodomain and extra terminal
ChIP	Chromatin immunoprecipitation
DKO	Double knockout (of *Ezh2* and *Runx1*)
DKOITD	Double knockout of *Ezh2* and *Runx1* combined with *Flt3-ITD*
DN	Double negative (for CD4 and CD8)
DP	Double positive (for CD48 and CD150 in Chapter 3; for CD4 and CD8 in Chapters 5 and 6).
ETP	Early thymic progenitor
ETP-ALL	Early thymic progenitor acute lymphoblastic leukaemia
FACS	Fluorescence activated cell sorting
GMP	Granulocyte-monocyte progenitor
GSEA	Gene set enrichment analysis
HSC	Haematopoietic stem cell
ITD	*Flt3-ITD*; Internal tandem duplication
KO	Knockout
Lin	Lineage
LK	Lineage negative, Kit positive (murine myeloid progenitor compartment)
LMPP	Lymphoid-primed multipotent progenitor
LSC	Leukaemia stem cell

LSK	Lineage negative, Sca1 positive, Kit positive (murine haematopoietic stem and progenitor compartment)
MDS	Myelodysplastic syndrome
MEP	Myeloid-erythroid progenitor
MPAL	Mixed phenotype acute leukaemia
MPP	Multipotent progenitor
MRD	Minimal residual disease
NSG	Non-obese diabetic, severe combined immune deficiency, common gamma chain-null (immune-deficient inbred mouse strain)
PRC2	Polycomb repressive complex 2
qPCR	Quantitative polymerase chain reaction
SCF	Stem cell factor
T-ALL	T-cell acute lymphoblastic leukaemia
TCR	T-cell receptor
TSS	Transcription start site
WBC	White blood cell
WT	Wild-type

Chapter 1: Introduction

1: Haematopoiesis

1.1: Haematopoietic stem and progenitor cells

Haematopoiesis is the process by which blood cells are produced. Mature blood cells can be divided into the key classes of erythrocytes, platelets (produced by megakaryocytes), myeloid cells and lymphocytes (Orkin, 2000; Orkin and Zon, 2008). Erythrocytes transport oxygen, platelets are involved in blood vessel repair and clotting, while myeloid cells and lymphocytes (collectively known as white blood cells or leukocytes) mediate immunity. While erythrocytes and platelets are relatively homogenous, the functional classes of myeloid cells and lymphocytes are very much more diverse. Generally, innate immunity is mediated by myeloid cells, of which the most common types are monocytes and neutrophils. Monocytes can act as phagocytic cells, engulfing and digesting foreign particles and bacteria, and are able to differentiate into macrophages, which are also phagocytes and involved in antigen presentation. Neutrophils are another type of phagocyte and are also known as granulocytes due to the granular appearance of their cytoplasm.

Lymphocytes are divided broadly into B-cells and T-cells, abundant in the bone marrow and thymus respectively. These cells mediate adaptive immunity and can be subdivided into many functional classes owing to the complexity of their task. The key function of B-cells is the production of antibodies, protein molecules that specifically recognise one cognate non-self antigen and label foreign cells for targeting by other members of the immune system. T-cells also recognise one specific antigen via their T-cell receptor (TCR), and act to both destroy foreign cells and stimulate the activity of B-cells.

Haematopoiesis is organised as a hierarchy, with the various mature blood cell lineages at the bottom, and haematopoietic stem cells (HSCs) at the top (Figure 1.01). HSCs are defined by the key

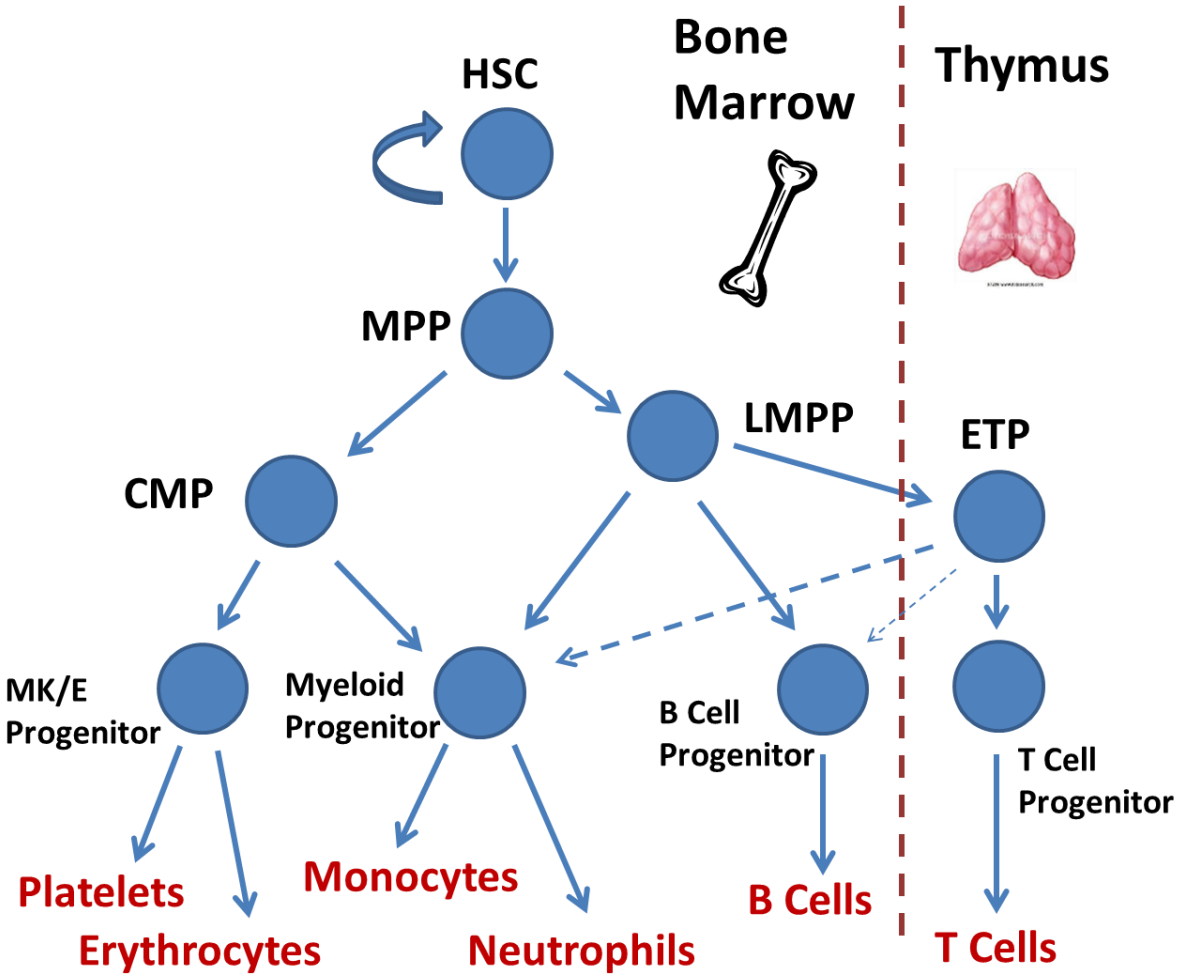


Figure 1.01

Simplified representation of the haematopoietic hierarchy. Arrows indicate which stem and progenitor populations give rise to which others; dashed arrows indicate reduced differentiation potential. HSC haematopoietic stem cell, MPP multipotent progenitor, CMP common myeloid progenitor, MK/E megakaryocyte/erythroid, LMPP lymphoid-primed multipotent progenitor, ETP early thymic progenitor.

properties of self-renewal and multipotency; that is they are able both to create identical copies of themselves and give rise to cells of all the haematopoietic lineages (Orkin and Zon, 2008). In between HSCs and mature blood cells are progenitor and precursor cells. These are cells which have lost the property of self-renewal characteristic of HSCs and have become committed to production of mature blood cells of one or more lineage.

Production of large numbers of mature haematopoietic cells necessitates multiple rounds of cell division, with each division requiring replication of the entire genome. It is during DNA replication that cells are most vulnerable to acquisition of damaging mutations, the accumulation of which inside a single cell could lead to malignant transformation (Gaillard et al., 2015). Progenitor cells have thus been selected by evolution to possess a pre-programmed, finite lifespan, such that any mutations picked up during the progress of a progenitor cell down the haematopoietic hierarchy are eventually lost from the organism as most mature cells are short lived (Corces-Zimmerman and Majeti, 2014). In contrast, mutations acquired by an HSC may be retained for the duration of the individual's lifespan, with any further mutations gained by the progeny of the HSC increasing the risk of malignant transformation. HSCs are therefore largely quiescent in steady state, adult haematopoiesis, only entering the cell cycle rarely (Nakamura-Ishizu et al., 2014). This strict coupling of self-renewal with quiescence within the haematopoietic system is thought to be a fundamental mechanism of cancer prevention (Wang and Dick, 2005).

However, this model does not explain the fact that during foetal development, HSCs multiply very rapidly in order to produce enough of themselves to last the lifespan of the animal, and yet do not seem to acquire a correspondingly high number of mutations (Bowie et al., 2006). Instead, mutations appear to be acquired steadily over the course of adult life (Sperling et al., 2017). The mechanisms underlying this discrepancy are unknown.

The haematopoietic hierarchy is usually displayed as a branching tree, with progenitor cells forming the branch points (Figure 1.01). Just beneath the HSC is the multipotent progenitor (MPP), which

while unable to self-renew retains the ability to produce cells of all lineages. Further down the hierarchy are committed progenitors, able to generate progeny of only one lineage. In between these extremes are progenitors with the ability to produce cells of some lineages but not others. The relationships between these oligopotent progenitors are difficult to decipher and have sometimes been controversial (Richie Ehrlich et al., 2011; Schlenner and Rodewald, 2010). However, this is an important area of research since it is within these stem and progenitor cell populations that acute leukaemias are thought to originate (Wang and Dick, 2005). Loss of multipotency as HSCs move towards commitment to a single lineage is thought to occur gradually through a process of lineage restriction (Luc et al., 2008).

1.2: Thymopoiesis

Thymopoiesis is the process of T-cell production, which occurs in the thymus. It consists of maturation over a defined sequence of developmental stages, from undifferentiated haematopoietic progenitor cells to immune-competent T-cells (Koch and Radtke, 2011). While this is a continuous process requiring constant replenishment of the immature thymocyte populations, the thymus itself contains no HSCs and is instead continually fed by a stream of progenitor cells from the bone marrow (Schwarz and Bhandoola, 2006).

In the mouse, the cells which enter the thymus having transited from the bone marrow through the bloodstream are known as thymus-seeding progenitors. Upon their establishment within the thymus they become known as early thymic progenitors (ETPs). These are the most primitive of the T-cell progenitor populations and belong to the double negative (DN) class of thymocytes, which have not yet started to express CD4 or CD8 on the cell surface. Murine DN cells are divided into four subclasses based on expression of CD25 and CD44, which mature sequentially from the DN1 to DN4 stages (Figure 1.02). These early stages of T-cell differentiation involve lineage restriction, which is

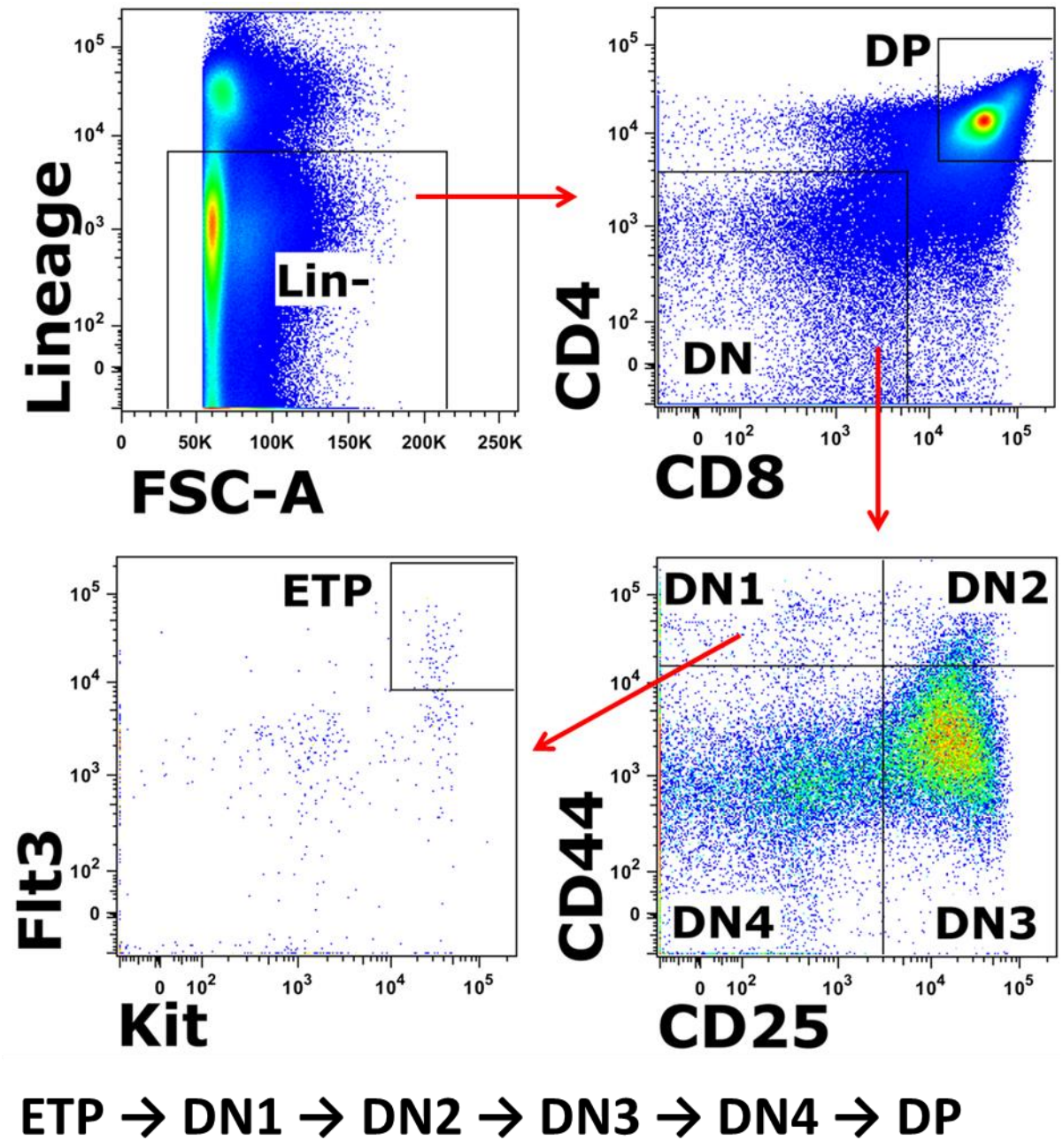


Figure 1.02

FACS plots illustrating staging of murine thymocyte progenitors. FSC-A forward scatter area (correlates with cell size), Lin- lineage negative (lineage cocktail consists of CD3e, Gr1, NK1.1, CD19, B220, CD11c), DN CD4/CD8 double negative, DP CD4/CD8 double positive, ETP early thymic progenitor.

the shutting down of the potential for differentiation along alternative lineage pathways (Rothenberg, 2011). DN1 cells retain the ability to produce cells of the myeloid as well as T-cell lineages, whereas by the time the DN3 stage is reached thymocytes are committed to the T-cell lineage. Notch signalling induced by the thymic micro-environment is important in this process, both for promoting T-cell lineage transcription and inhibiting transcription of other lineages (Yui and Rothenberg, 2014). The transcription factor *PU.1* is associated with retention of myeloid lineage differentiation potential, and the downregulation of its expression coincides with commitment to the T-cell lineage (Laiosa et al., 2006). The haematopoietic transcriptional regulator *RUNX1* is at least partially responsible for repression of *PU.1* at this stage (Huang et al., 2008; Zarnegar et al., 2010). The DN1 and DN2 stages, prior to T-cell lineage commitment, are also associated with rapid proliferation (Lu et al., 2005).

The T-cell receptor (TCR), responsible for the recognition of specific antigens, is essential for the immunogenic function of mature T-cells. In most T-cells (known as αβ T-cells), it consists of a heterodimeric complex of an alpha and a beta chain, and associates with the CD3 co-receptor (Clambey et al., 2014). The alpha and beta chains are both variable, expressed only after rearrangement of their encoding DNA sequence to produce an extreme structural diversity of antigen-binding sites. Once the DN3 stage is reached and thymocytes have become committed to the T-cell lineage, the pre-T cell receptor (pre-TCR) begins to be expressed. This consists of the TCR β-chain, which has already undergone rearrangement, paired with an unarranged α-chain. Each cell expresses a near-unique pre-TCR, and must next undergo β-selection in order to remove cells with a non-functional receptor. Cells passing the β-selection checkpoint advance to the DN4 stage, then upregulate both CD4 and CD8 to become DP cells. At this point, the alpha TCR chain is rearranged and a full αβ-TCR is expressed. DP thymocytes then undergo positive selection, where cells expressing TCRs which react either too weakly or too strongly with the major histocompatibility complex (MHC) are removed. Either CD4 or CD8 is then downregulated and cells are committed to becoming $CD4^+$ single-positive or $CD8^+$ single-positive cells. At the single-positive stage, thymocytes

undergo negative selection, where TCRs reacting too strongly with self-antigens are eliminated. Following negative selection the maturation process is complete, and the T-cells exit the thymus (Koch and Radtke, 2011).

2: The Lympho-Myeloid Pathway and Early Thymic Progenitors

2.1: The lympho-myeloid pathway

The lymphoid-primed multipotent progenitor (LMPP) was identified in 2005 as a bone marrow-resident oligopotent progenitor population with the ability to produce cells of the myeloid and lymphoid, but not megakaryocytic or erythroid, lineages (Adolfsson et al., 2005). This was in contrast to the previously accepted model of haematopoiesis in which the earliest lineage decision made by MPPs was a strict bifurcation between the myeloid, megakaryocytic and erythroid lineages on the one side, and the lymphoid lineage on the other (Orkin and Zon, 2008). The existence of the LMPP implies that this model is incomplete, since cells of the myeloid lineage may be produced through either of two differentiation pathways: the common myeloid progenitor pathway and the "lympho-myeloid" pathway. The clinical implications of this revised model of haematopoiesis were made clear in 2011, with the discovery that many acute myeloid leukaemias (AMLs) are propagated by a leukaemic stem cell (LSC) population bearing transcriptional hallmarks of the LMPP (Goardon et al., 2011). Since achieving cure of these leukaemias necessitates eradication of this LSC population (discussed in section 6.2 below), improved understanding of the lympho-myeloid pathway has important clinical implications.

2.2: Early thymic progenitors

Further implications of the lympho-myeloid pathway became evident in 2008, with the finding that the most primitive thymocyte progenitor populations found within the thymus retain the potential to produce myeloid as well as T-lymphoid cells (Bell and Bhandoola, 2008; Wada et al., 2008). These cells, known as early thymic progenitors (ETPs), like LMPPs are unable to produce cells of the megakaryocytic or erythroid lineages. This implies that the lympho-myeloid pathway is a major, and possibly the only, contributor to T-cell lymphopoiesis. In 2011 it was further shown that ETPs are closely related at the transcriptional and functional level to LMPPs, suggesting that LMPPs may act as thymus seeding progenitors (Luc et al., 2012). This was particularly pertinent in light of the identification in 2009 of early thymic progenitor acute lymphoblastic leukaemia (ETP-ALL), a molecularly and clinically distinct subgroup of T-ALL showing a transcriptional signature resembling that of the ETP (Coustan-Smith et al., 2009). It was suggested that ETP-ALLs initiate within ETP, though as yet experimental evidence for this hypothesis has been lacking. ETP-ALL is described in more detail in section 7 below.

3: Transcriptional Control of Haematopoiesis

3.1: The PRC2 complex in normal haematopoiesis

Within the nucleus, DNA is kept compact and organised through packaging into nucleosomes. These consist of a structural core of histone proteins around which DNA is wound, and unstructured N-terminal histone tails. These tails contain key residues which can be modified post-translationally, influencing the transcriptional status of the nearby DNA (Bannister and Kouzarides, 2011). Many of these modifiable residues and their functions are highly conserved through evolution – one of these being histone 3 lysine 27 (H3K27), which can exist in mono-, di- or trimethylated (H3K27me3), or acetylated (H3K27ac) forms (Pasini et al., 2010; Tie et al., 2009). H3K27me3 is associated with

silencing of transcription of nearby genes, while H3K27ac is conversely associated with active transcription (Wang et al., 2008).

The repressive H3K27me3 epigenetic mark is deposited by the polycomb repressive complex 2 (PRC2). The mammalian core PRC2 is composed of three subunits: EED, SUZ12, and EZH1 or EZH2 (Di Croce and Helin, 2013). EZH1 and EZH2 are close homologs and either one of them can comprise the catalytic subunit of PRC2, performing the deposition of methyl groups onto H3K27 via their SET domain. EED and SUZ12 are required for structural stability of the complex as well as activation of the catalytic activity of EZH1/2. In addition, EED is able to bind previously deposited H3K27me3, helping to establish a feed-forward mechanism of PRC2 activity for stable transcriptional repression. The complex can also bind to a number of other non-core subunits, such as JARID2 and AEBP2, which may function in recruitment to specific target genes or regulation of its activity (Holoch and Margueron, 2017; Margueron and Reinberg, 2011).

The ability of H3K27 to form either a repressive or an activating mark of gene expression means that it can act as a transcriptional switch (Pasini et al., 2010). Repressed genes associated with H3K27me3 can become activated via demethylation followed by acetylation of H3K27. Removal of the methyl groups is carried out by UTX and JMJD3 (Agger et al., 2007), while acetylation is performed by CBP and P300 (Pasini et al., 2010). Conversely, deacetylation of H3K27ac in preparation for methylation by PRC2 is carried out by the nucleosomal remodelling and deacetylation (NuRD) complex (Reynolds et al., 2012). Changes in the activity of any of the components of this switch mechanism, for example through mutations, can lead to deregulated transcription.

The H3K27me3 mark is found distributed throughout the genome in all cell types. H3K27me3 domains can be very large (>100 kb), or localised around gene bodies and the transcription start site (TSS) (Margueron and Reinberg, 2011). Despite its ubiquitous role in regulation of gene expression, the mechanisms of recruitment of PRC2 to specific target regions in specific cell types, and of

repression of transcription by the H3K27me3 mark, are still incompletely understood (Holoch and Margueron, 2017).

PRC2 plays an important role in the regulation of gene activity during cellular differentiation (Margueron and Reinberg, 2011). During differentiation, large-scale transcriptional programmes need to be rapidly switched on and off as cells transition from one organismal function to another. In embryonic stem cells, H3K27me3 co-localises with the activating mark H3K4me3 at so-called bivalent domains. These are thought to represent transcriptionally poised states, ready to be rapidly activated or repressed depending on the lineage fate (Aloia et al., 2013).

The function of the PRC2 complex has been shown to be important in many aspects of haematopoiesis. Several studies have demonstrated its importance in HSC function (Hidalgo et al., 2012; Kamminga et al., 2006; Lee et al., 2015; Mochizuki-Kashio et al., 2011; Xie et al., 2014). Interestingly, the relative importance of the two interchangeable catalytic subunits, EZH1 and EZH2, in HSC function has been suggested to change during development from the embryo to the adult. *Ezh1* appears to be crucial in adult but not foetal HSCs, while *Ezh2* is dispensable for adult HSC maintenance but may be more important in foetal HSCs (Hidalgo et al., 2012; Mochizuki-Kashio et al., 2011). The precise role of PRC2 in the emergence and maintenance of HSCs remains unclear, but it appears to be required for suppression of differentiation and maintenance of a slow-cycling state (Xie et al., 2014).

PRC2 also plays an essential role in lymphocyte differentiation. Mice in which *Ezh2* is conditionally knocked out in HSCs fail to produce mature B and T cells (Su et al., 2003; Su et al., 2005). In both cases, this is due to a failure of differentiating cells to progress past defined maturation stages, highlighting the importance of the complex in transcriptional control of differentiation. Knockout of *Suz12* targeted to early lymphoid progenitors using *Rag1Cre* produces a similar phenotype (Lee et al., 2015).

3.2: RUNX1 in normal haematopoiesis

Runt-related transcription factor 1 (*RUNX1*; also known as *AML1* or *CBFα2*) is a master transcriptional regulator of haematopoiesis, which can act as either a transcriptional activator or repressor of a broad range of target genes (de Bruijn and Dzierzak, 2017). It is required for the initial emergence of HSCs from the aorta-gonad-mesonephros region during foetal development, though not for their maintenance thereafter (Chen et al., 2009).

Despite being dispensable for adult HSC function, *RUNX1* is essential for their differentiation through several haematopoietic lineages. Mice in which inactivation of *Runx1* is targeted to adult HSCs display reduced numbers of mature B and T lymphocytes as well as platelets (Ichikawa et al., 2004). The loss of mature T-cells was shown to be due to a maturation block between the DN2 and DN3 stages, which also led to a loss of overall thymus cellularity (Growney et al., 2005; Ichikawa et al., 2004). A similar block was also observed in early B-cell maturation (Growney et al., 2005). Histological analysis of the bone marrow revealed a decrease in numbers of mature megakaryocytes, explaining the lack of platelets in the peripheral blood (Growney et al., 2005). Paradoxically, *in vitro* colony assays showed that megakaryocytes from *Runx1* knockout mice in fact showed increased proliferative potential compared to wild-type.

These mice also showed development of a mild myeloproliferative phenotype (Growney et al., 2005). Bone marrow and spleen tissue was hypercellular compared to wild-type controls, which could be accounted for by an expansion of Mac1-expressing myeloid cells. In addition, haematopoietic stem and progenitor cells were expanded in number approximately 3-fold, but their repopulating potential in competitive transplantation assays was reduced compared to wild-type cells. Despite conferring a myeloproliferative phenotype with perturbation of normal HSC function,

adult HSC-targeted Runx1 inactivation did not lead to development of acute malignancy in the absence of collaborating mutations (Growney et al., 2005).

RUNX1 mutations are some of the most common, and well-characterised, in haematological malignancies. These are discussed in section 8.2 below.

4: Intracellular Signalling Pathways

4.1: The RAS signalling pathway

The RAS pathway transmits extracellular proliferative signals from the cell membrane to the nucleus, where transcriptional responses are initiated. It comprises numerous components which vary between cell types and may respond to signals in different ways depending on the context (Steelman et al., 2011). In general, binding of the cognate extracellular ligand to a cell surface receptor causes activation of the receptor, for example through dimerization and intermolecular autophosphorylation (Blume-Jensen and Hunter, 2001). This causes recruitment of guanine nucleotide exchange factors (GEFs), which activate membrane-bound RAS protein by loading it with GTP. RAS:GTP then recruits RAF to the membrane where it is activated by a membrane-bound kinase. Activated RAF phosphorylates MEK, which in turn phosphorylates ERK, which is then able to translocate to the nucleus and modulate the activity of transcriptional regulators (Steelman et al., 2011).

In cancer cells, mutations in many components of this pathway cause constitutive signalling, leading to accelerated cell proliferation. However, mutations can also occur in regulators of the pathway. For example, NF1 is a negative regulator of RAS signalling, which works by activating the GTPase function of RAS (Steelman et al., 2011). This causes hydrolysis of the RAS-bound GTP to GDP,

inactivating RAS until it can be replaced by a GEF protein. Inactivating mutations in *NF1* therefore result in increased RAS activity (Steelman et al., 2011).

In many contexts, signalling pathways crosstalk with each other, increasing the complexity of signal transmission and modulation. The JAK-STAT pathway is closely associated with RAS signalling, since it can be activated by some of the same cell surface receptors, such as KIT and FLT3 (Cook et al., 2014). Here, binding of the extracellular ligand causes activation of the intracellular JAK protein, which enacts dimerization of STAT via phosphorylation. STAT dimers are then able to enter the nucleus where they activate transcription (Vainchenker and Constantinescu, 2013).

4.2: FLT3

The cytokine receptor FLT3 is a common target of activating signalling pathway mutations in ETP-ALL (Zhang et al., 2012b). Interestingly, mutations in *FLT3* are very rare in non-ETP T-ALLs (Zhang et al., 2012b) and are much more common in AML (around 41% of patients) (Metzeler et al., 2016). *FLT3* activating mutations fall into two categories: tyrosine kinase domain mutations (11% of AML patients) and internal tandem duplications (*FLT3-ITD*; 30% of AML patients). *FLT3-ITD* mutations are constitutively activating and have been long recognised as conferring a poor prognosis in AML (Abu-Duhier et al., 2000; Kiyoi et al., 1999; Kottaridis et al., 2001). They consist of insertions of between 3 and around 400 nucleotides in the juxtamembrane or tyrosine kinase 1 domain of *FLT3*, and are thought to disrupt the auto-inhibitory function of the juxtamembrane domain (Breitenbuecher et al., 2009). This leads to constitutive proliferative signalling via the RAS and STAT5 pathways, conferring growth factor independence to malignant cells (Hayakawa et al., 2000; Mizuki et al., 2000). *FLT3* mutations were present in 9 out of 64 ETP-ALL patients (14%) in the study conducted by Zhang et al. (2012). Of these, seven were ITDs and two were tyrosine kinase domain mutations. *FLT3* mutations are therefore a common mechanism of signalling pathway activation in ETP-ALL.

The molecular mechanisms underlying the effects of FLT3 activating mutations on haematopoiesis have been explored using mouse models. In 2002, Kelly et al. transduced the murine pro-B cell line Ba/F3 with human *FLT3-ITD* constructs derived from an AML patient (Kelly et al., 2002). The authors then transplanted these cells into wild-type mice and observed development of a myeloproliferative disease, characterised by overproduction of mature neutrophils. This group later went on to clone one of these human *FLT3-ITD* constructs into the endogenous murine *Flt3* locus, resulting in expression of the mutant protein only in cells that would normally express *Flt3* (Lee et al., 2007). These mice also developed a myeloproliferative phenotype which was more severe in homozygous $Flt3^{ITD/ITD}$ mice than heterozygous $Flt3^{ITD/+}$. In addition, the authors examined the haematopoietic stem and progenitor compartment of these mice, noting a 2-fold increase in the numbers of these cells which could be accounted for by an increase in the proportion of cells actively cycling, and a suppression of apoptosis. Our group later demonstrated that this *Flt3-ITD* allele confers a myeloid differentiation bias to these expanded stem and progenitor cells (Mead et al., 2013).

Together, these findings from mouse models demonstrate that *Flt3-ITD* increases the myeloid output of haematopoietic stem and progenitor cells, leading to a myeloproliferative phenotype which is non-lethal when expressed from the endogenous *Flt3* locus. On its own, *Flt3-ITD* is insufficient to induce acute leukaemia.

5: Myelodysplastic Syndrome

5.1: Clinical features and genetic causes of myelodysplastic syndrome

Myelodysplastic syndrome (MDS) comprises a heterogenous group of disorders of the haematopoietic system, characterised by cytopaenia of the peripheral blood abnormal morphology (dysplasia) of cells of the myeloid lineages within the bone marrow. These features are caused by proliferation of immature haematopoietic progenitor cells which are defective in maturation. This

can lead to fatigue, bleeding and increased susceptibility to infections as a result of anaemia, thrombocytopaenia and neutropaenia respectively (Sperling et al., 2017). MDS is also associated with high risk of transformation to acute myeloid leukaemia (AML).

The incidence of MDS is around 5 patients per 100,000 per year (Cogle, 2015; Ma et al., 2007). The incidence increases with age, and long-term survival rates are low (<50% 3-year survival, which decreases with age). Patients can be treated with DNA hypomethylating agents such as azacytidine, some groups can show sensitivity to lenalidomide, and patients with highly proliferative myeloid blast cells can be treated with cytotoxic drugs such as cytaribine (Sekeres and Cutler, 2014). Less elderly patients may also be suitable for allogeneic stem cell transplantation (allo-SCT), the only potentially curative treatment. However, none of these treatments show strong efficacy, and even after allo-SCT, disease-propagating MDS stem cells can persist, causing relapse (Tehranchi et al., 2010). Greater understanding of the mechanisms driving the disease is therefore required, so that new approaches for therapy can be developed.

Though some classes of MDS can result from inherited mutations, most are caused by acquisition of somatic mutations within the haematopoietic system. The mutations which are recurrently found in MDS generally fall into a small number of discrete categories: splicing factors (e.g. *SF3B1* and *SRSF2*), epigenetic regulators (e.g. *EZH2*, *ASXL1*, *DNMT3A* and *TET2*), members of the cohesin complex (e.g. *STAG2*) and transcriptional regulators (e.g. *RUNX1* and *GATA2*). The tumour suppressor gene *TP53* is also mutated in some MDS patients, and mutations in signalling pathway components such as *NRAS* are also found and associated with transformation to AML (Sperling et al., 2017).

Cytogenetic abnormalities are also found in MDS. The most common of these is deletion of chromosome 5q, present in around 25% of MDS patients (Komrokji et al., 2013), and which exists as the only genetic abnormality in around 3% (known as isolated 5q- syndrome) (Papaemmanuil et al., 2013). The commonly deleted region includes the genes encoding the ribosomal protein RPS14, the protein kinase CK1α, and the micro-RNAs *miR-145* and *miR-146a*. 5q- patients show sensitivity to the

drug lenalidomide, which acts via degradation of CK1α leading to activation of P53 and apoptosis. It is the haploinsufficency of CK1α resulting from loss of chromosome 5q that allows the threshold level of degradation to be reached for activation of apoptosis (Sperling et al., 2017). 5q- patients are associated with a relatively good prognosis and low risk of transformation to AML (Sperling et al., 2017).

Deletion of chromosome 7q is also common in MDS, and these cases are associated with a poor prognosis. *EZH2* is one of the genes lost in 7q- MDS, and point mutations in this gene are also independently associated with poor prognosis (Bejar et al., 2011).

These recurrent mutations have also been shown to occur in characteristic sequences over the progression of the disease. Some, such as those in the splicing regulators *SF3B1* and *U2AF1*, are known to occur early, while mutations in transcriptional regulators such as *RUNX1* tend to be acquired later (Papaemmanuil et al., 2013). Additionally, the identity of the mutations acquired early in disease progression can impact upon which mutations are acquired later, since some combinations of mutations are commonly found together while others are mutually exclusive. This could either result from functional redundancy (if two mutations produce the same phenotypic effect, so that acquisition of the second confers no advantage), or synthetic lethality (if loss of both genes results in cell death). As an example, *SF3B1* and *SRSF2* splicing factor mutations occur early and are mutually exclusive. *SRSF2*, but not *SF3B1*, mutations are in turn mutually exclusive with *EZH2* mutations (Papaemmanuil et al., 2013). The mechanism underlying the mutual exclusivity of *SRSF2* and *EZH2* mutations was shown to involve *SRSF2* inactivation-dependent inclusion of a poison exon into the *EZH2* mRNA transcript, resulting in nonsense-mediated decay (Kim et al., 2015). As well as providing insight into the relationship between the mutations in the context of MDS, these findings illustrate the importance of *EZH2* as a tumour suppressor in the haematopoietic system.

5.2: Clonal progression of MDS

Development of MDS occurs through a number of defined stages. The initiating event consists of one or more mutations occurring in an HSC, which confers a clonal advantage over normal HSCs (Cazzola et al., 2013). Over time, this leads to almost all haematopoiesis deriving from this premalignant HSC clone, although there may be no clinical symptoms at this stage. This is known as clonal haematopoiesis and can in fact be found in many healthy adults, who are at increased risk of later development of haematological malignancy (Steensma et al., 2015). Acquisition of further mutations by HSCs within the premalignant clone leads to development of subclones, which may in turn expand to dominate the bone marrow. Eventually these mutations bring about the defects in production of mature haematopoietic cells characteristic of MDS. Because non-mutated HSCs are by this stage largely absent from the bone marrow, residual normal haematopoiesis is insufficient to prevent onset of clinical symptoms. Although this clonal dominance phenomenon is central to the pathogenesis of MDS, the process by which it occurs is poorly understood. New models of the disease are therefore required, which can provide insight into the mechanisms underlying the gain of a competitive advantage by a malignant clone.

It has been shown that the progressive acquisition of mutations in MDS occurs exclusively within the malignant counterparts of normal HSCs, rather than in progenitor cells (Woll et al., 2014). This makes intuitive sense given that HSCs are the only self-renewing haematopoietic cells, and therefore any mutations acquired within other cells will eventually be lost, unless they result in aberrant gain of self-renewal. The observation that certain mutations are commonly mutated early in the progression of the disease, while others occur later, raises the possibility that there is a defined order in which sets of mutations have to occur, in order not to become extinguished. Early mutations such as in splicing factors may confer a clonal advantage to HSCs without compromising their haematopoietic output, while late mutations such as in transcription factors may reduce

mature cell production without terminating the clone, but only with the co-operation of the earlier mutations.

HSCs may therefore act simply as a “vehicle” for acquisition of mutations by haematopoietic cells, regardless of whether the biological effect of the mutations occurs within HSCs themselves. Alternatively, HSCs may possess intrinsic properties that make them the only cells capable of driving MDS, and the same mutations acquired by progenitor cells would not trigger the same phenotype even if they could persist. To explore these possibilities, models are required which allow targeting of MDS-causing mutations to different haematopoietic populations in order to compare the resulting phenotypes.

5.3: Mouse models of MDS

Several attempts have been made to generate mouse models of MDS (Beachy and Aplan, 2010; Zhou et al., 2015). Each of these displays some features characteristic of the human disease, but due to the heterogeneity of human MDS, there can be no “definitive” MDS mouse model.

One of the earliest to be described was the *NUP98-HOXD13* mouse model (Lin et al., 2005; Pineault et al., 2003). This is a rare fusion gene found in a patient with therapy-related AML (Raza-Egilmez et al., 1998), meaning that its relevance for MDS is questionable. Nonetheless, mice expressing the *NUP98-HOXD13* construct in haematopoietic tissues develop phenotypic features characteristic of human MDS, such as peripheral blood leukopaenia and dysplasia, combined with increased apoptosis of bone marrow cells. Additionally, many of these mice go on to develop acute leukaemias of various phenotypes: megakaryocytic, erythroid, lymphoid or myeloid.

In 2010, Barlow et al. developed a model of 5q- MDS, by conditional deletion of a region of mouse chromosome 18 corresponding to part of the commonly deleted region in human 5q- syndrome (Barlow et al., 2010). These mice develop key features of the human disease, including anaemia,

dysplasia of the erythroid and megakaryocyte lineages, and reduced numbers of erythroid and myeloid progenitor cells caused by increased apoptosis.

TET2 is one of the most commonly mutated genes in MDS (22% of patients) (Papaemmanuil et al., 2013), and two groups have generated MDS mouse models by combining *Tet2* deletion with other mutations. Haematopoietic-specific deletion of *Tet2* combined with either *Asxl1* (Abdel-Wahab et al., 2013) or *Ezh2* (Muto et al., 2013) was also shown to result in development of MDS phenotypes. These mice show increased numbers of stem and progenitor cells and go on to develop anaemia and leukopaenia.

In 2014, Sashida et al. showed that transplantation of LSK haematopoietic stem and progenitor cells transduced with a dominant negative *Runx1*-mutant construct, followed by conditional inactivation of *Ezh2*, resulted in development of an MDS-like phenotype (Sashida et al., 2014). This was characterised by leukopaenia and anaemia, and appeared to result in an extrinsic suppressive effect on the function of normal HSCs. These findings are interesting in light of the significant association of *EZH2* and *RUNX1* mutations in MDS patients. However, the experimental setup precluded targeting of the mutations to specific haematopoietic populations, so the relative importance of HSCs and progenitor cells in the development of *Ezh2* and *Runx1*-induced MDS remains unexplored.

6: Leukaemia

6.1: Acute leukaemia

Acute leukaemias are cancers of the white blood cells (leukocytes). They are characterised by uncontrolled proliferation of immature, non-functional haematopoietic progenitor/precursor cells. These malignant cells have gained both aberrant self-renewal and differentiation stage arrest, making them unable to perform their normal functions and wasting bone marrow resources on their

continued production. Acute leukaemias are classified based on their morphology, immunophenotype and genetics (Arber et al., 2016).

Acute myeloid leukaemia (AML) shows phenotypic characteristics of the myeloid lineage and is most common in older patients (Dores et al., 2012). As well as occurring in patients without prior history of haematological malignancy (*de novo* AML), it commonly evolves from MDS (secondary AML). This occurs when the immature progenitor cells that are blocked in differentiation acquire greatly increased proliferative capacity, saturating the bone marrow and spilling out into the peripheral blood. Meanwhile acute lymphocytic leukaemia (ALL), originating from either the B (B-ALL) or T (T-ALL) lymphocyte lineages, arises more frequently in children than in adults (Dores et al., 2012). Though treatment strategies have improved radically in recent decades and 5-year survival rates are now over 80% for patients aged under 20 years in the United States for most ALL subtypes (Dores et al., 2012), the short and long-term consequences of intensive ALL treatment can be severe.

Development of new treatment strategies which do not rely on intensive chemotherapy is therefore required. For example, molecularly targeted therapies, which bind to and modulate the activity of specific proteins, can potentially specifically inhibit the growth of malignant cells while leaving normal cells unaffected (Huang et al., 2014). One of the most successful examples is Imatinib, an inhibitor of the fusion protein kinase BCR-ABL specifically expressed by leukaemia cells in chronic myeloid leukaemia and some ALLs (Longo, 2017). This compound therefore inhibits the proliferation of malignant leukaemia cells, while leaving normal haematopoiesis relatively intact (Zitvogel et al., 2016).

Another highly successful example is the use of all-trans retinoic acid (ATRA) to treat acute promyelocytic leukaemia. Without treatment this one of the most aggressive subtypes of leukaemia, but since the introduction of ATRA in the early 1990s cure rates have become very high (Coombs et al., 2015). ATRA works by binding to and targeting for degradation the PML-RARα fusion protein uniquely expressed by promyelocytic leukaemia cells. Rather than induce cell death directly, this

instead causes their differentiation into mature granulocytes, effectively eliminating the disease-causing malignant progenitor population (de Thé and Chen, 2010).

Targeted therapies are a promising approach for treating haematological malignancies in a way that minimises off-target effects and risk for the patient. Development of these therapies requires thorough understanding of the mechanisms driving the disease, including the specific cell types responsible for their propagation. These disease-propagating cells are known as leukaemia stem cells.

6.2: Leukaemia stem cells

The concept of the cancer stem cell is now well-established in a number of contexts, particularly leukaemia (Clevers, 2011; Magee et al., 2012). This concept states that the malignant cells constituting a leukaemia are heterogenous, with only a proportion of them, the leukaemia stem cells (LSCs) possessing the ability to maintain and propagate the leukaemia. These cells are the cause of relapse following either cytotoxic treatment or allogeneic transplantation. These LSCs must be eliminated in order to achieve cure of the disease in a patient, and cases of relapse after a period of treatment-induced remission are directly caused by survival of as few as one single LSC. Importantly, elimination of the LSCs may also be sufficient to cure the disease, raising the possibility that specific targeting of these cells could dispense with the need for cytotoxic treatment.

The LSC hypothesis was first experimentally confirmed in 1994 with the demonstration that though the surface phenotype of bulk acute myeloid leukaemia (AML) is heterogenous, only $CD34^{+}CD38^{-}$ cells carry the potential to engraft in immune-deficient mice (Lapidot et al., 1994). Since normal human HSCs are also $CD34^{+}CD38^{-}$, this suggested that AML initiates within HSCs, and further that the leukaemic HSCs give rise to a hierarchy of more differentiated (though not fully mature) progeny in a similar way to normal HSCs (Bonnet and Dick, 1997). More recent experiments have shown that

while self-renewing HSCs act as a reservoir within which mutations can accumulate, LSC potential can also be detected within malignant cells more closely resembling progenitor populations (Goardon et al., 2011; Taussig et al., 2010). These results suggest that pre-malignant mutations may accumulate within an HSC clone, with the final steps towards leukaemogenesis occurring within progenitor cells produced by this clone (Shlush et al., 2014).

While initiation of AML may require accumulation of mutations within HSCs, this appears not to be the case for T-cell acute lymphoblastic leukaemia (T-ALL) (Aifantis et al., 2008). T-ALLs from different patients can be sub-categorised by their gene expression signature, with each subgroup showing similarity to a distinct stage of normal thymopoiesis (Ferrando et al., 2002). This suggests that T-ALLs initiate within thymocyte progenitors at various stages of T-cell differentiation, and not within HSCs. This hypothesis is in agreement with studies from mouse models of T-ALL. Overexpression of the *Lmo2* oncogene specifically targeted to thymus cells expressing *CD2* resulted in development of T-ALL at around 10 months of age (McCormack et al., 2010). This provides experimental evidence that thymocyte cells, as opposed to HSCs, are able to give rise to leukaemia.

As well as acute lymphoid and myeloid leukaemias, there are rare cases of leukaemia which cannot be reliably classified into either of these categories. These are known as biphenotypic leukaemias, or mixed phenotype acute leukaemias (MPALs), since they show expression of markers of multiple lineages (Matutes et al., 2011; Weinberg and Arber, 2010). These cases are rare, accounting for only 2-5% of all acute leukaemias (Weinberg and Arber, 2010). However, the patterns of lineage marker co-expression are of particular interest. A 2011 study of 100 biphenotypic leukaemia patients (Matutes et al., 2011) showed that co-expression of B-lymphoid with myeloid markers, or T-lymphoid with myeloid markers, were by far the most common combinations (59 and 35 cases respectively). In contrast, co-expression of B-lymphoid with T-lymphoid markers was very rare (4 cases), and myeloid with both B and T-lymphoid rarer still (2 cases). Co-expression of markers of either of the lymphoid lineages with erythroid or megakaryocyte markers was never observed. These

observations are consistent with initiation of these leukaemias within progenitor cells of the lympho-myeloid pathway, which share myeloid and lymphoid differentiation potential but not erythroid or megakaryocyte (Figure 1.01). However, initiation of leukaemias within lympho-myeloid progenitor populations has never been experimentally verified.

One lympho-myeloid progenitor population is the ETP, and in 2009 of a subtype of acute leukaemia was identified which appears to originate from these cells. This was named ETP-ALL and is described in the following section.

7: Early Thymic Progenitor Leukaemia

7.1: The identification of early thymic progenitor leukaemia

ETP-ALL was initially identified by unsupervised hierarchical clustering of transcriptional microarray data from childhood T-ALLs, using a set of genes previously shown to be differentially expressed in ETPs (Coustan-Smith et al., 2009). Around 15% of T-ALLs showed a distinct gene expression profile highly similar to ETPs, with aberrant expression of HSC and/or myeloid surface markers. These markers included CD117 (Kit; expressed by HSCs), CD34 (expressed by human HSCs), HLA-DR (expressed by antigen-presenting cells such as macrophages), CD13, CD33, CD11b and CD65 (all expressed by myeloid cells). A number of surface markers usually associated with T-ALL were found to be absent or only weakly expressed by these leukaemias; these included CD1a, CD8 and CD5. Analysis of the TCR gene loci did however reveal rearrangement in most cases, confirming the T-cell lineage origin of these leukaemias. The gene expression profile displayed by these cases showed strong upregulation of a number of genes expressed by normal mouse ETPs, which included *LYL1*, *ERG*, *CD44*, *GATA2*, *CEBPA* and *SPI1* (PU.1). ETP leukaemias also showed higher genetic instability than other T-ALLs, with significantly more DNA copy number alterations.

It was further shown that ETP leukaemias showed a higher risk of treatment resistance compared to T-ALL as a whole (Coustan-Smith et al., 2009). Detectable minimal residual disease (MRD) after the first phase of induction chemotherapy was present in a much higher proportion of ETP leukaemia patients (13 of 13 ETP patients compared to 55 of 91 non-ETP patients; P=0.0037), and the levels of MRD were also significantly higher (greater than 5% in 10 of 13 ETP patients compared to 12 of 91 non-ETP patients; P<0.0001). These findings remained similar at the end of induction therapy. This was accompanied by significantly reduced survival of ETP leukaemia patients (10-year overall survival rates were 19% for ETP compared to 84% for non-ETP patients; P<0.0001). Further understanding of the molecular mechanisms underlying this treatment resistance and poor prognosis is therefore required.

Subsequent studies called into doubt the relevance of the ETP leukaemia classification for prognostic stratification of T-ALL patients. In a study published in 2014, Patrick et al. found that presence of the ETP phenotype conferred only a non-significantly inferior overall survival rate and higher relapse rate among T-ALL patients (Patrick et al., 2014). The authors speculated that this discrepancy with previous findings may have been due to a higher proportion of ETP leukaemia patients being transferred to a more intensive treatment regimen, after showing slower initial responses to therapy. These findings are therefore consistent with ETP-ALL showing higher treatment resistance compared to non-ETP T-ALL.

Another study published in 2016 conversely found a worse prognosis for ETP-ALL patients, in agreement with the initial description of the disease subtype (Jain et al., 2016). The most recent study, meanwhile, followed adult ETP-ALL patients and is in agreement with the 2014 study in concluding that escalation of treatment intensity in response to monitoring of MRD brings survival rates of ETP-ALL patients in line with those of T-ALL as a whole (Bond et al., 2017). The authors of this study comment that patients in the 2016 study received heterogenous treatment regimens, and a low rate of allogeneic stem cell transplants (allo-SCT). Bond et al. specifically investigated the rates

of allo-SCT among patients in their study, finding that ETP-ALL patients were more likely to receive allo-SCT as a result of poor initial treatment response (48.9% of ETP patients compared to 28.3% of non-ETP patients; P=0.008). Together, these studies appear to agree that ETP-ALL shows increased treatment resistance compared to non-ETP T-ALL, but that higher intensity treatment and allo-SCT can abrogate the inferior survival rates initially reported when the disease was first identified. Nonetheless, development of more effective, targeted therapies would reduce the need for patients to undergo these intensive treatment regimens. A large number of patients still experience treatment failure (5-year overall survival was 59.6% for ETP-ALL patients in the Bond et al. study), so there remains an unmet need for novel treatments.

7.2: The genetic landscape of ETP leukaemia

In 2012, a landmark study comprehensively analysed the gene expression and mutational landscape of ETP-ALL (Zhang et al., 2012b). The authors performed whole genome sequencing on 12 ETP-ALL patients aged 2-18 years, and then analysed the frequency of all identified mutations in a separate cohort of 52 ETP-ALL and 42 non-ETP T-ALL patients. They also analysed gene expression of a subset of both groups by RNA-sequencing. Interestingly, although the overall gene expression profile and clinical features of ETP-ALL were found to be fairly homogenous and distinct from non-ETP T-ALL, no single genetic lesion was common to all cases. The authors found that ETP-ALL mutations tended to fall into a small number of distinct classes: those affecting signalling pathways, those in transcriptional regulators associated with haematopoietic development, and those in epigenetic modifiers.

Mutations in signalling pathway components were found to be much more common in ETP-ALL than non-ETP T-ALL (present in 67% of ETP patients versus 19% of non-ETP patients; P<0.0001) (Zhang et al., 2012b). The spectrum of signalling-related genes mutated was also distinct between the two

groups. Mutations in *NOTCH1*, which are a common feature of T-ALLs in general, were much less frequent in ETP-ALL (present in 16% of ETP patients versus 43% of non-ETP patients; P=0.0031). Conversely, *FLT3* mutations (discussed in more detail in section 4.2) were present in 14% of ETP-ALL patients but no non-ETP patients (P=0.0108). Mutations in NRAS, meanwhile, were found in both groups (17% ETP versus 10% non-ETP; P=0.3943).

Recurrent mutations in haematopoietic transcriptional regulators included those in *RUNX1*, *IKZF1*, *ETV6* and *GATA3* (Zhang et al., 2012b). Again, some of these were preferentially found in ETP-ALL as opposed to non-ETP T-ALL, such as *ETV6* (33% ETP versus 10% non-ETP; P=0.005) and *GATA3* (9% ETP versus 0% non-ETP; P=0.0797). Interestingly, *GATA3* mutations had not previously been detected in haematological malignancies, and they appear to be specific to ETP-ALL. *GATA3* encodes a transcription factor required for maturation of thymocytes (Yui and Rothenberg, 2014). *RUNX1* mutations are discussed in section 8.2 below.

Mutations detected in epigenetic modifiers included those in the PRC2 components *EZH2*, *EED* and *SUZ12*, the histone trimethylase *SETD2* and the histone acetyltransferase *EP300*. As discussed in section 8.1 below, mutations in PRC2 components were all predicted to be inactivating, and while *EZH2* mutations were present in both groups (16% ETP versus 12% non-ETP, P=0.5587), those in *EED* and *SUZ12* were relatively specific to ETP-ALL (*EED* 13% ETP versus 7% non-ETP, P=0.0208; *SUZ12* 17% ETP versus 2% non-ETP, P=0.0030).

Analysis of the global gene expression profile of ETP-ALL compared to non-ETP T-ALL revealed upregulation of HSC and myeloid lineage-associated gene transcription, in agreement with the previously characterised surface marker expression of the disease (Coustan-Smith et al., 2009; Zhang et al., 2012b). The authors also observed upregulation of a leukaemic stem cell gene expression signature associated with poor outcome in acute myeloid leukaemia patients (Eppert et al., 2011). Together, these findings indicate that ETP-ALL shows gene expression characteristic of undifferentiated haematopoietic stem and progenitor cells, and also displays several features more

reminiscent of AML than T-ALL. These included frequent mutations in RAS pathway components such as *NRAS* and *FLT3*, and upregulation of myeloid-associated gene transcription.

7.3: Mouse models of ETP leukaemia

The existence of such "ETP leukaemias," leukaemias showing a gene expression signature indicative of initiation within the ETP or an ETP-like population, combined with the findings that ETPs are closely related to LMPPs and that LMPPs are able to become LSCs, strongly suggest that ETPs are themselves a potential leukaemia-initiating progenitor population. However, this cannot be proven through the analysis of human ETP leukaemia samples and testing this possibility would require disease models. As yet, there have as yet been no published experimental studies confirming initiation or propagation of leukaemia by ETPs. A number of studies have demonstrated transformation of immature thymocytes via transgenic knockout or overexpression of various genes (Danis et al., 2016; Goossens et al., 2015; Treanor et al., 2014; Yu et al., 2012), but all fall short of establishing the LSC potential of ETPs. These are described sequentially below.

Yu et al. (2012) showed that T-cell lymphomas induced by inactivation of *Tcf7*, which is essential for thymocyte maturation (Schilham et al., 1998), display a gene expression signature similar to that of human ETP-ALL. Notably, ETP-ALLs show reduced expression of *TCF7* compared to non-ETP T-ALLs. However, this study used germline deletion of *Tcf7* and transplantation of thymocyte cells was not performed, so transformation of ETPs themselves could not be confirmed.

Treanor et al. (2014) transduced CD4/CD8 double negative (DN) thymocytes from $Arf^{-/-}$ mice with constitutively active *Il7r* mutant constructs, and transplanted these cells into $Rag1^{-/-}\gamma_c^{-/-}$ immune-deficient recipient mice. Il7r is a cytokine receptor signalling protein essential for early lymphoid differentiation, and mutations in *IL7R* have been found in ETP-ALL (Cramer et al., 2016; Zhang et al., 2012b). $Arf^{-/-}$ donor mice and $Rag1^{-/-}\gamma_c^{-/-}$ recipients were used to improve thymocyte engraftment

upon transplantation. Cultured *Il7r*-mutant DN cells showed a block in maturation beyond the DN2 stage *in vitro*, and transplanted recipients developed leukaemias displaying myeloid or erythroid surface phenotypes, and gene expression mirroring that of ETP-ALL. This study highlights the importance of signalling pathway activation in generation of the ETP leukaemia phenotype, but the model relies on transplantation of the whole DN thymocyte fraction after a period of *in vitro* culture, meaning that transformation of ETPs in the physiological setting cannot be confirmed.

Goossens et al. (2015) identified recurrent mutational activation of the ZEB2 transcription factor in ETP-ALLs, and generated a mouse model of overexpression of *Zeb2* combined with loss of *p53* in endothelial and haematopoietic cells. These mice developed T-cell leukaemias showing some gene expression associated with ETP-ALL. However, as in the first model, mutations were not targeted specifically to lymphoid cells, so the involvement of ETPs in the resulting leukaemias is not clear. Additionally, while *ZEB2* is indeed overexpressed in ETP-ALL relative to non-ETP T-ALL, *P53* mutations have not been found.

Danis et al. (2016) transduced *Cdkn2a*$^{-/-}$*Ezh2*$^{fl/fl}$ Lineage^{-}Sca1^{+}Kit^{+} (LSK) stem and progenitor cells with a constitutively active mutant *Nras*, together with *Cre* to inactivate *Ezh2*. These cells were then cultured in T-cell differentiation-promoting conditions for 14 days before transplantation into wild-type recipient mice. *CDKN2A* is a cell cycle regulator often deleted in ETP-ALL, though less frequently than non-ETP T-ALL (25% of ETP-ALLs versus 81% of non-ETP T-ALLs (Zhang et al., 2012b)). Mutations in components of the *RAS* pathway, such as *NRAS*, are common in ETP-ALL. Recipient mice developed leukaemias showing expression of the myeloid markers Mac1 and Gr1 together with the early thymocyte-associated marker CD44. While the mutations introduced in this model are relevant to ETP-ALL, the evidence presented for an overlap in phenotype with the human disease is somewhat weak. Additionally, and in common with the other ETP-ALL mouse models, mutations were not targeted to ETPs and their transformation was not demonstrated.

In summary, while introducing some combinations of clinically relevant mutations and demonstrating overlap in phenotypic features and gene expression with human ETP-ALL, none of these ETP leukaemia mouse models have specifically targeted mutations to ETPs or demonstrated leukaemia-propagating potential of cells within the stringently defined ETP compartment. Doing so is an important step towards validation of the concept of ETP-LSC leukaemia, with consequences for future patient stratification and treatment.

8: EZH2 and RUNX1 in malignant haematopoiesis

8.1: The PRC2 complex in malignant haematopoiesis

Consistent with its crucial role in multiple aspects of haematopoiesis, mutations in PRC2 components are common in haematological malignancies (Comet et al., 2016). Mutations in the catalytic component *EZH2* are the most common, and are also found in a wide range of non-haematological tumours (Comet et al., 2016). *EZH2* mutations were initially believed to be exclusively oncogenic, after overexpression of the gene was reported in epithelial malignancies (Kleer et al., 2003; Varambally et al., 2002). This view was supported by the identification of somatic mutations of *EZH2* in lymphoma patients, which were shown to increase its catalytic activity (Morin et al., 2010; Yap et al., 2011). However, two studies published in 2010 identified inactivation of *EZH2* in 6-13% of patients with myeloid malignancies (Ernst et al., 2010; Nikoloski et al., 2010). These mutations were identified in patients with acquired uniparental disomy (aUPD) of chromosome 7q. This occurs when a cell inherits two copies of one allele and loses the other allele due to errors in mitosis, meaning that the cell effectively becomes homozygous for any mutations present on the duplicated chromosome. In humans, the *EZH2* gene is found on chromosome 7q, and aUPD patients were found to possess mutations in the gene. The mutations resulted in premature stop codons or missense mutations in highly conserved catalytic regions, which were confirmed by functional analysis to lead

to loss of H3K27 trimethylase activity. Some patients also showed deletion of the entire *EZH2* locus, with inactivating mutations present in the other allele. Mutations of *EZH2* in myeloid malignancies therefore result in complete loss of function of the protein. These results established that *EZH2* is able to act as a tumour suppressor, as well as an oncogene, depending on the cellular context.

A subsequent study published in 2012 found that genomic deletions and inactivating mutations affecting the PRC2 components *EZH2* and *SUZ12* are also present in 25% of T-ALL patients (Ntziachristos et al., 2012). This was complemented by the description of a mouse model of haematopoietic-specific deletion of *Ezh2* which developed T-cell leukaemias (Simon et al., 2012). Detailed structural and computational analysis of the *EZH2* mutations found in ETP-ALL, which overlapped with those found in myeloid malignancies and non-ETP T-ALL, also provided conclusive evidence that these conferred loss of function through disruption of cofactor binding sites or the catalytic domain (Zhang et al., 2012b).

Zhang et al. also provided evidence that *EZH2* mutations in ETP-ALL, similarly to myeloid malignancies, can result in homozygous loss of function, as opposed to haploinsufficiency (Zhang et al., 2012b). 3 of 10 ETP-ALL patients carrying *EZH2* mutations were shown to have deletion of one allele combined with inactivating mutation of the other allele. This was also the case for 1 of the 8 *EED*-mutant patients.

8.2: RUNX1 in malignant haematopoiesis

Like *EZH2*, inactivating mutations of *RUNX1* are present in a range of haematological malignancies (Sood et al., 2017). The human *RUNX1* gene was first identified as the target on chromosome 21 of the t(8;21) recurrent translocation in AML, present in around 18% of patients and resulting in the fusion protein RUNX1-RUNX1T1 (also known as AML1-ETO) (Miyoshi et al., 1991). A large number of different fusion partners of *RUNX1* have since been identified in AML (Sood et al., 2017). These

fusion proteins retain the DNA-binding Runt domain of RUNX1, but not its transactivation domain. This means that it is able to bind to RUNX1 target genes but unable to activate their transcription. A further consequence is that the fusion protein competes with the remaining wild-type RUNX1 allele for binding to its target loci, resulting in dominant-negative repression of *RUNX1* function. The fusion partners may contribute other functions, such as recruitment of transcriptional co-factors to RUNX1 target loci resulting in modulation of gene expression (Sood et al., 2017).

Somatic point mutations of RUNX1 are found in around 9% of MDS patients, and are associated with a poor prognosis (Bejar et al., 2011). These mutations cluster within the Runt domain and result in disruption of DNA binding (Osato, 2004). However, binding of the mutant protein to the RUNX1 co-factor CBFβ is not disrupted, which is thought to lead to dominant negative inhibition of the wild-type RUNX1 allele via sequestration of CBFβ (Osato, 2004). Additionally, in ETP-ALL, deletion of one *RUNX1* allele combined with inactivating mutation of the other allele was detected in 3 of 10 patients (Zhang et al., 2012b). This evidence implicates *RUNX1* somatic mutations as causing complete loss-of-function in both MDS and ETP-ALL.

8.3: Association of EZH2 and RUNX1 inactivating mutations

Mutations in EZH2 and RUNX1 have been shown to be significantly associated in both MDS (Papaemmanuil et al., 2013) and ETP-ALL (Zhang et al., 2012b). In the 2013 study, Papaemmanuil et al. performed targeted next-generation sequencing of genomic DNA for 111 genes previously shown to be implicated in pathogenesis of myeloid malignancy, in 738 MDS patients. The authors identified oncogenic variants in 43 genes and performed sophisticated analyses to determine the variant allele frequencies of each mutation for each patient. As described in section 5.1, this allowed them to infer the relative timing of acquisition of mutations and the phylogenetic relationships between individual malignant clones. The authors also performed an analysis to find genes which were frequently co-

mutated within the same patient, or conversely mutually exclusive. This revealed a significant association of *EZH2* and *RUNX1* mutations. The mechanisms underlying this association, however, are unknown.

Zhang et al. performed a similar analysis of the relative associations of mutations found in ETP-ALL, in their panel of 12 whole-genome sequencing and 52 targeted sequencing patient samples (Zhang et al., 2012b). Similarly to MDS, this revealed a significant association of mutations in *EZH2* and *RUNX1*.

Together, these studies have shown that the mutations in EZH2 and RUNX1 that are found in both MDS and ETP-ALL confer loss of function, and often (or potentially always) act to completely abolish the function of the genes, either through deletion or mutational inactivation of both alleles, or a dominant negative effect of both alleles. Moreover, mutations in these two genes are significantly associated, that is patients showing mutation of one gene are more likely to show mutation of the other. However, the molecular mechanisms underlying this association are unknown. Related to this, it is unknown whether the impact of the combination of these two mutations differs depending on the specific haematopoietic stem and progenitor populations in which they occur. To explore these questions, new mouse models are required. These could potentially provide valuable insights into the pathogenesis of these diseases.

9: BET Inhibition

9.1: BET proteins

The bromodomain and extra terminal domain (BET) family of proteins are a class of epigenetic readers. These are proteins that recognise and bind to specific chromatin modifications in order to recruit modulators of transcription to appropriate genomic loci. In mice and humans, the BET family

consists of four members: BRDT, BRD2, BRD3 and BRD4. Each contains two conserved bromodomains, which recognise and bind to acetylated residues on chromatin N-terminal tails, such as H3K27ac (Filippakopoulos et al., 2012). They then recruit transcriptional apparatus including P-TEFb and Mediator in order to activate transcription (Shi and Vakoc, 2014).

Pharmacological inhibition of the BET proteins has been implicated as a potential therapeutic strategy for haematological malignancies. BET inhibitors were first described in 2010 and 2011, with the molecules JQ1 and I-BET151 respectively (Dawson et al., 2011; Filippakopoulos et al., 2010). Both of these molecules selectively bind to the bromodomains of the BET family, inhibiting their ability to recognise acetylated chromatin. Treatment of leukaemia cell lines with BET inhibitors results in downregulation of *MYC* and *BCL2*-associated transcription, resulting in proliferation arrest (Dawson et al., 2011; Zuber et al., 2011). It was demonstrated using RNAi screening in an AML mouse model that among the BET proteins, Brd4 in particular was essential for leukaemic growth (Zuber et al., 2011). Additionally, while I-BET151 targets all four BET proteins fairly evenly, JQ1 binds especially strongly to BRD4 (Dawson et al., 2011; Filippakopoulos et al., 2010). *BRD4* has therefore been particularly well studied in the context of haematological malignancy (Shi and Vakoc, 2014).

9.2: BET inhibition in PRC2-inactivated malignancy

It has also been suggested that BET inhibition could act to reverse the transcriptional consequences of loss-of-function mutations in EZH2 and other components of PRC2 (De Raedt et al., 2014). At gene loci which are normally kept silenced by PRC2-mediated deposition of H3K27me3, inactivation of the complex can lead to loss of this repressive mark. Activity of histone acetyltransferases such as CBP and P300 can then replace it with H3K27ac (Pasini et al., 2010). One mechanism by which this chromatin mark activates transcription is through recruitment of BET proteins (Filippakopoulos et

al., 2012). This means that inhibition of their function could potentially reverse oncogenic transcription upregulated as a result of PRC2 inactivation, such as in ETP-ALL (Zhang et al., 2012b).

This mechanism was first proposed using a mouse model of myelin peripheral nerve sheath tumours, in which inactivation of the PRC2 component *Suz12* was combined with mutational activation of the Ras signalling pathway (De Raedt et al., 2014). Interestingly, the combination of PRC2 inactivation with RAS pathway activation is particularly common in ETP-ALL (Zhang et al., 2012b). Using BET inhibition to target oncogenic transcription is therefore a potential novel avenue for treatment of this therapy-resistant disease.

10: Hypotheses

10.1: The type of target cell which acquires Ezh2 and/or Runx1 mutation will impact on the biology and clinical picture of the resulting malignant phenotype

Given the observation that inactivating mutations of *EZH2* and *RUNX1* are significantly associated in both MDS and ETP-ALL (Papaemmanuil et al., 2013; Zhang et al., 2012b), we first aimed to investigate the impact of combined inactivation of *Ezh2* and *Runx1* in the haematopoietic system of mice. We wished to explore the different phenotypes generated by inactivation of the two genes either individually or in combination, and by targeting of their inactivation to different haematopoietic populations. Is inactivation of *Ezh2* and *Runx1* sufficient to induce MDS? Is targeting of HSCs necessary for the development of an MDS phenotype? Can ETP leukaemia, rather than MDS, also be initiated following inactivation of *Ezh2* and *Runx1*? Could these two very distinct malignant phenotypes result from targeting of the same mutations to different haematopoietic populations?

To address these questions we utilised floxed alleles of *Ezh2* and *Runx1*, which upon Cre-mediated recombination excise functionally essential domains: the SET methyltransferase domain of Ezh2 (Su

et al., 2003) and the Runt DNA-binding domain of Runx1 (Growney et al., 2005). This results in translation of non-functional protein products from the mRNA transcripts. We used inducible or tissue-specific expression of Cre from three different lines: *Mx1Cre*, *Flt3Cre* and *Rag1Cre*.

Mx1Cre is activated by interferon signalling, which can be induced by intra-peritoneal administration of the synthetic double-stranded RNA molecule Poly(IC) (Kuhn et al., 1995). This results in a high frequency of Cre activation in the liver and haematopoietic cells, including HSCs.

Flt3Cre expression is confined to the haematopoietic system, and is activated in almost all MPPs but not HSCs (Benz et al., 2008; Boyer et al., 2011; Buza-Vidas et al., 2011). This means that almost all mature haematopoietic cells have passed through a Flt3Cre-active stage and show recombination of floxed alleles. HSCs, meanwhile, are exclusively non-recombined.

Rag1Cre becomes active in early lymphoid progenitors (McCormack et al., 2003). Since the Rag1 protein is essential for rearrangement of the immunoglobulin and TCR loci, all mature lymphocytes have passed through a *Rag1Cre*-active stage and show recombination of floxed alleles. In addition, a small proportion of mature myeloid cells (but no cells of the erythroid or megaryocyte lineages) show recombination (Boiers et al., 2013). This indicates that *Rag1Cre* activation initiates in lympho-myeloid progenitors, which mainly give rise to lymphoid cells but retain the potential to produce myeloid cells. These are the progenitor cells which seed the thymus, becoming ETPs upon their arrival (Luc et al., 2012).

10.2: ETPs are a LSC population in ETP leukaemias

Since it has been found that in ETP-ALL, mutations in epigenetic regulators such as *EZH2* and transcriptional regulators such as *RUNX1* are frequently combined with signalling pathway mutations such as in the RAS pathway (Zhang et al., 2012b), we aimed to determine whether combination of *Ezh2* and *Runx1* inactivation in early lymphoid progenitors with the *Flt3-ITD* mutation could

transform ETPs into an LSC population. Though ETP-ALL is hypothesised to originate from transformed ETPs due to its gene expression profile (Coustan-Smith et al., 2009), this has not been experimentally verified and there is as yet no evidence that ETPs are capable of becoming leukaemia-propagating cells.

To address this we first studied the cellular and molecular impacts of *Rag1Cre*-mediated *Ezh2* and *Runx1* inactivation on thymocyte progenitor populations, using flow cytometry analysis as well as gene expression analysis of ETPs. We then combined *Ezh2* and *Runx1* inactivation with constitutive Ras pathway signalling via the *Flt3-ITD* mutation. We used this model to explore whether this combination of mutations is sufficient to induce leukaemogenic transformation of ETPs.

10.3: BET inhibition will reverse PRC2-inactivation associated oncogenic transcriptional programs in ETP leukaemia

Experimental evidence has been obtained implicating BET inhibition as a viable therapeutic strategy for PRC2-inactivated malignancy (De Raedt et al., 2014). We therefore aimed to explore this in mouse and human PRC2-inactivated ETP leukaemias. We used ChIP-sequencing for H3K27me3 and H3K27ac combined with gene expression analysis using RNA-sequencing to investigate the mechanisms of altered gene expression in our *Ezh2*-inactivated mouse ETP leukaemia model, as well as a human *EZH2*-null ETP leukaemia cell line. We then explored the therapeutic efficacy of BET inhibition *in vitro* and *in vivo*, in both our mouse model and PRC2-mutant patient-derived ETP leukaemia xenografts.

Chapter 2: Materials and Methods

1: Animals

1.1: Mouse strains

All mice were bred and maintained in accordance with UK Home Office regulations. $Ezh2^{fl/+}$ (Su et al., 2003), $Runx1^{fl/+}$ (Growney et al., 2005), $Mx1Cre^{tg/+}$ (Kuhn et al., 1995), $Flt3Cre^{tg/+}$ (Benz et al., 2008), $Rag1^{Cre/+}$ (McCormack et al., 2003) and $Flt3^{ITD/+}$ mice (Lee et al., 2007) have been previously described. All mice, except for NSG recipient mice (Shultz et al., 2005), were on a pure C57BL/6 background.

Genotyping of mice was performed by extraction of DNA using the EZNA Tissue DNA Extraction Kit (Omega Bio-Tek), followed by PCR using Phire Hot Start II DNA Polymerase (Thermo Fisher). See table below for primer sequences.

Primers used for mouse genotyping:

Allele	**Primer name**	**Primer sequence**
Ezh2	Ezh2-F	CTGCTCTGAATGGCAACTCC
	Ezh2-R	GGTCACTGGGTTCTTTCTAAGG
Runx1	Runx1-F	GGTGATGGTCAGAGTGAAGC
	Runx1-R	GAGTCCCAGCTGTCAATTCC
Generic Cre (Mx1Cre, Flt3Cre, Rag1Cre)	GenCre-F	CGTTTTCTGAGCATACCTGGA
	GenCre-R	ATTCTCCCACCGTCAGTACG
Flt3-ITD	Flt3ITD-F	AGGTACGAGAGTCAGCTGCAGATG
	Flt3ITD-R	TGTAAAGATGGAGTAAGTGCGGGT

1.2: Poly(IC) treatment

For *Mx1Cre*-induced deletion of *Ezh2* and *Runx1*, $Ezh2^{fl/fl}Runx1^{fl/fl}Mx1Cre^{tg/+}$ mice and $Ezh2^{fl/fl}Runx1^{fl/fl}Mx1Cre^{+/+}$ littermate controls were treated at 6-8 weeks of age with 3 doses of 200 µg of Poly(IC) (GE Healthcare) at 2-day intervals by intra-peritoneal injection. C57BL/6 CD45.1 recipient mice transplanted with $Ezh2^{fl/fl}Runx1^{fl/fl}Mx1Cre^{+/+}$ or $Ezh2^{fl/fl}Runx1^{fl/fl}Mx1Cre^{tg/+}$ bone marrow cells were treated in the same way 4 weeks following transplantation.

1.3: Transplantation experiments

For transplantation experiments, C57BL/6 CD45.1 recipient mice were lethally irradiated using a split dose of 2 x 5 Gy irradiations, 4 hours apart. Recipients were then transplanted via intravenous injection with the indicated CD45.2 cell populations, together with 250,000 (unless otherwise indicated) unfractionated CD45.1 bone marrow support cells. NSG mice were sublethally irradiated using a split dose of 2 x 1.25 Gy irradiations, 4 hours apart. Recipients were then transplanted via intravenous injection with 2-2.5 million ETP-ALL xenograft spleen cells.

Peripheral blood readouts were performed at 2-4 week intervals following transplantation. A maximum of 10% of the total blood volume of the mouse was taken from the tail vein per readout, and a maximum of 15% per 4-week period. This equated to around 100 µl for 4-weekly bleedings, and 70 µl for 2-weekly bleedings.

When analysis of haematological parameters was performed alongside flow cytometry analysis, 20 µl of blood was withdrawn and diluted 1/4 in PBS for analysis using a Sysmex KX-21N analyser. The remainder was processed as described below for analysis of leukocytes by flow cytometry.

2: Flow Cytometry

2.1: Sample preparation for flow cytometry

For analysis of peripheral blood, blood was collected into EDTA-coated tubes (Sigma) and kept at room temperature until processing. Haematological parameters were measured using a Sysmex KX-21N analyser. For analysis of leukocytes by flow cytometry, blood was mixed in a 1:1 ratio with 1% dextran solution and incubated at 37 C for 20-25 min to separate leukocytes from erythrocytes and platelets. The supernatant, containing leukocytes, was withdrawn and spun at 1000 g for 4 min. The cell pellet was resuspended in 200 µl ammonium chloride solution for 1-2 min to lyse residual erythrocytes, before the reaction was quenched with 1 ml PBS containing 5% FCS (HyClone) (PBS/FCS). Cells were pelleted again and resuspended in 25 µl antibody staining cocktail. Antibody clones and suppliers are detailed in the tables below. Following incubation on ice for 15 min, cells were washed with 1 ml PBS/FCS before pelleting and resuspension in 100 µl PBS/FCS, and kept on ice for analysis by flow cytometry. No more than 1 hour before analysis, 7AAD or DAPI was added to the cell suspension at optimised concentrations for discrimination of live from dead cells.

For analysis of bone marrow, spleen and thymus, tissues were extracted from culled mice and kept on ice until processing. Bone marrow cells were extracted from femurs and tibias into 5-8 ml PBS/FCS by crushing using a pestle and mortar. Cells were suspended by pipetting and filtered before being counted using the Sysmex KX-21N. Single-cell suspensions of spleen and thymus were made by crushing against a filter using a 2 ml syringe plunger and washing through with 5-8 ml PBS/FCS. Cells were counted in the same way as bone marrow cells.

For analysis of E14.5 foetal liver, embryos were dissected to collect whole foetal livers in 1 ml Dulbecco's PBS (Thermo Fisher) containing 10% FCS (DPBS/FCS), and tail tissue for embryo genotyping. Foetal livers were kept on ice until processing. Single cell suspensions were obtained by passing foetal livers through 18 G and then 25 G needles using a 1 ml syringe. Cells were counted in the same way as bone marrow cells.

Bone marrow, spleen and thymus cells in single-cell suspension were prepared for flow cytometry analysis by first pelleting at 500 g for 5 min and resuspending in 12.5-50 µl Fc block solution (or more when large numbers of cells were required for sorting) for 10 min on ice. 2X concentrated antibodies were added in a 1:1 ratio and incubated for 15 min on ice. Antibody clones and suppliers are detailed in the tables below. Cells were then washed with 1 ml PBS/FCS before being pelleted and resuspended in 100-200 µl PBS/FCS and kept on ice for flow cytometry analysis. For sorting, cells were resuspended in larger volumes at lower cell concentrations. Thymus and foetal liver cells were filtered before analysis to prevent clogging. No more than 1 hour before analysis, 7AAD or DAPI was added to the cell suspension at optimised concentrations for discrimination of live from dead cells.

2.2: Phospho-flow cytometry

Unfractionated bone marrow cells were cultured in Opti-MEM (Gibco) at a concentration of 5 million cells per ml, without serum or cytokines for 2 hours, followed by stimulation with 8 ng/ml mSCF (PeproTech) for 10 minutes. For MEK inhibitor experiments, either 0.05% DMSO or 2.5 µM PD-901 (Cambridge Biosciences) were added to the media for the entire culture period. Immediately after SCF stimulation, cells were fixed by adding 100 µl of 16% paraformaldehyde (Electron Microscopy Sciences) to 1 ml cells suspended in culture media and incubating for 10 min at 22ºC. Cells were then pelleted by spinning at 500 g for 5 min before being resuspended in 25 ul Fc block and incubated on ice for 10 min. Pacific blue-conjugated Sca1 antibody at 2X concentration was added in a 1:1 ratio and incubated on ice for 15 min. Cells were washed with 1 ml PBS/FCS and pelleted, then permeabilised by resuspending in 300 µl ice-cold acetone for 15 minutes. Two PBS/FCS washes were performed to remove acetone, before cells were resuspended in 50 µl staining cocktail containing phospho-S6 (APC conjugate), Kit (APC-eF780 conjugate) and lineage markers (B220, Ter119, Mac1 and Gr1, all conjugated to PE-Cy5) on ice for 15 min. Antibody clones and suppliers are detailed in the tables below. Cells were then washed and resuspended in 150 µl PBS/FCS for flow cytometry analysis. $Lin^{-}Kit^{+}$ cells were gated for analysis of phospho-S6.

2.3: Flow cytometry antibody clones and suppliers

Antibodies used for lineage analysis of peripheral blood, bone marrow and spleen:

Antigen	Conjugate	Clone	Supplier
CD45.2	AF700	104	BioLegend
Mac1	AF700	M1/70	eBioscience
CD3e	APC	145-2C11	eBioscience
CD48	APC	HM48-1	BioLegend
Kit	APC	2B8	eBioscience
Mac1	APC	M1/70	BioLegend
CD4	APC-eF780	RM4-5	eBioscience
CD8	APC-eF780	53-6.7	eBioscience
Ter119	BV650	TER-119	BioLegend
CD45.1	FITC	A20	BioLegend
CD5	FITC	53-7.3	Becton Dickinson
Mac1	FITC	M1/70	eBioscience
Gr1	Pacific Orange	RB6-8C5	Caltag Medsystems
Flt3	PE	A2F10	BioLegend
Mac1	PE	M1/70	eBioscience
CD19	PE-Cy7	1D3	eBioscience

Antibodies used for analysis of bone marrow and spleen haematopoietic stem/progenitor cells:

Antigen	Conjugate	Clone	Supplier
CD45.2	AF700	104	BioLegend
CD48	APC	HM48-1	BioLegend

Antigen	Conjugate	Clone	Supplier
Phospho-S6	APC	D57.2.2E	Cell Signaling
Kit	APC-eF780	2B8	eBioscience
Sca1	BV605	D7	BioLegend
CD45.1	FITC	A20	BioLegend
Sca1	Pacific Blue	E13-161.7	BioLegend
B220	PE-Cy5	RA3-6B2	BioLegend
CD4	PE-Cy5	RM4-5	BioLegend
CD5	PE-Cy5	53-7.3	BioLegend
CD8a	PE-Cy5	53-6.7	BioLegend
Gr1	PE-Cy5	RB6-8C5	BioLegend
Mac1	PE-Cy5	M1/70	BioLegend
Ter119	PE-Cy5	TER-119	BioLegend
CD150	PE-Cy7	TC15-12F12.2	BioLegend

Antibodies used for thymus analysis and sorting:

Antigen	Conjugate	Clone	Supplier
CD4	AF700	RM4-5	eBioscience
B220	APC	RA3-6B2	BioLegend
CD11c	APC	N418	eBioscience
CD19	APC	1D3	eBioscience
CD3e	APC	145-2C11	eBioscience
Gr-1	APC	RB6-8C5	BioLegend
NK1.1	APC	PK136	eBioscience
Kit	APC-eF780	2B8	eBioscience

Antigen	Conjugate	Clone	Supplier
CD44	FITC	IM7	eBioscience
Flt3	PE	A2F10	BioLegend
CD8	PE-Cy7	53-6.7	eBioscience
CD25	PerCP-Cy5.5	PC61	BioLegend

Antibodies used for sorting bone marrow $Lin^{-}Kit^{+}$ progenitor cells:

Antigen	Conjugate	Clone	Supplier
CD45.2	AF700	104	BioLegend
Kit	APC	2B8	eBioscience
CD4	APC-eF780	RM4-5	eBioscience
CD8	APC-eF780	53-6.7	eBioscience
CD45.1	FITC	A20	BioLegend
Nk1.1	Pacific Blue	PK136	BioLegend
Mac1	PE	M1/70	eBioscience
CD19	PE-Cy7	1D3	eBioscience
B220	PE-Texas Red	RA3-6B2	Becton Dickinson

Antibodies used for chimerism analysis of patient-derived xenografts:

Antigen	Conjugate	Clone	Supplier
Mouse CD45	PE-Texas Red	30-F11	Invitrogen
Human CD45	AF700	HI30	BioLegend
Human CD34	APC	518	BioLegend

2.4: Configurations of instruments used for flow cytometry

LSRII, AriaII

Laser	Filter	Conjugate
Red (640nm)	670/14	APC
	730/45	AF700
	780/60	APC-eF780
Blue (488nm)	525/50	FITC
Violet (405nm)	450/50	DAPI, Pacific Blue
	525/50	Pacific Orange
	605/20	BV605
Green (532nm)	575/25	PE
	610/20	Prodium Iodide, PE-Texas Red
	685/35	7AAD, PE-Cy5
	710/40	PerCP-Cy5.5
	780/60	PE-Cy7

LSRFortessa, LSRFortessa X20, AriaIII, Aria Fusion

Laser	Filter	Conjugate
Red (633nm)	670/20	APC
	730/45	AF700
	780/60	APC-eF780
Blue (488nm)	530/30	FITC
Violet (405nm)	450/50	DAPI, Pacific Blue
	525/50	Pacific Orange
	610/20	BV605
Yellow/Green (561nm)	585/15	PE
	610/20	Prodium Iodide, PE-Texas Red

Laser	Filter	Conjugate
	670/30	7AAD, PE-Cy5
	710/50	PerCP-Cy5.5
	780/60	PE-Cy7

2.5: Surface marker definitions of stem and progenitor populations

Bone marrow and spleen

Lineage negative (Lin^-): $CD4^-CD5^-CD8a^-B220^-Mac1^-Gr1^-Ter119^-$

LSK: $Lin^-Kit^+Sca1^+$

HSC: $Lin^-Kit^+Sca1^+CD150^+CD48^-$

MPP: $Lin^-Kit^+Sca1^+CD150^-CD48^+$

DP: $Lin^-Kit^+Sca1^+CD150^+CD48^+$

LMPP: $Lin^-Kit^+Sca1^+Flt3^{high}$ (Top 25% LSK cells by Flt3 expression)

LK: $Lin^-Kit^+Sca1^-$

Foetal liver

Lineage negative (Lin^-): $CD3e^-NK1.1^-CD19^-B220^-F4/80^-Gr1^-Ter119^-$

LSK: $Lin^-Kit^+Sca1^+$

HSC: $Lin^-Kit^+Sca1^+CD150^+CD48^-$

MPP: $Lin^-Kit^+Sca1^+CD150^-CD48^+$

DP: $Lin^-Kit^+Sca1^+CD150^+CD48^+$

Thymus

Lineage negative (Lin^-): $CD3e^-CD19^-B220^-NK1.1^-CD11c^-Gr1^-$

DN: $Lin^-CD4^-CD8^-$

DN1: $Lin^-CD4^-CD8^-CD44^+CD25^-$

ETP: $Lin^-CD4^-CD8^-CD44^+CD25^-Kit^+Flt3^+$

DN2: $Lin^-CD4^-CD8^-CD44^+CD25^+$

DN3: $Lin^-CD4^-CD8^-CD44^-CD25^+$

DN4: $Lin^-CD4^-CD8^-CD44^-CD25^-$

DP: $Lin^-CD4^+CD8^+$

SP: $Lin^-CD4^+CD8^-$ or $Lin^-CD4^-CD8^+$

3: RNA-Sequencing

3.1: Sample preparation for RNA-sequencing

ETPs were prepared for RNA-seq as previously described (Ramskold et al., 2012), using the SMARTer Ultra Low RNA kit for Illumina Sequencing (Clontech). One hundred cells were sorted directly into lysis buffer supplemented with RNase Inhibitor (Clontech).

Preparation of ETP RNA for sequencing was performed by Alice Giustacchini and Tiphaine Bouriez-Jones:

> Preparation of cDNA libraries was performed as directed by the SMARTer Clontech kit using 15 cycles of amplification. Amplified cDNA libraries were sheared using the Covaris AFA system. The NEBNext DNA Library Prep Reagent Set for Illumina (New England Biolabs) was then used to repair ends, adenylate 3' ends, ligate genomic adaptors and amplify fragments with Genomic DNA Sample Prep oligos (Illumina). Samples were sequenced using the Illumina HiSeq 2000 platform, generating 50 bp single-end reads.

$Mac1^+$ BM cells were prepared for RNA-seq using Smart-Seq2 as previously described (Picelli et al., 2014). One hundred cells were sorted directly into 4 µl of lysis buffer containing 0.2% Triton X-100 (Sigma-Aldrich), RNase inhibitor (Clontech), 2.5 mM dNTPs (Thermo Fisher) and 2.5 µM oligo-$dT_{30}VN$ primer (Biomers.net).

Preparation of Mac1$^+$ bone marrow cell RNA for sequencing was performed by Neil Ashley:

> cDNA was generated using SuperScript II (Invitrogen), pre-amplified using KAPA HiFi HotStart ReadyMix (KAPA Biosystems) and purified using Agencourt AMPure XP (Beckman Coulter). Tagmentation and library preparation was performed using the Nextera XT DNA Library Preparation Kit (Illumina). Samples were sequenced using the Illumina NextSeq 500 platform, generating 75 bp paired-end reads.

LOUCY and JURKAT RNA was extracted from 500,000 cells per sample using the RNeasy Mini Kit (Qiagen). Ten nanograms of RNA per sample was prepared for sequencing using Smart-Seq2 as for Mac1$^+$ BM cells above, but generating 75 bp single-end reads. For RNA-seq following inhibitor treatment, 500,000 cells per sample were first cultured for 24 hours with either 0.05% DMSO, 250 nM JQ1 (MedChem Express) or 500 nM I-BET151 (Sigma-Aldrich).

3.2: ETP and Mac1$^+$ bone marrow cell RNA-sequencing data analysis

Analysis of mouse ETP and Mac1+ bone marrow cell RNA-sequencing data was performed by Nikolaos Barkas:

> Mouse ETP and Mac1$^+$ BM cell RNA-sequencing data quality was assessed with FastQC (version 0.10.10) [http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]. Reads were aligned to the *in silico* generated transcriptome of the mm10 mouse genome using Tophat2 (version 2.0.10) (Kim et al., 2013) and Bowtie2 (version 2.1.00) (Langmead and Salzberg, 2012) as the underlying aligner. Non-uniquely mapping reads were filtered using samtools (version 0.1.19) (Li et al., 2009). Read alignments were summarized to gene identifiers using the htseq-count tool (version 0.5.4p5) (Anders et al., 2015) against the mm10 UCSC transcriptome annotation. Differential expression analysis was performed using the generalized linear model function (McCarthy et al., 2012) of edgeR (version 3.10.4) (Robinson et al., 2010). Genes with fewer than 2 reads in at least 3 samples were excluded

from subsequent analysis. RPKM value calculation was performed using the rpkm function of edgeR.

3.3: LOUCY and JURKAT RNA-sequencing data analysis

LOUCY and JURKAT reads were aligned to the human genome (hg38) using STAR (version 2.5.1b) (Dobin et al., 2013) and the following parameters:

Parameter	**Argument**	**Function**
--genomeDir	/databank/igenomes/Homo_sapiens /UCSC/hg38/Sequence/STAR/	Location of genome directory
--outSAMtype	BAM SortedByCoordinate	Output file type (sorted BAM)
--readFilesCommand	zcat	Input files (gzip compressed)
--runThreadN	4	Number of threads to use
--readFilesIn	…/fastq.gz,…/fastq.gz	Location of input files
--outFileNamePrefix	...SampleXX_	Location and prefix of output files

Reads mapping to gene exons were counted using the Subread (version 1.4.5-p1) featureCounts tool (Liao et al., 2014) and UCSC hg38 annotation, with the following parameters:

Parameter	**Argument**	**Function**
-t	exon	Feature type to count (exons)
-g	gene_id	Attribute to assign counts to (gene name)
-T	5	Number of threads to use
-a	/databank/igenomes/Homo_sapiens/UCSC /hg38/Annotation/Genes/genes.gtf	Annotation file location
-o	…/LOUCY-JURKAT_featurecounts.txt	Output file location

EdgeR (version 3.10.5) (Robinson et al., 2010) was used to generate the MDS plot for all samples, RPKM values for all samples and all genes, and fold change values between genotypes and treatments for all genes. Genes detected at less than 1 count per million (cpm) in at least 4 samples were discarded from analysis.

3.4: Gene set enrichment analysis

Gene set enrichment analysis was performed using GSEA software (v 2.2.0) (Mootha et al., 2003; Subramanian et al., 2005). The Hallmark gene sets were obtained from MSigDB (Subramanian et al., 2005) and converted to mouse gene names using the Biomart (Durinck et al., 2009) R package prior to analysis. ETP-ALL gene sets were derived from the top 500 up- or downregulated genes in ETP-ALL relative to non-ETP T-ALL (Zhang et al., 2012b), and converted to mouse gene names using Biomart. For ETP-ALL GSEA, previously published RNA-seq data was obtained from a mouse Notch1 mutation-induced T-ALL (Ntziachristos et al., 2014) and wild-type mouse DP thymocytes (Zhang et al., 2012c), and compared to DKOITD $Mac1^{+}$ BM cells or DKOITD ETPs respectively. ETP-ALL gene sets were compared directly to untreated LOUCY and JURKAT cell gene expression using human gene names.

Custom gene expression signatures for early T-cells, myeloid cells and HSCs were generated by Nikolaos Barkas as follows:

> Previously published data (Luc et al., 2012; Seita et al., 2012) were used to perform differential expression comparisons between ETPs and HSCs, and DN2 and HSC (Luc et al., 2012), and between DN1 and HSCs, and GMPs and HSCs (Seita et al., 2012). Gene sets up-regulated to each of the following cell types were generated by combining genes from the above comparisons as follows: early T-cells (ETP against HSC, DN2 against HSC and DN1 against HSC); Myeloid (GMPs against HSCs). The early T-cell gene set consisted of all genes upregulated in early T-cells but not myeloid cells, while the myeloid gene set consisted of all genes upregulated in myeloid cells but not early T-cells. The HSC gene set consisted of all

genes upregulated in HSCs relative to any other cell type, and not present in the early T-cell or myeloid sets.

The PRC2 target gene set (De Raedt et al., 2014), Pre-GM and Pre-MegE gene sets (Sanjuan-Pla et al., 2013) have been previously described.

4: Quantitative PCR

4.1: Multiplex qPCR

Multiplex quantitative real time PCR analysis was performed using the BioMark 96.96 Dynamic Array platform (Fluidigm) and TaqMan gene expression assays (Life Technologies; see below for full list of assays used) as previously described (Tehranchi et al., 2010). CellsDirect One-Step qRT-PCR kit (Invitrogen) was used for cDNA synthesis and pre-amplification of target genes. Cells were FACS-sorted directly into 5 µl (single cells) or 10 µl (bulk; 50 cells) of lysis/pre-amplification buffer. Ten microliters of this buffer consisted of: 2.5 µl Taqman assay mix containing all assays at 0.2x dilution, 5 µL CellsDirect 2x Reaction mix (Invitrogen), 1.2 µL CellsDirect RT/Taq mix (Invitrogen), 1.2 µL TE buffer and 0.1 µL SUPERase-In RNase Inhibitor (Ambion). Reverse transcription and target pre-amplification were performed as follows: 50ºC for 15 min; 95ºC for 2 min; 22 cycles of 95ºC for 15 s and 60ºC for 4 min. Pre-amplified product was diluted 1:5 in TE buffer and analyzed using the default M96 RT-PCR protocol.

For analysis of deletion efficiency of *Ezh2* and *Runx1* in bulk (50-cell) samples of HSCs, LMPPs and ETPs, Ct values for Taqman probes specific for transcripts from non-deleted *Ezh2* and *Runx1* alleles were normalised to those for *Gapdh* to generate dCt values. 2^{-dCt} values were plotted to measure level of expression of non-deleted *Ezh2* and *Runx1* alleles in cells from WT, DKO and DKOITD mice. *Actb* expression normalised to *Gapdh* was used to confirm largely uniform expression of *Gapdh* across samples.

4.2: Analysis of single cell qPCR data

Single cell qPCR data analysis was performed by Nikolaos Barkas:

> Cells were filtered so as to only retain cells that expressed *Actb*, *B2M, Gapdh* and *Hprt* with Ct values lower than 21. Furthermore at least 70 assays were required to result in amplification per cell for it to be included in the analysis. Cells with consistently low amplification (mean non drop-out Ct > 20) were excluded from the analysis. Absolute expression was calculated as 2^{-Ct}. Ct values were standardized per gene, after excluding measurements that did not give rise to any amplification, for the purposes of visualization. Significance of differential expression was assessed by means of the Kolmogorov-Smirnov non-parametric test on the raw Ct values after FDR multiple testing correction.

4.3: Analysis of bulk qPCR data

Bulk ETP qPCR analysis was performed by Nikolaos Barkas:

> Bulk qPCR analysis was performed in R (version 3.2.2). Assays with Ct values over 30 and reactions that failed quality control by the instrument were excluded. Assays with undetermined values in more than two samples were excluded from the analysis. The mean Ct value of technical replicates was obtained and used for further analysis. Delta Ct normalization was performed using *Gapdh* as reference.

Lists of lineage-associated genes used in analysis of RNA-seq and qPCR data were derived from previous publications (Boiers et al., 2013; Grover et al., 2014; Luc et al., 2012; Mead et al., 2013; Riddell et al., 2014; Roy et al., 2012; Sanjuan-Pla et al., 2013).

4.4: Taqman probes used for gene expression analysis of ETPs

Target gene	**Assay ID**

Target gene	Assay ID
Actb	Mm00607939_s1
Anxa1	Mm00440225_m1
Arpp21	Mm00473630_m1
B2M	Mm00437762_m1
Bcl11a	Mm00479358_m1
Bmi1	Mm03053308_g1
Ccl9	Mm00441260_m1
Ccr7	Mm01301785_m1
Ccr9	Mm02620030_s1
Cd160	Mm00444461_m1
Cebpb	Mm00843434_s1
Cish	Mm01230623_g1
Cited2	Mm01188099_g1
Cpa3	Mm00483940_m1
Csf1r	Mm01266652_m1
Csf2ra	Mm00438331_g1
Csf3r	Mm00432735_m1
Ctsc	Mm00515580_m1
Ctsg	Mm00456011_m1
Cxcr4	Mm01996749_s1
Dtx1	Mm00492297_m1
Dyrk1a	Mm00432934_m1
Erg	Mm01214246_m1
Esam	Mm00518378_m1

Target gene	Assay ID
Etv6	Mm00468390_m1
Ezh2 - WT and Deleted	Mm00468449_m1
Ezh2 - WT only	Mm00468464_m1
F2r	Mm00438851_m1
Fcer1g	Mm02343757_m1
Fcgr3	Mm00438882_m1
Fes	Mm01318102_m1
Fli1	Mm00484410_m1
Flt3	Mm01210882_m1
Fos	Mm00487425_m1
Gapdh	Mm99999915_g1
Gata2	Mm00492301_m1
Gata3	Mm00484683_m1
Gfi1	Mm00515855_m1
Gfi1b	Mm00492318_m1
Hes1	Mm01342805_m1
Hlf	Mm00723157_m1
Hoxa9	Mm00439364_m1
Hprt	Mm01545399_m1
Id1	Mm00775963_g1
Id2	Mm00711781_m1
Igf2r	Mm00439576_m1
Ikzf1	Mm01187882_m1
Il2ra	Mm00434261_m1

Target gene	Assay ID
Il3ra	Mm00434273_m1
IL4	Mm00445259_m1
Il7r	Mm00434295_m1
Itgam	Mm00434455_m1
Jag1	Mm00496902_m1
Jun	Mm00495062_s1
Kit	Mm00445212_m1
Lat	Mm00456761_m1
Lck	Mm00802897_m1
Lef1	Mm00550265_m1
Lmo2	Mm01281680_m1
Ly6a	Mm00726565_s1
Ly6d	Mm00521959_m1
Lyz1	Mm00657323_m1
Lyz2	Mm01612741_m1
Mcl1	Mm00725832_s1
Mef2c	Mm01340842_m1
Meis1	Mm00487664_m1
Mpl	Mm00440310_m1
Mpo	Mm01298424_m1
Ms4a6b	Mm00728225_m1
Ms4a6c	Mm00459296_m1
Myb	Mm00501741_m1
Myc	Mm00487804_m1

Target gene	**Assay ID**
Mycn	Mm00476449_m1
Nfe2	Mm00801891_m1
Notch1	Mm00435249_m1
Nrarp	Mm00482529_s1
Pbx1	Mm04207617_m1
Pecam1	Mm01242584_m1
Pim1	Mm00435712_m1
Pim2	Mm00454579_m1
Prkcd	Mm00440891_m1
Ptcra	Mm00478363_m1
Rag1	Custom assay
	Forward: TGTGGAGCAAGGTAGCTTAGC
	Reverse: TCATCGGGTGCAGAACTGAAG
	Reporter: CATGGCTGCCTCCTTG
Rag2	Mm00501300_m1
Runx1 - WT and Deleted	Mm01213405_m1
Runx1 - WT only	Mm01213404_m1
Runx3	Mm00490666_m1
Selp	Mm01295931_m1
Socs1	Mm00782550_s1
Socs2	Mm00850544_g1
Spi1	Mm00488142_m1
Sterile *IgH*	Custom assay

Target gene	Assay ID
	Forward: GGACTTTGGGATGGGTTTGGTT
	Reverse: CCCTGGTCCTAGACATCAGAGTAAT
	Reporter: CCCAGATGAAGGGCTAC
Tcf12	Mm00441699_m1
Tcf3	Mm01175588_m1
Tcf7	Mm00493445_m1
Zfp467	Mm00465678_g1

5: Whole Exome Sequencing

5.1: Sample preparation for whole exome sequencing

For whole exome sequencing, genomic DNA was extracted from bone marrow cells using the DNeasy Blood & Tissue Kit (Qiagen). DNA was prepared for sequencing using the SureSelectXT Mouse All Exon Kit (Agilent). Sequencing was performed using the Illumina HiSeq 4000 platform, generating 75 bp paired-end reads.

5.2: Whole exome sequencing data analysis

Analysis of whole exome sequencing data was performed by Noushin Farnoud and Elli Papaemmanuil.

6: Chromatin Immunoprecipitation

6.1: Sample preparation for ChIP

ChIP and ChIP-sequencing library preparation was performed by Wen Hao Neo:

For WT and DKOITD bone marrow ChIP-seq, three samples of WT and five samples of DKOITD, each sample consisting of cells from separate mice, were used. WT samples were

first enriched for Kit expressing cells using CD117 microbeads and LS columns (Miltenyi Biotec) as directed. 5 million cells per sample were used for ChIP-seq. For H3K27me3 and H3K27ac ChIP, cells were fixed with 1% formaldehyde for 10 min. Fixed chromatin samples were fragmented using an S220 Focused-ultrasonicator (Covaris) to an average size of 200-500 bp. Immunoprecipitation was then performed using a mixture of Protein A and Protein G Dynabeads (Life Technologies). Antibodies used were: anti-H3K27me3 (07-449, Merck MilliPore) and anti-H3K27ac (C15410196, Diagenode). ChIP samples were quantified relative to inputs as previously described (Milne et al., 2009). One input control sample was used for each genotype or cell line. Samples were prepared for sequencing using the NEBNext Ultra II DNA Library Prep Kit for Illumina, with NEBNext Multiplex Oligos for Illumina (New England BioLabs) as directed. Sequencing was performed using the Illumina NextSeq 500 platform, generating 40 bp paired-end reads.

6.2: Mouse bone marrow ChIP-seq data analysis

For WT and DKOITD ChIP-seq analysis, reads were aligned to the mouse genome (mm10) using Bowtie2 (version 2.2.5) (Langmead and Salzberg, 2012) and the following parameters:

Parameter	**Argument**	**Function**
-q		Input files are fastq format
-N	1	Number of mismatches allowed in seed sequence
-x	/databank/igenomes/Mus_musculus/UCSC/mm10/Sequence/Bowtie2Index/genome	Genome index location
-1	…/_R1_001.fastq	Input read pair 1
-2	…/_R2_001.fastq	Input read pair 2
-S	…/SampleXX.sam	Output file

Aligned reads were converted to BAM format, sorted and indexed using Samtools (version 1.3). Peak calling was performed using the Macs2 callpeak function (version 2.0.10) (Zhang et al., 2008) and the following parameters:

Parameter	**Argument**	**Function**
-t	…/SampleXX.bam	Input treatment (sample) file location
-c	…/InputXX.bam	Input control (input) file location
-f	BAMPE	Input files are BAM paired end format
-g	mm	Genome size (Mus musculus)
--broad		Call broad regions of strong signal as single peaks
-n	SampleXX_H3K27xx	Name of output file
-B		Keep bedgraph files (used for generating signal tracks, see below)
--outdir	…/SampleXX/	Output directory

Signal tracks (comparing to input) for genome visualisation in the UCSC genome browser (http://genome.ucsc.edu/) were generated using the Macs2 bdgcmp function and the following parameters:

Parameter	**Argument**	**Function**
-t	…/SampleXX_treat_pileup.bdg	Input treatment (sample) bedgraph file
-c	…/SampleXX_input_control.bdg	Input control (input) bedgraph file
-m	ppois	Method for calculating signal enrichment over input control (Poisson p-value).

DiffBind (version 1.14.6) (Stark and Brown, 2011) was used to identify regions associated with different levels of H3K27ac and H3K27me3 in DKOITD versus WT samples. This used an edgeR analysis to compare number of reads (normalised to library size) at each peak location called by

Macs2 in all samples, to determine whether the signal at that peak was significantly different between the genotypes.

Identified peaks were mapped to nearby transcription start sites using the Homer annotatePeaks function.pl (version 4.7) (Heinz et al., 2010) and the default mm10 annotation. Lists of genes showing peaks within 2.5 kb of the transcription start site were obtained for peaks upregulated or downregulated in DKOITD relative to WT cells for H3K27me3 and H3K27ac. Predicted genes containing the name "Rik" were excluded.

Genes showing both loss of H3K27me3 and gain of H3K27ac in DKOITD relative to WT cells were used to create a geneset for GSEA comparing gene expression in WT and DKOITD $Mac1^{+}$ bone marrow cells. Significance of overlap of genes showing reduced H3K27me3 and increased H3K27ac with genes showing upregulated expression in mouse bone marrow cells was assessed using the R "phyper" function. The number of genes detected in either genotype by RNA-sequencing (11,172) was used as the total number of genes.

For generation of histograms showing average read density at all transcription start sites genome-wide, WT and DKOITD samples were merged to create one BAM file per histone modification per genotype. Tag directories for each merged sample were created using the Homer makeTagDirectory function. Histogram tables were then created using the Homer annotatePeaks.pl function and the following parameters:

Parameter	**Argument**	**Function**
tss		Transcription start site-centred analysis
mm10		Mouse genome annotation
-size	4000	4 kb window around transcription start sites

Parameter	Argument	Function
-hist	50	Create histogram of read densities with 50 bp bins
-d	WT-H3K27xx/ DKOITD-H3K27xx/	Locations of tag directories

To generate histograms of average read density and heatmaps of read density per gene, Homer annotatePeaks.pl was run using the same parameters except -ghist instead of -hist. This list was filtered using RefSeq IDs of genes found by DiffBind analysis to show reduced H3K27me3 at the transcription start site. Histograms were produced by averaging read density at each 50 bp bin, and heatmaps were produced using the heatmap.3 function of the GMD package (Zhao et al., 2011) for R.

6.3: LOUCY and JURKAT ChIP-seq analysis

For LOUCY and JURKAT ChIP-seq analysis, reads were aligned to the human genome (hg38) using Bowtie2 (version 2.2.5) (Langmead and Salzberg, 2012) and the following parameters:

Parameter	Argument	Function
-q		Input files are fastq format
-N	1	Number of mismatches allowed in seed sequence
-x	/databank/igenomes/Homo_sapiens/UCSC /hg38/Sequence/Bowtie2Index/genome	Genome index location
-1	…/_R1_001.fastq	Input read pair 1
-2	…/_R2_001.fastq	Input read pair 2
-S	…/SampleXX.sam	Output file

Aligned reads were converted to BAM format, sorted and indexed using Samtools (version 1.3). Peak calling was performed using the Macs2 callpeak function (version 2.0.10) (Zhang et al., 2008) and the following parameters:

Parameter	**Argument**	**Function**
-t	…/SampleXX.bam	Input treatment (sample) file location
-c	…/InputXX.bam	Input control (input) file location
-f	BAMPE	Input files are BAM paired end format
-g	hs	Genome size (Homo sapiens)
--broad		Call broad regions of strong signal as single peaks
-n	SampleXX_H3K27xx	Name of output file
-B		Keep bedgraph files (used for generating signal tracks, see below)
--outdir	…/SampleXX/	Output directory

Identified peaks were mapped to nearby transcription start sites using the Homer annotatePeaks function.pl (version 4.7) (Heinz et al., 2010) and the hg38 annotation data (NCBI RefSeq Curated) downloaded from the UCSC Table Browser (Karolchik et al., 2004). Lists of genes showing H3K27ac peaks within 1 kb of the transcription start site in either LOUCY sample but neither JURKAT sample, or in either JURKAT sample but neither LOUCY sample were obtained and used for GSEA.

7: Morphological Analysis and Histology

7.1: Morphological analysis

Peripheral blood smears were obtained from 3.5 µl of blood, and bone marrow cytospins were obtained by centrifugation of 100,000 cells at 500 rpm for 10 min. Slides were stained in May-Grünwald solution (Sigma) for 5 min followed by Giemsa solution (Sigma) for 20 min. Morphological

analysis of peripheral blood leukocytes and bone marrow cells was performed by Adam Mead. One hundred cells per sample were classified according to morphology.

7.2: Histology and immunohistochemistry

Processing of formalin-fixed mouse tissues for hematoxylin and eosin staining, and cutting of unstained slides for immunohistochemistry, was performed by the Oxford Centre for Histopathology Research (OCHRe):

> Mouse tissue was fixed in neutral buffered formalin and processed to paraffin using standard techniques. Three-micron sections were cut and stained with hematoxylin and eosin.

Staining and analysis of histological slides for immunohistochemistry was performed by Elizabeth Soilleux:

> Serial sections were immunostained for CD3 using polyclonal rabbit serum raised against a synthetic peptide corresponding to a region showing 100% homology between mouse and human CD3ε (Jones et al., 1993), at the manufacturer's "ready-to-use" concentration (IR503; Dako). Immunostaining for TdT was performed using polyclonal rabbit anti-TdT (ab85148; Abcam) at a 1:200 dilution. Staining was performed on a Ventana Discovery Ultra Immunostainer (Ventana Medical Systems). Briefly, following deparaffinisation, sections were heated to 95ºC for 36 min in CC1 solution (a tris-based buffer with slightly alkaline pH) and incubated in primary antibody at 36ºC for 32 min. Bound rabbit antibody was detected using the Omnimap anti-rabbit DAB kit. All reagents apart from the primary antibody were obtained from Ventana Medical Systems and used according to the manufacturer's instructions. Immunostaining of WT mouse thymus was used as a positive control for both CD3 and TdT. Rabbit anti-FITC antibody (71-1900; Life Technologies) was used as a negative

control on representative histological sections. Stained sections were photographed with a Nikon DS-FI1 camera with a Nikon DS-L2 control unit and an Olympus BX40 microscope.

8: *In Vitro* Proliferation Assays

8.1: Granulocyte-monocyte assays

For granulocyte-monocyte assays, 75 ETPs were sorted into 1.5 ml of X-Vivo 15 (SLS Life Science) supplemented with 10% FCS (HyClone), 10^{-4} M 2-mercaptoethanol (Sigma-Aldrich), 1% penicillin/streptomycin (PAA Laboratories) and the following cytokines: 2 ng/ml murine stem cell factor (mSCF; PeproTech), 5 ng/ml human FLT3 ligand (hFL; Immunex), 5 ng/ml human thrombopoietin (hTPO; PeproTech), 5 ng/ml murine interleukin 3 (mIL-3; PeproTech), 10 ng/ml human granulocyte colony-stimulating factor (hG-CSF; Neopogen) and 10 ng/ml murine granulocyte-macrophage colony-stimulating factor (hGM-CSF; Immunex). Twenty microliters were transferred to each well of a 60-well Terasaki plate (VWR), seeding on average one cell per well by Poisson distribution. Density of cell growth in each well was evaluated after 8 days of culture.

8.2: Colony assays

For colony assays, single cells were sorted directly into 60-well Terasaki plates containing media as defined above and including either 0.05% DMSO or equivalent volumes of JQ1 (MedChem Express) or I-BET151 (Sigma-Aldrich) to final concentrations as indicated. Density of cell growth in each well was evaluated after 8 days of culture, with each well containing a colony derived from a single cell. Colonies of 50 cells or more were defined as highly proliferative. The proportion of wells containing highly proliferative colonies for one 60-well plate was used as a single replicate.

8.3: MTS assays

For MTS assays, 500 cells (ETPs), 20,000 cells (cell lines) or 250,000 cells (patient-derived xenografts) were cultured in 100 µl of media (see below) in 96-well plates for 72 hours (ETPs and cell lines) or 48

hours (xenografts) before measuring proliferation using the CellTiter 96 AQueous Non-Radioactive Cell Proliferation Assay (Promega) as directed. MTS solution and PMS solution were mixed in a ratio of 2 ml MTS : 100 µl PMS, and 20 µl of this mixture was added to each well. Cells were then incubated for 2 hours (ETPs and cell lines) or 4 hours (xenografts) before absorbance at 490 nm and 650 nm was measured in each well. For analysis, absorbance at 650 nm (background due to debris) was subtracted from absorbance at 490 nm (signal, proportional to number of viable cells) in each well. Average absorbance in media-only control wells on each plate was then subtracted from absorbance in sample wells. Absorbance in each sample well was then divided by average absorbance in DMSO vehicle control wells for the corresponding genotype, cell line or patient to generate normalised measurements of proliferation.

For ETPs, MTS assay media consisted of Opti-MEM (Gibco) supplemented with 10% FCS (HyClone), 10^{-4} M 2-mercaptoethanol (Sigma-Aldrich), 1% penicillin/streptomycin (PAA Laboratories) and the following cytokines: 10 ng/ml mSCF (PeproTech), 5 ng/ml hFL (Immunex), 10 ng/ml hTPO (PeproTech), 20 ng/ml mIL-3 (PeproTech), 20 ng/ml hG-CSF (Neopogen), 20 ng/ml mGM-CSF (Immunex) and 10 ng/ml human colony stimulating factor 1 (hCSF1; Cetus). For cell lines, media consisted of RPMI-1640 (Gibco) supplemented with 10% FCS (HyClone) and 1% penicillin/streptomycin (PAA Laboratories). For patient-derived xenografts, media consisted of RPMI-1640 (Gibco) supplemented with 10% FCS (HyClone), 1% penicillin/streptomycin (PAA Laboratories), 10 ng/ml hSCF (PeproTech), 10 ng/ml hFL (Immunex), 10 ng/ml hTPO (PeproTech), 10 ng/ml hG-CSF (Neopogen), 10 ng/ml hGM-CSF (Berlex), 10 ng/ml hIL-6 (PeproTech) and 10 ng/ml hIL-7 (PeproTech). Either 0.05% DMSO or equivalent volumes of JQ1 (MedChem Express) or I-BET151 (Sigma-Aldrich) were added to final concentrations as indicated.

Human cell lines were purchased from ATCC and cultured in RPMI-1640 (Gibco) supplemented with 10% FCS (HyClone) and 1% penicillin/streptomycin (PAA Laboratories).

Patient-derived xenograft cells were obtained from spleens of leukemic recipient mice (Maude et al., 2015). FACS analysis was used to confirm >85% human chimerism of $CD45^+$ cells before cells were used for MTS assays.

Clinical characteristics and genetic lesions of ETP leukaemia xenograft samples (Maude et al., 2015):

Sample no.	**Age (years)**	**Gender**	**WBC ($x10^9$/L)**	**Genetic Lesions**
ETP-ALL 01	8	M	66.1	CTCF SV, DNM2 K557_K558 > K, JAK3 M511I, WT1 R370fs
ETP-ALL 05	15	M	41.8	PTPN11 A72V
ETP-ALL 12	16	M	251	EZH2 S651L, *RUNX1* T148fs, NOTCH1 S2492*, *IKZF1* SV, PHF6 N147fs, SUZ12 C350R, WT1 C350R L564_S568, *SH2B3* SV
ETP-ALL 13	3	M	21.5	EED S259F

Informed consent for use of diagnostic specimens for future research was obtained from patients or their guardians according to the Declaration of Helsinki (Maude et al., 2015).

9: BET Inhibition *In Vivo*

9.1: JQ1 administration *in vivo*

For *in vivo* administration, a 100 mg/ml stock of JQ1 (MedChem Express) in DMSO was diluted 1/20 into a pre-filtered 10% w/v solution of 2-hydroxypropyl-β-cyclodextrin (Cambridge Bioscience) in water, and administered daily by intra-peritoneal injection at a dose of 45 mg/kg. Control mice received a similar volume of diluted DMSO. Each mouse received a maximum of 36 intra-peritoneal injections as specified by Home Office approvals. Mice were weighed on a daily basis to monitor weight loss resulting from treatment. Any mouse losing > 10% of its starting weight was removed

from treatment until its weight recovered to > 90% of starting weight. Any mouse losing > 20% of its starting weight was culled.

9.2: I-BET151 administration *in vivo*

Transplantation of DKOITD bone marrow cells for *in vivo* administration of I-BET151 was performed by George Giotopoulos:

> C57BL/6 CD45.1 recipient mice were sub-lethally irradiated (a single dose of 5.5 Gy) before intra-venous transplantation of approximately 1.8 million DKOITD bone marrow cells re-suspended in Hanks' Balanced Salt Solution (Sigma). Peripheral blood was collected from the saphenous vein into EDTA-coated tubes (Sarstedt). Analysis of haematological parameters was performed using a Vet abc counter (Scil Animal Care). These in vivo studies used a modified murine feed containing 0.018% (180ppm) I-BET151 (or control feed). All mice had ad libitum access to feed.

Chapter 3: Ezh2 and Runx1 Inactivating Mutations Collaborate to Induce Myelodysplastic Syndrome

1: Introduction

Cancer stem cells are the tumour-propagating population of a malignancy. These cells may represent only a small proportion of the overall tumour bulk but must be eliminated in order to cure the disease. Insights into the biology of cancer stem cells may therefore lead to novel therapeutic approaches (Clevers, 2011; Magee et al., 2012). One key aspect of cancer stem cell biology relates to the normal, healthy cells from which they derive. In AML and MDS it has become a well-established paradigm that the cancer stem cells responsible for disease initiation bear phenotypic similarities to normal HSCs (Shlush et al., 2014; Woll et al., 2014), strongly suggesting that these diseases originate from acquisition of mutations by HSCs. However, the extent to which progenitor cell populations, which unlike HSCs do not already possess intrinsic self-renewal, are able to become transformed into tumour-propagating cancer stem cells is less well understood (Magee et al., 2012). We therefore sought to address the question of whether the diverse malignant phenotypes found in human diseases could be in part determined by the characteristics of the normal haematopoietic populations from which they arise.

We decided to approach this by using conditional targeting of the same mutations to different haematopoietic stem and progenitor cell populations. It has been established that in both MDS (Papaemmanuil et al., 2013) and ETP-ALL (Zhang et al., 2012b), inactivating mutations in *EZH2* and *RUNX1* are significantly associated; that is mutations in these two genes co-occur in the same patient more frequently than would be expected by chance. MDS has been shown to derive from sequential acquisition of mutations within the $CD34^{+}CD38^{-}$ stem/progenitor compartment of human bone marrow (Woll et al., 2014). Conversely, ETP-ALL is hypothesised based on its gene expression profile to originate within ETPs, the most primitive thymocyte progenitor population (Coustan-Smith

et al., 2009). These observations led us to hypothesise that co-mutation of *EZH2* and *RUNX1* can lead to different malignant phenotypes depending on the haematopoietic stem or progenitor populations in which their inactivation occurs.

In order to explore this possibility, we used conditional targeting of inactivation of *Ezh2* and *Runx1* within the haematopoietic system of mice. We utilised three separate *Cre* lines to target their deletion to either the entire haematopoietic system, including HSCs, of adult mice (*Mx1Cre*) (Kuhn et al., 1995), MPPs and all their progeny but excluding HSCs (*Flt3Cre*) (Boyer et al., 2011; Buza-Vidas et al., 2011), or early lymphoid progenitors (*Rag1Cre*) (McCormack et al., 2003).

Individual inactivation of either *Ezh2* or *Runx1* within HSCs has been previously shown to induce profound phenotypic changes to the haematopoietic system, including myeloproliferation and blocks in B-cell and T-cell lymphopoiesis and megakaryopoiesis (Growney et al., 2005; Ichikawa et al., 2004; Su et al., 2003; Su et al., 2005). We therefore hypothesised that co-inactivation of both genes in HSCs would lead to further, possibly malignant, phenotypic changes reflecting those seen in MDS patients. These could include acquisition of a competitive advantage by *Ezh2* and *Runx1*-deleted haematopoietic cells over their non-deleted counterparts, and development of peripheral blood cytopaenias resulting from decreased output of bone marrow progenitor cells.

Experimental evidence has been gained from patient bone marrow samples that MDS originates from and is exclusively propagated by malignant counterparts of normal HSCs (Woll et al., 2014). This led us to hypothesise that targeting of *Ezh2* and *Runx1* inactivation to non-HSC progenitor populations such as MPPs would not lead to development of MDS. We also hypothesised that initiation of an ETP-ALL phenotype would require targeting of inactivation to early lymphoid progenitors.

2: Results

2.1: Inactivation of *Ezh2* and *Runx1* in adult HSCs strongly impacts haematopoiesis

We first used *Mx1Cre* to inactivate *Ezh2* and/or *Runx1* in all haematopoietic cells, including HSCs (Kuhn et al., 1995), at 6-8 weeks of age (Figure 3.01A). To achieve this we generated four genotypes of mice for Poly(IC)-treatment: $Ezh2^{fl/fl}Runx1^{fl/fl}Mx1Cre^{+/+}$ (Mx1-WT), $Ezh2^{fl/fl}Runx1^{+/+}Mx1Cre^{tg/+}$ (Mx1-Ezh2KO), $Ezh2^{+/+}Runx1^{fl/fl}Mx1Cre^{tg/+}$ (Mx1-Runx1KO), $Ezh2^{fl/fl}Runx1^{fl/fl}Mx1Cre^{tg/+}$ (Mx1-DKO). We confirmed high efficiency of deletion of *Ezh2* and *Runx1* in HSCs and MPPs using qPCR primers designed to specifically amplify mRNA transcribed from non-deleted alleles (Figure 3.01B).

We performed phenotypic analysis of these mice 6 weeks post-activation of *Mx1Cre*. Analysis of haematological parameters of the peripheral blood showed that knockout of either *Ezh2* or *Runx1* resulted in reduced white blood cell counts and mild anaemia (Figure 3.02A). In addition, Mx1-Runx1KO but not Mx1-Ezh2KO mice showed a marked reduction in platelet numbers (Figure 3.02A). These results are consistent with the previously documented block in megakaryocyte maturation seen in Runx1 knockout mice, resulting in thrombocytopenia (Ichikawa et al., 2004). Knockout of either *Ezh2* or *Runx1* in adult HSCs causes lymphocytopenia due to blocks in the maturation of both B-cells and T-cells (Su et al., 2003; Su et al., 2005). Mild anaemia is also observed in these mice. Double knockout of both *Ezh2* and *Runx1* resulted in a reduction in white blood cell and platelet counts similar to those seen in Mx1-Runx1KO mice (Figure 3.02A). The anaemia in these mice was more profound than in either of the single knockouts, with a number of mice developing quite severe anaemia at this timepoint (Figure 3.02A). Thymus cellularity was markedly reduced in all three experimental genotypes (Figure 3.02B), consistent with impaired T-cell maturation previously demonstrated following loss of either *Ezh2* or *Runx1* (Ichikawa et al., 2004; Su et al., 2005).

We next performed lineage analysis of white blood cells in the peripheral blood (Figure 3.02A) and bone marrow (Figure 3.02B). Numbers of $CD19^+$ B-cells were markedly reduced in the peripheral blood and bone marrow of all three knockout genotypes (Figure 3.03A-B). Surprisingly, $CD4^+/CD8^+$ T-cell numbers were not reduced, probably reflecting the long lifespan of mature T-cells *in vivo* (Figure

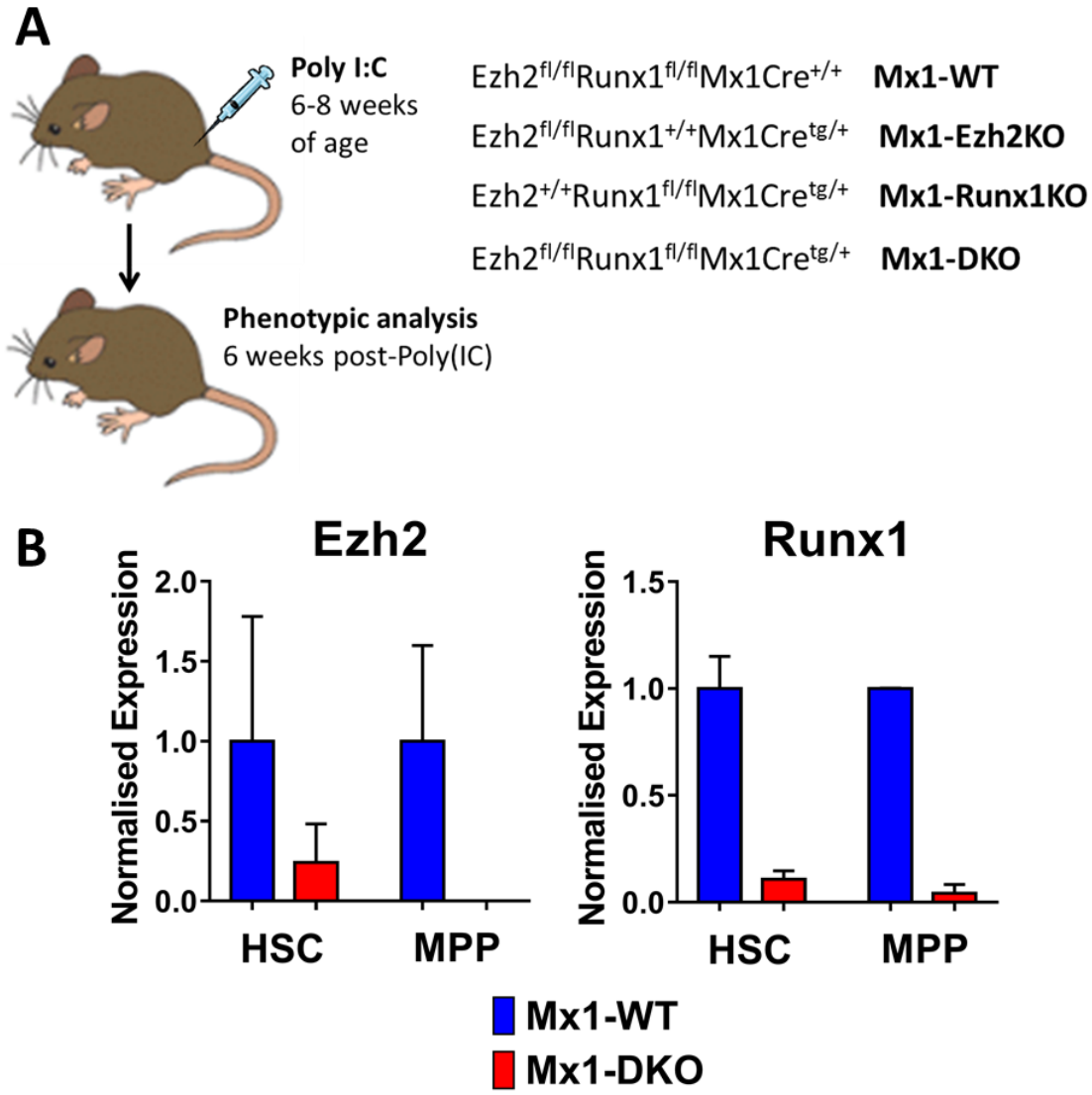


Figure 3.01

(A) Experimental setup and genotypes of experimental mice for Ezh2 and/or Runx1 inactivation in HSCs of adult mice using Mx1Cre. (B) Determination of efficiency of deletion by Mx1Cre of conditionally targeted Ezh2 and Runx1 alleles, in $Lin^{-}Kit^{+}Sca1^{+}CD150^{+}CD48^{-}$ HSCs and $Lin^{-}Kit^{+}Sca1^{+}CD150^{-}CD48^{+}$ MPPs from 6-8 week old Mx1-WT and Mx1-DKO mice. Relative expression levels of Ezh2 and Runx1 mRNA are shown; Ct values were normalised to Hprt. Amplicons were designed to target exons removed by Cre recombination. Mx1-WT n=2, Mx1-DKO n=2; each replicate represents a sample of 50 cells. Error bars represent standard error of the mean. Deletion efficiency experiments were performed by Wen Hao Neo.

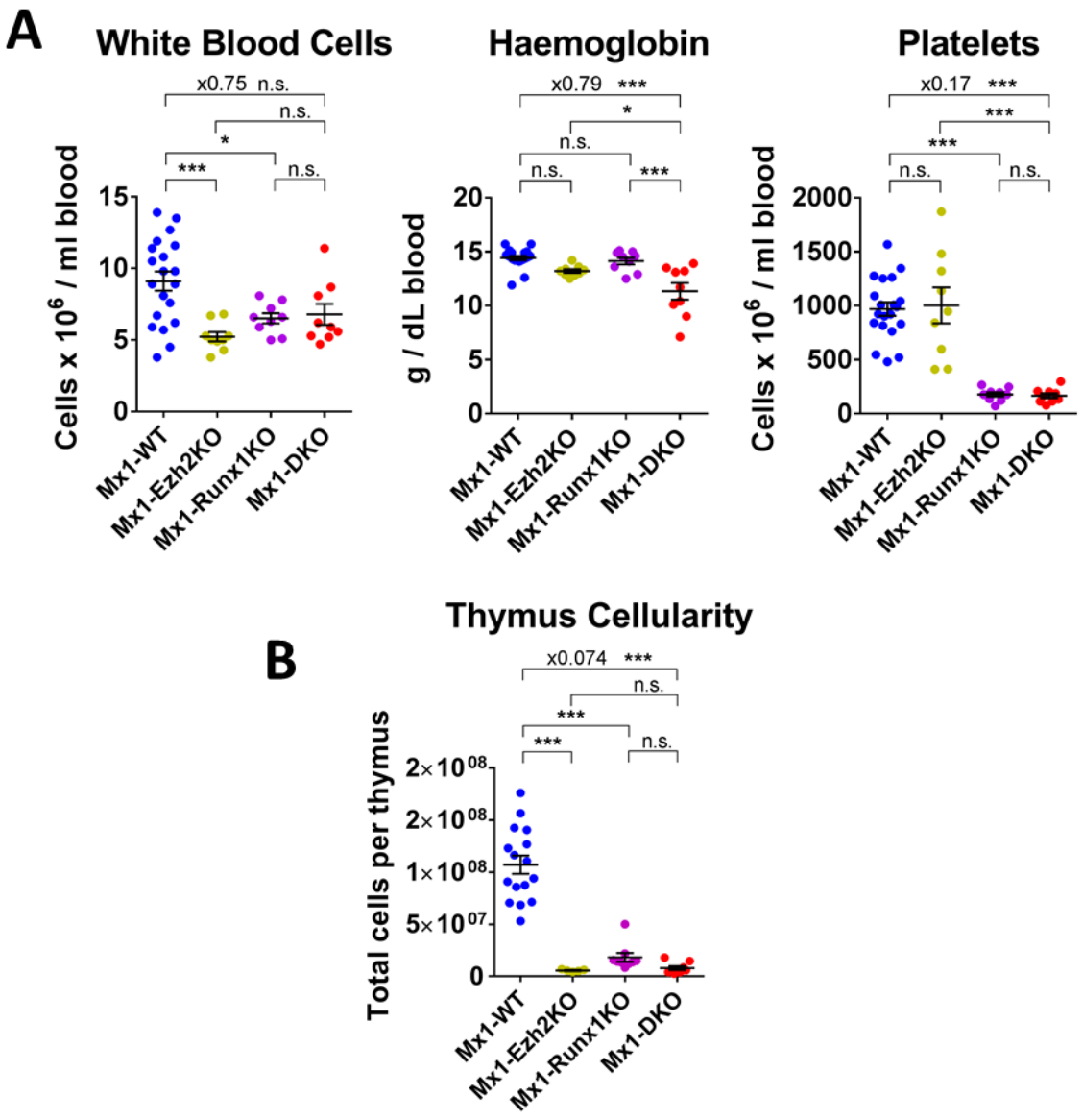


Figure 3.02

(A) Haematological parameters of peripheral blood 6 weeks post-Poly(IC) in Mx1-WT (n=20), Mx1-Ezh2KO (n=9), Mx1-Runx1KO (n=9) and Mx1-DKO (n=8) mice. (B) Total thymus cellularity 6 weeks post-Poly(IC) in Mx1-WT (n=16), Mx1-Ezh2KO (n=5), Mx1-Runx1KO (n=9) and Mx1-DKO (n=8) mice. Error bars represent standard error of the mean. P-values are by ANOVA followed by multiple comparisons testing; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

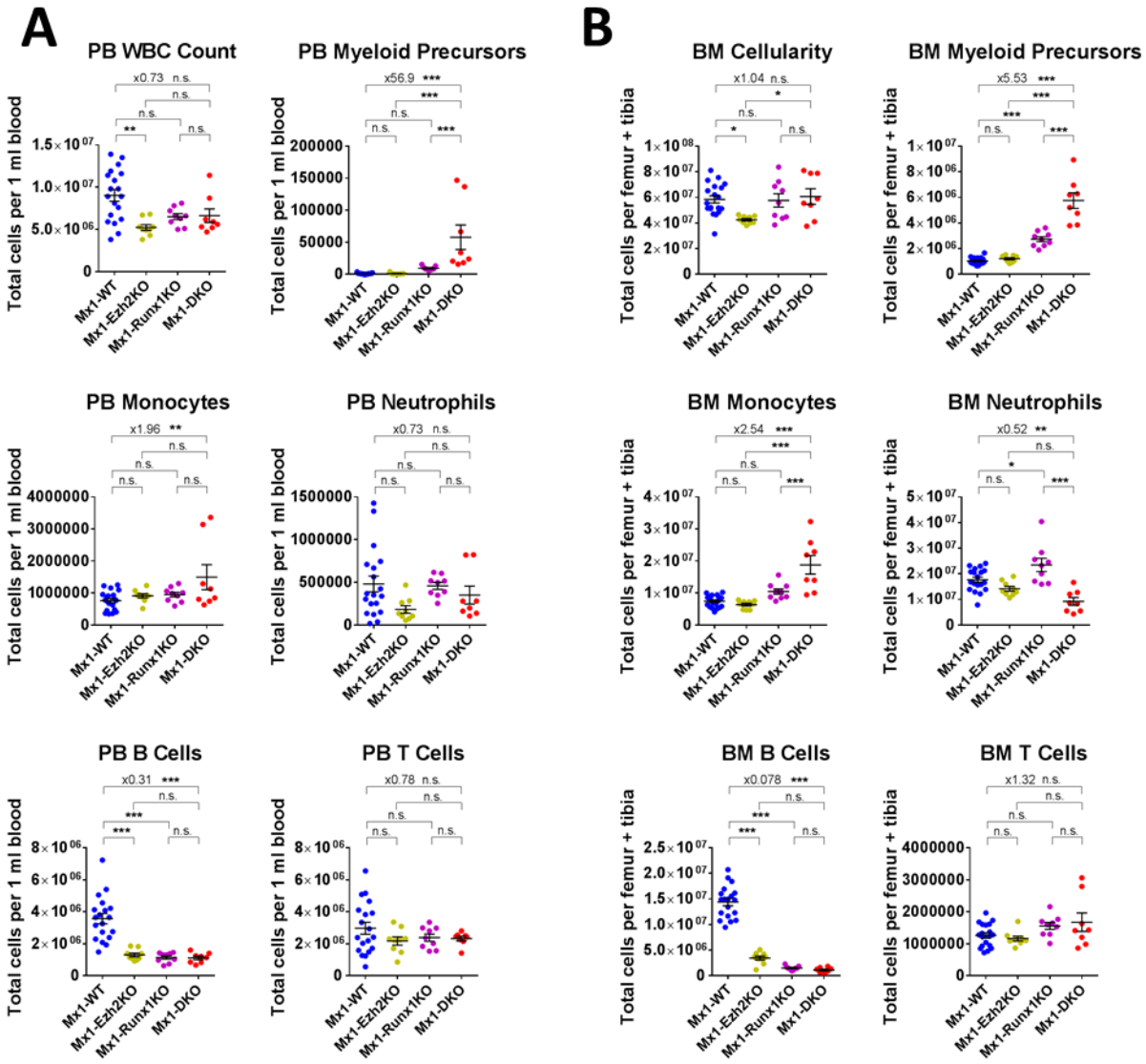


Figure 3.03

(A) Absolute numbers per ml of indicated leukocyte populations in the peripheral blood 6 weeks post-Poly(IC) in Mx1-WT (n=19), Mx1-Ezh2KO (n=9), Mx1-Runx1KO (n=9) and Mx1-DKO (n=8) mice. (B) Absolute numbers per femur and tibia of indicated bone marrow populations 6 weeks post-Poly(IC) in Mx1-WT (n=19), Mx1-Ezh2KO (n=9), Mx1-Runx1KO (n=9) and Mx1-DKO (n=8) mice. B-cells $CD19^{+}$, T-cells $CD4^{+}$ or $CD8^{+}$, monocytes $Mac1^{+}Gr1^{-}$, neutrophils $Mac1^{+}Gr1^{+}$, myeloid precursors $Kit^{+}Mac1^{low}$. Error bars represent standard error of the mean. P-values are by ANOVA followed by multiple comparisons testing; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

3.03A-B). Numbers of $Mac1^{+}Gr1^{+}$ neutrophils in the peripheral blood were not significantly affected in any knockout genotype (Figure 3.03A). In the bone marrow, Mx1-Runx1KO neutrophils were slightly increased, while neutrophils in Mx1-DKO mice were conversely decreased (Figure 3.03B). Peripheral blood and bone marrow $Mac1^{+}Gr1^{-}$ monocytes were increased only in Mx1-DKO mice (Figure 3.03A-B). Numbers of immature $Kit^{+}Mac1^{low}$ myeloid precursor cells (Bereshchenko et al., 2009) were slightly increased in the bone marrow of Mx1-Runx1KO mice, and much more strongly increased in Mx1-DKO mice (Figure 3.03B). These cells were also very markedly expanded in the peripheral blood of Mx1-DKO mice (Figure 3.03A).

We next analysed the $Lineage^{-}Kit^{+}Sca1^{+}$ haematopoietic stem and progenitor compartment in these mice (Figure 3.04). We observed a marked increase in total numbers of LSK cells in Mx1-Runx1KO and Mx1-DKO mice, but not Mx1-Ezh2KO. This was primarily caused by an expansion of LSK $CD48^{+}CD150^{-}$ MPPs. This expansion of $CD48^{+}$ cells within the LSK compartment of Mx1-Runx1KO mice has been previously documented, where it was shown that numbers of functional HSCs were not affected (Cai et al., 2011). In keeping with this, numbers of LSK $CD48^{-}CD150^{+}$ HSCs in Mx1-Runx1KO mice were normal. In Mx1-Ezh2KO mice, total numbers of LSK cells were unaffected compared to wild-type (Figure 3.04). However, there was a marked reduction in numbers of HSCs. This reduction in HSCs was even more severe in Mx1-DKO mice. Numbers of LSK $Flt3^{high}$ LMPPs were not significantly affected in any of the knockout genotypes (Figure 3.04).

Together, these data demonstrate that targeted inactivation of *Ezh2* and *Runx1* in adult HSCs has profound consequences on the haematopoietic system. The overall phenotype of Mx1-DKO mice combines features of Mx1-Ezh2KO and Mx1-Runx1KO mice, but also displays a number of unique changes. These were a more profound anaemia than either single knockout, an increase in bone marrow myeloid precursors and monocytes combined with a decrease in neutrophils, and a very marked reduction in HSCs. These results are indicative of near-complete blocks in lymphopoiesis and megakaryopoiesis, combined with reduced differentiation in the myeloid lineage and an

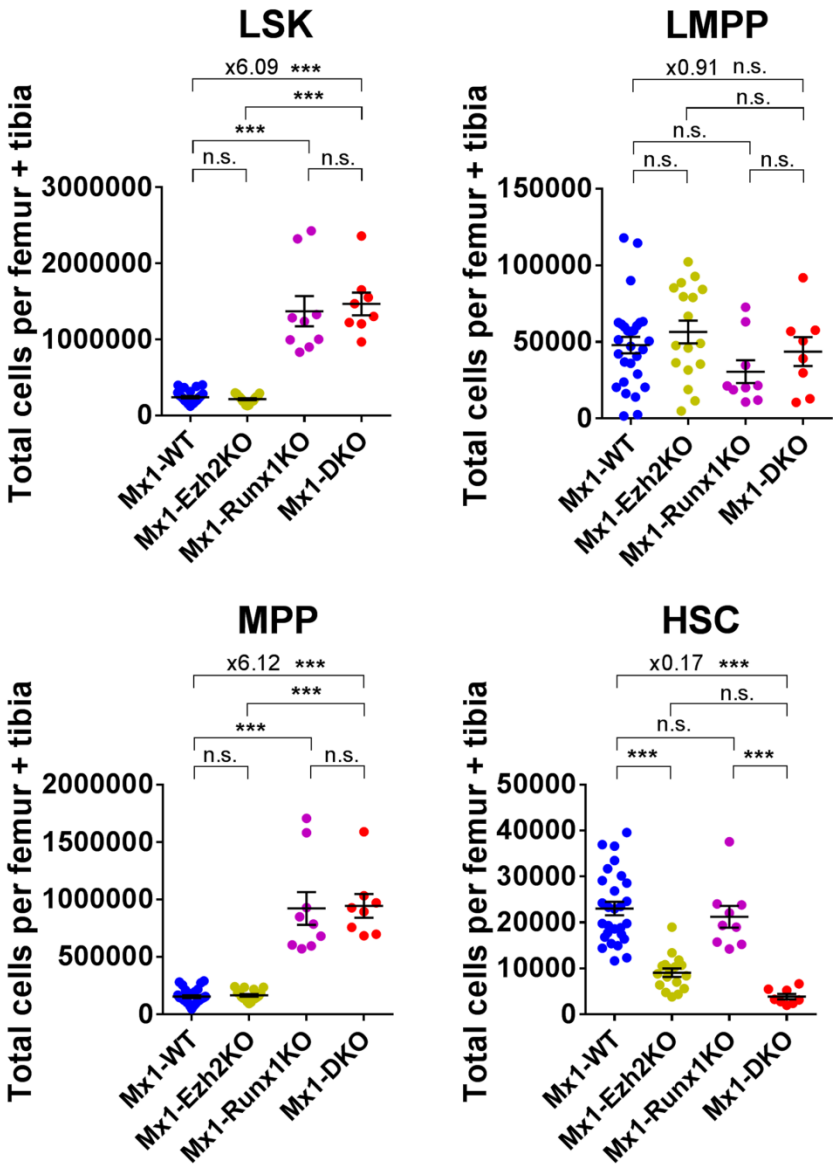


Figure 3.04

Absolute numbers per femur and tibia of indicated bone marrow stem and progenitor populations 6 weeks post-Poly(IC) in Mx1-WT (n=28), Mx1-Ezh2KO (n=17), Mx1-Runx1KO (n=9) and Mx1-DKO (n=8) mice. LSK Lineage$^-$Sca1$^+$Kit$^+$, LMPP Lineage$^-$Sca1$^+$Kit$^+$Flt3^{high}, MPP Lineage$^-$Sca1$^+$Kit$^+$CD48$^+$CD150$^-$, HSC Lineage$^-$Sca1$^+$Kit$^+$CD48$^-$CD150$^+$. Error bars represent standard error of the mean. P-values are by ANOVA followed by multiple comparisons testing; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

accumulation of myeloid precursors. Mx1-Runx1KO mice show an expansion of MPPs which is also present in Mx1-DKO mice, while Mx1-Ezh2KO mice show a reduction of HSCs which are further reduced in Mx1-DKO mice. These data could be consistent with development of an MDS phenotype.

2.2: Inactivation of *Ezh2* and *Runx1* in HSCs leads to development of myelodysplastic syndrome

To explore whether this was the case, we decided to investigate the long-term effects of *Ezh2* and *Runx1* inactivation in the haematopoietic system. To do this we followed up Mx1-WT, Mx1-Ezh2KO, Mx1-Runx1KO and Mx1-DKO mice to 52 weeks after Poly(IC) treatment at 6-8 weeks of age (Figure 3.05A). For disease-free survival analysis, mice that were either found dead without analysis or culled showing greater than 25 million leukocytes per millilitre of blood and/or less than 6.5 grams of haemoglobin per decilitre of blood were counted as events. Mice that were culled showing peripheral blood haematological parameters outside these constraints were censored. We chose to censor these mice because the immune suppression caused by loss of mature B and T cells may cause an increased susceptibility to infection. Setting a cut-off in the haemoglobin count allowed us to select for mice becoming sick as a result of haematological phenotypes.

No decrease in survival was observed in Mx1-WT or Mx1-Runx1KO mice (Figure 3.05A). Mx1-Ezh2KO mice meanwhile showed reduced survival (median latency 46.1 weeks; Figure 3.05A). Development of $CD3^+$ T-ALL has previously been reported in $Ezh2^{fl/fl}Mx1Cre^{tg/+}$ mice (Simon et al., 2012), and we confirmed onset of leukaemia as the cause of death in some of our Mx1-Ezh2KO mice. However, peripheral blood leukocytes in some of these mice also showed a $CD19^+Mac1^{low}$ phenotype, indicative of a biphenotypic B-cell/myeloid acute leukaemia, not previously described in this model (Figure 3.05B).

In contrast, Mx1-DKO mice displayed a markedly different phenotype and reduced survival compared to Mx1-Ezh2KO mice (median latency 32.3 weeks; Figure 3.05A). With one exception,

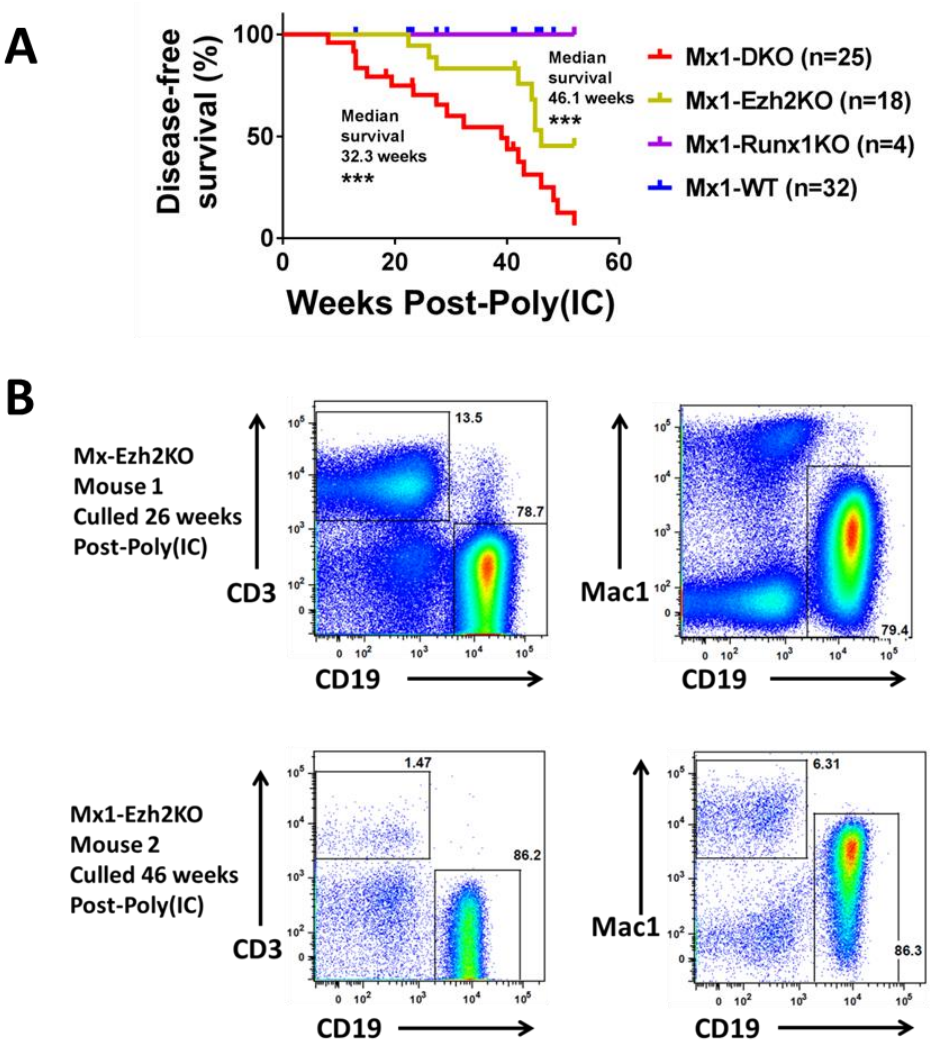


Figure 3.05

(A) Kaplan-Meier disease-free survival curves for Mx1-WT, Mx1-Ezh2KO, Mx1-Runx1KO and Mx1-DKO mice following Poly(IC) treatment at 6-8 weeks of age. Mice found dead without analysis or culled showing > 25 x 10^6 leukocytes per ml blood and/or < 6.5 g haemoglobin per dl blood were counted as events. Mice culled showing leukocyte and haemoglobin counts outside these constraints were censored. (B) FACS plots showing peripheral blood leukocyte staining for CD3, CD19 and Mac1 in two Mx1-Ezh2KO mice culled showing 104.5 x 10^6 (Mouse 1) and 41.2 x 10^6 (Mouse 2) leukocytes per ml blood. Figures indicate percentage of total viable ($7AAD^-$) cells falling into the respective gates. P-values are by logrank test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

these mice did not develop leukaemia and instead displayed pancytopenia in the peripheral blood at the time of culling (Figure 3.06A). Erythrocytes and leukocytes in the peripheral blood showed irregular morphology, and an accumulation of undifferentiated myeloid precursor cells could be observed in the bone marrow (Figure 3.06B). Quantitative morphological analysis of bone marrow cells revealed an increase in immature myeloid precursor cells and granulocytes, and a decrease in erythroid precursors (Figure 3.06C). Flow cytometry analysis of the bone marrow (Figure 3.07A) confirmed the increase in $Kit^{+}Mac1^{low}$ myeloid precursors (though non-significant) and decrease in $Ter119^{+}$ erythroid precursors. $Mac1^{+}Gr1^{-}$ monocytes were increased, although $Mac1^{+}Gr1^{+}$ neutrophils were non-significantly reduced. $CD19^{+}$ B-cells were reduced as expected. Together, these data demonstrate that inactivation of *Ezh2* and *Runx1* in HSCs induces an MDS-like phenotype, with high penetrance and shorter latency than previously described MDS mouse models (Abdel-Wahab et al., 2013; Lin et al., 2005; Sashida et al., 2014; Watanabe-Okochi et al., 2008).

The one Mx1-DKO mouse which developed leukaemia was culled at 42 weeks post-Poly(IC). Analysis of the peripheral blood revealed a similar $CD19^{+}Mac1^{low}$ leukocyte phenotype to that seen in Mx1-Ezh2KO mice (Figure 3.07B).

2.3: *Ezh2* and *Runx1*-inactivated haematopoiesis displays a competitive advantage over normal haematopoiesis, combined with an extrinsic suppression of wild-type HSCs

A key feature of MDS is that the malignant HSC clone displays a competitive advantage over normal HSC clones. To explore whether *Ezh2* and *Runx1*-inactivated haematopoiesis could show a similar competitive advantage over wild-type haematopoiesis, we transplanted equal numbers of wild-type CD45.1 bone marrow cells and either Mx1-WT or Mx1-DKO CD45.2 cells, before Poly(IC) treatment, into lethally irradiated CD45.1 recipient mice. Four weeks following transplantation we analysed the relative contribution of CD45.1 and CD45.2 cells to peripheral blood cell lineages, and then treated

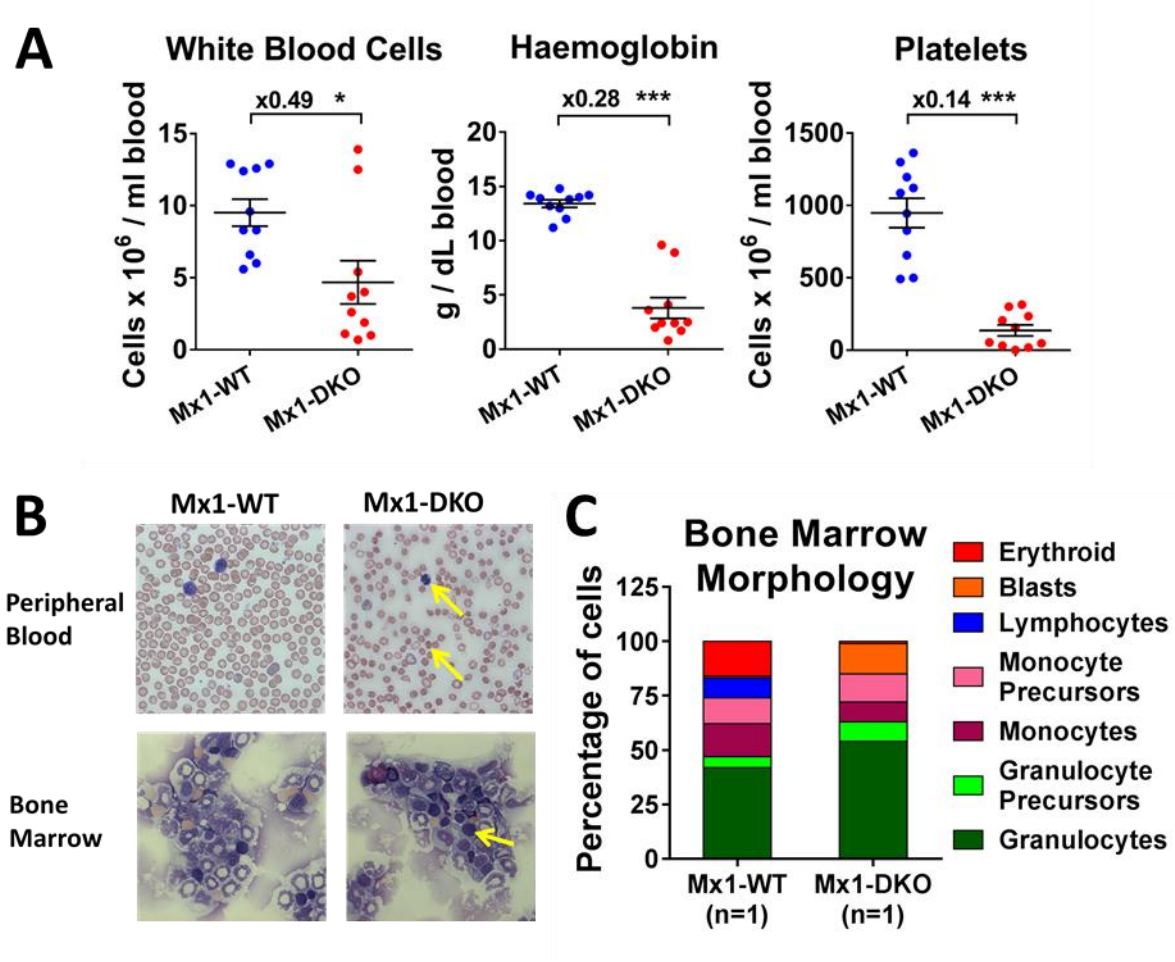


Figure 3.06

(A) Haematological parameters of follow-up Mx1-WT (n=10) and Mx1-DKO (n=10) mice at time of culling following Poly(IC) treatment at 6-8 weeks of age. One Mx1-DKO mouse showing 67.1 x10^6 leukocytes per ml blood was excluded (see Figure 3.07B). (B) Representative images of peripheral blood smears and bone marrow cytospins a sick Mx1-DKO mouse and Mx1-WT control. Arrows indicate irregular morphology of erythrocytes and leukocytes in peripheral blood, and accumulation of undifferentiated myeloid precursor cells in bone marrow. (C) Morphological analysis of bone marrow cells from a sick Mx1-DKO mouse and an Mx1-WT littermate control at time of culling. One hundred cells were examined from each mouse (performed by Adam Mead). Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

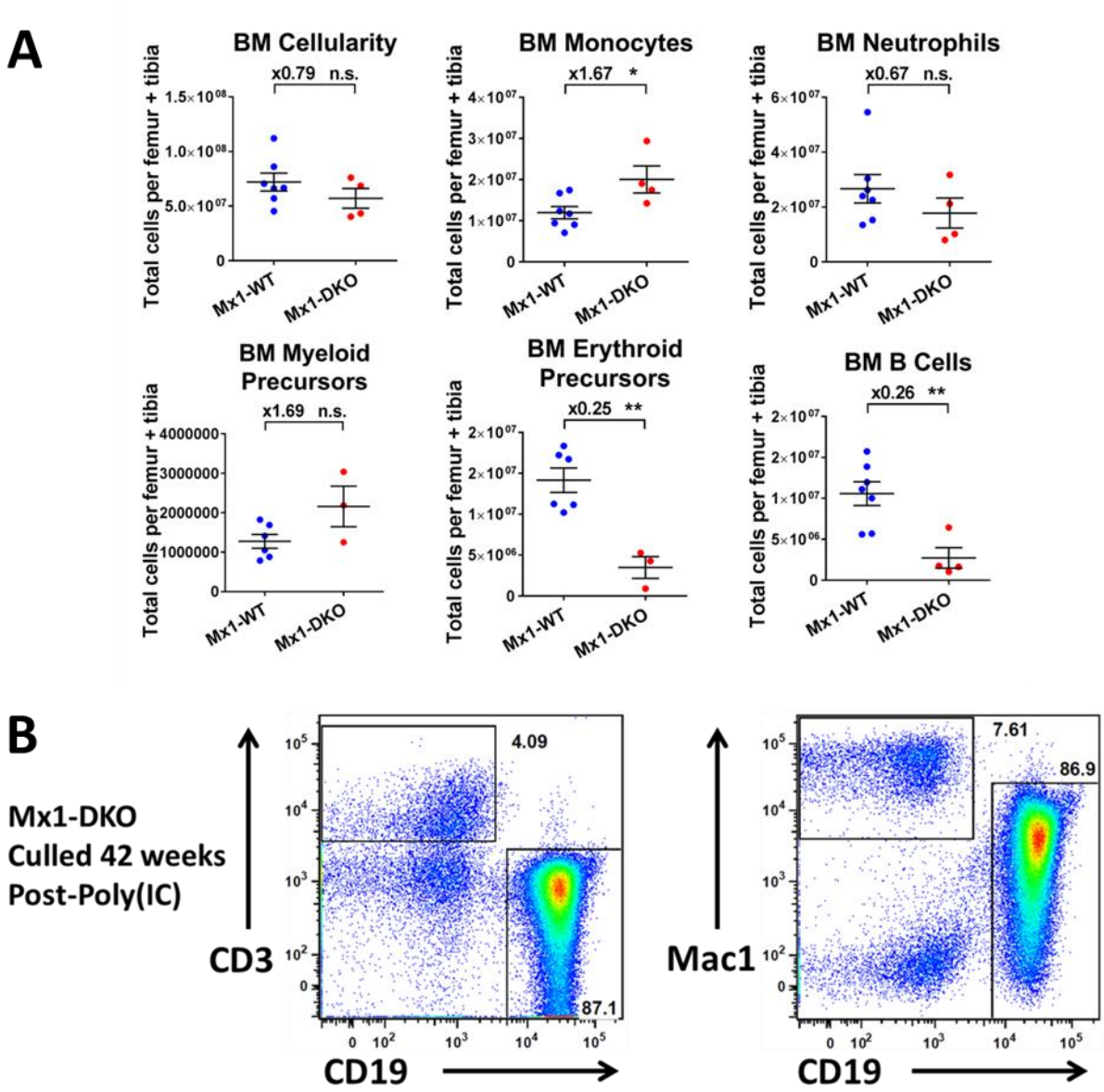


Figure 3.07

(A) Absolute numbers per femur and tibia of indicated bone marrow populations at time of culling in follow-up Mx1-WT (n=7) mice and Mx1-DKO (n=4) mice. Monocytes $Mac1^{+}Gr1^{-}$, neutrophils $Mac1^{+}Gr1^{+}$, myeloid precursors $Kit^{+}Mac1^{low}$, erythroid precursors $Ter119^{+}$, B-cells $CD19^{+}$. (B) FACS plots showing peripheral blood leukocyte staining for CD3, CD19 and Mac1 in one Mx1-DKO mouse culled showing 67.1 leukocytes per ml blood. Figures indicate percentage of total viable ($7AAD^{-}$) cells falling into the respective gates. Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

recipients with Poly(IC) to activate *Mx1Cre*. We then performed further analysis of the peripheral blood every four weeks (Figure 3.08A).

We found that Mx1-DKO bone marrow cells showed an increased contribution to peripheral blood myeloid cells (monocytes and neutrophils) over time relative to Mx1-WT cells (Figure 3.08B). Since mature myeloid cells are short-lived, these data indicate a competitive advantage of *Ezh2* and *Runx1*-inactivated over wild-type haematopoietic cells.

To explore the mechanisms underlying this competitive advantage, 16 weeks after Poly(IC)-treatment we culled the mice and analysed the relative contribution of CD45.2 cells to the bone marrow (Figure 3.09A), peripheral blood (Figure 3.09B) and spleen (Figure 3.09C). Overall contribution of CD45.2 cells to the bone marrow was high in both genotypes, but slightly higher in Mx1-DKO recipient mice compared to Mx1-WT recipient mice. Monocyte numbers were increased in all tissues, while neutrophils were not significantly affected. The fact that Mx1-DKO cells showed an increased overall contribution to the bone marrow compared to Mx1-WT cells supports our hypothesis that *Ezh2* and *Runx1* inactivation confers a competitive advantage to DKO bone marrow cells in a transplantation setting.

We also analysed $Kit^{+}Mac1^{low}$ myeloid precursor cells and found that, as in primary mice (Figure 3.02B), total numbers of these cells were markedly increased in the bone marrow of Mx1-DKO compared to Mx1-WT recipient mice (Figure 3.09A), and were almost entirely CD45.2. These cells were mainly absent in the peripheral blood and spleen of Mx1-WT recipient mice, but were present in substantial numbers in Mx1-DKO recipient mice and also mainly CD45.2 (Figures 3.09B and 3.09C).

Analysis of T-cells showed that, while total numbers were comparable in the peripheral blood (Figure 3.09B) and spleen (Figure 3.09C) of Mx1-DKO compared to Mx1-WT recipient mice, they were significantly reduced in the bone marrow (Figure 3.09A). These results are consistent with a reduced potential of *Ezh2* and *Runx1*-inactivated HSCs to produce mature T-cells.

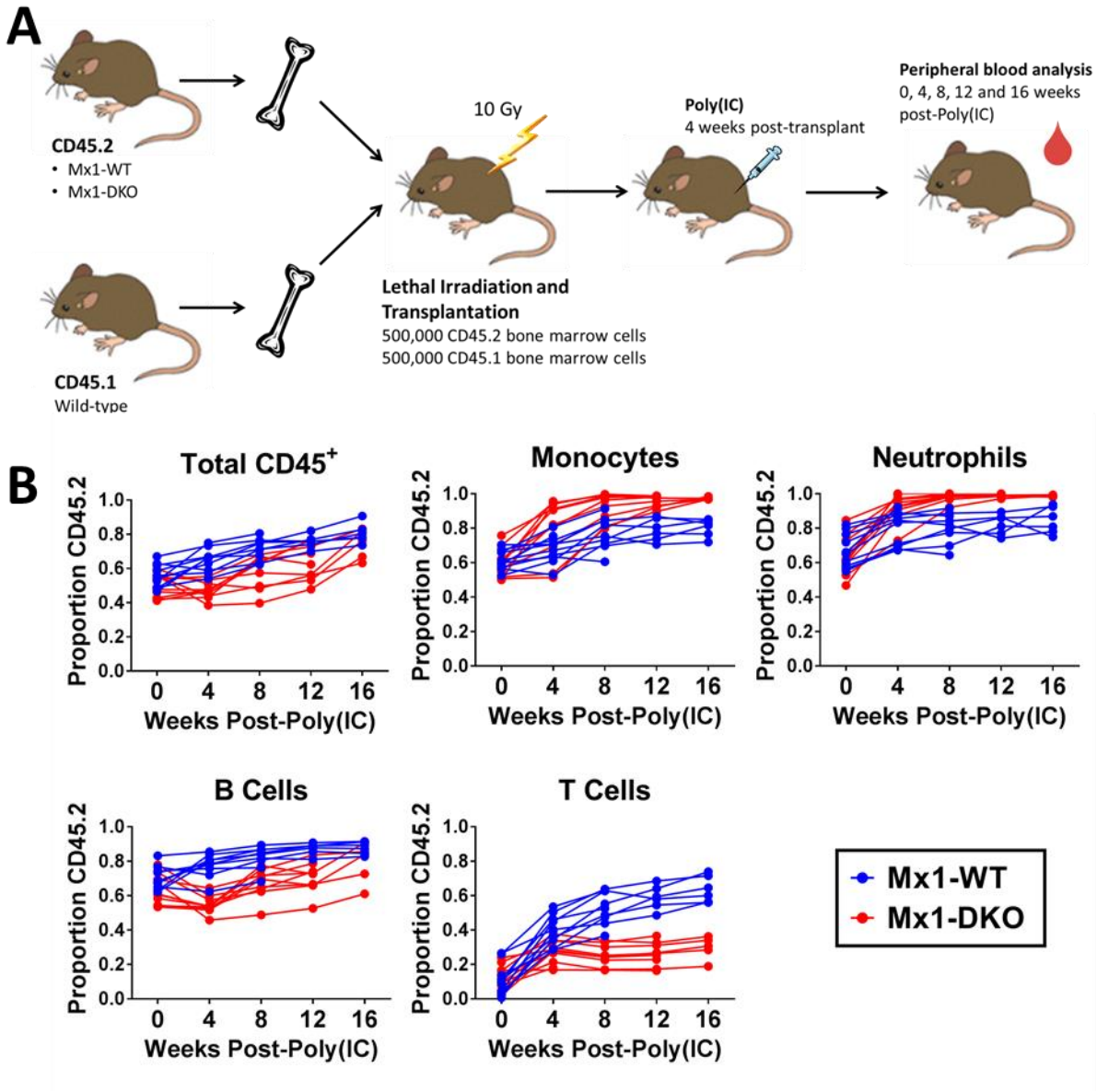


Figure 3.08

(A) Experimental setup for transplantation of equal numbers of CD45.1 wild-type and CD45.2 Mx1-WT or Mx1-DKO bone marrow cells followed by inactivation of Ezh2 and Runx1. (B) Relative contribution of CD45.2 cells to indicated leukocyte populations in the peripheral blood of wild-type CD45.1 recipient mice following transplantation of CD45.2 Mx1-WT or Mx1-DKO bone marrow cells and activation of Mx1Cre using Poly(IC). Each line represents an individual recipient mouse; Mx1-WT n=10, Mx1-DKO n=10. Monocytes $Mac1^{+}Gr1^{-}$, neutrophils $Mac1^{+}Gr1^{+}$, B-cells $CD19^{+}$, T-cells $CD4^{+}$ or $CD8^{+}$.

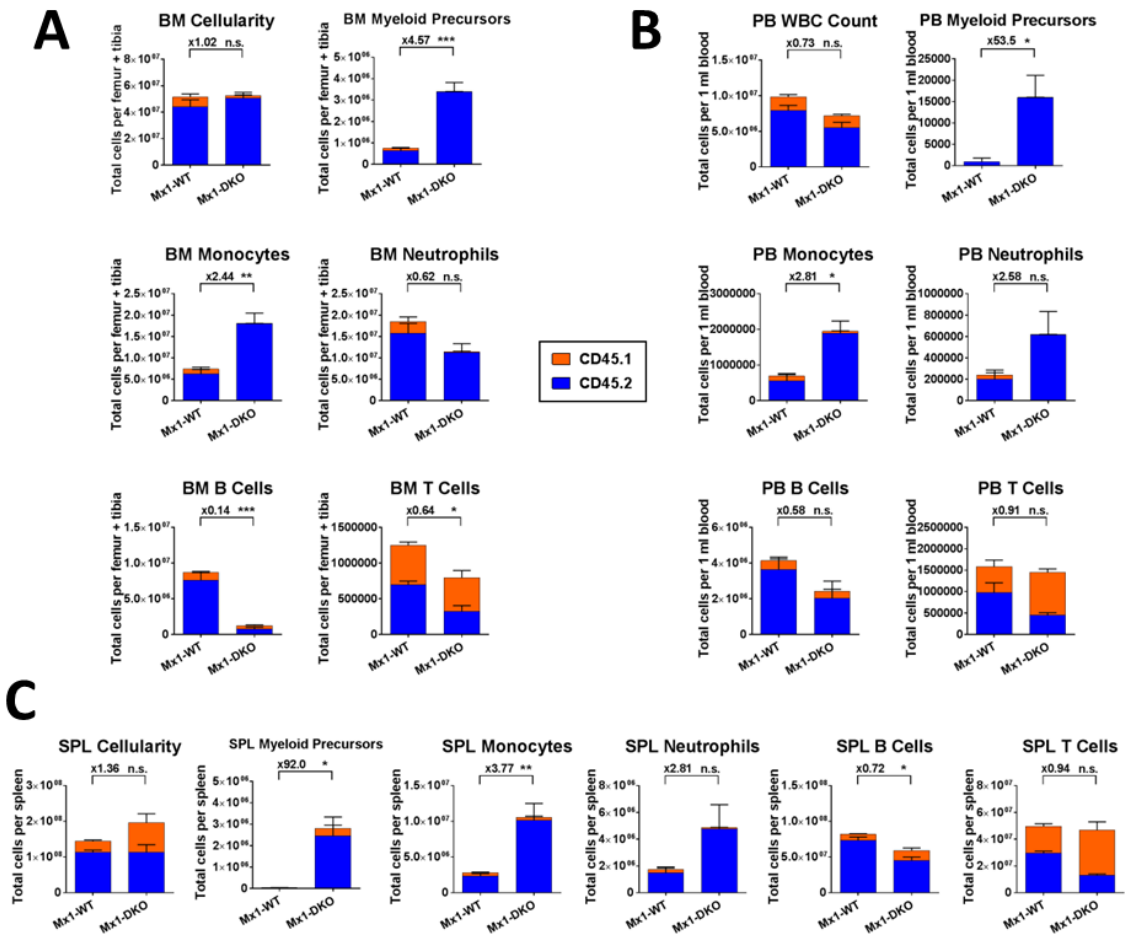


Figure 3.09

(A) Absolute numbers per femur and tibia of CD45.1 and CD45.2 cells contributing to indicated bone marrow populations 16 weeks post-Poly(IC) treatment in Mx1-WT (n=6) and Mx1-DKO (n=7) bone marrow recipient mice. (B) Absolute numbers per ml of CD45.1 and CD45.2 cells contributing to indicated leukocyte populations in the peripheral blood 16 weeks post-Poly(IC) treatment in Mx1-WT (n=6) and Mx1-DKO (n=6) bone marrow recipient mice. (C) Absolute numbers per spleen of CD45.1 and CD45.2 cells contributing to indicated splenocyte populations 16 weeks post-Poly(IC) treatment in Mx1-WT (n=6) and Mx1-DKO (n=5) bone marrow recipient mice. B-cells $CD19^{+}$, T-cells $CD4^{+}$ or $CD8^{+}$, monocytes $Mac1^{+}Gr1^{-}$, neutrophils $Mac1^{+}Gr1^{+}$, myeloid precursors $Kit^{+}Mac1^{low}$. Error bars represent standard error of the mean. P-values are by T-test comparing summed CD45.1 and CD45.2 cells; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

In the peripheral blood and spleen we found that total numbers of B-cells were non-significantly reduced in Mx1-DKO compared to Mx1-WT recipient mice (Figure 3.09B-C). Analysis of the bone marrow, however, revealed a marked reduction in total numbers of B-cells in Mx1-DKO recipient mice, mainly attributable to reduced CD45.2 cells (Figure 3.09A). This suggests that production of B-cells is indeed severely impaired in *Ezh2* and *Runx1*-inactivated bone marrow, and that the more normal numbers of mature CD45.2 B-cells present in the peripheral blood and spleen of Mx1-DKO mice were persisting cells from before Poly(IC) treatment, although this was not formally confirmed by deletion analysis.

We then analysed contribution of CD45.1 and CD45.2 cells to the bone marrow LSK stem and progenitor compartment of recipient mice (Figure 3.10A). We found that, as in primary mice (Figure 3.04), total numbers of LSK cells were markedly increased in Mx1-DKO compared to Mx1-WT recipient mice, and also that these cells were almost entirely CD45.2 (Figure 3.10A). This was also the case for MPPs and LMPPs, suggesting that this LSK expansion is mainly due to proliferation of *Ezh2* and *Runx1*-inactivated multipotent progenitor populations. In contrast, total numbers of HSCs were strongly reduced in Mx1-DKO compared to Mx1-WT recipient mice (Figure 3.10A), even more extensively than in primary mice (Figure 3.04). Interestingly, in Mx1-WT recipient mice close to 50% of HSCs were CD45.2 (Figure 3.10A), as would be expected given that we transplanted equal numbers of CD45.1 and CD45.2 bone marrow cells. It is not clear why most mature populations showed higher contribution of CD45.2 relative to CD45.1 cells even in the Mx1-WT control mice (Figure 3.09). In Mx1-DKO recipient mice, while total HSC numbers were reduced, almost all remaining HSCs were CD45.1 (Figure 3.10A). Furthermore, numbers of CD45.1 HSCs were reduced in Mx1-DKO compared to Mx1-WT recipient mice, indicating an extrinsic suppression of wild-type HSCs by *Ezh2* and *Runx1*-inactivated bone marrow cells.

We also analysed LSK cells in the spleen of recipient mice. Total LSK cells, LMPPs and MPPs were largely absent in Mx1-WT recipient spleens, but were present in substantial numbers in Mx1-DKO

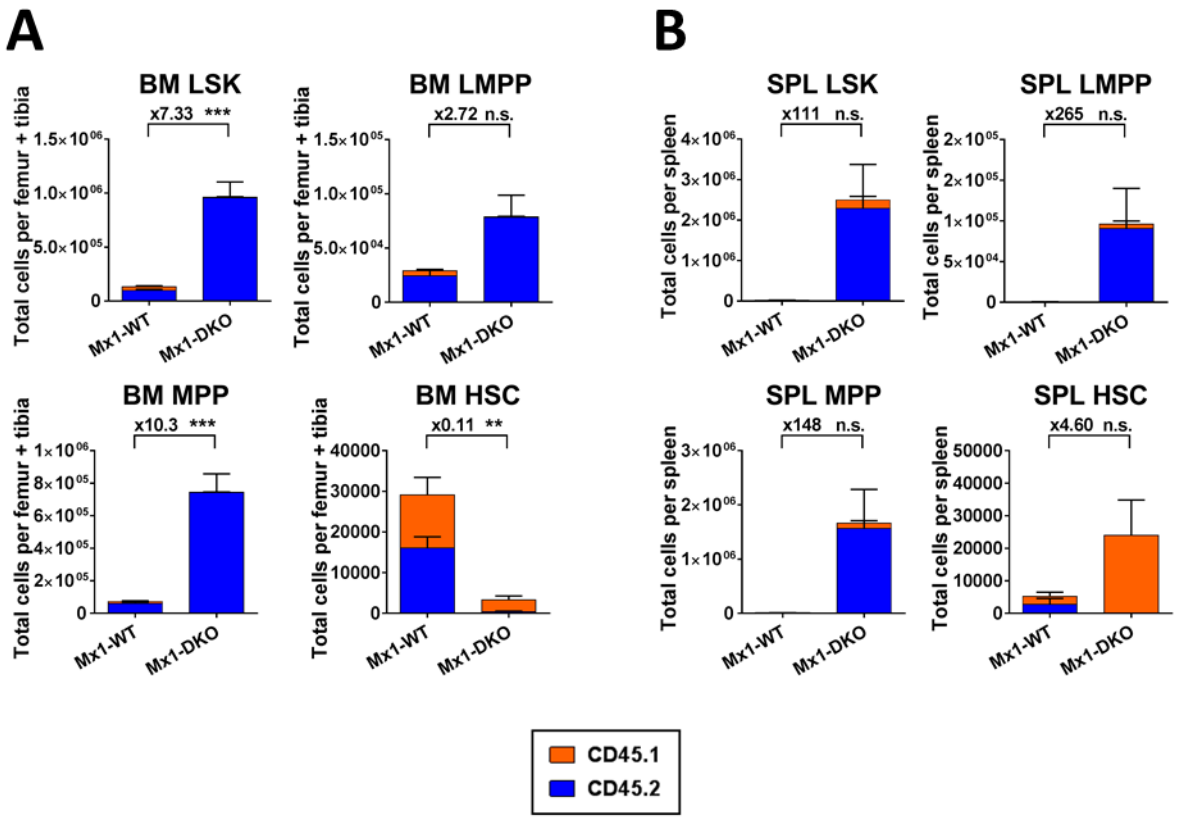


Figure 3.10

(A) Absolute numbers per femur and tibia of CD45.1 and CD45.2 cells contributing to indicated bone marrow stem and progenitor populations 16 weeks post-Poly(IC) treatment in Mx1-WT (n=6) and Mx1-DKO (n=7) bone marrow recipient mice. (B) Absolute numbers per spleen of CD45.1 and CD45.2 cells contributing to indicated splenocyte stem and progenitor populations 16 weeks post-Poly(IC) treatment in Mx1-WT (n=4) and Mx1-DKO (n=4) bone marrow recipient mice. LSK Lineage$^-$Sca1$^+$Kit$^+$, LMPP Lineage$^-$Sca1$^+$Kit$^+$Flt3^{high}, MPP Lineage$^-$Sca1$^+$Kit$^+$CD48$^+$CD150$^-$, HSC Lineage$^-$Sca1$^+$Kit$^+$CD48$^-$ CD150$^+$. Error bars represent standard error of the mean. P-values are by T-test comparing summed CD45.1 and CD45.2 cells; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

recipient spleens and were mainly CD45.2 (Figure 3.10B). This indicates a spilling over of Mx1-DKO progenitor cells from the bone marrow, where they are strongly increased in number, to extramedullary sites. Interestingly, numbers of HSCs in the spleen were also increased in Mx1-DKO compared to Mx1-WT recipient mice, and as in the bone marrow they were almost all CD45.1 (Figure 3.10B). This is indicative of extramedullary stress haematopoiesis resulting from migration of wild-type CD45.1 HSCs to the spleen, which is known to occur in response to disruption of the bone marrow by malignant haematopoiesis (Kim, 2010).

Together, these data demonstrate that *Ezh2* and *Runx1* inactivation confers a competitive advantage to haematopoietic cells. This may be achieved by an extrinsic suppression of wild-type HSCs, combined with a proliferative expansion of early progenitor populations including MPPs and myeloid precursors. Since Mx1-DKO HSCs were strongly reduced in primary mice and almost completely absent in transplanted recipient mice, we wondered whether the MDS phenotype which later develops in Mx1-DKO mice could be propagated by transformed progenitor cells, rather than HSCs. To explore this question, we decided to make use of a *Cre* line which targets deletion to MPPs but not HSCs.

2.4: Phenotypic consequences are similar whether inactivation of *Ezh2* and *Runx1* is targeted to MPPs or HSCs

In *Flt3Cre* mice, almost all haematopoietic cells pass through a *Flt3Cre*-activated stage as they leave the HSC state, but crucially *Cre* is not expressed in HSCs (Boyer et al., 2011; Buza-Vidas et al., 2011). This allowed us to examine the impact of targeting deletion of *Ezh2* and *Runx1* to MPPs (and therefore all mature haematopoietic cells) but not HSCs. We initially hypothesised that, since MDS in humans is thought to be propagated by HSCs and not progenitors (Woll et al., 2014), mice in which *Ezh2* and *Runx1* inactivation was targeted to MPPs would develop a distinct phenotype.

We analysed Ezh2$^{fl/fl}$Runx1$^{fl/fl}$Flt3Cre$^{tg/+}$ (Flt3-DKO) mice together with Ezh2$^{fl/fl}$Runx1$^{fl/fl}$Flt3Cre$^{+/+}$ (Flt3-WT) littermate controls at 6-8 weeks of age (Figure 3.11A). We first confirmed that *Ezh2* and *Runx1* were deleted at high efficiency in MPPs but not HSCs using qPCR (Figure 3.11B). Despite the lack of deletion in HSCs, we found that the overall phenotype of Flt3-DKO mice was very similar to Mx1-DKO mice. Total white blood cell counts were lower in Flt3-DKO compared to Flt3-WT mice (Figure 3.12A). There was no significant reduction in haemoglobin, but platelet numbers were decreased (Figure 3.12B). Thymus cellularity was markedly decreased in all knockout genotypes (Figure 3.12B). Peripheral blood B and T cells were decreased, and monocytes were increased while neutrophils were not significantly affected (Figure 3.13A). As with *Mx1Cre*, Kit^{+}Mac1^{low} myeloid precursor cells were markedly increased in the peripheral blood compared to Flt3-WT mice (Figure 3.13A). Also similarly to *Mx1Cre*, Flt3-DKO bone marrow showed reduced B-cells, increased monocytes and markedly increased Kit^{+}Mac1^{low} myeloid precursor cells in the bone marrow compared to Flt3-WT mice (Figure 3.13B). T-cell numbers were not significantly affected, while neutrophils were slightly decreased (Figure 3.13B).

In the LSK stem and progenitor compartment, we saw the same marked increase in total LSK and MPP numbers in Flt3-DKO compared to Flt3-WT mice as we had previously observed with *Mx1Cre* (Figure 3.14). Interestingly, while we still saw a decrease in phenotypic HSCs in Flt3-DKO compared to Flt3-WT mice, this decrease was not as striking as that observed in Mx1-DKO compared to Mx1-WT mice (x0.45 versus x0.17; Figures 3.14 and 3.04). This may reflect the fact that HSCs in Flt3-DKO mice are not *Ezh2* and *Runx1*-inactivated and therefore would not be expected to be reduced in number. The fact that we do still observe a decrease may reflect the fact that the phenotypically-defined HSC population includes many cells that are not true HSCs; it has been estimated that only around 50% of these cells are in fact true HSCs (Kiel et al., 2005). Alternatively, it is possible that there may be an extrinsic suppression of wild-type HSCs by *Ezh2* and *Runx1*-inactivated non-HSCs within the bone marrow, as suggested by our *Mx1Cre* competitive transplant experiments, and also as observed in other models of myeloid malignancies (Schepers et al., 2013; Zhang et al., 2012a).

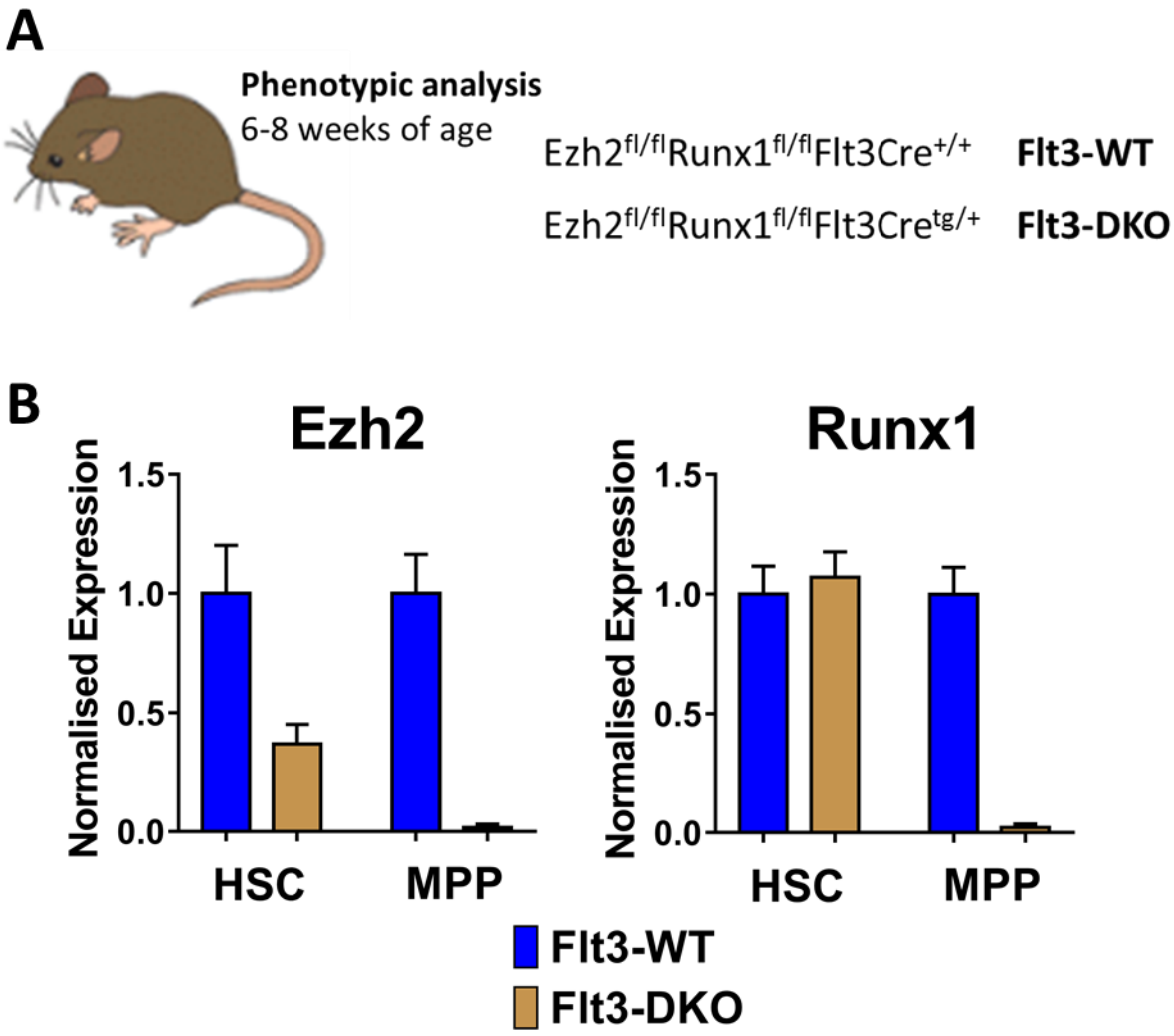


Figure 3.11

(A) Experimental setup and genotypes of experimental mice for Ezh2 and/or Runx1 inactivation in MPPs using Flt3Cre. (B) Determination of efficiency of deletion by Flt3Cre of conditionally targeted Ezh2 and Runx1 alleles, in $Lin^{-}Kit^{+}Sca1^{+}CD150^{+}CD48^{-}$ HSCs and $Lin^{-}Kit^{+}Sca1^{+}CD150^{-}CD48^{+}$ MPPs from 6-8 week old Mx1-WT and Mx1-DKO mice. Relative expression levels of Ezh2 and Runx1 mRNA are shown; Ct values were normalised to Hprt. Amplicons were designed to target exons removed by Cre recombination. Flt3-WT n=6, Flt3-DKO n=6; each replicate represents a sample of 50 cells. Error bars represent standard error of the mean. Deletion efficiency experiments were performed by Wen Hao Neo.

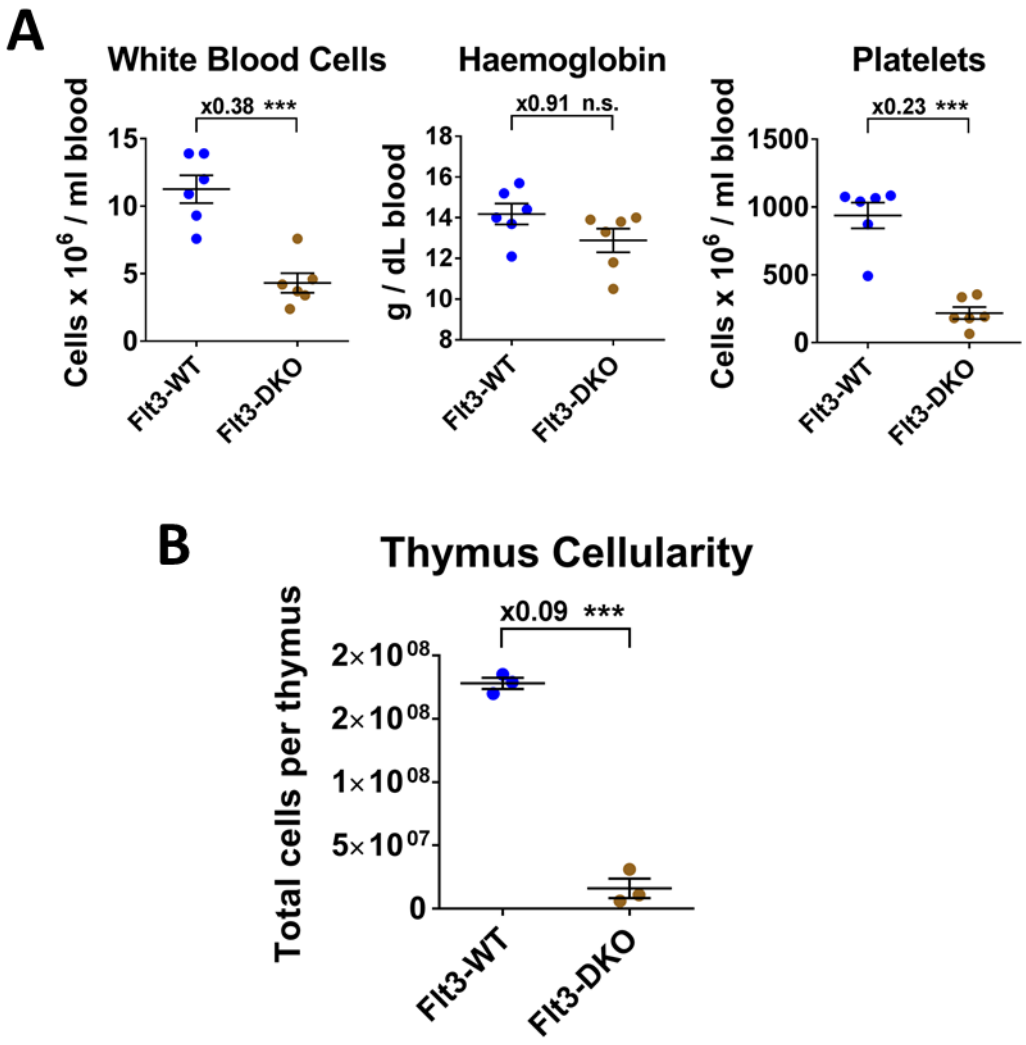


Figure 3.12

(A) Haematological parameters of peripheral blood at 6-8 weeks of age in Flt3-WT (n=6) and Flt3-DKO (n=6) mice. B) Total thymus cellularity at 6-8 weeks of age in Flt3-WT (n=3) and Flt3-DKO (n=3) mice. Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001. Experiments were performed with help from Wen Hao Neo.

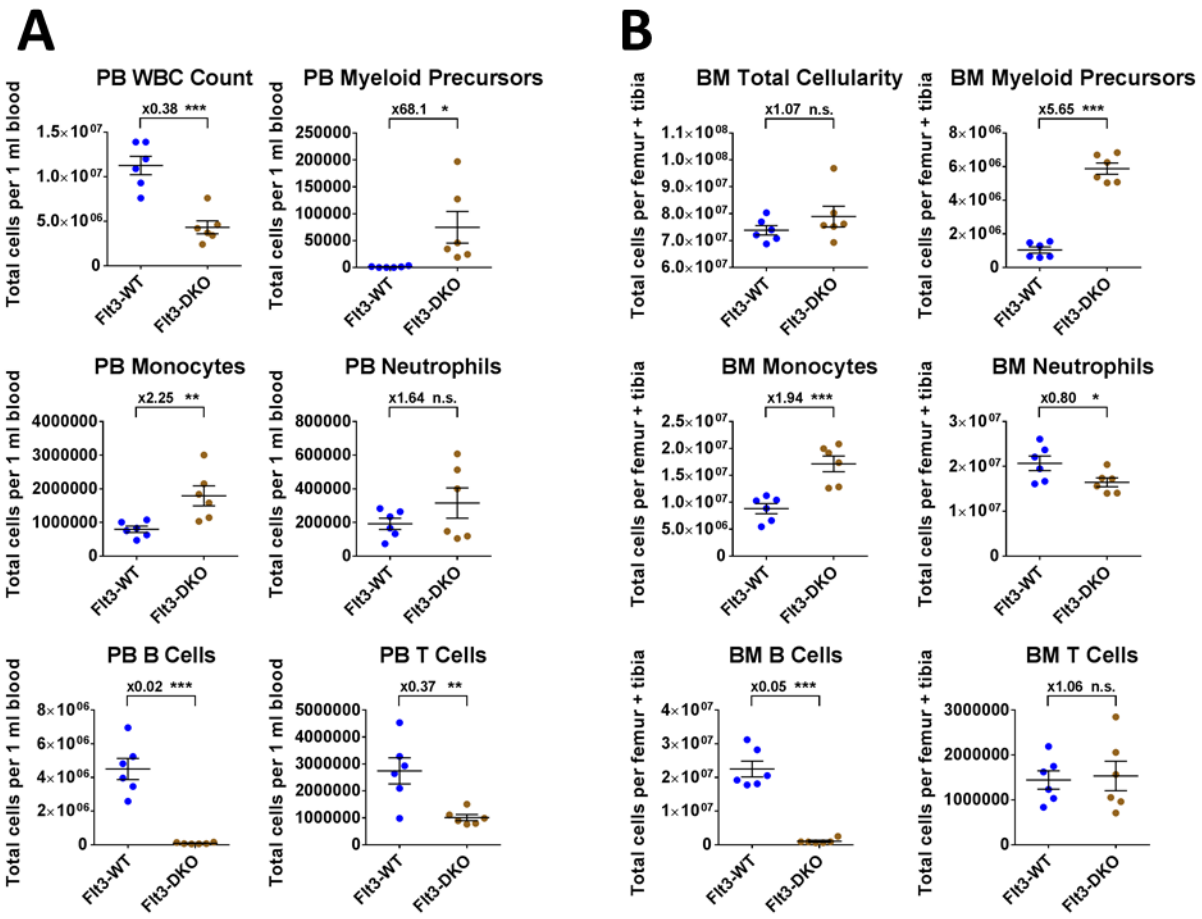


Figure 3.13

(A) Absolute numbers per ml of indicated leukocyte populations in the peripheral blood at 6-8 weeks of age in Flt3-WT (n=6) and Flt3-DKO (n=6) mice. (B) Absolute numbers per femur and tibia of indicated bone marrow populations at 6-8 weeks of age in Flt3-WT (n=6) and Flt3-DKO (n=6) mice. B-cells $CD19^{+}$, T-cells $CD4^{+}$ or $CD8^{+}$, monocytes $Mac1^{+}Gr1^{-}$, neutrophils $Mac1^{+}Gr1^{+}$, myeloid precursors $Kit^{+}Mac1^{low}$. Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001. Experiments were performed with help from Wen Hao Neo.

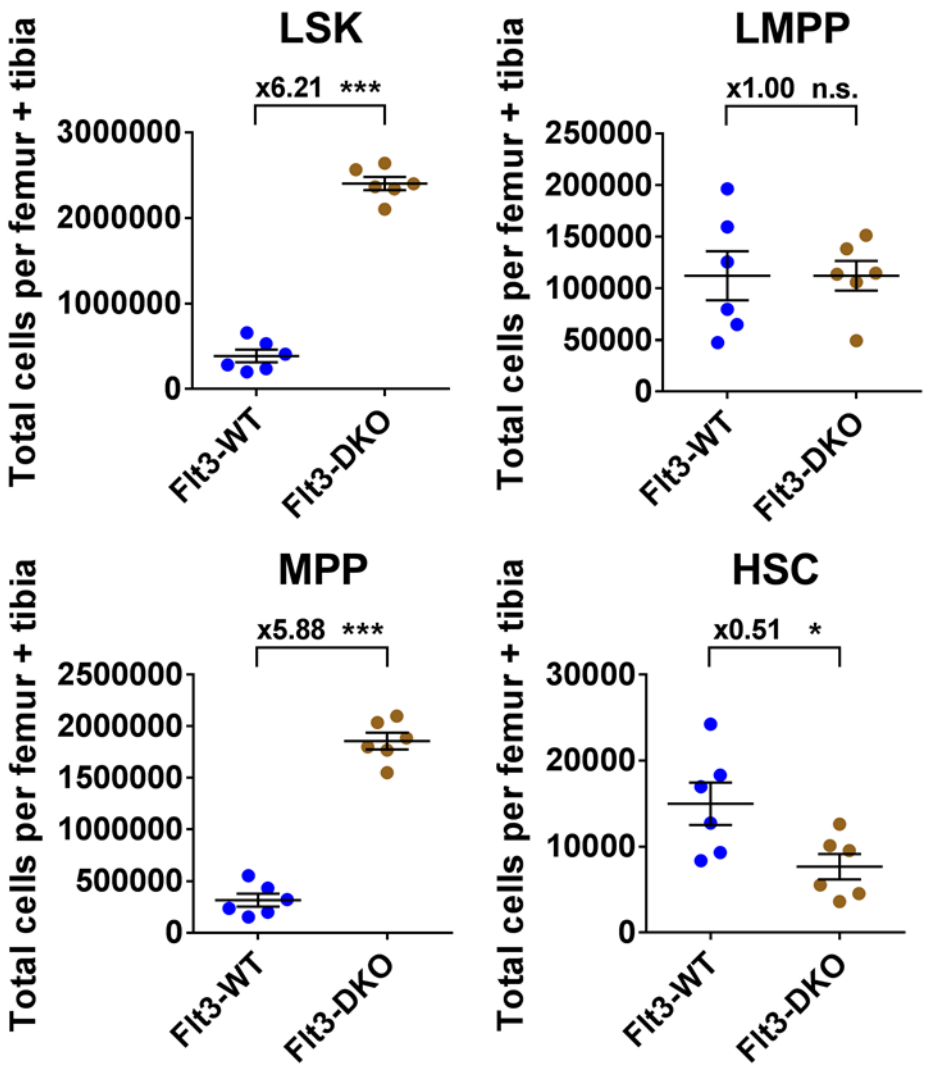


Figure 3.14

Absolute numbers per femur and tibia of indicated bone marrow stem and progenitor populations at 6-8 weeks of age in Flt3-WT (n=6) and Flt3-DKO (n=6) mice. LSK Lineage$^-$Sca1$^+$Kit$^+$, LMPP Lineage$^-$ Sca1$^+$Kit$^+$Flt3^{high}, MPP Lineage$^-$Sca1$^+$Kit$^+$CD48$^+$CD150$^-$, HSC Lineage$^-$Sca1$^+$Kit$^+$CD48$^-$CD150$^+$. Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001. Experiments were performed with help from Wen Hao Neo.

Having established that inactivation of *Ezh2* and *Runx1* in HSCs induces MDS, we decided to follow up Flt3-DKO mice to determine the outcome of inactivation in MPPs. Direct comparison of Flt3-DKO and Mx1-DKO mice by age (rather than weeks post-Poly(IC) treatment) showed that there was no significant difference in survival between the two genotypes (Figure 3.15A). Peripheral blood analysis of Flt3-DKO mice at time of culling revealed a similar pancytopenia phenotype to Mx1-DKO mice (Figure 3.15B). These results suggest that inactivation of *Ezh2* and *Runx1* in MPPs is sufficient to induce an MDS-like phenotype.

We also performed transplantation of equal numbers of Flt3-WT or Flt3-DKO CD45.2 bone marrow cells and wild-type CD45.1 bone marrow cells, to determine whether Flt3-DKO bone marrow carries a competitive advantage over wild-type bone marrow (Figure 3.16A). This showed similar results to *Mx1Cre* (Figure 3.08B), indicating that inactivation of *Ezh2* and *Runx1* in MPPs is sufficient to confer a competitive advantage over normal bone marrow.

Analysis of bone marrow stem and progenitor cells 16 weeks post-transplantation (Figure 3.16B) demonstrated a similar extrinsic suppression of wild-type CD45.1 HSCs to that observed in Mx1-DKO recipients (Figure 3.10A). LSK, MPP and LMPP numbers in Flt3-DKO recipients were expanded, similarly to those in Mx1-DKO recipients. Interestingly, while HSC numbers were reduced in Flt3-DKO recipients, those remaining were mainly CD45.2 (Figure 3.16B). This was in contrast to Mx1-DKO recipients whose HSCs at the final timepoint were mainly CD45.1 (Figure 3.10A). This reflects the fact that Flt3-DKO HSCs are non-deleted, and despite being reduced in number were able to remain present in the bone marrow.

Together, these data demonstrate that inactivation of *Ezh2* and *Runx1* in multipotent progenitors using *Flt3Cre* induces a similar phenotype to that observed using *Mx1Cre*. Numbers of mature lymphoid cells are reduced and mice show mild anaemia at 6 weeks of age. Upon follow-up, Flt3-DKO mice develop a similar MDS phenotype to Mx1-DKO mice, with a similar latency. In the bone marrow, MPPs are expanded and transplantation of bone marrow cells demonstrates an extrinsic

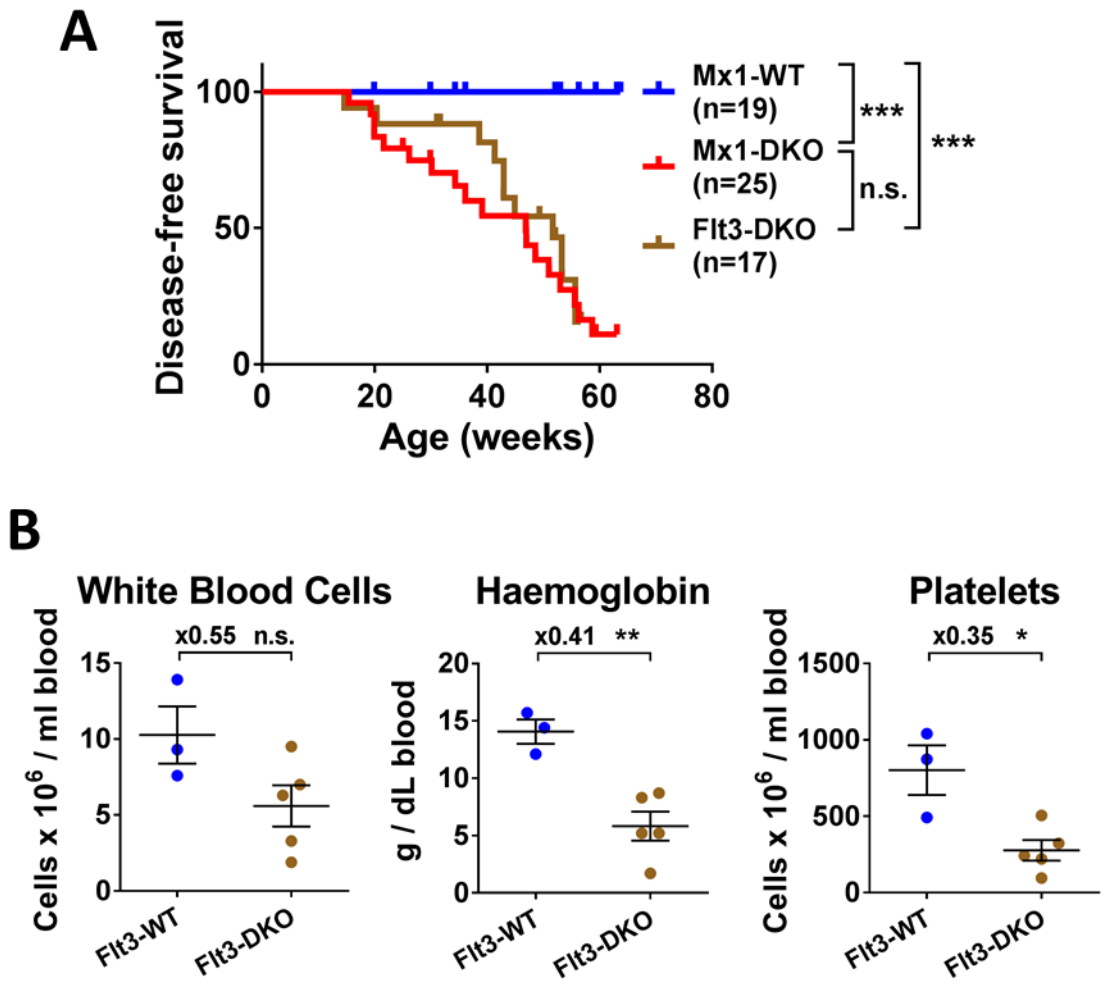


Figure 3.15

(A) Kaplan-Meier disease-free survival curves for Mx1-WT, Mx1-DKO and Flt3-DKO mice. Mx1-WT and Mx1-DKO mice were treated with Poly(IC) at 6-8 weeks of age. Mice found dead without analysis or culled showing > 25 x 10^6 leukocytes per ml blood and/or < 6.5 g haemoglobin per dl blood were counted as events. Mice culled showing leukocyte and haemoglobin counts outside these constraints were censored. (B) Haematological parameters of follow-up Flt3-WT (n=3) and Flt3-DKO (n=5) mice at time of culling. One Flt3-DKO mouse showing 35.6 x10^6 leukocytes per ml blood was excluded. Error bars represent standard error of the mean. P-values are by logrank test (A) or T-test (B); n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001. Experiments were performed with help from Wen Hao Neo.

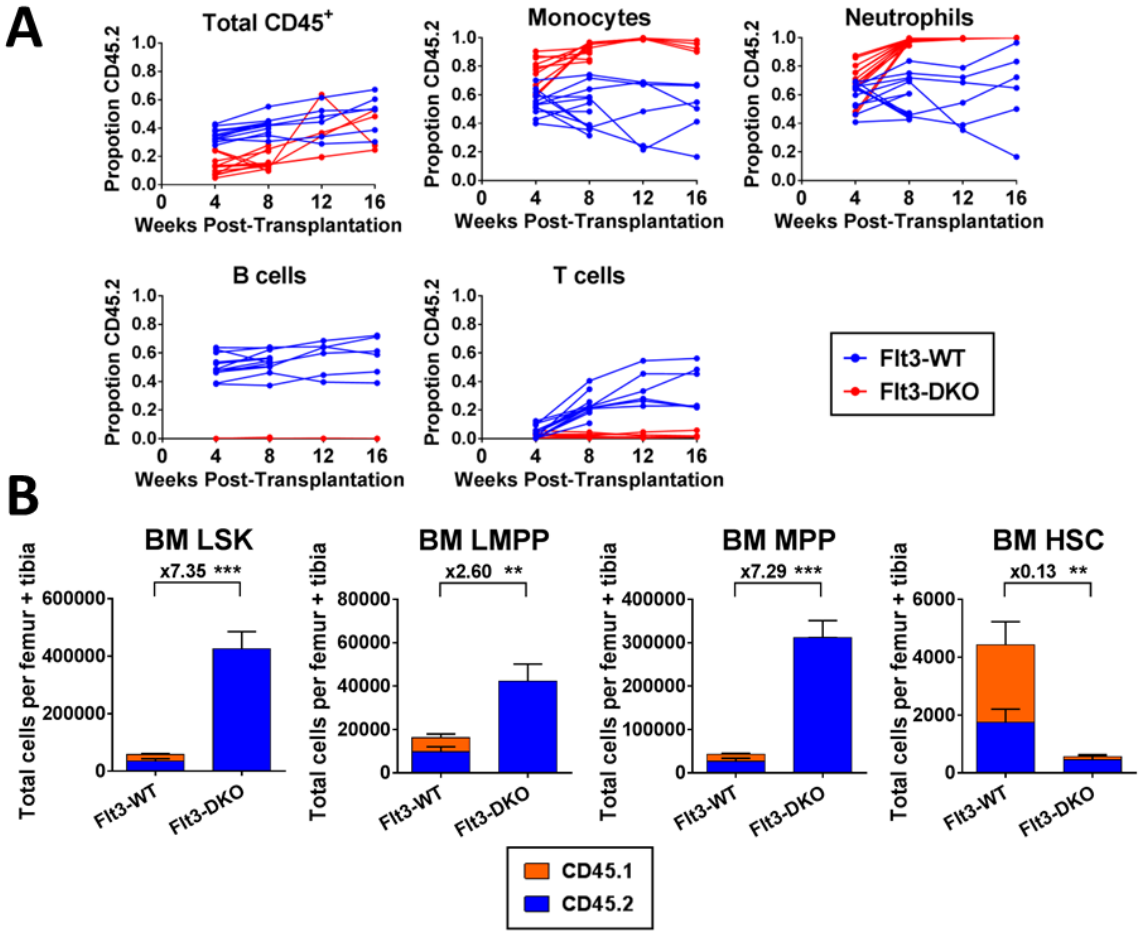


Figure 3.16

(A) Relative contribution of CD45.2 cells to indicated leukocyte populations in the peripheral blood of wild-type CD45.1 recipient mice following transplantation of CD45.2 Flt3-WT or Flt3-DKO bone marrow cells. Each line represents an individual recipient mouse; Flt3-WT n=12, Flt3-DKO n=12. Monocytes $Mac1^{+}Gr1^{-}$, neutrophils $Mac1^{+}Gr1^{+}$, B-cells $CD19^{+}$, T-cells $CD4^{+}$ or $CD8^{+}$. (B) Absolute numbers per femur and tibia of CD45.1 and CD45.2 cells contributing to indicated bone marrow stem and progenitor populations 16 weeks post-transplantation in Flt3-WT (n=6) and Flt3-DKO (n=4) bone marrow recipient mice. LSK $Lineage^{-}Sca1^{+}Kit^{+}$, LMPP $Lineage^{-}Sca1^{+}Kit^{+}Flt3^{high}$, MPP $Lineage^{-}Sca1^{+}Kit^{+}CD48^{+}CD150^{-}$, HSC $Lineage^{-}Sca1^{+}Kit^{+}CD48^{-}CD150^{+}$. Error bars represent standard error of the mean. P-values are by T-test comparing summed CD45.1 and CD45.2 cells; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001. Experiments were performed with help from Wen Hao Neo.

suppression of wild-type HSCs, as is the case in Mx1-DKO. These findings indicate that inactivation of *Ezh2* and *Runx1* in HSCs is not required to mediate the development of MDS or suppression of wild-type HSCs. This raises the question of which cells are responsible for propagating these phenotypes: HSCs, MPPs, or other progenitor populations? To answer this question, we decided to perform transplantation of purified stem and progenitor populations from Mx1-DKO mice.

2.5: *Ezh2* and *Runx1* inactivation confers self-renewal to multipotent progenitor populations

Having established that *Ezh2* and *Runx1*-inactivated bone marrow cells cause an extrinsic suppression of wild-type HSCs, leading to reduced overall output and extra-medullary haematopoiesis, we next decided to investigate which bone marrow populations might be responsible for this phenotype. As Flt3-DKO mice develop a very similar phenotype to Mx1-DKO mice despite lack of recombination in the HSC compartment, including outcompeting normal HSCs, we reasoned that disease propagating cells may not reside within the HSC compartment. To test this, we sorted HSCs, MPPs, LSK $CD48^{+}CD150^{+}$ early progenitors (DPs) and $Lin^{-}Kit^{+}Sca1^{-}$ myeloid progenitors (LKs) from CD45.2 Mx1-WT or Mx1-DKO mice 6 weeks post-Poly(IC) treatment, and transplanted these populations into wild-type CD45.1 recipient mice (Figure 3.17A). We then analysed peripheral blood every 4 weeks and culled the recipients for analysis of the bone marrow at 16 weeks post-transplant.

Among the recipients transplanted with Mx1-WT cells (Figure 3.17B), 4 of 4 HSC and 1 of 4 DP recipients showed sustained, high-level (>25%) contribution to mature myeloid cells (monocytes and neutrophils) over 16-weeks. Mx1-WT MPPs showed low-level contribution to mature myeloid cells in some recipients, which reduced with time, while Mx1-WT LK cells did not show any contribution. These findings are consistent with previous characterisation of these populations in wild-type mice (Oguro et al., 2013), except for our observation that DP cells produced long-term reconstitution in

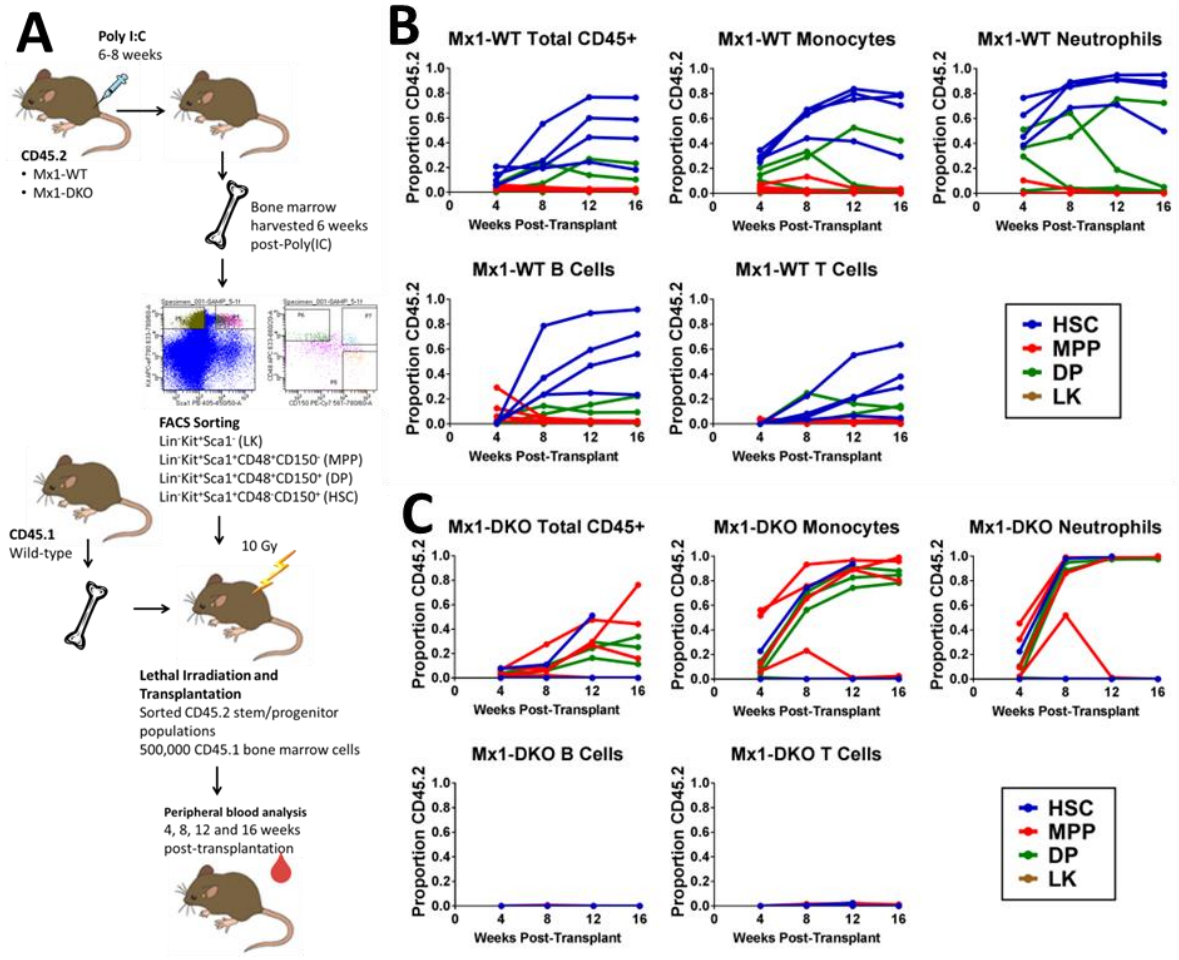


Figure 3.17

(A) Experimental setup for transplantation of stem and progenitor populations from CD45.2 Mx1-WT and Mx1-DKO mice after activation of Mx1Cre using Poly(IC), together with CD45.1 bone marrow support cells. (B-C) Relative contribution of CD45.2 cells to indicated leukocyte populations in the peripheral blood of wild-type CD45.1 recipient mice following transplantation of CD45.2 Mx1-WT (B) or Mx1-DKO (C) stem and progenitor populations after Mx1-Cre activation. Each line represents an individual recipient mouse. HSC Lineage$^-$Sca1$^+$Kit$^+$CD48$^-$CD150$^+$ (n=2 per genotype, 100 cells transplanted), MPP Lineage$^-$Sca1$^+$Kit$^+$CD48$^+$CD150$^-$ (n=2 per genotype, 502-750 Mx1-WT or 4347-8000 Mx1-DKO cells transplanted), DP Lineage$^-$Sca1$^+$Kit$^+$CD48$^+$CD150$^+$ (n=2 per genotype, 100 Mx1-WT or 742-2447 Mx1-DKO cells transplanted), LK Lineage$^-$Sca1$^-$Kit$^+$ (n=2 per genotype, 8000 cells transplanted). Monocytes Mac1$^+$Gr1$^-$, neutrophils Mac1$^+$Gr1$^+$, B-cells CD19$^+$, T-cells CD4$^+$ or CD8$^+$.

one recipient. Another recipient of these cells also showed high contribution to neutrophils for the first 8 weeks before falling rapidly. This suggests that the DP population may contain some medium-term repopulating activity in these mice, or alternatively that a small number of HSCs fell into the DP gate during FACS sorting. It has been shown that CD48 expression can be transiently upregulated by HSCs upon proliferation (Venezia et al., 2004). This would position them in the DP compartment and could explain our observed infrequent reconstitution potential by these cells. It is possible that Poly(IC) treatment was responsible for inducing proliferation of HSCs, since this is a known consequence of this procedure (Essers et al., 2009).

Analysis of recipients transplanted with Mx1-DKO cells, meanwhile, showed a very different picture (Figure 3.17C). Only 1 of 4 HSC-transplanted recipients showed high-level contribution to mature myeloid cells at any of the analysis timepoints. In contrast, 3 of 4 MPP and 3 of 4 DP-transplanted recipients showed high-level, long-term reconstitution which did not appear to be falling by the end of the experiment. No Mx1-DKO cells produced mature B-cells or T-cells, in contrast to Mx1-WT cells and in agreement with our previous findings (Figure 3.02). These results suggest that phenotypically defined MPPs possess self-renewal potential in *Ezh2* and *Runx1*-inactivated mice, while HSCs are severely functionally impaired.

At 16 weeks post-transplant we culled the recipients and analysed the bone marrow (Figure 3.18). The one CD45.2-reconstituted recipient transplanted with Mx1-DKO HSCs was found dead at 14 weeks post-transplant, precluding bone marrow analysis. Reconstitution of mature lineages in remaining recipient mice largely mirrored those in the peripheral blood (Figure 3.18A). CD45.2-reconstituted recipients transplanted with Mx1-DKO cells all showed an expansion of $Kit^{+}Mac1^{low}$ myeloid precursor cells, which were almost entirely CD45.2 (Figure 3.18A). In the stem and progenitor compartment, 2 of the 3 reconstituted Mx1-DKO MPP and 2 of the 3 reconstituted Mx1-DKO DP transplanted recipients showed an expansion of MPPs and DPs (Figure 3.18B). Very few phenotypic HSCs were detected in engrafted recipients of MPP and DP DKO cells.

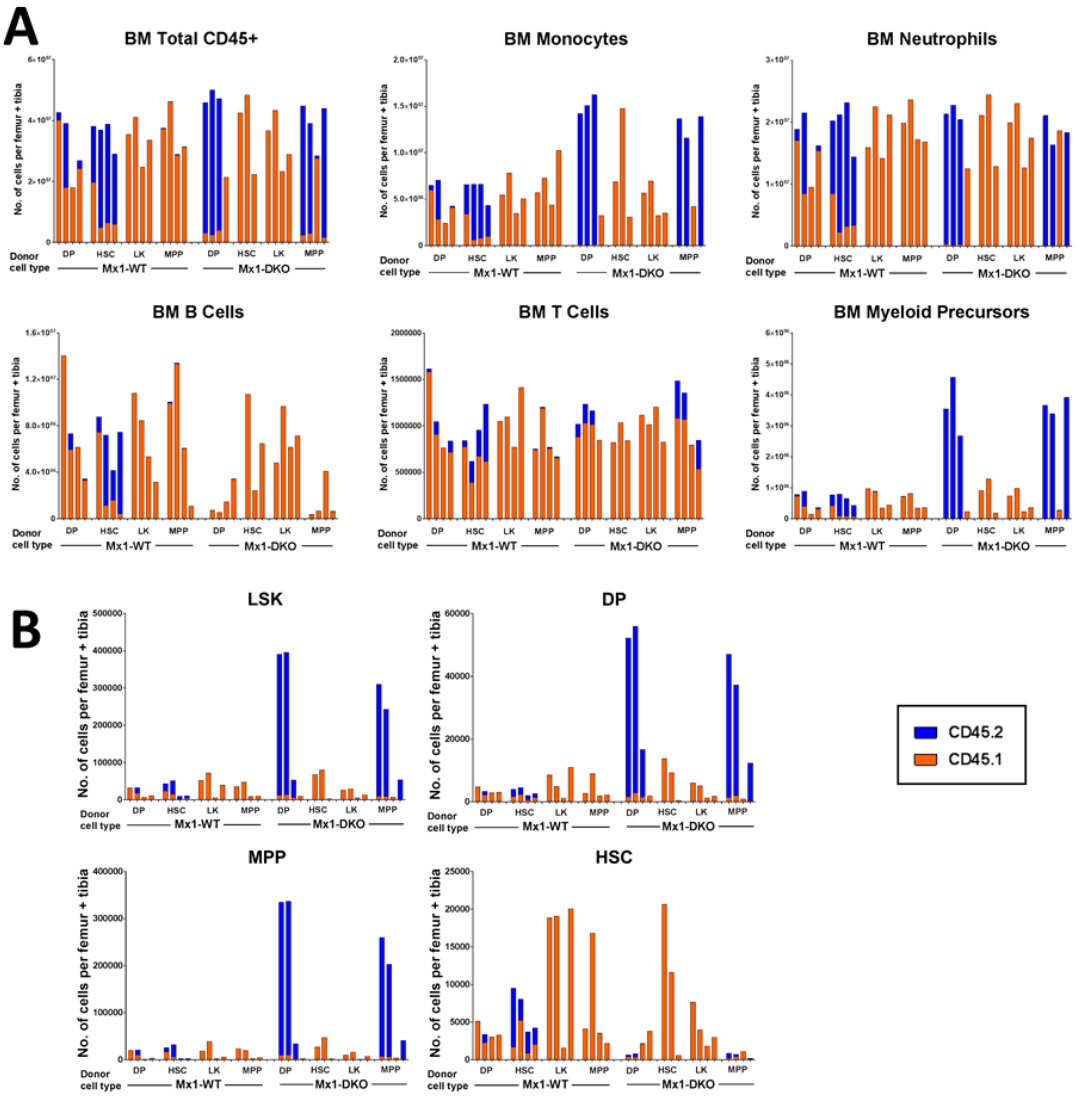


Figure 3.18

(A-B) Absolute numbers per femur and tibia of CD45.1 and CD45.2 cells contributing to indicated bone marrow mature cell (A) and stem and progenitor (B) populations 16 weeks post-transplant in wild-type CD45.1 recipient mice transplanted with CD45.2 Mx1-WT or Mx1-DKO bone marrow stem and progenitor populations. Each column represents an individual mouse, with the transplanted cell population and genotype indicated underneath. Monocytes $Mac1^{+}Gr1^{-}$, neutrophils $Mac1^{+}Gr1^{+}$, B-cells $CD19^{+}$, T-cells $CD4^{+}$ or $CD8^{+}$, myeloid precursors $Kit^{+}Mac1^{low}$, LSK $Lineage^{-}Sca1^{+}Kit^{+}$, HSC $Lineage^{-}Sca1^{+}Kit^{+}CD48^{-}CD150^{+}$, MPP $Lineage^{-}Sca1^{+}Kit^{+}CD48^{+}CD150^{-}$, DP $Lineage^{-}Sca1^{+}Kit^{+}CD48^{+}CD150^{+}$, LK $Lineage^{-}Sca1^{-}Kit^{+}$.

Together, these data show that: (1) Ezh2 and Runx1 inactivation in HSCs and/or MPPs leads to the development of an MDS-like phenotype, including a cell-extrinsic suppression of wild-type HSCs, (2) this MDS-like disease is propagated by phenotypic MPPs and not HSCs, and (3) there is no clear hierarchy between MPPs and DPs in Mx1-DKO mice; each population is able to give rise to the other.

This experimental approach does not distinguish between the possibilities that either: (1) *Ezh2* and *Runx1* inactivation in MPPs confers aberrant self-renewal to these cells, or (2) *Ezh2* and *Runx1* inactivation in self-renewing HSCs causes them to aberrantly gain an MPP surface phenotype (Cai et al., 2011). We reasoned that the latter possibility was less likely, given that this would require both gain of CD48 and loss of CD150 for HSCs to acquire an aberrant MPP phenotype. To help to determine which of these non-mutually exclusive possibilities are correct, we performed a similar experiment using Flt3-DKO foetal liver cells. At the E13.5 stage, the LSK stem and progenitor phenotype of Flt3-DKO mice is similar to WT mice, with no MPP expansion or HSC suppression (Figure 3.19A). We reasoned that this, combined with the fact that *Flt3Cre* does not activate in true HSCs (Boyer et al., 2011; Buza-Vidas et al., 2011), would preclude HSCs having aberrantly gained an MPP surface phenotype as a result of *Ezh2* and *Runx1* inactivation at this stage of ontogeny.

We therefore sorted HSCs, MPPs and DPs from Flt3-WT and Flt3-DKO E13.5 foetal livers and transplanted them into wild-type recipients (Figure 3.19B). Recipients of Flt3-WT cells showed stable CD45.2 reconstitution only when HSCs were transplanted (Figure 3.20A). Flt3-DKO cells were sorted from two embryos and surprisingly, one of these produced reconstitution from HSCs but not MPPs or DPs, while the other embryo produced stable reconstitution from both MPPs and DPs but not HSCs (Figure 3.20B). When recipient bone marrow was analysed at 16 weeks post-transplant, the engrafted Flt3-DKO HSCs had given rise to all LSK populations, while Flt3-DKO MPPs and DPs gave rise to MPPs and DPs but not HSCs (Figure 3.21). HSC numbers were reduced in all engrafted Flt3-DKO recipients (Figure 3.21). These results support the hypothesis that *Ezh2* and *Runx1* inactivation confers aberrant self-renewal to multipotent progenitors. The finding that these transplanted

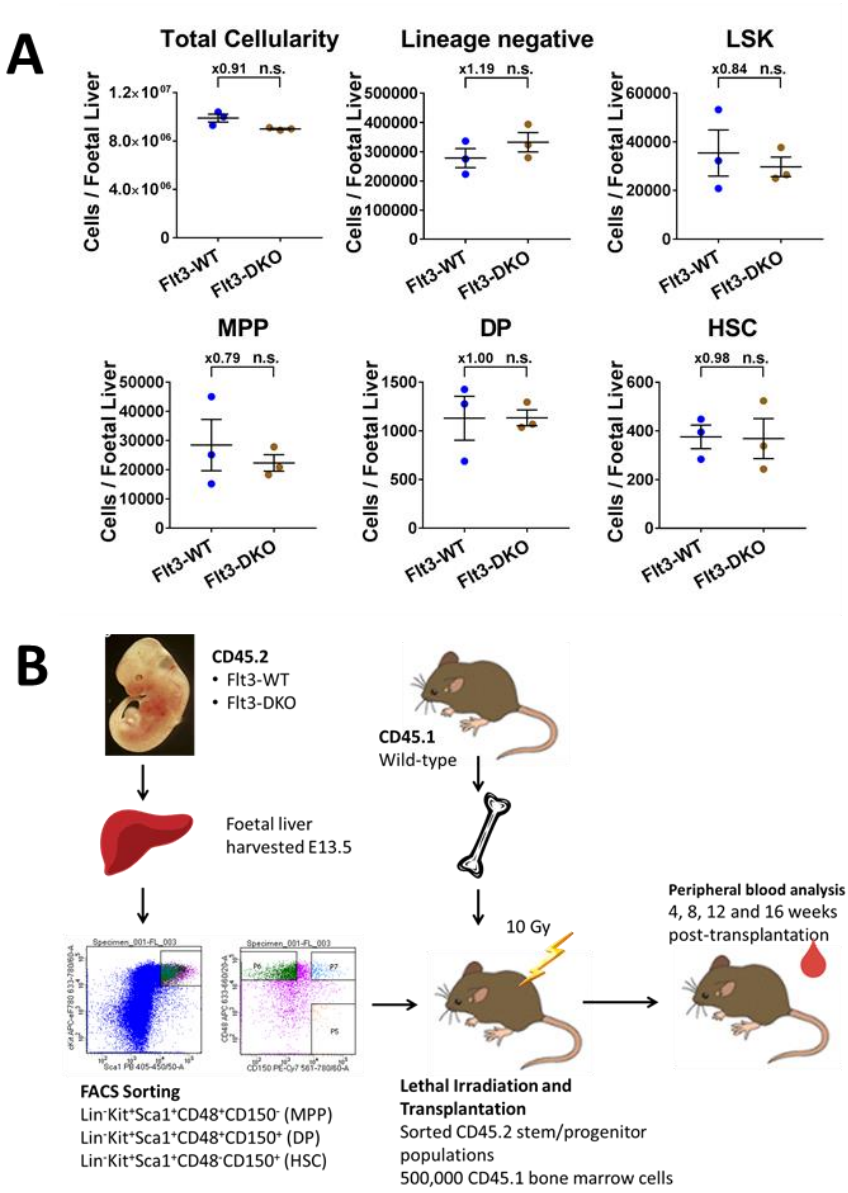


Figure 3.19

(A) Absolute numbers per foetal liver of indicated stem and progenitor populations in E13.5 Flt3-WT (n=3) and Flt3-DKO (n=3) embryos. LSK Lineage$^-$Sca1$^+$Kit$^+$, HSC Lineage$^-$Sca1$^+$Kit$^+$CD48$^-$CD150$^+$, MPP Lineage$^-$Sca1$^+$Kit$^+$CD48$^+$CD150$^-$, DP Lineage$^-$Sca1$^+$Kit$^+$CD48$^+$CD150$^+$. (B) Experimental setup for transplantation of stem and progenitor populations from CD45.2 Flt3-WT or Flt3-DKO E13.5 foetal livers, together with CD45.1 bone marrow support cells. Error bars represent standard error of the mean. P-values are by T-test comparing summed CD45.1 and CD45.2 cells; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

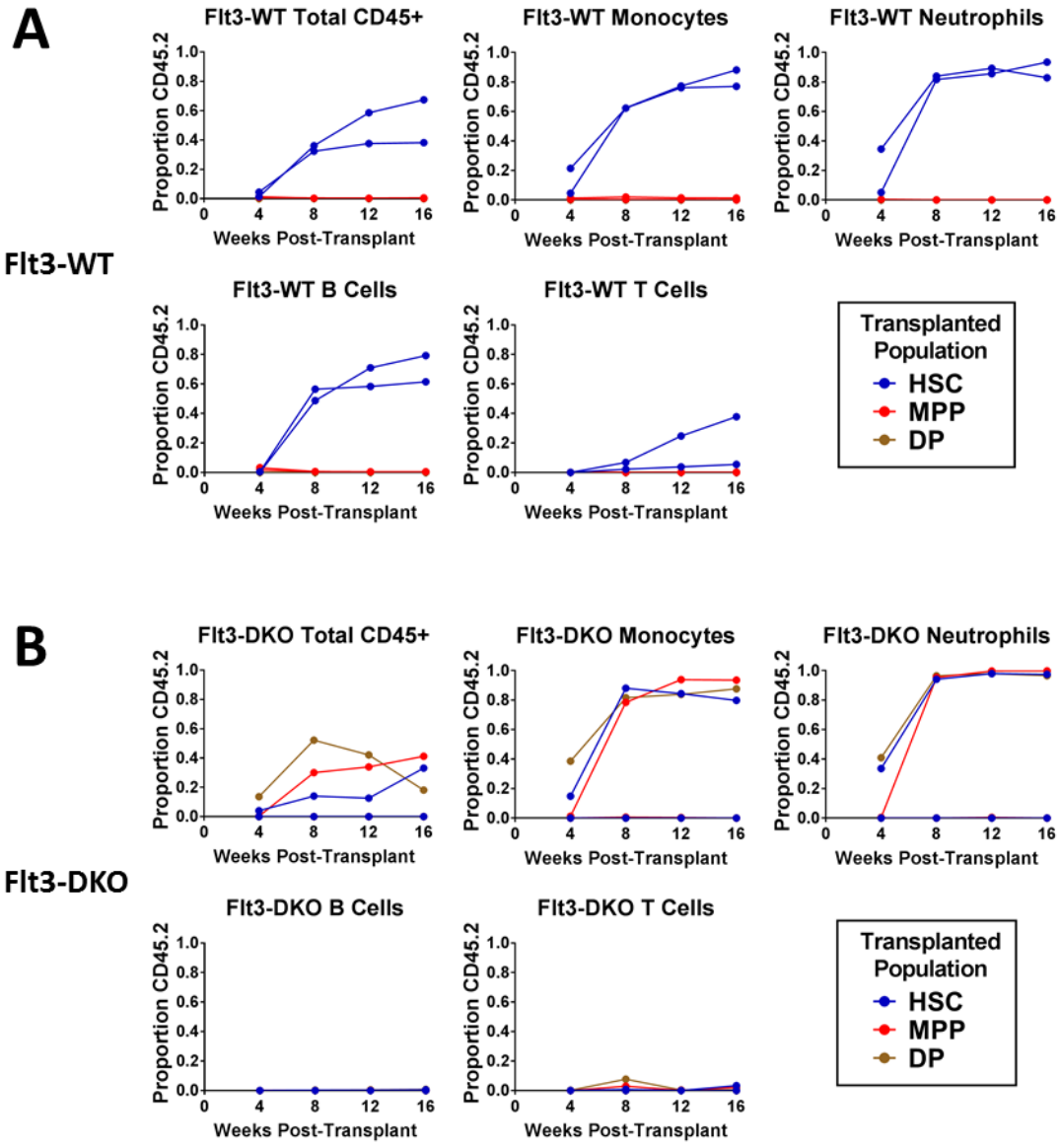


Figure 3.20

(A-B) Relative contribution of CD45.2 cells to indicated leukocyte populations in the peripheral blood of wild-type CD45.1 recipient mice following transplantation of CD45.2 Flt3-WT (A) or Flt3-DKO (B) E13.5 foetal liver stem and progenitor populations. Each line represents an individual recipient mouse. HSC Lineage$^-$Sca1$^+$Kit$^+$CD48$^-$CD150$^+$ (n=2 per genotype, 53-84 cells transplanted), MPP Lineage$^-$Sca1$^+$Kit$^+$CD48$^+$CD150$^-$ (n=2 per genotype, 2959-4729 cells transplanted), DP Lineage$^-$ Sca1$^+$Kit$^+$CD48$^+$CD150$^+$ (n=2 per genotype, 181-348 cells transplanted). Monocytes Mac1$^+$Gr1$^-$, neutrophils Mac1$^+$Gr1$^+$, B-cells CD19$^+$, T-cells CD4$^+$ or CD8$^+$.

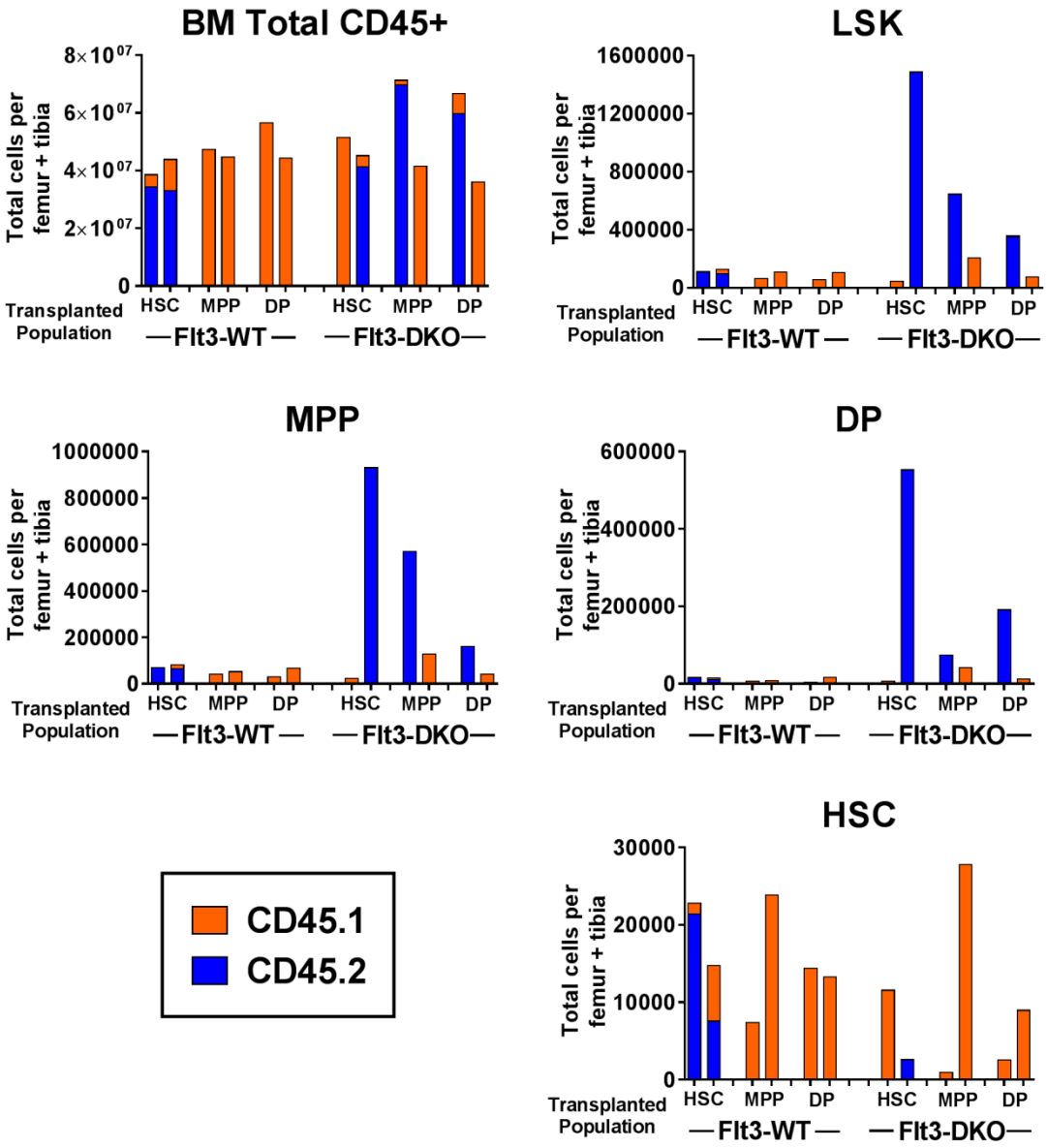


Figure 3.21

Absolute numbers per femur and tibia of CD45.1 and CD45.2 cells contributing to indicated bone marrow stem and progenitor populations 16 weeks post-transplant in wild-type CD45.1 recipient mice transplanted with CD45.2 Flt3-WT or Flt3-DKO E13.5 foetal liver stem and progenitor populations. Each column represents an individual mouse, with the transplanted cell population and genotype indicated underneath. LSK Lineage$^-$Sca1$^+$Kit$^+$, HSC Lineage$^-$Sca1$^+$Kit$^+$CD48$^-$CD150$^+$, MPP Lineage$^-$Sca1$^+$Kit$^+$CD48$^+$CD150$^-$, DP Lineage$^-$Sca1$^+$Kit$^+$CD48$^+$CD150$^+$.

multipotent progenitors do not give rise to phenotypically defined HSCs, together with previous evidence that *Flt3Cre* does not activate in true HSCs (Boyer et al., 2011; Buza-Vidas et al., 2011), provides further support that these aberrantly self-renewing progenitor cells are not simply HSCs that have gained deregulated surface marker expression.

2.6: Targeting of *Ezh2* and *Runx1* inactivation to early lymphoid progenitors does not result in suppression of HSCs or development of MDS

We next decided to test the impact of inactivation of *Ezh2* and *Runx1* in a more differentiated progenitor population, to see whether this apparent ability to suppress HSCs was specific to *Ezh2* and *Runx1*-inactivated HSCs and MPPs. We used *Rag1Cre*, which activates in early lymphoid progenitors (McCormack et al., 2003). In *Rag1Cre* mice, all mature lymphoid cells have passed through a *Rag1Cre*-active stage (Boiers et al., 2013), since Rag1 is essential for rearrangement of TCR and immunoglobulin genes (Mombaerts et al., 1992). However, a small proportion (about 2.2%) of mature myeloid cells in the adult bone marrow have also passed through a *Rag1Cre*-active stage (Boiers et al., 2013), since it first becomes active in LMPPs and ETPs which contribute to myelopoiesis as well as lymphopoiesis. *Rag1Cre* does not become active in any cells which produce erythroid or megakaryocyte progeny (Boiers et al., 2013).

We analysed $Ezh2^{fl/fl}Runx1^{fl/fl}Rag1^{+/+}$ (Rag1-WT), $Ezh2^{fl/fl}Runx1^{+/+}Rag1^{Cre/+}$ (Rag1-Ezh2KO), $Ezh2^{+/+}Runx1^{fl/fl}Rag1^{Cre/+}$ (Rag1-Runx1KO) and $Ezh2^{fl/fl}Runx1^{fl/fl}Rag1^{Cre/+}$ (Rag1-DKO) mice at 6-8 weeks of age (Figure 3.22A). We confirmed using qPCR that deletion of *Ezh2* and *Runx1* was very low in HSCs and LMPPs, but high in ETPs (Figure 3.22B). Total leukocyte counts in the peripheral blood were reduced in Rag1-Runx1KO and Rag1-DKO mice as expected, but surprisingly not in Rag1-Ezh2KO mice (Figure 3.23A). Anaemia was milder than in *Mx1Cre* knockout mice (Figures 3.23A and 3.01B). In contrast to *Mx1Cre* and *Flt3Cre* mice, no thrombocytopenia was observed in Rag1-Runx1KO or

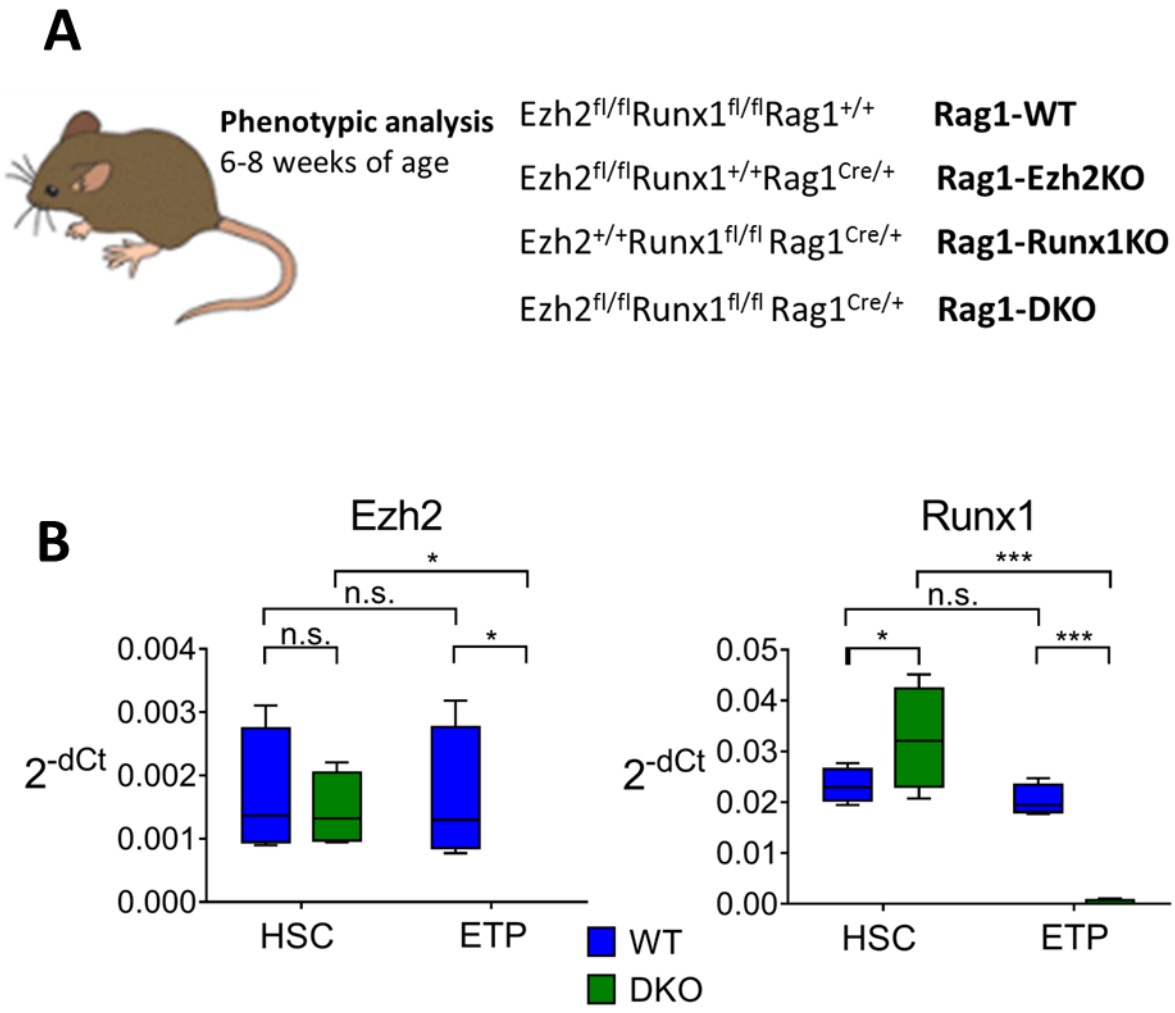


Figure 3.22

(A) Experimental setup and genotypes of experimental mice for Ezh2 and/or Runx1 inactivation in early lymphoid progenitors using Rag1Cre. (B) Determination of efficiency of deletion by Rag1Cre of conditionally targeted Ezh2 and Runx1 alleles, in $Lin^{-}Kit^{+}Sca1^{+}CD150^{+}CD48^{-}$ HSCs, $Lin^{-}Kit^{+}Sca1^{+}Flt3^{high}$ LMPPs and $Lin^{-}CD4^{-}CD8^{-}CD44^{+}CD25^{-}Kit^{+}Flt3^{+}$ ETPs from 6-8 week old WT and DKO mice. Relative expression levels of Ezh2 and Runx1 mRNA are shown; Ct values were normalised to Gapdh. Amplicons were designed to target exons removed by Rag1Cre recombination. n=4 for each cell type; each replicate represents a sample of 50 cells. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

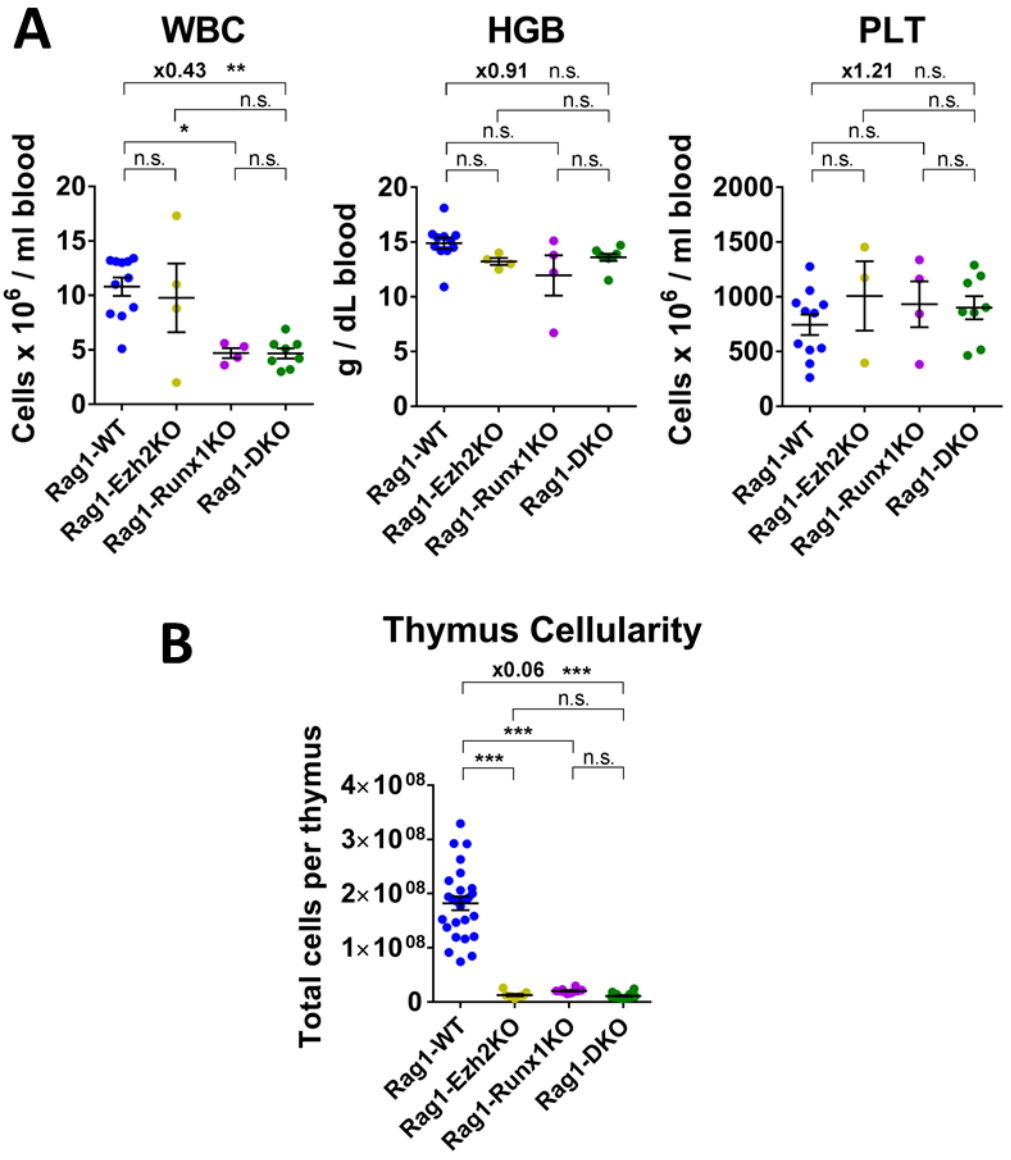


Figure 3.23

(A) Haematological parameters of peripheral blood at 6-8 weeks of age in Rag1-WT (n=11), Rag1-Ezh2KO (n=4), Rag1-Runx1KO (n=4) and Rag1-DKO (n=8) mice. (B) Total thymus cellularity at 6-8 weeks of age in Rag1-WT (n=26), Rag1-Ezh2KO (n=9), Rag1-Runx1KO (n=10) and Rag1-DKO (n=18) mice. Error bars represent standard error of the mean. P-values are by ANOVA followed by multiple comparisons testing; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

Rag1-DKO mice (Figure 3.23A). This reflects the fact that deletion was not targeted to cells of the megakaryocyte lineage using *Rag1Cre*. As in *Mx1Cre* and *Flt3Cre* mice, thymus cellularity was reduced in all *Rag1Cre* knockout genotypes (Figure 3.23B). Analysis of mature lineages in the peripheral blood (Figure 3.24A) and bone marrow (Figure 3.24B) showed overall a less severe phenotype to *Mx1Cre* and *Flt3Cre* mice. As expected, B-cells were severely reduced in all knockout genotypes, while T-cells were reduced in the peripheral blood but not the bone marrow. Unlike Mx1-DKO and Flt3-DKO mice, monocytes and neutrophils were not significantly affected in Rag1-DKO mice in either the bone marrow or peripheral blood (Figure 3.24A-B). $Kit^{+}Mac1^{low}$ myeloid precursors were increased in the bone marrow of Rag1-DKO mice (Figure 3.24B), but markedly less than in Mx1-DKO (Figure 3.02B) or Flt3-DKO mice (Figure 3.13B).

Analysis of the bone marrow LSK stem and progenitor compartment of *Rag1Cre* mice (Figure 3.25A) also demonstrated overt differences from *Mx1Cre* and *Flt3Cre* mice. While MPPs were increased compared to Rag1-WT in Rag1-DKO mice (x2.27), this expansion was much milder than in *Mx1Cre* (x6.12; Figure 3.04) and *Flt3Cre* mice (x10.2; Figure 3.14). This finding can be explained by the fact that *Rag1Cre* becomes active in only a proportion of cells falling into the MPP gate. However, HSCs, which were markedly decreased in Mx1-DKO and Flt3-DKO mice, were in fact slightly increased in Rag1-DKO mice. The finding that HSCs are not reduced in number in Rag1-DKO mice is interesting, since it suggests that the suppression of HSCs seen in Mx1-DKO and Flt3-DKO mice is not mediated by any of the progeny of early lymphoid progenitors. However it is surprising that HSC numbers were increased in Rag1-DKO relative to Rag1-WT mice. This may simply reflect increased numbers of non-HSCs falling into the HSC gate.

Follow-up of Rag1-Ezh2KO, Rag1-Runx1KO and Rag1-DKO mice to 52 weeks of age revealed no significant difference in disease-free survival of knockout genotypes compared to Rag1-WT controls (Figure 3.25B). These results show that targeting of *Ezh2* and *Runx1* deletion to early lymphoid progenitors, rather than HSCs or MPPs, does not result in development of MDS.

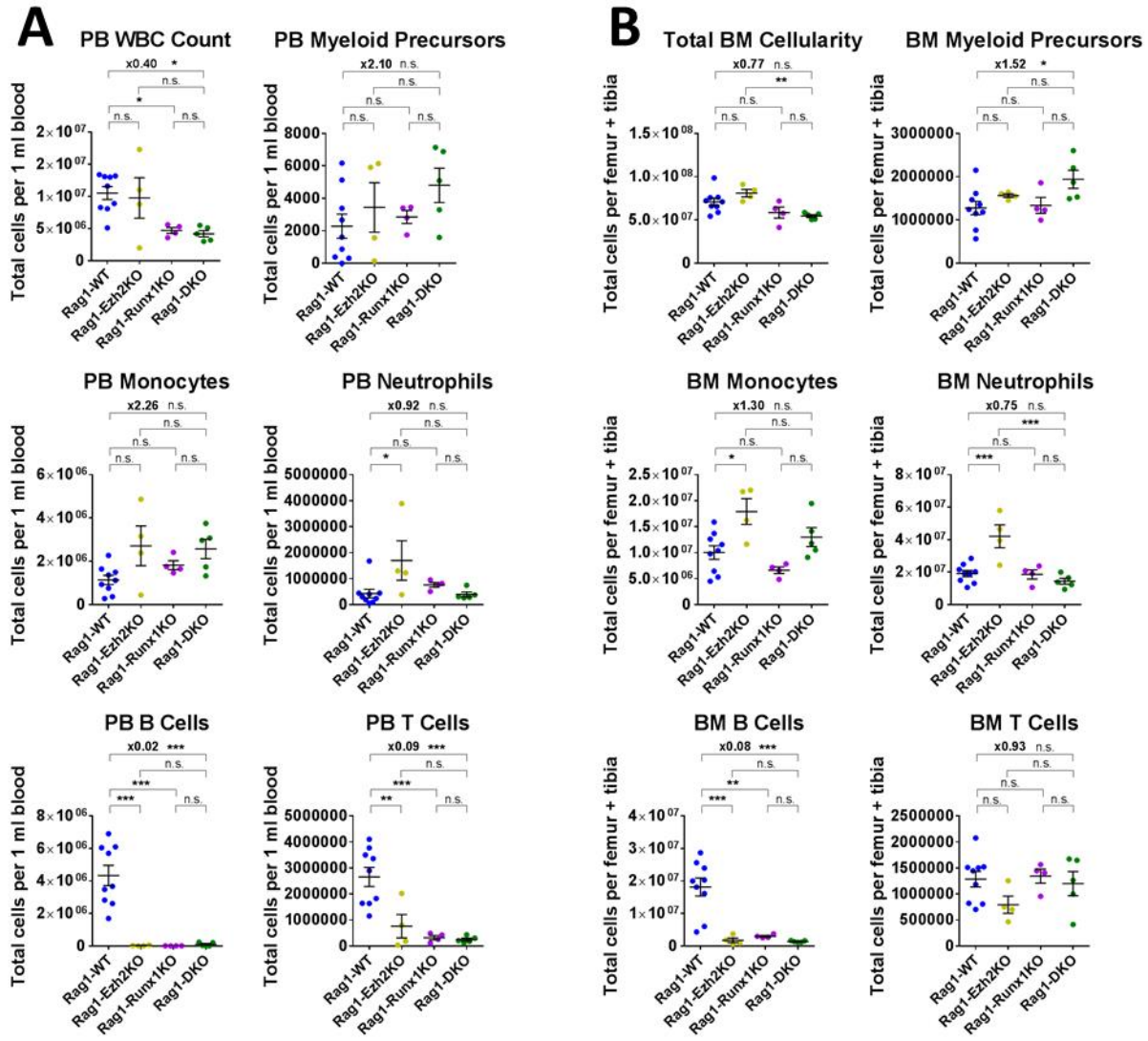


Figure 3.24

(A) Absolute numbers per ml of indicated leukocyte populations in the peripheral blood at 6-8 weeks of age in Rag1-WT (n=9), Rag1-Ezh2KO (n=4), Rag1-Runx1KO (n=4) and Rag1-DKO (n=5) mice. (B) Absolute numbers per femur and tibia of indicated bone marrow populations at 6-8 weeks of age in Rag1-WT (n=9), Rag1-Ezh2KO (n=4), Rag1-Runx1KO (n=4) and Rag1-DKO (n=5) mice. B-cells $CD19^{+}$, T-cells $CD4^{+}$ or $CD8^{+}$, monocytes $Mac1^{+}Gr1^{-}$, neutrophils $Mac1^{+}Gr1^{+}$, myeloid precursors $Kit^{+}Mac1^{low}$. Error bars represent standard error of the mean. P-values are by ANOVA followed by multiple comparisons testing; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

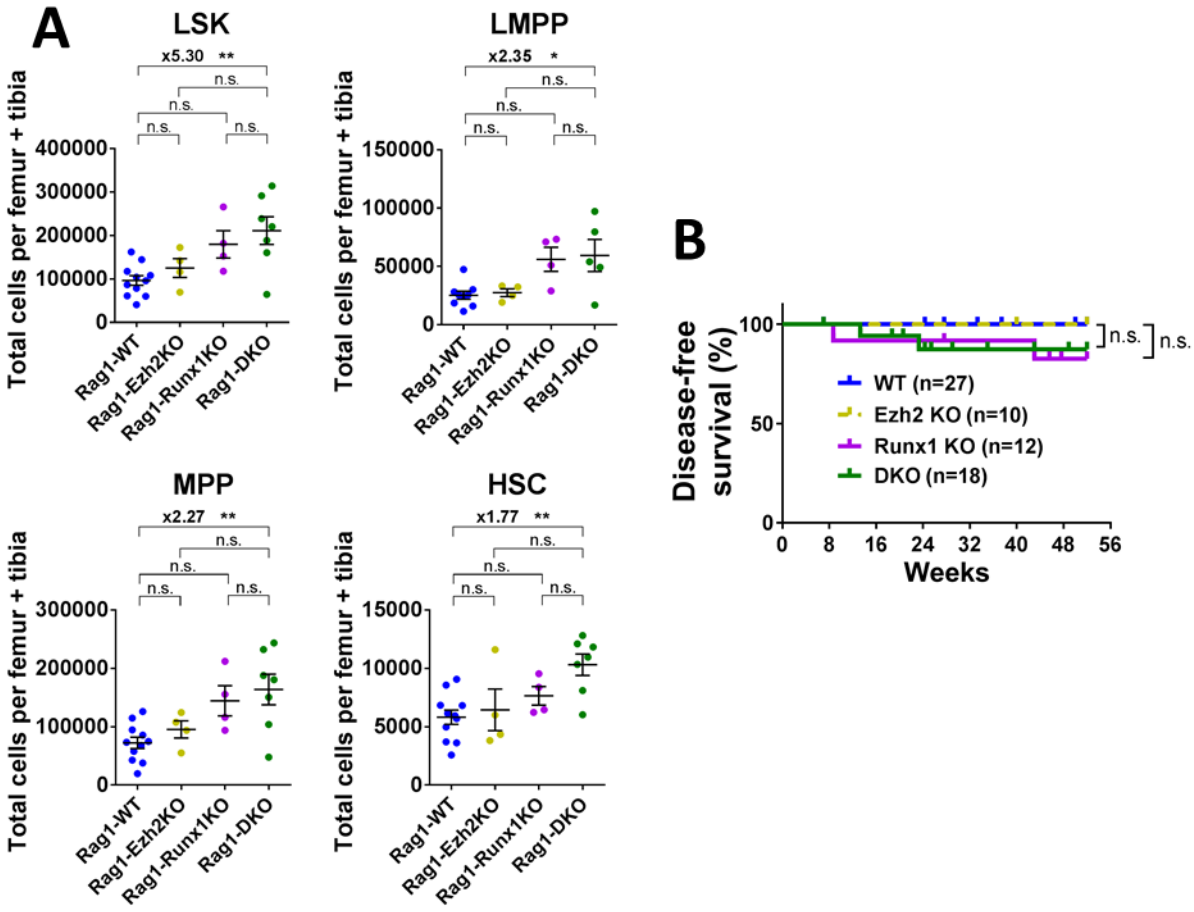


Figure 3.25

(A) Absolute numbers per femur and tibia of indicated bone marrow stem and progenitor populations at 6-8 weeks of age in Rag1-WT (n=11), Rag1-Ezh2KO (n=4), Rag1-Runx1KO (n=4) and Rag1-DKO (n=7) mice. (B) Kaplan-Meier disease-free survival curves for Rag1-WT, Rag1-Ezh2KO, Rag1-Runx1KO and Rag1-DKO mice. Mice found dead without analysis or culled showing > 25 x 10^6 leukocytes per ml blood and/or < 6.5 g haemoglobin per dl blood were counted as events. Mice culled showing leukocyte and haemoglobin counts outside these constraints were censored. Error bars represent standard error of the mean. P-values are by ANOVA followed by multiple comparisons testing (A) or logrank test (B); n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

3: Discussion

3.1: Development of myelodysplastic syndrome upon inactivation of *Ezh2* and *Runx1* in HSCs

Inactivating mutations of *EZH2* and *RUNX1* are significantly associated in MDS as well as ETP-ALL (Papaemmanuil et al., 2013; Zhang et al., 2012b). We therefore investigated the impact of combined inactivation of *Ezh2* and *Runx1* in the haematopoietic system of mice. In Mx1-DKO mice at six weeks of age, gross phenotypic changes were consistent with previous reports describing inactivation of either gene separately (Ichikawa et al., 2004; Su et al., 2003; Su et al., 2005). These included loss of mature B and T-cells, reduced thymus cellularity, thrombocytopenia and mild anaemia. One of the most striking phenotypic changes specific to double knockout mice was an increase in $Kit^{+}Mac1^{low}$ myeloid precursors cells in the bone marrow, which were also present in the peripheral blood where these cells are not normally present. This may be related to the development of MDS that occurs subsequently, since accumulation of immature myeloid precursors (with increased apoptosis and reduced differentiation potential) is a hallmark of MDS (Sperling et al., 2017).

The other major phenotypic change present in Mx1-DKO mice at 6-8 weeks was an expansion of MPPs combined with a suppression of HSCs. Analysis of single knockout mice provided insights into how inactivation of *Ezh2* and *Runx1* collaborate to produce this phenotype. Mx1-Ezh2KO mice showed decreased HSC numbers but no MPP expansion, while Mx1-Runx1KO mice showed expanded MPPs but normal HSC numbers. This finding may help to explain why inactivating mutations of *EZH2* and *RUNX1* are significantly associated in MDS.

Our Mx1-DKO mouse model, in which all HSCs within the animal are simultaneously targeted for inactivation of *Ezh2* and *Runx1*, contrasts with human MDS which originates from a single HSC. This HSC, through acquisition of one or more mutations, clonally expands at the expense of other HSCs (Cazzola et al., 2013). Broadly, there are two mechanisms by which mutant HSCs could displace

normal HSCs during this process. Firstly, the mutant HSC clone may simply self-renew faster than normal HSC clones, so that eventually there are more mutant HSCs and their progenitors competing for the limited amount of space within the bone marrow. Or secondly, mutant HSCs or their progeny may actively suppress normal HSC function, for example by "remodelling" bone marrow niches to make them more supportive of the mutant HSC clone (Schepers et al., 2013). Our mouse model provides insights into how *Ezh2* and *Runx1* inactivation might be involved in both of these mechanisms. As few as 742 transplanted Mx1-DKO DP cells were able to contribute to almost 100% of mature myeloid cells in the peripheral blood within only 8 weeks. This is suggestive of very rapid proliferation of the mutant clone at the expense of normal haematopoiesis. We also observed a clear extrinsic suppression of wild-type phenotypic HSC numbers by Mx1-DKO and Flt3-DKO bone marrow cells in all of our transplantation experiments. We did not explore the molecular mechanisms by which this suppression of HSC numbers occurs, or whether the remaining wild-type HSCs are functionally impaired, but this is a very interesting question for future research. This could provide insights into the mechanisms by which MDS clones exert a clonal advantage over normal HSCs.

3.2: Propagation of *Ezh2* and *Runx1*-inactivated MDS by multipotent progenitor cells

When we targeted inactivation of *Ezh2* and *Runx1* to MPPs using *Flt3Cre*, we observed development of a similar MDS-like phenotype with an identical latency to Mx1-DKO mice. According to the hypothesis that MDS mutations occur sequentially specifically in HSCs (Woll et al., 2014), it was possible that Flt3-DKO mice might not have developed an MDS phenotype if HSCs were essential for development of clonal dominance. This apparent contradiction could be explained by the fact that in humans, haematopoietic stem cells are less well defined than in mice. The mouse $Lin^-Kit^+Sca^+CD48^-CD150^+$ compartment contains around 50% true stem cells (Kiel et al., 2005), whereas only around 5% of the human stem cell compartment, $Lin^-CD34^+CD38^-CD90^+CD45RA^-$, are true stem cells (Majeti

et al., 2007; Notta et al., 2011). Many of the cells defined as HSCs in human studies may therefore in fact correspond to the MPPs targeted by *Flt3Cre* in our mouse model. If so, our data suggest that human MDS could be propagated by progenitor cells within the phenotypically defined HSC compartment, which have gained aberrant self-renewal through mutation. This is potentially an important conceptual advance in the field, but further work will be required to investigate these observations in further depth. This could include characterisation at the molecular level, for example using RNA-sequencing, of the disease-propagating Mx1-DKO and Flt3-DKO MPPs. Comparison of these cells with their WT counterparts, as well as with HSCs, will allow exploration of the mechanisms underlying gain of self-renewal by these cells.

Our model demonstrates that acquisition of MDS propagating potential can be conferred by inactivation of only two genes, *Ezh2* and *Runx1*. In humans, loss of both of these through mutation would nonetheless be unlikely to occur within a single non-self-renewing progenitor cell. It is more likely that an HSC clone would act as the self-renewing pool for mutations to accumulate, even if the malignant effects of those mutations are manifested only in downstream progenitors. Our *Mx1Cre* transplantation data indicate that non-HSCs within the LSK stem/progenitor compartment are more readily able to propagate MDS than HSCs, perhaps consistent with the increased proliferative activity of the normal counterparts of these cells. These findings are therefore consistent with HSCs acting as the initiating population of human MDS, while MPPs possess the most potential to propagate the disease.

3.3: Inactivation of *Ezh2* and *Runx1* in early lymphoid progenitors cannot initiate MDS

Inactivation of *Ezh2* and *Runx1* in early lymphoid progenitors rather than HSCs using *Rag1Cre* induced a less severe overall phenotype at 6-8 weeks of age. Unlike Mx1-DKO mice, Rag1-DKO mice showed a complete absence of thrombocytopenia, consistent with lack of *Rag1Cre* activation in any

cells of the megakaryocyte or erythroid lineages (Boiers et al., 2013). $Kit^{+}Mac1^{low}$ myeloid precursors were increased but not as significantly as in Mx1-DKO mice, probably reflecting the only partial contribution (2.2%) of *Rag1Cre*-expressing early lymphoid progenitors to the myeloid lineage (Boiers et al., 2013). Follow-up of Rag1-DKO mice demonstrated that inactivation of *Ezh2* and *Runx1* in early lymphoid progenitors does not lead to development of MDS. This is consistent with findings from patient samples that MDS stem cell potential resides only in cells displaying a normal HSC/MPP surface phenotype, and not more restricted progenitor populations such as granulocyte-macrophage progenitors (GMPs) or megakaryocyte-erythroid progenitors (MEPs) (Woll et al., 2014). When we introduced inactivating mutations of *Ezh2* and *Runx1* to early lymphoid progenitors – a lympho-myeloid restricted progenitor population – these cells were unable to initiate MDS.

In summary, we have investigated the impact of targeting inactivation of *Ezh2* and *Runx1* to different haematopoietic stem and progenitor populations. Inactivation in HSCs or MPPs but not early lymphoid progenitors produces an expansion of multipotent progenitor populations combined with suppression of phenotypic HSCs. This suppression of HSCs could also be performed extrinsically by *Ezh2* and *Runx1*-inactivated bone marrow cells upon wild-type HSCs. *Ezh2* and *Runx1* inactivation in HSCs or MPPs, but not early lymphoid progenitors, leads to the development of MDS. These findings provide insight into the interaction of *EZH2* and *RUNX1* inactivating mutations in MDS, and the haematopoietic cell populations responsible for propagation of the disease in human patients. We anticipate that this will prove to be a powerful, tractable model system to further explore mechanisms of clonal dominance in MDS.

We next wished to use our mouse models to explore the observation that inactivating mutations of *EZH2* and *RUNX1* are significantly associated not only in MDS, but also in ETP-ALL. These leukaemias are hypothesised to originate within ETPs, the most primitive thymocyte progenitor population. The development of MDS in Mx1-DKO and Flt3-DKO mice precludes modelling of leukaemia, however

Rag1-DKO mice showed efficient deletion of *Ezh2* and *Runx1* in ETPs while not developing MDS. We therefore decided to investigate ETP-ALL using *Rag1Cre* mediated inactivation of *Ezh2* and *Runx1*, as described in Chapter 4.

Chapter 4: Ezh2 and Runx1 Inactivating Mutations Collaborate to Initiate Lympho-Myeloid Leukaemia in Early Thymic Progenitors

1: Introduction

We established in Chapter 3 that targeting of *Ezh2* and *Runx1* inactivation to HSCs, MPPs or early lymphoid progenitors results in perturbed haematopoiesis, but that only targeting to HSCs or MPPs leads to development of an MDS phenotype. Because *EZH2* and *RUNX1* inactivating mutations are significantly associated in ETP-ALL (Zhang et al., 2012b), as well as MDS, we decided to use *Rag1Cre* to study the impact of *Ezh2* and *Runx1* inactivation in early lymphoid progenitors without inducing MDS.

Rag1Cre becomes active in the bone marrow, in lympho-myeloid progenitor cells that retain the potential to generate myeloid progeny (Boiers et al., 2013; McCormack et al., 2003). It is therefore an ideal tool for investigating the role of the lympho-myeloid pathway in leukaemia. ETP-ALL is thought to represent a leukaemia originating from the ETP, a lympho-myeloid progenitor population (Coustan-Smith et al., 2009). Both *EZH2* and *RUNX1* show biallelic inactivation in a proportion of ETP-ALL patients (Zhang et al., 2012b). In addition, both genes are required for the maturation of T-cells in mice (Growney et al., 2005; Su et al., 2005), demonstrating involvement in transcriptional regulation of differentiation in this haematopoietic lineage.

Our aims for this work were firstly to determine the impact of *Ezh2* and *Runx1* inactivation on ETPs, and to investigate how this impact relates to human ETP-ALL. We then aimed to generate a mouse model of ETP leukaemia, by combining other relevant mutations such as in signalling pathways, and determine experimentally whether ETPs are able to become an LSC population. Finally we aimed to use this ETP leukaemia model to investigate novel therapeutic possibilities for ETP-ALL, which shows resistance to conventional therapy (Coustan-Smith et al., 2009; Jain et al., 2016).

In this Chapter, *Rag1Cre* is the only *Cre* line used for inactivation of *Ezh2* and *Runx1*. Therefore, mouse genotypes that were referred to as Rag1-WT, Rag1-Ezh2KO, Rag1-Runx1KO and Rag1-DKO in Chapter 3 are in this Chapter referred to simply as WT, Ezh2-KO, Runx1-KO and DKO respectively.

2: Results

2.1: Inactivation of *Ezh2* and *Runx1* in early lymphoid progenitors causes expansion of ETPs

We generated mice with deletion of *Ezh2* and/or *Runx1* targeted to early lymphoid progenitors using *Rag1Cre*. This produced four experimental genotypes (Figure 4.01A): WT ($Ezh2^{fl/fl}Runx1^{fl/fl}Rag1^{+/+}$), Ezh2-KO ($Ezh2^{fl/fl}Runx1^{+/+}Rag1^{Cre/+}$), Runx1-KO ($Ezh2^{+/+}Runx1^{fl/fl}Rag1^{Cre/+}$) and DKO ($Ezh2^{fl/fl}Runx1^{fl/fl}Rag1^{Cre/+}$).

We previously analysed the peripheral blood and bone marrow phenotype of *Rag1Cre*-knockout mice at 6-8 weeks of age (Figures 3.23A and 3.24). We also observed a decreased in overall thymus cellularity in all knockout genotypes (Figure 3.23B). Given that *EZH2* and *RUNX1* inactivating mutations are significantly associated in ETP-ALL, which is thought to originate from ETPs, we decided to investigate the effects of their inactivation on early thymocyte development. Example FACS plots for all four genotypes are shown in Figure 4.01B.

We found that despite the reduced overall thymus cellularity, DKO mice showed a marked expansion of stringently defined ($Lineage^{-}CD4^{-}CD8^{-}CD44^{+}CD25^{-}Kit^{+}Flt3^{+}$) ETPs (Figure 4.02A). This expansion was specific to DKO mice and not present in either of the single knockouts. It has previously been shown that inactivation of *Ezh2* in HSCs results in a block in thymocyte maturation at the DN3-DN4 stage (Su et al., 2005), and that of Runx1 at the DN2-DN3 stage (Growney et al., 2005). We confirmed these observations in *Rag1-Cre* Ezh2-KO and Runx1-KO mice (Figure 4.02B). In DKO mice, DN1 and

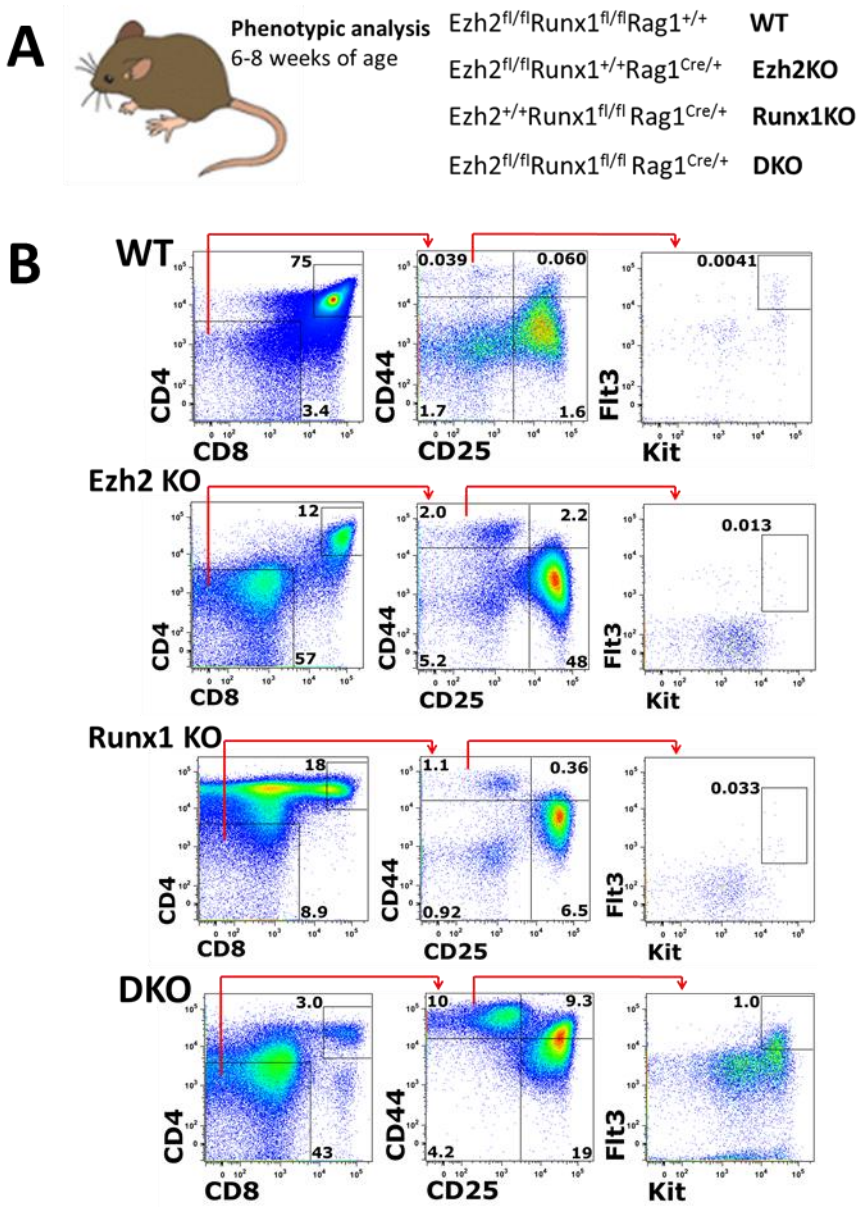


Figure 4.01

(A) Experimental setup and genotypes of experimental mice for Ezh2 and/or Runx1 inactivation in early lymphoid progenitors using Rag1Cre. (B) FACS gating strategies and representative plots for WT, Ezh2-KO, Runx1-KO and DKO thymuses. Figures indicate mean percentage of total thymocytes falling into each gate across all experiments.

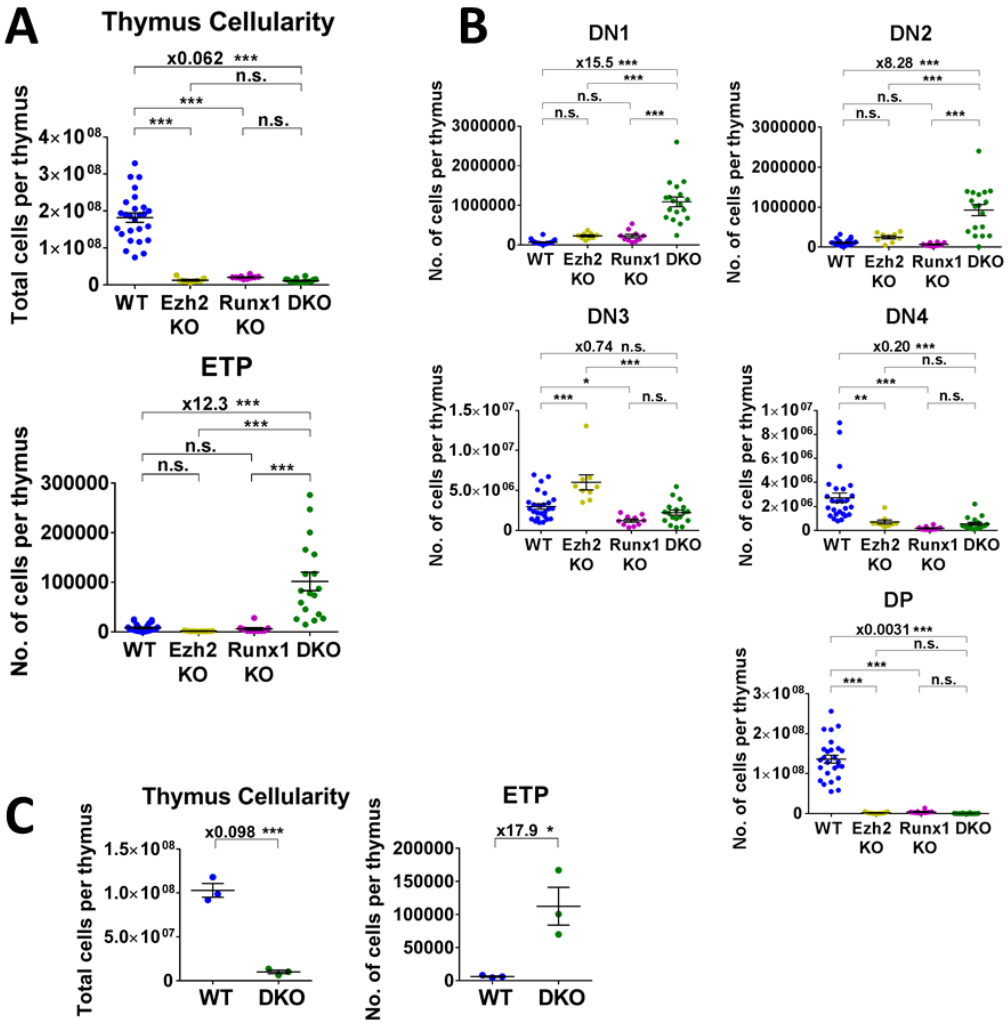


Figure 4.02

(A) Total thymus cellularity and absolute numbers of ETPs per thymus at 6-8 weeks of age for WT (n=26), Ezh2-KO (n=9), Runx1-KO (n=10) and DKO (n=18) mice. (B) Absolute numbers of indicated thymocyte populations per thymus at 6-8 weeks of age for WT (n=26), Ezh2-KO (n=9), Runx1-KO (n=10) and DKO (n=18) mice. (C) Total thymus cellularity and absolute numbers of ETPs per thymus at 6 months of age for WT (n=3) and DKO (n=3) mice. ETP $Lin^-CD4^-CD8^-CD44^+CD25^-Kit^+Flt3^+$, DN1 $Lin^-CD4^-CD8^-CD44^+CD25^-$, DN2 $Lin^-CD4^-CD8^-CD44^+CD25^+$, DN3 $Lin^-CD4^-CD8^-CD44^-CD25^+$, DN4 $Lin^-CD4^-CD8^-CD44^-CD25^-$, DP $Lin^-CD4^+CD8^+$. Error bars represent standard error of the mean. P-values are by ANOVA followed by multiple comparisons testing (A-B) or T-test (C); n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

DN2 cells were expanded relative to WT, but less so than ETPs. DN3 and further populations (DN4 and DP) were reduced, consistent with a block in maturation at the DN2-DN3 stage.

We next analysed the thymuses of WT and DKO mice at 6 months of age, to determine whether the expansion of ETPs we observed at 6 weeks was still present. In wild-type mice (as well as humans), ETPs are known to reduce in number with age along with overall thymus cellularity (Linton and Dorshkind, 2004). This was indeed the case for our WT control mice, but absolute numbers of ETPs in DKO mice remained at similar levels compared to 6 weeks of age, making their proportional increase compared to WT even greater (Figure 4.02C).

These data indicate that inactivation of *Ezh2* and *Runx1* collaborate to induce a selective expansion of ETPs which does not diminish with age, combined with a block in thymocyte maturation. This is an interesting observation in light of the fact that *EZH2* and *RUNX1* inactivating mutations are positively correlated in ETP-ALL, which is hypothesised to originate from ETPs. Since there is in fact no experimental evidence that ETPs are able to give rise to leukaemia, we decided to use our mouse model to explore this possibility. We started by investigating gene expression in the expanded ETP population of DKO mice.

2.2: Expanded DKO ETPs display gene expression patterns characteristic of ETP-ALL

We performed RNA-sequencing on bulk samples of 100 ETPs from four WT and four DKO mice at 6-8 weeks of age. Semi-supervised hierarchical clustering (using the top 3000 most variable genes) grouped the eight samples by genotype as expected (Figure 4.03A). Analysing significantly differentially expressed genes, we found that more were upregulated (238) in DKO relative to WT than downregulated (68). This is in keeping with the role of *Ezh2* as a transcriptional repressor and could be visualised clearly with supervised hierarchical clustering using only the 50 most significantly differentially expressed genes (Figure 4.03B). Gene set enrichment analysis (GSEA) revealed strong

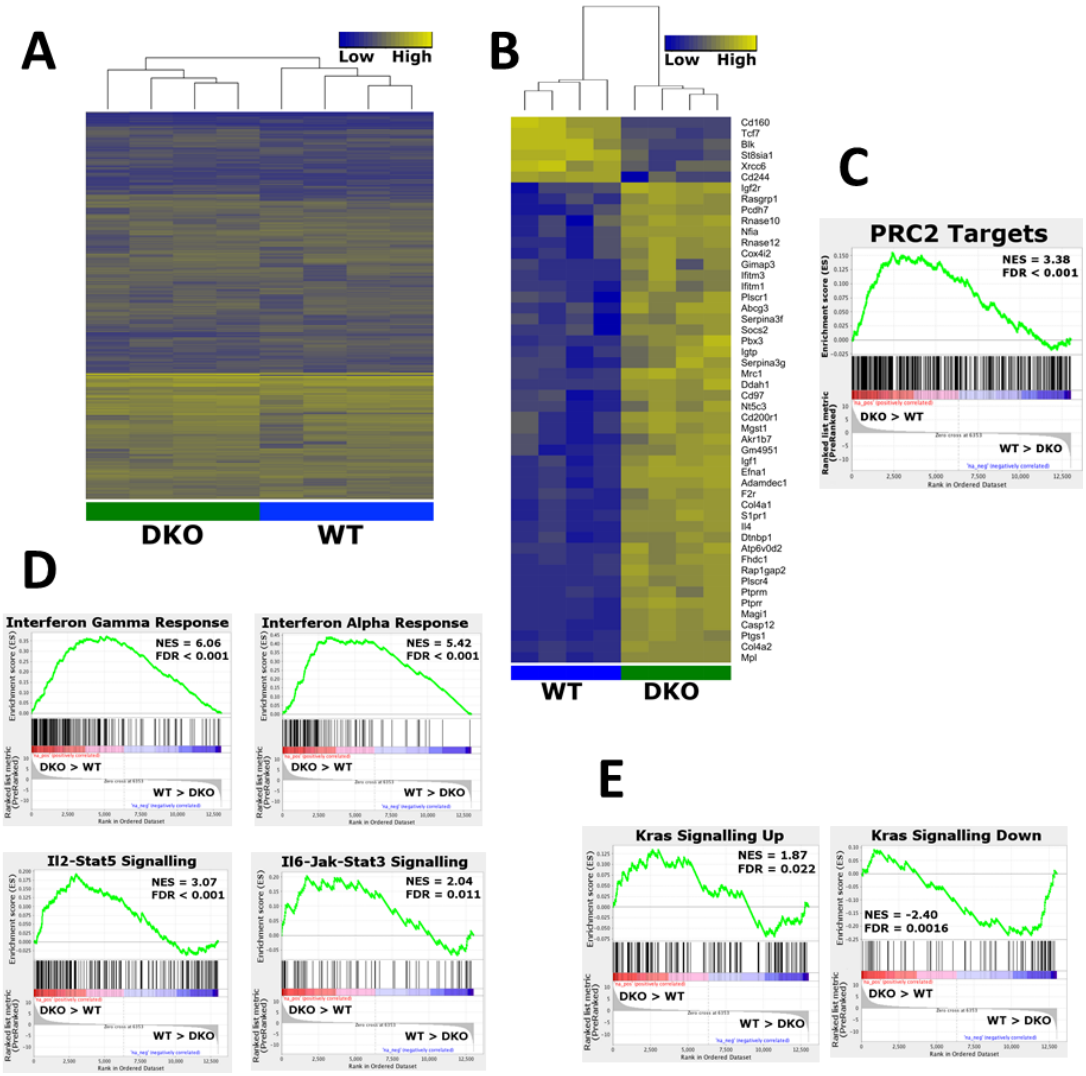


Figure 4.03

(A) Hierarchical clustering and heatmap showing expression levels, scaled per gene, of the 3000 most variable genes between WT and DKO ETPs (WT n=4, DKO n=4). Each replicate represents a sample of 100 purified ETPs from an individual mouse. (B) Hierarchical clustering and heatmap showing expression levels, scaled per gene, of the 50 most significantly differentially expressed genes between WT and DKO ETPs (WT n=4, DKO n=4). Each replicate represents a sample of 100 purified ETPs from an individual mouse. (C) GSEA comparing WT and DKO ETPs for PRC2 target genes. (D) GSEA comparing WT and DKO ETPs for target genes associated with the indicated signalling pathways. (E) GSEA comparing WT and DKO ETPs for genes upregulated (left) or downregulated (right) by Kras signalling. Data analysis was performed by Nikolaos Barkas.

upregulation of a set of genes associated with repression by PRC2 (Figure 4.03C), confirming the importance of loss of *Ezh2*.

We next used the Hallmark GSEA database (Liberzon et al., 2015) to investigate the transcriptional programmes dysregulated in DKO ETPs. This revealed strong upregulation of a number of signalling pathway-associated gene sets in DKO relative to WT ETPs (Figure 4.03D). These included interferon alpha and interferon alpha, IL2-Stat5 and IL6-Jak-Stat3 signalling. Genes associated with upregulated Kras signalling were relatively enriched in DKO ETPs, while genes associated with downregulated Kras signalling were conversely enriched in WT ETPs (Figure 4.03E). Together these findings indicate an aberrant activation of signalling pathways in the expanded DKO ETP population. Upregulated signalling and mutations in signalling pathway components are a key feature of ETP-ALL (Zhang et al., 2012b).

Another key feature of ETP-ALL is upregulated HSC and myeloid lineage-associated gene expression (Zhang et al., 2012b). This was mirrored in DKO relative to WT ETPs (Figure 4.04A). Conversely, gene expression associated with the megakaryocyte and erythroid lineages was not significantly changed (Figure 4.04B). This indicates that ETPs, which normally do not show any megakaryocyte/erythroid potential, were not aberrantly gaining expression of all lineage pathways indiscriminately upon loss of *Ezh2* and *Runx1*. Additionally, we did not observe upregulation of Notch signalling associated gene expression in DKO ETPs (Figure 4.04C). Upregulated Notch pathway signalling via mutations in *Notch* itself is a common feature of many T-ALLs, but not ETP-ALL (Zhang et al., 2012b).

We then used multiplex qPCR to verify some of the gene expression changes observed in ETPs in our RNA-sequencing data, and to further investigate gene expression changes in Ezh2-KO and Runx1-KO ETPs. We were able to validate the expression patterns seen for a number of HSC, myeloid and lymphoid-associated genes (Figure 4.04D). The most significantly upregulated gene in DKO ETPs was *Mpl* (Figures 4.04D and 4.03B), which has also been shown to be upregulated in ETP-ALLs compared to non-ETP T-ALLs (Zhang et al., 2012b). MPL is the surface receptor for thrombopoietin and is thus

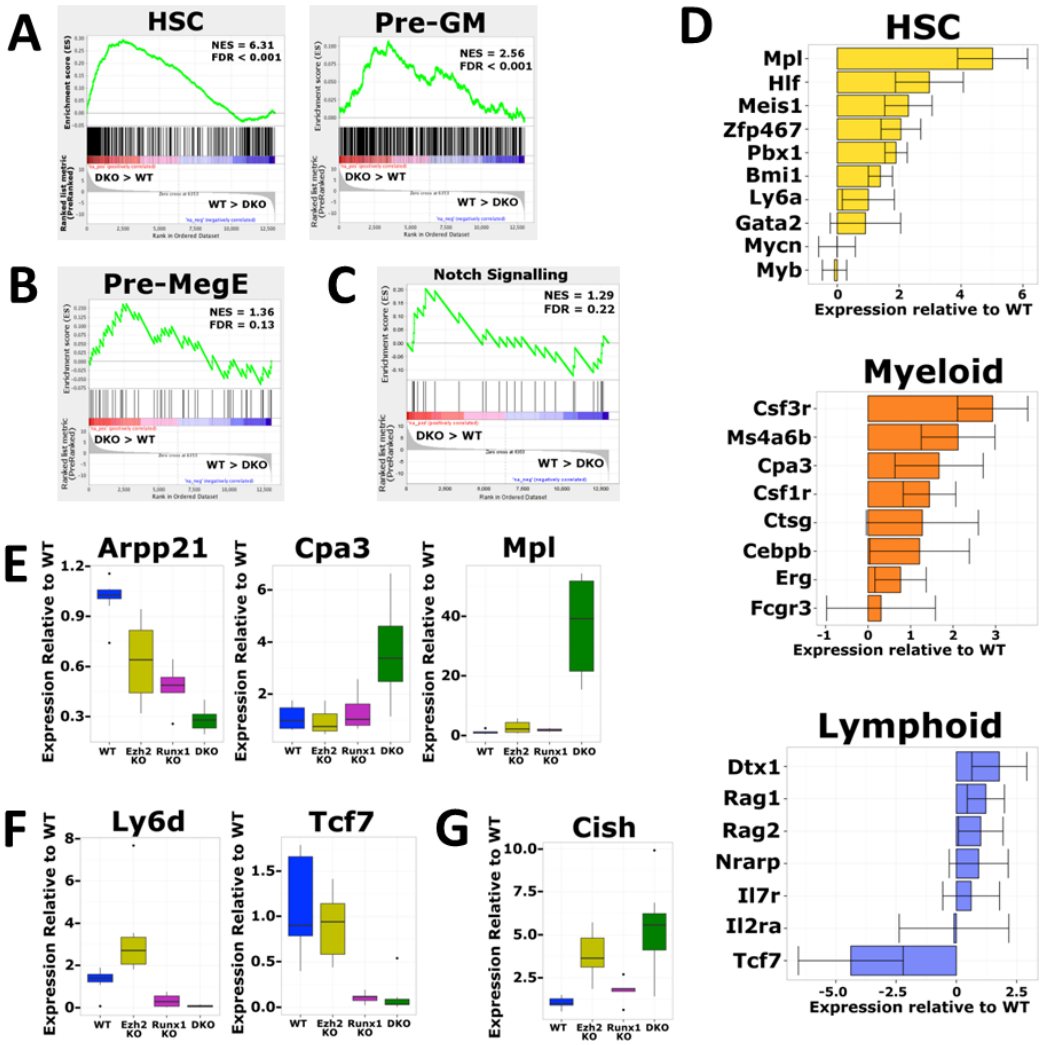


Figure 4.04

(A) GSEA comparing WT and DKO ETPs for HSC (left) and pre-GM (myeloid; right) lineage-associated genes. (B) GSEA comparing WT and DKO ETPs for Pre-MegE (megakaryocyte/erythroid) lineage-associated genes. (C) GSEA comparing WT and DKO ETPs for Notch signalling pathway associated genes. (D) Relative expression level of HSC, myeloid and lymphoid lineage-associated genes in DKO ETPs compared to WT ETPs (log2 scale) as determined by qPCR. (E-G) Relative expression level of indicated genes in WT, Ezh2-KO, Runx1-KO and DKO ETPs shown relative to the gene expression level in WT ETPs as determined by qPCR (WT n=8, Ezh2-KO n=6, Runx1-KO n=5, DKO n=7; each replicate represents a sample of 100 purified ETPs from an individual mouse). Error bars represent standard deviation (D) or min and max values (E-G). Data analysis was performed by Nikolaos Barkas.

critical in megakaryopoiesis, but it is also expressed by HSCs and is important in their maintenance and regulation of quiescence (Qian et al., 2007; Yoshihara et al., 2007).

Lymphoid gene expression was largely unaffected in DKO ETPs (Figure 4.04D). A notable exception to this was *Tcf7*, which was strongly downregulated in DKO relative to WT ETPs. *Tcf7* (also known as *TCF-1*) is a transcription factor important in normal thymocyte differentiation (Schilham et al., 1998). Its expression is however markedly lower in ETP-ALL relative to non-ETP T-ALL (Zhang et al., 2012b), and germline knockout in mice has been shown to induce development of lymphomas showing some transcriptional features of ETP-ALL (Yu et al., 2012).

We observed various patterns of gene expression in Ezh2-KO and Runx1-KO ETPs relative to DKO ETPs (Figures 4.04E-G). For some genes, such as *Arpp21*, *Cpa3* and *Mpl*, knockout of either *Ezh2* or *Runx1* alone had only minor effects on gene expression, while the combination of both had a large effect (Figure 4.04E). For other genes, the major contributor to expression in DKO ETPs appeared to be either *Runx1* (for example *Ly6d* and *Tcf7*; Figure 4.04F) or *Ezh2* (for example *Cish*; Figure 4.04G). *Cish* expression is induced by cytokine signalling, and is involved in downregulation of signalling as part of a negative feedback loop (Matsumoto et al., 1997). It is very strongly overexpressed in ETP-ALL compared to non-ETP T-ALL (Zhang et al., 2012b).

These examples illustrate the complex interplay between these two transcription factors and reflect the fact that the phenotypic changes seen in DKO mice are more than simply additive combinations of the phenotypes seen in the single knockouts.

2.3: HSC and myeloid genes are co-expressed with lymphoid genes within single DKO ETPs

The above RNA-sequencing and qPCR analysis was performed at the bulk level, using RNA extracted from a collection of 50-100 cells per sample. This showed that HSC and myeloid associated gene transcription was upregulated on average in DKO ETPs compared to WT ETPs and that lymphoid

expression was on average largely unaffected. This does not prove however that HSC and myeloid transcription is upregulated in the same cells that also retain normal lymphoid expression. The upregulation of HSC and myeloid transcription observed in DKO ETPs could rather reflect contamination of the ETP surface phenotype by a myeloid progenitor population in DKO thymuses. To determine whether this was the case, we decided to perform qPCR analysis at the single cell level.

We performed single cell qPCR analysis on WT and DKO ETPs for a panel of HSC, myeloid, lymphoid and signalling associated genes. Visualisation of these data in a heatmap showed that DKO ETPs displayed upregulation of HSC and myeloid-associated gene expression, but no loss of lymphoid-associated expression in the same cells (Figure 4.05A). We then determined for each cell the number of these genes for which expression was detected. This demonstrated that individual DKO ETPs showed expression of increased numbers of both HSC and myeloid genes compared to WT ETPs (Figure 4.05B). However, expression of increased numbers of HSC (Figure 4.05C) or myeloid (Figure 4.05D) genes by DKO ETPs corresponded with expression of equivalent numbers of lymphoid genes to WT ETPs. This supports our hypothesis that the expanded ETP population in DKO mice represents a *bona fide* lymphoid progenitor population which is considerably expanded in number and displays increased HSC and myeloid transcriptional priming.

2.4: Combination of *Flt3-ITD* with inactivation of *Ezh2* and *Runx1* generates acute lympho-myeloid leukaemia

In addition to mutations in transcriptional and epigenetic regulators such as *RUNX1* and *EZH2*, ETP-ALLs show a high frequency of mutations in signalling pathway components (Zhang et al., 2012b). The RAS pathway is a frequent target of these mutations, for example via activating mutations in the cytokine receptor *FLT3*. We therefore decided to combine the *Flt3-ITD* mutation with inactivation of *Ezh2* and *Runx1* targeted to early lymphoid progenitors, to determine whether this would be

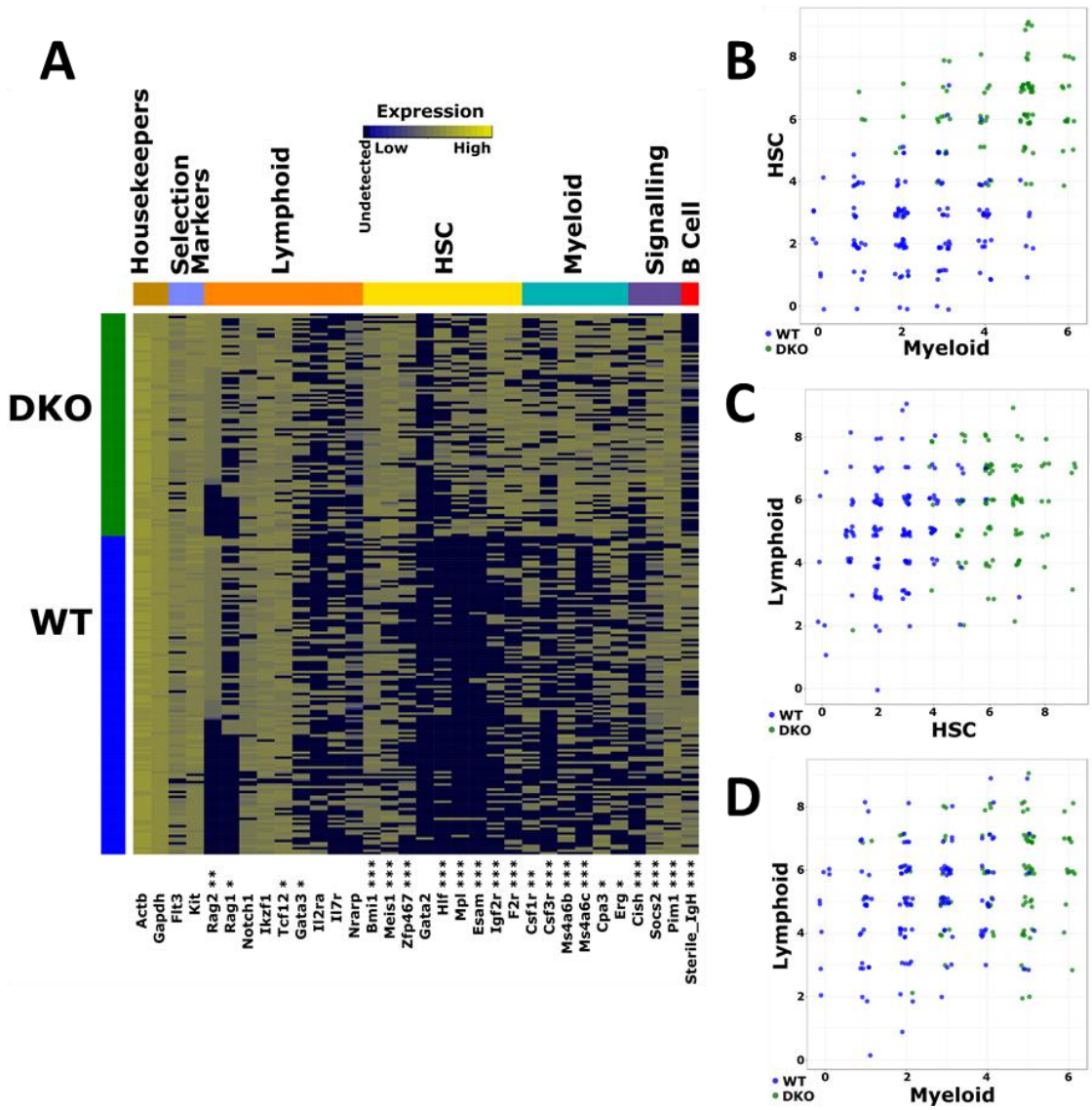


Figure 4.05

(A) Heatmap showing co-expression of indicated genes within single WT (n=129) and DKO (n=91) ETPs, as determined by single-cell qPCR. Expression levels shown are scaled per gene. Single ETPs were collected from 12 WT mice and 9 DKO mice. FDR q-values are by T-test based on expression levels; * < 0.05, ** < 0.01, *** < 0.001. (B-D) Scatter plots showing level of co-expression of myeloid with HSC (B), HSC with lymphoid (C) and myeloid with lymphoid (D) genes within single WT (n=129) and DKO (n=91) ETPs. Each dot represents a single cell, for which are plotted the number of genes showing detectable expression from the panel of HSC, myeloid and lymphoid lineage-associated genes shown in heatmap (A). Data analysis was performed by Nikolaos Barkas.

sufficient to induce leukaemia. This produced ITD ($Ezh2^{fl/fl}Runx1^{fl/fl}Rag1^{+/+}Flt3^{ITD/+}$) and DKOITD ($Ezh2^{fl/fl}Runx1^{fl/fl}Rag1^{Cre/+}Flt3^{ITD/+}$) experimental mice (Figure 4.06A).

We next performed phenotypic analysis of WT, DKO, ITD and DKOITD mice at 6-8 weeks of age. Peripheral blood showed significantly increased leukocyte numbers, reduced haemoglobin and reduced platelet numbers in DKOITD mice compared to all other genotypes (Figure 4.06B). The expanded leukocytes mainly showed Mac1 expression (Figure 4.07A). DKOITD bone marrow cellularity was not significantly changed (Figure 4.07B), while spleens were significantly larger than in other genotypes (Figure 4.07C). Lineage analysis of the bone marrow (Figure 4.07D) revealed increased $Kit^{+}Mac1^{low}$ myeloid precursors and $Mac1^{+}Gr1^{-}$ myeloid cells, but decreased mature $Mac1^{+}Gr1^{+}$ neutrophils compared to WT, DKO and ITD mice. As in DKO mice, $CD19^{+}$ B Cells were markedly reduced in number compared to WT or ITD, while $CD4^{+}$/$CD8^{+}$ T cells numbers were highly variable. $Ter119^{+}$ erythroid precursors were very strongly reduced in number compared to WT and ITD mice (DKO mice were not analysed). Together, these data are indicative of development of leukaemia in DKOITD but not DKO or ITD mice as early as 6-8 weeks of age, displaying a mainly $Mac1^{+}$ surface phenotype.

Upon follow-up, we found that DKOITD mice showed markedly reduced survival compared to any of the other genotypes (median 9.3 weeks; Figure 4.08A). Analysis of sick mice revealed profound leukocytosis, anaemia and thrombocytopenia (Figure 4.08B). Morphological analysis of peripheral blood leukocytes demonstrated a near total absence of lymphoid cells coupled with an increase in immature blast cells and $Mac1^{+}$ precursors (Figure 4.09A; images of both peripheral blood smears and bone marrow cytospins are shown Figure 4.09B). These data indicate that DKOITD mice succumb to an aggressive acute leukaemia at a young age and with full penetrance.

Despite the expression of surface Mac1 by both peripheral blood leukocytes (Figure 4.07A) and bone marrow cells (Figure 4.07D), immunohistochemical analysis of the bone marrow revealed expression of cytoplasmic CD3 and TdT in DKOITD but not WT mice (Figure 4.09C). FACS analysis showed that

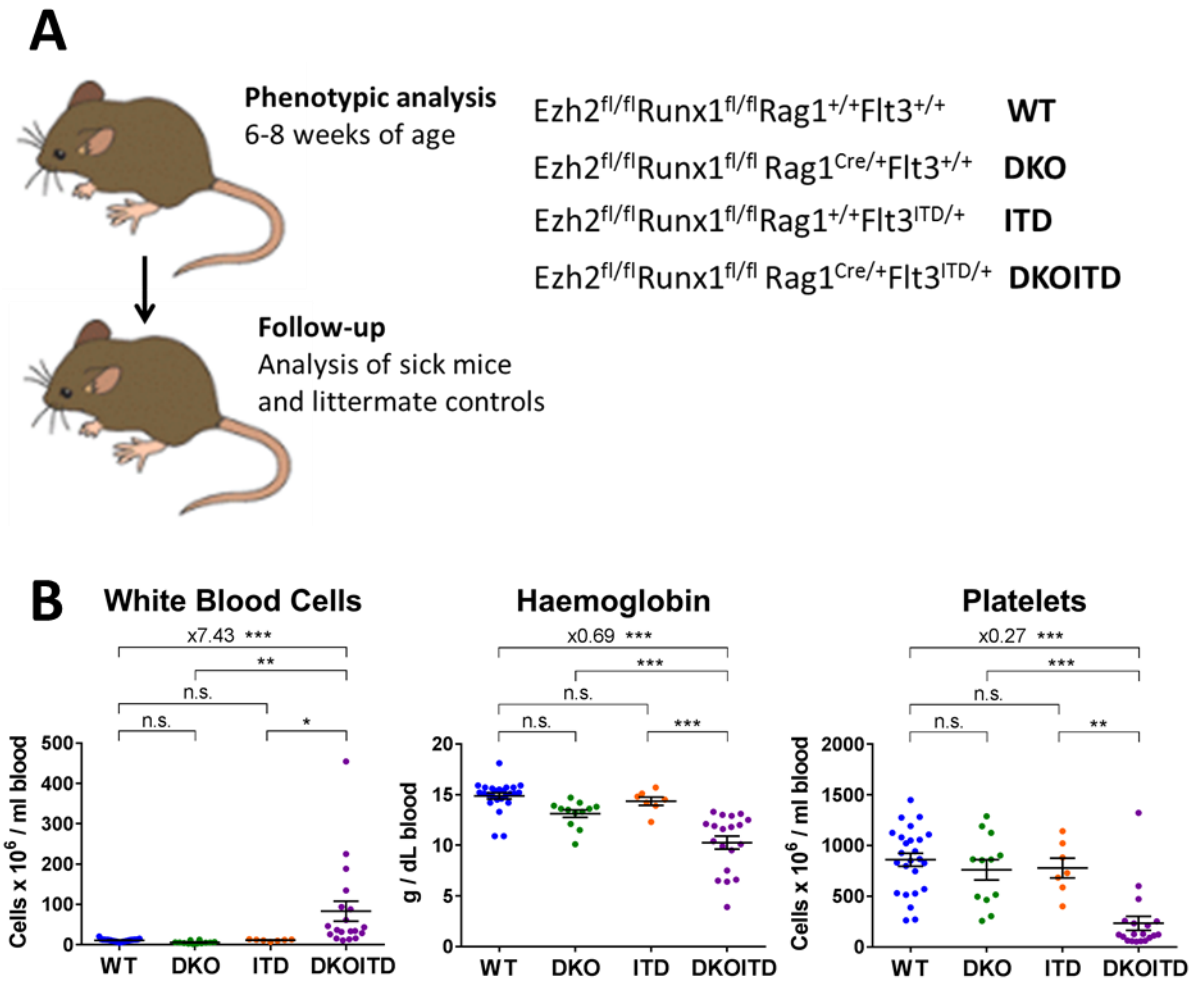


Figure 4.06

(A) Experimental setup and genotypes of experimental mice for Ezh2 and Runx1 inactivation in early lymphoid progenitors using Rag1Cre combined with constitutive Ras-pathway signalling using Flt3-ITD. (B) Haematological parameters of peripheral blood at 6-8 weeks of age in WT (n=25), DKO (n=12), ITD (n=7) and DKOITD (n=19) mice. Error bars represent standard error of the mean. P-values are by ANOVA followed by multiple comparisons testing; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

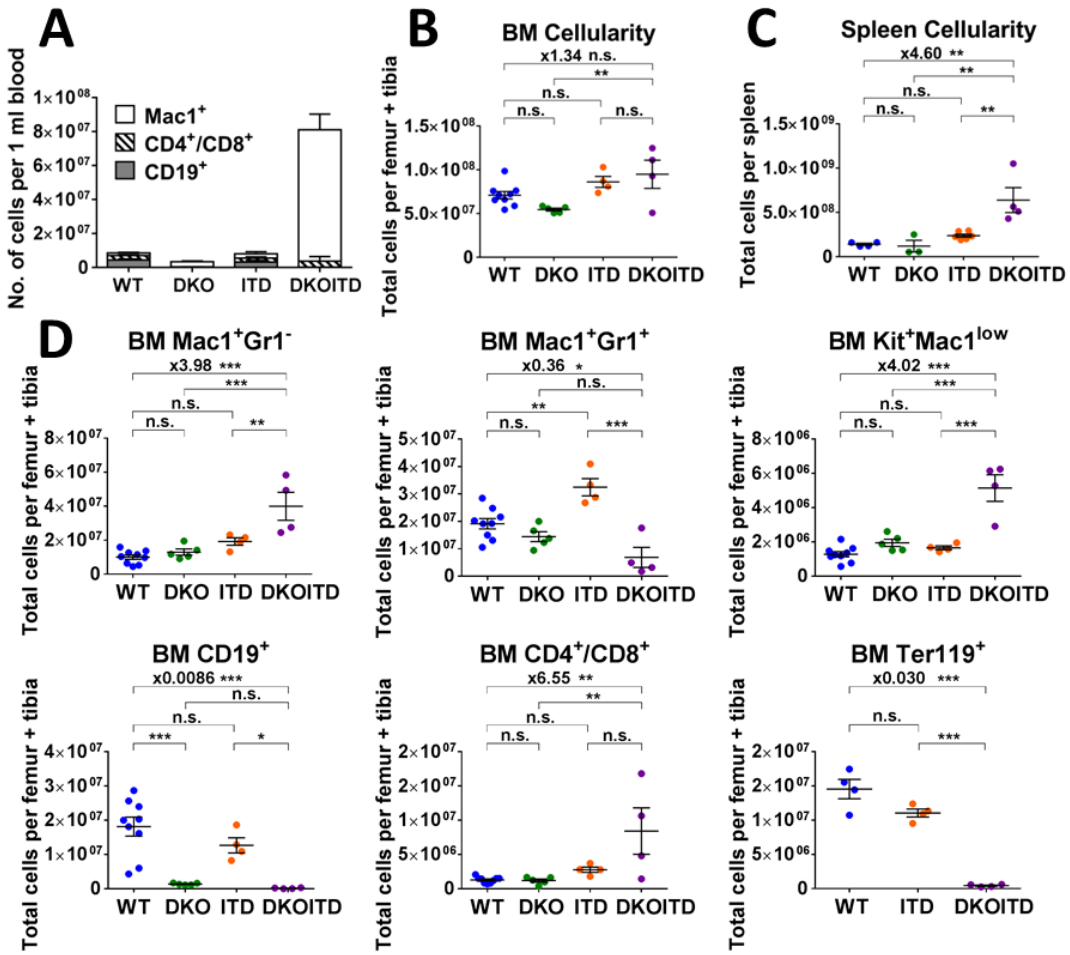


Figure 4.07

(A) Absolute numbers of $CD19^+$ (B-cells), CD4 and/or $CD8^+$ (T-cells) and $Mac1^+$ (myeloid cells) cells per millilitre of blood at 6-8 weeks of age in WT (n=9), DKO (n=5), ITD (n=4) and DKOITD (n=4) mice. (B) Overall BM cellularity per femur and tibia at 6-8 weeks of age in WT (n=9), DKO (n=5), ITD (n=4) and DKOITD (n=4) mice. (C) Overall spleen cellularity at 6-8 weeks of age in WT (n=4), DKO (n=3), ITD (n=6) and DKOITD (n=4) mice. (D) Absolute numbers per femur and tibia of indicated bone marrow populations at 6-8 weeks of age in WT (n=9), DKO (n=5), ITD (n=4) and DKOITD (n=4) mice. $Mac1^+Gr1^-$ myeloid cells (excluding neutrophils), $Mac1^+Gr1^+$ neutrophils, Kit^+Mac1^{low} myeloid precursors, $CD19^+$ B-cells, $CD4^+/CD8^+$ T-cells, $Ter119^+$ erythroid precursors. Error bars represent standard error of the mean. P-values are by ANOVA followed by multiple comparisons testing; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

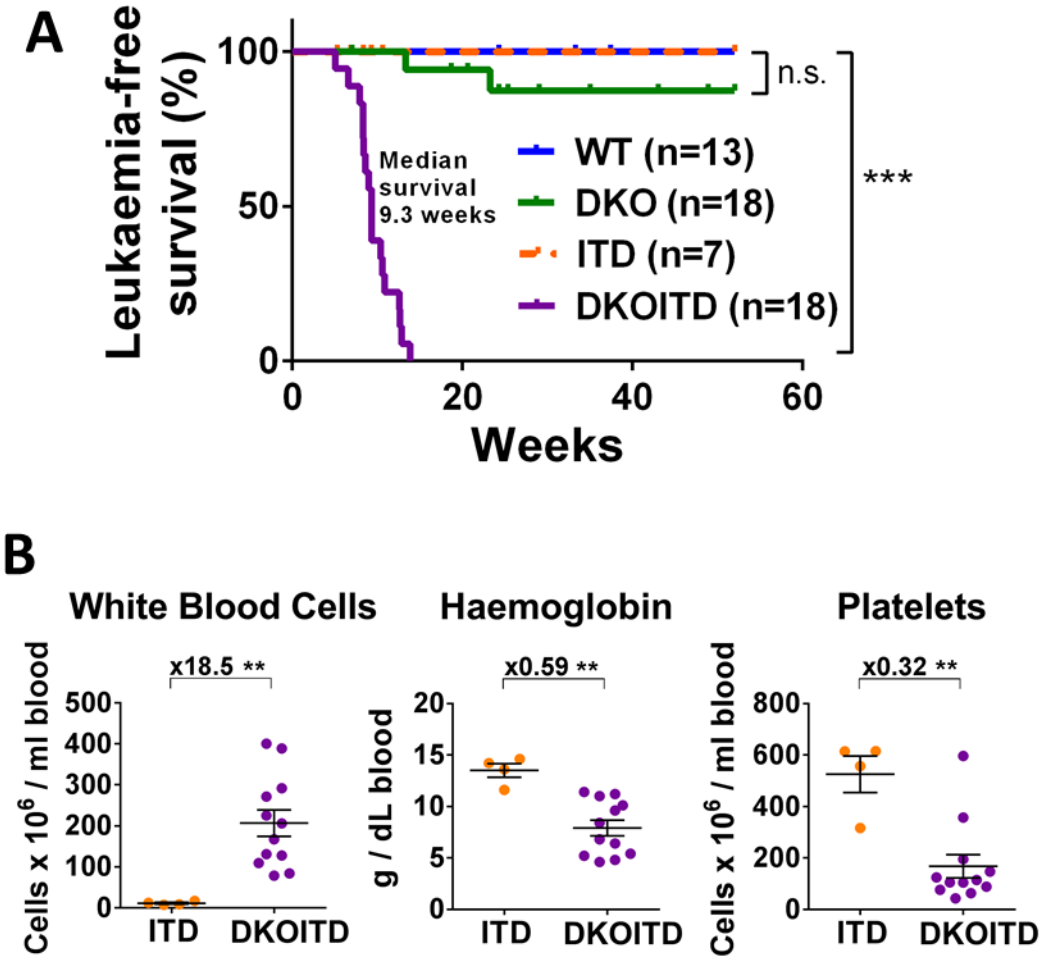


Figure 4.08

(A) Kaplan-Meier leukaemia-free survival curves for Mx1-WT, Mx1-Ezh2KO, Mx1-Runx1KO and Mx1-DKO mice following Poly(IC) treatment at 6-8 weeks of age. Mice found dead without analysis or culled showing > 25 x 10^6 leukocytes per ml blood were counted as events. Mice culled showing leukocyte counts below this constraint were censored. (B) Haematological parameters of peripheral at time of culling in leukaemic DKOITD (n=12) and ITD littermate control (n=4) mice. Error bars represent standard error of the mean. P-values are by logrank test (A) or T-test (B); n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

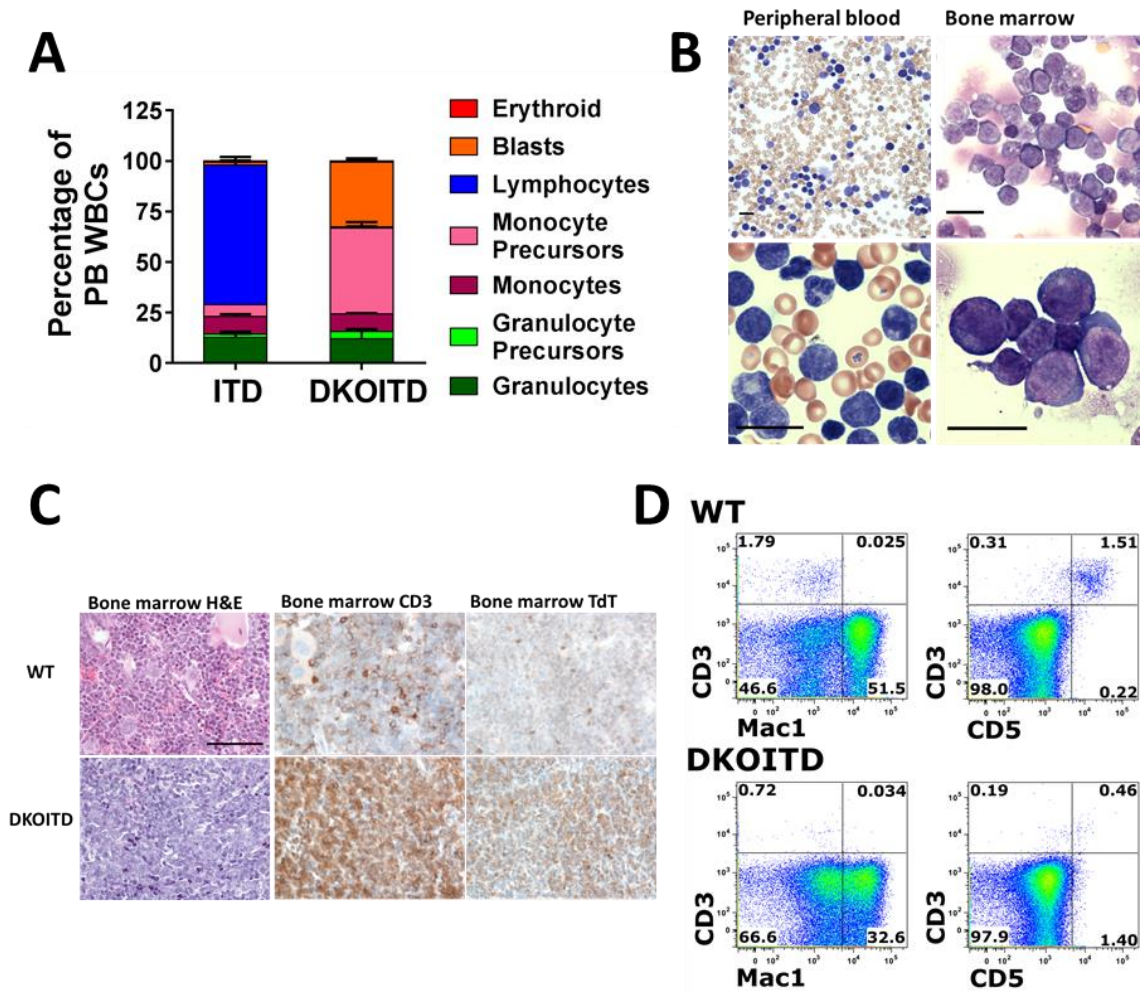


Figure 4.09

(A) Morphological analysis of peripheral blood leukocytes from leukaemic DKOITD (n=3) and ITD littermate control (n=3) mice. One hundred cells were classified for each sample. Classification was performed by Adam Mead. (B) Representative images of peripheral blood smears and bone marrow cytospins from leukaemic DKOITD mice. Scale bars represent 20 µm. (C) Representative images of histological sections of bone marrow from WT and leukaemic DKOITD mice stained for hematoxylin and eosin (H&E), CD3 or TdT. Scale bar represents 100 µm and pertains to all panels. (D) Representative FACS plots showing (left panels) lack of co-expression of Mac1 with surface CD3 and (right panels) positive co-expression of CD5 with surface CD3 in both WT and leukaemic DKOITD bone marrow. Figures represent the mean percentage of total bone marrow cells falling into each gate across all experiments. Error bars represent standard error of the mean.

DKOITD Mac1$^+$ bone marrow cells did not express CD3 on the cell surface (Figure 4.09D). The lymphocyte-specific DNA polymerase TdT is a general marker of ALL, while expression of cytoplasmic but not surface CD3 is an established marker of T-ALL deriving from immature thymocyte progenitors (Chiaretti et al., 2014). This suggests that DKOITD leukaemia cells co-express early T-lineage with myeloid gene transcription, indicative of initiation within a lympho-myeloid progenitor population such as ETPs.

To further explore this possibility, we performed RNA-sequencing of Mac1-expressing bone marrow cells from WT and DKOITD mice. This revealed strong downregulation of myeloid-associated gene expression in DKOITD relative to WT cells (Figure 4.10A), together with upregulation of genes associated with early T-cell differentiation (Figure 4.10B). This is indicative of a switch away from myeloid transcriptional programmes and towards early T-cell programmes in the Mac1$^+$ bone marrow cells in DKOITD mice. Consistent with this, a number of lymphoid-associated genes showed strong upregulation in the DKOITD cells (Figure 4.10C). Among these, *CD7* is specifically associated with the T-cell lineage and is expressed by both ETP-ALL and non-ETP T-ALL (Coustan-Smith et al., 2009), while *Il2ra* (*CD25*), *Il7ra* and *Id3* are all expressed during early thymocyte differentiation (Yui and Rothenberg, 2014).

We also saw increased expression of a number of stem and progenitor associated genes (Figure 4.10D), as well as leukaemia-associated oncogenes (Figure 4.10E) in the DKOITD cells. Among these, *Rasgrp1* is an activator of Ras pathway signalling and has been implicated in T-ALL leukaemogenesis in a mouse model (Oki et al., 2012). *Ccnd2* is a cell cycle regulator found to be stablised in *Runx1*-mutant leukaemias (promoting cell cycle progression) (Faber et al., 2016), while *Tspan3* has been implicated in development and propagation of AML (Kwon et al., 2015).

We next compared the gene expression of DKOITD Mac1$^+$ bone marrow cells to published expression data from ETP-ALL. Zhang et al. published fold-change data for differentially expressed genes between ETP-ALL and non-ETP T-ALL (Zhang et al., 2012b). To compare this with our data, we

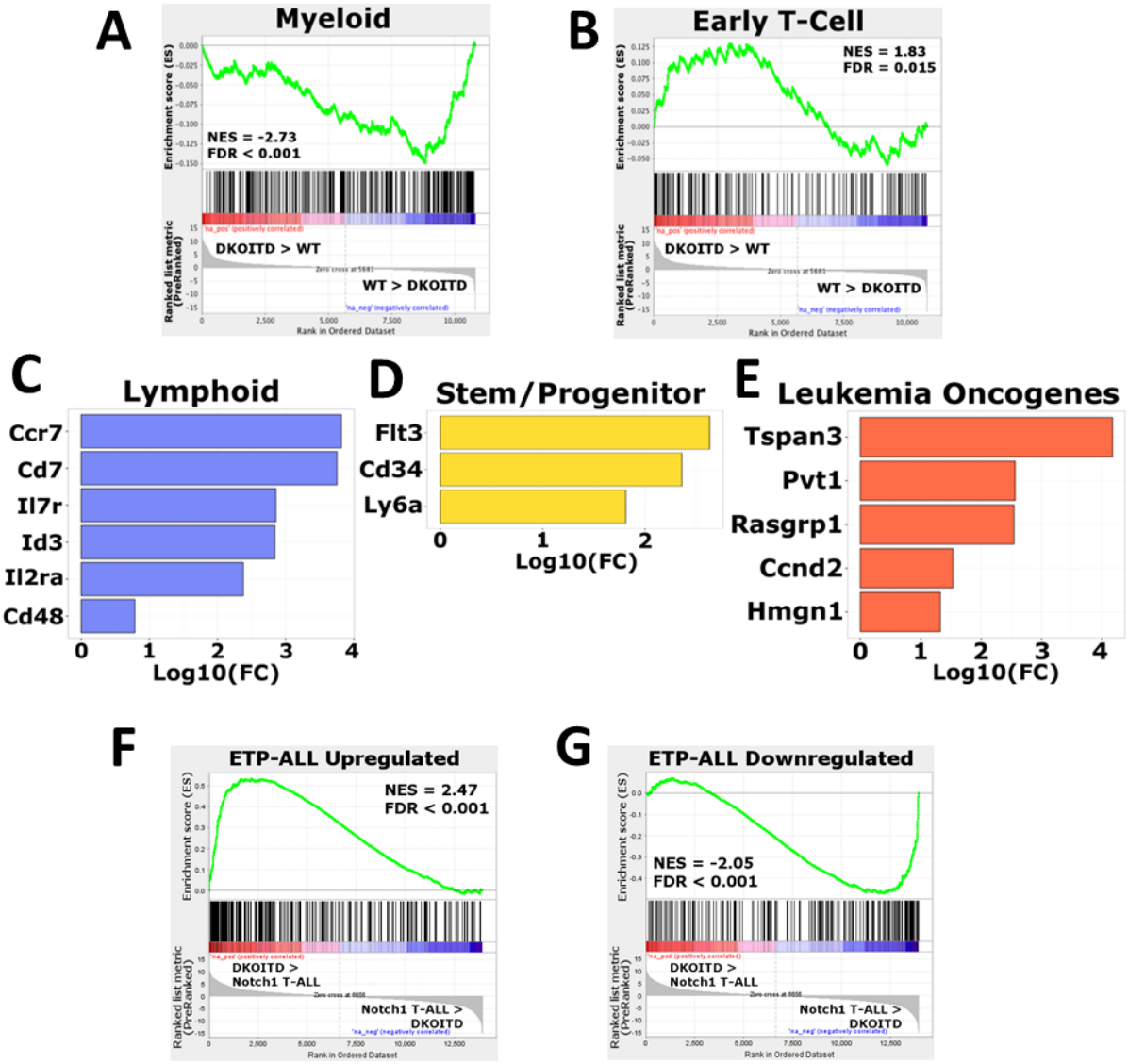


Figure 4.10

(A-B) GSEA comparing WT and DKOITD Mac1$^+$ bone marrow cells for (A) myeloid lineage associated genes and (B) early T-cell lineage associated genes. (C-E) Expression in DKOITD relative to WT Mac1$^+$ bone marrow cells of (C) lymphoid lineage-associated genes, (D) stem/progenitor-associated genes and (E) leukaemia oncogenes, as determined by RNA-sequencing. (F-G) GSEA comparing Notch1 mouse T-ALL and DKOITD Mac1$^+$ bone marrow cells for genes upregulated (F) or downregulated (G) in human ETP-ALL relative to non-ETP T-ALL. Data analysis was performed by Nikolaos Barkas.

obtained published data from a *Notch1*-induced mouse model of T-ALL (Ntziachristos et al., 2014) and produced fold-changes between bone marrow cells from these mice and our DKOITD bone marrow cells. We then created a set of genes upregulated in ETP-ALL relative to non-ETP T-ALL. This gene set was strongly enriched in DKOITD Mac1^{+} bone marrow cells relative to *Notch1*-induced T-ALL bone marrow cells (Figure 4.10F). Performing the same analysis using genes downregulated in ETP-ALL relative to non-ETP T-ALL resulted in negative enrichment (Figure 4.10G). These findings confirm that DKOITD leukaemia displays an ETP-ALL-like gene expression signature.

The fact that the DKOITD leukaemia latency was very short (median 9.3 weeks) suggested that the combination of *Ezh2* and *Runx1* inactivation with Flt3-ITD was sufficient to induce leukaemia without requiring additional mutations. To determine whether this was the case, we performed whole exome sequencing on genomic bone marrow cell DNA from eight DKOITD leukaemias and two related WT control mice. We also sequenced bone marrow DNA from eight wild-type C57BL/6 mice from an unrelated colony, to help with detection of somatically acquired variants in the leukaemia samples. The pipeline used for detection of potentially relevant mutations is shown in Figure 4.11. Using this approach, we detected variants present in at least one DKOITD sample, but no wild-type samples, in 23 cancer-associated genes (Figure 4.11). We then manually reviewed these variants to exclude those present in the germline, and those unlikely to be oncogenic (see Appendix). There was only one somatically acquired leukaemia-associated mutation found, in one DKOITD sample, in the gene *Gata3*. *Gata3* is a transcription factor essential for multiple stages of thymocyte differentiation (Hosoya et al., 2010). Mutations in *GATA3* have been found in ETP-ALL, and appear to be specific to this leukaemia subtype (Zhang et al., 2012b). However, the reads supporting this variant were all of poor sequencing quality. The fact that no other mutations were detected confirms that the combination of *Ezh2* and *Runx1* inactivation in early lymphoid progenitors with *Flt3-ITD* is sufficient to induce leukaemia.

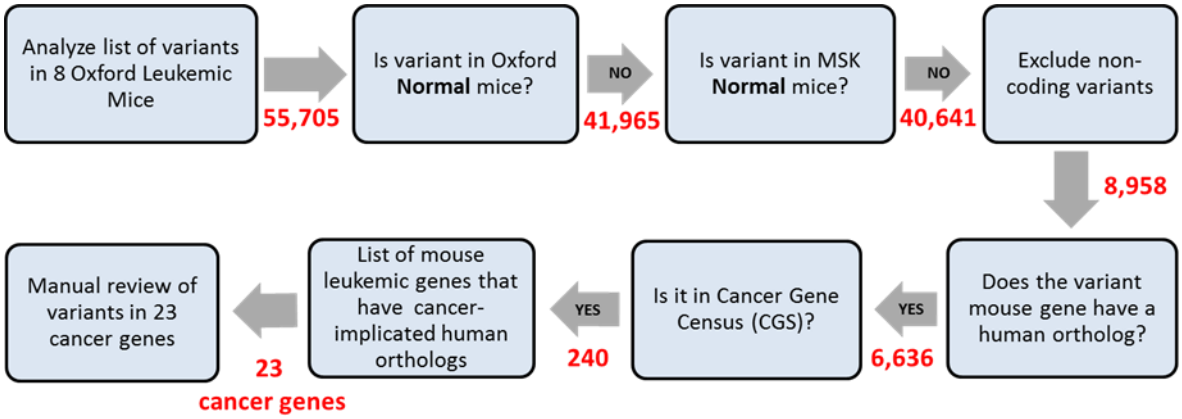


Figure 4.11

Overview of analysis of data from exome sequencing of DKOITD leukaemias. Genomic DNA was extracted from bone marrow cells of 8 DKOITD leukaemic mice, 2 WT mice from the same colony (Oxford Normal mice), and 8 WT mice from an unrelated C57BL/6 colony (MSK Normal mice). Data analysis was performed by Noushin Farnoud and Elli Papaemmanuil.

Together, these findings demonstrate that the leukaemia which develops in DKOITD mice is a lympho-myeloid leukaemia, displaying co-expression of lymphoid and myeloid transcriptional programmes. Notably, its global gene expression signature mirrors that of ETP-ALL, and it shows upregulation of early T-cell associated transcription, suggesting that it may arise from a lympho-myeloid progenitor population such as ETPs. If so, this would provide experimental evidence that lympho-myeloid progenitors can indeed give rise to leukaemias, such as ETP-ALL, which show phenotypic characteristics of multiple haematopoietic lineages.

2.5: DKOITD lympho-myeloid leukaemia can be propagated by ETPs

To explore the question of whether DKOITD leukaemia may arise from lympho-myeloid progenitors, we first analysed thymocyte development in ITD and DKOITD mice in comparison to WT and DKO at 6-8 weeks of age. DKOITD thymuses showed a comparable phenotype to DKO, with a similar reduction in overall cellularity and on average an even greater expansion of ETPs (Figure 4.12A), although this difference was not statistically significant. Thymocyte development appeared to be blocked at an even earlier stage compared to DKO, with no expansion of DN2 cells compared to WT and strongly reduced DN3 cell numbers (Figure 4.12B). ITD thymuses showed a similar phenotype to WT, though DN3 and DN4 cell numbers were slightly reduced. It has been shown previously that homozygous *Flt3*$^{ITD/ITD}$ mice show reduced numbers of ETPs (Mead et al., 2013), and we found that heterozygous ITD mice also showed a non-significant reduction (x0.05) in ETPs compared to WT. This data indicates that ITD does not confer large-scale phenotypic changes to either WT or DKO thymuses, but that DKOITD ETPs may be even more proliferative and impaired in differentiation than DKO ETPs.

We therefore decided to investigate the functional properties of DKOITD ETPs, in order to address whether these cells may possess sufficient proliferative potential to act as leukaemia-propagating

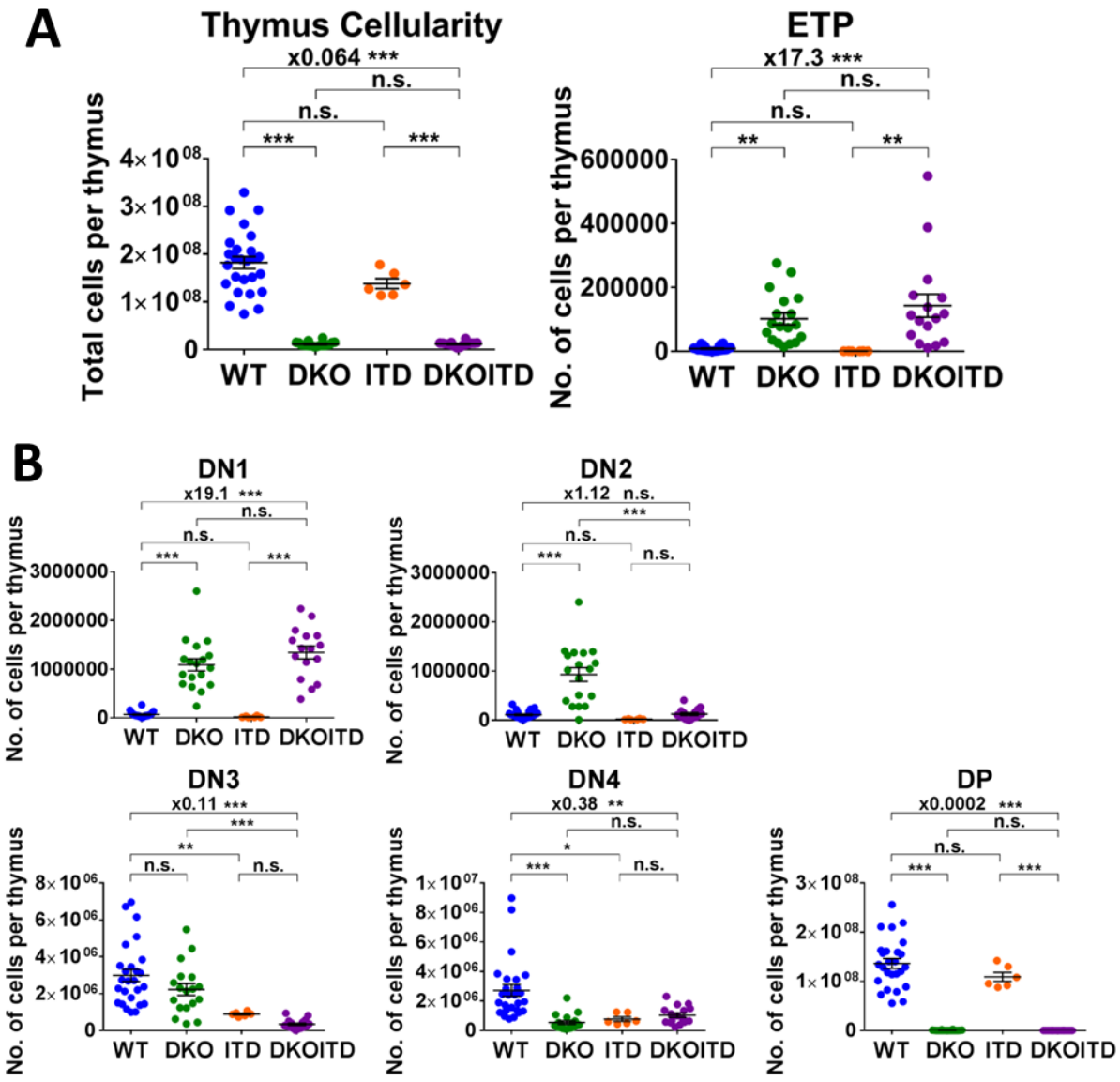


Figure 4.12

(A) Total thymus cellularity and absolute numbers of ETPs per thymus at 6-8 weeks of age for WT (n=26), DKO (n=18), ITD (n=6) and DKOITD (n=16) mice. (B) Absolute numbers of indicated thymocyte populations per thymus at 6-8 weeks of age for WT (n=26), DKO (n=18), ITD (n=6) and DKOITD (n=16) mice. ETP $Lin^-CD4^-CD8^-CD44^+CD25^-Kit^+Flt3^+$, DN1 $Lin^-CD4^-CD8^-CD44^+CD25^-$, DN2 $Lin^-CD4^-CD8^-CD44^+CD25^+$, DN3 $Lin^-CD4^-CD8^-CD44^-CD25^+$, DN4 $Lin^-CD4^-CD8^-CD44^-CD25^-$, DP $Lin^-CD4^+CD8^+$. Error bars represent standard error of the mean. P-values are by ANOVA followed by multiple comparisons testing; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

cells. We first performed granulocyte-monocyte *in vitro* proliferation assays on purified ETPs from WT, DKO and DKOITD mice. In this assay, single cells are sorted into individual wells containing media supplemented with myeloid differentiation-promoting cytokines. The amount of cell growth within each well is then scored after a culture period of 8 days. This provides information on the clonogenicity of the sorted cell population as well as the proliferative potential of individual clones. We found that DKO ETPs showed similar clonogenicity to WT ETPs, but appeared to be slightly more proliferative although this was not significant (Figure 4.13A). DKOITD ETPs in contrast were markedly more clonogenic and proliferative than either DKO or WT ETPs (Figure 4.13A). This suggests that, though DKO ETPs are greatly expanded compared to WT ETPs, the *Flt3-ITD* mutation is required to confer a significant increase to their proliferative potential.

Having established that DKOITD ETPs show markedly enhanced proliferative potential, we decided to transplant ETPs from DKOITD mice into wild-type recipients to determine whether they are able to act as a leukaemia-propagating population. We first showed that unfractionated bone marrow (Figure 4.13B) or thymus (Figure 4.13C) cells were both able to produce stable engraftment upon transplantation into wild-type recipients. Engrafted cells showed expression of Mac1 but not CD19. We therefore sorted ETPs, as well as DN2 and CD4/CD8 double positive and single positive cells from DKOITD thymuses and transplanted them into lethally irradiated recipient mice together with wild-type bone marrow support cells (Figure 4.14A). We found that only ETPs, and not the more mature thymocyte populations, were able to engraft in the recipient mice (Figure 4.14B). The engrafting cells in the peripheral blood showed uniform expression of Mac1, but not CD19, CD4 or CD8 (Figure 4.14B). The recipients went on to develop the same leukaemia phenotype as primary DKOITD mice, showing leukocytosis (in 2 of 4 engrafted recipients), anaemia and thrombocytopaenia (Figure 4.15A), myeloid blast-like morphology (Figure 4.15B) and expression of cytoplasmic CD3 in the bone marrow (Figure 4.15C). Analysis of the bone marrow stem and progenitor compartment in recipient mice revealed a marked expansion of LSK, MPP and LMPP populations, which were almost entirely composed of donor DKOITD ETP-derived cells (Figure 4.16A). Interestingly however, DKOITD donor

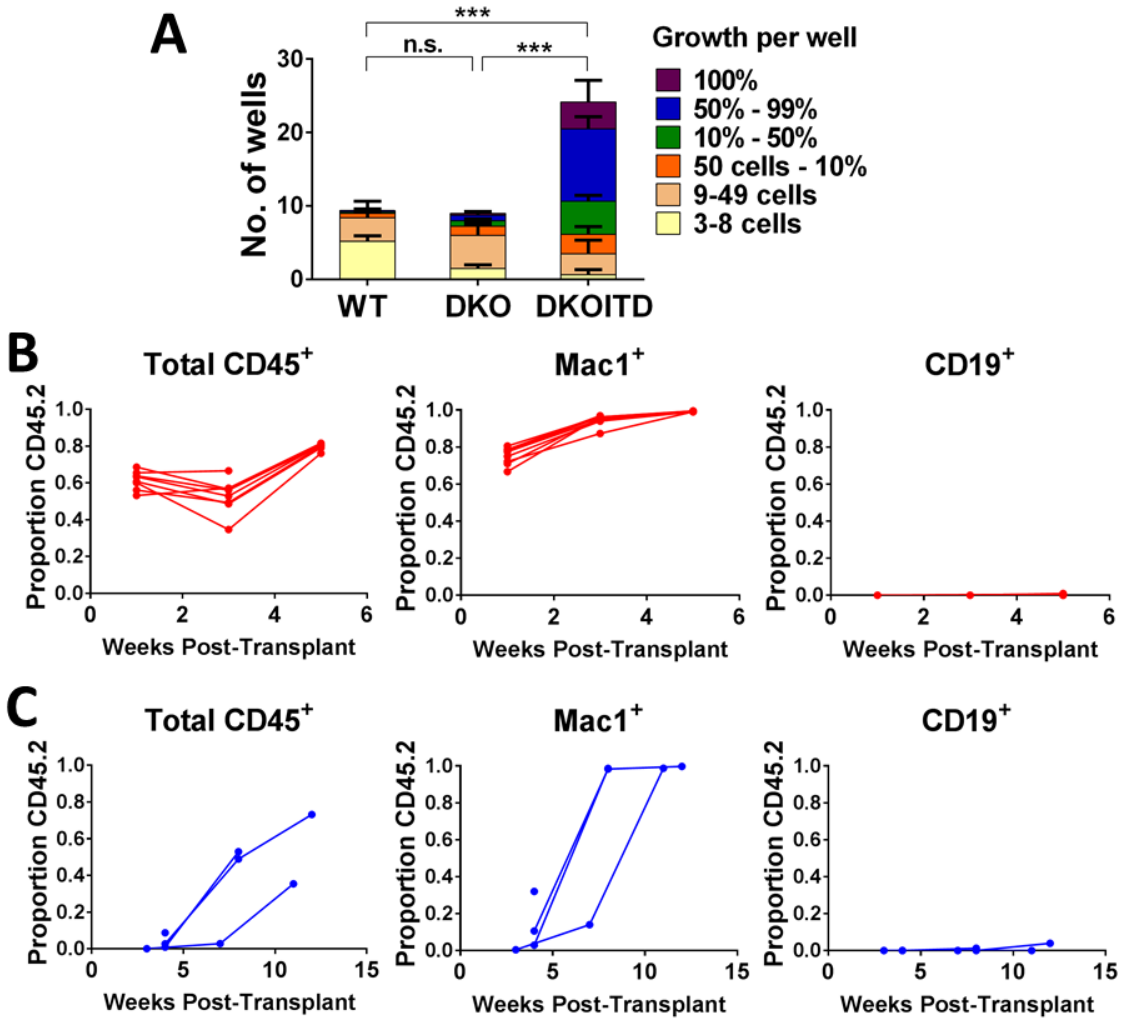


Figure 4.13

(A) In vitro granulocyte-monocyte assay of myeloid proliferative potential of ETPs from WT (n=5), DKO (n=4) and DKOITD (n=3) mice. Error bars represent standard error of the mean. P-values are by T-test comparing number of wells per plate showing high proliferation (>50 cells); n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001. (B-C) Level of CD45.2 chimerism of total white blood cells, myeloid ($Mac1^+$) and B cells ($CD19^+$) in peripheral blood of CD45.1 recipient mice post-transplantation of unfractionated CD45.2 DKOITD bone marrow (B; n=8, 2 million BM cells transplanted per recipient) or thymus (C; n=4, 2.4-9.2 million thymus cells transplanted per recipient) cells with CD45.1 support bone marrow cells. Each line represents an individual recipient mouse.

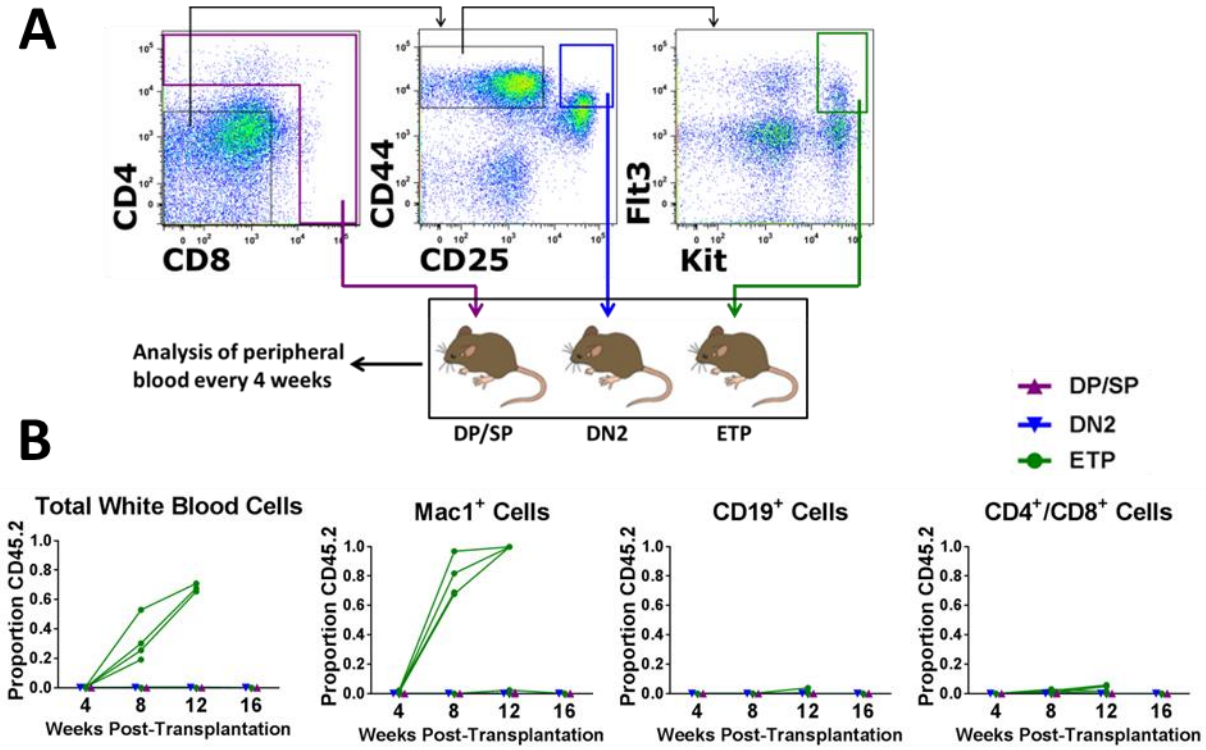


Figure 4.14

(A) Experimental design showing thymocyte population sorting strategy from CD45.2 DKOITD mice for transplantation into wild-type CD45.1 recipients. Recipient mice were lethally irradiated before intravenous injection of ETPs (range 413-6880 cells), DN2 cells (range 557-8005) or DP/SP cells (range 412-1462) with 150,000-250,000 CD45.1 unfractionated bone marrow support cells. (B) Level of CD45.2 chimerism of total white blood cells, myeloid ($Mac1^+$), B ($CD19^+$) and T (CD4 and/or $CD8^+$) cells in peripheral blood of CD45.1 recipient mice post-transplantation of FACS-purified CD45.2 DKOITD thymocyte populations with CD45.1 support bone marrow cells (DP/SP n=7, DN2 n=5, ETP n=12). Each line represents an individual recipient mouse. ETP $Lin^-CD4^-CD8^-CD44^+CD25^-Kit^+Flt3^+$, DN2 $Lin^-CD4^-CD8^-CD44^+CD25^+$, DP $Lin^-CD4^+CD8^+$, SP $Lin^-CD4^+CD8^-$ or $Lin^-CD4^-CD8^+$.

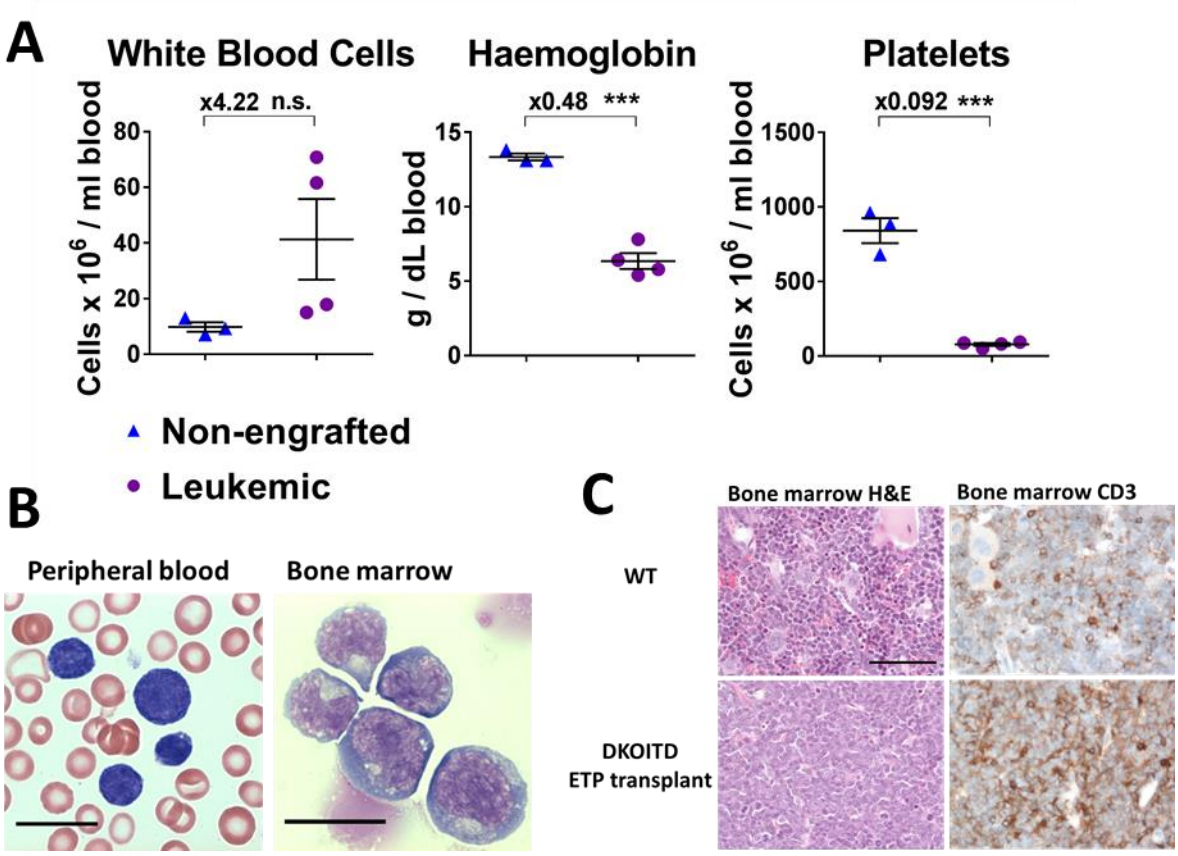


Figure 4.15

(A) Haematological parameters of peripheral blood from mice transplanted with DKOITD ETPs: 16 weeks post-transplantation for recipient mice showing no CD45.2 engraftment (n=3), and upon development of leukaemia for recipient mice showing CD45.2 engraftment (n=4). Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001. (B) Representative peripheral blood smear and bone marrow cytospin images from a leukaemic DKOITD ETP transplant recipient. Scale bars represent 20 µm. (C) Representative images of histological sections of bone marrow from WT and leukaemic DKOITD ETP transplant recipient mice stained for hematoxylin and eosin (H&E) or CD3. Scale bar represents 100 µm and pertains to all panels.

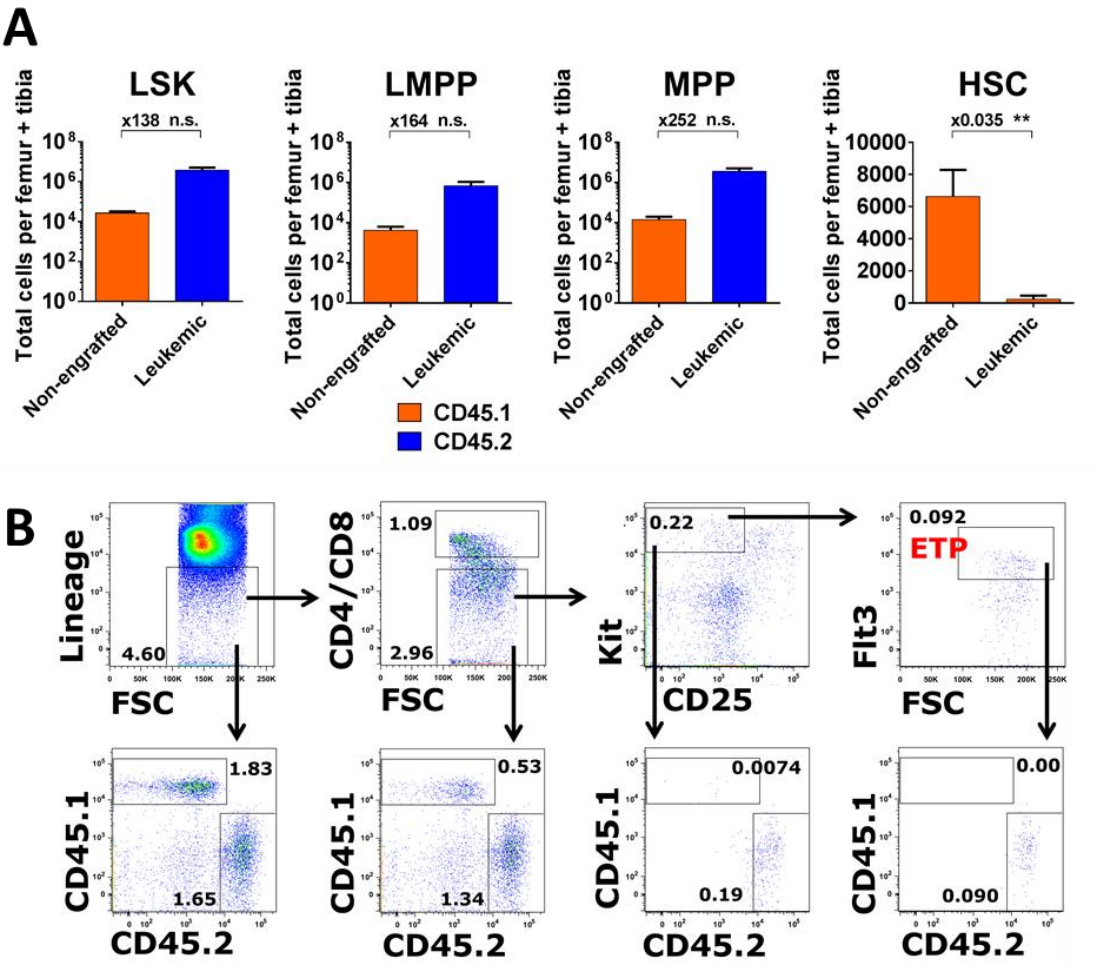


Figure 4.16

(A) Absolute numbers per femur and tibia of CD45.1 and CD45.2 cells contributing to indicated bone marrow stem and progenitor populations 16 weeks post-transplantation for recipient mice showing no CD45.2 engraftment (n=3), and upon development of leukaemia for recipient mice showing CD45.2 engraftment (n=4). LSK Lineage$^-$Sca1$^+$Kit$^+$, LMPP Lineage$^-$Sca1$^+$Kit$^+$Flt3^{high}, MPP Lineage$^-$ Sca1$^+$Kit$^+$CD48$^+$CD150$^-$, HSC Lineage$^-$Sca1$^+$Kit$^+$CD48$^-$CD150$^+$. Error bars represent standard error of the mean. P-values are by T-test comparing summed CD45.1 and CD45.2 cells; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001. (B) Representative FACS plots showing relative contribution of CD45.1 and CD45.2 cells to thymocyte populations in leukaemic DKOITD ETP-transplant recipient mice. Figures represent the mean percentage of total thymocytes falling into each gate across all replicates (n=4).

cells did not produce phenotypic HSCs. While HSCs were greatly reduced in number, they were entirely wild-type host or support cell-derived (Figure 4.16A).

Analysis of the thymuses of recipient mice at the time of culling revealed that early stages of thymocyte maturation showed significant contribution from donor-derived cells (Figure 4.16B). Recipient ETPs were in fact almost all derived from transplanted DKOITD ETPs. This indicates that DKOITD ETPs, while also propagating a leukaemia that spreads throughout the haematopoietic tissues, infiltrate the recipient thymus and displace the wild-type host early-stage thymocytes.

2.6: DKOITD ETPs and bone marrow stem/progenitor cells display upregulated Ras pathway signalling and cell cycle dysregulation

To further investigate the mechanisms underlying the enhanced proliferative potential of DKOITD ETPs, we performed RNA-sequencing on four samples of 100 ETPs from separate DKOITD mice and compared this data to WT and DKO ETPs. Semi-supervised hierarchical clustering using the 3000 most variable genes grouped the four WT samples separately from the DKO and DKOITD samples, but did not separate all of the DKO samples from the DKOITD samples (Figure 4.17A). This indicates that the addition of *Flt3-ITD*, while inducing greatly enhanced proliferative potential, did not confer large-scale transcriptional changes to DKO ETPs. This was supported by the fact that only 31 genes were differentially expressed in DKOITD versus DKO ETPs, compared to 306 in DKO versus WT and 481 in DKOITD versus WT (Figure 4.17B). When clustering using only differentially expressed genes, all samples did group together by genotype (Figure 4.17C).

To further explore the relationships between HSCs and lympho-myeloid progenitors in our experimental mice, we next performed single cell qPCR analysis of WT, DKO and DKOITD HSCs, LMPPs and ETPs. We used principal component analysis to compare the transcriptional profiles of these cells (Figure 4.18). We found that, in agreement with our RNA-seq data, DKO and DKOITD ETPs

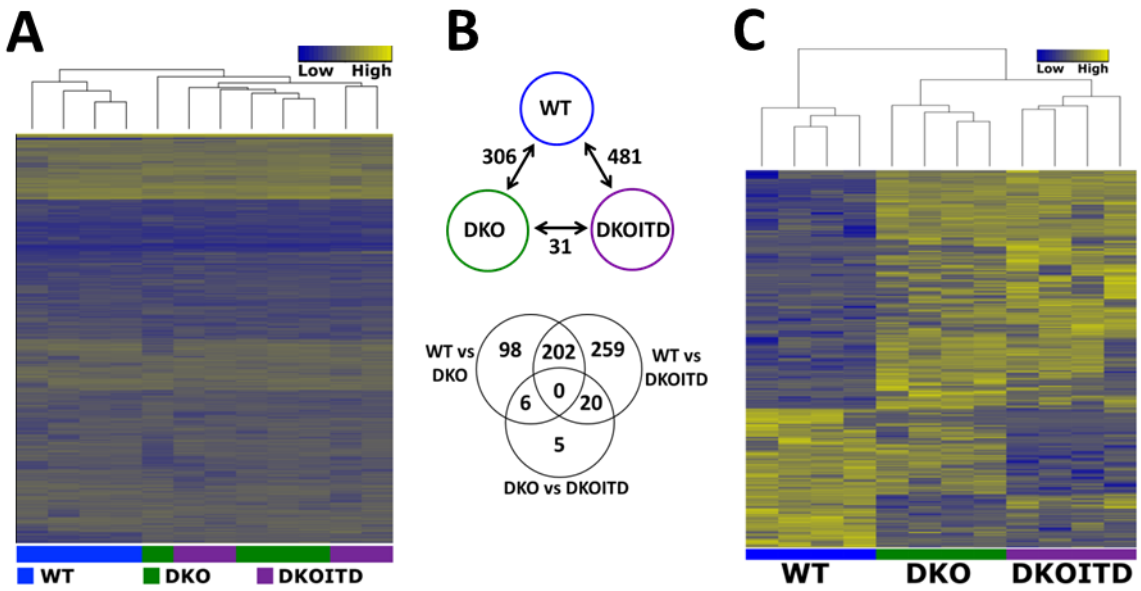


Figure 4.17

(A) Hierarchical clustering and heatmap showing expression levels, scaled per gene, of the 3000 most variable genes between WT, DKO and DKOITD ETPs (WT n=4, DKO n=4, DKOITD n=4). Each replicate represents a sample of 100 purified ETPs from an individual mouse. (B) (Top) Numbers of significantly differentially expressed genes between WT, DKO and DKOITD ETPs as determined by RNA-sequencing, and (bottom) overlap of genes within each pairwise comparison (WT n=4, DKO n=4, DKOITD n=4). (C) Hierarchical clustering and heatmap showing expression levels, scaled per gene, of all significantly differentially expressed genes between WT, DKO and DKOITD ETPs (WT n=4, DKO n=4, DKOITD n=4). Each replicate represents a sample of 100 purified ETPs from an individual mouse. Data analysis was performed by Nikolaos Barkas.

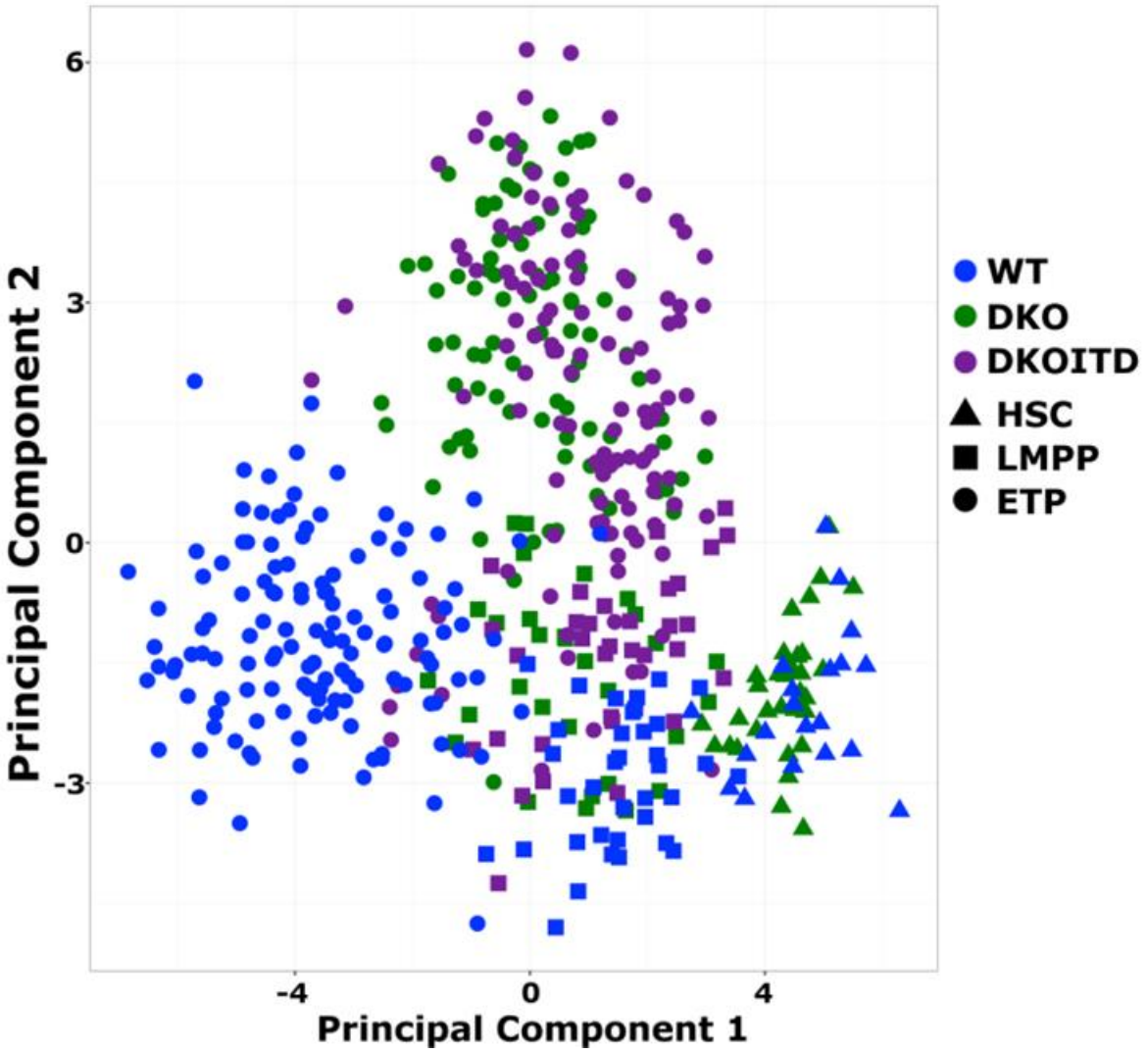


Figure 4.18

Principal component analysis of single-cell qPCR data from WT HSCs (n=20), LMPPs (n=36) and ETPs (n=129), DKO HSCs (n=33), LMPPs (n=31) and ETPs (n=91), and DKOITD LMPPs (n=33) and ETPs (n=122). Cells were collected from 12 WT, 9 DKO and 8 DKOITD mice. Data analysis was performed by Nikolaos Barkas.

clustered together and separately from WT ETPs. WT and DKO HSCs clustered together, reflecting the lack of *Rag1Cre* activation in these cells. Phenotypic HSCs are very rare in DKOITD bone marrow and we did not attempt to isolate these cells for transcriptional analysis. LMPPs of all genotypes were located in between the HSC and ETP clusters, reflecting their relationship to these cells within the haematopoietic hierarchy (Figure 1.01). Similarly to ETPs, DKO and DKOITD LMPPs tended to cluster together and separately from WT LMPPs, although there was more overlap between the clusters than for ETPs. This may reflect intermediate levels of *Rag1Cre* activation in these cells (Figure 4.06B). Together, this analysis confirms that DKOITD ETPs are broadly similar at the transcriptional level to DKO ETPs, despite the addition of the *Flt3-ITD* mutation and enhanced proliferative potential. This provides support for the phenotypic ETP population in DKOITD thymuses representing genuine ETPs which have gained aberrant self-renewal and sufficient proliferative potential to propagate leukaemia.

To investigate transcriptional differences that may exist between DKOITD and DKO ETPs, we used the Hallmark GSEA database (Liberzon et al., 2015) to compare the RNA-sequencing profiles from these populations. The most significant result was a negative enrichment for genes involved in the G2M cell cycle checkpoint (Figure 4.19A). This indicates that genes important for controlling progression through the cell cycle were downregulated in DKOITD relative to DKO ETPs, helping to explain why DKOITD ETPs were markedly more proliferative than DKO ETPs (Figure 4.13A). In agreement with this data, qPCR analysis showed that a number of signalling genes were upregulated in DKOITD relative to DKO ETPs (Figure 4.19B). This was also the case for some myeloid (Figure 4.19C) and HSC genes (Figure 4.19D).

We next questioned whether DKOITD ETPs showed a gene expression profile reflecting that of ETP-ALL. We performed differential expression analysis of DKOITD ETPs relative to a previously published dataset from wild-type CD4/CD8 double positive thymocytes (Zhang et al., 2012c). We then performed GSEA using our ETP-ALL gene set. This demonstrated that DKOITD ETPs are highly

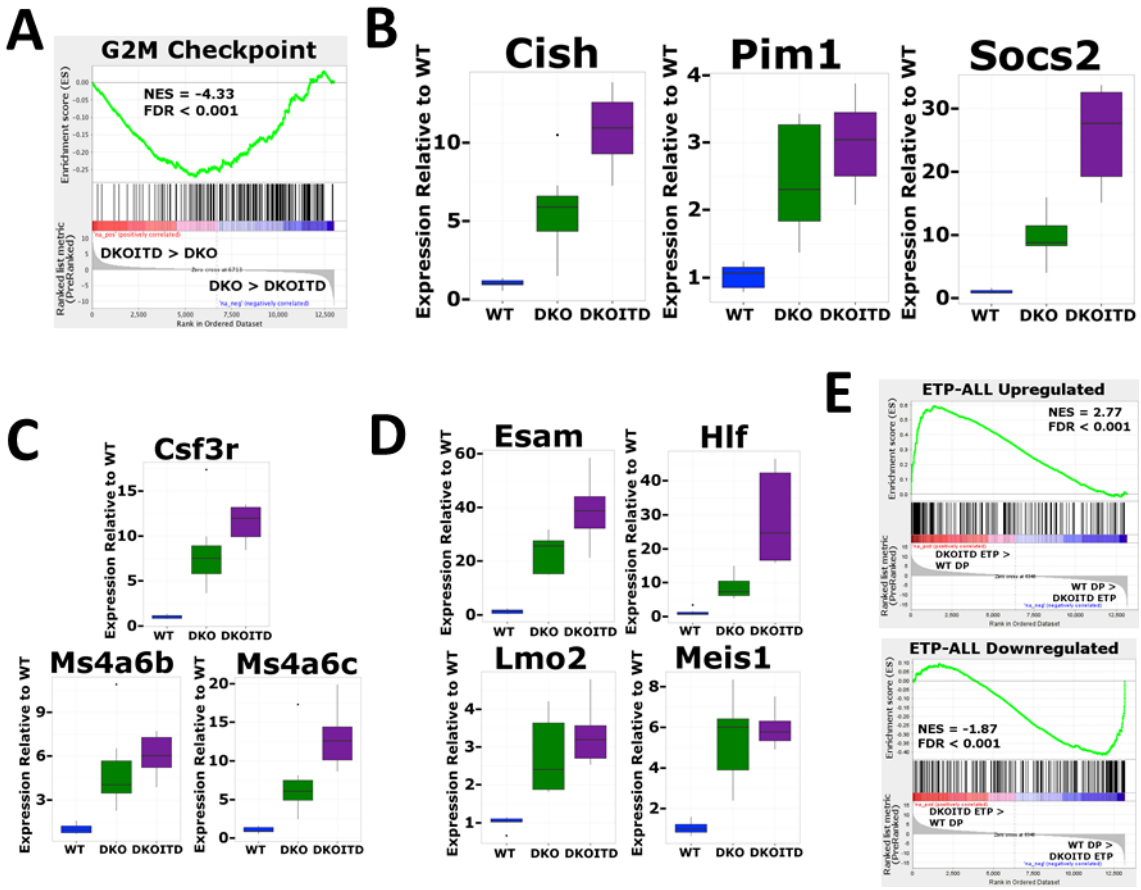


Figure 4.19

(A) GSEA comparing DKO and DKOITD ETPs for G2M checkpoint-associated gene expression. (B-D) Expression relative to WT of (B) signalling pathway, (C) myeloid lineage and (D) HSC associated genes in DKO and DKOITD ETPs as determined by qPCR. WT n=8, DKO n=7, DKOITD n=6; each replicate represents a sample of 100 purified ETPs from an individual mouse. Error bars represent min and max values. (E) GSEA comparing WT DP thymocytes and DKOITD ETPs for genes upregulated (top) or downregulated (bottom) in human ETP-ALL relative to non-ETP T-ALL. Data analysis was performed by Nikolaos Barkas.

enriched for genes upregulated in ETP-ALL, and negatively enriched for downregulated genes (Figure 4.19E).

In order to functionally validate our findings that signalling pathways are upregulated in DKOITD cells, we performed phospho-flow cytometry. We cultured freshly isolated WT, DKO, ITD and DKOITD bone marrow cells in serum-free conditions for 2 hr, then stimulated with stem cell factor (SCF) for 10 min. Following SCF stimulation, cells were immediately fixed, permeabilised and stained for lineage markers, Kit and phospho-S6. The S6 ribosomal protein has been shown to be phosphorylated in response in RAS pathway signalling (Roux et al., 2007), such as via SCF stimulation. We found that in WT, DKO and ITD lineage-negative, Kit-positive stem and progenitor cells, there was a clear shift in phospho-S6 staining between unstimulated and SCF-stimulated cells. However, DKOITD cells showed strongly increased basal levels of phospho-S6 staining in the absence of SCF stimulation (Figure 4.20A). This indicates that DKOITD bone marrow stem and progenitor cells show enhanced basal levels of Ras-pathway signalling compared to the other genotypes. To confirm that this increased phospho-S6 staining was caused by Ras pathway activation, we performed the same experiment including treatment with PD-901, a compound that inhibits RAS pathway signalling via inhibition of the MEK1 and MEK2 signalling transducers (Sebolt-Leopold and Herrera, 2004). When PD-901 was added to the culture media, staining for phospho-S6 in unstimulated DKOITD cells was reduced to the level of staining in unstimulated ITD cells without inhibitor (Figure 4.20B). This confirms that the enhanced basal signalling observed in DKOITD bone marrow progenitor cells was specific to the Ras pathway.

In summary, our findings demonstrate that inactivation of *Ezh2* and *Runx1* targeted to early lymphoid progenitors using *Rag1Cre* results in a marked expansion of ETPs, which show transcriptional features similar to human ETP-ALL including co-expression of lymphoid, myeloid and HSC-associated genes. Upon addition of *Flt3-ITD*, these ETPs are transformed into a leukaemia-

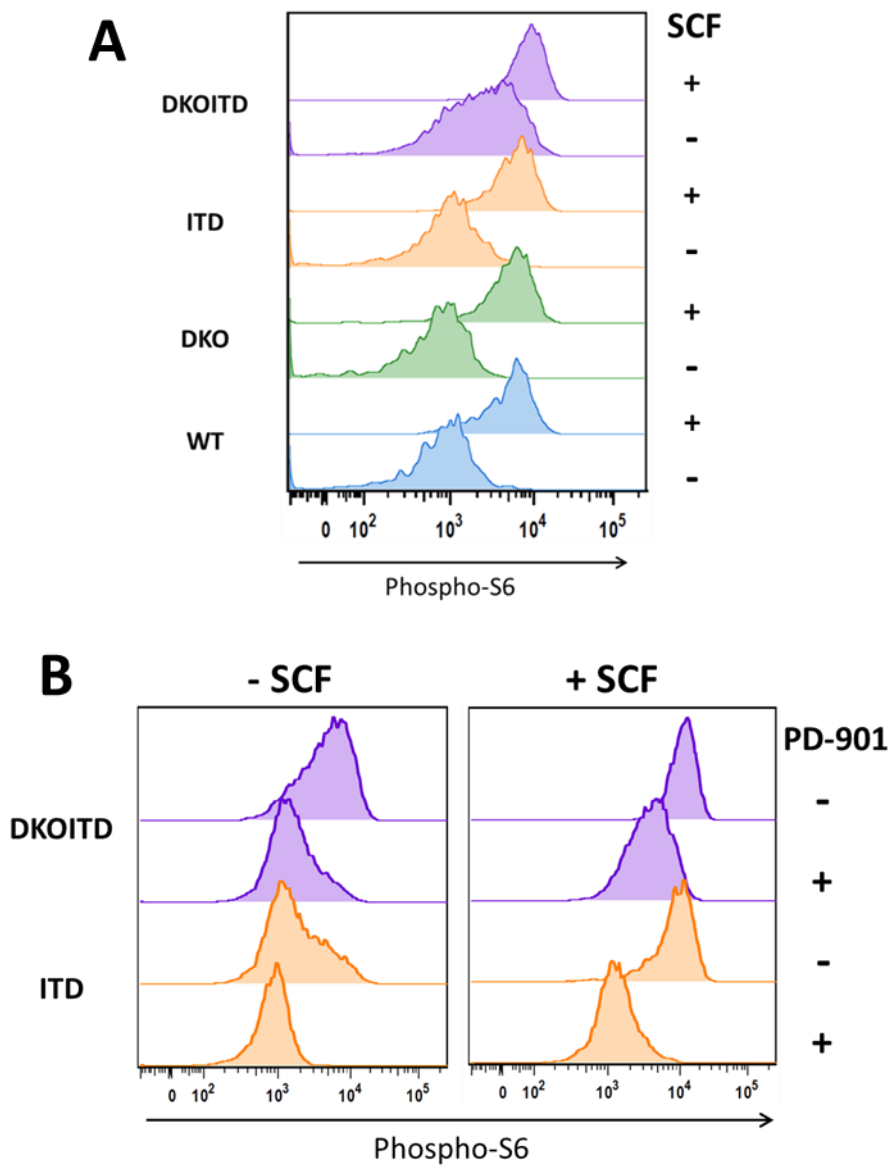


Figure 4.20

(A) Analysis of phospho-S6 in WT, DKO, ITD and DKOITD $Lin^{-}Kit^{+}$ bone marrow cells, either unstimulated or stimulated for 10 min with 8 ng/ml SCF. Representative histograms from 3 independent experiments are shown. (B) Analysis of phospho-S6 in ITD and DKOITD $Lin^{-}Kit^{+}$ bone marrow cells treated for 2 hr with DMSO or 2.5 µM PD-901 (MEK inhibitor) and then either unstimulated or stimulated for 10 min with 8 ng/ml SCF. Representative histograms from 3 independent experiments are shown.

propagating population, and the resulting leukaemia shows transcriptional and phenotypic lympho-myeloid characteristics as well as enhanced signalling via the Ras pathway. This study provides experimental evidence directly linking the phenotype of a leukaemia with its cell of origin, in this case lympho-myeloid progenitors.

3. Discussion

3.1: Establishment of a novel mouse model of ETP-ALL

We have established a novel mouse leukaemia model through inactivation of *Ezh2* and *Runx1*, and constitutive activation of *Flt3*. *Ezh2* and *Runx1* were inactivated specifically in early lymphoid progenitors using *Rag1Cre* (McCormack et al., 2003), while *Flt3-ITD* was knocked-in to the *Flt3* locus, becoming expressed only in cells that would normally express *Flt3* (Lee et al., 2007).

Several mouse leukaemia models have been published purporting to represent ETP leukaemias. Yu et al. used germline knockout of *Tcf7*, which is downregulated in ETP-ALL relative to non-ETP T-ALL, and showed that these mice developed T-cell lymphomas (Yu et al., 2012). Treanor et al. took thymocytes from *Arf* knockout mice, transduced them with constitutively active *Il7r* and transplanted them into immune deficient mice (Treanor et al., 2014). Goossens et al. used overexpression of *Zeb2* combined with *p53* knockout in endothelial and haematopoietic tissues (Goossens et al., 2015). Finally, Danis et al. combined *Ezh2* and *Cdkn2a* knockout with *Nras* activation in cultured LSK cells, followed by transplantation (Danis et al., 2016). All of these approaches resulted in malignancies showing some phenotypic and transcriptional similarities to ETP-ALL. However, none of them used conditional knockout approaches to target mutations specifically to ETPs and not HSCs. Additionally, none showed a selective expansion of ETPs or demonstrated that ETPs had gained LSC potential. Our DKOITD model is therefore unique in

establishing a close connection between ETPs and the leukaemia phenotype, and also in demonstrating leukaemia-propagating potential of ETPs themselves.

Our strategy of conditional knockout using *Rag1Cre* combined with knock-in of Flt3-ITD to the endogenous locus also carries the advantage of greater tractability compared to previous models. No viral transduction or transplantation steps are required for generation of leukaemic mice, and DKOITD mice succumb to the disease relatively soon after reaching adulthood. Transplantation of DKOITD bone marrow cells produces reliable engraftment at consistent levels, enabling preclinical studies of therapeutic compounds as explored further in Chapter 5.

3.2: The relevance of the lympho-myeloid pathway in initiation of acute leukaemia

Inactivation of *Ezh2* and *Runx1* was targeted specifically to early lymphoid progenitors for two main reasons: avoidance of the bone marrow failure phenotype which develops when these mutations are targeted to HSCs (Chapter 3); and to determine whether leukaemia may be initiated when these mutations are activated in progenitors and not HSCs.

AML was for some time believed to be initiated within HSCs and solely propagated by malignant counterparts of these cells (Bonnet and Dick, 1997; Ishikawa et al., 2007). However, more detailed sub-fractionation of the human stem/progenitor compartment showed that in many AML patients, LSC activity in fact exists within populations showing phenotypic and molecular characteristics of LMPPs and granulocyte-monocyte progenitors (GMPs) (Goardon et al., 2011). In T-ALL, evidence from gene expression analysis suggests that leukaemias from different patients initiate within progenitor populations at different stages of thymocyte maturation, implying that HSCs are not the target cell of leukaemic transformation (Aifantis et al., 2008; Ferrando et al., 2002). However, analysis of the gene expression of a leukaemia cannot provide a definitive answer as to the initiating cell population, since mutations acquired may confer transcriptional patterns uncharacteristic of the

true initiating cell. For example, our own data show that ETPs gain aberrant expression of HSC-associated transcription upon loss of *Ezh2* and *Runx1*, despite inactivation being specifically targeted to non-HSCs.

This raises the question of whether the lineage potential of the initiating progenitor population may have an impact on the molecular characteristics and clinical outcome of the resulting leukaemia, or whether these factors are simply a function of the specific combination of mutations acquired. The most primitive ETPs in the mouse thymus, expressing *Kit* and *Flt3*, retain the lineage potentials of the bone marrow LMPP: B-lymphoid and myeloid as well as T-lymphoid (Luc et al., 2012). As these cells mature they lose first B-lymphoid and then myeloid potential. It is notable therefore that ETP-ALLs display both myeloid and T-lymphoid gene expression, strongly implying that these leukaemias initiate within ETPs (Coustan-Smith et al., 2009; Zhang et al., 2012b). The clinical characteristics of ETP-ALL, moreover, provide evidence that the initiating population carries implications for the progression of the disease. Evidence has been published arguing both for (Coustan-Smith et al., 2009; Jain et al., 2016) and against (Bond et al., 2017; Patrick et al., 2014) ETP-ALLs carrying an inferior survival risk compared to T-ALLs as a whole, but they are clearly more resistant to chemotherapy and require stronger treatment to achieve remission. It has been suggested that this may be partially due to their HSC-like gene expression (Coustan-Smith et al., 2009), a consequence of initiation within a very primitive progenitor population.

The gene expression of multiple lineage pathways displayed by ETP-ALL means that it has some features of a “biphenotypic” leukaemia. The term biphenotypic leukaemia, or mixed phenotype acute leukaemia (MPAL), can also refer to rare cases of leukaemias which cannot be placed within any of the standard classifications due to expression of phenotypic markers belonging to multiple lineages (Matutes et al., 2011; Weinberg and Arber, 2010). ETP-ALL does not meet the formal criteria for this disease entity, but it is notable that biphenotypic leukaemias almost always co-express either T-lymphoid or B-lymphoid antigens with myeloid antigens. Co-expression of B-

lymphoid with T-lymphoid antigens is rare, and co-expression of lymphoid with megakaryocyte or erythroid antigens is not observed (Matutes et al., 2011; Weinberg and Arber, 2010). This suggests that these leukaemias may in some or all cases arise within oligopotent progenitor populations and that there are distinct pathways of haematopoietic differentiation along which transformation can occur, some combinations of lineage potentials being allowed and some not, and that this reflects the lineage restriction of the progenitor cell in which the leukaemia originated. This highlights the relevance of the lympho-myeloid pathway in the development of leukaemia, since it explains the existence of acute leukaemias expressing markers of both the lymphoid and myeloid lineages and predicts that these leukaemias need not initiate within HSCs.

Our DKOITD ETP leukaemia model confirms the importance of the lympho-myeloid pathway in initiation of a subset of acute leukaemias in several ways. Firstly, targeting of *Ezh2* and *Runx1* mutations specifically to lympho-myeloid progenitors, and not HSCs, was sufficient to initiate the leukaemia and not a bone marrow failure phenotype. Secondly, the leukaemia showed co-expression of myeloid and lymphoid (as well as HSC) lineage-associated transcription, arguing that the phenotypic and functional characteristics of the initiating population do indeed impact upon the resulting leukaemia. Thirdly, we were able to demonstrate LSC activity within ETPs, a lympho-myeloid progenitor population.

3.3: Experimental confirmation that acute leukaemia can be propagated by ETPs

ETP-ALL was identified as a subset of T-ALL showing a gene expression profile suggesting that it initiates within ETPs (Coustan-Smith et al., 2009). However, experimental evidence that leukaemogenic mutations targeted to ETPs can initiate transformation within these cells has thus far been lacking. This is an important step toward validating the concept of “ETP leukaemias.” Based on our current knowledge of ETPs and the haematopoietic tree, a number of predictions can be made

about the characteristics of a bona fide ETP leukaemia. Firstly, it would show transcriptional and/or cell surface expression of both the T-lymphoid and myeloid lineages, since ETPs retain differentiation potential for both of these pathways (Bell and Bhandoola, 2008; Wada et al., 2008). The most primitive ETPs have also been shown to retain limited B-lymphoid potential (Luc et al., 2012), but whether B-cell leukaemias might arise within an ETP leukaemia population remains untested. Secondly, transcriptional analysis using multidimensional scaling methods such as principal component analysis would place leukaemia cells closer to normal ETPs than any other progenitor population. Thirdly, LSCs would show phenotypic characteristics of normal ETPs, meaning that purified ETPs would have the potential to propagate the leukaemia.

We have demonstrated that DKOITD leukaemia fully satisfies at least the first and third of these criteria. Even without *Flt3-ITD*, DKO ETPs showed upregulation of HSC and myeloid transcriptional pathways by RNA-sequencing, and co-expression of HSC, myeloid and lymphoid transcription within single cells by qPCR. These features were retained in DKOITD ETPs, which did not show large-scale transcriptional differences to DKO ETPs. Furthermore, DKOITD leukaemias showed both myeloid and early T-cell transcriptional and protein expression in the bone marrow. Finally, stringently purified ETPs from thymuses of DKOITD mice were capable of propagating the leukaemia within wild-type recipients, demonstrating their LSC potential.

Concerning the second criteria listed above, it can be argued that DKOITD leukaemia does not fulfil the property of resembling at the transcriptional level normal ETPs more closely than any other progenitor population. DKOITD ETPs, which can act as LSCs in this leukaemia, clustered more closely to WT LMPPs than WT ETPs using principal component analysis of single-cell qPCR data. This could be explained by the fact that *Rag1Cre* first becomes active in a small proportion of LMPPs (Boiers et al., 2013), causing inactivation of *Ezh2* and *Runx1* in these cells, while *Flt3-ITD* is already expressed before this stage (though not in true HSCs) (Mead et al., 2017). The very short latency and full penetrance of DKOITD leukaemia, together with the results of our exome sequencing analysis,

indicate that these three mutations are sufficient to initiate the leukaemia. However, our data do not allow us to distinguish between two models of where this leukaemia initiation occurs. First, DKOITD leukaemias may initiate within *bone fide* ETPs in the thymus. Second, leukaemia may initiate within LMPPs in the bone marrow, followed by migration to the thymus. DKOITD ETPs could therefore simply represent transformed LMPPs that have migrated from the BM to the thymus. However, the fact that DKOITD ETPs do form a distinct cluster from DKOITD LMPPs argues that the picture is not this simple; ETPs remain an independent though related population to LMPPs in DKOITD mice. This supports the hypothesis that DKOITD ETPs are indeed a true LSC population, providing the first experimental confirmation that ETPs can become LSCs.

This leaves open the question of whether DKOITD leukaemia is truly able to initiate within ETPs, rather than only be propagated by ETPs. Since *Rag1Cre* first becomes active in only a proportion of cells upstream of ETPs, it is possible that there are ETPs which are WT upon entry to the thymus, activate *Rag1Cre* and only then transform into LSCs. This would prove convincingly that acute leukaemia can indeed initiate within ETPs, but we are unable to confirm whether this definitely occurs using this model. An alternative approach could be to use a *Cre* line which activates only in thymocyte progenitors and not bone marrow cells, such as *LckCre*. However, this could not confirm that transformation occurs within true ETPs rather than subsequent populations, since ETPs have by definition only very recently arrived in the thymus. Our data therefore come as close to experimentally proving the initiation of acute leukaemia within ETPs as is currently feasible.

In summary, we have developed a novel mouse model of ETP leukaemia by targeting *Ezh2* and *Runx1* inactivation to early lymphoid progenitors using *Rag1Cre*, in combination with *Flt3-ITD*. Mice show a marked expansion of stringently-defined ETPs within the thymus, which are able to act as LSCs by propagating the leukaemia in wild-type recipient mice. The leukaemia shows co-expression of lymphoid, myeloid and HSC-associated genes, and an overall transcriptional profile closely

mirroring that of human ETP-ALL. Our findings provide experimental support for the concept of ETP leukaemia as a subset of T-ALL and establish a tractable model for preclinical studies of ETP-ALL.

Chapter 5: EZH2-Inactivated Early Thymic Progenitor Leukaemia Can Be Targeted Using BET Inhibition

1: Introduction

As described in Chapter 4, we established a novel mouse model of ETP leukaemia using conditional inactivation of *Ezh2* and *Runx1* in early lymphoid progenitors combined with constitutively active Ras pathway signalling via the *Flt3-ITD* mutation knocked in to the endogenous *Flt3* locus. We decided to use this model to investigate potential novel therapeutic strategies for human ETP-ALL, which has been associated with resistance to conventional treatment (Coustan-Smith et al., 2009; Jain et al., 2016). To do so we chose to explore the transcriptional consequences of loss of *Ezh2* in this model as well as in human ETP leukaemia. We then decided to test whether PRC2-inactivated ETP leukaemias were sensitive to BET inhibition. This particular focus of study is based on a previous report in nerve sheath tumours that PRC2 loss conferred sensitivity to BET inhibition, specifically mediated through reversal of Ras pathway activation, which was notably also a prominent feature of our mouse model of ETP leukaemia (De Raedt et al., 2014).

De Raedt et al showed that in a mouse model of myelin peripheral nerve sheath tumours driven by activation of the Ras pathway via NF1 mutation, additional inactivation of PRC2 confers sensitivity to the BET inhibitor class of compounds (De Raedt et al., 2014). This was proposed to occur because PRC2 inactivation causes loss of H3K27me3 at functionally relevant loci, leading to reciprocal gain of H3K27ac. This activates gene transcription in part via recruitment of BET proteins such as Brd4, which binds to H3K27ac and recruits transcriptional machinery (Filippakopoulos et al., 2012; Loven et al., 2013; Shi and Vakoc, 2014). Inhibition of Brd4 binding to acetylated chromatin may therefore in some contexts reverse oncogenic gene transcription associated with PRC2 inactivation. Exploring whether this mechanism could also occur in PRC2-inactivated ETP leukaemia was a key aim for the part of our project described in this chapter.

A number of BET inhibitors have now been developed. Two of the earliest were JQ1 and I-BET151 (Dawson et al., 2011; Filippakopoulos et al., 2010). These compounds are structurally distinct, but both bind competitively to the chromatin binding region of BET proteins. This prevents their recruitment to acetylated chromatin such as at H3K27ac. The therapeutic benefit of BET inhibitors, which has been demonstrated in several cancer models, is thought to derive in part from a global downregulation of MYC-associated proliferative transcription (Dawson et al., 2011; Delmore et al., 2011; Loven et al., 2013; Zuber et al., 2011).

In this Chapter, *Rag1Cre* is the only *Cre* line used for inactivation of *Ezh2* and *Runx1*. Therefore, mouse genotypes that were referred to as Rag1-WT, Rag1-Ezh2KO, Rag1-Runx1KO and Rag1-DKO in Chapter 3 are in this Chapter referred to simply as WT, Ezh2-KO, Runx1-KO and DKO respectively.

2: Results

2.1: Loss of H3K27me3 and gain of H3K27ac correlates with upregulated gene expression in DKOITD relative to WT bone marrow cells

To gain insight into the epigenetic mechanisms underlying gene expression changes in our DKOITD mouse model, we decided to perform chromatin immunoprecipitation followed by sequencing (ChIP-sequencing) in WT and DKOITD bone marrow cells, and cross-compare these data to our previously obtained RNA-sequencing data from $Mac1^+$ bone marrow cells, which account for the majority of bone marrow mononuclear cells in leukaemic mice (Figure 5.01A). We performed ChIP-sequencing for H3K27me3 and H3K27ac in bone marrow cells from three WT and five DKOITD leukaemic mice. We used magnetic bead-based selection for Kit expression in WT samples to enrich for stem and progenitor cells. DKOITD bone marrow has a high proportion of Kit^+ cells without selection, so we used unsorted bone marrow for DKOITD samples. This meant that DKOITD samples would contain a mixture of cells in which *Rag1Cre*-mediated inactivation of *Ezh2* and *Runx1* had

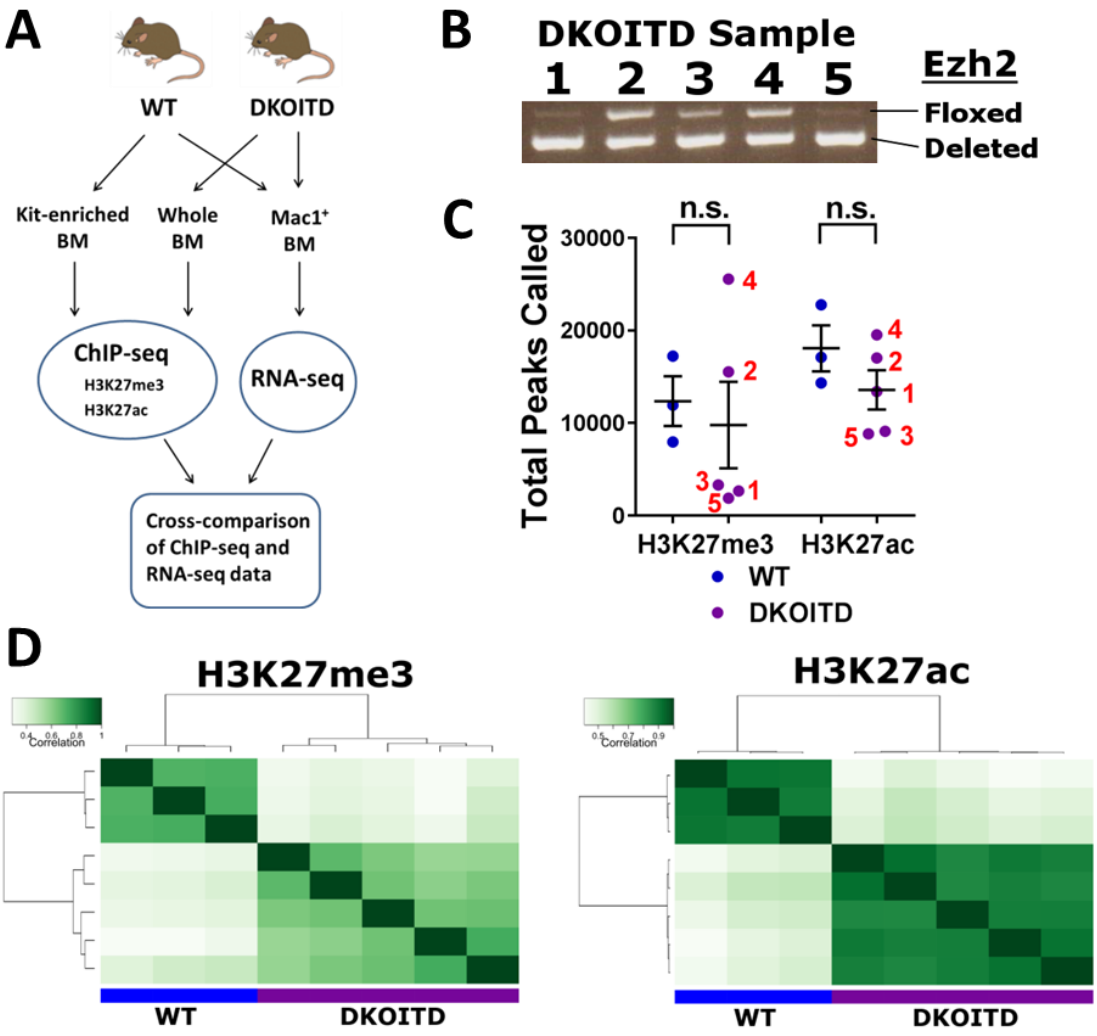


Figure 5.01

(A) Experimental design for ChIP-seq in WT and DKOITD bone marrow cells and cross-comparison with RNA-sequencing data. ChIP-seq n=3 per histone mark for WT, n=5 per histone mark for DKOITD; RNA-sequencing n=3 per genotype. (B) PCR analysis of relative abundance of floxed and deleted *Ezh2* alleles in DKOITD samples. Sample numbers correspond to those in panel (C). (C) Total number of peaks called genome-wide by Macs2 for each ChIP-seq sample. Red figures indicate DKOITD sample numbers shown in panel (B). (D) Hierarchical clustering and correlation between ChIP-seq samples, using peaks called by Macs2 combined with read density at each peak. Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

taken place, and cells in which *Rag1Cre* had never been active. For this reason we selected DKOITD mice which had already reached the stage of overt leukaemia, since we could assume that at this stage a high proportion of the non-recombined bone marrow cells had been displaced by leukaemic, *Ezh2* and *Runx1* inactivated cells. To confirm this, we performed PCR analysis of the *Ezh2* locus to estimate the relative abundance of floxed and deleted *Ezh2* alleles in each DKOITD sample (Figure 5.01B). This revealed some variability between the samples, but in all cases the band corresponding to the deleted *Ezh2* allele was considerably brighter than that of the floxed (non-deleted) allele.

Our initial analysis of peaks called relative to input controls using Macs2 (Zhang et al., 2008) showed that there was no significant difference in the numbers of H3K27me3 or H3K27ac peaks called per sample between WT and DKOITD (Figure 5.01C). It might have been expected that loss of *Ezh2* would lead to a global reduction in H3K27me3 and therefore fewer peaks called in the DKOITD samples compared to WT. However, H3K27me3 peaks may simply be reduced in size in DKOITD samples, or alternatively H3K27me3 may be redistributed through the genome. There may also be a degree of functional redundancy with Ezh1, which is able to take the place of Ezh2 within the PRC2 complex in some haematopoietic cell contexts (Hidalgo et al., 2012; Mochizuki-Kashio et al., 2015; Mochizuki-Kashio et al., 2011).

We therefore used DiffBind (Stark and Brown, 2011) to perform an analysis to determine regions of the genome where H3K27me3 and H3K27ac are significantly gained or lost in WT samples relative to DKOITD samples. DiffBind takes all regions of the genome covered by a Macs2-called peak in any of the samples and performs an EdgeR analysis, as used for RNA-sequencing data (Robinson et al., 2010), to determine which of these regions contain significantly more reads in one genotype than the other. Hierarchical clustering using peaks called by Macs2 combined with read count information at each peak clustered samples by genotype for both H3K27me3 and H3K27ac (Figure 5.01D).

We next performed DiffBind's EdgeR analysis to find peaks showing significantly different intensity of H3K27ac or H3K27me3 between WT and DKOITD samples. Genome-wide, approximately the

same number of peaks showed increased H3K27ac in WT relative to DKOITD as showed decreased H3K27ac, and the same was true for H3K27me3 (Figure 5.02A). To extract functional information from these differential peaks we mapped them to nearby transcription start sites (TSSs) using Homer (Heinz et al., 2010) and discarded all regions further than 2.5 kb from a TSS. Using this data we generated lists of all genes showing significantly increased or decreased H3K27me3 or H3K27ac within a 5 kb window of the TSS in DKOITD samples relative to WT. We found that there were approximately twice as many genes showing decreased H3K27me3 in DKOITD relative to WT as genes showing increased H3K27me3, and twice as many genes showing increased H3K27ac as genes showing decreased H3K27ac (Figure 5.02B). About half of the genes showing decreased H3K27me3 also showed increased H3K27ac (Figure 5.02C). These data are consistent with decreased H3K27me3 as a result of *Ezh2* inactivation occurring more prominently at TSSs than in the genome as a whole. They also indicate that loss of H3K27me3 is associated with a reciprocal gain in H3K27ac at TSSs.

We next sought to validate the concept that loss of H3K27me3 is associated with a reciprocal increase in H3K27ac at functionally relevant loci, even in the absence of mutational activation of histone acetyltransferase proteins. To do this, we combined sequencing reads from all three WT and all five DKOITD samples and inspected the distribution of H3K27me3 around TSSs genome-wide, after normalising for library sizes. This revealed a loss of H3K27me3 within 2 kb up and downstream of TSSs in DKOITD relative to WT cells (Figure 5.02D). Doing the same for H3K27ac also showed a general loss around TSSs genome-wide (Figure 5.02D). However, selecting only genes showing a significant loss of H3K27me3 revealed a general increase in H3K27ac at the TSSs of these genes (Figures 5.02E and F). This confirms that loss of H3K27me3 is associated with a reciprocal increase in H3K27ac at TSSs in bone marrow stem and progenitor cells. The increase in H3K27ac could be a direct consequence of decreased H3K27me3 (Pasini et al., 2010), although our data do not demonstrate a causal relationship.

A

All differential peaks	H3K27ac	H3K27me3
Increased in DKOITD	5313	995
Decreased in DKOITD	5534	1020

B

Genes with differential peaks at TSS	H3K27ac	H3K27me3
Increased in DKOITD	2643	247
Decreased in DKOITD	1245	517

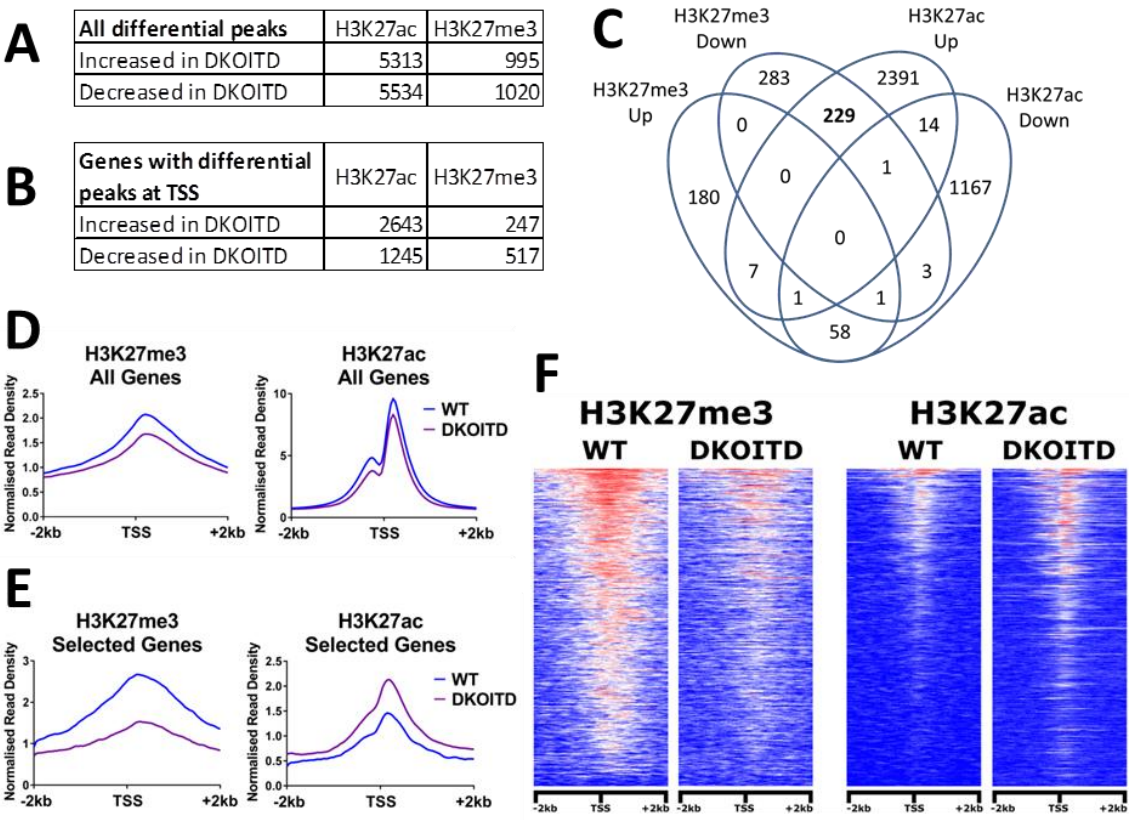


Figure 5.02

(A) Number of peaks genome-wide showing significantly different H3K27ac and H3K27me3 signal between WT (n=3) and DKOITD (n=5) samples as determined by DiffBind analysis. (B) Number of genes with peaks within 2.5 kb of the TSS showing significantly different H3K27ac and H3K27me3 signal between WT (n=3) and DKOITD (n=5) samples. (C) Overlap of genes shown in (B); each gene may have multiple differential peaks within 2.5 kb of the TSS. “Up” indicates increased H3K27me3 or H3K27ac signal in DKOITD relative to WT; “Down” indicates decreased signal in DKOITD relative to WT. (D) Normalised read density within a 4 kb window of all TSSs genome-wide for pooled WT and DKOITD libraries. (E) Normalised read density within a 4 kb window of all TSSs showing decreased H3K27me3 in DKOITD relative to WT as determined by DiffBind analysis, for pooled WT and DKOITD libraries. (F) Heatmaps, for pooled WT and DKOITD libraries, showing read density within a 4 kb window of the TSS for all genes showing decreased H3K27me3 in DKOITD relative to WT as determined by DiffBind analysis, ordered by average H3K27me3 read density in WT.

We next decided to determine whether loss of H3K27me3 and gain of H3K27ac near the TSS correlates with upregulated gene expression in DKOITD cells. To do this we cross-compared our ChIP-sequencing data with our RNA-sequencing data from WT and DKOITD $Mac1^{+}$ bone marrow cells. We found that of the 230 genes showing both loss of H3K27me3 and gain of H3K27ac, 62 showed significantly upregulated gene expression, compared with only 1 gene showing downregulated expression (Figure 5.03A). This overlap was highly significant when comparing to the overlap that would be expected by chance ($p=2.44 \times 10^{-12}$ using the hypergeometric distribution and number of genes detected by RNA-sequencing as the total number of genes). Examples of gene loci showing this pattern very clearly are displayed in Figure 5.03B. Notably, Flt3 showed very clear loss of H3K27me3 and gain of H3K27ac at the TSS, which correlated with strong upregulation of expression (Figures 5.03B and 4.10D). This provides an example of an epigenetic mutation (*Ezh2*) causing increased expression of a collaborating signalling mutation (*Flt3-ITD*), potentially driving further proliferative signalling.

When using the 230 genes showing loss of H3K27me3 and gain of H3K27ac as a gene set for GSEA, we found that these genes were highly enriched among genes upregulated in DKOITD relative to WT cells using RNA-sequencing (Figure 5.03C). Together these results validate the hypothesis that loss of H3K27me3 at TSSs as a result of *Ezh2* inactivation in DKOITD cells is associated with a reciprocal increase in H3K27ac, which correlates with upregulated gene expression. We hypothesised that the genes upregulated as a result of this epigenetic switch could be important oncogenic drivers of leukaemogenesis in this mouse model. Since BET inhibitors prevent recruitment of transcriptional apparatus at acetylated chromatin loci, they could potentially act to reduce upregulation of this oncogenic gene expression. Additionally, BET inhibition has previously shown preclinical efficacy in a PRC2-deficient nerve sheath tumour model, which like our DKOITD model was associated with upregulated Ras pathway signalling. We therefore decided to explore whether BET inhibition could be used as a therapeutic strategy for PRC2-deficient ETP leukaemias.

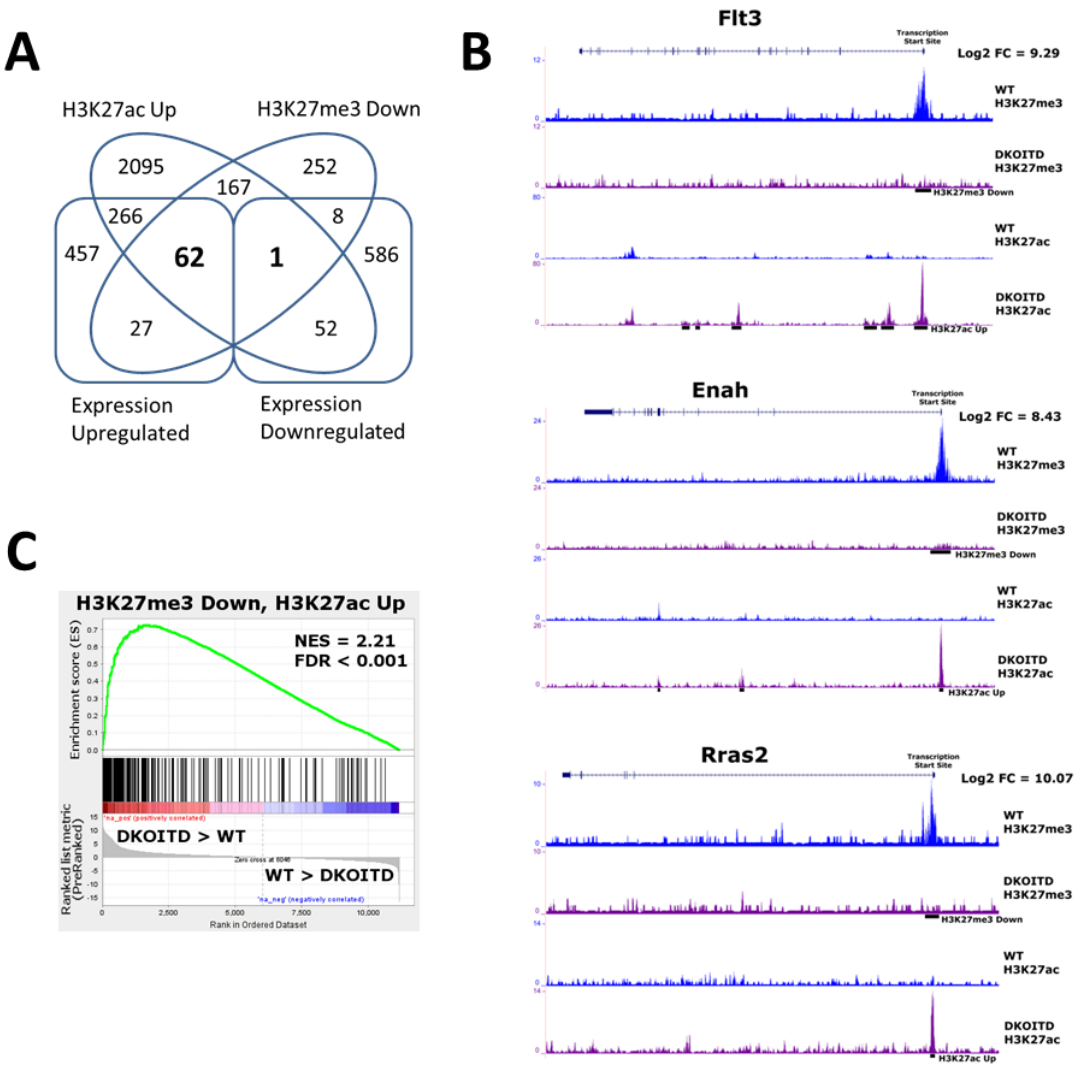


Figure 5.03

(A) Overlap of genes showing significant differential H3K27me3 or H3K27ac signal in DKOITD relative to WT as determined by ChIP-seq and DiffBind analysis, with differentially expressed genes in Mac1^{+} bone marrow cells as determined by RNA-sequencing. “Up” indicates increased ChIP-seq signal or gene expression in DKOITD relative to WT; “Down” indicates decreased ChIP-seq signal or gene expression in DKOITD relative to WT. (B) Genome browser tracks showing H3K27me3 and H3K27ac signal at the *Flt3*, *Enah* and *Rras2* loci in WT and DKOITD bone marrow. Fold change in expression in DKOITD relative to WT Mac1^{+} bone marrow cells is indicated for each gene. (C) GSEA comparing WT and DKOITD Mac1+ bone marrow cells for genes which lose H3K27me3 and gain H3K27ac (n=230) in DKOITD relative to WT bone marrow cells.

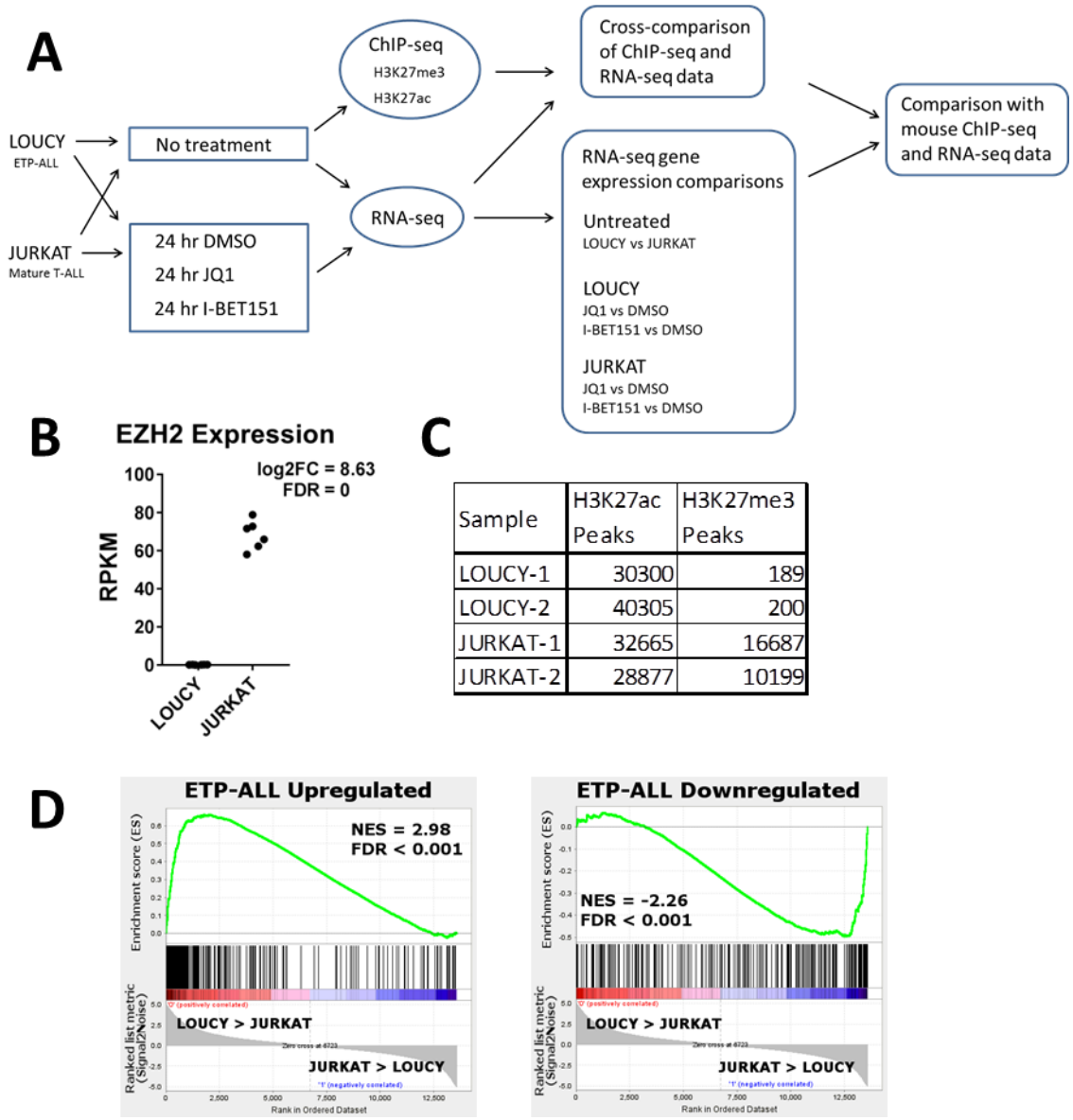
2.2: BET inhibition targets oncogenic transcription in human *EZH2*-inactivated ETP leukaemia

We first decided to investigate whether active transcription associated with lack of H3K27me3 and presence of H3K27ac could be downregulated by BET inhibition. For these experiments we used LOUCY cells, a human ETP leukaemia cell line. This is a T-ALL cell line, established before the identification of ETP-ALL, which has since been shown using gene expression analysis to be likely to have been derived from an ETP-ALL patient (Van Vlierberghe et al., 2011). It has also been shown to lack *EZH2* expression (Nagel et al., 2010), making it particularly relevant to our mouse ETP leukaemia model. We also chose to use JURKAT cells as a mature non-ETP T-ALL cell line for comparison with LOUCY.

The experiments we performed using the LOUCY and JURKAT cell lines are outlined in Figure 5.04A. We performed RNA-sequencing and ChIP-sequencing for H3K27me3 and H3K27ac in untreated cells, and RNA-sequencing on cells treated for 24 hr with DMSO, JQ1 or I-BET151.

We first validated that LOUCY cells do not express *EZH2*, in contrast to JURKAT cells, using our RNA-sequencing data (Figure 5.04B), and also, in contrast to the mouse leukaemias, that LOUCY cells display minimal H3K27me3 genome-wide (Figure 5.04C). We also validated previous findings that LOUCY cells display an ETP-ALL gene expression signature (Van Vlierberghe et al., 2011), by comparing their gene expression to JURKAT and performing GSEA. This confirmed that LOUCY cells show high expression of genes upregulated in ETP-ALL relative to non-ETP T-ALL, and low expression of downregulated genes (Figure 5.04D). LOUCY cells therefore represent a good cell-line model of human *EZH2*-inactivated ETP leukaemia.

To determine how the overall transcriptional profiles of untreated and BET inhibitor-treated LOUCY and JURKAT cells relate to each other, we performed multidimensional scaling analysis on all of the RNA-sequencing samples (Figure 5.05A). This showed that, as expected, the first component



Sample	H3K27ac Peaks	H3K27me3 Peaks
LOUCY-1	30300	189
LOUCY-2	40305	200
JURKAT-1	32665	16687
JURKAT-2	28877	10199

Figure 5.04

(A) Experimental design for ChIP-seq and RNA-seq experiments in LOUCY and JURKAT cells. ChIP-seq n=2 per cell line per histone mark, RNA-seq n=6 per cell line (untreated cells) and n=4 per condition (DMSO, JQ1 and I-BET151-treated LOUCY and JURKAT cells). (B) RPKM expression levels of *EZH2* in LOUCY and JURKAT cells. (C) Total number of peaks called genome-wide by Macs2 for each ChIP-seq sample. (D) GSEA comparing LOUCY and JURKAT cells for genes upregulated (left; n=332) or downregulated (right; n=278) in human ETP-ALL relative to non-ETP T-ALL.

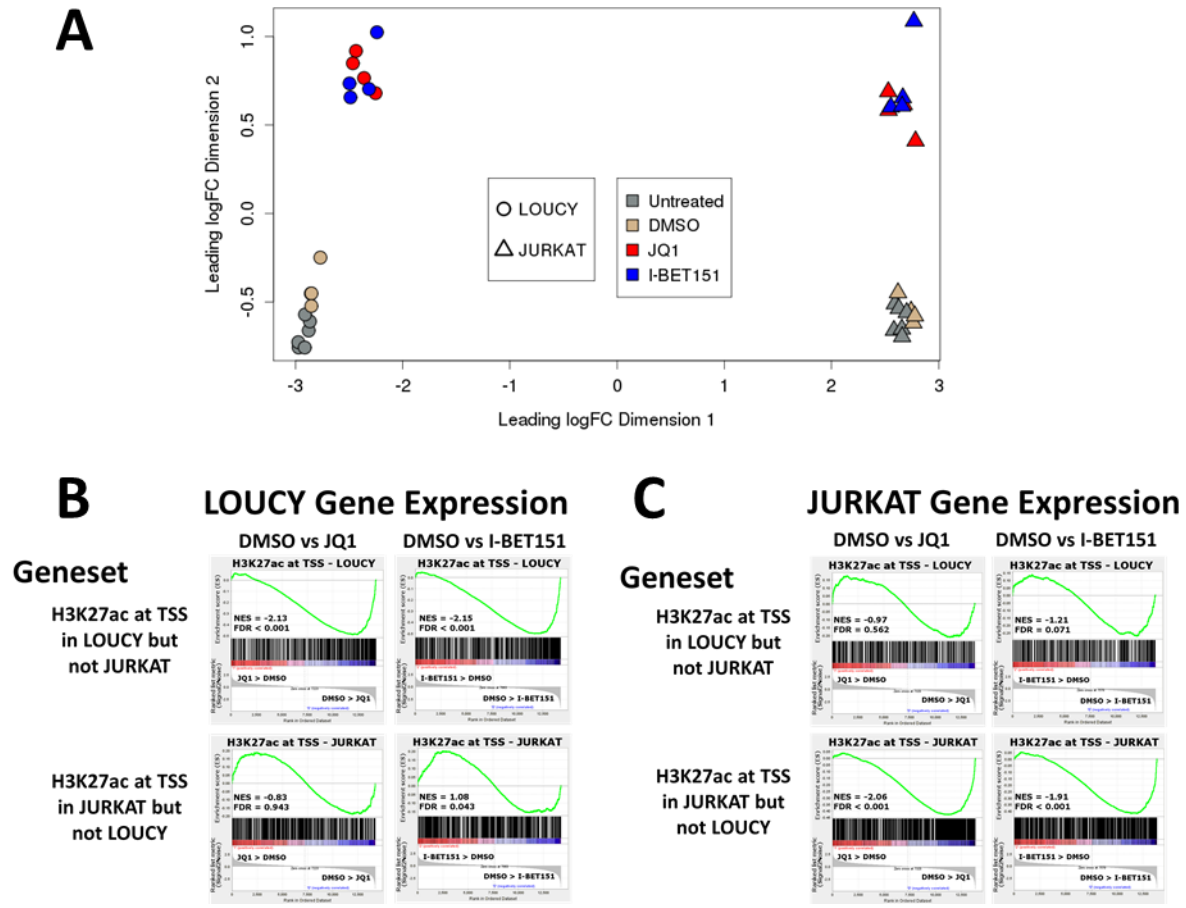


Figure 5.05

(A) Multidimensional scaling analysis of RNA-seq data from untreated LOUCY and JURKAT cells (n=6 per cell line) and DMSO, JQ1 and I-BET151-treated LOUCY and JURKAT cells (n=4 per cell line per condition). Cells were treated for 24 hr with 0.05% DMSO, 250 nM JQ1 or 500 nM I-BET151. (B-C) GSEA plots comparing gene expression in cells treated with DMSO and either JQ1 (left panels) or I-BET151 (right panels). Two genesets were used: genes showing an H3K27ac peak within 1 kb of the TSS in LOUCY but not JURKAT (n=796), and genes showing the same in JURKAT but not LOUCY (n=1044). (B) shows gene expression in LOUCY cells and (C) shows gene expression in JURKAT cells.

separated all LOUCY samples from all JURKAT samples. The second component separated untreated and DMSO-treated samples from JQ1 and I-BET151-treated samples, for both LOUCY and JURKAT. This result confirms that DMSO treatment did not cause significant transcriptional changes, while JQ1 and I-BET151 caused generally similar transcriptional changes.

We next wished to use the LOUCY model to investigate whether BET inhibition reduces the transcription of genes associated with H3K27ac at the TSS, since genes showing this epigenetic mark were upregulated in our DKOITD model and we reasoned were likely to be key drivers of oncogenesis. We selected genes showing H3K27ac at the TSS in LOUCY but not JURKAT cells and performed GSEA using a comparison of DMSO-treated and BET inhibitor-treated LOUCY cells. This showed that H3K27ac-associated genes were strongly downregulated by both JQ1 and I-BET151 (Figure 5.05B). These genes were not downregulated in JURKAT cells (Figure 5.05C). Performing the same comparison for genes associated with H3K27ac in JURKAT but not LOUCY cells conversely showed downregulation by BET inhibitors in JURKAT but not LOUCY (Figures 5.05B and C). This demonstrates that BET inhibition causes a selective downregulation of genes associated with H3K27ac at the TSS.

BET inhibition has been proposed to reduce proliferation by inhibition of the MYC pathway (Dawson et al., 2011; Delmore et al., 2011; Loven et al., 2013; Zuber et al., 2011). Performing GSEA using the Hallmark database (Liberzon et al., 2015) for gene expression in JQ1-treated relative to DMSO-treated LOUCY cells indeed revealed downregulation of genes associated with the MYC pathway, as well as a number of signalling pathways including KRAS, Interferon alpha and gamma, IL2-STAT5 and IL6-JAK-STAT3 (Figure 5.06A). All of these gene sets were also downregulated in JQ1-treated JURKAT cells, with the exception of IL6-JAK-STAT3 (Figure 5.06B). This could indicate a mechanism of action of BET inhibition specific to PRC2-inactivated ETP-ALL. Notably, a study of ETP-ALL xenograft samples previously showed strong STAT3 activation in ETP-ALLs compared to non-ETP T-ALLs (Maude et al., 2015).

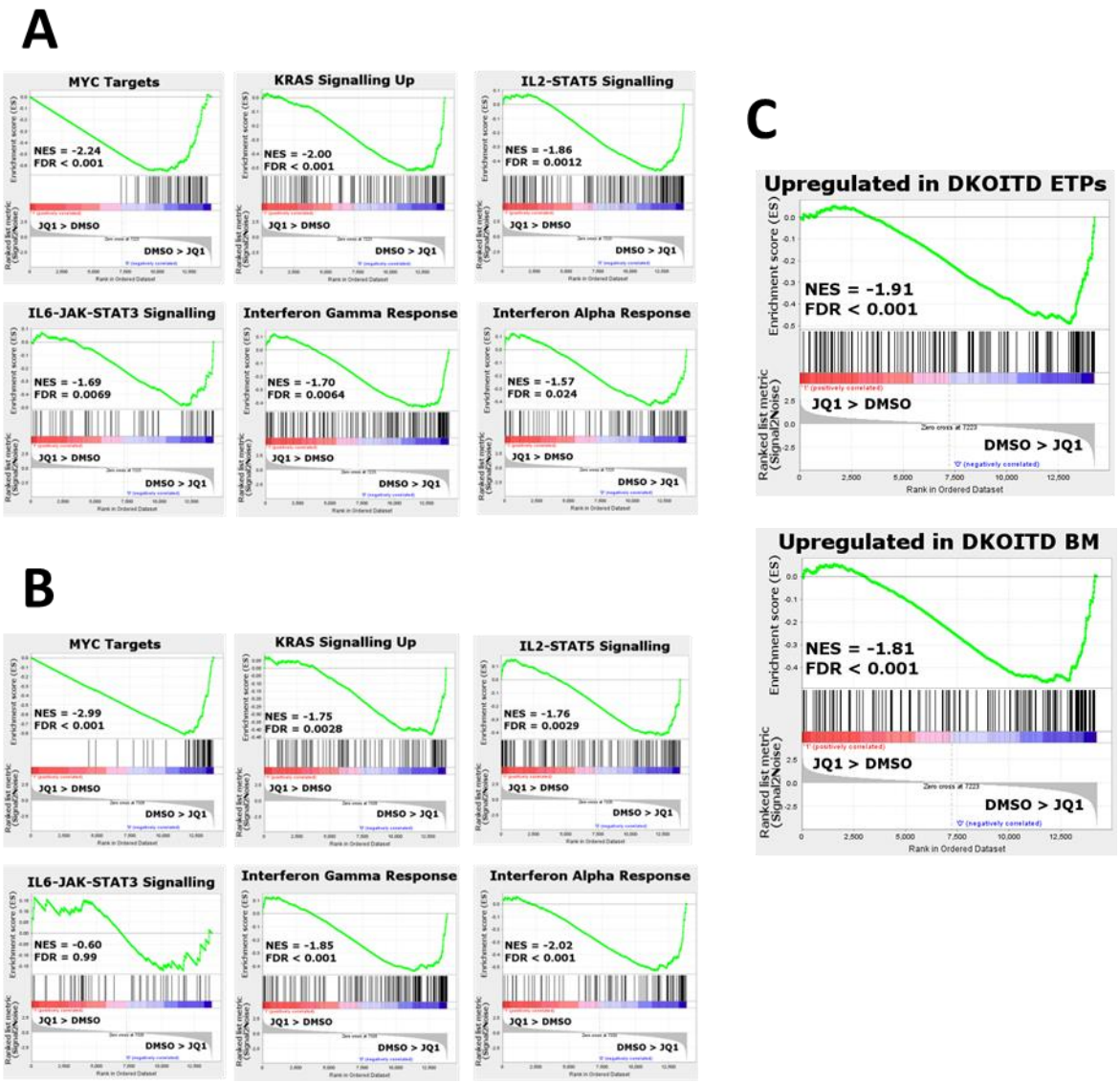


Figure 5.06

(A-B) GSEA plots comparing DMSO and JQ1-treated LOUCY (A) or JURKAT (B) cells for indicated Hallmark genesets. No enrichment was detected for IL6-JAK-STAT3 signalling in JURKAT cells. (C) GSEA plots comparing DMSO and JQ1-treated LOUCY cells for genes upregulated in (top) DKOITD relative to WT ETPs (logFC>2; n=262 genes) and (bottom) DKOITD relative to WT Mac1^{+} bone marrow cells (logFC>8; n=219 genes).

We previously observed that these signalling pathway gene sets were upregulated in DKO relative to WT ETPs (Figures 4.03D and E). This raised the possibility that BET inhibition could downregulate PRC2 inactivation-associated oncogenic gene expression in DKOITD leukaemia. To investigate this further, we compared the genes most strongly downregulated by JQ1 in LOUCY cells with the expression profiles of DKOITD relative to WT cells. This showed that JQ1-downregulated genes were strongly enriched among genes overexpressed both in DKOITD relative to WT ETPs, and DKOITD relative to WT $Mac1^+$ bone marrow cells (Figure 5.06C). These results suggest that BET inhibition could be a viable method of reversing aberrant gene expression associated with PRC2 inactivation. This could potentially therefore also target aberrant leukaemia proliferation by downregulating oncogenic gene expression, in both mouse DKOITD ETP leukaemia and human PRC2-inactivated ETP-ALL.

2.3: BET inhibition selectively targets proliferation of DKOITD leukaemia

We investigated whether DKOITD leukaemia could be therapeutically targeted using BET inhibition. We first tested the effect of increasing concentrations of JQ1 on sorted ETPs from DKOITD mice (Figure 5.07A), and compared to a number of human leukaemia cell lines (Figure 5.07B). The NF1, THP-1 and HL-60 cell lines have been previously shown to be highly sensitive to BET inhibition, while K562 and JURKAT have been shown to be relatively resistant (Zuber et al., 2011). We confirmed these data and also found that LOUCY cells were highly sensitive to BET inhibition using JQ1. The anti-proliferative effect of JQ1 on DKOITD ETPs was comparable to that of the sensitive human cell lines (IC_{50} = 49 nM for ETPs, IC_{50} = 57-95 nM for sensitive cell lines). This suggests that DKOITD LSCs could be targeted using BET inhibition.

We next decided to make a direct comparison of the sensitivity to BET inhibition of DKOITD relative to WT cells, to determine whether these compounds might selectively target leukaemia cells. Since

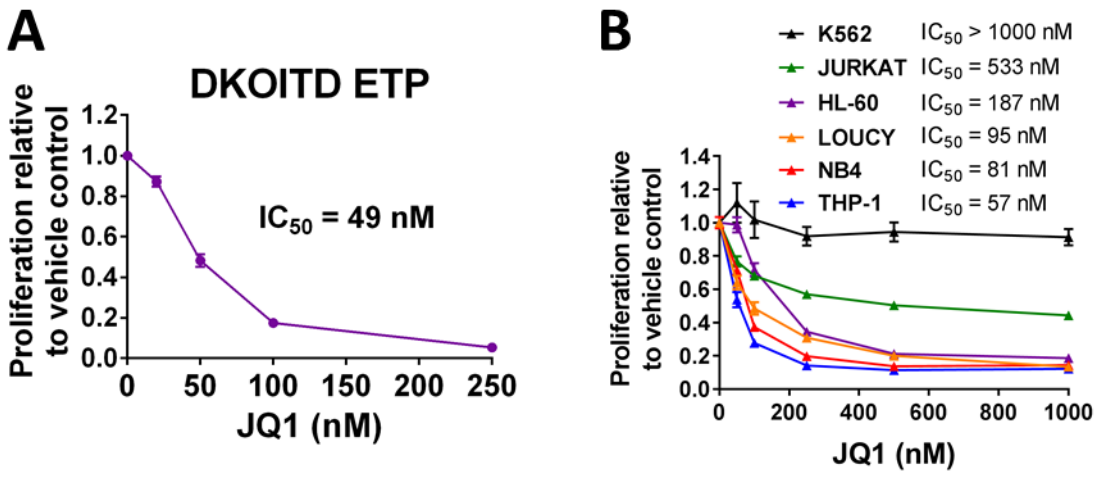


Figure 5.07

(A) MTS proliferation assay of DKOITD ETPs cultured with increasing concentrations of JQ1, shown relative to DMSO control (n=3 for each concentration). (B) MTS proliferation assay of leukaemia cell lines cultured with increasing concentrations of JQ1, shown relative to DMSO control (n=6 for each concentration). Data were collected from two independent experiments. Error bars represent standard error of the mean.

WT ETPs are far less abundant (Figure 4.10A) and proliferative (Figure 4.11A) than DKOITD ETPs, we decided to compare Lin^-Kit^+ bone marrow stem and progenitor cells from WT mice with those from leukaemic DKOITD ETP transplant recipient mice. This showed that proliferation of DKOITD Lin^-Kit^+ bone marrow cells was significantly more strongly reduced by both JQ1 (Figure 5.08A) and I-BET151 (Figure 5.08B) than that of the equivalent WT cells. These data show that DKOITD leukaemia is selectively targeted by BET inhibition *in vitro*.

To confirm this result *in vivo*, we transplanted DKOITD bone marrow cells into wild-type recipient mice and performed daily administration of either JQ1 or DMSO over a 4-week period. Probably in part due to the fact that chimerism of CD45.2 DKOITD relative to CD45.1 wild-type peripheral blood $Mac1^+$ cells was already very high at the start of the treatment period, we were unable to demonstrate an effect of JQ1 treatment on the increase in peripheral blood chimerism over time, or an alleviation of anaemia or thrombocytopaenia (Figure 5.09A). However, when analysing mice at the time of culling, either due to showing signs of malignancy or reaching the end of the treatment period, we found that JQ1 caused a significant reduction in the leukaemic burden compared to DMSO. This could be seen when measuring the peripheral blood white blood cell counts, total bone marrow cellularity per femur and tibia, and total spleen cellularity (Figure 5.09B).

We then performed a similar experiment using I-BET151 rather than JQ1, and an oral rather than intra-peritoneal administration route. Mice were given *ad libitum* access to the compound-containing feed (or control feed). This allowed for a more continuous dosing of the compound than could be achieved with daily injections, since the half-lives of JQ1 and I-BET151 *in vivo* are very short (around 1.5 hours) (Dawson et al., 2011; Zuber et al., 2011). I-BET151-treated mice in this experiment showed a similar reduction in leukaemic burden at the end of the treatment period, as measured by peripheral blood white blood cell counts, bone marrow cellularity and spleen weight (Figure 5.10). These data confirm that DKOITD mouse ETP leukaemia can be targeted by BET inhibition *in vivo*.

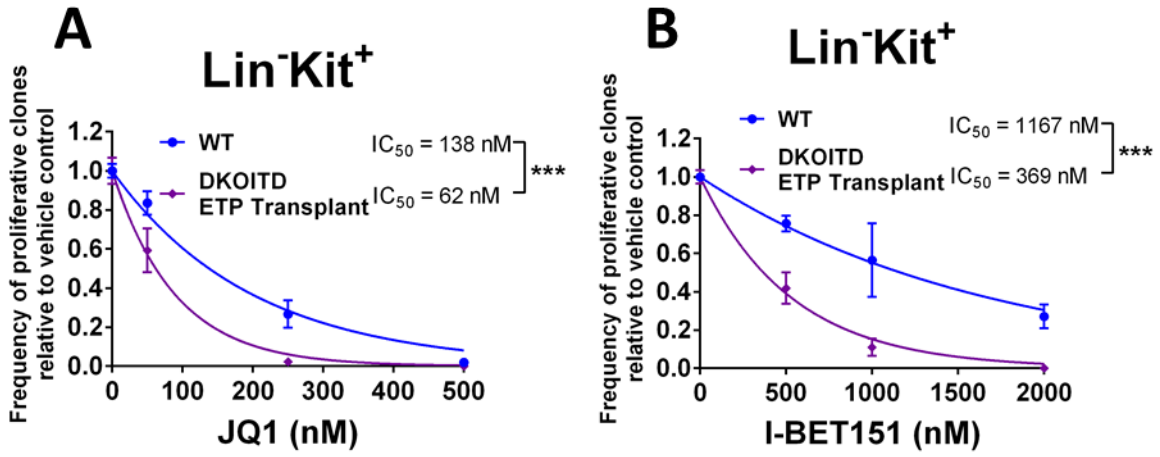


Figure 5.08

(A) Single-cell proliferation assay of CD45.2 Lin^-Kit^+ BM progenitor cells FACS-sorted from recipient mice transplanted with WT BM or DKOITD ETP cells and cultured with increasing concentrations of JQ1. Frequency of clones showing high proliferation at each concentration is shown relative to DMSO control (WT n=3-8 and DKOITD n=4-10 for each concentration). Cells were collected from three WT BM and four DKOITD ETP recipient mice. (B) Single-cell proliferation assay of CD45.2 Lin^-Kit^+ BM progenitor cells FACS-sorted from recipient mice transplanted with WT BM or DKOITD ETP cells and cultured with increasing concentrations of I-BET151. Frequency of clones showing high proliferation at each concentration is shown relative to DMSO control (WT n=3-5 and DKOITD n=4-8 for each concentration). Data were collected from two WT BM and four DKOITD ETP recipient mice. Error bars represent standard error of the mean. P-values are by F-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

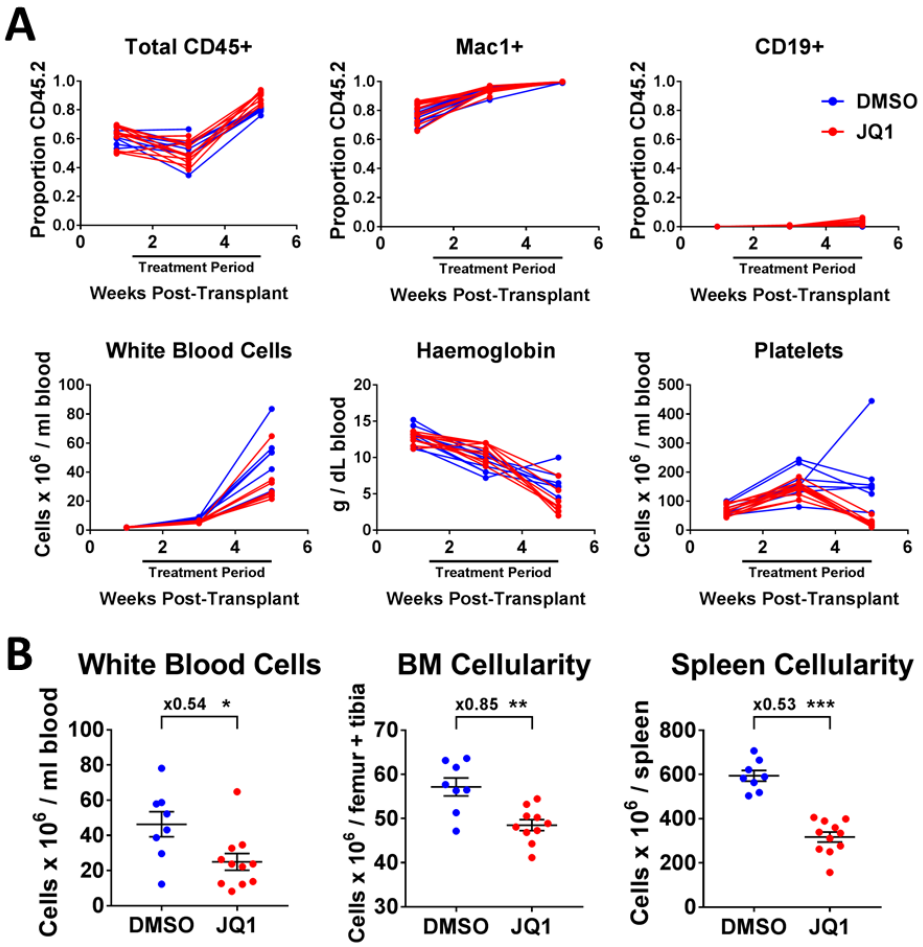


Figure 5.09

(A) Relative contribution of CD45.2 cells to indicated leukocyte populations (top row) and haematological parameters (bottom row) of the peripheral blood of wild-type CD45.1 recipient mice following transplantation of 2 million CD45.2 DKOITD bone marrow cells and treatment with 45 mg/kg/day JQ1 or a similar volume of DMSO for the indicated time period (28 days). Each line represents an individual recipient mouse; DMSO n=8, JQ1 n=12. (B) White blood cell counts and total bone marrow and spleen cellularity of wild-type mice transplanted with DKOITD bone marrow and treated with 45 mg/kg/day JQ1 or a similar volume of DMSO for 28 days. DMSO n=8, JQ1 n=11. One JQ1 mouse was excluded after being withdrawn from treatment due to weight loss. Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

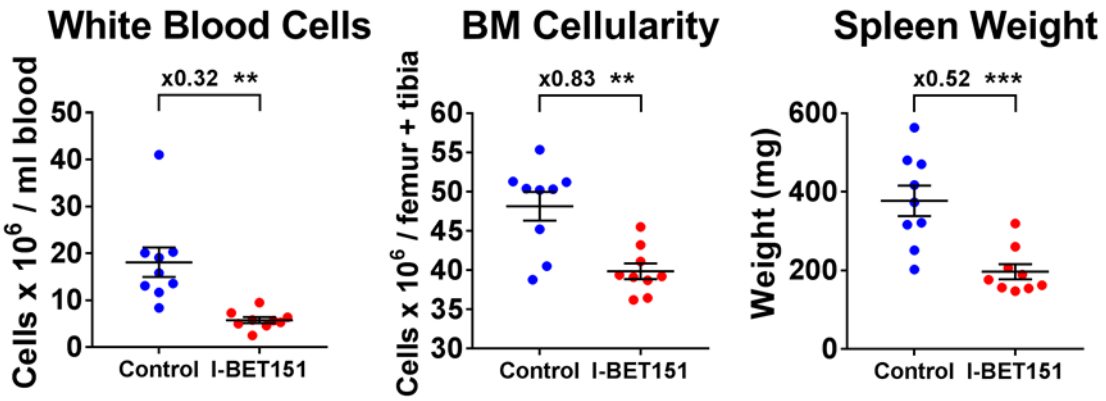


Figure 5.10

White blood cell counts, total bone marrow cellularity and spleen weights of wild-type mice transplanted with DKOITD bone marrow and administered modified murine feed containing 0.018% (180ppm) I-BET151 or control feed. Control n=9, I-BET151 n=9. This experiment was performed by George Giotopoulos. Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

2.4: Human PRC2-inactivated ETP leukaemia is sensitive to BET inhibition *in vitro* and *in vivo*

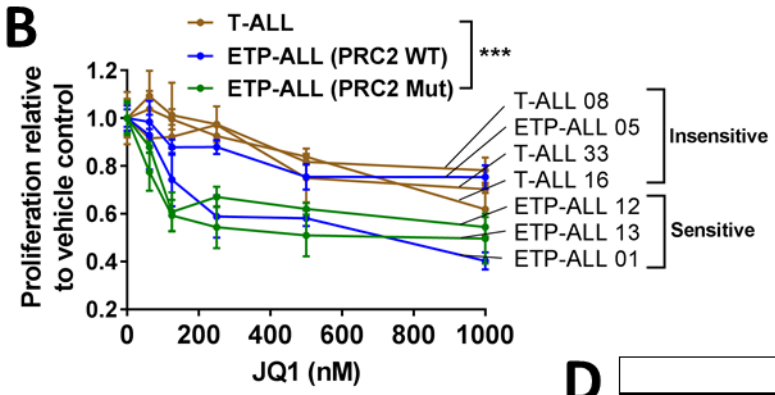
Having established that inactivation of the PRC2 component *Ezh2* confers sensitivity of DKOITD ETP leukaemia to BET inhibition, we decided to investigate whether BET inhibition might be a viable therapeutic strategy for human ETP-ALLs carrying inactivating mutations in PRC2. To do this, we obtained patient-derived xenograft samples of six ETP-ALLs and three non-ETP T-ALLs (Maude et al., 2015). Data on the mutations carried by each sample were available for the ETP-ALL samples, but not the non-ETP T-ALLs. Two of the ETP-ALL samples carried PRC2 inactivating mutations: ETP-ALL 13 carried a mutation in *EED*, while ETP-ALL 12 carried mutations in *EZH2* as well as *RUNX1* (Figure 5.11A, adapted from Maude et al., 2015, Table 1).

We first performed *in vitro* proliferation assays using these samples to test their sensitivity to JQ1 and I-BET151. Two of the PRC2 wild-type ETP-ALL samples did not show sufficient proliferation over the culture period to produce reliable data for the assay and were excluded from analysis. We found that the remaining samples fell into two distinct groups: sensitive or insensitive to BET inhibition (Figures 5.11B and C). These groups were consistent between JQ1 (Figure 5.11B) and I-BET151 (Figure 5.11C) treatment, although the data were less clear for I-BET151. The samples were split across two independent experiments, and the sample sensitivity was independent of the experiment number (Figure 5.11D). All of the non-ETP T-ALL samples belonged to the insensitive group, while three of the ETP-ALLs were sensitive and one was insensitive. Interestingly, both of the PRC2-mutant ETP-ALLs were highly sensitive to BET inhibition. This suggested that our data validating BET inhibition as a therapeutic strategy for PRC2-inactivated mouse ETP leukaemia could be applied to human ETP-ALL.

To confirm these findings *in vivo*, we transplanted cells from the sample ETP-ALL 12, carrying mutations in *EZH2* and *RUNX1* (Figure 5.11A), into NSG mice. We performed peripheral blood flow

A

Sample no.	Age (years)	Gender	WBC ($x10^9$/L)	Genetic Lesions
ETP-ALL 01	8	M	66.1	CTCF SV, DNM2 K557_K558 > K, JAK3 M511I, WT1 R370fs
ETP-ALL 05	15	M	41.8	PTPN11 A72V
ETP-ALL 12	16	M	251	EZH2 S651L, *RUNX1* T148fs, NOTCH1 S2492*, *IKZF1* SV, PHF6 N147fs, SUZ12 C350R, WT1 C350R L564_S568, *SH2B3* SV
ETP-ALL 13	3	M	21.5	EED S259F



D

	Experiment 1	Experiment 2
Insensitive Samples	T-ALL 16	ETP-ALL 05
	T-ALL 33	T-ALL 08
Sensitive Samples	ETP-ALL 01	ETP-ALL 12
	ETP-ALL 13	

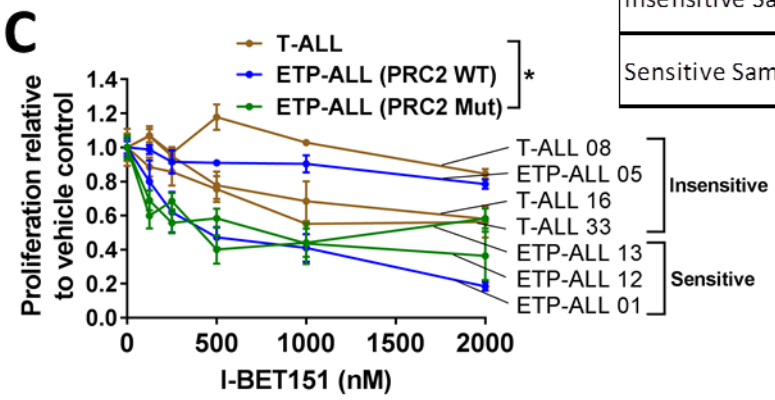


Figure 5.11

(A) Clinical characteristics and genetic lesions of ETP-ALL xenograft samples. Adapted from Maude et al., *Blood* 2015, Table 1. (B-C) MTS proliferation assay of T-ALL and ETP-ALL patient-derived xenografts cultured with increasing concentrations of JQ1 (B) or I-BET151 (C), shown relative to DMSO control (n=3 for each concentration). Data were collected across two independent experiments. (D) Samples used in each experiment. Error bars represent standard error of the mean. P-values are by F-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

cytometry readouts every two weeks, staining leukocytes for human CD45 and mouse CD45 in order to monitor progression of the human leukaemia. We waited until leukaemia engraftment was confirmed by detection of at least 1% human cells among peripheral blood leukocytes before beginning treatment with either JQ1 or DMSO control. This was so that we could rule out any observed effect of the treatment being caused by inhibition of leukaemia engraftment in the mice, rather than inhibition of progression of the leukaemia itself. We could also confirm that the leukaemia was at approximately the same level of progression in all recipient mice at the start of the treatment period. At 4 weeks post-transplant all mice showed robust engraftment, so at 5 weeks post-transplant we began administration of 45 mg/kg/day of JQ1 or a similar volume of DMSO by intra-peritoneal injection. We weighed all experimental mice daily in order to ensure that the treatment protocol was not having adverse effects on the overall health of the mice.

We found that JQ1 treatment markedly slowed progression of the human leukaemia relative to DMSO treatment (Figure 5.12A), with only a minor effect on body weight (Figure 5.12B). Mice were treated continuously for a period of five weeks and were all culled at the end of the treatment period (this is the maximum length of daily intra-peritoneal injection that is allowed on our animal licence). By the end of the treatment period, human to mouse cell chimerism in the peripheral blood was markedly lower in JQ1-treated mice than in DMSO-treated mice (x0.33, $P<0.001$; Figure 5.12A). The total number of leukocytes per millilitre of peripheral blood was also much lower in JQ1-treated mice, which could be accounted for mainly by the reduced number of human cells (x0.20, $P<0.001$; Figure 5.12C, left panel). Overall bone marrow cellularity was also lower in JQ1-treated mice, with about half as many human cells per femur and tibia as in DMSO controls (x0.49, $P<0.001$; Figure 5.12C, middle panel). All mice in fact showed hypocellular bone marrow (12-16 million cells per femur and tibia, compared to around 50-60 million in healthy wild-type mice, see for example Figure 3.03B). DMSO-treated mice displayed pronounced splenomegaly resulting from infiltration of human leukaemia cells (Figure 5.12C, right panel). This was strongly reduced in JQ1-treated mice (x0.17,

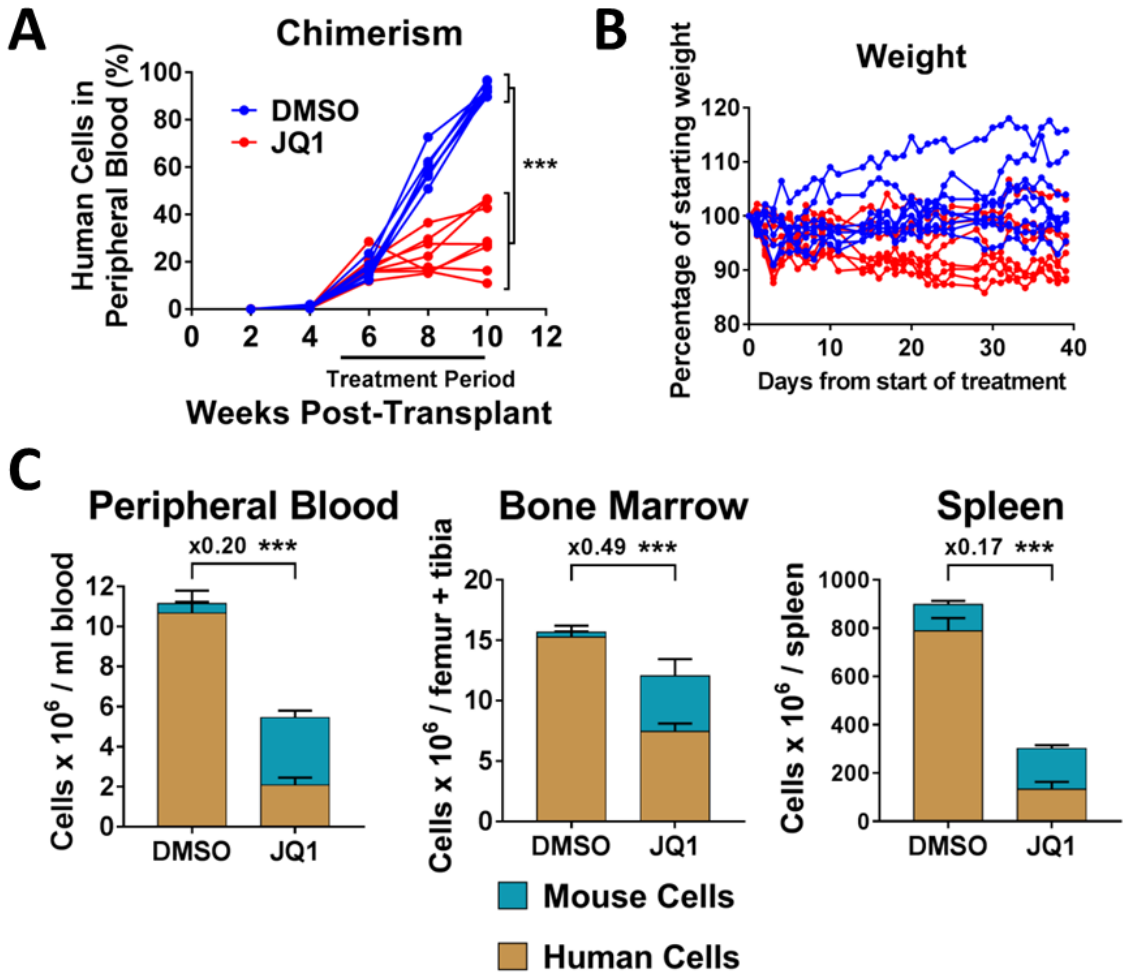


Figure 5.12

(A) Peripheral blood chimerism curve for NSG mice transplanted with a human *EZH2* and *RUNX1* mutant ETP-ALL xenograft (sample ETP-ALL 12) and treated with 45 mg/kg/day JQ1 or a similar volume of DMSO for the indicated time period (36 days). DMSO n=8, JQ1 n=8. (B) Weight of recipient mice following start of treatment with either DMSO or JQ1. (C) White blood cell counts, bone marrow and spleen cellularity, and composition of mouse and human cells at the end of the treatment period. Fold changes and p-values compare absolute numbers of human cells. DMSO n=8, JQ1 n=8. Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

P<0.001). Together, these data demonstrate a marked slowing of leukaemia progression leading to a profound reduction in leukaemic burden as a result of JQ1 treatment.

We then repeated this experiment using sample ETP-ALL 13, the other PRC2-mutant BET inhibitor-sensitive sample, carrying a mutation in *EED* (Figure 5.11A). This sample took longer to establish in the mice, and we did not begin treatment until 8 weeks post-transplant. We again observed a striking difference in human to mouse cell chimerism in peripheral blood leukocytes over time between the two groups (Figure 5.13A), with only a minor effect of JQ1 treatment on overall body weight (Figure 5.13B). In this sample, progression of the leukaemia was in fact completely abrogated by JQ1 treatment (Figure 5.13A), causing an even greater reduction in chimerism at the end of the treatment period (x0.031, P<0.001) compared to the previous sample.

Because this sample proliferated more slowly in recipient mice than the previous sample, the leukaemia was not as advanced in DMSO control mice by the end of the treatment period (Figures 5.13A and 5.12A), and overall white blood cell counts were comparable between the two groups (Figure 5.13C, left panel). There was nonetheless a strong reduction in absolute numbers of human cells in the peripheral blood (x0.023, P<0.001). Overall bone marrow cellularities were in this experiment higher in JQ1-treated mice than DMSO-treated mice at the end of the treatment period (Figure 5.13C, middle panel), and bone marrow was less hypocellular than in the previous experiment (Figure 5.12C, middle panel), consistent with the less advanced stage of the leukaemia at the time of culling. Absolute numbers of human cells were however much lower (x0.16, P<0.001). Splenomegaly was observed in DMSO control mice (Figure 5.13C, right panel), though to a lesser extent than in the previous experiment (Figure 5.12C, right panel). This was not the case in JQ1-treated mice, and absolute numbers of human cells were very strongly reduced in recipient spleens (x0.0039, P<0.001).

These *in vivo* data confirm BET inhibition as a novel, viable potential therapeutic strategy for PRC2-inactivated human ETP leukaemia. Daily JQ1 treatment markedly slows the proliferation of human

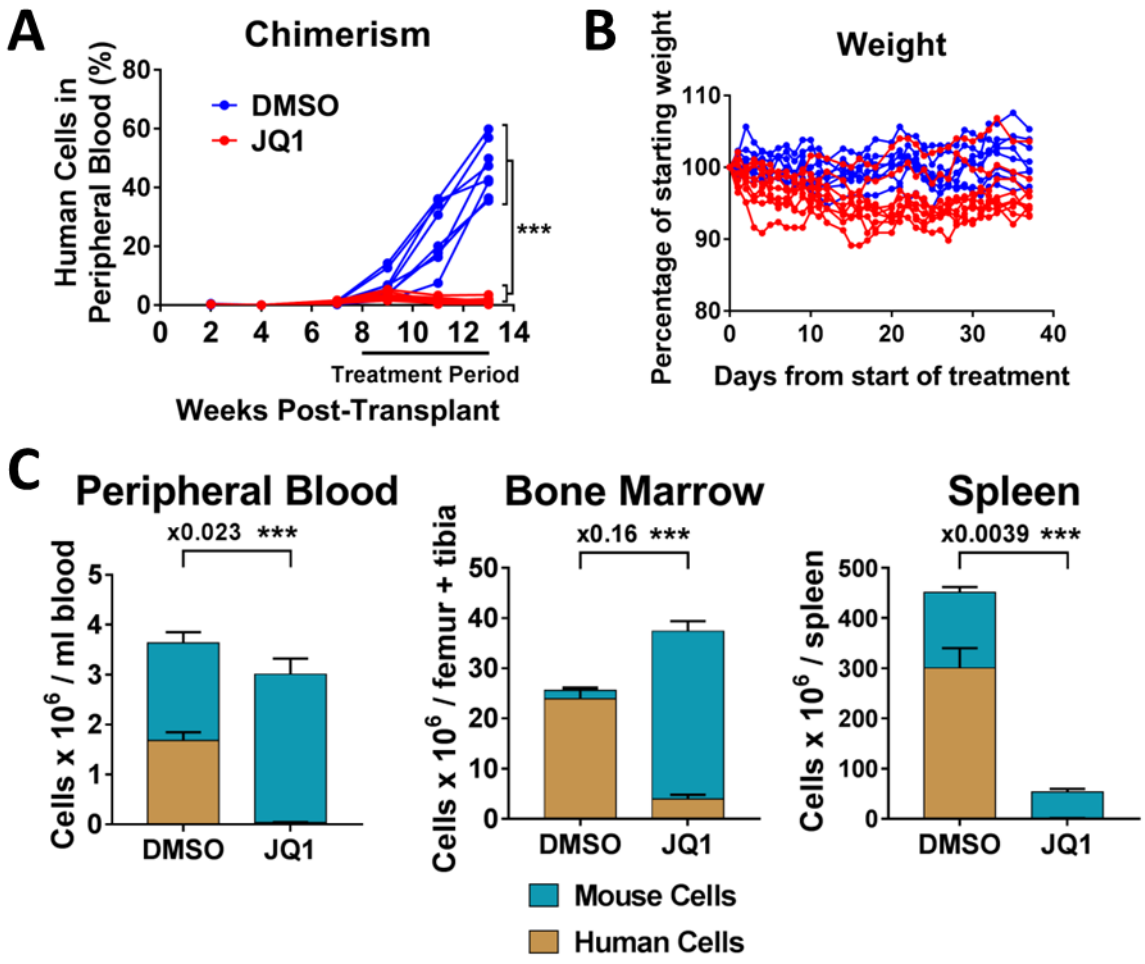


Figure 5.13

(A) Peripheral blood chimerism curve for NSG mice transplanted with a human *EED* mutant ETP-ALL xenograft (sample ETP-ALL 13) and treated with 45 mg/kg/day JQ1 or a similar volume of DMSO for the indicated time period (36 days). DMSO n=8, JQ1 n=9. (B) Weight of recipient mice following start of treatment with either DMSO or JQ1. (C) White blood cell counts, bone marrow and spleen cellularity, and composition of mouse and human cells at the end of the treatment period. Fold changes and p-values compare absolute numbers of human cells. DMSO n=8, JQ1 n=9. Error bars represent standard error of the mean. P-values are by T-test; n.s. (not significant) ≥ 0.05, * < 0.05, ** < 0.01, *** < 0.001.

leukaemias in NSG mice with minor effects on overall body weight, leading to strongly reduced leukaemic burden after 5 weeks of treatment.

3: Discussion

3.1: Epigenetic and transcriptional changes induced by *Ezh2* inactivation in mouse bone marrow

ETP-ALL has been demonstrated to show resistance to conventional treatment (Coustan-Smith et al., 2009; Jain et al., 2016). Some studies have reported survival of ETP-ALL patients being no worse than other T-ALL patients, but only when using intensified chemotherapy and/or allogeneic stem cell transplantation (Bond et al., 2017; Patrick et al., 2014). Novel therapies specifically targeting leukaemia cells are therefore required. Zhang et al. showed that mutations in PRC2 components, such as *EZH2*, confer a particularly poor prognosis among T-ALL patients (Zhang et al., 2012b). Most of the patients carrying these mutations fell into the ETP-ALL subgroup, which in this study was associated with a poor prognosis. Even after adjustment for ETP status, presence of PRC2 mutations showed a small negative effect on survival, which was more significant among younger patients (aged 1-9 years). Exploring the mechanisms by which PRC2 inactivation contributes to leukaemogenesis could therefore potentially deliver novel methods of overcoming the treatment resistance of ETP-ALL.

We decided to investigate the possibility of using BET inhibition to target PRC2-inactivated ETP leukaemia. Our rationale was that malignancies showing loss of function of PRC2 may be dependent on upregulation of oncogenic transcription at gene loci which have lost the repressive H3K27me3 histone mark. Aberrant transcription of these genes may be partially driven by BRD4 binding to H3K27ac and recruiting transcriptional machinery. This mechanism was proposed by De Raedt et al. and validated in a mouse model of peripheral nerve sheath tumours (De Raedt et al., 2014). These tumours were driven by inactivation of the PRC2 component *Suz12* combined with activation of the

Ras pathway through loss of the negative regulator *Nf1*, and inactivation of *p53*. Notably, though *TP53* mutations have not been found in ETP-ALL, mutational activation of the RAS pathway is common (Zhang et al., 2012b). This model is therefore potentially relevant to ETP-ALLs showing inactivation of PRC2 components (around 42% of patients) (Zhang et al., 2012b). Our DKOITD mouse model is also highly relevant given that it features inactivation of *Ezh2* and activation of the Ras pathway via *Flt3-ITD*. De Raedt et al. found that loss of *Suz12* conferred sensitivity to BET inhibition in the context of Ras pathway activation, and we decided to test whether this was also the case in DKOITD leukaemia and human ETP-ALL.

We first used ChIP-sequencing and RNA-sequencing experiments to demonstrate that upregulated gene expression in DKOITD leukaemia correlates with loss of H3K27me3 and gain of H3K27ac at the TSS relative to WT. The concept of the H3K27 residue acting as an "epigenetic switch," with the trimethylated state associated with transcriptional repression and the acetylated state with activation, was first demonstrated in *Drosophila* (Tie et al., 2009) and validated in mouse ES cells (Pasini et al., 2010; Reynolds et al., 2012). Importantly, this mechanism implies that aberrant activation of gene expression following loss of the trimethylated state would rely on epigenetic readers of the acetylated state. De Raedt et al. therefore suggested that inhibiting these readers could reverse the aberrantly upregulated transcription. An alternative approach could be to inhibit demethylases of H3K27me3, and such inhibitors have indeed been developed (Kruidenier et al., 2012). However, this approach would require caution, since inappropriately high levels of H3K27me3, through overexpression or activating mutations of *EZH2*, have also been associated with malignancy (Comet et al., 2016; Souroullas et al., 2016). Inhibition of readers of H3K27ac, such as BRD4, is therefore a potential therapeutic avenue for PRC2-deficient malignancies. Our data provide *in vitro* and *in vivo* evidence using both mouse and human models that further clinical exploration of this strategy would be worthwhile.

Our epigenetic analyses of DKOITD leukaemia were performed in bone marrow cells rather than ETPs, chiefly because of the difficulty in obtaining enough thymocyte cells for ChIP experiments. A wild-type mouse thymus at 6 weeks of age typically contains around 8000 ETPs (Figure 4.02A), but at most 1000-2000 could be expected to be obtained in a FACS experiment. Our bone marrow ChIP-seq experiments, in contrast, were performed using 5 million cells per sample. Using WT and DKOITD ETPs, or perhaps less purified progenitor populations, might have provided interesting insights into the mechanisms underlying LSC transformation. As techniques for performing ChIP-sequencing on very low cell numbers improve (Sundaram et al., 2016), this may become a more viable option.

By comparing WT and DKOITD bone marrow cells, we looked for regions of the genome showing significant loss of H3K27me3 in WT relative to DKOITD, and separately did the same for regions showing gain of H3K27ac. We then made lists of genes which had either of these regions within 2.5 kb upstream or downstream of their TSS. The modification state of H3K27 near the TSS has been shown to correlate with their transcriptional status (Pasini et al., 2010). Our approach does not necessarily select for regions which show both loss of H3K27me3 and gain of H3K27ac at exactly the same site, but we nonetheless assume that gain of H3K27ac is a direct result of loss of H3K27me3. This may not be a valid assumption in all cases, since the two regions may be up to 5 kb away from each other. Additionally, since our samples are taken from separate mice, we are not in fact showing "loss" and "gain" as *Ezh2* is removed, but rather distinct epigenetic states existing in very different contexts of global transcription. The concurrent deletion of *Runx1* is likely to have large-scale influences on transcription in these haematopoietic cells, including potentially that of epigenetic regulators, and we have not taken this factor into account in our analyses. However, performing these analyses in an unbiased, genome-wide manner minimises the potential confounding effects of these assumptions on our overall conclusions.

We find that genes showing both loss of H3K27me3 and gain of H3K27ac at the TSS (n=230) are far more likely to show significantly upregulated gene expression (n=62) than significantly

downregulated (n=1). The majority of these genes however are neither significantly up nor downregulated (n=167). This could be explained by a lack of sensitivity in RNA-sequencing causing a failure of small changes in expression to be determined as significant. It could also mean that for many genes, changes in H3K27 state at the TSS does not in fact affect gene expression. This may not be surprising considering that there are many other mechanisms of regulation of gene transcription, and the state of H3K27 may not always be a driver of expression changes. This is supported by the fact that over half the genes which were significantly upregulated (457 out of 812) did not show loss of H3K27me3 or gain of H3K27ac. Changes in gene expression among these genes could have been driven by transcriptional regulators downstream of *Ezh2*, or by the loss of *Runx1* or gain of *Flt3-ITD* in these cells. Our data nonetheless support a major role for the H3K27 switch in deregulated gene expression in this mouse leukaemia model.

3.2: Transcriptional responses to BET inhibition in *EZH2*-inactivated ETP leukaemia

We next used the LOUCY and JURKAT cell lines to explore the potential importance of this H3K27 switch mechanism, and transcriptional responses to BET inhibition, in human *EZH2*-inactivated ETP leukaemia. We found that genes showing H3K27ac at the TSS were downregulated in response to BET inhibition, providing evidence that this could be a viable method of reversing gene expression changes caused by loss of *EZH2*. However, this downregulation of H3K27ac-associated genes was not specific to LOUCY; genes associated with H3K27ac in JURKAT were also downregulated by BET inhibition in these cells. Our data may therefore simply reflect the general mechanism of action of BET inhibitors, rather than a specific sensitivity of ETP leukaemias to their effects. However, the sets of genes associated with H3K27ac and downregulated by BET inhibition were very different in LOUCY and JURKAT cells, and proliferation of LOUCY cells was much more sensitive to BET inhibition. This suggests that LOUCY cells are likely to be dependent on oncogenic transcriptional pathways which are not active in JURKAT cells, and which can be downregulated by BET inhibition. Our data indicate

that these may be active as a result of the PRC2 inactivation in LOUCY cells. Gene set enrichment analysis identified IL6-JAK-STAT3 signalling as a pathway downregulated by BET inhibition in LOUCY but not JURKAT, which would be worth exploring further in future studies.

A related line of experiments which would be interesting to perform is ChIP for the Brd4 protein in WT and DKOITD mouse cells. This could potentially identify specific genes which are associated with H3K27me3 in WT cells, then switch to H3K27ac and become associated with Brd4 in DKOITD cells. Gene expression analysis before and after BET inhibitor treatment could then identify candidate genes driving the reversal of oncogenic proliferation in BET inhibitor-treated DKOITD cells.

3.3: Sensitivity of mouse and human PRC2-inactivated ETP leukaemia to BET inhibition *in vitro* and *in vivo*

We then moved on to demonstrate a selective inhibition of proliferation of DKOITD relative to WT cells by BET inhibition *in vitro*. This is an important experiment, since it demonstrates a potential therapeutic window for these compounds in targeting leukaemia but not normal cells. Our *in vitro* data for this and other experiments supports a cytostatic, rather than cytotoxic, effect of JQ1 and I-BET151. This is in agreement with previous findings and carries implications for how such compounds might be used clinically. We only tested bone marrow $Lin^{-}Kit^{+}$ cells, which may include both LSCs and more proliferative leukaemia cells. However, our *in vitro* data indicate that DKOITD ETP LSCs are very proliferative compared to WT ETPs, and that this proliferation is also sensitive to BET inhibition.

We then tested the effectiveness of JQ1 treatment against DKOITD leukaemia *in vivo*. We were unable to demonstrate a survival advantage or show that JQ1 treatment slows the progression of the leukaemia through chimerism in the peripheral blood. This may have been due to the already advanced stage of the disease at the point where we began the treatment. Of note, previous *in vivo*

studies of BET inhibitors in AML mouse models have demonstrated only marginal survival advantages (Dawson et al., 2011; Zuber et al., 2011), and one replication study found that these were not reproducible (Shan et al., 2017). We were however able to show a reduction in the leukaemic burden even when starting treatment at this late stage, demonstrating clear evidence of *in vivo* efficacy of JQ1 treatment as a single agent. We confirmed this finding using another BET inhibitor, I-BET151, and a different method of administration, which allowed for more continuous dosing and a longer treatment period.

Having demonstrated *in vivo* efficacy of BET inhibition in DKOITD leukaemia, we were fortunate to be able to extend these studies to human ETP-ALL samples. We first tested the sensitivity to BET inhibition of a range of samples *in vitro*, including both PRC2-mutant and PRC2-wild-type ETP-ALLs. Our finding that the seven samples appeared to fall into two distinct classes –sensitive and insensitive – is interesting, as it implies that individual patients could be pre-stratified as likely or unlikely to respond to BET inhibition. This would require understanding of the mechanisms conferring sensitivity, which was one of the motivations for our mouse studies. Among the ETP-ALLs, both PRC2-mutant samples showed high sensitivity to BET inhibition, while one of the two PRC2-wild-type samples was also sensitive. We unfortunately did not have information on the mutations carried by the three tested non-ETP T-ALL samples, which were all less sensitive to BET inhibition. PRC2 mutations are less common in non-ETP T-ALLs (12% of patients) compared to ETP-ALLs (42% patients), so it is probable that these three samples are PRC2-wild type (Zhang et al., 2012b). Clearly these data, on a small number of samples, are insufficient to conclude that all ETP-ALL samples carrying PRC2 inactivating mutations will show sensitivity to BET inhibition, and we are unable to explain the mechanisms underlying the high sensitivity of one of the PRC2-wild-type samples. Nonetheless, these data provide support for our hypothesis that PRC2 inactivation confers sensitivity to BET inhibition in the context of human as well as mouse ETP leukaemia.

Our *in vivo* data showed a clear therapeutic effect of JQ1 treatment on both PRC2-mutant ETP-ALL samples. Leukaemia progression was markedly slowed, and in fact completely halted in one sample, with no acute detrimental effects on the overall health of the animals. In contrast to our DKOITD *in vivo* data, we started JQ1 treatment before mice had developed a very high leukaemia burden, although we did confirm leukaemias had engrafted and were proliferating in all treated mice. These data could be highly clinically relevant in cases where residual disease persists after treatment, where BET inhibitors could potentially be used in controlling relapse. It would be interesting to repeat these experiments using samples which were insensitive to BET inhibition *in vitro*, to determine whether they are also insensitive *in vivo*. If so, this could demonstrate that BET inhibition would only be of clinical benefit to patients carrying certain mutations. This would carry important implications for interpretation of results of clinical trials, which are currently ongoing in AML. It should also be noted that the compounds we have used in our studies, JQ1 and I-BET151, may be unsuitable for clinical use due to their low *in vivo* stability (Dawson et al., 2011; Zuber et al., 2011). Development of more stable compounds will therefore be necessary, investment in which may be more forthcoming if clear benefits can be demonstrated for a well-defined subset of patients, whether for ETP-ALL or other leukaemias. Further efforts to understand the mechanisms underlying sensitivity of some combinations of mutations to BET inhibitors, such as those we have undertaken using our DKOITD model, are therefore required for these compounds to be successful.

In summary, we have used our DKOITD ETP leukaemia mouse model to investigate the potential therapeutic efficacy of BET inhibitors in PRC2-inactivated ETP leukaemias. We found that inactivation of *Ezh2* caused upregulation of oncogenic gene expression via loss of H3K27me3 and reciprocal gain of H3K27ac. This conferred sensitivity to BET inhibition, both *in vitro* and *in vivo*. This sensitivity was also demonstrated by human PRC2-inactivated ETP-ALLs. These data support BET inhibition as a novel treatment strategy for this therapy-resistant disease.

Chapter 6: Conclusions and Future Directions

1: Summary of Findings

Inactivating mutations in *EZH2* and *RUNX1* are significantly associated in both MDS and ETP-ALL (Papaemmanuil et al., 2013; Zhang et al., 2012b), but the molecular mechanisms underlying this association is unknown. These haematological malignancies, with markedly different clinical phenotypes, are thought to arise within distinct stem and progenitor populations, raising the possibility that the identity of the cell of origin of a malignancy can impact upon the resulting phenotype. We explored this possibility by targeting inactivation of *Ezh2* and *Runx1* to different haematopoietic stem and progenitor populations in mice.

We found that targeting of *Ezh2* and *Runx1* inactivation to HSCs or MPPs, inevitably also including early lymphoid progenitors, resulted in development of an MDS phenotype. Mirroring the mechanisms by which the disease is thought to develop in human MDS patients (Cazzola et al., 2013), *Ezh2* and *Runx1*-inactivated bone marrow cells displayed a competitive advantage over wild-type bone marrow cells. This was accompanied by an extrinsic suppression of wild-type HSCs by *Ezh2* and *Runx1*-inactivated haematopoiesis. These findings establish a novel mouse model of MDS, as well as providing insight into how the malignant clone establishes dominance within the bone marrow of MDS patients.

Transplantation experiments determined that MPPs were able to engraft and propagate the double knockout phenotype more efficiently than HSCs, whether *Ezh2* and *Runx1* inactivation was targeted to HSCs or MPPs. It has been previously shown that in MDS patients, all disease-propagating activity resides within the phenotypic defined HSCs (Woll et al., 2014), seemingly at odds with our findings. However, human haematopoietic stem and progenitor populations are less well-defined than in mice, and the phenotypically defined HSC compartment in fact contains a large proportion of non-

HSCs, including MPPs. Our data therefore in fact remain consistent with findings in human MDS, but consistent with a different interpretation of the previous data.

Having noted that targeting of *Ezh2* and *Runx1* inactivation to early lymphoid progenitors using *Rag1Cre* does not result in development of MDS, we used this model to explore the collaboration of these two mutations in ETP-ALL. Thymuses of *Ezh2* and *Runx1*-inactivated mice showed an expansion of ETPs, which displayed increased expression of HSC, myeloid and signalling pathway-associated gene transcription. These are characteristic features of ETP-ALL, providing insight into how mutations in these two genes may lead to development of this malignancy. Importantly, ETP-ALL has thus far only been hypothesised based on its gene expression profile to initiate with ETPs, and experimental models have not been able to demonstrate that these cells are in fact capable of becoming a leukaemia-propagating population.

Despite the presence of this expanded ETP population displaying ETP leukaemia-associated transcription, these mice did not show development of a leukaemia phenotype. Since mutations in signalling pathway components such as *FLT3* are another characteristic feature of ETP-ALL, we therefore combined the *Flt3-ITD* mutation with inactivation of *Ezh2* and *Runx1* in early lymphoid progenitors. These mice developed an aggressive acute leukaemia showing expression of both myeloid and lymphoid-associated transcription, and which crucially could be propagated by the expanded ETP population. Both ETPs and bone marrow cells in leukaemic mice showed gene expression profiles mirroring ETP-ALL. These findings establish a novel mouse model of ETP leukaemia, and show for the first time that ETPs carry the potential for transformation into a leukaemic stem cell population. This is also of broader relevance for the lympho-myeloid haematopoietic pathway, since we present a leukaemia model arising from within lympho-myeloid progenitors which displays phenotypic characteristics of these lineages.

We then utilised our mouse model to explore therapeutic targeting of PRC2-inactivated ETP leukaemia. BET inhibition has been proposed as a mechanism of reversing oncogenic transcription

associated with PRC2 mutational inactivation, since this may be dependent on binding of BET proteins to aberrantly deposited H3K27ac (De Raedt et al., 2014). We found that, in agreement with this model, upregulated gene transcription in DKOITD bone marrow cells was strongly associated with loss of H3K27me3 and gain of H3K27ac compared to WT cells. Further, in the LOUCY human ETP leukaemia cell line, which showed loss of PRC2 function and high sensitivity to BET inhibition, genes associated with H3K27ac were preferentially downregulated by treatment with BET inhibitors. These genes also overlapped strongly with those upregulated in DKOITD ETPs and bone marrow cells compared to WT. These findings validate BET inhibition as a method of reversing potentially oncogenic upregulated transcription associated with inactivation of PRC2 in ETP leukaemia.

We therefore tested the sensitivity of *Ezh2* and *Runx1*-inactivated, *Flt3-ITD* ETP leukaemia to BET inhibitors. *In vitro*, proliferation of DKOITD leukaemia cells was more sensitive to these compounds than the equivalent cells from WT mice, suggesting that BET inhibition could selectively target leukaemia *in vivo* without affecting normal haematopoiesis. We tested this using transplantation experiments and found that treatment with either of two BET inhibitor compounds *in vivo* significantly reduced the leukaemic burden of recipients of DKOITD leukaemia cells.

Having established the therapeutic potential of BET inhibition in our mouse PRC2-inactivated ETP leukaemia model, we went on to test a number of human ETP-ALL xenograft samples. PRC2-mutant patient ETP leukaemias showed high sensitivity to treatment with BET inhibitors *in vitro*, in contrast to non-ETP T-ALL samples. We then performed *in vivo* treatment of recipient mice transplanted with either of two PRC2-mutant ETP-ALL samples. This demonstrated remarkable efficacy of BET inhibition against these leukaemias, with no measurable impact on normal haematopoiesis or the overall health of the mice. These pre-clinical findings confirm BET inhibition as a viable therapeutic strategy for selective targeting of PRC2-inactivated ETP leukaemias.

2: Future Directions

2.1: A novel *Ezh2* and *Runx1*-inactivated MDS mouse model

We have established a novel mouse model of MDS by targeting of *Ezh2* and *Runx1* inactivation to HSCs and MPPs. A previous study also reported development of an MDS phenotype by combining *Ezh2* inactivation with exogenous expression of a dominant negative *Runx1* construct (Sashida et al., 2014), but the experimental setup used precluded targeting of the mutations to specific haematopoietic populations. Our model will allow us to further explore the impact of *Ezh2* and *Runx1* mutations in HSCs as compared to MPPs, and how this impacts upon the biology of the resulting malignancy.

In our Mx1-DKO and Flt3-DKO mice, MPPs and DP cells were transformed into self-renewing, MDS-propagating cancer stem cells. We would like to further explore the transcriptional mechanisms underlying this observation, since the molecular pathways leading to MDS clone development in human patients remain poorly understood. RNA-sequencing of HSCs and MPPs from Mx1-WT and Mx1-DKO mice could identify candidate genes responsible for the aberrant acquisition of self-renewal by MPPs and also the loss of HSCs in this model. This could potentially lead to molecular targeting of the disease-propagating potential of human MDS clones.

We also observed an extrinsic suppression of wild-type HSCs by *Ezh2* and *Runx1*-inactivated malignant haematopoiesis in our transplantation experiments. This is a key finding, since the mechanisms by which normal haematopoiesis is suppressed in MDS patients are not well understood. We could therefore utilise our mouse model to explore this further. It has been shown that myeloproliferative neoplasia clones are able to remodel HSC niches in order to support malignant haematopoiesis at the expense of normal HSC function (Schepers et al., 2013), and we could perform analysis of bone marrow niche populations to explore whether this occurs in Mx1-DKO mice. We could also perform RNA-sequencing on wild-type HSCs exposed to Mx1-DKO

haematopoiesis and compare to non-exposed wild-type HSCs, to investigate the mechanistic impacts of the malignant MDS clone on normal HSCs.

2.2: A novel *Ezh2* and *Runx1*-inactivated ETP leukaemia model

We showed that inactivation of *Ezh2* and *Runx1* was sufficient to induce an expansion of ETPs displaying transcriptional features of ETP leukaemia, including upregulated signalling pathways such as Kras. However, the upregulated signalling-related transcription was insufficient to induce leukaemic transformation of these cells in the absence of the *Flt3-ITD* mutation. DKO ETPs may therefore represent a "pre-malignant" state, primed for transformation by a single additional mutation. Our finding that signalling pathways may become aberrantly activated in pre-malignant cells in the absence of direct mutational activation of the pathway is interesting and warrants additional experimental investigation. We would like to explore the target genes of *Ezh2* and *Runx1* driving this signalling-associated transcription by performing ChIP in ETPs. This would require very limited numbers of cells but may be feasible using recently developed techniques (Sundaram et al., 2016). It may be possible to identify targets of *Ezh2* and/or *Runx1* which become aberrantly expressed upon their inactivation, and which could potentially be targeted therapeutically.

We generated RNA-sequencing datasets for both ETPs and $Mac1^{+}$ bone marrow cells from our DKOITD mutant mice and wild-type controls. We focused our analysis of these data on global gene expression changes, rather than identifying specific genes which might be driving the leukaemic phenotype. We would therefore like to select a small number of genes for more detailed mechanistic studies, in order to provide further insights into the molecular mechanisms underlying DKOITD leukaemia. These may be applicable to human ETP leukaemias, and we would be able to explore our findings in our ETP-ALL xenograft samples. One candidate gene that would be appropriate for these studies is *Mpl*, which was the most significantly upregulated gene in both DKO

and DKOITD ETPs relative to WT. Its expression is also upregulated in ETP-ALL relative to non-ETP T-ALL (Zhang et al., 2012b), and it was strongly downregulated by BET inhibition in LOUCY cells. MPL is a receptor tyrosine kinase signalling protein and is critical for the function of HSCs (Qian et al., 2007; Yoshihara et al., 2007). It may therefore play a role in the pathogenesis DKOITD and human ETP leukaemia, and we would like to explore its function in more detail.

2.3: Therapeutic efficacy of BET inhibitors in PRC2-inactivated ETP leukaemia

We found that DKOITD bone marrow cells show upregulated expression of genes associated with aberrant loss of H3K27me3 and gain of H3K27ac. We also showed that many of the genes upregulated in DKOITD ETPs and bone marrow cells become downregulated by BET inhibition in LOUCY cells. Further identification of the specific genes and pathways responsible for the sensitivity of DKOITD leukaemia to BET inhibition could provide clinically relevant insights into the therapeutic mechanisms of these compounds. In this regard we could perform ChIP for the Brd4 protein in WT and DKOITD bone marrow cells to identify candidate genes driving leukaemogenesis which become downregulated upon treatment.

Our *in vitro* and particularly *in vivo* data displayed strong pre-clinical therapeutic potential of BET inhibition for treatment of PRC2-inactivated ETP-ALL. These findings may lead to initiation of clinical trials for this subset of patients. Our findings may also extend to patients with other haematological cancers such as AML who have inactivating mutations in PRC2 pathway components. For unknown reasons, mutations in *EZH2* are rare in *de novo* AMLs (those patients with no prior history of myeloid malignancy), but are common in secondary AML (evolving from MDS) (Sperling et al., 2017). Since the mutational landscape of ETP-ALL appears to have many features in common with AML as opposed to other T-ALLs (Zhang et al., 2012b), the therapeutic efficacy of BET inhibitors which we have observed in human PRC2-inactivated ETP-ALL may well be transferable to AML.

3: Conclusions

These studies have contributed to our understanding of the mechanisms by which mutations arising in different haematopoietic stem and progenitor populations induce distinct malignant phenotypes. By targeting *Ezh2* and *Runx1* inactivation to HSCs and MPPs, we have generated a novel mouse model of MDS which will be used for future studies into the molecular mechanisms driving this disease. Using this model we have provided new insights into the populations which drive propagation of this disease.

Using combination of *Flt3-ITD* with inactivation of *Ezh2* and *Runx1* targeted to early lymphoid progenitors, we generated a novel mouse model of ETP leukaemia. We have provided the first experimental evidence that this disease can be propagated by ETPs, a crucial step towards development of leukaemia stem cell-targeted therapy. This mouse model will be used for future studies into the mechanisms underlying the development of ETP leukaemia and its characteristic resistance to cytotoxic therapy.

Using this ETP leukaemia mouse model we have provided insights into the transcriptional and epigenetic mechanisms underlying the efficacy of BET inhibition in the context of PRC2-inactivated haematological malignancy. Finally, using mouse and human *in vitro* and *in vivo* models, we have demonstrated that these compounds show great promise for targeted therapy of PRC2-inactivated ETP leukaemia.

Appendix

List of cancer-associated variants detected by exome sequencing of eight DKOITD leukaemias (see Figure 4.11).

Gene	Implicated in ETP-ALL?	No. of unique variants	Variant	No. of leukemic mice	Likely germline?	Nearby variant found in human cancer?	ETP-ALL associated variant
Acsl3	No	1	p.I667V	5	Likely germline variant		
Akap9	No	2	p.F1119V	2	Likely germline variant		
			p.Q1138K	1		Nearby variant not found in human cancer	
Atm	No	3	p.A2119T	5	Likely germline variant		
			p.Q1078H	5	Likely germline variant		
			p.R32C	5	Likely germline variant		
Axin2	No	1	p.? Splice site variant	1		Nearby variant not found in human cancer	
Blm	No	2	p.V1298L	6	Likely germline variant		
			p.L134P	6	Likely germline variant		
Cars	No	1	p.S771G	5	Likely germline variant		
Creb3l2	No	1	p.P222L	1		Nearby variant found in endometrium carcinoma	
Cxcr4	No	1	p.V216I	1		Nearby variant not found in human cancer	
Eif4a2	No	11	p.V224M	1		Nearby variant found in large intestine carcinoma	
			p.R230S	2	Likely germline variant		
			p.R234*	2	Likely germline variant		
			p.Q249R	1		Nearby variant found in endometrium carcinoma	

Gene	Implicated in ETP-ALL?	No. of unique variants	Variant	No. of leukemic mice	Likely germline?	Nearby variant found in human cancer?	ETP-ALL associated variant
			p.R366Q	1		Nearby variant found in skin carcinoma	
			p.T378I	1		Nearby variant not found in human cancer	
			p.D381Y	1		Nearby variant not found in human cancer	
			p.R386C	2	Likely germline variant		
			p.D387E	2	Likely germline variant		
			p.I388N	2	Likely germline variant		
			p.I388M	2	Likely germline variant		
Fanca	No	14	p.M943T	3	Likely germline variant		
			p.M833V	3	Likely germline variant		
			p.L599F	3	Likely germline variant		
			p.V474M	3	Likely germline variant		
			p.V466I	3	Likely germline variant		
			p.I323T	3	Likely germline variant		
			p.N310D	3	Likely germline variant		
			p.M147V	3	Likely germline variant		
			p.R80L	3	Likely germline variant		
			p.R80C	3	Likely germline variant		
			p.L43V	3	Likely germline variant		
			p.T23A	3	Likely germline variant		

Gene	Implicated in ETP-ALL?	No. of unique variants	Variant	No. of leukemic mice	Likely germline?	Nearby variant found in human cancer?	ETP-ALL associated variant
			p.P15S	3	Likely germline variant		
			p.P2L	2	Likely germline variant		
Fcgr2b	No	1	p.S297I	4	Likely germline variant		
Flt3	Yes	6	p.L760S	8	Flt3-ITD knockin construct		
			p.F692L	8	Flt3-ITD knockin construct		
			p.D597E	8	Flt3-ITD knockin construct		
			p.L586P	8	Flt3-ITD knockin construct		
			p.P585S	8	Flt3-ITD knockin construct		
			p.A67T	8	Flt3-ITD knockin construct		
Foxa1	No	1	p.H389R	2	Likely germline variant		
Gata3	Yes	1	p.I350N	1			Poor sequence quality of reads supporting variant
Hsp90ab1	No	1	p.A650T	4	Likely germline variant		
Klk1b1	No	1	p.L226I	6	Likely germline variant		
Kmt2c	No	1	p.T3760I	2	Likely germline variant		
Ncor2	No	1	p.I1107L	1		Nearby variant not found in human cancer	
Nf1	Yes	1	p.D1062E	3	Likely germline variant		
Numa1	No	8	p.A539V	5	Likely germline variant		
			p.N627S	5	Likely germline variant		

Gene	Implicated in ETP-ALL?	No. of unique variants	Variant	No. of leukemic mice	Likely germline?	Nearby variant found in human cancer?	ETP-ALL associated variant
			p.Q938R	5	Likely germline variant		
			p.I976M	5	Likely germline variant		
			p.A1093T	5	Likely germline variant		
			p.S1094R	5	Likely germline variant		
			p.S1099P	5	Likely germline variant		
			p.V1270A	5	Likely germline variant		
Nup98	No	1	p.V1330I	5	Likely germline variant		
Tmprss2	No	2	p.S302G	3	Likely germline variant		
			p.L75P	3	Likely germline variant		
Zfp521	No	1	p.G1095C	2	Likely germline variant		

Bibliography

Abdel-Wahab, O., Gao, J., Adli, M., Dey, A., Trimarchi, T., Chung, Y. R., Kuscu, C., Hricik, T., Ndiaye-Lobry, D., Lafave, L. M.*, et al.* (2013). Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. J Exp Med *210*, 2641-2659.

Abu-Duhier, F. M., Goodeve, A. C., Wilson, G. A., Gari, M. A., Peake, I. R., Rees, D. C., Vandenberghe, E. A., Winship, P. R., and Reilly, J. T. (2000). FLT3 internal tandem duplication mutations in adult acute myeloid leukaemia define a high-risk group. Br J Haematol *111*, 190-195.

Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., Bryder, D., Yang, L., Borge, O. J., Thoren, L. A.*, et al.* (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential: a revised road map for adult blood lineage commitment. Cell *121*, 295-306.

Agger, K., Cloos, P. A. C., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A. E., and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature *449*, 731-734.

Aifantis, I., Raetz, E., and Buonamici, S. (2008). Molecular pathogenesis of T-cell leukaemia and lymphoma. Nat Rev Immunol *8*, 380-390.

Aloia, L., Di Stefano, B., and Di Croce, L. (2013). Polycomb complexes in stem cells and embryonic development. Development *140*, 2525-2534.

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., Bloomfield, C. D., Cazzola, M., and Vardiman, J. W. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood *127*, 2391-2405.

Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res *21*, 381-395.

Barlow, J. L., Drynan, L. F., Hewett, D. R., Holmes, L. R., Lorenzo-Abalde, S., Lane, A. L., Jolin, H. E., Pannell, R., Middleton, A. J., Wong, S. H.*, et al.* (2010). A p53-dependent mechanism underlies macrocytic anemia in a mouse model of human 5q- syndrome. Nat Med *16*, 59-66.

Beachy, S. H., and Aplan, P. D. (2010). Mouse models of myelodysplastic syndromes. Hematol Oncol Clin North Am *24*, 361-375.

Bejar, R., Stevenson, K., Abdel-Wahab, O., Galili, N., Nilsson, B., Garcia-Manero, G., Kantarjian, H., Raza, A., Levine, R. L., Neuberg, D., and Ebert, B. L. (2011). Clinical effect of point mutations in myelodysplastic syndromes. N Engl J Med *364*, 2496-2506.

Bell, J. J., and Bhandoola, A. (2008). The earliest thymic progenitors for T cells possess myeloid lineage potential. Nature *452*, 764-767.

Benz, C., Martins, V. C., Radtke, F., and Bleul, C. C. (2008). The stream of precursors that colonizes the thymus proceeds selectively through the early T lineage precursor stage of T cell development. J Exp Med *205*, 1187-1199.

Bereshchenko, O., Mancini, E., Moore, S., Bilbao, D., Mansson, R., Luc, S., Grover, A., Jacobsen, S. E., Bryder, D., and Nerlov, C. (2009). Hematopoietic stem cell expansion precedes the generation of committed myeloid leukemia-initiating cells in C/EBPalpha mutant AML. Cancer Cell *16*, 390-400.

Blume-Jensen, P., and Hunter, T. (2001). Oncogenic kinase signalling. Nature *411*, 355-365.

Boiers, C., Carrelha, J., Lutteropp, M., Luc, S., Green, J. C., Azzoni, E., Woll, P. S., Mead, A. J., Hultquist, A., Swiers, G.*, et al.* (2013). Lymphomyeloid contribution of an immune-restricted progenitor emerging prior to definitive hematopoietic stem cells. Cell Stem Cell *13*, 535-548.

Bond, J., Graux, C., Lhermitte, L., Lara, D., Cluzeau, T., Leguay, T., Cieslak, A., Trinquand, A., Pastoret, C., Belhocine, M.*, et al.* (2017). Early Response-Based Therapy Stratification Improves Survival in Adult Early Thymic Precursor Acute Lymphoblastic Leukemia: A Group for Research on Adult Acute Lymphoblastic Leukemia Study. J Clin Oncol *35*, 2683-2691.

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

Bowie, M. B., McKnight, K. D., Kent, D. G., McCaffrey, L., Hoodless, P. A., and Eaves, C. J. (2006). Hematopoietic stem cells proliferate until after birth and show a reversible phase-specific engraftment defect. J Clin Invest *116*, 2808-2816.

Boyer, S. W., Schroeder, A. V., Smith-Berdan, S., and Forsberg, E. C. (2011). All hematopoietic cells develop from hematopoietic stem cells through Flk2/Flt3-positive progenitor cells. Cell Stem Cell *9*, 64-73.

Breitenbuecher, F., Schnittger, S., Grundler, R., Markova, B., Carius, B., Brecht, A., Duyster, J., Haferlach, T., Huber, C., and Fischer, T. (2009). Identification of a novel type of ITD mutations located in nonjuxtamembrane domains of the FLT3 tyrosine kinase receptor. Blood *113*, 4074-4077.

Buza-Vidas, N., Woll, P., Hultquist, A., Duarte, S., Lutteropp, M., Bouriez-Jones, T., Ferry, H., Luc, S., and Jacobsen, S. E. (2011). FLT3 expression initiates in fully multipotent mouse hematopoietic progenitor cells. Blood *118*, 1544-1548.

Cai, X., Gaudet, J. J., Mangan, J. K., Chen, M. J., De Obaldia, M. E., Oo, Z., Ernst, P., and Speck, N. A. (2011). Runx1 loss minimally impacts long-term hematopoietic stem cells. PLoS One *6*, e28430.

Cazzola, M., Della Porta, M. G., and Malcovati, L. (2013). The genetic basis of myelodysplasia and its clinical relevance. Blood *122*, 4021-4034.

Chen, M. J., Yokomizo, T., Zeigler, B. M., Dzierzak, E., and Speck, N. A. (2009). Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. Nature *457*, 887-891.

Chiaretti, S., Zini, G., and Bassan, R. (2014). Diagnosis and Subclassification of Acute Lymphoblastic Leukemia. Mediterranean Journal of Hematology and Infectious Diseases *6*, e2014073.

Clambey, E. T., Davenport, B., Kappler, J. W., Marrack, P., and Homann, D. (2014). Molecules in medicine mini review: the alphabeta T cell receptor. J Mol Med (Berl) *92*, 735-741.

Clevers, H. (2011). The cancer stem cell: premises, promises and challenges. Nat Med *17*, 313-319.

Cogle, C. R. (2015). Incidence and Burden of the Myelodysplastic Syndromes. Curr Hematol Malig Rep *10*, 272-281.

Comet, I., Riising, E. M., Leblanc, B., and Helin, K. (2016). Maintaining cell identity: PRC2-mediated regulation of transcription and cancer. Nat Rev Cancer *16*, 803-810.

Cook, A. M., Li, L., Ho, Y., Lin, A., Stein, A., Forman, S., Perrotti, D., Jove, R., and Bhatia, R. (2014). Role of altered growth factor receptor-mediated JAK2 signaling in growth and maintenance of human acute myeloid leukemia stem cells. Blood *123*, 2826-2837.

Coombs, C. C., Tavakkoli, M., and Tallman, M. S. (2015). Acute promyelocytic leukemia: where did we start, where are we now, and the future. Blood Cancer J *5*, e304.

Corces-Zimmerman, M. R., and Majeti, R. (2014). Pre-leukemic evolution of hematopoietic stem cells: the importance of early mutations in leukemogenesis. Leukemia *28*, 2276-2282.

Coustan-Smith, E., Mullighan, C. G., Onciu, M., Behm, F. G., Raimondi, S. C., Pei, D., Cheng, C., Su, X., Rubnitz, J. E., Basso, G.*, et al.* (2009). Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. Lancet Oncol *10*, 147-156.

Cramer, S. D., Aplan, P. D., and Durum, S. K. (2016). Therapeutic targeting of IL-7Rα signaling pathways in ALL treatment. Blood *128*, 473-478.

Danis, E., Yamauchi, T., Echanique, K., Zhang, X., Haladyna, J. N., Riedel, S. S., Zhu, N., Xie, H., Orkin, S. H., Armstrong, S. A.*, et al.* (2016). Ezh2 Controls an Early Hematopoietic Program and Growth and Survival Signaling in Early T Cell Precursor Acute Lymphoblastic Leukemia. Cell Rep *14*, 1953-1965.

Dawson, M. A., Prinjha, R. K., Dittmann, A., Giotopoulos, G., Bantscheff, M., Chan, W. I., Robson, S. C., Chung, C. W., Hopf, C., Savitski, M. M.*, et al.* (2011). Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature *478*, 529-533.

de Bruijn, M., and Dzierzak, E. (2017). Runx transcription factors in the development and function of the definitive hematopoietic system. Blood *129*, 2061-2069.

De Raedt, T., Beert, E., Pasmant, E., Luscan, A., Brems, H., Ortonne, N., Helin, K., Hornick, J. L., Mautner, V., Kehrer-Sawatzki, H.*, et al.* (2014). PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies. Nature *514*, 247-251.

de Thé, H., and Chen, Z. (2010). Acute promyelocytic leukaemia: novel insights into the mechanisms of cure. Nat Rev Cancer *10*, 775-783.

Delmore, J. E., Issa, G. C., Lemieux, M. E., Rahl, P. B., Shi, J., Jacobs, H. M., Kastritis, E., Gilpatrick, T., Paranal, R. M., Qi, J.*, et al.* (2011). BET bromodomain inhibition as a therapeutic strategy to target c-Myc. Cell *146*, 904-917.

Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. Nat Struct Mol Biol *20*, 1147-1155.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Dores, G. M., Devesa, S. S., Curtis, R. E., Linet, M. S., and Morton, L. M. (2012). Acute leukemia incidence and patient survival among children and adults in the United States, 2001-2007. Blood *119*, 34-43.

Durinck, S., Spellman, P. T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc *4*, 1184-1191.

Eppert, K., Takenaka, K., Lechman, E. R., Waldron, L., Nilsson, B., van Galen, P., Metzeler, K. H., Poeppl, A., Ling, V., Beyene, J.*, et al.* (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. Nat Med *17*, 1086-1093.

Ernst, T., Chase, A. J., Score, J., Hidalgo-Curtis, C. E., Bryant, C., Jones, A. V., Waghorn, K., Zoi, K., Ross, F. M., Reiter, A.*, et al.* (2010). Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. Nat Genet *42*, 722-726.

Essers, M. A. G., Offner, S., Blanco-Bose, W. E., Waibler, Z., Kalinke, U., Duchosal, M. A., and Trumpp, A. (2009). IFN[agr] activates dormant haematopoietic stem cells in vivo. Nature *458*, 904-908.

Faber, Z. J., Chen, X., Gedman, A. L., Boggs, K., Cheng, J., Ma, J., Radtke, I., Chao, J. R., Walsh, M. P., Song, G.*, et al.* (2016). The genomic landscape of core-binding factor acute myeloid leukemias. Nat Genet *48*, 1551-1556.

Ferrando, A. A., Neuberg, D. S., Staunton, J., Loh, M. L., Huard, C., Raimondi, S. C., Behm, F. G., Pui, C. H., Downing, J. R., Gilliland, D. G.*, et al.* (2002). Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell *1*, 75-87.

Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J. P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Muller, S., Pawson, T.*, et al.* (2012). Histone recognition and large-scale structural analysis of the human bromodomain family. Cell *149*, 214-231.

Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W. B., Fedorov, O., Morse, E. M., Keates, T., Hickman, T. T., Felletar, I.*, et al.* (2010). Selective inhibition of BET bromodomains. Nature *468*, 1067-1073.

Gaillard, H., Garcia-Muse, T., and Aguilera, A. (2015). Replication stress and cancer. Nat Rev Cancer *15*, 276-289.

Goardon, N., Marchi, E., Atzberger, A., Quek, L., Schuh, A., Soneji, S., Woll, P., Mead, A., Alford, K. A., Rout, R.*, et al.* (2011). Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer Cell *19*, 138-152.

Goossens, S., Radaelli, E., Blanchet, O., Durinck, K., Van der Meulen, J., Peirs, S., Taghon, T., Tremblay, C. S., Costa, M., Farhang Ghahremani, M.*, et al.* (2015). ZEB2 drives immature T-cell lymphoblastic leukaemia development via enhanced tumour-initiating potential and IL-7 receptor signalling. Nat Commun *6*, 5794.

Grover, A., Mancini, E., Moore, S., Mead, A. J., Atkinson, D., Rasmussen, K. D., O'Carroll, D., Jacobsen, S. E., and Nerlov, C. (2014). Erythropoietin guides multipotent hematopoietic progenitor cells toward an erythroid fate. J Exp Med *211*, 181-188.

Growney, J. D., Shigematsu, H., Li, Z., Lee, B. H., Adelsperger, J., Rowan, R., Curley, D. P., Kutok, J. L., Akashi, K., Williams, I. R.*, et al.* (2005). Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. Blood *106*, 494-504.

Hayakawa, F., Towatari, M., Kiyoi, H., Tanimoto, M., Kitamura, T., Saito, H., and Naoe, T. (2000). Tandem-duplicated Flt3 constitutively activates STAT5 and MAP kinase and introduces autonomous cell growth in IL-3-dependent cell lines. Oncogene *19*, 624-631.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Hidalgo, I., Herrera-Merchan, A., Ligos, J. M., Carramolino, L., Nunez, J., Martinez, F., Dominguez, O., Torres, M., and Gonzalez, S. (2012). Ezh1 is required for hematopoietic stem cell maintenance and prevents senescence-like cell cycle arrest. Cell Stem Cell *11*, 649-662.

Holoch, D., and Margueron, R. (2017). Mechanisms Regulating PRC2 Recruitment and Enzymatic Activity. Trends Biochem Sci *42*, 531-542.

Hosoya, T., Maillard, I., and Engel, J. D. (2010). From the cradle to the grave: activities of GATA-3 throughout T-cell development and differentiation. Immunol Rev *238*, 110-125.

Huang, G., Zhang, P., Hirai, H., Elf, S., Yan, X., Chen, Z., Koschmieder, S., Okuno, Y., Dayaram, T., Growney, J. D.*, et al.* (2008). PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. Nat Genet *40*, 51-60.

Huang, M., Shen, A., Ding, J., and Geng, M. (2014). Molecularly targeted cancer therapy: some lessons from the past decade. Trends Pharmacol Sci *35*, 41-50.

Ichikawa, M., Asai, T., Saito, T., Seo, S., Yamazaki, I., Yamagata, T., Mitani, K., Chiba, S., Ogawa, S., Kurokawa, M., and Hirai, H. (2004). AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. Nat Med *10*, 299-304.

Ishikawa, F., Yoshida, S., Saito, Y., Hijikata, A., Kitamura, H., Tanaka, S., Nakamura, R., Tanaka, T., Tomiyama, H., Saito, N*., et al.* (2007). Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. Nat Biotech *25*, 1315-1321.

Jain, N., Lamb, A. V., O'Brien, S., Ravandi, F., Konopleva, M., Jabbour, E., Zuo, Z., Jorgensen, J., Lin, P., Pierce, S*., et al.* (2016). Early T-cell precursor acute lymphoblastic leukemia/lymphoma (ETP-ALL/LBL) in adolescents and adults: a high-risk subtype. Blood *127*, 1863-1869.

Jones, M., Cordell, J. L., Beyers, A. D., Tse, A. G., and Mason, D. Y. (1993). Detection of T and B cells in many animal species using cross-reactive anti-peptide antibodies. J Immunol *150*, 5429-5435.

Kamminga, L. M., Bystrykh, L. V., de Boer, A., Houwer, S., Douma, J., Weersing, E., Dontje, B., and de Haan, G. (2006). The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. Blood *107*, 2170-2179.

Karolchik, D., Hinrichs, A. S., Furey, T. S., Roskin, K. M., Sugnet, C. W., Haussler, D., and Kent, W. J. (2004). The UCSC Table Browser data retrieval tool. Nucleic Acids Res *32*, D493-496.

Kelly, L. M., Liu, Q., Kutok, J. L., Williams, I. R., Boulton, C. L., and Gilliland, D. G. (2002). FLT3 internal tandem duplication mutations associated with human acute myeloid leukemias induce myeloproliferative disease in a murine bone marrow transplant model. Blood *99*, 310-318.

Kiel, M. J., Yilmaz, O. H., Iwashita, T., Terhorst, C., and Morrison, S. J. (2005). SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. Cell *121*, 1109-1121.

Kim, C. H. (2010). Homeostatic and pathogenic extramedullary hematopoiesis. Journal of blood medicine *1*, 13-19.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Kim, E., Ilagan, Janine O., Liang, Y., Daubner, Gerrit M., Lee, Stanley C. W., Ramakrishnan, A., Li, Y., Chung, Young R., Micol, J.-B., Murphy, M. E.*, et al.* (2015). SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. Cancer Cell *27*, 617-630.

Kiyoi, H., Naoe, T., Nakano, Y., Yokota, S., Minami, S., Miyawaki, S., Asou, N., Kuriyama, K., Jinnai, I., Shimazaki, C.*, et al.* (1999). Prognostic implication of FLT3 and N-RAS gene mutations in acute myeloid leukemia. Blood *93*, 3074-3080.

Kleer, C. G., Cao, Q., Varambally, S., Shen, R., Ota, I., Tomlins, S. A., Ghosh, D., Sewalt, R. G., Otte, A. P., Hayes, D. F.*, et al.* (2003). EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. Proc Natl Acad Sci U S A *100*, 11606-11611.

Koch, U., and Radtke, F. (2011). Mechanisms of T cell development and transformation. Annu Rev Cell Dev Biol *27*, 539-562.

Komrokji, R. S., Padron, E., Ebert, B. L., and List, A. F. (2013). Deletion 5q MDS: molecular and therapeutic implications. Best Pract Res Clin Haematol *26*, 365-375.

Kottaridis, P. D., Gale, R. E., Frew, M. E., Harrison, G., Langabeer, S. E., Belton, A. A., Walker, H., Wheatley, K., Bowen, D. T., Burnett, A. K.*, et al.* (2001). The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials. Blood *98*, 1752-1759.

Kruidenier, L., Chung, C.-w., Cheng, Z., Liddle, J., Che, K., Joberty, G., Bantscheff, M., Bountra, C., Bridges, A., Diallo, H.*, et al.* (2012). A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. Nature *488*, 404-408.

Kuhn, R., Schwenk, F., Aguet, M., and Rajewsky, K. (1995). Inducible gene targeting in mice. Science *269*, 1427-1429.

Kwon, H. Y., Bajaj, J., Ito, T., Blevins, A., Konuma, T., Weeks, J., Lytle, N. K., Koechlein, C. S., Rizzieri, D., Chuah, C.*, et al.* (2015). Tetraspanin 3 Is Required for the Development and Propagation of Acute Myelogenous Leukemia. Cell Stem Cell *17*, 152-164.

Laiosa, C. V., Stadtfeld, M., Xie, H., de Andres-Aguayo, L., and Graf, T. (2006). Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBP alpha and PU.1 transcription factors. Immunity *25*, 731-744.

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M. A., and Dick, J. E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature *367*, 645-648.

Lee, B. H., Tothova, Z., Levine, R. L., Anderson, K., Buza-Vidas, N., Cullen, D. E., McDowell, E. P., Adelsperger, J., Frohling, S., Huntly, B. J.*, et al.* (2007). FLT3 mutations confer enhanced proliferation and survival properties to multipotent progenitors in a murine model of chronic myelomonocytic leukemia. Cancer Cell *12*, 367-380.

Lee, S. C., Miller, S., Hyland, C., Kauppi, M., Lebois, M., Di Rago, L., Metcalf, D., Kinkel, S. A., Josefsson, E. C., Blewitt, M. E.*, et al.* (2015). Polycomb repressive complex 2 component Suz12 is required for hematopoietic stem cell function and lymphopoiesis. Blood *126*, 167-175.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, Jill P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. Cell Systems *1*, 417-425.

Lin, Y. W., Slape, C., Zhang, Z., and Aplan, P. D. (2005). NUP98-HOXD13 transgenic mice develop a highly penetrant, severe myelodysplastic syndrome that progresses to acute leukemia. Blood *106*, 287-295.

Linton, P. J., and Dorshkind, K. (2004). Age-related changes in lymphocyte development and function. Nat Immunol *5*, 133-139.

Longo, D. L. (2017). Imatinib Changed Everything. New England Journal of Medicine *376*, 982-983.

Loven, J., Hoke, H. A., Lin, C. Y., Lau, A., Orlando, D. A., Vakoc, C. R., Bradner, J. E., Lee, T. I., and Young, R. A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell *153*, 320-334.

Lu, M., Tayu, R., Ikawa, T., Masuda, K., Matsumoto, I., Mugishima, H., Kawamoto, H., and Katsura, Y. (2005). The earliest thymic progenitors in adults are restricted to T, NK, and dendritic cell lineage and have a potential to form more diverse TCRbeta chains than fetal progenitors. J Immunol *175*, 5848-5856.

Luc, S., Buza-Vidas, N., and Jacobsen, S. E. (2008). Delineating the cellular pathways of hematopoietic lineage commitment. Semin Immunol *20*, 213-220.

Luc, S., Luis, T. C., Boukarabila, H., Macaulay, I. C., Buza-Vidas, N., Bouriez-Jones, T., Lutteropp, M., Woll, P. S., Loughran, S. J., Mead, A. J.*, et al.* (2012). The earliest thymic T cell progenitors sustain B cell and myeloid lineage potential. Nat Immunol *13*, 412-419.

Ma, X., Does, M., Raza, A., and Mayne, S. T. (2007). Myelodysplastic syndromes: incidence and survival in the United States. Cancer *109*, 1536-1542.

Magee, J. A., Piskounova, E., and Morrison, S. J. (2012). Cancer stem cells: impact, heterogeneity, and uncertainty. Cancer Cell *21*, 283-296.

Majeti, R., Park, C. Y., and Weissman, I. L. (2007). Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell Stem Cell *1*, 635-645.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343-349.

Matsumoto, A., Masuhara, M., Mitsui, K., Yokouchi, M., Ohtsubo, M., Misawa, H., Miyajima, A., and Yoshimura, A. (1997). CIS, a cytokine inducible SH2 protein, is a target of the JAK-STAT5 pathway and modulates STAT5 activation. Blood *89*, 3148-3154.

Matutes, E., Pickl, W. F., van't Veer, M., Morilla, R., Swansbury, J., Strobl, H., Attarbaschi, A., Hopfinger, G., Ashley, S., Bene, M. C.*, et al.* (2011). Mixed-phenotype acute leukemia: clinical and laboratory features and outcome in 100 patients defined according to the WHO 2008 classification. Blood *117*, 3163-3171.

Maude, S. L., Dolai, S., Delgado-Martin, C., Vincent, T., Robbins, A., Selvanathan, A., Ryan, T., Hall, J., Wood, A. C., Tasian, S. K.*, et al.* (2015). Efficacy of JAK/STAT pathway inhibition in murine xenograft models of early T-cell precursor (ETP) acute lymphoblastic leukemia. Blood *125*, 1759-1767.

McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res *40*, 4288-4297.

McCormack, M. P., Forster, A., Drynan, L., Pannell, R., and Rabbitts, T. H. (2003). The LMO2 T-cell oncogene is activated via chromosomal translocations or retroviral insertion during gene therapy but has no mandatory role in normal T-cell development. Mol Cell Biol *23*, 9003-9013.

McCormack, M. P., Young, L. F., Vasudevan, S., de Graaf, C. A., Codrington, R., Rabbitts, T. H., Jane, S. M., and Curtis, D. J. (2010). The Lmo2 oncogene initiates leukemia in mice by inducing thymocyte self-renewal. Science *327*, 879-883.

Mead, A. J., Kharazi, S., Atkinson, D., Macaulay, I., Pecquet, C., Loughran, S., Lutteropp, M., Woll, P., Chowdhury, O., Luc, S*., et al.* (2013). FLT3-ITDs instruct a myeloid differentiation and transformation bias in lymphomyeloid multipotent progenitors. Cell Rep *3*, 1766-1776.

Mead, A. J., Neo, W. H., Barkas, N., Matsuoka, S., Giustacchini, A., Facchini, R., Thongjuea, S., Jamieson, L., Booth, C. A. G., Fordham, N*., et al.* (2017). Niche-mediated depletion of the normal hematopoietic stem cell reservoir by Flt3-ITD-induced myeloproliferation. J Exp Med *214*, 2005-2021.

Metzeler, K. H., Herold, T., Rothenberg-Thurley, M., Amler, S., Sauerland, M. C., Gorlich, D., Schneider, S., Konstandin, N. P., Dufour, A., Braundl, K*., et al.* (2016). Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. Blood *128*, 686-698.

Milne, T. A., Zhao, K., and Hess, J. L. (2009). Chromatin immunoprecipitation (ChIP) for analysis of histone modifications and chromatin-associated proteins. Methods Mol Biol *538*, 409-423.

Miyoshi, H., Shimizu, K., Kozu, T., Maseki, N., Kaneko, Y., and Ohki, M. (1991). t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. Proceedings of the National Academy of Sciences *88*, 10431-10434.

Mizuki, M., Fenski, R., Halfter, H., Matsumura, I., Schmidt, R., Muller, C., Gruning, W., Kratz-Albers, K., Serve, S., Steur, C.*, et al.* (2000). Flt3 mutations from patients with acute myeloid leukemia induce transformation of 32D cells mediated by the Ras and STAT5 pathways. Blood *96*, 3907-3914.

Mochizuki-Kashio, M., Aoyama, K., Sashida, G., Oshima, M., Tomioka, T., Muto, T., Wang, C., and Iwama, A. (2015). Ezh2 loss in hematopoietic stem cells predisposes mice to develop heterogeneous malignancies in an Ezh1-dependent manner. Blood *126*, 1172-1183.

Mochizuki-Kashio, M., Mishima, Y., Miyagi, S., Negishi, M., Saraya, A., Konuma, T., Shinga, J., Koseki, H., and Iwama, A. (2011). Dependency on the polycomb gene Ezh2 distinguishes fetal from adult hematopoietic stem cells. Blood *118*, 6553-6561.

Mombaerts, P., Iacomini, J., Johnson, R. S., Herrup, K., Tonegawa, S., and Papaioannou, V. E. (1992). RAG-1-deficient mice have no mature B and T lymphocytes. Cell *68*, 869-877.

Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E.*, et al.* (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet *34*, 267-273.

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

Muto, T., Sashida, G., Oshima, M., Wendt, G. R., Mochizuki-Kashio, M., Nagata, Y., Sanada, M., Miyagi, S., Saraya, A., Kamio, A.*, et al.* (2013). Concurrent loss of Ezh2 and Tet2 cooperates in the pathogenesis of myelodysplastic disorders. J Exp Med *210*, 2627-2639.

Nagel, S., Venturini, L., Marquez, V. E., Meyer, C., Kaufmann, M., Scherr, M., MacLeod, R. A., and Drexler, H. G. (2010). Polycomb repressor complex 2 regulates HOXA9 and HOXA10, activating ID2 in NK/T-cell lines. Mol Cancer *9*, 151.

Nakamura-Ishizu, A., Takizawa, H., and Suda, T. (2014). The analysis, roles and regulation of quiescence in hematopoietic stem cells. Development *141*, 4656-4666.

Nikoloski, G., Langemeijer, S. M. C., Kuiper, R. P., Knops, R., Massop, M., Tonnissen, E. R. L. T. M., van der Heijden, A., Scheele, T. N., Vandenberghe, P., de Witte, T.*, et al.* (2010). Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. Nat Genet *42*, 665-667.

Notta, F., Doulatov, S., Laurenti, E., Poeppl, A., Jurisica, I., and Dick, J. E. (2011). Isolation of Single Human Hematopoietic Stem Cells Capable of Long-Term Multilineage Engraftment. Science *333*, 218-221.

Ntziachristos, P., Tsirigos, A., Van Vlierberghe, P., Nedjic, J., Trimarchi, T., Flaherty, M. S., Ferres-Marco, D., da Ros, V., Tang, Z., Siegle, J.*, et al.* (2012). Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. Nat Med *18*, 298-301.

Ntziachristos, P., Tsirigos, A., Welstead, G. G., Trimarchi, T., Bakogianni, S., Xu, L., Loizou, E., Holmfeldt, L., Strikoudis, A., King, B.*, et al.* (2014). Contrasting roles of histone 3 lysine 27 demethylases in acute lymphoblastic leukaemia. Nature *514*, 513-517.

Oguro, H., Ding, L., and Morrison, S. J. (2013). SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. Cell Stem Cell *13*, 102-116.

Oki, T., Kitaura, J., Watanabe-Okochi, N., Nishimura, K., Maehara, A., Uchida, T., Komeno, Y., Nakahara, F., Harada, Y., Sonoki, T.*, et al.* (2012). Aberrant expression of RasGRP1 cooperates with gain-of-function NOTCH1 mutations in T-cell leukemogenesis. Leukemia *26*, 1038-1045.

Orkin, S. H. (2000). Diversification of haematopoietic stem cells to specific lineages. Nat Rev Genet *1*, 57-64.

Orkin, S. H., and Zon, L. I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell *132*, 631-644.

Osato, M. (2004). Point mutations in the RUNX1//AML1 gene: another actor in RUNX leukemia. Oncogene *23*, 4284-4296.

Papaemmanuil, E., Gerstung, M., Malcovati, L., Tauro, S., Gundem, G., Van Loo, P., Yoon, C. J., Ellis, P., Wedge, D. C., Pellagatti, A.*, et al.* (2013). Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood *122*, 3616-3627; quiz 3699.

Pasini, D., Malatesta, M., Jung, H. R., Walfridsson, J., Willer, A., Olsson, L., Skotte, J., Wutz, A., Porse, B., Jensen, O. N., and Helin, K. (2010). Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. Nucleic Acids Res *38*, 4958-4969.

Patrick, K., Wade, R., Goulden, N., Mitchell, C., Moorman, A. V., Rowntree, C., Jenkinson, S., Hough, R., and Vora, A. (2014). Outcome for children and young people with Early T-cell precursor acute lymphoblastic leukaemia treated on a contemporary protocol, UKALL 2003. Br J Haematol *166*, 421-424.

Picelli, S., Faridani, O. R., Björklund, Å. K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat Protocols *9*, 171-181.

Pineault, N., Buske, C., Feuring-Buske, M., Abramovich, C., Rosten, P., Hogge, D. E., Aplan, P. D., and Humphries, R. K. (2003). Induction of acute myeloid leukemia in mice by the human leukemia-specific fusion gene NUP98-HOXD13 in concert with Meis1. Blood *101*, 4529-4538.

Qian, H., Buza-Vidas, N., Hyland, C. D., Jensen, C. T., Antonchuk, J., Mansson, R., Thoren, L. A., Ekblom, M., Alexander, W. S., and Jacobsen, S. E. (2007). Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. Cell Stem Cell *1*, 671-684.

Ramskold, D., Luo, S., Wang, Y. C., Li, R., Deng, Q., Faridani, O. R., Daniels, G. A., Khrebtukova, I., Loring, J. F., Laurent, L. C.*, et al.* (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol *30*, 777-782.

Raza-Egilmez, S. Z., Jani-Sait, S. N., Grossi, M., Higgins, M. J., Shows, T. B., and Aplan, P. D. (1998). NUP98-HOXD13 gene fusion in therapy-related acute myelogenous leukemia. Cancer Res *58*, 4269-4273.

Reynolds, N., Salmon-Divon, M., Dvinge, H., Hynes-Allen, A., Balasooriya, G., Leaford, D., Behrens, A., Bertone, P., and Hendrich, B. (2012). NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. EMBO J *31*, 593-605.

Richie Ehrlich, L. I., Serwold, T., and Weissman, I. L. (2011). In vitro assays misrepresent in vivo lineage potentials of murine lymphoid progenitors. Blood *117*, 2618-2624.

Riddell, J., Gazit, R., Garrison, B. S., Guo, G., Saadatpour, A., Mandal, P. K., Ebina, W., Volchkov, P., Yuan, G. C., Orkin, S. H., and Rossi, D. J. (2014). Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. Cell *157*, 549-564.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Rothenberg, E. V. (2011). T cell lineage commitment: identity and renunciation. J Immunol *186*, 6649-6655.

Roux, P. P., Shahbazian, D., Vu, H., Holz, M. K., Cohen, M. S., Taunton, J., Sonenberg, N., and Blenis, J. (2007). RAS/ERK signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation. J Biol Chem *282*, 14056-14064.

Roy, A., Cowan, G., Mead, A. J., Filippi, S., Bohn, G., Chaidos, A., Tunstall, O., Chan, J. K., Choolani, M., Bennett, P.*, et al.* (2012). Perturbation of fetal liver hematopoietic stem and progenitor cell development by trisomy 21. Proc Natl Acad Sci U S A *109*, 17579-17584.

Sanjuan-Pla, A., Macaulay, I. C., Jensen, C. T., Woll, P. S., Luis, T. C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Bouriez Jones, T.*, et al.* (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature *502*, 232-236.

Sashida, G., Harada, H., Matsui, H., Oshima, M., Yui, M., Harada, Y., Tanaka, S., Mochizuki-Kashio, M., Wang, C., Saraya, A.*, et al.* (2014). Ezh2 loss promotes development of myelodysplastic syndrome but attenuates its predisposition to leukaemic transformation. Nat Commun *5*, 4177.

Schepers, K., Pietras, E. M., Reynaud, D., Flach, J., Binnewies, M., Garg, T., Wagers, A. J., Hsiao, E. C., and Passegue, E. (2013). Myeloproliferative neoplasia remodels the endosteal bone marrow niche into a self-reinforcing leukemic niche. Cell Stem Cell *13*, 285-299.

Schilham, M. W., Wilson, A., Moerer, P., Benaissa-Trouw, B. J., Cumano, A., and Clevers, H. C. (1998). Critical involvement of Tcf-1 in expansion of thymocytes. J Immunol *161*, 3984-3991.

Schlenner, S. M., and Rodewald, H. R. (2010). Early T cell development and the pitfalls of potential. Trends Immunol *31*, 303-310.

Schwarz, B. A., and Bhandoola, A. (2006). Trafficking from the bone marrow to the thymus: a prerequisite for thymopoiesis. Immunol Rev *209*, 47-57.

Sebolt-Leopold, J. S., and Herrera, R. (2004). Targeting the mitogen-activated protein kinase cascade to treat cancer. Nat Rev Cancer *4*, 937-947.

Seita, J., Sahoo, D., Rossi, D. J., Bhattacharya, D., Serwold, T., Inlay, M. A., Ehrlich, L. I., Fathman, J. W., Dill, D. L., and Weissman, I. L. (2012). Gene Expression Commons: an open platform for absolute gene expression profiling. PLoS One *7*, e40321.

Sekeres, M. A., and Cutler, C. (2014). How we treat higher-risk myelodysplastic syndromes. Blood *123*, 829-836.

Shan, X., Fung, J. J., Kosaka, A., and Danet-Desnoyers, G. (2017). Replication Study: Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Elife *6*.

Shi, J., and Vakoc, C. R. (2014). The mechanisms behind the therapeutic activity of BET bromodomain inhibition. Mol Cell *54*, 728-736.

Shlush, L. I., Zandi, S., Mitchell, A., Chen, W. C., Brandwein, J. M., Gupta, V., Kennedy, J. A., Schimmer, A. D., Schuh, A. C., Yee, K. W*., et al.* (2014). Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature *506*, 328-333.

Shultz, L. D., Lyons, B. L., Burzenski, L. M., Gott, B., Chen, X., Chaleff, S., Kotb, M., Gillies, S. D., King, M., Mangada, J*., et al.* (2005). Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. J Immunol *174*, 6477-6489.

Simon, C., Chagraoui, J., Krosl, J., Gendron, P., Wilhelm, B., Lemieux, S., Boucher, G., Chagnon, P., Drouin, S., Lambert, R*., et al.* (2012). A key role for EZH2 and associated genes in mouse and human adult T-cell acute leukemia. Genes Dev *26*, 651-656.

Sood, R., Kamikubo, Y., and Liu, P. (2017). Role of RUNX1 in hematological malignancies. Blood *129*, 2070-2082.

Souroullas, G. P., Jeck, W. R., Parker, J. S., Simon, J. M., Liu, J. Y., Paulk, J., Xiong, J., Clark, K. S., Fedoriw, Y., Qi, J.*, et al.* (2016). An oncogenic Ezh2 mutation induces tumors through global redistribution of histone 3 lysine 27 trimethylation. Nat Med *22*, 632-640.

Sperling, A. S., Gibson, C. J., and Ebert, B. L. (2017). The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia. Nat Rev Cancer *17*, 5-19.

Stark, R., and Brown, G. (2011). DiffBind: differential binding analysis of ChIP-Seq peak data. http://bioconductororg/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBindpdf.

Steelman, L. S., Franklin, R. A., Abrams, S. L., Chappell, W., Kempf, C. R., Basecke, J., Stivala, F., Donia, M., Fagone, P., Nicoletti, F.*, et al.* (2011). Roles of the Ras/Raf/MEK/ERK pathway in leukemia therapy. Leukemia *25*, 1080-1094.

Steensma, D. P., Bejar, R., Jaiswal, S., Lindsley, R. C., Sekeres, M. A., Hasserjian, R. P., and Ebert, B. L. (2015). Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood *126*, 9-16.

Su, I. H., Basavaraj, A., Krutchinsky, A. N., Hobert, O., Ullrich, A., Chait, B. T., and Tarakhovsky, A. (2003). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat Immunol *4*, 124-131.

Su, I. H., Dobenecker, M. W., Dickinson, E., Oser, M., Basavaraj, A., Marqueron, R., Viale, A., Reinberg, D., Wulfing, C., and Tarakhovsky, A. (2005). Polycomb group protein ezh2 controls actin polymerization and cell signaling. Cell *121*, 425-436.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Sundaram, A. Y., Hughes, T., Biondi, S., Bolduc, N., Bowman, S. K., Camilli, A., Chew, Y. C., Couture, C., Farmer, A., Jerome, J. P.*, et al.* (2016). A comparative study of ChIP-seq sequencing library preparation methods. BMC Genomics *17*, 816.

Taussig, D. C., Vargaftig, J., Miraki-Moud, F., Griessinger, E., Sharrock, K., Luke, T., Lillington, D., Oakervee, H., Cavenagh, J., Agrawal, S. G.*, et al.* (2010). Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34(-) fraction. Blood *115*, 1976-1984.

Tehranchi, R., Woll, P. S., Anderson, K., Buza-Vidas, N., Mizukami, T., Mead, A. J., Astrand-Grundstrom, I., Strombeck, B., Horvat, A., Ferry, H.*, et al.* (2010). Persistent malignant stem cells in del(5q) myelodysplasia in remission. N Engl J Med *363*, 1025-1037.

Tie, F., Banerjee, R., Stratton, C. A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M. O., Scacheri, P. C., and Harte, P. J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development *136*, 3131-3141.

Treanor, L. M., Zhou, S., Janke, L., Churchman, M. L., Ma, Z., Lu, T., Chen, S. C., Mullighan, C. G., and Sorrentino, B. P. (2014). Interleukin-7 receptor mutants initiate early T cell precursor leukemia in murine thymocyte progenitors with multipotent potential. J Exp Med *211*, 701-713.

Vainchenker, W., and Constantinescu, S. N. (2013). JAK/STAT signaling in hematological malignancies. Oncogene *32*, 2601-2613.

Van Vlierberghe, P., Ambesi-Impiombato, A., Perez-Garcia, A., Haydu, J. E., Rigo, I., Hadler, M., Tosello, V., Della Gatta, G., Paietta, E., Racevskis, J.*, et al.* (2011). ETV6 mutations in early immature human T cell leukemias. J Exp Med *208*, 2571-2579.

Varambally, S., Dhanasekaran, S. M., Zhou, M., Barrette, T. R., Kumar-Sinha, C., Sanda, M. G., Ghosh, D., Pienta, K. J., Sewalt, R. G., Otte, A. P.*, et al.* (2002). The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature *419*, 624-629.

Venezia, T. A., Merchant, A. A., Ramos, C. A., Whitehouse, N. L., Young, A. S., Shaw, C. A., and Goodell, M. A. (2004). Molecular Signatures of Proliferation and Quiescence in Hematopoietic Stem Cells. PLoS Biology *2*, e301.

Wada, H., Masuda, K., Satoh, R., Kakugawa, K., Ikawa, T., Katsura, Y., and Kawamoto, H. (2008). Adult T-cell progenitors retain myeloid potential. Nature *452*, 768-772.

Wang, J. C., and Dick, J. E. (2005). Cancer stem cells: lessons from leukemia. Trends Cell Biol *15*, 494-501.

Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Peng, W., Zhang, M. Q., and Zhao, K. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet *40*, 897-903.

Watanabe-Okochi, N., Kitaura, J., Ono, R., Harada, H., Harada, Y., Komeno, Y., Nakajima, H., Nosaka, T., Inaba, T., and Kitamura, T. (2008). AML1 mutations induced MDS and MDS/AML in a mouse BMT model. Blood *111*, 4297-4308.

Weinberg, O. K., and Arber, D. A. (2010). Mixed-phenotype acute leukemia: historical overview and a new definition. Leukemia *24*, 1844-1851.

Woll, Petter S., Kjällquist, U., Chowdhury, O., Doolittle, H., Wedge, David C., Thongjuea, S., Erlandsson, R., Ngara, M., Anderson, K., Deng, Q.*, et al.* (2014). Myelodysplastic Syndromes Are Propagated by Rare and Distinct Human Cancer Stem Cells In Vivo. Cancer Cell *25*, 794-808.

Xie, H., Xu, J., Hsu, J. H., Nguyen, M., Fujiwara, Y., Peng, C., and Orkin, S. H. (2014). Polycomb repressive complex 2 regulates normal hematopoietic stem cell function in a developmental-stage-specific manner. Cell Stem Cell *14*, 68-80.

Yap, D. B., Chu, J., Berg, T., Schapira, M., Cheng, S. W., Moradian, A., Morin, R. D., Mungall, A. J., Meissner, B., Boyle, M.*, et al.* (2011). Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. Blood *117*, 2451-2459.

Yoshihara, H., Arai, F., Hosokawa, K., Hagiwara, T., Takubo, K., Nakamura, Y., Gomei, Y., Iwasaki, H., Matsuoka, S., Miyamoto, K.*, et al.* (2007). Thrombopoietin/MPL Signaling Regulates Hematopoietic Stem Cell Quiescence and Interaction with the Osteoblastic Niche. Cell Stem Cell *1*, 685-697.

Yu, S., Zhou, X., Steinke, F. C., Liu, C., Chen, S. C., Zagorodna, O., Jing, X., Yokota, Y., Meyerholz, D. K., Mullighan, C. G.*, et al.* (2012). The TCF-1 and LEF-1 transcription factors have cooperative and opposing roles in T cell development and malignancy. Immunity *37*, 813-826.

Yui, M. A., and Rothenberg, E. V. (2014). Developmental gene networks: a triathlon on the course to T cell identity. Nat Rev Immunol *14*, 529-545.

Zarnegar, M. A., Chen, J., and Rothenberg, E. V. (2010). Cell-type-specific activation and repression of PU.1 by a complex of discrete, functionally specialized cis-regulatory elements. Mol Cell Biol *30*, 4922-4939.

Zhang, B., Ho, Y. W., Huang, Q., Maeda, T., Lin, A., Lee, S. U., Hair, A., Holyoake, T. L., Huettner, C., and Bhatia, R. (2012a). Altered microenvironmental regulation of leukemic and normal stem cells in chronic myelogenous leukemia. Cancer Cell *21*, 577-592.

Zhang, J., Ding, L., Holmfeldt, L., Wu, G., Heatley, S. L., Payne-Turner, D., Easton, J., Chen, X., Wang, J., Rusch, M.*, et al.* (2012b). The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. Nature *481*, 157-163.

Zhang, J. A., Mortazavi, A., Williams, B. A., Wold, B. J., and Rothenberg, E. V. (2012c). Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. Cell *149*, 467-482.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., and Liu, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhao, X., Valen, E., Parker, B. J., and Sandelin, A. (2011). Systematic clustering of transcription start site landscapes. PLoS One *6*, e23409.

Zhou, T., Kinney, M. C., Scott, L. M., Zinkel, S. S., and Rebel, V. I. (2015). Revisiting the case for genetically engineered mouse models in human myelodysplastic syndrome research. Blood *126*, 1057-1068.

Zitvogel, L., Rusakiewicz, S., Routy, B., Ayyoub, M., and Kroemer, G. (2016). Immunological off-target effects of imatinib. Nat Rev Clin Oncol *13*, 431-446.

Zuber, J., Shi, J., Wang, E., Rappaport, A. R., Herrmann, H., Sison, E. A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M.*, et al.* (2011). RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature *478*, 524-528.