

**The molecular and cellular basis for oncogene
collaboration in acute myeloid leukaemia**



Cristina Assunta Di Genua

Jesus College

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2018

ABSTRACT

Acute myeloid leukaemia (AML) is an aggressive malignancy of the bone marrow caused by the uncontrollable proliferation of immature myeloid cells that fail to terminally differentiate. Mutations in AML are acquired hierarchically, where pre-leukaemic mutations are acquired within haematopoietic stem cells (HSCs); while secondary mutations, such as signalling mutations, are acquired within downstream pre-leukaemic progenitor cells. Why certain mutations are pre-leukaemic and thus acquired within the HSCs, compared to others, and the mechanism by which pre-leukaemic and secondary mutations collaborate, are still poorly understood.

Utilising knock-in mutant mouse models we showed that HSCs expressing the pre-leukaemic fusion protein Aml1-ETO conferred a competitive advantage. However, activated K-Ras in Aml1-ETO-expressing HSCs had a marked detrimental effect, leading to cell cycle activation and loss of self-renewal associated gene expression. This provides a mechanism for the observed absence of a signalling mutation in pre-malignant HSCs.

The specific association between the biallelic *CEBPA* and zinc finger 1 (ZF1) *GATA2* mutations is observed in AML as well as acute erythroid leukaemia (AEL). We show the mutations collaborate to generate a model of AEL with bi-lineage transformation of the myeloid and erythroid lineages. This is due to the biallelic *Cebpa* mutations reprogramming the neutrophil-monocyte progenitors (NMPs) to acquire ectopic erythroid potential, while the *Gata2* mutation synergises with the *Cebpa* mutations to impair erythroid differentiation progression. The investigation of ZF1 *Gata2* mutation in isolation at both heterozygous and homozygous level showed the ZF1 mutation results in loss of *Gata2* function only when the mutation is homozygous. As ZF1 *GATA2* mutations are heterozygous in patients, this data supports the findings that ZF1 *GATA2* and *CEBPA* mutations must work synergistically to cause leukaemia development.

ACKNOWLEDGMENTS

I would like to thank my supervisors Prof. Claus Nerlov and Prof. Adam Mead for the opportunity to carry out my D.Phil in their labs, and for their guidance and support through my studies.

I am also thankful to all my colleagues at the Weatherall Institute of Molecular Medicine for teaching me and helping me through my PhD. I would especially like to thank Dr. Ruggiero Norfo for all his help in my first year that enabled me to make a great start on my PhD. I am also thankful to my thesis committee, the FACS, genome engineering, virus production, sequencing and single cell facilities, as well as the computational biology research group, for their help with my projects.

I am also very grateful to my family, friends, and most of all my partner Rob, for their support during my PhD.

TABLE OF CONTENT

ABSTRACT.....	2
ACKNOWLEDGMENTS.....	3
TABLE OF CONTENT.....	4
LIST OF FIGURES	8
LIST OF TABLES.....	11
COMMONLY USED ABBREVIATIONS	12
CHAPTER ONE: Introduction.....	15
1.1 Haematopoiesis.....	15
1.2 Acute myeloid leukaemia	17
1.2.1 <i>Mutations in AML.....</i>	20
1.2.2 <i>Leukaemic stem cell</i>	23
1.2.3 <i>Clonal evolution in AML</i>	24
1.2.4 <i>Role of AML1-ETO as a pre-leukaemic mutation in AML</i>	26
1.2.5 <i>Role of biallelic CEBPA mutations as pre-leukaemic mutations in AML</i>	32
1.2.6 <i>Role of K-RAS mutations as secondary mutations in AML.....</i>	38
1.2.7 <i>Role of zinc finger 1 GATA2 mutations as secondary mutations in AML with biallelic CEBPA mutations</i>	41
1.3 Modelling and analysis of co-occurring mutations in AML	47
1.3.1 <i>Lentiviral and retroviral transduction</i>	48
1.3.2 <i>Knock-in and knock-out mouse models</i>	51
1.3.3 <i>Gene expression and mutation analysis of AML patient samples</i>	53
1.4 Acute erythroid leukaemia	55
1.5 Aims of the present work	58

CHAPTER TWO: Materials and methods.....	61
2.1 Mice.....	61
2.1.1 Husbandry.....	61
2.1.2 Mouse strains.....	61
2.2 Genotyping	62
2.3 Timed mating.....	64
2.4 Cell preparation	64
2.4.1 Bone marrow.....	64
2.4.2 c-Kit enrichment of BM cells	64
2.4.3 Spleen	65
2.4.4 Peripheral blood	66
2.4.5 Peritoneal cells.....	66
2.4.6 Foetal liver cells.....	67
2.4.7 Freezing cells	67
2.4.8 Thawing frozen cells.....	68
2.5 Flow cytometry and fluorescence-activated cell sorting	68
2.5.1 Antibody staining and analysis.....	68
2.5.2 HSC cell cycle	73
2.5.3 Fluorescence-activated cell sorting.....	73
2.5.4 Instrument and analysis.....	73
2.6 Transplantation assays.....	74
2.7 In vitro assays.....	75
2.7.1 Serial replating	75
2.7.2 Cell culture for colony gene expression and morphology analysis	75
2.7.3 Myeloid cell culture	76
2.8 Cell morphology.....	76

2.9	Gene expression analysis.....	77
2.9.1	<i>RNA-sequencing</i>	77
2.9.2	<i>Multiplex quantitative real-time PCR analysis</i>	78
2.10	Chromatin immunoprecipitation-sequencing data analysis.....	79
2.11	Statistical analysis.....	80
CHAPTER THREE: A study of the pre-leukaemic fusion protein Aml1-ETO in combination with a signalling mutation in K-Ras.....		81
3.1	Introduction.....	81
3.2	Results	82
3.2.1	<i>Aml1-ETO ameliorates key features of the MPN phenotype caused by K-RasG12D.....</i>	<i>82</i>
3.2.2	<i>K-RasG12D reverses the HSC expansion associated with Aml1-ETO.....</i>	<i>92</i>
3.2.3	<i>K-RasG12D induces loss of quiescence in Aml1-ETO HSCs.....</i>	<i>96</i>
3.3	Discussion	100
3.3.1	<i>AM and KM recapitulate the phenotype in previous published data</i>	<i>100</i>
3.3.2	<i>K-RasG12D is detrimental to normal and pre-leukaemic HSCs</i>	<i>101</i>
3.3.3	<i>Identification of Gzmb and Gja1 as important genes for HSCs</i>	<i>104</i>
3.3.4	<i>Conclusion.....</i>	<i>105</i>
CHAPTER FOUR: A study of the biallelic Cebpa mutations in combination with a novel zinc finger 1 Gata2 mutation.....		107
4.1	Introduction.....	107
4.2	Results	108
4.2.1	<i>Generation of a genetic model of Cebpa and Gata2 mutant AML</i>	<i>108</i>
4.2.2	<i>Biallelic Cebpa and ZF1 Gata2 mutations synergise during leukaemogenesis.....</i>	<i>109</i>

4.2.3	<i>Mice with combined biallelic Cebpa and ZF1 Gata2 mutation develop an erythroleukaemia</i>	114
4.2.4	<i>Cebpa and Gata2 mutant AEL is sustained by LSCs with a NMP immune-phenotype</i>	124
4.2.5	<i>Biallelic Cebpa mutations introduce an ectopic erythroid lineage potential in NMPs</i>	134
4.2.6	<i>ZF1 Gata2 mutation lead to accumulation of erythroid progenitors at the CFU-E stage</i>	140
4.3	Discussion	142
4.3.1	<i>Neomorphic neutrophil-erythroid progenitor act as a bi-potent LSC in AEL</i>	142
4.3.2	<i>Cebpa^{K/L};Gata2^{G320D/+} as an AEL model</i>	144
4.3.3	<i>Conclusion</i>	145
CHAPTER FIVE: Characterising the effect of Gata2G320D on haematopoiesis		146
5.1	Introduction.....	146
5.2	Results	146
5.2.1	<i>Gata2 homozygous mutation is detrimental to HSC function</i>	146
5.2.2	<i>Gata2 homozygous mutation results in a loss of mast cell development</i> ..	150
5.2.3	<i>Gata2 homozygous mutation leads to block of mast cell potential at the basophil-mast cell progenitor stage</i>	157
5.3	Discussion	161
5.3.1	<i>Gata2 homozygous mutant phenotype is similar to Gata2 knock-out</i>	161
5.3.2	<i>Comparison with previous functional studies of Gata2G320D</i>	163
5.4	Conclusion.....	163
CONCLUSION		165
BIBLIOGRAPHY		166

LIST OF FIGURES

Figure 1.1 Revised model of the haematopoietic hierarchy	16
Figure 1.2 Nine distinct groups of mutations in AML	21
Figure 1.3 Co-mutations between the nine distinct groups of mutations in AML.....	22
Figure 1.4 Hierarchy of mutation acquisition in AML	26
Figure 1.5 AML1-ETO domains	27
Figure 1.6 Biallelic <i>CEBPA</i> mutations	36
Figure 1.7 <i>GATA2</i> mutations identified in haematological malignancies and disorders ...	46
Figure 3.1 Experimental design	83
Figure 3.2 Aml1-ETO ameliorates key features of the MPN phenotype caused by K- RasG12D.....	86
Figure 3.3 Aml1-ETO ameliorates the increase in Mac1 ⁺ Gr1 ^{lo} myeloid cells in the PB caused by K-RasG12D	87
Figure 3.4 Aml1-ETO causes a decrease in lymphoid cells in the PB	88
Figure 3.5 Aml1-ETO reverses some of the myelo-erythroid progenitor cell phenotypes caused by K-RasG12D	91
Figure 3.6 Phenotypic HSC expansion caused by Aml1-ETO is reversed by K-RasG12D	93
Figure 3.7 Functional HSC expansion caused by Aml1-ETO is reversed by K-RasG12D95	
Figure 3.8 AKM HSCs are characterised by loss of quiescence and HSC-associated gene expression	97
Figure 3.9 RNA-sequencing reveals distinct molecular signature in AKM HSCs	99
Figure 4.1 Experimental design	109
Figure 4.2 Biallelic <i>Cebpa</i> combined with <i>Gata2</i> mutation causes an increase in myeloid reconstitution overtime in the PB	112
Figure 4.3 Biallelic <i>Cebpa</i> synergises with <i>Gata2</i> mutation to cause a more rapid disease development consistent with AML.....	113

Figure 4.4 Myeloid and erythroid blast morphology in ~40% of KLG cases.....	116
Figure 4.5 KL and KLG leukaemic mice contain myeloblasts and an expansion of the immature c-Kit ⁺ Mac1 ^{lo} cells in the BM and spleen.....	117
Figure 4.6 KLG-E mice contain erythroblasts in the BM and spleen.....	118
Figure 4.7 KLG-E mice develop a more aggressive disease than KLG-M mice	120
Figure 4.8 Secondary transplantation of bulk leukaemic KLG-E cells replicates the original bi-lineage disease phenotype	123
Figure 4.9 LK cells expanded in leukaemic KLG mice	124
Figure 4.10 Prominent NMP population present in KL and KLG leukaemic mice	127
Figure 4.11 L-NMPs act as the LSCs in both KLG-M and KLG-E mice re-capitulating their original diseases	129
Figure 4.12 Suppression of myeloid and erythroid differentiation programmes in leukaemic KLG mice	133
Figure 4.13 Biallelic <i>Cebpa</i> mutations leads to a loss of myeloid and an enrichment in erythroid gene expression.....	135
Figure 4.14 Biallelic <i>Cebpa</i> mutations lead to a loss of myeloid and an increase in erythroid gene expression in NMPs at the single cell level.....	137
Figure 4.15 Biallelic <i>Cebpa</i> mutations lead to ectopic erythroid potential in pre-leukaemic KLG NMPs.....	139
Figure 4.16 <i>Gata2</i> mutation synergises with biallelic <i>Cebpa</i> mutations resulting in an accumulation of EMkMPPs and CFU-Es.....	141
Figure 5.1 Schematic representation of WT, targeted allele, and constitutive knock-in allele after Flp recombination	147
Figure 5.2 The effect of <i>Gata2</i> mutation on blood cells and BM counts.....	148
Figure 5.3 <i>Gata2</i> homozygous mutation causes a loss of phenotypic HSCs	149
Figure 5.4 <i>Gata2</i> homozygous mutation causes a loss of functional HSCs.....	150
Figure 5.5 <i>Gata2</i> mutation does not affect the myelo-erythroid progenitors	151
Figure 5.6 Trend towards a decrease in EoMPs with <i>Gata2</i> mutation	152

Figure 5.7 <i>Gata2</i> mutation does not affect B and T cells.....	153
Figure 5.8 <i>Gata2</i> homozygous mutation results in a loss of mast cells	154
Figure 5.9 <i>Gata2</i> mutation results in a loss of mast cell generation <i>in vitro</i>	155
Figure 5.10 Gene expression analysis of lineage commitment genes in <i>Gata2</i> homozygous mutant EoMPs.....	156
Figure 5.11 Block in mast cell development at the BMCP stage in <i>Gata2</i> ^{G320D/G320D} mice	158
Figure 5.12 <i>Gata2</i> homozygous mutation leads to a loss of mast cell protease gene expression	160

LIST OF TABLES

Table 1.1 FAB classification of AML.....	18
Table 1.2 WHO classification of AML and related neoplasms	19
Table 2.1 Genotyping primers.....	63
Table 2.2 Antibodies used for FACS staining.....	70
Table 4.1 Bulk transplantation of leukaemic KLG-M and KLG-E cells	121
Table 4.2 Transplantation of purified L-NMP from KLG-M and KLG-E mice, and L-EoMP and L-EB from KLG-E mice, into secondary recipients	130
Table 4.3 Percentage of erythroblasts in LKCD45 ⁻ cells from KLG-E BM.....	131

COMMONLY USED ABBREVIATIONS

4-OHT: 4-hydroxytamoxifen

7AAD: 7-amino-actinomycin-D

AEL: Acute erythroid leukaemia

AKM: *Aml1*^{ETO/+}; *Kras*^{G12D/+}; *Mx1-Cre*^{tg/+}

AM: *Aml1*^{ETO/+}; *Mx1-Cre*^{tg/+}

AML: Acute myeloid leukaemia

ANOVA: Analysis of variance

BM: Bone marrow

BMCP: Basophil-mast cell progenitor

CFU-E: Colony forming unit-erythrocyte

ChIP: Chromatin immunoprecipitation

CLP: Common lymphoid progenitor

CML: Chronic myeloid leukaemia

CMML: Chronic myelomonocytic leukaemia

CMP: Common myeloid progenitor

DMSO: Dimethyl sulfoxide

EB: Erythroblast

EMkMPP: Erythroid-megakaryocyte primed multipotent progenitor

ENU: *N*-ethyl-*N*-nitrosourea

EoMP: Eosinophil-mast cell progenitor

ES: Embryonic stem

FAB: French-American-British

FACS: Fluorescence-activated cell sorting

FCS: Foetal calf serum

FL: Foetal liver

F-MuLV: Friend murine leukaemia virus

G: *Gata2*^{G320D/+}

GMP: Granulocyte-monocyte progenitor

GSEA: Gene set enrichment analysis

HSC: Haematopoietic stem cell

HSPC: Haematopoietic stem and progenitor cell

IMDM: Iscove's modified Dulbecco's medium

JMML: Juvenile myelomonocytic leukaemia

KL: *Cebpa*^{K313KK/Lp30}

KLG: *Cebpa*^{K313KK/Lp30}; *Gata2*^{G320D/+}

KM: *Kras*^{G12D/+}; *Mx1-Cre*^{tg/+}

L-EB: Leukaemic erythroblast

L-EoMP: Leukaemic eosinophil-mast cell progenitor

LK: Lineage⁻c-Kit⁺

L-MB: Leukaemic myeloblast

LMPP: Lymphoid-primed multipotent progenitor

L-NMP: Leukaemic neutrophil-monocyte progenitor

LSC: Leukaemic stem cell

LSK: Lineage⁻Sca1⁺c-Kit⁺

MB: Myeloblast

MDS: Myelodysplastic syndrome

MkP: Megakaryocyte progenitor

MPN: Myeloproliferative neoplasm

MPP: Multipotent progenitor

NA: Not applicable

NK: Natural killer

NMP: Neutrophil-monocyte progenitor

PB: Peripheral blood

PBS: Phosphate-buffered saline

Poly(I:C): Polyinosinic:polycytidylic acid

Pre-CFU-E: Pre-colony forming unit-erythrocyte

Pre-GM: Pre-granulocyte-monocyte

Pre-MegE: Pre-megakaryocyte-erythroid

Pre-NM: Pre-neutrophil-monocyte

RBC: Red blood cell

RPKM: Reads per kilobase of transcript per million

SCF: Stem cell factor

T-ALL: T-cell acute lymphoblastic leukaemia

TSS: Transcription start site

VAF: Variant allele frequency

WBC: White blood cell

WHO: World Health Organization

WT: Wild type

ZF: Zinc finger

CHAPTER ONE: Introduction

1.1 Haematopoiesis

Haematopoiesis is a hierarchical process of forming all the different blood cells. The process is highly regulated with the haematopoietic stem cells (HSCs) at the top of the hierarchy where all blood cells are derived from. The HSCs have two key properties: multipotency, which is the ability to differentiate into all the blood cells; and self-renewal, which allows HSCs to replicate without losing their multipotency (Nemeth & Bodine, 2007).

Originally it was thought that the first lineage commitment step during haematopoiesis was the complete separation of myelopoiesis and lymphopoiesis through the common myeloid progenitor (CMP), defined as $\text{Lin}^- \text{Sca}^- \text{c-Kit}^+ \text{CD34}^+ \text{Fc}\gamma\text{RII/III}^{\text{low}}$, and the common lymphoid progenitor (CLP), defined as $\text{Lin}^- \text{Sca}^{\text{Int}} \text{c-Kit}^{\text{Int}} \text{Il7}\alpha^+$ (Seita & Weissman, 2010). However, Adolfsson *et al.*, found that $\text{Lin}^- \text{Sca}^+ \text{c-Kit}^+$ (LSK) $\text{CD34}^+ \text{Flt3}^+$ cells had lymphoid and myeloid potential but lost the ability to produce megakaryocyte and erythroid cells. This suggests that the HSC lineage restriction starts with the loss of megakaryocyte and erythroid potential through the lymphoid-primed multipotent progenitor (LMPP) (Adolfsson *et al.*, 2005). The CMP is also thought to be a heterogenous population (Nutt, Metcalf, D'Amico, Polli, & Wu, 2005) and has since been subdivided into different myelo-erythroid

progenitors (Pronk et al., 2007). In this model all myeloid cells are generated through the same myeloid-restricted pre-granulocyte-monocyte progenitor (pre-GM), defined as $\text{Lin}^- \text{Sca1}^- \text{c-Kit}^+ \text{CD41}^+ \text{Fc}\gamma\text{RII/III}^- \text{CD150}^- \text{CD105}^-$, however it has recently been shown that there are distinct myeloid progenitor differentiation pathways; one which expresses *Gata1* that generates mast cell, eosinophils, megakaryocytes and erythroid cells, and another pathway that lacks *Gata1* expression that generates monocytes, neutrophils and lymphocytes (Drissen et al., 2016) (Figure 1.1).

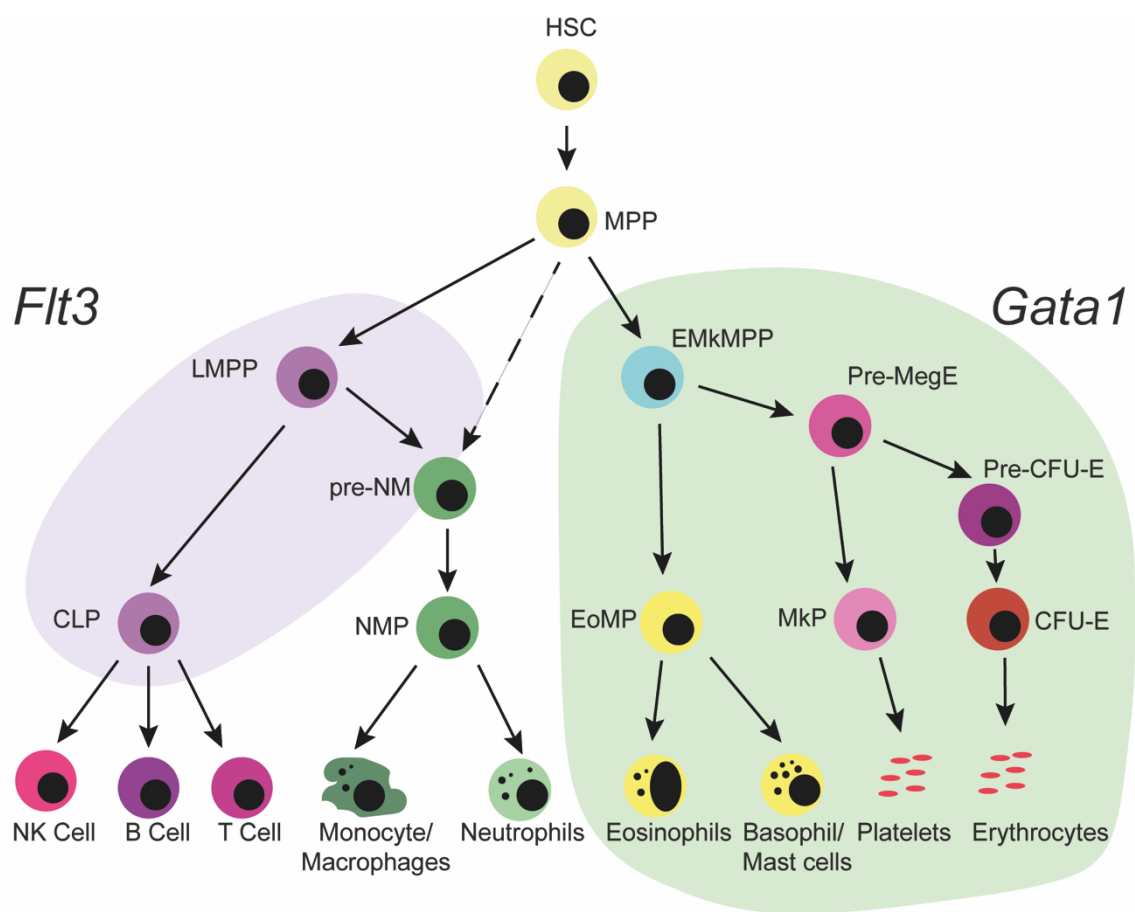


Figure 1.1 Revised model of the haematopoietic hierarchy

A schematic of the revised model of the haematopoietic hierarchy based on the early branch point generating the LMPPs and erythroid-megakaryocyte primed multipotent progenitor (EMkMPP), where the monocyte/macrophage and neutrophil lineage along with the lymphoid lineage separate away from the eosinophil, basophil/mast cell and

megakaryocyte/erythroid lineages according to their *Gata1* expression. The *Gata1* and *Flt3* expression domains are indicated. There is no direct evidence at present for the Flt3⁺ fraction of the pre-neutrophil-monocyte (pre-NM) derivation from the multipotent progenitors (MPPs), therefore the pathway remains theoretical and is indicated by the dash arrow. HSC, haematopoietic stem cell; LMPP, lymphoid-primed multipotent progenitor; CLP, common lymphoid progenitor; NMP, neutrophil-monocyte progenitor; EoMP, Eosinophil-mast cell progenitor; pre-MegE, pre-megakaryocyte-erythroid; MkP, megakaryocyte progenitor; Pre-CFU-E, pre-colony forming unit-erythrocyte; CFU-E, colony forming unit-erythrocyte. Adapted from (Drissen et al., 2016).

1.2 Acute myeloid leukaemia

Acute myeloid leukaemia (AML) leads to the haematopoietic hierarchy becoming perturbed resulting in the uncontrollable proliferation of the immature myeloid cells that fail to terminally differentiate (Estey & Dohner, 2006; Sykes, Kokkaliaris, Milsom, Levine, & Majeti, 2015). This is then accompanied by multilineage cytopenias (Dohner, Weisdorf, & Bloomfield, 2015). The lack of normal blood cells including red blood cells (RBC), platelets and granulocytes eventually contributes to the morbidity in AML (Thomas & Majeti, 2017). Chemotherapy is still the main form of treatment for patients and outcome has not significantly improved in the last 30 years with five year overall survival rate for patients < 60 years between 35-40%, and a median overall survival of approximately one years (Dohner et al., 2010; Dohner et al., 2015).

AML compromises a heterogenous group of disorders (Lowenberg, Downing, & Burnett, 1999). The French-American-British (FAB) Cooperative Group initially developed a classification system to recognise the heterogeneity in patients based on morphology and cytochemical characteristics (Bennett et al., 1976, 1991) (Table 1.1).

Table 1.1 FAB classification of AML

FAB	
Classification	Phenotype
M0	Minimally differentiated leukaemia
M1	Myeloblastic leukaemia without maturation
M2	Myeloblastic leukaemia with maturation
M3	Promyelocytic leukaemia
M4	Myelomonocytic leukaemia
M5	Monocytic leukaemia
M6	Erythroleukaemia
M7	Megakaryoblastic leukaemia

Advances in detecting cytogenetic and molecular abnormalities in combination with a better understanding of disease biology have led to a new classification provided by the World Health Organization (WHO) that integrated genetic, immunophenotypic, biological and clinical features to define specific AML subcategories (Arber et al., 2016) (Table 1.2). However, WHO authors have still struggled to incorporate all gene mutations recently discovered in AML (Cancer Genome Atlas Research, 2013).

Table 1.2 WHO classification of AML and related neoplasms

WHO AML and related neoplasms
AML with recurrent genetic abnormalities
AML with t(8;21)(q22;q22.1); <i>AML1-ETO</i>
AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); <i>CBFB-MYH11</i>
APL with <i>PML-RARA</i>
AML with t(9;11)(p21.3;q23.3); <i>MLLT3-KMT2A</i>
AML with t(6;9)(p23;q34.1); <i>DEK-NUP214</i>
AML with inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2); <i>GATA2, MECOM</i>
AML (megakaryoblastic) with t(1;22)(p13.3;q13.3); <i>RBM15-MKL1</i>
Provisional entity: AML with <i>BCR-ABL1</i>
AML with mutated <i>NPM1</i>
AML with biallelic mutations of <i>CEBPA</i>
Provisional entity: AML with mutated <i>RUNX1</i>
AML with myelodysplasia-related changes
Therapy-related myeloid neoplasms
AML, NOS (not otherwise specified)
AML with minimal differentiation
AML without maturation
AML with maturation
Acute myelomonocytic leukaemia
Acute monoblastic/monocytic leukaemia
Pure erythroid leukaemia
Acute megakaryoblastic leukaemia
Acute basophilic leukaemia

Acute panmyelosis with myelofibrosis

Myeloid sarcoma

Myeloid proliferations related to Down syndrome

Transient abnormal myelopoiesis (TAM)

Myeloid leukaemia associated with Down syndrome

1.2.1 Mutations in AML

Mutations identified in AML were originally categorized into two classes. Class I mutations conferred a proliferation and survival advantage and would occur in genes encoding receptor and non-receptor tyrosine kinases, such as *JAK2*, *KIT* and *FLT3*, as well as genes encoding GTPase activity, such as *NRAS* and *KRAS*. Class II mutations would impair differentiation and promote self-renewal, and would occur in genes encoding transcription factors. Gene function would be disrupted by either creating fusion proteins such as PML-RAR α , AML1-ETO, CBF β -SMMHC or point mutations such as *CEBPA* and *AML1* (Gilliland & Griffin, 2002). Gilliland and Griffin proposed a two-hit model where mutations from both classes are necessary for AML to occur. This is supported by the observation that class I (Braun et al., 2004; Chu et al., 2012) and class II mutations (Castilla et al., 1999; Higuchi et al., 2002) alone are unable to cause leukaemic transformation in mouse models. The model also proposed that the mutational classes were mutually exclusive and predicted that only one mutation in each class would be required for AML to develop (Gilliland & Griffin, 2002). Advances in genome sequencing have shown that the two-hit model is oversimplified, with novel mutations occurring in genes that are involved in epigenetic control, histone modification and splicing that do not fit within either class (Conway O'Brien, Prideaux, & Chevassut, 2014). Analysis of 200 *de novo* AML patient samples has identified nine categories of mutations defined by their

biological function: transcription-factor fusions (25%), *NPM1* (27%), tumour-suppressors (16%), DNA-methylation-related genes (44%), activated signalling genes (59%), chromatin-modifying genes (30%), myeloid transcription-factor genes (22%), cohesion-complex genes (13%), and spliceosome-complex genes (14%) (Figure 1.2). On average patients had five mutations in genes that are recurrently mutated in AML. Between the distinct groups are both co-occurring mutations, such as transcription factor fusions with activated signalling genes, and mutually exclusive groups such as myeloid transcription factor and tumour suppressors (Cancer Genome Atlas Research, 2013) (Figure 1.3). Co-occurring mutations may suggest collaboration between different pathways, whereas mutually exclusive mutations may suggest mutations in such pathways drive initiation or development of AML (S. J. Chen, Shen, & Chen, 2013).

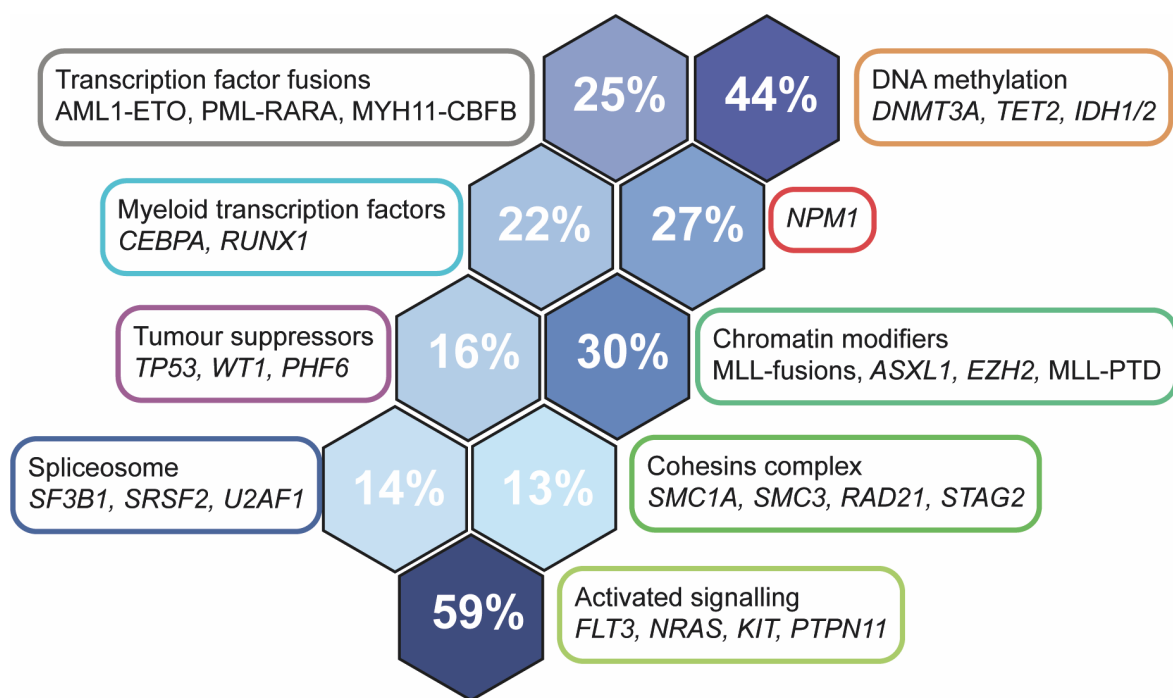


Figure 1.2 Nine distinct groups of mutations in AML

Nine distinct functional groups of mutations identified in AML (Cancer Genome Atlas Research, 2013). Proportion of AML patients with mutations within each functional group are indicated as a percentage. Adapted from (Grove & Vassiliou, 2014).

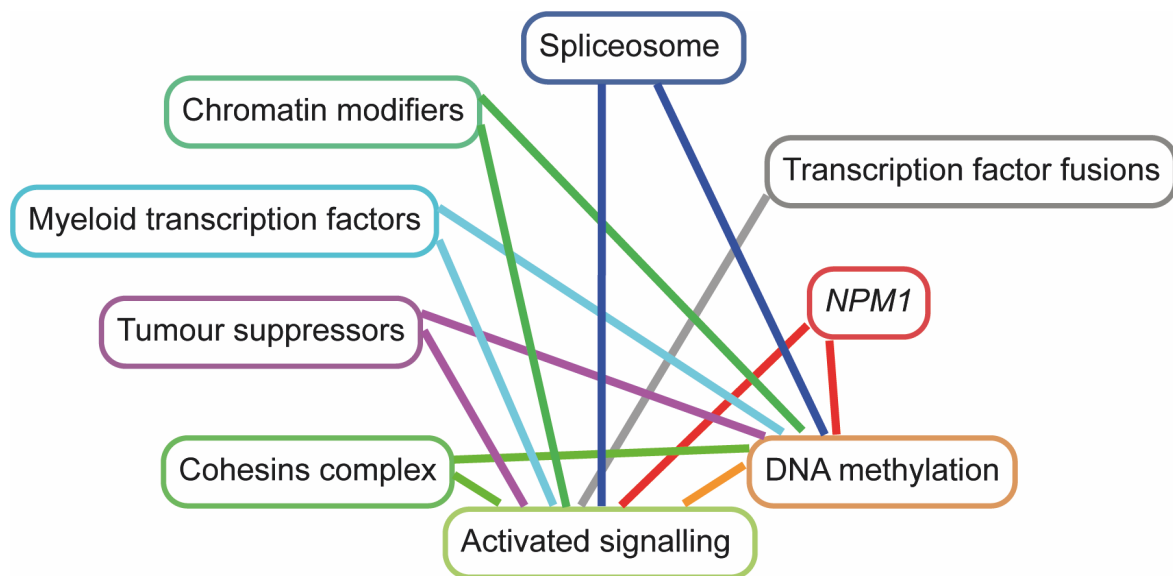


Figure 1.3 Co-mutations between the nine distinct groups of mutations in AML

Schematic showing the cooperation and mutual exclusivity of mutations between the nine different mutation functional groups in AML. The lines that connect the distinct categories reflect the association between the different groups. Adapted from (S. J. Chen et al., 2013).

The use of genetic data to assist with disease classification and clinical practices is currently being investigated. Papaemmanuil *et al.*, investigated 1540 patients that were treated in prospective trials by targeted sequencing of 111 genes and cytogenetic analysis, and found a pattern of co-mutations which classified AML patients into 11 categories that had distinct clinical phenotypes and outcomes. (Papaemmanuil et al., 2016). This study not only illustrated the association of mutations within different functional groups but also the association between specific mutations such as t(8;21) with *KIT*, and biallelic *CEBPA* and *GATA2* mutations.

1.2.2 Leukaemic stem cell

It has been shown that AML occurs in a hierarchical structure with a population of self-renewing leukaemic cells at the top of the hierarchy that maintains the disease long-term. These cells are termed the leukaemic stem cells (LSCs) which are capable of initiating and sustaining the disease and give rise to the non-LSC blasts that do not have the capability of self-renewal (Thomas & Majeti, 2017). From transplantation of AML patient samples into NOD/SCID mice it was originally thought that LSCs resided within the CD34⁺CD38⁻ subpopulation of leukaemic cells (Bonnet & Dick, 1997). As the LSCs were immunophenotypically similar to HSCs it was initially proposed that the LSCs arise from the HSCs. However, further studies using more permissible models such as NOD/SCID/IL2R- γ null mice showed LSC activity in immunophenotypic progenitor populations, such as the LMPPs and granulocyte-monocyte progenitors (GMPs), but also in CD34⁻ granulocyte-monocyte precursors downstream of the progenitors. In keeping with their immunophenotype the LSC populations also resembled their normal counterpart in gene expression compared to HSCs, however multiple genes normally expressed in HSCs were also expressed (Goardon et al., 2011). This suggests the LSCs can arise from progenitors that acquire self-renewal properties. As self-renewal is not a normal progenitor property, the acquisition of self-renewal in AML LSCs is a result of genetic, and possibly epigenetic, changes.

Mouse model supports this idea as MLL-AF9-transformed GMPs, and committed myeloid progenitors containing the knock-in allele of the N-terminal mutation in *Cebpa* have also demonstrated that LSCs can reside within the GMPs. Leukaemic GMP gene expression was similar to their normal counterpart but with an activation of stem and self-renewal associated gene signatures (Kirstetter et al., 2008; Krivtsov et al., 2006).

1.2.3 Clonal evolution in AML

AML has been shown to be a highly complex and dynamic disease characterised by multiple somatic mutations with competing clones that evolve over time (Dohner et al., 2017). The clonal evolution and order of mutations has been studied by sequencing AML samples. Using deep sequencing methods of mixed tumour cell populations, the variant allele frequency (VAF) was used to infer the clonal evolution, for example if the VAF of a heterozygous mutation was ~50% of a bulk tumour sample then the mutation must be present in all the cells, called the founding clone. A VAF of ~20% would suggest the mutation was found in a subclone that must contain the mutations found in the founding clone (Welch et al., 2012). Analysis of sequential AML samples from patients that relapse or develop secondary AML also revealed the order of mutations. Mutations such as *STAG2*, *U2AF1*, and *DNMT3A* were found in primary samples and persisted at relapse or disease progression, indicating they were an initiating mutation within the founding clone. Mutations in *PTPN11*, *FLT3* and *KRAS* were only present in the serial sample and therefore a secondary mutation (Ding et al., 2012; Pellagatti et al., 2016; Walter et al., 2012). Finally, the sequencing of purified HSC and progenitor populations from AML patients revealed that balanced translocations and mutations in epigenetic regulators, such as AML1-ETO, *DNMT3A* and *TET2*, were frequently acquired within the HSC compartment as early events in disease evolution, generating pre-leukaemic stem cells which have a clonal advantage over non-mutated HSCs. The pre-leukaemic stem cells are non-malignant as they were still able to generate non-malignant lineages such as the B and T cells. Mutations found in pre-leukaemic HSCs may also persist after therapy, and lead to clonal expansion during remission to cause relapse (Dohner et al., 2017).

Recently, studies of elderly health individuals has identified that recurrent pre-leukaemic mutations, such as *DNMT3A*, *ASXL1*, and *TET2*, as well as splicing factor genes, such as *SF3B1* and *SRSF2*, were associated with clonal haematopoietic expansion, termed clonal

haematopoiesis of indeterminant potential, and correlated with an increased risk of developing a haematological neoplasm (Busque et al., 2012; Jaiswal et al., 2014; Xie et al., 2014).

Activated signalling mutations of genes such as *FLT3*, *KIT* and *KRAS* occur as late events that were detected in the transformed leukaemic progenitors and rarely detected in the pre-leukaemic HSC compartment (Corces-Zimmerman & Majeti, 2014; Jan et al., 2012; Miyamoto, Weissman, & Akashi, 2000; Shlush et al., 2014).

Collectively these studies show that AML acquire mutation hierarchy in two ways; temporally as well as in cell type in the haemopoietic hierarchy. Mutations such as epigenetic regulators and balanced translocations are acquired as the initial hit within the HSCs establishing the pre-leukaemic stem cell that harbours a selective growth advantage. Subsequently, additional mutations in activated signalling genes such as *KRAS* and *FLT3* are acquired with the pre-leukaemic progenitor cells which lead to leukaemic transformation within the progenitor cell and establishment of the LSC. The LSCs are then capable of maintaining the disease due to their acquired self-renewal capacity, and produce the leukaemic blasts (Figure 1.4).

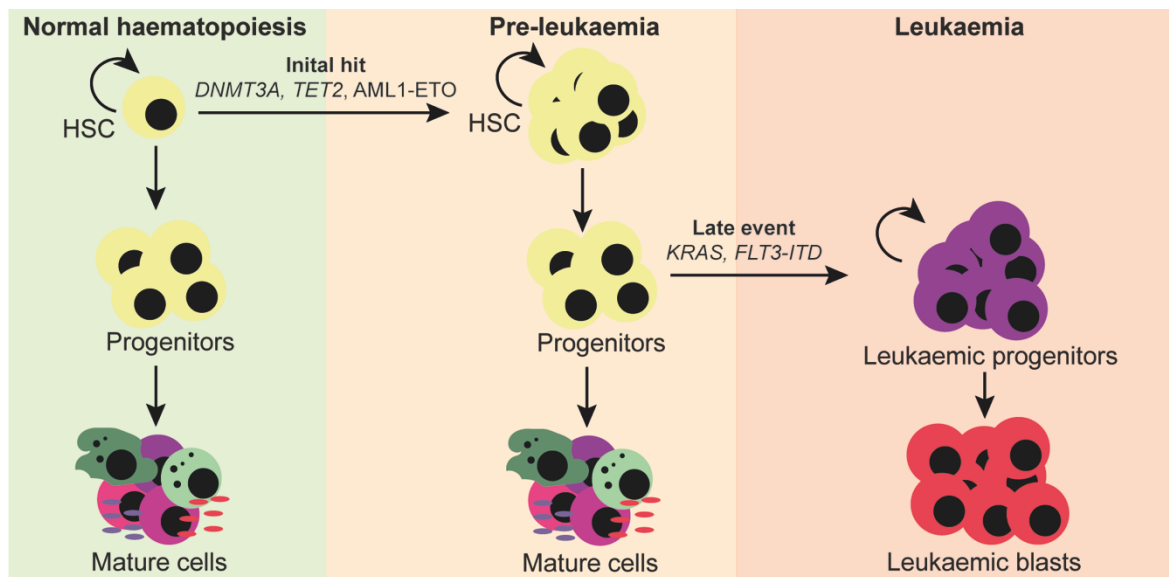


Figure 1.4 Hierarchy of mutation acquisition in AML

Schematic depicting the hierarchical and temporal acquisition of mutations. An initial mutation in genes involved in epigenetic regulation such as *DNMT3A* and *TET2* as well as balanced translocations such as AML1-ETO are acquired in the HSC. This leads to the establishment of the pre-leukaemic stem cell which still has the capacity to produce all the mature cells but has a clonal advantage over non-leukaemic HSCs. Late events in genes involved in signalling such *KRAS* and *FLT3-ITD* are acquired in the pre-leukaemic progenitors leading to the generation of the leukaemic progenitors with an acquired self-renewal capacity, which can then produce the leukaemic blasts.

1.2.4 Role of AML1-ETO as a pre-leukaemic mutation in AML

The t(8;21) translocation results from the fusion between AML1 (also known as RUNX1) and ETO (also known as CBFA2T1 and MTG8). The chimeric protein consists of the intact runt-homology domain (RHD) of AML1 fused with almost the entire coding region of ETO (Ahn et al., 1998) (Figure 1.5).

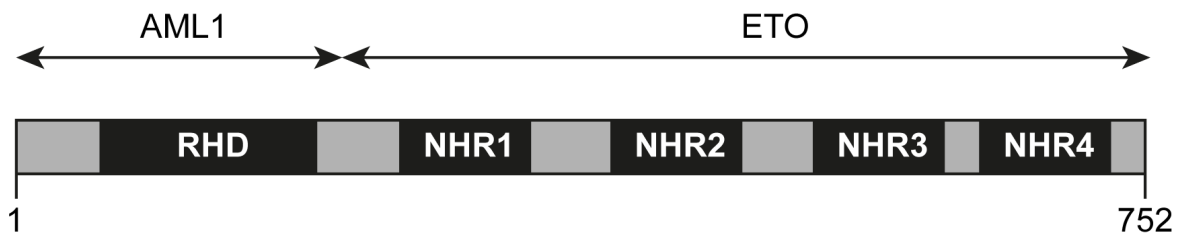


Figure 1.5 AML1-ETO domains

RHD, runt-homology domain; NHR1-4, nervy homology regions.

AML1 is a member of the core binding factor family of transcription factors and plays a critical role in regulating haematopoiesis. AML1 contains a runt-homology domain that mediates DNA binding, as well as heterodimerisation with the core binding factor β subunit, which facilitates efficient DNA binding (Ogawa et al., 1993).

Aml1 knock-out studies have shown that Aml1 is indispensable for the establishment of definitive haematopoiesis, as *Aml1*^{-/-} embryos died around E12.5 with lack of foetal liver (FL) haematopoiesis (Okuda, van Deursen, Hiebert, Grosveld, & Downing, 1996). Chen *et al.*, used conditional deletion of Aml1 using vascular endothelial cadherin (*Vec*)-Cre and *Vav*-Cre to further determine the role of Aml1. They found conditional knock-out of Aml1 using *Vec*-Cre yielded similar results to Aml1 null and also found a decrease in c-Kit cells suggesting Aml1 is required for HSC formation. In contrast conditional knock-out using *Vav*-Cre showed no foetal lethality indicating Aml1 is not required to maintain the HSCs once or before *Vav* is expressed (M. J. Chen, Yokomizo, Zeigler, Dzierzak, & Speck, 2009).

Conditional knock-out using *Mx1*-Cre demonstrated that Aml1 is important for maturation of megakaryocytes, B and T cells as the development of these cells were impaired, whereas myeloid and erythroid differentiation were maintained (Ichikawa et al., 2004).

There has been some contradictions in the literature over the effect of loss of Aml1 in HSCs. Ichikawa *et al.*, observed an increase LSKCD34⁺ cells and an increase in colony replating compared to wild type (WT), however in competitive transplantation there was no significant difference between WT and *Aml1*^{-/-} granulocytes and monocytes over time (Ichikawa *et al.*, 2004). Jacob *et al.*, found an initial increase in LSKCD34⁺Flt3⁺ cells in 10 week old mice, however HSCs declined with age, and limited dilution assays indicated lower frequency of competitive repopulating units in *Aml1*^{-/-} compared to WT (Jacob *et al.*, 2010). Cai *et al.*, attempted to resolve this issue and found Aml1 loss had minimal effect on LT-HSCs with either a moderate decrease in LT-HSCs when contribution to peripheral blood (PB) was scored, or no significant difference when contribution to bone marrow (BM) was assessed (Cai *et al.*, 2011).

Together, these studies show AML1 is important for definitive haematopoiesis, and maturation of megakaryocytes, B and T cells in adult BM with minimal effects on LT-HSCs.

AML1 mutations have been found in both familial and sporadic haematological malignancies. Germline mutations have been found in familial platelet disorder which is a autosomal dominant disorder, characterised by thrombocytopenia and a predisposition to developing AML, caused by loss-of-function mutation in *AML1* (Michaud *et al.*, 2002). Several studies have reported somatic mutations in *AML1* in patients with primary myelodysplastic syndrome (MDS), therapy-related MDS, and MDS that progressed to AML (C. Y. Chen *et al.*, 2007; Christiansen, Andersen, & Pedersen-Bjergaard, 2004; S. C. Tsai *et al.*, 2015). *AML1* is one of the most frequent mutations in MDS occurring in approximately 10% of cases (Osato, 2004). Mutations in *AML1* occur frequently in cytogenetically normal AML occurring between 13-33% with a broad spectrum of missense, nonsense and frameshift changes distributed throughout the gene, with high

frequency in the RHD (Greif, Konstandin, et al., 2012; Schnittger et al., 2011; Tang et al., 2009).

The function of the fusion partner ETO in t(8;21) AML patients is not well understood. Homozygous mutant of Eto in mice results in gastrointestinal defects but no haematopoietic defects (Calabi, Pannell, & Pavloska, 2001). ETO contains nervy homology regions (NHR) 1-4 (Davis, McGhee, & Meyers, 2003). Our understanding of the function of ETO is mainly based on studying the activity of ETO in the context of the AML1-ETO fusion. NHR1 is homologous to several TATA binding protein-associated factors and has been shown to inhibit transcriptional activation by E proteins, by preventing recruitment of p300/CREB-binding protein coactivators (J. Zhang, Kalkum, Yamamura, Chait, & Roeder, 2004). NHR2 interacts with ETO family members and is involved in oligomerisation. Oligomerisation is an important process in leukaemia development as it contributes to the inhibition of AML1-ETO in granulocyte differentiation and enhances the clonogenicity of murine BM cells. NHR2 also interacts with corepressors mSin3A, histone deacetylases (HDACs) HDAC1 and HDAC2 (Liu et al., 2006). NHR3 is a α -helical domain (H. T. Yang et al., 2004) and has been shown to interact with the regulatory subunit of type II cyclin AML-dependent protein kinase (Fukuyama et al., 2001). NHR4 is also known as a myeloid-Nery-DEAF-1 domain, and has two putative non-DNA binding zinc fingers and mediates protein-protein interaction with corepressors, including N-Cor, SMRT and mSin3A, which are able to recruit HDACs to mediate transcriptional repression (Gelmetti et al., 1998; Lutterbach, Sun, Schuetz, & Hiebert, 1998). Together, the data suggests that ETO, in AML1-ETO, functions as a transcriptional corepressor. However it has also been shown that p300 binds to NHR1 to facilitate transcription (L. Wang et al., 2011).

As the AML1 DNA binding domain remains intact in AML1-ETO, this allows AML1-ETO to bind to AML1 target gene promoters. As ETO recruits corepressors to AML1 target genes,

this causes down-regulation of AML1 targets such as MDR-1, GM-CSF receptor, IL3, c-fos, TCR β enhancer, immunoglobulin α promotor, and p14^{ARF} (Peterson & Zhang, 2004).

AML1-ETO occurs in approximately 7% of adult AML patients (Dohner et al., 2017; Grimwade et al., 2010). Several lines of evidence suggest that *AML1-ETO* is acquired in pre-leukaemic HSCs. First, *AML1-ETO* mRNA could still be detected in AML patients that had been in clinical remission for up to 150 months (Miyamoto et al., 1996). Further evidence was seen in children, and twin studies, where *AML1-ETO* was detected *in utero*, and the children did not develop leukaemic until later in life, in some cases occurring more than 10 years after birth (Wiemels et al., 2002). *AML1-ETO* also remains stable in patients that relapsed, while additional mutations were highly dynamic with mutations both gained and lost at relapse (Krauth et al., 2014).

Constitutive knock-in *Aml1*^{ETO/+} mice largely replicated homozygous knock-out of Aml1 as Aml1-ETO was embryonically lethal with complete absence of definitive haematopoiesis and lethal haemorrhages. This supporting data indicates that Aml1-ETO represses Aml1 transcriptional activity. In contrast to Aml1 knock-out, Aml1-ETO FLs contained abnormal multi-lineage haematopoietic progenitors that despite the inability to expand *in vivo* had an enhanced self-renewal capacity *in vitro* (Okuda et al., 1998; Yergeau et al., 1997).

To avoid embryonic lethality of Aml1-ETO, knock-in inducible transgenic models of Aml1-ETO were developed. Rhoades *et al.*, used a tetracycline-inducible system and showed that in the absence of tetracycline Aml1-ETO was expressed and mice lived into adulthood, however mice did not develop a leukaemia. Mice had no abnormal haematopoiesis, however when cells were serially replated *in vitro*, Aml1-ETO expressing cells continued to replate and generate early progenitor and early myeloid cell colonies (Rhoades et al., 2000). Yuan *et al.*, controlled the expression of Aml1-ETO using the human *MRP8* promotor, which allows Aml1-ETO to be specifically expressed in myeloid

cells. Mice had no abnormal haematopoiesis and only when treated with the DNA-alkylating agent *N*-ethyl-*N*-nitrosourea (ENU) did mice develop a leukaemia, indicating that additional mutations are required with Aml1-ETO for leukaemic transformation (Yuan et al., 2001).

Retroviral expression of Aml1-ETO in LSK murine BM cells and transplantation into lethally irradiated mice showed an increase in primitive myeloblasts, however only by 10%, and therefore not considered a leukaemia. An increase in the LSK population was also observed (de Guzman et al., 2002).

Using a conditional knock-in model of Aml1-ETO using *Mx1*-Cre found a striking increase in serial replating ability *in vitro*, with no abnormal haematopoiesis. Only when treated with ENU did mice develop a leukaemia (Higuchi et al., 2002).

In summary, AML1-ETO is a pre-leukaemic mutation acquired in the HSC which confers an enhanced self-renewal capacity. The fusion of ETO to the AML1 DNA binding domain allow corepressors to be recruited and therefore down-regulates AML1 target gene expression. AML1-ETO alone is insufficient to cause leukaemia and requires additional mutations for leukaemic transformation.

Sequencing studies of t(8;21) patients has shown high frequency of mutations in *KIT* (25%), *NRAS* (~20%), *ZBTB7A* (~20%), *ASXL2* (~20%), *ASXL1* (10%) and cohesion mutations (total ~20%) including; *RAD21* (~10%), *SMC1A* (~5%) and *SMC3* (~5%) (Dohner et al., 2017; Grimwade et al., 2010). Other co-occurring mutations that occur less frequently include *EZH2* (~5%), *KDM6A* (~5%), *MGA* (5%), *DHX15* (~5%), *FLT3-ITD* (5%), *KRAS* (4.3%), *FLT3-TKD* (4.3%), *CBL* (4.3%), *IDH2* (3.6%), *JAK2* (2.9%) and *IDH1* (0.7%) (Dohner et al., 2017; Grimwade et al., 2010; Krauth et al., 2014).

1.2.5 Role of biallelic *CEBPA* mutations as pre-leukaemic mutations in AML

C/EBP α is a transcription factor important for granulopoiesis. In the haemopoietic system C/EBP α is mainly expressed in myeloid lineage cells but is also expressed in <10% of LT-HSCs (Wolfler et al., 2010). It is an intronless gene that translates into two isoforms using alternative start sites producing either the full-length C/EBP α (42kDa; p42) or the shorter N-terminal truncated isoform (30KDa; p30). P30 lacks the transactivation domain 1 (TAD1) that regulates transcriptional activation (Nerlov & Ziff, 1995) and has been implicated in C/EBP α -mediated repression of E2F that controls cell cycle progression and represses proliferation (Slomiany, D'Arigo, Kelly, & Kurtz, 2000). As p30 lacks TAD1 it therefore cannot form a contact with proteins involved in transcription (Pabst & Mueller, 2009). C-terminus of C/EBP α mediates dimerization which is required for DNA binding (Friedman, 2015).

The ratio of p30 to p42 is regulated by extrinsic factors. When there are abundant nutrients and growth factors the eukaryotic translation initiation factors eIF2a/eIF3E are activated, which increases the p30/p42 ratio. Increased p30 promotes proliferation in favourable conditions (Calkhoven, Muller, & Leutz, 2000).

Knock-out of *Cebpa* is perinatally lethal and neonates die shortly after birth due to defects in liver development caused by abnormal hepatic glycogenesis that leads to an imbalance of energy homeostasis (N. D. Wang et al., 1995). Analysis of *Cebpa* knock-out newborn mice or the FL haematopoietic system revealed normal levels of haemoglobin, platelets, monocytes and lymphocytes but no neutrophils and eosinophils. There was also a loss of the granulocyte colony stimulating factor receptor, which is a *Cebpa* target. As a consequence myeloid progenitors from *Cebpa* null FLs did not respond to granulocyte colony stimulating factor *in vivo*. The inability to produce granulocytes was a cell intrinsic

effect as transplantation of *Cebpa* knock-out cells could reconstitute lymphoid cells but not neutrophilic cells (D. E. Zhang et al., 1997).

Using both FL *Cebpa* knock-out and conditional *Cebpa* knock-out in adult BM, Zhang *et al.*, showed that loss of *Cebpa* caused a block in early granulopoiesis preventing the formation of the GMPs. This caused an accumulation of immature myeloid blasts, however no AML developed (P. Zhang et al., 2004).

There has been some controversy in the literature regarding the impact of loss of *Cebpa* on the HSCs. Zhang *et al.*, and Ye *et al.*, report that loss of *Cebpa* leads to an enhanced competitive advantage in competitive transplantation and an increase in HSC number both phenotypically and in limited dilution assays (Ye et al., 2013; P. Zhang et al., 2004). This contrasts Hasemann *et al.*, results that demonstrated reduced phenotypic and functional HSCs (Hasemann et al., 2014). The results may differ due to analysis of conditional knock-out mice at different time points after induction using polyinosinic:polycytidylic acid [poly(I:C)]. Analyse of mice at an early time point after poly(I:C) in Zhang *et al.*, and Ye *et al.*, studies may mean HSCs were actively cycling due to interferon signalling (Essers et al., 2009). Timing of poly(I:C) also differed between the studies. In Ye *et al.*, study donor cells were transplanted into recipients followed by poly(I:C) induction, whereas in Hasemann *et al.*, study donor cells first underwent poly(I:C) deletion before transplantation. Loss of *Cebpa* before transplantation may have functionally compromised the HSCs during engraftment resulting in lower reconstitution (Pulikkan, Tenen, & Behre, 2017). All studies however report a loss of HSCs in G0 and an increase of HSCs in G1 and S/G2M. Studies of the deletion of the myeloid-specific (+37) *Cebpa* enhancer showed a loss of HSCs (Avellino et al., 2016; Guo, Cooper, & Friedman, 2016). Therefore it is likely that reduced expression of *Cebpa* is detrimental to HSC frequency.

Homodimer of C/EBP α is required for granulopoiesis while C/EBP α heterodimer with AP-1 proteins, such as c-Jun and c-Fos, in collaboration with PU.1, leads to activation of monocyte specific genes (Friedman, 2015). C/EBP α primes and activates myeloid gene expression through forming a complex with PU.1, ERG, c-MYB and RUNX1, and binds to promoters or enhancers of myeloid genes such as *CSF3R*, *IL-6R*, *CEBPE*, *GFI1*, *KLF5*, *MPO*, and *ELA2* in immature or mature granulocytes and immature monocytes (Federzoni et al., 2014; Friedman, 2002, 2007; Ma, Hong, Guo, Ghiaur, & Friedman, 2014; P. Zhang et al., 1998). C/EBP α also drives myeloid differentiation by repressing non-myeloid genes. Expression of *Cebpa* in non-myeloid cells converted lymphoid phenotype of B and T cell progenitors into myeloid progenitors that differentiated into monocytes and granulocytes (Di Tullio et al., 2011; Laiosa, Stadtfeld, Xie, de Andres-Aguayo, & Graf, 2006). *Cebpa* knock-out studies have also shown T cell genes such as *Cd7* and *Lck* are up-regulated, and when *Cebpa* expression is re-introduced, T cell genes were attenuated and drove myeloid differentiation (Wouters et al., 2007).

There are two classes of *CEBPA* mutations in AML; C-terminal mutations that affect the DNA-binding and/or dimerisation (Gombart et al., 2002), and N-terminal mutations that are located between the start codon for the p42 and p30 isoform; which leads to exclusive expression of the p30 isoform from the mutated allele (Ohlsson, Schuster, Hasemann, & Porse, 2016).

Monoallelic *CEBPA* mutations occur in 30% of patients and are heterogenous in additional mutation, gene expression, and prognosis. The other 70% of patients have biallelic *CEBPA* mutations, which is a distinct subtype of AML recognised in the WHO classification of AML (Arber et al., 2016), associated with a good prognosis.

Biallelic *CEBPA* mutation occurs in ~4% of AML patients (Dohner et al., 2017; Grimwade et al., 2010) and usually involves an N-terminal mutation in one allele and a C-terminal mutation on the other allele. This causes only WT p30 to be expressed on the allele with the N-terminal mutation causing a dominant negative effect. The C-terminal mutation leads to the expression of p30 and p42 with defective DNA-binding site. Therefore only WT p30/p30 homodimer is produced (Figure 1.6). The other types of biallelic *CEBPA* mutations are very rare, such as two N-terminal mutations, or an N-terminal mutation with a frameshift/nonsense mutation in the central part of *CEBPA*. Both mutations result in only p30/p30 homodimer. The other rare form of biallelic *CEBPA* mutations are two C-terminal mutations (Avellino & Delwel, 2017; Fasan et al., 2014).

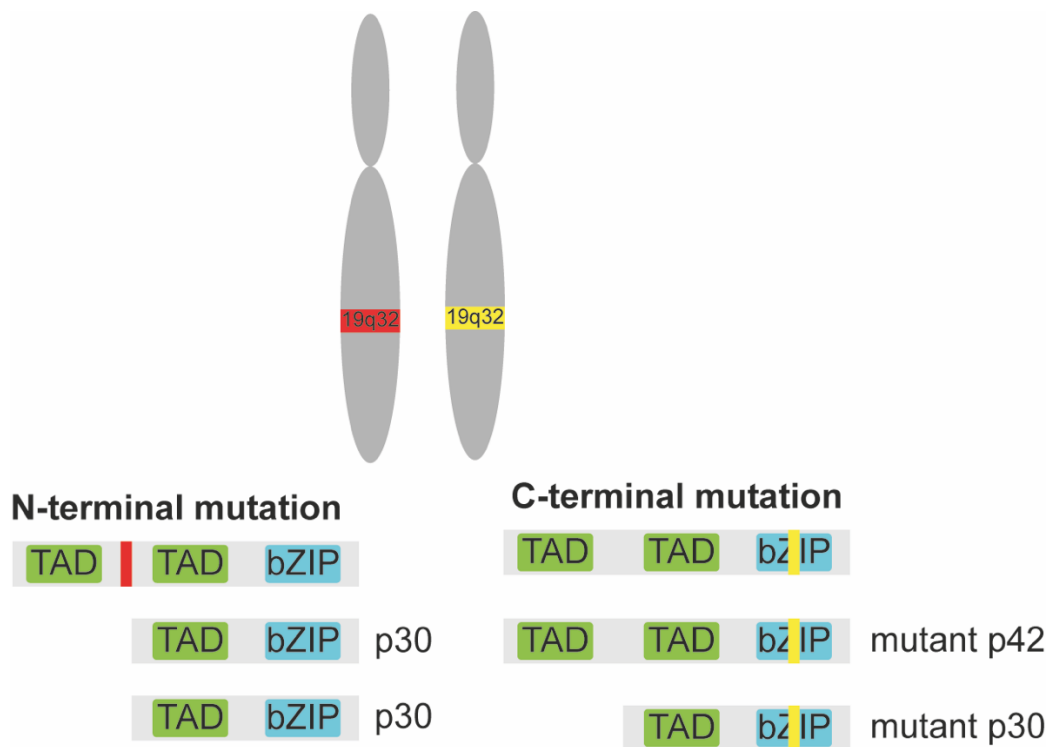


Figure 1.6 Biallelic *CEBPA* mutations

Schematic of the most common biallelic *CEBPA* mutations. N-terminal mutation prevents p42 expression, only allowing p30 isoform to be expressed. C-terminal mutation in the basic leucine zipper domain (bZIP) DNA-binding domain means mutant p42 and mutant p30 contain a defective DNA-binding domain. Therefore only WT p30 is produced.

Adapted from (Avellino & Delwel, 2017).

The biallelic *CEBPA* mutations occur in sporadic and familial AML. Germline mutations have been found in families with autosomal dominant inheritance of AML and eosinophilia. In families carrying the *CEBPA* mutations the germline mutation is always an N-terminal mutation while the C-terminal mutation is an acquired mutation suggesting the N-terminal mutation is the pre-leukaemic mutation (Carmichael et al., 2010).

The *Cebpa* +42 enhancer in mice (+37 in human) is a known target for AML1-ETO, CBFβ-MYH11, and EVI1 in EVI1-transformed leukaemias. *CEBPA* is often down-regulated in these AMLs therefore binding to the *CEBPA* enhancer may result in down-

regulation of *CEBPA* (Avellino & Delwel, 2017). However, the *Cebpa* knock-out mouse model did not develop a leukaemia suggesting some level of C/EBP α is required for leukaemia development (P. Zhang et al., 2004). Interestingly, Ye *et al.*, demonstrated when *Cebpa* is knocked out of MLL-AF9 transduced cells, AML does not develop due to a block in granulopoiesis. However, if IL-3 and GM-CSF were injected into the mice granulocyte differentiation is restored and AML transformation is re-established (Ye et al., 2015). This indicates a complete block in myeloid differentiation prevents AML transformation and sufficient GMPs formation is required for transformation.

The Nerlov group initially designed a knock-in mouse model of the N-terminal mutation by placing a LoxP-flanked neomycin resistance cassette site in the *Cebpa* locus between the p42 and p30 initiation codon. Upon Cre-mediated deletion of the cassette, the remaining LoxP site introduced a nonsense codon in the p42 isoform initiation site leading to only p30 isoform being expressed from the mutant allele, mice were termed *Cebpa*^{Lp30/+} or L/+. Homozygous N-terminal mutation, termed LL, caused a full penetrant AML within one year and the establishment of LSCs in the committed myeloid cells (Lin⁻Mac1⁺c-Kit⁺) that had an enhanced self-renewal *in vitro* and were capable of transplanting disease *in vivo* (Kirstetter et al., 2008). Gene expression analysis showed LSCs from LL mice had an overlapping gene signature with MLL-AF9-transformed GMPs (Krivtsov et al., 2006). This study suggests p30 allows sufficient GMPs to be produced and loss of p42 allows for self-renewal to occur in the committed myeloid cells.

The biallelic *Cebpa* mutant model was then established by developing a C-terminal mutant mouse by creating a knock-in allele mimicking a common C-terminal mutation K313KK, mice were termed *Cebpa*^{K313K/+} or K/+. N-terminal and C-terminal mutations were then combined, termed *Cebpa*^{K313K/Lp30} or KL. KL mice had a shorter latency compared to LL mice. Interestingly, homozygous C-terminal mutation, termed KK mice, had a longer latency compared to both KL and LL mice. This was due to the severe and

persistent block in the formation of committed myeloid progenitors, which is the cell population required for leukaemic transformation. The N-terminal and C-terminal mutations were shown to collaborate as C-terminal mutation was shown to cause pre-leukaemic HSC expansion, thereby increasing the number of committed myeloid progenitors, while the N-terminal mutation preserved the formation of GMP to allow LSCs to form in the committed myeloid cells (Bereshchenko et al., 2009). Together, this shows loss of p42 causes loss of proliferation control leading to cell expansion and the gain of self-renewal in the myeloid committed cells, while loss of DNA binding causes loss of granulopoiesis and loss of normal HSC function.

Biallelic *CEBPA* mutant AML has a distinct secondary mutational spectrum. There is an under-representation of mutations that are common in AML such as *ASXL1*, *RUNX1* and *TET2* while mutation in *GATA2* (~30%) and *WT1* (~20%) are enriched (Fasan et al., 2014). The association between biallelic *CEBPA* and *GATA2* is statistically significant, FWER < 0.05 (family-wise error rate) (Papaemmanuil et al., 2016). Other commonly co-occurring mutations include *NRAS* (~30%) and *CSF3R* (~20%) (Dohner et al., 2017; Grimwade et al., 2010).

1.2.6 Role of K-RAS mutations as secondary mutations in AML

RAS proteins are a family of guanine nucleotide-binding proteins involved in cell proliferation, differentiation and survival (Downward, 2003). The three members of RAS family are HRAS, KRAS and NRAS, all three are widely expressed however KRAS is expressed ubiquitously. This is reflected in mouse knock-out where *Hras* and *Nras* are not required for normal development, whereas *Kras* was shown to be essential as *Kras* mutant mice die between E12-14 with FL defects and anaemia (Johnson et al., 1997).

RAS proteins are signal-switch molecules that modulate cell fate by cycling between inactive GDP-bound and active GTP-bound confirmation. The process of cycling from inactive to active state is catalysed by nucleotide exchange via guanine nucleotide exchange factor (GEF). Nucleotide hydrolysis via GTPase activating protein (GAP) results in RAS signalling termination (Downward, 2003).

Active RAS can bind and activate several families of effector enzymes that, promoting cell proliferation. RAS activates RAF/MEK/ERK pathway which results in stimulating expression of the ETS family of transcription factors e.g. ELK1, as well as cell-cycle regulators such as D-type cyclins (Pruitt & Der, 2001). RAS also activates the type I phosphatidylinositol 3-HO kinase (PI(3)K) pathway which in turn activates AKT, which has an anti-apoptotic function and has a role in survival signalling (Khwaja, Rodriguez-Viciano, Wennstrom, Warne, & Downward, 1997). PI3K activation also leads to stimulating RAC which regulates actin cytoskeleton and activates NF- κ B to promote cell survival (Lambert et al., 2002). RAS also activates the RALGDS pathway which inhibits FORKHEAD transcription factors that are implicated in promoting cell-cycle arrest through induction of p27, and also promotes apoptosis through expression of BIM (De Ruiter, Burgering, & Bos, 2001).

RAS mutations occurs in approximately 30% of myeloid malignancies including AML, myeloproliferative neoplasm (MPN), MDS, chronic myelomonocytic leukaemia (CMML) and juvenile myelomonocytic leukaemia (JMML) (Van Etten & Shannon, 2004). Almost all the mutations are in codons 12, 13 and 61 (Bos, 1989). Substitution at residues G12 and G13 prevents the formation of van de Waals bonds between RAS and GAP perturbing the normal orientation of the catalytic glutamine Q61 in RAS; resulting in the attenuation of GTP hydrolysis (Scheffzek et al., 1997), persistence of GTP-bound RAS, and ultimately the constant downstream activation of RAS effector pathways (Pylayeva-Gupta, Grabocka, & Bar-Sagi, 2011).

RAS mutations frequently co-occur with t(8;21) (*NRAS* = 12.9%, *KRAS* = 4.3%) (Krauth Leukemia 2014). *RAS* mutations are thought to be secondary or late event mutations as the sequencing of AML patients samples that have achieved remission has shown *RAS* mutations are unstable and often lost at subsequent relapse, with the gain of a novel activated signalling mutation (e.g. *FLT3-ITD*) (Krauth Leukemia 2014).

The Braun group conducted a series of experiments using the conditional knock-in mouse model *Kras*^{G12D/+} using *Mx1-Cre*. They found *Kras* mutant mice developed a fully penetrant fatal MPN characterised by splenomegaly, leucocytosis and anaemia. There was no expansion of the immature blasts therefore no AML developed. *Kras*G12D-expressing cells were capable of forming colonies *in vitro* without exogenous cytokines and when exposed to GM-CSF and IL-3, *Kras*G12D-expressing cells were hypersensitive with a dramatic increase in proliferation (Braun et al., 2004). They found that Lineage^{c-Kit}⁺ (LK) *Kras*G12D-expressing cells resulted in an increase in phosphorylated STAT5, ERK and S6 under both basal conditions and under exposure to GM-CSF; but were attenuated in response to stem cell factor (SCF) (Van Meter et al., 2007). By sorting HSCs and CMPs from *Kras*^{G12D/+} mice, they were able to show that only the *Kras*G12D-expressing HSCs were able to confer the MPN phenotype. However, the *Kras* mutation had a negative impact on the HSCs, resulting in a reduced number of quiescent HSCs that also had higher cyclin D1 expression. Limited dilution assay also showed a 10-fold decrease in LT-HSCs. The self-renewal transformed population only arose when Notch1 cooperating mutation occurred within the thymus, which lead to mice developing a T-cell acute lymphoblastic leukaemia (T-ALL) (Sabnis et al., 2009). Therefore K-RasG12D alone is not able to transform the myeloid progenitor, and myeloproliferation is only sustained by the HSC.

Collectively, these studies suggest that *KRAS* mutations are acquired as late events within the progenitor cell and are undetectable in the HSC in AML. K-RasG12D causes

myeloproliferation that is propagated only by the HSCs and K-RasG12D does not result in leukaemia transformation of the progenitor cells.

1.2.7 Role of zinc finger 1 *GATA2* mutations as secondary mutations in AML with biallelic *CEBPA* mutations

GATA2 is a transcription factor important for both definitive and adult haematopoiesis.

Gata2 is expressed in the yolk sac, FL, placenta, HSCs, early erythroid cells, mast cells and megakaryocytes (Minegishi et al., 1999; Y. K. Ng, George, Engel, & Linzer, 1994; Orlic, Anderson, Biesecker, Sorrentino, & Bodine, 1995; Robert-Moreno, Espinosa, de la Pompa, & Bigas, 2005; F. Y. Tsai et al., 1994; F. Y. Tsai & Orkin, 1997).

Knock-out studies of *Gata2* showed that it is essential during embryogenesis as embryos did not survive past E11.5. Embryos were viable at E9.5 and contained normal primitive erythroid cells, however the total number of blood cells were reduced and embryos were pale despite containing primitive erythroid cells. Therefore *Gata2* null embryos were anaemic and failed to survive beyond the stage of primitive haematopoiesis (F. Y. Tsai et al., 1994). An *in vitro* embryonic stem (ES) cell differentiations assay demonstrated that *Gata2*-expression is required for expansion of multipotent haematopoietic progenitors and the formation of mast cells, but was dispensable for the terminal differentiation of erythroid cells and macrophages (F. Y. Tsai & Orkin, 1997).

De Pater *et al.*, further dissected the function of *Gata2* during embryogenesis using *Vec-Cre* and *Vav-Cre* to understand the effects of loss of *Gata2* both before HSC generation (using *Vec-Cre*), and after HSC generation (using *Vav-Cre*). Conditional knock-out of *Gata2* using *Vec-Cre* showed *Gata2* was important for the endothelial-to-haematopoietic

transition during generation of HSCs, and the use of *Vav-Cre* showed *Gata2* is required for HSC survival (de Pater et al., 2013).

Haploinsufficiency of *Gata2* in adult mice showed no overt abnormalities however there was a decrease in LSKCD34⁺ cells in the BM and a decrease in competitive transplantability (Ling et al., 2004; Rodrigues et al., 2005).

Over-expression studies have shown that HSCs become quiescent, reduce proliferation and block haematopoietic reconstitution (Persons et al., 1999). Therefore the balance of *Gata2* expression is important in HSC function.

Collectively, these studies demonstrate that *Gata2* is integral for generation of HSCs in the aorta-gonad-mesonephros and survival of HSCs into adulthood. *Gata2* is not required for terminal differentiation of erythroid or monocytes but is important for mast cell formation, and the balance of *Gata2* expression is important for HSC function.

GATA2 transcription is regulated by several genes such as *CEBPA*, *HOXA9*, *ETS1* (Lulli et al., 2006), *BMP4* (Maeno et al., 1996), *NOTCH1* (Kumano et al., 2001), *SFPI1* (Walsh et al., 2002), and *EVI1* (Yuasa et al., 2005), by cytokines such as IL-1 and TNF α (La Ferla, Reimann, Jelkmann, & Hellwig-Burgel, 2002), but also by *GATA2* itself, as well as *GATA1* (Vicente, Conchillo, Garcia-Sanchez, & Otero, 2012). *GATA2* is involved in a complex heptad of transcription factors that specify early lineage commitment including *FLI1*, *ERG*, *RUNX1*, *SCL*, *LMO2* and *LYL1*. The heptad is associated with genes that are highly expressed in haematopoietic stem and progenitor cells (HSPCs), but was also shown to prime genes for subsequent expression in lineage cells (Beck et al., 2013).

GATA2 and *GATA1* binding sites overlap to regulate target gene expression and the balance between proliferation and differentiation (Dore, Chlon, Brown, White, & Crispino,

2012). Chromatin occupancy exchange between GATA2 and GATA1 occurs during differentiation in approximately one third of GATA1 bound sites, termed the GATA switch, and which likely lead to activation or repression of gene expression. In general GATA1 and GATA2 act opposingly in switch target genes, including GATA2 and KIT which are both repressed during erythroid differentiation (Grass et al., 2003; Jing et al., 2008).

Disruption in the balance of GATA2 through alteration of its protein structure or activity can contribute to leukaemia development. Gene expression analysis of AML patient samples has shown *GATA2* was over-expressed and correlated with both *FLT3-ITD* and *NPM1* mutations, as well as *WT1* or *EVI1* over-expression (Vicente, Vazquez, et al., 2012). *GATA2* has also been shown to be over-expressed in *de novo* paediatric AML (Luesink et al., 2012), however other studies have instead shown that *GATA2* was down-regulated in AML patients (Bonadies et al., 2011). Interestingly, in AML with inv(3)/t(3;3), it has been shown that upon chr 3q rearrangement, a *GATA2* distal enhancer ectopically activates *EVI1* expression and removes the enhancer from its normal genomic location, leading to functional haploinsufficiency of *GATA2* (Groschel et al., 2014). Together, these studies indicate *GATA2* can be both up-regulated and down-regulated in AML.

Sequencing studies have recently identified *GATA2* mutations in haematological neoplasms. Germline mutations in *GATA2* lead to autosomal dominant inheritance of haematological malignancies pre-disposing to MDS and/or AML (Hahn et al., 2011). Recently, germline mutations in *GATA2* were also reported in paediatric MDS patients (Wlodarski et al., 2016). Patients with *GATA2* germline mutations can develop a range of symptoms including Emberger syndrome, which is characterised by lymphedema and a pre-disposition to MDS/AML (Kazenwadel et al., 2012; Mansour et al., 2010; Ostergaard et al., 2011), immunodeficiency disorders such as monocytopenia and mycobacterial infection (MonoMac) syndrome (Hsu et al., 2011; Vinh et al., 2010), dendritic cell, monocyte, B and natural killer (NK) lymphoid (DCML) deficiency syndrome (Bigley et al.,

2011; Dickinson et al., 2011), chronic neutropenia (Pasquet et al., 2013), NK cell deficiency (Mace et al., 2013) and aplastic anaemia (Ganapathi et al., 2015); all of which have a pre-disposition to developing MDS/AML. Germline mutations are scattered throughout the gene however the following are the three major types: truncating mutations found either before or within zinc finger (ZF) 2 which account for ~60% of cases, missense mutation within ZF2 (~30%), and non-coding variants within intron 4 which affects the +9.5kb regulatory site that contains E-box, GATA, and ETS elements (~10%) (Wlodarski, Collin, & Horwitz, 2017) (Figure 1.7). Germline mutations in these diseases seem to cause a loss-of-function of GATA2 and GATA2 haploinsufficiency (Hsu et al., 2011).

Somatic *GATA2* mutations are found in sporadic myeloid leukaemias such as chronic myeloid leukaemia (CML) in blast crisis (S. J. Zhang et al., 2008), monocytic AML (Yan et al., 2011), and cytogenetically normal AML with mono- or biallelic *CEBPA* mutations (Fasan et al., 2013; Green et al., 2013; Greif, Dufour, et al., 2012; Marceau-Renaut et al., 2015). The association between biallelic *CEBPA* and *GATA2* mutations has also been observed in the acute erythroid leukaemia (AEL) subtype called erythroleukaemia (Ping et al., 2017). Somatic mutations are distinct from germline mutations as the majority occur within the ZF1. Somatic mutations outside of ZF1 include L359V, which are only seen in CML with blast crisis (S. J. Zhang et al., 2008), and R362Q which are seen in AML (Fasan et al., 2013; Green et al., 2013; Greif, Dufour, et al., 2012; Marceau-Renaut et al., 2015) (Figure 1.7).

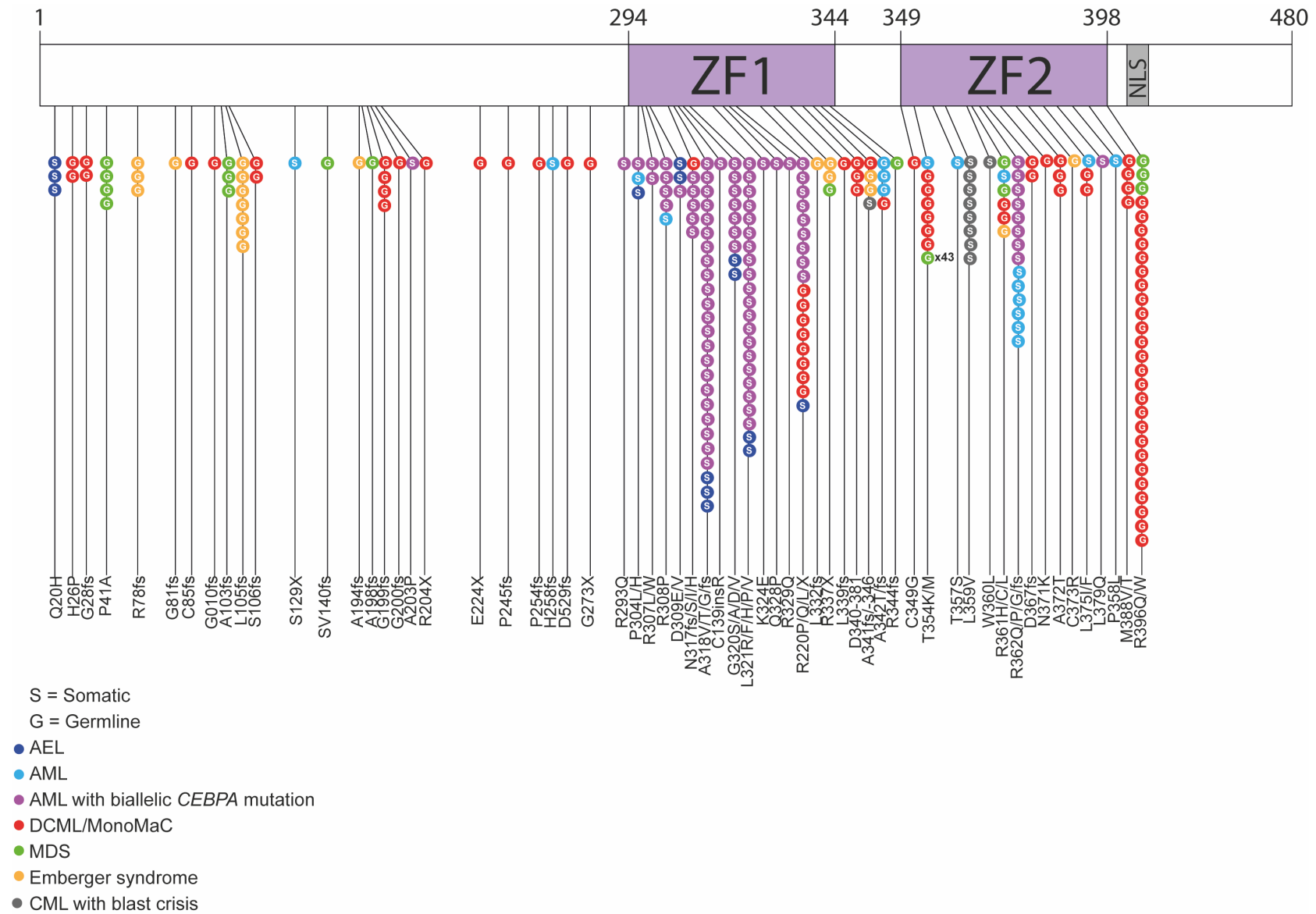


Figure 1.7 *GATA2* mutations identified in haematological malignancies and disorders

Schematic depicting a summary of *GATA2* mutations identified in haematological malignancies and disorders. Coloured dots indicate the haematological malignancy or disorder associated with the mutation location. Germline mutations in T354 were found in 43 MDS patient samples. S, somatic mutation; G, germline mutation; ZF1, zinc finger 1; ZF2, zinc finger 2; NLS, nuclear localisation signal (Collin, Dickinson, & Bigley, 2015; Marceau-Renaut et al., 2015; Ping et al., 2017; Wlodarski et al., 2016).

Sequencing remission and relapse samples of patients with biallelic *CEBPA* and *GATA2* mutations at diagnosis showed one patient retained the *GATA2* mutation at relapse, while in the remaining three patients the *GATA2* mutation was absent at relapse (Fasan et al., 2013). Analysis of patients with familial *CEBPA* mutation also revealed the initial germline heterozygous mutation in the N-terminus of *CEBPA*, and progression to AML was associated with the acquisition of the C-terminal mutation on the remaining *CEBPA* allele as well as a *GATA2* mutation (Green et al., 2013). Together, these studies suggest *GATA2* mutations are secondary to the biallelic *CEBPA* mutations.

Functional studies have mainly involved reporter gene assays in HEK 293T cells and have been performed on P304H, A318T, G320D, L321F, L321P, and R330X which are found in ZF1; and A350_N351ins8, T354M, T355del, T358N, C373R, R361L, and R396Q which are found in ZF2 (Cortes-Lavaud et al., 2015; Greif, Dufour, et al., 2012; Hahn et al., 2011; Niimi et al., 2013; Ostergaard et al., 2011; Ping et al., 2017). These studies have shown the mutations cause loss-of-function with reduced DNA binding ability and/or transactivation activity. The point mutation G320D in ZF1 was shown to enhance transactivation activity of *GATA2*, however with a reduced capacity to enhance *CEBPA* dependent activation of transcription (Greif, Dufour, et al., 2012). Conversely, L359V

found in CML blast crisis has been described as a gain-of-function mutation, given that L359V has an enhanced transactivation activity and has enhanced inhibitory effects on PU.1 activity (Greif, Dufour, et al., 2012; S. J. Zhang et al., 2008).

In summary from the current functional studies, the majority of *GATA2* mutations in both ZF1 and ZF2 lead to a loss-of-function mutation and haploinsufficiency, however ZF1 is involved in protein-protein interaction while ZF2 controls transcription through protein-DNA binding (Hirabayashi, Wlodarski, Kozyra, & Niemeyer, 2017). Somatic ZF1 mutations are mainly found in AML patients with pre-leukaemic biallelic *CEBPA* mutation and are likely to be acquired as secondary mutations, while ZF2 germline mutations would act as the initiating mutation to cause pre-disposition to MDS/AML. It is therefore possible that the mutations in the two ZFs have different effects on *GATA2* function.

1.3 Modelling and analysis of co-occurring mutations in AML

Model systems provide a platform to investigate the genetic basis of AML by identifying or validating cooperating genetic events that can lead to leukaemogenesis. They play an important role in the understanding of AML biology, demonstrating how leukaemia transformation occurs, identifying LSC populations, and drug screening. These could result in better understanding of leukaemia development and the discovery of novel druggable targets and biomarkers that could be used to specifically target LSC or pre-leukaemic stem cells.

As certain functional groups of mutations have been shown to co-occur and to be mutually exclusive, it suggests that common pathways are involved for leukaemia development. Accurate modelling of commonly co-occurring mutations could identify important dysregulated pathways that are also important in other mutation combinations.

Model techniques include lentiviral and retroviral transduction and transgenic mouse models. These techniques have the ability to perform controlled experiments with defined genotypes. Further, the analysis of patient samples containing specific mutations of interest can offer a useful way to understand AML biology. Xenotransplantation can be used to test patient samples functionally and identify LSC populations. Using a combination of DNA and RNA sequencing of patient samples elucidates the genetic and transcriptional basis of AML development, however the genetic and biological heterogeneity between patients can make analysing patient samples challenging.

1.3.1 Lentiviral and retroviral transduction

Lentiviral and retroviral transduction of both murine and human cells have been previously used to study mutation combinations. Numerous studies have looked at the interaction of AML1-ETO with co-occurring mutations.

AML1-ETO has been retrovirally co-transduced with KRASG12D, FLT3-length mutation and C-KIT mutation in murine BM cells (Schessl et al., 2005; Y. Y. Wang et al., 2011; Zhao et al., 2014). In all models when murine cells are co-transduced and transplanted into mice they result in a more aggressive leukaemia compared to single-transduced mutant recipients. All co-transduced recipients however had a long latency (C-KITN882K = 177.5 days, KRASG12D = 115 days, FLT3-length mutation = 233 days). AML1-ETO co-transduced with C-KITN882K and KRASG12D both developed a myeloid disease, however some mice with AML1-ETO combined with FLT3-length mutation developed a myeloblastic leukaemia while the remaining mice developed a lymphoblastic leukaemia. Transducing bulk BM cells in AML1-ETO with FLT3-length mutation may indicate that lymphoid committed stage cells were hit by the retrovirus, and combined with the long latency this could suggest that additional secondary mutations were acquired *in vivo*, resulting in the lymphoblastic leukaemias (Schessl et al., 2005).

Micol *et al.*, transduced *Asx12* null murine BM cells with AML1-ETO and showed mice develop a leukaemia, although with a long latency, as mice become moribund between ~150-300 days post-transplantation. They also monitored conditional knock-in mice containing Aml1-ETO combined with *Asx12* knock-out for leukaemia development, and found mice had a longer latency with a median survival of 301.5 days and incomplete penetrance. Only when *Aml1^{ETO/+};Asx12^{-/-};Mx1-Cre* BM cells were transduced with NRASG12D did mice develop a very rapid leukaemia with a median survival of 16.8 days. *Aml1^{ETO/+};Mx1-Cre* BM cells transduced with NRASG12D also had a rapid leukaemia development with a median survival of 38.5 days (Micol *et al.*, 2017). This study demonstrates the key differences between using knock-in alleles and retroviral transduction, as the retroviral transduction caused a more aggressive disease with complete penetrance in *Asx12^{-/-}* BM cells transduced with AML1-ETO, compared to *Aml1^{ETO/+};Asx12^{-/-};Mx1-Cre* knock-in BM cells. In addition an overexpression of NRASG12D was the main causal factor of the accelerated disease in *Aml1^{ETO/+};Mx1-Cre* and *Aml1^{ETO/+};Asx12^{-/-};Mx1-Cre* BM cells transduced with NRASG12D.

Finally, Chou *et al.*, performed a retroviral transduction of CD34⁺ human AML1-ETO expressing cord blood with NRASG12D. They found when AML1-ETO and NRASG12D were combined there was an enhanced colony replating ability *in vitro* compared to AML1-ETO alone. CD34⁺ cells expressing AML1-ETO had a down-regulation of Bcl-2 and when combined with NRASG12D resulted in Bcl-2 restoration accompanied by a decrease in apoptosis. These findings were confirmed by ectopic expression of Bcl-2 in AML1-ETO expressing cells, which resulted in a reduction in apoptosis. When AML1-ETO expressing cells transduced with NRASG12D were transplanted into NOD/SCID mice there was an increase in engraftment compared to AML1-ETO expressing cells alone, however the mice did not develop a leukaemia (Chou, Wunderlich, Griesinger, & Mulloy, 2011).

To my knowledge no retroviral or lentiviral experiments have been performed to study *CEBPA* mutations alone or in combination with co-occurring mutations. Together these studies show that when mutations that co-occur in AML are combined using retroviral transduction a more aggressive disease is observed.

The use of lentiviral and retroviral transduction has many advantages, as mice can be generated rapidly with multiple mutation combinations easily established. This has been further developed with the recent advances in gene editing using CRISPR-Cas9 technology. Heckl *et al.*, demonstrated that a combination of guide RNAs can be delivered into a single HSC using a lentiviral vector, with up to five genes modified, leading to clonal outgrowth and myeloid malignancy and thereby more closely modelling the complexity of AML (Heckl *et al.*, 2014).

There are however some key issues with using lentiviral and retroviral transduction. Expression is often driven by a strong viral promoter, resulting in gene expression that may not be physiologically relevant. Overexpression of a transgene delivered by a retroviral vector could alter the biological outcome of the transgene *in vivo*. For example low expression, instead of high expression, of PML/RAR α is critical for leukaemogenesis as high expression was shown to be toxic (Westervelt *et al.*, 2003).

The AML1-ETO retroviral transduction studies discussed above demonstrate the variability observed when using retroviral transduction methods, with the outcome of the experiments influenced by a combination of the virus titres, transduction protocols and retroviral constructs (Ren, 2004). There is also the potential for retroviral integration mutagenesis which can lead to oncogenic mutations that play a role in leukaemogenesis. Du *et al.*, demonstrated that retroviral transduction of murine BM cells with a replication-

incomplete murine stem cell virus that only expressed *neo*, could immortalise murine cells *in vitro* (Du, Jenkins, & Copeland, 2005).

The majority of the AML1-ETO retroviral transduction studies also transduced bulk murine BM cells and are thereby transducing not only stem but also progenitor cells, as well as cells from various different lineages. Krivtsov *et al.*, demonstrated MLL-AF9-transduced HSCs were more likely than transduced GMPs to initiate AML in mice (Krivtsov *et al.*, 2013). Transduction of bulk BM cells also leads to unwanted targeting of lymphoid committed cells progressing to a lymphoid leukaemia as seen in retroviral transduction of AML1-ETO and FLT3-length mutation (Schessl *et al.*, 2005). This highlights the importance of targeting the mutant gene expression to specific cell populations of interest.

1.3.2 Knock-in and knock-out mouse models

Mouse models can provide a powerful tool to accurately model mutations. In particular, conditional expression of gene mutation can introduce genetic mutations in a cell in a type-specific, time-controlled, and dosage-relevant manner (Ren, 2004).

Many previous studies have used knock-in and knock-out alleles to study pre-leukaemic mutations in combination with common co-occurring mutations. These include: *Tet2*^{-/-}; *Flt3*^{ITD/+} (Shih *et al.*, 2015), *Dnmt3a*^{-/-}; *Flt3*^{ITD/+} (L. Yang *et al.*, 2016), *Cebpa*^{K/L}; *Flt3*^{ITD/+} (Reckzeh *et al.*, 2012), *Dnmt3a*^{-/-}; *Kras*^{G12D/+} (Chang *et al.*, 2015), and *Dnmt3a*^{R878H/+}; *Npm1*^c; *Flt3*^{ITD/+} (Guryanova *et al.*, 2016).

All studies show that upon combination lead to a more aggressive disease. The LSC population varies between the models, residing within the LSK, LK and GMP compartments. All studies also demonstrated that pre-leukaemic mutation leads to enhanced self-renewal alone. When the pre-leukaemic mutation was combination with the

co-occurring mutation studied, resulted in the acquired self-renewal within the LSC population. Some studies show that the mutations were able to synergise during leukaemia development, such as *Tet2^{-/-};Flt3^{ITD/+}*, which were shown to have an increase in hypermethylation compared to single mutants alone (Shih et al., 2015). Conversely, *Dnmt3a^{-/-};Flt3^{ITD/+}* showed an increase in hypomethylation compared to single mutant alone (L. Yang et al., 2016). However, the direct mechanism that causes the mutation to synergise was not fully explored.

The *Cebpa^{K/L}* mouse model demonstrates how two mutations synergise to cause leukaemia development. The C-terminal mutation was shown to cause pre-leukaemic HSC expansion, thereby increasing the number of committed myeloid progenitors, while the N-terminal mutation preserved the formation of GMP to allow LSCs to form in the committed myeloid cells (Bereshchenko et al., 2009). When *Cebpa^{K/L}* is combined with *Flt3^{ITD/+}* there is a further increase in myeloid progenitors, leading to a more aggressive disease compared to *Cebpa^{K/L}* (Reckzeh et al., 2012).

There are currently no mouse models of *GATA2* mutations, however *GATA2* haploinsufficiency has been investigated with *inv(3)(q21q26)*. The Yamamoto group developed a 3q21q26 bacterial artificial chromosome transgenic model. The model recapitulated the *inv(3)(q21;q26)* as the inversion led to the overexpression of *EVI1*, caused by the *GATA2* distal enhancer being brought into close proximity to *EVI1*, resulting in leukaemogenesis. Removal of the *GATA2* enhancer results in mice no longer developing a leukaemia (Yamazaki et al., 2014). In the human disease *inv(3)(q21;q26)* repositions the *GATA2* enhancer to cause overexpression of *EVI1* and haploinsufficiency of *GATA2* (Groschel et al., 2014). As the bacterial artificial chromosome transgenic model retains the endogenous murine *Gata2* gene, they then combine the 3q21q26 model with *Gata2* heterozygous knock-out to recapitulate both *EVI1* misexpression and *Gata2* haploinsufficiency. Combined mice developed a leukaemia significantly faster than

3q21q26 mice, demonstrating Gata2 haploinsufficiency is critical in contributing to the progression of 3q21q26 leukaemogenesis (Katayama et al., 2017).

Conditional knock-in gene targeting allows genes to be modified in their endogenous locus, allowing them to be expressed at the normal physiological level. Genes can also be induced at a specific time and/or lineage. This has provided the most accurate way to genetically model mutations. Murine models also have the advantage of being able to quantify the disease over time through PB collection and BM aspiration.

However, the use of primary knock-in or knock-out mice means the mutations are induced in a large number of cells and even when a tissue-specific promoter is used to drive mutant gene expression it still dictates that every cell of the tissue expresses the mutant gene. Subsequently the process of clonal selection, which is an important aspect of leukaemia development, is not modelled and may bypass critical steps in leukaemia development. Competitive transplantation helps to address this issue, however it still does not recapitulate the clonal dynamics seen in humans. Using germline mutant mouse models also may not represent the true initiating events in human AML, and instead the presence of the mutation during development may affect aberrant targets and inappropriately model adult disease (Kohnken, Porcu, & Mishra, 2017).

1.3.3 Gene expression and mutation analysis of AML patient samples

Analysis of AML patient samples containing mutation combinations of interest can be performed by xenotransplantation. This technique has been critical for identification of LSCs, allowing them to be both phenotypically and functionally analysed as well as re-isolated for gene expression analysis. This technique was recently used to develop a prognostic score based on a 17-gene expression signature, which ultimately allowed the outcome of AML patients to be predicted (S. W. Ng et al., 2016).

Xenotransplantations however have their disadvantages. On average only 47% of AML samples engraft (Thomas & Majeti, 2017). There is also great variation in the frequency of LSCs between patient samples and between the mouse models used in xenografts. The immunological and interspecies effects in the models used is currently unknown and could have an effect on the molecular biology of the LSCs (Thomas & Majeti, 2017). Xenotransplants also lack the appropriate microenvironment for human cells. However, the recent advancement in the humanised ossicle system has increased the rate of engraftment to 87% (Reinisch et al., 2016).

Gene expression and DNA-methylation analysis of AML patient samples with specific mutation combinations has been used to support findings identified within mouse model studies (Shih et al., 2015; L. Yang et al., 2016). Gene expression analysis of biallelic *CEBPA* patient samples have shown them to possess a homogenous gene expression profile compared with patients with monoallelic *CEBPA*. This included down-regulation of *HOX* genes (Dufour et al., 2010). Lavallée *et al.*, developed a 95-gene signature for typical biallelic *CEBPA* and showed samples from patients with the distinct gene expression signature were more sensitive to JAK inhibitors compared to normal karyotype AML patients, as well as biallelic *CEBPA* samples that do not display the gene expression signature. This indicates that gene expression analysis of biallelic *CEBPA* patient samples could be used to identify patients that may be responsive to JAK inhibitors (Lavallee et al., 2016).

The use of AML patient samples means that the complexity of the disease remains intact, however this also incurs a disadvantage in that there is heterogeneity between patients, and further intra-tumour heterogeneity. Most gene expression and mutation analysis has been performed on bulk BM samples or on sorted populations of interest such as LSK cells, which is still a heterogenous population. New technologies have been introduced to

decipher the heterogeneity at the single cell level. For example, combining targeted mutational analysis with single cell RNA-sequencing has recently shown heterogeneity in CML stem cells that was previously unrecognised (Giustacchini et al., 2017). This technique presents a powerful tool in understanding pre-leukaemic and LSC biology in AML.

1.4 Acute erythroid leukaemia

AEL is an uncommon subtype of AML, representing <5% of all cases. AEL is characterised by predominant erythroid proliferation (Zuo, Polski, Kasyan, & Medeiros, 2010). The 2008 WHO criteria classified AEL into two subtypes: erythroleukaemia (erythroid/myeloid leukaemia) and pure erythroid leukaemia. The erythroleukaemia is the most common type of AEL and is defined as $\geq 50\%$ erythroid precursors and $\geq 20\%$ of myeloblasts in the non-erythroid component. Pure erythroid leukaemia is defined as $\geq 80\%$ erythroid precursors (Vardiman et al., 2009). AEL was designated as AML-M6 in the previous FAB classification system (Bennett et al., 1976).

In the most recent updated 2016 WHO criteria has now removed the erythroleukaemia subtype from the AML category. In the new classification myeloblasts are counted as a percentage of total BM. This change therefore means many of the cases previously classified as AEL with a blast count of <20%, would be reclassified as MDS (usually MDS with excess blasts). Cases with $\geq 50\%$ erythroid precursors and $\geq 20\%$ total myeloblasts would be classified as AML with myelodysplasia-related changes (Arber et al., 2016). Pure erythroid leukaemia remains as an AML. The change in criteria was based on the biological similarity between erythroleukaemia and MDS in terms of clinical presentation, morphology, mutations, and the requirement to generalise the way blasts are counted across all myeloid neoplasms (Hasserjian et al., 2010; Zuo et al., 2012). However, not

everyone agrees with the new classification. Several studies have shown that the presence of $\geq 50\%$ of erythroid precursors is a poor prognosis indicator in MDS (Mazzella et al., 2006; S. A. Wang et al., 2008), with some suggesting that calculating the percentage of BM blasts from the non-erythroid cells can improve prognostic assessment of MDS (Arenillas et al., 2016). Others are concerned that mixing erythroleukaemias and MDS with excess blasts, without an erythroid-rich subcategory may lead to patients not being sufficiently identified as possessing a more aggressive type of MDS (Calvo et al., 2016). However, the prognostic value of erythroid-rich diseases remains controversial (Arber, 2017).

Only a few studies have investigated the mutations acquired in AEL patients. The most common mutations are *TP53* (30%), *NPM1* (16%), *DNMT3A* (12%), *IDH1/2* (8%), *RUNX1* (7%), and *NRAS* (5%) (Cervera et al., 2016; Grossmann et al., 2013; Ping et al., 2017). Ping *et al.*, analysed six erythroleukaemia AEL patient samples by whole exome sequencing and identified four patients with *CEBPA* mutations and two patients with *GATA2* mutations. Using targeted sequencing 18/55 (33%) of AEL patients were found to have *CEBPA* mutations, of which 6/22 were biallelic mutations. They also found 13/58 (22%) *GATA2* mutations, of which four had a biallelic *CEBPA* mutation (Ping et al., 2017). When comparing the mutational spectrum to other haematological neoplasms, AEL has a lower frequency of *FLT3* mutations compared to AML overall. AEL has a high frequency of *TP53* mutations, which is comparable to therapy-related AML and AML with myelodysplastic-related changes; however AEL has a low frequency of *ASXL1* mutations which is a marker of dysplasia. There is also a similar frequency of spliceosome mutation compared to *de novo* AML. Together, this shows AEL has molecular features seen in both AML and MDS (Cervera et al., 2016). After the change in WHO classification of AEL, Cervera *et al.*, revisited the samples previously diagnosed as AEL and reclassified the samples as either MDS or AML. They then compared the gene mutations and prognostic markers between the reclassified MDS and AML. No difference in molecular profile or

overall survival were found between the two groups. They found the prognosis of the previously classified AEL samples were best predicted using molecular classification compared to morphological features. Cervera *et al.*, also screened for *GATA2* and *CEBPA* mutations, finding no *GATA2* mutations in the 51 cases analysed, and only found one sample with biallelic *CEBPA* mutations (Cervera *et al.*, 2017).

There are currently no mouse models of AEL. Murine AEL cell lines have been produced by activation or inactivation of genes using retroviral integration. Moreau-Gachelin *et al.*, demonstrated an erythroleukaemia cell line could be generated by injecting spleen focus forming virus (SFFV) into mice. Analysis of the proviral integration sites showed a common integration site for *Sfp1*, causing transcriptional activation of *Sfp1* (Moreau-Gachelin, Tavitian, & Tambourin, 1988). Injecting Friend murine leukaemia virus (F-MuLV) into newborn BALB/c or NIH/Swiss mice showed the most common integration site was for *Fli1*, which led to transcriptional activation of *Fli1* (Ben-David, Giddens, & Bernstein, 1990). Inactivation of p53 has also been a common, but unessential, step in leukaemic transformation induced by F-MuLV (Ben David, Prideaux, Chow, Benchimol, & Bernstein, 1988).

Both *Sfp1* (also known as PU.1) and *Fli1* have low expression in WT Pre-CFU-Es (Drissen *et al.*, 2016), and therefore over-expression of both *Spi1* and *Fli1* in an erythroid progenitor cell may block erythroid differentiation. *Sfp1* has been shown to interact directly with Gata1, a transcription factor required for erythroid differentiation. *Sfp1* represses Gata1-mediated transcription and blocks terminal differentiation in erythroleukaemia cells (Rekhtman, Radparvar, Evans, & Skoultchi, 1999).

1.5 Aims of the present work

Previous studies have shown that a single mutation alone is insufficient to cause a leukaemia and only when mutations are combined can it lead to leukaemia development. However, the molecular and cellular mechanism by which the mutations synergise still remains poorly understood. Specific associations between mutations also suggests the mutations must collaborate during leukaemogenesis. Knock-in mouse models provide the most accurate way to genetically model the mutations, and therefore the aim is to model pre-leukaemic mutations in combination with common co-occurring mutations using knock-in mouse models, to study the molecular and cellular mechanism that lead to AML development.

Studies have shown balanced translocations, such as AML1-ETO, and mutations in epigenetic regulators are acquired within the HSCs, generating pre-leukaemic stem cells (Corces-Zimmerman, Hong, Weissman, Medeiros, & Majeti, 2014; Jan et al., 2012; Miyamoto et al., 2000). In contrast, activated signalling mutations of genes such as *FLT3*, *KIT*, or *KRAS* occur as late events within transformed leukaemic progenitors, and are rarely detected in the pre-leukaemic HSC compartment (Corces-Zimmerman et al., 2014; Shlush et al., 2014). Knock-in mutations or knock-out of pre-leukaemic mutations, such as the biallelic *Cebpa* mutations, *Tet2* and *Dnmt3a* knock-out, have been studied in combination with the signalling mutation *Flt3-ITD* (Reckzeh et al., 2012; Shih et al., 2015; L. Yang et al., 2016). However, *Flt3* is not expressed at detectable levels in repopulating mouse HSCs (Adolfsson et al., 2001). Resultantly such modelling cannot measure the intrinsic effect of a signalling mutation on pre-leukaemic HSCs, while cell-extrinsic effects of *Flt3-ITD* would likely play a role (Mead et al., 2017). It is therefore unclear what the impact of a signalling mutation would be on the pre-leukaemic HSC.

It is hypothesised that the absence of signalling mutations in the HSC may reflect a detrimental impact of such mutations on pre-leukaemic HSCs. To address this question, the conditional mouse genetics were used to introduce Aml1-ETO and K-RasG12D separately or in combination into WT HSCs, to determine the effect of K-Ras activation on a well-defined pre-leukaemic HSC population.

Sequencing studies of AML patient samples have identified a specific association between the biallelic *CEBPA* and *GATA2* mutations suggesting the mutations synergise during leukaemia development (Papaemmanuil et al., 2016). This association is also observed in the erythroleukaemia subtype of AEL which is characterised by bi-lineage transformation of both the myeloid and erythroid lineages (Ping et al., 2017). These observations suggest that the biallelic *CEBPA* and ZF1 *GATA2* mutations synergise during leukaemogenesis, and in particular this combination may contribute to both a myeloid and erythroid leukaemia. Myeloid leukaemias are generated through transformation of the neutrophil granulocyte lineages (Goardon et al., 2011; Kirstetter et al., 2008; Krivtsov et al., 2006). Murine studies have shown that neutrophil differentiation occurs through progenitors committed to neutrophil/monocyte lineages (NMPs), where *Gata2* expression is virtually absent (Drissen et al., 2016). Conversely, erythroid lineage progenitors express high levels of *Gata2* but lack *Cebpa* expression (Pronk et al., 2007). This raises two key questions; 1) how, and within which cell type, does transformation occur when both *CEBPA* and *GATA2* are mutated, and 2) if a bi-lineage leukaemia develops, is it maintained by a single bi-potent, or by two distinct lineage-restricted leukaemia propagating cell populations? These distinctions are of considerable importance as there are currently no models of AEL and no mutations identified as causative of AEL.

The biological basis for the association between biallelic *CEBPA* and *GATA2* mutations and role in leukaemogenesis is currently unknown. Therefore, to determine if biallelic

CEBPA and ZF1 *GATA2* mutations synergise during leukaemogenesis, and in particular if they induce an AEL, a genetically accurate model of the combined mutations was generated using the well-established biallelic *Cebpa* knock-in mouse model (Bereshchenko et al., 2009) in combination with a novel ZF1 point mutation knock-in *Gata2*^{G320D}, which is found to be mutated in both AML and AEL (Fasan et al., 2013; Greif, Dufour, et al., 2012; Hou et al., 2015; Ping et al., 2017).

There are also no current mouse models of *GATA2* point mutations. Phenotypic variation and separate clustering of the germline, and somatic *GATA2* mutations in haematological neoplasms, suggest the mutations in ZF1 may affect the function of *GATA2* differently to mutations in ZF2. To understand the effect of ZF1 *Gata2* mutation on haematopoiesis, the *Gata2*^{G320D} mouse model was studied at both heterozygous and homozygous level.

CHAPTER TWO: Materials and methods

2.1 Mice

2.1.1 Husbandry

All mouse lines were maintained on a pure C57Bl/6J genetic background. All mice were bred and maintained in accordance with UK Home Office regulations. Experiments were conducted following approval by the University of Oxford Animal Welfare and Ethical Review Body (project license numbers 30/3103 and 30/3359).

2.1.2 Mouse strains

Conditional knock-in mice expressing Aml1-ETO (*Aml1^{ETO/+}*) (Higuchi et al., 2002) and oncogenic K-Ras (*Kras^{G12D/+}*) (Jackson et al., 2001), either individually or combined (*Aml1^{ETO/+};Kras^{G12D/+}*), were crossed to the *Mx1*-Cre mouse line (Kuhn, Schwenk, Aguet, & Rajewsky, 1995).

Novel constitutive knock-in mice expressing Gata2G320D (*Gata2^{G320D/+}*) were generated by Cyagen Biosciences Inc, California, USA. The G320D mutation was introduced into exon four in the 3' homologous arm of the target vector (BAC clone RP23-81F7 or RP23-108J7) by site-directed mutagenesis. The target vector contained a Neo cassette flanked by Frt sites and DTA was used for negative selection. The neo cassette was removed upon Flp-mediated recombination. C57Bl/6J ES cells were used for gene targeting. *Gata2^{G320D/+}* mice were then bred to homozygosity.

Gata2^{G320D/+} mice were combined with constitutive knock-in mice expressing a mutation in the C-terminal of C/EBPα (*Cebpa*^{K313KK/+}) (Bereshchenko et al., 2009).

Gata2^{G320D/+}; *Cebpa*^{K313KK/+} were then time-mated with constitutive knock-in mice expressing a mutation in the N-terminal of C/EBPα (*Cebpa*^{Lp30/+}) (Kirstetter et al., 2008), to produce triple transgenic mutant FL cells (*Cebpa*^{K313KK/Lp30}; *Gata2*^{G320D/+}).

2.2 Genotyping

DNA was isolated from either ear notches using E.Z.N.A tissue DNA kit (Omega Biotek, Georgia, USA), from cells using QIAamp DNA minikit (Qiagen, Hilden, Germany), or from embryo tail tips using Phire Animal Tissue Direct PCR Kit (Thermo Fisher Scientific, Massachusetts, USA). Isolated DNA was kept at -20°C. Genotyping was performed using primers in Table 2.1. Sex genotyping was performed using a PCR on the *Sry* gene using primers in Table 2.1.

Table 2.1 Genotyping primers

Genotype	Forward Primer	Reverse Primer	WT Primer
<i>KrasG12D</i>	AGCTAGCCACCATGGCTTGAGTAAGTCTGCA	CCTTTACAAGCGCACGCAGACTGTAGA	GTCGACAAGCTCATGCGGGTG
<i>Aml1-ETO</i>	TGCCTGTGACAGCACTGATG	CTGTGGTAGGTGGCGACTTG	ACAGATACGCACCTCTACCG
<i>Gata2G320D</i>	GTAGCCCTCTCTTCCGTGTT	GCCACAGTTGTCCAGGATTG	
<i>CebpaLp30</i>	GCACCAGACGTCTATAGAC	CTGTCGGCTGTGCTGGAAGA	
<i>CebpaK313KK</i>	CGCTGGTGATCAAACAAGAG	CTCGTTGCTGTTCTTGTCCA	
<i>Cre</i>	CGTTTTCTGAGCATACCTGGA	ATTCTCCCACCGTCAGTACG	
<i>Sry</i>	CATCGGAGGGCTAAAGTGTC	GCTGGTGGTGGTCATAGAAC	

2.3 Timed mating

To produce *Gata2*^{G320D/+};*Cebpa*^{K313KK/Lp30} FL cells, *Cebpa*^{Lp30/+} females were placed in a cage with a *Gata2*^{G320D/+};*Cebpa*^{K313KK/+} stud male overnight. *Cebpa*^{Lp30/+} females were then checked for plugs the following morning. Embryos were harvested 14.5 days after plugging.

2.4 Cell preparation

2.4.1 Bone marrow

Tibias and femurs were harvested and placed in phosphate-buffered saline (PBS) + 5% foetal calf serum (FCS). Bones were then crushed using a pestle and mortar and filtered through a 50µM sterile filter. 80µl was then used to perform a cell count on a Sysmex KX-21N (Sysmex, Milton Keynes, UK). Cells were then aliquoted according to the number of cells needed for analysis, transplantation, or freezing, and spun down at 500g for 5mins at 4°C.

2.4.2 c-Kit enrichment of BM cells

c-Kit enrichment of BM cells was performed by positive selection using MACS magnetics separation (MACS Miltenyi Biotec, Bergisch Gladbach, Germany). BM cells were magnetically labelled using CD117 MicroBeads (MACS Miltenyi Biotec) at 2.5µl/100 million cells. Cells were incubated on ice for 20mins. Cells were then washed with PBS+5%FCS and spun down at 500g for 5mins at 4°C and resuspended in 3ml of PBS+5%FCS. LS columns MicroBeads (MACS Miltenyi Biotec) were placed on the MACS magnetic separator and first washed with 3ml of PBS+5%FCS. 50µM filter was

placed on top of the LS column. BM cells were then added to the LS column. The LS column was then washed three times using 3ml of PBS+5%FCS each time. LS column was then removed from the magnet and placed into a 15ml falcon. c-Kit positive cells were eluted by adding 5ml of PBS+5%FCS and forced through the LS column using the plunger. LS column was then washed using 3ml of PBS+5%FCS. 80µl of elute was used to perform a cell count on a Sysmex KX-21N. Cells were then aliquoted according to the number of cells needed for analysis and spun down at 500g for 5mins at 4°C.

2.4.3 Spleen

Spleen was harvested and placed in PBS+5%FCS. Spleen was initially weighted and then crushed using a 5ml syringe plunger through a 70µM sterile cell strainer. Strainer was then washed with 5ml of PBS+5%FCS. 80µl was then used to perform a cell count on a Sysmex KX-21N. Cells were then aliquoted according to the number of cells needed for analysis, transplantation or freezing, and spun down at 500g for 5mins at 4°C.

To analyse the basophil-mast cell progenitor (BMCP) in the spleen, the spleen was crushed using a 5ml syringe plunger through a 70µM sterile cell strainer. The strainer was then washed with 5ml of PBS+5%FCS. Cells were then spun down at 500g for 5mins at 4°C. Supernatant was removed and the pellet was resuspended in 100µl PBS+5%FCS, then 200µl of 2% dextran sulphate sodium salt was added (Sigma-Aldrich, Missouri, USA), and incubated at 37°C for 20mins. Upper phase was then taken and washed using PBS+5%FCS at 3800rpm for 5mins at 4°C. Supernatant was removed and pellet was dissolved in 200µl of ammonium chloride (Stem Cell Technologies, Vancouver, Canada) and incubated for 2mins to lyse the RBCs. Cells were washed with 1ml of PBS+5%FCS, and spun down at 3800rpm for 5mins at 4°C.

2.4.4 Peripheral blood

100-120µl of PB was taken from the tail vein of the mice, or taken from a cardiac puncture when mice were terminated. PB was collected in either EDTA or heparin coated microvette tubes depending on whether RBCs or platelets were analysed, respectively. PB parameters were taken by making a 1:5 dilution using 16µl of blood and 64µl of PBS and was analysed using a Sysmex KX-21N.

RBCs were first isolated from PB by taking 2µl of blood and 48µl of PBS+5%FCS was added. 50µl of 2x antibody cocktail was then added for fluorescence-activated cell sorting (FACS) staining.

Platelets were isolated from PB by spinning the PB at 1000rpm for 10mins at room temperature. 10µl of plasma was collected and 40µl of PBS+5%FCS was added. 50µl of 2x antibody cocktail was added for FACS staining.

White blood cell (WBCs) were isolated from the blood by taking 100µl of blood and adding 100µl of PBS+5%FCS. 200µl of 2% dextran sulphate sodium was added and incubated at 37°C for 20mins. Upper phase was then taken and washed using PBS+5%FCS at 3800rpm for 5mins at 4°C. Supernatant was removed and pellet was dissolved in 200µl of ammonium chloride and incubated for 2mins to lyse the RBCs. Cells were washed with 1ml of PBS+5%FCS, spun down at 3800rpm for 5mins at 4°C, and resuspended in 1x antibody cocktail for FACS analysis.

2.4.5 Peritoneal cells

Peritoneal cells were isolated by first injecting 5ml of PBS+5%FCS into the peritoneum of the mouse. A syringe was then used to recover the PBS+5%FCS in the peritoneum and

put into a 15ml falcon tube. The abdomen was then cut open and the remaining PBS+5%FCS was recovered using a Pasteur pipet. Cells were then spun down at 500g for 5mins at 4°C for analysis.

2.4.6 Foetal liver cells

E14.5 embryos were removed from the culled pregnant female and placed in a 50ml falcon containing 25ml of PBS+10%FCS. In a petri dish each embryo was separated using forceps. Yolk sac was then cut open to release the embryo. The embryo was then separated from the placenta. Schedule one of the embryo was performed by cutting the hindbrain using scissors. The tip of the tail was cut and placed in an empty 1.5ml Eppendorf tube for DNA extraction. FL was then isolated from each embryo and placed in PBS+10%FCS. A single cell suspension of the FL was made by using a 1ml syringe with an 18G needle to gently break up the FL. The needle was then replaced with a 27G needle to break up the FL further. The cells were then filtered through a 35µM cell strainer. Cells were then spun down at 300g for 5mins at 4°C and resuspended in freezing media. See section 2.4.7 for freezing method.

2.4.7 Freezing cells

10×10^6 BM cells or spleen cells were aliquoted and spun down at 500g for 5mins at 4°C. Supernatant was removed and resuspended in 500µl of FCS. Cells were then transferred into cryovials and 500µl of FCS+20% dimethyl sulfoxide (DMSO) was added to make a final concentration of 10% DMSO. Cryovials were then placed in a freezing container and placed in the -80°C freezer. Cryovials were then transferred into liquid nitrogen storage the following day.

2.4.8 Thawing frozen cells

Cryovials were thawed rapidly in 37°C water bath. Cells were then transferred into a 15ml falcon tube. 1ml of Iscove's modified Dulbecco's medium (IMDM)+20%FCS was used to rinse the cryovial. 13ml of IMDM+20%FCS was added dropwise into the 15ml falcon containing the cells. The cells were then spun down at 300g for 6mins at room temperature. Supernatant was then removed and resuspended in 10mls of IMDM+20%FCS and spun down at 300g for 6mins at room temperature. Supernatant was then removed and resuspended in 1ml of PBS+5%FCS. Cells were then counted using a haemocytometer.

2.5 Flow cytometry and fluorescence-activated cell sorting

2.5.1 Antibody staining and analysis

Details of mouse antibodies and viability dyes used for flow cytometry are shown in Table 2.2. All antibodies were used at pre-determined optimal concentrations. From BM and spleen 15×10^6 cells were used for HSC staining panel, 10×10^6 cells were used for myelo-erythroid staining panel, 5×10^6 cells were used for WBC staining panel, 5×10^6 cells were used for erythroblast panel. In stains where anti-FcγRII/III antibody was not included, cells were incubated with Fc-block. Gates were set using a combination of fluorescence minus one controls and populations known to be negative for the antigen.

HSPCs were defined as the following: HSC, LSKCD150⁺Flt3⁻ or LSKCD150⁺CD48⁻; MPP, LSKCD150⁻CD48⁺; and LMPP, LSKCD150⁻Flt3⁺. Myelo-erythroid progenitors were defined as the following: MkP, LKCD41⁺CD150⁺; pre-GM, LKCD41⁻FcγRII/III⁻CD105⁻CD150⁻; EMkMPP, LKCD41⁻FcγRII/III⁻CD105⁻CD150⁻CD55⁺; pre-NM, LKCD41⁻FcγRII/III⁻CD105⁻

CD150⁻CD55⁻; GMP, LKCD41⁻CD150⁻FcγRII/III⁺; EoMP, LKCD41⁻CD150⁻
 FcγRII/III⁺CD55⁺; NMP, LKCD41⁻CD150⁻FcγRII/III⁺CD55⁻; pre-MegE, LKCD41⁻FcγRII/III⁻
 CD105⁻CD150⁺; pre-CFU-E, LKCD41⁻FcγRII/III⁻CD105⁺CD150⁺; and CFU-E, LKCD41⁻
 FcγRII/III⁻CD105⁺CD150⁻. Myeloid cells were defined as the following: PB myeloid cell
 reconstitution, CD19⁻CD4⁻CD8a⁻NK1.1⁻Mac1⁺; BMCP in the spleen LKβ7-
 integrin⁺FcγRII/III⁺; mast cells in the peritoneum CD4⁻CD8a⁻Ter119⁻CD19⁻CD11b^{lo}Ly6G⁻
 SiglecF⁺FcεR1α⁺; eosinophils in the PB and peritoneum CD4⁻CD8a⁻Ter119⁻CD19⁻
 CD11b^{lo}Ly6G⁻SiglecF⁺; monocytes in the PB and peritoneum, CD4⁻CD8a⁻Ter119⁻CD19⁻
 CD11b⁺Ly6G⁻ and neutrophils in the PB and peritoneum CD4⁻CD8a⁻Ter119⁻CD19⁻
 CD11b^{lo}Ly6G⁺. B cells were defined as follows: PB B cell reconstitution, Mac1⁻NK1.1⁻
 CD19⁺CD4⁻CD8a⁻; B cells in PB and peritoneum (myeloid panel); Ter119⁻CD19⁺CD4⁻
 CD8a⁻. T cells were defined as follows: PB T cell reconstitution, Mac1⁻NK1.1⁻CD19⁻
 CD4⁺CD8a⁺; T cells in PB and peritoneum (myeloid panel), Ter119⁻CD19⁻CD4⁺CD8a⁺.
 Platelets were defined as CD150⁺CD41⁺Ter119⁻. RBCs were defined as CD150⁻CD41⁻
 Ter119⁺. Terminal erythroid differentiation populations were defined as the following:
 proerythroblasts (stage I), Ter119^{-/lo}CD71⁺; basophilic erythroblasts (stage II),
 Ter119⁺CD71⁺; polychromatic erythroblasts (stage III), Ter119⁺CD71^{mid}; and
 orthochromatic erythroblasts (stage IV) Ter119⁺CD71⁻. Leukaemic erythroblasts were
 defined as CD71⁺Ter119^{-/lo}c-Kit⁺. Leukaemic myeloblasts were defined as c-Kit^{lo}Mac1⁺.
 Leukaemic immature myeloid cells were defined as c-Kit⁺Mac1^{lo}.

Table 2.2 Antibodies used for FACS staining

Antibody details				Chapter 3: AKM project				Chapter 4: AEL project				Chapter 5: Gata2 homozygous mutant project						All
Antigen	Conjugate	Company	Cat- alogue No.	HSC	HSC cell cycle	HSC RNA- sequencing /LSK sort	Myelo- erythroid progenitors	Myelo- erythroid progenitor with CD55	HSC	Erythro- blast	Platelet and RBC	Myelo- erythroid progenitors	Myelo- erythroid progenitor with CD55	BMCP	Mye- loid	<i>In vitro</i> cell culture	HSC	WBC
7AAD	PE-Cy5	Cayman Chemical	11397	Y			Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y
CD4	APC-eF780	eBioscience	47-0042- 82												Y			Y
CD8a	APC-eF780	eBioscience	47-0081- 82												Y			Y
NK1.1	PB	BioLegend	108722															Y
Gr1	PO	Caltag Medsystem	RM3030															Y
CD19	PE-Cy7	BioLegend	25-0183- 82												Y			Y
Mac1	APC	BioLegend	101212							Y								Y
CD45.1	PE	eBioscience	12-0453- 83															Y
CD45.2	AF700	BioLegend	109822	Y	Y	Y	Y	Y	Y	Y						Y		Y
FcyR	PE	eBioscience	12-0161- 81				Y					Y						
CD4	PE-Cy5	BioLegend	100514				Y	Y	Y			Y	Y	Y			Y	
CD8a	PE-Cy5	BioLegend	100710				Y	Y	Y			Y	Y	Y			Y	
Ter119	PE-Cy5	BioLegend	116210	Y			Y	Y	Y			Y	Y	Y	Y		Y	
Mac1	PE-Cy5	BioLegend	101210				Y	Y	Y			Y	Y	Y			Y	
CD19	PE-Cy5	eBioscience	15-0193- 83	Y			Y											
Gr1	PE-Cy5	BioLegend	108410	Y			Y	Y	Y			Y	Y	Y			Y	
CD150	APC	BioLegend	115910				Y	Y				Y	Y					

cKit	APC-eF780	eBioscience	47-1171-82	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y		Y	Y	
Sca1	BV786	BD	563991				Y					Y						
CD105	BV421	BD	562760				Y					Y						
CD41	PE-Cy7	eBioscience	25-0411-82				Y					Y						
CD45.1	BV650	BioLegend	110736	Y	Y	Y	Y	Y	Y	Y						Y		
CD48	APC	BioLegend	103412	Y					Y								Y	
F4/80	PE-Cy5	BioLegend	123112	Y														
CD3e	PE-Cy5	BioLegend	100310	Y														
NK1.1	PE-Cy5	BioLegend	108716	Y														
CD150	PE-Cy7	BioLegend	115914	Y	Y	Y			Y	Y	Y						Y	
Sca1	PB	BioLegend	108120	Y	Y	Y			Y								Y	
CD34	Biotin	eBioscience	13-0341-85	Y													Y	
Streptavidin	PE-TxRed	BD	551487	Y				Y					Y				Y	
Flt3	PE	BioLegend	135306	Y	Y	Y			Y								Y	
CD45.1	FITC	BioLegend	110706		Y													
Ki-67	PE	BD	556027		Y													
Flt3	APC	BioLegend	135309		Y													
Sca1	PerCP-Cy5.5	eBioscience	45-5981-80		Y													
DAPI	BV421	Thermo	D3571		Y													
F4/80	APC	BioLegend	123116			Y												
CD3e	APC	eBioscience	17-0031-83			Y												
NK1.1	APC	eBioscience	17-5941-82			Y												
Ter119	APC	eBioscience	17-5921-81			Y												
Gr1	APC	BioLegend	108412			Y												

CD19	APC	eBioscience	17-0193-82			Y												
CD5	PE-Cy5	BioLegend	100610					Y	Y			Y	Y	Y			Y	
B220	PE-Cy5	BioLegend	103210					Y	Y			Y	Y	Y			Y	
FcγR	PE-Cy7	eBioscience	25-0161-82					Y					Y	Y				
Sca1	BV605	BioLegend	108133					Y					Y					
CD105	Biotin	eBioscience	13-1051-85					Y					Y					
CD41	BV421	BioLegend	133911					Y		Y			Y					
CD55	PE	BioLegend	131804					Y					Y					
β7-integrin	PE	BD	557498											Y				
FcεR1a	APC	eBioscience	17-5898-82												Y	Y		
Ly6G	PB	BioLegend	127612												Y	Y		
SiglecF	PE	BD	552126													Y		
Mac1	AF700	eBioscience	56-0112-82												Y			
Mac1	PE-Cy7	eBioscience	25-0223-81													Y		
CD71	PE	BioLegend	113808							Y								
Ter119	PerCP-Cy5.5	eBioscience	45-5921-82							Y	Y							
CD41	PE	eBioscience	12-0411-83								Y							

Y = used in staining panel

2.5.2 HSC cell cycle

BM cells were c-Kit-enriched and stained following initial Fc-block incubation. Stained cells were then fixed and permeabilised using BD cytofix/cytoperm fixation and permeabilisation solution (BD Biosciences, California, USA). Cells were stained with Ki-67 PE (BD Biosciences) overnight. Cells were then stained with 4',6-diamidino-2-phenylindole (DAPI) (0.5µg/ml) (Thermo Fisher Scientific) for one hour before analysis.

2.5.3 Fluorescence-activated cell sorting

BM cells were first c-Kit-enriched for *Aml1*^{ETO/+};*Kras*^{G12D/+} HSC RNA-sequencing, *Aml1*^{ETO/+};*Kras*^{G12D/+} LSK cell serial replating, and *Gata2*^{G320D/G320D} EoMP cell multiplex quantitative real-time PCR analysis. 40x10⁶ BM cells were used and were not c-Kit-enriched for *Gata2*^{G320D/G320D} BMCP RNA-sequencing, *Gata2*^{G320D/G320D} EMkMPP/EoMP cell culture, *Gata2*^{G320D/+};*Cebpa*^{K313KK/Lp30} RNA-sequencing and multiplex quantitative real-time PCR analysis. For *Gata2*^{G320D/+};*Cebpa*^{K313KK/Lp30} secondary transplants all BM cells recovered from thawed vials were used for sorting. In stains where anti-FcγRII/III antibody was not included, cells were incubated with Fc-block. Gates were set using a combination of fluorescence minus one controls and populations known to be negative for the antigen.

2.5.4 Instrument and analysis

Cell acquisition and analysis were performed on a BD LSRFortessa (BD Biosciences) using BD FACSDiva™ software (BD Biosciences). Cell sorting was performed on a BD FACSAriaII cell sorter (BD Biosciences). Analysis was performed using Flowjo software version 10.0.8 (Flowjo LLC, Oregon, USA).

2.6 Transplantation assays

For *Aml1*^{ETO/+};*Kras*^{G12D/+} project, congenic CD45.1 WT C57Bl/6J adult mice (7-13 weeks old) were utilised as recipients. Competitive transplantations were performed using 2.5×10^5 CD45.2 FL cells and 1×10^6 CD45.1 WT BM competitor cells into lethally irradiated recipients (2 x 500rads). Recombination was induced by subcutaneous injection of poly(I:C) four weeks post-transplantation, and mice culled eight weeks post-poly(I:C) and haematopoietic phenotype analysed.

Serial transplantations were performed by co-transplanting 1.25×10^5 CD45.2 FL cells with 5×10^6 CD45.1 WT BM competitor cells into lethally irradiated primary recipients. Bulk secondary and tertiary transplants were then performed by transplanting 3×10^6 BM cells from primary and secondary recipients into lethally irradiated recipients at eight weeks post-poly(I:C) for secondary transplants or 12 weeks post-transplantation for tertiary transplants. Tertiary transplanted mice were analysed 12 weeks post-transplantation.

For *Gata2*^{G320D/G320D} project, congenic CD45.1/CD45.2 WT C57Bl/6J adult mice (7-12 weeks old) were utilised as recipients. Competitive transplantations were performed, using 5×10^5 CD45.2 BM cells and 5×10^5 CD45.1 WT BM competitor cells transplanted into lethally irradiated recipients. Bulk secondary transplants were performed by transplanting 10×10^6 BM cells, from primary transplanted mice 17-18 weeks post-transplantation, into lethally irradiated recipients.

For *Gata2*^{G320D/+};*Cebpa*^{K313KK/Lp30} project, CD45.1/CD45.2-*Gata1*-EGFP adult mice (8-13 weeks old) were utilised as recipients (Drissen et al., 2016). Competitive transplantations were performed by using 2.5×10^5 CD45.2 FL cells and 2.5×10^5 CD45.1/CD45.2-*Gata1*-EGFP BM competitor cells into lethally irradiated recipients. For pre-leukaemia analysis mice were culled at six weeks post-transplantation. For leukaemia analysis mice were

monitored up to one year post-transplantation. Mice were culled earlier if mice became hunched with pale paws, PB WBC count was $\geq 15 \times 10^9/L$, and RBC count was $\leq 7 \times 10^{12}/L$.

For secondary transplants lethally irradiated CD45.1/CD45.2-*Gata1*-EGFP adult mice (8-13 weeks old) were utilised as recipients. Bulk secondary transplants were performed by transplanting $5-7.5 \times 10^5$ BM cells from *Gata2*^{G320D/+}; *Cebpa*^{K313KK/Lp30} that develop a leukaemia with 2.5×10^5 CD45.1/CD45.2-*Gata1*-EGFP BM cells for radioprotection, into CD45.1/CD45.2-*Gata1*-EGFP lethally irradiated recipients. For secondary transplants using sorted cell populations, all cells that were collected from the sort were split into two recipients with 2.5×10^5 CD45.1/CD45.2-*Gata1*-EGFP BM cells for radioprotection, into CD45.1/CD45.2-*Gata1*-EGFP lethally irradiated recipients.

2.7 *In vitro* assays

2.7.1 Serial replating

100 CD45.2 LSK BM cells were sorted from mice transplanted with 2.5×10^5 FL cells and 1×10^6 CD45.1 WT BM competitor cells eight weeks post-poly(I:C). Cells were seeded into 1ml of methylcellulose medium (Methocult, M3434, Stem Cell Technologies) and incubated in 37°C, in 5% CO₂, with $\geq 95\%$ humidity. Colonies (≥ 30 cells) were counted after eight days. Cells were resuspended and replated at 1×10^4 cells per 1ml of methylcellulose medium. Cells were then counted and replated after 6-7 days.

2.7.2 Cell culture for colony gene expression and morphology analysis

100 *Cebpa*^{K313KK/Lp30}; *Gata2*^{G320D/+} or 300 WT CD45.2 LKFcγRII/III⁺ BM cells were sorted from mice transplanted with 2.5×10^5 CD45.2 FL cells and 2.5×10^5 CD45.1/CD45.2 BM competitor cells six weeks post-transplantation. Cells were seeded into 1ml of

methylcellulose medium (Methocult, M3434, STEMCELL Technologies) and incubated in 37°C, in 5% CO₂, with ≥95% humidity. After 8 days colonies (≥30 cells) were counted and colonies were picked by taking 1µl of cells from the colony and cells were resuspended in a well of a 96 well plate containing 200µl PBS+5%FCS. Cell suspension was then split into two separate plates. Both plates were spun down at 500g for 5mins at 4°C.

Supernatant was then removed. One plate was resuspended in 15µl of lysis buffer containing 14.85µl of CellDirect 2x reaction mix (Thermo Fisher Scientific) and 0.15µl of SUPERase-In RNase Inhibitor (Thermo Fisher Scientific), then flash frozen on dry ice and stored at -80°C. See 2.9.2 section for multiplex quantitative real-time PCR analysis method. The second plate was resuspended in 20µl of PBS+20%FCS to be used to make a cytospin. See 2.8 section for cytospin method.

2.7.3 Myeloid cell culture

To test myeloid potential, 100 EoMP BM cells were sorted and cultured in IMDM with L-glutamine (Invitrogen, California, USA), 20% heat-inactivated FCS (Thermo Fisher Scientific), penicillin/streptomycin (Invitrogen) and 0.1µM β-mercaptoethanol (Sigma-Aldrich), supplemented with 20ng/ml mSCF (Peprotech, New Jersey, USA), 20ng/ml mL-3 (Peprotech), 10ng/ml mGM-CSF (Peprotech), 50ng/ml mL-5 (R&D Systems), and 50ng/ml mL-9 (R&D Systems, Minnesota, USA) at 37 °C, 5% CO₂. Cells were analysed after seven days of culture.

2.8 Cell morphology

Cytospin was made using a cytospin centrifuge (Thermo Fisher Scientific) on superfrost microscope slides (Thermo Fisher Scientific). 10x10⁴ BM or spleen cells were spun at 1000rpm for 10mins and left to air dry.

Air-dried cytopsins and blood smears slides were first stained with May-Grüwald stain (Sigma-Aldrich) for 5mins. Slides were then washed with distilled water (Sigma-Aldrich) and then stained with 1:5 dilution of Giemsa stain (Sigma-Aldrich) for 20mins. Slides were then washed in distilled water and air-dried. One drop of Pertex mouting medium (Cell Path, Powys, UK) was placed on each slide using a pasture pipette, and a cover slip placed on top. Slides were analysed using an Olympus BX60 microscope (Olympus, Tokyo, Japan) at 40x or 100x objective using immersion oil (Immersol, ZEISS, Oberkochen, Germany).

2.9 Gene expression analysis

2.9.1 RNA-sequencing

50 or 100 cells per biological replicate were sorted into 4µl of lysis buffer containing; 0.2% Triton X-100 (Sigma-Aldrich, Missouri, USA), 2.5µM OligodT (Biomers, Ulm, Germany), 2.5mM dNTPs (Thermo Fisher Scientific), RNase Inhibitor 20U (Takara Bio USA, Inc, California, USA) and ERCC spike-in 1:4x10⁶ (Thermo Fisher Scientific). cDNA synthesis and PCR amplification were performed based on the published Smart-seq 2 protocol (Picelli et al., 2013) with some modifications. SMARTScribe RT enzyme (Takara Bio USA, Inc) was used in the RT mix (50U) and SeqAMP enzyme (Takara Bio USA, 50U) was used for the PCR step for 18 cycles. cDNA traces were bead-purified using Ampure XP beads (Beckman Coulter, California, USA). cDNA was evaluated using either a high sensitivity DNA kit (Agilent, California, USA) on an Agilent 2100 Bioanalyzer, or using a high sensitivity NGS fragment analysis kit (Advanced Analytical, Milton Keynes, UK) on a Fragment Analyzer. cDNA was quantified either using PicoGreen (Thermo Fisher Scientific) or using a Qubit (Invitrogen). Normalised cDNA traces were used for library preparation using a miniaturised version of the Nextera XT Kit (Illumina, California, USA).

After tagmentation and 12 cycles of barcoding PCR, tagmented libraries were purified using AmpureXP beads, evaluated using a high sensitivity DNA kit on an Agilent 2100 Bioanalyzer and quantified using a Qubit. Finally, libraries were pooled and sequenced on four lanes on a NextSeq 500 (Illumina), using 75 bp single-end reads.

Following quality control analysis with the fastQC package (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>), reads were aligned using STAR (Dobin et al., 2013) against the mm10 mouse reference genome. Gene expression levels were quantified as read counts using the featureCounts function from the Subread package (Liao, Smyth, & Shi, 2013) with default parameters. The read counts were used for the identification of global differential gene expression between specified populations and/or genotypes using either the edgeR package (Robinson, McCarthy, & Smyth, 2010) or the DESeq2 package (Love, Huber, & Anders, 2014). Reads per kilobase of transcript per million (RPKM) values were then calculated. Genes were considered differentially expressed between populations and/or genotypes if they had an adjusted p-value (padj) of less than 0.05. The pheatmap function was used to generate a heatmap, and prcomp function was used to generate a principle component analysis, in R statistical programming environment (www.r-project.org). Venn-diagrams were generated using Venny (<http://bioinfogp.cnb.csic.es/tools/venny/>). Gene-set enrichment analysis (GSEA) was performed using GSEA software (Mootha et al., 2003; Subramanian et al., 2005).

2.9.2 Multiplex quantitative real-time PCR analysis

Multiplex quantitative real-time PCR was performed on single cells or picked colonies from methylcellulose cell culture. CellDirect One-Step qPT-PCR kit (Thermo Fisher Scientific) was used according to manufacturer's protocol for preparation and amplification of cDNA. The BioMark 192.24 Dynamic Array platform (Fluidigm, California,

USA) and Taqman assays (Thermo Fisher Scientific) were used to perform the multiplex quantitative real-time PCR according to the manufacturer's instructions.

Ct values were generated using the BioMark Real-Time PCR analysis software (Fluidigm). Each amplification curve for each gene and each cell was visually inspected on the BioMark Real-Time PCR analysis software. Any outliers that were not automatically detected from the software were manually changed to fail. Data analysis was then performed in R statistical programming environment. Ct values of all assays marked as 'Fail' were set as undetected (Ct = 999). A histogram was generated using Microsoft Office Excel (Microsoft, Washington, USA) to analyse the Ct values for the housekeeping genes. Cut-offs for the housekeeping genes Ct values were set in accordance to the histogram analysis. Cells that had a housekeeping gene Ct value that did not meet the cut-off, or were undetected, were removed from analysis. Ct values were then normalised to the housekeeping gene. If more than one housekeeping gene was used in the assay then a mean was calculated of the housekeeping genes. Ct values were normalised to the mean of the housekeeping genes. $2^{-(\text{Normalised Ct})}$ was then used for analysis. $2^{-(\text{Normalised Ct})}$ values for each gene were visually inspected and outliers removed. Statistical significance between genotypes was performed using a Mann-Whitney test. Pheatmap function was used to generate a heatmap. Lineage priming analysis was performed by calculating the sum of all the myeloid and megakaryocyte/erythroid genes expressed in each cell. Scatterplots were then generated using the sum of the myeloid vs. megakaryocyte/erythroid genes using ggplot2 function in R statistical programming environment.

2.10 Chromatin immunoprecipitation-sequencing data analysis

Previously published chromatin immunoprecipitation (ChIP)-sequencing data for Gata2 in primary murine mast cells (Calero-Nieto et al., 2014; Moignard et al., 2013) was analysed

by mapping the reads to mm10 mouse reference genome using Bowtie2 (Langmead & Salzberg, 2012). SAMtools was then used to convert sam files to bam files (H. Li et al., 2009). Peaks were called using MACS2 with default parameters (Y. Zhang et al., 2008). Peaks were annotated using Homer (Heinz et al., 2010). R statistical programming environment was used to generate bubble plots representing transcription factor binding sites around differentially expressed genes in *Gata2*^{G320D/G320D} BMCPs using the top 5000 most significant peaks from the transcription factor ChIP-sequencing data, and differentially expressed genes with padj < 0.1 from the RNA-sequencing data.

2.11 Statistical analysis

Unless otherwise indicated, statistical significance of differences between genotypes was determined using an analysis of variance (ANOVA) for multiple comparisons following Tukey's multiple comparisons test. Data sets that were shown to not have a normal distribution using a D'Agostino-Pearson omnibus normality test, the non-parametric Kruskal-Wallis test was then used for multiple comparison following Dunn's multiple comparison test. Event-free survival was analysed using Log-rank (Mantel-Cox) test. Graphpad prism version 7 was used to analyse the statistical differences and generate graphs.

CHAPTER THREE: A study of the pre-leukaemic fusion protein Aml1-ETO in combination with a signalling mutation in K-Ras

3.1 Introduction

Different functional groups such as transcription factor fusions and activated signalling mutations have been found to co-occur in AML patients (Cancer Genome Atlas Research, 2013). Mechanistic studies using mouse models support mutational classes that co-occur in AML patients and can collaborate to cause AML development (Reckzeh et al., 2012; Shih et al., 2015; L. Yang et al., 2016). More recently, analyses of purified HSCs and progenitor populations from AML patients have revealed that leukaemia initiating mutations, which include balanced translocations (e.g. AML1-ETO), and mutations in epigenetic regulators, are frequently acquired within the HSC compartment as early events in disease evolution, generating pre-leukaemic stem cells (Corces-Zimmerman et al., 2014; Jan et al., 2012; Miyamoto et al., 2000; Shlush et al., 2014). In contrast, activated signalling mutations of genes such as *FLT3*, *KIT* or *KRAS* occur as late events within transformed leukaemic progenitors, and are rarely detected in the pre-leukaemic HSC compartment (Corces-Zimmerman et al., 2014; Shlush et al., 2014). The impact of an activated signalling mutation acquired in the pre-leukaemic HSC remains unclear. This is of considerable importance for the understanding of why signalling mutations are absent from the pre-leukaemic HSC compartment.

Experimental approaches investigating collaboration between Aml1-ETO and activated Ras have mainly used retroviral expression of oncogenes, showing that such mutations collaborate to cause an acute leukaemia (Chou et al., 2011; Micol et al., 2017; Zhao et al., 2014). However, retroviral transduction can lead to expression at non-physiological levels as well as ectopic expression within the cellular hierarchy (Reckzeh et al., 2012), and

does not accurately address the effect of the mutant proteins on HSC function. Other approaches using conditional knock-in mutations or gene knock-out to study the collaboration between pre-leukaemic mutations such as the biallelic *Cebpa* mutations, *Tet2* and *Dnmt3a* knock-out, in combination with the signalling mutation *Flt3-ITD*, have resulted in leukaemic transformation (Reckzeh et al., 2012; Shih et al., 2015; L. Yang et al., 2016). However, *Flt3* is not expressed at detectable levels in repopulating mouse HSCs (Adolfsson et al., 2001). Therefore, such modelling cannot measure the intrinsic effect of a signalling mutation on pre-leukaemic HSCs, while cell-extrinsic effects of *Flt3-ITD* would be likely to play a role (Mead et al., 2017). Therefore it remains unclear what the impact of signalling mutations would be on the pre-leukaemic HSC in this regard.

It is hypothesised that the absence of signalling mutations in the HSC may reflect a detrimental impact of such mutations on pre-leukaemic HSCs. To address this question we investigated the combination of the pre-leukaemic fusion protein Aml1-ETO in combination with the signalling mutation K-RasG12D. Although *KRAS* mutations are not the most common co-occurring mutations found in t(8;21) patients, they occur in approximately 4% of patients with AML1-ETO (Krauth et al., 2014). Conditional mouse genetics were used to introduce Aml1-ETO and K-RasG12D, separately or in combination, both expressed from their endogenous loci, into WT HSCs, to determine the effect of K-Ras activation on a well-defined pre-leukaemic HSC population.

3.2 Results

3.2.1 Aml1-ETO ameliorates key features of the MPN phenotype caused by K-RasG12D

To study the effect of combining Aml1-ETO with mutant K-Ras, whilst avoiding the previously described spontaneous recombination in primary *Kras*^{G12D/+};*Mx1-Cre*^{tg/+} mice (Sabnis et al., 2009), E14.5 FL cells were generated with the necessary genotypes: *Aml1*^{ETO/+};*Mx1-Cre*^{tg/+} (AM genotype); *Kras*^{G12D/+};*Mx1-Cre*^{tg/+} (KM genotype); *Aml1*^{ETO/+};*Kras*^{G12D/+};*Mx1-Cre*^{tg/+} (AKM genotype); and *Mx1-Cre*^{tg/+} controls (CON genotype); all CD45.2 allotype. To create a scenario where competitive advantage and disadvantage of mutant HSCs could be observed, mice were generated where the haematopoietic system was partially repopulated by experimental cells and partially by WT competitor cells (CD45.1 allotype). This was done by competitively transplanting 2.5×10^5 FL cells (CD45.2) and 1×10^6 WT BM competitor cells (CD45.1) into lethally irradiated CD45.1 recipients. Recombination was induced four weeks post-transplantation by subcutaneous injection of poly(I:C). Long-term analysis of primary transplanted mice lead to both KM and AKM developing a T-ALL (data not shown), therefore the haematopoietic phenotype was analysed eight weeks post-poly(I:C) before T-ALL development occurred (Figure 3.1).

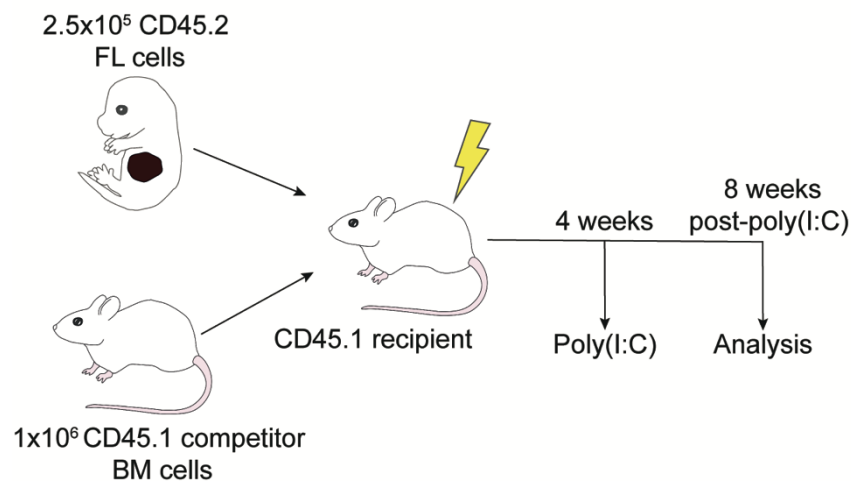


Figure 3.1 Experimental design

Schematic of *in vivo* competitive transplant experiment. 2.5×10^5 FL cells (CD45.2) and 1×10^6 WT BM competitor cells (CD45.1) were transplanted into lethally irradiated CD45.1

recipients. Recombination was induced four weeks post-transplantation by subcutaneous injection of poly(I:C). Mice were analysed eight weeks post-poly(I:C).

AM displayed a mild leukopenia as there was a decrease in WBC count compared to CON (Figure 3.2A). There was no difference in haemoglobin levels, BM cellularity, spleen weight or in morphology compared to CON (Figure 3.2B-D, F). However, AM showed a 27% decrease in platelet count compared to CON (Figure 3.2E).

At eight weeks post-poly(I:C) KM became moribund. In line with previous results (Braun et al., 2004; I. T. Chan et al., 2004), KM developed a MPN, consisting of leucocytosis as there was an increase in WBC count compared to CON, anaemia as there was a decrease in haemoglobin levels compared to CON, increase in BM cellularity, increase in spleen weight, thrombocytopenia as there was a decrease in platelet count compared to CON, and blast cells seen in the blood smears (Figure 3.2A-F).

Strikingly, AKM did not result in a more aggressive disease, but ameliorated key features of the phenotype associated with K-RasG12D. This included restoration of WBC count and haemoglobin levels to CON levels (Figure 3.2A-B), reduction in spleen weight compared to KM (Figure 3.2D) and loss of blast cells found in the blood smears compared to KM (Figure 3.2F). However, platelet count was further reduced compared to CON (Figure 3.2E), and AKM still showed an increase in BM cellularity (Figure 3.2C).

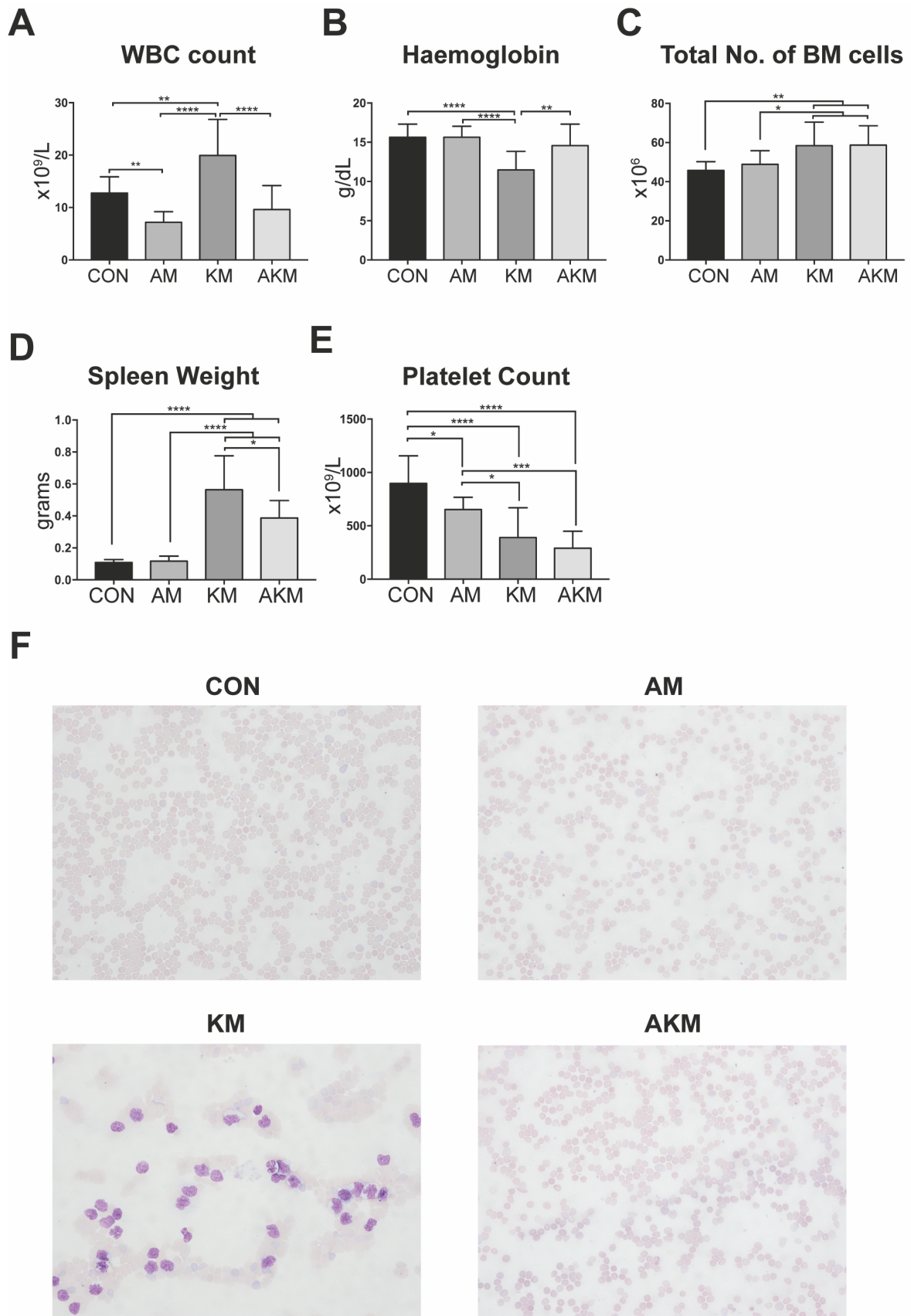


Figure 3.2 Aml1-ETO ameliorates key features of the MPN phenotype caused by K-RasG12D

(A) PB WBC count. (B) PB haemoglobin levels. (C) Total number of BM cells per tibia and femur. (D) spleen weight. (E) PB platelet count. (F) Microscopy of blood smears taken with a 40x magnification. Analysis of CON (n = 13), AM (n = 14), KM (n = 12), and AKM (n = 14). Results were generated in three independent experiments. The results were analysed using multiple comparison ANOVA and are presented as the mean $\pm$ SD.

*P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

To determine which mature cells were affected by Aml1-ETO and K-RasG12D, myeloid, B and T cells were analysed in the PB by flow cytometry (Figure 3.3A and 3.4A). AM had no difference in Mac1⁺Gr1^{lo} myeloid cells in the PB compared to CON. KM had an increase in Mac1⁺Gr1^{lo} myeloid cells in the PB compared to CON. There was a reduction in the expansion of Mac1⁺Gr1^{lo} myeloid cells caused by K-RasG12D when combined with Aml1-ETO (Figure 3.3B-C).

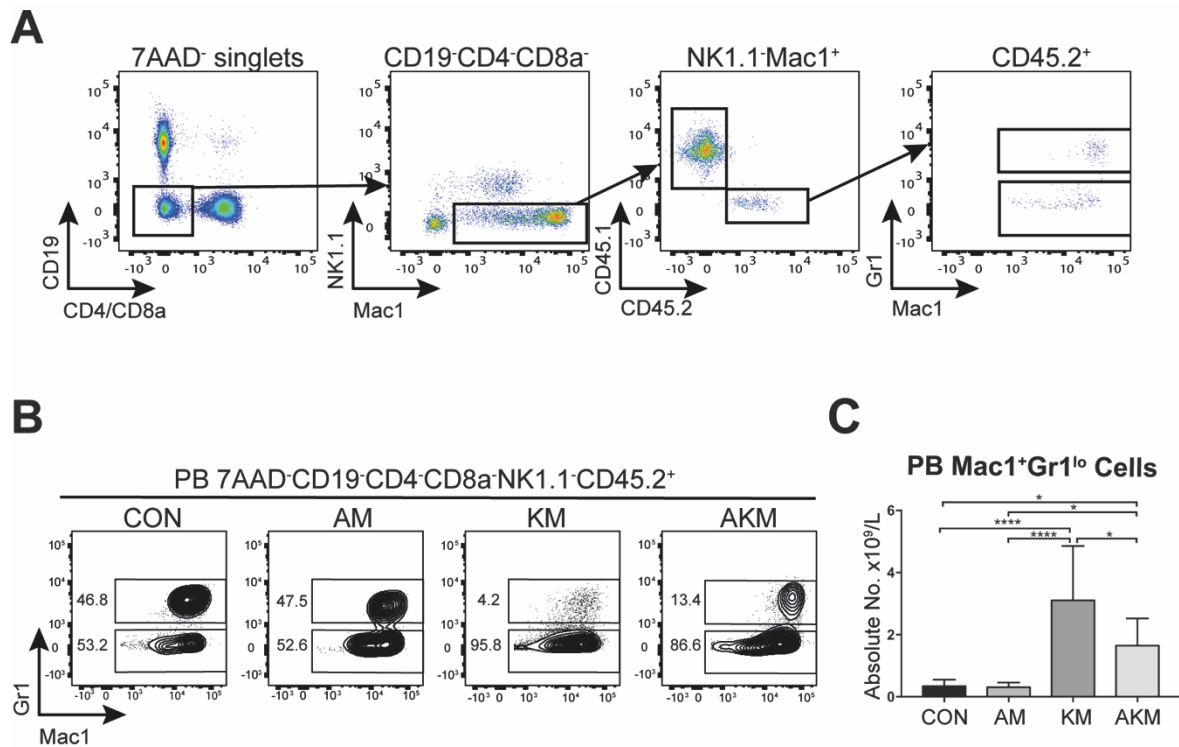


Figure 3.3 Aml1-ETO ameliorates the increase in Mac1⁺Gr1^{lo} myeloid cells in the PB caused by K-RasG12D

(A) Gating strategy for analysing myeloid cells in the PB. (B) Representative FACS plots of Mac1⁺Gr1⁺ and Mac1⁺Gr1^{lo} myeloid cells as a percentage of 7AAD⁻CD19⁻CD4⁻CD8a⁻ NK1.1⁺CD45.2⁺ cells across all experiments in the PB. (C) Mac1⁺Gr1^{lo} myeloid cells as an absolute number in the PB. Analysis of CON (n = 13), AM (n = 14), KM (n = 12), and AKM (n = 14). Results were generated in three independent experiments. The results were analysed using multiple comparison ANOVA and are presented as the mean ± SD.

*P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

When analysing the mature lymphoid cells in the PB, a decrease in B and T cells in the PB of AM compared to CON was identified (Figure 3.4B-C). This may explain the leukopenia seen in the WBC count (Figure 3.2A). KM had a decrease in B cells but no difference in T cells in the PB compared to CON (Figure 3.4B-C). AKM had a decrease in B and T cells in the PB compared to CON (Figure 3.4B-C).

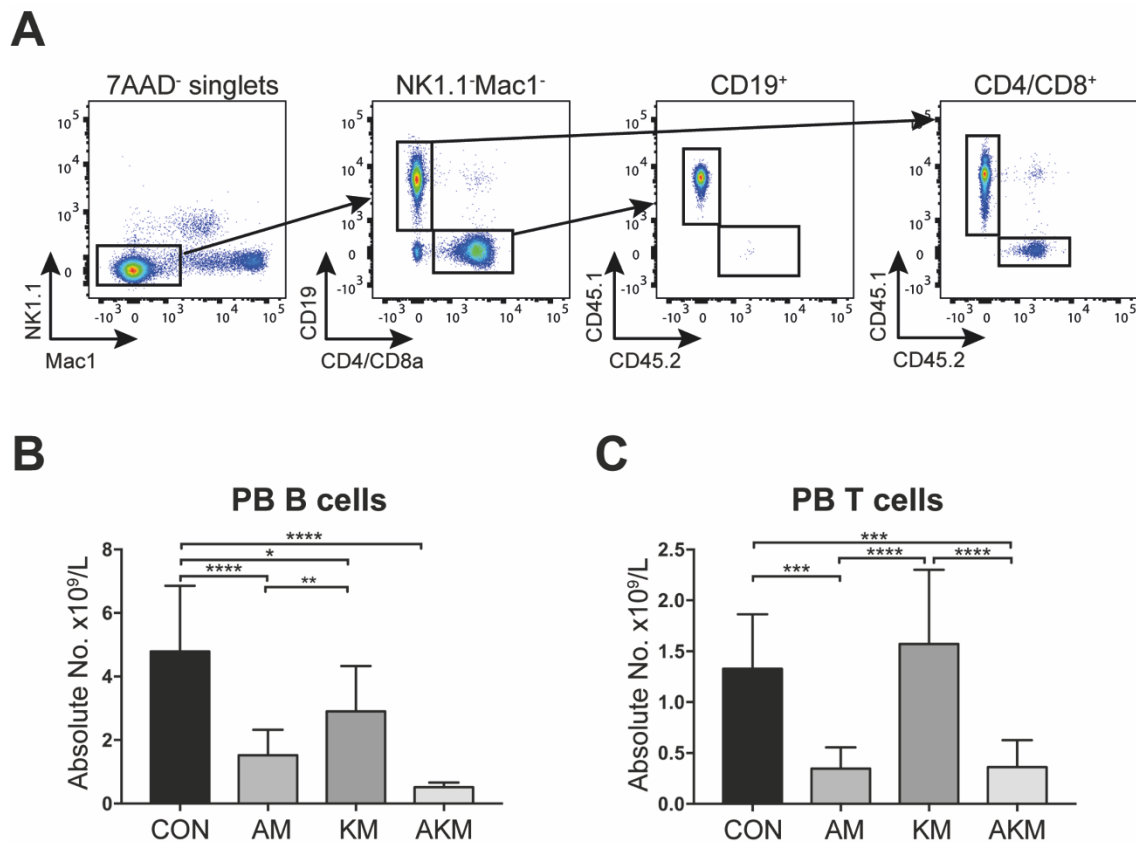


Figure 3.4 Aml1-ETO causes a decrease in lymphoid cells in the PB

(A) Gating strategy for analysing B and T cells in the PB. (B) B cells as an absolute number in the PB. (C) T cells as absolute number in the PB. Analysis of CON (n = 13), AM (n = 14), KM (n = 12), and AKM (n = 14). Results were generated in three independent experiments. The results were analysed using multiple comparison ANOVA and are presented as the mean $\pm$ SD. *P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

The myelo-erythroid progenitors were then analysed by flow cytometry in the BM and spleen (Figure 3.5A). Both AM (2-fold, $P = 0.0093$) and KM (8-fold, $P < 0.0001$) had a decrease in the number of MkPs, and when combined resulted in a further reduction (15-fold, $P < 0.0001$) in the BM compared to CON (Figure 3.5B). This paralleled the decrease in platelet counts found in the PB (Figure 3.2E). AM also had an increased number of MkPs in the spleen compared to CON (Figure 3.5C).

There was no difference in the number of BM pre-MegE, pre-CFU-E and CFU-E, in KM and AKM compared to CON (Figure 3.5D, F, H). However, KM has an increase in the pre-MegE, pre-CFU-E and CFU-E in the spleen compared to CON (Figure 3.5E, G, I). The increase in erythroid progenitors in the spleen in KM may indicate stress erythropoiesis. AKM also has an increase in pre-MegE, pre-CFU-E in the spleen, however the increase in CFU-E seen in KM is ameliorated in AKM (Figure 3.5E, G, I). This may suggest that the stress erythropoiesis was reversed in AKM at the CFU-E stage resulting in normal haemoglobin levels (Figure 3.2B).

There was no difference in pre-GM and GMPs in the BM of KM and AKM compared to CON (Figure 3.5J, L). In the spleen however there is an increase in KM pre-GMs and GMPs compared to CON (9-fold increase, $P < 0.0001$; Figure 3.5K, M). The increase in pre-GMs was reversed in AKM compared to KM (3-fold decrease, $P < 0.0001$; Figure 3.5K). The reduction of pre-GMs to CON levels in AKM may explain why the myeloproliferative phenotype in AKM mice is ameliorated, however there was still an increase in GMPs in the spleen compared to CON (Figure 3.5M).

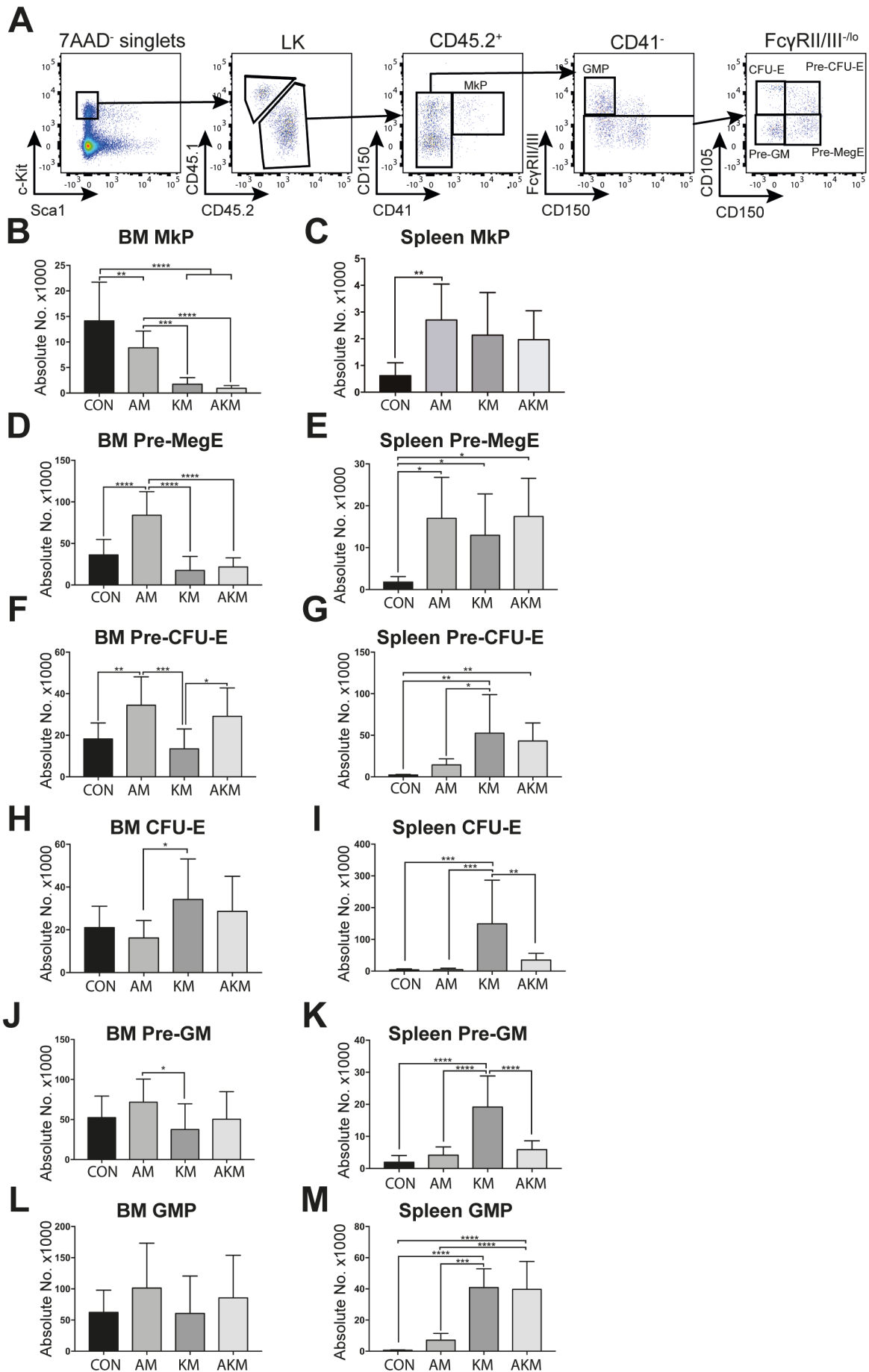


Figure 3.5 Aml1-ETO reverses some of the myelo-erythroid progenitor cell phenotypes caused by K-RasG12D

(A) Myelo-erythroid progenitor gating strategy. Absolute number of MkP (B), pre-MegE (D), pre-CFU-E (F), CFU-E (H), pre-GM (J), and GMP (L) in the BM from CON (n = 13), AM (n = 14), KM (n = 12), and AKM (n = 14). Results were generated in three independent experiments. Absolute number of MkP (C), pre-MegE (E), pre-CFU-E (G), CFU-E (I), Pre-GM (K), and GMP (M) in the spleen from CON (n = 8), AM (n = 10), KM (n = 9), and AKM (n = 10). Results were generated in two independent experiments, were analysed using multiple comparison ANOVA, and are presented as the mean $\pm$ SD. *P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

Together, these data demonstrates that when Aml1-ETO is co-expressed with K-RasG12D the myeloproliferative phenotype caused by K-RasG12D is ameliorated rather than enhanced. This is evidenced by the reduction of the WBC count, spleen weight, PB myeloid Mac1⁺Gr1^{lo} cells, spleen CFU-E and spleen pre-GMs and increase in haemoglobin in AKM compared to KM. Importantly, AKM mice did not become moribund at eight weeks post-poly(I:C) compared to KM mice.

3.2.2 K-RasG12D reverses the HSC expansion associated with Aml1-ETO

As the co-expression of Aml1-ETO and K-RasG12D was insufficient to induce acute leukaemic transformation, this provided an ideal model to study the impact of these mutations on pre-leukaemic HSCs. It was reasoned that the myeloproliferative phenotype may have been ameliorated when Aml1-ETO and K-RasG12D were co-expressed, due to loss of disease propagating HSCs. Subsequently phenotypic HSCs were analysed eight weeks post-poly(I:C). Phenotypic HSCs were defined as LSKFlt3⁻CD150⁺ (Figure 3.6A). The SLAM marker CD48 was not used as it was previously reported that CD48 expression is dysregulated in *Aml1* deficient HSCs (Cai et al., 2011).

There was an expansion of AM HSC compared to CON (5-fold increase, $P < 0.0001$; Figure 3.6B-C). There was no significant difference in KM HSC number compared to CON. However, the expansion caused by Aml1-ETO was reversed in AKM HSCs (3-fold decrease, $P < 0.0001$; Figure 3.6B-C).

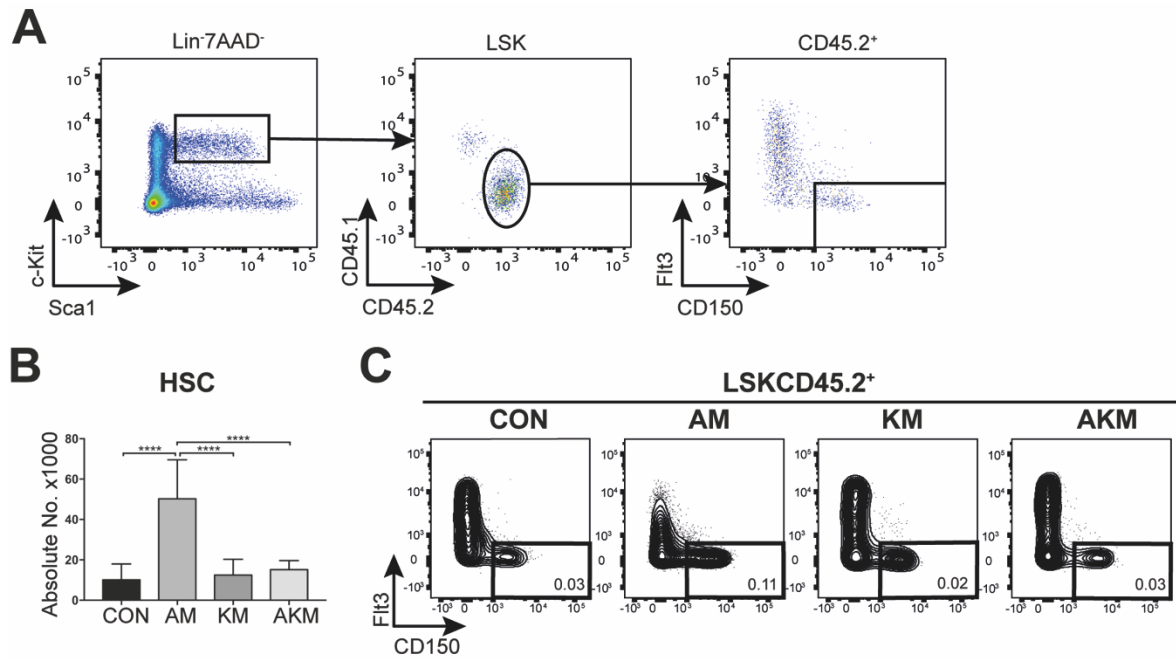


Figure 3.6 Phenotypic HSC expansion caused by Aml1-ETO is reversed by K-RasG12D

(A) HSC gating strategy. (B) HSCs as an absolute number in the BM. (C) Representative FACS plots showing gating used to quantify HSCs as a percentage of the BM mononuclear cells across all experiments. Analysis of CON (n = 11), AM (n = 14), KM (n = 13), and AKM (n = 14). Results were generated in three independent experiments. The results were analysed using multiple comparison ANOVA and are presented as the mean $\pm$ SD. *P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

To determine if K-RasG12D is detrimental to the function of Aml1-ETO-expressing HSCs, serial transplantations were performed. Secondary AM transplants showed an increase in myeloid reconstitution over time and a 10-fold increase in phenotypic HSC number compared to CON (Figure 3.7A-B), indicating a competitive advantage possibly due to an enhanced self-renewal capacity. However, AM enhanced self-renewal was not sustained in tertiary transplants (Figure 3.7A,C). KM and AKM secondary transplants showed a decreased myeloid reconstitution and HSC number (Figure 3.7A-C).

Furthermore, AM LSK cells showed increased replating potential *in vitro* (Figure 3.7D), in keeping with previous reports (Higuchi et al., 2002). In contrast, AKM abrogated this enhanced replating potential (Figure 3.7D).

Collectively, these results show Aml1-ETO expression is associated with increased self-renewal of HSCs, but in the presence of K-RasG12D, HSCs are at a competitive disadvantage, consistent with functional impairment of AKM HSCs.

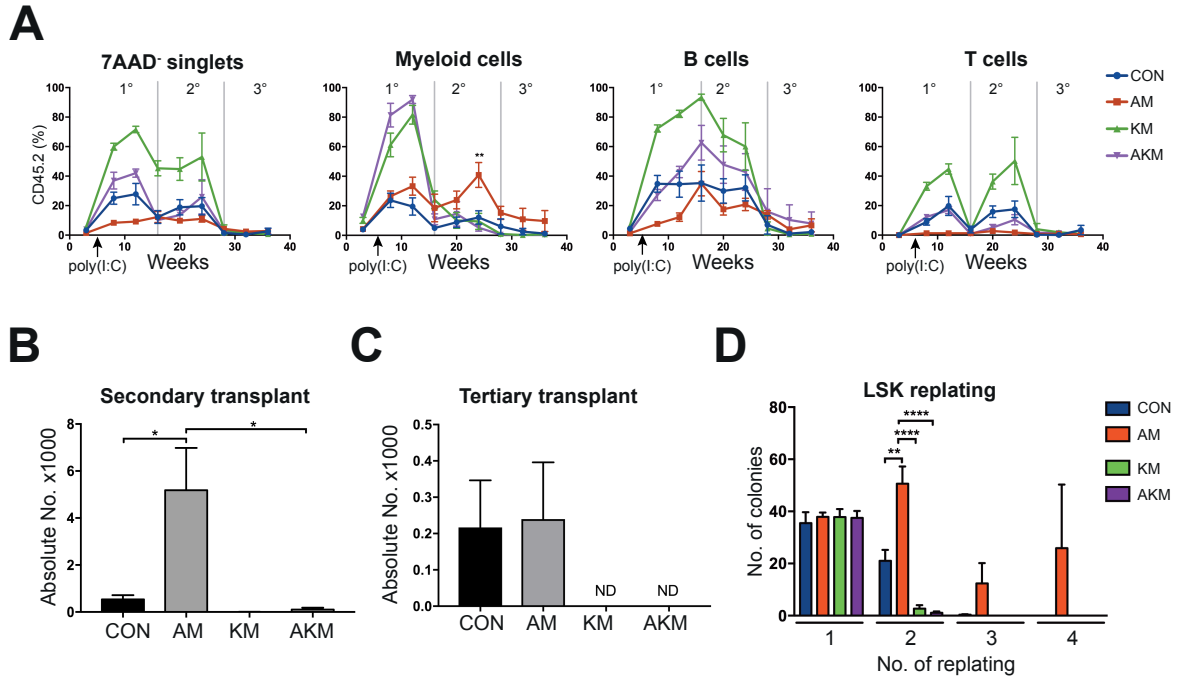


Figure 3.7 Functional HSC expansion caused by Aml1-ETO is reversed by K-RasG12D

(A) CD45.2 reconstitution as a percentage of 7AAD⁻ singlets, myeloid (7AAD⁻CD19⁻CD4⁻CD8a⁻NK1.1⁻), B cells (7AAD⁻NK1.1⁻Mac1⁻CD19⁺) and T cell (7AAD⁻NK1.1⁻Mac1⁻CD4⁺CD8a⁺) compartments in primary, secondary and tertiary transplantations. (B) Absolute number of CD45.2⁺ HSCs in secondary recipients of CON (n = 9 recipient mice in n = 2 independent experiments), AM (n = 10 recipient mice in n = 3 independent experiments), KM (n = 4 recipient mice in n = 2 independent experiments), and AKM (n = 7 recipient mice in n = 2 independent experiments). (C) Absolute number of CD45.2⁺ HSCs in tertiary recipients of CON (n = 5 recipient mice in n = 2 independent experiments), AM (n = 7 recipient mice in n = 3 independent experiments), KM (n = 3 recipient mice in n = 2 independent experiments), and AKM (n = 6 recipient mice in n = 2 independent experiments). (D) Replating efficiency of CD45.2⁺ LSK BM cells. Average number of colonies is shown for 5-6 biological replicates per genotype in two independent experiments. The results were analysed using multiple comparison ANOVA. The results are presented as the mean $\pm$ SEM. *P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

3.2.3 K-RasG12D induces loss of quiescence in Aml1-ETO HSCs

To gain molecular insight into the functional impairment of KM HSCs, RNA-sequencing was carried out of AM, KM, AKM and CON CD45.2⁺LSKCD150⁺Flt3⁻ cells from competitively transplanted recipients eight weeks post-poly(I:C) (n = 5-6 replicates per genotype). GSEA revealed a marked enrichment in E2F and Myc associated gene expression in AKM compared to AM, likely indicating higher levels of cell cycle activity (Figure 3.8A-B). This was accompanied by loss of HSC self-renewal-associated gene expression (Figure 3.8C), in keeping with the observed functional impairment of HSCs. Cell cycle activation was confirmed by flow cytometry, showing a 3-fold decrease in quiescent (G0) AKM HSCs, compared to AM HSCs (Figure 3.8D-E).

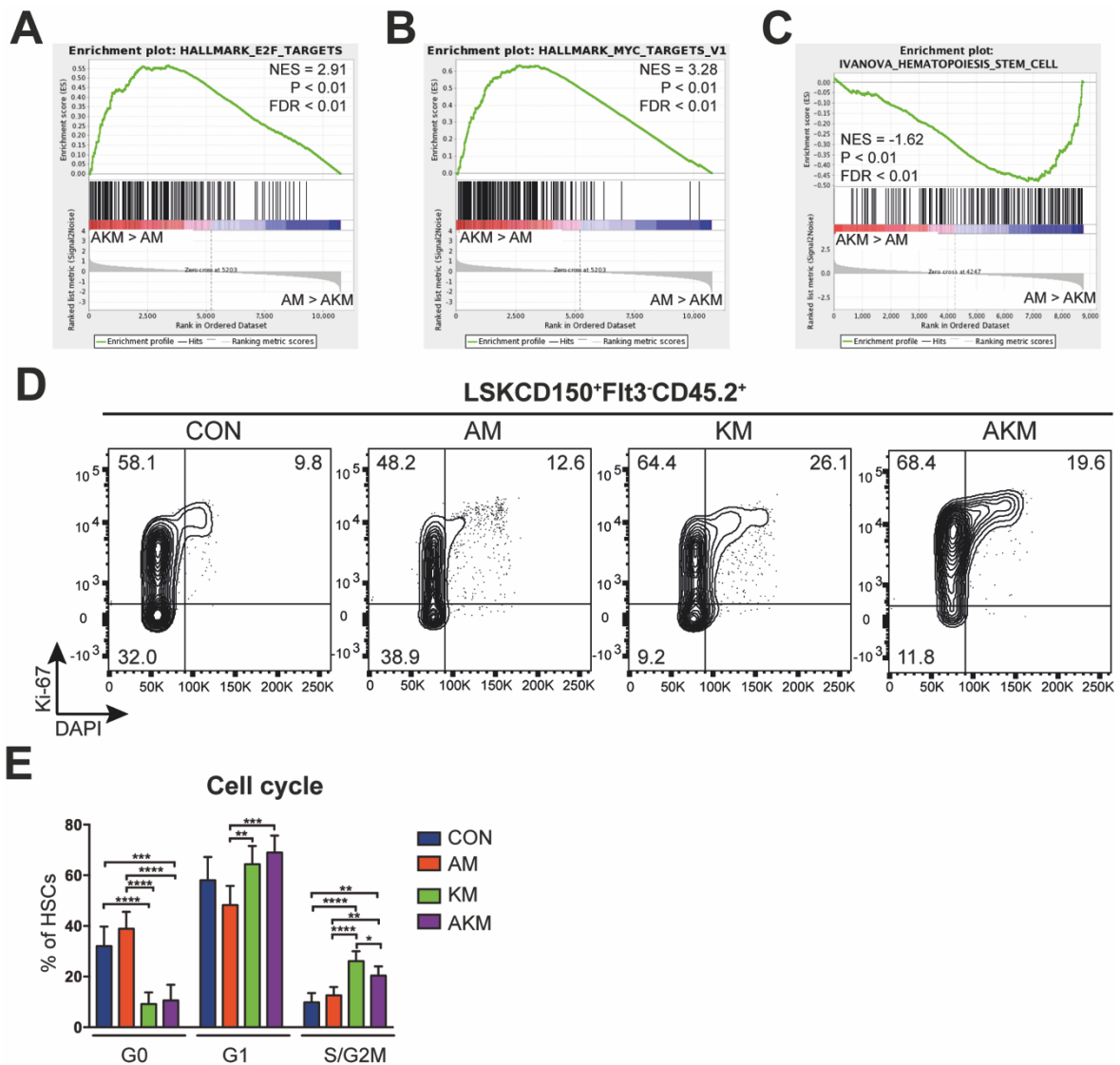


Figure 3.8 AKM HSCs are characterised by loss of quiescence and HSC-associated gene expression

(A-C) Bulk CD45.2⁺LSKCD150⁺Flt3⁻ cells were subjected to RNA-sequencing (5-6 biological replicates per genotype in two independent experiments). GSEA analysis of AKM vs. AM HSCs for E2F targets (A) Myc targets (B) and HSC gene signature (C). (D) Representative FACS plots showing cell cycle analysis of CD45.2⁺LSKCD150⁺Flt3⁻ phenotypic HSCs from CON (n = 5 recipient mice in n = 2 independent experiments), AM (n = 10 recipient mice in n = 3 independent experiments), KM (n = 5 recipient mice in n = 2 independent experiments) and AKM (n = 6 recipient mice in n = 3 independent experiments). (E) Percentage of BM HSCs at each cell cycle stage. The results were

analysed using multiple comparison ANOVA. The results are presented as the mean $\pm$ SD. *P<0.05 **P<0.01 ***P<0.001.

To identify genes that may be involved in the loss of quiescence and HSC function, differential gene expression analysis was performed. Aml1-ETO caused an up-regulation of 52 genes and down-regulation of 36 genes in phenotypic HSCs compared to CON (Figure 3.9A-B). Among the down-regulated genes were *Gja1* and *Gzmb* (Figure 3.9C-D). K-RasG12D caused more extensive disruption of gene expression, with up-regulation of 389 genes and down-regulation of 526 genes compared to CON (Figure 3.9A-B). Among the up-regulated genes were MAPK pathway genes such as *Etv4* and *Ccnd1* (Figure 3.9E-F), in keeping with activation of this signalling pathway by activated K-RasG12D (Pratilas et al., 2009).

AKM HSCs showed up-regulation of 319 genes and down-regulation of 482 genes compared to CON (Figure 3.9A-B). A large proportion of these genes were only regulated in the presence of both mutations (Figure 3.9A-B), indicating that the two mutations together collaborated to induce a distinct pattern of gene expression that only partially overlapped with K-RasG12D or Aml1-ETO regulated genes (Figure 3.9A-B, G). *Gja1* and *Gzmb* were up-regulated in AKM HSCs when compared to AM (Figure 3.9C-D, G).

Collectively, these results suggest K-RasG12D is detrimental to HSCs harbouring Aml1-ETO causing both a loss of HSC-associated gene expression and a loss of quiescence.

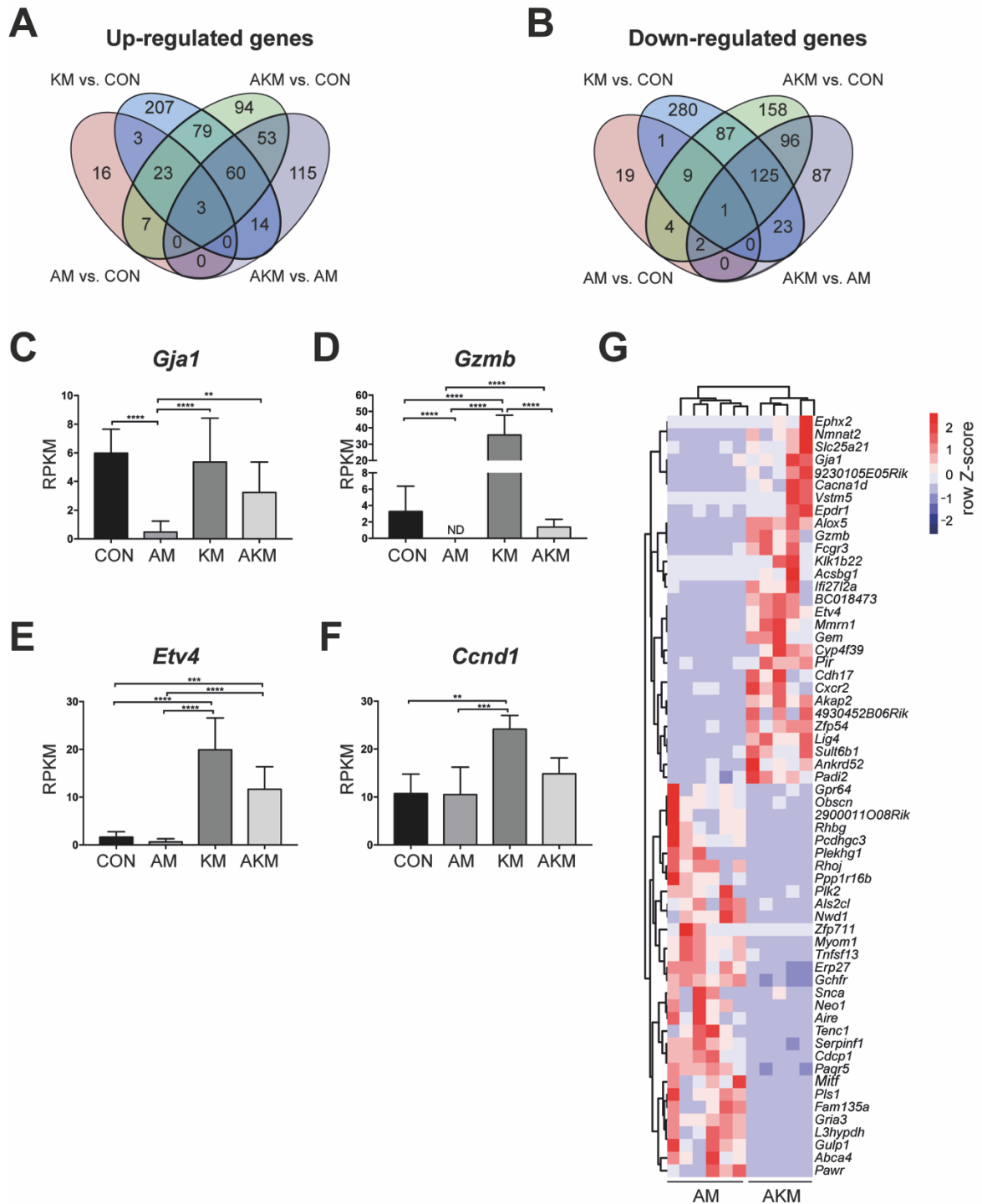


Figure 3.9 RNA-sequencing reveals distinct molecular signature in AKM HSCs

(A-B) Venn-diagram of significantly up-regulated (A) and down-regulated genes (B) in HSCs identified by RNA-sequencing. (C-F) RPKM of selected genes identified from RNA-sequencing. RPKM and FDR were generated using edgeR package. The results are presented as the mean $\pm$ SD. * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ **** $P < 0.0001$; (G) Heatmap

depicting the RPKM values of the top 30 significantly up-regulated and down-regulated genes in AKM HSCs compared to AM (FDR < 0.05).

3.3 Discussion

There is an observed absence of mutations in signalling pathway genes, such as *KRAS*, in pre-leukaemic HSCs from AML patients (Corces-Zimmerman et al., 2014; Shlush et al., 2014). It was hypothesised that the observed absence may be due to such mutations having a detrimental impact on not only normal, but also pre-leukaemic HSCs. To test this hypothesis conditional mouse genetics was used to study the pre-leukaemic fusion protein Aml1-ETO in combination with activated K-Ras. While Aml1-ETO improved the repopulating capacity of HSCs, K-RasG12D had a markedly detrimental effect on Aml1-ETO-expressing pre-leukaemic HSCs leading to their eventual depletion, likely due to loss of quiescence and HSC-associated gene expression. The loss of disease-propagating HSCs also likely underlies the amelioration of the K-RasG12D-induced myeloproliferative phenotype when the mutations were combined.

3.3.1 AM and KM recapitulate the phenotype in previous published data

AM mice recapitulated the main findings of previous published data on the *Aml1*^{ETO/+} mouse model. Higuchi *et al.*, demonstrated a dramatic serial replating ability, with AM cells replating up to 24 passages with ~47% of colonies that could generate new colonies when replated. This is consistent with this study's results, where ~36% of colonies generated new colonies when replated. However, only 4 passages were performed with 3/4 plates experiencing a decrease in colony number over time, therefore the replating ability was not as enhanced as that observed in the previous study. Consistently, there was no difference in RBCs, however Higuchi *et al.*, did not find any difference in

nucleated blood cell count or platelet count. This differs to this study's results as a decrease in WBC count and decrease in platelet count were found. The decrease in WBC count in AM may have resulted in a decrease in B and T cells which is consistent with Aml1 knock-out mice (Ichikawa et al., 2004). Aml1 null mice also have a defect in megakaryocyte maturation. This demonstrates that AM has similar features to the Aml1 knock-out model.

KM mice also recapitulates previous published data on the *Kras*^{G12D/+} mouse model, as mice develop a leucocytosis, anaemia, splenomegaly, increased number of GMPs in the spleen and expansion of the Mac1⁺Gr1^{lo} myeloid cells in the BM and spleen (Braun et al., 2004; I. T. Chan et al., 2004).

3.3.2 K-RasG12D is detrimental to normal and pre-leukaemic HSCs

This study complements the findings of Sabnis *et al.*, showing K-RasG12D is detrimental to normal HSCs (Sabnis et al., 2009). They showed KM HSCs, defined as LSKFlt3⁻, were less quiescent compared to WT with 50% of KM HSCs in G0 whereas there were 80% of WT HSCs in G0. This study showed 9% KM HSCs and 32% of WT HSCs in G0. There were less quiescent HSCs in both KM and WT HSCs in this study compared to Sabnis *et al.*, study. This may be due to the competitive transplant setting in this study, whereas Sabnis *et al.*, analysed primary non-transplanted mice. Sabnis *et al.*, also report a decrease in KM LSKFlt3⁻ HSCs two weeks after poly(I:C) compared to WT. In this study there was no difference in KM HSC number compared to WT. The HSCs were however analysed eight weeks post-poly(I:C). The decrease in HSCs seen by Sabnis *et al.*, may reflect the early time point analysed after poly(I:C) as it has previously been shown that acute poly(I:C) treatment can cause quiescent HSCs to proliferate, which diminishes the quiescent HSC compartment (Essers et al., 2009). Sabnis *et al.*, were not able to study the long-term fitness of KM HSCs due to lethality caused by K-RasG12D-induced

myeloproliferation. By performing serial transplant eight weeks post-poly(I:C) in this study it meant mice did not develop a MPN, and long-term fitness of KM HSCs could be analysed showing that HSCs are in fact lost in secondary transplantation.

This study also complements other previously published data on the negative impact that signalling mutations have upon HSCs. Previous studies have shown N-RasG12D increases cell division and reduces the self-renewal in a subset of HSCs. However, this negative impact is counteracted as N-RasG12D also increases the self-renewal potential and reduces division in another subset of HSCs. This bimodal effect allows NRasG12D-expressing HSCs to outcompete WT HSCs (Q. Li et al., 2013). *Ptpn11*^{D61Y/+} mouse model also shows a loss in phenotypic LT-HSCs with a reduction in G0 LSK. When performing bulk secondary transplant mice lose engraftment, indicating a loss of functional HSCs (G. Chan et al., 2009). Studies on Flt3-ITD also show a depletion of LT-HSCs (Chu et al., 2012). However as Flt3 is not expressed in the HSCs, a recent study has shown that the effect of Flt3-ITD on HSCs may be cell-extrinsic, delivered by disrupting HSC-supporting BM stromal cells and increasing inflammation-associated gene expression (Mead et al., 2017).

This study also demonstrates the negative impact K-RasG12D-expression has on pre-leukaemic HSCs harbouring Aml1-ETO. The HSCs expansion, enhanced self-renewal, and competitive advantage caused by Aml1-ETO-expression is lost when combined with K-RasG12D-expression.

The results in this study differ to other models which demonstrate that Ras mutations collaborate with pre-leukaemic mutation to develop AML, such as the *Dnmt3a*^{-/-};*Kras*^{G12D/+} (Chang et al., 2015) and the *Cbfb*^{MYH11/+};*NRas*^{G12D/+} (Xue, Pulikkan, Valk, & Castilla, 2014); wherein both models develop a more aggressive disease when the mutations are combined. Both models lead to transformation of the myeloid progenitor, but there is no

transformation seen in AKM, even after bulk serial transplantation. Both models do however result in a loss of LT-HSCs which is consistent with the results in this study. The lack of transformation of myeloid progenitors in AKM mice uniquely allowed the study of pre-leukaemic stem cells' functionality in the absence of progenitor cell transformation as seen in the other models.

The more aggressive phenotype seen in *Dnmt3a*^{-/-};*Kras*^{G12D/+} mice may be due to mice developing a T-ALL as well as an MPN. Both *Kras*^{G12D}-expression and loss of *Dnmt3a* have been previously shown to develop a T-ALL (Mayle et al., 2015; Sabnis et al., 2009). *Dnmt3a* loss recapitulates T-ALL better than AML as almost half of T-ALL patients have biallelic *DNMT3A* mutations (L. Yang, Rau, & Goodell, 2015), whereas heterozygous *DNMT3A* mutations are the most common type in AML (Kim et al., 2013). It may be of further investigative value to determine whether the *Dnmt3a*^{R878H/+} knock-in mouse would yield similar results in combination with *Kras*^{G12D/+}, as seen in the *Aml1*^{ETO/+};*Kras*^{G12D/+} combination, given that *Dnmt3a*^{R878H} results in expansion of the HSCs and enhances self-renewal capacity, but does not cause leukaemia (Guryanova et al., 2016).

Another potential consideration for the lack of transformation in AKM mice is the requirement of additional mutation for leukaemia development. It has been shown that t(8;21) patients have a higher number of acquired recurrent mutations compared to mutations acquired in other AML subtypes (8 mutations vs. 5 mutations) (Cancer Genome Atlas Research, 2013). *KRAS* mutations are also not the most frequent mutations in t(8;21) patients occurring in only 4% of t(8;21) patients. *KIT* is the most common mutation (~25%), and *NRAS* and *ASXL2* occur in ~20% of t(8;21) patients. This may reflect an enhanced cooperative effect of *KIT*, *NRAS* and *ASXL2* with AML1-ETO. Micol et al., combined heterozygous knock-out of *ASXL2* with *Aml1*^{ETO/+} and showed only 50% of mice became moribund with a long latency (average median survival of 301.5 days). These

results further illustrate the necessity of additional mutations for transformation with Aml1-ETO (Micol et al., 2017).

3.3.3 Identification of *Gzmb* and *Gja1* as important genes for HSCs

RNA-sequencing revealed genes that may underlie the observed HSC phenotype of AM and AKM mice. *Gzmb* is a serine protease that has recently been reported to be important in HSC function. Knock-out of *Gzmb* confers enhanced self-renewal to HSCs in a cell-intrinsic manner. *Gzmb* deficient mice also had a better survival rate after administering 5-FU (Carnevali et al., 2014).

Gja1 encodes gap junction channel protein connexin 43 found on HSCs. *Gja1* deficient HSCs were shown to be more quiescent after 5-FU treatment. *Gja1* deficient HSCs also developed an accumulation of reactive oxygen species (ROS) (Taniguchi Ishikawa et al., 2012). ROS levels in HSCs have been shown to play an important role in haematopoietic reconstitution (Lewandowski et al., 2010). CD34+ BM cells from AML patients with *AML1-ETO* had a lower expression of *GJA1* compared to WT CD34+ BM cells (El-Hadya, Almasryc, & ElHefneyb, 2012).

Together down-regulation of *GZMB* and *GJA1* may be important for the *AML1-ETO* associated competitive advantage of HSCs, and could potentially be important for HSC persistence after chemotherapy. Importantly, expression of both *Gzmb* and *Gja1* were increased in the presence of K-RasG12D, indicating that re-expression of these genes may have contributed to the loss of HSC function and self-renewal.

3.3.4 Conclusion

K-RasG12D-expression is detrimental to both normal and pre-leukaemic HSCs causing loss of phenotypic and functional HSCs, loss of quiescence and HSC-associated gene expression. Down-regulation of *Gzmb* and *Gja1* may be important for pre-leukaemic HSC expansion and re-expression by K-RasG12D could cause loss of HSC quiescence.

Mutations are acquired in a stepwise manner in AML, and the consequence of the type and order of the mutations acquired will make the HSC more or less likely to facilitate subsequent evolution to leukaemia (Grove & Vassiliou, 2014). These findings help to provide a cellular and molecular basis for the observed patterns of clonal evolution during AML development. HSCs that acquire *Aml1-ETO* gain a competitive advantage, leading to an expansion in HSC number, increasing the pool of cells available to acquire additional mutation that could eventually promote leukaemia development. HSCs that acquire *Kras* mutation, either alone or in combination with *Aml1-ETO*, are depleted. These findings help to explain why signalling mutations such as *KRAS* are not observed within pre-leukaemic HSC in AML patients and usually occur as a late event in leukaemogenesis.

To further dissect the importance of temporal and cellular context of mutation acquisition in AML, a dual-recombinase system could be used (Schonhuber et al., 2014). Using *Mx1-Cre* to activate both *Aml1-ETO* and K-RasG12D expression means both mutations are expressed simultaneously in the HSC. To determine if there is a different effect if K-RasG12D was expressed after *Aml1-ETO* a different recombinase could be used to activate K-RasG12D expression in the HSC.

It would also be interesting to determine if the re-expression of *Gzmb* or *Gja1* were responsible for the loss of quiescence in AKM HSCs. This could be achieved by knock-

out of *Gzmb* or *Gja1* in AKM FL cells using CRISPR-Cas9 technology and performing a competitive transplant to determine if loss of *Gzmb* and *Gja1* restores HSC reconstitution. Alternatively, *Gzmb* or *Gja1* could be over-expressed in AM FL cells and a competitive transplant performed to determine if over-expression leads to loss of HSC enhanced self-renewal. Only a few genes such as *HOXA* gene cluster have been shown to possess shared dysregulation by pre-leukaemic mutations (Abdel-Wahab et al., 2012; Abdel-Wahab et al., 2013; Krivtsov et al., 2006; Yan et al., 2011). It would be interesting to see if *Gzmb* or *Gja1* are also dysregulated by other pre-leukaemic mutations.

CHAPTER FOUR: A study of the biallelic *Cebpa* mutations in combination with a novel zinc finger 1 *Gata2* mutation

4.1 Introduction

Sequencing studies in AML patients have shown patterns of co-mutations between functional groups, such as the association between transcription factor fusions and activated signalling, and tumour suppressors with DNA methylation (Papaemmanuil et al., 2016). In addition, very specific associations between mutations have been identified such as the association between biallelic *CEBPA* and in ZF1 *GATA2* mutations, where the *GATA2* mutation occurs in ~30% of AML patients with the biallelic *CEBPA* (Dohner et al., 2017; Fasan et al., 2013; Green et al., 2013; Greif, Dufour, et al., 2012; Marceau-Renaut et al., 2015; Papaemmanuil et al., 2016). This association has also been observed in the erythroleukaemia subtype of AEL which is characterised by bi-lineage transformation of both the myeloid and erythroid lineages (Ping et al., 2017). These observations suggest that the biallelic *CEBPA* and ZF1 *GATA2* mutations synergise during leukaemogenesis, and in particular the combination may contribute to both a myeloid and erythroid leukaemia. Myeloid leukaemias are generated through transformation of the neutrophil granulocyte lineages (Goardon et al., 2011; Kirstetter et al., 2008; Krivtsov et al., 2006). Murine studies have shown neutrophil differentiation occurs through the neutrophil-monocyte progenitor (NMPs), where *Gata2* expression is virtually absent (Drissen et al., 2016). Conversely, erythroid lineage progenitors express high levels of *Gata2* but lack *Cebpa* expression (Pronk et al., 2007). This raises the questions: 1) Which cell type is transformed in leukaemias where both biallelic *CEBPA* and *GATA2* are mutated? 2) If a bi-lineage leukaemia develops is it maintained by a single bi-potent, or by two distinct lineage-restricted leukaemia propagating cell populations?

To determine if biallelic *CEBPA* and ZF1 *GATA2* mutations synergise during leukaemogenesis, and in particular induce an AEL, a genetically accurate model of the combined mutations was generated using the well-established biallelic *Cebpa* knock-in mouse model (Bereshchenko et al., 2009) in combination with a novel ZF1 point mutation knock-in *Gata2*G320D, which is mutated in both AML and AEL (Fasan et al., 2013; Greif, Dufour, et al., 2012; Hou et al., 2015; Ping et al., 2017).

4.2 Results

4.2.1 Generation of a genetic model of *Cebpa* and *Gata2* mutant AML

To determine if the biallelic *Cebpa* and *Gata2* mutations synergise during leukaemia development the previously described N-terminal *Cebpa* knock-in mutation [*Cebpa*^L (Kirstetter et al., 2008)] and C-terminal *Cebpa* knock-in mutation [*Cebpa*^K (Bereshchenko et al., 2009)] were combined with a heterozygous ZF1 *Gata2* mutation (*Gata2*^{G320D/+}). To avoid the perinatal lethality of the biallelic *Cebpa* mutations, E14.5 FL cells were generated with the necessary genotypes: *Cebpa*^{+/+};*Gata2*^{+/+} (WT) *Cebpa*^{+/+};*Gata2*^{G320D/+} (G genotype), *Cebpa*^{K/L};*Gata2*^{+/+} (KL genotype), and *Cebpa*^{K/L};*Gata2*^{G320D/+} (KLG). A competitive transplant was performed using 2.5x10⁵ experimental FL cells (CD45.2 allotype) combined with 2.5x10⁵ CD45.1 or CD45.1/2 WT competitor cells into lethally irradiated recipients (CD45.1/2 allotype). Where indicated CD45.1 or CD45.1/2 allotypes were combined with the *Gata1*-EGFP transgene (Drissen et al., 2016), that efficiently labels platelets and erythroid cells (Carrelha et al., 2018), allowing experimental erythroid lineage cells (EGFP-) to be distinguished from competitor and recipient WT erythroid cells (EGFP+) (Figure 4.1). PB was analysed every four weeks up to 24 weeks post-transplantation and terminal analysis was performed when mice had a RBC count was ≤

$7 \times 10^{12}/L$, WBC count was $\geq 15 \times 10^9/L$ and were visibly hunched/pale, or at 52 weeks post-transplantation (Figure 4.1).

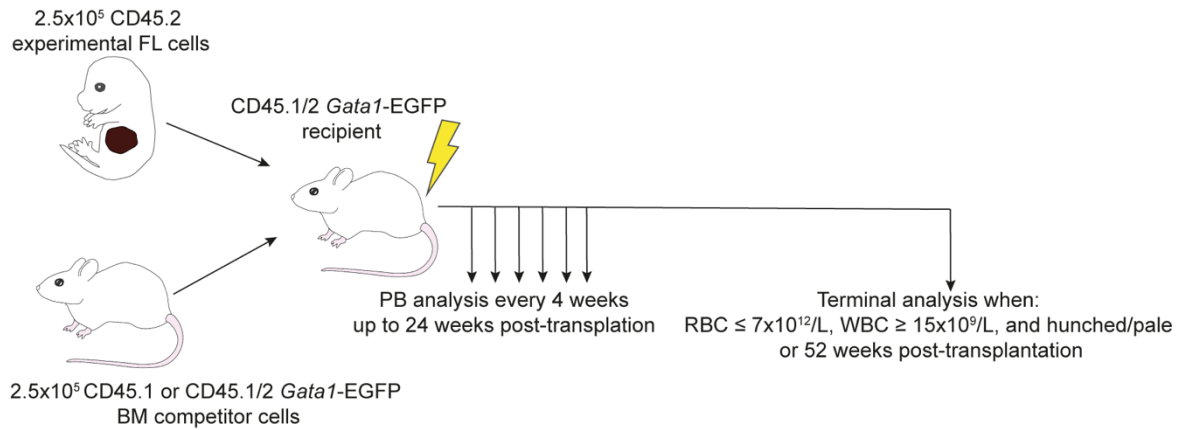


Figure 4.1 Experimental design

Schematic of *in vivo* competitive transplantation experiment. 2.5×10^5 CD45.2 experimental FL cells were co-transplanted with 2.5×10^5 either CD45.1 or CD45.1/2 *Gata1*-EGFP BM competitor cells into lethally irradiated (2×500 rads) CD45.1/2 *Gata1*-EGFP recipient mice. PB analysis was performed every four weeks up to 24 weeks post-transplantation. Mice were culled for terminal analysis when the RBC count was $\leq 7 \times 10^{12}/L$, WBC count was $\geq 15 \times 10^9/L$ and mice were physically hunched and pale, or mice were culled at 52 weeks post-transplantation.

4.2.2 Biallelic *Cebpa* and ZF1 *Gata2* mutations synergise during leukaemogenesis

Reconstitution was analysed for the CD45.2, myeloid, B, and T cells in the PB over time, as previously described (Figure 3.3A and 3.4A) as well as analysis of the erythrocytes and platelets (Figure 4.2A-B). A comparable overall reconstitution of PB CD45.2⁺ cells by all four genotypes prior to the onset of leukaemia development, except for at eight and 16 weeks post-transplantation of KL compared to KLG (Figure 4.2C). KLG showed an increase in myeloid reconstitution overtime compared to KL (Figure 4.2D) with no

significant difference in B and T cell contribution in KLG compared to KL (Figure 4.2E-F).

In addition, both KL and KLG showed a more rapid initial reconstitution of erythrocytes (Figure 4.2G) but not of platelets (Figure 4.2H).

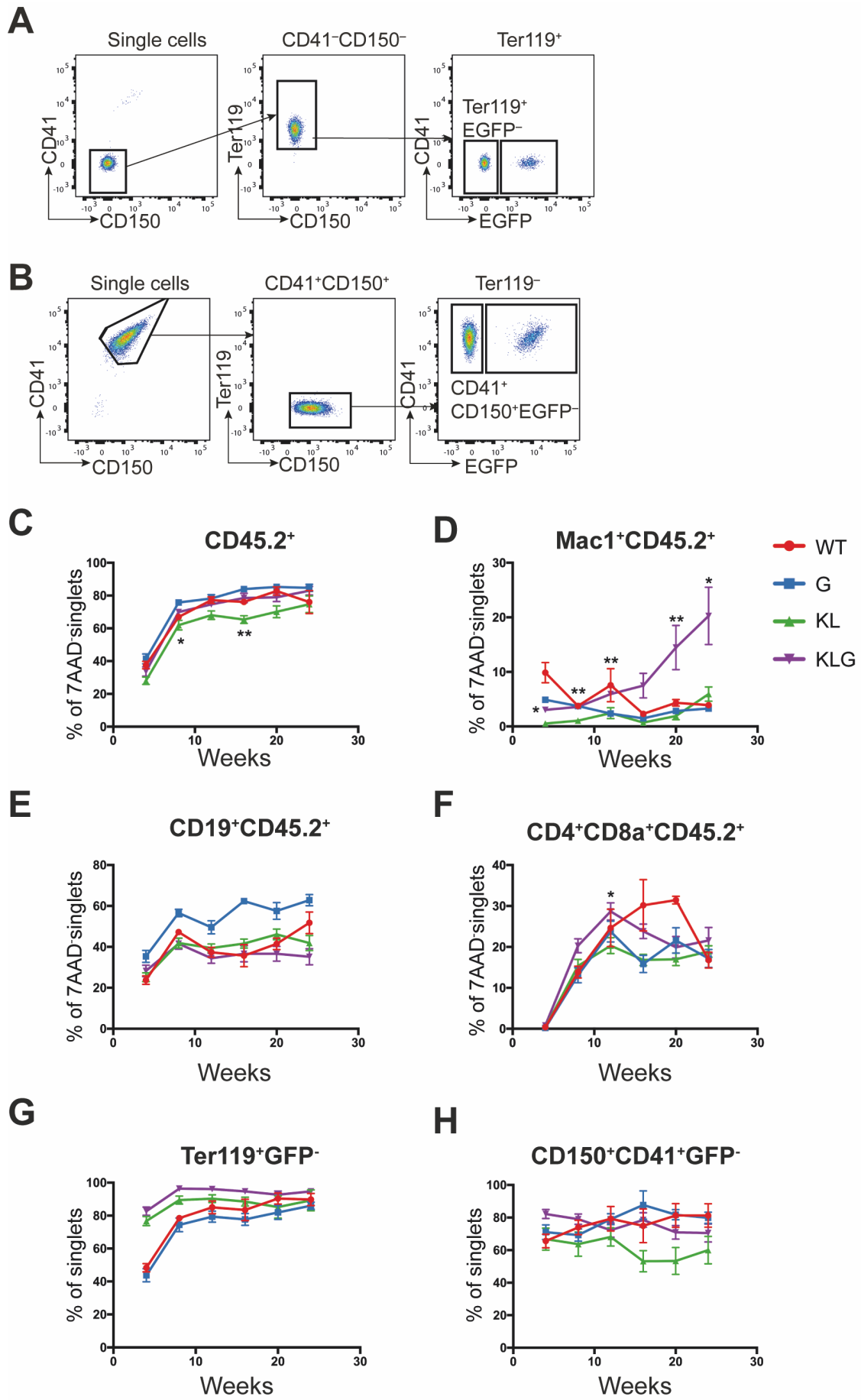


Figure 4.2 Biallelic *Cebpa* combined with *Gata2* mutation causes an increase in myeloid reconstitution overtime in the PB

(A, B) Gating strategy erythrocytes (A) and platelets (B). PB reconstitution over time in donor derived cells. Total CD45.2⁺ cells (C), myeloid cells (Mac1⁺; D), B cells (CD19⁺; E), and T cells (CD4⁺CD8⁺; F) as a percentage of 7AAD⁻singlets. Analysis of WT (n = 9 recipient mice), G (n = 9 recipient mice), KL (n = 13 recipient mice) and KLG (n = 18 recipient mice) from 2-4 independent experiments. (G-H) Experimental (EGFP-) reconstitution of erythrocytes (Ter119⁺; G) and platelets (CD150⁺CD41⁺; H) as a percentage of singlets. Analysis of WT (n = 9), G (n = 7), KL (n = 11), and KLG (n = 14) from 2-3 independent experiments. The results are presented as the mean $\pm$ SEM. The results were analysed using a multiple comparison ANOVA. Statistical significance of KLG vs KL comparison are shown. *P<0.05 **P<0.01.

Consistent with a higher myeloid lineage output, KLG mice developed a more rapid disease, with mice becoming moribund starting at five months post-transplantation with an average latency of eight months, while KL mice develop a leukaemia on average two months later with an average latency of 10 months ($P = 0.030$; Figure 4.3A). No leukaemia was observed in WT or G mice. Moribund mice were characterised by an elevated WBC count, anaemia shown by a decrease in RBC count, thrombocytopenia shown by a decrease in platelet count, and splenomegaly, all consistent with AML (Figure 4.3B-E).

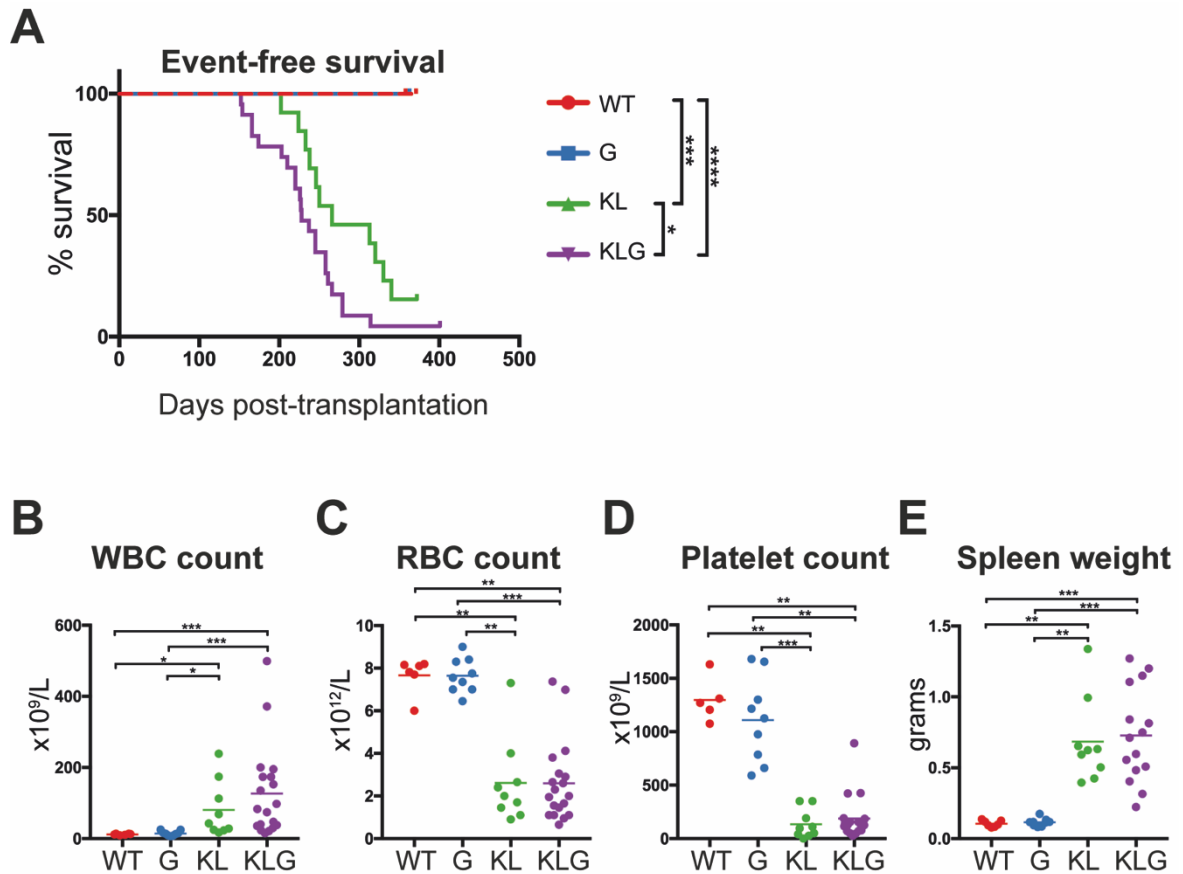


Figure 4.3 Biallelic *Cebpa* synergises with *Gata2* mutation to cause a more rapid disease development consistent with AML

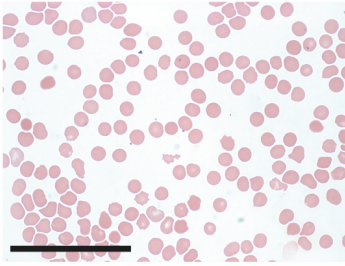
(A) Event-free survival. Survival curve was analysed using a log-rank (Mantel-Cox) test. (B) WBC count. (C) RBC count. (D) Platelet count. (E) Spleen weight. (B-E) Parameters were measured during terminal analysis. Analysis of WT (n = 9), G (n = 9), KL (n = 13), and KLG (n = 18) from 2-4 independent experiments. A D'Agostino & Pearson normality test was first used to determine if data fell into a normal distribution. If data did not have a normal distribution then a multiple comparison Kruskal-Wallis test was performed. If the data had a normal distribution then a multiple comparison ANOVA was performed. The results are presented as the mean. *P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

4.2.3 Mice with combined biallelic *Cebpa* and ZF1 *Gata2* mutation develop an erythroleukaemia

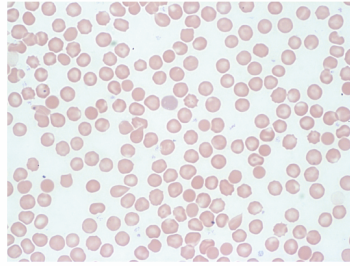
Examination of PB smears from leukaemic mice showed the presence of leukaemia blasts. However, while KL blasts were consistent with a myeloid phenotype, 38% (5/13) of KLG mice contained both myeloid and erythroid blasts in the PB (termed KLG-E mice), while the remaining mice showed only myeloid blasts in the PB (termed KLG-M mice) (Figure 4.4A). The same pattern was also observed in the bone marrow (Figure 4.4B) and spleen (Figure 4.4C).

A

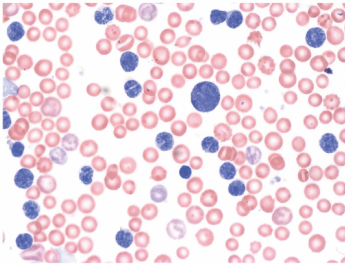
WT



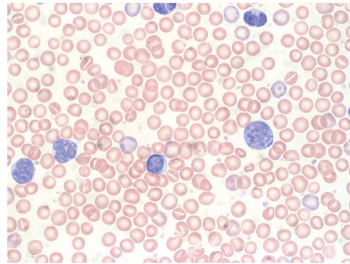
G



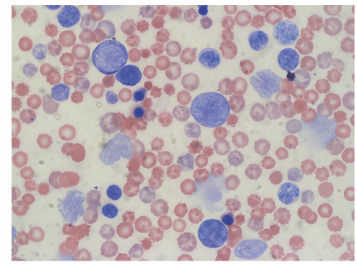
KL



KLG - M

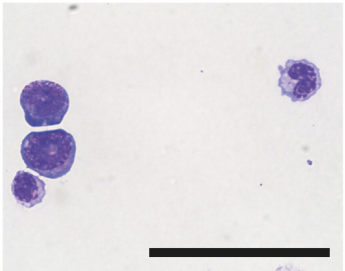


KLG - E

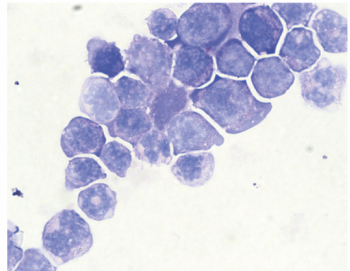


B

KLG - M

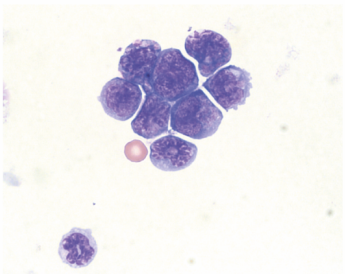


KLG - E



C

KLG - M



KLG - E

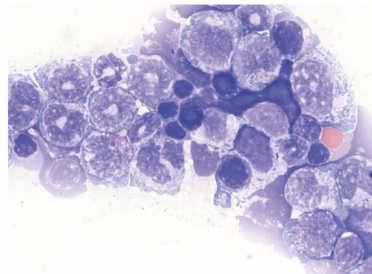


Figure 4.4 Myeloid and erythroid blast morphology in ~40% of KLG cases

(A) PB smears from terminal analysed mice. (B) BM cytospin and (C) spleen cytospin of terminal analysed mice. Analysis of three replicates per genotype from 2-4 independent experiments. Scale bar, 50 μ M.

Analysis of myeloid cells by flow cytometry (Figure 4.5A) showed an expansion of mutant CD45.2 myeloblasts (c-Kit^{lo}Mac1⁺) and in particular immature c-Kit⁺Mac1^{lo} cells in all KL and KLG leukaemic mice in the BM (Figure 4.5B-C) and spleen (Figure 4.5D-E).

However, it should be noted that not all the c-Kit^{lo}Mac1⁺ cells will be myeloblasts as some more mature cells were seen in the BM and spleen cytospins (Figure 4.4B-C).

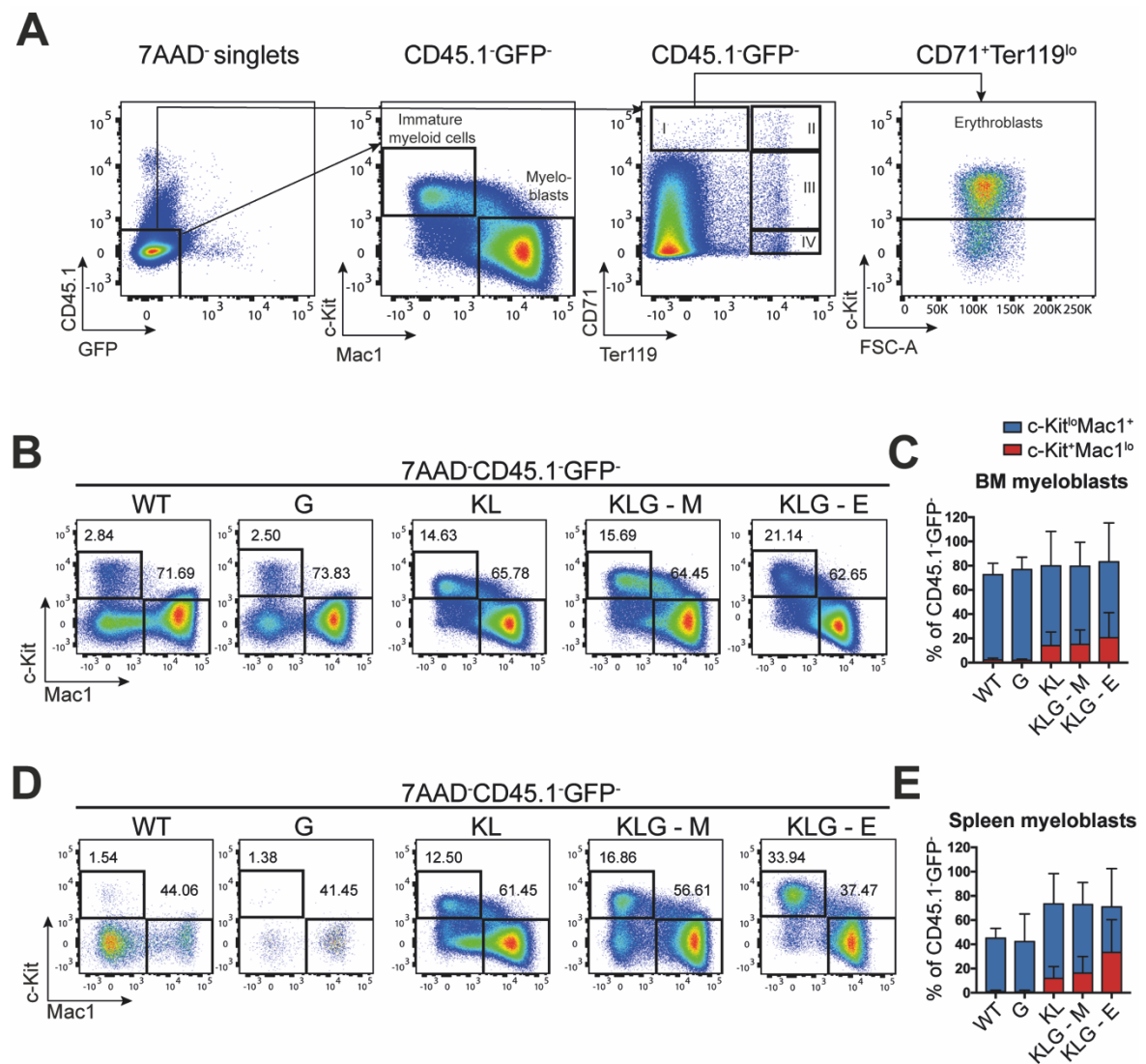


Figure 4.5 KL and KLG leukaemic mice contain myeloblasts and an expansion of the immature c-Kit^{lo}Mac1^{lo} cells in the BM and spleen

(A) Gating strategy for myeloid and erythroid lineage. CD45.1⁺ and GFP⁺ cells were gated out to remove all competitor and recipient cells, including erythroid lineage competitor and recipient cells that would be GFP⁺. Remaining cells are donor derived cells. (B) Representative FACS plots of experimental myeloid cells in the BM as a percentage of BM 7AAD⁻CD45.1⁻GFP⁻ cells across all experiments. (C) c-Kit^{lo}Mac1⁺ and c-Kit⁺Mac1^{lo} cells as a percentage of CD45.1⁻GFP⁻ cells in the BM. (D) Representative FACS plots of experimental myeloid cells in the spleen as a percentage of spleen 7AAD⁻CD45.1⁻GFP⁻ cells across all experiments. (E) c-Kit^{lo}Mac1⁺ and c-Kit⁺Mac1^{lo} cells as a percentage of CD45.1⁻GFP⁻ cells in the spleen. Analysis of WT (n = 7), G (n = 9), KL (n = 3), KLG-M (n = 8), KLG-E (n = 5) from 2-3 independent experiments.

Analysis of the different erythroid differentiation stages were also performed using flow cytometry (Socolovsky et al., 2001) (Figure 4.5A). KL and KLG-M mice showed a loss of all four erythroid stages in the BM and spleen (Figure 4.6A-D). However, in KLG-E mice there was an expansion of the stage I pro-erythroblasts in the BM (19-fold; $P = 0.001$) and to an even greater extent in the spleen compared to WT (73-fold; $P = 0.003$) (Figure 4.6A-D). These immature erythroblasts also expressed high levels of c-Kit in the BM and spleen (Figure 4.6E-H). There was also a loss of stage III-IV erythroid progenitors in KLG-E mice in the BM and spleen (Figure 4.6A-D) suggesting the accumulation of leukaemic stage I pro-erythroblasts were blocked in differentiation. Comparing the survival of KLG-M and KLG-E mice showed KLG-E mice developed a more aggressive disease than KLG-M mice ($P = 0.02$; Figure 4.7).

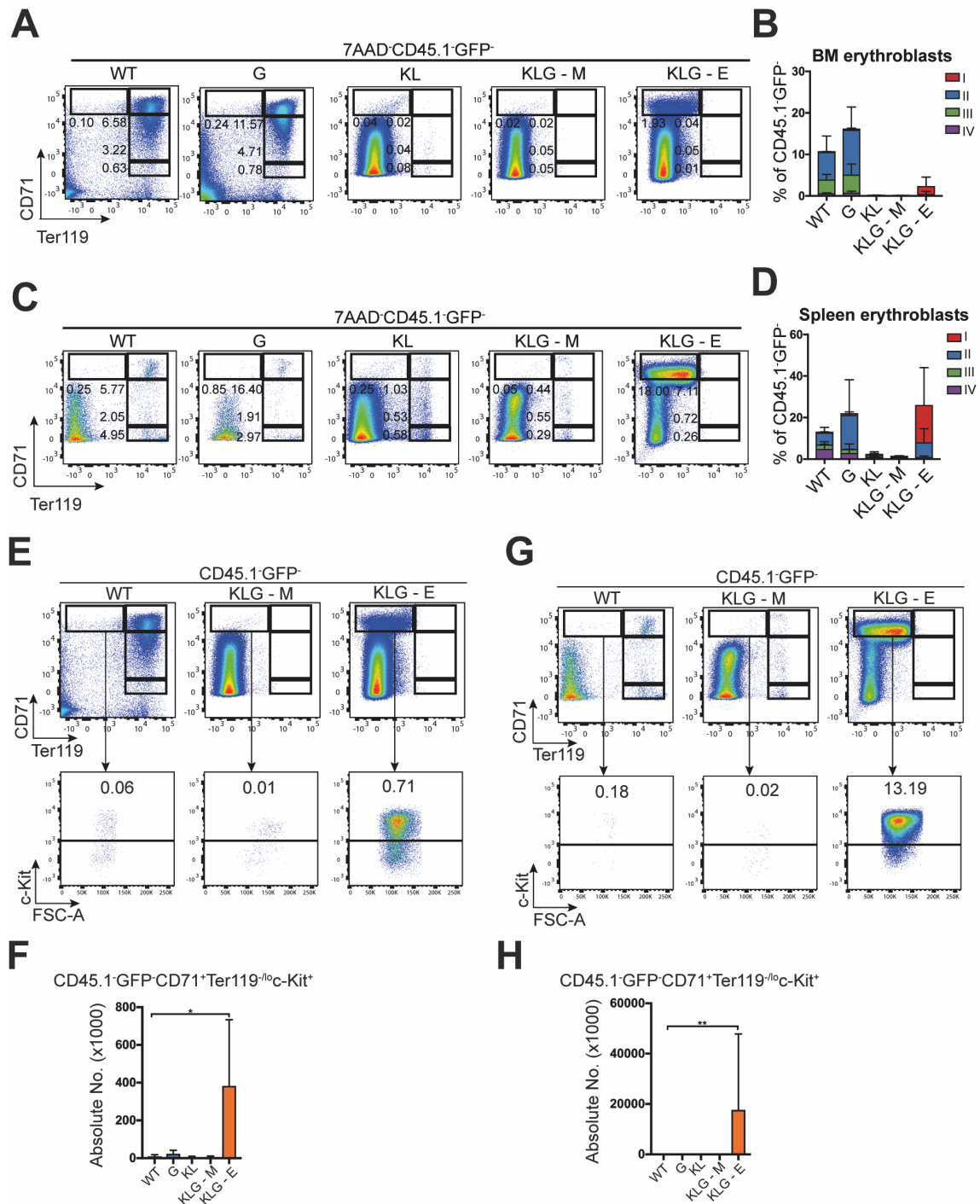


Figure 4.6 KLG-E mice contain erythroblasts in the BM and spleen

(A) Representative FACS plots of experimental stage I-IV erythroblast cells in the BM as a percentage of BM 7AAD⁻CD45.1⁻GFP⁻ cells across all experiments. (B) Stage I-IV erythroblast cells as a percentage of CD45.1⁻GFP⁻ cells in the BM. (C) Representative FACS plots of experimental stage I-IV erythroblast cells in the spleen as a percentage of spleen 7AAD⁻CD45.1⁻GFP⁻ cells across all experiments. (D) Stage I-IV erythroblast cells

as a percentage of CD45.1⁻GFP⁻ cells in the spleen. (E) Representative FACS plots of c-Kit⁺ stage I pro-erythroblasts in the BM as a percentage of BM 7AAD⁻CD45.1⁻GFP⁻ CD71⁺Ter119^{-/lo} cells across all experiments. (F) c-Kit⁺ stage I pro-erythroblasts as an absolute number in the BM. (G) Representative FACS plots of c-Kit⁺ stage I pro-erythroblasts in the spleen as a percentage of spleen 7AAD⁻CD45.1⁻GFP⁻CD71⁺Ter119^{-/lo} cells across all experiments. (F) c-Kit⁺ stage I pro-erythroblasts as an absolute number in the spleen. Analysis of WT (n = 7), G (n = 9), KL (n = 3), KLG-M (n = 8), KLG-E (n = 5) from 2-3 independent experiments. The results were analysed using a Mann-Whitney U-test. The results are presented as the mean ± SD. *P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

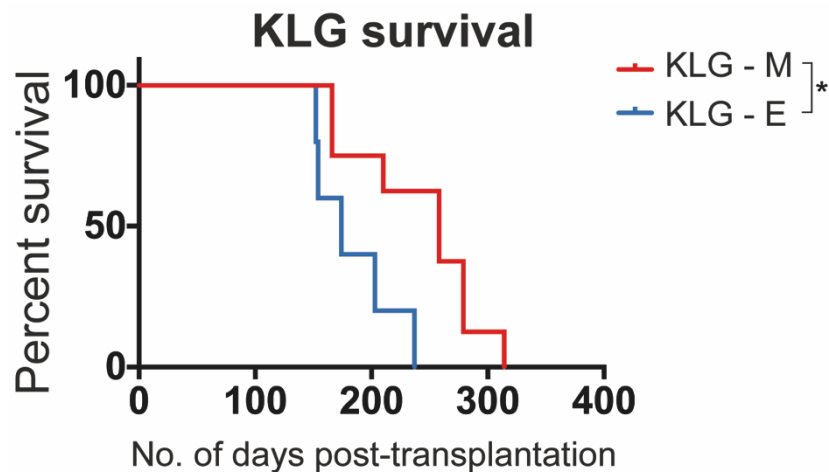


Figure 4.7 KLG-E mice develop a more aggressive disease than KLG-M mice

Analysis of KLG-M (n = 8) and KLG-E (n = 5). Survival curve was analysed using a log-rank (Mantel-Cox) test. *P<0.05.

Transplantation of bulk KLG-M leukaemia BM cells into lethally irradiated recipients generated a pure myeloid leukaemia within eight weeks (Table 4.1; Figure 4.8A-H); whereas mice transplanted with bulk KLG-E leukaemia cells developed a leukaemia faster, an average of five weeks latency, and accumulated both erythroid and myeloid blasts in the BM and spleen (Figure 4.8A-H), replicating the original disease phenotype.

Together these data suggests that ~60% of KLG mice develop a pure myeloid leukaemia while the remaining ~40% of KLG mice develop an AEL of the erythroleukaemia subtype displaying a bi-lineage phenotype with both myeloid and erythroid blasts in the PB, BM, and spleen. The erythroleukaemia is transplantable and more aggressive than AML.

Table 4.1 Bulk transplantation of leukaemic KLG-M and KLG-E cells

Sample No.	Phenotype	No. of cells	Developed leukaemia? (Y/N)	No. of weeks latency
97	KLG - M	750,000	Y	8
99	KLG - M	750,000	Y	8
136	KLG - M	500,000	Y	4
137	KLG - M	500,000	Y	4
138	KLG - M	500,000	Y	17
100	KLG - E	750,000	Y	6
101	KLG - E	750,000	Y	8
102	KLG - E	750,000	Y	7
121	KLG - E	500,000	Y	5
122	KLG - E	500,000	Y	4
123	KLG - E	500,000	Y	4
139	KLG - E	500,000	Y	3
140	KLG - E	500,000	Y	3
141	KLG - E	500,000	Y	3

Y = Yes, N = No

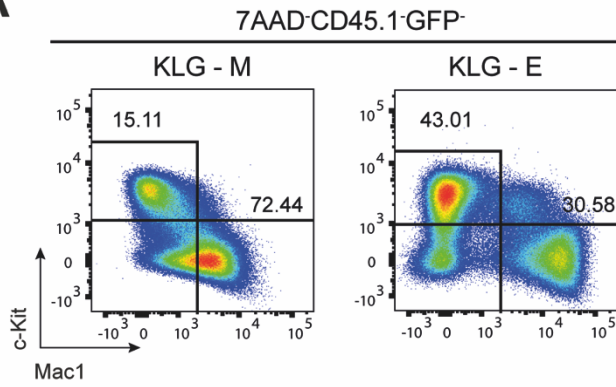
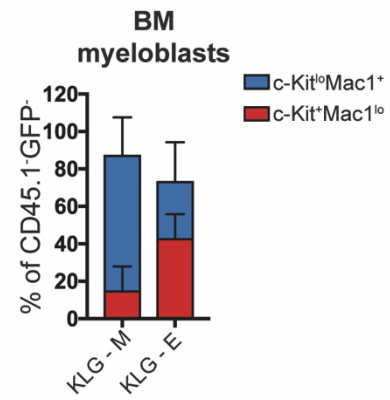
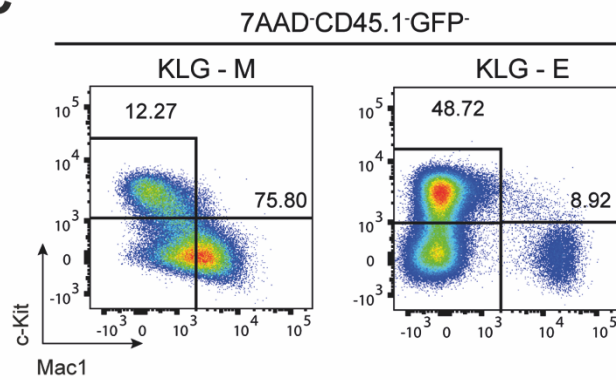
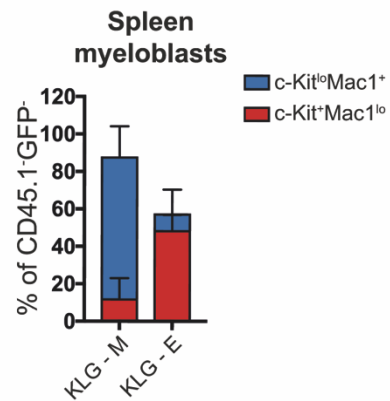
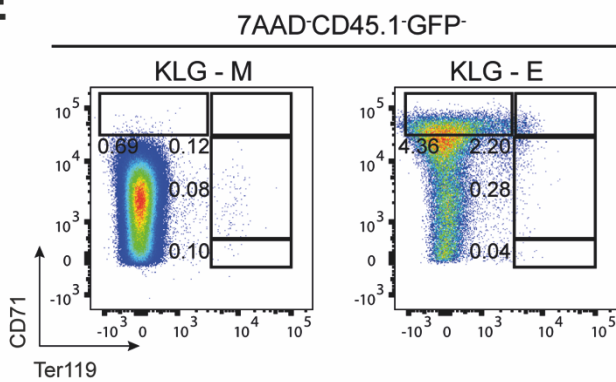
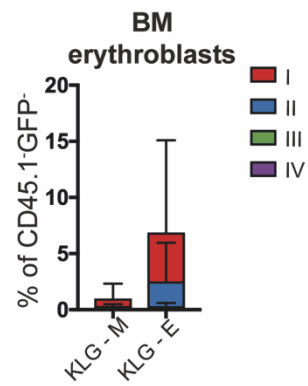
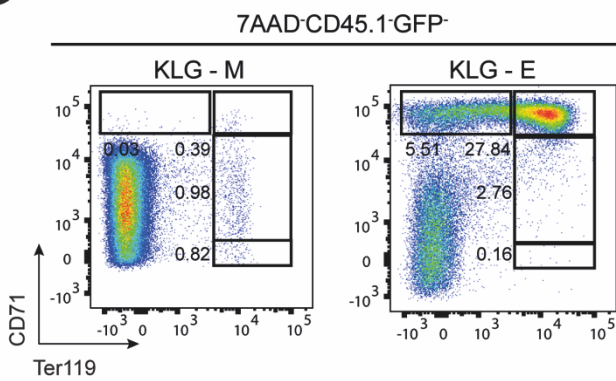
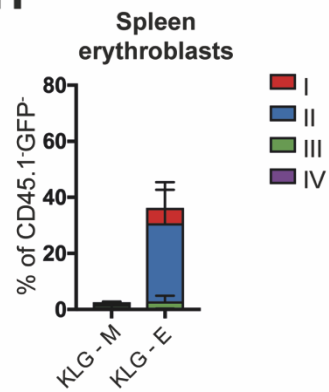
A**B****C****D****E****F****G****H**

Figure 4.8 Secondary transplantation of bulk leukaemic KLG-E cells replicates the original bi-lineage disease phenotype

(A) Representative FACS plots of experimental myeloid cells in the BM as a percentage of BM 7AAD⁻CD45.1⁻GFP⁻ cells across all experiments. (B) c-Kit^{lo}Mac1⁺ and c-Kit⁺Mac1^{lo} cells as a percentage of CD45.1⁻GFP⁻ cells in the BM. (C) Representative FACS plots of experimental myeloid cells in the spleen as a percentage of spleen 7AAD⁻CD45.1⁻GFP⁻ cells across all experiments. (D) c-Kit^{lo}Mac1⁺ and c-Kit⁺Mac1^{lo} cells as a percentage of CD45.1⁻GFP⁻ cells in the spleen. (E) Representative FACS plots of experimental stage I-IV erythroblast cells in the BM as a percentage of BM 7AAD⁻CD45.1⁻GFP⁻ cells across all experiments. (F) Stage I-IV erythroblast cells as a percentage of CD45.1⁻GFP⁻ cells in the BM. (G) Representative FACS plots of experimental stage I-IV erythroblast cells in the spleen as a percentage of spleen 7AAD⁻CD45.1⁻GFP⁻ cells across all experiments. (H) Stage I-IV erythroblast cells as a percentage of CD45.1⁻GFP⁻ cells in the spleen. Analysis of KLG-M (n = 5) and KLG-E (n = 5) from three independent experiments.

4.2.4 *Cebpa* and *Gata2* mutant AEL is sustained by LSCs with a NMP immune-phenotype

To determine if both the myeloid and erythroid blasts in KLG-E mice derived from the same transplantable LSC, the CD45.2 HSPC compartment was first analysed to identify the putative LSCs. There was no difference in BM LSK HSPC compartment, whereas there was a significant increase in LK cells in KLG-M mice compared to WT (24-fold; $P = 0.049$) and a 15-fold increase in LK cells in KLG-E mice compared to WT; however the difference was not statistically different (Figure 4.9A-B).

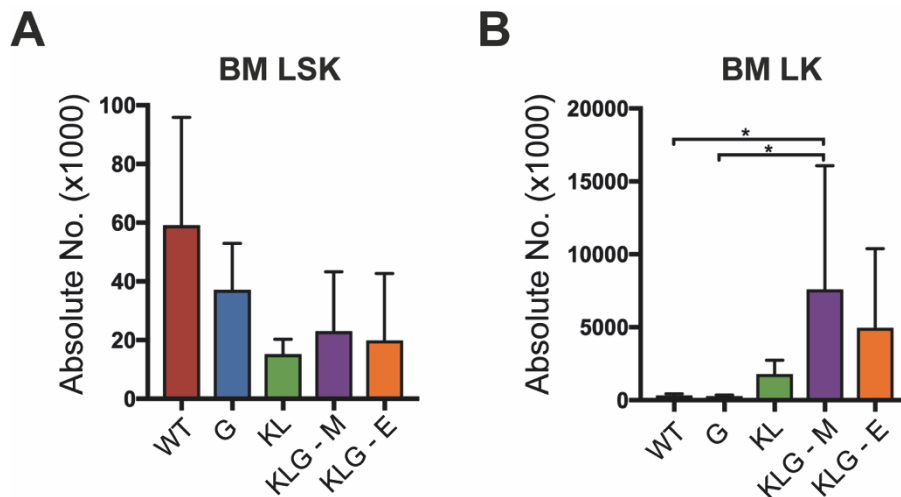


Figure 4.9 LK cells expanded in leukaemic KLG mice

LSK (A) and LK cells (B) in the BM at terminal analysis. Analysis of WT ($n = 7$), G ($n = 9$), KL ($n = 5$), KLG-M ($n = 8$), KLG-E ($n = 5$) from 2-3 independent experiments. The results were analysed using a multiple comparison ANOVA. The results are presented as the mean $\pm$ SD. * $P < 0.05$.

Flow cytometry analysis was then performed using the most recently published myelo-erythroid progenitor definitions that lie within the LK compartment (Drissen et al., 2016; Pronk et al., 2007) (Figure 4.10A). This showed CD45.2⁺ LK cells from WT and G mice displayed a normal distribution of myelo-erythroid progenitors, as previously defined (Figure 4.10B). In contrast, the LK compartment in leukaemic KL and KLG mice consisted principally of LKCD41⁻CD150⁺FcγRII/III⁺CD55⁻ cells, which are the immuno-phenotype of the NMPs (Figure 4.10B). We also observed a significant LKCD41⁺CD150⁻ population in both KL and KLG leukaemic mice (Figure 4.10C). These cells are normally rare and phenotypically heterogenous, however in the leukaemic mice they were predominantly FcγRII/III⁺CD55⁻, similar to the NMPs (Figure 4.10C). Therefore leukaemic NMPs (L-NMPs) were defined as LKFcγRII/III⁺CD55⁻ which included both CD41⁺ and CD41⁻ cells.

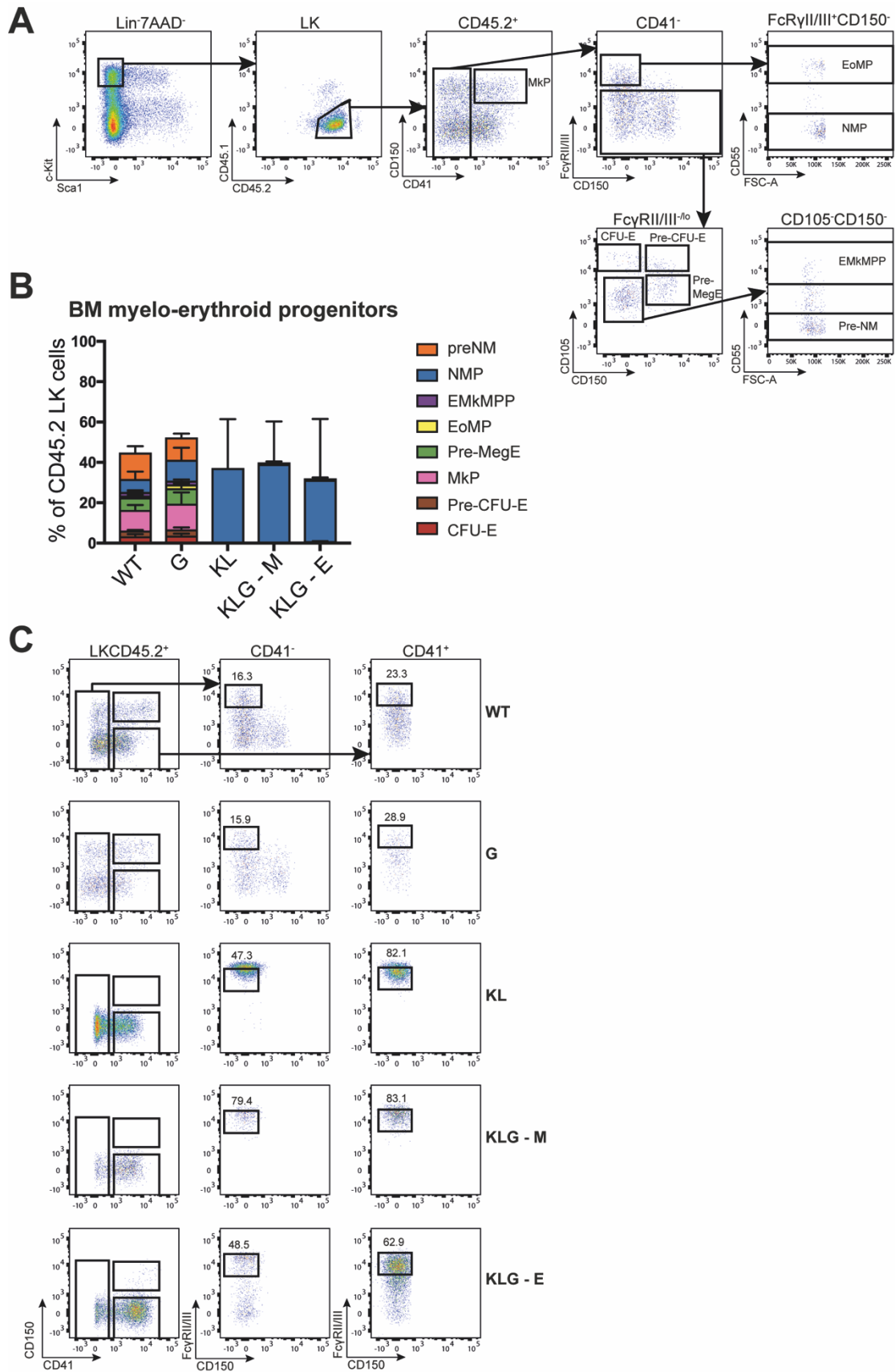


Figure 4.10 Prominent NMP population present in KL and KLG leukaemic mice

(A) Gating strategy used for the myelo-erythroid progenitors. (B) Myelo-erythroid progenitors in the BM of terminal analysed mice as a percentage of CD45.2⁺ LK cells. (C) Representative FACS plots of CD150⁻CD41⁺ cells analysed for the presence of FcγRII/III marker. Analysis of WT (n = 7), G (n = 9), KL (n = 5), KLG-M (n = 8), KLG-E (n = 5) from 2-3 independent experiments.

L-NMPs from KLG-M and KLG-E mice were then sorted and transplanted into secondary recipients (Figure 4.11A). Transplanted L-NMPs from KLG-M and KLG-E mice consistently resulted in rapid disease development (4/4 and 8/9 recipients engrafted respectively, Table 4.2) and in all cases re-capitulated the phenotype of the original disease, with mice transplanted with L-NMPs from KLG-M leukaemic mice developing a pure myeloid disease, while mice transplanted with L-NMPs from KLG-E leukaemic mice instead developed a bi-lineage disease with myeloblasts and erythroblasts in the BM and spleen (Figure 4.11B-E). Purified and transplanted LKFcγRII/III⁺CD55⁺ cells (termed L-EoMPs based on their phenotypic similarity to eosinophil-mast cell progenitors), and LKCD45⁻ cells that contain erythroblasts (Table 4.3 and Figure 4.11A), from KLG-E mice, showed the populations could initiate leukaemia and re-establish the bi-lineage phenotype; however at a lower frequency and/or longer latency compared to L-NMPs (Table 4.2; Figure 4.11B-E). It was therefore concluded that the main LSC population in both KLG-M and KLG-E are the L-NMPs, and from KLG-E mice the L-NMPs could re-establish both transformed erythroid and myeloid blasts in secondary transplants.

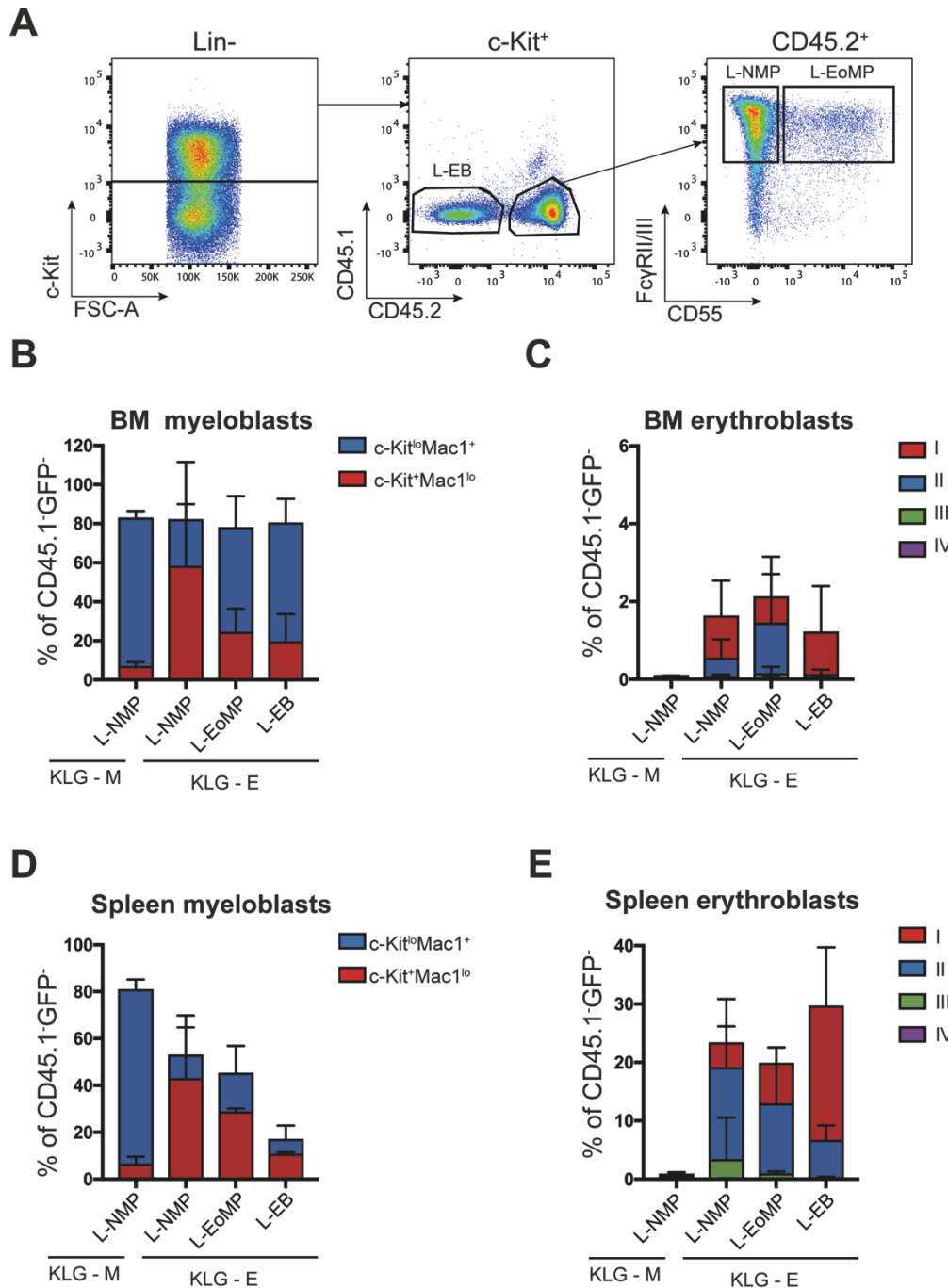


Figure 4.11 L-NMPs act as the LSCs in both KLG-M and KLG-E mice re-capitulating their original diseases

(A) Gating strategy for purifying L-NMPs from KLG-M and KLG-E mice, and L-EoMPs and L-EBs from KLG-E mice for secondary transplants. (B) c-Kit^{lo}Mac1⁺ and c-Kit⁺Mac1^{lo} cells as a percentage of CD45.1⁺GFP⁻ cells in the BM. (C) Stage I-IV erythroblast cells as a percentage of CD45.1⁺GFP⁻ cells in the BM. (D) c-Kit^{lo}Mac1⁺ and c-Kit⁺Mac1^{lo} cells as a

percentage of CD45.1⁺GFP⁺ cells in the spleen. (E) Stage I-IV erythroblast cells as a percentage of CD45.1⁺GFP⁺ cells in the spleen. Analysis of KLG-M L-NMP (n = 2), KLG-E L-NMP (n = 5), KLG-E L-EoMP (n = 2), KLG-E L-EB (n = 2) from three independent experiments.

Table 4.2 Transplantation of purified L-NMP from KLG-M and KLG-E mice, and L-EoMP and L-EB from KLG-E mice, into secondary recipients

Sample No.	Phenotype	Population	No. of cells	Developed leukaemia? (Y/N)	No. of weeks latency
107	KLG - M	L-NMP	2,243	Y	9
108	KLG - M	L-NMP	2,243	Y	9
149	KLG - M	L-NMP	8,934	Y	8
151	KLG - M	L-NMP	8,934	Y	8
105	KLG - E	L-NMP	8,509	Y	6
106	KLG - E	L-NMP	8,509	Y	6
126	KLG - E	L-NMP	1,345	N	NA
127	KLG - E	L-NMP	1,345	Y	5
132	KLG - E	L-NMP	29,500	Y	4
133	KLG - E	L-NMP	29,500	Y	4
147	KLG - E	L-NMP	23,032	Y	5
152	KLG - E	L-NMP	17,394	Y	4
155	KLG - E	L-NMP	17,394	Y	4
103	KLG - E	L-EoMP	1,738	N	NA
104	KLG - E	L-EoMP	1,738	N	NA
124	KLG - E	L-EoMP	40	N	NA
125	KLG - E	L-EoMP	40	N	NA
134	KLG - E	L-EoMP	91	Y	5
135	KLG - E	L-EoMP	91	Y	4
148	KLG - E	L-EoMP	288	N	NA
128	KLG - E	L-EB	1,545	Y	16
129	KLG - E	L-EB	1,545	Y	6
130	KLG - E	L-EB	122	N	NA
131	KLG - E	L-EB	122	N	NA
154	KLG - E	L-EB	260	N	NA
156	KLG - E	L-EB	260	N	NA

Y = Yes, N = No, NA = Not applicable

Table 4.3 Percentage of erythroblasts in LKCD45⁺ cells from KLG-E BM

KLG-E sample type	Sample ID	CD71 ⁺ Ter119 ^{lo} c-Kit ⁺ (% of LKCD45 ⁺)
BM	66	80.86
	67	96.94
	70	0.36
	115	47.35
Spleen	66	75.67
	67	91.62
	70	42.39
	115	69.99

In order to determine the molecular basis of the differentiation block generated in the myeloid and erythroid lineage in KLG-E mice, RNA-sequencing was performed on pre-leukaemic and leukaemic populations. 100 purified NMPs, c-Kit^{lo}Mac1⁺ cells (MB), and CD71⁺Ter119^{lo}c-Kit⁺ cells (EB) were sorted from all genotypes at the pre-leukaemic stage, six weeks post-transplantation. NMP and MB populations were collected from KL and KLG genotypes at the leukaemic stage and EB cells were collected from KLG-E mice at the leukaemic stage. Leukaemic populations were denoted by an 'L-' before the name of the population. Clustering of these populations using principle component analysis showed that non-leukaemic MBs, EBs and NMPs clustered according to cell identity (Figure 4.12A). The transformed MB and EB populations clustered closer to the NMPs, consistent with a more progenitor-like state (Figure 4.12A). Using GSEA, pre-CFU-E genes were shown to be down-regulated in KLG L-EBs compared to pre-leukaemic KLG EBs (Figure 4.12B), and the pre-GM gene set was enriched in KLG L-EBs compared to pre-leukaemic KLG EBs (Figure 4.12C). Neutrophil differentiation-specific genes were down-regulated in L-MBs from both KLG-M and KLG-E mice compared to pre-leukaemic KLG MB (Figure 4.12D-E). Together this indicates the block in differentiation along the myeloid and erythroid lineage in KLG mice, correlated with suppression of the respective differentiation programmes at the molecular level, leading to a progenitor-like state.

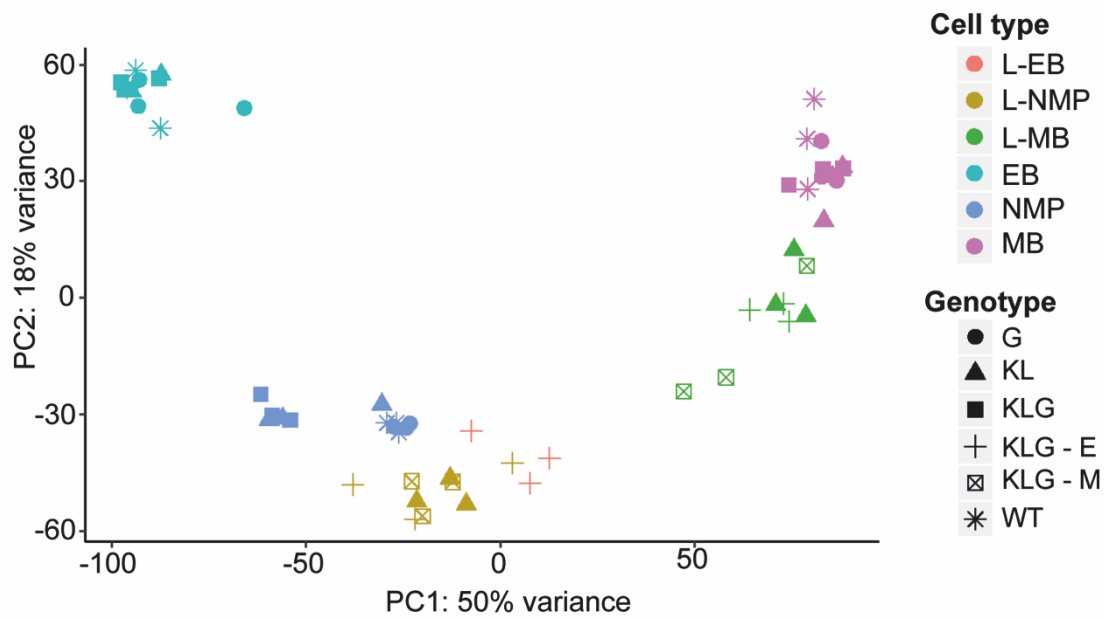
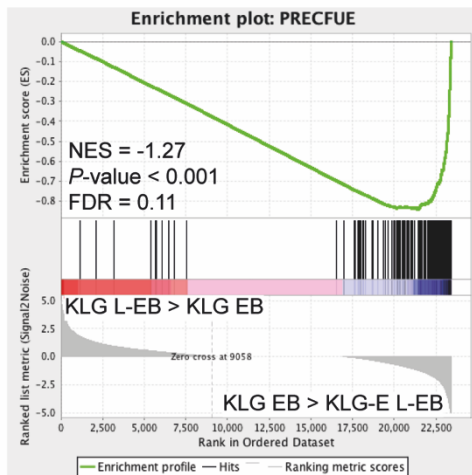
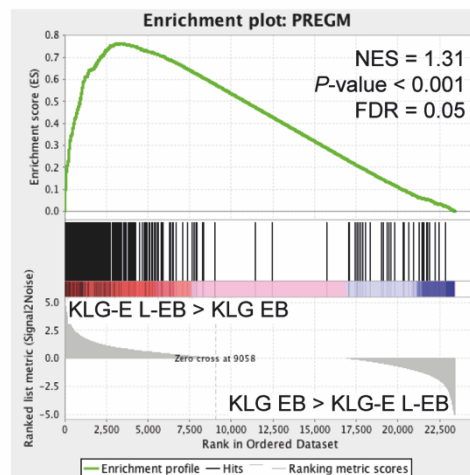
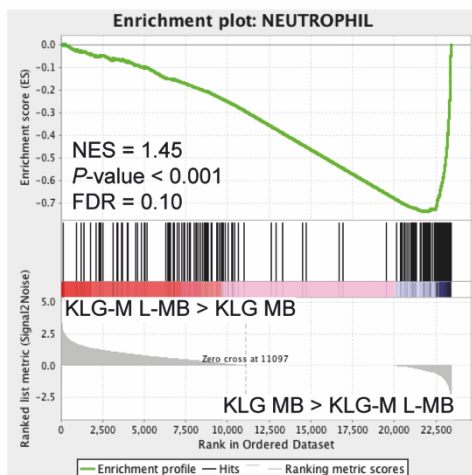
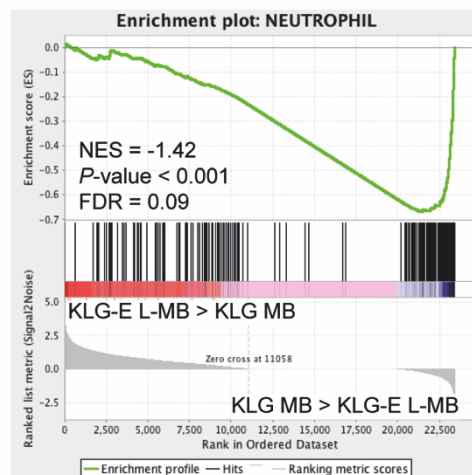
A**B****C****D****E**

Figure 4.12 Suppression of myeloid and erythroid differentiation programmes in leukaemic KLG mice

100 purified NMPs, c-Kit^{lo}Mac1⁺ (MB), CD71⁺Ter119^{lo}c-Kit⁺ (EB) were sorted from all genotypes at the pre-leukaemic stage six weeks post-transplantation. NMPs and MBs were collected from KL and KLG genotypes at the leukaemic stage and EBs were collected from KLG-E mice at the leukaemic stage. Leukaemic populations were denoted by a 'L-' before the name of the population. Three replicates were generated per population per genotype, except for KL EBs where there were two replicates. (A) Principle component analysis of all samples using the top 500 highly variable genes. (B-C) GSEA of KLG-E L-EB vs. KLG EB for pre-CFU-E genes and (B) and pre-GM genes (C). (D) GSEA of KLG-M L-MB vs. KLG MB for neutrophil genes. (E) GSEA of KLG-E L-MB vs. KLG MB for neutrophil genes.

4.2.5 Biallelic *Cebpa* mutations introduce an ectopic erythroid lineage potential in NMPs

NMPs are normally restricted to neutrophil and monocyte fates, and do not have detectable erythroid lineage potential (Drissen et al., 2016). However, Bereshchenko *et al.*, have previously demonstrated that pre-leukaemic LSKCD150⁺Flt3⁻ HSCs from KL mice were enriched in erythroid and depleted of myeloid gene expression compared to WT HSCs (Bereshchenko et al., 2009). To determine if a similar effect was present in *Cebpa* mutant NMPs, GSEA was first performed on KL and KLG NMPs on pre-MegE and pre-GM gene sets. Indeed, there was a significant loss of genes associated with pre-GM cells, and an up-regulation of genes associated pre-MegE cells in KL NMPs compared to WT NMPs, however not statistically significant, indicating a loss of myeloid and a possible up-regulation in erythroid gene expression (Figure 4.13A-B). The same pattern of gene expression was also observed in KLG NMPs compared to G NMPs (Figure 4.13C-D). This indicated that the biallelic *Cebpa* mutations are responsible for the altered gene expression in NMPs.

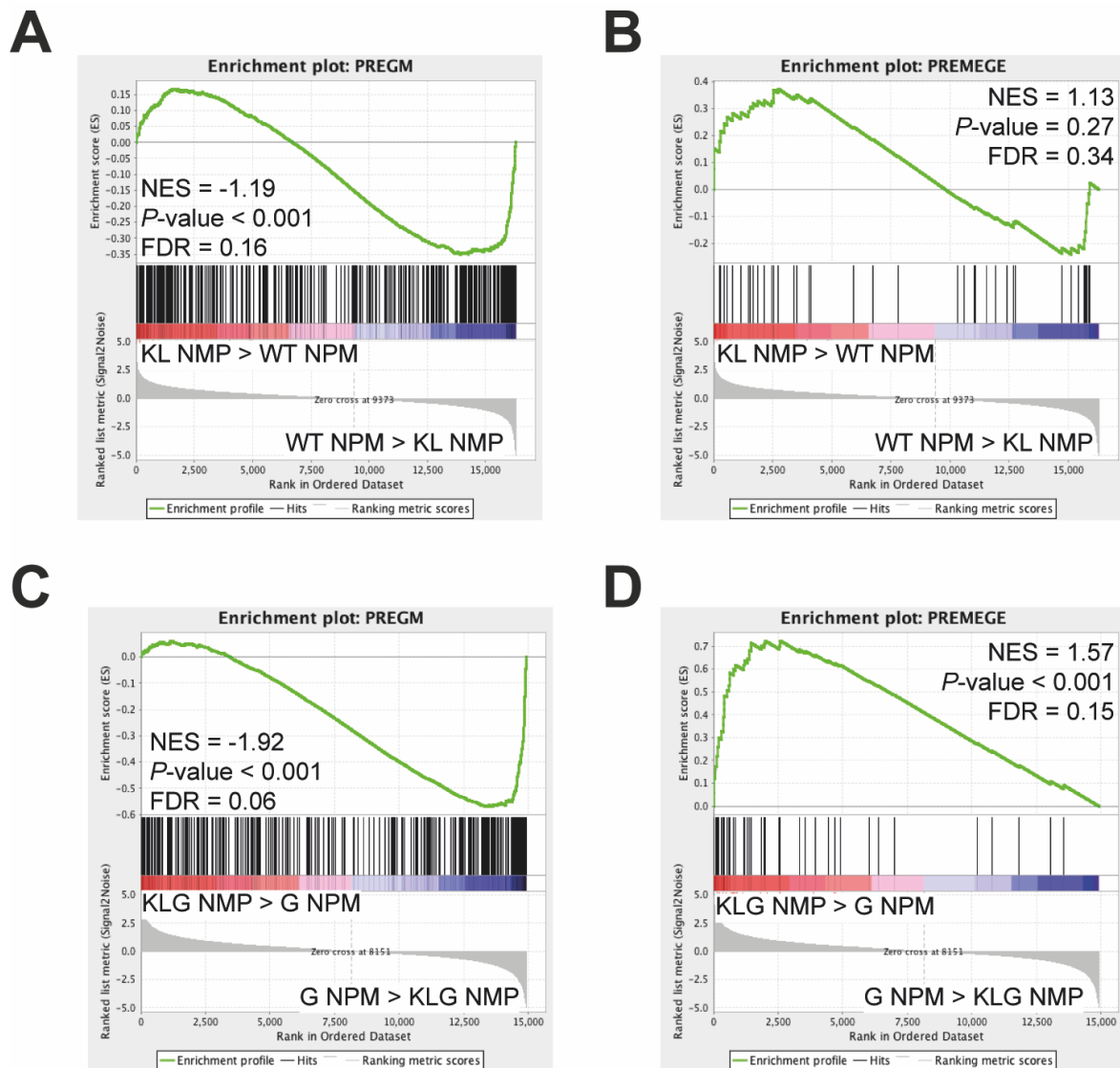


Figure 4.13 Biallelic *Cebpa* mutations leads to a loss of myeloid and an enrichment in erythroid gene expression

GSEA of KL NMPs vs WT NMPs for pre-GM (A) and pre-MegE gene set (B). GSEA of KLG NMPs vs G NMPs for pre-GM (C) and pre-MegE gene set (D). Three replicates per genotype.

Myeloid and erythroid lineage programming in NMPs was further investigated by analysing gene expression of transcription factor genes important for myeloid or erythroid priming. Pre-leukaemic KL and KLG NMPs expressed fewer myeloid and more erythroid transcription factor genes compared to WT and G at the single cell level (Figure 4.14).

Interestingly, KL and KLG NMPs showed co-expression of both myeloid and erythroid transcription factor genes indicating KL and KLG NMPs are both myeloid and erythroid primed, which contrasted to WT and G NMPs which both predominantly only expressed myeloid genes (Figure 4.14B-E).

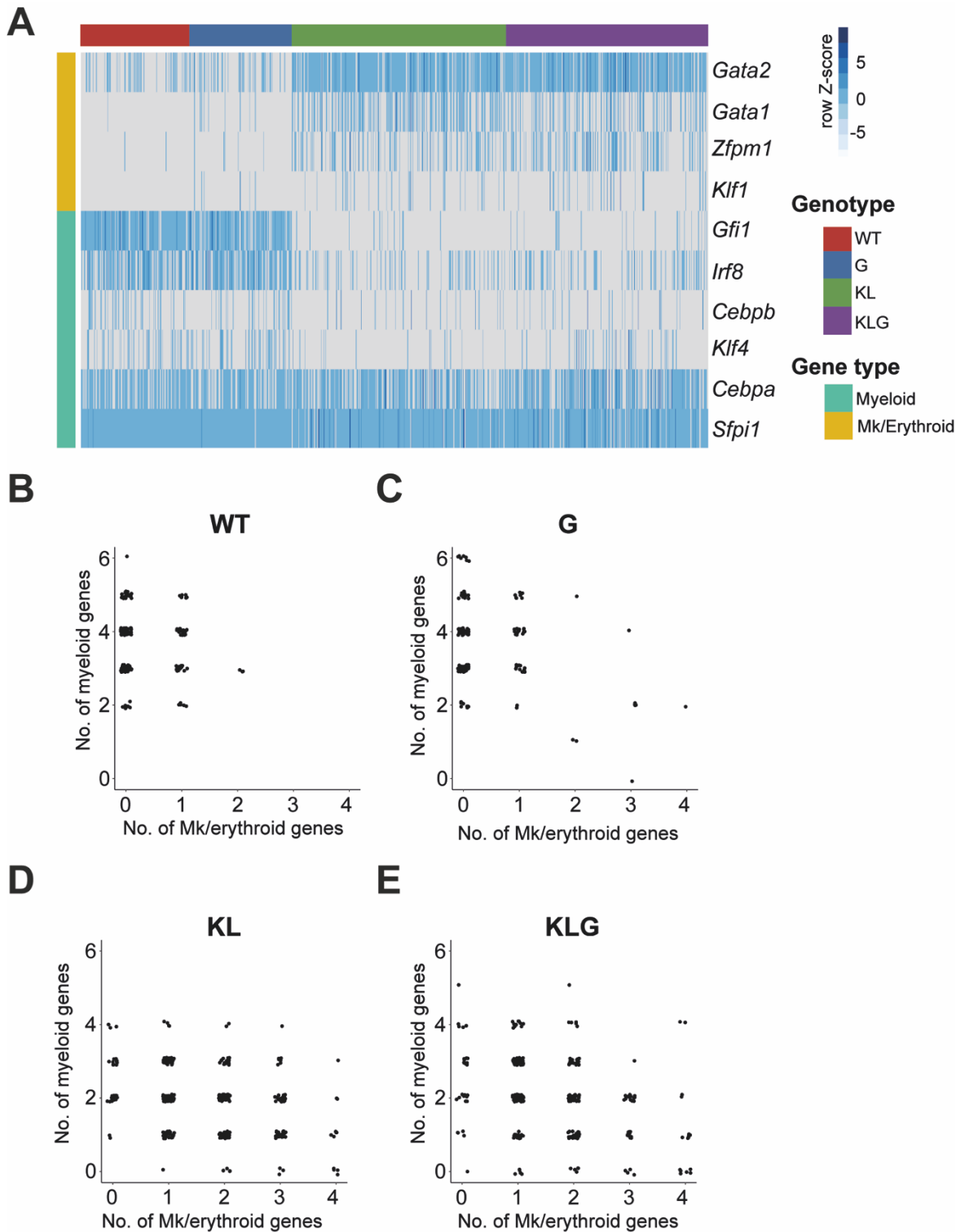


Figure 4.14 Biallelic *Cebpa* mutations lead to a loss of myeloid and an increase in erythroid gene expression in NMPs at the single cell level

(A) Multiplex quantitative real-time PCR of myeloid and megakaryocyte(Mk)/erythroid transcription factor genes on single NMPs taken at the pre-leukaemic stage six weeks post-transplantation. Heatmap depicting $2^{-\Delta Ct}$ normalised to *Hprt*. (B-E) Scatterplot depicting the number of myeloid and Mk/erythroid transcription factor genes expressed per NMP for WT (B), G (C), KL (D) and KLG (E). Analysis of single NMPs of WT (n = 192), G (n = 192), KL (n = 384), and KLG (n = 384).

Pre-leukaemic NMPs from KLG and WT genotypes were then tested functionally in methylcellulose cell culture. Gene expression and morphology were analysed from colonies derived from single NMPs. As expected WT NMPs generated neutrophils and monocytes as shown by morphology analysis and expression of only myeloid genes (Figure 4.15A-C). This contrasts KLG NMPs that generated colonies with erythroid and immature myeloid morphology (Figure 4.15A). Gene expression analysis of KLG NMPs also showed co-expression of multiple neutrophil- and erythroid-specific genes (Figure 4.15B-C). This data suggests both neutrophil and erythroid lineage cells derived from a single KLG NMP.

Together, this shows phenotypic NMPs carrying the biallelic *Cebpa* mutations are transcriptionally and functionally re-programmed to acquire ectopic erythroid potential, generating a neomorphic neutrophil-erythroid progenitor population. This suggests that myeloid and erythroid blasts are likely to derive from a bi-potent LSC with a NMP immunophenotype.

4.2.6 ZF1 *Gata2* mutation lead to accumulation of erythroid progenitors at the CFU-E stage

Although the biallelic *Cebpa* mutation introduces ectopic erythroid potential in the NMPs, only mice carrying both the biallelic *Cebpa* and *Gata2* mutations develop an AEL, suggesting the mutations must synergise to cause AEL. As AEL is characterised by an erythroid differentiation block, erythroid progenitor cells were analysed by flow cytometry at the pre-leukaemic stage, six weeks post-transplantation, to determine if the *Gata2* mutation impairs the progression of progenitors along the erythroid differentiation pathway.

There was no difference in any of the myelo-erythroid progenitors in G mice compared to WT (Figure 4.16A-H). KL mice had an increase in absolute number of pre-NMs and an increase in some, but not all, of the *Gata1*-expressing myelo-erythroid progenitors including; the pre-MegE, MkP, and pre-CFU-E compared to WT (Figure 4.16E-G). KLG mice had a further increase in pre-NMs compared to KL (2-fold, $P = 0.006$; Figure 4.16A), and there was a significant expansion of all *Gata1*-expressing myelo-erythroid progenitors, except for pre-CFU-Es, compared to WT (Figure 4.16C-F, H). By normalising the size of the progenitor populations to WT mice, the results showed a selective expansion of the EMkMPPs and CFU-Es in the KLG mice that is not seen in KL mice (Figure 4.16I). This data suggests in the presence of the biallelic *Cebpa* mutation, the *Gata2* mutation leads to impaired progression of the myelo-erythroid progenitor differentiation at two distinct stages.

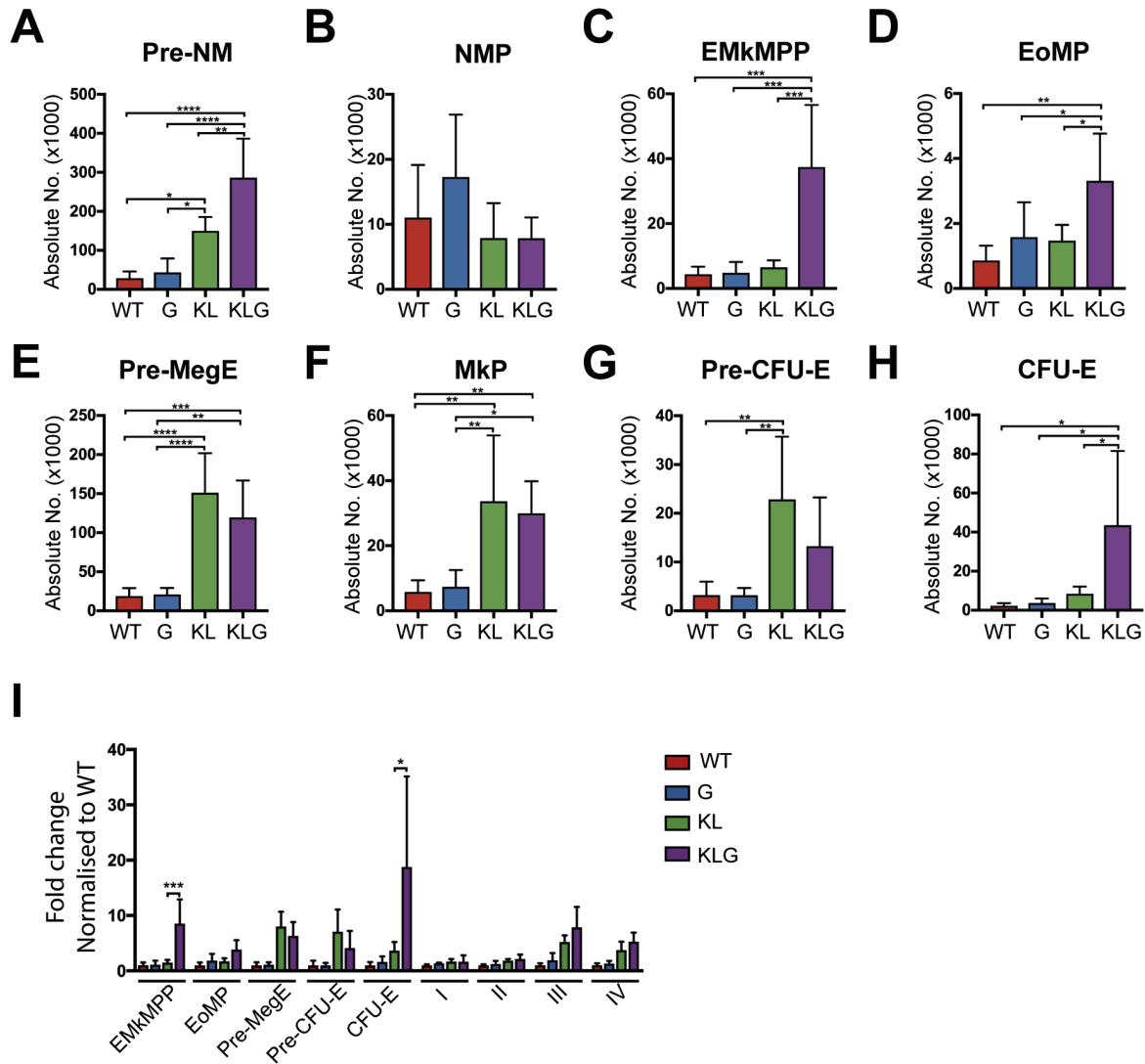


Figure 4.16 *Gata2* mutation synergises with biallelic *Cebpa* mutations resulting in an accumulation of EMkMPPs and CFU-Es

Analysis of myelo-erythroid progenitors at the pre-leukaemic stage six weeks post-transplantation. (A-H) Absolute numbers of myelo-erythroid progenitors in the BM. (A) Pre-NM. (B) NMP. (C) EMkMPP. (D) EoMP. (E) Pre-MegE. (F) MkP. (G) Pre-CFU-E. (H) CFU-E. (I) Myelo-erythroid progenitors and stage I-IV erythroblasts normalised to WT mice. Myelo-erythroid progenitor analysis was performed on 4-6 replicates from two independent experiments. Stage I-IV erythroblast analysis was performed on 3-4 replicates from one experiment. The results were analysed using a multiple comparison ANOVA. The results are presented as the mean $\pm$ SD. * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ **** $P < 0.0001$.

4.3 Discussion

The specific association between biallelic *CEBPA* and *GATA2* mutation identified in AML and bi-lineage AEL patients, suggests the mutations synergise to cause leukaemia development. To determine if the mutations synergise, and more specifically develop a bi-lineage leukaemia with myeloid and erythroid blasts, a knock-in mouse model containing the biallelic *Cebpa* and ZF1 *Gata2* mutations was generated to study leukaemia development. This study shows that the mutations synergise leading to a more aggressive disease. ~60% of KLG mice develop a pure AML while the remaining ~40% develop a bi-lineage AEL with myeloid and erythroid blasts. The AEL is sustained by a bi-potent LSC that was generated by the biallelic *Cebpa* mutations reprogramming the NMPs to induce ectopic erythroid potential, creating a neomorphic neutrophil-erythroid progenitor. The *Gata2* mutation synergises with the biallelic *Cebpa* mutation to impair erythroid differentiation progression at the EMkMPP and CFU-E stages.

4.3.1 Neomorphic neutrophil-erythroid progenitor act as a bi-potent LSC in AEL

The cells that sustain AML are transformed progenitors committed to myeloid lineage differentiation, including LMPPs and GMPs (Goardon et al., 2011). These cells normally give rise to neutrophils and monocytes, but lack the capability to generate erythroid lineage cells. In the mouse the equivalent cells are LMPPs and NMPs, and consistent with the human findings the main LSC in *Cebpa* mutant leukaemias have an NMP phenotype. Normally, *Gata2* is not expressed in LMPPs or NMPs (Drissen et al., 2016). However, both bulk RNA-sequencing and single cell quantitative real-time PCR showed the biallelic *Cebpa* mutations lead to an up-regulation of *Gata2* expression in both KL and KLG phenotypic NMPs. Ectopic *Gata2* expression may be a crucial step in AEL development as it has been previously shown that ectopic expression of GATA2 promotes

erythroblast proliferation and block terminal erythroid differentiation in the erythroblast cell line HD3 (Briegel et al., 1993).

The up-regulation of *Gata2* in *Cebpa* mutant NMPs was part of a more general lineage reprogramming. Megakaryocyte/erythroid gene expression were generally up-regulated in KL and KLG NMPs compared to WT and G NMP, with concurrent down-regulation of myeloid gene expression. As previously mentioned up-regulation of megakaryocyte/erythroid gene expression in KL HSCs has been reported (Bereshchenko et al., 2009). Interestingly, biallelic C-terminal *Cebpa* mutations (KK) were also investigated for leukaemia development and only 25% of mice developed a myeloid disease with the remaining mice developing a more immature phenotype with undetectable myeloid infiltration, characterised by an accumulation of immature erythroid cells. KK HSCs also had a stronger up-regulation of erythroid genes compared to KL. This indicates that the C-terminal mutation may be specifically involved in erythroid transformation. However, in this study there was no difference in the expression of megakaryocyte/erythroid genes between KL and KLG NMPs, therefore an additive effect of the *Gata2* mutation to further up-regulate erythroid gene expression does not seem to be the cause of AEL development in KLG mice.

Up-regulation of erythroid gene expression was accompanied by the acquisition of erythroid lineage potential in KLG NMPs. By analysing both morphology and gene expression from colonies derived from single NMPs, this study demonstrated that a single NMP contained both myeloid and erythroid potential, suggesting the bi-lineage AEL leukaemia derived from a single bi-potent LSC rather than two distinct LSC populations. Mixed lineage leukaemias are also bi-phenotypic as they present features of acute lymphoblastic leukaemia and AML. This can be bi-lineage with a myeloid and lymphoid population or bi-phenotypic with the co-expression of lymphoid and myeloid antigens and gene expression, or can also contain both bi-lineage and bi-phenotypic features. The MLL

fusion is a common initiating event in mixed lineage leukaemias (Alexander et al., 2018; Ottersbach et al., 2018). Mouse models of MLL fusions have suggested that the LMPP, which has both myeloid and lymphoid potential, could be the cell-of-origin for MLL-driven leukaemias (Barrett et al., 2016; Cozzio et al., 2003; Stavropoulou et al., 2018). This differs to the bi-lineage leukaemia in this study as the reprogramming of the NMPs leads to an acquisition of erythroid potential, presenting a novel mechanism of bi-lineage leukaemia development.

Sequencing studies have shown the association between specific mutations, however few studies have shown how mutations synergise in leukaemia development. This study shows ectopic erythroid potential in NMPs was insufficient to induce AEL in the absence of ZF1 *Gata2* mutation. The erythroid blasts observed in KLG-E mice were arrested at the CD71⁺Ter119^{lo}c-Kit⁺ pro-erythroblast stage, which corresponds closely to the phenotypic CFU-E. An accumulation of phenotypic CFU-Es in pre-leukaemic KLG mice was observed. This may be an important step for AEL development which is only seen in KLG mice, suggesting the mutations synergise to cause arrest at CFU-E stage of erythroid differentiation. Further analysis of the RNA-sequencing data on the EB and L-EBs may give insight into the specific gene expression changes that lead to the arrest in erythroid differentiation by the combined mutations.

4.3.2 *Cebpa*^{KL};*Gata2*^{G320D/+} as an AEL model

To my knowledge this is the first demonstration of an AEL model displaying both morphological and phenotypic AEL characteristics seen in AEL erythroleukaemia (erythroid/myeloid) patients. As mentioned in chapter one, the recently revised WHO classification of myeloid neoplasms has removed the erythroleukaemia subtype from the AML category and is now reclassified as MDS (Arber et al., 2016). This study shows the AEL developed in KLG-E mice was more aggressive than mice with AML, suggesting the

erythroleukaemia that develops in this model closer resembles an AML subtype rather than an MDS.

Only one paper to date has shown *GATA2* and *CEBPA* co-mutated in AEL (Ping et al., 2017). However, there are currently limited mutational studies performed on AEL patients samples, and most of the previous studies did not include *GATA2* in their screening panels. Analysis of larger cohorts may reveal *GATA2* and *CEBPA* as important causative mutations in AEL.

4.3.3 Conclusion

Biallelic *Cebpa* and *Gata2* mutations synergise during leukaemia development. At the pre-leukaemic stage the biallelic *Cebpa* mutation reprograms the NMPs to acquire ectopic erythroid potential, while the *Gata2* mutation synergises with the *Cebpa* mutation to impair erythroid differentiation progression. LSCs are formed within the phenotypic NMPs and due to the ectopic erythroid potential creates a bi-potent neomorphic neutrophil-erythroid progenitor that generates both myeloid and erythroid blasts. Further investigation is required to confirm these findings in human AEL samples.

AEL only develops when the *Cebpa* and *Gata2* mutations are combined, however why some KLG mice develop an AEL while the remaining mice develop an AML remains unclear. Gene expression analysis of L-NMPs from KLG-M and KLG-E mice showed minimal changes in differentially expressed genes (data not shown). This suggests other mechanisms may be involved. Both C/EBP α and GATA2 have been shown to cause epigenetic alterations (Hasemann et al., 2014; Kitajima et al., 2006) therefore epigenetic dysregulation may play a possible role in AEL development and would need to be investigated in future studies.

CHAPTER FIVE: Characterising the effect of Gata2^{G320D} on haematopoiesis

5.1 Introduction

Phenotypic variation and separate clustering of the germline and somatic mutations in *GATA2* in haematological neoplasms, suggests somatic mutations in ZF1 may affect the function of *GATA2* differently to germline mutations in ZF2. Functional studies of *GATA2* mutations have so far only been studied by expression of mutant *GATA2* in HEK 293T cells (Cortes-Lavaud et al., 2015; Greif, Dufour, et al., 2012; Hahn et al., 2011; Niimi et al., 2013; Ostergaard et al., 2011; Ping et al., 2017; S. J. Zhang et al., 2008), therefore a *Gata2*^{G320D} mouse model was generated and studied at both heterozygous and homozygous level to understand the effects of a ZF1 *Gata2* mutation on haematopoiesis.

5.2 Results

5.2.1 *Gata2* homozygous mutation is detrimental to HSC function

To assess the effect of mutant *Gata2* on haematopoiesis the *Gata2*^{G320D/+} mouse model was established (Figure 5.1) and bred mice to homozygosity. Pups were born at the expected Mendelian ratio, however *Gata2*^{G320D/G320D} mice were smaller in size compared to WT and *Gata2*^{G320D/+} mice, with some female *Gata2*^{G320D/G320D} mice developing a defective reproductive system rendering them unable to mate. Also all *Gata2*^{G320D/G320D} mice developed an eye defect (Cristina Di Genua, observation). Due to the eye defect mice were culled and analysed at 4-5 weeks of age.

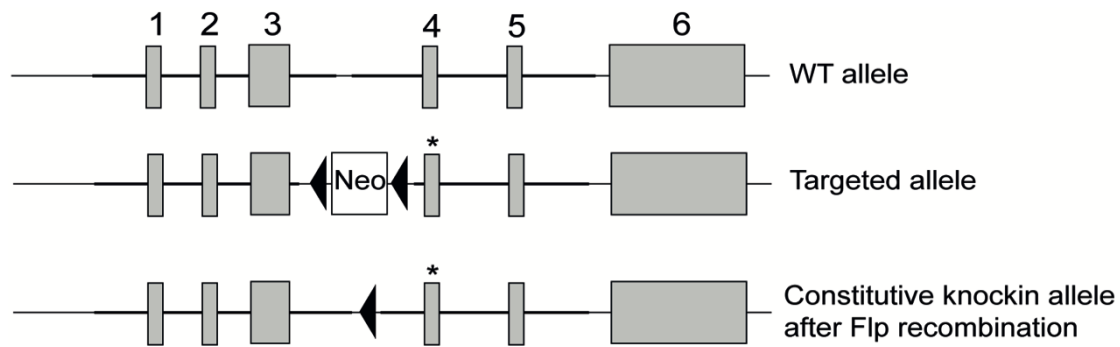


Figure 5.1 Schematic representation of WT, targeted allele, and constitutive knock-in allele after Flp recombination

Closed boxes, exons; open box, neomycin cassette; triangles, Frt-sites; asterisk, mutation.

To evaluate the haematopoietic effect of the *Gata2* mutation at both heterozygous and homozygous level blood cells and BM counts were first performed. There was no difference in WBC count or BM cellularity in *Gata2*^{G320D/+} mice, however there was a decrease in haemoglobin levels compared to WT ($P = 0.0023$) (Figure 5.2A-C). There was also no difference in WBC count in *Gata2*^{G320D/G320D} mice compared to WT, however there was a decrease in BM cellularity ($P = 0.0241$) and a decrease in haemoglobin levels compared to WT (Figure 5.2A-C).

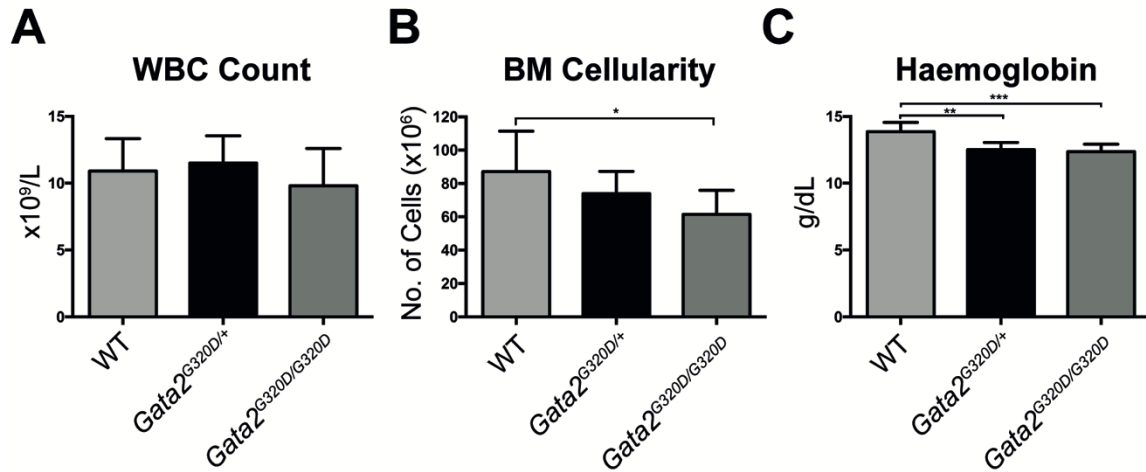


Figure 5.2 The effect of *Gata2* mutation on blood cells and BM counts

(A) PB WBC count analysed in two independent experiments (n = 5). (B) Total BM cellularity analysed in four independent experiments (n = 8-12). (C) PB haemoglobin analysed in three independent experiments (n = 7). The results were analysed using a multiple comparison ANOVA. The results are presented as the mean $\pm$ SD. *P<0.05 **P<0.01 ***P<0.001.

The decrease in BM cellularity could indicate a decrease in HSCs therefore the HSPCs were analysed by flow cytometry (Figure 5.3A). There was a decrease in LSK in *Gata2*^{G320D/+} mice compared to WT ($P = 0.0021$; Figure 5.3B). This decrease was due to a decrease in the LMPP compartment ($P = 0.0388$; Figure 5.3C). *Gata2*^{G320D/G320D} mice had a 2-fold decrease in the LSK cell ($P = < 0.0001$; Figure 5.3B). Within the LSK compartment there was a decrease in all HSPCs analysed including a 1.5-fold decrease in HSCs (Figure 5.3C-E).

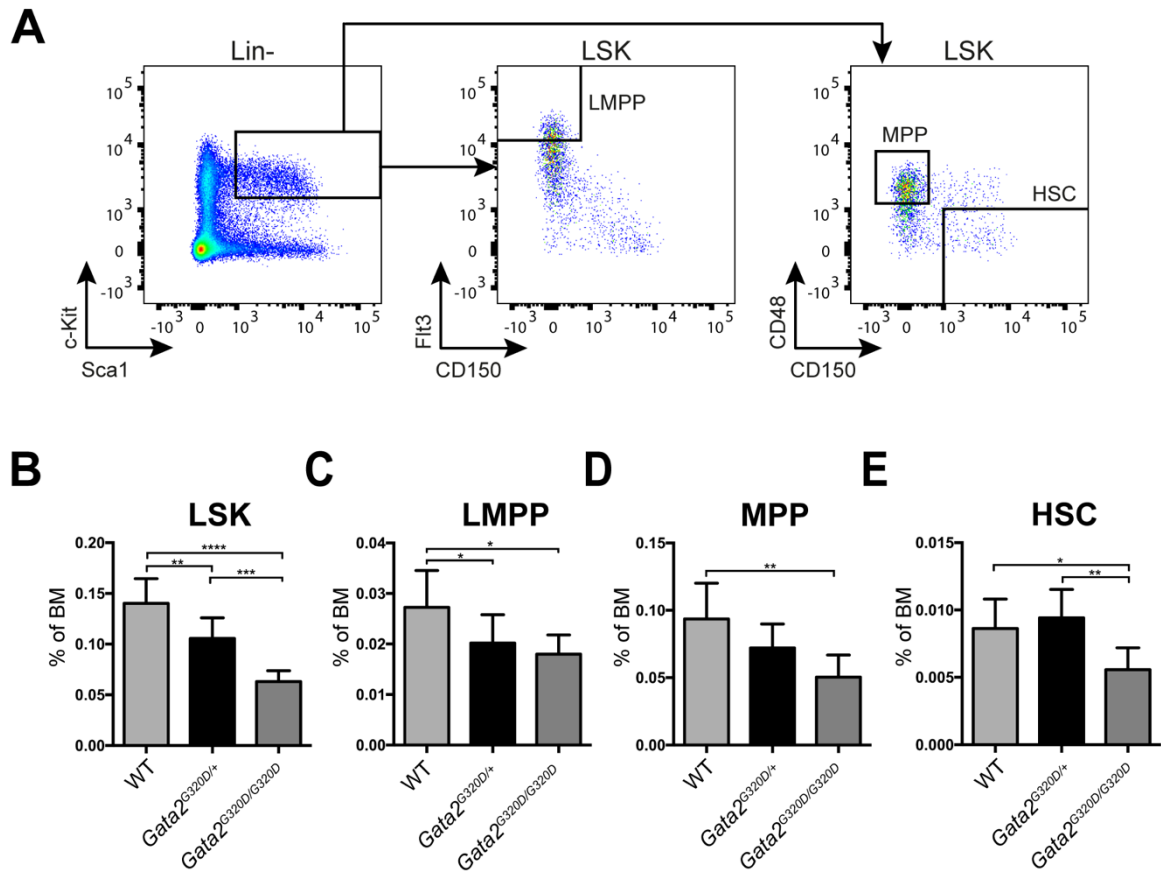


Figure 5.3 *Gata2* homozygous mutation causes a loss of phenotypic HSCs

(A) Schematic of HSC gating strategy, (B) LSK, (C) LMPPs, (D) MPPs defined as LSKCD150⁺CD48⁺, and (E) HSCs defined as LSKCD150⁺CD48⁻, as a percentage of the BM. Four independent experiments, 8-12 biological replicate per genotype. The results were analysed using multiple comparison ANOVA. The results are presented as the mean $\pm$ SD. *P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

To determine if the decrease in phenotypic HSCs correlated with decrease in HSC function a competitive transplant was performed. There was no difference in reconstitution over time in *Gata2*^{G320D/+} BM cells compared to WT. However, *Gata2*^{G320D/G320D} BM cells were outcompeted by WT BM competitor cells as there was a decrease in reconstitution over time in primary and secondary transplant in myeloid, B and T cells (Figure 5.4).

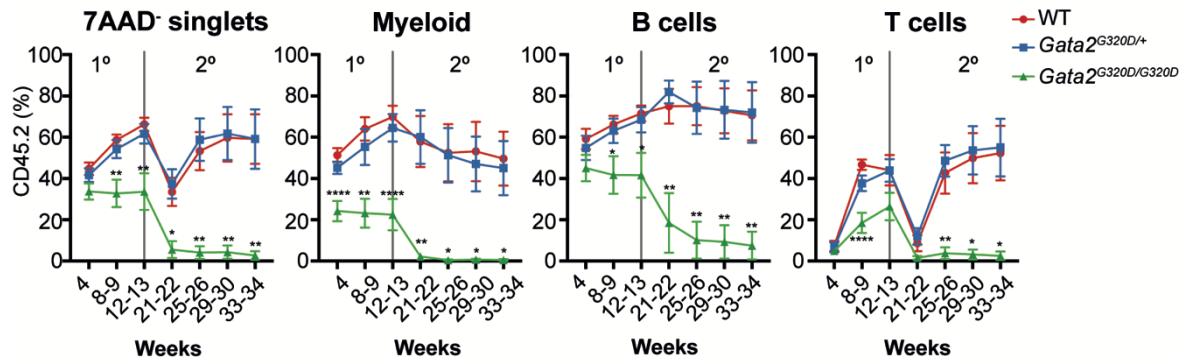


Figure 5.4 *Gata2* homozygous mutation causes a loss of functional HSCs

500,000 BM cells (CD45.2 allotype) were transplanted with 500,000 BM competitor cells (CD45.1 allotype) into lethally irradiated (2 x 500rads) recipients (CD45.1/2 allotypes). Secondary transplant was performed 17-18 weeks post-transplantation by transplanting 10×10^6 BM cells into lethally irradiated recipients. Percentage reconstitution of CD45.2⁺ cells in 7AAD⁻ singlets, myeloid, B and T cell compartments. Two independent experiments, 8-10 replicates per genotype for primary transplants. Six replicates per genotype for secondary transplants. The results were analysed using multiple comparison ANOVA. The results are presented as the mean $\pm$ SEM. Statistical significance between WT and *Gata2*^{G320D/G320D} are shown. *P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

Together, this data suggests that *Gata2* mutation at both heterozygous and homozygous level decreases haemoglobin levels, however not to a level that can be considered anaemic, and loss of phenotypic and functional HSCs is caused only by the homozygous mutation.

5.2.2 *Gata2* homozygous mutation results in a loss of mast cell development

The myelo-erythroid progenitors were analysed by flow cytometry (Figure 3.5A). There was no significant difference between the myelo-erythroid progenitors in *Gata2*^{G320D/+} and *Gata2*^{G320D/G320D} BM cells compared to WT (Figure 5.5).

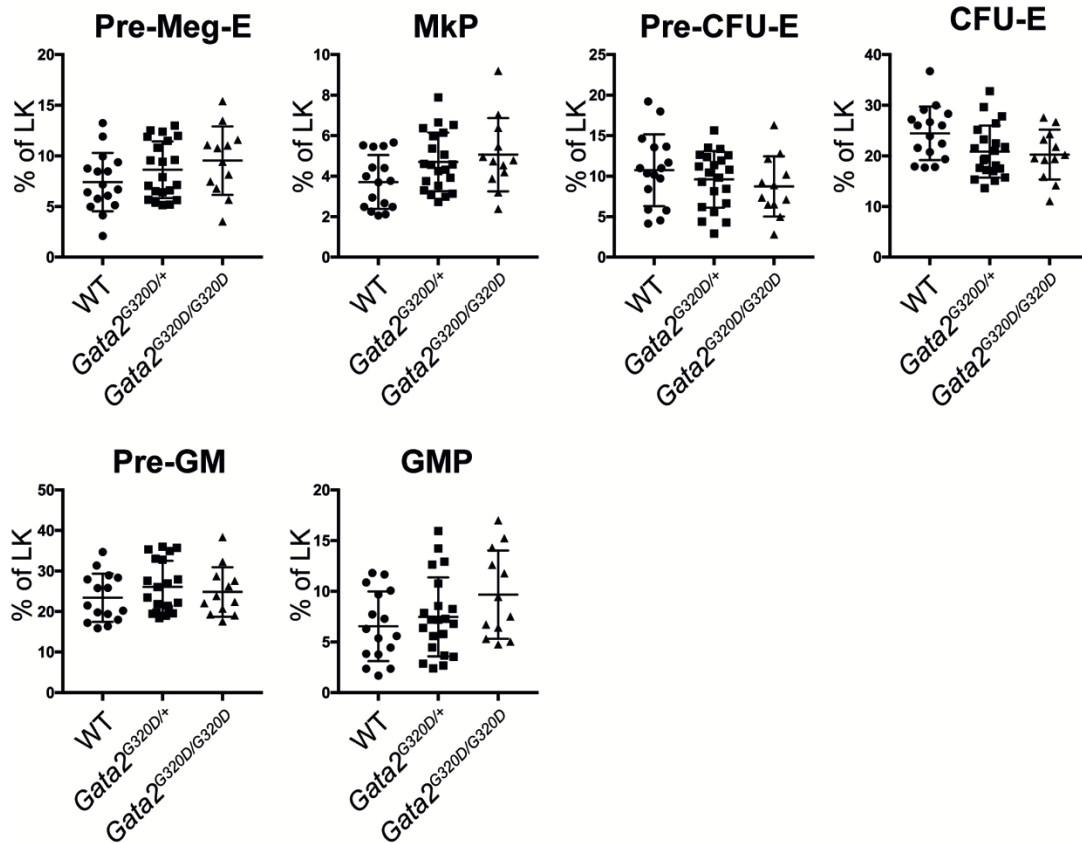


Figure 5.5 *Gata2* mutation does not affect the myelo-erythroid progenitors

Myelo-erythroid progenitor cells as a percentage of the LK cells. Analysis of WT (n = 16), *Gata2*^{G320D/+} (n = 21), *Gata2*^{G320D/G320D} (n = 12) mice from seven independent experiments. The results were analysed using a multiple comparison ANOVA. The results are presented as the mean ± SD.

GATA switch is important for lineage specification (Grass et al., 2003), therefore mutations in *Gata2* may inhibit the GATA switch process (Bresnick, Lee, Fujiwara, Johnson, & Keles, 2010) and may affect the differentiation of *Gata1*-expressing haematopoietic myeloid progenitor and mature cells. The *Gata1*-expressing myeloid progenitors (EMkMPP and EoMP) and myeloid progenitors that lack *Gata1*-expression (pre-NM and NMP) were first analysed (Figure 4.10A). There was no difference in frequency of pre-NM, NMP and EMkMPP cells in *Gata2*^{G320D/+} or *Gata2*^{G320D/G320D} BM cells

compared to WT (Figure 5.6A-C). However, there was a trend towards a decrease in frequency of EoMPs (Figure 5.6D).

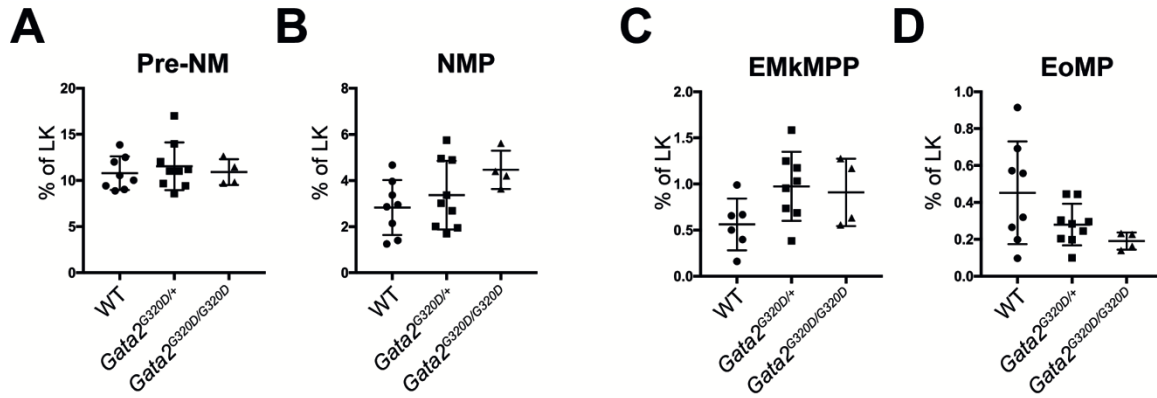


Figure 5.6 Trend towards a decrease in EoMPs with *Gata2* mutation

(A-D) Myeloid progenitors as a percentage of LK cells of (A) Pre-NMs, (B) NMPs, (C) EMkMPPs, and (D) EoMPs. Analysis of WT (n = 8), *Gata2*^{G320D/+} (n = 9), *Gata2*^{G320D/G320D} (n = 4) mice from three independent experiments. The results were analysed using a multiple comparison ANOVA. The results are presented as the mean ± SD.

The *Gata1*-expressing mature myeloid cells (mast cells and eosinophils), and the mature myeloid cells that lack *Gata1*-expression (monocytes and neutrophils), as well as B and T cells, were then analysed in the PB and peritoneum (Figure 5.7A). There was no difference in the frequency of B and T cells in the PB and peritoneum in *Gata2*^{G320D/+} and *Gata2*^{G320D/G320D} mice compared to WT (Figure 5.7B-C).

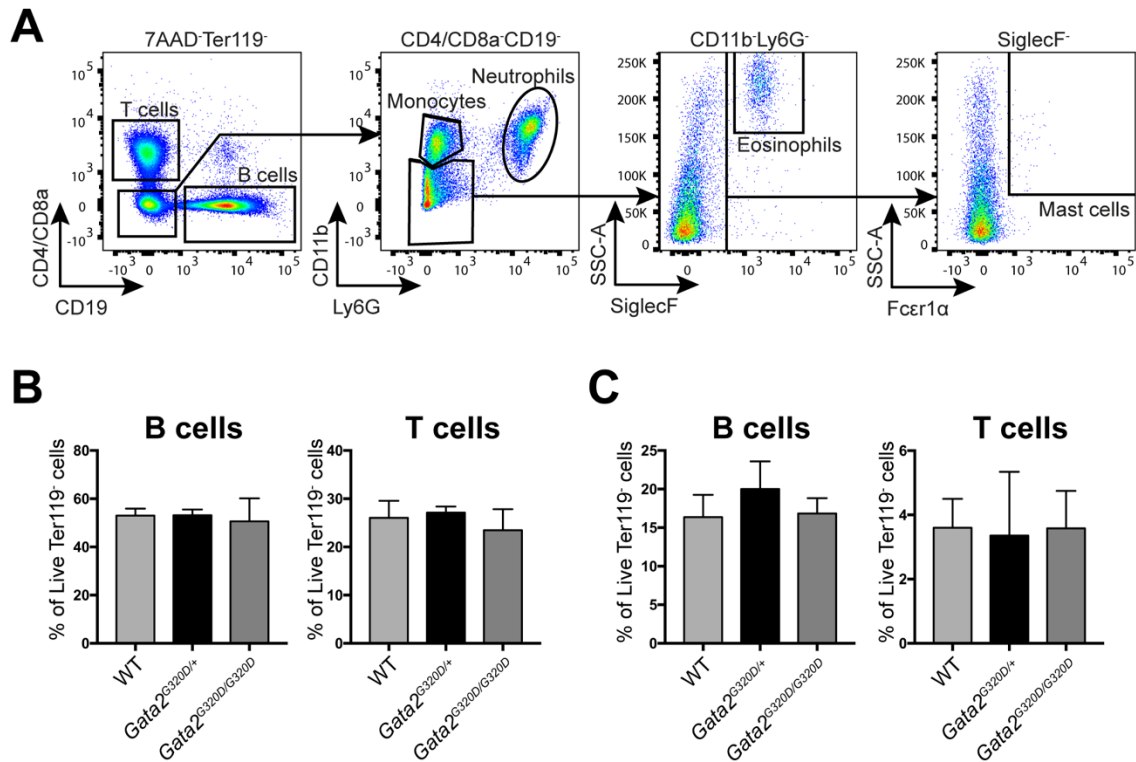


Figure 5.7 *Gata2* mutation does not affect B and T cells

(A) Myeloid, B and T cell gating strategy. (B-C) B and T cells in the PB (B) and peritoneum (C). Two independent experiments, 3-4 biological replicates per genotype. The results were analysed using a multiple comparison ANOVA. The results are presented as the mean $\pm$ SD.

There was no difference in mature myeloid cells in the PB (Figure 5.8A-C). However, in the peritoneum there was a decrease in mast cells in *Gata2*^{G320D/+} compared to WT (2-fold, $P = 0.0003$) (Figure 5.8D), but no difference in eosinophils, monocytes, and neutrophils (Figure 5.8E-G). Interestingly, there was a complete loss of mast cells in *Gata2*^{G320D/G320D} compared to WT ($P < 0.0001$; Figure 5.8D) and also a 2-fold decrease in eosinophils compared to WT ($P = 0.013$; Figure 5.8E), however there was no difference in monocyte and neutrophils (Figure 5.8F-G). These results suggest the *Gata2* mutation leads to a loss of *Gata1*-expressing myeloid cells.

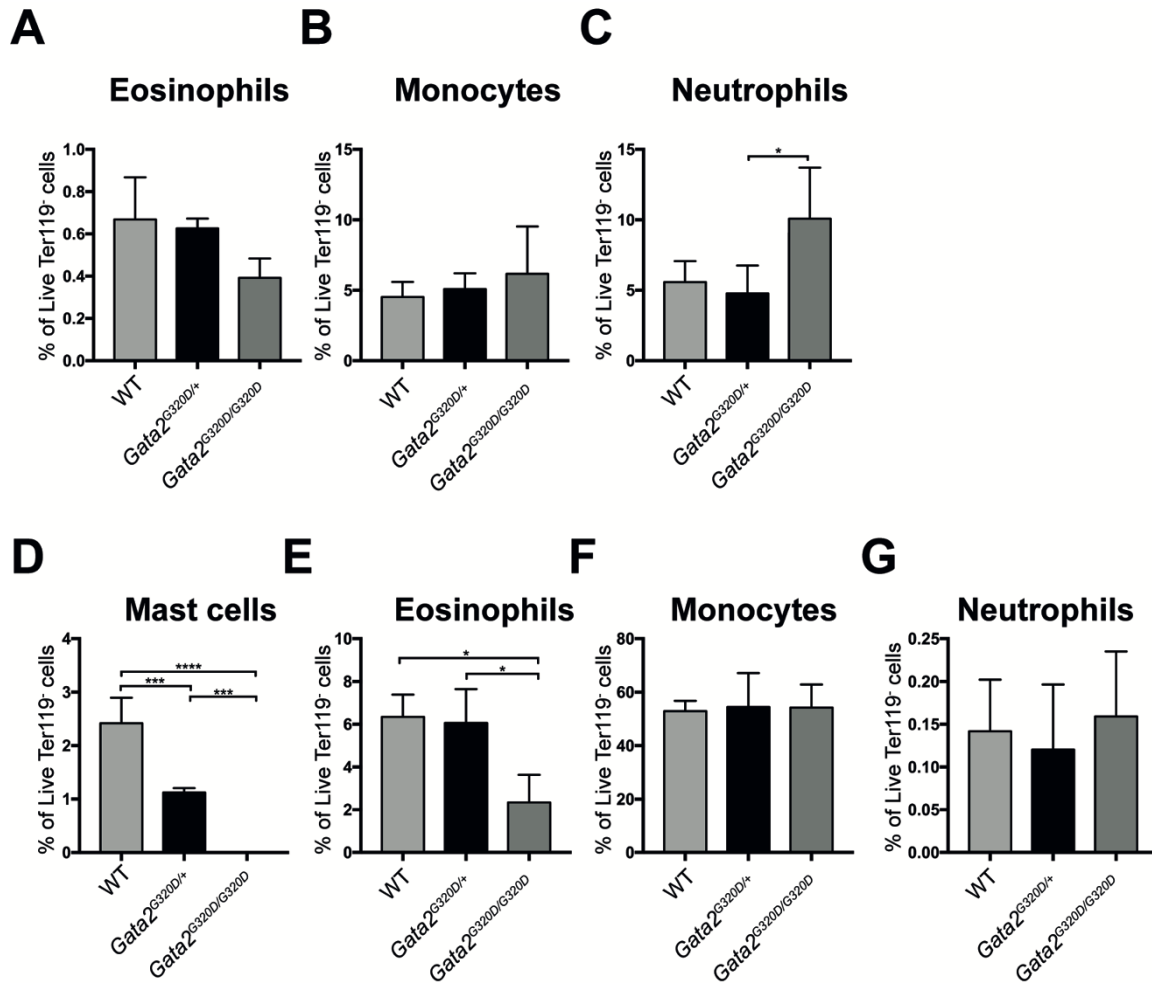


Figure 5.8 *Gata2* homozygous mutation results in a loss of mast cells

(A) Eosinophils, (B) monocytes, and (C) neutrophils in the PB. (D) Mast cells, (E) eosinophils, (F) monocytes, and (G) neutrophils in the peritoneum. Two independent experiments, 3-4 biological replicates per genotype. The results were analysed using multiple comparison ANOVA. The results are presented as the mean $\pm$ SD. *P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

The ability of EMkMPPs and EoMPs to generate myeloid cells was then tested *in vitro* (Figure 5.9A). There was a 2-fold decrease in mast cells generated by *Gata2*^{G320D/+} EMkMPPs compared to WT ($P = 0.0167$; Figure 5.9B) and an 18-fold decrease in mast cells generated by *Gata2*^{G320D/G320D} EMkMPPs compared to WT ($P < 0.0001$; Figure 5.9B). There was also complete loss of mast cells by *Gata2*^{G320D/G320D} EoMPs (Figure 5.9C).

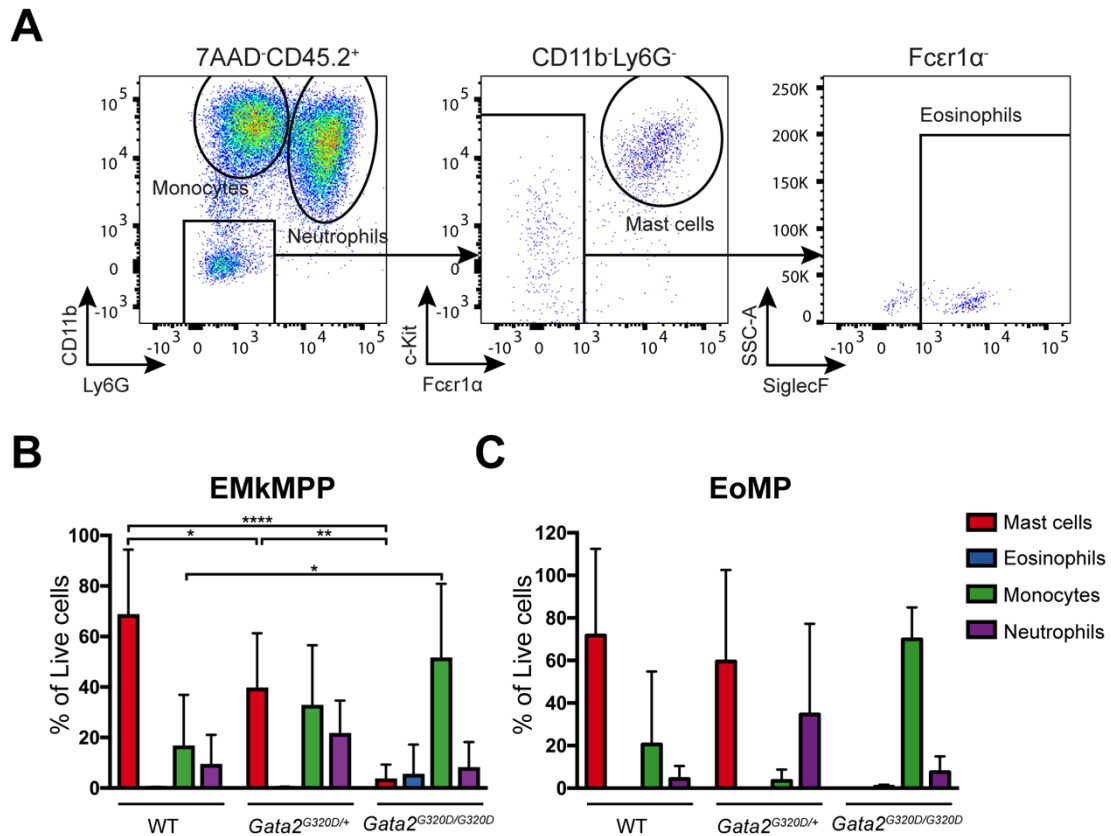


Figure 5.9 *Gata2* mutation results in a loss of mast cell generation *in vitro*

100 EMkMPP or EoMP cells were sorted and cultured for eight days and myeloid cells were analysed by FACS. (A) Gating strategy used to analyse myeloid cells generated. (B) Myeloid cells generated from cultured EMkMPPs. (C) Myeloid cells generated from cultured EoMPs. Two independent experiments. EMkMPP: WT (n = 10), *Gata2*^{G320D/+} (n = 9), *Gata2*^{G320D/G320D} (n = 6); EoMP: WT (n = 7), *Gata2*^{G320D/+} (n = 6), *Gata2*^{G320D/G320D} (n = 2). The results were analysed using multiple comparison ANOVA. The results are presented as the mean ± SD. *P<0.05 **P<0.01 ***P<0.001 ****P<0.0001.

To test whether the loss in mast cell generation was due to altered gene expression of lineage commitment, single cell multiplex quantitative real-time PCR was performed on WT and *Gata2*^{G320D/G320D} EoMPs. Transcription factor genes and genes expressed in mature cells were selected for analysis for myeloid, erythroid and lymphoid lineages. Only 2/23 genes were significantly different between *Gata2*^{G320D/G320D} and WT EoMP cells

(Figure 5.10A-C). EoMPs had a significant decrease in *Sfpi1* and *Ms4a2* expression in *Gata2*^{G320D/G320D} compared to WT (Figure 5.10B-C). There was no significant difference in *Gata2*, *Gata1* or *Cebpa* expression.

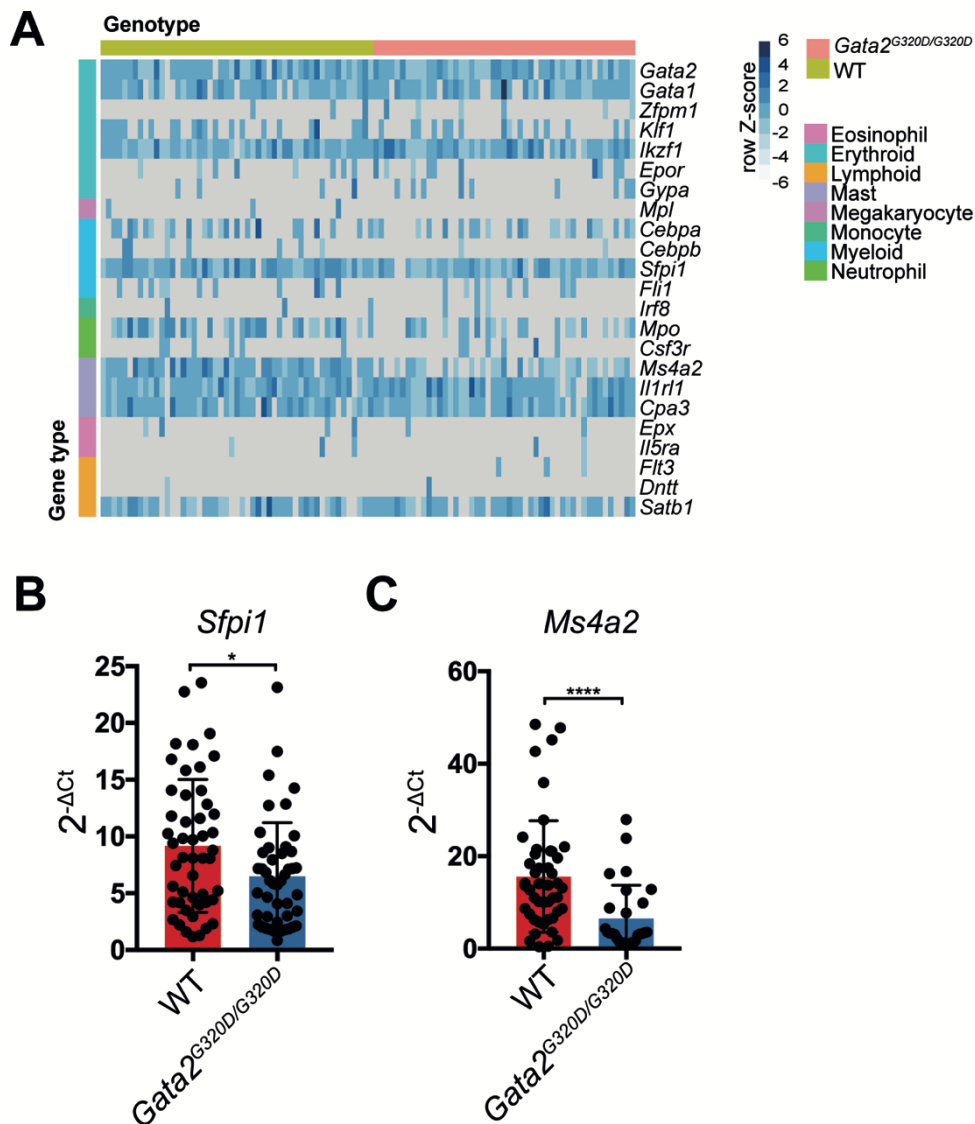


Figure 5.10 Gene expression analysis of lineage commitment genes in *Gata2* homozygous mutant EoMPs

Single EoMPs (n = 64) were sorted from WT and *Gata2*^{G320D/G320D} BM and used to perform a multiplex quantitative real-time PCR. (A) Heatmap depicting $2^{-\Delta Ct}$ normalised to *Hprt*. (B-C) $2^{-\Delta Ct}$ normalised to *Hprt* for *Sfpi1* (B) and *Ms4a2* (C). The results were analysed using a

Mann-Whitney test. The results are presented as the mean $\pm$ SD. * $P < 0.05$ ** $P < 0.01$
*** $P < 0.001$ **** $P < 0.0001$.

5.2.3 *Gata2* homozygous mutation leads to block of mast cell potential at the basophil-mast cell progenitor stage

To determine the stage of the haematopoietic hierarchy mast cell differentiation was blocked in *Gata2*^{G320D/G320D} mice, the basophil-mast cell progenitor (BMCP) in the spleen (Arinobu et al., 2005) was analysed by flow cytometry (Figure 5.11A). There was no difference in the frequency of BMCPs in *Gata2*^{G320D/+} compared to WT, however there was a 2-fold increase in BMCP in *Gata2*^{G320D/G320D} compared to WT ($P = 0.0246$; Figure 5.11B-C). These findings suggest that mast cell development was blocked at the BMCP stage leading to an accumulation of BMCPs.

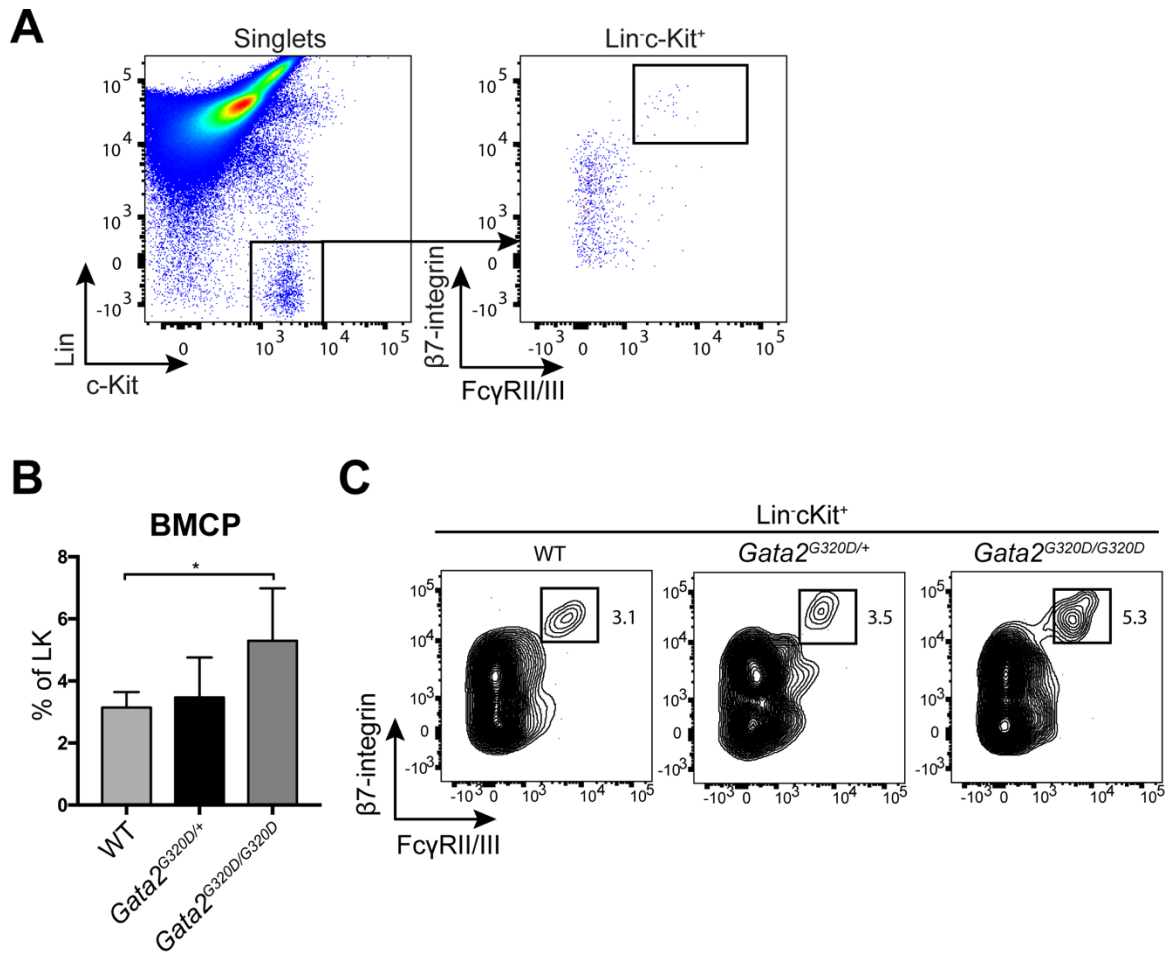


Figure 5.11 Block in mast cell development at the BMCP stage in *Gata2*^{G320D/G320D} mice

(A) BMCP gating strategy. (B) Spleen BMCPs as a percentage of LK cells. (C) Representative FACS plots of BMCPs as a percentage of LK cells. Three independent experiments 4-7 biological replicates per genotype. The results were analysed using multiple comparison ANOVA. The results are presented as the mean $\pm$ SD. * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ **** $P < 0.0001$.

To gain a molecular insight into the block of mast cell potential, 50 BMCPs from three WT and *Gata2*^{G320D/G320D} mice were sorted for RNA-sequencing. 25 down-regulated genes and two up-regulated genes were identified in *Gata2*^{G320D/G320D} mice compared to WT. Of the down-regulated genes, five encoded mast cell proteases (*Mcpt1*, *Mcpt2*, *Mcpt4*, *Mcpt8*, *Mcpt-ps1*) (Figure 5.12A). Using previously published chromatin immunoprecipitation (ChIP)-sequencing data of Gata2 in primary murine mast cells (Moignard et al., 2013), Gata2 was shown to directly bind within 2KB of the transcription start site (TSS) for *Mcpt4* and *Mcpt-ps1* (Figure 5.12B). This indicates that *Gata2* homozygous mutation leads to loss of binding to *Mcpt4* and *Mcpt-ps1* and subsequently to the down-regulation of their gene expression.

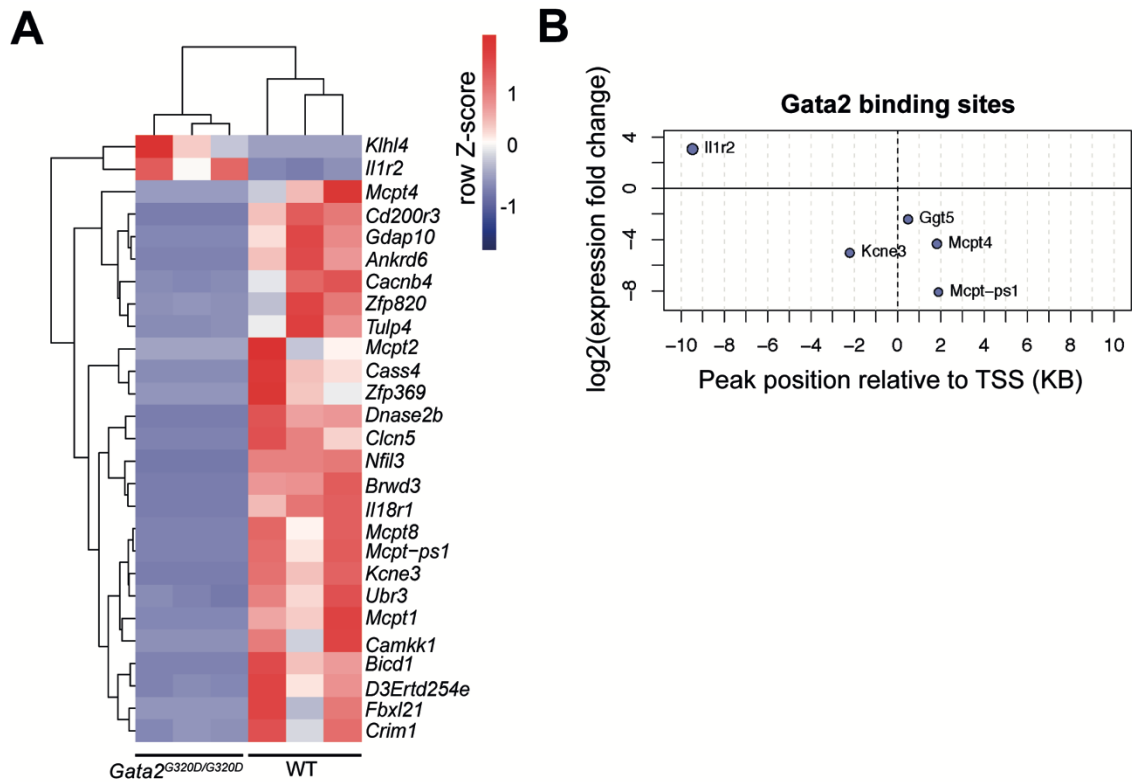


Figure 5.12 *Gata2* homozygous mutation leads to a loss of mast cell protease gene expression

(A) Heatmap illustrating the hierarchical clustering of *Gata2*^{G320D/G320D} and WT spleen BMCPs (n = 3). Differentially expressed genes shown as RPKM (padj < 0.05). (B) Bubble plot representation of Gata2 binding sites around differentially regulated genes, identified by RNA-sequencing. Distance from the TSS (x-axis) and log₂ fold change in gene expression between *Gata2*^{G320D/G320D} and WT (y-axis). Size of bubble correlates with size of peak.

Collectively, these results suggest *Gata2* homozygous mutation leads a block in mast cell development at the BMCP stage with a loss of Gata2 binding to mast cell genes and subsequent loss of mast cell gene expression.

5.3 Discussion

Mutations in *GATA2* have been recently identified in numerous haematological neoplasms. *Gata2* knock-out has been used to study the effects of loss of *Gata2* function however, there are currently no studies on the effect of *Gata2* mutations on haematopoiesis *in vivo*. This study shows, similar to *Gata2* knock-out, ZF1 *Gata2* homozygous mutation leads to loss of HSC function and mast cell development.

5.3.1 *Gata2* homozygous mutant phenotype is similar to *Gata2* knock-out

Consistent with findings from the *Gata2* knock-out (de Pater et al., 2013; F. Y. Tsai et al., 1994; F. Y. Tsai & Orkin, 1997), *Gata2*^{G320D/G320D} mice also display a loss of phenotypic and functional HSCs, and mast cells, *in vivo* and *in vitro*. This indicates that mutations in ZF1 may behave as a loss-of-function mutation. *Gata2*^{G320D/G320D} mice also displayed urogenital and eye defects. It has been previously reported that *Gata2* loss can cause multiple excretory and sex organ developmental defects during embryogenesis with females displaying abnormally developed uteri and vaginas (Zhou et al., 1998), however there has been no previous reports of eye defects in *Gata2* null mice.

Ohmori *et al.*, studied an inducible *Gata2* knock-out of BM derived mast cells *in vitro*. They took cKit⁺FcεRIα⁺ cells from murine BM and treated with 4-hydroxytamoxifen (4-OHT) to induce *Gata2* knock-out *in vitro*. After 4-OHT treatment they found a decrease in mast cell and basophil marks and an increase in monocyte, neutrophil and eosinophil markers. In this study there was also a decrease in mast cells, but also a decrease in eosinophils, and no difference in monocytes or neutrophils in *Gata2* mutant mice compared to WT *in vivo*.

Ohmori *et al.*, showed *in vitro* culture of BM derived mast cells, after treatment, were able to differentiate into macrophage and neutrophils but not eosinophils. This indicates that the cells lose potential to make *Gata1*-expressing myeloid cells (mast cell/basophils and eosinophils) and gain the ability to make myeloid cells that lack *Gata1*-expression (monocytes and neutrophils). This is also true in *Gata2*^{G302D/G320D} mice, as there is a loss of *Gata1*-expressing mast cells and eosinophils, and an increase in monocyte generation *in vitro*.

Ohmori *et al.*, showed BM derived mast cells after treatment had a decrease in expression of *Gata1*, *Gata2*, and *Kit*, and an increase in expression of *Cebpa*, *Sfpi1*, *Itgam*, and *Mpo*. When analysing the gene expression in *Gata2*^{G320D/G320D} EoMPs in this study, there was no difference in *Cebpa* and *Mpo* in *Gata2*^{G320D/G320D} EoMPs compared to WT. Conversely, this study showed a decrease in *Sfpi1* expression in *Gata2*^{G320D/G320D} EoMPs compared to WT. *Sfpi1* has been shown to cooperate with *Gata2* in mast cell generation (Walsh et al., 2002), therefore loss of mast cell development may start to occur in the EoMPs through down-regulation of *Sfpi1* in *Gata2*^{G320D/G320D} mice.

Finally, Ohmori *et al.*, showed a down-regulation of mast cell specific genes including *Mcpt4*, *Mcpt9*, *Mcpt2*, *Tpsg1* and *Ms4a2* when treated (Ohmori et al., 2015). This is in accordance with this study as *Gata2*^{G320D/G320D} BMCPs showed a decrease in mast cell proteases compared to WT, and *Ms4a2* was found to be down-regulated in *Gata2*^{G320D/G320D} EoMPs compared to WT.

In summary, this study supports the idea that *Gata2*^{G320D} when homozygous may result in loss-of-function. Previous studies have shown in *Gata2* heterozygous knock-out mice, and heterozygous ZF2 *GATA2* mutations in humans, display a haploinsufficiency phenotype (Cortes-Lavaud et al., 2015; Ling et al., 2004; Ostergaard et al., 2011; Rodrigues et al., 2005), however *Gata2*^{G320D/+} mice were relatively unperturbed with no

effect on HSC function and no alteration in *Gata2* gene expression. Therefore ZF1 *Gata2* heterozygous mutation does not result in haploinsufficiency.

5.3.2 Comparison with previous functional studies of Gata2G320D

Using a reporter gene assay in HEK 293T cells Greif *et al.*, demonstrated G320D may lead to increased GATA2 activation, but showed a reduced capacity to enhance CEBPA-dependent activation of transcription. These findings are inconsistent with this study as *Gata2*^{G320D/G320D} did not result in a difference in *Gata2* or *Cebpa* gene expression in EoMPs or BMCPs compared to WT.

5.4 Conclusion

Gata2G320D results in loss of Gata2 function only when the mutation is homozygous as there is no difference in haematopoiesis when mutated at heterozygous level.

Considering *GATA2* mutations in AML patients with biallelic *CEBPA* mutations are heterozygous (Greif, Dufour, et al., 2012), this data supports the findings in chapter four that *GATA2* and *CEBPA* mutations must work synergistically to cause leukaemic development. Heterozygous ZF2 *GATA2* mutations results in haploinsufficiency. As there is no significant effect of heterozygous ZF1 *GATA2* mutation on haematopoiesis suggests that the ZF1 and ZF2 effect Gata2 function differently.

It remains unclear how G320D causes loss-of-function of Gata2. Gene expression analysis of EoMPs and BMCPs did not reveal major changes in gene expression of important transcription factors for lineage specification, other than the down-regulation of *Spfi1* in EoMPs. As the ZF1 main function is protein-protein interaction, ChIP analysis

may reveal loss of important protein interactions that contribute to loss of mast cell development.

CONCLUSION

The study presented in chapter three demonstrates the detrimental effect activated K-Ras has on HSCs expressing the pre-leukaemic fusion protein Aml1-ETO. These findings provide a mechanistic basis for the observed absence of *KRAS*, and potentially other signalling mutations, in the pre-leukaemic HSC compartment, helping to explain a fundamental aspect of AML biology.

The studies presented in chapters four and five provide a model of AEL, demonstrating that *Cebpa* mutations reprogram the NMPs to acquire ectopic erythroid potential, while the *Gata2* mutation synergises with the *Cebpa* mutations to impair erythroid differentiation progression. Bi-lineage transformation results from erythroid and myeloid blasts generated from the neomorphic neutrophil-erythroid progenitor. This study demonstrates how a bi-potent LSC can be formed by repressing the natural cell fate and also re-establishing a new program in the NMPs. This novel model of AEL may also offer a route to better understanding of the human disease. Identification and characterisation of the LSCs is an important step in developing new therapies that can eliminate this crucial population that sustains the disease. This model could also be used to identify novel pathways or surface markers that can subsequently be used for targeted treatment.

Together these studies also highlight the importance of using knock-in models to accurately model disease. In the case of the *Aml1*^{ETO/+};*Kras*^{G12D/+} model, the study emphasises the importance of the temporal and cellular context of the mutations necessary for leukaemia development. To my knowledge, no studies so far have been able to fully recapitulate all aspects of leukaemia development. However the recent advances in CRISPR-Cas9 genome editing technology will continue to provide more precise tools for modelling leukaemia.

BIBLIOGRAPHY

- Abdel-Wahab, O., Adli, M., LaFave, L. M., Gao, J., Hricik, T., Shih, A. H., . . . Levine, R. L. (2012). ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell*, 22(2), 180-193. doi:10.1016/j.ccr.2012.06.032
- Abdel-Wahab, O., Gao, J., Adli, M., Dey, A., Trimarchi, T., Chung, Y. R., . . . Levine, R. L. (2013). Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. *J Exp Med*, 210(12), 2641-2659. doi:10.1084/jem.20131141
- Adolfsson, J., Borge, O. J., Bryder, D., Theilgaard-Monch, K., Astrand-Grundstrom, I., Sitnicka, E., . . . Jacobsen, S. E. (2001). Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity*, 15(4), 659-669.
- Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., . . . Jacobsen, S. E. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell*, 121(2), 295-306. doi:10.1016/j.cell.2005.02.013
- Ahn, M. Y., Huang, G., Bae, S. C., Wee, H. J., Kim, W. Y., & Ito, Y. (1998). Negative regulation of granulocytic differentiation in the myeloid precursor cell line 32Dcl3 by ear-2, a mammalian homolog of Drosophila seven-up, and a chimeric leukemogenic gene, AML1/ETO. *Proc Natl Acad Sci U S A*, 95(4), 1812-1817.
- Alexander, T. B., Gu, Z., Iacobucci, I., Dickerson, K., Choi, J. K., Xu, B., . . . Mullighan, C. G. (2018). The genetic basis and cell of origin of mixed phenotype acute leukaemia. *Nature*. doi:10.1038/s41586-018-0436-0
- Arber, D. A. (2017). Revisiting erythroleukemia. *Curr Opin Hematol*, 24(2), 146-151. doi:10.1097/MOH.0000000000000314
- Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., . . . Vardiman, J. W. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*, 127(20), 2391-2405. doi:10.1182/blood-2016-03-643544
- Arenillas, L., Calvo, X., Luno, E., Senent, L., Alonso, E., Ramos, F., . . . Florensa, L. (2016). Considering Bone Marrow Blasts From Nonerythroid Cellularity Improves the Prognostic Evaluation of Myelodysplastic Syndromes. *J Clin Oncol*, 34(27), 3284-3292. doi:10.1200/JCO.2016.66.9705
- Arinobu, Y., Iwasaki, H., Gurish, M. F., Mizuno, S., Shigematsu, H., Ozawa, H., . . . Akashi, K. (2005). Developmental checkpoints of the basophil/mast cell lineages in adult murine hematopoiesis. *Proc Natl Acad Sci U S A*, 102(50), 18105-18110. doi:10.1073/pnas.0509148102
- Avellino, R., & Delwel, R. (2017). Expression and regulation of C/EBPalpha in normal myelopoiesis and in malignant transformation. *Blood*, 129(15), 2083-2091. doi:10.1182/blood-2016-09-687822
- Avellino, R., Havermans, M., Erpelinck, C., Sanders, M. A., Hoogenboezem, R., van de Werken, H. J., . . . Delwel, R. (2016). An autonomous CEBPA enhancer specific for myeloid-lineage priming and neutrophilic differentiation. *Blood*, 127(24), 2991-3003. doi:10.1182/blood-2016-01-695759

- Barrett, N. A., Malouf, C., Kapeni, C., Bacon, W. A., Giotopoulos, G., Jacobsen, S. E. W., . . . Ottersbach, K. (2016). Mll-AF4 Confers Enhanced Self-Renewal and Lymphoid Potential during a Restricted Window in Development. *Cell Rep*, 16(4), 1039-1054. doi:10.1016/j.celrep.2016.06.046
- Beck, D., Thoms, J. A., Perera, D., Schutte, J., Unnikrishnan, A., Knezevic, K., . . . Pimanda, J. E. (2013). Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes. *Blood*, 122(14), e12-22. doi:10.1182/blood-2013-03-490425
- Ben David, Y., Prideaux, V. R., Chow, V., Benchimol, S., & Bernstein, A. (1988). Inactivation of the p53 oncogene by internal deletion or retroviral integration in erythroleukemic cell lines induced by Friend leukemia virus. *Oncogene*, 3(2), 179-185.
- Ben-David, Y., Giddens, E. B., & Bernstein, A. (1990). Identification and mapping of a common proviral integration site Fli-1 in erythroleukemia cells induced by Friend murine leukemia virus. *Proc Natl Acad Sci U S A*, 87(4), 1332-1336.
- Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A., Gralnick, H. R., & Sultan, C. (1976). Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol*, 33(4), 451-458.
- Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A., Gralnick, H. R., & Sultan, C. (1991). Proposal for the recognition of minimally differentiated acute myeloid leukaemia (AML-MO). *Br J Haematol*, 78(3), 325-329.
- Bereshchenko, O., Mancini, E., Moore, S., Bilbao, D., Mansson, R., Luc, S., . . . Nerlov, C. (2009). Hematopoietic stem cell expansion precedes the generation of committed myeloid leukemia-initiating cells in C/EBPalpha mutant AML. *Cancer Cell*, 16(5), 390-400. doi:10.1016/j.ccr.2009.09.036
- Bigley, V., Haniffa, M., Doulatov, S., Wang, X. N., Dickinson, R., McGovern, N., . . . Collin, M. (2011). The human syndrome of dendritic cell, monocyte, B and NK lymphoid deficiency. *J Exp Med*, 208(2), 227-234. doi:10.1084/jem.20101459
- Bonadies, N., Foster, S. D., Chan, W. I., Kvinlaug, B. T., Spensberger, D., Dawson, M. A., . . . Gottgens, B. (2011). Genome-wide analysis of transcriptional reprogramming in mouse models of acute myeloid leukaemia. *PLoS One*, 6(1), e16330. doi:10.1371/journal.pone.0016330
- Bonnet, D., & Dick, J. E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med*, 3(7), 730-737.
- Bos, J. L. (1989). ras oncogenes in human cancer: a review. *Cancer Res*, 49(17), 4682-4689.
- Braun, B. S., Tuveson, D. A., Kong, N., Le, D. T., Kogan, S. C., Rozmus, J., . . . Shannon, K. M. (2004). Somatic activation of oncogenic Kras in hematopoietic cells initiates a rapidly fatal myeloproliferative disorder. *Proc Natl Acad Sci U S A*, 101(2), 597-602. doi:10.1073/pnas.0307203101
- Bresnick, E. H., Lee, H. Y., Fujiwara, T., Johnson, K. D., & Keles, S. (2010). GATA switches as developmental drivers. *J Biol Chem*, 285(41), 31087-31093. doi:10.1074/jbc.R110.159079
- Briegleb, K., Lim, K. C., Plank, C., Beug, H., Engel, J. D., & Zenke, M. (1993). Ectopic expression of a conditional GATA-2/estrogen receptor chimera arrests erythroid differentiation in a hormone-dependent manner. *Genes Dev*, 7(6), 1097-1109.

- Busque, L., Patel, J. P., Figueroa, M. E., Vasanthakumar, A., Provost, S., Hamilou, Z., . . . Levine, R. L. (2012). Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat Genet*, 44(11), 1179-1181. doi:10.1038/ng.2413
- Cai, X., Gaudet, J. J., Mangan, J. K., Chen, M. J., De Obaldia, M. E., Oo, Z., . . . Speck, N. A. (2011). Runx1 loss minimally impacts long-term hematopoietic stem cells. *PLoS One*, 6(12), e28430. doi:10.1371/journal.pone.0028430
- Calabi, F., Pannell, R., & Pavloska, G. (2001). Gene targeting reveals a crucial role for MTG8 in the gut. *Mol Cell Biol*, 21(16), 5658-5666. doi:10.1128/MCB.21.16.5658-5666.2001
- Calero-Nieto, F. J., Ng, F. S., Wilson, N. K., Hannah, R., Moignard, V., Leal-Cervantes, A. I., . . . Gottgens, B. (2014). Key regulators control distinct transcriptional programmes in blood progenitor and mast cells. *EMBO J*, 33(11), 1212-1226. doi:10.1002/embj.201386825
- Calkhoven, C. F., Muller, C., & Leutz, A. (2000). Translational control of C/EBPalpha and C/EBPbeta isoform expression. *Genes Dev*, 14(15), 1920-1932.
- Calvo, X., Arenillas, L., Luno, E., Senent, L., Arnan, M., Ramos, F., . . . Florensa, L. (2016). Erythroleukemia shares biological features and outcome with myelodysplastic syndromes with excess blasts: a rationale for its inclusion into future classifications of myelodysplastic syndromes. *Mod Pathol*, 29(12), 1541-1551. doi:10.1038/modpathol.2016.146
- Cancer Genome Atlas Research, N. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*, 368(22), 2059-2074. doi:10.1056/NEJMoa1301689
- Carmichael, C. L., Wilkins, E. J., Bengtsson, H., Horwitz, M. S., Speed, T. P., Vincent, P. C., . . . Scott, H. S. (2010). Poor prognosis in familial acute myeloid leukaemia with combined biallelic CEBPA mutations and downstream events affecting the ATM, FLT3 and CDX2 genes. *Br J Haematol*, 150(3), 382-385. doi:10.1111/j.1365-2141.2010.08204.x
- Carnevali, L. S., Scognamiglio, R., Cabezas-Wallscheid, N., Rahmig, S., Laurenti, E., Masuda, K., . . . Trumpp, A. (2014). Improved HSC reconstitution and protection from inflammatory stress and chemotherapy in mice lacking granzyme B. *J Exp Med*, 211(5), 769-779. doi:10.1084/jem.20131072
- Carrelha, J., Meng, Y., Kettyle, L. M., Luis, T. C., Norfo, R., Alcolea, V., . . . Jacobsen, S. E. W. (2018). Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature*, 554(7690), 106-111. doi:10.1038/nature25455
- Castilla, L. H., Garrett, L., Adya, N., Orlic, D., Dutra, A., Anderson, S., . . . Liu, P. P. (1999). The fusion gene Cbfb-MYH11 blocks myeloid differentiation and predisposes mice to acute myelomonocytic leukaemia. *Nat Genet*, 23(2), 144-146. doi:10.1038/13776
- Cervera, N., Carbuccia, N., Garnier, S., Guille, A., Adelaide, J., Murati, A., . . . Gelsi-Boyer, V. (2016). Molecular characterization of acute erythroid leukemia (M6-AML) using targeted next-generation sequencing. *Leukemia*, 30(4), 966-970. doi:10.1038/leu.2015.198
- Cervera, N., Carbuccia, N., Mozziconacci, M. J., Adelaide, J., Garnier, S., Guille, A., . . . Gelsi-Boyer, V. (2017). Revisiting gene mutations and prognosis of ex-M6a-acute

- erythroid leukemia with regard to the new WHO classification. *Blood Cancer J*, 7(8), e594. doi:10.1038/bcj.2017.68
- Chan, G., Kalaitzidis, D., Usenko, T., Kutok, J. L., Yang, W., Mohi, M. G., & Neel, B. G. (2009). Leukemogenic Ptpn11 causes fatal myeloproliferative disorder via cell-autonomous effects on multiple stages of hematopoiesis. *Blood*, 113(18), 4414-4424. doi:10.1182/blood-2008-10-182626
- Chan, I. T., Kutok, J. L., Williams, I. R., Cohen, S., Kelly, L., Shigematsu, H., . . . Gilliland, D. G. (2004). Conditional expression of oncogenic K-ras from its endogenous promoter induces a myeloproliferative disease. *J Clin Invest*, 113(4), 528-538. doi:10.1172/JCI20476
- Chang, Y. I., You, X., Kong, G., Ranheim, E. A., Wang, J., Du, J., . . . Zhang, J. (2015). Loss of Dnmt3a and endogenous Kras(G12D/+) cooperate to regulate hematopoietic stem and progenitor cell functions in leukemogenesis. *Leukemia*, 29(9), 1847-1856. doi:10.1038/leu.2015.85
- Chen, C. Y., Lin, L. I., Tang, J. L., Ko, B. S., Tsay, W., Chou, W. C., . . . Tien, H. F. (2007). RUNX1 gene mutation in primary myelodysplastic syndrome--the mutation can be detected early at diagnosis or acquired during disease progression and is associated with poor outcome. *Br J Haematol*, 139(3), 405-414. doi:10.1111/j.1365-2141.2007.06811.x
- Chen, M. J., Yokomizo, T., Zeigler, B. M., Dzierzak, E., & Speck, N. A. (2009). Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature*, 457(7231), 887-891. doi:10.1038/nature07619
- Chen, S. J., Shen, Y., & Chen, Z. (2013). A panoramic view of acute myeloid leukemia. *Nat Genet*, 45(6), 586-587. doi:10.1038/ng.2651
- Chou, F. S., Wunderlich, M., Griesinger, A., & Mulloy, J. C. (2011). N-Ras(G12D) induces features of stepwise transformation in preleukemic human umbilical cord blood cultures expressing the AML1-ETO fusion gene. *Blood*, 117(7), 2237-2240. doi:10.1182/blood-2010-01-264119
- Christiansen, D. H., Andersen, M. K., & Pedersen-Bjergaard, J. (2004). Mutations of AML1 are common in therapy-related myelodysplasia following therapy with alkylating agents and are significantly associated with deletion or loss of chromosome arm 7q and with subsequent leukemic transformation. *Blood*, 104(5), 1474-1481. doi:10.1182/blood-2004-02-0754
- Chu, S. H., Heiser, D., Li, L., Kaplan, I., Collector, M., Huso, D., . . . Small, D. (2012). FLT3-ITD knockin impairs hematopoietic stem cell quiescence/homeostasis, leading to myeloproliferative neoplasm. *Cell Stem Cell*, 11(3), 346-358. doi:10.1016/j.stem.2012.05.027
- Collin, M., Dickinson, R., & Bigley, V. (2015). Haematopoietic and immune defects associated with GATA2 mutation. *Br J Haematol*, 169(2), 173-187. doi:10.1111/bjh.13317
- Conway O'Brien, E., Prideaux, S., & Chevassut, T. (2014). The epigenetic landscape of acute myeloid leukemia. *Adv Hematol*, 2014, 103175. doi:10.1155/2014/103175
- Corces-Zimmerman, M. R., Hong, W. J., Weissman, I. L., Medeiros, B. C., & Majeti, R. (2014). Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proc Natl Acad Sci U S A*, 111(7), 2548-2553. doi:10.1073/pnas.1324297111

- Corces-Zimmerman, M. R., & Majeti, R. (2014). Pre-leukemic evolution of hematopoietic stem cells: the importance of early mutations in leukemogenesis. *Leukemia*, 28(12), 2276-2282. doi:10.1038/leu.2014.211
- Cortes-Lavaud, X., Landecho, M. F., Maicas, M., Urquiza, L., Merino, J., Moreno-Miralles, I., & Otero, M. D. (2015). GATA2 germline mutations impair GATA2 transcription, causing haploinsufficiency: functional analysis of the p.Arg396Gln mutation. *J Immunol*, 194(5), 2190-2198. doi:10.4049/jimmunol.1401868
- Cozzio, A., Passegue, E., Ayton, P. M., Karsunky, H., Cleary, M. L., & Weissman, I. L. (2003). Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors. *Genes Dev*, 17(24), 3029-3035. doi:10.1101/gad.1143403
- Davis, J. N., McGhee, L., & Meyers, S. (2003). The ETO (MTG8) gene family. *Gene*, 303, 1-10.
- de Guzman, C. G., Warren, A. J., Zhang, Z., Gartland, L., Erickson, P., Drabkin, H., . . . Klug, C. A. (2002). Hematopoietic stem cell expansion and distinct myeloid developmental abnormalities in a murine model of the AML1-ETO translocation. *Mol Cell Biol*, 22(15), 5506-5517.
- de Pater, E., Kaimakis, P., Vink, C. S., Yokomizo, T., Yamada-Inagawa, T., van der Linden, R., . . . Dzierzak, E. (2013). Gata2 is required for HSC generation and survival. *J Exp Med*, 210(13), 2843-2850. doi:10.1084/jem.20130751
- De Ruiter, N. D., Burgering, B. M., & Bos, J. L. (2001). Regulation of the Forkhead transcription factor AFX by Ral-dependent phosphorylation of threonines 447 and 451. *Mol Cell Biol*, 21(23), 8225-8235. doi:10.1128/MCB.21.23.8225-8235.2001
- Di Tullio, A., Vu Manh, T. P., Schubert, A., Castellano, G., Mansson, R., & Graf, T. (2011). CCAAT/enhancer binding protein alpha (C/EBP(alpha))-induced transdifferentiation of pre-B cells into macrophages involves no overt retrodifferentiation. *Proc Natl Acad Sci U S A*, 108(41), 17016-17021. doi:10.1073/pnas.1112169108
- Dickinson, R. E., Griffin, H., Bigley, V., Reynard, L. N., Hussain, R., Haniffa, M., . . . Collin, M. (2011). Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. *Blood*, 118(10), 2656-2658. doi:10.1182/blood-2011-06-360313
- Ding, L., Ley, T. J., Larson, D. E., Miller, C. A., Koboldt, D. C., Welch, J. S., . . . DiPersio, J. F. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*, 481(7382), 506-510. doi:10.1038/nature10738
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., . . . Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15-21. doi:10.1093/bioinformatics/bts635
- Dohner, H., Estey, E., Grimwade, D., Amadori, S., Appelbaum, F. R., Buchner, T., . . . Bloomfield, C. D. (2017). Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*, 129(4), 424-447. doi:10.1182/blood-2016-08-733196
- Dohner, H., Estey, E. H., Amadori, S., Appelbaum, F. R., Buchner, T., Burnett, A. K., . . . European, L. (2010). Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood*, 115(3), 453-474. doi:10.1182/blood-2009-07-235358

- Dohner, H., Weisdorf, D. J., & Bloomfield, C. D. (2015). Acute Myeloid Leukemia. *N Engl J Med*, 373(12), 1136-1152. doi:10.1056/NEJMra1406184
- Dore, L. C., Chlon, T. M., Brown, C. D., White, K. P., & Crispino, J. D. (2012). Chromatin occupancy analysis reveals genome-wide GATA factor switching during hematopoiesis. *Blood*, 119(16), 3724-3733. doi:10.1182/blood-2011-09-380634
- Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. *Nat Rev Cancer*, 3(1), 11-22. doi:10.1038/nrc969
- Drissen, R., Buza-Vidas, N., Woll, P., Thongjuea, S., Gambardella, A., Giustacchini, A., . . . Nerlov, C. (2016). Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat Immunol*, 17(6), 666-676. doi:10.1038/ni.3412
- Du, Y., Jenkins, N. A., & Copeland, N. G. (2005). Insertional mutagenesis identifies genes that promote the immortalization of primary bone marrow progenitor cells. *Blood*, 106(12), 3932-3939. doi:10.1182/blood-2005-03-1113
- Dufour, A., Schneider, F., Metzeler, K. H., Hoster, E., Schneider, S., Zellmeier, E., . . . Spiekermann, K. (2010). Acute myeloid leukemia with biallelic CEBPA gene mutations and normal karyotype represents a distinct genetic entity associated with a favorable clinical outcome. *J Clin Oncol*, 28(4), 570-577. doi:10.1200/JCO.2008.21.6010
- El-Hadya, S. B. M., Almasry, E., & ElHefneyb, A. M. (2012). Detection of Cx43 and P27 in acute myeloid leukemia patients with t(8;21). *The Egyptian Society of Haematology*, 37(2), 1110-1067.
- Essers, M. A., Offner, S., Blanco-Bose, W. E., Waibler, Z., Kalinke, U., Duchosal, M. A., & Trumpp, A. (2009). IFN α activates dormant haematopoietic stem cells in vivo. *Nature*, 458(7240), 904-908. doi:10.1038/nature07815
- Estey, E., & Dohner, H. (2006). Acute myeloid leukaemia. *Lancet*, 368(9550), 1894-1907. doi:10.1016/S0140-6736(06)69780-8
- Fasan, A., Eder, C., Haferlach, C., Grossmann, V., Kohlmann, A., Dicker, F., . . . Schnittger, S. (2013). GATA2 mutations are frequent in intermediate-risk karyotype AML with biallelic CEBPA mutations and are associated with favorable prognosis. *Leukemia*, 27(2), 482-485. doi:10.1038/leu.2012.174
- Fasan, A., Haferlach, C., Alpermann, T., Jeromin, S., Grossmann, V., Eder, C., . . . Schnittger, S. (2014). The role of different genetic subtypes of CEBPA mutated AML. *Leukemia*, 28(4), 794-803. doi:10.1038/leu.2013.273
- Federzoni, E. A., Humbert, M., Torbett, B. E., Behre, G., Fey, M. F., & Tschan, M. P. (2014). CEBPA-dependent HK3 and KLF5 expression in primary AML and during AML differentiation. *Sci Rep*, 4, 4261. doi:10.1038/srep04261
- Friedman, A. D. (2002). Transcriptional regulation of granulocyte and monocyte development. *Oncogene*, 21(21), 3377-3390. doi:10.1038/sj.onc.1205324
- Friedman, A. D. (2007). Transcriptional control of granulocyte and monocyte development. *Oncogene*, 26(47), 6816-6828. doi:10.1038/sj.onc.1210764
- Friedman, A. D. (2015). C/EBP α in normal and malignant myelopoiesis. *Int J Hematol*, 101(4), 330-341. doi:10.1007/s12185-015-1764-6
- Fukuyama, T., Sueoka, E., Sugio, Y., Otsuka, T., Niho, Y., Akagi, K., & Kozu, T. (2001). MTG8 proto-oncoprotein interacts with the regulatory subunit of type II cyclic AMP-dependent protein kinase in lymphocytes. *Oncogene*, 20(43), 6225-6232. doi:10.1038/sj.onc.1204794

- Ganapathi, K. A., Townsley, D. M., Hsu, A. P., Arthur, D. C., Zerbe, C. S., Cuellar-Rodriguez, J., . . . Calvo, K. R. (2015). GATA2 deficiency-associated bone marrow disorder differs from idiopathic aplastic anemia. *Blood*, *125*(1), 56-70. doi:10.1182/blood-2014-06-580340
- Gelmetti, V., Zhang, J., Fanelli, M., Minucci, S., Pelicci, P. G., & Lazar, M. A. (1998). Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. *Mol Cell Biol*, *18*(12), 7185-7191.
- Gilliland, D. G., & Griffin, J. D. (2002). The roles of FLT3 in hematopoiesis and leukemia. *Blood*, *100*(5), 1532-1542. doi:10.1182/blood-2002-02-0492
- Giustacchini, A., Thongjuea, S., Barkas, N., Woll, P. S., Povinelli, B. J., Booth, C. A. G., . . . Mead, A. J. (2017). Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nat Med*, *23*(6), 692-702. doi:10.1038/nm.4336
- Goardon, N., Marchi, E., Atzberger, A., Quek, L., Schuh, A., Soneji, S., . . . Vyas, P. (2011). Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. *Cancer Cell*, *19*(1), 138-152. doi:10.1016/j.ccr.2010.12.012
- Gombart, A. F., Hofmann, W. K., Kawano, S., Takeuchi, S., Krug, U., Kwok, S. H., . . . Koefler, H. P. (2002). Mutations in the gene encoding the transcription factor CCAAT/enhancer binding protein alpha in myelodysplastic syndromes and acute myeloid leukemias. *Blood*, *99*(4), 1332-1340.
- Grass, J. A., Boyer, M. E., Pal, S., Wu, J., Weiss, M. J., & Bresnick, E. H. (2003). GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc Natl Acad Sci U S A*, *100*(15), 8811-8816. doi:10.1073/pnas.1432147100
- Green, C. L., Tawana, K., Hills, R. K., Bodor, C., Fitzgibbon, J., Inglott, S., . . . Gale, R. E. (2013). GATA2 mutations in sporadic and familial acute myeloid leukaemia patients with CEBPA mutations. *Br J Haematol*, *161*(5), 701-705. doi:10.1111/bjh.12317
- Greif, P. A., Dufour, A., Konstandin, N. P., Ksienzyk, B., Zellmeier, E., Tizazu, B., . . . Bohlander, S. K. (2012). GATA2 zinc finger 1 mutations associated with biallelic CEBPA mutations define a unique genetic entity of acute myeloid leukemia. *Blood*, *120*(2), 395-403. doi:10.1182/blood-2012-01-403220
- Greif, P. A., Konstandin, N. P., Metzeler, K. H., Herold, T., Pasalic, Z., Ksienzyk, B., . . . Bohlander, S. K. (2012). RUNX1 mutations in cytogenetically normal acute myeloid leukemia are associated with a poor prognosis and up-regulation of lymphoid genes. *Haematologica*, *97*(12), 1909-1915. doi:10.3324/haematol.2012.064667
- Grimwade, D., Hills, R. K., Moorman, A. V., Walker, H., Chatters, S., Goldstone, A. H., . . . National Cancer Research Institute Adult Leukaemia Working, G. (2010). Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood*, *116*(3), 354-365. doi:10.1182/blood-2009-11-254441
- Groschel, S., Sanders, M. A., Hoogenboezem, R., de Wit, E., Bouwman, B. A. M., Erpelinck, C., . . . Delwel, R. (2014). A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell*, *157*(2), 369-381. doi:10.1016/j.cell.2014.02.019

- Grossmann, V., Bacher, U., Haferlach, C., Schnittger, S., Potzinger, F., Weissmann, S., . . . Haferlach, T. (2013). Acute erythroid leukemia (AEL) can be separated into distinct prognostic subsets based on cytogenetic and molecular genetic characteristics. *Leukemia*, 27(9), 1940-1943. doi:10.1038/leu.2013.144
- Grove, C. S., & Vassiliou, G. S. (2014). Acute myeloid leukaemia: a paradigm for the clonal evolution of cancer? *Dis Model Mech*, 7(8), 941-951. doi:10.1242/dmm.015974
- Guo, H., Cooper, S., & Friedman, A. D. (2016). In Vivo Deletion of the Cebpa +37 kb Enhancer Markedly Reduces Cebpa mRNA in Myeloid Progenitors but Not in Non-Hematopoietic Tissues to Impair Granulopoiesis. *PLoS One*, 11(3), e0150809. doi:10.1371/journal.pone.0150809
- Guryanova, O. A., Shank, K., Spitzer, B., Luciani, L., Koche, R. P., Garrett-Bakelman, F. E., . . . Levine, R. L. (2016). DNMT3A mutations promote anthracycline resistance in acute myeloid leukemia via impaired nucleosome remodeling. *Nat Med*, 22(12), 1488-1495. doi:10.1038/nm.4210
- Hahn, C. N., Chong, C. E., Carmichael, C. L., Wilkins, E. J., Brautigan, P. J., Li, X. C., . . . Scott, H. S. (2011). Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet*, 43(10), 1012-1017. doi:10.1038/ng.913
- Hasemann, M. S., Lauridsen, F. K., Waage, J., Jakobsen, J. S., Frank, A. K., Schuster, M. B., . . . Porse, B. T. (2014). C/EBPalpha is required for long-term self-renewal and lineage priming of hematopoietic stem cells and for the maintenance of epigenetic configurations in multipotent progenitors. *PLoS Genet*, 10(1), e1004079. doi:10.1371/journal.pgen.1004079
- Hasserjian, R. P., Zuo, Z., Garcia, C., Tang, G., Kasyan, A., Luthra, R., . . . Wang, S. A. (2010). Acute erythroid leukemia: a reassessment using criteria refined in the 2008 WHO classification. *Blood*, 115(10), 1985-1992. doi:10.1182/blood-2009-09-243964
- Heckl, D., Kowalczyk, M. S., Yudovich, D., Belizaire, R., Puram, R. V., McConkey, M. E., . . . Ebert, B. L. (2014). Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. *Nat Biotechnol*, 32(9), 941-946. doi:10.1038/nbt.2951
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., . . . 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(4), 576-589. doi:10.1016/j.molcel.2010.05.004
- Higuchi, M., O'Brien, D., Kumaravelu, P., Lenny, N., Yeoh, E. J., & Downing, J. R. (2002). Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. *Cancer Cell*, 1(1), 63-74.
- Hirabayashi, S., Wlodarski, M. W., Kozyra, E., & Niemeyer, C. M. (2017). Heterogeneity of GATA2-related myeloid neoplasms. *Int J Hematol*, 106(2), 175-182. doi:10.1007/s12185-017-2285-2
- Hou, H. A., Lin, Y. C., Kuo, Y. Y., Chou, W. C., Lin, C. C., Liu, C. Y., . . . Tien, H. F. (2015). GATA2 mutations in patients with acute myeloid leukemia-paired samples analyses show that the mutation is unstable during disease evolution. *Ann Hematol*, 94(2), 211-221. doi:10.1007/s00277-014-2208-8
- Hsu, A. P., Sampaio, E. P., Khan, J., Calvo, K. R., Lemieux, J. E., Patel, S. Y., . . . Holland, S. M. (2011). Mutations in GATA2 are associated with the autosomal dominant and

- sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. *Blood*, 118(10), 2653-2655. doi:10.1182/blood-2011-05-356352
- Ichikawa, M., Asai, T., Saito, T., Seo, S., Yamazaki, I., Yamagata, T., . . . 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(3), 299-304. doi:10.1038/nm997
- Jackson, E. L., Willis, N., Mercer, K., Bronson, R. T., Crowley, D., Montoya, R., . . . Tuveson, D. A. (2001). Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev*, 15(24), 3243-3248. doi:10.1101/gad.943001
- Jacob, B., Osato, M., Yamashita, N., Wang, C. Q., Taniuchi, I., Littman, D. R., . . . Ito, Y. (2010). Stem cell exhaustion due to Runx1 deficiency is prevented by Evi5 activation in leukemogenesis. *Blood*, 115(8), 1610-1620. doi:10.1182/blood-2009-07-232249
- Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P. V., Mar, B. G., . . . Ebert, B. L. (2014). Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*, 371(26), 2488-2498. doi:10.1056/NEJMoa1408617
- Jan, M., Snyder, T. M., Corces-Zimmerman, M. R., Vyas, P., Weissman, I. L., Quake, S. R., & Majeti, R. (2012). Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Sci Transl Med*, 4(149), 149ra118. doi:10.1126/scitranslmed.3004315
- Jing, H., Vakoc, C. R., Ying, L., Mandat, S., Wang, H., Zheng, X., & Blobel, G. A. (2008). Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Mol Cell*, 29(2), 232-242. doi:10.1016/j.molcel.2007.11.020
- Johnson, L., Greenbaum, D., Cichowski, K., Mercer, K., Murphy, E., Schmitt, E., . . . Jacks, T. (1997). K-ras is an essential gene in the mouse with partial functional overlap with N-ras. *Genes Dev*, 11(19), 2468-2481.
- Katayama, S., Suzuki, M., Yamaoka, A., Keleku-Lukwete, N., Katsuoka, F., Otsuki, A., . . . Yamamoto, M. (2017). GATA2 haploinsufficiency accelerates EVI1-driven leukemogenesis. *Blood*, 130(7), 908-919. doi:10.1182/blood-2016-12-756767
- Kazenwadel, J., Secker, G. A., Liu, Y. J., Rosenfeld, J. A., Wildin, R. S., Cuellar-Rodriguez, J., . . . Harvey, N. L. (2012). Loss-of-function germline GATA2 mutations in patients with MDS/AML or MonoMAC syndrome and primary lymphedema reveal a key role for GATA2 in the lymphatic vasculature. *Blood*, 119(5), 1283-1291. doi:10.1182/blood-2011-08-374363
- Khwaja, A., Rodriguez-Viciana, P., Wennstrom, S., Warne, P. H., & Downward, J. (1997). Matrix adhesion and Ras transformation both activate a phosphoinositide 3-OH kinase and protein kinase B/Akt cellular survival pathway. *EMBO J*, 16(10), 2783-2793. doi:10.1093/emboj/16.10.2783
- Kim, S. J., Zhao, H., Hardikar, S., Singh, A. K., Goodell, M. A., & Chen, T. (2013). A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. *Blood*, 122(25), 4086-4089. doi:10.1182/blood-2013-02-483487
- Kirstetter, P., Schuster, M. B., Bereshchenko, O., Moore, S., Dvinge, H., Kurz, E., . . . Nerlov, C. (2008). Modeling of C/EBPalpha mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells. *Cancer Cell*, 13(4), 299-310. doi:10.1016/j.ccr.2008.02.008

- Kitajima, K., Tanaka, M., Zheng, J., Yen, H., Sato, A., Sugiyama, D., . . . Nakano, T. (2006). Redirecting differentiation of hematopoietic progenitors by a transcription factor, GATA-2. *Blood*, 107(5), 1857-1863. doi:10.1182/blood-2005-06-2527
- Kohnken, R., Porcu, P., & Mishra, A. (2017). Overview of the Use of Murine Models in Leukemia and Lymphoma Research. *Front Oncol*, 7, 22. doi:10.3389/fonc.2017.00022
- Krauth, M. T., Eder, C., Alpermann, T., Bacher, U., Nadarajah, N., Kern, W., . . . Schnittger, S. (2014). High number of additional genetic lesions in acute myeloid leukemia with t(8;21)/RUNX1-RUNX1T1: frequency and impact on clinical outcome. *Leukemia*, 28(7), 1449-1458. doi:10.1038/leu.2014.4
- Krivtsov, A. V., Figueroa, M. E., Sinha, A. U., Stubbs, M. C., Feng, Z., Valk, P. J., . . . Armstrong, S. A. (2013). Cell of origin determines clinically relevant subtypes of MLL-rearranged AML. *Leukemia*, 27(4), 852-860. doi:10.1038/leu.2012.363
- Krivtsov, A. V., Twomey, D., Feng, Z., Stubbs, M. C., Wang, Y., Faber, J., . . . Armstrong, S. A. (2006). Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature*, 442(7104), 818-822. doi:10.1038/nature04980
- Kuhn, R., Schwenk, F., Aguet, M., & Rajewsky, K. (1995). Inducible gene targeting in mice. *Science*, 269(5229), 1427-1429.
- Kumano, K., Chiba, S., Shimizu, K., Yamagata, T., Hosoya, N., Saito, T., . . . Hirai, H. (2001). Notch1 inhibits differentiation of hematopoietic cells by sustaining GATA-2 expression. *Blood*, 98(12), 3283-3289.
- La Ferla, K., Reimann, C., Jelkmann, W., & Hellwig-Burgel, T. (2002). Inhibition of erythropoietin gene expression signaling involves the transcription factors GATA-2 and NF-kappaB. *FASEB J*, 16(13), 1811-1813. doi:10.1096/fj.02-0168fje
- Laiosa, C. V., Stadtfeld, M., Xie, H., de Andres-Aguayo, L., & 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(5), 731-744. doi:10.1016/j.immuni.2006.09.011
- Lambert, J. M., Lambert, Q. T., Reuther, G. W., Malliri, A., Siderovski, D. P., Sondek, J., . . . Der, C. J. (2002). Tiam1 mediates Ras activation of Rac by a PI(3)K-independent mechanism. *Nat Cell Biol*, 4(8), 621-625. doi:10.1038/ncb833
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods*, 9(4), 357-359. doi:10.1038/nmeth.1923
- Lavallee, V. P., Kros, J., Lemieux, S., Boucher, G., Gendron, P., Pabst, C., . . . Sauvageau, G. (2016). Chemo-genomic interrogation of CEBPA mutated AML reveals recurrent CSF3R mutations and subgroup sensitivity to JAK inhibitors. *Blood*, 127(24), 3054-3061. doi:10.1182/blood-2016-03-705053
- Lewandowski, D., Barroca, V., Duconge, F., Bayer, J., Van Nhieu, J. T., Pestourie, C., . . . Romeo, P. H. (2010). In vivo cellular imaging pinpoints the role of reactive oxygen species in the early steps of adult hematopoietic reconstitution. *Blood*, 115(3), 443-452. doi:10.1182/blood-2009-05-222711
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078-2079. doi:10.1093/bioinformatics/btp352
- Li, Q., Bohin, N., Wen, T., Ng, V., Magee, J., Chen, S. C., . . . Morrison, S. J. (2013). Oncogenic Nras has bimodal effects on stem cells that sustainably increase competitiveness. *Nature*, 504(7478), 143-147. doi:10.1038/nature12830

- Liao, Y., Smyth, G. K., & Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*, 41(10), e108. doi:10.1093/nar/gkt214
- Ling, K. W., Ottersbach, K., van Hamburg, J. P., Oziemlak, A., Tsai, F. Y., Orkin, S. H., . . . Dzierzak, E. (2004). GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J Exp Med*, 200(7), 871-882. doi:10.1084/jem.20031556
- Liu, Y., Cheney, M. D., Gaudet, J. J., Chruszcz, M., Lukasik, S. M., Sugiyama, D., . . . Bushweller, J. H. (2006). The tetramer structure of the Nrvy homology two domain, NHR2, is critical for AML1/ETO's activity. *Cancer Cell*, 9(4), 249-260. doi:10.1016/j.ccr.2006.03.012
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 15(12), 550. doi:10.1186/s13059-014-0550-8
- Lowenberg, B., Downing, J. R., & Burnett, A. (1999). Acute myeloid leukemia. *N Engl J Med*, 341(14), 1051-1062. doi:10.1056/NEJM199909303411407
- Luesink, M., Hollink, I. H., van der Velden, V. H., Knops, R. H., Boezeman, J. B., de Haas, V., . . . Jansen, J. H. (2012). High GATA2 expression is a poor prognostic marker in pediatric acute myeloid leukemia. *Blood*, 120(10), 2064-2075. doi:10.1182/blood-2011-12-397083
- Lulli, V., Romania, P., Morsilli, O., Gabbianelli, M., Pagliuca, A., Mazzeo, S., . . . Marziali, G. (2006). Overexpression of Ets-1 in human hematopoietic progenitor cells blocks erythroid and promotes megakaryocytic differentiation. *Cell Death Differ*, 13(7), 1064-1074. doi:10.1038/sj.cdd.4401811
- Lutterbach, B., Sun, D., Schuetz, J., & Hiebert, S. W. (1998). The MYND motif is required for repression of basal transcription from the multidrug resistance 1 promoter by the t(8;21) fusion protein. *Mol Cell Biol*, 18(6), 3604-3611.
- Ma, O., Hong, S., Guo, H., Ghiaur, G., & Friedman, A. D. (2014). Granulopoiesis requires increased C/EBPalpha compared to monopoiesis, correlated with elevated Cebpa in immature G-CSF receptor versus M-CSF receptor expressing cells. *PLoS One*, 9(4), e95784. doi:10.1371/journal.pone.0095784
- Mace, E. M., Hsu, A. P., Monaco-Shawver, L., Makedonas, G., Rosen, J. B., Dropulic, L., . . . Orange, J. S. (2013). Mutations in GATA2 cause human NK cell deficiency with specific loss of the CD56(bright) subset. *Blood*, 121(14), 2669-2677. doi:10.1182/blood-2012-09-453969
- Maeno, M., Mead, P. E., Kelley, C., Xu, R. H., Kung, H. F., Suzuki, A., . . . Zon, L. I. (1996). The role of BMP-4 and GATA-2 in the induction and differentiation of hematopoietic mesoderm in *Xenopus laevis*. *Blood*, 88(6), 1965-1972.
- Mansour, S., Connell, F., Steward, C., Ostergaard, P., Brice, G., Smithson, S., . . . Lymphoedema Research, C. (2010). Emberger syndrome-primary lymphedema with myelodysplasia: report of seven new cases. *Am J Med Genet A*, 152A(9), 2287-2296. doi:10.1002/ajmg.a.33445
- Marceau-Renaut, A., Guihard, S., Castaigne, S., Dombret, H., Preudhomme, C., & Cheok, M. (2015). Classification of CEBPA mutated acute myeloid leukemia by GATA2 mutations. *Am J Hematol*, 90(5), E93-94. doi:10.1002/ajh.23949

- Mayle, A., Yang, L., Rodriguez, B., Zhou, T., Chang, E., Curry, C. V., . . . Goodell, M. A. (2015). Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood*, 125(4), 629-638. doi:10.1182/blood-2014-08-594648
- Mazzella, F. M., Smith, D., Horn, P., Cotelingam, J. D., Rector, J. T., Shrit, M. A., . . . Schumacher, H. R. (2006). Prognostic significance of pronormoblasts in erythrocyte predominant myelodysplastic patients. *Am J Hematol*, 81(7), 484-491. doi:10.1002/ajh.20563
- Mead, A. J., Neo, W. H., Barkas, N., Matsuoka, S., Giustacchini, A., Facchini, R., . . . Jacobsen, S. E. W. (2017). Niche-mediated depletion of the normal hematopoietic stem cell reservoir by Flt3-ITD-induced myeloproliferation. *J Exp Med*, 214(7), 2005-2021. doi:10.1084/jem.20161418
- Michaud, J., Wu, F., Osato, M., Cottles, G. M., Yanagida, M., Asou, N., . . . Scott, H. S. (2002). In vitro analyses of known and novel RUNX1/AML1 mutations in dominant familial platelet disorder with predisposition to acute myelogenous leukemia: implications for mechanisms of pathogenesis. *Blood*, 99(4), 1364-1372.
- Micol, J. B., Pastore, A., Inoue, D., Duployez, N., Kim, E., Lee, S. C., . . . Abdel-Wahab, O. (2017). ASXL2 is essential for haematopoiesis and acts as a haploinsufficient tumour suppressor in leukemia. *Nat Commun*, 8, 15429. doi:10.1038/ncomms15429
- Minegishi, N., Ohta, J., Yamagiwa, H., Suzuki, N., Kawauchi, S., Zhou, Y., . . . Yamamoto, M. (1999). The mouse GATA-2 gene is expressed in the para-aortic splanchnopleura and aorta-gonads and mesonephros region. *Blood*, 93(12), 4196-4207.
- Miyamoto, T., Nagafuji, K., Akashi, K., Harada, M., Kyo, T., Akashi, T., . . . Niho, Y. (1996). Persistence of multipotent progenitors expressing AML1/ETO transcripts in long-term remission patients with t(8;21) acute myelogenous leukemia. *Blood*, 87(11), 4789-4796.
- Miyamoto, T., Weissman, I. L., & Akashi, K. (2000). AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation. *Proc Natl Acad Sci U S A*, 97(13), 7521-7526.
- Moignard, V., Macaulay, I. C., Swiers, G., Buettner, F., Schutte, J., Calero-Nieto, F. J., . . . Gottgens, B. (2013). Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nat Cell Biol*, 15(4), 363-372. doi:10.1038/ncb2709
- Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., . . . Groop, L. C. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*, 34(3), 267-273. doi:10.1038/ng1180
- Moreau-Gachelin, F., Tavitian, A., & Tambourin, P. (1988). Spi-1 is a putative oncogene in virally induced murine erythroleukaemias. *Nature*, 331(6153), 277-280. doi:10.1038/331277a0
- Nemeth, M. J., & Bodine, D. M. (2007). Regulation of hematopoiesis and the hematopoietic stem cell niche by Wnt signaling pathways. *Cell Res*, 17(9), 746-758. doi:10.1038/cr.2007.69
- Nerlov, C., & Ziff, E. B. (1995). CCAAT/enhancer binding protein-alpha amino acid motifs with dual TBP and TFIIB binding ability co-operate to activate transcription in both yeast and mammalian cells. *EMBO J*, 14(17), 4318-4328.

- Ng, S. W., Mitchell, A., Kennedy, J. A., Chen, W. C., McLeod, J., Ibrahimova, N., . . . Wang, J. C. (2016). A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature*, *540*(7633), 433-437. doi:10.1038/nature20598
- Ng, Y. K., George, K. M., Engel, J. D., & Linzer, D. I. (1994). GATA factor activity is required for the trophoblast-specific transcriptional regulation of the mouse placental lactogen I gene. *Development*, *120*(11), 3257-3266.
- Niimi, K., Kiyoi, H., Ishikawa, Y., Hayakawa, F., Kurahashi, S., Kihara, R., . . . Naoe, T. (2013). GATA2 zinc finger 2 mutation found in acute myeloid leukemia impairs myeloid differentiation. *Leuk Res Rep*, *2*(1), 21-25. doi:10.1016/j.lrr.2013.02.002
- Nutt, S. L., Metcalf, D., D'Amico, A., Polli, M., & Wu, L. (2005). Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. *J Exp Med*, *201*(2), 221-231. doi:10.1084/jem.20041535
- Ogawa, E., Inuzuka, M., Maruyama, M., Satake, M., Naito-Fujimoto, M., Ito, Y., & Shigesada, K. (1993). Molecular cloning and characterization of PEBP2 beta, the heterodimeric partner of a novel Drosophila runt-related DNA binding protein PEBP2 alpha. *Virology*, *194*(1), 314-331. doi:10.1006/viro.1993.1262
- Ohlsson, E., Schuster, M. B., Hasemann, M., & Porse, B. T. (2016). The multifaceted functions of C/EBPalpha in normal and malignant haematopoiesis. *Leukemia*, *30*(4), 767-775. doi:10.1038/leu.2015.324
- Ohmori, S., Moriguchi, T., Noguchi, Y., Ikeda, M., Kobayashi, K., Tomaru, N., . . . Ohneda, K. (2015). GATA2 is critical for the maintenance of cellular identity in differentiated mast cells derived from mouse bone marrow. *Blood*, *125*(21), 3306-3315. doi:10.1182/blood-2014-11-612465
- Okuda, T., Cai, Z., Yang, S., Lenny, N., Lyu, C. J., van Deursen, J. M., . . . Downing, J. R. (1998). Expression of a knocked-in AML1-ETO leukemia gene inhibits the establishment of normal definitive hematopoiesis and directly generates dysplastic hematopoietic progenitors. *Blood*, *91*(9), 3134-3143.
- Okuda, T., van Deursen, J., Hiebert, S. W., Grosveld, G., & Downing, J. R. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell*, *84*(2), 321-330.
- Orlic, D., Anderson, S., Biesecker, L. G., Sorrentino, B. P., & Bodine, D. M. (1995). Pluripotent hematopoietic stem cells contain high levels of mRNA for c-kit, GATA-2, p45 NF-E2, and c-myb and low levels or no mRNA for c-fms and the receptors for granulocyte colony-stimulating factor and interleukins 5 and 7. *Proc Natl Acad Sci U S A*, *92*(10), 4601-4605.
- Osato, M. (2004). Point mutations in the RUNX1/AML1 gene: another actor in RUNX leukemia. *Oncogene*, *23*(24), 4284-4296. doi:10.1038/sj.onc.1207779
- Ostergaard, P., Simpson, M. A., Connell, F. C., Steward, C. G., Brice, G., Woollard, W. J., . . . Mansour, S. (2011). Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). *Nat Genet*, *43*(10), 929-931. doi:10.1038/ng.923
- Ottersbach, K., Sanjuan-Pla, A., Torres-Ruiz, R., Bueno, C., Velasco-Hernández, T., & Menendez, P. (2018). The "Never-Ending" Mouse Models for MLL- Rearranged Acute Leukemia Are Still Teaching Us. *HemaSphere*, *2*(4), e57.
- Pabst, T., & Mueller, B. U. (2009). Complexity of CEBPA dysregulation in human acute myeloid leukemia. *Clin Cancer Res*, *15*(17), 5303-5307. doi:10.1158/1078-0432.CCR-08-2941

- Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., . . . Campbell, P. J. (2016). Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med*, 374(23), 2209-2221. doi:10.1056/NEJMoa1516192
- Pasquet, M., Bellanne-Chantelot, C., Tavitian, S., Prade, N., Beaupain, B., Larochelle, O., . . . Delabesse, E. (2013). High frequency of GATA2 mutations in patients with mild chronic neutropenia evolving to MonoMac syndrome, myelodysplasia, and acute myeloid leukemia. *Blood*, 121(5), 822-829. doi:10.1182/blood-2012-08-447367
- Pellagatti, A., Roy, S., Di Genua, C., Burns, A., McGraw, K., Valletta, S., . . . Boultonwood, J. (2016). Targeted resequencing analysis of 31 genes commonly mutated in myeloid disorders in serial samples from myelodysplastic syndrome patients showing disease progression. *Leukemia*, 30(1), 247-250. doi:10.1038/leu.2015.129
- Persons, D. A., Allay, J. A., Allay, E. R., Ashmun, R. A., Orlic, D., Jane, S. M., . . . Nienhuis, A. W. (1999). Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. *Blood*, 93(2), 488-499.
- Peterson, L. F., & Zhang, D. E. (2004). The 8;21 translocation in leukemogenesis. *Oncogene*, 23(24), 4255-4262. doi:10.1038/sj.onc.1207727
- Picelli, S., Bjorklund, A. K., Faridani, O. R., Sagasser, S., Winberg, G., & Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods*, 10(11), 1096-1098. doi:10.1038/nmeth.2639
- Ping, N., Sun, A., Song, Y., Wang, Q., Yin, J., Cheng, W., . . . Chen, S. (2017). Exome sequencing identifies highly recurrent somatic GATA2 and CEBPA mutations in acute erythroid leukemia. *Leukemia*, 31(1), 195-202. doi:10.1038/leu.2016.162
- Pratilas, C. A., Taylor, B. S., Ye, Q., Viale, A., Sander, C., Solit, D. B., & Rosen, N. (2009). (V600E)BRAF is associated with disabled feedback inhibition of RAF-MEK signaling and elevated transcriptional output of the pathway. *Proc Natl Acad Sci U S A*, 106(11), 4519-4524. doi:10.1073/pnas.0900780106
- Pronk, C. J., Rossi, D. J., Mansson, R., Attema, J. L., Norddahl, G. L., Chan, C. K., . . . Bryder, D. (2007). Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell*, 1(4), 428-442. doi:10.1016/j.stem.2007.07.005
- Pruitt, K., & Der, C. J. (2001). Ras and Rho regulation of the cell cycle and oncogenesis. *Cancer Lett*, 171(1), 1-10.
- Pulikkan, J. A., Tenen, D. G., & Behre, G. (2017). C/EBPalpha deregulation as a paradigm for leukemogenesis. *Leukemia*, 31(11), 2279-2285. doi:10.1038/leu.2017.229
- Pylayeva-Gupta, Y., Grabocka, E., & Bar-Sagi, D. (2011). RAS oncogenes: weaving a tumorigenic web. *Nat Rev Cancer*, 11(11), 761-774. doi:10.1038/nrc3106
- Reckzeh, K., Bereshchenko, O., Mead, A., Rehn, M., Kharazi, S., Jacobsen, S. E., . . . Cammenga, J. (2012). Molecular and cellular effects of oncogene cooperation in a genetically accurate AML mouse model. *Leukemia*, 26(7), 1527-1536. doi:10.1038/leu.2012.37
- Reinisch, A., Thomas, D., Corces, M. R., Zhang, X., Gratzinger, D., Hong, W. J., . . . Majeti, R. (2016). A humanized bone marrow ossicle xenotransplantation model enables improved engraftment of healthy and leukemic human hematopoietic cells. *Nat Med*, 22(7), 812-821. doi:10.1038/nm.4103
- Rekhtman, N., Radparvar, F., Evans, T., & Skoultschi, A. I. (1999). Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. *Genes Dev*, 13(11), 1398-1411.

- Ren, R. (2004). Modeling the dosage effect of oncogenes in leukemogenesis. *Curr Opin Hematol*, 11(1), 25-34.
- Rhoades, K. L., Hetherington, C. J., Harakawa, N., Yergeau, D. A., Zhou, L., Liu, L. Q., . . . Zhang, D. E. (2000). Analysis of the role of AML1-ETO in leukemogenesis, using an inducible transgenic mouse model. *Blood*, 96(6), 2108-2115.
- Robert-Moreno, A., Espinosa, L., de la Pompa, J. L., & Bigas, A. (2005). RBPjkappa-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. *Development*, 132(5), 1117-1126. doi:10.1242/dev.01660
- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139-140. doi:10.1093/bioinformatics/btp616
- Rodrigues, N. P., Janzen, V., Forkert, R., Dombkowski, D. M., Boyd, A. S., Orkin, S. H., . . . Scadden, D. T. (2005). Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood*, 106(2), 477-484. doi:10.1182/blood-2004-08-2989
- Sabnis, A. J., Cheung, L. S., Dail, M., Kang, H. C., Santaguida, M., Hermiston, M. L., . . . Braun, B. S. (2009). Oncogenic Kras initiates leukemia in hematopoietic stem cells. *PLoS Biol*, 7(3), e59. doi:10.1371/journal.pbio.1000059
- Scheffzek, K., Ahmadian, M. R., Kabsch, W., Wiesmuller, L., Lautwein, A., Schmitz, F., & Wittinghofer, A. (1997). The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science*, 277(5324), 333-338.
- Schessler, C., Rawat, V. P., Cusan, M., Deshpande, A., Kohl, T. M., Rosten, P. M., . . . Buske, C. (2005). The AML1-ETO fusion gene and the FLT3 length mutation collaborate in inducing acute leukemia in mice. *J Clin Invest*, 115(8), 2159-2168. doi:10.1172/JCI24225
- Schnittger, S., Dicker, F., Kern, W., Wendland, N., Sundermann, J., Alpermann, T., . . . Haferlach, T. (2011). RUNX1 mutations are frequent in de novo AML with noncomplex karyotype and confer an unfavorable prognosis. *Blood*, 117(8), 2348-2357. doi:10.1182/blood-2009-11-255976
- Schonhuber, N., Seidler, B., Schuck, K., Veltkamp, C., Schachtler, C., Zukowska, M., . . . Saur, D. (2014). A next-generation dual-recombinase system for time- and host-specific targeting of pancreatic cancer. *Nat Med*, 20(11), 1340-1347. doi:10.1038/nm.3646
- Seita, J., & Weissman, I. L. (2010). Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip Rev Syst Biol Med*, 2(6), 640-653. doi:10.1002/wsbm.86
- Shih, A. H., Jiang, Y., Meydan, C., Shank, K., Pandey, S., Barreyro, L., . . . Levine, R. L. (2015). Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. *Cancer Cell*, 27(4), 502-515. doi:10.1016/j.ccell.2015.03.009
- Shlush, L. I., Zandi, S., Mitchell, A., Chen, W. C., Brandwein, J. M., Gupta, V., . . . Dick, J. E. (2014). Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature*, 506(7488), 328-333. doi:10.1038/nature13038
- Slomiany, B. A., D'Arigo, K. L., Kelly, M. M., & Kurtz, D. T. (2000). C/EBPalpha inhibits cell growth via direct repression of E2F-DP-mediated transcription. *Mol Cell Biol*, 20(16), 5986-5997.

- Socolovsky, M., Nam, H., Fleming, M. D., Haase, V. H., Brugnara, C., & Lodish, H. F. (2001). Ineffective erythropoiesis in Stat5a(-/-)5b(-/-) mice due to decreased survival of early erythroblasts. *Blood*, 98(12), 3261-3273.
- Stavropoulou, V., Almosaileakh, M., Royo, H., Spetz, J.-F., Juge, S., Brault, L., . . . Schwaller, J. (2018). A Novel Inducible Mouse Model of MLL-ENL-driven Mixed-lineage Acute Leukemia. *HemaSphere*, 2(4), e51.
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., . . . 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(43), 15545-15550. doi:10.1073/pnas.0506580102
- Sykes, S. M., Kokkaliaris, K. D., Milsom, M. D., Levine, R. L., & Majeti, R. (2015). Clonal Evolution of Preleukemic hematopoietic stem cells in acute myeloid leukemia. *Exp Hematol*. doi:10.1016/j.exphem.2015.08.012
- Tang, J. L., Hou, H. A., Chen, C. Y., Liu, C. Y., Chou, W. C., Tseng, M. H., . . . Tien, H. F. (2009). AML1/RUNX1 mutations in 470 adult patients with de novo acute myeloid leukemia: prognostic implication and interaction with other gene alterations. *Blood*, 114(26), 5352-5361. doi:10.1182/blood-2009-05-223784
- Taniguchi Ishikawa, E., Gonzalez-Nieto, D., Ghiaur, G., Dunn, S. K., Ficker, A. M., Murali, B., . . . Cancelas, J. A. (2012). Connexin-43 prevents hematopoietic stem cell senescence through transfer of reactive oxygen species to bone marrow stromal cells. *Proc Natl Acad Sci U S A*, 109(23), 9071-9076. doi:10.1073/pnas.1120358109
- Thomas, D., & Majeti, R. (2017). Biology and relevance of human acute myeloid leukemia stem cells. *Blood*, 129(12), 1577-1585. doi:10.1182/blood-2016-10-696054
- Tsai, F. Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., . . . Orkin, S. H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*, 371(6494), 221-226. doi:10.1038/371221a0
- Tsai, F. Y., & Orkin, S. H. (1997). Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood*, 89(10), 3636-3643.
- Tsai, S. C., Shih, L. Y., Liang, S. T., Huang, Y. J., Kuo, M. C., Huang, C. F., . . . Liang, D. C. (2015). Biological Activities of RUNX1 Mutants Predict Secondary Acute Leukemia Transformation from Chronic Myelomonocytic Leukemia and Myelodysplastic Syndromes. *Clin Cancer Res*, 21(15), 3541-3551. doi:10.1158/1078-0432.CCR-14-2203
- Van Etten, R. A., & Shannon, K. M. (2004). Focus on myeloproliferative diseases and myelodysplastic syndromes. *Cancer Cell*, 6(6), 547-552. doi:10.1016/j.ccr.2004.12.004
- Van Meter, M. E., Diaz-Flores, E., Archard, J. A., Passegue, E., Irish, J. M., Kotecha, N., . . . Braun, B. S. (2007). K-RasG12D expression induces hyperproliferation and aberrant signaling in primary hematopoietic stem/progenitor cells. *Blood*, 109(9), 3945-3952. doi:10.1182/blood-2006-09-047530
- Vardiman, J. W., Thiele, J., Arber, D. A., Brunning, R. D., Borowitz, M. J., Porwit, A., . . . Bloomfield, C. D. (2009). The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*, 114(5), 937-951. doi:10.1182/blood-2009-03-209262

- Vicente, C., Conchillo, A., Garcia-Sanchez, M. A., & Odero, M. D. (2012). The role of the GATA2 transcription factor in normal and malignant hematopoiesis. *Crit Rev Oncol Hematol*, 82(1), 1-17. doi:10.1016/j.critrevonc.2011.04.007
- Vicente, C., Vazquez, I., Conchillo, A., Garcia-Sanchez, M. A., Marcotegui, N., Fuster, O., . . . Odero, M. D. (2012). Overexpression of GATA2 predicts an adverse prognosis for patients with acute myeloid leukemia and it is associated with distinct molecular abnormalities. *Leukemia*, 26(3), 550-554. doi:10.1038/leu.2011.235
- Vinh, D. C., Patel, S. Y., Uzel, G., Anderson, V. L., Freeman, A. F., Olivier, K. N., . . . Holland, S. M. (2010). Autosomal dominant and sporadic monocytopenia with susceptibility to mycobacteria, fungi, papillomaviruses, and myelodysplasia. *Blood*, 115(8), 1519-1529. doi:10.1182/blood-2009-03-208629
- Walsh, J. C., DeKoter, R. P., Lee, H. J., Smith, E. D., Lancki, D. W., Gurish, M. F., . . . Singh, H. (2002). Cooperative and antagonistic interplay between PU.1 and GATA-2 in the specification of myeloid cell fates. *Immunity*, 17(5), 665-676.
- Walter, M. J., Shen, D., Ding, L., Shao, J., Koboldt, D. C., Chen, K., . . . Graubert, T. A. (2012). Clonal architecture of secondary acute myeloid leukemia. *N Engl J Med*, 366(12), 1090-1098. doi:10.1056/NEJMoa1106968
- Wang, L., Gural, A., Sun, X. J., Zhao, X., Perna, F., Huang, G., . . . Nimer, S. D. (2011). The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. *Science*, 333(6043), 765-769. doi:10.1126/science.1201662
- Wang, N. D., Finegold, M. J., Bradley, A., Ou, C. N., Abdelsayed, S. V., Wilde, M. D., . . . Darlington, G. J. (1995). Impaired energy homeostasis in C/EBP alpha knockout mice. *Science*, 269(5227), 1108-1112.
- Wang, S. A., Tang, G., Fadare, O., Hao, S., Raza, A., Woda, B. A., & Hasserjian, R. P. (2008). Erythroid-predominant myelodysplastic syndromes: enumeration of blasts from nonerythroid rather than total marrow cells provides superior risk stratification. *Mod Pathol*, 21(11), 1394-1402. doi:10.1038/modpathol.2008.142
- Wang, Y. Y., Zhao, L. J., Wu, C. F., Liu, P., Shi, L., Liang, Y., . . . Chen, S. J. (2011). C-KIT mutation cooperates with full-length AML1-ETO to induce acute myeloid leukemia in mice. *Proc Natl Acad Sci U S A*, 108(6), 2450-2455. doi:10.1073/pnas.1019625108
- Welch, J. S., Ley, T. J., Link, D. C., Miller, C. A., Larson, D. E., Koboldt, D. C., . . . Wilson, R. K. (2012). The origin and evolution of mutations in acute myeloid leukemia. *Cell*, 150(2), 264-278. doi:10.1016/j.cell.2012.06.023
- Westervelt, P., Lane, A. A., Pollock, J. L., Oldfather, K., Holt, M. S., Zimonjic, D. B., . . . Ley, T. J. (2003). High-penetrance mouse model of acute promyelocytic leukemia with very low levels of PML-RARalpha expression. *Blood*, 102(5), 1857-1865. doi:10.1182/blood-2002-12-3779
- Wiemels, J. L., Xiao, Z., Buffler, P. A., Maia, A. T., Ma, X., Dicks, B. M., . . . Greaves, M. (2002). In utero origin of t(8;21) AML1-ETO translocations in childhood acute myeloid leukemia. *Blood*, 99(10), 3801-3805.
- Wlodarski, M. W., Collin, M., & Horwitz, M. S. (2017). GATA2 deficiency and related myeloid neoplasms. *Semin Hematol*, 54(2), 81-86. doi:10.1053/j.seminhematol.2017.05.002
- Wlodarski, M. W., Hirabayashi, S., Pastor, V., Stary, J., Hasle, H., Masetti, R., . . . Ewog, M. D. S. (2016). Prevalence, clinical characteristics, and prognosis of GATA2-related

- myelodysplastic syndromes in children and adolescents. *Blood*, 127(11), 1387-1397; quiz 1518. doi:10.1182/blood-2015-09-669937
- Wolfler, A., Danen-van Oorschot, A. A., Haanstra, J. R., Valkhof, M., Bodner, C., Vroegindeweij, E., . . . Touw, I. P. (2010). Lineage-instructive function of C/EBPalpha in multipotent hematopoietic cells and early thymic progenitors. *Blood*, 116(20), 4116-4125. doi:10.1182/blood-2010-03-275404
- Wouters, B. J., Jorda, M. A., Keeshan, K., Louwers, I., Erpelinck-Verschueren, C. A., Tielemans, D., . . . Delwel, R. (2007). Distinct gene expression profiles of acute myeloid/T-lymphoid leukemia with silenced CEBPA and mutations in NOTCH1. *Blood*, 110(10), 3706-3714. doi:10.1182/blood-2007-02-073486
- Xie, M., Lu, C., Wang, J., McLellan, M. D., Johnson, K. J., Wendl, M. C., . . . Ding, L. (2014). Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med*, 20(12), 1472-1478. doi:10.1038/nm.3733
- Xue, L., Pulikkan, J. A., Valk, P. J., & Castilla, L. H. (2014). NrasG12D oncoprotein inhibits apoptosis of preleukemic cells expressing Cbfbeta-SMMHC via activation of MEK/ERK axis. *Blood*, 124(3), 426-436. doi:10.1182/blood-2013-12-541730
- Yamazaki, H., Suzuki, M., Otsuki, A., Shimizu, R., Bresnick, E. H., Engel, J. D., & Yamamoto, M. (2014). A remote GATA2 hematopoietic enhancer drives leukemogenesis in inv(3)(q21;q26) by activating EVI1 expression. *Cancer Cell*, 25(4), 415-427. doi:10.1016/j.ccr.2014.02.008
- Yan, X. J., Xu, J., Gu, Z. H., Pan, C. M., Lu, G., Shen, Y., . . . Chen, S. J. (2011). Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet*, 43(4), 309-315. doi:10.1038/ng.788
- Yang, H. T., Wu, D. H., Xue, X. Y., Liang, W. X., Miao, X. Y., Pang, H., & Chen, S. J. (2004). Cloning, expression, purification and crystallization of NHR3 domain from acute myelogenous leukemia-related protein AML1-ETO. *Acta Biochim Biophys Sin (Shanghai)*, 36(8), 566-570.
- Yang, L., Rau, R., & Goodell, M. A. (2015). DNMT3A in haematological malignancies. *Nat Rev Cancer*, 15(3), 152-165. doi:10.1038/nrc3895
- Yang, L., Rodriguez, B., Mayle, A., Park, H. J., Lin, X., Luo, M., . . . Goodell, M. A. (2016). DNMT3A Loss Drives Enhancer Hypomethylation in FLT3-ITD-Associated Leukemias. *Cancer Cell*, 29(6), 922-934. doi:10.1016/j.ccell.2016.05.003
- Ye, M., Zhang, H., Amabile, G., Yang, H., Staber, P. B., Zhang, P., . . . Tenen, D. G. (2013). C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence. *Nat Cell Biol*, 15(4), 385-394. doi:10.1038/ncb2698
- Ye, M., Zhang, H., Yang, H., Koche, R., Staber, P. B., Cusan, M., . . . Tenen, D. G. (2015). Hematopoietic Differentiation Is Required for Initiation of Acute Myeloid Leukemia. *Cell Stem Cell*, 17(5), 611-623. doi:10.1016/j.stem.2015.08.011
- Yergeau, D. A., Hetherington, C. J., Wang, Q., Zhang, P., Sharpe, A. H., Binder, M., . . . Zhang, D. E. (1997). Embryonic lethality and impairment of haematopoiesis in mice heterozygous for an AML1-ETO fusion gene. *Nat Genet*, 15(3), 303-306. doi:10.1038/ng0397-303
- Yuan, Y., Zhou, L., Miyamoto, T., Iwasaki, H., Harakawa, N., Hetherington, C. J., . . . Zhang, D. E. (2001). AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations. *Proc Natl Acad Sci U S A*, 98(18), 10398-10403. doi:10.1073/pnas.171321298

- Yuasa, H., Oike, Y., Iwama, A., Nishikata, I., Sugiyama, D., Perkins, A., . . . Morishita, K. (2005). Oncogenic transcription factor Evi1 regulates hematopoietic stem cell proliferation through GATA-2 expression. *EMBO J*, 24(11), 1976-1987. doi:10.1038/sj.emboj.7600679
- Zhang, D. E., Zhang, P., Wang, N. D., Hetherington, C. J., Darlington, G. J., & Tenen, D. G. (1997). Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice. *Proc Natl Acad Sci U S A*, 94(2), 569-574.
- Zhang, J., Kalkum, M., Yamamura, S., Chait, B. T., & Roeder, R. G. (2004). E protein silencing by the leukemogenic AML1-ETO fusion protein. *Science*, 305(5688), 1286-1289. doi:10.1126/science.1097937
- Zhang, P., Iwama, A., Datta, M. W., Darlington, G. J., Link, D. C., & Tenen, D. G. (1998). Upregulation of interleukin 6 and granulocyte colony-stimulating factor receptors by transcription factor CCAAT enhancer binding protein alpha (C/EBP alpha) is critical for granulopoiesis. *J Exp Med*, 188(6), 1173-1184.
- Zhang, P., Iwasaki-Arai, J., Iwasaki, H., Fenyus, M. L., Dayaram, T., Owens, B. M., . . . Tenen, D. G. (2004). Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity*, 21(6), 853-863. doi:10.1016/j.immuni.2004.11.006
- Zhang, S. J., Ma, L. Y., Huang, Q. H., Li, G., Gu, B. W., Gao, X. D., . . . Chen, S. J. (2008). Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. *Proc Natl Acad Sci U S A*, 105(6), 2076-2081. doi:10.1073/pnas.0711824105
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhoutte, J., Johnson, D. S., Bernstein, B. E., . . . Liu, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol*, 9(9), R137. doi:10.1186/gb-2008-9-9-r137
- Zhao, S., Zhang, Y., Sha, K., Tang, Q., Yang, X., Yu, C., . . . Cui, S. (2014). KRAS (G12D) cooperates with AML1/ETO to initiate a mouse model mimicking human acute myeloid leukemia. *Cell Physiol Biochem*, 33(1), 78-87. doi:10.1159/000356651
- Zhou, Y., Lim, K. C., Onodera, K., Takahashi, S., Ohta, J., Minegishi, N., . . . Engel, J. D. (1998). Rescue of the embryonic lethal hematopoietic defect reveals a critical role for GATA-2 in urogenital development. *EMBO J*, 17(22), 6689-6700. doi:10.1093/emboj/17.22.6689
- Zuo, Z., Medeiros, L. J., Chen, Z., Liu, D., Bueso-Ramos, C. E., Luthra, R., & Wang, S. A. (2012). Acute myeloid leukemia (AML) with erythroid predominance exhibits clinical and molecular characteristics that differ from other types of AML. *PLoS One*, 7(7), e41485. doi:10.1371/journal.pone.0041485
- Zuo, Z., Polski, J. M., Kasyan, A., & Medeiros, L. J. (2010). Acute erythroid leukemia. *Arch Pathol Lab Med*, 134(9), 1261-1270. doi:10.1043/2009-0350-RA.1