The role of AML secretome and bone marrow microenvironment in establishment and maintenance of leukaemia initiating cells

Xiying Ren

Green Templeton College

Radcliffe Department of Medicine

University of Oxford

United Kingdom

A thesis presented for the degree of

Doctor of Philosophy in Medical Sciences

January 2, 2023

Abstract

Acute myeloid leukaemia (AML) is the most common acute leukaemia in adults and is characterised by the accumulation of immature malignant precursors in haematopoietic and other tissues. Even though next-generation sequencing has dramatically advanced our understanding of the genomic landscape of AML and identified many leukaemia-associated gene signatures, intensive chemotherapy remains the mainstay of the general therapeutic strategy without much progression over decades. In addition, AML still has the worst overall survival among adult haematological malignancies. The prognosis and survival for elderly patients who are too frail to be eligible for intensive chemotherapy can be even poorer. Therefore, the identification of novel AML dependencies is urgently needed to develop less toxic therapeutic strategies that can potentially eliminate AML cells, especially the self-renewing leukaemia initiating cells (LICs).

One potential therapeutic target could be the cell surface receptors on LICs, which covey autocrine and paracrine signals existing in the bone marrow (BM) microenvironment, as there is considerable evidence that de-regulated BM microenvironmental factors could contribute to AML development. In order to identify cell surface receptors essential for LIC maintenance, I performed *in vivo* CRISPR-based screens focused on the cytokine and chemokine receptor genes in the context of a murine model of bi-allelic *CEBPA* mutant AML. This murine AML model, as opposed to the commonly used *MLL-AF9* model, is strictly BM microenvironment-dependent, which allows better characterisation of the interplay between LICs and the microenvironmental factors. By doing this, I identified and successfully validated 2 cytokine receptor genes, *Ccr1* and *Il7r*, as *in vivo* AML dependencies. Of note, these 2 genes are highly overexpressed across different subtypes of human AML cells, indicating their potential use as therapeutic targets for patients with AML.

Besides, in light of previous evidence suggesting the aged BM microenvironment's implication in AML propagation, I sought to investigate further the biological basis for the aged BM microenvironment's ability to facilitate AML progression. Several inflammation-related process networks were simultaneously enriched in BM stromal cells during ageing and AML development. In particular, I demonstrated that both physiological ageing and *CEBPA* mutant AML progression induced elevated protein levels of the pro-inflammatory CCR1 ligands, CCL3 and CCL5, in the BM cavity. This implied the role of CCR1 signalling as an ageing-associated AML dependency. Indeed, expansion of the murine *Cebpa* mutant AML cells in the unperturbed aged BM was significantly impaired by a selective CCR1 inhibitor. This suggests that physiological ageing harnessed CCR1 ligands as a critical mediator in age-associated AML predisposition, and targeting CCR1 can potentially benefit elderly patients

with AML.

Acknowledgements

I would like to thank my supervisor Prof Claus Nerlov for his guidance throughout my PhD, his willingness to allow me to study in his lab 4 years ago when I barely had the experience of working in the field of haematology, and his consistent trust in my ability and competency since the very beginning of my PhD.

I am grateful to Dr Simona Valletta for being my co-supervisor, mentor, colleague, and best friend in the Nerlov Lab. Without her help both in science and life, I cannot imagine how I would get through this whole PhD.

I want to thank Prof Adam Mead for serving as my secondary supervisor, who inspired me to focus on human AML transcriptomes while searching for candidate genes to be included in my second CRISPR screen.

I am thankful to all the current and past members of the Nerlov Lab, in particular to Dr Alexander Thomas for generating the amazing RNAseq data set, without which none of my studies would be possible, and Dr Cristina Di Genua for setting a super high standard for how a PhD student should be working, and Dr Yiran Meng for never saying NO to me whenever I needed help. I am also thankful to my thesis committee, my transfer and confirmation examiners, the FACS, genome engineering, virus production, and sequencing facilities, and the BMS staff for their help with my projects. Especially, I am lucky enough to have worked with Dr Philip Hublitz and Dr Ryan Beveridge from the first day of my PhD. Their expertise is invaluable.

I am also very grateful to my friends, Qian, Ziqi, and Rong. Without their accompany throughout this 4-year journey, it would not be so memorable and unique. I must mention Qian in particular, who not only helped me enormously with building my bioinformatics toolkit but also offered me a place to stay during the pandemic.

Last but not least, I owe a debt of gratitude to my mom, who has always supported me since I decided to apply for a PhD. Even though she, at that moment, did not understand why a medical student like me would need to spend 4 years working with mice, she still encouraged me to do whatever I believed was the best for me.

Declaration

The library preparation for bulk RNA sequencing of murine cells, as shown in Figures 3.1 and 5.1, was done by Alexander Thomas. The raw RNA sequencing read counts, differentially expressed gene lists, and Metacore perturbed process network lists used for plotting Figures 3.2, 3.3, 3.4, 3.12, 5.2, 5.3, 5.4, and 5.8A-D were initially generated by Dr Alexander Thomas. The downstream analysis of these RNA sequencing data, such as the hypergeometric test and gene-set enrichment analysis, was done by myself.

Contents

List of Figures **x**

List of Tables **xiii**

Commonly Used Abbreviations **xiv**

1 Chapter 1: Introduction **1**

1.1 Acute myeloid leukaemia (AML) 1

1.1.1 Incidence of AML among adults 1

1.1.2 AML pathogenesis 2

1.1.3 Leukaemic stem cells (LSCs) in AML 3

1.1.4 Current therapeutic strategies in patients with AML 5

1.1.5 Emerging therapeutic strategies in patients with AML 7

1.1.6 Survival of patients with AML 8

1.2 Murine models of AML 8

1.2.1 A murine model of bi-allelic *CEBPA* mutant AML 8

1.2.2 Other mouse models of AML, including xenograft mouse models 11

1.3 The BM microenvironment and myeloid malignancies 12

1.3.1 The BM microenvironment and its implications in normal haematopoiesis 12

1.3.2 Myeloid malignancies can originate in the BM niche 14

1.3.3 AML cells can cause niche remodelling 14

1.3.4 The BM microenvironment is a pro-inflammatory environment during myeloid malignancies 16

1.4 AML in the elderly 17

1.4.1 Characteristics of AML in the elderly . . . 17
1.4.2 HSC ageing and AML . . . 18
1.4.3 Ageing of the BM microenvironment and AML . . . 19
1.5 CRISPR-Cas9 genome editing and CRISPR screens . . . 21
1.5.1 CRISPR-Cas9 as a tool for genome editing . . . 21
1.5.2 Pooled CRISPR-Cas9 screens . . . 23
1.5.3 Applications of CRISPR-based screens in leukaemia research . 27

2 Chapter 2: Methods 29

2.1 Animals . . . 29
2.2 Murine leukaemia model . . . 29
2.3 Cell preparation . . . 30
2.3.1 BM . . . 30
2.3.2 c-Kit enrichment of BM cells . . . 30
2.3.3 Spleen . . . 31
2.3.4 PB . . . 31
2.3.5 BM stromal cell isolation . . . 31
2.3.6 Freezing cells . . . 32
2.3.7 Thawing frozen cells . . . 33
2.4 Flow cytometry and Fluorescence-activated cell sorting (FACS) . . . 33
2.4.1 Antibody staining and flow cytometry analysis . . . 33
2.4.2 FACS . . . 34
2.4.3 Instrument and analysis . . . 35
2.5 RNAseq library preparation . . . 35
2.6 Candidate gene selection for the CRISPR screens . . . 35
2.7 Lentiviral sgRNA library generation . . . 41
2.8 *In vivo* CRISPR screens . . . 41
2.9 qRT-PCR . . . 43
2.10 Multiplex qRT-PCR analysis . . . 44
2.11 ELISA . . . 45
2.12 Quantification of CRISPR editing . . . 45
2.13 *In vivo* validation of *Ccr1* and *Il7r* as dependencies of *Cebpa* mutant LICs . . . 46

2.14 Hydrodynamic injection 46
2.15 Proliferation assay 47
2.16 Drug treatment of leukaemic mice 47
2.17 Quantification and statistical analysis 48
2.17.1 Bulk RNAseq analysis 48
2.17.2 Pooled CRISPR screen data analysis 49
2.17.3 The Cancer Genome Atlas (TCGA) data analysis 50
2.17.4 Beat AML data analysis 50
2.17.5 Multiplex qRT-PCR analysis 51
2.17.6 Statistical analysis 51

3 Chapter 3: Identification of cytokine receptors potentially essential for LIC survival using *in vivo* CRISPR-Cas9 dropout screens 52
3.1 Introduction and aims 52
3.2 Cytokine receptors can potentially be used as a druggable target to eliminate the *Cebpa* mutant LICs 56
3.3 An *in vivo* CRISPR dropout screen focused on genes consistently upregulated throughout *Cebpa* mutant AML development in mice . . 61
3.3.1 Selection of candidate genes included in the screen and sgRNA library design 61
3.3.2 Development of the *in vivo* screen 62
3.3.3 Dynamic change in sgRNA distributions during leukaemia development 64
3.3.4 *Ccr1* was identified as the top hit 67
3.4 An *in vivo* CRISPR dropout screen focused on cytokine receptor genes simultaneously overexpressed in murine and human AML cells . 73
3.4.1 Selection of candidate genes included in the screen and sgRNA library design 73
3.4.2 Development of the *in vivo* screen 74
3.4.3 Dynamic evolution of sgRNA library representations during leukaemia development 77
3.4.4 *Il7r* was identified as a gene potentially essential for LIC maintenance 79

3.5 Discussion 83

4 Chapter 4: Validation of CCR1 and IL7R as dependencies in murine *Cebpa* mutant AML cells 87

4.1 Introduction and aims 87

4.2 Investigating on-target editing efficiencies of the selected *Ccr1*- and *Il7r*-targeting sgRNAs 90

4.3 Knocking out *Ccr1* or *Il7r* depleted the murine *Cebpa* mutant LICs . 94

4.4 High *CCR1* expression was associated with poor prognosis in human AML 99

4.5 CCR1 ligand overexpression was not able to promote *Cebpa* mutant LIC propagation 103

4.6 *IL7R* was overexpressed in human AML samples 107

4.7 *In vivo* IL7R-blocking antibody treatment did not impair *Cebpa* mutant LIC propagation 109

4.8 Discussion 113

5 Chapter 5: Identification of cytokine signalling contributing to age-related AML predisposition 117

5.1 Introduction and aims 117

5.2 AML- and age-induced transcriptional changes to BM stromal cells are correlated 121

5.3 IL-6 signalling was not involved in increased AML progression during ageing 127

5.4 CCR1 signalling facilitated AML progression in the aged BM 130

5.5 Discussion 137

6 Conclusions 140

References 142

List of Figures

1.1 Structure of the CEBPA translation isoforms 9
1.2 Schematic of the BM microenvironment 13
1.3 Schematic of the CRISPR-Cas9 nuclease system 22
1.4 Workflow of pooled CRISPR screens 24

3.1 Gene expression profiling of bi-allelic *Cebpa* mutant leukaemia cells throughout AML development 56
3.2 Cytokine signalling gene expression in AML cells 58
3.3 Inflammation-related gene sets are enriched in bi-allelic *Cebpa* mutant leukaemia cells 60
3.4 Selection of candidate genes included in the first CRISPR screen . . . 62
3.5 Gating of sgRNA-expressing AML cells in the first CRISPR screen . . 64
3.6 Schematic representation of the first CRISPR-Cas9 dropout screen . . 65
3.7 Representation of the sgRNA library in AML cells and NIH3T3 cells during the first screen 67
3.8 *Ccr1* was identified as the top hit by MAGeCK RRA 68
3.9 CRISPR scores (CS) of genes in the first screen 70
3.10 *Ccr1* and the genes encoding CCR1 ligands were upregulated during AML 71
3.11 mRNA expression of cytokine receptor genes included in the first screen in the leukaemia cells 72
3.12 Selection of candidate genes included in the second CRISPR screen . 74
3.13 Gating of sgRNA-expressing AML cells in the second CRISPR screen 75
3.14 Schematic representation of the second CRISPR-Cas9 dropout screen 76

3.15 Representation of the sgRNA Library in AML cells and NIH3T3 cells during the second screen 79
3.16 *Il7r* was identified as the top hit by MAGeCK RRA 80
3.17 *Il7r* was highly overexpressed in the LICs 82

4.1 On-target modification of NIH3T3 cells using the selected *Ccr1*- and *Il7r*-targeting sgRNAs 91
4.2 ICE analysis of indels for the selected *Ccr1*- and *Il7r*-targeting sgRNAs in NIH3T3 cells 92
4.3 Schematic showing Sanger sequencing of the sgRNA-targeted loci . . 93
4.4 Schematic of *in vivo* validation of CCR1 and IL7R as AML dependencies 95
4.5 Representative flow cytometry showing percentages of GFP+ AML cells 48/72 hours post-transduction 96
4.6 CCR1 and IL7R were *in vivo* dependencies of *Cebpa* mutant LICs . . 97
4.7 ICE analysis of the *Ccr1* and *Il7r* loci in the sgRNA-expressing BM leukaemia cells 98
4.8 Increased CCR1 signalling is conserved between murine and TCGA human *CEBPA* mutant AML 100
4.9 Increased CCR1 signalling is conserved between murine and Beat AML human *CEBPA* mutant AML 102
4.10 Schematic of validating the role of CCR1 ligands in AML progression *in vivo* 104
4.11 Overexpression of CCR1 ligands by hydrodynamic injection did not promote *Cebpa* mutant LIC propagation *in vivo* 105
4.12 CCL3, CCL6, and CCL9 could not increase *Cebpa* mutant LIC viability *ex vivo* 106
4.13 Overexpression of *IL7R* in human AML samples 108
4.14 *In vivo* IL7R-blocking antibody treatment in AML mice 110
4.15 Blocking IL7R *in vivo* did not impair *Cebpa* mutant LIC propagation 111

5.1 Schematic of profiling leukaemic and aged BM stromal cells by bulk RNAseq 121

5.2 A significant number of up-DEGs were shared by leukaemic and aged CBM VECs 122
5.3 A significant number of Metacore perturbed networks were shared by leukaemic and aged CBM VECs 124
5.4 AML-induced molecular perturbations in CBM VECs pre-exist in aged BM 126
5.5 Overexpression of IL-6 by hydrodynamic injection could not remodel the young BM niche into an AML-permissive environment like the old 128
5.6 Blocking IL-6-signalling *in vivo* could not counteract the age-related increased AML propagation 130
5.7 CCR1 ligands constitutively presented in aged BM at levels similar to those generated by AML conditioning 131
5.8 Aged BM stromal cells were not the main cellular source of increased CCR1 ligands in the aged BM 132
5.9 Gating of BM stromal cell populations 134
5.10 CCR1-signalling could support AML progression in the aged BM . . 135

List of Tables

2.1 Mouse antibodies and viability dyes used for flow cytometry analysis and cell sorting . . . 34
2.2 Candidate genes included in the first CRISPR screen . . . 36
2.3 Sequences of sgRNAs used in the first screen . . . 37
2.4 Candidate genes included in the second CRISPR screen . . . 38
2.5 Sequences of sgRNAs used in the second screen . . . 38
2.6 Primers for sgRNA Amplification . . . 43
2.7 Taqman assays . . . 44
2.8 Primers for quantification of CRISPR editing . . . 45
2.9 Plasmids for hydrodynamic injections . . . 46

3.1 KEGG cytokine signalling genes overexpressed in LICs and LBs . . . 59
3.2 Sorted sgRNA-expressing populations for sequencing during the first screen . . . 66
3.3 Sorted sgRNA-expressing populations for sequencing during the second screen . . . 77

5.1 Metacore perturbed networks shared by LICs, leukaemic CBM VECs, and aged CBM VECs . . . 125

Commonly Used Abbreviations

AML: Acute myeloid leukaemia

B-ALL: B-cell acute lymphoblastic leukaemia

BL: Bone-lining

BM: Bone marrow

CAR: CXCL12-abundant reticular

CBM: Central bone marrow

CML: Chronic myelogenous leukaemia

CS: CRISPR gene score

CSC: Cancer stem cell

DEG: Differentially expressed gene

DMEM: Dulbecco Modified Eagle Medium

DMSO: Dimethyl sulfoxide

DSB: DNA double-strand break EC: Endothelial cell

FACS: Fluorescence-activated cell sorting

FCS: Foetal calf serum

FL: Foetal liver

gDNA: Genomic DNA

GMP: Granulocyte-macrophage progenitor

GSEA: Gene-set enrichment analysis

GTEx: Genotype-Tissue Expression Project

HCT: Haematopoietic cell transplantation

HDR: Homology-directed repair

HSC: Haematopoietic stem cell

HSPC: Haematopoietic stem and progenitor cell

ICE: Inference of CRISPR Edits

IMDM: Iscove's Modified Dulbecco's Medium

LB: Leukaemic myeloblast

LEPR+: Leptin receptor-expressing

LIC: Leukaemia initiating cell

LSC: Leukaemic stem cell

MDS: Myelodysplastic syndrome

MM: Multiple myeloma

MMEJ: Microhomology-mediated end joining

MOI: Multiplicity of infection

MPN: Myeloproliferative neoplasm

MPNA: Metacore process network analysis

MPP: Multipotent progenitor

MSC: Mesenchymal stromal cell

NGS: Next-generation sequencing

NHEJ: Nonhomologous end joining

OB: Osteoblast

OS: Overall survival

PB: Peripheral blood

PBS: Phosphate-buffered saline

PCA: Principal component analysis

PVC: Perivascular stromal cell

qRT-PCR: Quantitative Real-time PCR Analysis

RNAi: RNA interference

RNAseq: RNA sequencing

RPKM: Reads per kilobase of transcript per million

RRA: Robust ranking algorithm

RSEM: RNA-Seq by Expectation-Maximisation

SCF: Stem cell factor

sgRNA: Single-guide RNA

shRNA: Short hairpin RNA

siRNA: Small interfering RNA

SNS: Sympathetic nervous system

T-ALL: T-cell acute lymphoblastic leukaemia

TCGA: The Cancer Genome Atlas

TPM: Transcripts per million

TPO: Thrombopoietin

VEC: Vascular endothelial cell

WBC: White blood cell

WT: Wild-type

Chapter 1

Introduction

1.1 Acute myeloid leukaemia (AML)

AML is a form of myeloid malignancy characterised by clonal expansion of abnormal, immature myeloid precursors, resulting in infiltration of the bone marrow (BM), spleen, peripheral blood (PB) and other tissues [1]. These malignant cells can proliferate in an uncontrollable fashion but are not able to terminally differentiate into mature red cells, neutrophils, monocytes, and platelets [2]. The lack of normal cells eventually leads to BM failure, which is the most common cause of death in AML [2].

1.1.1 Incidence of AML among adults

Even though AML only accounted for less than 1% of all new cancer cases in the UK in 2016-2018 (`https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml#heading-Zero`), it is the commonest myeloid malignancy among adults [3, 4]. The age-adjusted incidence of AML is around 4-5 per 100,000 person-years in the developed world and is slowly increasing over time. In addition, among all subtypes of leukaemia, AML contributes to the highest percentage of mortalities [4].

1.1.2 AML pathogenesis

The pathogenesis of AML is believed to be initiated by a series of genetic perturbations in the haematopoietic stem and progenitor cells (HSPCs). There are various types of these leukaemia-driving genetic changes, such as chromosomal translocations, aneuploidies, and gene mutations [5]. Although AML is heterogeneous in terms of the driver event, general mechanisms underlying leukaemic transformation have been identified, which include impaired myeloid differentiation, increased cell survival, acquisition of self-renewal, uncontrolled proliferation, altered metabolic properties, epigenetic dysregulation, and specific BM niche interactions [6–8].

Originally, the leukaemic genetic changes were classified into 2 types. The first type (class 1) mutations result in the activation of signal transduction pathways [2, 9], conferring a proliferative and/or survival advantage of haematopoietic progenitors. Class 1 mutations can occur in genes including *KIT*, *FLT3*, *RAS*, etc. Class 2 mutations occur in genes encoding transcription factors or components of the transcriptional coactivation complex, which result in impaired differentiation and/or aberrant acquisition of self-renewal capabilities [2, 9]. These include transcription-factor fusions such as *MYH11-CBFB* and *RUNX1-RUNX1T1* and point mutations such as *CEBPA*. It has been shown in mouse models that class 1 mutations alone (such as *K-ras* mutations) only induce a myeloproliferative disorder but not a full-blown AML [10], and class 2 mutations leading to myeloid maturation block alone are not sufficient to cause leukaemia [11]. Likewise, in patients with AML, mutations within each class rarely occur in the same patient, whereas mutations between different classes often occur together in the same AML patient [9, 11]. These findings indicate that mutations from at least 2 classes are required for AML development as the two-hit model proposed by Gilliland and colleagues [12]. In addition, studies of patients with AML relapse demonstrated that the class 2 mutations were generally conserved in the relapse tumours. In contrast, the class 1 mutations frequently differ between the initial and relapse tumours [13]. This suggests that the relatively conserved class 2 mutations occur early in the disease trajectory, and the class 1 mutations are subsequent events. In addition, this implies that the development of AML follows specific and ordered patterns of driver mutation acquisition [5].

However, this two-hit hypothesis has been greatly challenged, especially by the development of whole-genome sequencing, which enables the identification of novel leukaemia-initiating genes. For example, novel driver genes regulating RNA splicing (such as *SRSF2* and *SF3B1*) or encoding epigenetic modifiers (such as *DNMT3A* and *IDH1/2*) have been identified [5]. In 2013, a study performed by the Cancer Genome Atlas Research Network identified 9 categories of genetic mutations that were related to AML pathogenesis. These include mutations in the gene encoding nucleophosmin (NPM1), tumour-suppressor genes, DNA-methylation-related genes, signalling genes, chromatin-modifying genes, myeloid transcription-factor genes, cohesin-complex genes, spliceosome-complex genes, and transcription-factor fusions [14]. In reality, mutations in genes encoding signalling proteins (class 1 mutations) have only been found in 59% of AML patients included in the study, even though it was believed to be a requirement of AML pathogenesis according to the traditional two-hit model [14]. In addition, it is discernible that many of the recently identified mutations do not fit within either of the 2 traditional classes. Following this study, by performing genetic profiling, including cytogenetic analyses and sequencing of 111 genes to samples from 1540 AML patients, Papaemmanuil and colleagues developed a new genetic classification of AML based on the pattern of co-occurrences of the previously described mutations, which included 11 genetic subgroups and had prognostic implications [5]. Co-mutation of specific genes has been shown to be associated with the prognosis [5]. These discoveries demonstrate the biological complexity of AML and refute the use of the oversimplified two-hit hypothesis to model leukaemogenesis. Also, the advanced understanding of the mutational profile and heterogeneity of AML prompts the need for discovering common AML characteristics and developing novel AML therapies that would transcend different genetic subgroups of AML.

1.1.3 Leukaemic stem cells (LSCs) in AML

A hierarchical organisation of cells carrying oncogenic mutations exist in cancers, where self-renewing cancer stem cells (CSCs) reside at the top of the hierarchy. The CSCs are able to sustain long-term maintenance of cancer and drive relapse [15, 16]. LSCs in AML have been a paradigm for the field of CSCs. The LSCs can initiate and

maintain AML by generating the whole population of non-LSC AML blasts which have limited self-renewal capacity [17, 18]. The AML LSCs were originally defined by Dick and colleagues as the only population that could engraft immunocompromised mice. They also showed that the AML LSCs, similar to haematopoietic stem cells (HSCs) and multipotent progenitors (MPPs), were solely CD34+CD38- [19]. However, the original functional and immunophenotypic definitions of the AML LSCs have been challenged in recent years. First, due to recent improvements in the immunocompromised mouse strains, the proportion of AML samples that engraft and the engraftment level both have increased. This results in the inconsistent estimation of the frequencies of AML LSCs in patient samples [17]. In addition, AML cells with leukaemia-initiating capacity have also been identified in the CD34+CD38+ [20] and CD34- fractions [21, 22]. This suggests the heterogeneity in the phenotype of LSCs. Understanding of the AML LSCs was further advanced by RNA sequencing (RNAseq), which showed that both CD34+ and CD34- LSCs were similar to progenitors/precursors instead of HSCs in most primary human AML patients. This indicated that leukaemic transformation occurred at the progenitor/precursor stage, imparting self-renewal capacity to the progenitors/precursors [22, 23]. Even though AML LSCs do not arise from HSCs directly in most primary human AML patients, extensive evidence has shown that early mutations probably happen in HSCs, giving rise to expanded clonal populations of pre-leukaemic HSCs in patients, from which AML evolves [24]. These early mutations are often found in genes encoding epigenetic modifiers such as *DNMT3A*, *ASXL1*, *IDH1/2*, and *TET2*, which are not sufficient for overt leukaemia and need additional mutations, which occur later, to drive leukaemogenesis [5].

LSCs or leukaemia-initiating cells (LICs) have also been identified in mouse models of AML. In a knockin mouse model of biallelic *CEBPA* mutations, transplantable LICs were identified only from a myeloid progenitor fraction (Sca-1-Mac-1^{lo}c-Kit+) rather than from a more differentiated fraction (Sca-1-Mac-1^{hi}c-Kit-) or HSC-containing fraction (Sca-1+) [13]. Using a mouse model of MLL-AF9 leukaemia, Krivtsov and colleagues found that LSCs were enriched in a granulocyte-macrophage progenitor (GMP)-like population [25]. Even though the MLL-AF9 leukaemia could be initiated by retroviral transduction of the MLL-AF9 oncogene into either murine GMPs

or HSCs, the LSCs were always enriched in the GMP-like population [26]. Using a similar mouse model of MLL-AF9 leukaemia, Somervaille and colleagues detected leukaemia-initiating activity in cells expressing myeloid lineage-specific antigens, downstream of the GMPs [27]. Despite the differences in the findings of Krivtsov and Somervaille, both of them suggested LSCs were more similar to progenitors/precursors rather than HSCs [25, 27]. This is consistent with the conclusion drawn by using patient AML samples as previously described.

It is crucial to characterise the LSCs as AML patient survival should be more closely related to the properties of LSCs than the blasts according to the CSC model [16, 18]. Indeed, in AML patients, a high percentage of CD34+CD38- cells at diagnosis directly correlated with a high minimal residual disease frequency and poor survival [28]. Gentles and colleagues also developed an LSC gene signature and found that high expression of the LSC gene signature was associated with poorer outcomes among AML patients with either normal karyotypes or chromosomal abnormalities [29]. In addition, the elimination of the LSCs should be essential for the cure of AML as according to the CSC model any remaining LSCs that survive the therapy would be able to drive relapse [30]. Hence, it is important to develop AML therapies that can either selectively eradicate LSCs or eradicate both the AML blasts as well as the LSCs [7]. In agreement with the putative CSC model, Shlush and colleagues demonstrated that therapy-resistant AML cells already existed at diagnosis and AML relapse originated from either LSCs or AML cells with a committed immunophenotype that retained strong stemness transcriptional signatures [31]. It has also been shown that pre-leukaemic HSCs can survive chemotherapy, which may contribute to AML relapse [1, 24].

1.1.4 Current therapeutic strategies in patients with AML

The standard therapeutic strategy in patients with AML has not changed substantially in decades [32]. Initial assessment evaluates whether a patient is eligible for intensive induction chemotherapy. Although age is associated with a higher risk of treatment-related mortality after intensive therapy, it has been shown that elderly patients who are medically fit may still benefit from the intensive induction chemotherapy. It has been recommended that older patients should still be con-

sidered as candidates for intensive induction chemotherapy unless they have poor performance status and significant comorbidities [32]. The combination of 7 days of cytarabine and 3 days of anthracycline (the "7+3" regimen) remains the standard induction therapy although better regimens have been proposed [1, 33]. It leads to a complete remission rate of 60-85% in patients who are 60 years or younger and 40-60% in patients older than 60 years of age [1]. Most patients who are in complete remission ultimately have a relapse without consolidation therapy. Post-induction consolidation therapies comprise intensive chemotherapy and haematopoietic cell transplantation (HCT). Consolidation chemotherapy includes several cycles of intermediate or high doses of cytarabine [34]. Whether higher doses of cytarabine are superior to intermediate doses and the optimal number of cycles remains unknown, especially for patients with older age. Currently, it is recommended that patients who are older than 60 years of age and with favourable European LeukemiaNet (ELN) genetic risk should receive 2-3 cycles of intermediate-dose cytarabine; older patients with intermediate or unfavourable genetic risk are less likely to benefit from the cytarabine-based consolidation regimen [1].

Allogeneic HCT is reserved for patients unlikely to achieve long-term remission with other conventional therapies. The procedure has a strong anti-leukaemic effect because of pre-transplant chemoradiotherapy conditioning and an immune-mediated graft-versus-leukaemia effect [1]. The selection of candidates for HCT is complicated, based on the patient's genetic risk, response to induction chemotherapy, patient fitness, donor factors, etc [32, 34]. Reduced-intensity conditioning allows older patients aged up to 75 years to undergo HCT, but, in reality, a large number of elderly patients may still not be appropriate candidates for HCT due to comorbidities, inability to achieve a complete remission, etc [32]. Even though allogeneic HCT provides the strongest anti-leukaemic effect, relapse after transplantation is still common and can cause dismal outcomes. In particular, a recent study suggested relapse after transplantation was associated with dysregulation of pathways that influence immune function [35].

1.1.5 Emerging therapeutic strategies in patients with AML

Patients with AML are still in great need of novel therapies, given the limitation of current treatments. Increased understanding of the heterogeneity of AML also partly explains why only a subpopulation of AML patients can achieve prolonged remission with the "7+3" regimen followed by consolidation therapy. The development of novel and less cytotoxic AML treatment is particularly important for patients at high risk of relapse and elderly patients who are ineligible for intensive therapy. Novel targets of AML treatment involved diverse pathways promoting the survival of AML cells, such as pro-apoptotic pathways, cell cycle regulation, immune evasion, interaction with the microenvironment, etc [36, 37]. Thanks to a better characterisation of the LSCs, several novel AML treatments targeting LSCs have also been described, which may eradicate this chemotherapy-resistant and relapse-driving population [7]. First, LSC-specific immunophenotypes have been identified, and antibody-based treatment can be used to target LSCs. For example, anti-CD123 and anti-CD33 antibodies are under investigation [7, 34]. Secondly, LSCs have distinct metabolic and epigenetic properties, based on which LSC-targeting treatment may be developed [7]. For example, *IDH1/2* mutations result in the overproduction of the oncogenic metabolite 2-hydroxyglutarate and subsequent epigenetic changes [38]. Several small molecule inhibitors of mutant IDH have been approved for treating patients with relapsed or refractory AML harbouring *IDH1/2* mutations [39]. Besides, interactions between LSCs and the BM microenvironment may be targeted to suppress LSC growth. There is considerable evidence suggesting that BM microenvironmental changes can contribute to the initiation of haematological malignancies, and leukaemia can induce BM microenvironment remodelling, which, in turn, promotes leukaemia development [40]. For example, NF-$\varkappa$B signalling has been shown to be upregulated in LICs, suggesting that TNF-α in the BM microenvironment could be a promising therapeutic target [41]. However, currently, no therapy targeting the AML BM microenvironment, but several inhibitors that may block the interaction between LSCs and the niche cells are under study such as a small-molecule inhibitor of E-selectin [42]. These novel therapeutic strategies can be used in combination with current chemotherapy, eradicating the bulk AML cells and relatively rare LSCs simultaneously [7, 36].

1.1.6 Survival of patients with AML

Among all types of acute leukaemia, AML has the shortest survival with a 5-year overall survival (OS) of 24.0% as of 2016 [4]. Despite recent progress in understanding the pathophysiology of AML and novel therapeutic targets being proposed, a significant decrease in the mortality rate of AML has not been observed during 2005–2016 [4]. The older patients (60 years and older) have a 5-year survival rate of < 10-15% according to data from MD Anderson [33]. This indicates there is still a significant gap between the findings from the bench and improved clinical outcomes from the bedside. Further translational studies are still needed to improve the survival of patients with AML, especially elderly patients.

1.2 Murine models of AML

1.2.1 A murine model of bi-allelic *CEBPA* mutant AML

CEBPA mutations exist in 7-15% of normal karyotype AML patients [43]. Among AML patients with normal karyotype, biallelic *CEBPA* mutations are associated with improved survival compared with monoallelic mutations and *CEBPA*-wild-type [44]. The recent World Health Organization (WHO) classification of tumours of the haematopoietic and lymphoid tissues has classified AML with biallelic mutations of *CEBPA* as an independent disease entity [45]. When Papaemmanuil and colleagues were classifying AML into mutually exclusive genomic subtypes based on patterns of co-mutation, AML with biallelic *CEBPA* mutations was also identified as a distinct genomic subtype, which made up 4% of the 1540 patients studied [5].

CEBPA is a transcription factor that coordinates cellular differentiation with cellular proliferation [46]. CEBPA has 2 translation isoforms: a full-length form, known as p42, and a shorter form, known as p30. These isoforms share the same carboxyl terminal, which contains the basic region-leucine zipper (BR-LZ) DNA-binding domain, but the p30 lacks the amino-terminal 117 amino acids [47]. CEBPA can directly activate the transcription of lineage-specific genes, such as granulocyte-specific genes and eosinophil-specific genes [48, 49]. Only the full-length p42 can interact with TBP/TFIIB and CBP/p300. But both CEBPA isoforms can interact

with SWI/SNF, and other transcription factors such as PU.1, GATA1 and RUNX1 (by the BR-LZ domain) (Figure 1.1) [47]. As a result, the p30 is still able to induce eosinophil lineage commitment, despite failing to induce terminal eosinophil differentiation [50]. In terms of CEBPA's ability to block cell proliferation, it has been shown that it is due to CEBPA-mediated repression of the E2F complex, which can regulate genes required for cell-cycle progression [51]. Since the p30 lacks the E2F-repressing activity, the ability to arrest cell proliferation seems specific to the p42 [47].

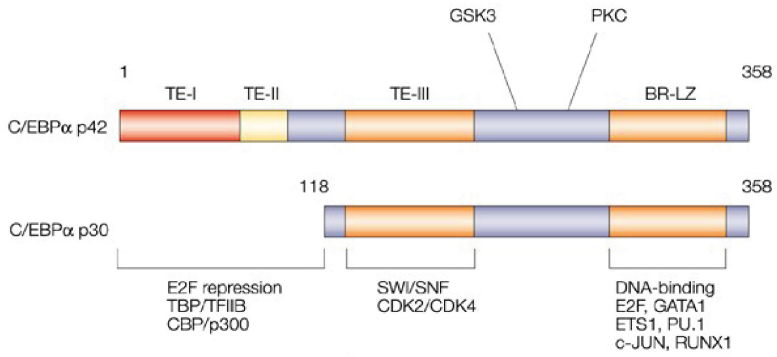

Figure 1.1. Structure of the CEBPA translation isoforms

Schematic depicting the location of specific functional domains within the CEBPA protein. The basic region leucine zipper (BR-LZ) DNA-binding domain has been highlighted and the transcription factors it can interact with were also shown. Besides, the 3 transactivation elements (TE-I, TE-II and TE-III) are labelled, and the transcription apparatus they can interact with is listed. The numbers represent the position of the relevant start/stop codon within the open reading frame. Adapted from (Nerlov, 2004)[47].

There are 2 types of *CEBPA* mutations that have been described in AML. The first type of mutation is a frameshift at the N-terminal of CEBPA that is unique to the p42, preventing expression of the p42 but allowing expression of the p30 [52]. The second type is C-terminal mutations that affect the BR-LZ domain, leading to a deficiency in the DNA-binding ability of the mutant protein [53]. Most AML patients with *CEBPA* mutations carry biallelic mutations, and most of the biallelic mutations are 2 heterozygous mutations [54]. Usually, one allele contains the C-terminal mutation in the BR-LZ DNA-binding domain, whereas the other contains the N-terminal 'p30-only' mutation [47]. To better understand how the N-terminal and C-terminal *CEBPA* mutations collaborate to promote leukaemogenesis, the Nerlov laboratory previously generated a genetically accurate murine model of bi-allelic *CEBPA* mutant AML by combining knock-in mutations that mimic the

common N- and C-terminal mutations, respectively [13, 55]. The *Cebpa* mutant AML was initiated by transplanting foetal liver (FL) cells carrying the bi-allelic *Cebpa* knock-in mutation together with wild-type (WT) competitor BM cells into lethally irradiated recipient mice [13]. Using this murine model, the Nerlov laboratory demonstrated that the C-terminal *Cebpa* mutation could induce preleukaemic expansion of HSCs/MPPs. In contrast, the N-terminal *Cebpa* mutation contributed to residual myeloid lineage commitment, which explained why the combination of C- and N-terminal *CEBPA* mutations were commonly found in *CEBPA* mutant AML patients [13].

Even though this murine model of *CEBPA* mutant AML can accurately model the clinical pattern of biallelic *CEBPA* mutations, it still has several limitations which prevent its further use in studying the disease trajectory in a reproducibly way and better understanding the dynamic change of BM microenvironment throughout leukaemogenesis. In detail, the recipient mice became moribund after a long latency of 20-50 weeks after transplantation without consistent kinetics. In addition, to allow more efficient engraftment of the *Cebpa* mutant FL cells, the recipient mice had to be lethally irradiated, causing significant damage to the BM microenvironment [13]. To overcome these hurdles, the Nerlov laboratory further developed this murine model by transplanting purified *Cebpa* mutant LICs (Sca-1-Mac-1^{lo}c-Kit+) into sub-lethally irradiated recipient mice. In this way, AML reproducibly progressed over 4 weeks to a terminal phase that included leukocytosis and anaemia, loss of normal haematopoietic cells in BM, spleen infiltration by leukaemic cells and splenomegaly [56].

Additionally, sub-lethal irradiation minimised damage to the BM microenvironment, particularly to the BM stromal cells [56]. More importantly, when the LICs were purified from this murine model, they completely lost the ability to re-initiate leukaemia development after 48 hours of *ex vivo* culture, demonstrating their need for an adequately conditioned *in vivo* environment to maintain their leukaemogenic potential [56]. Therefore, this murine model of *CEBPA* mutant AML is kinetically tractable and strictly dependent on the BM microenvironment to maintain its malignant phenotype. It allows AML-induced BM microenvironmental changes to be characterised

during distinct phases of disease development.

1.2.2 Other mouse models of AML, including xenograft mouse models

Modelling of AML in the mouse has largely relied on introducing AML-driving oncogenes into the HSPCs. One of the commonly used viral transduction-based models is a murine model of *MLL-AF9* leukaemia. In this model, leukaemia is initiated by transplanting HSCs or progenitors transduced by MLL-AF9-encoding retroviruses into sub-lethally irradiated recipient mice [25, 26]. Viral vectors encoding AML-associated *MOZ-TIF2*, *MLL-ENL*, *AML1-ETO*, and *MLL-GAS7* fusion genes have also been used to transform murine HSPCs and initiate AML in the mouse [57].

Another way of modelling human AML in the mouse is by transplanting human-derived AML cells into immune-compromised recipient mice. The xenografts can be derived from either human primary AML cells or AML cell lines [58]. Since maintenance of primary human AML cells for even a few days remain challenging [59], the patient-derived xenograft (PDX) mouse model enables recapitulation of the histology, genomic pattern, and heterogeneity of the original disease [60]. Most importantly, the PDX mouse model can be used to evaluate the efficacy of novel therapeutic strategies and standard chemotherapies *in vivo* and explore the mechanisms by which patient AML cells survive specific therapies [61, 62]. The development of more immune-compromised mouse strains such as the severe combined immunodeficient (SCID)-non-obese diabetic (NOD)-IL2rγ^{null} (NSG) strain facilitated the engraftment of primary AML patient samples, allowing more widespread use of the PDX mouse model of AML [63]. One limitation of such a PDX mouse model of AML is that some human-specific BM microenvironmental factors can not be provided by the mouse microenvironment [64]. This might be circumvented by the use of material scaffolds to create a humanised microenvironment or co-transplanting patient-derived mesenchymal stromal cells (MSCs) with the AML cells [64].

1.3 The BM microenvironment and myeloid malignancies

1.3.1 The BM microenvironment and its implications in normal haematopoiesis

Niches are local microenvironments that maintain and regulate stem cells [65]. HSCs also reside in a specific BM microenvironment, which is known as the HSC niche. The HSC niche comprises both BM stromal cells and haematopoietic cells such as megakaryocytes, which regulate the quiescence, proliferation and differentiation of HSCs (Figure 1.2) [66]. The BM niche can be divided into 2 anatomically different niches, the central bone marrow (CBM) niche, which is located in the inner BM, and the bone-lining (BL) niche, which is proximal to the bone surface. The perivascular stromal cells (PVCs) and vascular endothelial cells (VECs) are the main cellular components of the BM niche, which are also the main source of SCF and CXCL12, 2 critical secreted factors for HSC maintenance [67, 68]. The PVCs were also known as CXCL12-abundant reticular (CAR) cells [69], which were shown to largely overlap with leptin receptor-expressing (LEPR+) cells [70]. The PVCs have also been identified using the expression of other marker genes, including Nestin [71] and NG2 [72]. Recent single-cell RNAseq revealed that there was considerable heterogeneity and overlapping among these BM niche cell types. For instance, LEPR+ cells could also be found in VECs, and a large number of the sinusoidal VECs minimally expressed CXCL12 and SCF. In addition, it was shown that LEPR+, NG2+, and Nestin+ cells were molecularly distinct, even though LEPR, Nestin, and NG2 had all been identified as markers of PVCs [73]. Other cellular components of the BM niche include the osteoblasts (OBs), sympathetic nerves, macrophages, and megakaryocytes [66, 74]. Even though a clear delineation and precise function of distinct cellular components of the BM niche remains controversial, several imaging studies have shown that in both the CBM and BL niche, HSCs localise adjacent to the endothelial cells and their associated PVCs (a so-called perivascular niche) [69, 72] . This strongly supports the role of the BM niche, especially the perivascular niche, in HSC maintenance. By contrast, the OBs expressed SCF and CXCL12 at

much lower levels and few HSCs were in contact with the OBs, which refuted that the OBs were an important HSC regulator [66, 72, 75].

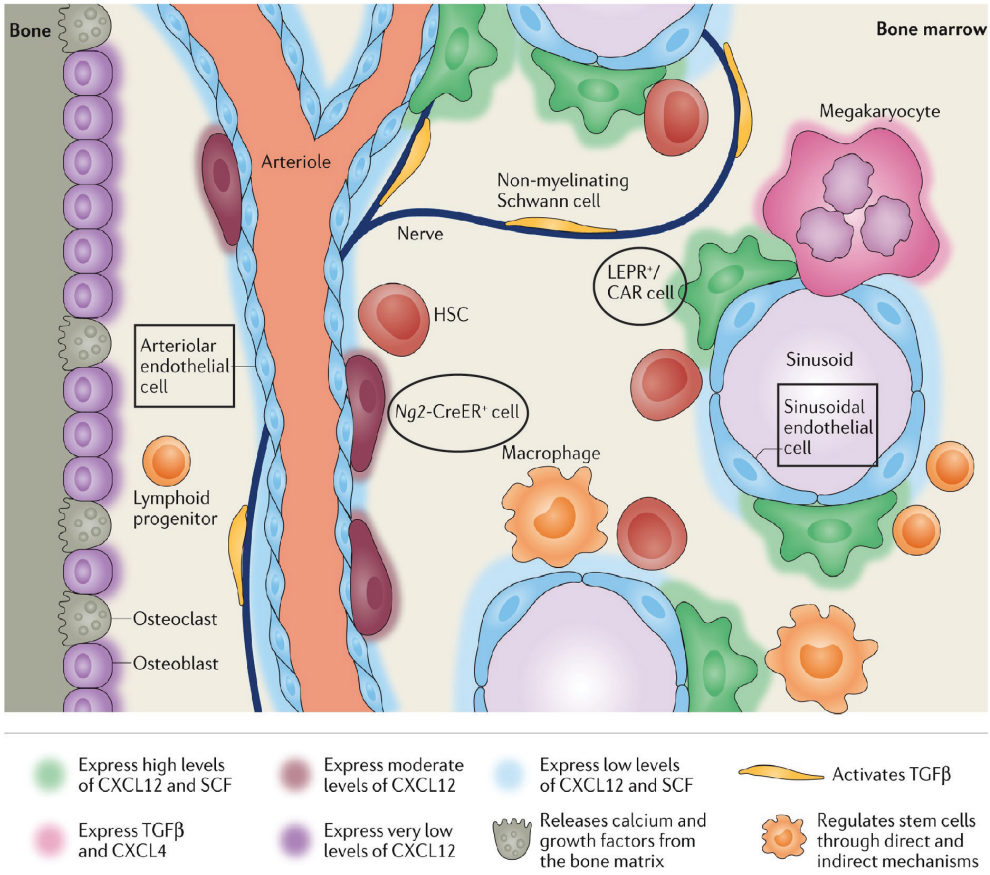

Figure 1.2. Schematic of the BM microenvironment

Common haematopoietic and non-haematopoietic cells that reside in the BM are shown. The perivascular stromal cells (PVCs) are in black circles, including CXCL12-abundant reticular cells (CAR cells)/leptin receptor-expressing cells (LEPR+ cells) and NG2+ cells. Vascular endothelial cells (VECs) are in black squares, including sinusoidal and arteriolar endothelial cells. Adapted from (Crane et al., 2017) [74].

Similar to the HSCs, the maintenance of LSCs is also dependent on their local niche, which could contribute to disease progression, resistance to therapy, and relapse [76]. Two mechanisms whereby the BM niche contributes to leukaemogenesis have been proposed: primary functional or genetic changes in the BM niche may induce leukaemia; the niche which has been remodelled by malignant haematopoietic cells may further support leukaemia progression and survival [77].

1.3.2 Myeloid malignancies can originate in the BM niche

There is considerable evidence that genetic modification of the BM niche cells can induce myeloid lineage defects such as myelodysplastic syndromes (MDS) and myeloproliferative neoplasms (MPN) that can progress to transplantable leukaemia. In 2007, Walkley and colleagues observed that retinoic acid receptor γ deficiency in the BM niche induced MPN, revealing the capability of the BM niche to be the initiator of haematological malignancies [78]. In 2010, Raaijmakers and colleagues deleted *Dicer1*, specifically in mouse osteoprogenitors, which resulted in MDS and AML [79]. This again supports the BM niche as the site of the initiating event that leads to haematological malignancies. Subsequent studies demonstrated that the genetically modified BM niche cells contributed to disease development by producing pro-inflammatory cytokines or chemokines such as TNF-α, G-CSF, and CCL3 [78, 80, 81].

In terms of patients with MDS or AML, Blau and colleagues also detected chromosomal aberrations in mesenchymal stromal cells isolated from the patients and the genetic alterations in the stromal cells were distinct from those in the leukaemic cells [82, 83]. Even though this is not direct evidence that proves the pathogenesis of MDS and leukaemia in human patients can be initiated by genetic alterations in the BM niche cells, this still raises the possibility that BM microenvironmental changes may precede and induce the leukaemic transformation of the haematopoietic cells.

1.3.3 AML cells can cause niche remodelling

It is well-established that AML progression induces BM microenvironmental changes and such niche changes are able to favour AML growth but impair normal haematopoiesis [84, 85]. At the cellular level, changes to different types of BM niche cells have been observed in different studies. In 2014, Hanoun and colleagues used a murine model of MLL-AF9 AML and showed that the development of leukaemia led to increased sinusoidal VEC frequencies, expansion of Nestin-GFP+ PVCs, and disruption of sympathetic nervous system (SNS) nerves. In particular, they showed that the deletion of β2 adrenergic receptors in the BM niche promoted leukaemic BM infiltration, indicating the remodelled BM niche (the SNS network at least) had the ability to

support AML development [86]. In 2014, Krevvata and colleagues showed that the number of OBs decreased significantly during AML, and the depletion of OBs was shown to be supportive of AML progression by using both AML patient samples and mouse models [87]. In 2018, Duarte and colleagues also used the mouse model *MLL-AF9* AML, where they observed a significant reduction in endosteal VECs by both imaging and flow cytometry, even though the total number of VECs in the overall BM stroma was preserved. In addition, they suggested that normal HSC numbers can be increased, and chemotherapy efficiency could be improved by rescuing this loss of endosteal VECs [85]. This uncovers the change in BM vasculature during AML and its implication in leukaemogenesis. The importance of VECs for AML progression and chemoresistance was further supported by the finding that the direct interaction of AML cells with endothelial VCAM-1 (via CD98) [88] and E-selectin [89] supports leukaemia progression and chemoresistance. Collectively, these findings suggest that AML cells can induce adaptive changes to the BM microenvironment they reside in, which in turn support their expansion and survival.

The underlying mechanisms by which the malignant cells remodel the BM niche have been widely investigated. First, cytokines or chemokines expressed by the malignant haematopoietic cells can remodel the BM nice. For instance, Duarte and colleagues showed that CXCL2 and TNF, which were upregulated in the endosteal AML cells, contributed to the remodelling of endosteal vessels [85]. Frisch and colleagues also observed that an increased expression level of CCL3 in malignant haematopoietic cells could play a role in the AML-related inhibition of OBs [90]. Besides, Schepers and colleagues identified a role of thrombopoietin (TPO) and CCL3 in chronic myelogenous leukaemia (CML)-associated expansion of osteoblastic lineage cells in a mouse model of CML [91].

Second, exosomes secreted by the malignant haematopoietic cells can transform the BM nice. Exosomes are microvesicles containing mRNA, miRNA and angiogenic proteins which can be secreted by both normal and malignant cells [92]. In 2018, Kumar and colleagues demonstrated that AML-derived exosomes could target the mesenchymal stromal progenitors and lead to their decreased osteoblast lineage differentiation potential [84]. Third, Passaro and colleagues demonstrated that an

AML-induced hypoxic environment was able to drive an expansion of the VECs, alter vascular permeability in the BM, and change the transcriptome of VECs. This supports AML-induced hypoxia as one of the mechanisms by which AML cells remodel the BM niche cells [93].

1.3.4 The BM microenvironment is a pro-inflammatory environment during myeloid malignancies

The pro-inflammatory milieu is a common feature of the BM microenvironment conditioned by the various types of myeloid malignancy. Elevated levels of pro-inflammatory cytokines and chemokines such as IL-1β, IL-6, TNF-α, and CCL3 have been observed in the BM or PB of patients with AML [94–97]. The source of the increased level of pro-inflammatory secreted factors can be both the non-haematopoietic BM niche cells and malignant haematopoietic cells. First, as described above, several studies have implicated the aberrant production of pro-inflammatory factors by the genetically altered BM stromal cells as a contributing mechanism of haematological malignancy initiation [78, 80, 81]. Second, the AML-exposed human BM stromal cells showed increased expression of pro-inflammatory factors such as IL-1β, IL-8, and IL-6 [98, 99]. A similar trend of increased production of pro-inflammatory cytokines such as IL-6 and TNF-α by the BM stromal cells has also been identified in patients with MDS [100]. More importantly, AML cells overproduce a similar range of pro-inflammatory cytokines and chemokines as the leukaemic BM stromal cells: Steinbach and colleagues showed that *CCL23* was over 50-fold overexpressed in the AML cells compared to normal BM cells using Taqman assays [101]; Wang and colleagues also reported overexpression of *CCL3* in human AML cells [97]. Besides, in 2020, Zhang and colleagues demonstrated overexpression of IL-6 in human AML cells at both the mRNA and protein levels [102]. These pro-inflammatory factors were shown to be capable of promoting AML cell expansion whilst inducing BM failure [94, 97, 102]. Similar phenomena have been observed in AML mouse models as well. For instance, the Nerlov laboratory has performed bulk RNAseq in AML cells and AML-exposed BM stromal cells isolated from the murine model of *CEBPA* mutant AML. By comparing the transcriptome of leukaemia cells and leukaemic niche cells to the non-leukaemic controls, we also observed the upreg-

ulation of different cytokine and cytokine receptor genes during AML development [56]. More details of the RNAseq result are described in the following result chapters.

The above findings not only confirm the role of the leukaemic BM niche as a pro-inflammatory environment but also suggest that AML-associated BM niche modifications can generate AML-promoting paracrine signals that mimic autocrine signalling by the AML cells themselves. This emphasises the engagement of the AML secretome and BM microenvironment in leukaemogenesis. Therefore, the deregulated AML secretome and pro-inflammatory BM microenvironment may represent a potential therapeutic target for AML.

1.4 AML in the elderly

1.4.1 Characteristics of AML in the elderly

AML is a disease of elderly adults. According to the statistics from the National Comprehensive Cancer Care Network (NCCN), the median age of patients with AML at diagnosis is 67 years, and one-third of the patients are older than 75 years. Similarly, the Surveillance, Epidemiology, and End Results (SEER) Cancer Statistics Review reports that the age-adjusted incidence of AML in adults aged 65 years or older was 20.1 per 100,000 person-years, whereas that in adults aged 65 years or younger was just 2.0 per 100,000 person-years. Not only does the incidence of AML increase significantly with age, but also elderly AML patients have poorer outcomes compared to younger patients. In 2006, Appelbaum and colleagues showed that AML patients' response rates (to induction therapy) and median OS dropped with age [103]. In 2009. Buchner and colleagues demonstrated that age was an independent risk factor for an inferior overall/relapse-free survival and remission duration [104]. The SEER statistics also revealed an ˜10% decrease in 5-year OS with each increasing decade of life in patients diagnosed with AML.

The poor outcome of elderly AML patients can be explained by both the disease-related and patient-related features of AML in the older population. Regarding the disease-related features, the incidence of unfavourable karyotypes was higher in elderly patients, while the incidence of the favourable karyotype decreased with age

[103, 104]. The frequency of AML secondary to MDS, MPN, and treatment-related myeloid neoplasms increases with age, which also contributes to the inferior outcome [105]. Besides, AML in the elderly has more frequent expression of the multidrug resistance glycoprotein MDR1, which is independently associated with AML-related poor outcomes [106]. In terms of the patient-related features, elderly patients with AML have more comorbidities and poorer performance status than younger patients. It has been reported that the performance status of AML patients decreases with age, and poor performance status is independently associated with an adverse prognosis in older patients with AML [107, 108]. In addition, because of comorbidities and poor performance status in elderly AML patients, a large proportion of them are not able to tolerate standard chemotherapy and are treated with supportive care alone (including hydroxyurea) or less-intensive treatment approaches such as hypomethylating agents and low-dose cytarabine [1, 4]. Indeed, elderly AML adults are more likely to experience treatment-related mortality (TRM) [105, 109, 110].

However, the age-associated poor prognosis of AML patients can not be fully explained by the increased incidence of high-risk cytogenetics, inferior performance and intolerance of intensive chemotherapy. Both Appelbaum and Buchner have shown that among similarly-treated patients from the same cytogenetic risk group, the elderly still experience lower OS than the younger [103, 104]. This may suggest that age-related differences extrinsic to leukaemic cells, such as systemic factors or the BM microenvironment, may contribute to AML therapy resistance and disease relapse. Additionally, identifying the molecular basis for the increased ability of an aged, compared to the young, BM microenvironment to propagate the AML cells would be critical to developing better treatment for elderly AML patients.

1.4.2 HSC ageing and AML

Both HSC-intrinsic and HSC-extrinsic changes during ageing can explain the age-associated rise in AML incidence. In this section, I will review and discuss the cell-intrinsic mechanisms involved in HSC ageing and their implication for AML development. First, AML development depends on the stepwise acquisition of a series of somatic mutations along one's lifetime, which is directly linked to ageing. Several studies have shown that haematopoietic clones harbouring early leukaemia-initiating

mutations, such as *DNMT3A* and *TET2*, are more frequently found in healthy older adults [111, 112]. This phenomenon is known as age-related clonal haematopoiesis, which is strongly associated with haematological malignancies [111, 113]. To illustrate, the mutated HSC clones can outgrow the other WT clones, and the increased clonal size facilitates subsequent AML-associated mutations to occur, leading to AML development eventually [114–116]. Desai and colleagues focused on the association between AML and clonal haematopoiesis and demonstrated that healthy individuals with pre-existing mutations in genes recurrently mutated in haematological malignancies had statistically increased odds of developing AML; similar observations have also been reported by Abelson and colleagues [117, 118].

Second, autophagy decreases in the aged HSCs, and several studies suggested that impaired autophagy could be associated with AML development. Simon and colleagues demonstrated that the deletion of key autophagy genes, *Atg7* and *Atg5*, led to a lethal pre-leukaemic phenotype in mice [119, 120]. In human AML samples, several autophagy genes are mutated or expressed at lower levels, which further supports autophagy's involvement in AML development [121].

Third, ageing-associated cancer incidence has traditionally been associated with a less efficient DNA damage response (DDR). It has been shown that aged HSCs present with more gH2AX foci, a DNA damage marker, compared to young HSCs [122, 123]. Therefore, it seems plausible that the aged HSCs are more likely to experience leukaemic transformation due to their deficient DDR. However, a more recent study demonstrated that aged HSCs could repair DNA damage as efficiently as their young counterparts, questioning the impact of ageing on DNA damage and genomic instability [124]. Therefore, it remains undetermined whether changes in DDR efficiency during ageing can explain the age-related AML predisposition.

1.4.3 Ageing of the BM microenvironment and AML

In this section, I will summarise and discuss the HSC-extrinsic changes during ageing and how they may contribute to increased AML incidence. Several BM microenvironmental alterations have already been identified during physiological ageing. Regarding endothelium, the number of type H endothelial cells (ECs) and distal

arteriolar ECs, both of which were found in the endosteal niche, reduced during age; the absolute number of total ECs remained unchanged or increased significantly according to inconsistent findings from different labs, which indicated a vasculature expansion in the CBM niche anyway [125–127]. The frequency of CBM perivascular MSCs also increased significantly during ageing [125]. Regarding the endosteum, a decrease in the frequency of OBs was identified [128]. Besides the changes to the BM niche cell frequencies, ageing-associated elevated BM levels of multiple pro-inflammatory cytokines and chemokines, including IL-1β, TNF-α, IL-6, TGF-β, CCL3, and CCL5, have been reported in mice and humans [125, 128–132]. The increased BM extracellular fluid levels of these cytokines can be due to either high local production by the BM niche cells or distant overproduction of circulating cytokines [130, 131]. These changes could contribute to the age-associated decline in lymphocyte and erythrocyte production, and the resulting platelet-myeloid bias of the cellular output [125, 128, 130].

Since BM niche alterations can initiate certain myeloid malignancies, as shown using mouse models [77], the age-related BM microenvironmental changes might play a role in AML initiation and progression in humans as well, leading to the increased AML prevalence with advancing age. Another piece of evidence that supports the aged BM microenvironment contributing to the increased AML incidence with age is that the changes in the BM niche induced by AML and ageing are largely similar. In detail, the numbers of OBs and endosteal VECs decrease during ageing and AML; protein levels of pro-inflammatory factors, such as IL-6 and TNF-α, increase in both the aged and leukaemic BM. In agreement with this, the Nerlov laboratory found that the AML-exposed and aged BM niches were transcriptional and functionally alike by performing gene expression profiling and transplantation assays [56]. However, the specific molecular mechanism by which the aged BM niche supports AML development remains to be determined.

1.5 CRISPR-Cas9 genome editing and CRISPR screens

1.5.1 CRISPR-Cas9 as a tool for genome editing

The CRISPR-Cas9 nuclease system was originally identified as a microbial immune system, which has subsequently been repurposed for genome editing in mammalian cells by heterologous expression of the CRISPR components after certain optimisation [133]. In this system, a nuclease, Cas9, is guided by 18-20nt RNA molecules to specific genomic loci and induces DNA double-strand breaks (DSBs) [134]. The 18-20nt guide RNA sequence is required to be fused with a scaffold RNA sequence known as the trans-activating crRNA (tracrRNA) to create a chimeric, single-guide RNA (sgRNA) [134]. The DNA target of any guide RNA must be directly upstream of a 5'-NGG protospacer-adjacent motif (PAM), and Cas9 can be directed toward almost any target DNA of interest by altering the guide RNA sequence within the sgRNA as long as the target DNA is immediately preceding a PAM (Figure 1.3) [134, 135]. Therefore, the CRISPR-Cas9 system is significantly easier to design/-customise compared to other genome editing technologies, including zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) [135]. One of the widely used systems for delivering components of the CRISPR-Cas9 system is the DNA plasmid encoding both the Cas9 protein and the guide RNA. The plasmids can be introduced into cells by electroporation, hydrodynamic injection, viral vectors, etc [136].

Once the Cas9-mediated DSB occurs, it will usually be repaired by either the indel-forming nonhomologous end joining (NHEJ) or homology-directed repair (HDR) upon simultaneous introduction of a donor DNA template [133, 138]. During NHEJ, which is now known as classical NHEJ, DSBs are repaired by direct ligation of DNA either perfectly or imperfectly [139, 140]. The latter will generate deletions or insertions most of which occur at the 3' end of the target DNA sequence and are usually small in size (< 10 bp) [138, 141]. Therefore, NHEJ can be harnessed to disrupt target protein expression, as NHEJ events occurring within the coding region can result in frameshifts and premature stop codons, leading to subsequent gene

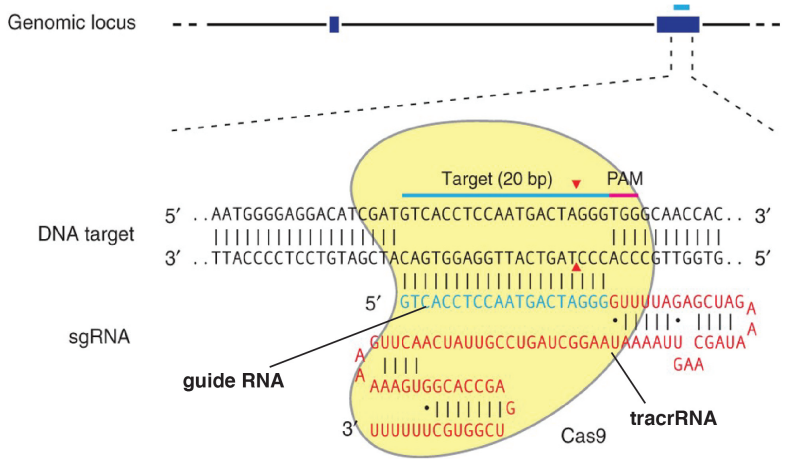

Figure 1.3. Schematic of the CRISPR-Cas9 nuclease system

The Cas9 (yellow) is guided to genomic DNA by a sgRNA, which is a fusion of a 20nt guide RNA sequence and a scaffold RNA sequence (also known as tracrRNA). The target DNA must be directly upstream of a 5'-NGG protospacer-adjacent motif (PAM) in order to be targeted by the sgRNA. Cut sites of the Cas9 (red triangles) are usually at the 3' end of the target DNA sequence. Adapted from (Ran et al, 2013) [137].

knockout [137]. Simultaneous introduction of multiple sgRNAs targeting different genomic loci of the same gene can result in larger deletions within the target gene of interest, facilitating the desired gene knockout [138]. More recently, non-canonical forms of NHEJ that coexist with the classical NHEJ, namely alternative end-joining (alt-EJ) or microhomology-mediated end joining (MMEJ), have also been identified in mammalian cells, including murine and human cells [140, 142]. MMEJ is believed to be highly error-prone and has been correlated with an increased frequency of larger deletions [139, 143].

Presumably, the specificity of the CRISPR-Cas9 system is strictly controlled by the 18-20nt guide RNA sequence, which only binds to and guides Cas9 to the complementary DNA sequence that lies next to the PAM. However, it has been shown that off-target mutations can happen in the DNA sequence that differs by up to 5 positions from the intended on-target sequence. Besides, mismatches occurring in the 5' half of a guide RNA are better tolerated than those occurring in the 3' half [144]. Hsu and colleagues also showed that the specificity of the CRISPR-Cas9 system was sequence- and locus-dependent, and they proposed several rules of guide RNA design to maximise the specificity of the Cas9 [145]. For example, guide RNA sequences should not have potential off-target sequences which have fewer than

three mismatches [145]. As the understanding of Cas9-mediated off-target mutations increases, more studies on guide RNA design and on-target/off-target efficiency prediction have been published, which greatly facilitate the selection of more active and specific guide RNAs [146–149].

1.5.2 Pooled CRISPR-Cas9 screens

Forward genetic screens are a powerful tool to explore whether a specific gene contributes to a specific phenotype, biological process, or disease. Forwards genetic screens involve perturbation of the expression of candidate genes and selecting the cells with the desired phenotype. This is followed by mapping the phenotype to the genotype of the selected cells [150]. Such forward genetic screens are largely based on RNA interference (RNAi) technology and the previously described CRISPR-Cas9 technology. Compared to RNAi-based screens, CRISPR-based screens are considered superior, especially during large-scale screens, for several reasons: 1) small interfering RNA (siRNA) binding can be non-specific, leading to extensive phenotypic changes derived from unintended interactions between siRNA and mRNA; 2) RNAi can only lead to knockdown instead of knockout of the gene, and the knockdown efficiencies dan vary [150–152]. Indeed, Evers and colleagues demonstrated that the CRISPR-based screens outperformed the RNAi-based screens with a lower false-discovery rate and better consistency across different cell lines, even though CRISPR technology also has off-target effects [153].

CRISPR screens can be conducted as an arrayed or pooled screen. For arrayed screens, sgRNAs are individually introduced to cells that sit in separate culture wells in multi-well plates. For pooled screens, a library of multiple sgRNAs is simultaneously introduced to a bulk of cells, resulting in stable expression of only 1 sgRNA in most cells (Figure 1.4). It is conceivable that compared to arrayed screens, pooled screens are less costly and labour-intensive [151, 154]. In addition, pooled CRISPR screens can be performed *in vitro*, *ex vivo*, and *in vivo*. For *in vivo* pooled CRISPR screens, sgRNAs can be introduced to cells *in vitro*/*ex vivo* followed by transplantation or sgRNA-carrying plasmids/lentiviruses can be introduced to mice directly [155]. Pooled CRISPR screens can be further categorised into positive or negative selection screens. Positive selection screens are aimed at identifying genetic

perturbations that allow cells to survive specific selective pressure, such as drugs, toxins, or pathogens [150]. In contrast, the goal of negative selection screens is to identify sgRNA-induced genetic perturbations that cause cells to be depleted during selection. This allows the identification of essential genes, which could potentially be used as druggable targets for cancers [150, 154].

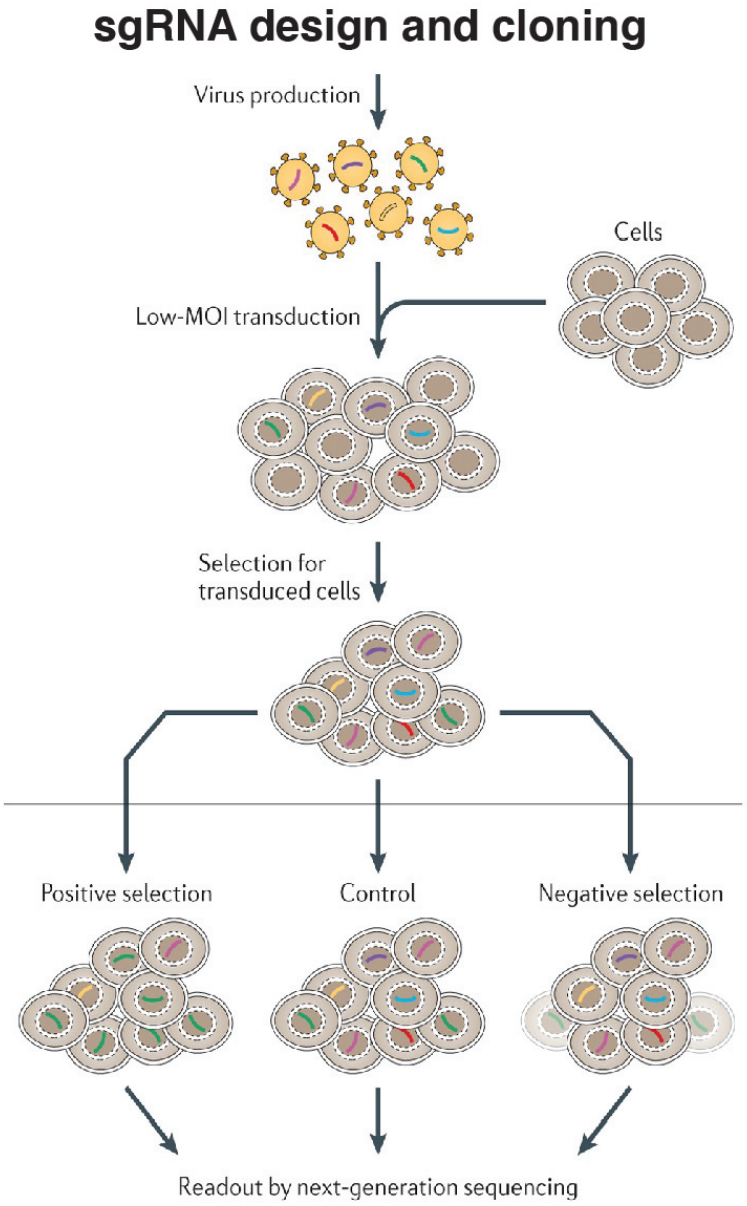

Figure 1.4. Workflow of pooled CRISPR screens

First, sgRNAs are designed and synthesised, followed by cloning into the lentiviral vector of choice. This generates a plasmid sgRNA library, which leads to a lentiviral sgRNA library with a similar representation of sgRNAs after virus packaging. The lentiviral sgRNA library is then applied to cells at a low multiplicity of infection (MOI). sgRNA-carrying cells are isolated before and after selection (usually by flow cytometry), and sgRNAs are read out using next-generation sequencing (NGS). Adapted from (Shalem et al., 2015) [150].

In pooled screens, sgRNA libraries are usually introduced to cells using viral transduction. Therefore, after designing the sgRNAs for each gene of interest and positive/negative control gene, which is usually based on published papers or computational tools [146–149], the sgRNAs are cloned into plasmid backbones. As previously described, on-target activity and off-target effects differ among different sgRNAs targeting the same gene. To mitigate this impact on screening results (e.g. generating false-positive hits), at least 4 sgRNAs are usually designed and included in the library for each gene [151, 154]. Subsequently, the plasmid library is packaged into lentivirus, retrovirus, or adeno-associated virus (AAV), while lentiviral sgRNA libraries are the most commonly used [154]. Instead of designing and constructing the sgRNA library from scratch, many genome-wide and targeted sgRNA libraries are also commercially available, for example, through Addgene (`http://www.addgene.org/pooled-library/#crispr`). The lentiviral sgRNA library is then delivered to the cells at the onset of a screen, and the lentiviral library needs to be applied at a low multiplicity of infection (MOI) to increase the chance of introducing only 1 sgRNA to most cells [150, 154]. Finally, next-generation sequencing (NGS) is used to read out and compare the number of sgRNAs in the cells before and after a selection pressure is applied. Screening hits are then identified using statistical analyses followed by functional validation of individual hits [150, 151, 154]. In a successful screen, the distribution of sgRNAs in the cells after the selection pressure is applied should be more skewed than at baseline. If control sgRNAs are included in the library, they can also be used to help interpret whether the screen is successful or not [154].

Sufficient coverage of the sgRNA library needs to be kept throughout all steps of the screen, including the transduction of cells using the lentiviral sgRNA library, upon transplantation and after engraftment of cells, isolation of sgRNA-carrying cells, and PCR amplification of the sgRNA sequences. A high representation of the sgRNA library is imperative to avoid picking up false-positive hits due to random depletion of sgRNAs [151, 154]. Genome-wide CRISPR screens have been successfully performed *in vivo* with approximately 400-fold coverage of the sgRNA library by different research groups [156, 157]. Besides, Strezoska and colleagues compared the reproducibility and ability to identify primary hits of short hairpin

RNA (shRNA)-based screens, which were performed with 100- and 500-fold coverage of the shRNA library. They demonstrated that even though higher coverage of the shRNA library significantly increased the reproducibility of the screen, the screen with 100-fold shRNA representation was able to detect a large proportion of the hits identified by the screen done with 500-fold shRNA representation [158]. Considering CRISPR-based screens allow lower coverage than RNAi-base screens due to their higher specificity, CRISPR screens with a 100-fold representation of sgRNAs may be able to identify the hits with more prominent changes. In contrast, a 500-fold or higher representation of sgRNAs can be necessary if small changes are expected [151, 158].

It can be technically challenging to maintain high coverage of the sgRNA library during *ex vivo* and *in vivo* CRISPR screens. First, as cells need to be transduced at a low MOI, a large number of cells need to be used during infection in order to reach a high sgRNA representation. This can be infeasible when a large-scale screen such as a genome-wide screen is performed, especially when primary cells are used. Second, in *in vivo* screens involving transplantations, the engraftment efficiency and time of engraftment of the cells can be difficult to predict [154]. To increase the sgRNA representation and engraftment efficiency in each recipient mouse, more cells need to be transplanted per mouse. This can cause the recipient mice to become moribund before the desired phenotypes can be observed or the knockouts happen completely [151, 154]. Of course, it is not necessary to maintain a 400- to 500-fold coverage of the sgRNA library in every recipient mouse, since the cells from different recipient mice could be computationally or physically pooled together as single biological replicates to reach the intended sgRNA representation [156, 159]. However, this means a number of mice need to be transplanted and sacrificed at the same time, especially when the sgRNA library is large. Therefore, focused CRISPR screens with limited sgRNA library sizes have been more commonly performed by many research groups *in vivo*[160–163], even though an increasing number of large-scale *in vivo* screens began to emerge [156, 164–166].

1.5.3 Applications of CRISPR-based screens in leukaemia research

Even though the genomic landscape of AML and AML-related gene overexpression has been better delineated during recent years, it remains to be determined whether a specific mutation or aberrantly overexpressed gene is functionally important and can be used as a druggable target to address the unmet clinical challenges [5, 14, 155]. CRISPR screens appear to be a powerful tool to address this question, which allows the identification of genes that are essential for AML cell survival (known as AML dependencies) or genes involved in differentiation blockade in leukaemia cells [162, 163].

The first CRISPR screens in AML cells were performed *in vitro*, using common leukaemia cell lines. Such *in vitro* CRISPR-based screens allowed comprehensive investigation into genes essential to AML cells by virtue of the easy accessibility of AML cell lines [167, 168]. Indeed, novel AML dependencies and potential therapeutic targets, such as *KAT2A*, have been successfully identified [167]. Wang and colleagues also harnessed the screen to study synthetic lethal interactions and uncovered genes specifically required by the oncogenic Ras [168]. Even though such *in vitro* screens could generate many interesting datasets and findings, they failed to recapitulate the interactions between AML cells and the surrounding BM microenvironmental factors, which, as described above, are actively involved in AML development. This means that the context-specific AML dependencies can hardly be identified using such *in vitro* screens, especially when many of the AML cell lines in use are cytokine-independent [167, 168].

More recently, several *in vivo* CRISPR screens focused on AML have been published. Due to the technical challenges of reaching high sgRNA representation *in vivo*, they often used a focused sgRNA library, which means only the essentiality of genes involved in specific biological processes (cell surface receptor signalling, etc.) was investigated [161, 162]. To my knowledge, the first *in vivo* genome-wide screen in murine models of AML was published very recently [164]. Even though such *in vivo* screens should in theory outperform the *in vitro* screens by allowing interactions between genetically perturbed AML cells and the surrounding BM niche, they still

have limitations (to be discussed more in detail in Chapter 3). An *in vivo* CRISPR screen which can more faithfully study the effect of genetic perturbations on AML cells in a strictly BM microenvironment-dependent context is warranted.

Chapter 2

Methods

2.1 Animals

All experimental procedures and mouse breeding and maintenance were in accordance with UK Home Office regulations. All experiments were approved by the Oxford University Clinical Medicine Ethical Review Committee. All mice used were from the C57Bl6/J background. For transplantations, recipients were of the CD45.1/2 allotype or the CD45.1 allotype; donor mice were of the CD45.2 allotype. To generate heterozygous CD45.1/2 recipients, WT C57BL/6 CD45.1 mice were bred with WT C57BL/6 CD45.2 mice (Envigo, Huntingdon). Young mice were 2-3 months old and old mice were 23-25 months old. When multiple experimental groups were analyzed, mice were allocated so that each group was evenly matched for the age range. No statistical methods were used to predetermine the experimental sample size.

2.2 Murine leukaemia model

Double transgenic mutant FL cells ($Cebpa^{\mathrm{K313KK/Lp30}}$) were previously generated by combining constitutive knock-in mice expressing a mutation in the C-terminal of C/EBPα ($Cebpa^{\mathrm{K313KK/+}}$) [13] with constitutive knock-in mice expressing a mutation in the N-terminal of C/EBPα ($Cebpa^{\mathrm{Lp30/+}}$) [55]. The FL cells (CD45.2 allotype) were competitively transplanted into lethally irradiated CD45.1/2 recipients along

with CD45.1 WT competitor BM cells to develop and propagate C/EBPα mutant AML in the recipient mice. When the recipient mice became moribund, *Cebpa* mutant LICs (CD45.2 allotype) were enriched by harvesting the BM cells, followed by c-Kit enrichment using the AutoMACS system (Miltenyi) [13, 56].

Unless otherwise indicated, all experiments involving a murine *Cebpa* mutant leukaemia model were initiated by intravenously transplanting 25,000 c-Kit-enriched LICs into sub-lethally irradiated (520 rads) CD45.1/2+ or CD45.1+ recipient mice. Using this procedure, AML reproducibly progressed over 4 weeks to a terminal phase of leukocytosis, anaemia, BM and spleen infiltration by leukemic cells and splenomegaly [56]. Control (mock transplanted) mice for LIC transplantations were sub-lethally irradiated and time-matched but not transplanted. In some experiments, 100,000-200,000 c-Kit-enriched lentivirus-transduced LICs were co-transplanted with 1 million CD45.1/2 or CD45.1 BM cells into CD45.1/2 or CD45.1 lethally irradiated (2 x 500rads) recipients. This led to the recipient mice becoming moribund in 3-4 weeks, similar to the kinetics seen in the sub-lethally irradiated mice transplanted with 25,000 c-Kit-enriched LICs.

2.3 Cell preparation

2.3.1 BM

BM cells were isolated by crushing bones from the tibias and femurs (also pelves and sternums in some experiments). BM cells were then filtered through a 50μM sterile filter, and a cell count was performed on a Pentra Haematology Analyzer (Pentra ES60, Horiba).

2.3.2 c-Kit enrichment of BM cells

c-Kit enrichment of BM cells was performed by positive selection using MACS magnetic separation (MACS Miltenyi Biotec, Bergisch Gladbach, Germany). BM cells were magnetically labelled using CD117 MicroBeads (MACS Miltenyi Biotec) at 2.5μL/100 million cells and incubated on ice for 20 minutes. LS columns (MACS Miltenyi Biotec) were placed on the MACS magnetic separator. BM cells were then

added to the LS column. The LS column was washed three times and then removed from the magnet. c-Kit positive cells were eluted by adding 5mL of phosphate-buffered saline (PBS)+5% foetal calf serum (FCS) and forced through the LS column using the plunger. LS column was then washed using 3mL of PBS+5%FCS. 95μL of elute was used to perform a cell count on a NucleoCounter®NC-3000™.

2.3.3 Spleen

Splenic cells were isolated by crushing the spleen using a 1mL syringe plunger through a 70μM sterile cell strainer. The strainer was then washed with 5mL of PBS+5%FCS. 80μL was used to perform a cell count on a Pentra Haematology Analyzer.

2.3.4 PB

Around 50-60μL of PB was taken from the tail vein of the mice. PB was collected in heparin-coated microvette tubes (Sarstedt). PB parameters were taken by making a 1:5 dilution using 16μL of blood and 64μL of PBS and were analysed using a Pentra Haematology Analyzer. White blood cells (WBCs) were isolated by resuspending the blood in 10mL of lysis buffer containing ammonium chloride (150mM) and potassium bicarbonate (10mM) for 7-8 minutes at room temperature. 1mL of 10X PBS was then added to stop the lysis reaction, and samples were centrifuged at 500g for 5 minutes at 4°C. Cells were washed with 1mL of PBS+5%FCS and centrifuged at 500g for 5 minutes at 4°C.

2.3.5 BM stromal cell isolation

Protocols used to isolate CBM and BL stromal cell populations were adapted from [169–171]. To isolate CBM stromal cells, dissected bones were crushed twice in a pestle and mortar containing 5mL PBS+5% FCS buffer, aspirated and passed through a 50μm mesh filter into a fresh 15mL Falcon. After crushing, remaining bone fragments were kept for isolating BL niche cells. The filtered BM solution was counted using a NucleoCounter®NC-3000™, centrifuged (5 minutes, 500g, 4°C) and resuspended in 5mL Dulbecco Modified Eagle Medium (DMEM) containing

10% FCS, 200U/mL DNaseI (Stemcell Technology) and 10mg of collagenase type IV (Worthington Biochemical) for 20 minutes at 37°C on a rotating plate. After incubation, cells were washed with PBS+5% FCS and centrifuged (5 minutes, 500g, 4°C). The pellet was resuspended in 5mL of ammonium chloride solution (StemCell Technology) for 10 minutes at room temperature. The lysis reaction was neutralized with 5mL of PBS+5% FCS, centrifuged (5 minutes, 500g, 4°C), resuspended in 5mL PBS+5% FCS and counted. For CD45+ cell depletion cells were resuspended in 90µL PBS+5% FCS and 10µL anti-CD45 microbeads (Miltenyi) per 10X10^6 cells. The solution was gently agitated every 2 minutes for a total of 20 minutes at 4°C and passed through a MACS LS column (Miltenyi). Effluent cells were collected by centrifugation (5 minutes, 500g, 4°C).

Remaining bone fragments left after CBM stromal cells isolation were minced into smaller fragments using scissors, resuspended in 5mL of PBS+5% FCS and centrifuged (5 minutes, 500g, 4°C) and resuspended in 5mL DMEM containing 10% FCS, 200U/mL DNaseI (Stemcell Technology) with 15mg of collagenase type I (Worthington Biochemical) for 45 minutes at 37°C on a rotating plate. Digested bone fragments were allowed to settle and the top layer was aspirated, passed through a 70µm mesh filter into a 50mL Falcon tube and stored at 4°C. Collagenase treatment was repeated on the remaining bone material to ensure maximum BL niche cell recovery. The resulting cell solution was washed with 10mL PBS+5% FCS by centrifugation (5 minutes, 500g, 4°C).

2.3.6 Freezing cells

10-20x10^6 BM cells were aliquoted and centrifuged at 500g for 5 minutes at 4°C. The supernatant was discarded, and cells were resuspended in 1mL of IMDM+20% FCS+10% dimethyl sulfoxide (DMSO). Cells were then transferred into cryovials, placed in a freezing container and placed in the -80°C freezer. Cryovials were moved into liquid nitrogen storage the following day.

2.3.7 Thawing frozen cells

Cryovials were thawed rapidly in a 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. 9mL of IMDM+20% FCS was added dropwise into the 15mL falcon containing the cells. The cells were then centrifuged at 300g for 6 minutes at room temperature. The supernatant was then removed and re-suspended in 10mL of IMDM+20% FCS and centrifuged at 300g for 6 minutes at room temperature. The supernatant was then removed, and cells were resuspended in 1mL of PBS+5% FCS. Cells were then counted using a haemocytometer or a NucleoCounter®NC-3000™.

2.4 Flow cytometry and Fluorescence-activated cell sorting (FACS)

2.4.1 Antibody staining and flow cytometry analysis

Details of mouse antibodies and viability dyes used for flow cytometry are shown in Table 2.1. All antibodies were used at pre-determined optimal concentrations, determined by titration. From BM, $10x10^6$ cells were used for the LIC staining panel. From the spleen, $5x10^6$ cells were used for the LIC staining panel. From PB, all the WBCs isolated were used for the WBC staining panel. In stains where anti-FcγRII/III antibody was not included, cells were incubated with purified anti-FcγRII/III antibody to prevent non-specific antibody binding. Gates were set using a combination of fluorescence minus one control and populations known to be negative/positive for the antigen. Leukaemic myeloblasts (LBs) in the BM/spleen were defined as CD45.2+Lineage-Sca-1-c-Kit-CD11b^{hi}. LICs in the BM/spleen were defined as CD45.2+Lineage-Sca-1-c-Kit+CD11b^{lo}. Leukaemia cells in the PB were defined as CD45.2+CD19-CD4-CD8a-NK1.1-CD11b+. CBM VECs were defined as Ter119-CD71-CD45-CD31+. BL VECs were defined as Ter119-CD71-CD45-CD31+. OBs were defined as Ter119-CD71-CD45-CD31-ALCAM+.

Antigen and Fluorophore	Supplier	Catalogue Number	Application
7-Aminoactinomycin D (7AAD)	Biotium	40037	Cell viability
Anti-mouse CD16/32	eBioscience	14-0161-81	Fc receptor block
Anti-mouse NK1.1 BV650	BioLegend	108735	PB staining
Anti-mouse CD19 PE-Cy7	eBioscience	25-0193-82	PB staining
Anti-mouse CD4 ACP-eF780	eBioscience	47-0042-82	PB staining
Anti-mouse CD8a APC-eF780	eBioscience	47-0081-82	PB staining
Anti-mouse CD11b APC	eBioscience	17-0112-83	PB staining
Anti-mouse Gr1 PO	ThermoFisher	RM3030	PB staining
Anti-mouse CD45.1 AF700	BioLegend	110724	Allotype discrimination during LIC and PB staining
Anti-mouse CD45.2 PB	Biolegend	109819	Allotype discrimination during LIC and PB staining
Anti-mouse Ter119 PE-Cy5	BioLegend	116210	Lineage staining during LIC and GMP staining; Stromal cell staining
Anti-mouse CD4 PE-Cy5	BioLegend	100514	Lineage staining during LIC and GMP staining
Anti-mouse CD8a PE-Cy5	BioLegend	100716	Lineage staining during LIC and GMP staining
Anti-mouse CD5 PE-Cy5	BioLegend	100610	Lineage staining during LIC and GMP staining
Anti-mouse Gr1 PE-Cy5	BioLegend	108410	Lineage staining during GMP staining
Anti-mouse CD11b PE-Cy5	BioLegend	101210	Lineage staining during GMP staining
Anti-mouse B220 PE-Cy5	BioLegend	103210	Lineage staining during GMP staining
Anti-mouse Sca-1 BV786	BD Horizon	563991	LIC staining; GMP staining
Anti-mouse c-Kit ACP-eF780	eBioscience	47-1171-82	LIC staining; GMP staining
Anti-mouse CD11b (Mac1) PE-Cy7	eBioscience	25-0112-81	LIC staining
Anti-mouse CD41 BV421	BioLegend	133911	GMP staining
Anti-mouse CD150 APC	BioLegend	115910	GMP staining
Anti-mouse CD16/32 PE-Cy7	eBioscience	25-0161-82	GMP staining
Anti-mouse CCR1 PE	R&D Systems	FAB5986P	CCR1 staining
Rat IgG2B PE Isotype Control	R&D Systems	IC013P	Isotype control for CCR1 staining
Anti-mouse CD127 PE	BioLegend	135009	IL7RA staining
Rat IgG2a, kappa PE Isotype Control	BioLegend	400507	Isotype control for IL7RA staining
Anti-mouse CD71 FITC	Biolegend	113806	Stromal cell staining
Anti-mouse CD45 AF700	eBioscience	56-0451-82	Stromal cell staining
Anti-mouse CD31 PE-Cy7	eBioscience	25-0311-82	Stromal cell staining
Anti-mouse Sca-1 BV605	BioLegend	108133	Stromal cell staining
Anti-mouse ALCAM BIOTIN	R&D Systems	BAF1172	Stromal cell staining
Anti-mouse Streptavidin BV421	BioLegend	405226	Secondary staining during stromal cell staining

Table 2.1. Mouse antibodies and viability dyes used for flow cytometry analysis and cell sorting

The antigen, fluorophore, supplier, catalogue number, and application of each antibody used in this study are shown.

2.4.2 FACS

To sort GFP+ (coexpressed with sgRNA) leukaemic cells for CRISPR screens or indel analyses, all BM cells from the secondarily transplanted leukemic mice were used. In some experiments, the leukaemic BM cells recovered from thawed vials were used for sorting. To sort leukaemia cells for quantitative Real-Time PCR

Analysis (qRT-PCR), all BM cells from the secondarily transplanted leukaemic mice were used. To sort healthy GMPs for qRT-PCR analysis, BM cells from the mock transplanted mice were first c-Kit-enriched, and all enriched cells were used for sorting. To sort CBM and BL stromal cells from young and aged mice for multiplex qRT-PCR, all collagenase-digested CBM and BL stromal cells from each mouse were used. In stains where anti-FcγRII/III antibody was not included, cells were incubated with purified anti-FcγRII/III antibody to prevent non-specific antibody binding. Gates were set using a combination of fluorescence minus one controls and populations known to be negative/positive for the antigen.

2.4.3 Instrument and analysis

Cell acquisition and analysis were performed on a BD LSRFortessa (BD Biosciences) using BD FACSDivaTM software (BD Biosciences). Cell sorting was performed on a BD FACSAriaIII cell sorter or a BD FACSAria Fusion (BD Biosciences). Analysis was performed using Flowjo software version 10.6.2 (Flowjo LLC, Oregon, USA).

2.5 RNAseq library preparation

Leukaemia (LICs and LBs) and BM stromal cell populations were isolated and FACS sorted from the *Cebpa* mutant leukaemia mice at 2, 3, and 4 weeks after LIC transplantation. GMPs and BM niche cells purified from time-matched, mock transplanted mice were used as control. cDNA synthesis and PCR amplification were performed based on the Smart-seq2 protocol with modifications [172]. Finally, different samples were pooled and sequenced on an Illumina HiSeq4000 (50bp single-end reads). More details of how the sequencing library was prepared can be found in Thomas, 2018 [56]. RNAseq library preparation for comparing young and aged BM niche cell populations has been described in Valletta et al., 2020 [130].

2.6 Candidate gene selection for the CRISPR screens

For the first screen, candidate genes were selected solely based on RNAseq data of murine *Cebpa* mutant LICs (vs normal GMPs) [56]. First, cytokine receptor genes

that were overexpressed (log2 fold change > 1, adjusted p-value < 0.05) in LICs at more than 1 time point of leukaemia development and had an reads per kilobase of transcript per million (RPKM) value > 5 were included (details to be described in Chapter 3.3). This included *Ccr1*, *Il17rb*, *Il22ra2*, *Il2rg*, and *Tgfbr1*. In addition, another two similarly overexpressed non-receptor genes, *Gdf3* and *Slc39a4*, were included. SgRNAs were designed for each of the 7 candidate genes (Table 2.2) based on the sgRNAs in the Mouse version 2 (v2) CRISPR library developed by Koike-Yusa and colleagues [167]. The Mouse v2 library consists of 90,230 guide sequences targeting a total of 18,424 mouse genes, which were designed using a customised pipeline as described in Tzelepis K, 2016 [167]. Since there were up to 5 guides for each gene in the Mouse v2 library, all of the guides targeting the 7 candidate genes were selected, resulting in a library of 35 unique sgRNAs. Sequences of the sgRNAs are provided in Table 2.3.

Candidate Genes	Time Points	Log2 fold change	Adjusted p-value	RPKM in LICs
Ccr1	Week 2	3.845114223	5.39E-31	48.55067762
Ccr1	Week 3	3.548441319	3.44E-28	35.95582818
Ccr1	Week 4	3.865946843	1.65E-29	50.5869161
Il22ra2	Week 2	3.814077452	7.48E-22	117.9842688
Il22ra2	Week 3	2.687286274	1.19E-10	45.3509545
Il22ra2	Week 4	3.59959718	3.06E-11	171.1062273
Il17rb	Week 2	1.994869264	2.18E-07	25.6596996
Il17rb	Week 3	1.852159784	4.85E-06	22.64327947
Il17rb	Week 4	1.711024542	2.68E-05	20.79815395
Tgfbr1	Week 2	1.277245151	0.002490904	9.784453796
Tgfbr1	Week 3	1.155922082	0.000940726	8.87740816
Tgfbr1	Week 4	1.11734502	0.008806368	8.606358517
Il2rg	Week 2	2.340238382	3.91E-20	833.2353602
Il2rg	Week 3	0.85398573	0.003793577	297.0620362
Il2rg	Week 4	1.902246627	1.24E-14	602.2652959
Gdf3	Week 2	3.180506361	2.64E-13	97.11988081
Gdf3	Week 3	3.559779957	1.22E-08	142.6206425
Gdf3	Week 4	3.129845131	1.80E-13	92.28504707
Slc39a4	Week 2	7.561869743	7.51E-55	32.95190144
Slc39a4	Week 3	9.175376474	5.26E-53	38.53863334
Slc39a4	Week 4	8.247016203	4.19E-76	77.68612372

Table 2.2. Candidate genes included in the first CRISPR screen

The RPKM value (LICs), log2 fold change (LICs vs GMPs), and adjusted p-value (LICs vs GMPs) of the genes included in the CRISPR screen at different time points are shown.

For the second screen, candidate genes were selected based on the RNAseq data of murine *Cebpa* mutant LICs and gene expression profiles of human AML samples

sgRNA ID	Target gene	Guide sequence	Application
Ccr1 a	*Ccr1*	GGTAGCAAATATCAGACGCA	The first screen, validation
Ccr1 b	*Ccr1*	GGTGAGTGAACTCCCACT	The first screen
Ccr1 c	*Ccr1*	GGCTTAAAAAAGTATAAGGC	The first screen
Ccr1 d	*Ccr1*	GGTATCTGTCAATCGTCAAC	The first screen
Ccr1 e	*Ccr1*	GGAACACTAGAGAATACA	The first screen
Slc39a4 a	*Slc39a4*	GGACGATTACCTGGCCACAC	The first screen
Slc39a4 b	*Slc39a4*	GGCCTTTTCACACGGCCCGT	The first screen
Slc39a4 c	*Slc39a4*	GGACGGCCTGTTAAATACGC	The first screen
Slc39a4 d	*Slc39a4*	GGTCCTGAATACGGATAG	The first screen
Slc39a4 e	*Slc39a4*	GGTCCCAATATCACGCTGCA	The first screen
Gdf3 a	*Gdf3*	GGCGATCCCACCCTAAGGT	The first screen
Gdf3 b	*Gdf3*	GGCTGTCTGTGGTTCAGGAC	The first screen
Gdf3 c	*Gdf3*	GGCAGGTTATAGTAGGACC	The first screen
Gdf3 d	*Gdf3*	GGAGGACTTATGCTACGTGA	The first screen
Gdf3 e	*Gdf3*	GGAAGACGCGAGGCACAGGT	The first screen
Il17rb a	*Il17rb*	GGTTCAACGGGACACGACAT	The first screen
Il17rb b	*Il17rb*	GGAGGGAGTTTACCTGGCG	The first screen
Il17rb c	*Il17rb*	GGACTTGCTTTGCCGCCGGA	The first screen
Il17rb d	*Il17rb*	GGCAAACTTACCGTCTGCC	The first screen
Il17rb e	*Il17rb*	GGTCTTGACGAGTTCCACT	The first screen
Il22ra2 a	*Il22ra2*	GGTATTTTGCACTGGCAAGC	The first screen
Il22ra2 b	*Il22ra2*	GGAAGATAAAGTTGACTGC	The first screen
Il22ra2 c	*Il22ra2*	GGTACGAGCTGTATTACGGG	The first screen
Il22ra2 d	*Il22ra2*	GGGTTATAGTCACGACCGG	The first screen
Il22ra2 e	*Il22ra2*	GGCAAAGATGCGTTAACTC	The first screen
Il2rg a	*Il2rg*	GGCAATACTTGGTGCAGTAC	The first screen
Il2rg b	*Il2rg*	GGTTCTGTACAGCTCGCCTC	The first screen
Il2rg c	*Il2rg*	GGTATAGTGCAGCGTGAGGT	The first screen
Il2rg d	*Il2rg*	GGAGTACATGAATTGCACT	The first screen
Il2rg e	*Il2rg*	GGAGCACTGAACCTCTGGAA	The first screen
Tgfbr1 a	*Tgfbr1*	GGAATAGAACTCCCAACTAC	The first screen
Tgfbr1 b	*Tgfbr1*	GGAAAGCAGTCAGCTGGCCT	The first screen
Tgfbr1 c	*Tgfbr1*	GGCTCTTCATTTGGCACACGG	The first screen
Tgfbr1 d	*Tgfbr1*	GGAAAGGGCGATCTAGTGA	The first screen
Tgfbr1 e	*Tgfbr1*	GGTTTGGCGAGGCAAATGGC	The first screen

Table 2.3. Sequences of sgRNAs used in the first screen

The target gene, guide sequence, and application of each sgRNA used in this first screen are shown.

published by Valk and colleagues simultaneously [173]. In brief, cytokine receptor genes that were overexpressed both in human AML cells with a log2 fold change > 1 (compared with CD34+ normal HSPCs) and in mouse LICs with a log2 fold change > 0 (compared with normal GMPs) at any time point of leukaemia development were included (details to be described in Chapter 3.4). This included *Ccr1*, *Csf2ra*, *Ifngr1*, *Il17ra*, *Il18rap*, *Il21r*, *Il3ra*, *Il5ra*, *Il7r*, and *Pdgfrb* (Table 2.4). 5 sgRNAs

were designed and included for each of them except for *Ccr1*, as previously described. 5 sgRNAs targeting *Rpa1* were included as the control for positive regulators of leukaemia cell proliferation [160]. 10 sgRNAs targeting *Rosa26*, *Renilla luciferase*, etc., were used as the negative control. This resulted in a library of 60 unique sgRNAs. Sequences of the sgRNAs are provided in Table 2.5.

Candidate Genes	Dataset	Time Points	Log2 fold change (LICs vs GMPs)	Adjusted p-value	Log2 fold change (Human AML vs Healthy)
Ccr1	Mouse *Cebpa* mutant LICs	Week 2	3.845114223	5.39E-31	NA
Ccr1	Mouse *Cebpa* mutant LICs	Week 3	3.548441319	3.44E-28	NA
Ccr1	Mouse *Cebpa* mutant LICs	Week 4	3.865946843	1.65E-29	NA
CCR1	Human AML	NA	NA	NA	1.624967167
Csf2ra	Mouse *Cebpa* mutant LICs	Week 2	0.802121397	0.035119799	NA
Csf2ra	Mouse *Cebpa* mutant LICs	Week 3	1.06134143	1.15E-05	NA
CSF2RA	Human AML	NA	NA	NA	1.711747336
Ifngr1	Mouse *Cebpa* mutant LICs	Week 2	0.921993766	1.88E-05	NA
Ifngr1	Mouse *Cebpa* mutant LICs	Week 4	0.742306846	7.16E-05	NA
IFNGR1	Human AML	NA	NA	NA	1.176602808
Il17ra	Mouse *Cebpa* mutant LICs	Week 3	0.941524973	0.013998571	NA
IL17RA	Human AML	NA	NA	NA	1.056634876
Il18rap	Mouse *Cebpa* mutant LICs	Week 2	2.091635166	0.000455019	NA
Il18rap	Mouse *Cebpa* mutant LICs	Week 3	1.992389872	0.000889393	NA
IL18RAP	Human AML	NA	NA	NA	1.977783346
Il21r	Mouse *Cebpa* mutant LICs	Week 3	1.868410429	2.76E-05	NA
IL21R	Human AML	NA	NA	NA	1.292215905
Il3ra	Mouse *Cebpa* mutant LICs	Week 3	2.004588431	5.34E-07	NA
IL3RA	Human AML	NA	NA	NA	1.220007244
Il5ra	Mouse *Cebpa* mutant LICs	Week 2	3.564081342	3.11E-08	NA
Il5ra	Mouse *Cebpa* mutant LICs	Week 3	5.254706064	3.64E-11	NA
Il5ra	Mouse *Cebpa* mutant LICs	Week 4	1.910088138	0.008237508	NA
IL5RA	Human AML	NA	NA	NA	1.74249467
Il7r	Mouse *Cebpa* mutant LICs	Week 3	2.488340855	0.007869541	NA
IL7R	Human AML	NA	NA	NA	3.040447529
Pdgfrb	Mouse *Cebpa* mutant LICs	Week 2	1.70754387	0.00744601	NA
Pdgfrb	Mouse *Cebpa* mutant LICs	Week 3	3.386429132	5.14E-13	NA
Pdgfrb	Mouse *Cebpa* mutant LICs	Week 4	1.354010669	0.044230762	NA
PDGFRB	Human AML	NA	NA	NA	1.088658686

Table 2.4. Candidate genes included in the second CRISPR screen

The log2 fold change (LICs vs GMPs and human AML cells vs healthy CD34+ cells), and adjusted p-value (LICs vs GMPs) of the genes included in the CRISPR screen are shown. For the values retrieved from the RNAseq data of mouse Cebpa *mutant LICs, only the values with an adjusted p-value < 0.05 are shown.*

Table 2.5. Sequences of sgRNAs used in the second screen

sgRNA ID	Target gene	Guide sequence	Application

Csf2ra a	*Csf2ra*	GGATGCATGGCCGACCCCG	The second screen
Csf2ra b	*Csf2ra*	GGCGAGTACCGGACGTCCGC	The second screen
Csf2ra c	*Csf2ra*	GGCCACGTACCGAGACTCGC	The second screen
Csf2ra d	*Csf2ra*	GGCCGTGACGTTCCCCGCGG	The second screen
Csf2ra e	*Csf2ra*	GGTCCGACCACACCGGACGC	The second screen
Ifngr1 a	*Ifngr1*	GGCTCGGAGAGATTACCCGA	The second screen
Ifngr1 b	*Ifngr1*	GGTAGAACATTCGTCGGTAC	The second screen
Ifngr1 c	*Ifngr1*	GGTATTCCCAGCATACGAC	The second screen
Ifngr1 d	*Ifngr1*	GGTATAGTCGAATGTGTAAC	The second screen
Ifngr1 e	*Ifngr1*	GGTGGAGCATAACCGGAGTG	The second screen
Il17ra a	*Il17ra*	GGCACCTCCTGCGCGCAGAC	The second screen
Il17ra b	*Il17ra*	GGATGAGGCCATACACCCAC	The second screen
Il17ra c	*Il17ra*	GGCGCCGATCAAGAGAAACA	The second screen
Il17ra d	*Il17ra*	GGGACTAATTCTCCGTGC	The second screen
Il17ra e	*Il17ra*	GGCGAGTGGACTTCACCCTG	The second screen
Il18rap a	*Il18rap*	GGAGTGGAAGTCCACTAG	The second screen
Il18rap b	*Il18rap*	GGTATCTCTTAGGAGCCCCC	The second screen
Il18rap c	*Il18rap*	GGCCAGTACTGCCAAGAAGT	The second screen
Il18rap d	*Il18rap*	GGTACACTGGACGTAGAGCT	The second screen
Il18rap e	*Il18rap*	GGACATACCACTTTATCAC	The second screen
Il21r a	*Il21r*	GGAATTGAGACAAGCGCATA	The second screen
Il21r b	*Il21r*	GGCAATGTGACGGACCAGTC	The second screen
Il21r c	*Il21r*	GGTATCATAGCGTCCTGAGA	The second screen
Il21r d	*Il21r*	GGCTCCAACTACGTGCTGAG	The second screen
Il21r e	*Il21r*	GGTGACCAAGCTGATCTCAG	The second screen
Il3ra a	*Il3ra*	GGCGCTGACATCACGACA	The second screen
Il3ra b	*Il3ra*	GGTTCGGACCCCGCTCCCAT	The second screen
Il3ra c	*Il3ra*	GGACGGGTGACGTGCACTAC	The second screen
Il3ra d	*Il3ra*	GGTTCTGGCGTGATGTGCGG	The second screen
Il3ra e	*Il3ra*	GGACACTCGCGGTTGTGGGC	The second screen
Il5ra a	*Il5ra*	GGTGCAACGAAGGGACACT	The second screen
Il5ra b	*Il5ra*	GGAAGTTAAATTCGTAACCG	The second screen
Il5ra c	*Il5ra*	GGCTTTGCAGCTAGCGTG	The second screen

Il5ra d	*Il5ra*	GGTGTGACCCCCCTTCATGA	The second screen
Il5ra e	*Il5ra*	GGAACTTGAGCTAATCCAG	The second screen
Il7r a	*Il7r*	GGTGGCCTACCGCCCAGCAA	The second screen, validation
Il7r b	*Il7r*	GGAACGACTTTCAGGTCAGA	The second screen
Il7r c	*Il7r*	GGTATGGCTATAAACACAAT	The second screen
Il7r d	*Il7r*	GGCAATATATGTGTGAAGCT	The second screen
Il7r e	*Il7r*	GGCTTCCATCCACTTCCAAC	The second screen
Pdgfrb a	*Pdgfrb*	GGTCGTAGGGTACGTGTAGG	The second screen
Pdgfrb b	*Pdgfrb*	GGACGTACCCCCGCATGA	The second screen
Pdgfrb c	*Pdgfrb*	GGTGGAGTGCCCTCCCGCAT	The second screen
Pdgfrb d	*Pdgfrb*	GGAATCCGCTGTCGTGGCCG	The second screen
Pdgfrb e	*Pdgfrb*	GGTGACGGTAAGGACCACTA	The second screen
Rpa1 a	*Rpa1*	GGCTCACTTGGACTGGTACG	The second screen
Rpa1 b	*Rpa1*	GGCAGCAGGCCAGTTCCGGC	The second screen
Rpa1 c	*Rpa1*	GGAACACCCTGAAAGACGGC	The second screen
Rpa1 d	*Rpa1*	GGCTGAACACCCTGGTCGA	The second screen
Rpa1 e	*Rpa1*	GGCCTGGAGTAATTCCCGGG	The second screen
Rosa26 a	*Rosa26*	GGCCCATCTTCTAGAAAGAC	The second screen
Rosa26 b	*Rosa26*	G GTCTTTCTAGAAGATGGGC	The second screen
Rosa26 c	*Rosa26*	GGTCCAGTCTTTCTAGAAGAT	The second screen, validation
Rosa26 d	*Rosa26*	GG AGTCTTTCTAGAAGATGGG	The second screen
Rosa26 e	*Rosa26*	GGCTCCAGTCTTTCTAGAAGA	The second screen
GAL4 a	*GAL4*	GGAGTGCACGATAGGGCAGTA	The second screen
GAL4 b	*GAL4*	GGCACACTGTGATAACAGAAA	The second screen
luciferase a	*luciferase*	GGTGAATTGCTCAACAGTAT	The second screen
luciferase b	*luciferase*	GGATTCTAAAACGGATTACC	The second screen
Jennifer Doudna (Scrambled)	NA	GGAGTAAAACCTCTACAAATG	The second screen

The target gene, guide sequence, and application of each sgRNA used in this second screen are shown.

2.7 Lentiviral sgRNA library generation

The pL-CRISPR.SFFV.GFP vector was obtained from Addgene (#57827) [174]. All the sgRNAs used in this study (35 for the first screen; 60 for the second screen) were cloned into the BsmBI-digested GFP vector. For the first screen, a GFP-expressing lentiviral sgRNA library was generated by transfecting HEK293T cells (ATCC, Virginia, USA) with the viral constructs encoding the 35 sgRNAs (as in Table 2.3), plasmid pMD2.G and plasmid psPAX2 using PEI Pro (Polyplus Transfection, New York, USA). As a control for transduction efficiency, sgRNAs targeting *Myod1* (RPKM = 0 in LICs and LBs) were cloned into pL-CRISPR.SFFV.tRFP (Addgene #57826) and the resulting virus co-transduced with the GFP-expressing virus. For the second screen, a GFP-expressing lentiviral sgRNA library was generated by transfecting HEK293T cells with the viral constructs encoding the 60 sgRNAs (as in Table 2.5), plasmid pMD2.G and plasmid psPAX2. Harvests were collected 48 and 72 hours post-transfection, combined and concentrated by ultracentrifugation (2 hours at 98,000g, 4°C).

2.8 *In vivo* CRISPR screens

c-Kit+ BM cells were freshly enriched from primarily transplanted leukaemic mice as previously described and placed into culture (IMDM, with 0.05% BSA, 100U/m penicillin, 100µg/mL streptomycin, 150µM 2-mercaptoethanol, mSCF at 50ng/mL, mIL-3 at 10ng/mL, and mIL-6 at 10ng/mL). The c-Kit-enriched LICs were transduced with concentrated lentiviral supernatant in the presence of 4µg/mL Polybrene. The cells were incubated at 37°C, 5% CO2, for 8 hours. Spinfection with the lentiviral supernatant was performed at 500g for 1 hour at room temperature in some of the experiments.

For the first screen, c-Kit-enriched LICs were transduced with the GFP lentiviral sgRNA library at an MOI of 20-30, defined as the titre on HEK293T cells divided by the number of c-Kit-enriched leukaemia cells. This generated an infection rate of ca. 20% (tested 48h after transduction started by flow cytometry). After the 8-hour transduction, 100,000-150,000 cells (the equivalent of at least 400 transduced cells

per sgRNA) were co-transplanted with 1 million WT CD45.1/2 BM cells into each CD45.1/2 lethally irradiated recipient. 4-5 weeks post-transplantation, mice were culled when they became moribund. SgRNA-expressing LICs (CD45.2+GFP+c-Kit+CD11b^{lo}) and LBs (CD45.2+GFP+c-Kit-CD11b^{hi}) were individually sorted from the BM for genomic DNA (gDNA) extraction. For the second screen, c-Kit-enriched LICs were transduced with the GFP lentiviral sgRNA libraries at an MOI of 60-70. This generated an infection rate of ca. 35% (tested 48h after transduction started by flow cytometry). After the 8-hour transduction, 200,000 cells (the equivalent of at least 1,000 transduced cells per sgRNA) were co-transplanted with 1 million WT CD45.1 BM cells into each CD45.1 lethally irradiated recipient. 3-4 weeks post-transplantation, mice were culled when they became moribund. SgRNA-expressing LICs and LBs transduced were individually sorted from the BM for gDNA extraction.

WT NIH3T3 cells (ATCC, Virginia, USA) were transduced using the same lentiviral libraries at an MOI of 1 overnight. This generated an infection rate of ca. 20-25% (tested 48-72h post-transduction). SgRNA-expressing NIH3T3 cells (GFP+) were sorted at 3 and 14 days post-transduction for gDNA extraction, which was used as input. At least 500-fold coverage of the sgRNA library in NIH3T3 cells was maintained throughout the process.

A QIAamp DNA micro kit (Qiagen, Maryland, USA) was used for gDNA extraction from the FACS sorted cell pellets, followed by quantification using a Qubit dsDNA Assay Kit (Invitrogen). For each individually sorted sample, an amount of gDNA equivalent to ˜500 cells per sgRNA (For the first screen, 115ng per LICs or LBs sample; for the second screen, 180ng per LICs or LBs sample) was used for PCR amplification. This helped preserve full sgRNA library complexity. sgRNA sequences were amplified with primers designed around the sgRNA region with Nextera XT compatible overhangs, which allowed for indexing. A stagger sequence was included in the forward primers between Nextera XT compatible overhangs and the forward sequence [175]. The primers for sgRNA amplification are listed in Table 2.6. PCR products were purified using Ampure XP beads, evaluated using a DNA 1000 kit on an Agilent 2100 Bioanalyzer, and quantified using a Qubit dsDNA As-

say Kit. A second PCR was performed to add Nextera XT Index Kit v2 (Illumina) sequencing adapters to the amplicons. PCR products were purified using AmpureXP beads, evaluated using a DNA 1000 kit on an Agilent 2100 Bioanalyzer, and quantified using a Qubit dsDNA Assay Kit. Finally, the samples were pooled and sequenced on a MiSeq (Illumina, 150 bp paired-end reads) to investigate the change in the sgRNA representation between the leukaemic and input NIH3T3 samples.

Gene/Genomic region	Assay	Sequence
sgRNA amplification Forward	PCR	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGtatatcttgtggaaaggacg
sgRNA amplification Forward	PCR	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAtatatcttgtggaaaggacg
sgRNA amplification Forward	PCR	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGAtatatcttgtggaaaggacg
sgRNA amplification Forward	PCR	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCGAtatatcttgtggaaaggacg
sgRNA amplification Forward	PCR	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTCGAtatatcttgtggaaaggacg
sgRNA amplification Forward	PCR	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATCGAtatatcttgtggaaaggacg
sgRNA amplification Forward	PCR	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGATCGAtatatcttgtggaaaggacg
sgRNA amplification Reverse	PCR	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGgtggatgaatactgccatttgtc

Table 2.6. Primers for sgRNA Amplification

Sequences of oligonucleotides used for sgRNA library preparation (PCR) are shown.

2.9 qRT-PCR

Total RNA was extracted using an RNeasy Plus Micro Kit (QIAGEN, Valencia, CA) from FACS sorted LICs, LBs, and normal GMPs. LICs and LBs were sorted 4 weeks after LIC transplantation; normal GMPs were sorted from the mock transplanted mice 4 weeks after irradiation. cDNA was synthesized using the Superscript III First-Strand Synthesis System (Invitrogen, Grand Island, NY). qRT-PCR reactions were performed using Taqman assays (Thermo Fisher Scientific) (Table 2.7) and TaqMan universal PCR master mix according to the manufacturer's instructions (Applied Biosystems, Hammonton, NJ). A TaqMan probe for mouse *B2m* (Mm00437762_m1) was used as an endogenous control. Ct values were generated using a QuantStudio 7 Flex Real-Time PCR System and then normalised to *B2m*. $2^{-\Delta Ct}$ values were then used for analysis.

Gene	Assay	Assay ID
B2m	Taqman probe	Mm00437762_m1
Gapdh	Taqman probe	Mm99999915_g1
Ccr1	Taqman probe	Mm01216147_m1
Il2rg	Taqman probe	Mm01229219_g1
Il17rb	Taqman probe	Mm00444709_m1
Tgfbr1	Taqman probe	Mm00436964_m1
Il22ra2	Taqman probe	Mm01192969_m1
Il7r	Taqman probe	Mm00434295_m1
Il5ra	Taqman probe	Mm00434284_m1
Il18rap	Taqman probe	Mm00516053_m1
Ifngr1	Taqman probe	Mm01192836_m1
Il21r	Taqman probe	Mm00600317_m1
Csf2ra	Taqman probe	Mm00438331_g1
Pdgfrb	Taqman probe	Mm00435553_m1
Il17ra	Taqman probe	Mm00434214_m1
Il3ra	Taqman probe	Mm00434273_m1
Ccl3	Taqman probe	Mm00441259_g1
Ccl5	Taqman probe	Mm01302427_m1
Ccl6	Taqman probe	Mm01302419_m1
Ccl9	Taqman probe	Mm00441260_m1

Table 2.7. Taqman assays

The identity of TaqMan probes used for qRT-PCR in this study is shown.

2.10 Multiplex qRT-PCR analysis

Multiplex qRT-PCR was performed on CBM VECs, OBs, and BL VECs isolated from young and aged mice (100 cells per replicate). CellDirect One-Step qRT-PCR kit (Thermo Fisher Scientific) was used according to the manufacturer's protocol for the preparation and amplification of cDNA. 10μL of preamplification buffer consisted of 2.5μL TaqMan assay mix containing all assays at 0.2X dilution, 5μL CellsDirect 2X Reaction mix (Invitrogen), 1.2μL CellsDirect RT/Taq mix (Invitrogen), 1.1μL TE buffer and 0.2μL SUPERaseIn RNase Inhibitor (Ambion). Targeted cDNA pre-amplification was performed during 20 cycles and preamplified product was diluted 1:5 in TE buffer before further processing. The BioMark 192.24 Dynamic Array platform (Fluidigm, California, USA) and Taqman Gene Expression Assays (Thermo Fisher Scientific) were used to perform the multiplex qRT-PCR according to the manufacturer's instructions. Taqman Gene Expression Assays used in this study were listed in Table 2.7.

2.11 ELISA

Femurs and tibias from mice were flushed with 200μL of PBS; the flushed BM suspension was centrifuged (5 minutes, 500g, 4°C), and the BM supernatant was collected, aliquoted and kept at -80°C. CCL3, CCL6, and CCL5 BM supernatant levels were determined using the Mouse CCL3 Quantikine ELISA Kit (R&D systems), Mouse CCL6/C10 DuoSet ELISA (R&D systems), and Mouse CCL5/RANTES Quantikine ELISA Kit (R&D systems) according to the manufacturer's instructions, respectively.

2.12 Quantification of CRISPR editing

The QIAamp DNA Micro Kit was used for gDNA extraction from FACS sorted GFP+ NIH3T3 cells, GFP+ LICs, GFP+ LBs, and non-edited cells (without transduction). Genomic regions flanking the sgRNA binding regions were PCR amplified. PCR products were run on a 2% agarose gel, followed by gel extraction using a Monarch DNA Gel Extraction Kit (NEB). Purified PCR products were then subjected to Sanger sequencing. Sanger sequencing files were analysed using Synthego's ICE (short for Inference of CRISPR Edits) tool (`https://ice.synthego.com/`) to calculate the CRISPR editing efficiency. The primers for amplifying and sequencing each sgRNA binding region are listed in Table 2.8.

Genomic region	Assay	Sequence
Ccr1 sgRNA a binding site Forward	PCR	CTAGCCATCTTAGCTTC
Ccr1 sgRNA a binding site Reverse	PCR	GTGGATGGAGATATAGAAC
Ccr1 sgRNA c binding site Forward	PCR	CTAGTGTTCATCATTGGAGTG
Ccr1 sgRNA c binding site Reverse	PCR	GCTTACTCTGCTCACACTG
Il7r sgRNA a binding site Forward	PCR	CTGCTGACCTCCTGTTGTC
Il7r sgRNA a binding site Reverse	PCR	CAAGGAGGCTATCTAGGAAGAT
Il7r sgRNA d binding site Forward	PCR	GGTAGATGGCCTGCTTTAG
Il7r sgRNA d binding site Reverse	PCR	CCTCCTCACCATTAACTTGC

Table 2.8. Primers for quantification of CRISPR editing

Sequences of oligonucleotides used for indel analysis of Ccr1 and Il7r alleles after sgRNA-mediated gene editing are shown.

2.13 *In vivo* validation of *Ccr1* and *Il7r* as dependencies of *Cebpa* mutant LICs

c-Kit+ BM cells were freshly enriched from primarily transplanted leukaemic mice and transduced with GFP lentiviral vectors encoding the sgRNAs targeting *Ccr1*, *Il7r*, or *Rosa26*. sgRNA sequences used in this validation experiment can be found in Tables 2.3 and 2.5. After 8 hours of transduction, the cells were washed by centrifuge. An aliquot of cells from each infection group was re-plated and cultured for up to 3 days. The percentage of GFP+ cells in culture was assessed by flow cytometry at 48 and 72 hours post-transduction. The rest of the cells (200,000 cells per secondary recipient mouse) were co-transplanted with 1 million WT CD45.1 BM cells into lethally irradiated CD45.1 recipient mice directly after the 8-hour transduction. The mice were culled 3 weeks post-transplantation, and the engrafted cells were analysed through flow cytometry.

2.14 Hydrodynamic injection

Krebs-Ringer modified buffer (pH 7.4) with 0.05% BSA was used as the final working buffer for all hydrodynamic injections as previously described [130]. The injection volume was ˜8-12% of the mouse's body weight [130]. After being generally anaesthetised, mice were tail vein-injected in the working buffer with vectors expressing the genes of interest or an empty pCMV-entry vector as control. All the vectors used were purchased from OriGene (Table 2.9). The injected mice were kept overnight or 72 hours before being transplanted with leukaemic cells.

Product Name	Catalogue Number	Supplier	Application	
Ccl3 (NM_011337) Mouse Tagged ORF Clone	MR225554	Origene	Overexpressing CCL3	
Ccl6 (NM_009139) Mouse Tagged ORF Clone	MR200502	Origene	Overexpressing CCL6	
Ccl9 (NM_011338) Mouse Tagged ORF Clone	MR223660	Origene	Overexpressing CCL9	
Il6 (NM_031168) Mouse Tagged ORF Clone	MR227281	Origene	Overexpressing IL-6	
pCMV6-Entry Mammalian Expression Vector	PS100001	Origene	Empty control	

Table 2.9. Plasmids for hydrodynamic injections

The product name, catalogue number, supplier, and application of each plasmid used for hydrodynamic injections in this study are shown.

2.15 Proliferation assay

c-Kit enriched *Cebpa* mutant LICs were maintained in 96-well flat-bottom plates in 0.1mL IMDM supplemented with 0.5% BSA, 100U/m penicillin, 100µg/mL streptomycin, 150µM 2-mercaptoethanol, mSCF (50ng/mL), mIL-3 (10ng/mL), and mIL-6 (10ng/mL) at 37°C in 5% CO2. 25,000 viable cells were plated in each well. CCL3 (Peprotech), CCL6 (Peprotech), or CCL9 (Biolegend) were added to the culture at a final concentration of 50ng/mL. The number of living cells was measured using the CellTiter 96®AQueous One Solution Cell Proliferation Assay (MTS) (Promega, G3582) after 24 and 48 hours as instructed by the manufacturer's instructions.

2.16 Drug treatment of leukaemic mice

For IL-7R-blocking experiments, young CD45.1 recipient mice were intraperitoneally injected with 500µg monoclonal anti-IL7R antibody (A7R34; BioXCell) or 500µg isotype control antibody (2A3; BioXCell) 2 times a week for a total of 3 weeks. The treatment was started the next day following sub-lethal irradiation and transplantation of 25,000 purified LICs. All mice were culled 21 days after LIC transplantation for flow cytometry analysis.

For IL-6-blocking experiments, aged CD45.1 recipient mice were intraperitoneally injected with 100µg of mouse IL-6 neutralising antibody (10F9; InvivoGen) or Rat IgG2a isotype control (C1-2A7; InvivoGen) dissolved in 1mL of sterile saline solution at 1-week intervals [130]. A total of 3 treatments were given to each mouse. 1 week after the last treatment, the mice were transplanted with 25,000 purified LICs without irradiation. All mice were culled 28 days after LIC transplantation for flow cytometry analysis.

BL5923 (kindly provided by Novartis) was dissolved in 0.5% hydroxyethyl cellulose [176, 177] and kept at 4°C for no more than a week. Aged CD45.1 recipient mice were transplanted with 25,000 purified LICs and treated twice daily by oral gavage at a 50mg/kg dose for 10 consecutive days starting at 10 days after LIC transplantation. Control mice were treated with the relevant solvent in the same manner. All mice were culled 21 days after LIC transplantation for flow cytometry analysis.

2.17 Quantification and statistical analysis

2.17.1 Bulk RNAseq analysis

FASTQ files were inspected using FastQC (v0.10.1) for quality control analysis followed by Nextera adapter sequence removal using TrimGalore! (v0.4.1; `http://www.bioinformatics.babraham.ac.uk/projects/trim_galore`) with a stringency setting of 3. Then, reads were aligned using STAR (v2.4.2a) [178] against the mm9 mouse reference genome. Gene expression levels were quantified as read counts using featureCounts (v1.4.5-p1) [179] and the UCSC mm9 annotation file. All output files were quality assessed using MultiQC (v0.7) [180]. RNAseq data supporting this study's findings have been deposited in the Gene Expression Omnibus under accession no. GSE210357 (gene expression profiles of *Cebpa* mutant LICs and LBs), GSE210363 (gene expression profiles of AML-exposed BM stromal cells), and GSE130899 (gene expression profiles of aged BM stromal cells). The read counts were imported into RStudio (R project, Revolution Analytics), and the DESeq2 (v1.14.1) [181] package was used for normalisation and differential gene expression analysis. Differentially expressed genes (DEGs) between samples from the leukaemia and mock transplanted mice were defined as the ones with an adjusted p-value < 0.05. RPKM values were calculated using the EdgeR::rpkm function (v2.16.5) [182]. To generate a heatmap, ln(RPKM+1) expression values were used as input; euclidean distance and wards clustering methods were used to cluster samples according to their expression of known gene sets. DEGs were ranked by log2 fold change and used as the input for pre-ranked gene-set enrichment analysis (GSEA). When GSEA was performed without pre-ranking, the transcripts per million (TPM) value for each gene was used as input. GSEA was performed using GSEA software (v4.0.3) [183]. Hallmark gene sets and KEGG gene sets were downloaded from the Molecular Signatures Database (`http://www.gsea-msigdb.org/gsea/msigdb/collections.jsp`). Ranked DEGs were used as the input for Metacore process network analysis (MPNA), which was performed using Metacore v6.35 (Thomson Reuters, `www.portal.genego.com`) to determine enriched process networks. All Metacore output displayed in the study has a p-value < 0.05.

For testing the probability of achieving the obtained overlap of DEGs (from the

pool of all expressed genes with RPKM values > 1: 15399 genes for the CBM VECs, 14778 genes for the perivascular cells) and perturbed networks (from the pool of all analysed networks: 159 networks in total), p-values were calculated using the hypergeometric test for 2-way overlap and random draw simulation (10,000,000 iterations) for 3-way overlap using Rstudio. Venn diagrams were generated using the VennDiagram R package.

2.17.2 Pooled CRISPR screen data analysis

Raw FASTQ sequencing data files were first demultiplexed using Illumina indices. For each sample, sequencing reads were searched for sequencing matching the designed sgRNAs and the numbers of reads for each sgRNA were counted using a custom-written script in R statistical programming environment. For mapping, a maximum of one mismatch was allowed in the 19-20bp spacer sequence. Finally, the total reads for all sgRNAs across different samples were normalised. After normalisation, for all sgRNAs, a count of 1 was added as a pseudo count. The change in sgRNA representation during leukaemia development was determined by comparing read counts from the end GFP+ leukaemic samples to input GFP+ NIH3T3 samples sorted 3 days post-transduction. The MAGeCK robust ranking algorithm (RRA) was used to calculate the enrichment and depletion of sgRNAs and genes [184, 185]. Figures were generated in RStudio. To identify enrichment and depletion of sgRNAs that were not specific to leukaemia development, read counts from GFP+ NIH3T3 samples sorted 14 days post-transduction were compared to input GFP+ NIH3T3 samples sorted 3 days post-transduction.

For the first screen, the normalised data were also analysed by calculating CRISPR gene scores (CS) [168]. First, the input abundance of the 35 experimental sgRNAs was calculated by averaging the sgRNA representation of the 3 GFP+ NIH3T3 replicates sorted 3 days post-transduction. Then, the log2 fold change in the abundance of each experimental sgRNA was calculated between each of the 3 sorted GFP+ leukaemic replicates and the input and averaged across the leukaemic replicates. Finally, the CS for the candidate genes were defined as the average log2 fold change in the abundance of all sgRNAs targeting a given gene.

2.17.3 The Cancer Genome Atlas (TCGA) data analysis

mRNAseq (RNA-Seq by Expectation-Maximisation (RSEM) values) and clinical data for the TCGA-LAML cohort were downloaded from cBioportal (`https://www.cbioportal.org`); raw RNAseq read counts were retrieved from the National Cancer Institute Genomic Data Commons (GDC) (`https://gdc.cancer.gov/about-data/publications/laml_2012`). *CEBPA* mutation status and the cytogenetic subtype were extracted from the downloaded clinical data. The raw RNAseq read counts of patients with *CEBPA* mutations were imported into RStudio and used as input for differential gene expression using DESeq2. Principal component analysis (PCA) was carried out on the output of DESeq2 (DEGs with an adjusted p-value $<$ 0.05) to assess sample clustering using the prcomp function. GSEA was performed without pre-ranking, using TPM values for each gene as input. When comparing gene expression levels between TCGA-LAML samples and Genotype-Tissue Expression Project (GTEx) normal BM samples, TPM values free of computational batch effects were downloaded from the UCSC Toil RNAseq Recompute Compendium (`https://xenabrowser.net/hub/`) [186, 187].

For survival analysis, overall survival data from the fields, "OS_MONTHS" and "OS_STATUS", were extracted from the downloaded TCGA-LAML clinical data. The AML patients were segregated based on the *CCR1* and *IL7R* mRNA expression levels (RSEM values downloaded from cBioportal). Kaplan-Meier survival curves and log-rank test were performed using Rstudio.

2.17.4 Beat AML data analysis

Raw RNAseq read counts, log2 normalised RPKM values of *CCR1*, *CCL3*, *CCL23*, and *IL7R*, and detailed clinical annotations for the Beat AML cohort were downloaded from the Beat AML data viewer (Vizome, `http://www.vizome.org/aml/`) [188]. Differential gene expression, PCA and GSEA analyses were performed on *de novo* AML patients with *CEBPA* mutations as described above. For survival analysis, we accessed the overall survival data of all the *de novo* AML patients from the fields, "OverallSurvival" and "VitalStatus" from the downloaded clinical data.

2.17.5 Multiplex qRT-PCR analysis

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 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). Ct values were then normalised to the geometric mean of the 2 housekeeping genes, *B2m* and *Gapdh*. $2^{-\Delta Ct}$ was then used for analysis. Differential gene expression statistical significance between young and aged BM stromal cells was performed using an unpaired Student's t-test.

2.17.6 Statistical analysis

Unless otherwise indicated, the statistical significance of differences between groups was calculated using an unpaired two-tailed Student's t-test. A Mann-Whitney test was performed if data did not conform to a normal distribution. Prism version 8 (`http://www.graphpad.com`) was used for statistical analysis and result visualisation. On graphs, error bars represent standard deviation. For all graphs, a p-value of less than 0.05 was considered statistically significant, * indicates p-value < 0.05, **p-value < 0.01, ***p-value < 0.001, and ****p-value < 0.001.

Chapter 3

Identification of cytokine receptors potentially essential for LIC survival using *in vivo* CRISPR-Cas9 dropout screens

3.1 Introduction and aims

AML is the most common myeloid malignancy in adults [2]. The prognosis for elderly AML patients is extremely poor: the estimated median overall survival of elderly AML patients (>= 65 years old) is around 4 months [4]. Even for younger patients (<= 60 years old), AML is cured in only 35-40% of them [1]. Therefore, there is an urgent need to discover novel therapeutic targets and develop better-tolerated treatments, especially for elderly AML patients who are too frail to receive intensive chemotherapy. Ideally, the novel treatment should be able to eradicate LSCs, the elimination of which is supposed to be essential for the cure of AML according to the CSC model [30].

There is considerable evidence suggesting that the initiation and progression of myeloid malignancies are influenced by the interplay between malignant haematopoietic cells and the surrounding BM environmental factors. First, several studies have shown that genetic modification of BM niche cells can induce myeloproliferative disorders and even transplantable AML [78, 80, 81, 189]. This is associated with

the genetically altered BM stromal cells, which induce increased production of pro-inflammatory cytokines and chemokines such as TNF-α, G-CSF and CCL3 [80, 81]. In addition, AML cells can produce a similar range of pro-inflammatory cytokines and chemokines, including IL-1β, TNF-α, CCL3 and CCL23, which may support AML cell growth [85, 94, 97, 101]. Importantly, primary human AML cells, especially the engraftable LSCs, are extremely difficult to be maintained and expanded ex vivo over a long period, even if the culture is supplemented with a combination of standard cytokines including SCF, TPO, IL-3, IL-6, etc [59, 190, 191]. Taken together, AML cells may need a specific microenvironment with specific paracrine and autocrine signals, which are usually pro-inflammatory, to maintain their leukaemogenic potential. One potential therapeutic target for AML could be the cytokine receptors that are druggable using monoclonal antibodies or small molecule inhibitors. In particular, the ones preferentially expressed on AML cells, especially the LSCs, can be promising. Thus, the question I will address in this chapter is which cytokine receptor(s) could be essential for the survival of AML cells, especially the leukaemia-propagating LSCs that have the capability of self-renewal and are refractory to standard therapies [17].

CRISPR screens have been widely used to identify genes essential for tumour progression in various cancers, including AML [155, 192]. However, many CRISPR screens focused on AML were performed using AML cell lines [167, 168]. Such *in vitro* screens can hardly reflect the native interaction between AML cells and the relevant BM microenvironment where they reside, let alone identify essential cytokine receptor genes. More recently, a couple of *in vivo* screens in murine models of AML have emerged, potentially facilitating the study of interactions between AML cells and the recipient microenvironmental factors [161, 162, 164]. However, they still have obvious limitations.

The study of Lin and colleagues was not focused on cell surface receptors at all and involved screening for PDX samples that remained transplantable after a period of *ex vivo* culture, which might not be environmental factor-dependent anymore [162]. Besides, transplanting human AML cells into a murine BM microenvironment limited the possibility of comprehensively studying the interaction between the AML

cell surface receptors and environmental factors since not all human cytokine receptors can interact with the associated murine cytokines as human CCR1 and murine CCL3 do [193].

The study of Ramakrishnan and colleagues used a murine MLL-AF9 AML model, where leukaemia was initiated by viral transduction of the GMPs with the oncogenic *MLL-AF9* fusion gene [161]. The *MLL-AF9* leukaemia cells, unlike most primary human samples, remained highly transplantable after up to 4 days of *ex vivo* culture [59, 161]. This indicates that the murine *MLL-AF9* LSCs are not as dependent on a specific BM microenvironment as primary human LSCs, and the *MLL-AF9*-driven model may not be optimal in terms of recapitulating the complex BM microenvironmental milieu in human AML.

Most recently, Mercier and colleagues published the first *in vivo* genome-wide CRISPR screens in murine models of AML, which provided a valuable pipeline for conducting large-scale CRISPR screens *in vivo* [164]. They successfully identified a number of hits which were specific to the *in vivo* setting and had never been identified through *in vitro* screens. This corroborated the unique role of *in vivo* CRISPR screens in discovering AML dependencies. However, similar to the screen done by Ramakrishnan and colleagues, the murine AML cells (driven by the *MLL-AF9* fusion or *HOXA9/MEIS1*) used for transplantation remained engraftable after as long as 7 days of *ex vivo* culture. This again questions whether these AML cells are BM-microenvironment-dependent and whether such screens are able to discover genes essential for conveying signalling coming from the surrounding environment. Collectively, a kinetically tractable and strictly environmental factor-dependent AML mouse model is needed to more accurately investigate the interplay between AML cells and the BM microenvironment. This model would facilitate the identification of cytokine receptor genes essential for the LSCs during *in vivo* CRISPR screens.

The Nerlov laboratory has developed an AML mouse model by transplanting 25,000 LICs that harboured biallelic *Cebpa* mutations into sub-lethally irradiated recipient mice [13, 56]. In this model, the transplantable LIC was identified as a transformed CD11b^{lo}c-Kit+ myeloid progenitor [13]. The transplanted mice reproducibly progressed to a terminal phase of AML in 4 weeks. More specifically, the 2-, 3- and 4-

week time points represented 3 distinct stages of leukaemia progression as described in Thomas A, 2018 [56]. 2 weeks after LIC transplantation represented an early expansion phase with no overt leukaemia but with leukaemia cells readily detectable in haematopoietic organs. 3 weeks after LIC transplantation represented a late expansion phase with leukocytosis and leukaemia cells highly abundant in haematopoietic organs, which are still expanding. 4 weeks after LIC transplantation represented a terminal phase where the normal HSPCs were almost depleted and both leukocytosis and anaemia were evident [56]. Such kinetic tractability allows studying the genes essential for leukaemogenesis in a reproducible manner and identifying genes critical to different stages of AML development. More importantly, the LICs purified from the leukaemia mice completely lost their leukaemogenic potential after 48 hours of *ex vivo* culture when mSCF, mIL-3, and mIL-6 were supplemented, as described in Thomas A, 2018 [56]. This demonstrated the *Cebpa* mutant LICs' need for a proper *in vivo* BM microenvironment to maintain the leukaemogenic potential, just like the primary human LSCs. Therefore, our AML model (as opposed to the more commonly used *MLL-AF9* model) can be reckoned to be stringently dependent on the BM microenvironment in order to maintain its malignant phenotype. Hence, in this chapter, the *Cebpa* mutant AML mouse model was used to perform *in vivo* CRISPR dropout screens to identify which cytokine receptor(s) could be essential for LIC maintenance.

3.2 Cytokine receptors can potentially be used as a druggable target to eliminate the *Cebpa* mutant LICs

LICs and LBs were previously sorted at 2, 3, and 4 weeks after LIC transplantation and profiled using bulk RNAseq by the Nerlov laboratory, as outlined in Thomas A, 2018 (Figure 3.1A) [56]. In the meantime, cells from mock transplanted mice were purified and profiled using bulk RNAseq as control. Since no direct equivalent to the LICs and LBs exists in the mock-transplanted control mice, normal GMPs were used as the reference point (Figure 3.1B) to understand which genes or signalling pathways were upregulated in the AML cells and potentially essential for their maintenance [13, 56].

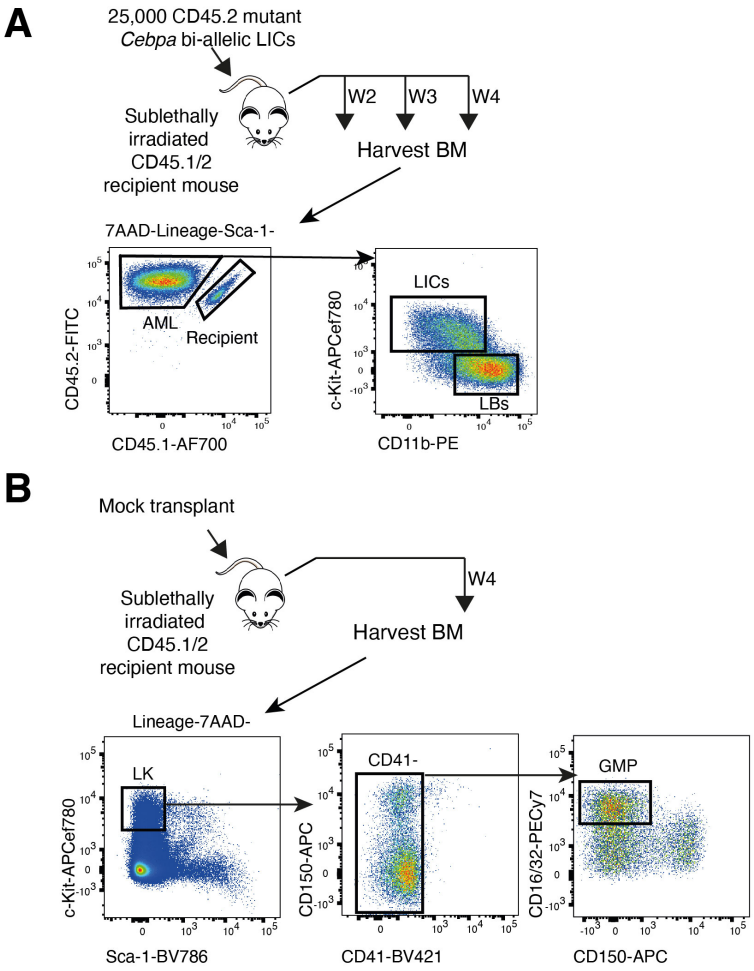

Figure 3.1. Gene expression profiling of bi-allelic* Cebpa *mutant leukaemia cells throughout AML development

(A) Graphical representation of isolating Cebpa mutant LICs and LBs for bulk RNAseq. Briefly, 25,000 CD45.2 bi-allelic Cebpa mutant LICs were transplanted into sub-lethally irradiated recipients (CD45.1/2). After 2, 3, or 4 weeks post-transplantation, mice were euthanised for leukaemia cell isolation and RNAseq library preparation.
(B) Graphical representation of isolating healthy GMPs for bulk RNAseq. Briefly, recipient mice (CD45.1/2) were sub-lethally irradiated without LIC transplantation. 4 weeks post-transplantation, mice were euthanised for GMP isolation and RNAseq library preparation.

First, DESeq2 was used to identify DEGs by comparing the LIC and healthy GMP samples sorted at 2, 3, and 4 weeks after LIC transplantation, as described in Thomas A, 2018 [56]. Interestingly, for both the LICs and LBs, the number of DEGs peaked 3 weeks after LIC transplantation, when the BM leukaemic burden was already high and continued to increase [56]. Consistent with the hypothesis that cytokine receptors can be used as a potential target for AML, a number of cytokine and cytokine receptor genes were overexpressed in the LICs and LBs compared with the healthy GMPs. These genes included *Il2rg*, *Ccr1*, *Ccl3*, *Tnf*, etc (Figure 3.2) [56]. In total, over 20 cytokine receptors were upregulated in the LICs during AML progression; a complete list of the cytokine and cytokine receptor genes overexpressed in the LICs and LBs at different time points can be found in Table 3.1.

To further interrogate how AML cell function changed throughout AML progression, the Nerlov laboratory also performed MPNA to identify perturbed molecular networks, as described in Thomas A, 2018 [56]. In line with previous studies suggesting the leukaemic BM microenvironment is inflamed [77], several inflammation pathways were increased in the leukaemia cells. This included IL-6 signalling, IL-2 signalling, TREM1 signalling, etc. In addition to the leukaemia cells, the Nerlov laboratory also looked into the transcriptional changes to the BM stromal cells during AML development by bulk RNAseq [56]; more details will be described in Chapter 5. Not surprisingly, MPNA showed that inflammation-related process networks such as IL-6 signalling and IL-2 signalling were also enriched in the AML-exposed CBM VECs, which further supported the possibility of eliminating AML cells by inhibiting specific cytokine signalling, especially by targeting the pro-inflammatory cytokine receptors highly expressed on the LICs.

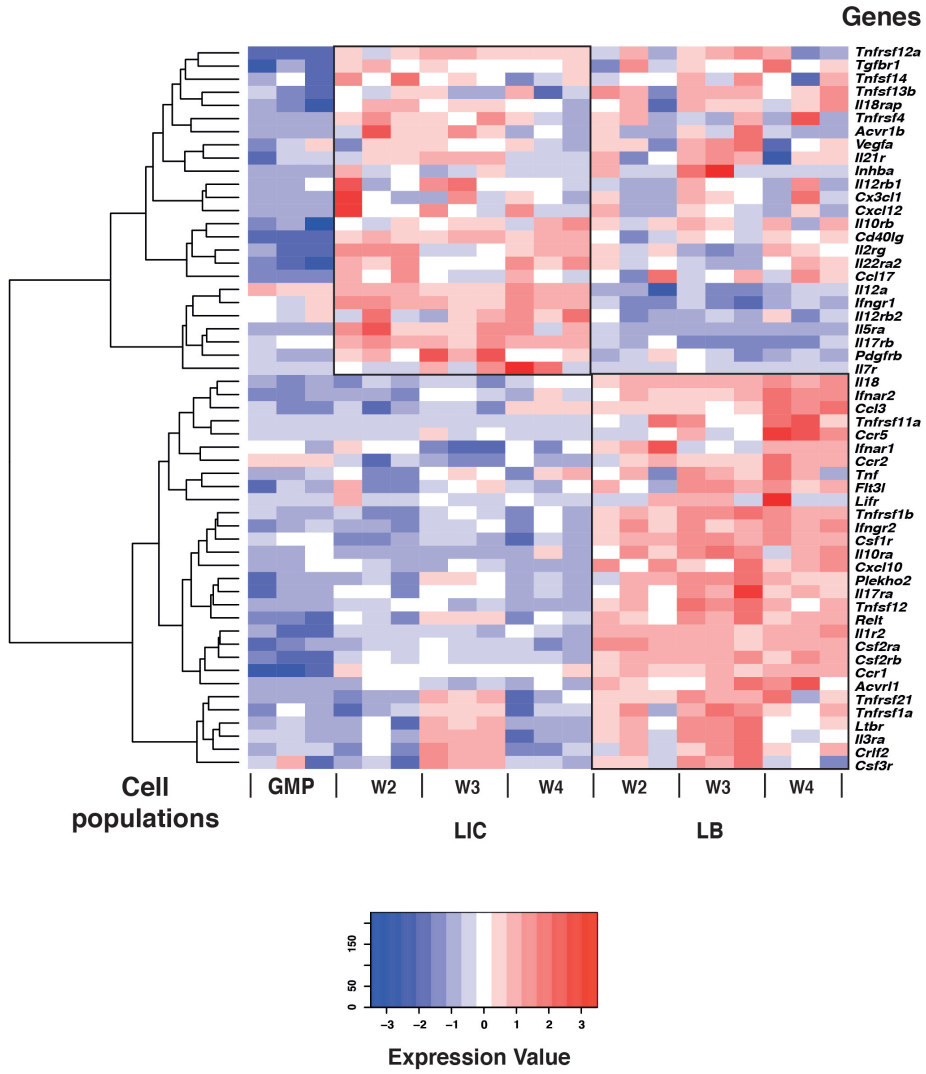

Figure 3.2. Cytokine signalling gene expression in AML cells

Heatmap showing ln(RPKM+1) expression values of up-DEGs from the KEGG cytokine-cytokine receptor interaction gene set in the LICs, LBs, and GMPs isolated from LIC and mock-transplanted mice at the indicated time points. The up-DEGs are defined as DEGs of LIC and LBs with an adjusted p-value < 0.05 and a log2 fold change > 0 at any time point compared to GMPs. The expression values are centred on the mean value for each gene.

To investigate the transcriptional changes to the *Cebpa* mutant LICs during leukaemia progression at the gene set level, I also performed GSEA with or without pre-ranking to the RNAseq data, focused on the inflammation-related gene sets. Since the 3-week time point transplantation seemed to be a critical time point during AML progression with the highest number of DEGs, I was particularly interested in the gene set enriched in LICs at 3 weeks after LIC transplantation. For the week 3 LICs, when analysed with pre-ranking, the KEGG cytokine-cytokine receptor gene set and the Hallmark TNFα signalling via NFκB gene set were significantly enriched (Fig-

Cell types	Time points	Genes
LICs	Week 2	*Cd40lg*, *Ccr1*, *Il22ra2*, *Il5ra*, *Il2rg*, *Tnfrsf12a*, *Ccl17*, *Il18rap*, *Il17rb*, *Pdgfrb*, *Tnfsf14*, *Tnfrsf4*, *Tgfbr1*, *Csf2rb*, *Ifngr1*, *Csf2ra*
	Week 3	*Cd40lg*, *Cxcl12*, *Il5ra*, *Tnfrsf12a*, *Ccr1*, *Pdgfrb*, *Tnfrsf4*, *Acvr1b*, *Il22ra2*, *Il7r*, *Relt*, *Il12rb1*, *Tnf*, *Il3ra*, *Plekho2*, *Il18rap*, *Il21r*, *Il17rb*, *Crlf2*, *Ltbr*, *Tnfrsf21*, *Tgfbr1*, *Csf2ra*, *Csf2rb*, *Il17ra*, *Il2rg*, *Il10rb*, *Ifnar2*
	Week 4	*Cd40lg*, *Ccr1*, *Il22ra2*, *Tnfrsf12a*, *Ccl3*, *Il5ra*, *Il2rg*, *Il17rb*, *Il12rb2*, *Pdgfrb*, *Il12a*, *Tgfbr1*, *Il10rb*, *Ifngr1*, *Csf2rb*
LBs	Week 2	*Ccr1*, *Cd40lg*, *Acvrl1*, *Il18*, *Tnfrsf12a*, *Tnfrsf11a*, *Il22ra2*, *Ccl3*, *Csf2ra*, *Il2rg*, *Ifngr2*, *Il17ra*, *Tnfrsf21*, *Ifnar2*, *Ifnar1*
	Week 3	*Cd40lg*, *Cxcl12*, *Tnfrsf12a*, *Acvrl1*, *Ccr1*, *Tnfsf12*, *Tnfrsf11a*, *Relt*, *Lifr*, *Tnf*, *Il18*, *Plekho2*, *Il3ra*, *Inhba*, *Cx3cl1*, *Tnfrsf1b*, *Ccr5*, *Flt3l*, *Ccl3*, *Il10ra*, *Tnfsf13b*, *Il1r2*, *Il18rap*, *Ltbr*, *Cxcl10*, *Csf1r*, *Il17ra*, *Csf2rb*, *Csf2ra*, *Crlf2*, *Il21r*, *Tnfrsf1a*, *Vegfa*, *Il22ra2*, *Tnfrsf21*, *Ifngr2*, *Csf3r*, *Tgfbr1*, *Ifnar2*, *Il10rb*
	Week 4	*Ccr5*, *Cd40lg*, *Ccl3*, *Tnfrsf11a*, *Ccr1*, *Acvrl1*, *Il18*, *Il22ra2*, *Ccl17*, *Tnf*, *Relt*, *Tnfrsf12*, *Tnfsf12*, *Plekho2*, *Csf1r*, *Il18rap*, *Flt3l*, *Il10ra*, *Csf2ra*, *Tnfrsf1b*, *Csf2rb*, *Il2rg*, *Tgfbr1*, *Ifnar2*, *Ifngr2*, *Il17ra*, *Tnfrsf21*, *Ccr2*, *Il10rb*, *Ifnar1*

Table 3.1. KEGG cytokine signalling genes overexpressed in LICs and LBs

Up-DEGs of LICs and LBs from the KEGG cytokine-cytokine receptor interaction gene set at 2, 3, and 4 weeks post-transplantation are shown. Healthy GMPs isolated from mock-transplanted mice are used as controls. All DEGs shown have an adjusted p-value < 0.05 and a log2 fold change > 0.

ure 3.3A). When the week 3 LICs were analysed without pre-ranking, the KEGG chemokine signalling gene set and the KEGG JAK-STAT signalling gene set were significantly enriched (Figure 3.3B). The JAK-STAT signalling pathway is downstream of various cytokine receptors, such as the IL-6 receptor, playing an essential role in inflammation [194].

Taken together, all the data above suggests that cytokine secretion in the BM environment where the AML cells reside and proliferate is de-regulated. The AML-exposed BM environment is seemingly pro-inflammatory, with several inflammation-

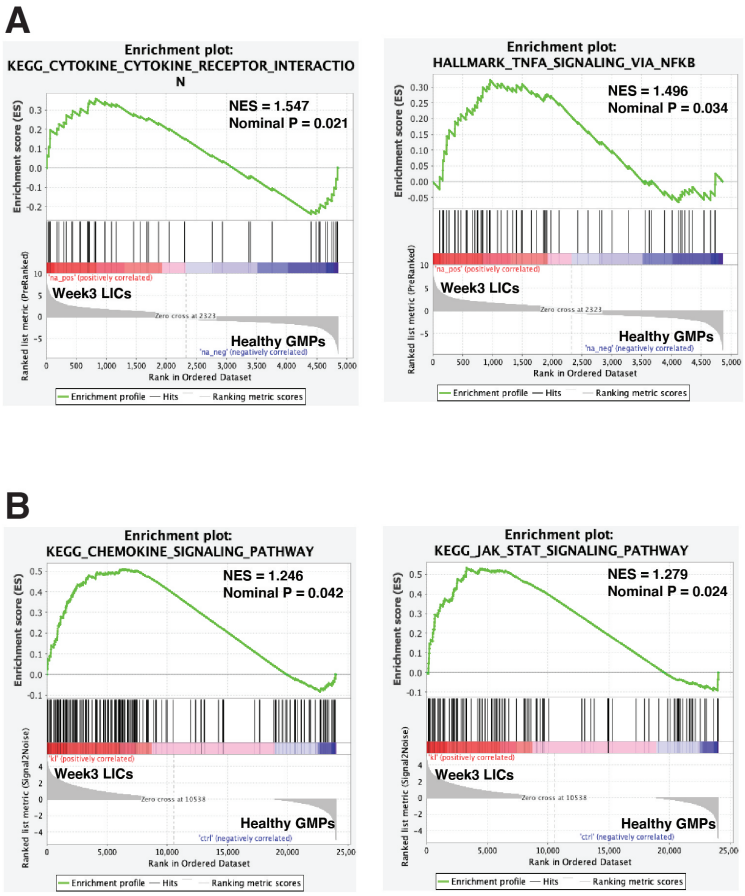

Figure 3.3. Inflammation-related gene sets are enriched in bi-allelic* Cebpa *mutant leukaemia cells

(A) GSEA (with pre-ranking) comparing LICs isolated 3 weeks after LIC transplantation to healthy GMPs using the KEGG cytokine-cytokine receptor gene set and the Hallmark TNF-α signalling via NFκB gene set. The RNAseq data used are from Figure 3.1. The normalised enrichment score (NES) and p-value are indicated.
(B) GSEA (without pre-ranking) comparing LICs isolated 3 weeks after LIC transplantation to healthy GMPs using the KEGG chemokine signalling gene set and the KEGG JAK-STAT signalling gene set. The RNAseq data used are from Figure 3.1. The NES and p-value are indicated.

related signalling pathways being activated and potentially functional in propagating AML. Therefore, it is conceivable that specific cytokine receptor genes, especially those highly expressed in the LICs, could be essential for LIC maintenance and potentially be used as druggable targets to eliminate the LICs. As previously mentioned, *in vivo* CRISPR dropout screens would be a powerful tool to identify the cytokine receptor essential for LIC survival/maintenance. Besides, a high representation of the sgRNA library is critical for the success of a given CRISPR screen, which helps avoid random depletion of the sgRNA and the ensuing false positive hits

[151]. Therefore, in order to maintain a high representation of the sgRNA library by reducing the library size, I performed 2 independent *in vivo* CRISPR dropout screens to examine the role of all cytokine receptor genes of interest demonstrated in Figure 3.2 [151, 154]. Each screen was focused on a different set of genes that were significantly upregulated by the LICs, and a mini-pool lentiviral library comprising 40-60 sgRNAs was used in individual screens.

3.3 An *in vivo* CRISPR dropout screen focused on genes consistently upregulated throughout *Cebpa* mutant AML development in mice

3.3.1 Selection of candidate genes included in the screen and sgRNA library design

As previously mentioned, I sought to identify *in vivo* dependencies in AML by performing two CRISPR dropout screens using the murine *Cebpa* mutant AML model. While selecting candidate genes included in the first screen, I focused on genes, especially the cytokine receptor genes, which were significantly overexpressed in the *Cebpa* mutant LICs during AML progression. I focused on the LICs rather than the LBs as the LICs are the population that can initiate leukaemia and, therefore, are more representative of the human LSCs.

DEG analysis of the LICs revealed a number of overexpressed genes; of particular interest were the cytokine receptor genes, *Ccr1*, *Il22ra2*, *Il5ra*, *Il2rg*, *Tnfrsf12a*, *Il17rb*, *Pdgfrb*, *Tgfbr1*, *Il18rap*, and *Tnfrsf4*, which were highly overexpressed (adjusted p-value < 0.05, log2 fold change > 1) at 2 or 3 time points. Presumably, the cytokine receptor genes overexpressed at multiple stages of AML progression are more likely to be an AML dependency than those overexpressed at just 1 single time point. Then, I excluded *Il5r*, *Tnfrsf12a*, *Il18rap*, *Tnfrsf4*, and *Pdgfrb* from the current screen because of their relatively low RNA expression level (RPKM value $>$ 5) when compared to the other candidates. In addition to these cytokine receptor genes, for comparison, I included 2 non-receptor genes, *Gdf3* and *Slc39a4*. They

were similarly overexpressed by the LICs (adjusted p-value < 0.05, log2 fold change > 1, RPKM value > 5). *Gdf3* was the only secreted factor gene overexpressed by the LICs throughout the disease trajectory; *Slc39a4*, encoding a zinc transporter, was the only gene overexpressed by both the LICs and the AML-exposed CBM VECs at all the time points. Collectively, 7 genes were included in the current screen, including *Ccr1*, *Il22ra2*, *Il2rg*, *Il17rb*, *Tgfbr1*, *Gdf3*, and *Slc39a4* (Figure 3.4). 5 sgRNAs were designed for each of them according to previous protocols as outlined in Tzelepis et al., 2016 [167], leading to a total of 35 experimental sgRNAs (Table 2.3).

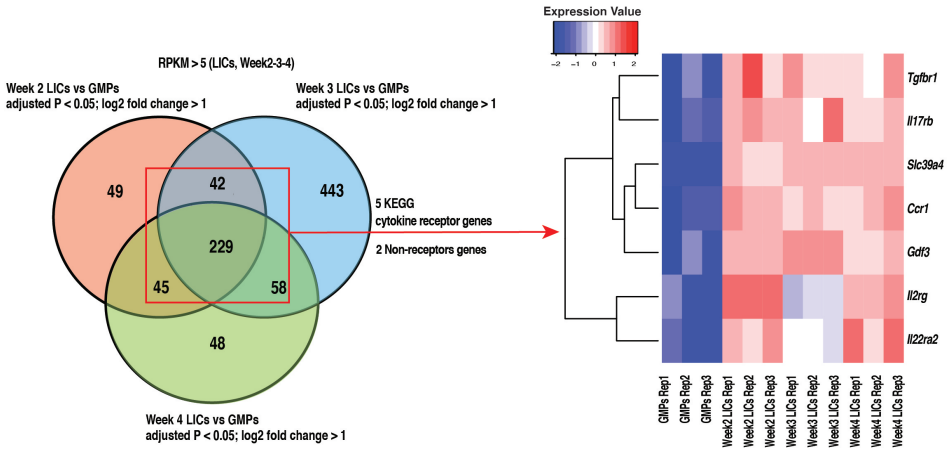

Figure 3.4. Selection of candidate genes included in the first CRISPR screen
Venn diagram showing the intersect of up-DEGs with an RPKM value > 5 and a log2 fold change > 1 in LICs. From such up-DEGs at two or more time points filtering on KEGG cytokine receptor-encoding genes in LICs identified 5 genes. These and 2 non-receptor genes were included in the screen. The heatmap shows ln(RPKM+1) expression values centred on the mean value for each gene.

3.3.2 Development of the *in vivo* screen

To functionally examine whether these candidate genes were essential for LIC maintenance *in vivo*, the GFP-expressing lentiviral sgRNA library (encoding the 35 experimental sgRNAs) was introduced into the c-Kit-enriched LICs, followed by 100,000-150,000 LICs being transplanted into each lethally irradiated mouse with competitor cells after an 8-hour transduction. All the LICs used for lentiviral transduction were freshly purified from the BM of primarily transplanted recipient mice. After the 8-hour transduction, an aliquot of the LICs was kept *ex vivo* for 48 hours instead of being transplanted to the secondary recipients to readout transduction

efficiency by flow cytometry [195]. The transduction protocol led to a transduction efficiency of ca. 20% after 48 hours (Figure 3.5A), equating to 400-500 sgRNA-expressing cells/sgRNA being transplanted into each secondary recipient mouse. According to previous work, such coverage of the sgRNA library should suffice [156, 157, 163]. 2 independent infection replicate experiments were performed. The secondary recipient mice were sacrificed when they became hunched with pale paws or when CD45.2+ leukaemia cell engraftment in the PB was $> 60\%$. In the first experiment, 3 mice were culled at 4 weeks post-transplantation. In the second experiment, 2 mice were culled at 5 weeks post-transplantation. Hence, the kinetics of the original sub-lethal irradiated model, from which the bulk RNAseq of LICs was performed (as shown in Figure 3.1), were well preserved.

After the secondarily transplanted mice were sacrificed, sgRNA-expressing LICs and LBs (GFP+) in the BM were isolated by cell sorting followed by genomic DNA purification (Figure 3.5B). Then, sgRNAs were amplified by PCR and subjected to Illumina sequencing, which was used to read out the sgRNA representation in the leukaemia cells and to identify the sgRNAs that were depleted during leukaemia progression. In parallel, WT NIH3T3 cells were transduced with the same lentiviral libraries and collected at 3 days post-transduction by sorting the GFP+ cells. NIH3T3 is a murine fibroblast cell line whose culture does not need external supplementation of cytokines, suggesting its independence of functional cytokine receptors [196]. Therefore, the sgRNA representation within these cells was used as an input reference for the sgRNA representation at the beginning of the experiment. In addition, the transduced NIH3T3 cells were collected at 14 days post-transduction to identify any sgRNAs of which the depletion was not specific to leukaemia cells. A schematic representation of how the screen was carried out is shown in Figure 3.6. Details of the source of materials for each sequenced sample can be found in Table 3.2.

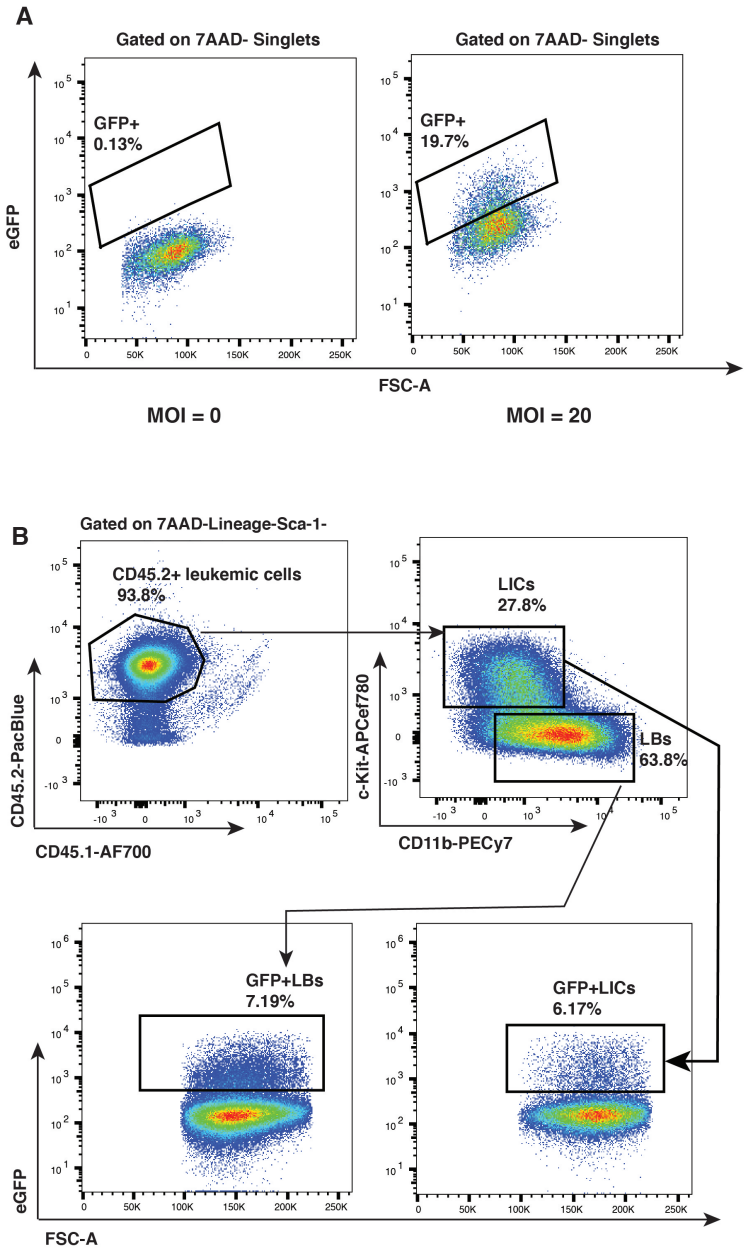

Figure 3.5. Gating of sgRNA-expressing AML cells in the first CRISPR screen
(A) Representative flow cytometry of sgRNA-expressing GFP+ leukaemia cells at 48 hours post-transduction (right). The GFP+ gate was set based on MOI=0 control (left).
(B) Representative flow cytometry of sgRNA-expressing GFP+ LICs (bottom right) and LBs (bottom left) isolated from the BM 4 weeks after LIC transplantation.

3.3.3 Dynamic change in sgRNA distributions during leukaemia development

To have a general impression of how the experimental sgRNA representation changed over time in both the leukaemic and non-leukaemic settings, the overall distributions of experimental sgRNAs from all the GFP+ samples that had been sequenced (Day

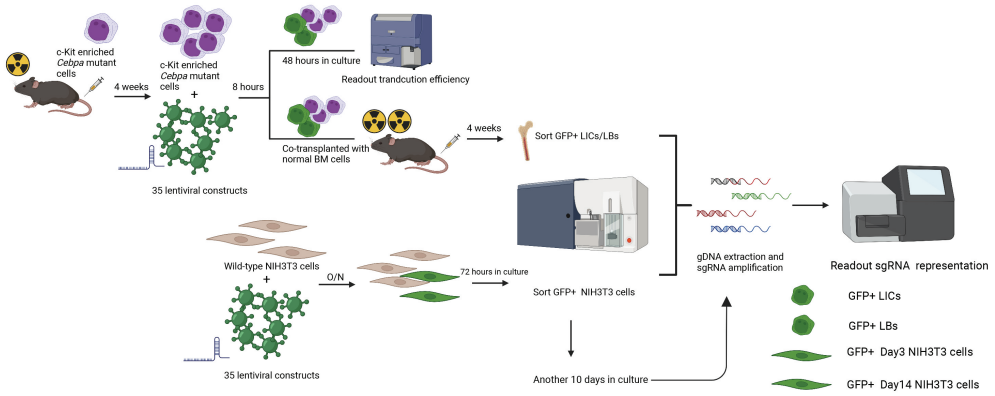

Figure 3.6. Schematic representation of the first CRISPR-Cas9 dropout screen
Freshly isolated LICs were transduced with an EGFP-expressing lentiviral sgRNA library targeting the 7 candidate genes (5 sgRNAs per gene) shown in Figure 3.4 and transplanted into recipient mice (N = 5, 2 infection experiments, 100-150,000 cells per mouse) after an 8-hour incubation. The sgRNA-expressing LICs and LBs were re-isolated 4-5 weeks after transplantation based on EGFP fluorescence, and the sgRNAs were retrieved from their gDNA. In parallel, sgRNAs were retrieved from WT NIH3T3 cells 3 days after transduction with the same lentiviral sgRNA library as input control. sgRNAs were also retrieved from WT NIH3T3 cells 14 days after transduction to identify non-specific dropouts. sgRNA representations were determined by NGS and compared using the MAGeCK RRA.

3 NIH3T3 cells, Day 14 NIH3T3 cells, LICs and LBs) were analysed (Figure 3.7). The pattern of experimental sgRNA distributions did not change significantly between Day 3 and Day 14 NIH3T3 cells. This indicated that all the candidate genes included in the screen were dispensable for the survival of the non-leukaemic cells and supported the use of Day 3 NIH3T3 cells as input reference.

Regarding the sgRNA representation in the LICs, all 35 experimental sgRNAs were detected in the first experiment, whereas 35-45% of experimental sgRNAs were not detected in the second experiment. Besides, in the second experiment, the sgRNA distribution was strongly skewed within each sample: a random sgRNA contributed to over 90% of the total reads. This suggested that the sgRNA library coverage was insufficient in each transplanted mouse. Since the number of sgRNA-expressing LICs transplanted should have included enough cells for a 500-fold coverage, this was possibly caused by inefficient engraftment of the transduced LICs after transplantation. Indeed, the mice from the second experiment (which became moribund at week 5) were 1 week behind the mice from the first experiment (which became moribund at week 4) in terms of AML progression, suggesting a lower number of

Sample ID	Source of materials
1	GFP+ LICs sorted at endpoint, experiment 1, mouse 1
2	GFP+ LBs sorted at endpoint, experiment 1, mouse 1
3	GFP+ LICs sorted at endpoint, experiment 1, mouse 2
4	GFP+ LBs sorted at endpoint, experiment 1, mouse 2
5	GFP+ LICs sorted at endpoint, experiment 1, mouse 3
6	GFP+ LBs sorted at endpoint, experiment 1, mouse 3
7	GFP+ LICs sorted at endpoint, experiment 2, mouse 1
8	GFP+ LBs sorted at endpoint, experiment 2, mouse 1
9	GFP+ LICs sorted at endpoint, experiment 2, mouse 2
10	GFP+ LBs sorted at endpoint, experiment 2, mouse 2
11	GFP+ NIH3T3 cells sorted 3 days post-transduction, replicate 1
12	GFP+ NIH3T3 cells sorted 3 days post-transduction, replicate 2
13	GFP+ NIH3T3 cells sorted 3 days post-transduction, replicate 3
14	GFP+ NIH3T3 cells sorted 14 days post-transduction, replicate 1
15	GFP+ NIH3T3 cells sorted 14 days post-transduction, replicate 2

Table 3.2. Sorted sgRNA-expressing populations for sequencing during the first screen

The cellular source of gDNA used for sgRNA amplification and NGS for identifying sgRNA/genes depleted during AML progression is shown.

implanted cells. Noticeably, the sgRNA distributions in the mice from the first experiment were also skewed compared to those in the NIH3T3 samples. But they were more balanced than those in the mice from the second experiment, indicating that a reasonable coverage of the experimental sgRNA library was reached and specific candidate genes may negatively impact LIC maintenance. Hence, the data from the second experiment were excluded from downstream analyses.

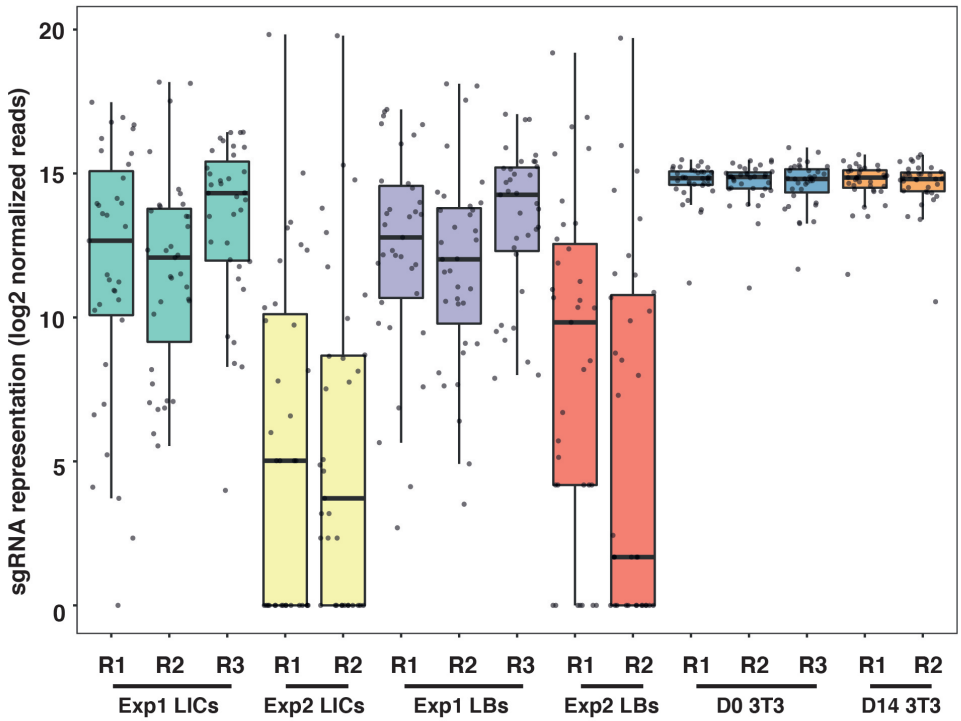

Figure 3.7. Representation of the sgRNA library in AML cells and NIH3T3 cells during the first screen

Boxplot of the normalised sgRNA read counts for the GFP+ LICs, GFP+ LBs, Day 3 NIH3T3 cells, and Day 14 NIH3T3 cells as in Figure 3.6. Individual dots overlaid on each boxplot indicate read counts for the 35 sgRNAs included in the lentiviral sgRNA library.

3.3.4 *Ccr1* was identified as the top hit

First, the total sgRNA reads in each sample were normalised across all the samples that had been sequenced. To identify the sgRNAs and genes that were the most depleted during leukaemia development, the MAGeCK RRA was used to compare sgRNA read counts between the end GFP+ LICs and the input Day 3 GFP+ NIH3T3 cells [184, 185] and calculate log2 fold change of the sgRNA representation. The MAGeCK RRA has been widely used to analyse CRISPR screen data and identify top hits [163, 167]. As expected, the log2 fold change of the sgRNA representation in the two biologically related populations, GFP+ LICs and GFP+ LBs (vs input), was well correlated (Pearson correlation coefficient = 0.886) (Figure 3.8A). This indicated good reproducibility of the screening protocol. The MAGeCK pipeline was also used to calculate each candidate gene's depletion score for the LICs. Of the 7 candidate genes included in the screen, *Ccr1* and *Il2rg* were the top 2 dropout genes in the LICs, with the lowest depletion score (Figure 3.8B). For *Ccr1*

and *Il2rg*, 3 corresponding sgRNAs were depleted (log2 fold change < -1), whereas no sgRNA was enriched (log2 fold change > 1). In contrast, for the least depleted genes, *Il22ra2* and *Tgfbr1*, 3 corresponding sgRNAs were enriched (log2 fold change > 1), whereas no sgRNA was depleted (log2 fold change < -1) (Figure 3.8C).

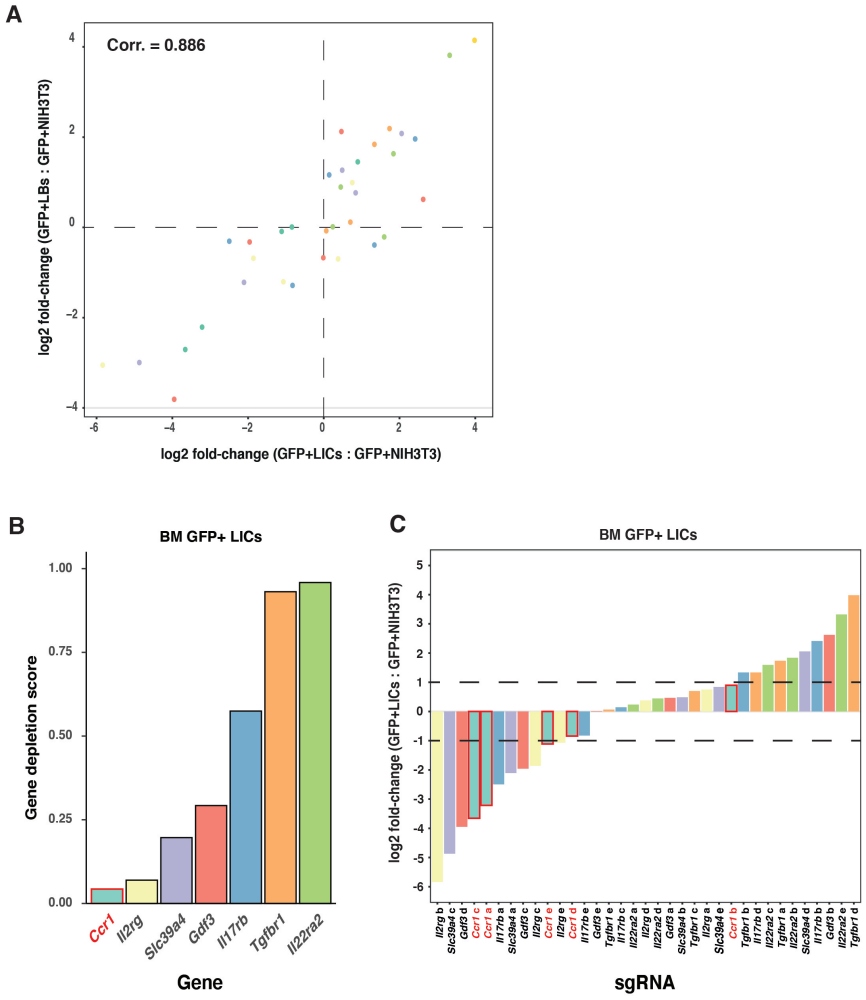

Figure 3.8. Ccr1 was identified as the top hit by MAGeCK RRA

(A) Scatterplot of log2 fold change of normalised sgRNA read counts (vs input control) in LICs and LBs for all sgRNAs in the screen. Only mice from the first infection experiment were included. The log2 fold change was calculated by MAGeCK RRA. Different colour dots indicate sgRNAs targeting different genes. Pearson correlation coefficient of log2 fold change of normalised sgRNA read counts (vs input control) between BM LICs and LBs is shown.
(B) Histogram showing the negative selection score and ranking (calculated by MAGeCK RRA) of the candidate genes included in the screen. Lower scores indicate a higher level of gene depletion in the BM LICs vs the NIH3T3 input control.
(C) Waterfall plot showing the log2 ratio (calculated by MAGeCK RRA) of sgRNA representation in the LICs versus NIH3T3 input control. The bars with red borders represent sgRNAs targeting Ccr1.

To further support the depletion scores calculated by the MAGeCK pipeline, a CS

for each candidate gene was also calculated, which had already been used to help identify essential genes during CRISPR-based screens. The CS for each gene was defined as the average log2 fold change in the representation of all sgRNAs targeting the same gene between the input and endpoint LICs [168]. The log2 fold change for each sgRNA was calculated by first comparing the sgRNA representation in each murine LIC sample and the input sgRNA representation, followed by averaging the log2 fold change values across different mouse replicates. Then, the averaged log2 fold change for each sgRNA was used for CS calculation (Figure 3.9A). Eventually, all the candidate genes included in the screen were ranked based on their CS (Figure 3.9B). Consistent with the depletion score generated by the MAGeCK, *Ccr1* and *Il2rg* had the lowest CS among all the genes studied, indicating them as the 2 top hits; *Tgfbr1* and *Il22ra2* had the highest CS, indicating their low depletion levels in the LICs.

qRT-PCR was then used to more accurately quantify the expression level of the cytokine receptor genes included in the screen by comparing leukaemia cells (sorted 4 weeks after LIC transplantation) and healthy GMPs (sorted 4 weeks after mock transplantation). Of note, *Ccr1* was highly overexpressed in both the LICs (20-fold increase, p-value < 0.0001) and LBs (160-fold increase, p-value $= 0.0252$) (Figure 3.10A). CCR1 expression in the LICs and LBs was also confirmed using flow cytometry (Figure 3.10B). Furthermore, overexpression of the genes encoding common CCR1 ligands, *Ccl3* (28-fold increase, p-value < 0.0001), *Ccl6* (7-fold increase, p-value < 0.0001), and *Ccl9* (1.5-fold increase, p-value $= 0.0105$) was seen in LBs (Figure 3.10C-E). Only the gene expression of these 3 CCR1 ligands was tested using qRT-PCR as they are the ones of which the mRNA level was upregulated in the LICs/LBs (vs GMPs) according to the bulk RNAseq data. The protein level of CCL3 was significantly increased in AML-conditioned BM at different stages of AML progression (week 3, 58-fold increase, p-value < 0.0001; week 4, 95-fold increase, p-value $= 0.0004$), indicating the consistent existence of activated CCR1 signalling throughout the disease trajectory (Figure 3.10F). The protein level of CCL6 was also measured, but a significant increase in the leukaemic BM fluid was not observed (Figure 3.10G). Taken together, the strong leukaemia-induced overexpression of CCR1 and its key ligands supports the potential use of CCR1 as a

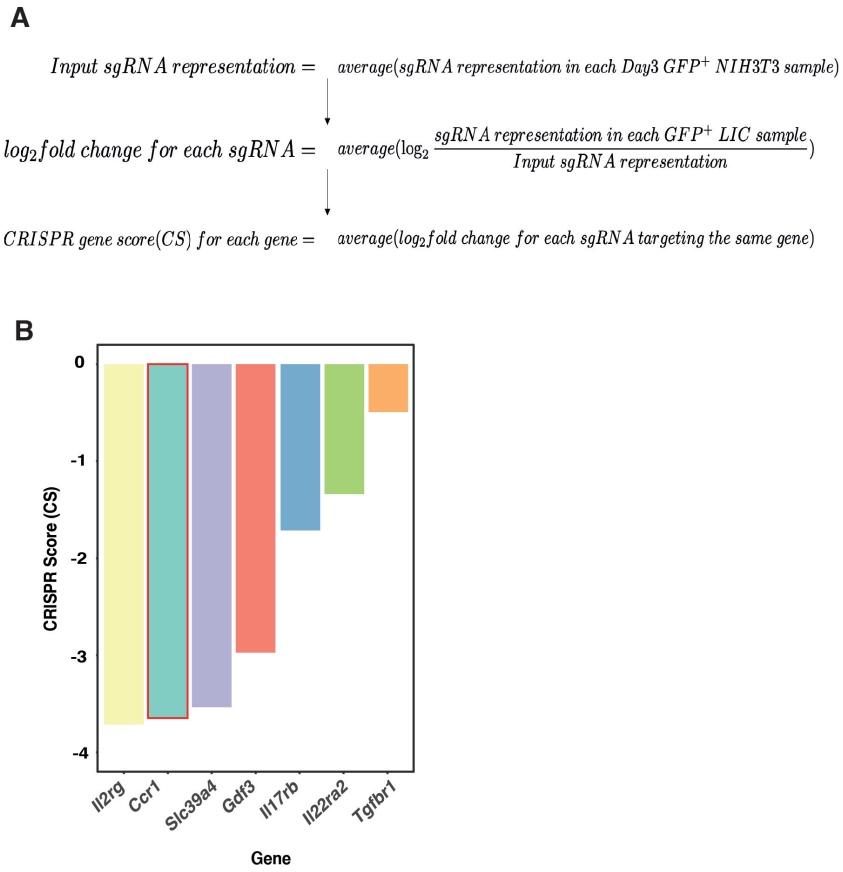

Figure 3.9. CRISPR scores (CS) of genes in the first screen

(A) Equation for calculating the CS for each candidate gene.
(B) Histogram showing the CS and ranking of the candidate genes included in the screen. Lower scores indicate a higher level of gene depletion in the BM LICs vs the NIH3T3 input control.

druggable target for the LICs.

Regarding other cytokine receptor genes included in the screen, overexpression of *Il2rg* in the leukaemia cells was also observed by qRT-PCR, albeit to a lower extent (Figure 3.11A). Interestingly, despite being identified as highly upregulated genes by the RNAseq, the expression of *Il17rb* and *Tgfbr1* was not increased in the LICs according to the qRT-PCR result (Figure 3.11B-C). This might explain why they were less depleted than *Ccr1* and *Il2rg* during the screen and less indispensable for LIC maintenance. In addition, the Nerlov laboratory has treated mice transplanted with *Cebpa* mutant LICs with a potent TGF-β receptor 1 inhibitor, where no significant reduction in the BM LIC engraftment was observed [56]. This further refutes *Tgfbr1* as a strong AML dependency, as suggested by the MAGeCK gene ranking and CS. However, the qRT-PCR verified the significant overexpression of *Il22ra2* in

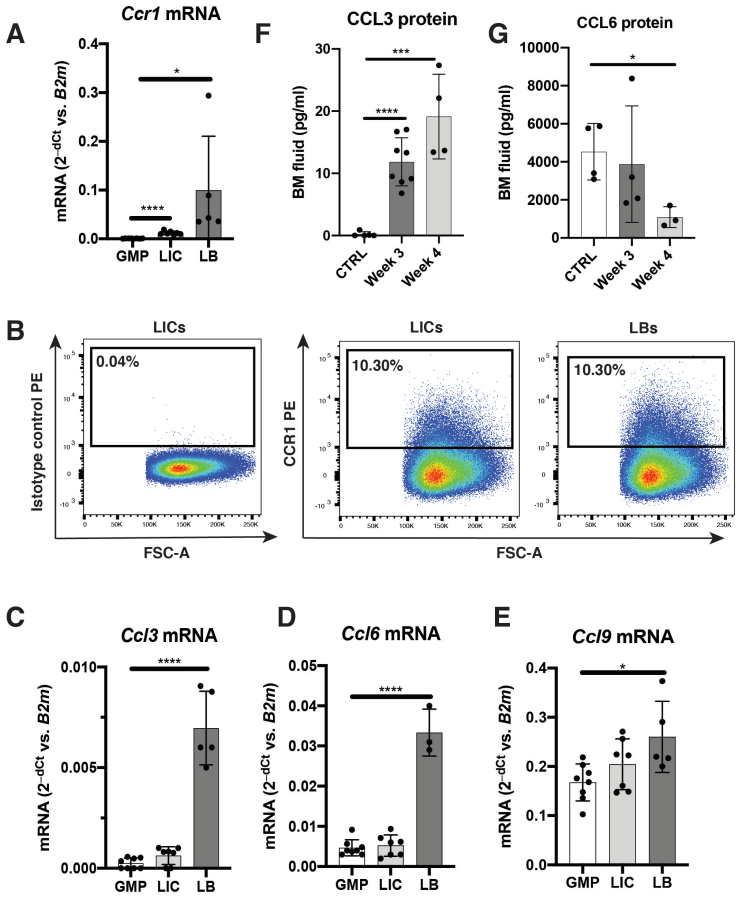

Figure 3.10.* Ccr1 *and the genes encoding CCR1 ligands were upregulated during AML

(A) Histogram showing Ccr1 mRNA expression levels in LICs (N=7) and LBs (N=5) isolated 4 weeks after LIC transplantation, compared to normal GMPs (N=8). Values shown are $2^{-\Delta Ct}$ normalised to B2m. Data from 2 independent experiments. Each biological replicate was the average of 2 technical replicates. P-values were calculated using the Student's t-test.

(B) Representative flow cytometry showing CCR1 expression on the BM LICs (middle) and LBs (right) 4 weeks after LIC transplantation. The positive gate was set based on isotype control antibody staining (left).

(C) Histogram as in (A) showing Ccl3 mRNA expression levels in LICs (N=7) and LBs (N=5) isolated 4 weeks after LIC transplantation, compared to normal GMPs (N=8).

(D) Histogram as in (A) showing Ccl6 mRNA expression levels in LICs (N=7) and LBs (N=3) isolated 4 weeks after LIC transplantation, compared to normal GMPs (N=8).

(E) Histogram as in (A) showing Ccl9 mRNA expression levels in LICs (N=7) and LBs (N=5) isolated 4 weeks after LIC transplantation, compared to normal GMPs (N=8).

(F) Histogram showing the levels of CCL3 protein in BM fluid from mock-transplanted mice (N=5) and LIC transplanted mice at the 3-week (N=8) and 4-week (N=4) time points. Protein levels were measured by ELISA. Data from 2 independent experiments. P-values were calculated using the Student's t-test.

(G) Histogram as in (F) showing the levels of CCL6 protein in BM fluid from mock-transplanted mice (N=4) and LIC transplanted mice at the 3-week (N=5) and 4-week N=3) time points.

the leukaemia cells (Figure 3.11D), even though it was not ranked highly as an *in vivo* hit during the dropout screen.

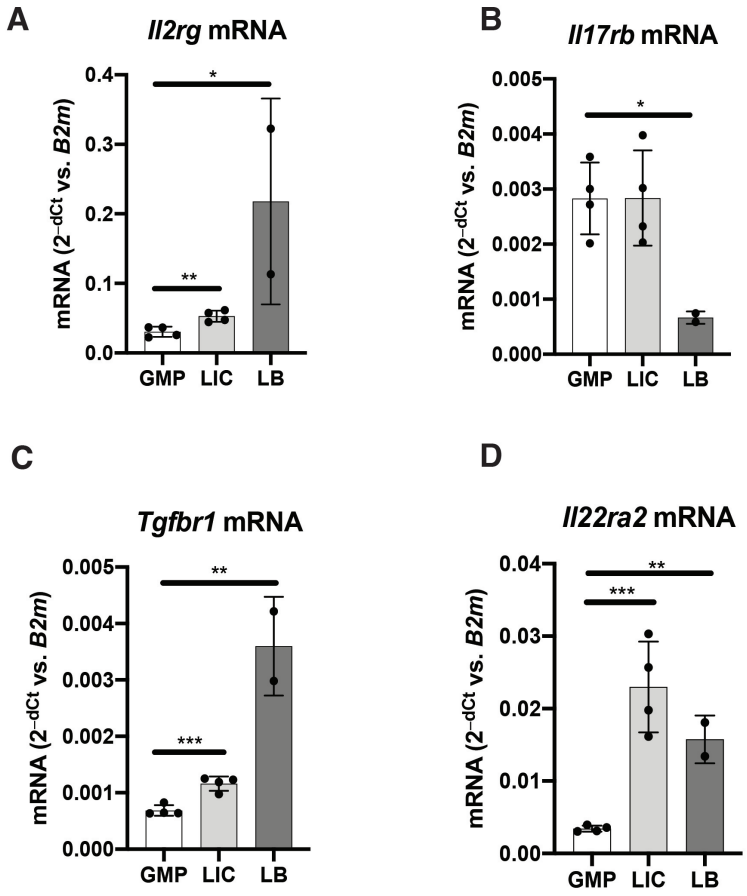

Figure 3.11. mRNA expression of cytokine receptor genes included in the first screen in the leukaemia cells

(A) Histogram showing Il2rg mRNA expression levels in LICs (N=4) and LBs (N=2) isolated 4 weeks after LIC transplantation, compared to normal GMPs (N=4). Values shown are $2^{-\Delta Ct}$ normalised to B2m. Data from 1 experiment. Each biological replicate was the average of 2 technical replicates. P-values were calculated using the Student's t-test.
(B) Histogram as in (A) showing Il17rb mRNA expression levels.
(C) Histogram as in (A) showing Tgfbr1 mRNA expression levels.
(D) Histogram as in (A) showing Il22ra2 mRNA expression levels.

Given the high degree of depletion of *Ccr1* during the screen and its high expression level in the AML cells, I decided to focus on this gene for further validation. Actually, CCR1 signalling has been implicated in the development of several types of haematological malignancies other than AML, such as CML and multiple myeloma (MM) [197, 198].

3.4 An *in vivo* CRISPR dropout screen focused on cytokine receptor genes simultaneously over-expressed in murine and human AML cells

3.4.1 Selection of candidate genes included in the screen and sgRNA library design

The second screen was focused on the cytokine receptor genes overexpressed in the murine *Cebpa* mutant LICs and human AML cells simultaneously. Gene expression values of human AML cells were retrieved from Valk et al., 2004 [173]. The DEG analysis of murine *Cebpa* mutant LICs identified 24 cytokine receptor genes over-expressed (adjusted p-value < 0.05) at one or more time points after murine *Cebpa* mutant LIC transplantation (Table 3.1). Among them, 10 genes (*Ccr1*, *Il7r*, *Il21r*, *Il18rap*, *Il5ra*, *Csf2ra*, *Il3ra*, *Ifngr1*, *Pdgfrb*, and *Il17ra*) were also overexpressed in the human AML cells (log2 fold change of the expression level > 1 between the AML cells and CD34+ normal HSPCs) (Figure 3.12) [173]. Interestingly, *Ccr1*, the top hit from the previous screen, was the only cytokine receptor gene meeting the selection criteria for both screens. To minimise the sgRNA library size, *Ccr1* was not included in the current screen. For the other 9 cytokine receptor genes, 5 sgRNAs were designed for each as described above, and 5 sgRNA targeting the universally essential gene *Rpa1* were also included as the positive control [160]. Besides, 10 non-targeting or non-essential gene-targeting (e.g. *Rosa26*-targeting) sgRNAs were included as the negative control. This led to a sgRNA library of 60 guides, and they were cloned into the previously described GFP-expressing lentiviral vector. Sequences and details of all the sgRNAs included in this screen can be found in Table 2.5.

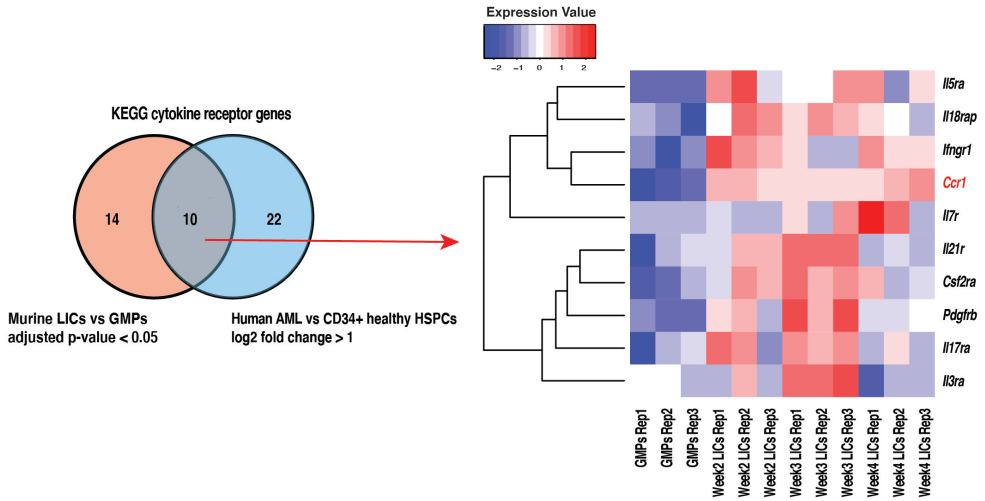

Figure 3.12. Selection of candidate genes included in the second CRISPR screen

Venn diagram showing the overlap of KEGG cytokine receptor-encoding genes upregulated in murine LICs (with an adjusted p-value < 0.05 and a log2 fold change of gene expression > 1 at any time point, healthy GMPs were used as controls) (from Table 3.1) and in human AML cells (with a log2 fold change of gene expression > 1, CD34+ normal HSPCs were used as controls) (data from Valk PJ, et al. 2004 [173]). The heatmap shows ln(RPKM+1) expression values centred on the mean value for each gene.

3.4.2 Development of the *in vivo* screen

To functionally assess which of the 9 candidate genes were essential for *in vivo* LIC maintenance, the GFP-expressing lentiviral sgRNA library was introduced into the c-Kit-enriched LICs, followed by 200,000 LICs being transplanted into each lethally irradiated mouse with competitor cells after an 8-hour transduction. All the LICs used for lentiviral transduction were freshly purified from the BM of primarily transplanted recipient mice. Transduction efficiency was assessed 48 hours post-transduction as previously described. 2 independent infection replicate experiments were performed: in the first experiment, only ca. 5.2% of the LICs were GFP+, whereas, in the second experiment, ca. 31.8% of the LICs were GFP+ at 48 hours post-transduction (Figure 3.13A). The secondarily transplanted mice were monitored as previously described to decide when they would be sacrificed: in the first experiment, 7 mice were transplanted and culled at week 4; in the second experiment, 3 mice were transplanted and culled at week 3. After the mice were sacrificed, sgRNA-expressing LICs and LBs (GFP+) in the BM were isolated by FACS, followed by gDNA purification and Illumina sequencing library preparation

(Figure 3.13B).

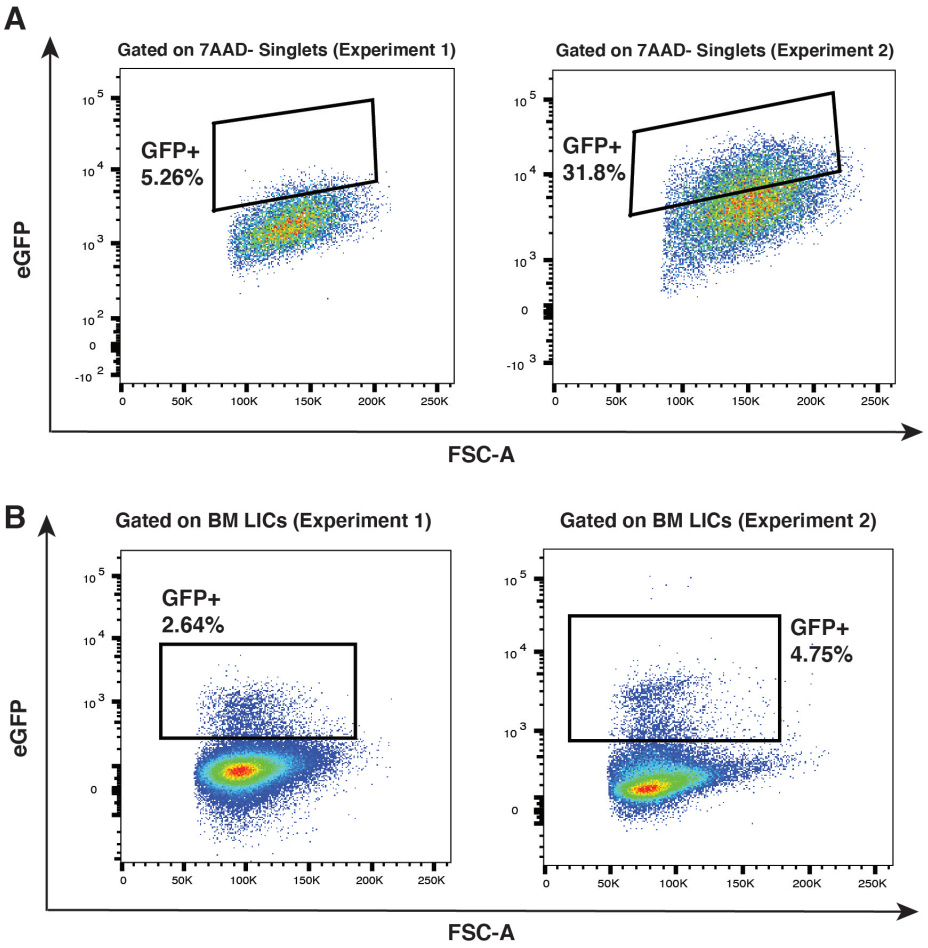

Figure 3.13. Gating of sgRNA-expressing AML cells in the second CRISPR screen

(A) Representative flow cytometry of sgRNA-expressing GFP+ leukaemia cells at 48 hours post-transduction during the first (left) and second (right) infection experiments. The GFP+ gate was set based on MOI=0 control.
(B) Representative flow cytometry of sgRNA-expressing GFP+ LICs isolated from the BM 4 weeks after LIC transplantation during the first (left) and second (right) infection experiments.

It was conceivable that the low transduction efficiency during the first experiment would lead to insufficient coverage of the sgRNA library in each recipient mouse (ca. 170-fold coverage). To compensate for this issue, during the library preparation for deep sequencing of the sgRNA representation, the 7 mice were randomly assigned to 2 groups, each with 3 or 4 mice. Then, an equal amount of gDNA from mice within each group was pooled together before the PCR reactions for sgRNA amplification. This would increase the sgRNA coverage in each sample when sequenced and avoid

an aberrantly skewed sgRNA library representation [156, 162]. In the meantime, leftover gDNA from 2 of the 7 mice was processed without pooling for comparison. Considering the relatively high transduction efficiency in the second experiment, which would lead to a good sgRNA coverage in each secondarily transplanted mouse (ca. 1000-fold coverage), the sorted BM samples were individually processed without pooling between different mice during library preparation. Like the first screen, WT NIH3T3 cells were transduced with the same sgRNA-expressing lentiviral library. GFP+ NIH3T3 cells were purified by FACS 3 and 14 days after transduction. They were used as the input reference and to identify any hit not specific to the leukaemia cells. A schematic representation of how the screen was performed is shown in Figure 3.14. Details of the source of materials for each sample that was sequenced can be found in Table 3.3.

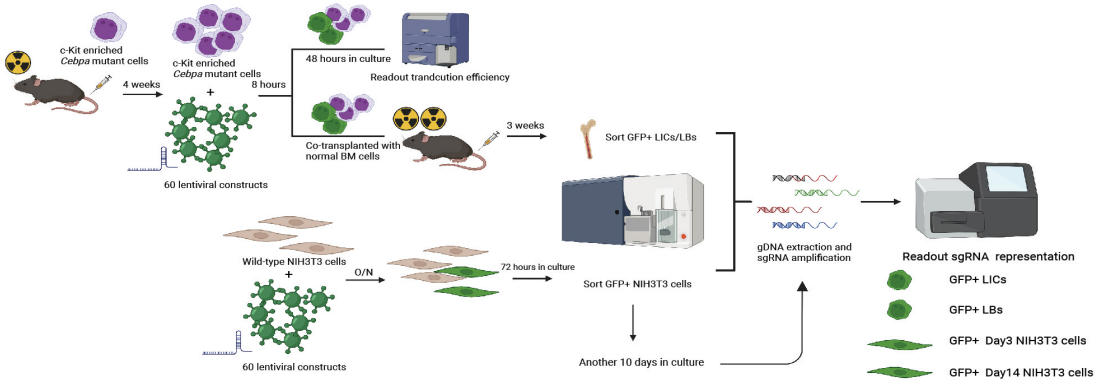

Figure 3.14. Schematic representation of the second CRISPR-Cas9 dropout screen

Freshly isolated LICs were transduced with an EGFP-expressing lentiviral sgRNA library, including 44 sgRNAs targeting the 9 candidate genes shown in Figure 3.12, 10 negative control sgRNAs, and 5 positive sgRNA targeting Rpa1. After 8-hour incubation, the LICs were transplanted into recipient mice (N = 7, 2 infection experiments, 200,000 cells per mouse). The sgRNA-expressing LICs and LBs were re-isolated 3-4 weeks after transplantation based on EGFP fluorescence, and the sgRNAs were retrieved from their genomic DNA. In parallel, sgRNAs were retrieved from WT NIH3T3 cells 3 days after transduction with the same lentiviral sgRNA library as input control. sgRNAs were also retrieved from WT NIH3T3 cells 14 days after transduction to identify non-specific dropouts. sgRNA representations were determined by NGS and compared using the MAGeCK RRA.

Sample ID	Source of materials
1	GFP+ LICs sorted at endpoint, experiment 1, a pool of mouse 1, 2, and 4 (R1)
2	GFP+ LICs sorted at endpoint, experiment 1, a pool of mouse 3, 5, 7, and 8 (R2)
3	GFP+ LICs sorted at endpoint, experiment 1, mouse 4 alone (R3)
4	GFP+ LICs sorted at endpoint, experiment 1, mouse 7 alone (R4)
5	GFP+ LICs sorted at endpoint, experiment 2, mouse 1
6	GFP+ LICs sorted at endpoint, experiment 2, mouse 2
7	GFP+ LICs sorted at endpoint, experiment 2, mouse 3
8	GFP+ LBs sorted at endpoint, experiment 1, a pool of mouse 1, 2, and 4 (R1)
9	GFP+ LBs sorted at endpoint, experiment 1, a pool of mouse 3, 5, 7, and 8 (R2)
10	GFP+ LBs sorted at endpoint, experiment 1, mouse 4 alone (R3)
11	GFP+ LBs sorted at endpoint, experiment 1, mouse 7 alone (R4)
12	GFP+ LBs sorted at endpoint, experiment 2, mouse 1
13	GFP+ LBs sorted at endpoint, experiment 2, mouse 2
14	GFP+ LBs sorted at endpoint, experiment 2, mouse 3
15	GFP+ NIH3T3 cells sorted 3 days post-transduction, replicate 1
16	GFP+ NIH3T3 cells sorted 3 days post-transduction, replicate 2
17	GFP+ NIH3T3 cells sorted 3 days post-transduction, replicate 3
18	GFP+ NIH3T3 cells sorted 14 days post-transduction, replicate 1
19	GFP+ NIH3T3 cells sorted 14 days post-transduction, replicate 2
20	GFP+ NIH3T3 cells sorted 14 days post-transduction, replicate 3

Table 3.3. Sorted sgRNA-expressing populations for sequencing during the second screen

The cellular source of gDNA used for sgRNA amplification and NGS for identifying sgRNA/genes depleted during AML progression is shown.

3.4.3 Dynamic evolution of sgRNA library representations during leukaemia development

The overall sgRNA distributions in all the sequenced samples were examined after mapping the sequencing reads to the sgRNA sequences (Figure 3.15). Compared with the input reference (Day 3 NIH3T3 cells), the pattern of experimental sgRNA distributions did not change significantly in the Day 14 NIH3T3 cells. The only exception was the positive control *Rpa1*-targeting sgRNAs, which were significantly depleted in the day 14 samples. This indicates that the cytokine receptor genes included were not essential for NIH3T3 cell survival and the validity of using them

as non-dropout control. Then, I looked into the 2 samples from the first experiment that were processed without pooling. As expected, the sgRNA distributions were strongly skewed: 1 or 2 sgRNAs contributed to more than 80% of the total mapped reads, and some sgRNAs were not recovered.

Regarding the 2 samples from the first experiment that were sequenced with pooling, it was evident that the pattern of sgRNA distributions was less skewed. This implied improved coverage of the sgRNA library after pooling different mice together. However, the top 5 most abundant sgRNAs still made up more than 70% of the total reads, and all the negative control sgRNAs were depleted (constituting less than 0.2% of the total reads). This indicated insufficient sgRNA library coverage and the existence of accidental dropout. As a result, all the samples from the first infection experiment had to be excluded from downstream analyses. For the samples from the second experiment, where the transduction efficiency was high, the sgRNA distribution was relatively balanced, suggesting improved coverage of the sgRNA library. Besides, more significant changes in the sgRNA representation could still be observed when compared with the Day 14 NIH3T3 cells, indicating that specific candidate genes were exclusively essential for LIC survival.

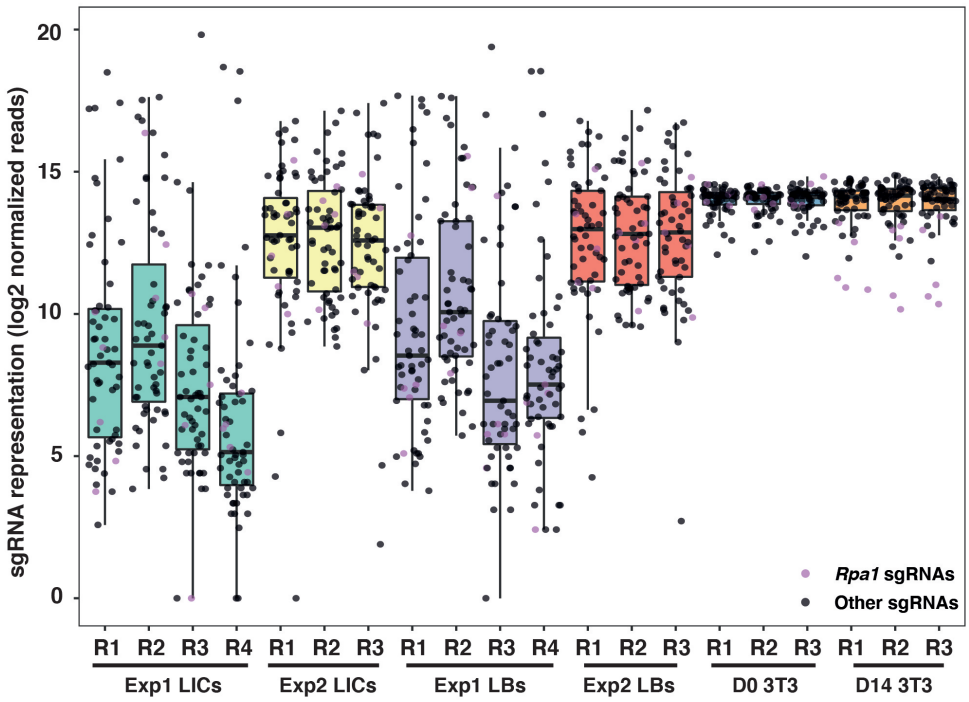

Figure 3.15. Representation of the sgRNA Library in AML cells and NIH3T3 cells during the second screen

Boxplot of the sgRNA normalised read counts for the GFP+ LICs, GFP+ LBs, Day 3 NIH3T3 cells, and Day 14 NIH3T3 cells as in Figure 3.14. Individual dots in grey overlaid on each boxplot indicate read counts for the cytokine receptor-targeting and Rpa1-targeting sgRNAs; individual dots in purple indicate read counts for the 10 negative control sgRNAs included in the screen.

3.4.4 *Il7r* was identified as a gene potentially essential for LIC maintenance

As all the sgRNAs were cloned into the GFP-expressing lentiviral vector, sgRNA representations in the LICs and input reference (Day 3 NIH3T3 cells) were compared using MAGeCK RRA after normalising total read counts across different samples. Log2 fold change of the sgRNA representation between the leukaemia cells and input reference was also calculated using the MAGeCK RRA. As expected, the log2 fold change of the sgRNA representation in the LICs and LBs was correlated (Pearson correlation coefficient = 0.794) (Figure 3.16A). This showed good reproducibility of the screening protocol. According to the depletion score calculated by the MAGeCK pipeline, *Il7r* was the most depleted gene in the LICs, followed by *Rpa1*, the positive control gene (Figure 3.16B). For *Il7r*, 2 corresponding sgRNAs were significantly depleted (log2 fold change < -1), whereas no sgRNA was enriched (log2 fold change

> 1). Notably, the negative controls (*Rosa26*, etc.) were less depleted than *Il7r* and *Rpa1*, which showed the validity of the screening protocol (Figure 3.16C).

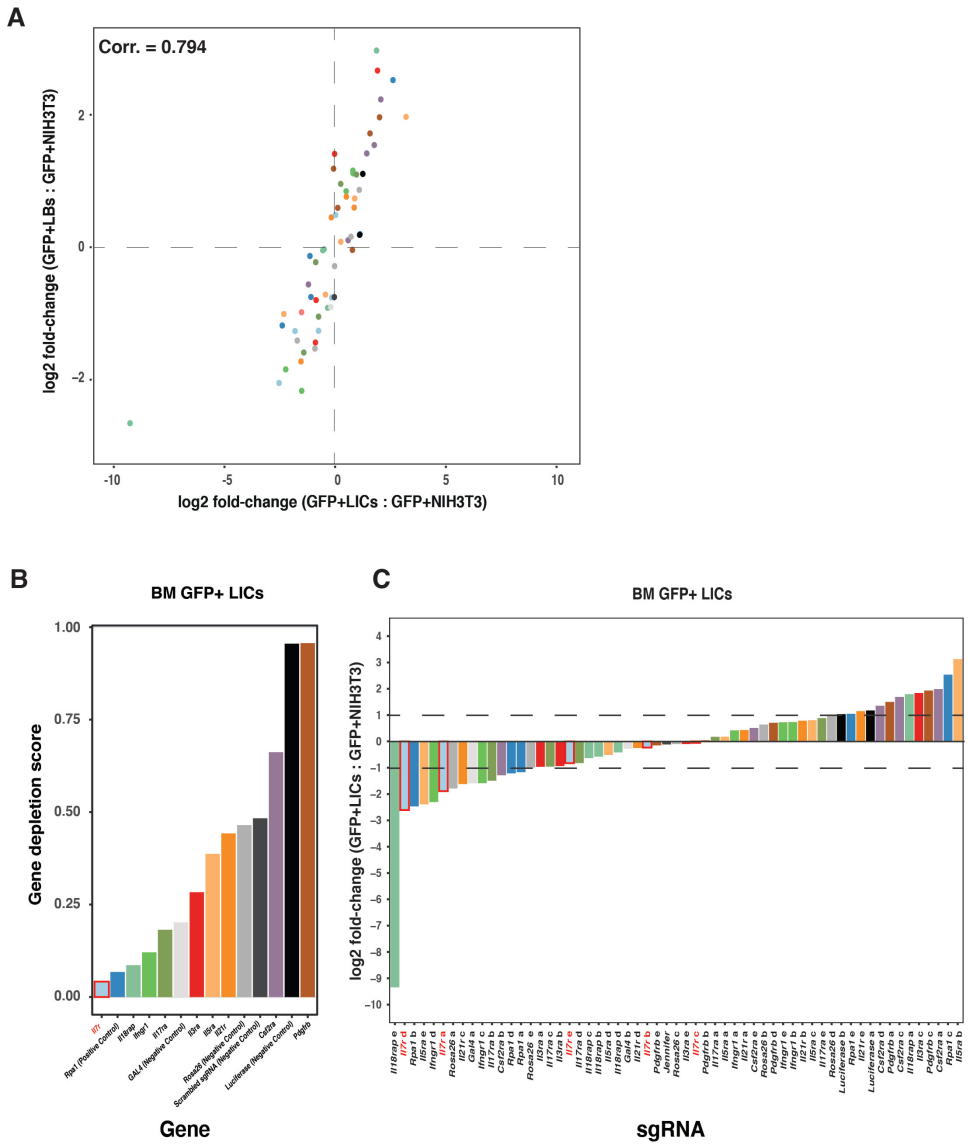

Figure 3.16.* Il7r *was identified as the top hit by MAGeCK RRA

(A) Scatterplot of log2 fold change of normalised sgRNA read counts (vs input control) in LICs and LBs for all sgRNAs in the screen. Only mice from the second infection experiment were included. The log2 fold change was calculated by MAGeCK RRA. Different colour dots indicate sgRNAs targeting different genes. Pearson correlation coefficient of log2 fold change of normalised sgRNA read counts (vs input control) between BM LICs and LBs is shown.
(B) Histogram showing the negative selection score and ranking (calculated by MAGeCK RRA) of the targeted genes included in the screen. Lower scores indicate a higher level of gene depletion in the BM LICs vs the NIH3T3 input control.
(C) Waterfall plot showing the log2 ratio (calculated by MAGeCK RRA) of sgRNA representation in the LICs versus NIH3T3 input control. The bars with red borders represent sgRNAs targeting Il7r.

Overexpression of *Il7r* in the LICs (28-fold increase, p-value = 0.0006) and LBs (4.5-fold increase, p-value = 0.0077) was confirmed by qRT-PCR (Figure 3.17A). The other cytokine receptor genes included in the screen were either overexpressed to a lesser extent or not overexpressed in the LICs at all, which might explain why LIC maintenance was not dependent on them as much as *Il7r* (Figure 3.17B-I). Unlike CCR1, whose high expression level was detected on the surface of ca. 10% of the LICs by flow cytometry, only ca. 0.1% of the LICs were IL7R positive when measured using flow cytometry (Figure 3.17J). This was consistent with the low $2^{-\Delta Ct}$ value of *Il7r* when measured using qRT-PCR (ca. 0.001, normalised to *B2m*) in the week 4 LICs, even though it was significantly higher than the healthy GMPs. For reference, this value for *Ccr1* was 0.012 in the week 4 LICs. Nonetheless, I decided to select *Il7r* from the current screen for further validation considering its strong depletion in LICs and significant overexpression in the *Cebpa* mutant AML cells (vs healthy GMPs).

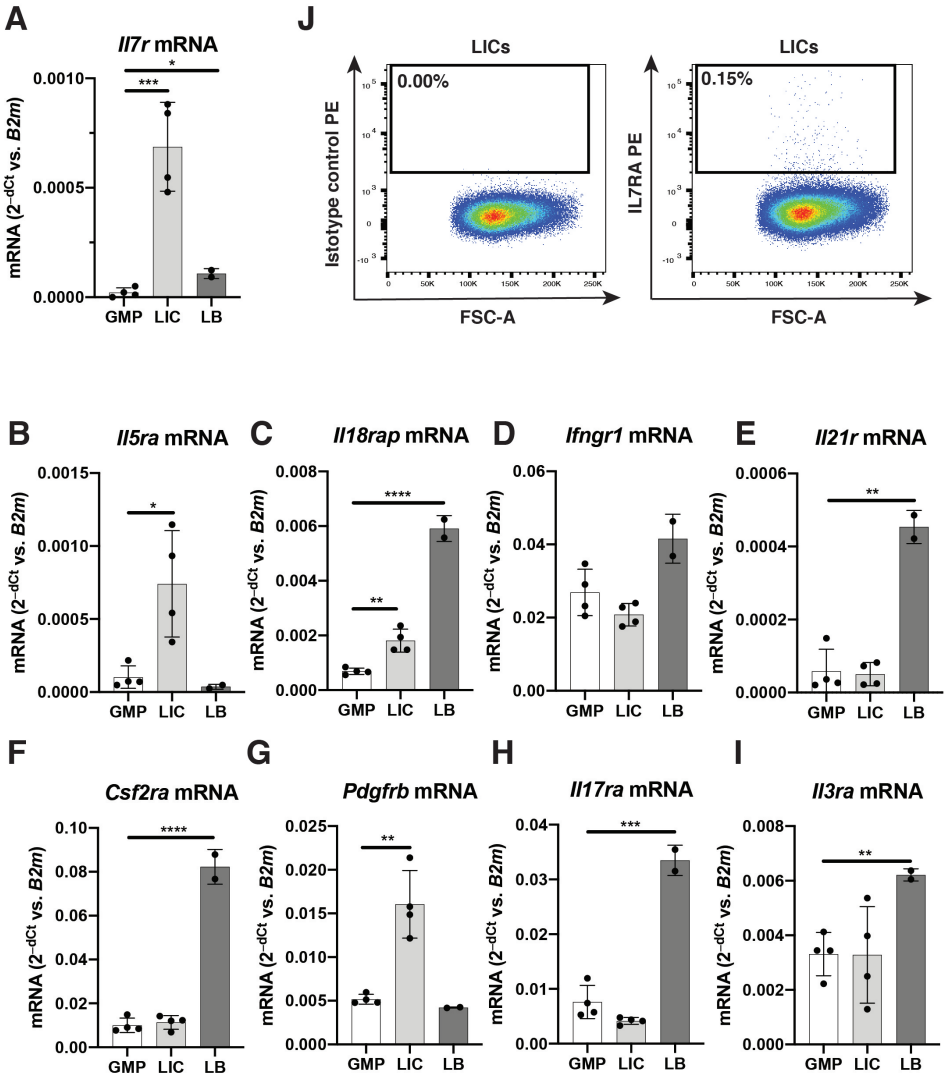

Figure 3.17.* Il7r *was highly overexpressed in the LICs

(A) Histogram showing Il7r mRNA expression levels in LICs (N=4) and LBs (N=2) isolated 4 weeks after LIC transplantation, compared to normal GMPs (N=4). Values shown are $2^{-\Delta Ct}$ normalised to B2m. Data from 1 experiment. Each biological replicate was the average of 2 technical replicates. P-values were calculated using the Student's t-test.
(B) Histogram as in (A) showing Il5ra mRNA expression levels.
(C) Histogram as in (A) showing Il18rap mRNA expression levels.
(D) Histogram as in (A) showing Ifngr1 mRNA expression levels.
(E) Histogram as in (A) showing Il21r mRNA expression levels.
(F) Histogram as in (A) showing Csf2ra mRNA expression levels.
(G) Histogram as in (A) showing Pdgfrb mRNA expression levels.
(H) Histogram as in (A) showing Il17ra mRNA expression levels.
(I) Histogram as in (A) showing Il3ra mRNA expression levels.
(J) Representative flow cytometry showing IL7RA expression on the BM LICs (right) 4 weeks after LIC transplantation. The positive gate was set based on isotype control antibody staining (left).

3.5 Discussion

Here, by conducting *in vivo* CRISPR screens in the strictly BM microenvironment-dependent murine model of AML, I identified 2 cytokine receptor genes (*Ccr1* and *Il7r*) as novel AML dependencies. To my knowledge, neither of them has previously been reported to function as AML dependencies. Several genome-wide CRISPR screens have been performed in human AML cell lines, but none identified CCR1 or IL7R as top AML dependencies [167, 168, 192, 199]. This can be partly explained by the fact that most of the AML cell lines are cytokine-independent, such as MOLM-13 and OCI-AML2, or only need a simple panel of cytokines to be supplemented, such as OCI-AML5 [167, 168]. Such simplified *in vitro* systems can hardly reflect the complex *in vivo* context where the interactions between AML cell surface receptors and the environmental factors exist and function. Additionally, it has been shown that CRISPR screens performed in the same cell line cultured in the presence or absence of a specific cytokine can identify different dependencies [168]. This further demonstrated the need for performing CRISPR screens *in vivo* so that leukaemogenesis can be more accurately modelled as it happens in patients [155].

A CRISPR screen focused on cell surface receptors was performed in an *MLL-AF9* AML mouse model and showed that CXCR4-signalling was essential for *MLL-AF9* AML [161]. It included a sgRNA library targeting 96 cell surface genes overexpressed in leukaemic GMPs (vs healthy GMPs). But it did not identify *Ccr1* or *Il7r* as top AML dependencies. This may be explained by several reasons. First, given that AML is a heterogeneous disease, *MLL-AF9* and *CEBPA* mutant leukaemia can have distinct biological features and rely on different cytokine/chemokine signallings [5]. This is supported by the fact that only 5 of the 96 genes overexpressed by the *MLL-AF9* leukaemia cells met the candidate gene selection criteria presented in this chapter [161]. The 5 genes were *Ccr1*, *Tgfbr1*, *Csf2ra*, *Il3ra*, and *Ifngr1*. Second, as previously mentioned, one key feature of primary human AML cells was the significant loss of leukaemia-propagating LSCs during *ex vivo* culture, suggesting them being sensitive to the deprivation of the native BM microenvironment. In contrast, the c-Kit-enriched *MLL-AF9* leukaemia cells used in the study of Ramakrishnan and colleagues remained highly transplantable after 2 sequential

transductions during up to 96 hours of *ex vivo* culture, where only IL-3, IL-6, and SCF were supplemented [161]. However, the *Cebpa* mutant LICs used in the current study had to be transplanted 8 hours after the transduction started otherwise the cells would lose their leukaemogenic potential significantly. This suggests that the *Cebpa* mutant LICs are more BM microenvironment-dependent than the *MLL-AF9* leukaemia cells, and our mouse model of AML can be a more appropriate context to identify essential cytokine receptors. Finally, even though CXCR4, the receptor for CXCL12, was identified as an essential gene for *MLL-AF9* AML development, its role was proved independent of CXCL12-CXCR4 interaction and may be attributed to constitutive activation of CXCR4 signalling [161]. This also questions the validity of using the *MLL-AF9* AML mouse model to study the interactions between AML cells and the autocrine/paracrine signals coming from the BM microenvironment.

One challenge hindering many *in vivo* CRISPR screens performed in AML mouse models is the limitation on the number of cells that can be transplanted and successfully engrafted, which is vital for a high sgRNA library representation and the success of the screen [151]. To increase the robustness of the screen and avoid false-positive hits due to random depletion of sgRNAs, our *Cebpa* mutant AML mouse model was further developed by using lethal irradiation. This allowed more LICs to be transplanted and potentially engrafted per mouse, leading to a higher representation of a given sgRNA library. Despite this, to keep the kinetics of AML development and biological features similar to the original sub-lethally irradiated model, no more than 200,000 c-Kit-enriched LICs were transplanted into each recipient mouse. In addition, to decrease the chance of delivering multiple sgRNAs into each cell, the intended transduction rate was kept at 20-40% [141, 157]. Therefore, in an ideal scenario where all the LICs transplanted successfully engrafted, each mouse would only harbour 40,000-80,000 LICs carrying sgRNAs. This demonstrated the necessity to limit the sgRNA library size in order to reach a high sgRNA representation [151]. As a result, 2 independent mini-pool libraries of no more than 60 sgRNA were used. Since multiple sgRNAs are usually designed for each gene to decrease the chance of picking up false-positive hits caused by off-target effects [151], only a limited number of candidate genes were studied at each time, and strict candidate gene selection criteria were employed.

As the actual number of transduced LICs that engrafted was unknown, it was difficult to estimate the actual coverage of the sgRNA in each mouse and ensure that the current sgRNA library size was appropriate. It would be helpful to perform an experiment by transducing the LICs with a sgRNA library constructed by the same lentiviral constructs encoding 60 neutral sgRNAs. If the 60 sgRNAs could all be recovered and a balanced distribution of the sgRNAs could be observed at the endpoint, the relevant screening condition would be deemed optimal [162]. However, such an optimisation experiment was not performed, and I have unfortunately not got solid evidence to show that the sgRNA library size is optimal. In spite of the above efforts to maximise the coverage of the sgRNA library in each mouse, some of the infection replicates still showed a strongly biased distribution of the sgRNA library. This was probably due to the unpredictable occurrence of insufficient sgRNA library coverage, which can be attributed to varying engraftment efficiency and transducibility of LICs from batch to batch. This again reflects the complexity of performing *in vivo* CRISPR screens by transplanting primary cells [155]. However, in the mouse replicates where a less biased distribution of the sgRNA library was observed, the robustness of the screen was confirmed by introducing the positive and negative control sgRNAs.

Because of the limited number of LICs that can be transplanted per mouse and the need for a low transduction efficiency, while introducing the sgRNA library into the LICs, it would be rather challenging and labour-demanding to perform screens involving larger sgRNA libraries or even a genome-wide screen *in vivo*. An alternative to splitting one sgRNA library into several smaller ones would be pooling different mice transplanted with the transduced LICs together [151]. This can be done by either pooling sgRNA reads from different mice together computationally or pooling gDNA/sorted cell samples from different mice together physically [156, 159, 162, 164]. It would be interesting to design one single lentiviral library comprising sgRNAs targeting all the candidate genes included in the 2 screens above and pool GFP+LICs from multiple mice as one biological replicate to investigate whether *Ccr1* and *Il7r* remain to be the top hits. Besides, it should be noteworthy that the BM stromal cells are more severely damaged in the lethally-irradiated mouse model than in the original sub-lethal irradiated mode (with only 25,000 LICs

transplanted), where damage to BM stromal cells is minimised [200, 201]. This difference in the environment where the *Cebpa* mutant LICs proliferate may alter their gene expression profile and their interaction with environmental factors. It would also be interesting to perform the screens in the sub-lethal irradiated mouse model, where the molecular profiling of LICs and BM stromal cells was initially done. This, of course, requires more mice to be transplanted with the transduced LICs and pooled at the endpoint to reach a reasonable representation of the sgRNA library.

Chapter 4

Validation of CCR1 and IL7R as dependencies in murine *Cebpa* mutant AML cells

4.1 Introduction and aims

Having identified *Ccr1* and *Il7r* as potential AML dependencies, I intended to individually validate their essential roles in AML development in this chapter. CCR1 (CC chemokine receptor type 1) is expressed in various human haematopoietic cells, including CD34+ BM progenitor cells, monocytes, T lymphocytes and natural killer cells [202]. Using CCR1 deficient mice, CCR1 has been implicated in the proliferation and migration of myeloid progenitor cells and neutrophil chemotaxis [203, 204]. CCR1 can interact with multiple chemokines such as CCL3, CCL5, CCL15 (CCL9 in mice) and CCL23 (CCL6 in mice) [205]. Among them, CCL3 is believed to be the most potent activator of CCR1 [198].

Immunohistologic analysis of human haematological malignancy samples also identified CCR1 expression in AML, MM, and lymphomas [206]; CCL3 levels have been found to be increased in the BM of AML patients [90], MM patients [207], and MPN mice [91]; elevated levels of CCL23 have been observed in both AML and MDS patients [208]. This is in line with our research investigating the *Cebpa* mutant AML mouse model, where upregulated gene expression of *Ccr1* and the genes encoding its core ligand was observed in the leukaemia cell population using bulk RNAseq

and qRT-PCR as shown in Chapter 3. There are several pieces of evidence that the CCL3-CCR1 signalling could promote haematological malignancy development. First, the expression of CCR1 and CCL3 has been shown to be essential for CML LIC maintenance, and CCR1 expression was associated with MM tumour dissemination [197, 198]. In addition, CCL3 derived from AML cells and MM plasma cells can impair the function of and decrease the number of OBs, leading to bone loss and disrupting the niche needed for normal haematopoiesis [90, 207]. Furthermore, CCL3 produced by the BM microenvironment can initiate MPN, possibly by generating a pro-inflammatory environment [81]. CCL3 in the AML environment can also suppress normal erythropoiesis [97]. However, there has not yet been clear evidence that intrinsic CCR1 signalling is essential for and directly involved in AML LIC maintenance.

The *IL7R* gene encodes the IL-7 receptor (IL-7R) alpha-chain (CD127), which forms the heterodimeric IL-7R with the common cytokine receptor gamma-chain encoded by the *IL2RG* gene [209]. CD127 expression has been found to be mainly in lymphoid progenitors and mature lymphoid cells, including T cells, pre-B cells and innate lymphoid cells [209, 210]. Upon IL-7 binding to the receptor, IL-7R and the downstream JAK-STAT signalling pathway can promote CD8+ and CD4+ memory T cell homeostasis; in mice, IL-7R signalling also contributes to B cell development [211]. Additionally, IL-7R signalling has been implicated in the oncogenesis of T-cell acute lymphoblastic leukaemia (T-ALL) [212]; IL7R gain-of-function mutation has been shown to be able to initiate B-cell acute lymphoblastic leukaemia (B-ALL) [213]. However, in myeloid cells, the function of CD127 has not been well elucidated yet. CD127 expression was observed in human monocytes many years ago. A recent study demonstrated that human inflammatory diseases such as COVID-19 and rheumatoid arthritis could induce high-level CD127 expression in a subset of inflammatory monocytes/macrophages [210]. Therefore, the upregulation of *Il7r* expression in the murine *Cebpa* LICs, as shown in Chapter 3, may be attributed to the pro-inflammatory BM microenvironment induced by AML cell engraftment. Potentially supporting this, a previous study demonstrated that AML cells could proliferate in response to IL-7, which robustly indicates that the biological significance of IL-7R signalling is not restricted to the lymphoid lineage and lymphocytic

leukaemias [214]. Further studies are still needed to understand the role of IL-7R signalling in AML LIC maintenance.

4.2 Investigating on-target editing efficiencies of the selected *Ccr1*- and *Il7r*-targeting sgRNAs

In order to investigate editing efficiencies of the *Ccr1*- and *Il7r*-targeting sgRNAs that were significantly depleted (p-value < 0.05, log2 fold change < -1) in the previous screens, WT NIH3T3 cells were transduced with the previously described GFP-expressing lentiviral constructs encoding the relevant sgRNAs. These sgRNAs had no adverse effect on the survival of the WT NIH3T3 cells, as discussed in Chapter 3. The NIH3T3 cells were incubated at 37°C for 72 hours post-transduction before GFP+ cell isolation and gDNA extraction. I have confirmed that WT NIH3T3 cells share the same sequences with primary mouse BM cells at the *Ccr1* or *Il7r* loci where the relevant sgRNAs bind. This allows base pairing between the NIH3T3 DNA and the sgRNAs used in the CRISPR screen.

Previous studies demonstrated that close to complete genomic modification required approximately 7-10 days across different gene targets [150, 152, 215], including HSPCs [174]. Therefore, NIH3T3 cells were used for indel analysis since they were much easier to be kept for a longer period of time after the transduction started. Conversely, the c-Kit-enriched primary *Cebpa* LICs can only be maintained *ex vivo* for up to 48 hours without significant loss of viability, even when not transduced [56]. This hinders the direct use of *Cebpa* LICs for investigating on-target editing efficiencies of the sgRNA. On the other hand, the use of NIH3T3 cells enables a large number of transduced cells to be harvested for more comprehensive indel analysis and allows gene editing to happen sufficiently due to the longer *in vitro* culture period [150, 216, 217].

Importantly, the on-target activity (indel profile) and specificity of sgRNAs rely on their sequences and the sequence of the target site, which has been shown to be nonrandom and relatively consistent across different cell types [142, 143, 148, 149, 199, 215, 218]. On the other hand, sgRNA efficacy depends less on local accessibility to the Cas9/sgRNA complex, epigenetic factors, or genomic context, which are relatively variable across cell types [142, 152, 216], even though local chromatin

accessibility does influence the binding of Cas9 to DNA [147, 219]. This greatly facilitates predicting the editing efficiency of each sgRNA of interest in the c-Kit-enriched primary *Cebpa* LICs by using the more malleable NIH3T3 cells, as indicated by van Overbeek and colleagues [142].

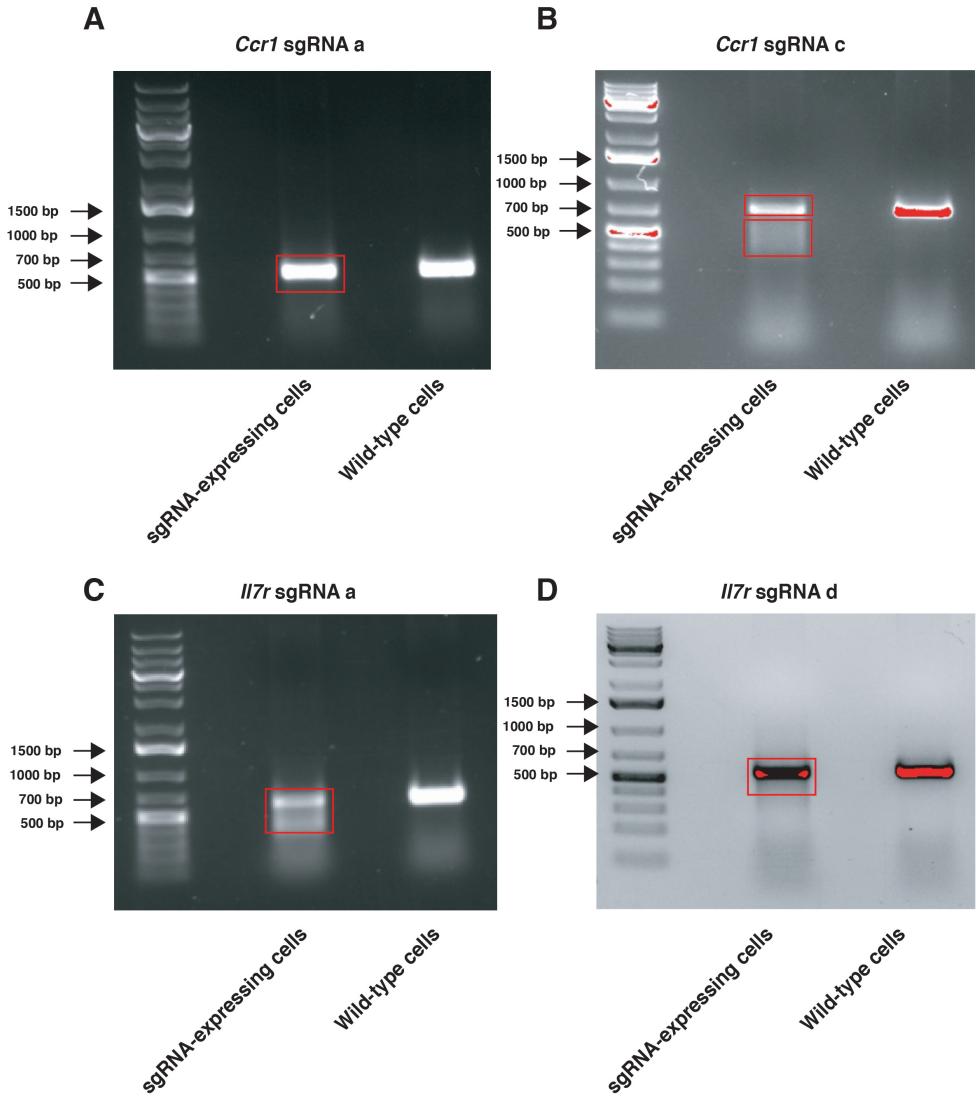

Figure 4.1. On-target modification of NIH3T3 cells using the selected* Ccr1- *and* Il7r*-targeting sgRNAs

(A) PCR gel showing amplification of the Ccr1 sgRNA a-targeted loci before (right) and after (left) introducing Ccr1 sgRNA a to WT NIH3T3 cells.
(B) PCR gel as in (A) showing amplification of the Ccr1 sgRNA c-targeted loci.
(C) PCR gel as in (A) showing amplification of the Il7r sgRNA a-targeted loci.
(D) PCR gel as in (A) showing amplification of the Il7r sgRNA d-targeted loci.

The genomic region flanking each sgRNA target locus was PCR amplified (the primer binding sites were at least 200 bases upstream or downstream of the sgRNA

target locus), followed by gel purification. See Table 2.8 for primer sequences for each sgRNA that was studied. Purified PCR products were sequenced by Sanger sequencing and analysed by the ICE analysis [220]. 2 bands with distinct sizes were seen after amplifying the target site of *Ccr1*-targeting sgRNA c. Hence, the upper and lower bands were gel purified, sequenced, and analysed separately (Figure 4.1B). The ICE analysis revealed that the upper band was mainly composed of *Ccr1* loci that were unedited, with an indel frequency of only around 15% (Figure 4.2 and Figure 4.3B). For the lower band, an indel frequency of 63% was observed (Figure 4.2 and Figure 4.3C). For the other guides, only one gel cut was purified and sequenced (Figure 4.1A, C-D). For *Ccr1*-targeting sgRNA a (Figures 4.2 and 4.3A), *Il7r*-targeting sgRNA a (Figures 4.2 and reffig:4.3D), and *Il7r*-targeting sgRNA d (Figures 4.2 and 4.3E), an editing efficiency of 43-80% was achieved. This demonstrated that all the significantly depleted guides selected by the previous screens were functional.

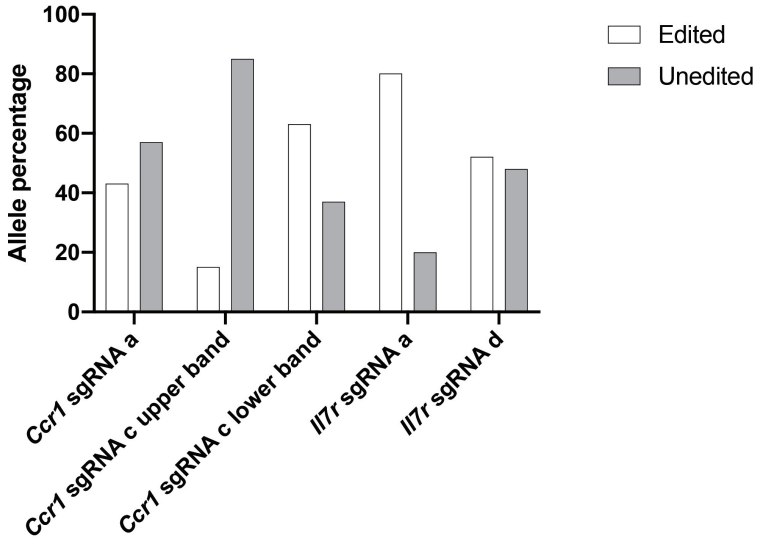

Figure 4.2. ICE analysis of indels for the selected* Ccr1- *and* Il7r-*targeting sgRNAs in NIH3T3 cells

Histogram showing on-target editing efficiencies of the selected Ccr1- and Il7r-targeting sgRNAs. To assess indels at targeted genomic regions, sgRNA-expressing GFP+ NIH3T3 cells were purified by FACS 72 hours after transduction. Non-edited NIH3T3 cells were used as control. After gDNA purification, targeted loci were amplified by PCR, followed by gel purification of the amplicons. Indel frequencies within individual amplicons (as in Figure 4.1) were determined by Sanger sequencing, followed by data analysis using Synthego's ICE tool.

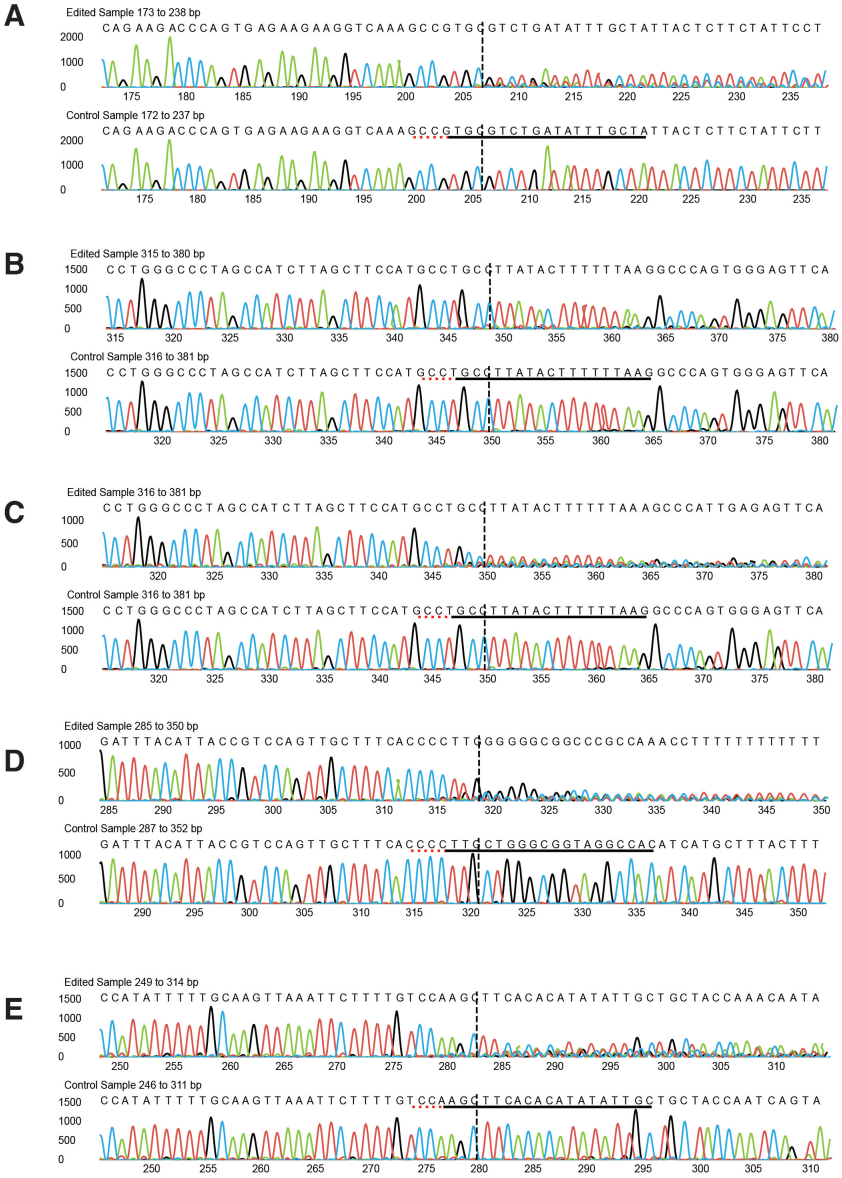

Figure 4.3. Schematic showing Sanger sequencing of the sgRNA-targeted loci

(A) Sanger sequencing of Ccr1 loci in NIH3T3 cells after lentiviral transduction of Ccr1 sgRNA a (vs without transduction). PCR product for sequencing was purified from Figure 4.1A.

(B) Sanger sequencing as in (A) for Ccr1 sgRNA c. PCR product for sequencing was purified from Figure 4.1B, the upper band.

(C) Sanger sequencing as in (A) for Ccr1 sgRNA c. PCR product for sequencing was purified from Figure 4.1B, the lower band.

(D) Sanger sequencing as in (A) for Il7r sgRNA a. PCR product for sequencing was purified from Figure 4.1C.

(E) Sanger sequencing as in (A) for Il7r sgRNA d. PCR product for sequencing was purified from Figure 4.1D.

Of note, the majority of NIH3T3 cells are triploid (26.02%) or tetraploid (60%), while *Cebpa* mutant LICs are diploid [13, 196]. As genome editing at each al-

lele happens independently in cells, it is conceivable that editing all alleles in each NIH3T3 cell is more difficult than in each *Cebpa* mutant LIC. As a result, more unedited/WT loci would remain in the NIH3T3 cells compared with the diploid *Cebpa* mutant LICs, even when the same sgRNAs are applied. In addition, previous studies showed that genomic cleavage would continue to increase beyond 72 hours post-transduction [216], especially when the sgRNA and Cas9 were constitutively expressed as in the current study [142]. Altogether, the on-target editing efficiency of the *Ccr1*- and *Il7r*- targeting sgRNAs should be higher in the diploid *Cebpa* mutant LICs that would stay in the recipient mice for over 3 weeks than in the WT NIH3T3 cells used for analysis here. Another murine cell line, which is diploid and originates from haematopoietic cells, may have been more similar to the *Cebpa* mutant LICs and more appropriate for indel analysis.

4.3 Knocking out *Ccr1* or *Il7r* depleted the murine *Cebpa* mutant LICs

After confirming the on-target functionality of the selected *Ccr1* and *Il7r* sgRNAs, I sought to validate *Ccr1* and *Il7r* as *in vivo* dependencies in AML individually. For each gene, the sgRNA with higher on-target editing efficiency based on the previous indel analysis (*Ccr1* sgRNA a and *Il7r* sgRNA a) was used for the *in vivo* validation experiment. A sgRNA targeting *Rosa26* (*Rosa26* sgRNA c) was used as the negative control. The schematic representation of the validation experiment can be found in Figure 4.4.

To study whether *Ccr1* or *Il7r* disruption could affect the maintenance of the *Cebpa* mutant LICs, c-Kit-enriched LICs were transduced with GFP-expressing lentiviral vectors encoding sgRNAs targeting *Ccr1* or *Il7r*. After an 8-hour transduction, 200,000 infected c-Kit-enriched LICs were co-transplanted into lethally irradiated recipient mice together with 1 million healthy competitor BM cells. In the meantime, the c-Kit-enriched LICs were transduced with the same sgRNA-expressing lentiviral vector targeting *Rosa26* and transplanted in the same manner for control. In parallel, an aliquot of cells under each transduction condition was kept *ex vivo* for

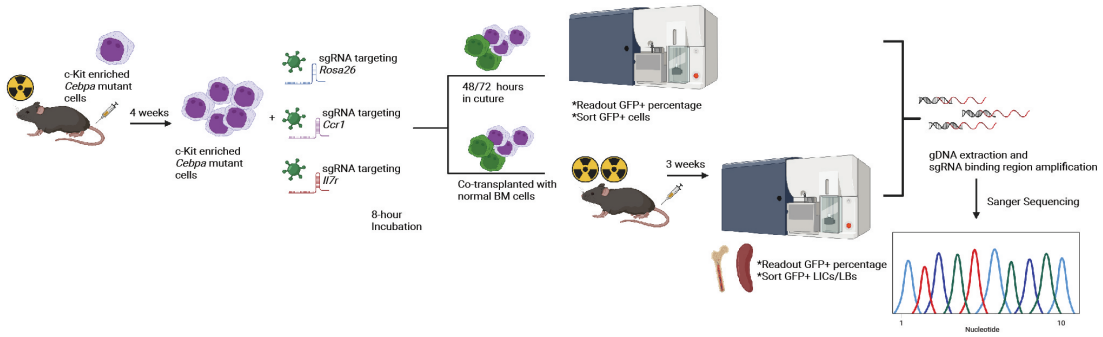

Figure 4.4. Schematic of* in vivo *validation of CCR1 and IL7R as AML dependencies

Freshly isolated LICs were transduced with EGFP-expressing lentiviral vectors encoding Ccr1 sgRNA a, Il7r sgRNA a, or Rosa26 sgRNA c and transplanted into recipient mice (N = 6, 2 infection experiments, 200,000 cells per mouse) after an 8-hour incubation. An aliquot of LICs was kept ex vivo to readout transduction efficiency 48 hours and 72 hours (as input GFP+ percentage) after transduction by flow cytometry. 3 weeks after transplantation, mouse BM and spleen cells were collected for evaluating the endpoint GFP+ percentage. When the mice were sacrificed, BM GFP+ LICs and LBs were also FACS sorted for indel analysis.

up to 72 hours to read out the GFP+ percentage, which was used as the input value for downstream analysis. 2 independent infection experiments were performed. Of note, the cells transduced with the *Ccr1* and *Il7r* sgRNA consistently had a higher 72-hour GFP+ percentage than those transduced with the *Rosa26* sgRNA (Figure 4.5B). During the second infection experiment, the GFP+ percentages were also determined at 48 hours post-transduction, where a similar difference in the GFP+ percentage across transduction conditions was observed (Figure 4.5A). Interestingly, the ratio of the GFP+ percentage in the cells transduced with the *Ccr1* or *Il7r* sgRNA to that in the cells transduced with the *Rosa26* sgRNA dropped from 48 hours (1.63 for *Ccr1*; 1.53 for *Il7r*) to 72 hours (1.25 for *Ccr1*; 1.18 for *Il7r*). This may indicate that increased *Ccr1* or *Il7r* disruption negatively impacts the *ex vivo* survival of the LICs, mirroring the screen result in Chapter 3.

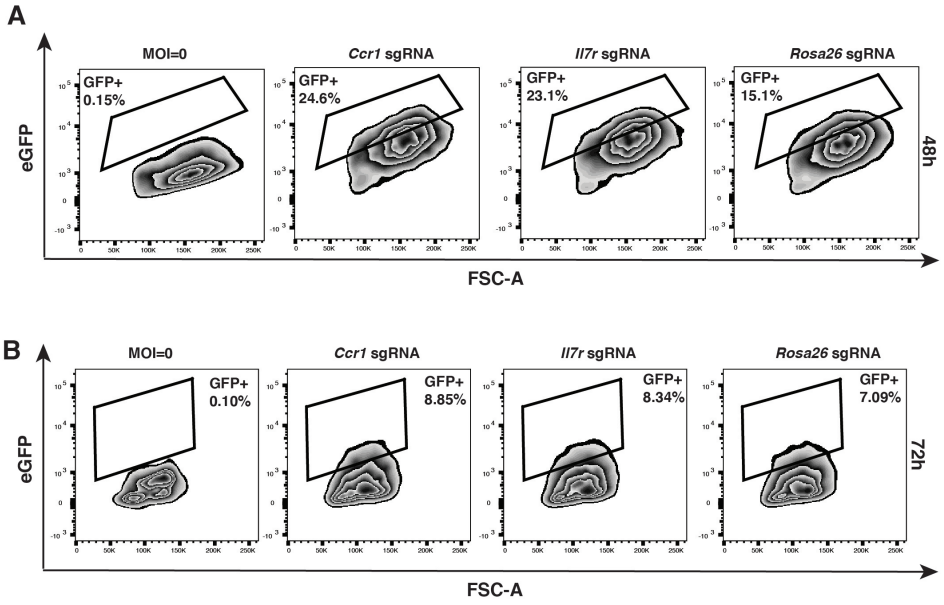

Figure 4.5. Representative flow cytometry showing percentages of GFP+ AML cells 48/72 hours post-transduction

(A) Representative flow cytometry of sgRNA-expressing GFP+ leukaemia cells at 48 hours post-transduction. The GFP+ gate (on 7AAD-Singlets) was set based on MOI=0 control (leftmost).
(B) Representative flow cytometry of sgRNA-expressing GFP+ leukaemia cells at 72 hours post-transduction. The GFP+ gate (on 7AAD-Singlets) was set based on MOI=0 control (leftmost).

At 3 weeks post-transplantation, all the recipient mice were culled to collect the BM and spleen. The endpoint percentage of GFP+ cells within the whole CD45.2+ leukaemia cell population was determined using flow cytometry and normalised to the 72-hour input GFP+ percentage (Figure 4.6A). This compensated for the difference in initial transduction efficiency across lentiviral constructs encoding different sgRNAs. In accordance with the results from the screens, a significant reduction in the GFP+ CD45.2+ leukaemia cells was observed in the BM and spleen after *Ccr1* and *Il7r* knockout (Figure 4.6B-C).

Not surprisingly, the GFP+ CD45.2+ leukaemia cells expressing the *Ccr1*- or *Il7r*-targeting sgRNA were not completely depleted in the BM when the mice reached the endpoint. To investigate why the GFP+ cells expressing the *Ccr1*- or *Il7r*-targeting sgRNAs were not eradicated, the GFP+ LICs and LBs were isolated by FACS; those expressing the *Rosa26*-targeting sgRNA were similarly isolated for control. After gDNA purification and PCR amplification of the *Ccr1* or *Il7r* sgRNA target

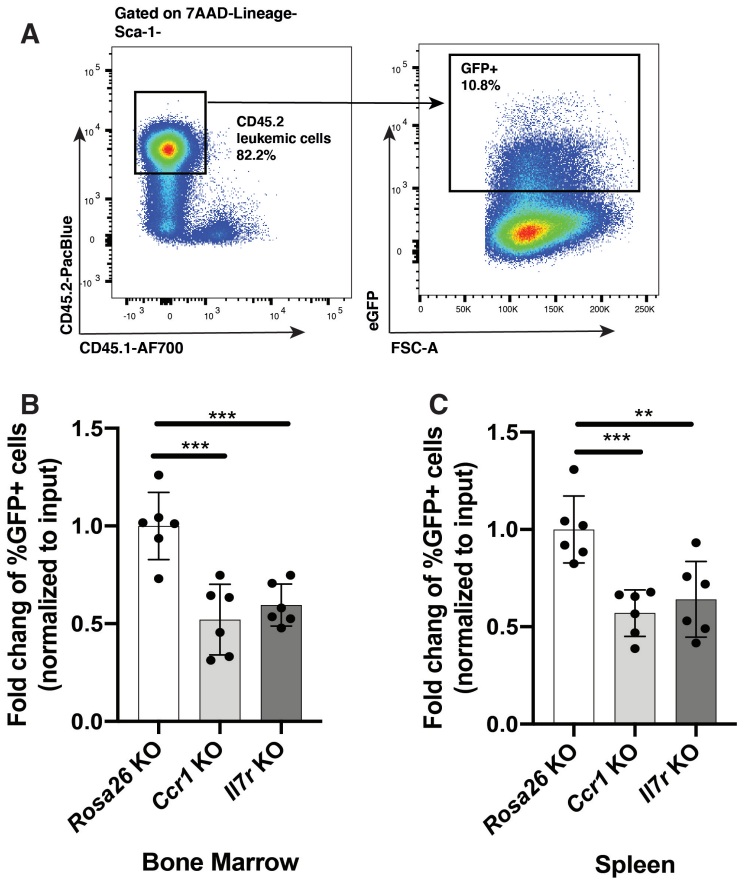

Figure 4.6. CCR1 and IL7R were* in vivo *dependencies of* Cebpa *mutant LICs
(A) Flow cytometry gating strategy used to identify sgRNA-expressing GFP+ leukaemia cells within the BM and spleen. The GFP+ gate was set based on MOI=0 control.
(B) Histogram showing the fraction of GFP+ cells within BM CD45.2 leukaemia cells 3 weeks after LIC transplantation. For each transduction condition, the endpoint GFP+ percentage was first normalised to the input GFP+ percentage (determined at 72 hours post-transduction as in Figure 4.5) and then normalised to the average GFP+ percentage among all the Rosa26 knockout mice (Rosa26 KO group = 1). P-values were calculated using Student's t-test.
(C) Histogram as in (B) showing the fraction of GFP+ cells within spleen CD45.2 leukaemia cells.

locus, the samples were analysed using the ICE analysis. A certain level of indels were observed in the leukaemia cells expressing the *Ccr1*-targeting (at *Ccr1* locus) and *Il7*-targeting (at *Il7r* locus) sgRNA. However, indels were rarely detected in the *Rosa26* sgRNA-expressing LICs and LBs (Figure 4.7A-B). This confirmed the activity and specificity of the sgRNAs used during the experiment. Notably, for both the *Ccr1*- and *Il7r*-knockout group, the majority of the GFP+ LICs that remained in the BM at the endpoint didn not retain the edited locus: WT *Ccr1* loci were observed in over 80% of the *Ccr1* sgRNA-expressing LICs; WT *Il7r* loci

were observed in approximately 95% of the *Il7r* sgRNA-expressing LICs (Figure 4.7A-B). Even though more edited loci were observed in the GFP+ LBs (Figure 4.7A-B), unedited loci in the GFP+ LBs were still significantly enriched during leukaemogenesis, given the baseline sgRNA editing efficiencies observed in the near-tetraploid NIH3T3 cells as shown in Figure 4.2. This indicated that the leukaemia cells harbouring WT *Ccr1* and *Il7r* loci outcompeted the edited cells. Therefore, it is likely that the GFP+ LICs in which *Ccr1* or *Il7r* was successfully disrupted were depleted, whereas the GFP+ LICs that somehow escaped the *Ccr1* or *Il7r* editing were able to survive, continue to proliferate and contribute to the GFP+ leukaemic population at the endpoint. These findings suggested the essential role of CCR1 and IL7R for LIC maintenance *in vivo*.

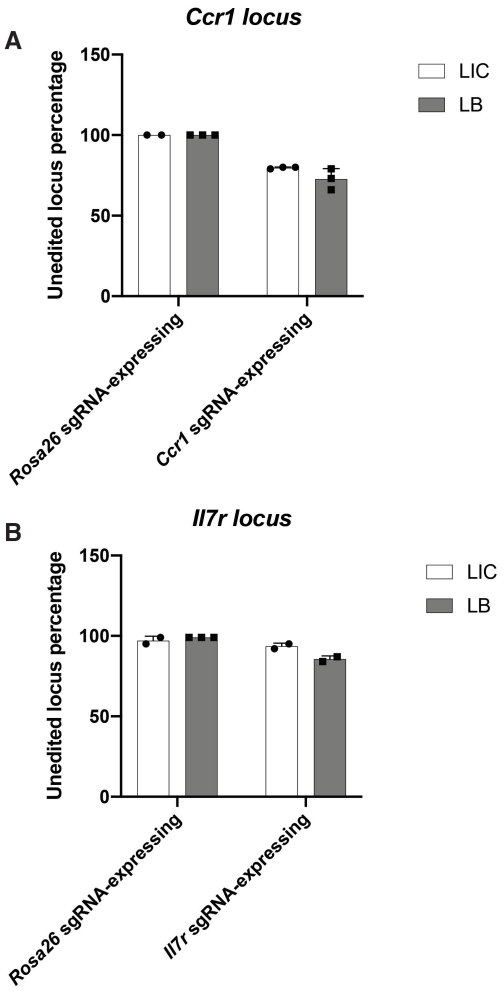

Figure 4.7. ICE analysis of the* Ccr1 *and* Il7r *loci in the sgRNA-expressing BM leukaemia cells

(A) Histogram showing percentages of unedited Ccr1 loci in the BM LICs (white bar) and LBs (grey bar) expressing the Rosa26-targeting (left) or Ccr1-targeting (right) sgRNAs. GFP+ sgRNA-expressing cells were FACS sorted after the mice were sacrificed, as in Figure 4.4. WT murine BM cells were used as control.
(B) Histogram as in (A) for unedited Il7r loci in leukaemia cells expressing the Rosa26-targeting (left) or Il7r-targeting (right) sgRNAs.

4.4 High *CCR1* expression was associated with poor prognosis in human AML

To gain insight into whether CCR1 plays a role in human AML development and can be potentially used as a therapeutic target for AML patients, RNAseq data from the TCGA AML samples and the GTEx healthy human BM tissues were retrieved and analysed [14, 186]. First, the expression level of *CCR1* and the genes encoding key CCR1 ligands, *CCL3* and *CCL23*, were examined. This revealed significant overexpression of *CCR1*, *CCL3*, and *CCL23* in human AML (Figure 4.8A). Significantly, elevated *CCR1* expression was associated with an adverse prognosis in AML patients (Figure 4.8B). Further genetic stratification of the AML samples showed that *CCR1*, *CCL3*, and *CCL23* overexpression occurred across different cytogenetic AML subtypes and in bi-allelic *CEBPA* mutant AML (Figure 4.8C-D). PCA analysis of the RNAseq data using the top 500 most variable genes showed that bi-allelic *CEBPA* mutant leukaemias expressing high (*CCR1* expression within the top 50% among all the TCGA AML samples) and low levels of *CCR1* (*CCR1* expression within the bottom 50% among all the TCGA AML samples) clustered separately (Figure 4.8E). GSEA of *CCR1* high- and low-expressing bi-allelic *CEBPA* mutant leukaemias revealed that expression of genes associated with cytokine and chemokine signalling was enriched in $CCR1^{\text{hi}}$ compared to $CCR1^{\text{lo}}$ bi-allelic *CEBPA* mutant leukaemias (Figure 4.8F). Since the number of DEGs (adjusted p-value $<$ 0.05) between $CCR1^{\text{hi}}$ and $CCR1^{\text{lo}}$ bi-allelic *CEBPA* mutant leukaemias were low (only 83), GSEA was only performed without pre-ranking.

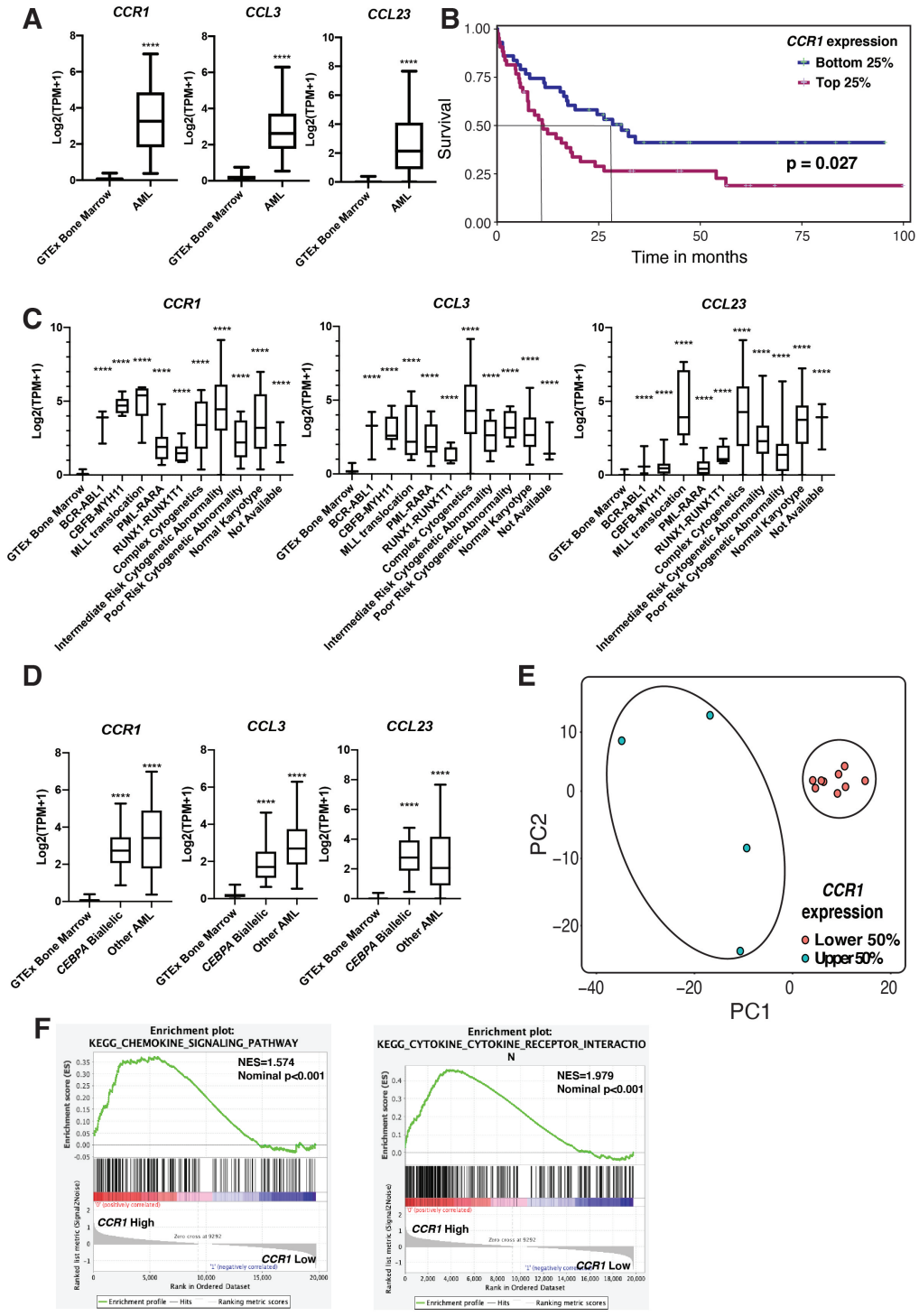

Figure 4.8. Increased CCR1 signalling is conserved between murine and TCGA human CEBPA mutant AML

(A) Box plots comparing CCR1, CCL3, and CCL23 mRNA expression (as log2(TPM+1)) between AML samples (TCGA-LAML RNAseq dataset) and normal BM samples (GTEx RNAseq dataset). P-values were calculated using Student's t-test.
(B) Survival data from the TCGA-LAML cohort for the patients with the highest and lowest quartiles of CCR1 mRNA expression (RSEM value). The p-value was calculated using the log-rank test.

(C) Box plots as in (A) comparing CCR1, CCL3, and CCL23 mRNA expression between AML samples separated by cytogenetic subgroup and normal BM samples.
(D) Box plots as in (A) comparing CCR1, CCL3, and CCL23 mRNA expression between AML samples grouped based on the presence or absence of bi-allelic CEBPA mutations and normal BM samples.
(E) PCA of RNAseq data from CEBPA mutant TCGA-LAML samples using the top 500 most variable genes. The CEBPA mutant AML samples were segregated based on CCR1 mRNA expression determined by the RSEM value (upper vs lower 50% among all the TCGA-LAML patients). The ovals were manually drawn to encompass the populations indicated.
(F) GSEA comparing CCR1-high vs CCR1-low CEBPA mutant TCGA-LAML samples using the KEGG chemokine (left) and cytokine (right) signalling pathway gene sets. Normalised enrichment score (NES) and p-value are shown.

To further support these findings, another publicly available human AML dataset (Beat AML) was similarly investigated [188]. When RNAseq data from the Beat AML *de novo* AML patients and healthy donors were analysed, overexpression of *CCR1*, *CCL3*, and *CCL23* was also observed across different subtypes of leukaemias, including the bi-allelic *CEBPA* mutant leukaemias (Figure 4.9A-B). The 6 bi-allelic *CEBPA* mutant leukaemia samples were grouped by their *CCR1* expression levels: the $CCR1^{hi}$ group included 4 AML samples; the $CCR1^{lo}$ group included 2 AML samples, of which the *CCR1* expression level was comparable to the control (Figure 4.9C). PCA analysis of the RNAseq data using the top 500 most variable genes demonstrated distinct clustering of the $CCR1^{hi}$ and $CCR1^{lo}$ bi-allelic *CEBPA* mutant leukaemias as well (Figure 4.9D). Furthermore, in agreement with the result from the TCGA AML samples, GSEA (without pre-ranking) of the Beat AML $CCR1^{hi}$ and $CCR1^{lo}$ bi-allelic *CEBPA* mutant leukaemias revealed enrichment of the cytokine and chemokine signalling genes in the $CCR1^{hi}$ leukaemias (Figure 4.9E).

To summarise, consistent with the finding in the murine *Cebpa* mutant AML model, the analysis of human AML datasets suggests that CCR1 signalling, after being activated by the environmental factors, directly contributes to the distinct molecular phenotype and adverse prognosis observed in *CCR1* high-expressing AML patients. Increased activation of intrinsic CCR1 signalling is likely a feature conserved between murine and human *CEBPA* mutant AML. Additionally, inhibiting CCR1 signalling in AML patients could potentially target the AML cells and decrease the BM leukaemic burden by disrupting the AML cells' interactions with the leukaemia-

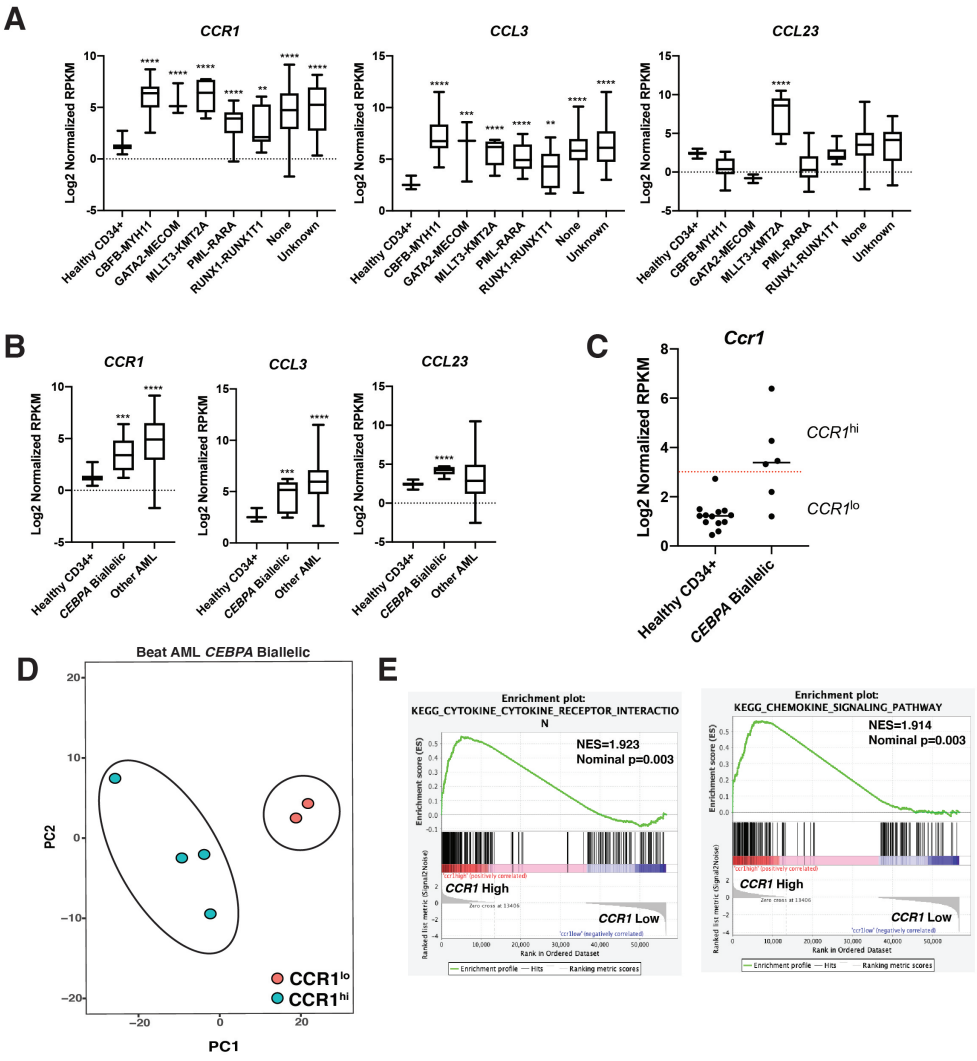

Figure 4.9. Increased CCR1 signalling is conserved between murine and Beat AML human* CEBPA *mutant AML

(A) Box plots comparing CCR1, CCL3, and CCL23 mRNA expression (as log2 normalised RPKM)) between Beat AML samples and healthy BM samples (CD34+ normal HSPCs). Beat AML samples were separated by cytogenetic subgroup. P-values were calculated using Student's t-test.

(B) Box plots as in (A) comparing CCR1, CCL3, and CCL23 mRNA expression between Beat AML samples grouped based on the presence or absence of bi-allelic CEBPA mutations and healthy BM samples.

(C) Scatter plot showing segregation of $CCR1^{hi}$ and $CCR1^{lo}$ CEBPA mutant Beat AML samples.

(D) PCA of RNAseq data from CEBPA mutant Beat AML samples using the top 500 most variable genes. The CEBPA mutant AML samples were segregated based on CCR1 mRNA expression as in (C). The ovals were manually drawn to encompass the populations indicated.

(E) GSEA comparing CCR1-high vs CCR1-low CEBPA mutant Beat AML samples using the KEGG chemokine (right) and cytokine (left) signalling pathway gene sets. Normalised enrichment score (NES) and p-value are shown.

supporting paracrine/autocrine signals.

4.5 CCR1 ligand overexpression was not able to promote *Cebpa* mutant LIC propagation

Based on the previous findings, I hypothesised that activating CCR1 and its downstream signalling might induce *in vivo* AML progression and promote *ex vivo Cebpa* mutant LIC survival. The Nerlov laboratory previously transplanted *Cebpa* mutant LICs into non-conditioned young and old recipient mice. We observed that LICs were engrafted and propagated efficiently in aged BM. However, little or no leukaemia engraftment was detected in the unperturbed young BM microenvironment [56]. This allowed me to use the non-conditioned young BM microenvironment as an unperturbed control and investigate whether the external provision of specific cytokines or chemokines, such as CCR1 ligands, could facilitate LIC engraftment in the otherwise not AML-propagating environment.

The hydrodynamic injection is a method for *in vivo* gene transfer performed by rapidly injecting a large volume of plasmid DNA solution into the mouse tail vein [221]. It has already been used to study the effect of cytokine overexpression *in vivo* [222]. To understand whether CCR1 ligands can promote *in vivo* AML development, young female mice were hydrodynamically injected with plasmids encoding CCL3, CCL6, and CCL9, followed by transplantation of *Cebpa* mutant LICs without conditioning. I chose to investigate the effect of these 3 CCR1 ligands because the genes encoding them were overexpressed in *Cebpa* mutant LICs/LBs when measured using bulk RNAseq, while genes encoding other CCR1 ligands were not (data not shown). Mice hydrodynamically injected with an equal amount of empty plasmid were transplanted in the same way and used as control.

Each CCR1 ligand-encoding plasmid was injected at 1μg since previous data from the Nerlov laboratory demonstrated that a high level of serum IL-6 expression was maintained for at least 7 days after hydrodynamic injection of 1μg IL-6-encoding plasmid [130]. Therefore, 1μg of plasmid for each CCR1 ligand for injection should be a reasonable starting point. The consequent 7-day duration of high serum tar-

get cytokine levels should be able to facilitate the early expansion phase of *Cebpa* mutant AML progression. In addition, studies showed that the serum level of the target cytokine gene peaked within 24 hours post-injection [221, 222]. Hence, after the hydrodynamic injection, the mice were left overnight for the expression of the injected plasmids to peak before being transplanted with the LICs. In this way, once transplanted, the LICs would be directly exposed to a CCR1 ligand-enriched environment. The schematic representation of the experiment can be found in Figure 4.10.

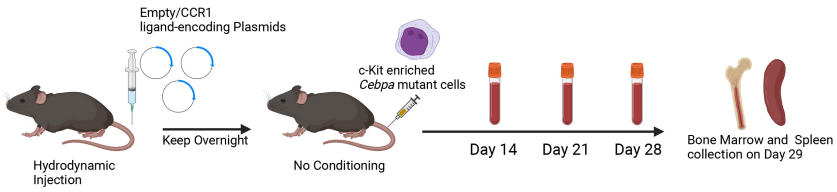

***Figure 4.10. Schematic of validating the role of CCR1 ligands in AML progression* in vivo**

Young CD45.1 female mice (2-3 months) were hydrodynamically injected with empty plasmids (3μg) or plasmids encoding CCL3, CCL6, and CCL9 (1μg/each, 3μg in total). After the injection, each mouse was transplanted with 25,000 purified CD45.2 bi-allelic Cebpa mutant LICs and monitored for up to 4 weeks.

PB samples were taken and analysed using flow cytometry at 2, 3, and 4 weeks after LIC transplantation (Figure 4.11A). Although minimal engraftment of CD45.2+ leukaemia cells was observed at the 2-week time point in most mice, regardless of injection groups, this engraftment was transient. No PB engraftment of the LICs was observed at week 3 and week 4 (Figure 4.11B-D). When all the mice were sacrificed and analysed 4 weeks after LIC transplantation, little or no BM leukaemia engraftment was observed in the mice injected with either the CCR1 ligand-encoding or empty plasmids (Figure 4.11E). A similar finding was also observed in the spleen (Figure 4.11F). It is possible that overexpression of the CCR1 ligands by hydrodynamic injection was insufficient in activating CCR1 signalling in the LICs, considering the high serum levels of CCR1 ligands probably could not be maintained throughout the 4 weeks. ELISA of the serum needs to be done in the hydrodynamically injected mice to assess the dynamic change of CCR1 ligand levels after the injection.

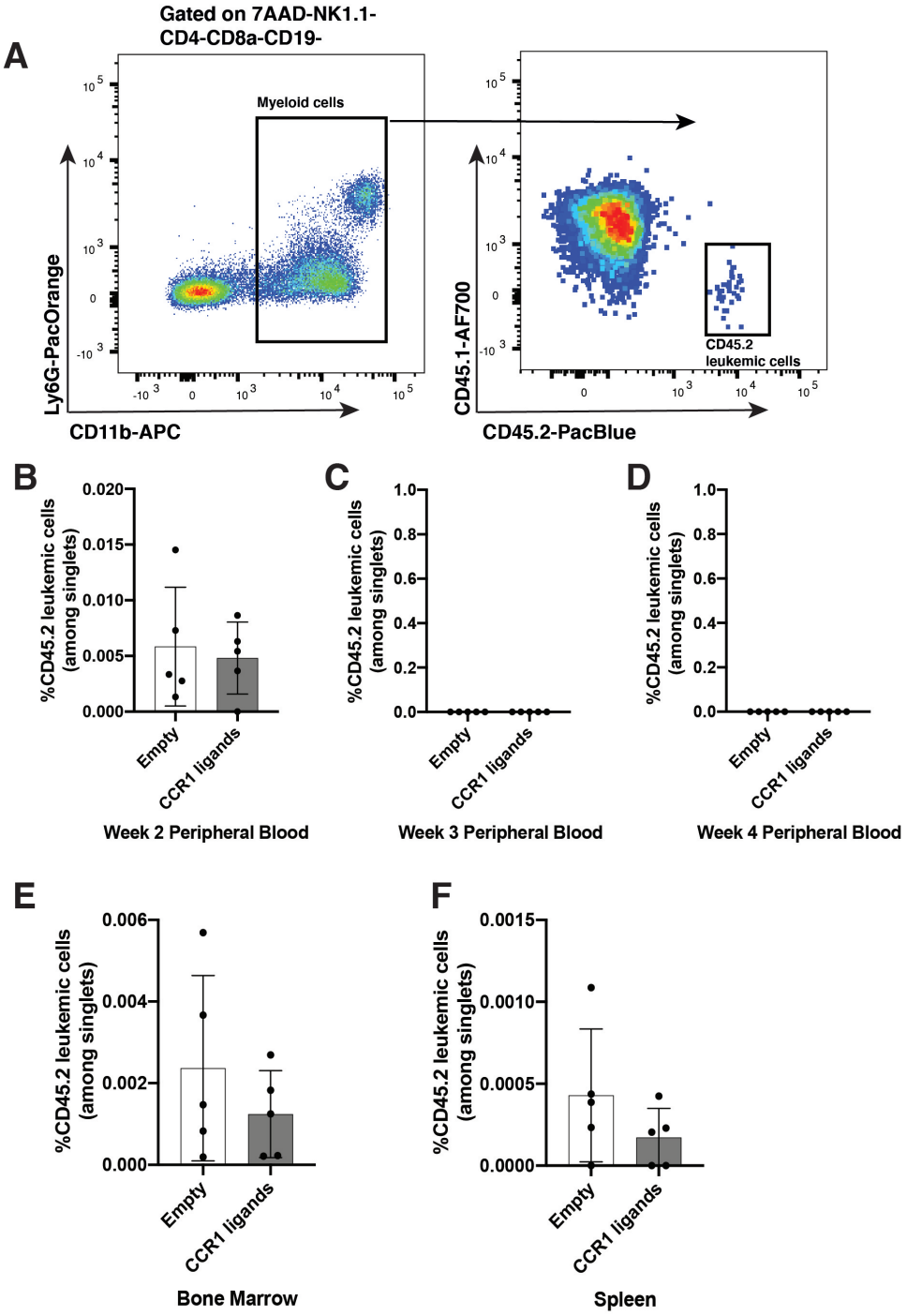

***Figure 4.11. Overexpression of CCR1 ligands by hydrodynamic injection did not promote* Cebpa *mutant LIC propagation* in vivo**

(A) Flow cytometry gating strategy used to identify CD45.2+ leukaemia cells within the peripheral blood.
(B) Histogram showing the frequency of CD45.2+ leukaemia cells in the peripheral blood of mice hydrodynamically injected with empty (N=5) or CCR1 ligand-encoding (N=5) plasmids 2 weeks after LIC transplantation. Data from 1 experiment.
(C) Histogram as in (B) showing the frequency of CD45.2+ leukaemia cells in the periph-

eral blood 3 weeks after LIC transplantation.
(D) Histogram as in (B) showing the frequency of CD45.2+ leukaemia cells in the peripheral blood 4 weeks after LIC transplantation.
(E) Histogram as in (B) showing the frequency of CD45.2+ leukaemia cells in the BM 4 weeks after LIC transplantation.
(F) Histogram as in (B) showing the frequency of CD45.2+ leukaemia cells in the spleen 4 weeks after LIC transplantation.

The MTS assay was used to investigate whether the CCR1 ligands are critical for *Cebpa* mutant LIC proliferation and survival in cells cultured *ex vivo*. *Cebpa* mutant LICs were cultured with 50ng/mL of CCL3, CCL6, or CCL9 and viability was assessed after 24 and 48 hours. However, when added alone, none of the cytokines increased the number of total viable LICs (Figure 4.12A-B). More studies are warranted to understand whether supplementation of the CCR1 ligands in *Cebpa* mutant LIC culture could change their phenotype (e.g. by examining the frequency of the leukaemia-initiating CD11b^{lo}c-Kit+ population using flow cytometry) or increase their engraftability (e.g. by transplanting the LICs with/without CCR1 ligand supplement into sub-lethally-irradiated recipient mice after the *ex vivo* culture).

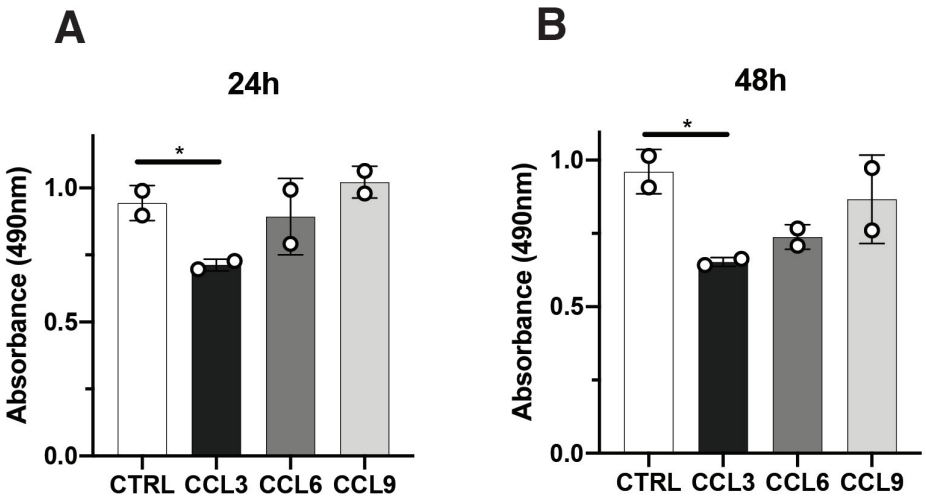

***Figure 4.12. CCL3, CCL6, and CCL9 could not increase* Cebpa *mutant LIC viability* ex vivo**

(A) Histogram showing the viability of Cebpa mutant LICs after being cultured ex vivo for 24 hours with or without recombinant CCR1 ligands (N=2 per culture condition). CCL3, CCL6, or CCL9 were added to the culture at a final concentration of 50ng/mL. Viability was measured using an MTS reagent. P-values were calculated using the Student's t-test.
(B) Histogram as in (A) after 48 hours of ex vivo culture.

4.6 *IL7R* was overexpressed in human AML samples

The expression level of *IL7R* was examined in the TCGA AML samples. Significant *IL7R* overexpression was observed in all the AML samples, including the bi-allelic *CEBPA* mutant leukaemias (Figure 4.13A). The increase in *IL7R* expression was also seen in the Beat AML *de novo* AML samples (Figure 4.13B). However, within the TCGA AML patients, a higher level of *IL7R* expression was not linked with worse OS (Figure 4.13C). To explore the molecular mechanism underlying IL7R being an AML dependency, the TCGA *CEBPA* mutant AML samples were segregated based on *IL7R* expression: *IL7R* high- and low-expressing denote the samples of which *IL7R* expression was within upper and lower 50% of all the TCGA-LAML patients, respectively. When PCA analysis was applied to the TCGA bi-allelic *CEBPA* mutant leukaemias, separate clustering of the *IL7R* high-expressing and *IL7R* low-expressing samples was observed (Figure 4.13D).

GSEA without pre-ranking comparing the *IL7R* high-expressing *CEBPA* mutant leukaemias and *IL7R* low-expressing *CEBPA* mutant revealed significant upregulation of the cytokine-signalling genes and borderline upregulation of the chemokine-signalling genes in the *IL7R* high-expressing samples as seen in the *CCR1* high-expressing samples (Figure 4.13E). Regarding the Beat AML samples, the 6 bi-allelic *CEBPA* mutant leukaemia samples were similarly divided into 2 groups by their *IL7R* expression levels: the $IL7R^{\text{hi}}$ group included 4 AML samples with *IL7R* expression levels significantly higher than the control; the $IL7R^{\text{lo}}$ group included 2 samples, of which the *IL7R* expression levels were similar to the control (Figure 4.13F). Clustering using PCA showed that the Beat AML *CEBPA* mutant leukaemias clustered according to the *IL7R* expression level (Figure 4.13G). However, upregulation of the cytokine- (p-value = 0.166) or chemokine-signalling genes (p-value = 0.645) was not observed in the $IL7R^{\text{hi}}$ *CEBPA* mutant leukaemias when GSEA was performed. Taken together, high *IL7R* expression in the AML is accompanied by the presence of distinct transcriptional changes, which may contribute to AML maintenance. However, the mechanism underlying this dependency of IL7R in AML cells remains unclear, which might be related to increased cytokine and

chemokine signalling.

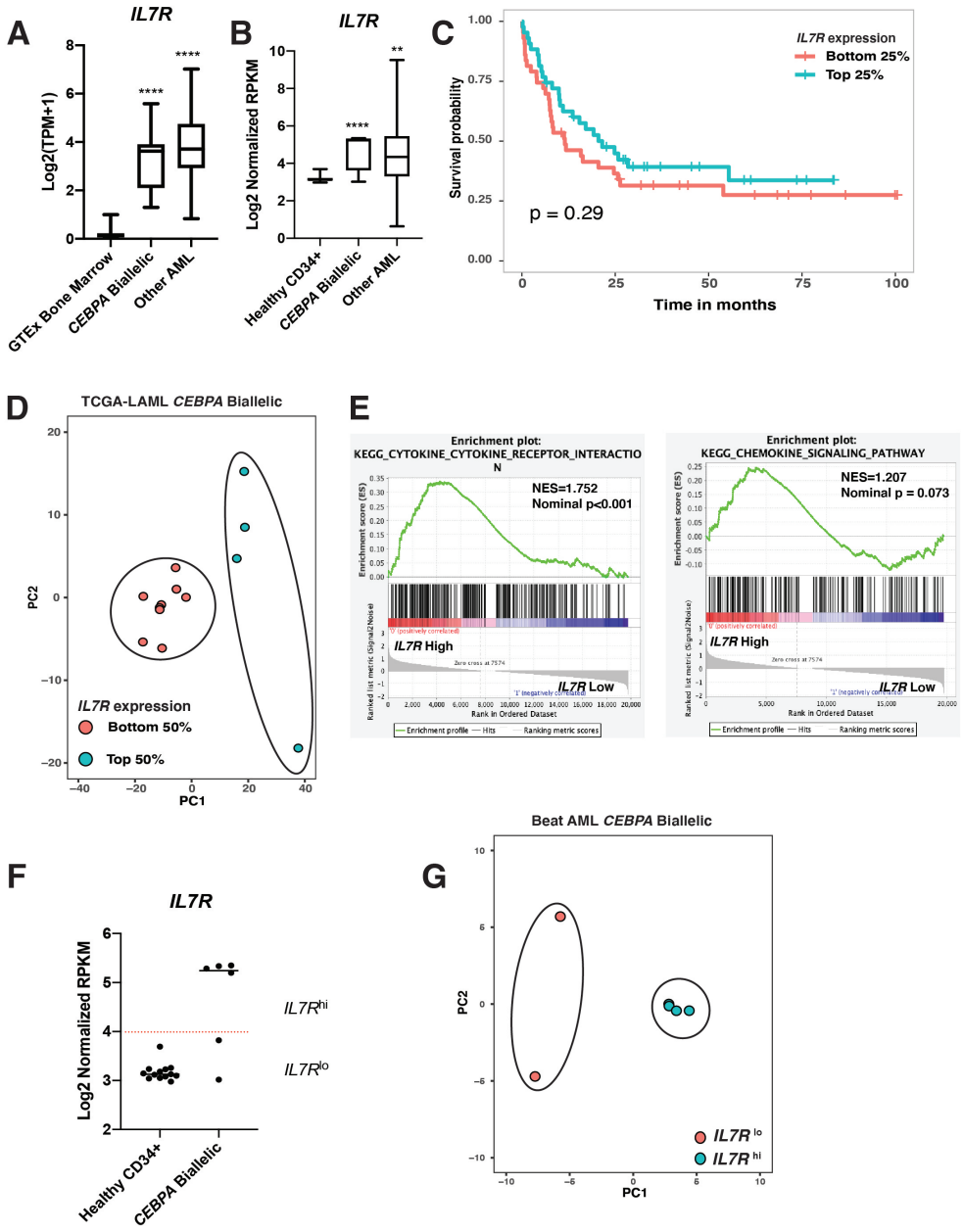

Figure 4.13. Overexpression of* IL7R *in human AML samples

(A) Box plots comparing IL7R mRNA expression (as log2(TPM+1)) between AML samples (TCGA-LAML RNAseq dataset) and normal BM samples (GTEx RNAseq dataset). The AML samples were grouped based on the presence or absence of bi-allelic CEBPA mutations. P-values were calculated using Student's t-test.
(B) Box plots comparing IL7R mRNA expression (as log2 normalised RPKM)) between Beat AML samples and healthy BM samples (CD34+ normal haematopoietic stem and progenitor cells). The AML samples were grouped based on the presence or absence of bi-allelic CEBPA mutations. P-values were calculated using Student's t-test.
(C) Survival data from the TCGA-LAML cohort for the patients with the highest and lowest quartiles of IL7R mRNA expression (RSEM value). The p-value was calculated

using the log-rank test.
(D) PCA of RNAseq data from CEBPA mutant TCGA-LAML samples using the top 500 most variable genes. The CEBPA mutant AML samples were segregated based on IL7R mRNA expression determined by the RSEM value (upper vs lower 50% among all the TCGA-LAML patients). The ovals were manually drawn to encompass the populations indicated.
(E) GSEA comparing IL7R-high vs IL7R-low CEBPA mutant TCGA-LAML samples using the KEGG chemokine (right) and cytokine (left) signalling pathway gene sets. Normalised enrichment score (NES) and p-value are shown.
(F) Scatter plot showing segregation of $IL7R^{hi}$ and $IL7R^{lo}$ CEBPA mutant Beat AML samples.
(G) PCA of RNAseq data from CEBPA mutant Beat AML samples using the top 500 most variable genes. The CEBPA mutant AML samples were segregated based on IL7R mRNA expression as in (F). The ovals were manually drawn to encompass the populations indicated.

4.7 *In vivo* IL7R-blocking antibody treatment did not impair *Cebpa* mutant LIC propagation

Here, I intended to assess whether inhibiting IL7R signalling can be used as a potential therapeutic strategy for AML. A monoclonal anti-IL7R antibody (A7R34) has already been demonstrated to be effective in prolonging the survival of mice transplanted with T-ALL cells when given over a time course of 3 weeks post-transplantation [212]. To explore whether this anti-IL7R antibody could convey an anti-leukaemic effect in AML, the *Cebpa* mutant AML mice were treated with the anti-IL7R monoclonal antibody or isotype control (500μg per treatment) twice a week for 3 consecutive weeks after LIC transplantation (Figure 4.14A). The antibodies were administered by intraperitoneal injections. PB was taken 20 days after transplantation for flow cytometry analysis (Figure 4.14B). All the mice were sacrificed 21 days post-transplantation for flow cytometry analysis to read out leukaemia cell engraftment.

Surprisingly, no significant difference in BM leukaemic engraftment was observed between the IL7R-blocking antibody-treated and control leukaemia mice (Figure 4.15A), even though *Il7r* disruption in the LICs through the *Il7r*-targeting sgRNA significantly reduced the BM and spleen leukaemic burden as shown in Figure 4.6. More strikingly, the leukaemic burden in the spleen significantly increased after

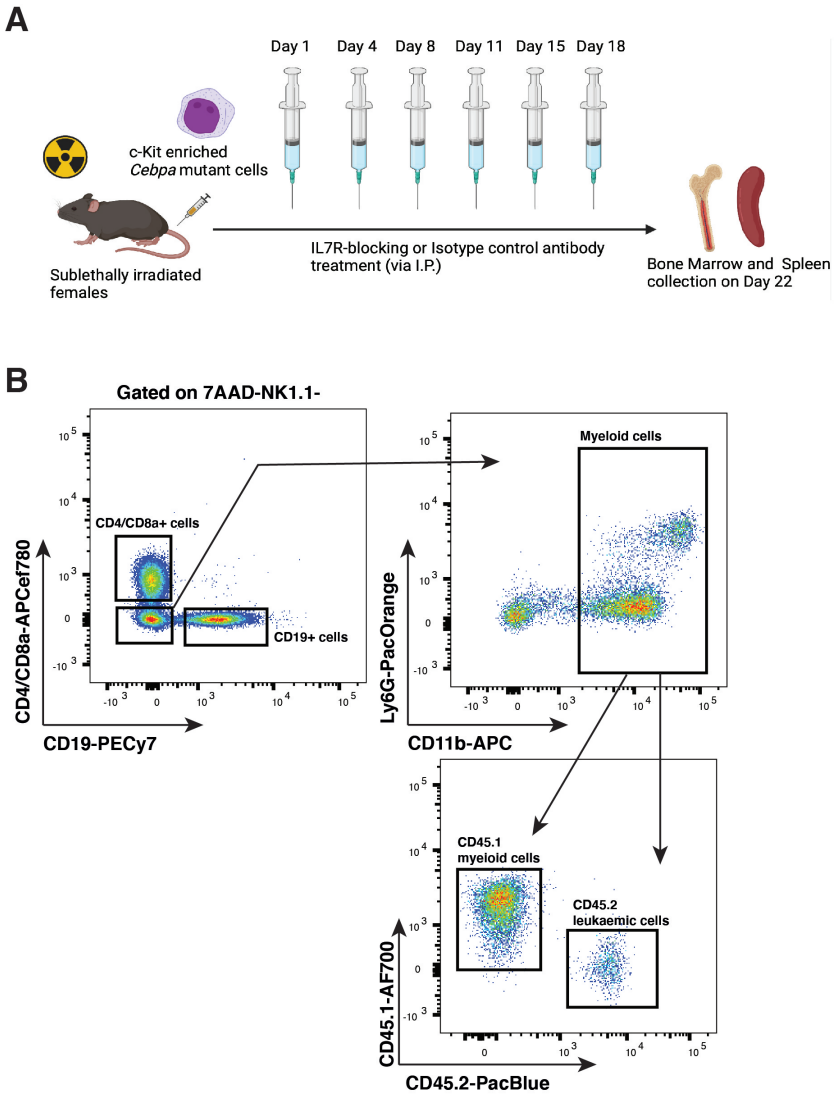

Figure 4.14.* In vivo *IL7R-blocking antibody treatment in AML mice

(A) Schematic of examining the effect of in vivo IL7R inhibition on AML progression. Sublethally-irradiated young CD45.1 female mice (2-3 months) were transplanted with 25,000 purified CD45.2 bi-allelic Cebpa mutant LICs, followed by administration of anti-IL7R monoclonal antibody or isotype control twice a week for 3 consecutive weeks.
(B) Flow cytometry gating strategy used to identify CD45.2+ leukaemic cells, normal CD45.1+ myeloid cells, CD4+/CD8a+ T cells and CD19+ B cells within the PB.

IL7R inhibition (Figure 4.15B). Similar to what was observed in the spleen, the leukaemic burden (CD45.2+CD11b+) also increased dramatically in PB after IL7R inhibition (Figure 4.15C), which was accompanied by a highly significant reduction in the PB frequency of mature T cells (CD8a+/CD4+) and B cells (CD19+) (Figure 4.15D-E). Interestingly, the PB frequency of recipient-derived normal myeloid cells (CD45.1+CD11b+) also increased significantly after IL7R inhibition (Figure 4.15F).

This indicated that the depletion of T and B cells in the PB of IL7R-inhibited mice was unlikely to be caused by the more progressive overcrowding of the BM space by AML blasts.

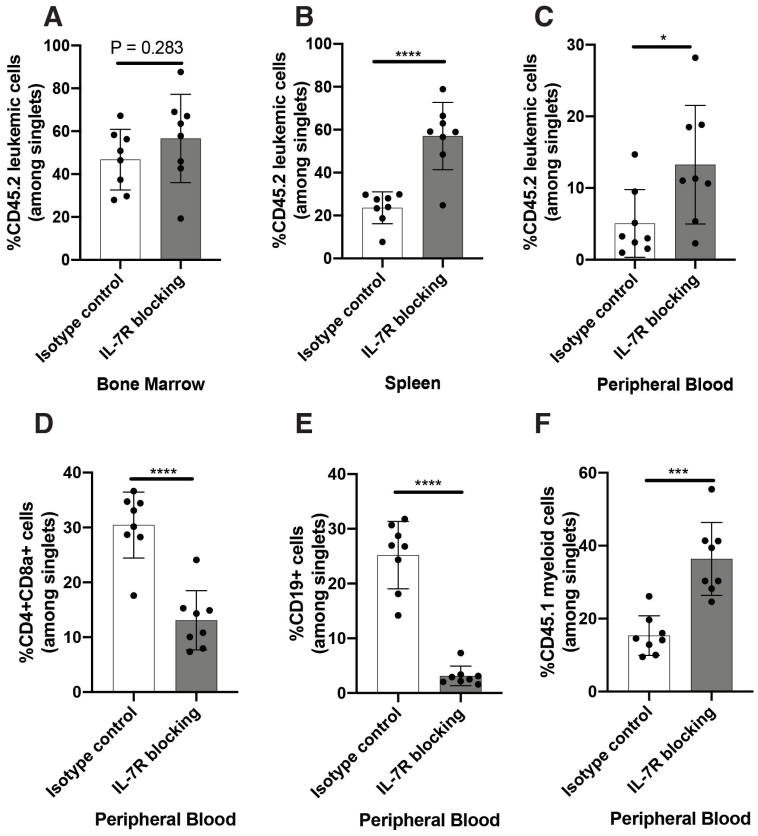

Figure 4.15. Blocking IL7R* in vivo *did not impair* Cebpa *mutant LIC propagation

(A) Histogram showing the frequency of CD45.2+ leukaemia cells in the BM of mice treated with the isotype control (N=8) or IL7R-blocking (N=8) antibody 3 weeks after LIC transplantation. Data from 2 experiments. P-values were calculated using Student's t-test.
(B) Histogram as in (A) showing the frequency of CD45.2+ leukaemia cells in the spleen.
(C) Histogram as in (A) showing the frequency of CD45.2+ leukaemia cells in the PB.
(D) Histogram as in (A) showing the frequency of CD4+/CD8a+ T cells in the PB.
(E) Histogram as in (A) showing the frequency of CD19+ B cells in the PB.
(F) Histogram as in (A) showing the frequency of normal CD45.1+ myeloid cells in the PB.

First, if the drop in PB T and B cell frequencies had been caused by the rise in leukaemic burden after IL7R inhibition, all recipient-derived normal haematopoietic cells, including the CD45.1+ myeloid cells, would have been effaced. Further decoupling the lower T and B cell levels after IL7R inhibition from advanced AML

progression, the level of lymphopenia has been shown to be independent of the AML disease burden [102]; a significant increase in the absolute number of total T cells in the PB was shown in AML patients as compared with healthy donors [223]. Moreover, the lifespan of mature T cells is relatively long [224, 225]. This argues against the possibility of the significantly lower T cell frequencies observed as early as 3 weeks after IL7R inhibition being caused by the leukaemia-driving T-cell lymphopoiesis impairment. Last but not least, deficiency of IL-7 and IL-7R has already been shown to impair murine T and B cell development [209]. The antibody against normal IL7R could be expected to target normal mature T cells and B cell precursors, which may lead to irreversible severe immunodeficiency in the mouse [226, 227]. Overall, I reckon that the depletion of T and B cells after IL7R inhibition was more a direct consequence of the IL7R-blocking antibody treatment than more advanced AML progression following IL7R inhibition.

This raised the possibility that the expansion of CD45.2+ leukaemia cells after IL7R inhibition was not a leukaemia-specific effect and did not necessarily mean that the IL7R-blocking antibody was pro-leukaemic. Instead, the expansion of CD45.2+ leukaemia cells observed in the spleen and PB could be attributed to the treatment-related depletion of T and B cells. Such depletion, in turn, resulted in relative enrichment of both the normal and malignant myeloid cells, including the CD45.2+ leukaemic ones. Another possibility is that IL7R inhibition causes immune deficiency in the mice. This can paradoxically promote LIC engraftment and expansion as seen in the NOD/SCID mouse, despite LIC's dependency on IL7R [228]. Taken together, despite IL7R being an *in vivo* AML dependency, which has been demonstrated by significant LIC depletion after CRISPR-mediated *Il7r* disruption, systemic inhibition of IL7R signalling using an IL7R-blocking antibody does not have a therapeutic effect in the AML mouse. This is possibly due to its immunosuppressive effects causing depletion of normal T and B cells. More studies are warranted to develop an *in vivo* therapeutic strategy that can target IL7R, more specifically in AML cells, while the normal immune system remains unaffected.

4.8 Discussion

In line with *Ccr1* being the top hit during the first CRISPR dropout screen, I showed in this chapter that disruption of *Ccr1* expression using CRISPR demonstrated that CCR1 was indispensable for AML LIC maintenance. In addition, the chemokine signalling gene signature was enriched in both the murine *Cebpa* mutant LICs and human CCR1^{hi} *CEBPA* mutant leukaemias. These findings, together with Chapter 3 showing CCR1 ligand upregulation during AML, further demonstrated that intrinsic CCR1 signalling is essential for AML LIC maintenance. This is consistent with a previous study showing the requirement of CCR1 expression for the maintenance of CML LICs [197]. However, the underlying mechanism of CCR1 being an *in vivo* AML dependency is still unclear [229]. To better elucidate how *Ccr1* disruption impacts the biology of the LICs, RNAseq of the *Ccr1* sgRNA-expressing LICs is needed to pinpoint the key regulators of LIC maintenance downstream of CCR1.

Surprisingly, *in vivo* overexpression of CCL3, CCL6, and CCL9 using hydrodynamic injection could not promote AML development. CCL3, CCL6, or CCL9 also failed to stimulate murine LIC proliferation or improve murine LIC survival *ex vivo*. One possible explanation is that to activate CCR1 signalling in the LICs, the CCR1 ligands must be derived from specific cell types, such as leukaemia or BM stromal cells. There is considerable evidence that the effect of specific BM microenvironmental factors is cellular source-dependent. For example, SCF is a critical factor that maintains HSCs; HSCs only require SCF expressed by VECs and PVCs rather than OBs or other cell types [67]. Hydrodynamic injection of the CCR1 ligand-expressing plasmids leads to target gene overexpression predominantly in the liver cells [221]. Therefore, this may be unable to upregulate CCR1 signalling in the LICs. Besides, it is unclear how long the high level of CCR1 ligands in the BM needs to be maintained to activate CCR1 signalling and propagate AML in the non-irradiated recipient mice. A high level of CCR1 ligands in the early phase of AML development only may not suffice. Therefore, a single dose of hydrodynamic injection may not be appropriate to maintain prolonged *in vivo* CCR1 ligand overexpression since the target protein level usually declines in a time-dependent manner after the injection [221]. Indeed, according to the ELISA result shown in Figure 3.10F, high

levels of BM CCL3 were still detectable 4 weeks after LIC transplantation. This suggests that a more sustained method of *in vivo* CCR1 ligand overexpression, such as administration of the recombinant cytokines regularly, is needed to promote AML development in the non-irradiated recipient [197]. It is also possible that the crucial role of CCR1 signalling in LIC maintenance is independent of CCR1 ligand activation, although upregulation of the ligands was observed in the leukaemia mice. This hypothesis still needs experimental validation. It is still too early to conclude that the interaction between CCR1 and its ligands is not essential for AML development.

Even though I only experimentally validated the *in vivo* role of CCR1 in the *Cebpa* mutant mouse model, it is possible that this finding could be extended to other subtypes of AML. For instance, CCR1 may be essential in *MLL-AF9* leukaemia as it is also overexpressed in the *MLL-AF9* leukaemia cells [161]. Therefore, it would be interesting to delete *Ccr1* expression in the *MLL-AF9* leukaemia cells and investigate whether this can similarly deplete the leukaemic burden. It is noteworthy that *Ccl3* knockout in the *MLL-AF9* leukaemia cells has been shown to delay leukaemogenesis, which implies the potential role of CCR1 signalling in *MLL-AF9* leukaemia development [230].

The fact that *CCR1* and the genes encoding its ligands are overexpressed in human AML samples of various cytogenetic and molecular genetic subtypes further implies the potential role of CCR1-signalling beyond *CEBPA* mutant AML. However, all my experimental data came from a murine AML model, and experiments validating the role of CCR1 in primary human AML cells are still lacking. It would be important to disrupt *CCR1* expression in primary human AML cells, followed by using cell culture or a xenograft model, to validate CCR1 as a dependency in human AML cells. Besides, a selective CCR1 antagonist (BL5923) was developed several years ago [176], which has a therapeutic effect on mice with diseases such as lupus nephritis [231] and disseminated colon cancer [177]. This implies the high translational potential of CCR1 as a therapeutic target for AML. Therefore, it would be interesting to treat the young mice transplanted with *Cebpa* mutant or *MLL-AF9* leukaemia cells with the selective CCR1 antagonist to investigate whether the leukaemic burden can be reduced or whether survival of the mice can be prolonged. Furthermore, to study

whether CCR1 inhibition can be a potential therapy for AML patients, xenograft models of human AML need to be treated with the CCR1 antagonist.

The significant depletion of BM and splenic leukaemia cells after *Il7r* knockout confirmed the essential role of IL7R in AML LIC maintenance. Interestingly, not only was *Il7r*, the gene encoding one of the subunits of IL-7R, identified as the top hit during the second CRISPR, but *Il2rg*, the gene encoding the other subunit of the heterodimeric IL-7R, was also ranked highly during the first CRISPR screen. This suggests that IL-7R signalling can be essential for AML LIC maintenance. Although previous studies have demonstrated that IL-7R signalling upregulation can contribute to T-ALL and B-ALL development [212, 213], this is the first time, to our knowledge, that IL-7R signalling is implicated in AML development *in vivo*. However, in contrast to the pro-survival effect of the IL7R-blocking antibody seen in the T-ALL mice [212], the IL7R-blocking antibody was not able to decrease the leukaemic burden in our *Cebpa* mutant AML mouse model. This could be due to the immunosuppressive effect of the IL7R-blocking antibody, which targets not only the IL7R-expressing LICs but also the normal T cells and progenitors of B cells [209]. Considering the last IL7R-blocking antibody treatment was at 18 days after LIC transplantation, and all the mice were culled at 21 days, it might be beneficial to keep the treated leukaemic mice longer after the last treatment. This would allow the suppressed immune cell populations to rebound and more AML-specific effects to emerge and be observed. In addition, considering IL7R expression is barely detectable in the *Cebpa* mutant AML LICs using flow cytometry, more solid evidence is still needed to illustrate the existence of IL-7R signalling in the LICs. For instance, *Cebpa* mutant LICs can be treated with recombinant IL-7 *ex vivo*. The recipient mouse can also be treated with IL-7 before and after LIC transplantation to investigate whether *in vivo* LIC proliferation is promoted. Besides, RNAseq can be performed after treating the leukaemia mice with recombinant IL-7 to study whether the IL-7R signalling pathway is activated in the LICs. Interestingly, similar to *CCR1*, *IL7R* expression was upregulated across different subtypes of human AML. Therefore, it would also be worthwhile to study the role of IL7R using other AML mouse models, such as the *MLL-AF9* leukaemia model.

Taken together, the *in vivo* CRISPR dropout screen pipeline, combined with the kinetically tractable and strictly microenvironment-dependent murine model of *CEBPA* mutant AML, facilitates the identification of *in vivo* AML dependencies. Using this strategy, 2 cell surface receptor genes (*Ccr1* and *Il7r*) essential for *CEBPA* mutant AML development have been identified and validated, both of which can be targeted using a small molecular inhibitor or a monoclonal antibody. Although the findings still need further experimental validation and also be studied in primary human AML samples, this work provides the proof of principle that targeting CCR1 and IL7R in AML cells, individually or in combination, may be promising treatments for AML. Such less cytotoxic therapy can be particularly beneficial for older or frail patients who are ineligible for intensive chemotherapy and HCT. The high expression levels of *CCR1* and *IL7R* across various subtypes of AML also indicate the potential of extending the current finding in *CEBPA* mutant AML to other subtypes of AML, such as AML with MLL-fusions.

Chapter 5

Identification of cytokine signalling contributing to age-related AML predisposition

5.1 Introduction and aims

This chapter is focused on delineating the association between AML- and age-induced BM microenvironmental changes and how this may contribute to ageing-related AML predisposition. As previously mentioned, AML is much more frequent in elderly patients (>= 65 years of age). Moreover, the survival of AML worsens significantly with advancing age: for AML patients, each increase of a decade of life leads to a decrease of around 10% in 5-year OS [4]. This can partly be explained by AML cell-intrinsic factors: the cytogenetic features of AML that are associated with a poor prognosis, such as chromosome 5 abnormalities and the monosomal karyotype, are more commonly seen in elderly AML patients; on the contrary, the frequency of favourable mutations, such as *NPM1* or *CEBPA* mutations, decreases significantly with age [232]. However, when comparing elderly (>65 years old) and younger (<56 years old) patients from the same cytogenetic risk group, the elderly patients still experienced poorer OS [103]. This suggests that, in addition to the AML cell-intrinsic factors, other factors also play a role in the poorer outcomes seen in elderly patients. On the one hand, many elderly patients with AML are not eligible for intensive induction chemotherapy due to their poor performance status

and significant comorbidities [233]. They are only given palliative treatment, such as best supportive care, or less intensive chemotherapy [4, 232, 234], which can be associated with their significantly poorer survival rate [105]. On the other hand, age-related BM changes in haematopoietic cell-extrinsic compartments, such as the BM stromal cells or secreted factors, could contribute to AML therapy resistance, disease relapse, and decreased survival observed in elderly AML patients. This is supported by the fact that outcomes in patients with the same cytogenetic abnormalities and treated using the same regimens still worsened with age [103]. Indeed, as described in Chapter 1, the BM microenvironment can regulate normal and malignant haematopoietic cells by both direct cell-cell interactions and secreted factors, and both the cellular components of the BM microenvironment and the secreted factors within the BM nice (secreted locally or distantly) undergo significant changes with age [129, 235].

AML progression can also induce BM microenvironmental changes, as ageing does. Regarding endothelium, using a murine model of *MLL-AF9* AML, it has been reported that AML increases or maintains CBM sinusoidal densities while the endosteal vessels and arterioles, which are found predominantly in the endosteal niche, are reduced [72, 85, 86]. A similar decrease in the endosteal vessels has also been observed in AML patients [85]. Besides, impaired adipogenesis and OB differentiation have been shown to be associated with AML progression using both AML patient samples and mice transplanted with AML cells [85, 87, 236]. Importantly, these phenotypic BM microenvironmental changes have been functionally implicated in leukaemia progression. In addition, increased protein levels of pro-inflammatory cytokines and chemokines, including TNF-α, CCL3, and IL-6, have been observed in both mouse models of AML and patients diagnosed with AML [85, 95, 97]. The source of the elevated cytokines can be either the AML cells or the remodelled BM stromal cells [85, 102, 237].

It is therefore evident that some of the AML- and age-induced BM microenvironmental changes are overlapping. At the cellular level, both AML and ageing can cause damage to the endosteal niche [85, 126]. Inflammation is a hallmark of the aged and AML BM microenvironment: an increase in pro-inflammatory cytokine and

chemokine levels has been observed in both the AML-conditioned and aged BM. For instance, increased BM IL-6 levels during ageing [129] could directly contribute to impaired erythroid progenitor function [130], and AML-derived IL-6 has also been shown to play a role in AML-induced anaemia [102]. Besides, IL-1 can promote AML cell expansion while suppressing normal progenitor growth, and chronic IL-1 exposure might contribute to functional HSC impairment during ageing [94, 238]. In this regard, there is a possibility that ageing and AML exposure have similar effects on the BM microenvironment, which leads to pre-existing AML-induced features in the aged niche and increases the ability of the aged BM to propagate AML. However, a direct link between age-related BM microenvironmental changes and AML progression is yet to be established.

To more systemically explore whether and how ageing and AML exposure have similar effects on the BM microenvironment and identify the overlapping features that may play a role in AML progression, the Nerlov laboratory performed systematic transcriptional profiling of aged (vs young) and AML-exposed (vs non-exposed normal) BM stromal cells at multiple time points during AML progression. Not surprisingly, there was significant conservation of process networks de-regulated during AML in the aged BM niche cells [56]. Furthermore, the possibility of AML-induced transcriptional changes pre-existing in the aged BM microenvironment and being AML-supportive was functionally corroborated by transplantation assays: superior engraftment of *Cebpa* mutant LICs was observed in the aged compared to young, non-irradiated recipient mice, as described in Thomas A, 2018 [56].

In light of the established link between age-related BM microenvironmental changes and AML progression, the next question to address is which specific age-associated changes to the BM niche, such as increased environmental factor(s) or activated signalling pathway(s), are critical to the superior AML propagation in the old BM. Identifying this specific factor or signalling pathway could provide insight into developing novel therapies for elderly AML patients, who are usually ineligible for intensive chemotherapy. In this chapter, I will focus on investigating the role of several pro-inflammatory cytokines/chemokines, including IL-6 and CCR1 ligands, in AML development during ageing. Eventually, I showed that CCR1 signalling rather

than IL-6 signalling constituted an AML-dependency that was increased during ageing, which, at least partly, explained the increased ability of AML to propagate in the aged BM.

5.2 AML- and age-induced transcriptional changes to BM stromal cells are correlated

Given the potential association between the aged and AML-exposed BM microenvironment, the Nerlov laboratory has been seeking to identify common transcriptomic features between the aged and AML-exposed BM niche cells, which could play a role in AML progression. Thus, bulk RNAseq of BM stromal cells harvested from young (8-12 weeks of age) and old (23-24 months of age) mice was performed by the Nerlov laboratory to characterise transcriptional changes that occurred to the BM microenvironment during physiological ageing [130]. As the bulk of leukaemia cells mainly occupies the central BM, which also contains most sinusoidal endothelial niches [239], the AML-exposed and control CBM niche cells (sorted from mock-transplanted mice) were similarly profiled using bulk RNAseq by the Nerlov laboratory. More details were described in Thomas A, 2018 [56]. The schematic representation of how the RNAseq was performed is shown in Figure 5.1.

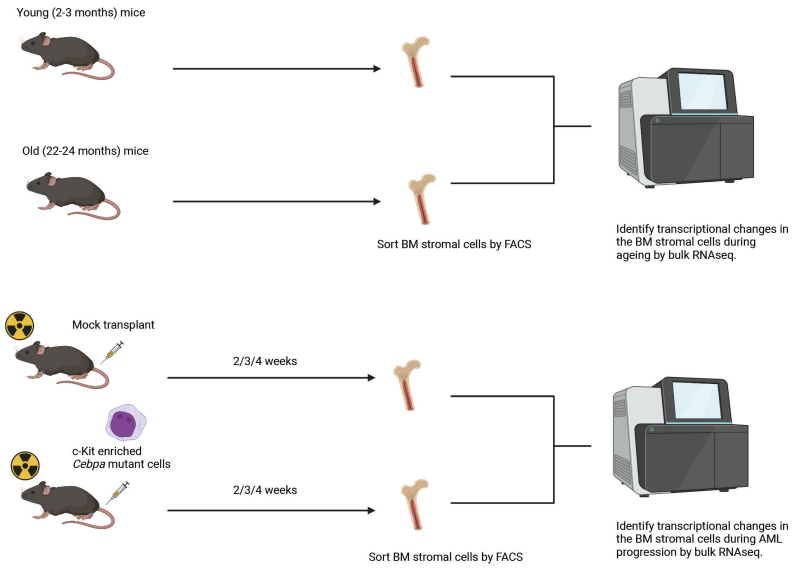

Figure 5.1. Schematic of profiling leukaemic and aged BM stromal cells by bulk RNAseq

To investigate BM microenvironmental changes that occur during normal ageing (top), SmartSeq2-based RNAseq of FACS purified BM stromal cells from both young (as control) and aged mice was performed. To investigate the functional interaction between AML and BM stromal cells over time (bottom), SmartSeq2-based RNAseq of FACS purified BM stromal cells from both mock (as control) and LIC transplanted mice was performed at the 2-, 3-, and 4-week time points of AML development.

In line with the hypothesis that some of the AML- and age-induced transcriptional changes to the BM microenvironment may overlap, the Nerlov laboratory found that the AML-exposed and aged CBM VECs were transcriptionally alike, as described in Thomas A, 2018 [56]. First, DESeq2 was used to identify DEGs by comparing the leukaemic and control samples at 2, 3, and 4 weeks after LIC transplantation. This way, any transcriptional changes associated with the sub-lethal irradiation protocol could be controlled. Interestingly, a highly significant number of upregulated DEGs (up-DEGs) were shared between the aged and AML-exposed CBM VECs throughout the disease trajectory (Figure 5.2A). Of note, the highest number of up-DEGs shared by the aged and AML-exposed CBM VECs was observed at week 3, a key time point when the BM leukaemic burden was high and continued to increase. In contrast, similar comparisons for CBM PVCs showed only marginal significant overlap at week 4 (Figure 5.2B).

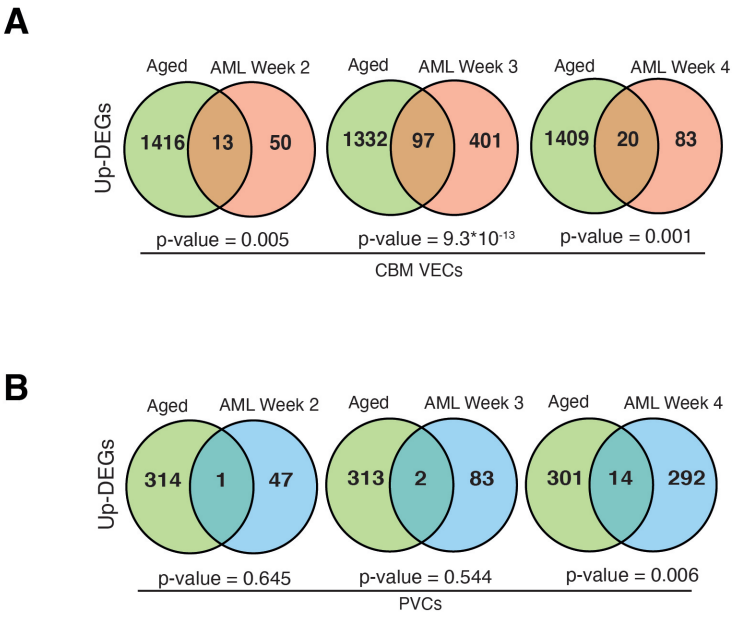

Figure 5.2. A significant number of up-DEGs were shared by leukaemic and aged CBM VECs

(A) Venn diagrams depicting the overlap of up-DEGs identified by DESeq2 analysis between aged CBM VECs and AML-exposed CBM VECs at different time points post-transplantation. DESeq2 analysis of aged (vs young) CBM VECs and AML-exposed (vs control) CBM VECs was performed using RNAseq data from Figure 5.1. P-values for the overlaps were calculated using a hypergeometric test.
(B) Analysis as in (A) of the overlap between up-DEGs in aged and AML-exposed and PVCs.

To further investigate whether CBM stromal cells were similarly affected by AML

progression and ageing, MPNA was used to identify perturbed molecular networks in the aged and AML-exposed CBM VECs by the Nerlov laboratory, as described in Thomas A, 2018 [56]. Consistent with the DEG analysis, a highly significant number of perturbed networks shared by the aged and AML-exposed CBM VECs was observed at week 3 (Figure 5.3A). Interestingly, the shared process networks between the aged and week 3 CBM VECs included IL-6 signalling and TGF-β/activin signalling, both of which had been identified by the Nerlov laboratory as critical contributors to haematopoietic ageing [130]. The shared network perturbations also included other inflammation-associated signallings, such as Amphoterin and TREM1 signalling. This is in agreement with previous findings demonstrating that both ageing and myeloid malignancies can increase inflammatory signalling in the BM microenvironment [77, 235], which might explain the predisposition to AML observed in the non-irradiated aged mice [56] and elderly population [105]. In contrast, for CBM PVCs, the overlap between the age- and AML-induced changes to process networks only showed borderline significance at all time points (Figure 5.3B). This indicates that ageing and AML exposure de-regulate similar inflammation-associated processes, specifically in CBM VECs.

In light of the above findings, it is reasoned that leukaemic/age-associated conditioning of the BM microenvironment that can alter the function of CBM VECs and promote leukaemia progression would manifest itself as de-regulated process networks that are preserved by both the CBM VECs and leukaemia cells. Indeed, further analysis demonstrated a highly significant overlap of perturbed process networks between the aged CBM VECs, the AML-exposed CBM VECs, and the LICs (Figure 5.3C). Importantly, IL-6 signalling, Amphoterin signalling, and TREM1 signalling remained as overlapping perturbed process networks, which further supported the role of pro-inflammatory signalling in age-related AML predisposition. A full list of the perturbed networks shared by the aged CBM VECs, AML-exposed CBM VECs, and LICs can be found in Table 5.1.

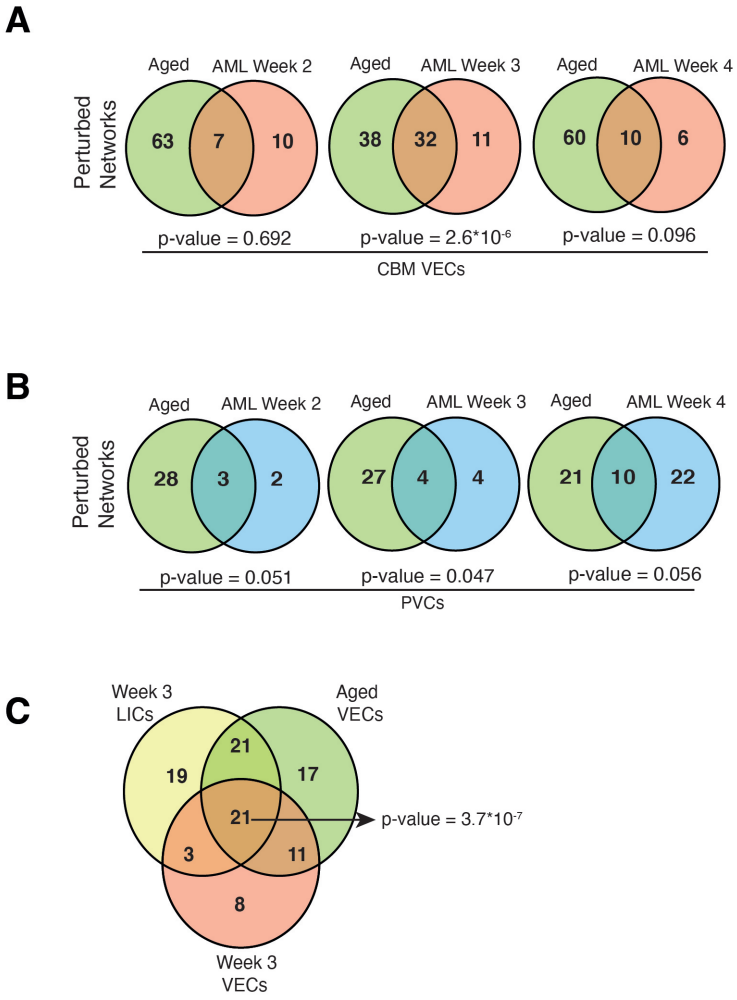

Figure 5.3. A significant number of Metacore perturbed networks were shared by leukaemic and aged CBM VECs

(A) Venn diagrams depicting the overlap of perturbed networks identified by Metacore analysis between aged CBM VECs and AML-exposed CBM VECs at different time points post-transplantation. Metacore perturbed network analysis of aged (vs young) CBM VECs and AML-exposed (vs control) CBM VECs was performed using RNAseq data from Figure 5.1. P-values for the overlaps were calculated using a hypergeometric test.
(B) Analysis as in (A) of the overlap between networks perturbed in aged and AML-exposed and PVCs.
(C) Venn diagram depicting the overlap of perturbed networks in week 3 LICs, week 3 CBM VECs, and aged CBM VECs. The p-value was calculated using a random draw simulation (10,000,000 iterations).

In addition, I performed non-ranked GSEA to characterise common transcriptional changes shared between the AML-exposed and aged CBM VECs. Not surprisingly, GSEA showed that the expression of up-DEGs of leukaemic CBM VECs was significantly enriched in the aged CBM VECs (Figure 5.4A). First, this demonstrates that a significant proportion of transcriptional changes to the AML-exposed CBM VECs

Networks
Immune response_Phagocytosis
Proliferation_Positive regulation cell proliferation
Cytoskeleton_Actin filaments
Cell adhesion_Integrin-mediated cell-matrix adhesion
Development_Blood vessel morphogenesis
Cell adhesion_Platelet aggregation
Cytoskeleton_Regulation of cytoskeleton rearrangement
Inflammation_IL-6 signaling
Cell adhesion_Integrin priming
Cytoskeleton_Intermediate filaments
Immune response_Phagosome in antigen presentation
Immune response_TCR signaling
Immune response_IL-5 signalling
Development_Hemopoiesis, Erythropoietin pathway
Signal Transduction_Cholecystokinin signaling
Proliferation_Lymphocyte proliferation
Translation_Regulation of initiation
Inflammation_Amphoterin signaling
Cell cycle_Mitosis
Inflammation_Protein C signaling
Inflammation_TREM1 signaling

Table 5.1. Metacore perturbed networks shared by LICs, leukaemic CBM VECs, and aged CBM VECs

Metacore perturbed process networks shared by week 3 LICs (vs GMPs), week 3 leukaemic CBM VECs (vs control CBM VECs), and aged CBM VECs (vs young CBM VECs) are shown.

pre-exist in the aged CBM VECs. Second, this, together with the fact that engraftment of the LICs is superior in the aged recipient mice, compared to the young, indicates that these pre-existing transcriptional changes to the old CBM VECs can functionally promote AML progression in the aged BM [56]. In addition, consistent with gene expression signatures associated with chemokine and cytokine signalling being increased in the AML-exposed CBM VECs, the cytokine signalling signature was significantly enriched in the aged CBM VECs (Figure 5.4B). Although statistically insignificant, the chemokine signalling signature was also enriched in the aged CBM VECs (Figure 5.4C). These results align with the previous findings demonstrating that both ageing and leukaemia can induce a pro-inflammatory BM environment [77, 235]. Hence, targeting specific cytokine/chemokine signalling, especially the pro-inflammatory cytokines, may counteract the predisposition to AML during ageing and be used as a novel therapy for elderly AML patients.

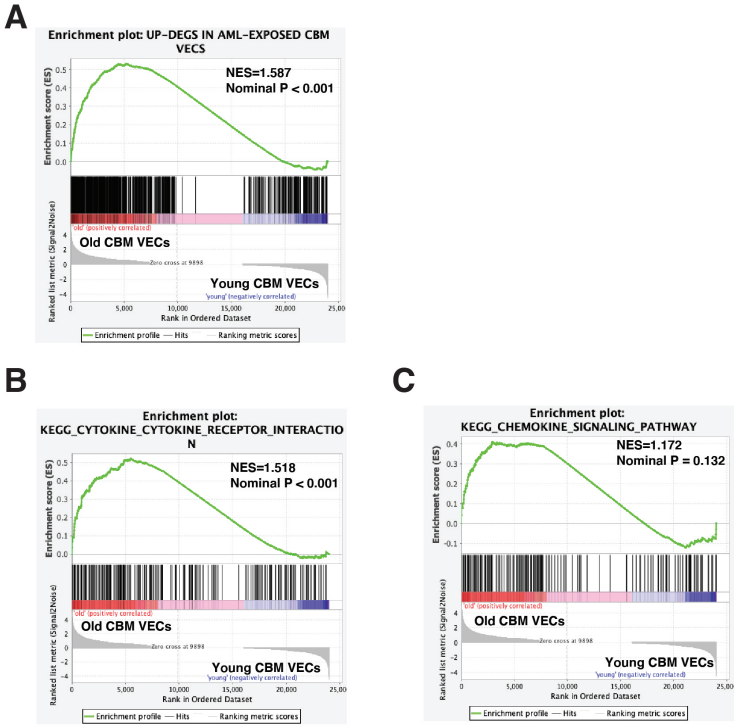

Figure 5.4. AML-induced molecular perturbations in CBM VECs pre-exist in aged BM

(A) GSEA comparing aged vs young CBM VECs using a customised gene set including up-DEGs of AML-exposed CBM VECs. Up-DEGs of AML-exposed CBM VECs (vs control) were defined as DEGs with an adjusted p-value < 0.05 and log2 fold change > 0 at the 2-, 3- or 4-week time points of AML development; genes with an RPKM value constitutively

lower than 1 in all the CBM VECs that had been sequenced were excluded. This led to a gene set of 624 genes. Normalised enrichment score (NES) and p-value are shown.
(B) GSEA comparing aged vs young CBM VECs using the KEGG cytokine signalling pathway gene set. Normalised enrichment score (NES) and p-value are shown.
(C) GSEA as in (B) using the KEGG chemokine signalling pathway gene set.

5.3 IL-6 signalling was not involved in increased AML progression during ageing

I next intended to investigate which pro-inflammatory cytokine(s) and the associated signalling pathway are critical to the increased AML progression in the old BM microenvironment. The first pro-inflammatory cytokine examined was IL-6, considering the MPNA showing that IL-6 signalling was simultaneously enriched in the LICs, the AML-exposed CBM VECs, and the aged CBM VECs. Analysis of human patient samples also showed significant overexpression of *IL6R* in TCGA-LAML samples compared to the control, and the higher *IL6R* expression level was associated with poorer OS in AML patients; high circulating levels of IL-6 were also associated with decreased OS in AML patients, which indicates the implication of increased IL-6 signalling in AML progression [237, 240]. Besides, the protein level of IL-6 was significantly elevated in the old BM [129], and IL-6 signalling was significantly enriched in the majority of BM stromal cell types, such as BL MSCs, BL VECs, CBM VECs, and CBM LEPR+ PVCs, during ageing [130]. This demonstrates that increased IL-6 signalling pre-exists in the old BM microenvironment. But the direct link between increased IL-6 signalling and AML progression during ageing is yet to be established.

First, I investigated whether upregulation of IL-6 signalling in the young mice could promote *Cebpa* mutant LICs engraftment in the young BM microenvironment, which is otherwise not AML-permissive. The hydrodynamic injection was used to activate IL-6 signalling in the young BM stromal cells, which could effectively elevate systemic IL-6 levels for at least 7 days and generate aged haematopoietic progenitor phenotypes in the young mice [130]. Therefore, to assess the role of elevated IL-6 signalling in AML progression, I hydrodynamically injected non-irradiated young mice with 1μg of empty or IL-6-encoding plasmids, allowing the IL-6 expression

level to increase and peak for 72 hours, followed by transplanting the mice with purified LICs (Figure 5.5A). However, I did not observe overt BM engraftment of the LICs in the recipients injected with either the empty or IL-6-encoding plasmids when the mice were sacrificed 28 days post-transplantation (Figure 5.5B). Even though detectable engraftment of the LICs was observed in the PB of one of the IL-6-overexpressing mice at 14 days post-transplantation, it no longer existed by 21 days post-transplantation (Figure 5.5C). Therefore, overexpression of IL-6 by the hydrodynamic injection was ineffective in remodelling the young BM into an AML growth-permissive microenvironment like the old.

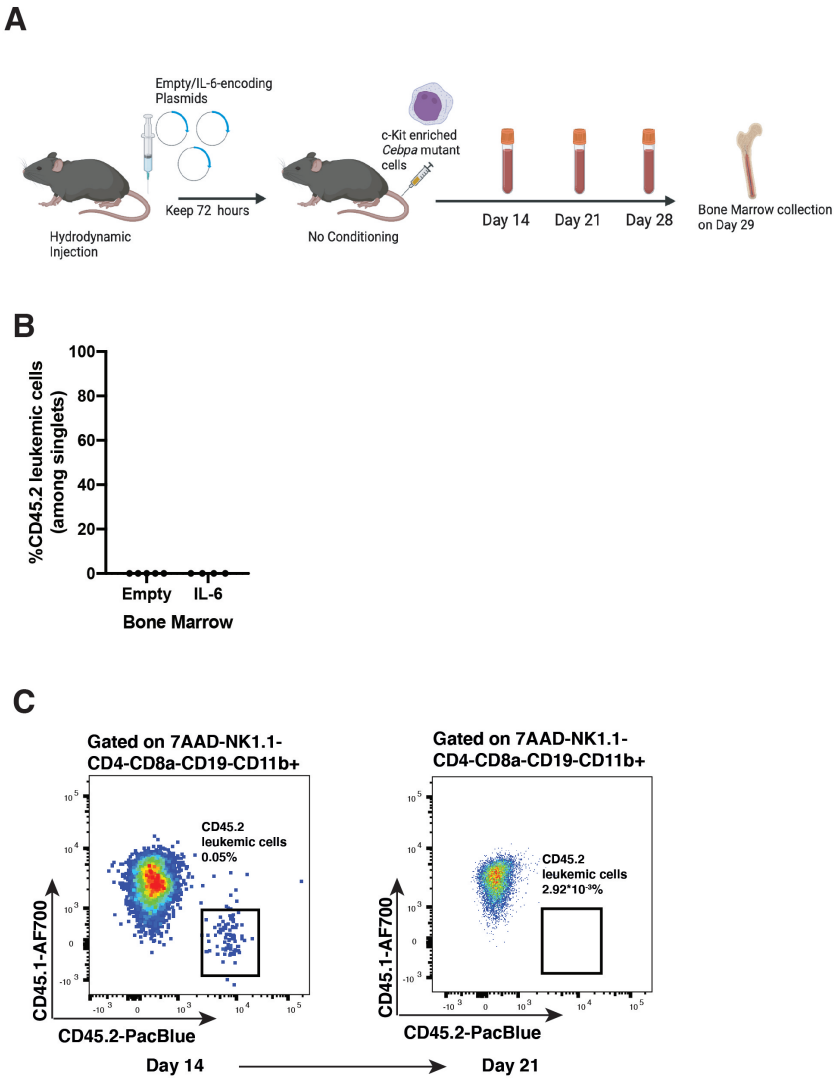

Figure 5.5. Overexpression of IL-6 by hydrodynamic injection could not remodel the young BM niche into an AML-permissive environment like the old

(A) Schematic of investigating the role of increased IL-6-signalling in AML progression in

vivo. Young CD45.1 female mice (2-3 months old) were hydrodynamically injected with empty (1μg) or IL-6-encoding (1μg) plasmids. 72 hours after the injection, each mouse was transplanted with 25,000 purified CD45.2 bi-allelic Cebpa mutant LICs and monitored for up to 4 weeks.
(B) Histogram showing the frequency of CD45.2+ leukaemia cells in the BM of mice hydrodynamically injected with empty (N=5) or IL-6-encoding (N=4) plasmids 4 weeks after LIC transplantation. Data from 1 experiment.
(C) Representative flow cytometry of CD45.2+ leukaemia cells in the PB 2 (left) and 3 (right) weeks after LIC transplantation. The flow cytometry of the only mouse with detectable CD45.2+ leukaemic engraftment in the PB (> 0.01% of 7AAD-Singlets) at 2 weeks is shown.

The Nerlov laboratory showed that treating the aged mice with a neutralising anti-IL-6 antibody inhibited the age-dependent IL-6 signalling in the BM stromal cells and reversed age-associated erythroid progenitor phenotypes. Therefore, it was hypothesised that IL-6 inhibition should reduce BM LIC engraftment in the non-conditioned aged mice if increased IL-6 signalling in the BM stromal cells was important to the superior engraftment of the LICs in the aged BM. To test this hypothesis, aged recipient mice were first intraperitoneally injected with the IL-6 neutralising or isotype control antibodies 3 times as previously described [130]. Then, the mice were transplanted with purified LICs 1 week after the last treatment and sacrificed 28 days after LIC transplantation (Figure 5.6A). Consistent with the finding described in Thomas A, 2018 [56], the LICs did engraft and propagate highly efficiently in the BM of almost all the aged mice. However, no significant difference in the BM or spleen LIC engraftment was observed between the differently treated mice (Figure 5.6B-C). Taken together, these results suggest that increased IL-6 signalling in the old BM stromal cells could not contribute to the increased ability of the aged BM microenvironment to propagate AML, despite IL-6 signalling's role in haematopoietic ageing.

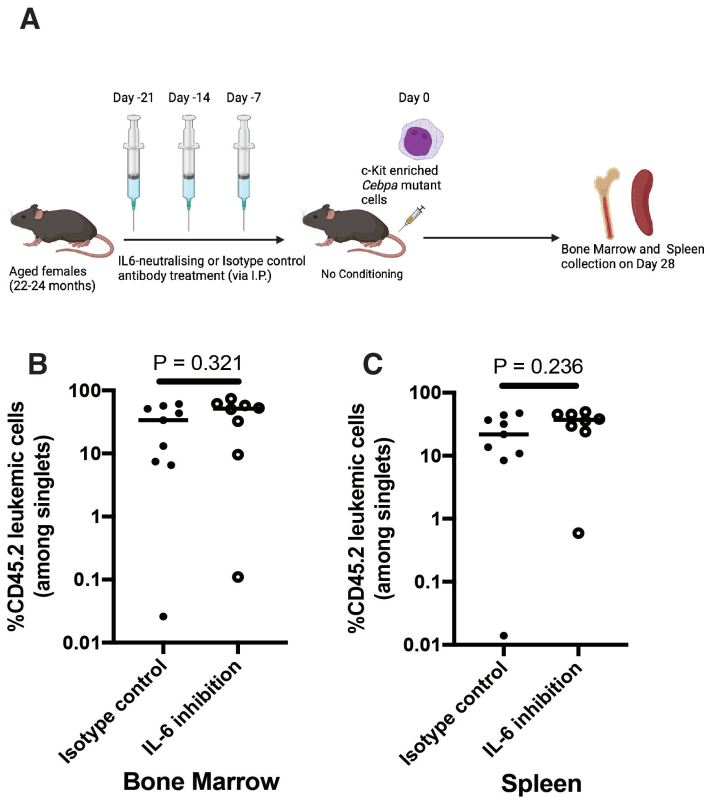

Figure 5.6. Blocking IL-6-signalling* in vivo *could not counteract the age-related increased AML propagation

(A) Schematic of examining the effect of in vivo IL-6-signalling inhibition on age-related AML propagation. Old CD45.1 female mice (22-24 months old) were treated with isotype control or IL-6 neutralising antibodies. 1 week after the last treatment, each mouse was transplanted with 25,000 purified CD45.2 bi-allelic Cebpa mutant LICs and monitored for up to 4 weeks.
(B) Scatter plot showing the frequency of CD45.2+ leukaemia cells in the BM of mice treated with isotype control (N=9) or IL-6 neutralising antibodies (N=8). Data from 2 experiments. The p-value was calculated using a Mann-Whitney test.
(C) Analysis as in (B) showing the frequency of CD45.2+ leukaemia cells in the spleen.

5.4 CCR1 signalling facilitated AML progression in the aged BM

Another pro-inflammatory cytokine signalling that could contribute to the increased AML growth permissiveness of the aged BM microenvironment is the CCR1 signalling. First, protein levels of key CCR1 ligands, including CCL3 and CCL5, have been shown to increase in the BM of *MLL-AF9* AML mice and old mice [97, 128, 235]. CCL3 levels have also been shown to increase in the elderly popu-

lation and AML patients [90, 241]. Consistent with this, I observed elevated CCL3 levels in both the AML-conditioned (vs mock transplanted) (Figure 3.10F) and old (vs young) BM fluid (Figure 5.7A). Elevated CCL5 levels were also observed in the AML-conditioned and old (vs young) BM fluid (Figure 5.7B). In addition, as described in previous chapters, CCR1 signalling is essential for LIC maintenance. These findings suggest that the aged BM microenvironment may promote AML progression by providing pre-existing high levels of CCR1 ligands, which could activate CCR1 signalling in the LICs and propagate them.

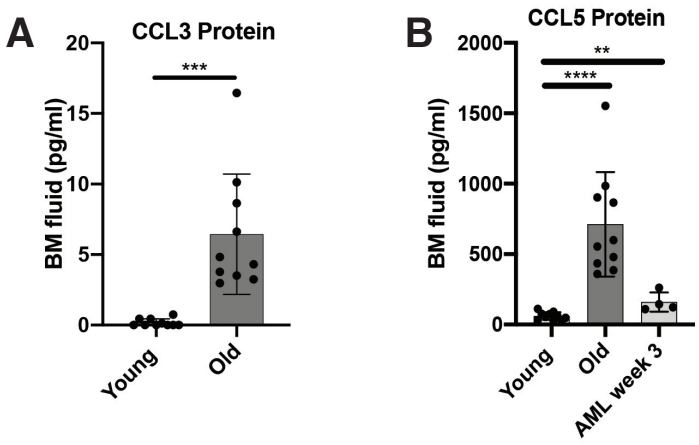

Figure 5.7. CCR1 ligands constitutively presented in aged BM at levels similar to those generated by AML conditioning

(A) Histogram showing the levels of CCL3 protein in BM fluid from young (2-3 months old) and aged (22-24 months old) mice (N=10/condition). Protein levels were measured by ELISA. Data are from 2 experiments. P-value was calculated using the Student's t-test. (B) Histogram as in (A) showing the levels of CCL5 protein in BM fluid from young (2-3 months old) (N=10), aged (22-24 months old) mice (N=10), and LIC transplanted mice (N=4) at the 3-week time point.

Considering that CCL3 and CCL5 are constitutively present in aged BM at levels similar to those generated by AML conditioning, an outstanding question that remains to be addressed is where these CCR1 ligands come from. RNAseq of the aged BM stromal cells showed significant overexpression of *Ccl3* and *Ccl6* in CBM VECs and OBs, respectively (Figure 5.8A, C). The expression level of *Ccl5* and *Ccl9* was also examined, but statistically significant overexpression was observed in none of the aged (vs young) BM stromal cell types (Figure 5.8B, D). This still raised the possibility that the aged BM stromal cells could generate an age-related CCR1 ligand-enriched BM environment by expressing and secreting increased levels of the

CCR1 ligands. Even though the primary cellular source of CCL3 in the BM is not clear yet and may not be the BM stromal cells [229], a previous study showed that old bone-derived stromal cells had increased *Ccl5* mRNA expression levels compared to the young [131]. Therefore, to investigate whether the aged BM stromal cells were indeed the primary cellular source of the high CCR1 ligand protein level in the aged BM cavity, multiplex qRT-PCR was performed to compare the mRNA expression level of key CCR1 ligand-encoding genes in old vs young BM stromal cells.

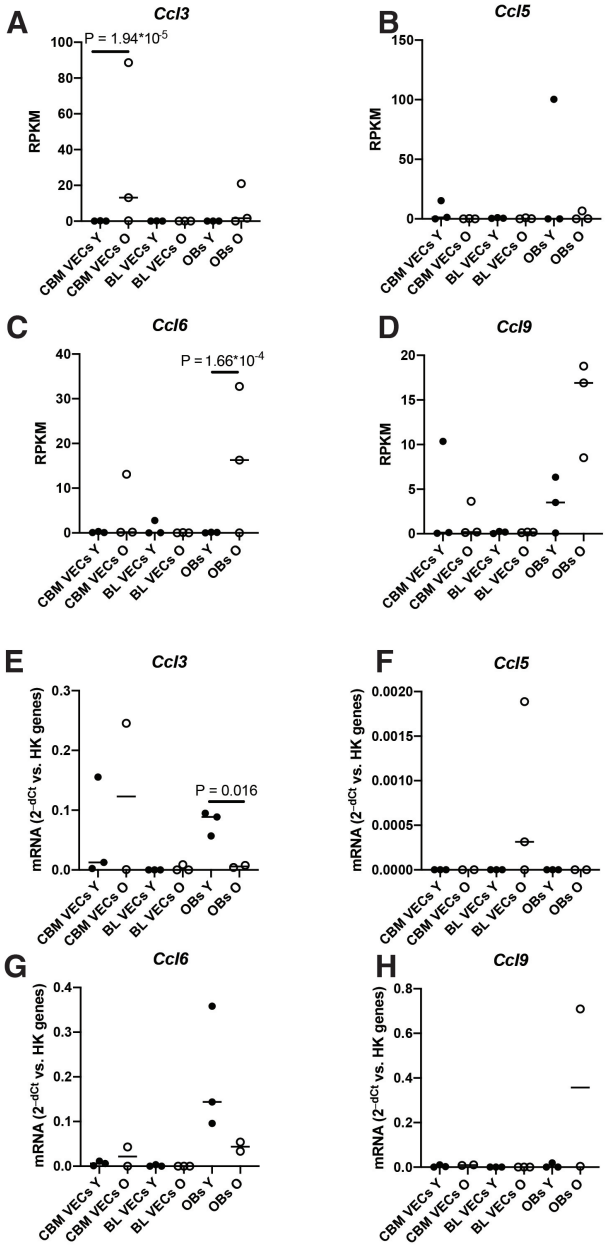

Figure 5.8. Aged BM stromal cells were not the main cellular source of increased CCR1 ligands in the aged BM

(A) Scatter plot comparing expression levels of Ccl3 measured by RNAseq (as in Figure 5.1) between the aged and young BM stromal cell populations. RPKM values are shown, N = 3 per indicated cell population. P-value was calculated by DESeq2.
(B) Scatter plot as in (A) for Ccl5.
(C) Scatter plot as in (A) for Ccl6.
(D) Scatter plot as in (A) for Ccl9.
(E) Scatter plot comparing expression levels of Ccl3 measured by multiplex qRT-PCR between the aged and young BM stromal cell populations. Values shown are $2^{-\Delta Ct}$ normalised to the geometric mean of the 2 housekeeping genes, B2m and Gapdh. N = 3 per indicated cell population from 1 experiment. Each biological replicate was the average of at least 2 technical replicates. P-values were calculated using the Student's t-test.
(F) Scatter plot as in (E) for Ccl5.
(G) Scatter plot as in (E) for Ccl6.
(H) Scatter plot as in (E) for Ccl9.

I focused on CBM VECs and OBs (Figure 5.9A-B) as the RNAseq result indicated that they were the only 2 BM stromal cell populations with age-related CCR1 ligand-encoding gene upregulation as shown in Figure 5.8A-D. However, the multiplex qRT-PCR showed no significant overexpression of any key CCR1 ligand-encoding gene in the aged BM stromal cell population (Figure 5.8E-H). This suggests that the aged BM stromal cells are not the predominant source of the elevated CCR1 ligands in the aged BM cavity. They may come from haematopoietic cell populations, such as macrophages or other distal sites. Besides, both CCL3 and CCL5 are part of the senescence-associated secretory phenotype (SASP) [242]. Therefore, CCL3 and CCL5 may come from senescent cells in the aged mouse.

To test the possibility that ageing provides an environment conducive to AML progression through, at least in part, upregulation of CCR1 ligands, I transplanted non-irradiated aged mice with purified LICs and allowed leukaemia cells to engraft and proliferate for 10 days. Then the mice were treated for 10 consecutive days with a selective CCR1 antagonist BL5923 (Novartis) or vehicle [176]. The CCR1 antagonist BL5923 has been shown to be effective in inhibiting CCR1 signalling and alleviating lupus nephritis and metastatic colon cancer [177, 231]. After the last treatment (21 days after LIC transplantation), the leukaemic burden in the PB, BM, and spleen was measured by flow cytometry (Figure 5.10A). In agreement with the hypothesis that increased CCR1 signalling in the aged BM microenvironment could contribute to AML progression, the total number of BM CD45.2+ leukaemia cells and LICs were significantly lower in the old mice treated with BL5923 (Figure

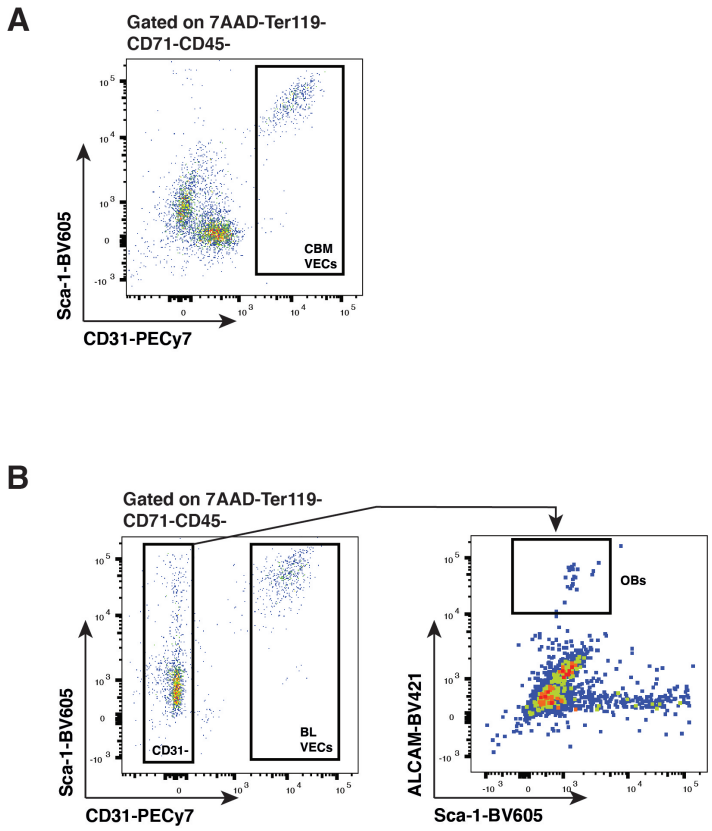

Figure 5.9. Gating of BM stromal cell populations

Flow cytometry gating strategies used to identify and sort CBM (A) and BL (B) niche cell populations.

5.10B-D). A significant reduction in the leukaemia cells was also observed in the spleen and PB (Figure 5.10E-F). This strongly demonstrated that CCR1 inhibition impaired the progression and lowered the leukaemic burden of AML in the aged BM microenvironment. To our knowledge, the effect of CCR1 inhibition has never been shown in an AML mouse model, even if pharmacological CCR1 inhibition was shown to have an anti-tumour effect in mice with MM, colon cancer, and breast cancer [243]. Chapter 4 showed that overexpression of key CCR1 ligands using the hydrodynamic injection could not facilitate AML progression in the non-irradiated young BM niche. This may suggest that mechanisms other than increased CCR1 signalling should also be involved in the increased ability of AML to propagate in the aged BM, which still warrants more investigation.

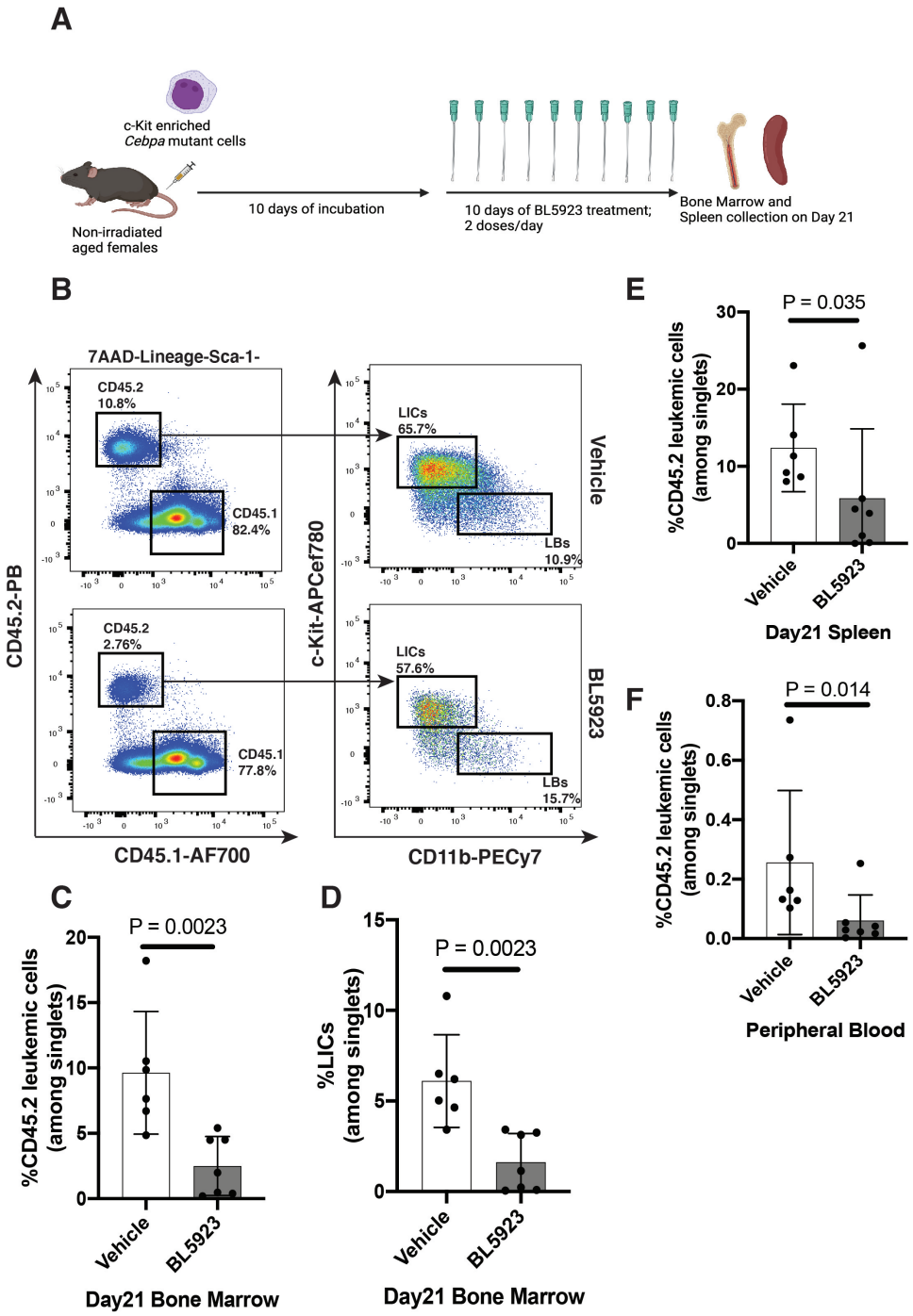

Figure 5.10. CCR1-signalling could support AML progression in the aged BM

(A) Workflow for evaluation of the role of CCR1-signalling in Cebpa mutant AML progression in aged mice. Non-irradiated aged mice (22-24 months old) were transplanted with 25,000 purified LICs. The leukaemia cells were allowed to engraft for 10 days before the treatment started.
(B) Representative flow cytometry plot comparing BM leukaemic burden in the aged leukaemia mice treated with vehicle or the selective CCR1 antagonist (BL5923). Percentages shown are gated populations as a percentage of the parental gate.
(C) Histogram showing the frequency of CD45.2+ leukaemia cells in the BM of vehicle

(N=6) or BL5923 (N=7) treated aged mice 3 weeks after LIC transplantation. Data from 2 experiments. The p-value was calculated using a Mann-Whitney test.
(D) Histogram as in (C) showing the frequency of LICs in the BM.
(E) Histogram as in (C) showing the frequency of CD45.2+ leukaemia cells in the spleen.
(F) Histogram as in (C) showing the frequency of CD45.2+ leukaemia cells in the PB.

5.5 Discussion

Even though AML- and age-induced BM microenvironmental changes have been widely studied and presumably share several common features, such as elevated levels of pro-inflammatory cytokines and disruption of endosteal vessels, a systemic comparison of the transcriptional alternations to the aged and AML-exposed BM niche cells is still lacking. By performing bulk RNAseq on the BM niche cells sorted from old or AML mice, we were able to identify a number of overlapping DEGs and de-regulated process networks, including several inflammation-associated signalling pathways. This implies the potential benefit of targeting aberrantly increased pro-inflammatory cytokines and the relevant signalling pathways for elderly AML patients.

IL-6 signalling has been implicated in the development of several types of myeloid malignancies other than AML: genetic or pharmacological blocking of IL-6 signalling was able to delay CML onset and prevent clonal haematopoiesis from progressing to CML-like disease [244, 245]. Regarding AML, it was shown that AML cell-secreted IL-6 was able to induce anaemia [102], but the role of IL-6 signalling in AML establishment and propagation still needs further investigation. Here, I demonstrated that inhibition of IL-6 signalling using an anti-IL-6 monoclonal antibody was not able to prevent *Cebpa* mutant AML progression in the aged mice. Likewise, the hydrodynamic injection of 1μg of IL-6-encoding plasmid failed to remodel the otherwise not AML-permissive young BM and support the expansion of the *Cebpa* mutant AML. As the *Cebpa* mutant LICs used for transplantations were already transformed [13], these results indicate that IL-6 signalling may not have implications for maintaining the leukaemogenic potential of the LICs or propagating them. These results also question the role of IL-6 signalling in age-related predisposition to AML despite the elevated IL-6 levels in the aged BM. Of course, it would be interesting to treat the aged leukaemia mice with IL-6-neutralising antibodies both before and after the LIC transplantation or use a more sustainable way of IL-6 overexpression, such as regular administration of recombinant IL-6, before jumping to a definitive conclusion. With respect to delineating the role of IL-6 signalling in leukaemogenesis, it would be worthwhile to transplant the preleukaemic FL cells carrying the *Cebpa*

mutations into recipient mice following IL-6 blockade or overexpression and investigate whether the onset of overt AML can be delayed or advanced [13]. However, given the long latency period of around 20 weeks in mice transplanted with the FL cells, such experiments may only be feasible when using the young recipient mice [13]. Given the effect of IL-6 blocking on improving the survival of human AML xenograft mice, albeit predominantly by improving AML-induced anaemia rather than directly suppressing AML progression [102], IL-6 blocking is still a promising therapeutic approach for elderly AML patients even though the role of IL-6 in AML establishment and propagation remains unclear.

Even if intrinsic CCR1 signalling had been shown to be essential for the maintenance of CML LICs and MM tumour dissemination [197, 198], it was not clear whether it had an effect on AML LIC maintenance [229]. The results described in the current Chapter and Chapter 4 unequivocally demonstrate that CCR1 signalling is essential for AML LIC maintenance and constitutes an AML-dependency that is increased during ageing, contributing directly to the increased ability of the transformed *Cebpa* mutant LICs to propagate in the aged BM. However, the impact of intrinsic CCR1 signalling on AML establishment remains obscure until more experiments are done (e.g. transplantation of CCR1-deficient FL cells carrying the *Cebpa* mutations). Nevertheless, the results underpin the potential benefit of CCR1 signalling inhibition for elderly AML patients. In particular, this can benefit frail elderly patients who would otherwise only receive less intensive chemotherapy or the best supportive care. Besides, similar to IL-6, CCL3 also contributes to AML-induced defective erythropoiesis [97], which further justifies the potential use of selective CCR1 inhibitors in AML patients. Targeting senescent cells may also constitute a therapeutic strategy in AML, as has previously been proposed for cancers in general [246]. Even though the result from this chapter was focused on the leukaemia-reducing effect of CCR1 inhibition in the aged leukaemia mice, considering the CCR1 dependency of the LICs shown in Chapter 4, the selective CCR1 inhibitor should presumably be able to decrease the leukaemic burden in young AML mice as well. Of course, further experiments are still warranted to confirm this. It would also be interesting to investigate the effect of combining standard chemotherapy and selective CCR1 inhibition in AML mice to investigate whether this combined treatment can further

reduce the leukaemic burden and prolong the leukaemia-free survival of AML mice.

Chapter 6

Conclusions

The studies presented in Chapters 3 and 4 identified and functionally validated 2 cytokine receptor genes, *Ccr1* and *Il7r*, as *in vivo* dependencies in a murine model of bi-allelic *CEBPA* mutant AML by virtue of CRISPR-Cas9-based screens. This suggests an intrinsic role for CCR1 and IL7R signalling in *CEBPA* mutant AML propagation. In addition, mirroring the increased expression of *Ccr1* and *Il7r* in murine *Cebpa* mutant AML cells, both *CCR1* and *IL7R* were overexpressed across various cytogenetic and molecular genetic subtypes of human AML, and high *CCR1* or *IL7R* expression in *CEBPA* mutant AML patients led to distinct gene expression profiles. Furthermore, high *CCR1* expression was correlated with poor OS, consistent with CCR1 signalling playing a general and intrinsic role in AML progression. These studies not only provide novel AML dependencies, which can potentially be used for targeting LSCs but also offer a platform for conducting targeted CRISPR-based screens in a strictly BM microenvironment-dependent context.

The study presented in Chapter 5 investigated the molecular basis for the increased ability of AML-propagating cells to expand in an aged, compared to a young, BM microenvironment and identified CCR1 signalling as a critical mediator of this. However, IL-6 signalling, which had been implicated in haematopoietic ageing and CML development, seemed to play a minimal, if any, role in the increased ability of AML to propagate in the aged BM. Therefore, CCR1 may constitute a molecular target in AML in the elderly.

Together, these studies highlight the role of BM microenvironmental factors, CCR1 ligands in particular, in *CEBPA* mutant AML development and age-related *CEBPA* mutant AML predisposition. However, more experimental studies are warranted to elucidate whether this finding can apply to other subtypes of AML. In addition, given the significant similarities between AML-conditioned and aged BM stromal cell transcriptomes, further investigation may identify additional BM microenvironmental AML dependencies that contribute to age-related AML predisposition.

References

[1] Hartmut Döhner, Daniel J Weisdorf, and Clara D Bloomfield. Acute myeloid leukemia. *New England Journal of Medicine*, 373(12):1136–1152, 2015.

[2] Elihu Estey and Hartmut Döhner. Acute myeloid leukaemia. *The Lancet*, 368(9550):1894–1907, 2006.

[3] Milena Sant, Claudia Allemani, Carmen Tereanu, Roberta De Angelis, Riccardo Capocaccia, Otto Visser, Rafael Marcos-Gragera, Marc Maynadié, Arianna Simonetti, Jean-Michel Lutz, et al. Incidence of hematologic malignancies in europe by morphologic subtype: results of the haemacare project. *Blood, The Journal of the American Society of Hematology*, 116(19):3724–3734, 2010.

[4] Rory M Shallis, Rong Wang, Amy Davidoff, Xiaomei Ma, and Amer M Zeidan. Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges. *Blood reviews*, 36:70–87, 2019.

[5] Elli Papaemmanuil, Moritz Gerstung, Lars Bullinger, Verena I Gaidzik, Peter Paschka, Nicola D Roberts, Nicola E Potter, Michael Heuser, Felicitas Thol, Niccolo Bolli, et al. Genomic classification and prognosis in acute myeloid leukemia. *New England Journal of Medicine*, 374(23):2209–2221, 2016.

[6] Myriam Alcalay, Annette Orleth, Carla Sebastiani, Natalia Meani, Ferdinando Chiaradonna, Cristina Casciari, Maria Teresa Sciurpi, Vania Gelmetti, Daniela Riganelli, Saverio Minucci, et al. Common themes in the pathogenesis of acute myeloid leukemia. *Oncogene*, 20(40):5680–5694, 2001.

[7] Daniel A Pollyea and Craig T Jordan. Therapeutic targeting of acute myeloid

leukemia stem cells. *Blood, The Journal of the American Society of Hematology*, 129(12):1627–1635, 2017.

[8] Xavier Roussel, Etienne Daguindau, Ana Berceanu, Yohan Desbrosses, Walid Warda, Mathieu Neto da Rocha, Rim Trad, Eric Deconinck, Marina Deschamps, and Christophe Ferrand. Acute myeloid leukemia: From biology to clinical practices through development and pre-clinical therapeutics. *Frontiers in Oncology*, 10:599933, 2020.

[9] Stefan Fröhling, Claudia Scholl, D Gary Gilliland, and Ross L Levine. Genetics of myeloid malignancies: pathogenetic and clinical implications. *Journal of Clinical Oncology*, 23(26):6285–6295, 2005.

[10] Iris T Chan, Jeffery L Kutok, Ifor R Williams, Sarah Cohen, Lauren Kelly, Hirokazu Shigematsu, Leisa Johnson, Koichi Akashi, David A Tuveson, Tyler Jacks, et al. Conditional expression of oncogenic k-ras from its endogenous promoter induces a myeloproliferative disease. *The Journal of clinical investigation*, 113(4):528–538, 2004.

[11] PJ Valk, David T Bowen, Marion E Frew, Anne C Goodeve, B Lowenberg, and John T Reilly. Second hit mutations in the rtk/ras signaling pathway in acute myeloid leukemia with inv (16). *Haematologica*, 89(1):106–106, 2004.

[12] Louise M Kelly and D Gary Gilliland. Genetics of myeloid leukemias. *Annual review of genomics and human genetics*, 3(1):179–198, 2002.

[13] Oxana Bereshchenko, Elena Mancini, Susan Moore, Daniel Bilbao, Robert Månsson, Sidinh Luc, Amit Grover, Sten Eirik W Jacobsen, David Bryder, and Claus Nerlov. Hematopoietic stem cell expansion precedes the generation of committed myeloid leukemia-initiating cells in c/ebpα mutant aml. *Cancer cell*, 16(5):390–400, 2009.

[14] Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *New England Journal of Medicine*, 368(22):2059–2074, 2013.

[15] Lars Velten, Benjamin A Story, Pablo Hernández-Malmierca, Simon Raffel, Daniel R Leonce, Jennifer Milbank, Malte Paulsen, Aykut Demir, Chelsea Szu-Tu, Robert Frömel, et al. Identification of leukemic and pre-leukemic stem cells by clonal tracking from single-cell transcriptomics. *Nature communications*, 12(1):1–13, 2021.

[16] Antonija Kreso and John E Dick. Evolution of the cancer stem cell model. *Cell stem cell*, 14(3):275–291, 2014.

[17] David Vetrie, G Vignir Helgason, and Mhairi Copland. The leukaemia stem cell: similarities, differences and clinical prospects in cml and aml. *Nature Reviews Cancer*, 20(3):158–173, 2020.

[18] Daniel Thomas and Ravindra Majeti. Biology and relevance of human acute myeloid leukemia stem cells. *Blood, The Journal of the American Society of Hematology*, 129(12):1577–1585, 2017.

[19] Tsvee Lapidot, Christian Sirard, Josef Vormoor, Barbara Murdoch, Trang Hoang, Julio Caceres-Cortes, Mark Minden, Bruce Paterson, Michael A Caligiuri, and John E Dick. A cell initiating human acute myeloid leukaemia after transplantation into scid mice. *Nature*, 367(6464):645–648, 1994.

[20] David C Taussig, Farideh Miraki-Moud, Fernando Anjos-Afonso, Daniel J Pearce, Kirsty Allen, Christopher Ridler, Debra Lillington, Heather Oakervee, Jamie Cavenagh, Samir G Agrawal, et al. Anti-cd38 antibody–mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. *Blood, The Journal of the American Society of Hematology*, 112(3):568–575, 2008.

[21] David C Taussig, Jacques Vargaftig, Farideh Miraki-Moud, Emmanuel Griessinger, Kirsty Sharrock, Tina Luke, Debra Lillington, Heather Oakervee, Jamie Cavenagh, Samir G Agrawal, et al. Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the cd34- fraction. *Blood, The Journal of the American Society of Hematology*, 115(10):1976–1984, 2010.

[22] Lynn Quek, Georg W Otto, Catherine Garnett, Ludovic Lhermitte, Dimitris Karamitros, Bilyana Stoilova, I-Jun Lau, Jessica Doondeea, Batchimeg Usukhbayar, Alison Kennedy, et al. Genetically distinct leukemic stem cells in human cd34- acute myeloid leukemia are arrested at a hemopoietic precursor-like stage. *Journal of Experimental Medicine*, 213(8):1513–1535, 2016.

[23] Nicolas Goardon, Emanuele Marchi, Ann Atzberger, Lynn Quek, Anna Schuh, Shamit Soneji, Petter Woll, Adam Mead, Kate A Alford, Raj Rout, et al. Coexistence of lmpp-like and gmp-like leukemia stem cells in acute myeloid leukemia. *Cancer cell*, 19(1):138–152, 2011.

[24] LI Shulush, S Zandi, A Mitchell, WC Chen, JM Brandwein, V Gupta, et al. Identification of preleukemic stem cells in acute leukemia. *Nature*, 506:328–333, 2014.

[25] Andrei V Krivtsov, David Twomey, Zhaohui Feng, Matthew C Stubbs, Yingzi Wang, Joerg Faber, Jason E Levine, Jing Wang, William C Hahn, D Gary Gilliland, et al. Transformation from committed progenitor to leukaemia stem cell initiated by mll–af9. *Nature*, 442(7104):818–822, 2006.

[26] Andrei V Krivtsov, Maria E Figueroa, Amit U Sinha, Matthew C Stubbs, Zhaohui Feng, Peter JM Valk, Ruud Delwel, Konstanze Döhner, Lars Bullinger, Andrew L Kung, et al. Cell of origin determines clinically relevant subtypes of mll-rearranged aml. *Leukemia*, 27(4):852–860, 2013.

[27] Tim CP Somervaille and Michael L Cleary. Identification and characterization of leukemia stem cells in murine mll-af9 acute myeloid leukemia. *Cancer cell*, 10(4):257–268, 2006.

[28] Anna Van Rhenen, Nicole Feller, Angèle Kelder, August H Westra, Elwin Rombouts, Sonja Zweegman, Marjolein A Van Der Pol, Quinten Waisfisz, Gert J Ossenkoppele, and Gerrit Jan Schuurhuis. High stem cell frequency in acute myeloid leukemia at diagnosis predicts high minimal residual disease and poor survival. *Clinical cancer research*, 11(18):6520–6527, 2005.

[29] Andrew J Gentles, Sylvia K Plevritis, Ravindra Majeti, and Ash A Alizadeh.

Association of a leukemic stem cell gene expression signature with clinical outcomes in acute myeloid leukemia. *Jama*, 304(24):2706–2715, 2010.

[30] Fumihiko Ishikawa, Shuro Yoshida, Yoriko Saito, Atsushi Hijikata, Hiroshi Kitamura, Satoshi Tanaka, Ryu Nakamura, Toru Tanaka, Hiroko Tomiyama, Noriyuki Saito, et al. Chemotherapy-resistant human aml stem cells home to and engraft within the bone-marrow endosteal region. *Nature biotechnology*, 25(11):1315–1321, 2007.

[31] Liran I Shlush, Amanda Mitchell, Lawrence Heisler, Sagi Abelson, Stanley WK Ng, Aaron Trotman-Grant, Jessie JF Medeiros, Abilasha Rao-Bhatia, Ivana Jaciw-Zurakowsky, Rene Marke, et al. Tracing the origins of relapse in acute myeloid leukaemia to stem cells. *Nature*, 547(7661):104–108, 2017.

[32] Hartmut Döhner, Elihu Estey, David Grimwade, Sergio Amadori, Frederick R Appelbaum, Thomas Büchner, Hervé Dombret, Benjamin L Ebert, Pierre Fenaux, Richard A Larson, et al. Diagnosis and management of aml in adults: 2017 eln recommendations from an international expert panel. *Blood, The Journal of the American Society of Hematology*, 129(4):424–447, 2017.

[33] Hagop Kantarjian, Tapan Kadia, Courtney DiNardo, Naval Daver, Gautam Borthakur, Elias Jabbour, Guillermo Garcia-Manero, Marina Konopleva, and Farhad Ravandi. Acute myeloid leukemia: current progress and future directions. *Blood cancer journal*, 11(2):1–25, 2021.

[34] Laura F Newell and Rachel J Cook. Advances in acute myeloid leukemia. *Bmj*, 375, 2021.

[35] Matthew J Christopher, Allegra A Petti, Michael P Rettig, Christopher A Miller, Ezhilarasi Chendamarai, Eric J Duncavage, Jeffery M Klco, Nicole M Helton, Michelle O'Laughlin, Catrina C Fronick, et al. Immune escape of relapsed aml cells after allogeneic transplantation. *New England Journal of Medicine*, 379(24):2330–2341, 2018.

[36] Ramya Nair, Alejandro Salinas-Illarena, and Hanna-Mari Baldauf. New strategies to treat aml: novel insights into aml survival pathways and combination therapies. *Leukemia*, 35(2):299–311, 2021.

[37] Eric S Winer and Richard M Stone. Novel therapy in acute myeloid leukemia (aml): moving toward targeted approaches. *Therapeutic advances in hematology*, 10:2040620719860645, 2019.

[38] Michael D Amatangelo, Lynn Quek, Alan Shih, Eytan M Stein, Mikhail Roshal, Muriel D David, Benoit Marteyn, Noushin Rahnamay Farnoud, Stephane de Botton, Olivier A Bernard, et al. Enasidenib induces acute myeloid leukemia cell differentiation to promote clinical response. *Blood, The Journal of the American Society of Hematology*, 130(6):732–741, 2017.

[39] Sung Choe, Hongfang Wang, Courtney D DiNardo, Eytan M Stein, Stéphane de Botton, Gail J Roboz, Jessica K Altman, Alice S Mims, Justin M Watts, Daniel A Pollyea, et al. Molecular mechanisms mediating relapse following ivosidenib monotherapy in idh1-mutant relapsed or refractory aml. *Blood advances*, 4(9):1894–1905, 2020.

[40] Koen Schepers, Timothy B Campbell, and Emmanuelle Passegué. Normal and leukemic stem cell niches: insights and therapeutic opportunities. *Cell stem cell*, 16(3):254–267, 2015.

[41] Yuki Kagoya, Akihide Yoshimi, Keisuke Kataoka, Masahiro Nakagawa, Keiki Kumano, Shunya Arai, Hiroshi Kobayashi, Taku Saito, Yoichiro Iwakura, Mineo Kurokawa, et al. Positive feedback between nf-κb and tnf-α promotes leukemia-initiating cell capacity. *The Journal of clinical investigation*, 124(2):528–542, 2014.

[42] Sophia K Khaldoyanidi, Antreas Hindoyan, Anthony Stein, and Marion Subklewe. Leukemic stem cells as a target for eliminating acute myeloid leukemia: Gaps in translational research. *Critical Reviews in Oncology/Hematology*, page 103710, 2022.

[43] Ulrike Bacher, Susanne Schnittger, Katja Macijewski, Vera Grossmann, Alexander Kohlmann, Tamara Alpermann, Andreas Kowarsch, Niroshan Nadarajah, Wolfgang Kern, Claudia Haferlach, et al. Multilineage dysplasia does not influence prognosis in cebpa-mutated aml, supporting the who

proposal to classify these patients as a unique entity. *Blood, The Journal of the American Society of Hematology*, 119(20):4719–4722, 2012.

[44] Annika Dufour, Friederike Schneider, Klaus H Metzeler, Eva Hoster, Stephanie Schneider, Evelyn Zellmeier, Tobias Benthaus, Maria-Cristina Sauerland, Wolfgang E Berdel, Thomas Büchner, et al. Acute myeloid leukemia with biallelic cebpa gene mutations and normal karyotype represents a distinct genetic entity associated with a favorable clinical outcome. *Journal of clinical oncology*, 28(4):570–577, 2010.

[45] Daniel A Arber, Attilio Orazi, Robert Hasserjian, Jürgen Thiele, Michael J Borowitz, Michelle M Le Beau, Clara D Bloomfield, Mario Cazzola, and James W Vardiman. The 2016 revision to the world health organization classification of myeloid neoplasms and acute leukemia. *Blood, The Journal of the American Society of Hematology*, 127(20):2391–2405, 2016.

[46] Claus Nerlov. The c/ebp family of transcription factors: a paradigm for interaction between gene expression and proliferation control. *Trends in cell biology*, 17(7):318–324, 2007.

[47] Claus Nerlov. C/ebpα mutations in acute myeloid leukaemias. *Nature Reviews Cancer*, 4(5):394–400, 2004.

[48] Daniel G Tenen, Robert Hromas, Jonathan D Licht, and Dong-Er Zhang. Transcription factors, normal myeloid development, and leukemia. *Blood*, 90(2):489–519, 1997.

[49] Kelly M McNagny, Michael H Sieweke, Gabriele Döderlein, Thomas Graf, and Claus Nerlov. Regulation of eosinophil-specific gene expression by a c/ebp–ets complex and gata-1. *The EMBO journal*, 17(13):3669–3680, 1998.

[50] Claus Nerlov, Kelly M McNagny, Gabriele Döderlein, Elisabeth Kowenz-Leutz, and Thomas Graf. Distinct c/ebp functions are required for eosinophil lineage commitment and maturation. *Genes & development*, 12(15):2413–2423, 1998.

[51] Beatrix A Slomiany, Kenneth L D'Arigo, Margaret M Kelly, and David T

Kurtz. C/ebpα inhibits cell growth via direct repression of e2f-dp-mediated transcription. *Molecular and cellular biology*, 20(16):5986–5997, 2000.

[52] Thomas Pabst, Beatrice U Mueller, Pu Zhang, Hanna S Radomska, Sailaja Narravula, Susanne Schnittger, Gerhard Behre, Wolfgang Hiddemann, and Daniel G Tenen. Dominant-negative mutations of cebpa, encoding ccaat/enhancer binding protein-α (c/ebpα), in acute myeloid leukemia. *Nature genetics*, 27(3):263–270, 2001.

[53] Adrian F Gombart, Wolf-K Hofmann, Seiji Kawano, Seisho Takeuchi, Utz Krug, Scott H Kwok, Renee J Larsen, Hiroya Asou, Carl W Miller, Dieter Hoelzer, et al. Mutations in the gene encoding the transcription factor ccaat/enhancer binding protein α in myelodysplastic syndromes and acute myeloid leukemias. *Blood, The Journal of the American Society of Hematology*, 99(4):1332–1340, 2002.

[54] Bas J Wouters, Bob Löwenberg, Claudia AJ Erpelinck-Verschueren, Wim LJ van Putten, Peter JM Valk, and Ruud Delwel. Double cebpa mutations, but not single cebpa mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. *Blood, The Journal of the American Society of Hematology*, 113(13):3088–3091, 2009.

[55] Peggy Kirstetter, Mikkel B Schuster, Oksana Bereshchenko, Susan Moore, Heidi Dvinge, Elke Kurz, Kim Theilgaard-Mönch, Robert Månsson, Thomas Å Pedersen, Thomas Pabst, et al. Modeling of c/ebpα mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells. *Cancer cell*, 13(4):299–310, 2008.

[56] Alexander William Thomas. *How does leukemia disrupt hematopoiesis?* PhD thesis, University of Oxford, 2018.

[57] Marwa Almosailleakh and Juerg Schwaller. Murine models of acute myeloid leukaemia. *International journal of molecular sciences*, 20(2):453, 2019.

[58] Cristina Mambet, Mihaela Chivu-Economescu, Lilia Matei, Laura Georgiana Necula, Denisa Laura Dragu, Coralia Bleotu, and Carmen Cristina Dia-

conu. Murine models based on acute myeloid leukemia-initiating stem cells xenografting. *World journal of stem cells*, 10(6):57, 2018.

[59] Caroline Pabst, Jana Krosl, Iman Fares, Geneviève Boucher, Réjean Ruel, Anne Marinier, Sébastien Lemieux, Josée Hébert, and Guy Sauvageau. Identification of small molecules that support human leukemia stem cell activity ex vivo. *Nature methods*, 11(4):436–442, 2014.

[60] Sung-Yup Cho, Wonyoung Kang, Jee Yun Han, Seoyeon Min, Jinjoo Kang, Ahra Lee, Jee Young Kwon, Charles Lee, and Hansoo Park. An integrative approach to precision cancer medicine using patient-derived xenografts. *Molecules and cells*, 39(2):77, 2016.

[61] Mark Wunderlich, Benjamin Mizukawa, Fu-Sheng Chou, Christina Sexton, Mahesh Shrestha, Yogen Saunthararajah, and James C Mulloy. Aml cells are differentially sensitive to chemotherapy treatment in a human xenograft model. *Blood, The Journal of the American Society of Hematology*, 121(12):e90–e97, 2013.

[62] Thomas Farge, Estelle Saland, Fabienne De Toni, Nesrine Aroua, Mohsen Hosseini, Robin Perry, Claudie Bosc, Mayumi Sugita, Lucille Stuani, Marine Fraisse, et al. Chemotherapy-resistant human acute myeloid leukemia cells are not enriched for leukemic stem cells but require oxidative metabolismenergetic control of in vivo chemoresistance in aml. *Cancer discovery*, 7(7):716–735, 2017.

[63] M Wunderlich, FS Chou, K_A Link, B Mizukawa, RL Perry, M Carroll, and JC Mulloy. Aml xenograft efficiency is significantly improved in nod/scid-il2rg mice constitutively expressing human scf, gm-csf and il-3. *Leukemia*, 24(10):1785–1788, 2010.

[64] Ander Abarrategi, Syed A Mian, Diana Passaro, Kevin Rouault-Pierre, William Grey, and Dominique Bonnet. Modeling the human bone marrow niche in mice: From host bone marrow engraftment to bioengineering approaches. *Journal of Experimental Medicine*, 215(3):729–743, 2018.

[65] Sean J Morrison and Allan C Spradling. Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell*, 132(4):598–611, 2008.

[66] Avital Mendelson and Paul S Frenette. Hematopoietic stem cell niche maintenance during homeostasis and regeneration. *Nature medicine*, 20(8):833–846, 2014.

[67] Lei Ding, Thomas L Saunders, Grigori Enikolopov, and Sean J Morrison. Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature*, 481(7382):457–462, 2012.

[68] Lei Ding and Sean J Morrison. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature*, 495(7440):231–235, 2013.

[69] Tatsuki Sugiyama, Hiroshi Kohara, Mamiko Noda, and Takashi Nagasawa. Maintenance of the hematopoietic stem cell pool by cxcl12-cxcr4 chemokine signaling in bone marrow stromal cell niches. *Immunity*, 25(6):977–988, 2006.

[70] Bo O Zhou, Rui Yue, Malea M Murphy, James G Peyer, and Sean J Morrison. Leptin-receptor-expressing mesenchymal stromal cells represent the main source of bone formed by adult bone marrow. *Cell stem cell*, 15(2):154–168, 2014.

[71] Simón Méndez-Ferrer, Tatyana V Michurina, Francesca Ferraro, Amin R Mazloom, Ben D MacArthur, Sergio A Lira, David T Scadden, Avi Ma'ayan, Grigori N Enikolopov, and Paul S Frenette. Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *nature*, 466(7308):829–834, 2010.

[72] Yuya Kunisaki, Ingmar Bruns, Christoph Scheiermann, Jalal Ahmed, Sandra Pinho, Dachuan Zhang, Toshihide Mizoguchi, Qiaozhi Wei, Daniel Lucas, Keisuke Ito, et al. Arteriolar niches maintain haematopoietic stem cell quiescence. *Nature*, 502(7473):637–643, 2013.

[73] Ninib Baryawno, Dariusz Przybylski, Monika S Kowalczyk, Youmna Kfoury, Nicolas Severe, Karin Gustafsson, Konstantinos D Kokkaliaris, Francois

Mercier, Marcin Tabaka, Matan Hofree, et al. A cellular taxonomy of the bone marrow stroma in homeostasis and leukemia. *Cell*, 177(7):1915–1932, 2019.

[74] Genevieve M Crane, Elise Jeffery, and Sean J Morrison. Adult haematopoietic stem cell niches. *Nature Reviews Immunology*, 17(9):573–590, 2017.

[75] Sean J Morrison and David T Scadden. The bone marrow niche for haematopoietic stem cells. *Nature*, 505(7483):327–334, 2014.

[76] Brian JP Huntly and D Gary Gilliland. Leukaemia stem cells and the evolution of cancer-stem-cell research. *Nature Reviews Cancer*, 5(4):311–321, 2005.

[77] Simón Méndez-Ferrer, Dominique Bonnet, David P Steensma, Robert P Hasserjian, Irene M Ghobrial, John G Gribben, Michael Andreeff, and Daniela S Krause. Bone marrow niches in haematological malignancies. *Nature Reviews Cancer*, 20(5):285–298, 2020.

[78] Carl R Walkley, Gemma Haines Olsen, Sebastian Dworkin, Stewart A Fabb, Jeremy Swann, Grant A McArthur, Susan V Westmoreland, Pierre Chambon, David T Scadden, and Louise E Purton. A microenvironment-induced myeloproliferative syndrome caused by retinoic acid receptor γ deficiency. *Cell*, 129(6):1097–1110, 2007.

[79] Marc HGP Raaijmakers, Siddhartha Mukherjee, Shangqin Guo, Siyi Zhang, Tatsuya Kobayashi, Jesse A Schoonmaker, Benjamin L Ebert, Fatima Al-Shahrour, Robert P Hasserjian, Edward O Scadden, et al. Bone progenitor dysfunction induces myelodysplasia and secondary leukaemia. *Nature*, 464(7290):852–857, 2010.

[80] Lin Wang, Huajia Zhang, Sonia Rodriguez, Liyun Cao, Jonathan Parish, Christen Mumaw, Amy Zollman, Malgorzata M Kamoka, Jian Mu, Danny Z Chen, et al. Notch-dependent repression of mir-155 in the bone marrow niche regulates hematopoiesis in an nf-κb-dependent manner. *Cell stem cell*, 15(1):51–65, 2014.

[81] Lei Dong, Wen-Mei Yu, Hong Zheng, Mignon L Loh, Silvia T Bunting, Melinda Pauly, Gang Huang, Muxiang Zhou, Hal E Broxmeyer, David T Scadden, et al. Leukaemogenic effects of ptpn11 activating mutations in the stem cell microenvironment. *Nature*, 539(7628):304–308, 2016.

[82] Olga Blau, Wolf-Karsten Hofmann, Claudia Dorothea Baldus, Gundula Thiel, Verena Serbent, Elke Schümann, Eckhard Thiel, and Igor Wolfgang Blau. Chromosomal aberrations in bone marrow mesenchymal stroma cells from patients with myelodysplastic syndrome and acute myeloblastic leukemia. *Experimental hematology*, 35(2):221–229, 2007.

[83] Olga Blau, Claudia Dorothea Baldus, Wolf-Karsten Hofmann, Gundula Thiel, Florian Nolte, Thomas Burmeister, Seval Türkmen, Ouidad Benlasfer, Elke Schümann, Annette Sindram, et al. Mesenchymal stromal cells of myelodysplastic syndrome and acute myeloid leukemia patients have distinct genetic abnormalities compared with leukemic blasts. *Blood, The Journal of the American Society of Hematology*, 118(20):5583–5592, 2011.

[84] B Kumar, M Garcia, L Weng, X Jung, JL Murakami, X Hu, T McDonald, A Lin, AR Kumar, DL DiGiusto, et al. Acute myeloid leukemia transforms the bone marrow niche into a leukemia-permissive microenvironment through exosome secretion. *Leukemia*, 32(3):575–587, 2018.

[85] Delfim Duarte, Edwin D Hawkins, Olufolake Akinduro, Heather Ang, Katia De Filippo, Isabella Y Kong, Myriam Haltalli, Nicola Ruivo, Lenny Straszkowski, Stephin J Vervoort, et al. Inhibition of endosteal vascular niche remodeling rescues hematopoietic stem cell loss in aml. *Cell stem cell*, 22(1):64–77, 2018.

[86] Maher Hanoun, Dachuan Zhang, Toshihide Mizoguchi, Sandra Pinho, Halley Pierce, Yuya Kunisaki, Julie Lacombe, Scott A Armstrong, Ulrich Dührsen, and Paul S Frenette. Acute myelogenous leukemia-induced sympathetic neuropathy promotes malignancy in an altered hematopoietic stem cell niche. *Cell stem cell*, 15(3):365–375, 2014.

[87] Maria Krevvata, Barbara C Silva, John S Manavalan, Marta Galan-Diez,

Aruna Kode, Brya Grace Matthews, David Park, Chiyuan A Zhang, Naomi Galili, Thomas L Nickolas, et al. Inhibition of leukemia cell engraftment and disease progression in mice by osteoblasts. *Blood, The Journal of the American Society of Hematology*, 124(18):2834–2846, 2014.

[88] Jeevisha Bajaj, Takaaki Konuma, Nikki K Lytle, Hyog Young Kwon, Jailal N Ablack, Joseph M Cantor, David Rizzieri, Charles Chuah, Vivian G Oehler, Elizabeth H Broome, et al. Cd98-mediated adhesive signaling enables the establishment and propagation of acute myelogenous leukemia. *Cancer cell*, 30(5):792–805, 2016.

[89] Valerie Barbier, Johanna Erbani, Corrine Fiveash, Julie M Davies, Joshua Tay, Michael R Tallack, Jessica Lowe, John L Magnani, Diwakar R Pattabiraman, Andrew C Perkins, et al. Endothelial e-selectin inhibition improves acute myeloid leukaemia therapy by disrupting vascular niche-mediated chemoresistance. *Nature communications*, 11(1):1–15, 2020.

[90] Benjamin J Frisch, John M Ashton, Lianping Xing, Michael W Becker, Craig T Jordan, and Laura M Calvi. Functional inhibition of osteoblastic cells in an in vivo mouse model of myeloid leukemia. *Blood, The Journal of the American Society of Hematology*, 119(2):540–550, 2012.

[91] Koen Schepers, Eric M Pietras, Damien Reynaud, Johanna Flach, Mikhail Binnewies, Trit Garg, Amy J Wagers, Edward C Hsiao, and Emmanuelle Passegué. Myeloproliferative neoplasia remodels the endosteal bone marrow niche into a self-reinforcing leukemic niche. *Cell stem cell*, 13(3):285–299, 2013.

[92] Johan Skog, Tom Würdinger, Sjoerd Van Rijn, Dimphna H Meijer, Laura Gainche, William T Curry, Bob S Carter, Anna M Krichevsky, and Xandra O Breakefield. Glioblastoma microvesicles transport rna and proteins that promote tumour growth and provide diagnostic biomarkers. *Nature cell biology*, 10(12):1470–1476, 2008.

[93] Diana Passaro, Alessandro Di Tullio, Ander Abarrategi, Kevin Rouault-Pierre, Katie Foster, Linda Ariza-McNaughton, Beatriz Montaner, Probir

Chakravarty, Leena Bhaw, Giovanni Diana, et al. Increased vascular permeability in the bone marrow microenvironment contributes to disease progression and drug response in acute myeloid leukemia. *Cancer cell*, 32(3):324–341, 2017.

[94] Alyssa Carey, Christopher A Eide, Laura Newell, Elie Traer, Bruno C Medeiros, Daniel A Pollyea, Michael W Deininger, Robert H Collins, Jeffrey W Tyner, Brian J Druker, et al. Identification of interleukin-1 by functional screening as a key mediator of cellular expansion and disease progression in acute myeloid leukemia. *Cell reports*, 18(13):3204–3218, 2017.

[95] JM Horacek, T Kupsa, M Vasatova, L Jebavy, and P Zak. Biochip array technology and evaluation of serum levels of multiple cytokines and adhesion molecules in patients with newly diagnosed acute myeloid leukemia. *Experimental oncology*, 2014.

[96] Beatriz Sanchez-Correa, Juan M Bergua, Carmen Campos, Inmaculada Gayoso, Maria Jose Arcos, Helena Bañas, Sara Morgado, Javier G Casado, Rafael Solana, and Raquel Tarazona. Cytokine profiles in acute myeloid leukemia patients at diagnosis: survival is inversely correlated with il-6 and directly correlated with il-10 levels. *Cytokine*, 61(3):885–891, 2013.

[97] Y Wang, A Gao, H Zhao, P Lu, H Cheng, F Dong, Y Gong, S Ma, Y Zheng, H Zhang, et al. Leukemia cell infiltration causes defective erythropoiesis partially through mip-1α/ccl3. *Leukemia*, 30(9):1897–1908, 2016.

[98] Matheus Rodrigues Lopes, João Kleber Novais Pereira, Paula de Melo Campos, João Agostinho Machado-Neto, Fabiola Traina, Sara T Olalla Saad, and Patricia Favaro. De novo aml exhibits greater microenvironment dysregulation compared to aml with myelodysplasia-related changes. *Scientific reports*, 7(1):1–12, 2017.

[99] Sara Civini, Ping Jin, Jiaqiang Ren, Marianna Sabatino, Luciano Castiello, Jianjian Jin, Huan Wang, Yuanlong Zhao, Francesco Marincola, and David Stroncek. Leukemia cells induce changes in human bone marrow stromal cells. *Journal of Translational Medicine*, 11(1):1–14, 2013.

[100] Eugenia Flores-Figueroa, Guillermo Gutiérrez-Espíndola, Juan José Montesinos, Rosa María Arana-Trejo, and Hector Mayani. In vitro characterization of hematopoietic microenvironment cells from patients with myelodysplastic syndrome. *Leukemia research*, 26(7):677–686, 2002.

[101] Daniel Steinbach, Alexander Schramm, Angelika Eggert, Masanori Onda, Kristin Dawczynski, Andreas Rump, Ira Pastan, Susann Wittig, Nadine Pfaffendorf, Astrid Voigt, et al. Identification of a set of seven genes for the monitoring of minimal residual disease in pediatric acute myeloid leukemia. *Clinical cancer research*, 12(8):2434–2441, 2006.

[102] Tian Yi Zhang, Ritika Dutta, Brooks Benard, Feifei Zhao, Raymond Yin, and Ravindra Majeti. Il-6 blockade reverses bone marrow failure induced by human acute myeloid leukemia. *Science translational medicine*, 12(538):eaax5104, 2020.

[103] Frederick R Appelbaum, Holly Gundacker, David R Head, Marilyn L Slovak, Cheryl L Willman, John E Godwin, Jeanne E Anderson, and Stephen H Petersdorf. Age and acute myeloid leukemia. *Blood*, 107(9):3481–3485, 2006.

[104] Thomas Büchner, Wolfgang E Berdel, Claudia Haferlach, Torsten Haferlach, Susanne Schnittger, Carsten Müller-Tidow, Jan Braess, Karsten Spiekermann, Joachim Kienast, Peter Staib, et al. Age-related risk profile and chemotherapy dose response in acute myeloid leukemia: a study by the german acute myeloid leukemia cooperative group. *Journal of Clinical Oncology*, 27(1):61–69, 2009.

[105] Nikolai A Podoltsev, Maximilian Stahl, Amer M Zeidan, and Steven D Gore. Selecting initial treatment of acute myeloid leukaemia in older adults. *Blood reviews*, 31(2):43–62, 2017.

[106] Catherine P Leith, Kenneth J Kopecky, John Godwin, Thomas McConnell, Marilyn L Slovak, I-Ming Chen, David R Head, Frederick R Appelbaum, and Cheryl L Willman. Acute myeloid leukemia in the elderly: assessment of multidrug resistance (mdr1) and cytogenetics distinguishes biologic subgroups with remarkably distinct responses to standard chemotherapy. a southwest

oncology group study. *Blood, The Journal of the American Society of Hematology*, 89(9):3323–3329, 1997.

[107] Hagop Kantarjian, Susan O'brien, Jorge Cortes, Francis Giles, Stefan Faderl, Elias Jabbour, Guillermo Garcia-Manero, William Wierda, Sherry Pierce, Jianqin Shan, et al. Results of intensive chemotherapy in 998 patients age 65 years or older with acute myeloid leukemia or high-risk myelodysplastic syndrome: predictive prognostic models for outcome. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 106(5):1090–1098, 2006.

[108] Keith Wheatley, Cassandra L Brookes, Andrew J Howman, Anthony H Goldstone, Donald W Milligan, Archibald G Prentice, Anthony V Moorman, Alan K Burnett, United Kingdom National Cancer Research Institute Haematological Oncology Clinical Studies Group, and Acute Myeloid Leukaemia Subgroup. Prognostic factor analysis of the survival of elderly patients with aml in the mrc aml11 and lrf aml14 trials. *British journal of haematology*, 145(5):598–605, 2009.

[109] Lagadinou D Eleni, Zoumbos C Nicholas, and Spyridonidis Alexandros. Challenges in treating older patients with acute myeloid leukemia. *Journal of oncology*, 2010, 2010.

[110] Heidi D Klepin and Lodovico Balducci. Acute myelogenous leukemia in older adults. *The oncologist*, 14(3):222–232, 2009.

[111] Siddhartha Jaiswal, Pierre Fontanillas, Jason Flannick, Alisa Manning, Peter V Grauman, Brenton G Mar, R Coleman Lindsley, Craig H Mermel, Noel Burtt, Alejandro Chavez, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *New England Journal of Medicine*, 371(26):2488–2498, 2014.

[112] Andrew L Young, Grant A Challen, Brenda M Birmann, and Todd E Druley. Clonal haematopoiesis harbouring aml-associated mutations is ubiquitous in healthy adults. *Nature communications*, 7(1):1–7, 2016.

[113] Giulio Genovese, Anna K Kähler, Robert E Handsaker, Johan Lindberg, Samuel A Rose, Samuel F Bakhoum, Kimberly Chambert, Eran Mick, Benjamin M Neale, Menachem Fromer, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood dna sequence. *New England Journal of Medicine*, 371(26):2477–2487, 2014.

[114] M Ryan Corces-Zimmerman, Wan-Jen Hong, Irving L Weissman, Bruno C Medeiros, and Ravindra Majeti. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proceedings of the National Academy of Sciences*, 111(7):2548–2553, 2014.

[115] Max Jan, Thomas M Snyder, M Ryan Corces-Zimmerman, Paresh Vyas, Irving L Weissman, Stephen R Quake, and Ravindra Majeti. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Science translational medicine*, 4(149):149ra118–149ra118, 2012.

[116] Thomas McKerrell and George S Vassiliou. Aging as a driver of leukemogenesis. *Science translational medicine*, 7(306):306fs38–306fs38, 2015.

[117] Pinkal Desai, Nuria Mencia-Trinchant, Oleksandr Savenkov, Michael S Simon, Gloria Cheang, Sangmin Lee, Michael Samuel, Ellen K Ritchie, Monica L Guzman, Karla V Ballman, et al. Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nature medicine*, 24(7):1015–1023, 2018.

[118] Sagi Abelson, Grace Collord, Stanley WK Ng, Omer Weissbrod, Netta Mendelson Cohen, Elisabeth Niemeyer, Noam Barda, Philip C Zuzarte, Lawrence Heisler, Yogi Sundaravadanam, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature*, 559(7714):400–404, 2018.

[119] Alexander S Watson, Thomas Riffelmacher, Amanda Stranks, Owen Williams, Jasper De Boer, Kelvin Cain, Marion MacFarlane, Joanna McGouran, Benedikt Kessler, Shivani Khandwala, et al. Autophagy limits proliferation and glycolytic metabolism in acute myeloid leukemia. *Cell death discovery*, 1(1):1–10, 2015.

[120] Monika Mortensen, Elizabeth J Soilleux, Gordana Djordjevic, Rebecca Tripp, Michael Lutteropp, Elham Sadighi-Akha, Amanda J Stranks, Julie Glanville,

Samantha Knight, Sten-Eirik W. Jacobsen, et al. The autophagy protein atg7 is essential for hematopoietic stem cell maintenance. *Journal of Experimental Medicine*, 208(3):455–467, 2011.

[121] Patrick Auberger and Alexandre Puissant. Autophagy, a key mechanism of oncogenesis and resistance in leukemia. *Blood, The Journal of the American Society of Hematology*, 129(5):547–552, 2017.

[122] Derrick J Rossi, David Bryder, Jun Seita, Andre Nussenzweig, Jan Hoeijmakers, and Irving L Weissman. Deficiencies in dna damage repair limit the function of haematopoietic stem cells with age. *Nature*, 447(7145):725–729, 2007.

[123] Shailaja Akunuru and Hartmut Geiger. Aging, clonality, and rejuvenation of hematopoietic stem cells. *Trends in molecular medicine*, 22(8):701–712, 2016.

[124] Johanna Flach, Sietske T Bakker, Mary Mohrin, Pauline C Conroy, Eric M Pietras, Damien Reynaud, Silvia Alvarez, Morgan E Diolaiti, Fernando Ugarte, E Camilla Forsberg, et al. Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature*, 512(7513):198–202, 2014.

[125] Ya-Hsuan Ho, Raquel Del Toro, Jose Rivera-Torres, Justyna Rak, Claudia Korn, Andrés García-García, David Macías, Cristina González-Gómez, Alberto Del Monte, Monika Wittner, et al. Remodeling of bone marrow hematopoietic stem cell niches promotes myeloid cell expansion during premature or physiological aging. *Cell stem cell*, 25(3):407–418, 2019.

[126] Anjali P Kusumbe, Saravana K Ramasamy, Tomer Itkin, Maarja Andaloussi Mäe, Urs H Langen, Christer Betsholtz, Tsvee Lapidot, and Ralf H Adams. Age-dependent modulation of vascular niches for haematopoietic stem cells. *Nature*, 532(7599):380–384, 2016.

[127] Maria Maryanovich, Ali H Zahalka, Halley Pierce, Sandra Pinho, Fumio Nakahara, Noboru Asada, Qiaozhi Wei, Xizhe Wang, Paul Ciero, Jianing Xu, et al. Adrenergic nerve degeneration in bone marrow drives aging of the hematopoietic stem cell niche. *Nature medicine*, 24(6):782–791, 2018.

[128] Novella Guidi, Mehmet Sacma, Ludger Ständker, Karin Soller, Gina Marka, Karina Eiwen, Johannes M Weiss, Frank Kirchhoff, Tanja Weil, Jose A Cancelas, et al. Osteopontin attenuates aging-associated phenotypes of hematopoietic stem cells. *The EMBO journal*, 36(7):840–853, 2017.

[129] Curtis J Henry, Matias Casás-Selves, Jihye Kim, Vadym Zaberezhnyy, Leila Aghili, Ashley E Daniel, Linda Jimenez, Tania Azam, Eoin N McNamee, Eric T Clambey, et al. Aging-associated inflammation promotes selection for adaptive oncogenic events in b cell progenitors. *The Journal of clinical investigation*, 125(12):4666–4680, 2015.

[130] Simona Valletta, Alexander Thomas, Yiran Meng, Xiying Ren, Roy Drissen, Hilal Sengül, Cristina Di Genua, and Claus Nerlov. Micro-environmental sensing by bone marrow stroma identifies il-6 and tgfβ1 as regulators of hematopoietic ageing. *Nature communications*, 11(1):1–13, 2020.

[131] Aysegul V Ergen, Nathan C Boles, and Margaret A Goodell. Rantes/ccl5 influences hematopoietic stem cell subtypes and causes myeloid skewing. *Blood, The Journal of the American Society of Hematology*, 119(11):2500–2509, 2012.

[132] Larisa V Kovtonyuk, Kristin Fritsch, Xiaomin Feng, Markus G Manz, and Hitoshi Takizawa. Inflamm-aging of hematopoiesis, hematopoietic stem cells, and the bone marrow microenvironment. *Frontiers in immunology*, 7:502, 2016.

[133] Le Cong, F Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D Hsu, Xuebing Wu, Wenyan Jiang, Luciano A Marraffini, et al. Multiplex genome engineering using crispr/cas systems. *Science*, 339(6121):819–823, 2013.

[134] Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A Doudna, and Emmanuelle Charpentier. A programmable dual-rna–guided dna endonuclease in adaptive bacterial immunity. *science*, 337(6096):816–821, 2012.

[135] Seung Woo Cho, Sojung Kim, Jong Min Kim, and Jin-Soo Kim. Targeted

genome engineering in human cells with the cas9 rna-guided endonuclease. *Nature biotechnology*, 31(3):230–232, 2013.

[136] Christopher A Lino, Jason C Harper, James P Carney, and Jerilyn A Timlin. Delivering crispr: a review of the challenges and approaches. *Drug delivery*, 25(1):1234–1257, 2018.

[137] FAFA Ran, Patrick D Hsu, Jason Wright, Vineeta Agarwala, David A Scott, and Feng Zhang. Genome engineering using the crispr-cas9 system. *Nature protocols*, 8(11):2281–2308, 2013.

[138] Prashant Mali, Luhan Yang, Kevin M Esvelt, John Aach, Marc Guell, James E DiCarlo, Julie E Norville, and George M Church. Rna-guided human genome engineering via cas9. *Science*, 339(6121):823–826, 2013.

[139] Ludovic Deriano and David B Roth. Modernizing the nonhomologous end-joining repertoire: alternative and classical nhej share the stage. *Annual review of genetics*, 47(1):433–55, 2013.

[140] Agnel Sfeir and Lorraine S Symington. Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? *Trends in biochemical sciences*, 40(11):701–714, 2015.

[141] Matthew C Canver, Maximilian Haeussler, Daniel E Bauer, Stuart H Orkin, Neville E Sanjana, Ophir Shalem, Guo-Cheng Yuan, Feng Zhang, Jean-Paul Concordet, and Luca Pinello. Integrated design, execution, and analysis of arrayed and pooled crispr genome-editing experiments. *Nature protocols*, 13(5):946–986, 2018.

[142] Megan van Overbeek, Daniel Capurso, Matthew M Carter, Matthew S Thompson, Elizabeth Frias, Carsten Russ, John S Reece-Hoyes, Christopher Nye, Scott Gradia, Bastien Vidal, et al. Dna repair profiling reveals nonrandom outcomes at cas9-mediated breaks. *Molecular cell*, 63(4):633–646, 2016.

[143] Michael Kosicki, Felicity Allen, Frances Steward, Kärt Tomberg, Yangyang Pan, and Allan Bradley. Cas9-induced large deletions and small indels are controlled in a convergent fashion. *Nature communications*, 13(1):1–11, 2022.

[144] Yanfang Fu, Jennifer A Foden, Cyd Khayter, Morgan L Maeder, Deepak Reyon, J Keith Joung, and Jeffry D Sander. High-frequency off-target mutagenesis induced by crispr-cas nucleases in human cells. *Nature biotechnology*, 31(9):822–826, 2013.

[145] Patrick D Hsu, David A Scott, Joshua A Weinstein, F Ran, Silvana Konermann, Vineeta Agarwala, Yinqing Li, Eli J Fine, Xuebing Wu, Ophir Shalem, et al. Dna targeting specificity of rna-guided cas9 nucleases. *Nature biotechnology*, 31(9):827–832, 2013.

[146] Maximilian Haeussler, Kai Schönig, Hélène Eckert, Alexis Eschstruth, Joffrey Mianné, Jean-Baptiste Renaud, Sylvie Schneider-Maunoury, Alena Shkumatava, Lydia Teboul, Jim Kent, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide rna selection tool crispor. *Genome biology*, 17(1):1–12, 2016.

[147] John G Doench, Ella Hartenian, Daniel B Graham, Zuzana Tothova, Mudra Hegde, Ian Smith, Meagan Sullender, Benjamin L Ebert, Ramnik J Xavier, and David E Root. Rational design of highly active sgrnas for crispr-cas9–mediated gene inactivation. *Nature biotechnology*, 32(12):1262–1267, 2014.

[148] Hiroko Koike-Yusa, Yilong Li, E-Pien Tan, Martin Del Castillo Velasco-Herrera, Kosuke Yusa, et al. Genome-wide recessive genetic screening in mammalian cells with a lentiviral crispr-guide rna library. *Nature biotechnology*, 32(3):267–273, 2014.

[149] John G Doench, Nicolo Fusi, Meagan Sullender, Mudra Hegde, Emma W Vaimberg, Katherine F Donovan, Ian Smith, Zuzana Tothova, Craig Wilen, Robert Orchard, et al. Optimized sgrna design to maximize activity and minimize off-target effects of crispr-cas9. *Nature biotechnology*, 34(2):184–191, 2016.

[150] Ophir Shalem, Neville E Sanjana, and Feng Zhang. High-throughput functional genomics using crispr–cas9. *Nature Reviews Genetics*, 16(5):299–311, 2015.

[151] Anne Schuster, Hélène Erasimus, Sabrina Fritah, Petr V Nazarov, Eric van Dyck, Simone P Niclou, and Anna Golebiewska. Rnai/crispr screens: from a pool to a valid hit. *Trends in biotechnology*, 37(1):38–55, 2019.

[152] Ophir Shalem, Neville E Sanjana, Ella Hartenian, Xi Shi, David A Scott, Tarjei S Mikkelsen, Dirk Heckl, Benjamin L Ebert, David E Root, John G Doench, et al. Genome-scale crispr-cas9 knockout screening in human cells. *Science*, 343(6166):84–87, 2014.

[153] Bastiaan Evers, Katarzyna Jastrzebski, Jeroen PM Heijmans, Wipawadee Grernrum, Roderick L Beijersbergen, and Rene Bernards. Crispr knockout screening outperforms shrna and crispri in identifying essential genes. *Nature biotechnology*, 34(6):631–633, 2016.

[154] Julia Joung, Silvana Konermann, Jonathan S Gootenberg, Omar O Abudayyeh, Randall J Platt, Mark D Brigham, Neville E Sanjana, and Feng Zhang. Genome-scale crispr-cas9 knockout and transcriptional activation screening. *Nature protocols*, 12(4):828–863, 2017.

[155] Ryan D Chow and Sidi Chen. Cancer crispr screens in vivo. *Trends in cancer*, 4(5):349–358, 2018.

[156] Jeevisha Bajaj, Michael Hamilton, Yutaka Shima, Kendall Chambers, Kyle Spinler, Eric L Van Nostrand, Brian A Yee, Steven M Blue, Michael Chen, David Rizzeri, et al. An in vivo genome-wide crispr screen identifies the rna-binding protein staufen2 as a key regulator of myeloid leukemia. *Nature cancer*, 1(4):410–422, 2020.

[157] Sidi Chen, Neville E Sanjana, Kaijie Zheng, Ophir Shalem, Kyungheon Lee, Xi Shi, David A Scott, Jun Song, Jen Q Pan, Ralph Weissleder, et al. Genome-wide crispr screen in a mouse model of tumor growth and metastasis. *Cell*, 160(6):1246–1260, 2015.

[158] Žaklina Strezoska, Abel Licon, Josh Haimes, Katie Jansen Spayd, Kruti M Patel, Kevin Sullivan, Katarzyna Jastrzebski, Kaylene J Simpson, Devin Leake, Anja van Brabant Smith, et al. Optimized pcr conditions and increased shrna

fold representation improve reproducibility of pooled shrna screens. *PloS one*, 2012.

[159] Tyler E Miller, Brian B Liau, Lisa C Wallace, Andrew R Morton, Qi Xie, Deobrat Dixit, Daniel C Factor, Leo JY Kim, James J Morrow, Qiulian Wu, et al. Transcription elongation factors represent in vivo cancer dependencies in glioblastoma. *Nature*, 547(7663):355–359, 2017.

[160] Chi-Chao Chen, Bo Li, Scott E Millman, Cynthia Chen, Xiang Li, John P Morris IV, Allison Mayle, Yu-Jui Ho, Evangelia Loizou, Hui Liu, et al. Vitamin b6 addiction in acute myeloid leukemia. *Cancer Cell*, 37(1):71–84, 2020.

[161] Ramprasad Ramakrishnan, Pablo Peña-Martínez, Puneet Agarwal, Maria Rodriguez-Zabala, Marion Chapellier, Carl Högberg, Mia Eriksson, David Yudovich, Mansi Shah, Mats Ehinger, et al. Cxcr4 signaling has a cxcl12-independent essential role in murine mll-af9-driven acute myeloid leukemia. *Cell reports*, 31(8):107684, 2020.

[162] Shan Lin, Clément Larrue, Nastassja K Scheidegger, Bo Kyung A Seong, Neekesh V Dharia, Miljan Kuljanin, Caroline S Wechsler, Guillaume Kugener, Amanda L Robichaud, Amy Saur Conway, et al. An in vivo crispr screening platform for prioritizing therapeutic targets in aml. *Cancer discovery*, 12(2):432–449, 2022.

[163] Fangxue Yan, Jinyang Li, Jelena Milosevic, Ricardo Petroni, Suying Liu, Zhennan Shi, Salina Yuan, Janice M Reynaga, Yuwei Qi, Joshua Rico, et al. Kat6a and enl form an epigenetic transcriptional control module to drive critical leukemogenic gene-expression programs. *Cancer discovery*, 12(3):792–811, 2022.

[164] Francois E Mercier, Jiantao Shi, David B Sykes, Toshihiko Oki, Maja Jankovic, Cheuk Him Man, Youmna S Kfoury, Elizabeth Miller, Shutao He, Alexander Zhu, et al. In vivo genome-wide crispr screening in murine acute myeloid leukemia uncovers microenvironmental dependencies. *Blood Advances*, 6(17):5072–5084, 2022.

[165] Xiaoqing Wang, Collin Tokheim, Shengqing Stan Gu, Binbin Wang, Qin Tang, Yihao Li, Nicole Traugh, Zexian Zeng, Yi Zhang, Ziyi Li, et al. In vivo crispr screens identify the e3 ligase cop1 as a modulator of macrophage infiltration and cancer immunotherapy target. *Cell*, 184(21):5357–5374, 2021.

[166] Matthew B Dong, Guangchuan Wang, Ryan D Chow, Lupeng Ye, Lvyun Zhu, Xiaoyun Dai, Jonathan J Park, Hyunu R Kim, Youssef Errami, Christopher D Guzman, et al. Systematic immunotherapy target discovery using genome-scale in vivo crispr screens in cd8 t cells. *Cell*, 178(5):1189–1204, 2019.

[167] Konstantinos Tzelepis, Hiroko Koike-Yusa, Etienne De Braekeleer, Yilong Li, Emmanouil Metzakopian, Oliver M Dovey, Annalisa Mupo, Vera Grinkevich, Meng Li, Milena Mazan, et al. A crispr dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell reports*, 17(4):1193–1205, 2016.

[168] Tim Wang, Haiyan Yu, Nicholas W Hughes, Bingxu Liu, Arek Kendirli, Klara Klein, Walter W Chen, Eric S Lander, and David M Sabatini. Gene essentiality profiling reveals gene networks and synthetic lethal interactions with oncogenic ras. *Cell*, 168(5):890–903, 2017.

[169] Satoru Morikawa, Yo Mabuchi, Yoshiaki Kubota, Yasuo Nagai, Kunimichi Niibe, Emi Hiratsu, Sadafumi Suzuki, Chikako Miyauchi-Hara, Narihito Nagoshi, Takehiko Sunabori, et al. Prospective identification, isolation, and systemic transplantation of multipotent mesenchymal stem cells in murine bone marrow. *Journal of Experimental Medicine*, 206(11):2483–2496, 2009.

[170] Yuka Nakamura, Fumio Arai, Hiroko Iwasaki, Kentaro Hosokawa, Isao Kobayashi, Yumiko Gomei, Yoshiko Matsumoto, Hiroki Yoshihara, and Toshio Suda. Isolation and characterization of endosteal niche cell populations that regulate hematopoietic stem cells. *Blood, The Journal of the American Society of Hematology*, 116(9):1422–1432, 2010.

[171] Craig L Semerad, Matthew J Christopher, Fulu Liu, Brenton Short, Paul J Simmons, Ingrid Winkler, Jean-Pierre Levesque, Jean Chappel, F Patrick

Ross, and Daniel C Link. G-csf potently inhibits osteoblast activity and cxcl12 mrna expression in the bone marrow. *Blood*, 106(9):3020–3027, 2005.

[172] Simone Picelli, Omid R Faridani, Åsa K Björklund, Gösta Winberg, Sven Sagasser, and Rickard Sandberg. Full-length rna-seq from single cells using smart-seq2. *Nature protocols*, 9(1):171–181, 2014.

[173] Peter JM Valk, Roel GW Verhaak, M Antoinette Beijen, Claudia AJ Erpelinck, Sahar Barjesteh van Waalwijk van Doorn-Khosrovani, Judith M Boer, H Berna Beverloo, Michael J Moorhouse, Peter J Van Der Spek, Bob Löwenberg, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. *New England Journal of Medicine*, 350(16):1617–1628, 2004.

[174] Dirk Heckl, Monika S Kowalczyk, David Yudovich, Roger Belizaire, Rishi V Puram, Marie E McConkey, Anne Thielke, Jon C Aster, Aviv Regev, and Benjamin L Ebert. Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using crispr-cas9 genome editing. *Nature biotechnology*, 32(9):941–946, 2014.

[175] Felix Krueger, Simon R Andrews, and Cameron S Osborne. Large scale loss of data in low-diversity illumina sequencing libraries can be recovered by deferred cluster calling. *PloS one*, 6(1):e16607, 2011.

[176] Laszlo Revesz, Birgit Bollbuck, Thomas Buhl, Janet Dawson, Roland Feifel, Richard Heng, Peter Hiestand, Helmut Sparrer, Achim Schlapbach, and Rudolf Waelchli. Bridged piperazines and piperidines as ccr1 antagonists with oral activity in models of arthritis and multiple sclerosis. *Letters in Drug Design & Discovery*, 3(10):689–694, 2006.

[177] Takanori Kitamura, Teruaki Fujishita, Pius Loetscher, Laszlo Revesz, Hiroki Hashida, Shinae Kizaka-Kondoh, Masahiro Aoki, and Makoto M Taketo. Inactivation of chemokine (cc motif) receptor 1 (ccr1) suppresses colon cancer liver metastasis by blocking accumulation of immature myeloid cells in a mouse model. *Proceedings of the National Academy of Sciences*, 107(29):13063–13068, 2010.

[178] Alexander Dobin, Carrie A Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R Gingeras. Star: ultrafast universal rna-seq aligner. *Bioinformatics*, 29(1):15–21, 2013.

[179] Yang Liao, Gordon K Smyth, and Wei Shi. featurecounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7):923–930, 2014.

[180] Philip Ewels, Måns Magnusson, Sverker Lundin, and Max Käller. Multiqc: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19):3047–3048, 2016.

[181] Michael I Love, Wolfgang Huber, and Simon Anders. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome biology*, 15(12):1–21, 2014.

[182] Mark D Robinson, Davis J McCarthy, and Gordon K Smyth. edger: a bioconductor package for differential expression analysis of digital gene expression data. *bioinformatics*, 26(1):139–140, 2010.

[183] Aravind Subramanian, Pablo Tamayo, Vamsi K Mootha, Sayan Mukherjee, Benjamin L Ebert, Michael A Gillette, Amanda Paulovich, Scott L Pomeroy, Todd R Golub, Eric S Lander, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550, 2005.

[184] Wei Li, Han Xu, Tengfei Xiao, Le Cong, Michael I Love, Feng Zhang, Rafael A Irizarry, Jun S Liu, Myles Brown, and X Shirley Liu. Mageck enables robust identification of essential genes from genome-scale crispr/cas9 knockout screens. *Genome biology*, 15(12):1–12, 2014.

[185] Binbin Wang, Mei Wang, Wubing Zhang, Tengfei Xiao, Chen-Hao Chen, Alexander Wu, Feizhen Wu, Nicole Traugh, Xiaoqing Wang, Ziyi Li, et al. Integrative analysis of pooled crispr genetic screens using mageckflute. *Nature protocols*, 14(3):756–780, 2019.

[186] John Vivian, Arjun Arkal Rao, Frank Austin Nothaft, Christopher Ketchum, Joel Armstrong, Adam Novak, Jacob Pfeil, Jake Narkizian, Alden D Deran, Audrey Musselman-Brown, et al. Toil enables reproducible, open source, big biomedical data analyses. *Nature biotechnology*, 35(4):314–316, 2017.

[187] Zefang Tang, Chenwei Li, Boxi Kang, Ge Gao, Cheng Li, and Zemin Zhang. Gepia: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic acids research*, 45(W1):W98–W102, 2017.

[188] Jeffrey W Tyner, Cristina E Tognon, Daniel Bottomly, Beth Wilmot, Stephen E Kurtz, Samantha L Savage, Nicola Long, Anna Reister Schultz, Elie Traer, Melissa Abel, et al. Functional genomic landscape of acute myeloid leukaemia. *Nature*, 562(7728):526–531, 2018.

[189] Aruna Kode, John S Manavalan, Ioanna Mosialou, Govind Bhagat, Chozha V Rathinam, Na Luo, Hossein Khiabanian, Albert Lee, Vundavalli V Murty, Richard Friedman, et al. Leukaemogenesis induced by an activating β-catenin mutation in osteoblasts. *Nature*, 506(7487):240–244, 2014.

[190] Hector Mayani, Eugenia Flores-Figueroa, and Antonieta Chávez-González. In vitro biology of human myeloid leukemia. *Leukemia research*, 33(5):624–637, 2009.

[191] Sawa Ito, A John Barrett, Amalia Dutra, Evgenia Pak, Samantha Miner, Keyvan Keyvanfar, Nancy F Hensel, Katayoun Rezvani, Pawel Muranski, Paul Liu, et al. Long term maintenance of myeloid leukemic stem cells cultured with unrelated human mesenchymal stromal cells. *Stem cell research*, 14(1):95–104, 2015.

[192] Fiona M Behan, Francesco Iorio, Gabriele Picco, Emanuel Gonçalves, Charlotte M Beaver, Giorgia Migliardi, Rita Santos, Yanhua Rao, Francesco Sassi, Marika Pinnelli, et al. Prioritization of cancer therapeutic targets using crispr–cas9 screens. *Nature*, 568(7753):511–516, 2019.

[193] Ronald P Gladue, Susan H Cole, Marsha L Roach, Laurie A Tylaska, Robin T Nelson, Richard M Shepard, John D McNeish, Kevin T Ogborne,

and Kuldeep S Neote. The human specific ccr1 antagonist cp-481,715 inhibits cell infiltration and inflammatory responses in human ccr1 transgenic mice. *The Journal of Immunology*, 176(5):3141–3148, 2006.

[194] Xiaoyi Hu, Maorong Fu, Xia Zhao, Wei Wang, et al. The jak/stat signaling pathway: From bench to clinic. *Signal Transduction and Targeted Therapy*, 6(1):1–33, 2021.

[195] Michelle Millington, Allison Arndt, Maureen Boyd, Tanya Applegate, and Sylvie Shen. Towards a clinically relevant lentiviral transduction protocol for primary human cd34+ hematopoietic stem/progenitor cells. *PloS one*, 4(7):e6461, 2009.

[196] Christine Leibiger, Nadezda Kosyakova, Hasmik Mkrtchyan, Michael Glei, Vladimir Trifonov, and Thomas Liehr. First molecular cytogenetic high resolution characterization of the nih 3t3 cell line by murine multicolor banding. *Journal of Histochemistry & Cytochemistry*, 61(4):306–312, 2013.

[197] Tomohisa Baba, Kazuhito Naka, Soji Morishita, Norio Komatsu, Atsushi Hirao, and Naofumi Mukaida. Mip-1α/ccl3-mediated maintenance of leukemia-initiating cells in the initiation process of chronic myeloid leukemia. *Journal of Experimental Medicine*, 210(12):2661–2673, 2013.

[198] Mara N Zeissig, Duncan R Hewett, Vasilios Panagopoulos, Krzysztof M Mrozik, L Bik To, Peter I Croucher, Andrew CW Zannettino, and Kate Vandyke. Expression of the chemokine receptor ccr1 promotes the dissemination of multiple myeloma plasma cells in vivo. *haematologica*, 106(12):3176, 2021.

[199] Robin M Meyers, Jordan G Bryan, James M McFarland, Barbara A Weir, Ann E Sizemore, Han Xu, Neekesh V Dharia, Phillip G Montgomery, Glenn S Cowley, Sasha Pantel, et al. Computational correction of copy number effect improves specificity of crispr–cas9 essentiality screens in cancer cells. *Nature genetics*, 49(12):1779–1784, 2017.

[200] Nicolas Severe, Nezihi Murat Karabacak, Karin Gustafsson, Ninib Baryawno, Gabriel Courties, Youmna Kfoury, Konstantinos D Kokkaliaris, Catherine

Rhee, Dongjun Lee, Elizabeth W Scadden, et al. Stress-induced changes in bone marrow stromal cell populations revealed through single-cell protein expression mapping. *Cell Stem Cell*, 25(4):570–583, 2019.

[201] Jean-Paul Abbuehl, Zuzana Tatarova, Werner Held, and Joerg Huelsken. Long-term engraftment of primary bone marrow stromal cells repairs niche damage and improves hematopoietic stem cell transplantation. *Cell stem cell*, 21(2):241–255, 2017.

[202] Shao-bo Su, Naofumi Mukaida, Jian-bin Wang, Hideki Nomura, and Kouji Matsushima. Preparation of specific polyclonal antibodies to a c-c chemokine receptor, ccr1, and determination of ccr1 expression on various types of leukocytes. *Journal of leukocyte biology*, 60(5):658–666, 1996.

[203] Hal E Broxmeyer, Scott Cooper, Giao Hangoc, Ji-Liang Gao, and Philip M Murphy. Dominant myelopoietic effector functions mediated by chemokine receptor ccr1. *The Journal of experimental medicine*, 189(12):1987–1992, 1999.

[204] Ji-Liang Gao, Thomas A Wynn, Yun Chang, Eric J Lee, Hal E Broxmeyer, Scott Cooper, H Lee Tiffany, Heiner Westphal, June Kwon-Chung, and Philip M Murphy. Impaired host defense, hematopoiesis, granulomatous inflammation and type 1–type 2 cytokine balance in mice lacking cc chemokine receptor 1. *The Journal of experimental medicine*, 185(11):1959–1968, 1997.

[205] Gwendal Lazennec and Ann Richmond. Chemokines and chemokine receptors: new insights into cancer-related inflammation. *Trends in molecular medicine*, 16(3):133–144, 2010.

[206] Matthew W Anderson, Shuchun Zhao, Weiyun Z Ai, Robert Tibshirani, Ronald Levy, Izidore S Lossos, and Yasodha Natkunam. Cc chemokine receptor 1 expression in human hematolymphoid neoplasia. *American journal of clinical pathology*, 133(3):473–483, 2010.

[207] Sonia Vallet, Samantha Pozzi, Kishan Patel, Nileshwari Vaghela, MariaTeresa Fulciniti, Petter Veiby, Teru Hideshima, Loredana Santo, Diana Cirstea, David T Scadden, et al. A novel role for ccl3 (mip-1α) in myeloma-induced

bone disease via osteocalcin downregulation and inhibition of osteoblast function. *Leukemia*, 25(7):1174–1181, 2011.

[208] Haydar Çelik, Katherine E Lindblad, Bogdan Popescu, Gege Gui, Meghali Goswami, Janet Valdez, Christin DeStefano, Catherine Lai, Julie Thompson, Jack Y Ghannam, et al. Highly multiplexed proteomic assessment of human bone marrow in acute myeloid leukemia. *Blood advances*, 4(2):367–379, 2020.

[209] Deng Chen, Ting-Xuan Tang, Hai Deng, Xiang-Ping Yang, and Zhao-Hui Tang. Interleukin-7 biology and its effects on immune cells: mediator of generation, differentiation, survival, and homeostasis. *Frontiers in Immunology*, page 5156, 2021.

[210] Bin Zhang, Yuan Zhang, Lei Xiong, Yuzhe Li, Yunliang Zhang, Jiuliang Zhao, Hui Jiang, Can Li, Yunqi Liu, Xindong Liu, et al. Cd127 imprints functional heterogeneity to diversify monocyte responses in inflammatory diseases. *Journal of Experimental Medicine*, 219(2):e20211191, 2022.

[211] Warren J Leonard, Jian-Xin Lin, and John J O'Shea. The γc family of cytokines: basic biology to therapeutic ramifications. *Immunity*, 50(4):832–850, 2019.

[212] Steven Goossens, Enrico Radaelli, Odile Blanchet, Kaat Durinck, Joni Van der Meulen, Sofie Peirs, Tom Taghon, Cedric S Tremblay, Magdaline Costa, Morvarid Farhang Ghahremani, et al. Zeb2 drives immature t-cell lymphoblastic leukaemia development via enhanced tumour-initiating potential and il-7 receptor signalling. *Nature communications*, 6(1):1–12, 2015.

[213] Afonso RM Almeida, João L Neto, Ana Cachucho, Mayara Euzébio, Xiangyu Meng, Rathana Kim, Marta B Fernandes, Beatriz Raposo, Mariana L Oliveira, Daniel Ribeiro, et al. Interleukin-7 receptor α mutational activation can initiate precursor b-cell acute lymphoblastic leukemia. *Nature communications*, 12(1):1–16, 2021.

[214] Werner Digel, Mathias Schmid, Gerhard Heil, Peter Conrad, Steven Gillis, and Franz Porzsolt. Human interleukin-7 induces proliferation of neoplastic cells

from chronic lymphocytic leukemia and acute leukemias. *Blood, The Journal of the American Society of Hematology*, 78(3), 1991.

[215] Felicity Allen, Luca Crepaldi, Clara Alsinet, Alexander J Strong, Vitalii Kleshchevnikov, Pietro De Angeli, Petra Páleníková, Anton Khodak, Vladimir Kiselev, Michael Kosicki, et al. Predicting the mutations generated by repair of cas9-induced double-strand breaks. *Nature biotechnology*, 37(1):64–72, 2019.

[216] Tim Wang, Jenny J Wei, David M Sabatini, and Eric S Lander. Genetic screens in human cells using the crispr-cas9 system. *Science*, 343(6166):80–84, 2014.

[217] F Ann Ran, Patrick D Hsu, Chie-Yu Lin, Jonathan S Gootenberg, Silvana Konermann, Alexandro E Trevino, David A Scott, Azusa Inoue, Shogo Matoba, Yi Zhang, et al. Double nicking by rna-guided crispr cas9 for enhanced genome editing specificity. *Cell*, 154(6):1380–1389, 2013.

[218] Wei Chen, Aaron McKenna, Jacob Schreiber, Maximilian Haeussler, Yi Yin, Vikram Agarwal, William Stafford Noble, and Jay Shendure. Massively parallel profiling and predictive modeling of the outcomes of crispr/cas9-mediated double-strand break repair. *Nucleic acids research*, 47(15):7989–8003, 2019.

[219] Xuebing Wu, David A Scott, Andrea J Kriz, Anthony C Chiu, Patrick D Hsu, Daniel B Dadon, Albert W Cheng, Alexandro E Trevino, Silvana Konermann, Sidi Chen, et al. Genome-wide binding of the crispr endonuclease cas9 in mammalian cells. *Nature biotechnology*, 32(7):670–676, 2014.

[220] Tim Hsiau, David Conant, Nicholas Rossi, Travis Maures, Kelsey Waite, Joyce Yang, Sahil Joshi, Reed Kelso, Kevin Holden, Brittany L Enzmann, et al. Inference of crispr edits from sanger trace data. *BioRxiv*, page 251082, 2019.

[221] F Liu, YK Song, and D Liu. Hydrodynamics-based transfection in animals by systemic administration of plasmid dna. *Gene therapy*, 6(7):1258–1266, 1999.

[222] Hideyo Hirai, Pu Zhang, Tajhal Dayaram, Christopher J Hetherington, Shinichi Mizuno, Jiro Imanishi, Koichi Akashi, and Daniel G Tenen. C/ebpβ

is required for'emergency'granulopoiesis. *Nature immunology*, 7(7):732–739, 2006.

[223] Rifca Le Dieu, David C Taussig, Alan G Ramsay, Richard Mitter, Faridah Miraki-Moud, Rewas Fatah, Abigail M Lee, T Andrew Lister, and John G Gribben. Peripheral blood t cells in acute myeloid leukemia (aml) patients at diagnosis have abnormal phenotype and genotype and form defective immune synapses with aml blasts. *Blood, The Journal of the American Society of Hematology*, 114(18):3909–3916, 2009.

[224] J Sprent and A Basten. Circulating t and b lymphocytes of the mouse: Ii. lifespan. *Cellular immunology*, 7(1):40–59, 1973.

[225] Ineke den Braber, Tendai Mugwagwa, Nienke Vrisekoop, Liset Westera, Ramona Mögling, Anne Bregje de Boer, Neeltje Willems, Elise HR Schrijver, Gerrit Spierenburg, Koos Gaiser, et al. Maintenance of peripheral naive t cells is sustained by thymus output in mice but not humans. *Immunity*, 36(2):288–297, 2012.

[226] Sarah D Cramer, Peter D Aplan, and Scott K Durum. Therapeutic targeting of il-7rα signaling pathways in all treatment. *Blood, The Journal of the American Society of Hematology*, 128(4):473–478, 2016.

[227] Angela Maria Savino and Shai Izraeli. Interleukin-7 signaling as a therapeutic target in acute lymphoblastic leukemia, 2017.

[228] R Lumkul, NC Gorin, MT Malehorn, GT Hoehn, R Zheng, B Baldwin, D Small, S Gore, D Smith, PS Meltzer, et al. Human aml cells in nod/scid mice: engraftment potential and gene expression. *Leukemia*, 16(9):1818–1826, 2002.

[229] Tomohisa Baba and Naofumi Mukaida. Role of macrophage inflammatory protein (mip)-1α/ccl3 in leukemogenesis. *Molecular & cellular oncology*, 1(1):e29899, 2014.

[230] Rhonda J Staversky, Mary A Georger, Marian Ackun-Farmmer, Michael W Becker, Danielle SW Benoit, Laura M Calvi, and Benjamin J Frisch. Ccl3 sig-

naling is essential for leukemia progression but dispensable for hematopoietic stem cell maintenance. *Blood*, 130:2603, 2017.

[231] Alexandre Bignon, Françoise Gaudin, Patrice Hémon, Hugo Tharinger, Katia Mayol, Thierry Walzer, Pius Loetscher, Michel Peuchmaur, Dominique Berrebi, and Karl Balabanian. Ccr1 inhibition ameliorates the progression of lupus nephritis in nzb/w mice. *The Journal of Immunology*, 192(3):886–896, 2014.

[232] Frédéric Peyrade, Lauris Gastaud, Daniel Ré, Sandrine Pacquelet-Cheli, and Antoine Thyss. Treatment decisions for elderly patients with haematological malignancies: a dilemma. *The lancet oncology*, 13(8):e344–e352, 2012.

[233] Jonathan A Webster and Keith W Pratz. Acute myeloid leukemia in the elderly: therapeutic options and choice. *Leukemia & lymphoma*, 59(2):274–287, 2018.

[234] Nigel H Russell. Improving outcomes for elderly patients with aml. *The Lancet. Oncology*, 13(11):1065–1066, 2012.

[235] Evgenia V Verovskaya, Paul V Dellorusso, and Emmanuelle Passegué. Losing sense of self and surroundings: hematopoietic stem cell aging and leukemic transformation. *Trends in molecular medicine*, 25(6):494–515, 2019.

[236] Allison L Boyd, Jennifer C Reid, Kyle R Salci, Lili Aslostovar, Yannick D Benoit, Zoya Shapovalova, Mio Nakanishi, Deanna P Porras, Mohammed Almakadi, Clinton JV Campbell, et al. Acute myeloid leukaemia disrupts endogenous myelo-erythropoiesis by compromising the adipocyte bone marrow niche. *Nature cell biology*, 19(11):1336–1347, 2017.

[237] Alexandra M Stevens, Jennifer M Miller, Jaime O Munoz, Amos S Gaikwad, and Michele S Redell. Interleukin-6 levels predict event-free survival in pediatric aml and suggest a mechanism of chemotherapy resistance. *Blood advances*, 1(18):1387–1397, 2017.

[238] Eric M Pietras, Cristina Mirantes-Barbeito, Sarah Fong, Dirk Loeffler, Larisa V Kovtonyuk, SiYi Zhang, Ranjani Lakshminarasimhan, Chih Peng Chin,

José-Marc Techner, Britta Will, et al. Chronic interleukin-1 exposure drives haematopoietic stem cells towards precocious myeloid differentiation at the expense of self-renewal. *Nature cell biology*, 18(6):607–618, 2016.

[239] Melih Acar, Kiranmai S Kocherlakota, Malea M Murphy, James G Peyer, Hideyuki Oguro, Christopher N Inra, Christabel Jaiyeola, Zhiyu Zhao, Katherine Luby-Phelps, and Sean J Morrison. Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. *Nature*, 526(7571):126–130, 2015.

[240] Yanjie Zhang, Hezhou Guo, Zhaoli Zhang, Wei Lu, Jiang Zhu, and Jun Shi. Il-6 promotes chemoresistance via upregulating cd36 mediated fatty acids uptake in acute myeloid leukemia. *Experimental Cell Research*, 415(1):113112, 2022.

[241] WAN Hui, Tie-Yong Qian, Xiao-Jing Hu, Ci-You Huang, and Wei-Feng Yao. Correlation of serum ccl3/mip-1α levels with disease severity in postmenopausal osteoporotic females. *Balkan Medical Journal*, 35(4):320–325, 2018.

[242] Larissa GP Langhi Prata, Inna G Ovsyannikova, Tamara Tchkonia, and James L Kirkland. Senescent cell clearance by the immune system: Emerging therapeutic opportunities. In *Seminars in immunology*, volume 40, page 101275. Elsevier, 2018.

[243] Valeria Mollica Poeta, Matteo Massara, Arianna Capucetti, and Raffaella Bonecchi. Chemokines and chemokine receptors: new targets for cancer immunotherapy. *Frontiers in immunology*, 10:379, 2019.

[244] Damien Reynaud, Eric Pietras, Keegan Barry-Holson, Alain Mir, Mikhail Binnewies, Marion Jeanne, Olga Sala-Torra, Jerald P Radich, and Emmanuelle Passegué. Il-6 controls leukemic multipotent progenitor cell fate and contributes to chronic myelogenous leukemia development. *Cancer cell*, 20(5):661–673, 2011.

[245] Zhigang Cai, Jonathan J Kotzin, Baskar Ramdas, Sisi Chen, Sai Nelanuthala, Lakshmi Reddy Palam, Ruchi Pandey, Raghuveer Singh Mali, Yan Liu,

Mark R Kelley, et al. Inhibition of inflammatory signaling in tet2 mutant preleukemic cells mitigates stress-induced abnormalities and clonal hematopoiesis. *Cell stem cell*, 23(6):833–849, 2018.

[246] Boshi Wang, Jaskaren Kohli, and Marco Demaria. Senescent cells in cancer therapy: friends or foes? *Trends in cancer*, 6(10):838–857, 2020.