Characterising *KMT2A* and *MLLT10* Rearranged Acute Lymphoblastic Leukaemia

Michelle Olivia Forgione

B.Sc (Adv) Biochemistry and Pharmacology

A thesis submitted in fulfilment of the

Doctor of Philosophy Degree in Sciences

Department of Molecular and Cellular Biology

School of Biological Sciences

University of Adelaide, Australia

January 2022

Table of Contents

Abstract

KMT2A-rearrangements (*KMT2A*r) and *MLLT10*-rearrangements (*MLLT10*r) are recurrently identified oncogenic lesions in acute lymphoblastic leukaemia (ALL). *KMT2A*r induce aggressive, high-risk ALL in patients of all ages, including >70% of infant ALL diagnoses. *MLLT10*r, most commonly *PICALM-MLLT10* and *DDX3X-MLLT10*, account for approximately 10% of new T-cell ALL (T-ALL) diagnoses, and are biologically and clinically poorly characterised. *KMT2A*r and *MLLT10*r leukaemic blasts share the important feature of *HOXA* cluster transcriptional dysregulation, suggesting a similar aetiology and subsequently similar therapeutic vulnerabilities. However, there is presently a lack of pre-clinical *in vitro* and *in vivo* models that accurately recapitulate the aggressive ALL induced by *KMT2A*r and *MLLT10*r, limiting the ability to understand the biological aetiology of these subtypes and screen for efficacious targeted therapies.

This thesis describes novel *in vitro* models of *KMT2A-AFF1*, *PICALM-MLLT10* and *DDX3X-MLLT10* T-ALL, and *KMT2A-AFF1* B-ALL. These models demonstrated the ability of *KMT2A-AFF1* to induce oncogenic changes in murine MOHITO cells. Specifically, *KMT2A-AFF1* expression induced T-ALL characterised by an immature immunophenotype, increased cellular proliferation and upregulation of *HOXA* cluster genes, reflecting the aggressive disease observed in patients.

For the first time, an *in vivo* transplantation model of *DDX3X-MLLT10* and *PICALM-MLLT10* T-ALL is also presented, generated by the expression of these fusions in CD4⁻CD8⁻ p19^Arf-null (Arf^-/-) thymocytes. When transplanted into sub lethally irradiated immunocompromised NSG mice, *MLLT10*r-expressing cells consistently engrafted to induce T-ALL. Differences were noted in the immunophenotype and dissemination patterns of engrafted cells between the two fusion cohorts, indicating that the 5' fusion partner plays an important role in *MLLT10*r.

The mutational landscape of non-infant *KMT2A*r B-ALL and *MLLT10*r ALL were investigated by interrogating patient leukaemic blasts for the presence of single nucleotide variants (SNVs) and copy number alterations (CNAs). An understanding of the complete genomic landscape of leukaemic blasts is required, in order to generate pre-clinical models that fully recapitulate ALL. Analysis revealed that non-infant *KMT2A*r B-ALL lacks typical leukaemia-associated alterations that are common in

other genomic subtypes of B-ALL, demonstrating a silent mutational landscape that parallels *KMT2A*r infant cohorts. However, known and novel SNVs within epigenetic regulatory genes were identified in *KMT2A*r B-ALL, that were less common or absent in the *BCR-ABL1* comparator B-ALL cohort, implying that epigenetic regulators may serve a functional role specifically in *KMT2A*r B-ALL. Recurrent alterations in leukaemia-associated genes were identified in *MLLT10*r T-ALL cases, suggesting that cooperative lesions play a role in *MLLT10*r leukaemogenesis, unlike in *KMT2A*r.

Therapeutic efficacy of the novel curaxin CBL0137 in infant *KMT2A*r ALL has recently been reported. Pre-clinical sensitivity studies in *in vitro* models demonstrated that functional p53 is critical in the mechanism of action of CBL0137. CBL0137 demonstrated efficacy across multiple genomic subtypes of acute leukaemia, including *KMT2A*r, only in the presence of wild-type *TP53*. Findings support that CBL0137 has clinical promise in p53 wild-type ALL, including *KMT2A*r ALL due to the low prevalence of *TP53* alterations across all ages in this subtype.

In summary, this thesis provides critical insight into the complex biology of *KMT2A*r and *MLLT10*r ALL, and presents the first described models of *KMT2A*r and *MLLT10*r T-ALL. Together, these findings contribute to improving our understanding of the biology of these high-risk subtypes, and the described models permit the pre-clinical exploration of targeted therapies in a T-ALL context, to ultimately improve outcomes for ALL patients.

Declaration of originality

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Michelle Olivia Forgione

January 2022

Acknowledgements

I would like to express my sincerest gratitude to my supervisors Prof. Deborah White, Dr. Barbara McClure and Dr. Laura Eadie for their guidance, support and mentorship throughout my candidature. It is an honour to have been guided through my PhD by three incredible women. Deb, I am inspired by your expertise and leadership. Your ability to see the bigger picture has ensured that my work was always kept on the right track. Barb, you are just amazing. Your generosity with your time and knowledge is both invaluable and admirable. You have built my confidence and abilities as a scientist right from the start, when you took me on as a nervous Undergraduate placement student. Laura, you are a fantastic mentor to me. I am so appreciative of your patience and encouragement with my mouse work, your remarkable attention-to-detail with my innumerable drafts, and your unwavering dedication to my success.

To my lab mates Paniz, Elly, Charlotte and Steph, thank you for your camaraderie over the years. You have always been there for me, whether it be to troubleshoot experiments or blow off steam after a long week. Dian, you are the greatest cheerleader a woman could ever ask for – thank you for all the pep talks and therapy sessions. Thanks to Steph and Bron for their tireless logistical contributions, Jacqui for her bioinformatic support and patience in helping me with R, and Randall for his assistance with flow cytometry (and for always giving me the black jellybeans). Caitlin, you are worth your weight in gold. Thank you for all your support in the lab, particularly for being an on-call expert in all things mouse-related, and also for the netball fun times.

Some special acknowledgements go out to my friends and family. To my best friend Lauren, you are truly unique. I am forever grateful for our bond, based on a healthy combination of dark humour, fierce resilience, pursuit of knowledge and understanding, and most importantly, good phở. Tom, your contribution to my wellbeing cannot be overstated; you are an invaluable sounding board and emotional support, and your sense of humour is phenomenal. Thank you for constantly going out of your way to meet me for mid-week coffee chats, always prepared to re-energise me with a good laugh. Cherian, a true kindred spirit of mine, thanks for your help with proofreading, and for taking such keen interest in my nerdy pursuits. Mum, I appreciate your constant encouragement, even though you don't understand why I'm still at

university. To my brother Peter, you were my first friend in life, and you understand me in a way that nobody else ever could. Thanks for being there for me, always down for a Subway or Gong Cha catch-up. To my soon-to-be husband who is also named Peter, thank you for navigating my emotions throughout this rollercoaster with such devoted gentleness and patience. Our relationship is a profound source of joy and comfort in my life, and I could not have picked a better person with whom to share my life.

Acute lymphoblastic leukaemia claimed the life of my father, John Forgione, on the 11th of June 2016, after seven years of gruelling treatment and devastating obstacles. Dad, thanks for always having my back. You were a man of great strength, warmth and wisdom. The final years of your life were filled with unspeakable suffering but also with great love and resilience, and this has inspired me to fight for the lives of other people diagnosed with cancer. I will continue to do my best to make you proud.

Publications

Forgione, M. O., McClure, B. J., Page, E. C., Yeung, D. T., Eadie, L.N. & White, D. L. *TP53* loss of function mutations reduce sensitivity of acute leukaemia to the curaxin CBL0137. *Oncology Reports* **57**, doi:10.3892/or.2022.8310

Forgione, M. O., McClure, B. J., Yeung, D. T., Eadie, L. N. & White, D. L. MLLT10 rearranged acute leukemia: incidence, prognosis and possible therapeutic strategies. *Genes, Chromosomes & Cancer* **59**, 709-721, doi:10.1002/gcc.22887 (2020).

Forgione, M. O., McClure, B. J., Eadie, L. N., Yeung, D. T. & White, D. L. KMT2A rearranged acute lymphoblastic leukaemia: Unravelling the genomic complexity and heterogeneity of this high-risk disease. *Cancer Letters* **469**, 410-418, doi:10.1016/j.canlet.2019.11.005 (2020).

Abstracts and presentations

MO Forgione, BJ McClure, LN Eadie and DL White. *CD44 and its ligand osteopontin are upregulated in KMT2A-AFF1 T-cell acute lymphoblastic leukaemia*. Blood21 Annual Scientific Meeting, September 2021, Virtual. **Poster presentation**.

MO Forgione, BJ McClure, LN Eadie and DL White. *KMT2A-rearranged acute lymphoblastic leukaemia: A complex clinical and biological challenge*. SAHMRI Precision Medicine Theme Meeting, September 2021, Adelaide, SA, Australia. **Invited oral presentation**.

MO Forgione, BJ McClure, LN Eadie and DL White. *KMT2A-rearranged acute lymphoblastic leukaemia: A complex clinical and biological challenge*. SAHMRI Cancer in SA Translational Meeting, August 2021, Adelaide, SA, Australia. **Invited oral presentation**.

MO Forgione, BJ McClure, LN Eadie and DL White. *CD44 and its ligand osteopontin are upregulated in KMT2A-AFF1 T-cell acute lymphoblastic leukaemia*. ANZCHOG Annual Scientific Meeting, June 2021, Virtual. **Poster presentation**.

MO Forgione, BJ McClure, LN Eadie and DL White. *CD44 and its ligand osteopontin are upregulated in KMT2A-AFF1 T-cell acute lymphoblastic leukaemia*. European Society of Haematology 2nd Translational Research E-Conference on Acute Lymphoblastic Leukaemia, May 2021, Virtual. **Poster presentation. Also selected for "poster walk" mini oral presentation session.**

MO Forgione, BJ McClure, LN Eadie and DL White. *CD44 and its ligand osteopontin are upregulated in KMT2A-AFF1 T-cell acute lymphoblastic leukaemia*. Australian Society for Medical Research (ASMR), November 2020, Adelaide, SA, Australia. **Oral presentation**.

MO Forgione, BJ McClure, LN Eadie and DL White. *Precision medicine in acute lymphoblastic leukaemia*. SAHMRI Scientific Seminar Series, May 2020, Adelaide, SA, Australia. **Invited oral presentation**.

MO Forgione, BJ McClure, LN Eadie, SL Heatley, J Rehn, J Breen, DT Yeung and DL White. *Single nucleotide variants in epigenetic regulators and the PI3K-RAS pathway found in KMT2A rearranged adults with acute lymphoblastic leukaemia*. New

Directions in Leukaemia Research Annual Meeting, March 2020, Brisbane, QLD, Australia. **Poster presentation.** *Cancelled due to COVID-19.*

MO Forgione, BJ McClure, SL Heatley, J Rehn, J Breen and DL White. *Single nucleotide variants in epigenetic regulators and the PI3K-RAS pathway found in KMT2A rearranged adults with acute lymphoblastic leukaemia.* SAHMRI Scientific Showcase, November 2019, Adelaide, SA, Australia. **Poster presentation.**

MO Forgione, BJ McClure, LN Eadie, SL Heatley, J Rehn, J Breen, DT Yeung and Deborah L. White. *The landscape of single nucleotide variants in epigenetic regulators found in KMT2A and MLLT10 rearranged acute lymphoblastic leukaemia.* Epigenetics Consortium of South Australia (EpiCSA), November 2019, Adelaide, SA, Australia. **Invited oral presentation**.

MO Forgione, BJ McClure, J Rehn, James Breen and Deborah L. White. *Single nucleotide variants in epigenetic regulators and the PI3K-RAS pathway found in KMT2A rearranged adults with acute lymphoblastic leukaemia.* University of Adelaide Postgraduate Consortium, October 2019, Adelaide, SA, Australia. **Poster presentation.**

MO Forgione, BJ McClure, LN Eadie, SL Heatley, J Rehn, J Breen, DT Yeung and DL White. *Single nucleotide variants within epigenetic regulatory genes found in patients with acute lymphoblastic leukaemia.* ASMR September 2019, Adelaide, SA, Australia. **Oral presentation**.

MO Forgione, BJ McClure, J Breen, and DL White. *Single nucleotide variants within epigenetic regulatory genes found in adult acute lymphoblastic leukaemia.* ASMR, September 2018, Adelaide, SA, Australia. **Poster presentation**.

MO Forgione, BJ McClure, J Breen, and DL White. *Single nucleotide variants within epigenetic regulatory genes found in patients with acute lymphoblastic leukaemia.* EpiCSA August 2018, Adelaide, SA, Australia. **Oral presentation**.

Scholarships and awards

ASMR Best Mini Oral Presentation Award

Awarded at the ASMR South Australia Annual Scientific Meeting (2020)

University of Adelaide Research Travel Scholarship

Awarded $750 to travel to Brisbane, Australia to present research at the New Directions in Leukaemia Research Annual Meeting (2020).

Cancelled due to COVID-19.

University of Adelaide Upgrade to PhD Scholarship (2019-2021)

University of Adelaide Master of Philosophy (No Honours) Scholarship (2018-2019)

EpiCSA "Epic Award" for Best Oral Presentation

Awarded at the EpiCSA Annual Research Meeting (2018)

Abbreviations

μg	micro gram (10^{-6} gram)
μM	micro molar (10^{-6} molar)
AGRF	Australian Genome Research Facility
ALL	acute lymphoblastic leukaemia
AML	acute myeloid leukaemia
AYA	adolescent and young adult
B-ALL	B-cell acute lymphoblastic leukaemia
BCR	breakpoint cluster region
BCR-ABL1	breakpoint cluster region-Abelson kinase 1 proto-oncogene
BET	bromodomain and extra-terminal motif
BM	bone marrow
bp	base pairs
C	celsius
cDNA	complementary deoxyribonucleic acid
CDX	cell line-derived xenograft
CML	chronic myeloid leukaemia
CNA	copy number alterations
CNS	central nervous system
CR	complete response
Ct	cycle threshold
DEPC	diethylpyrocarbonate
DLBCL	diffuse large B-cell lymphoma
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphates
EFS	event-free survival
FACS	fluorescence activated cell sorting
FBS	foetal bovine serum
FDA	U.S. Food and Drug Administration
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
gDNA	genomic DNA
GFP	green fluorescent protein
H&E	haematoxylin and eosin
H3K27me	histone 3 lysine 27 methylation/methyltransferase
H3K4me	histone 3 lysine 4 methylation/methyltransferase
H3K79me	histone 3 lysine 78 methylation/methyltransferase
HBSS	Hank's balanced salt solution
HCl	hydrochloric acid

HDAC	histone deacetylase
HF	high fidelity
HOX	homeobox
HR	high risk
iAMP21	intrachromosomal amplification of chromosome 21
IC50	inhibitory concentration 50%
IHC	immunohistochemistry
INDELs	insertion/deletion mutations
IL-2	interleukin 2
IL-3	interleukin 3
IL-7	interleukin 7
kb	Kilobase
kDa	Kilo Dalton
*KMT2A*r	*KMT2A*-rearranged/rearrangement
L	Litre
LIV	liver
M	Molar (Moles per litre)
MDS	myelodysplastic syndrome
min	minutes
mL	milli Litre
*MLLT10*r	*MLLT10*-rearranged/rearrangement
mM	milli Molar (10^{-3} molar)
MM	multiple myeloma
MNC	mononuclear cell
MPAL	mixed phenotype acute leukaemia
MRD	minimal residual disease
mRNA	messenger ribonucleic acid
MUT	mutated
NCBI	National Center for Biotechnology Information
NF-H_2O	nuclease-free H_2O
NHEJ	non-homologous end-joining
NHL	non-Hodgkin's lymphoma
nM	nano molar (10^{-9} Molar)
NSG	*NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/Sz*
PB	peripheral blood
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDX	patient-derived xenograft
Ph	Philadelphia chromosome positive ALL
PK	pharmacokinetic
PR	partial response

PRC2	polycomb repressor complex 2
PTD	partial tandem duplication
R/R	relapsed/refractory
RNA	ribonucleic acid
RO	reverse osmosis
rpm	revolutions per minute
RPMI	Roswell Park Memorial Institute (-1640 medium)
RQ-PCR	real-time quantitative polymerase chain reaction
RT	room temperature
RT-PCR	reverse transcription polymerase chain reaction
s	seconds
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	super elongation complex
SPL	spleen
SR	standard risk
T-ALL	T-cell acute lymphoblastic leukaemia
TBS	tris buffered saline
TBST	tris buffered saline with 0.1% Tween® 20
THY	thymus
TKI	tyrosine kinase inhibitor
Tris	tris(hydroxylmethyl)aminomethane
TRXG	trithorax group
UTR	untranslated region
UV	ultra-violet
VAF	variant allele frequency
WT	wild-type

Chapter 1 – Introduction

1.1 Incidence and outcomes of ALL

Acute lymphoblastic leukaemia (ALL) is a haematological malignancy characterised by uncontrolled proliferation of B- or T-lineage lymphoid progenitors in the blood, bone marrow or extramedullary sites such as the spleen. ALL is the most common childhood malignancy, with a peak between 2 and 5 years of age, and 60% of new diagnoses occurring in children[1]. Long-term clinical outcomes are favourable for 85-90% of childhood ALL cases; however, 10-15% of patients will eventually relapse and succumb to their disease making ALL a leading cause of non-traumatic mortality in children[2]. ALL also occurs in infants, adolescents and young adults (AYA) and adults, with a comparatively poor prognosis[2]. Infant ALL (<1 year of age) accounts for approximately 2.5% of paediatric ALL cases, and carries a four-year event-free survival rate of <45%[3]. Long-term survival rates in adults are similarly poor, with 30-40% of adults diagnosed with ALL surviving long-term[4].

ALL prognosis is stratified by several clinical parameters, where high-risk factors include high presenting leukocyte count ($\geq 50 \times 10^9$/L), minimal residual disease (MRD) response, and age (infant or ≥ 10 years of age), and the presence of leukaemia within the central nervous system (CNS)[5,6]. Genomic alterations are also important in ALL diagnosis and prognostic stratification, such as chromosomal translocations and aneuploidy, as determined by cytogenetic or genomic identification[2].

1.2 Biological features of ALL

1.2.1 B-lineage ALL (B-ALL)

B-lineage ALL (B-ALL) is the most common ALL subtype, accounting for 85% of paediatric ALL and 75% of adult ALL cases[7]. B lymphocytes are important in the adaptive immune system, and maturation primarily occurs in the bone marrow. Next generation sequencing has enabled the identification of a diverse range of genomic aberrations involved in B-ALL (**Figure 1.1**). Genomic lesions that drive B-ALL typically involve genes related to B-cell development, such as the transcription factors *ETV6, RUNX1, PAX5* and *IKZF1*. Other genomic subtypes of ALL involve chromosomal copy number alterations, such as hypodiploidy, hyperdiploidy and intrachromosomal

amplification of chromosome 21 (iAMP21). Epigenetic regulators may also be altered in the pathogenesis of B-ALL, such as rearrangements of *MEF2D* and *KMT2A*[8].
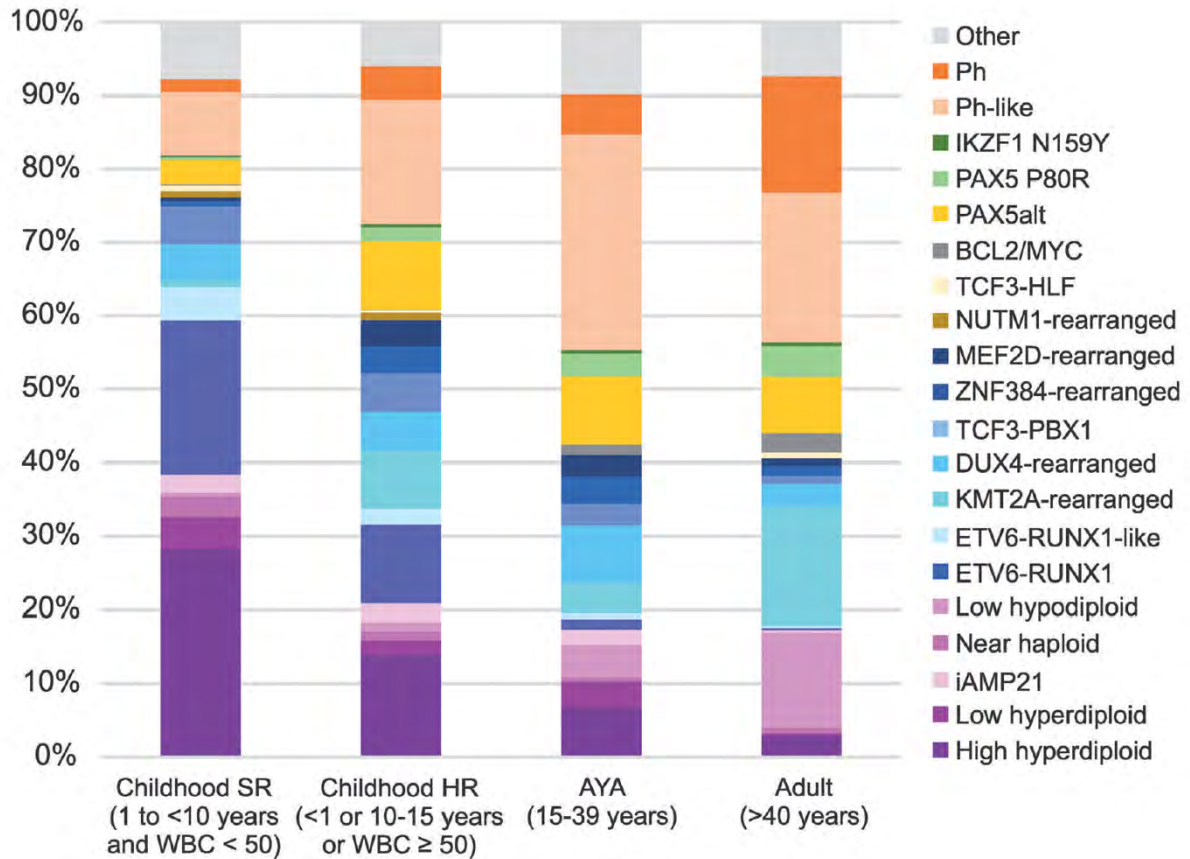
Figure 1.1 – The frequency of genomic subtypes in B-ALL varies depending on age

Stacked column charts depict genomic alterations present in B-ALL based on age group, divided into childhood standard risk (SR), childhood high-risk (HR), adolescent/young adult (AYA) and adult. Abbreviations: iAMP21, intrachromosomal amplification of chromosome 21; PAX5 alt, alterations occurring within the *PAX5* gene; Ph, Philadelphia chromosome positive ALL (also known as *BCR-ABL1* positive ALL); SR, standard-risk. Image used with permission from Inaba & Mulligan 2020[9].

1.2.2 T-lineage ALL (T-ALL)

T-ALL accounts for approximately 15-20% of new ALL diagnoses[10,11]. T lymphocytes are important in the adaptive immune system, and maturation primarily occurs in the thymus. Compared with B-ALL, patients diagnosed with T-ALL are typically older, more likely to be male, present with a higher white cell count, and have a greater incidence of CNS leukaemic involvement[11]. Next generation sequencing has revealed recurrent genomic aberrations in T-ALL (**Figure 1.2**). Genomic alterations that drive T-ALL are genetically diverse, including alterations that affect transcriptional and epigenetic regulators such as *KMT2A* and *MLLT10*, transcription factors such as *TAL1*, cell cycle regulators such as *CDKN2A/B*, and *NOTCH1* alterations.
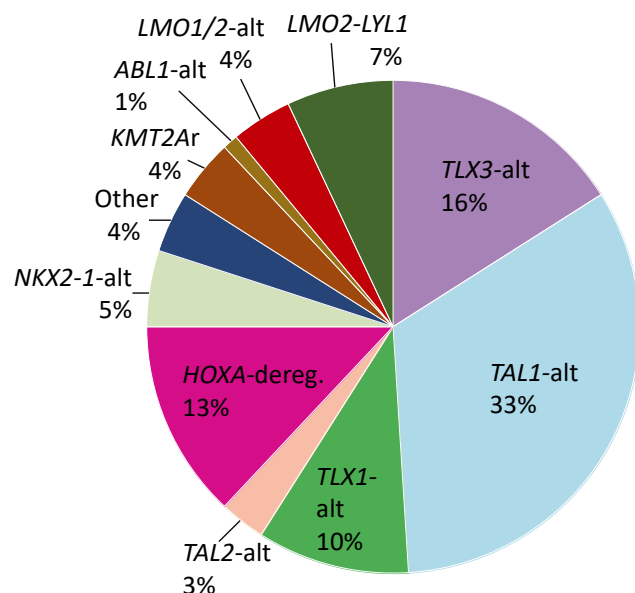
Figure 1.2 – Frequency of genomic subtypes in childhood T-ALL

Pie chart depicts genomic alterations that occur in newly-diagnosed paediatric (1-10 years of age) T-ALL. Common genomic aberrations include genes encoding basic helix-loop-helix transcription factors *TAL1*, *TAL2* and *LYL1*, and homeobox transcription factors such as *TLX1*, *TLX3*, and rearrangements directly involving *HOXA* genes such as *TRB-HOXA9* and *TRB-HOXA11*. The *HOXA*-deregulated subtype includes genomic aberrations not directly involving *HOXA* genes, but that result in aberrant expression of the *HOXA* gene cluster, such as *MLLT10*r and *SET-NUP214*. Image adapted from Teachey & Pui 2019[11].

1.2.3   ALL can be prognostically stratified based on genomic subtype

Cytogenetic and genomic features have an important role in predicting outcome, where recurrently identified genomic lesions are associated with favourable or unfavourable outcomes[8] (**Figure 1.3**). Risk stratification based on both clinical and genomic features informs treatment decisions, where high-risk patients may undergo intensified treatment regimens, potentially including haematopoietic stem cell transplant (HSCT), or the addition of targeted therapies directed towards specific genomic aberrations present. For example, tyrosine kinase inhibitors (TKIs) may be prescribed in ALL cases harbouring *ABL*-class fusions such as *BCR-ABL1* and *EBF1-PDGFRB*[12].
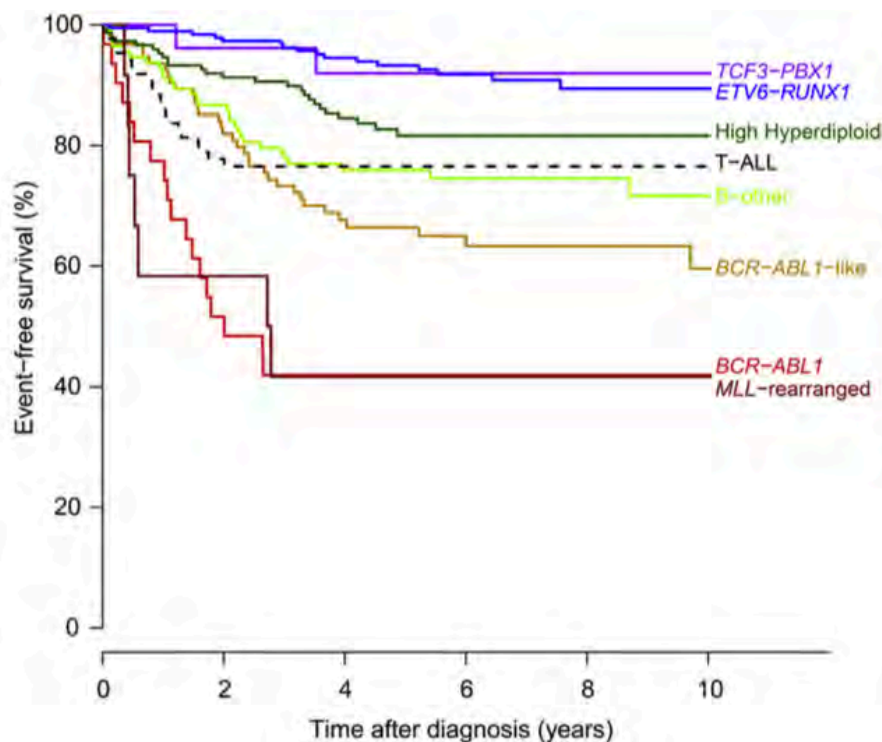
Figure 1.3 – Outcomes of non-relapsed paediatric ALL vary based on genomic subtype

Many recurrent genomic aberrations identified in ALL are associated with either poor or good prognosis. Some genomic subtypes, such as *TCF3-PBX1* and *ETV6-RUNX1*, are associated with favourable outcome. Conversely, the presence of *BCR-ABL1* or *MLL* (referred to in this thesis as *KMT2A*) rearrangements are associated with unfavourable long-term outcomes. Image used with permission from Boer & de Boer (2017)[13].

1.2.4   Genomic cooperation in ALL

Vogelstein's two-hit model of cancer proposes that leukaemia, and all other malignancies, develops after acquisition of a series of genomic alterations over time. An initiating alteration, also known as a gatekeeping mutation, provides an initial selective growth advantage[14]. The initiating alteration may occur sporadically or as an inherited susceptibility[15]. Secondary alterations are then acquired over time, promoting further growth advantage and clonal expansion, eventually resulting in malignancy[14]. Acute leukaemia requires the accumulation of fewer genomic alterations than adult solid tumours, due to the high renewal capacity of healthy haematopoietic tissue[14]. The initiating lesions commonly identified in ALL are discussed in sections **1.2.2-1.2.3**. Cooperative mutations in ALL are typically genes associated with cell cycle and tumour suppressors (*CDKN2A, CDKN2B, TP53*), cytokine receptor and kinases (*CRLF2, JAK1, JAK2, ABL1*), Ras signalling (*NRAS, KRAS, NF1*), and epigenetic regulators (*CREBBP*).[5]

1.3   KMT2A gene rearrangements in leukaemia

KMT2A-rearranged (*KMT2A*r) acute leukaemia is a high-risk genomic subtype that occurs in B-ALL, T-ALL, and acute myeloid leukaemia (AML). It represents a high-risk genomic subtype, regardless of immunophenotype and patient age. Despite extensive research spanning decades, the aetiology and biology underlying the aggressive phenotype of *KMT2A*r acute leukaemia remains largely elusive. In this review, the incidence, biology, aetiology and clinical manifestations of *KMT2A*r in acute leukaemia are discussed.

Statement of Authorship

Title of Paper	*KMT2A* rearranged acute lymphoblastic leukaemia: Unravelling the genomic complexity and heterogeneity of this high-risk disease
Publication Status	☒ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Michelle O. Forgione, Barbara J. McClure, Laura N. Eadie, David T. Yeung, Deborah L. White Published in *Cancer Letters*, 2020, Vol. 469, 410-418.

Principal Author

Name of Principal Author (Candidate)	Michelle Forgione		
Contribution to the Paper	Conceptualised idea, conducted research, constructed manuscript.		
Overall percentage (%)	80%		
Certification:	This paper is a literature review that I conducted during the period of my Higher Degree by Research candidature, and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	3/1/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:
    i.   the candidate's stated contribution to the publication is accurate (as detailed above);
    ii.   permission is granted for the candidate in include the publication in the thesis; and
    iii.   the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Barbara J. McClure		
Contribution to the Paper	Conceptualised idea, critical review of manuscript.		
Signature		Date	4/1/2022

Name of Co-Author	Laura N. Eadie		
Contribution to the Paper	Critical review of manuscript.		
Signature		Date	4/1/2022

Name of Co-Author	David T. Yeung		
Contribution to the Paper	Critical review of manuscript.		
Signature		Date	4/1/2022

Name of Co-Author	Deborah L. White		
Contribution to the Paper	Conceptualised idea, critical review of manuscript.		
Signature		Date	4/1/2022

Mini-review

KMT2A rearranged acute lymphoblastic leukaemia: Unravelling the genomic complexity and heterogeneity of this high-risk disease

Michelle O. Forgione[a,b,*,1], Barbara J. McClure[a,c], Laura N. Eadie[a,c], David T. Yeung[a,c,f], Deborah L. White[a,b,c,d,e]

[a] Cancer Program, Precision Medicine Theme, South Australian Health & Medical Research Institute (SAHMRI), Adelaide, SA, 5000, Australia
[b] School of Biological Sciences, University of Adelaide, SA, 5000, Australia
[c] Faculty of Health and Medical Science, University of Adelaide, Adelaide, SA, 5000, Australia
[d] Australian Genomics Health Alliance (AGHA), The Murdoch Children's Research Institute, Parkville, Vic, 3052, Australia
[e] Australian and New Zealand Children's Oncology Group (ANZCHOG), Clayton, Vic, 3168, Australia
[f] Department of Haematology, Royal Adelaide Hospital, SA, 5000, Australia

A B S T R A C T

KMT2A rearranged (KMT2Ar) acute lymphoblastic leukaemia (ALL) is a high-risk genomic subtype, with long-term survival rates of less than 60% across all age groups. These cases present a complex clinical challenge, with a high incidence in infants, high-risk clinical features and propensity for aggressive relapse.

KMT2A rearrangements are highly pathogenic leukaemic drivers, reflected by the high incidence of KMT2Ar ALL in infants, who carry few leukaemia-associated cooperative mutations. However, transgenic murine models of KMT2Ar ALL typically exhibit long latency and mature or mixed phenotype, and fail to recapitulate the aggressive disease observed clinically. Next-generation sequencing has revealed that KMT2Ar ALL also occurs in adolescents and adults, and potentially cooperative genomic lesions such as PI3K-RAS pathway variants are present in KMT2Ar patients of all ages.

This review addresses the aetiology of KMT2Ar ALL, with a focus on the cell of origin and mutational landscape, and how genomic profiling of KMT2Ar ALL patients in the era of next-generation sequencing demonstrates that KMT2Ar ALL is a complex heterogenous disease. Ultimately, understanding the underlying biology of KMT2Ar ALL will be important in improving long-term outcomes for these high-risk patients.

Background

Histone-lysine N-methyltransferase 2A (KMT2A) rearranged acute lymphoblastic leukaemia (KMT2Ar ALL) is a high-risk genomic subtype that affects more than 70% of new ALL diagnoses in infants (< 1 year of age) [1], 5–6% of paediatric cases [2] and 15% of adult cases [3]. In infants, the median age of diagnosis is 4 months [4], whereas in adults, incidence steadily increases with age, with a median age at diagnosis occurring between 38 and 43 years [3,5]. Patients with KMT2Ar ALL typically present with high-risk clinical characteristics including high white cell count and extramedullary involvement at diagnosis. The immunophenotype of KMT2Ar ALL is typically CD34+CD19+ pro-B or early pre-B cell with variable CD10 expression. Myeloid markers such

as CD15 and CD65 are often coexpressed, hence the original gene name and disease subtype "mixed lineage leukaemia" [6,7]. Disease relapse is common, after which prognosis is extremely poor, demonstrating the ongoing therapeutic challenges presented by this disease [1,8].

Despite extensive research, the aetiology of KMT2Ar ALL remains incompletely understood. For instance, transgenic murine models of KMT2Ar ALL typically exhibit long latency resulting in the expansion of mature or mixed phenotype clones, which fail to recapitulate the aggressive disease observed clinically [9–11]. The in utero origin of KMT2A rearrangement [12] and low frequency of somatic mutations in infants with KMT2Ar ALL [13] is well-characterised, but the initiating events remain poorly understood, particularly in older patients. Notably, the mutational landscape of non-infant KMT2Ar ALL remains

Fig. 1. Domain structure of wild-type KMT2A
KMT2A is proteolytically cleaved by TASPASE 1 into two covalently-linked subunits (KMT2A^N and KMT2A^C). Following proteolytic cleavage,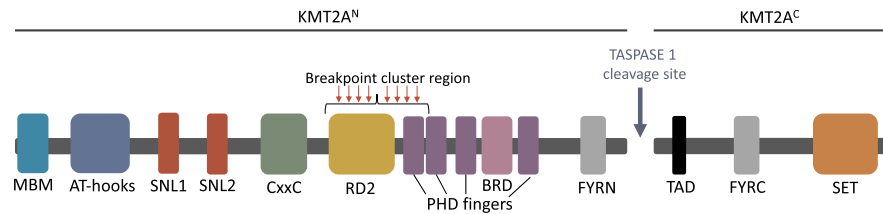 "FY-rich" N-terminal (FYRN) and C-terminal (FYRC) domains are responsible for interactions between KMT2A^N and KMT2A^C subunits, and subnuclear localisation of this complex. Several DNA and protein binding domains are located N-terminally, including the menin binding motif (MBM), DNA binding N-terminal AT-hooks and CxxC region, subnuclear localisation domains (SNL1 & 2) and repression domain (RD2). Four plant homeodomain (PHD) fingers and bromodomain (BRD) are important in post-translational regulation and mediation of protein-protein interactions. *KMT2A* fusion breakpoints are typically located within the 8.3 kb breakpoint cluster region spanning exons 9 to 14. KMT2A^C contains a transactivation domain (TAD) and the SET domain responsible for the H3K4 methyltransferase activity of KMT2A. Figure adapted from Muntean et al. 2012 [2].

unexplored, and how genomic variants influence leukaemia initiation, relapse and response to targeted therapies is an important consideration. This review will discuss the underlying cause of *KMT2A*r ALL across all age groups, with focus on the cell of origin and mutational landscape, to address the ongoing need to improve our understanding of this aggressive disease.

Wild-type *KMT2A*

The *KMT2A* gene (11q23.3), formerly known as mixed lineage leukaemia 1 (*MLL1/MLL/ALL-1/HRX/HTRX1*) encodes the 3969 amino-acid, 500 kDa KMT2A protein (Fig. 1) that functions as an epigenetic regulator of transcriptional initiation and elongation through histone 3 lysine 4 (H3K4) methylation (H3K4me) of target gene promoter regions [14]. KMT2A has a wide range of targets including regulators of haematopoietic cell proliferation and differentiation *Meis homeobox 1* (*MEIS1*) and the *homeobox A* (*HOXA*) gene cluster [15,16]. Deregulation of these genes initiates the development of leukaemia through inhibition of proper haematopoietic development [17].

KMT2A rearrangements in acute leukaemia

The term "mixed lineage leukaemia" was coined in the 1980s to describe cases of high-risk acute leukaemia where leukaemic blasts expressed both lymphoid and myeloid surface markers [18]. This phenomenon was associated with chromosomal translocations on the long arm of chromosome 11 band q23 (11q23), involving a specific gene that was subsequently given numerous names including *MLL* (*mixed lineage leukaemia*), *ALL-1* (*acute lymphoblastic leukaemia 1*) and *HRX* (human trithorax, the human homolog of *Drosophila trithorax, trx*) [19]. It was recently renamed *KMT2A* to reflect its lysine methyltransferase function, and to avoid confusion with the clinical term mixed lineage leukaemia (also known as mixed phenotype acute leukaemia, MPAL).

Cases of acute leukaemia bearing rearrangements of *KMT2A* have been recognised as a separate entity by the World Health Organisation (WHO) since the introduction of the WHO Classification of Neoplastic Diseases of the Haematopoietic and Lymphoid Tissues in 1999 [20]. The current classification recognises *KMT2A*r acute leukaemias under their respective immunophenotypes: acute myeloid leukaemia (AML) with the specific t(9;11)(p21.3;q23.3) translocation (*MLLT3-KMT2A*); MPAL with t(v;11q23.3) (*KMT2A*r) with any fusion partner; and B lymphoblastic leukaemia/lymphoma with t(v;11q23.3) (*KMT2A*r) with any fusion partner [21]. B-ALL and AML are the most commonly diagnosed neoplasms bearing *KMT2A* rearrangements, accounting for 6% [22] of new diagnoses of B-ALL, including 70% of infant ALL cases [1], and 4–14% of new AML diagnoses, including 35–50% of infant AML

cases [2]. For this reason, B-ALL and AML are the focus of most studies addressing the mechanisms of *KMT2A*r leukaemia, but it is important to note that the presence of a *KMT2A* rearrangement is an unfavourable prognostic factor for all subtypes of leukaemia, regardless of patient age [23,24].

MPAL is a comparatively rare, high-risk subtype of leukaemia. Immunophenotypically, blasts exhibit a mix of both myeloid and lymphoid features, or there may be co-existence of distinct clonal populations with separate immunophenotypes, most commonly mixed myeloid/B-cell [24,25]. *KMT2A* rearrangements are identified in 15% of MPAL cases overall, and is disproportionately common in infants [24]. The genomic alterations in MPAL have not been fully elucidated due to the rarity of the disease, and difficulty in diagnosis due to heterogeneity of cell surface marker expression [26]. However, a recent genomic survey suggested that *KMT2A*r MPAL has a distinct mutational pattern and low mutational burden [24].

T-ALL accounts for 10–15% of paediatric and 20–25% of adult ALL overall, and 4–8% of cases harbour a *KMT2A* rearrangement [27–29]. Infant T-ALL is extremely rare, but one study identified *KMT2A* rearrangement in three out of thirteen infants with T-ALL [30]. Fusion partners detected in *KMT2A*r T-ALL are shared with those commonly observed in B-ALL and AML, most commonly *MLLT4* and *MLLT1* (Fig. 2).

Types of genomic lesions involving *KMT2A*

Numerous genomic lesions involving *KMT2A* have been identified in acute leukaemia, including chromosomal translocations, internal tandem duplications, internal deletions and amplifications [16]. Chromosomal translocations are the most common genomic lesion involving *KMT2A* in acute leukaemia, resulting in various fusion genes that express an abnormally functioning fusion protein. *KMT2A* fusion breakpoints are typically located within the 8.3 kb breakpoint cluster region spanning exons 9 to 14 [16] (Fig. 1), conserving the non-specific DNA and protein-binding capacity of *KMT2A*, while the regulatory bromodomain, repression domain and PHD domains are partially or entirely lost. Importantly, the entire KMT2A^C fragment is truncated and therefore intrinsic H3K4 methyltransferase activity is lost. *KMT2A*r patients and *in vitro KMT2A*r models both exhibit widespread aberrant epigenetic dysregulation, particularly involving H3K4me and H3 lysine 79 methylation (H3K79me) [31,32]. H3K4me dysregulation is the direct consequence of aberrant *KMT2A* H3K4me activity, but *KMT2A* rearrangements also induce altered interactions with other histone writers such as DOT1L, resulting in aberrant H3K79me patterns.

More than 80 direct fusion partner genes of *KMT2A* have been identified across B-ALL, T-ALL and AML [14], and all are associated with poor outcomes, except for the rare *KMT2A-MLLT11* which confers

Fig. 2. Frequency of *KMT2A* fusion partners found in ALL and AML

Common *KMT2A*r ALL fusion partners in ALL (upper centre and right) include *AFF1* (formerly *AF4*), *MLLT1* (*ENL*), *MLLT3* (*AF9*), *MLLT10* (*AF10*)and *MLLT4* (*AF6*). All of these partner genes also occur in *KMT2A*r AML (upper left). *KMT2A-EPS15* occurs in both B-ALL and AML, but is more common in B-ALL. *ELL* and partial tandem duplications (PTD) of *KMT2A* are common in AML but rare in ALL. The remaining "other" category refers to unidentified or uncommon fusion partners, often identified in single cases. The frequency of fusion partners in B-ALL differs depending on patient age (lower panel), but *AFF1* remains the most common in all age groups. B-ALL and AML figures adapted from Meyer et al. 2018 [40] and Marschalek 2016 [16], and T-ALL figure adapted from Peterson et al. 2018 [23] and Meyer et al. 2018 [40].

a very good prognosis [33,34]. The mechanism behind the favourable prognosis of *KMT2A-MLLT11* is not known, but may relate to the function of *MLLT11*. *MLLT1 1* has roles in haematopoietic cell fate specification, and high expression of *MLLT11* is correlated with adverse outcomes in AML and myelodysplastic syndrome; this may suggest that any loss of function mutations in *MLLT11*, such as truncations, may be a favourable prognostic feature [35]. In addition, unlike other common *KMT2A* fusion partners, *MLLT1 1* has no known direct role in transcriptional initiation or elongation, which may be a contributing factor to the pathogenicity of *KMT2A* rearrangement with other fusion partners such as *KMT2A-AFF1* [36]. Data on *KMT2A-MLLT11* breakpoints or broader genomic profiling of *KMT2A-MLLT11* patients are not available, but would be useful to ascertain the oncogenic mechanisms of this fusion.

The frequency of common fusion partners in *KMT2A*r ALL and AML is summarised in Fig. 2. Many fusion partners, including *MLLT1, MLLT3* and *MLLT10*, are common in both ALL and AML, but there are key differences in the frequencies of some fusion partners [37]. For instance, *AFF1* is extremely rare in AML despite accounting for 60% of *KMT2A*r ALL cases, and partial tandem duplications (PTD) and *ELL* fusions are only common in AML (Fig. 2). The role of the fusion partner in lineage determination of *KMT2A*r acute leukaemia is unclear, and is likely to vary depending on the specific fusion partner [38]. Lin et al. 2017 demonstrated that *KMT2A-AFF1* haematopoietic stem or

progenitor cells (HSPCs) induced expansion of a CD19⁺ lymphoid population under myeloid culture conditions, but had compromised long-term renewal and incomplete oncogenic transformation, at secondary transplant, into myeloid-promoting NSG mice, whereas *KMT2A-MLLT3* efficiently resulted in the development of AML [39]. The inability of *KMT2A-AFF1* to induce complete oncogenic transformation of CD19⁺ lymphoid cells demonstrates that lymphoid lineage identity is linked with the full transforming capacity of this fusion. Comparatively, other fusions such as *KMT2A-MLLT3* that are common in both ALL and AML likely have a less lineage-specific mechanism of oncogenic transformation. The distribution of fusion partners in MPAL is less clear, due to the rarity of the disease.

The four most common *KMT2A*r ALL fusion partners account for more than 90% of cases, and all have shared roles in transcriptional regulation. *AFF1, MLLT1* and *MLLT3* are components of the transcriptional super elongation complex that recruits RNA polymerase II [36], and *MLLT10, MLLT1* and *MLLT3* are critical components of the DOT1L complex, enabling H3K79 methylation. Dysregulation of both of these complexes results in an aberrant gene expression profile that is characteristic of *KMT2A*r ALL, most notably upregulation of the haematopoietic developmental regulators *MEIS1* and the *HOXA* gene cluster [41–43]. Downregulation of protein transport regulators is also observed, that may cause aberrant cellular localisation of oncoproteins and tumour suppressors [42].

In *KMT2Ar* ALL, the fusion partner is the crucial link between truncated 5′ *KMT2A* and dysregulation of cellular functions, such as transcriptional elongation, enabling aberrant gene expression that results in leukaemogenesis. Experimental evidence supports that truncation of *KMT2A* alone is not sufficient for leukaemic transformation [44,45], but rare cases of 3′ *KMT2A* deletions (ie. *KMT2A* truncation without a fusion partner) have been reported in ALL [23,27,45]. In one study, 3′ *KMT2A* deletions were detected in two T-ALL cases (7% of the reported *KMT2Ar* T-ALL cases in the study cohort) by conventional cytogenetics. Both patients had additional T-ALL-associated genomic lesions; one with *TLX3* rearrangement and homozygous *CDKN2A* deletion, and the other with gain of *TAL1* [23]. This suggests that patients with 3' *KMT2A* deletions may require additional genomic events to induce leukaemia, but it is not known whether their gene expression profile, overall survival or response to therapy are similar to other *KMT2Ar* patients.

Incidence and outcomes

Long-term outcomes of *KMT2Ar* ALL are poor compared to other genomic subtypes of ALL, irrespective of patient age. Complete remission is initially achieved in most cases ($\geq$ 90% [46]) and 82.5–93% [5,7] for infants and adults respectively), but disease control is typically short-lived and relapse occurs rapidly, after which outcomes are dismal. Allogeneic stem cell transplant remains the standard of care in appropriate cases. For instance, of the 85 patients with *KMT2A-AFF1* rearrangements in the UKALLXII/ECOG 2993 study, 93% achieved remission; the 5 year overall survival rate was 56% and 67% respectively for patients who received sibling matched and matched unrelated donor allografts, versus 24% in those who did not. Indeed, many risk stratified combination chemotherapy protocols would classify *KMT2A* rearrangement as being adverse risk, and recommend either treatment intensification and/or transplant in first remission. *KMT2Ar* ALL patients also typically exhibit high minimal residual disease (MRD), which is correlated with poor outcomes [47]. Event-free survival (EFS) rates of *KMT2Ar* ALL from representative studies are provided in Table 1.

In utero origin of KMT2A rearrangements

The identification of *KMT2A* rearrangements in neonatal blood spots of infants with B- and T-ALL indicates that *KMT2A* fusions can arise *in utero* [12,30,54]. This is supported by cases of monochorionic twins with concordant leukaemia that share a *KMT2A* rearrangement with identical breakpoints, where the fusion likely originates in one foetus and is propagated to the other through the shared placenta [6,55]. Concordance rate is close to 100% in these cases, demonstrating the high pathogenicity of *KMT2A* rearrangements [56]. *KMT2Ar* ALL is considered to be a disease of short latency, as many cases are diagnosed days or months after birth, but *KMT2A* rearrangements are also

identified at birth in blood spots from patients diagnosed up to two years later [1,12]. The *in utero* origin of other ALL genomic drivers such as *ETV6-RUNX1* and *BCR-ABL1* have also been demonstrated in patients that develop ALL later in childhood (ie > 2 years), but these genomic subtypes are rarely diagnosed in infancy, indicating that these fusions require additional genomic hits or other triggers to induce leukaemia [57]. There is no published data available on the analysis of blood spots from older children and adults, and a negative result would be uninterpretable; the fusion may not be present at birth, or it may be present at an undetectably low frequency [58].

Cell of KMT2A fusion origin

Given the short disease latency between initiation of *KMT2A* rearrangement and disease development, high concordance rate amongst twins and high incidence of *KMT2Ar* ALL in infants, there is perhaps an intrinsic property of foetal haematopoietic stem and progenitor cells that confers susceptibility to leukaemogenic transformation by *KMT2A* rearrangements [59]. This is supported by a doxycycline inducible mouse model of *KMT2A-MLLT1*, where acute myeloid leukaemia (AML) developed more frequently and rapidly in foetal and neonatal mice than in juveniles and adults [60].

Transgenic murine models of *KMT2Ar* ALL are difficult to establish, and existing models typically develop mature B-lymphoid malignancies or mixed lymphoid-myeloid hyperplasia with long latency periods (median latency between 131 and 520 days) [9–11]. This may be attributed to developmental haematopoietic differences between mice and humans, or possibly an indication that these models are missing important elements, potentially the cell type used to model *KMT2A* rearrangement initiation, or an immune or environmental event.

It is generally accepted that *KMT2Ar* pro-B ALL initiates in a CD34$^+$CD19$^-$ lymphoid-restricted but non-lineage-committed progenitor that is capable of differentiation into a leukaemic blast with a B-lineage immunophenotype [56,61–63]. Cases of mixed phenotype acute leukaemia bearing *KMT2A* rearrangement provide supporting evidence that *KMT2A* rearrangements affect a non-committed progenitor cell type [24,25]. *KMT2A-AFF1* expression in human CD34$^+$ haematopoietic stem or progenitor cells (HSPCs) enhances haematopoietic repopulation and increases clonogenicity *in vitro* but *KMT2A-AFF1* expression alone was not sufficient to induce overt leukaemia *in vivo*. This suggests that either additional genomic hits are required to induce overt ALL, or that CD34$^+$ HSPCs are not an appropriate cellular model for *KMT2A-AFF1* ALL [64].

Two independent studies demonstrate expression of *KMT2A-AFF1* in pre-leukaemic mesenchymal stem cell (MSC) cultures expanded from the bone marrow of infants and paediatric cases of *KMT2Ar* pro-B ALL [65,66]. Importantly, Shalapour et al. also detected *ETV6-RUNX1, KMT2A-MLLT1* and *TCF3-PBX1* expression in MSCs from paediatric patients with these fusions. Conversely, Menendez et al. did not detect

Table 1
Reported five-year event-free survival (EFS) rates for ALL patients with *KMT2Ar* and non-*KMT2Ar* ALL.

EFS of *KMT2Ar* ALL	EFS of age group overall[a]	Cohort Group	Reference
20.9%[cd]	49.4%[c]	Infants (< 12 months of age), Interfant-06 protocol (n = 164)	Pieters et al., 2019 [48]
35.5%	42.3%	Infants (< 12 months of age), various protocols (n = 141)	Dreyer et al., 2015 [49]
36.9%[x]	46.4%	Infants (< 12 months of age), Interfant-99 protocol (n = 308)	Pieters et al., 2007 [50]
33.6%	41.7%	Infants (< 12 months of age), CCG 1953 protocol (n = 115)	Hilden et al., 2006 [51]
59.3%	87.3%	Paediatric (< 18 years of age), various protocols (n = 15)	Pui et al., 2014 [52]
34%	35–40% [53]	Adolescent and adult *KMT2A-AFF1* BCP-ALL (B-cell precursor ALL) (15–59 years of age), UKALLXII/ECOG 2993 (n = 85)	Marks et al., 2013 [5]
32.7%	35–40% [53]	Adult ($\geq$ 21 years of age), *KMT2Ar* BCP-ALL, various protocols (n = 94)	Roberts et al., 2017 [3]

[a] Including *KMT2Ar* cases.
[x] 4-year event-free survival.
[c] 6-year event-free survival.
[d] High-risk patients only (< 6 months of age plus white cell count > 300 $\times$ 10^9/L).

expression of *ETV6-RUNX1*, *BCR-ABL1,* hyperdiploid ALL or *KMT2A-MLLT3* in MSCs of ALL patients harbouring these genomic aberrations at diagnosis. The implication of fusion identification in MSCs is that fusions would be acquired in a very primitive cell of origin in the patients studied, or alternatively, that lymphoid-restricted leukaemic blast cells or leukaemia stem cells are capable of dedifferentiation through epigenetic reprogramming. Leukaemia stem cells are characterised by loss of cell identity, enhanced plasticity and altered self-renewal properties, and may contribute to the high rates of disease relapse and lineage switch observed in *KMT2A*r ALL [67,68].

It is not known whether the originating cell of *KMT2A*r ALL is shared across all age groups, and this is an important consideration when modelling this disease for the investigation of therapeutic targets.

Lineage plasticity and myeloid lineage switch

Aberrant myeloid marker expression at diagnosis and the propensity for myeloid lineage switch at relapse demonstrates the lineage plasticity of *KMT2A*r ALL. Myeloid lineage switch is rare in ALL but is associated with *KMT2A* rearrangements, particularly in response to CD19-directed bi-specific T-cell engaging therapy such as blinatumomab [8,69–71]. Lineage switch typically occurs rapidly after commencement of therapy (median of 15 days after commencement of blinatumomab [72]), and outcomes are extremely poor although there are cases of disease control with AML chemotherapy protocols [70,71,73]. A case study by Wolfl et al. further emphasises the lineage plasticity of *KMT2A*r ALL. They describe myeloid lineage switch in a five month-old infant with *KMT2A*r B-ALL following commencement of blinatumomab, followed by spontaneous reversion back to B-ALL 7 days after treatment cessation [74].

Lineage switch is evidently a mechanism of therapeutic evasion, but the exact means by which it occurs is unknown. It is difficult to distinguish between true lineage switch and selective cell culling through ALL-directed therapy, such that non-ALL clones proliferate due to a selective advantage. The short time from therapy initiation to emergence of lineage-altered clones, and cases of *KMT2A*r ALL lineage switch following non-lineage-discriminate chemotherapy suggests the latter as a more likely possibility [72,75]. A true lineage switch may result from a non-lineage-committed leukaemia stem cell that drives repopulation of lineage-switched leukaemic blasts [8,72], or the global epigenetic dysregulation induced by *KMT2A* rearrangements and the dysregulated roles of *KMT2A* in haematopoiesis and broader development may enable reprogramming of leukaemic blasts themselves.

The role of environmental exposures in *KMT2A*r ALL

ALL, and likely all cancers, arise from a complex combination of environmental exposures, inherited susceptibility and chance. The two-hit genomic model of childhood ALL proposes that an initiating alteration such as *ETV6-RUNX1* occurs *in utero* to induce the formation of a pre-leukaemic clone. Cooperative secondary mutations then occur, resulting in the eventual initiation of overt ALL [58].

There is growing epidemiological evidence to suggest that microbial exposures early in life play an important role in leukaemia onset in predisposed children. Greaves proposes the delayed infection hypothesis, where exposure to common infections in infancy are protective against B-ALL, whereas lack of exposure as infants results in dysregulated responses upon exposure to infective agents later in childhood, and this may promote or trigger the secondary genomic events that result in the development of ALL [58,76]. Epidemiological evidence suggests that exposure of infants to common infective agents through vaginal delivery, childcare, sibling exposure and breastfeeding promotes the development of a healthy microbiota, priming individuals for protection against common infections at an early age [58]. It is the absence of this early exposure that may result in a dysregulated response upon exposure to infection later in childhood, resulting in ALL. Greaves

describes this as a "paradox of progress", in that ALL is largely a consequence of incongruity between evolutionary programming of the immune system and the modern developed world's anti-germ lifestyle [58].

It has been historically accepted that *KMT2A*r pro-B infant ALL does not endorse the typical two-hit genomic model of cancer development, but rather that leukaemogenesis is essentially completed *in utero* through the acquisition of a *KMT2A* rearrangement which serves as a single genomic hit sufficient to induce ALL [58]. However, *KMT2A* rearrangements have been identified in neonatal blood spots from patients diagnosed up to two years of age, demonstrating a long disease latency between fusion origin and leukaemia development in these patients [12]. The identification of *RAS* pathway mutations in infant and paediatric *KMT2A*r patients at diagnosis and relapse [63] suggests that secondary genomic events may have an initiating role in some cases of *KMT2A*r ALL. There is therefore a potential role for environmental exposures such as infection in *KMT2A*r ALL, but this remains to be explored.

The role of ethnicity in onset of *KMT2A*r ALL

It is unclear whether incidence of *KMT2A*r ALL is influenced by ethnicity. In a Brazilian cohort of ALL cases aged 0–24 months, non-Caucasian patients had a higher incidence of *KMT2A* rearrangement than Caucasian infants [77]. Conversely, in a study by Sam et al., infants of African-American ethnicity with ALL were less likely to harbour a *KMT2A* rearrangement compared to Caucasian infants [78]. Interpretation of these studies is difficult due to low non-Caucasian sample sizes, but further research in this area would be beneficial to determine whether factors relating to ethnicity may contribute to the aetiology of *KMT2A*r ALL.

The mutational landscape of *KMT2A*r ALL

The very low mutational burden of infant *KMT2A*r acute leukaemia suggests that cooperative mutations are not required to induce overt leukaemia in these patients [13,63]. Non-*KMT2A*r infant ALL is rare but two studies observe a similarly low mutational burden in infants with ALL independent of *KMT2A* rearrangement status, suggesting that the mechanisms of leukaemogenesis are different between infants and non-infants independent of genomic status [13,63]. Regardless, the mutational landscape of infant ALL is distinct from ALL in children and adults, where alterations in other cancer-associated genes are common, such as *CDKN2A, IKZF1, KRAS* and *CREBBP* [22,79]. Comparatively, the presence of cooperative mutations in paediatric and adult *KMT2A*r ALL remains largely unexplored, so it is uncertain whether a low mutational burden is characteristic of *KMT2A*r ALL in non-infants.

***PI3K-RAS* pathway mutations**

Despite an overall low mutational burden, *KMT2A*r infant and paediatric ALL cases often present with a *PI3K-RAS* pathway mutation at diagnosis (Table 2), whereas only one study reports the frequency of *KRAS* and *NRAS* mutations in adult *KMT2A*r ALL, and this was much lower in adults than in paediatric cases (8% and 26% respectively) [80]. Differences in reported mutation frequencies (Table 2) are likely due to discrepancies in the panel of *PI3K-RAS* pathway genes included in analysis, and the inclusion of mutations with low variant allele frequency (VAF).

In infants, *PI3K-RAS* pathway mutations are often subclonal at diagnosis (VAF $\leq 30\%$). Subclonal variants are lost at relapse in up to 50% of cases, but in other instances expand into the dominant clone at relapse, suggesting that these mutations can provide a survival advantage but are not necessary for leukaemia maintenance [13,63,83,84]. *PI3K-RAS* variants are also present in a substantial portion of both infant non-*KMT2A*r and non-infant paediatric *KMT2A*r

Table 2
Frequency of *PI3K-RAS* pathway mutations among infant and paediatric cases of *KMT2A*r ALL.

Reference	Infant *KMT2A*r ALL	Infant non-*KMT2A*r ALL	Paediatric *KMT2A*r ALL
	Frequency of *PI3K-RAS* pathway mutations (%)		
Agraz-Doblas et al., 2019 [63]	38%	26%	Not reported
Liang et al., 2018 [81]	25.9% (*KRAS* only)	Not reported	Not reported
Fedders et al., 2017 [82]	22.1%	Not reported	13.9%
Trentin et al., 2016 [83]	77.2%	Not reported	33.3%
Andersson et al., 2015 [13]	47%	39%	50%
Emerenciano et al., 2015 [84]	21.7%	Not reported	Not reported
Driessen et al., 2013 [85]	14.3%	7.8%	Not reported

ALL, and Andersson et al. observe that *PI3K-RAS* pathway variants were more often present in the major clone at diagnosis in non-infant paediatric patients. Interestingly, Tamai et al. reports that the addition of *KRAS* p.G12D to a transgenic mouse model of *KMT2A-AFF1* leukaemia/lymphoma accelerated disease onset to 6 months, from a median of 14 months with *KMT2A-AFF1* alone [11].

The prognostic significance of *PI3K-RAS* pathway mutations is unclear. One study identifies the presence of a *RAS* pathway mutation in *KMT2A*r infants as an independent predictor for poor outcome, and was associated with high-risk clinical factors such as high white cell count at diagnosis and glucocorticoid resistance [85]. Others report no statistically significant differences in prednisolone response [83], clinical outcome [63,81,83] or diagnostic parameters such as percentage of blasts and white blood cell count [63]. The prognostic significance of *PI3K-RAS* pathway variants in adult *KMT2A*r ALL remains unexplored.

Overall, *PI3K-RAS* pathway mutations may provide a proliferative advantage but are not necessary for the maintenance of *KMT2A*r ALL in children and infants, as these variants are often not maintained at relapse. Further investigation of the incidence of *PI3K-RAS* pathway variants in adult *KMT2A*r ALL, and resultant effects on outcome, is required.

Epigenetic regulatory gene mutations in *KMT2A*r ALL

The role of the epigenome in cancer is an emerging topic of interest, and mutations in genes that encode direct epigenetic regulators are common in many paediatric malignancies including ALL, but are rare in infant ALL [86]. Widespread epigenetic dysregulation is a hallmark of *KMT2A*r ALL, resulting in aberrant gene expression. Specifically, global H3K79me [87] and hypomethylation of specific promoter regions including the *HOXA* gene cluster are observed [88]. Several classes of therapies targeted toward the epigenome are emerging as potential treatment avenues for various cancers including *KMT2A*r ALL. Andersson et al. observed that mutations in epigenetic regulators are present in 45% of paediatric *KMT2A*r cases (7–19 years of age) compared to 14% of an infant cohort. Frequently mutated genes included *CHD4, SETD2* and *CREBBP*, but the functional impact and prognostic significance of these variants remains mostly unknown in both childhood and adult *KMT2A*r ALL [13].

Treatment of *KMT2A*r ALL

Chemotherapy remains the standard of care for *KMT2A*r ALL, though outcomes are significantly inferior compared to cases without *KMT2A* rearrangement, especially in infant ALL. In the Children's Cancer Group study CCG 1953, for instance, the 5 year EFS for *KMT2A*r infant ALL was 34% vs 60% in cases without *KMT2A*r [51]. There may be subtle differences between cases of *KMT2A* rearrangements, where *MLLT11* may be more favourable vs *MLLT10* which is unfavourable [51,89]. In most cases, remission is initially achieved, only to be followed by rapid relapse, rather than true induction failures, suggesting that selection of chemotherapy-resistant clones is responsible for

treatment failure. As *KMT2A*r ALL is postulated to have derived from an early haematopoietic precursor, chemotherapy traditionally used in AML had been trialled as consolidation, though such an approach did not show additional benefit [48].

The role of allogeneic stem cell transplant as consolidation in infant ALL is yet to be clearly established. The benefits of long-term disease control through graft versus leukaemia effect is balanced by treatment-related mortality and long-term toxicity of this strategy [90]. Infants diagnosed at a very young age with a high white cell count at diagnosis and those with persistence of MRD may stand to benefit [49,89,91]. In contrast, allogeneic transplantation is recommended under most adult treatment algorithms for *KMT2A*r ALL.

***PI3K-RAS* pathway inhibition**

Mutations activating the *PI3K-RAS* pathway are common in *KMT2A*r ALL, and can synergise with *KMT2A* rearrangements to reduce leukaemia latency [11,15]. MEK inhibitors such as trametinib reduce leukaemic burden in bone marrow and delay progression of *RAS* mutant *KMT2A*r ALL in murine xenografts, and enhance prednisolone sensitivity of *KMT2A*r cell lines regardless of *RAS* mutation status [92,93]. MEK inhibitors have been trialled clinically for a range of solid tumours and haematological malignancies with varying degrees of success, but efficacy in humans with *KMT2A*r acute leukaemia is unknown. The *NRAS* p.Q61K mutated *KMT2A-AFF1* B-ALL cell line MI04 is sensitive to MEK/ERK inhibition, whereas *NRAS* wild-type B-ALL cell lines RS4;11 and SEM are insensitive, indicating that MEK/ERK pathway inhibition is a potential therapeutic option only for patients with specific genomic aberrations that confer sensitivity [83]. This emphasises the potential benefit of a precision medicine approach in the treatment of high-risk *KMT2A*r ALL patients.

FLT3 inhibitors

FLT3 overexpression is common in patients with *KMT2A*r ALL, even in cases lacking *FLT3* activating mutations [89,94]. Inhibitors of FLT3 have been successfully used in the treatment of AML with *FLT3* activating mutations [95,96]. The Children's Oncology Group AALL0631 study was the first trial to incorporate a FLT3 inhibitor for use in frontline treatment of *KMT2A*r infant ALL, but failed to demonstrate additional benefit of the first-generation FLT3 inhibitor lestaurtinib [89]. Second generation FLT3 inhibitors such as quizartinib may produce more effective results due to enhanced potency and a longer half-life *in vivo* [97]. One study evaluated quizartinib in four cases of *KMT2A*r ALL, including three infants and one adolescent, but no objective clinical responses or cytotoxic effects in bone marrow were observed. Further studies are required for full assessment of FLT3 inhibition using second-generation FLT3 inhibitors in *KMT2A*r ALL.

Therapies targeted towards the epigenome

Several classes of therapies targeted toward the epigenome are

currently in pre-clinical and clinical development for the treatment of *KMT2A*r ALL. These therapies exploit the widespread epigenetic dysregulation that underlies *KMT2A* rearrangement-mediated acute leukaemia. *DOT1L* is the direct target of pinometostat (EPZ-5676) [41], histone deacetylases (HDACs) are targeted by several HDAC inhibitors including panobinostat [98], birabresib (OTX-015) targets *BRD2, BRD3* and *BRD4* [99], and azacitidine is a hypomethylating agent [100]. Many of these agents have limited efficacy or substantial toxicities as single agents [99,101,102], but clinical trials of combination approaches such as pinometostat or azacitidine in combination with standard chemotherapy in *KMT2A*r acute leukaemia are underway (ClinicalTrials.gov identifiers: NCT03724084, NCT02828358).

Conclusions and perspectives

*KMT2A*r ALL is a complex disease with poor outcomes in patients of all ages. The cell of origin, fusion partner, cooperative genomic lesions and environmental factors are all likely to contribute to the initiation, disease course and therapeutic response of *KMT2A*r ALL. There is growing evidence to suggest that the cell of fusion gene origin is an important aetiological factor, and this may explain why existing *in vivo* models have failed to recapitulate the aggressive phenotype of *KMT2A*r ALL.

Much of *KMT2A*r ALL research has occurred in the setting of infants, despite evidence to suggest considerable differences between the aetiology of ALL in infants and non-infants, regardless of genomic subtype. *KMT2A*r ALL in adults is comparatively poorly characterised despite very poor long-term outcomes, and characterising the genomic profile of these patients would be beneficial to understand how the underlying biology of *KMT2A*r ALL differs depending on age of onset. Of particular significance, the mutational landscape of *KMT2A*r ALL in adults remains to be explored. It is possible that a different system should be used to study ALL in infants, and this has implications in understanding the aetiology and therapeutic sensitivities of *KMT2A*r ALL overall. Ultimately, understanding how *KMT2A*r ALL differs with patient age is a crucial factor in improving long-term outcomes for patients with this aggressive disease.

Authorship statement

M.O.F., B.J.M. and D.L.W. conceptualised the presented idea. M.O.F. constructed the manuscript in consultation with B.J.M., L.N.E., D.T.Y. and D.L.W. All authors provided critical feedback and helped shape the manuscript.

Declaration of competing interest

D.L.W. receives research support from BMS, and Honoraria from BMS and Amgen. D.T.Y. receives research support from BMS & Novartis, and Honoraria from BMS, Novartis, Pfizer and Amgen. None of these agencies have had a role in the preparation of this manuscript. All other authors declare no conflicts of interest.

Acknowledgements

References

[1] A.C. Winters, K.M. Bernt, MLL-rearranged leukemias-an update on science and clinical approaches, Front. Pediatr. 5 (2017) 4.

[2] A.G. Muntean, J.L. Hess, The pathogenesis of mixed-lineage leukemia, Annu. Rev. Pathol. 7 (2012) 283–301.

[3] K.G. Roberts, Z. Gu, D. Payne-Turner, K. McCastlain, R.C. Harvey, I.M. Chen, D. Pei, I. Iacobucci, M. Valentine, S.B. Pounds, L. Shi, Y. Li, J. Zhang, C. Cheng, A. Rambaldi, M. Tosi, O. Spinelli, J.P. Radich, M.D. Minden, J.M. Rowe, S. Luger, M.R. Litzow, M.S. Tallman, P.H. Wiernik, R. Bhatia, I. Aldoss, J. Kohlschmidt, K. Mrozek, G. Marcucci, C.D. Bloomfield, W. Stock, S. Kornblau, H.M. Kantarjian, M. Konopleva, E. Paietta, C.L. Willman, C.G. Mullighan, High frequency and poor outcome of philadelphia chromosome-like acute lymphoblastic leukemia in adults, J. Clin. Oncol. 35 (2017) 394–401.

[4] M. Kato, D. Hasegawa, K. Koh, K. Kato, J. Takita, J. Inagaki, H. Yabe, H. Goto, S. Adachi, A. Hayakawa, Y. Takeshita, A. Sawada, Y. Atsuta, K. Kato, Allogeneic haematopoietic stem cell transplantation for infant acute lymphoblastic leukaemia with KMT2A (MLL) rearrangements: a retrospective study from the paediatric acute lymphoblastic leukaemia working group of the Japan Society for Haematopoietic Cell Transplantation, Br. J. Haematol. 168 (2015) 564–570.

[5] D.I. Marks, A.V. Moorman, L. Chilton, E. Paietta, A. Enshaie, G. DeWald, C.J. Harrison, A.K. Fielding, L. Foroni, A.H. Goldstone, M.R. Litzow, S.M. Luger, A.K. McMillan, J. Racevskis, J.M. Rowe, M.S. Tallman, P. Wiernik, H.M. Lazarus, The clinical characteristics, therapy and outcome of 85 adults with acute lymphoblastic leukemia and t(4;11)(q21;q23)/MLL-AFF1 prospectively treated in the UKALLXII/ECOG2993 trial, Haematologica 98 (2013) 945–952.

[6] A. Sanjuan-Pla, C. Bueno, C. Prieto, P. Acha, R.W. Stam, R. Marschalek, P. Menéndez, Revisiting the biology of infant t (4; 11)/MLL-AF4 + B-cell acute lymphoblastic leukemia, Blood 126 (2015) 2676–2685.

[7] C. Motlló, J.M. Ribera, M. Morgades, I. Granada, P. Montesinos, S. Brunet, J. Bergua, M. Tormo, R. García-Boyero, J. Sarrá, E. Del Potro, C. Grande, P. Barba, T. Bernal, M.L. Amigo, J. Grau, J. Cervera, E. Feliu, Frequency and prognostic significance of t(v;11q23)/KMT2A rearrangements in adult patients with acute lymphoblastic leukemia treated with risk-adapted protocols, Leuk. Lymphoma 58 (2017) 145–152.

[8] A. Rayes, R.L. McMasters, M.M. O'Brien, Lineage switch in MLL-rearranged infant leukemia following CD19-directed therapy, Pediatr. Blood Cancer 63 (2016) 1113–1115.

[9] W. Chen, Q. Li, W.A. Hudson, A. Kumar, N. Kirchhof, J.H. Kersey, A murine Mll-AF4 knock-in model results in lymphoid and myeloid deregulation and hematologic malignancy, Blood 108 (2006) 669–677.

[10] M. Metzler, A. Forster, R. Pannell, M.J. Arends, A. Daser, M.N. Lobato, T.H. Rabbitts, A conditional model of MLL-AF4 B-cell tumourigenesis using invertor technology, Oncogene 25 (2006) 3093–3103.

[11] H. Tamai, K. Miyake, M. Takatori, N. Miyake, H. Yamaguchi, K. Dan, T. Shimada, K. Inokuchi, Activated K-Ras protein accelerates human MLL/AF4-induced leukemo-lymphomogenicity in a transgenic mouse model, Leukemia 25 (2011) 888–891.

[12] K.B. Gale, A.M. Ford, R. Repp, A. Borkhardt, C. Keller, O.B. Eden, M.F. Greaves, Backtracking leukemia to birth: identification of clonotypic gene fusion sequences in neonatal blood spots, Proc. Natl. Acad. Sci. 94 (1997) 13950–13954.

[13] A.K. Andersson, J. Ma, J. Wang, X. Chen, A.L. Gedman, J. Dang, J. Nakitandwe, L. Holmfeldt, M. Parker, J. Easton, The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias, Nat. Genet. 47 (2015) 330.

[14] D. Steinhilber, R. Marschalek, How to effectively treat acute leukemia patients bearing MLL-rearrangements? Biochem. Pharmacol. 147 (2018) 183–190.

[15] R. Ono, H. Kumagai, H. Nakajima, A. Hishiya, T. Taki, K. Horikawa, K. Takatsu, T. Satoh, Y. Hayashi, T. Kitamura, T. Nosaka, Mixed-lineage-leukemia (MLL) fusion protein collaborates with Ras to induce acute leukemia through aberrant Hox expression and Raf activation, Leukemia 23 (2009) 2197–2209.

[16] R. Marschalek, Systematic classification of mixed-lineage leukemia fusion partners predicts additional cancer pathways, Ann. Lab. Med. 36 (2016) 85–100.

[17] C.W. Chen, S.A. Armstrong, Targeting DOT1L and HOX gene expression in MLL-rearranged leukemia and beyond, Exp. Hematol. 43 (2015) 673–684.

[18] J. Mirro, T. Zipf, C. Pui, G. Kitchingman, D. Williams, S. Melvin, S. Murphy, S. Stass, Acute mixed lineage leukemia: clinicopathologic correlations and prognostic significance, Blood 66 (1985) 1115–1123.

[19] A.K.N. Chan, C.W. Chen, Rewiring the epigenetic networks in MLL-rearranged leukemias: epigenetic dysregulation and pharmacological interventions, Front. Cell. Dev. Biol. 7 (2019) 81.

[20] N. Harris, E. Jaffe, J. Diebold, G. Flandrin, H. Muller-Hermelink, J. Vardiman, T. Lister, C. Bloomfield, The world Health organization classification of neoplastic diseases of the hematopoietic and lymphoid Tissues, Ann. Oncol. 10 (1999) 1419–1432.

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

[22] H. Inaba, M. Greaves, C.G. Mullighan, Acute lymphoblastic leukaemia, The Lancet 381 (2013) 1943–1955.

[23] J.F. Peterson, L.B. Baughn, K.E. Pearce, C.M. Williamson, J.C. Benevides Demasi, R.M. Olson, T.A. Goble, R.G. Meyer, P.T. Greipp, R.P. Ketterling, KMT2A (MLL) rearrangements observed in pediatric/young adult T-lymphoblastic leukemia/lymphoma: a 10-year review from a single cytogenetic laboratory, Genes Chromosomes Cancer 57 (2018) 541–546.

[24] T.B. Alexander, Z. Gu, I. Iacobucci, K. Dickerson, J.K. Choi, B. Xu, D. Payne-Turner, H. Yoshihara, M.L. Loh, J. Horan, B. Buldini, G. Basso, S. Elitzur, V. de Haas, C.M. Zwaan, A. Yeoh, D. Reinhardt, D. Tomizawa, N. Kiyokawa, T. Lammens, B. De Moerloose, D. Catchpoole, H. Hori, A. Moorman, A.S. Moore, O. Hrusak, S. Meshinchi, E. Orgel, M. Devidas, M. Borowitz, B. Wood,

N.A. Heerema, A. Carrol, Y.L. Yang, M.A. Smith, T.M. Davidsen, L.C. Hermida, P. Gesuwan, M.A. Marra, Y. Ma, A.J. Mungall, R.A. Moore, S.J.M. Jones, M. Valentine, L.J. Janke, J.E. Rubnitz, C.H. Pui, L. Ding, Y. Liu, J. Zhang, K.E. Nichols, J.R. Downing, X. Cao, L. Shi, S. Pounds, S. Newman, D. Pei, J.M. Guidry Auvil, D.S. Gerhard, S.P. Hunger, H. Inaba, C.G. Mullighan, The genetic basis and cell of origin of mixed phenotype acute leukaemia, Nature 562 (2018) 373–379.

[25] K. Takahashi, F. Wang, K. Morita, Y. Yan, P. Hu, P. Zhao, A.A. Zhar, C.J. Wu, C. Gumbs, L. Little, S. Tippen, R. Thornton, M. Coyle, M. Mendoza, E. Thompson, J. Zhang, C.D. DiNardo, N. Jain, F. Ravandi, J.E. Cortes, G. Garcia-Manero, S. Kornblau, M. Andreeff, E. Jabbour, C. Bueso-Ramos, A. Takaori-Kondo, M. Konopleva, K. Patel, H. Kantarjian, P.A. Futreal, Integrative genomic analysis of adult mixed phenotype acute leukemia delineates lineage associated molecular subtypes, Nat. Commun. 9 (2018) 2670.

[26] N.J. Charles, D.F. Boyer, Mixed-phenotype Acute leukemia: diagnostic criteria and pitfalls, Arch. Pathol. Lab Med. 141 (2017) 1462–1468.

[27] K. Matlawska-Wasowska, H. Kang, M. Devidas, J. Wen, R.C. Harvey, C.K. Nickl, S.A. Ness, M. Rusch, Y. Li, M. Onozawa, C. Martinez, B.L. Wood, B.L. Asselin, I.M. Chen, K.G. Roberts, A. Baruchel, J. Soulier, H. Dombret, J. Zhang, R.S. Larson, E.A. Raetz, W.L. Carroll, K.P. Dunsmore, P.D. Aplan, M.L. Loh, C.G. Mullighan, S.P. Hunger, N.A. Heerema, A.J. Carroll, K.P. Dunsmore, S.S. Winter, MLL rearrangements impact outcome in HOXA-deregulated T-lineage acute lymphoblastic leukemia: a Children's Oncology Group Study, Leukemia 30 (2016) 1909–1912.

[28] Y. Liu, J. Easton, Y. Shao, J. Maciaszek, Z. Wang, M.R. Wilkinson, K. McCastlain, M. Edmonson, S.B. Pounds, L. Shi, X. Zhou, X. Ma, E. Sioson, Y. Li, M. Rusch, P. Gupta, D. Pei, C. Cheng, M.A. Smith, J.G. Auvil, D.S. Gerhard, M.V. Relling, N.J. Winick, A.J. Carroll, N.A. Heerema, E. Raetz, M. Devidas, C.L. Willman, R.C. Harvey, W.L. Carroll, K.P. Dunsmore, S.S. Winter, B.L. Wood, B.P. Sorrentino, J.R. Downing, M.L. Loh, S.P. Hunger, J. Zhang, C.G. Mullighan, The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia, Nat. Genet. 49 (2017) 1211–1218.

[29] L. Belver, A. Ferrando, The genetics and mechanisms of T cell acute lymphoblastic leukaemia, Nat. Rev. Cancer 16 (2016) 494–507.

[30] M.B. Mansur, F.W. van Delft, S.M. Colman, C.L. Furness, J. Gibson, M. Emerenciano, H. Kempski, E. Clappier, H. Cave, J. Soulier, M.S. Pombo-de-Oliveira, M. Greaves, A.M. Ford, Distinctive genotypes in infants with T-cell acute lymphoblastic leukaemia, Br. J. Haematol. 171 (2015) 574–584.

[31] M.G. Guenther, L.N. Lawton, T. Rozovskaia, G.M. Frampton, S.S. Levine, T.L. Volkert, C.M. Croce, T. Nakamura, E. Canaani, R.A. Young, Aberrant chromatin at genes encoding stem cell regulators in human mixed-lineage leukemia, Genes Dev. 22 (2008) 3403–3408.

[32] I. Chijimatsu, Y. Imanaka, D. Tomizawa, M. Eguchi, S. Nishimura, S. Karakawa, M. Miki, K. Hamamoto, N. Fujita, Azacitidine successfully maintained the second remission in an infant with KMT2A-rearranged acute lymphoblastic leukemia who relapsed after unrelated cord blood transplantation, Pediatr. Blood Cancer 64 (2017) 1–4.

[33] S.K. Tasian, S.P. Hunger, Genomic characterization of paediatric acute lymphoblastic leukaemia: an opportunity for precision medicine therapeutics, Br. J. Haematol. 176 (2017) 867–882.

[34] B.V. Balgobind, S.C. Raimondi, J. Harbott, M. Zimmermann, T.A. Alonzo, A. Auvrignon, H.B. Beverloo, M. Chang, U. Creutzig, M.N. Dworzak, E. Forestier, B. Gibson, H. Hasle, C.J. Harrison, N.A. Heerema, G.J. Kaspers, A. Leszl, N. Litvinko, L.L. Nigro, A. Morimoto, C. Perot, R. Pieters, D. Reinhardt, J.E. Rubnitz, F.O. Smith, J. Stary, I. Stasevich, S. Strehl, T. Taga, D. Tomizawa, D. Webb, Z. Zemanova, C.M. Zwaan, M.M. van den Heuvel-Eibrink, Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study, Blood 114 (2009) 2489–2496.

[35] M. Yamada, J. Clark, A. Iulianella, MLLT11/AF1q is differentially expressed in maturing neurons during development, Gene Expr. Patterns 15 (2014) 80–87.

[36] E. Smith, C. Lin, A. Shilatifard, The super elongation complex (SEC) and MLL in development and disease, Genes Dev. 25 (2011) 661–672.

[37] C. Meyer, J. Hofmann, T. Burmeister, D. Groger, T.S. Park, M. Emerenciano, M. Pombo de Oliveira, A. Renneville, P. Villarese, E. Macintyre, H. Cave, E. Clappier, K. Mass-Malo, J. Zuna, J. Trka, E. De Braekeleer, M. De Braekeleer, S.H. Oh, G. Tsaur, L. Fechina, V.H. van der Velden, J.J. van Dongen, E. Delabesse, R. Binato, M.L. Silva, A. Kustanovich, O. Aleinikova, M.H. Harris, T. Lund-Aho, V. Juvonen, O. Heidenreich, J. Vormoor, W.W. Choi, M. Jarosova, A. Kolenova, C. Bueno, M. Menendez, S. Wehner, C. Eckert, P. Talmant, S. Tondeur, E. Lippert, E. Launay, C. Henry, P. Ballerini, H. Lapillone, M.B. Callanan, J.M. Cayuela, C. Herbaux, G. Cazzaniga, P.M. Kakadiya, S. Bohlander, A. Ahlmann, J.R. Choi, P. Gameiro, D.S. Lee, J. Krauter, P. Cornillet-Lefebvre, G. Te Kronnie, B.W. Schafer, S. Kubetzko, C.N. Alonso, U. zur Stadt, R. Sutton, N.C. Venn, S. Izraeli, L. Trakhtenbrot, H.O. Madsen, P. Archer, J. Hancock, N. Cerveira, M.R. Teixeira, L. Lo Nigro, A. Moricke, M. Stanulla, M. Schrappe, S. Sedek, T. Szczepanski, C.M. Zwaan, E.A. Coenen, M.M. van den Heuvel-Eibrink, S. Strehl, M. Dworzak, R. Panzer-Grumayer, T. Dingermann, T. Klingebiel, R. Marschalek, The MLL recombinome of acute leukemias in 2013, Leukemia 27 (2013) 2165–2176.

[38] J. de Boer, V. Walf-Vorderwülbecke, O. Williams, In focus: MLL-rearranged leukemia, Leukemia 27 (2013) 1224–1228.

[39] S. Lin, R.T. Luo, M. Shrestha, M.J. Thirman, J.C. Mulloy, The full transforming capacity of MLL-Af4 is interlinked with lymphoid lineage commitment, Blood 130 (2017) 903–907.

[40] C. Meyer, T. Burmeister, D. Groger, G. Tsaur, L. Fechina, A. Renneville, R. Sutton, N.C. Venn, M. Emerenciano, M.S. Pombo-de-Oliveira, C. Barbieri Blunck,

B. Almeida Lopes, J. Zuna, J. Trka, P. Ballerini, H. Lapillonne, M. De Braekeleer, G. Cazzaniga, L. Corral Abascal, V.H.J. van der Velden, E. Delabesse, T.S. Park, S.H. Oh, M.L.M. Silva, T. Lund-Aho, V. Juvonen, A.S. Moore, O. Heidenreich, J. Vormoor, E. Zerkalenkova, Y. Olshanskaya, C. Bueno, P. Menendez, A. Teigler-Schlegel, U. Zur Stadt, J. Lentes, G. Gohring, A. Kustanovich, O. Aleinikova, B.W. Schafer, S. Kubetzko, H.O. Madsen, B. Gruhn, X. Duarte, P. Gameiro, E. Lippert, A. Bidet, J.M. Cayuela, E. Clappier, C.N. Alonso, C.M. Zwaan, M.M. van den Heuvel-Eibrink, S. Izraeli, L. Trakhtenbrot, P. Archer, J. Hancock, A. Moricke, J. Alten, M. Schrappe, M. Stanulla, S. Strehl, A. Attarbaschi, M. Dworzak, O.A. Haas, R. Panzer-Grumayer, L. Sedek, T. Szczepanski, A. Caye, L. Suarez, H. Cave, R. Marschalek, The MLL recombinome of acute leukemias in 2017, Leukemia 32 (2018) 273–284.

[41] E.M. Stein, G. Garcia-Manero, D.A. Rizzieri, R. Tibes, J.G. Berdeja, M.R. Savona, M. Jongen-Lavrenic, J.K. Altman, B. Thomson, S.J. Blakemore, S.R. Daigle, N.J. Waters, A.B. Suttle, A. Clawson, R. Pollock, A. Krivtsov, S.A. Armstrong, J. DiMartino, E. Hedrick, B. Löwenberg, M.S. Tallman, The DOT1L inhibitor pinometostat reduces H3K79 methylation and has modest clinical activity in adult acute leukemia, Blood 131 (2018) 2661–2669.

[42] H. Kang, N.D. Sharma, C.K. Nickl, M. Devidas, M.L. Loh, S.P. Hunger, K.P. Dunsmore, S.S. Winter, K. Matlawska-Wasowska, Dysregulated transcriptional networks in KMT2A- and MLLT10-rearranged T-ALL, Biomark. Res. 6 (2018) 27.

[43] R. Kostadinov, M. Farooqi, B. Yoo, E. Farrow, S. Kelley, M. Gibson, N. Miller, E.M. Guest, P. Brown, Landscape of somatic mutations and gene expression changes in relapsed infant MLL-rearranged acute lymphoblastic leukemia, Blood 128 (2016) 1735-1735.

[44] D. Mueller, M.P. García-Cuéllar, C. Bach, S. Buhl, E. Maethner, R.K. Slany, Misguided transcriptional elongation causes mixed lineage leukemia, PLoS Biol. 7 (2009) e1000249.

[45] R.C. Rao, Y. Dou, Hijacked in cancer: the KMT2 (MLL) family of methyltransferases, Nat. Rev. Cancer 15 (2015) 334–346.

[46] D. Tomizawa, Recent progress in the treatment of infant acute lymphoblastic leukemia, Pediatr. Int. 57 (2015) 811–819.

[47] R. Sutton, T. Law, N. Venn, C. Wadham, E. Mould, M. Norris, R. Kotecha, R. Marschalek, C. Meyer, L. Dalla-Pozza, G. Marshall, L. RB, T. Trahair, Comparison of MRD levels and gene expression patterns in MLL-R versus non-MLL infant ALL, Blood 128 (2016) 1740.

[48] R. Pieters, P. De Lorenzo, P. Ancliffe, L. Aversa, B. Brethon, A. Biondi, M. Campbell, G. Escherich, A. Ferster, R. Gardner, R. Kotecha, B. Lausen, C. Li, F. Locatelli, A. Attarbaschi, C. Peters, J. Rubnitz, L. Silverman, J. Stary, T. Szczepanski, A. Vora, M. Schrappe, M. Valsecchi, Outcome of infants younger than 1 Year with acute lymphoblastic leukemia treated with the interfant-06 protocol: results from an international phase III randomized study, J. Clin. Oncol. 37 (2019) 2246–2256.

[49] Z.E. Dreyer, J.M. Hilden, T.L. Jones, M. Devidas, N.J. Winick, C.L. Willman, R.C. Harvey, I.M. Chen, F.G. Behm, J. Pullen, B.L. Wood, A.J. Carroll, N.A. Heerema, C.A. Felix, B. Robinson, G.H. Reaman, W.L. Salzer, S.P. Hunger, W.L. Carroll, B.M. Camitta, Intensified chemotherapy without SCT in infant ALL: results from COG P9407 (Cohort 3), Pediatr. Blood Cancer 62 (2015) 419–426.

[50] R. Pieters, M. Schrappe, P. De Lorenzo, I. Hann, G. De Rossi, M. Felice, L. Hovi, T. LeBlanc, T. Szczepanski, A. Ferster, G. Janka, J. Rubnitz, L. Silverman, J. Stary, M. Campbell, C.-K. Li, G. Mann, R. Suppiah, A. Biondi, A. Vora, M.G. Valsecchi, A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial, The Lancet 370 (2007) 240–250.

[51] J.M. Hilden, P.A. Dinndorf, S.O. Meerbaum, H. Sather, D. Villaluna, N.A. Heerema, R. McGlennen, F.O. Smith, W.G. Woods, W.L. Salzer, H.S. Johnstone, Z. Dreyer, G.H. Reaman, Children's Oncology, Analysis of prognostic factors of acute lymphoblastic leukemia in infants: report on CCG 1953 from the Children's Oncology Group, Blood 108 (2006) 441–451.

[52] C.H. Pui, D. Pei, D. Campana, C. Cheng, J.T. Sandlund, W.P. Bowman, M.M. Hudson, R.C. Ribeiro, S.C. Raimondi, S. Jeha, S.C. Howard, D. Bhojwani, H. Inaba, J.E. Rubnitz, M.L. Metzger, T.A. Gruber, E. Coustan-Smith, J.R. Downing, W.H. Leung, M.V. Relling, W.E. Evans, A revised definition for cure of childhood acute lymphoblastic leukemia, Leukemia 28 (2014) 2336–2343.

[53] R. Bassan, D. Hoelzer, Modern therapy of acute lymphoblastic leukemia, J. Clin. Oncol. 29 (2011) 532–543.

[54] K. Fasching, S. Panzer, O.A. Haas, R. Marschalek, H. Gadner, E.R. Panzer-Grümayer, Presence of clone-specific antigen receptor gene rearrangements at birth indicates an in utero origin of diverse types of early childhood acute lymphoblastic leukemia, Blood 95 (2000) 2722–2724.

[55] M.F. Greaves, A.T. Maia, J.L. Wiemels, A.M. Ford, Leukemia in twins: lessons in natural history, Blood 102 (2003) 2321–2333.

[56] M.F. Greaves, J. Wiemels, Origins of chromosome translocations in childhood leukaemia, Nat. Rev. Cancer 3 (2003) 639–649.

[57] Y. Ma, S.E. Dobbins, A.L. Sherborne, D. Chubb, M. Galbiati, G. Cazzaniga, C. Micalizzi, R. Tearle, A.L. Lloyd, R. Hain, M. Greaves, R.S. Houlston, Developmental timing of mutations revealed by whole-genome sequencing of twins with acute lymphoblastic leukemia, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 7429–7433.

[58] M. Greaves, A causal mechanism for childhood acute lymphoblastic leukaemia, Nat. Rev. Cancer (2018) 471–484.

[59] I. Roberts, N.J. Fordham, A. Rao, B.J. Bain, Neonatal leukaemia, Br. J. Haematol. 182 (2018) 170–184.

[60] T.O. Owuor, S.N. Porter, A. Cluster, J. Ryoo, R. Patel, J.A. Magee, Fetal and neonatal hematopoietic progenitors are exquisitely sensitive to MLL-ENL-driven

leukemia initiation, Blood 130 (2017) 2473.

[61] C. Bueno, R. Montes, P. Catalina, R. Rodríguez, P. Menendez, Insights into the cellular origin and etiology of the infant pro-B acute lymphoblastic leukemia with MLL-AF4 rearrangement, Leukemia 25 (2010) 400–410.

[62] C. Malouf, K. Ottersbach, The fetal liver lymphoid-primed multipotent progenitor provides the prerequisites for the initiation of t(4;11) MLL-AF4 infant leukemia, Haematologica 103 (2018) e571–e574.

[63] A. Agraz-Doblas, C. Bueno, R. Bashford-Rogers, A. Roy, P. Schneider, M. Bardini, P. Ballerini, G. Cazzaniga, T. Moreno, C. Revilla, M. Gut, M.G. Valsecchi, I. Roberts, R. Pieters, P. De Lorenzo, I. Varela, P. Menendez, R.W. Stam, Unraveling the cellular origin and clinical prognostic markers of infant B-cell acute lymphoblastic leukemia using genome-wide analysis, Haematologica 104 (2019) 1176–1188.

[64] R. Montes, V. Ayllón, I. Gutierrez-Aranda, I. Prat, M. Hernández-Lamas, L. Ponce, S. Bresolin, G.t. Kronnie, M. Greaves, C. Bueno, P. Menendez, Enforced expression of MLL-AF4 fusion in cord blood CD34 + cells enhances the hematopoietic re-populating cell function and clonogenic potential but is not sufficient to initiate leukemia, Blood 117 (2011) 4746–4758.

[65] P. Menendez, P. Catalina, R. Rodriguez, G.J. Melen, C. Bueno, M. Arriero, F. Garcia-Sanchez, A. Lassaletta, R. Garcia-Sanz, J. Garcia-Castro, Bone marrow mesenchymal stem cells from infants with MLL-AF4 + acute leukemia harbor and express the MLL-AF4 fusion gene, J. Exp. Med. 206 (2009) 3131–3141.

[66] S. Shalapour, C. Eckert, K. Seeger, M. Pfau, J. Prada, G. Henze, T. Blankenstein, T. Kammertoens, Leukemia-associated genetic aberrations in mesenchymal stem cells of children with acute lymphoblastic leukemia, J. Mol. Med. (Berl.) 88 (2010) 249–265.

[67] E.N. Wainwright, P. Scaffidi, Epigenetics and cancer stem cells: unleashing, hi-jacking, and restricting cellular plasticity, Trends Cancer 3 (2017) 372–386.

[68] L.I. Shlush, S. Zandi, A. Mitchell, W.C. Chen, J.M. Brandwein, V. Gupta, J.A. Kennedy, A.D. Schimmer, A.C. Schuh, K.W. Yee, J.L. McLeod, M. Doedens, J.J. Medeiros, R. Marke, H.J. Kim, K. Lee, J.D. McPherson, T.J. Hudson, H.P.L.G.P.C., A.M. Brown, F. Yousif, Q.M. Trinh, L.D. Stein, M.D. Minden, J.C. Wang, J.E. Dick, Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia, Nature 506 (2014) 328–333.

[69] H. Sakaki, K. Kanegane, K. Nomura, K. Goi, K. Sugita, M. Miura, E. Ishii, T. Miyawaki, Early lineage switch in an infant acute lymphoblastic leukemia, Int. J. Hematol. 90 (2009) 653–655.

[70] C.L. Haddox, A.A. Mangaonkar, D. Chen, M. Shi, R. He, J.L. Oliveira, M.R. Litzow, A. Al-Kali, W.J. Hogan, M.A. Elliott, Blinatumomab-induced lineage switch of B-ALL with t(4;11)(q21;q23) KMT2A/AFF1 into an aggressive AML: pre- and post-switch phenotypic, cytogenetic and molecular analysis, Blood Canc. J. 7 (2017) e607.

[71] E. Jacoby, S.M. Nguyen, T.J. Fountaine, K. Welp, B. Gryder, H. Qin, Y. Yang, C.D. Chien, A.E. Seif, H. Lei, Y.K. Song, J. Khan, D.W. Lee, C.L. Mackall, R.A. Gardner, M.C. Jensen, J.F. Shern, T.J. Fry, CD19 CAR immune pressure in-duces B-precursor acute lymphoblastic leukaemia lineage switch exposing in-herent leukaemic plasticity, Nat. Commun. 7 (2016) 12320.

[72] J.G. Rossi, A.R. Bernasconi, C.N. Alonso, P.L. Rubio, M.S. Gallego, C.A. Carrara, M.R. Guitter, S.E. Eberle, M. Cocce, P.A. Zubizarreta, M.S. Felice, Lineage switch in childhood acute leukemia: an unusual event with poor outcome, Am. J. Hematol. 87 (2012) 890–897.

[73] U. Duffner, A. Abdel-Mageed, J. Younge, C. Tornga, K. Scott, J. Staddon, K. Elliott, J. Stumph, P. Kidd, The possible perils of targeted therapy, Leukemia 30 (2016) 1619–1621.

[74] M. Wolfl, M. Rasche, M. Eyrich, R. Schmid, D. Reinhardt, P.G. Schlegel, Spontaneous reversion of a lineage switch following an initial blinatumomab-in-duced ALL-to-AML switch in MLL-rearranged infant ALL, Blood Adv. 2 (2018) 1382–1385.

[75] A. Ittel, E. Jeandidier, C. Helias, N. Perrusson, C. Humbrecht, B. Lioure, I. Mazurier, C. Mayeur-Rousse, A. Lavaux, S. Thiebault, F. Lerintiu, C. Gervais, L. Mauvieux, First description of the t(10;11)(q22;q23)/MLL-TET1 translocation in a T-cell lymphoblastic lymphoma, with subsequent lineage switch to acute myelomonocytic myeloid leukemia, Haematologica 98 (2013) e166–e168.

[76] M. Greaves, Infection, immune responses and the aetiology of childhood leu-kaemia, Nat. Rev. Cancer 6 (2006) 193–203.

[77] M. Emerenciano, C. Meyer, M.B. Mansur, R. Marschalek, M.S. Pombo-de-Oliveira, L, Brazilian Collaborative Study Group of Infant Acute, the distribution of MLL breakpoints correlates with outcome in infant acute leukemia, Br. J. Haematol. 161 (2013) 224–236.

[78] T.N. Sam, J.H. Kersey, A.M. Linabery, K.J. Johnson, N.A. Heerema, J.M. Hilden, S.M. Davies, G.H. Reaman, J.A. Ross, MLL gene rearrangements in infant leukemia vary with age at diagnosis and selected demographic factors: a Children's Oncology Group (COG) study, Pediatr. Blood Cancer 58 (2012) 836–839.

[79] M. Burns, S.A. Armstrong, A. Gutierrez, Pathobiology of Acute Lymphoblastic Leukemia, Hematology 2018, pp. 1005-1019.e1011.

[80] C. Prelle, A. Bursen, T. Dingermann, R. Marschalek, Secondary mutations in t (4;11) leukemia patients, Leukemia 27 (2013) 1425–1427.

[81] D.C. Liang, S.H. Chen, H.C. Liu, C.P. Yang, T.C. Yeh, T.H. Jaing, I.J. Hung, J.Y. Hou, T.H. Lin, C.H. Lin, L.Y. Shih, Mutational status of NRAS, KRAS, and PTPN11 genes is associated with genetic/cytogenetic features in children with B-precursor acute lymphoblastic leukemia, Pediatr. Blood Cancer 65 (2018) e26786.

[82] H. Fedders, A. Alsadeq, J. Schmäh, F. Vogiatzi, M. Zimmermann, A. Möricke, L. Lenk, U.z. Stadt, M.A. Horstmann, R. Pieters, M. Schrappe, M. Stanulla, G. Cario,

[83] L. Trentin, S. Bresolin, E. Giarin, M. Bardini, V. Serafin, B. Accordi, F. Fais, C. Tenca, P. De Lorenzo, M.G. Valsecchi, G. Cazzaniga, G.T. Kronnie, G. Basso, Deciphering KRAS and NRAS mutated clone dynamics in MLL-AF4 paediatric leukaemia by ultra deep sequencing analysis, Sci. Rep. 6 (2016) 34449.

[84] M. Emerenciano, T.d.C. Barbosa, B. Almeida Lopes, C. Meyer, R. Marschalek, M.S. Pombo-de-Oliveira, Subclonality and prenatal origin of RAS mutations in KMT2A (MLL)-rearranged infant acute lymphoblastic leukaemia, Br. J. Haematol. 170 (2015) 268–271.

[85] E.M. Driessen, E.H. van Roon, J.A. Spijkers-Hagelstein, P. Schneider, P. de Lorenzo, M.G. Valsecchi, R. Pieters, R.W. Stam, Frequencies and prognostic impact of RAS mutations in MLL-rearranged acute lymphoblastic leukemia in infants, Haematologica 98 (2013) 937–944.

[86] R. Huether, L. Dong, X. Chen, G. Wu, M. Parker, L. Wei, J. Ma, M.N. Edmonson, E.K. Hedlund, M.C. Rusch, S.A. Shurtleff, H.L. Mulder, K. Boggs, B. Vadordaria, J. Cheng, D. Yergeau, G. Song, J. Becksfort, G. Lemmon, C. Weber, Z. Cai, J. Dang, M. Walsh, A.L. Gedman, Z. Faber, J. Easton, T. Gruber, R.W. Kriwacki, J.F. Partridge, L. Ding, R.K. Wilson, E.R. Mardis, C.G. Mullighan, R.J. Gilbertson, S.J. Baker, G. Zambetti, D.W. Ellison, J. Zhang, J.R. Downing, The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer gen-omes, Nat. Commun. 5 (2014) 3630.

[87] K.M. Bernt, N. Zhu, A.U. Sinha, S. Vempati, J. Faber, A.V. Krivtsov, Z. Feng, N. Punt, A. Daigle, L. Bullinger, R.M. Pollock, V.M. Richon, A.L. Kung, S.A. Armstrong, MLL-rearranged leukemia is dependent on aberrant H3K79 me-thylation by DOT1L, Cancer Cell 20 (2011) 66–78.

[88] A.K. Bergmann, G. Castellano, J. Alten, O. Ammerpohl, J. Kolarova, J. Nordlund, J.I. Martin-Subero, M. Schrappe, R. Siebert, DNA methylation profiling of pediatric B-cell lymphoblastic leukemia with KMT2A rearrangement identifies hypo-methylation at enhancer sites, Pediatr. Blood Cancer 64 (2017) e26251.

[89] P. Brown, R. Pieters, A. Biondi, How I treat infant leukemia, Blood 133 (2019) 205–214.

[90] E. Sison, P. Brown, Does hematopoietic stem cell transplantation benefit infants with acute leukemia? Hematology 1 (2013) 601–604.

[91] K. Schmiegelow, K. Müller, S.S. Mogensen, P.R. Mogensen, B.O. Wolthers, U.K. Stoltze, R. Tuckuviene, T. Frandsen, Non-infectious chemotherapy-associated acute toxicities during childhood acute lymphoblastic leukemia therapy, F1000Res. 6 (2017) 444.

[92] M. Kerstjens, E. Driessen, M. Willekes, S. Pinhanços, P. Schneider, R. Pieters, R. Stam, MEK inhibition is a promising therapeutic strategy for MLL-rearranged infant acute lymphoblastic leukemia patients carrying RAS mutations, Oncotarget 8 (2017) 14835–14846.

[93] M. Kerstjens, S.S. Pinhancos, P.G. Castro, P. Schneider, P. Wander, R. Pieters, R.W. Stam, Trametinib inhibits RAS mutant MLL-rearranged acute lymphoblastic leukemia at specific niche sites and reduces ERK phosphorylation in vivo, Haematologica 103 (2018) e147–e150.

[94] M.C. Chillón, M.T. Gómez-Casares, C.E. López-Jorge, C. Rodriguez-Medina, A. Molines, M.E. Sarasquete, M. Alcoceba, J.D. Miguel, C. Bueno, R. Montes, F. Ramos, J.N. Rodríguez, P. Giraldo, M. Ramírez, R. García-Delgado, J.L. Fuster, M. González-Díaz, P. Menendez, Prognostic significance of FLT3 mutational status and expression levels in MLL-AF4 + and MLL-germline acute lymphoblastic leu-kemia, Leukemia 26 (2012) 2360–2366.

[95] C.E. Annesley, P. Brown, The biology and targeting of FLT3 in pediatric leukemia, Front. Oncol. 4 (2014) 263.

[96] T.M. Cooper, J. Cassar, E. Eckroth, J. Malvar, R. Sposto, P. Gaynon, B.H. Chang, L. Gore, K. August, J.A. Pollard, S.G. DuBois, L.B. Silverman, J. Oesterheld, G. Gammon, D. Magoon, C. Annesley, P.A. Brown, A phase I study of quizartinib combined with chemotherapy in relapsed childhood leukemia: a therapeutic ad-vances in childhood leukemia & lymphoma (TACL) study, Clin. Cancer Res. 22 (2016) 4014–4022.

[97] S. Wander, M. Levis, A. Fathi, The evolving role of FLT3 inhibitors in acute myeloid leukemia: quizartinib and beyond, Ther. Adv. Hematol. 5 (2014) 65–77.

[98] P. Garrido Castro, E.H.J. van Roon, S.S. Pinhancos, L. Trentin, P. Schneider, M. Kerstjens, G. Te Kronnie, O. Heidenreich, R. Pieters, R.W. Stam, The HDAC inhibitor panobinostat (LBH589) exerts in vivo anti-leukaemic activity against MLL-rearranged acute lymphoblastic leukaemia and involves the RNF20/RNF40/WAC-H2B ubiquitination axis, Leukemia 32 (2018) 323–331.

[99] H. Dombret, C. Preudhomme, C. Berthon, E. Raffoux, X. Thomas, N. Vey, C. Gomez-Roca, M. Ethell, K. Yee, F. Bourdel, P. Herait, M. Michallet, C. Recher, C. Roumier, B. Quesnel, A phase 1 study of the BET-bromodomain inhibitor OTX015 in patients with advanced acute leukemia, Blood 124 (2014) 117.

[100] D.J. Stumpel, P. Schneider, E.H. van Roon, R. Pieters, R.W. Stam, Absence of global hypomethylation in promoter hypermethylated Mixed Lineage Leukaemia-rearranged infant acute lymphoblastic leukaemia, Eur. J. Cancer 49 (2013) 175–184.

[101] N. Shukla, C. Wetmore, M.M. O'Brien, L.B. Silverman, P. Brown, T.M. Cooper, B. Thomson, S.J. Blakemore, S. Daigle, B. Suttle, N.J. Waters, A.V. Krivstov, S.A. Armstrong, P.T. Ho, L. Gore, Final report of phase 1 study of the DOT1L inhibitor, pinometostat (EPZ-5676), Children with Relapsed or Refractory MLL-R Acute Leukemia, Blood, 128 2016 2780-2780.

[102] M. Van Veggel, E. Westerman, P. Hamberg, Clinical pharmacokinetics and phar-macodynamics of panobinostat, Clin. Pharmacokinet. 57 (2018) 21–29.

17

1.4 MLLT10 gene rearrangements in leukaemia

Rearrangements of the *MLLT10* gene are poorly characterised, despite accounting for approximately 10% of new T-ALL diagnoses, in addition to rare occurrences in AML and mixed phenotype acute leukaemia (MPAL). Here, the existing base of published evidence to characterise the molecular mechanisms and clinical manifestations of *MLLT10*r in leukaemia are reviewed, to identify gaps for future studies.

Statement of Authorship

Title of Paper	*MLLT10* rearranged acute leukemia: Incidence, prognosis, and possible therapeutic strategies
Publication Status	[X] Published [ ] Accepted for Publication [ ] Submitted for Publication [ ] Unpublished and Unsubmitted work written in manuscript style
Publication Details	Michelle O. Forgione, Barbara J. McClure, David T. Yeung, Laura N. Eadie, Deborah L. White Published in *Genes, Chromosomes & Cancer*, 2020, Vol. 59, pages 709-721.

Principal Author

Name of Principal Author (Candidate)	Michelle Forgione
Contribution to the Paper	Conceptualised idea, conducted research, constructed manuscript.
Overall percentage (%)	80%
Certification:	This paper is a literature review that I conducted during the period of my Higher Degree by Research candidature, and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 3/1/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

    i.   the candidate's stated contribution to the publication is accurate (as detailed above);

    ii.   permission is granted for the candidate in include the publication in the thesis; and

    iii.   the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Barbara J. McClure		
Contribution to the Paper	Critical review of manuscript.		
Signature		Date	4/1/2022

Name of Co-Author	David T. Yeung		
Contribution to the Paper	Critical review of manuscript.		
Signature		Date	4/1/2022
Name of Co-Author	Laura N. Eadie		
Contribution to the Paper	Critical review of manuscript.		
Signature		Date	4/1/2022

Name of Co-Author	Deborah L. White		
Contribution to the Paper	Critical review of manuscript.		
Signature		Date	4/1/2022

REVIEW ARTICLE

WILEY

MLLT10 rearranged acute leukemia: Incidence, prognosis, and possible therapeutic strategies

Michelle O. Forgione[1,2] | Barbara J. McClure[1,3] | David T. Yeung[1,3,6] | Laura N. Eadie[1,3] | Deborah L. White[1,2,3,4,5]

[1]Cancer Program, Precision Medicine Theme, South Australian Health and Medical Research Institute (SAHMRI), Adelaide, South Australia, Australia

[2]Faculty of Science, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia

[3]Faculty of Health and Medical Science, University of Adelaide, Adelaide, South Australia, Australia

[4]Australian Genomics Health Alliance (AGHA), The Murdoch Children's Research Institute, Parkville, Victoria, Australia

[5]Australian and New Zealand Children's Oncology Group (ANZCHOG), Hudson Institute, Clayton, Victoria, Australia

[6]Department of Haematology, Royal Adelaide Hospital, Adelaide, South Australia, Australia

Correspondence
Prof Deborah L. White, South Australian Health and Medical Research Institute (SAHMRI) North Terrace, Adelaide 5001, PO Box 11060, South Australia.
Email: deborah.white@sahmri.com

Funding information
Cancer Council South Australia, Grant/Award Numbers: Beat Cancer Project, Peter Nelson Research Fellow; Leukaemia Foundation; National Health and Medical Research Council; University of Adelaide: PhD Scholarship

Abstract

Rearrangements of the *MLLT10* gene occur in acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL), most commonly T-lineage ALL (T-ALL), in patients of all ages. *MLLT10* rearranged (*MLLT10*r) acute leukemia presents a complex diagnostic and therapeutic challenge due to frequent presentation of immature or mixed phenotype, and a lack of consensus regarding optimal therapy. Cases of *MLLT10*r AML or T-ALL bearing immature phenotype are at high risk of poor outcome, but the underlying molecular mechanisms and sensitivity to targeted therapies remain poorly characterized. This review addresses the incidence and prognostic significance of *MLLT10*r in acute leukemia, and how the aberrant gene expression profile of this disease can inform potential targeted therapeutic strategies. Understanding the underlying genomics of *MLLT10*r acute leukemia, both clinically and molecularly, will improve prognostic stratification and accelerate the development of targeted therapeutic strategies, to improve patient outcomes.

KEYWORDS

AF10, leukemia genomics, *MLLT10*, MLLT10 rearrangements, t(10;11)(p12-13;q14-21)

1 | INTRODUCTION

The *MLLT10* (Mixed lineage leukemia or myeloid/lymphoid leukemia, translocated to, 10, formerly known as *AF10*) gene was first described

as a fusion partner of *KMT2A* (*Histone-lysine N-methyltransferase 2A*, also known as *MLL*, *MLL1*, or acute lymphoblastic leukemia [*ALL1*]) in acute myeloid leukemia (AML).[1] Chimeric fusion genes involving *MLLT10* arise from chromosomal rearrangements occurring at the *MLLT10* locus on chromosome 10p12, and are recurrent events in ALL[2] and AML[3], and also in isolated cases of acute leukemia of ambiguous lineage,[4] lymphoblastic lymphoma[5] and myeloid sarcoma.[6] While there is no published evidence for *MLLT10* ITDs, point mutations involving *MLLT10* are reported across a range of malignancies.[7] However, the functional significance of these variants in tumor pathogenesis remains unknown. Current data indicates that *MLLT10*

20

rearrangements (*MLLT10*r) occur in approximately 10% of all T-lineage ALL (T-ALL) cases, 6% to 8% of AML cases,[7] and 3% to 3.7% of B-ALL cases[8,9] (Table 1 and Figure 1). This review will discuss the incidence and prognostic significance of *MLLT10*r in acute leukemia, and how the aberrant gene expression profile of this disease can inform potential targeted therapeutic strategies. For the purposes of this manuscript, the term *MLLT10*r refers to all fusion genes involving *MLLT10*, including *KMT2A-MLLT10*.

2 | INCIDENCE OF *MLLT10*r T-ALL

T-ALL accounts for 10% to 15% of pediatric and 20% to 25% of adult new ALL diagnoses,[20] and *MLLT10*r account for approximately 8% to 10% of all T-ALL cases, most commonly *PICALM-MLLT10*/t(10;11)(p12-13;q14-21) (6%-8%) (Table 1 and Figure 1).[2] Despite historically poor outcomes compared with B-ALL, intensification of modern chemotherapy protocols has improved long-term overall survival of T-ALL to 80% in children and 60% in adults, comparable with B-ALL outcomes.[20,21] However, outcomes after relapse or treatment failure remain poor, demonstrating an ongoing need to improve therapeutic options for T-ALL patients.[20] Like B-ALL and T-ALL is characterized by a heterogenous array of genomic aberrations, including transcription factor oncogenes (eg, *TAL1* and *TAL2*), *NOTCH1* pathway members, cell cycle regulators (eg, *CDKN2A*) and transcription factor tumor suppressors (eg, *RUNX1* and *GATA3*),[20] but the prognostic relevance of most genomic aberrations in T-ALL remains unclear.[22] To date, the World Health Organization (WHO) has not classified any genomic subtypes of T-ALL. Indeed, the only WHO-recognized subtypes of T-ALL are the provisional entities of early T-cell precursor lymphoblastic leukemia and natural killer (NK) cell lymphoblastic leukemia/lymphoma, based on immunophenotype.[22] The genomic diagnosis and

prognosis of T-ALL remains in its infancy, and ongoing large prospective cohort studies are required for this purpose.[22]

PICALM-MLLT10 was overrepresented within the early thymic precursor (ETP)-ALL cohort of one study, where 56% (9/16) of *PICALM-MLLT10* cases were of ETP immunophenotype, compared with 19% (33/209) of non-*PICALM-MLLT10* T-ALL cases (*P* = .002).[2] ETP-ALL is a subtype of T-ALL that exhibits a stem cell/immature myeloid-like immunophenotype, defined by expression of CD5, lack of CD1a/CD8, and expression of at least one stem cell or myeloid antigen (CD34, CD13, CD33, and/or CD117),[23] suggesting that *MLLT10*r are capable of inducing disease characterized by an early differentiation block. However, T-cell receptor (TCR) positive (TCR[pos]) disease is observed in 62% to 67% of pediatric and 25% to 28% of adult cases of *PICALM-MLLT10* T-ALL[2,12] (Table 2), indicative of differentiation arrest at a less immature stage. This observation indicates that *MLLT10*r T-ALL can present at varying stages of differentiation arrest, ranging from ETP-ALL to a more mature TCR[pos] phenotype. It is not known how *MLLT10*r is capable of inducing both TCR negative (TCR[neg]) and TCR[pos] cases of T-ALL, but it is not associated with the specific *PICALM-MLLT10* breakpoint.[2]

3 | PROGNOSIS IN ALL

All *KMT2A* rearrangements (*KMT2A*r), including *KMT2A-MLLT10*/t(10;11)(p12;q23), confer adverse prognosis in B-ALL, characterized by high risk of minimal residual disease, shorter time to relapse, reduced event-free survival (EFS), and overall survival (OS).[8] In T-ALL, *MLLT10*r do not confer overall increased risk of central nervous system (CNS) disease,[2,14] mediastinal involvement,[2] induction failure,[24] or inferior long-term EFS[24] or OS[14]. However, *MLLT10*r T-ALL may present with ETP and/or TCR[neg] immunophenotype, which is a high-risk subgroup when observed in *PICALM-MLLT10* T-ALL.

TABLE 1 Reported frequencies of *MLLT10* rearrangements in large cohort studies of acute leukemia

Malignancy	Age group	Incidence of *MLLT10*r[a] d (n)	References
AML[b]	All ages	1.7d (5/289) *PICALM-MLLT10* only	[3][a]
AML	0–30 y	5.7d (127/2226, any fusion partner) 1d (23/2226, excluding *KMT2A-MLLT10*)	[10][b]
B-ALL[c]	All ages	3.7d (51/1361) *KMT2A-MLLT10* only	[8][c]
T-ALL[d]	>16 y	7d (30/431) *PICALM-MLLT10* only	[2][a]
	≤16 y	6d (15/234) *PICALM-MLLT10* only	
T-ALL	0–18 y	7.5d (14/197) *PICALM-MLLT10* only	[11][a]
T-ALL	Pediatric (age range not specified)	7d (7/102) Any fusion partner	[12][d]
T-ALL	<15 y	8d (4/49) *PICALM-MLLT10* only	[13][a]
	≥15 y	10d (8/82) *PICALM-MLLT10* only	

Note: Fusion identification method varied.
[a]Fusions identified by RT-PCR and/or Sanger sequencing.
[b]Fusions identified by FISH and/or RNA sequencing.
[c]Fusions identified by either cytogenetic analyses, split-signal FISH or RT-PCR, in combination with long-distance inverse PCR.
[d]Fusions identified by FISH only.
Abbreviations: AML, acute myeloid leukemia; B-ALL, B-lineage acute lymphoblastic leukemia; T-ALL, T-lineage acute lymphoblastic leukemia.
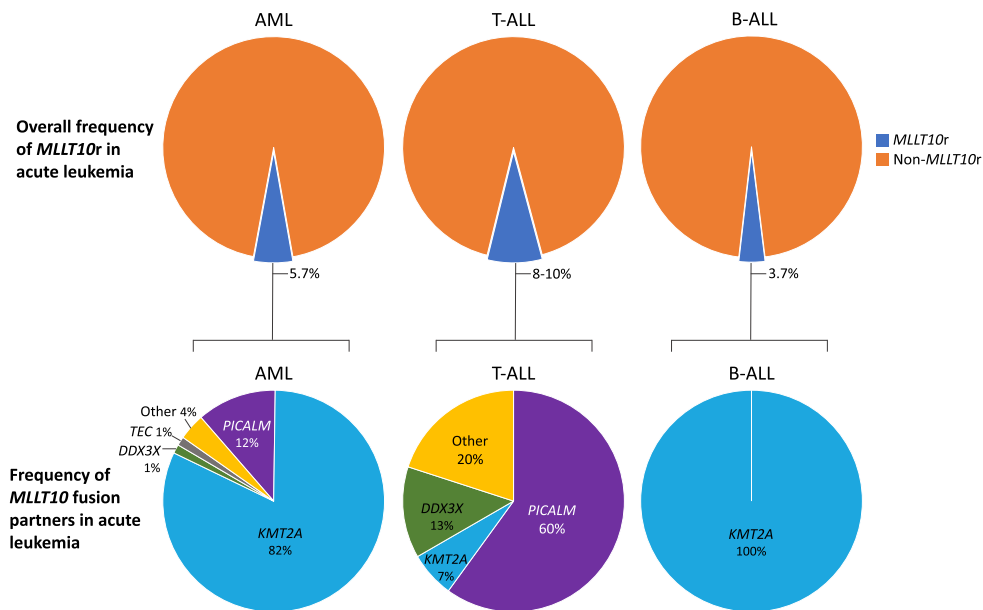
FIGURE 1 Frequency of *MLLT10r* and specific *MLLT10* fusion partners in ALL and AML *MLLT10r* accounts for 5.7% of AML[10], 8 to 10% of T-ALL[11,12] and 3.7% of B-ALL[8] new diagnoses overall (upper panel). The bottom panel depicts relative frequencies of different *MLLT10* fusion partners in AML, T-ALL, and B-ALL. *KMT2A* is the predominant *MLLT10* fusion partner identified in AML[10], whereas *PICALM-MLLT10*/t(10;11)(p12-13;q14-21) is the most common *MLLT10r* identified in T-ALL.[13] *KMT2A-MLLT10*/t(10;11)(p12;q23) is the only *MLLT10r* identified in B-ALL.[8] Rare fusion partners categorised as "other" are specified in Table 3. ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; T-ALL, T-lineage ALL [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Reported T-cell receptor status and survival of *PICALM-MLLT10* T-ALL

Treatment protocol (Cohort)	Frequency of [a]TCR[neg] phenotype % (n)		Overall survival at time of follow-up				
			Pediatric		Adult		
	Pediatric[a]	Adult	TCR[pos]	TCR[neg]	TCR[pos]	TCR[neg]	References
LALA-94/ GRAALL03/05 (adults), various protocols (pediatric)	33% (5/15)	72% (21/30)	80% (8/10)	40% (2/5)	57% (4/7)	0% (0/15)	[2][b]
			P value not provided		*P* = .0017		
LALA-94 (adult) / FRALLE93/00 (pediatric)	38% (3/8)	75% (9/12)	100% (5/5)	33% (1/3)	100% (3/3)	11% (8/9)	[13][c]
			P value not provided		*P* value not provided		

Note: Time of follow-up varied.
[a]For both studies, pediatric was defined as up to 15 years of age, and adults 16 years or over at diagnosis.
[b]Time of follow-up was 5 years for adults, and ranged from 6 to 115 months postinduction for pediatric.
[c]Time of follow-up ranged from 46 to 64 months postdiagnosis for adults, and 8 to 115 months postinduction for pediatric.
Abbreviation: TCR, T-cell Receptor.

These patients have poor long-term survival in comparison to TCR[pos] or mature immunophenotype *PICALM-MLLT10* disease (Table 2), and ETP phenotype, non-*PICALM-MLLT10* cases.[23]

Interestingly, the phenotype of *PICALM-MLLT10* T-ALL differs with age; adults with *PICALM-MLLT10* T-ALL are more likely to present with an immature TCR[neg] phenotype, whereas pediatric cases are typically of a mature TCRγδ status[2]. In adults, *PICALM-MLLT10* TCR[neg]

disease is characterized by poor 3 year EFS and OS when compared with both TCR[pos] *PICALM-MLLT10* and non-*PICALM-MLLT10* T-ALL.[2] Asnafi et al also identified *PICALM-MLLT10* TCR[neg] T-ALL as high risk for poor outcome in both adults and children, compared with TCR[pos] *PICALM-MLLT10* T-ALL.[12]

One study suggests that *HOXA* deregulation may underpin the poor outcomes observed in *MLLT10r* ETP-ALL, as a similar pattern is

observed for other genomic subtypes characterized by *HOXA* deregulation, such as *KMT2A*r. In this study, ETP immunophenotype predicted poor outcome exclusively in *HOXA* deregulated T-ALL, in a cohort of 190 T-ALL adults treated according to LALA-94 or GRAALL03/05 protocols.[23] Within the ETP cohort, patients presenting with high expression levels of *HOXA* genes (*HOXA*[high]) had worse outcomes compared with ETP-ALL cases that had normal expression levels of *HOXA* genes (*HOXA*[norm]) (OS 31.2% vs 74.2%; EFS 25% vs 60.8%). A TCR[neg]/ETP phenotype was also significantly more common in *HOXA*[high] than *HOXA*[norm] disease (TCR[neg] 57% vs 17%, and ETP phenotype 41% vs 15% respectively).[23] The *HOXA*[high] cohort included eight *PICALM-MLLT10* patients and four other *MLLT10* fusions, and 21 other cases harboring *TCRB-HOXA*, *SET-NUP214*, *KMT2A*r or *NUP98* rearrangement. This study indicates that high expression of *HOXA* may drive poor outcomes in ETP-ALL regardless of genomic subtype, rather than poor outcome being a characteristic specific to *MLLT10*r ETP-ALL, but further studies are required.

3.1 | Incidence of *MLLT10*r AML

AML is underpinned by genomic aberrations within a heterogeneous range of genes that can be used for prognostication of patients,[25] including signaling genes (eg, *FLT3* and *RAS* family members), DNA methylation-associated genes (eg, *DNMT3A* and *TET2*), myeloid transcription factors (eg, *RUNX1* and t[8;21]), and chromatin modifiers (eg, *ASXL1* and *KMT2A*r).[26,27] *MLLT10*r are rare in AML, predominantly *KMT2A-MLLT10* (5.7% of new AML diagnoses), and 1.5 to 2% of overall cases comprising fusion partners including *PICALM*, *DDX3X*, and *TEC*, and single cases of *DDX3Y*, *CEP164*, *NAP1L1*, *SCN2B*, *TREH*, and *XPO1* (Tables 1 and 3).[10]

3.2 | Prognosis of *MLLT10*r AML

While there are few studies assessing the prognosis of *MLLT10*r AML, there is a trend towards intermediate to poor outcomes.[3,10,28] In one study, *PICALM-MLLT10* AML cases were at high risk of extramedullary disease and CNS involvement.[3] Median OS of patients with *PICALM-MLLT10* (12.3 months) was between intermediate and high-risk cytogenetic subtypes of AML (35 months vs 7.6 months, intermediate and high-risk subtypes respectively).[3] A Children's oncology group study of 2226 children and young adults (0-30 years) diagnosed with AML and treated on AAML0531 or AAML1031 trial protocols identified that *MLLT10*r conferred a high risk of poor outcome (5 year EFS 18.6% vs 49% for non-*MLLT10*r), and there were no statistically significant differences between *KMT2A-MLLT10* and other *MLLT10*r (5 year EFS 19.5% vs 12.7% respectively).[10] Another study reported *KMT2A-MLLT10* as a predictor of poor outcome compared with other *KMT2A*r in an AML patient cohort (5-year EFS 31% vs 44% for *KMT2A*r overall), but this was not compared to cases of other *MLLT10*r[28].

3.3 | Acute leukemia of ambiguous lineage and varying stages of differentiation arrest

Mixed phenotype or acute undifferentiated leukemia harboring *MLLT10*r occur in rare cases.[4] Aberrant co-expression of myeloid markers in *MLLT10*r T-ALL, or mixed phenotypic presentation such as the presence of a mediastinal mass in *MLLT10*r AML, represent a diagnostic challenge.[29] An isolated case of *PICALM-MLLT10* T-lineage lymphoblastic lymphoma (T-LBL) with aberrant myeloid marker expression has also been reported.[5] The occurrence of *MLLT10*r across AML, B-ALL, acute leukemia of ambiguous lineage, ETP-ALL and T-ALL at varying stages of differentiation arrest raises questions about the etiology of *MLLT10*r acute leukemia presenting these various immunophenotypes.

3.4 | 5′ fusion partners of *MLLT10*

MLLT10 is invariably the 3′ gene fusion partner and the OM-LZ domain is conserved in all reported cases, whereas numerous 5′ partner genes have been reported (Table 3). It is unclear if the 5′ fusion partner contributes leukemic transformation by *MLLT10*r, or whether expression of truncated 3′ *MLLT10* is the sole leukemic driver in these fusions. Expression of reciprocal *PICALM-MLLT10* and *DDX3X-MLLT10*/t(X;10)(p11;p12) are identified in approximately 50% of cases, and expression does not correlate with any specific breakpoint location or fusion partner.[15,30,31] As *PICALM-MLLT10* and *DDX3X-MLLT10* typically present as simple translocations, it is unclear why reciprocal expression is only detected in 50% of cases.[15,31]

3.5 | *PICALM-MLLT10*

The t(10;11)(p12-13;q14-21) translocation was first identified in the U937 cell line, derived from the pleural effusion of a 37-year-old male with diffuse histiocytic lymphoma.[32,33] This translocation fused 3′ *MLLT10* to the novel *CALM* gene (Clathrin assembly lymphoid myeloid leukemia),[32] later named *PICALM* (Phosphatidylinositol binding clathrin assembly lymphoid myeloid leukemia) to reflect its integral function in clathrin-mediated endocytosis.[34,35] *PICALM-MLLT10* is typically a simple two-way reciprocal chromosomal translocation, and is the most common and well-studied *MLLT10* fusion. The molecular mechanism of leukemogenesis by *PICALM-MLLT10* is not entirely understood. Inactivating mutations in *PICALM* result in defective hematopoiesis and iron uptake in mice, suggesting an important role of *PICALM* within the hematopoietic system.[36,37] *PICALM-MLLT10* leukemia is also dependent on expression of *BMI1* and *HOXA5*, as genetic depletion of either target results in failure of *PICALM-MLLT10* to transform murine hematopoietic precursor cells.[37,38]

Two C-terminal domains of PICALM, the NES and clathrin-binding domain (Figure 2B), are potential contributors to the leukemogenic mechanism of *PICALM-MLLT10*. The PICALM NES mediates cytoplasmic

localization of PICALM and PICALM-MLLT10. Deletion of the NES results in localization of PICALM-MLLT10 to the nucleus resulting in total abrogation of the colony-forming potential of murine bone marrow cells transduced with *PICALM-MLLT10*. Conway et al[40] propose that cytoplasmic localization of PICALM-MLLT10 and PICALM-MLLT10-bound DOT1L, mediated by the NES of PICALM, is responsible for the global H3K79 hypomethylation observed in *MLLT10*r patients. Interestingly, the NES of DDX3X is also invariably retained in *DDX3X-MLLT10* (Figure 2C), but further functional studies are not available.[15]

The PICALM clathrin-binding domain is integral in clathrin-mediated endocytosis, and deletion of the clathrin-binding domain from PICALM-MLLT10 alters myeloid disease phenotype in murine xenotransplantation models.[43] Mice transplanted with transduced murine bone marrow cells expressing full-length *PICALM-MLLT10* consistently acquired AML, whereas *PICALM-MLLT10* lacking the clathrin binding domain instead induced disease with a preleukemic myeloproliferative phenotype.[43] Another study demonstrated that the fusion of the clathrin-binding domain of PICALM to the OM-LZ domain of MLLT10 was sufficient for leukemogenesis in murine bone marrow cells, but the transforming capacity of the MLLT10 OM-LZ domain alone was not evaluated.[46] Notably, the clathrin binding domain is partially truncated in approximately 50% of patients[2,12] (Figure 2B), suggesting the entire domain is not essential in all cases. To date however, it is not clear whether the breakpoint of PICALM-MLLT10 influences disease phenotype or outcomes in patients, as existing data is conflicting.[2,12] Further to this, other *MLLT10* fusion partners, such as *XPO1* and *NAP1L1*, do not contain a NES or clathrin-binding domain, and the role of these fusion partners in *MLLT10*r-mediated leukemogenesis remains unknown.

Both xenotransplantation and transgenic murine models of *PICALM-MLLT10* yield leukemia with a mixed or myeloid phenotype, reflecting the occurrence of *PICALM-MLLT10* in AML, T-ALL and acute leukemia of ambiguous lineage. Murine models reveal disease characterized by expression of myeloid markers (Mac1, Gr1, MPO) and B220 (lymphoid marker), infiltration of organs including the thymus, spleen and bone marrow, impaired myeloid and T-cell maturation,[47] and 8-fold or greater increase in expression of multiple members of the *HOXA* gene cluster (*HOXA5*, 7, 10, 11, and *MEIS1*).[47,48] Notably, laboratory studies to date model the disease in murine non-lineage-committed bone marrow progenitor cells, and comparable studies in thymic precursor cells may provide further insight into the pathogenesis of *MLLT10*r T-ALL. Furthermore, a long latency period was observed in a transgenic murine model of *PICALM-MLLT10* leukemia, where 40% to 50% of *PICALM-MLLT10*-expressing mice developed acute leukemia with a median latency of 12 months.[47] The long latency of transgenic mice may indicate the requirement of additional genomic lesions to induce leukemia.

Data describing cooperative genomic lesions in *MLLT10*r ALL is currently very limited. Variants in the histone methyltransferase *EZH2* were identified in 4/15 (27%) of *PICALM-MLLT10* acute leukemia patients (3 T-ALL and 1 AML), whereas none were identified in 12 matched non-*PICALM-MLLT10* T-ALLs.[49] *EZH2* mutations have been reported in B-cell lymphomas and chronic hematopoietic malignancies, but are rare in ALL and AML.[49,50] *PICALM-MLLT10* T-ALL is also less likely to co-occur with a *NOTCH1* or *FBXW7* mutation (46%

vs 69%, not statistically significant).[2] Consistent with data demonstrating global genome instability conferred by *PICALM-MLLT10*[40,51], a secondary cytogenetic abnormality or complex karyotype was identified in a significantly higher proportion of *PICALM-MLLT10* leukemia patient samples, compared with *KMT2A-MLLT3* and *PML-RARA*.[51] However, it is unclear whether this observation extends to a higher overall mutational burden for *PICALM-MLLT10* leukemia.

3.6 | DDX3X-MLLT10

The *DDX3X-MLLT10* fusion arises from the t(X;10)(p11;p12) reciprocal chromosomal translocation. It is the second most common *MLLT10*r in T-ALL, with an estimated incidence of 1% to3%,[13,15] and is comparatively rare in AML (Table 3).[10] DDX3X is a DEAD box RNA helicase involved in RNA transcription, splicing, mRNA transport, initiation of translation, and stress granule assembly.[18,52] Inactivating missense point mutations and truncating mutations of DDX3X are associated with other malignancies including medulloblastoma[53] and carcinomas of the lung, breast and uterus.[52] At present, the molecular mechanisms of *DDX3X-MLLT10* or inactivating *DDX3X* mutations have not been explored in any malignancy, but Epling et al[53] speculate that impaired translation of mRNA by defective DDX3X may contribute to tumorigenesis in pediatric medulloblastoma.

DDX3X-MLLT10 acute leukemia cases are rare; one report of four adult *DDX3X-MLLT10* T-ALL patients is the largest series to date.[15] Each fusion identified was in-frame, with variable breakpoints spanning *DDX3X* exons 3 to 6 and *MLLT10* exons 6 to 17 (Figure 2C). The MLLT10 OM-LZ domain, at least one MLLT10 NLS and the DDX3X NES were invariably retained, but the DDX3X DEAD box domain was consistently lost. An in-frame reciprocal *MLLT10-DDX3X* was identified in 2/4 cases, but is not known whether the presence of the reciprocal fusion influenced disease phenotype or outcome. Activating *NOTCH1* mutations were identified in all four patients within this case series, but further studies are required to establish the broader mutational landscape of *MLLT10*r leukemia.[15]

3.7 | KMT2A-MLLT10

KMT2A-MLLT10/t(10;11)(p12;q23) is the most common *MLLT10*r in B-ALL and AML, where it accounts for 80% of new *MLLT10*r diagnoses in AML[10], and is the only *MLLT10*r identified in B-ALL.[8] It is a comparatively rare entity in T-ALL (Figure 1 and Table 3). *KMT2A-MLLT10* a fascinating *MLLT10*r lesion for two reasons. First, unlike *PICALM-MLLT10* and *DDX3X-MLLT10*, *KMT2A-MLLT10* presents as a complex chromosomal rearrangement due to the opposing transcriptional orientations of *KMT2A* and *MLLT10*, typically involving an in-frame or out-of-frame three-way rearrangement.[8,45] It is also distinct from other *MLLT10*r because it involves the *KMT2A* gene. *KMT2A-MLLT10* therefore overlaps with two known genomic subtypes of acute leukemia, and is subsequently often categorized as both a *KMT2A*r[8,45] and *MLLT10*r[10]. It is likely that *KMT2A* is the dominant

TABLE 3 In-frame *MLLT10* (10p12) fusion partners identified in acute leukemia

5′ fusion partner	Chromosomal localization and orientation	Frequency in AML	Frequency in T-ALL	Frequency in B-ALL
PICALM	11q14.2 Centromeric	0.6%-1.7%[3,10]	6%-10%[2,12,14]	NA
DDX3X	Xp11.4 Centromeric	0.1%[10]	0.8%-3%[13,15]	NA
KMT2A	11q23.3 Telomeric	1.6%-4.7%[8,10]	Rare (isolated cases reported)[8,13,16,17]	3%-3.7%[8,9]
HNRNPH1	5q35.3 Centromeric	NA	one reported case[18]	NA
XPO1	2p15 Telomeric	one reported case[10]	one reported case[18]	NA
NAP1L1	12q21.1 Centromeric	one reported case[10]	three reported cases[13,19]	NA
TEC	4p12-p11 Telomeric	0.1%[10]	NA	NA
DDX3Y	Yq11.221 Telomeric	one reported case[10]	NA	NA
CEP164	11q23.3 Telomeric	one reported case[10]	NA	NA
SCN2B	11q23.3 Centromeric	one reported case[10]	NA	NA
TREH	11q23.3 Centromeric	one reported case[10]	NA	NA
CAPS2	12q21.1-q21.2 Centromeric	NA	one reported case[13]	NA
FAM171A1	10p13 Telomeric	NA	one reported case[13]	NA

Abbreviations: AML, acute myeloid leukemia; B-ALL, B-lineage acute lymphoblastic leukemia; T-ALL, T-lineage acute lymphoblastic leukemia.

leukemic driver in this fusion, based on the observation that no other *MLLT10*r occurs in B-ALL, whereas *KMT2A*r B-ALL occurs with over 70 different fusion partners.[44]

The pathogenesis of *KMT2A*r acute leukemia have been extensively studied, but there is a lack of data specifically investigating *KMT2A-MLLT10*. It is known that upregulation of several transcription factors including *HOXA9* and *EYA1* are important in *KMT2A*r acute leukemia, but it is not known whether there are any differences between *KMT2A-MLLT10* and other *KMT2A*r.[37] A murine xenotransplantation model using human CD34⁺ pluripotent stem cells transduced with *KMT2A-MLLT10* induced enhanced multilineage hematopoiesis, but failed to induce overt leukemia.[54] In addition, the same study revealed that co-expression of *KMT2A-MLLT10* and *KRAS* p.G12V rapidly induced AML, providing supporting evidence for the role of secondary genomic events in the development of *MLLT10*r and *KMT2A*r acute leukemia. The mutational landscape of *KMT2A*r acute leukemia is well-characterized overall. Of particular interest, variants within the *PI3K-RAS* pathway have been identified in cases of *KMT2A-MLLT10* AML and B-ALL across all age groups.[55-57] One study also observed that inactivating mutations of *PAX5* were more frequently associated with pediatric cases bearing *KMT2A-MLLT10* and *KMT2A-MLLT3*, compared with other *KMT2A*r.[55] Further functional research and large patient cohort studies would be beneficial to compare the molecular mechanisms and gene expression profile of *KMT2A-MLLT10* to other *MLLT10*r and *KMT2A*r, to inform classification of this fascinating fusion.

4 | RARE FUSION PARTNERS

A number of other 5′ fusion partners have been identified in few or isolated cases of *MLLT10*r leukemia (Table 3). Based on chromosomal localization and orientation, several fusion partner genes are able to form

simple reciprocal translocations with *MLLT10*, including *PICALM*, *DDX3X*, *HNRNPH1* and *NAP1L1*. Conversely, several fusion partners can only arise from complex rearrangements and three-way recombinations, including *KMT2A* and *XPO1* (Table 3). Several rare fusion partners have functions in nuclear export and mRNA processing, but further functional studies of these fusions are required. XPO1 mediates nuclear export of DDX3X-MLLT10 and PICALM-MLLT10 through direct binding to the NES conserved in these fusions.[58] *DDX3Y* is a paralog of *DDX3X*, and one case of *DDX3Y-MLLT10* has been reported to date.[10] *HNRNPH1* encodes an RNA processing factor with roles in the mediation of the nuclear export of some mRNAs.[59] The function of other *MLLT10* fusion partners vary broadly, but is not known whether the function of the 5′ fusion gene partner is important in *MLLT10*r function. *NAP1L1* has roles in nucleosome assembly and chromatin formation,[60] *TEC* encodes a non-receptor protein tyrosine kinase, and *FAM171A1* and *CEP164* are involved in cytoskeletal organization and DNA damage response.[61,62] It is unclear whether different *MLLT10* fusion partners vary functionally or influence patient outcomes, an important consideration in the prognostic stratification of patients, and when assessing the therapeutic sensitivity of *MLLT10*r leukemia.

It is also unknown whether cooperative genomic events may contribute to the development of acute leukemia with rare *MLLT10*r fusion partners. *NOTCH1* and *FBXW7* mutation status is variable in the few reported cases of *NAP1L1-MLLT10* and *HNRNPH1-MLLT10* T-ALL,[18,19] but research involving screening for co-occurring genomic aberrations in *MLLT10*r has been rare to date.

5 | THE ROLE OF *MLLT10* IN *MLLT10*R-MEDIATED LEUKEMOGENESIS

MLLT10 is a coactivator of DOT1L, the only known histone 3 lysine 79 (H3K79) methyltransferase in humans.[2,63] Both proteins are
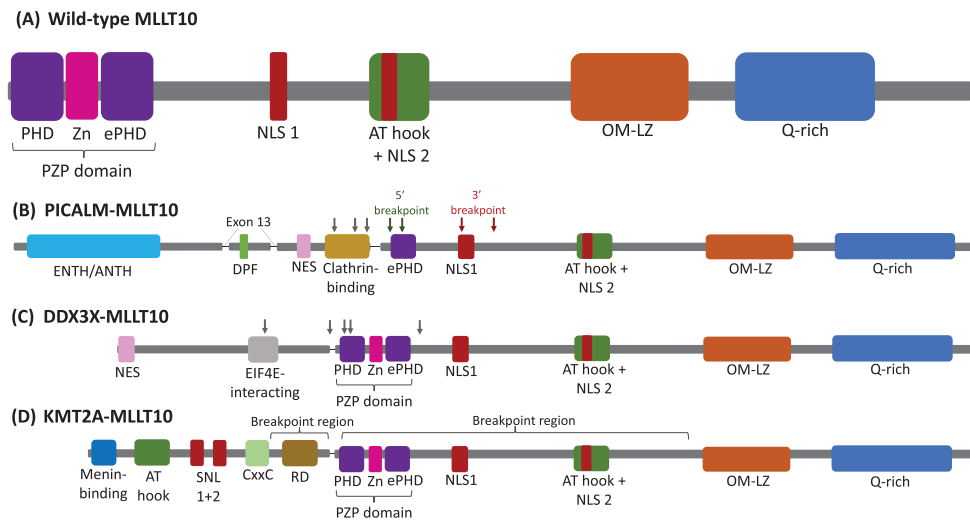
FIGURE 2 Functional domains of wild-type MLLT10 and common MLLT10 fusion proteins A, key functional domains present in the wild-type MLLT10 and B-D, common MLLT10 fusion proteins B, PICALM-MLLT10, C, DDX3X-MLLT10, and D, KMT2A-MLLT10 are shown. Grey arrows indicate recurrently reported breakpoint locations. A, Functional domains within MLLT10 include the plant homology domain (PHD) involved in interaction with histone proteins, extended PHD (ePHD) domain involved in mediation of homo-oligomerization of MLLT10, and an interspacing Zn-knuckle (Zn), together forming the PZP domain.[39] Other domains include nuclear localization signals (NLS), AT hook, Q-rich domain and the octapeptide motif-leucine zipper (OM-LZ) domain, the key functional domain of MLLT10.[2] Breakpoints within MLLT10r are widely variable, but consistently retain the OM-LZ domain.[15] B, Key domains of PICALM include the N-terminal ENTH/ANTH domain responsible for cell membrane association,[36] C-terminal nuclear export signal (NES), and clathrin binding domain.[40] Exon 13 is spliced out in 40% of PICALM transcripts (PICALM short form),[41] removing a DPF motif associated with endocytic adaptor AP2 binding, but this does not affect normal PICALM function and is variably present in PICALM-MLLT10 transcripts.[42] Breakpoints are typically located at the far C-terminus, and are typically categorized as either 5′ or 3′ breakpoints with respect to MLLT10, and conserve the NES and truncated or full-length clathrin-binding domain.[43] C, DDX3X-MLLT10 contains the invariably conserved NES of DDX3X, but the EIF4E-interacting domain required for modulation of translation by DDX3X is often truncated.[15] D, KMT2A-MLLT10 conserves the N-terminal domains of KMT2A responsible for protein interactions and non-specific DNA binding, including the menin-binding domain, AT hook, subnuclear localization (SNL) domains and CxxC region, but the repression domain (RD) involved in mediation of transcriptional repression is typically truncated.[44] KMT2A-MLLT10 breakpoints are widely variable, typically located within the 8.3 kb breakpoint cluster region of KMT2A spanning introns 9 to 14,[44] and have been identified across introns 1 to 18 in MLLT10, most frequently introns 8 and 9.[45] [Color figure can be viewed at wileyonlinelibrary.com]

essential components of the DOT1L complex responsible for H3K79 methylation of target gene promoter regions, to activate gene expression.[64,65] In MLLT10r acute leukemia, dysregulation of the interaction between MLLT10 and DOT1L results in aberrant recruitment of DOT1L, subsequent H3K79 methylation and target gene activation.[64] MLLT10r acute leukemias typically exhibit global H3K79 methylation, and hypermethylation at promoter regions of key leukemic genes such as the HOXA locus.[38,51]

MLLT10 fusions lacking the OM-LZ domain (Figure 2) are unable to transform retrovirally transduced murine bone marrow cells in vitro due to prevention of interaction with DOT1L, making it the only domain of MLLT10 with a known critical function in leukemogenesis.[38,46] No other domain of MLLT10 is known to be required for leukemogenesis, but the presence of the 5′ MLLT10 ePHD domain (Figure 2) may influence the stage at which differentiation arrest occurs. In T-ALL patients, 5′ breakpoints that conserve the ePHD domain of MLLT10 were more likely to occur in mature phenotype/TCR[pos] disease (ePHD conserved in 62% of cases) than in immature phenotype, TCR negative (TCR[neg]) cases (ePHD

conserved in 41% of cases).[2] However, this was not statistically significant, is yet to be validated experimentally, and is of unknown clinical importance.

5.1 | HOXA dysregulation in MLLT10r acute leukemia

Irrespective of the 5′ fusion partner, leukemic blasts of patients with MLLT10r are consistently characterized by high expression levels of the HOXA gene cluster, supporting the notion that MLLT10 is the predominant leukemic driver.[18,19,66] HOXA genes are homeodomain-containing transcription factors important in developmental regulation.[20,67] Most HOXA genes are transcriptionally silenced upon completion of embryonic development, but some remain important regulators of adult hematopoiesis, including HOXA9[67]. In both patients and laboratory models, increased expression of the HOXA cluster (typically HOXA3, HOXA5, HOXA7, HOXA9 and HOXA10) is a hallmark of acute leukemia bearing MLLT10r, as well as KMT2Ar involving any fusion partner,[10,31,46,49,68,69]

suggesting activation of common pathways in both subtypes. However, not all differentially expressed genes are important in leukemic initiation or maintenance. Okada et al[38] established that *HOXA5* upregulation by DOT1L is critical for leukemic transformation mediated by *PICALM-MLLT10* but not *KMT2A-MLLT10*. This study also demonstrated that *HOXA9* is not critical in *PICALM-MLLT10*-mediated leukemogenesis, whereas it is required for the transformation of murine bone marrow progenitors by *KMT2Ar*.[70,71] These studies indicate that *PICALM-MLLT10* and *KMT2A-MLLT10* have different mechanisms of leukemogenesis. Furthermore, current data demonstrates that in patients with *KMT2A-AFF1* B-ALL, expression of reciprocal *AFF1-KMT2A* is correlated with high levels of *HOXA* expression and more favourable outcomes. Comparatively, patients only expressing *KMT2A-AFF1* have low levels of *HOXA*, but high levels of *HOXB4* and *IRX1*, and poorer outcomes.[72] The relationship between reciprocal *MLLT10r* expression and *HOXA* expression is not known, or whether reciprocal fusion or *HOXA* expression influences outcome. Similar studies addressing the role of different *HOXA* genes in the transformation of other *MLLT10* fusions, such as *DDX3X-MLLT10*, would resolve whether other *MLLT10r* share a similar leukemogenic mechanism to that of *PICALM-MLLT10* or *KMT2A-MLLT10*.

5.2 | Expression of *MLLT10*-adjacent genes

Genes in close proximity to *MLLT10* on chromosome 10 demonstrate increased expression in *MLLT10r* patients[68,73] and in a laboratory model of *PICALM-MLLT10*[31], but not in cases of *KMT2Ar* with fusion partners other than *MLLT10*[68]. Specifically, increased expression of *COMMD3*, *BMI1*, *DNAJC1* and *SPAG6* are consistently observed across studies, all located centromeric to MLLT10 on chromosome 10 (Figure 3).[31,68,73] Genomic and pharmacological depletion of the polycomb group protein *BMI1* inhibits *PICALM-MLLT10*-driven AML[37], demonstrating a potential therapeutic target. The role of *DNAJC1*, *COMMD3* and *SPAG6* in *MLLT10r* acute leukemia is yet to be investigated, but both *COMMD3* and *SPAG6* have implicated roles in other malignancies. *COMMD3* levels are elevated in prostate cancer, and have been demonstrated to upregulate *C-MYC* expression, promoting proliferation in prostate cancer models.[74] *SPAG6* is a member of the cancer-testis antigen family primarily involved in regulation of the microtubule system, with secondary roles associated with inhibition of the Akt and p53 pathways.[75] High expression levels of *SPAG6* occur in numerous malignancies including

myelodysplastic syndrome, and genetic knockdown of *SPAG6* inhibits cell proliferation of myeloid leukemia cell lines, but pharmacological inhibitors of *SPAG6* are not yet available.[76]

5.3 | Other dysregulated pathways in *MLLT10r* acute leukemia

Understanding the pathways dysregulated in *MLLT10r* acute leukemia will guide the use of targeted therapies. Increased expression of genes involved in chromatin assembly and maintenance and DNA repair are observed in *PICALM-MLLT10* T-ALL.[31,51] *PICALM-MLLT10* expression in murine bone marrow cells induces genome-wide instability, likely due to widespread epigenetic changes such as reduced global H3K79 methylation.[40,51] Genome instability results in increased rates of DNA damage and subsequent increased expression of genes involved in DNA repair and chromatin assembly and maintenance.[40,51] It is speculated that widespread genome instability may underly the accumulation of cooperative mutations that lead to leukemic transformation by *PICALM-MLLT10*[51], but experimental data is lacking.

The class II homeobox gene *HHEX* is expressed in healthy hematopoietic stem cells, and is uniquely highly expressed in *MLLT10r* T-ALL, but not in T-ALL with *KMT2Ar*, *SET-NUP214* or *TCRB-HOXA*[18]. Other genes with increased expression in *MLLT10r* acute leukemia include *CD69* (T-cell activation antigen), *CDKN1B* (CDK inhibitor), *JUNB* (cancer-associated transcription factor) and *TCF8* (negative regulator of IL-2 expression).[73] Genes with low expression levels in *MLLT10r* T-ALL include *CDK4* (cyclin dependent kinase), *CSF3R* (G-CSF receptor), *FANCL* (ubiquitin E3 ligase mutated in Fanconi anemia) and *LMO4* (LIM protein expressed predominantly in T-cells and neural cells).[73] One study identified high expression of members of the Sonic hedgehog (Shh) signaling pathway in cases of *MLLT10r* and *KMT2Ar* T-ALL, and receptor and/or tyrosine kinase-mediated pathways such as *TGFB*, *PYK2*, *ERBB* and *PDGFR* in only *MLLT10r* cases.[68] Chen et al identified high expression of factors associated with pro-inflammatory cytokine signaling and innate immunity in a Tet-inducible model of *PICALM-MLLT10* AML. This was found to be driven through Janus kinase-Signal transducer and activator of transcription (JAK-STAT) and NF-κB signaling, specifically direct interaction of PICALM-MLLT10 with JAK1.[69] Evidently, various studies have identified a broad range of differentially expressed genes in *PICALM-MLLT10* acute leukemia. These findings should be validated in experimental models or independent patient cohorts, as some differentially expressed genes may represent
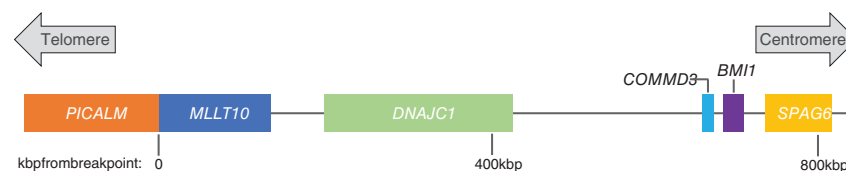

FIGURE 3 Genes in close proximity to MLLT10 on chromosome 10 *DNAJC1*, *COMMD3*, *BMI1* and *SPAG6* are genes located centromeric to *MLLT10* on chromosome 10, within 800 kilobase pairs (kbp) of the *PICALM-MLLT10* breakpoint. Figure adapted from Dik et al 2005.[73] [Color figure can be viewed at wileyonlinelibrary.com]

potential therapeutic targets, such *BMI1* and *JAK1*. It would also be interesting to compare the *MLLT10*r gene expression profile to other cases of T-ALL through large-cohort gene expression profiling studies, particularly whether *KMT2A-MLLT10* exhibits a gene expression profile more similar to that of other *MLLT10*r or *KMT2A*r.

5.4 | Treatment of *MLLT10*r acute leukemia

There are presently no targeted therapies available for the treatment of *MLLT10*r acute leukemia, and multiagent chemotherapy remains the standard of care for *MLLT10*r acute leukemias.[77] T-ALL patients more commonly present with higher risk clinical features as compared with B-ALL, such as hyperleukocytosis and extramedullary leukemic infiltration, and therefore undergo intensified treatment, which may include allogeneic hematopoietic stem cell transplantation.[77,78] Small studies indicate that allogeneic HSCT may provide disease-free survival benefit in TCR[neg] T-ALL or ETP-ALL *MLLT10*r patients, but larger cohort studies are required.[2,12,15] Data regarding therapeutic sensitivity of *MLLT10*r acute leukemia is lacking, owing to the rarity of these fusions, but DOT1L and BMI1 inhibition both represent promising targets in the treatment of *MLLT10*r acute leukemia, based on preclinical models.[37,79] Other potential candidates include histone deacetylase (HDAC) inhibitors and the DNA demethylating agent azacytidine. These agents have preclinical efficacy in *KMT2A*r leukemia,[68] and may have promise in *MLLT10*r cases, based on the similar gene expression profile shared by the two genomic subtypes.[68] The potential role of these targeted therapies in the treatment of acute leukemia remains unclear, but they are likely to play a supplementary role in optimizing the efficacy of existing multiagent chemotherapy regimens, rather than provide benefit as single agents.[80]

6 | DOT1L INHIBITORS

DOT1L is essential in both *KMT2A*r and *MLLT10*r-driven leukemogenesis,[64,81] and is the only known human H3K79 methyltransferase, representing a highly specific therapeutic target. Preclinical models demonstrate efficacy of DOT1L inhibition in primary *PICALM-MLLT10* AML cells[79] and in vitro AML models.[82] Early-stage clinical trials (clinicaltrials.gov IDs NCT01684150 and NCT02141828) established that the DOT1L inhibitor pinometostat had an acceptable toxicity profile and modest efficacy as a single agent in adults with *KMT2A*r acute leukemia.[83] Moving forward, clinical trials are assessing the efficacy of pinometostat in combination with standard chemotherapy (NCT03724084) or the DNA demethylating agent azacytidine in *KMT2A*r AML (NCT03701295). Preclinical testing is ongoing, evaluating the efficacy of DOT1L inhibitors in combination with other therapeutic agents. For example, dual inhibition of DOT1L and menin synergistically induces apoptosis, and inhibits proliferation of *KMT2A*r cell lines.[81] Further

preclinical work is required to validate the efficacy of DOT1L inhibition in *MLLT10*r AML and T-ALL, but DOT1L inhibition is currently the most advanced and promising therapeutic target for *MLLT10*r acute leukemia.

7 | BMI1 INHIBITORS

PICALM-MLLT10 AML is dependent on expression of *BMI1*, and murine and human models of *PICALM-MLLT10*-driven leukemias are sensitive to BMI1 inhibition.[31,37,73] BMI1 is a polycomb group protein, responsible for the epigenetic repression of developmentally important genes, including *CDKN2A* locus genes $p16^{INK4A}$ and $p14^{ARF}$, critical in the regulation of cell cycle progression and self-renewal of stem cells.[37,84] Both genetic and pharmacological inhibition of BMI1 impedes *PICALM-MLLT10*-driven myeloid transformation, and inhibits proliferation of human *PICALM-MLLT10* AML cells in vitro and in murine xenotransplantation models.[37] BMI1 is also overexpressed in a wide range of malignancies including multiple myeloma, glioblastoma, non-small cell lung cancer and carcinomas of the colon, liver, breast and prostate.[85] High expression levels of BMI1 in these diseases confers risk of metastasis and poor outcome.[86] Pharmacological inhibition with the BMI1 inhibitor PTC-209 is effective in preclinical models of head and neck squamous cell carcinoma,[87] glioblastoma[88] and multiple myeloma, both alone and in combination with epigenetic inhibitors targeting *EZH2* or bromodomain and extra-terminal motif (BET).[84] The efficacy of BMI1 inhibitors have not been compared in other specific genomic subtypes of acute leukemia, but PTC-596 induces apoptosis in AML progenitor cells.[89] No direct BMI1 inhibitor is yet to reach human clinical trials in any malignancy, but further investigation is warranted, given promising preclinical efficacy in a diverse range of tumors.

8 | JAK1 INHIBITORS

Chen et al[69] established that direct recruitment of *JAK1* by *PICALM-MLLT10* results in upregulation of JAK-STAT and NF-κB pathways. Cre-recombinase-mediated deletion of *JAK1* reduced proliferation and promoted differentiation of *PICALM-MLLT10* murine bone marrow progenitors in vitro, indicating that JAK1 is important in maintaining the undifferentiated state of *PICALM-MLLT10* leukemic cells. Targeting of undifferentiated leukemic precursor or stem cells is an important concept, as demonstrated by the role of arsenic trioxide and ATRA in *PML-RARA*-driven acute promyelocytic leukemia.[90] Treatment with two JAK1 inhibitors, itacitinib and filgotinib, both resulted in a dose-dependent decrease in proliferation, and increase in apoptosis of cells in vitro[69]. Itacitinib and filgotinib have passed phase I toxicity trials (NCT01905813 and NCT01419990 respectively), and are potential therapeutic targets that warrant further investigation in *PICALM-MLLT10* acute leukemia.[91]

9 | HDAC INHIBITORS

HDAC inhibitors have preclinical efficacy in a wide range of malignancies, including *KMT2A*r ALL.[80,92] In murine xenograft models of *KMT2A*r B-ALL, treatment with single-agent panobinostat (LBH589) reduced overall leukemic burden and significantly extended survival,[80] but a clinical trial assessing the use of single-agent panobinostat in relapsed ALL and AML (NCT00723203) was terminated due to lack of efficacy. Evidently, combinatorial therapeutic approaches are required to achieve disease control in these aggressive malignancies, but it is also possible that recruitment should be targeted to specific genomic subtypes of ALL and AML that have demonstrated preclinical sensitivity to HDAC inhibition.

Treatment of malignant cell lines with HDAC inhibitors induces broad epigenetic changes and loosening of chromatin, resulting in widespread changes in gene expression, increasing cell cycle inhibition and DNA damage.[93] The precise anti-cancer mechanisms of HDAC inhibitors are yet to be elucidated, but one study demonstrated that panobinostat elicits anti-leukemic activity through knockdown of the RNF20/RNF40/WAC E3 ligase complex, a critical pathway in *KMT2A*r leukemic maintenance.[80] It is unclear whether HDAC inhibition is effective in *MLLT10*r acute leukemia, but downregulation of polycomb group components including *BMI1* is observed in solid tumors in response to treatment with HDAC-inhibiting compounds such as trichostatin A and sodium butyrate.[86,94] These findings should be replicated in preclinical models of *MLLT10*r acute leukemia with clinically relevant HDAC inhibitors such as panobinostat and vorinostat.

9.1 | Other compounds with efficacy in *KMT2A*r acute leukemia

*KMT2A*r acute leukemia is well-studied, owing to the high incidence across B-ALL, T-ALL and AML (approximately 6%-7% in each disease subtype[44]), including a high proportion of infant diagnoses, and this subtype confers very poor outcomes across all ages.[95] Many therapeutic agents have preclinical efficacy in *KMT2A*r acute leukemia, and the high similarity in gene expression profiles of *KMT2A*r and *MLLT10*r acute leukemia suggest that these therapeutic targets may be effective in treating patients with *MLLT10*r.

Azacytidine is a DNA demethylating agent used in the treatment of myelodysplastic syndromes, with efficacy against promoter-hypermethylated infant *KMT2A*r leukemia.[96] A phase II clinical trial of azacytidine in combination with chemotherapy in infants with *KMT2A*r B-ALL is ongoing (NCT02828358). DNA methylation studies have not been performed in *MLLT10*r patients, but collection of these data may indicate whether DNA methylation inhibitors may be a potential therapeutic target for *MLLT10*r leukemia.

Several other inhibitors target direct interactors of complexes associated with *KMT2A* fusion proteins. BET inhibitors target BRD4 to prevent recruitment of the P-TEFb complex to inhibit aberrant gene transcription,[97] and inhibitors of menin-KMT2A interaction selectively induce apoptosis of *KMT2A*r cells.[98] The PARP1 inhibitor olaparib[99]

and CDK9 inhibitors[100] are also effective in preclinical models of *KMT2A*r AML. The efficacy of these therapeutics in *MLLT10*r acute leukemia is currently unknown. Therapeutic targeting of cooperative lesions in leukemia have also provided some success to date, such as MEK inhibition in *KMT2A*r ALL bearing pathogenic RAS-pathway mutations,[101] and research suggests further clinical trials are justified.

Treatment with single agents is unlikely to effectively control the aggressive nature of ALL or AML, and targeting only single epigenetic pathways may lead to non-genetic drug resistance, through remodelling of the signaling pathways upon which leukemic cells are dependent.[80] Instead, a combinatorial approach of therapies targeted towards different pathways, and/or in addition to conventional cytotoxic chemotherapeutics is likely to prove an optimal approach. Inhibitors targeting HDAC,[102] BET,[103,104] and menin[81] are ideal combinatorial agents in acute leukemia, inducing synergistic targeted effects in leukemic cells with either additional targeted therapies or conventional chemotherapeutic agents such as dexamethasone or vincristine. It is not known whether *MLLT10*r with different fusion partners respond differently to specific targeted therapies, or if specific inhibitors are effective only in certain disease phenotypes such as B-ALL, T-ALL, ETP-ALL or AML.

10 | CONCLUSIONS

Next-generation sequencing has recently identified numerous genomic subtypes within ALL and AML. The prognostic significance of many of these subtypes is still unclear, including *MLLT10*r, which represents one of the most common genomic aberrations in T-ALL. Current evidence suggests that TCR[neg] *PICALM-MLLT10* T-ALL and *MLLT10*r AML[10] are high-risk subgroups for poor outcomes, but larger cohort studies are required to fully characterize the clinical significance of *MLLT10*r in T-ALL and AML, to improve prognostic stratification of patients.[22]

The investigation of targeted therapies in *MLLT10*r disease is lacking, and further characterization of the molecular mechanisms behind *MLLT10*r leukemogenesis would prove beneficial in the development of targeted therapeutic strategies. Here, several candidate targets have been discussed, including inhibitors of DOT1L and BMI1, that exhibit promising preclinical potential in *MLLT10*r acute leukemia. Other targeted therapies such as azacytidine and inhibitors of BET, PARP and CDK9 should also be investigated, based on efficacy in *KMT2A*r acute leukemia, a well-characterized high-risk genomic subtype of ALL and AML. A precision medicine approach is undoubtedly the future of improving outcomes in acute leukemia, and it is imperative to inform treatment decisions with an understanding of both the molecular mechanisms and prognostic influence of genomic aberrations in this disease.

of these agencies have had a role in the preparation of this manuscript. All other authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

Michelle O. Forgione conceptualized the presented idea. Michelle O. Forgione constructed the manuscript in consultation with Barbara J. McClure, David T. Yeung, Laura N. Eadie, and Deborah L. White All authors provided critical feedback and helped shape the manuscript.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Michelle O. Forgione https://orcid.org/0000-0002-7020-7440
Barbara J. McClure https://orcid.org/0000-0002-5201-4127
David T. Yeung https://orcid.org/0000-0002-7558-9927
Laura N. Eadie https://orcid.org/0000-0003-1912-7602
Deborah L. White https://orcid.org/0000-0003-4844-333X

REFERENCES

1. Beverloo H, Le Coniat M, Wijsman J, et al. Breakpoint heterogeneity in t(10;11) translocation in AML-M4/M5 resulting in fusion of AF10 and MLL is resolved by fluorescent in situ hybridization analysis. *Cancer Res.* 1995;55(19):4220-4224.

2. Abdelali R, Asnafi V, Petit A, et al. The prognosis of CALM-AF10-positive adult T-cell acute lymphoblastic leukemias depends on the stage of maturation arrest. *Haematologica.* 2013;98(11):1711-1717.

3. Borel C, Dastugue N, Cances-Lauwers V, et al. PICALM-MLLT10 acute myeloid leukemia: a French cohort of 18 patients. *Leuk Res.* 2012;36(11):1365-1369.

4. Alexander TB, Gu Z, Iacobucci I, et al. The genetic basis and cell of origin of mixed phenotype acute leukaemia. *Nature.* 2018;562 (7727):373-379.

5. Khurana S, Melody ME, Ketterling RP, et al. Molecular and phenotypic characterization of an early T-cell precursor acute lymphoblastic lymphoma harboring PICALM-MLLT10 fusion with aberrant expression of B-cell antigens. *Cancer Genet.* 2020;240:40-44.

6. Naesens L, Devos H, Nollet F, Michaux L, Selleslag D. Mediastinal myeloid sarcoma with TP53 mutation preceding acute myeloid leukemia with a PICALM-MLLT10 fusion gene. *Acta Haematol.* 2018; 140(2):97-104.

7. Tate JG, Bamford S, Jubb HC, et al. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.* 2019;47(D1):D941-D947.

8. Meyer C, Burmeister T, Groger D, et al. The MLL recombinome of acute leukemias in 2017. *Leukemia.* 2018;32(2):273-284.

9. Marschalek R. Systematic classification of mixed-lineage leukemia fusion partners predicts additional cancer pathways. *Ann Lab Med.* 2016;36(2):85-100.

10. Ries R, Leonti A, Junius Triche T, et al. Structural variants involving MLLT10/AF10 are associated with adverse outcome in AML regardless of the partner gene - a COG/Tpaml study. *Blood.* 2019;134(S1):461.

11. Matlawska-Wasowska K, Harvey R, Heerema N, et al. AF10 fusion transcripts are identified by HOXA9/10 and MEIS1 overexpression in T-ALL: a report from Children's. *Oncology Group AALL0434 Blood.* 2013;122(21):2601.

12. Asnafi V, Radford-Weiss I, Dastugue N, et al. CALM-AF10 is a common fusion transcript in T-ALL and is specific to the TCR. *Neoplasia.* 2003;102(3):1000-1006.

13. Liu Y, Easton J, Shao Y, et al. The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nat Genet.* 2017;49(8):1211-1218.

14. Lo Nigro L, Mirabile E, Tumino M, et al. Detection of PICALM-MLLT10 (CALM-AF10) and outcome in children with T-lineage acute lymphoblastic leukemia. *Leukemia.* 2013;27(12):2419-2421.

15. Brandimarte L, La Starza R, Gianfelici V, et al. DDX3X-MLLT10 fusion in adults with NOTCH1 positive T-cell acute lymphoblastic leukemia. *Haematologica.* 2014;99(5):64-66.

16. Peterson JF, Baughn LB, Pearce KE, et al. KMT2A (MLL) rearrangements observed in pediatric/young adult T-lymphoblastic leukemia/lymphoma: a 10-year review from a single cytogenetic laboratory. *Genes Chromosomes Cancer.* 2018;57(11):541-546.

17. Peterson JF, Sukov WR, Pitel BA, et al. Acute leukemias harboring KMT2A/MLLT10 fusion: a 10-year experience from a single genomics laboratory. *Genes Chromosomes Cancer.* 2019;58(8): 567-577.

18. Brandimarte L, Pierini V, Giacomo DD, et al. New MLLT10 gene recombinations in pediatric T-acute lymphoblastic leukemia. *Blood.* 2013;121(25):5064-5067.

19. Bond J, Touzart A, Cieslak A, et al. NAP1L1-MLLT10 is a rare recurrent translocation that is associated with HOXA activation and poor treatment response in T-cell acute lymphoblastic leukaemia. *Br J Haematol.* 2015;174(3):470-473.

20. Belver L, Ferrando A. The genetics and mechanisms of T cell acute lymphoblastic leukaemia. *Nat Rev Cancer.* 2016;16(8): 494-507.

21. McMahon CM, Luger SM. Relapsed T cell ALL: current approaches and new directions. *Curr Hematol Malig Rep.* 2019;14 (2):83-93.

22. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood.* 2016;127(20):2391-2405.

23. Bond J, Marchand T, Touzart A, et al. An early Thymic precursor phenotype predicts outcome exclusively in HOXA-overexpressing adult T-cell acute lymphoblastic leukemia: a Group for Research in adult acute lymphoblastic leukemia study. *Haematologica.* 2015;101 (6):732-740.

24. Matlawska-Wasowska K, Kang H, Devidas M, et al. MLL rearrangements impact outcome in HOXA-deregulated T-lineage acute lymphoblastic leukemia: a Children's oncology group study. *Leukemia.* 2016;30(9):1909-1912.

25. De Kouchkovsky I, Abdul-Hay M. Acute myeloid leukemia: a comprehensive review and 2016 update. *Blood Cancer J.* 2016;6(7):e441.

26. Bullinger L, Döhner K, Döhner H. Genomics of acute myeloid leukemia diagnosis and pathways. *J Clin Oncol.* 2017;35(9):934-946.

27. Khwaja A, Bjorkholm M, Gale RE, et al. Acute myeloid leukaemia. *Nat Rev Dis Primers.* 2016;2(1):16010.

28. Balgobind BV, Raimondi SC, Harbott J, et al. Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. *Blood.* 2009;114 (12):2489-2496.

29. Savage NM, Kota V, Manaloor EJ, et al. Acute leukemia with PICALM-MLLT10 fusion gene: diagnostic and treatment struggle. *Cancer Genet Cytogenet.* 2010;202(2):129-132.

30. Bohlander S, Muschinsky V, Schrader K, et al. Molecular analysis of the CALM/AF10 fusion: identical rearrangements in acute myeloid leukemia, acute lymphoblastic leukemia and malignant lymphoma patients. *Leukemia.* 2000;14(1):93-99.

31. Mulaw MA, Krause A, Deshpande AJ, et al. CALM/AF10-positive leukemias show upregulation of genes involved in chromatin

assembly and DNA repair processes and of genes adjacent to the breakpoint at 10p12. *Leukemia*. 2012;26(5):1012-1019.

32. Dreyling M, Martinez-Climent J, Zheng M, Mao J, Rowley J, Bohlander S. The t(10;11)(p13;q14). In the U937 cell line results in the fusion of the AF10 gene and CALM, encoding a new member of the AP-3 clathrin assembly protein family. *Med Sci*. 1996;93(10): 4804-4809.

33. Sundström C, Nilsson K. Establishment and characterization of a human histiocytic lymphoma cell line (U-937). *Int J Cancer*. 1976;17 (5):565-577.

34. Tebar F, Bohlander S, Sorkin A. Clathrin assembly lymphoid myeloid leukemia (CALM) protein: localization in endocytic-coated pits, interactions with clathrin, and the impact of overexpression on clathrin-mediated traffic. *Mol Biol Cell*. 1999;10(8):2687-2702.

35. Ishikawa Y, Maeda M, Pasham M, et al. Role of the clathrin adaptor PICALM in normal hematopoiesis and polycythemia vera pathophysiology. *Haematologica*. 2015;100(4):439-451.

36. Scotland PB, Heath JL, Conway AE, et al. The PICALM protein plays a key role in iron homeostasis and cell proliferation. *PLoS One*. 2012; 7(8):e44252.

37. Barbosa K, Deshpande A, Chen BR, et al. Acute myeloid leukemia driven by the CALM-AF10 fusion gene is dependent on BMI1. *Exp Hematol*. 2019;74:42-51 e43.

38. Okada Y, Jiang Q, Lemieux M, Jeannotte L, Su L, Zhang Y. Leukaemic transformation by CALM-AF10 involves upregulation of Hoxa5 by hDOT1L. *Nat Cell Biol*. 2006;8(9):1017-1024.

39. Chen S, Yang Z, Wilkinson AW, et al. The PZP domain of AF10 senses unmodified H3K27 to regulate DOT1L-mediated methylation of H3K79. *Mol Cell*. 2015;60(2):319-327.

40. Conway AE, Scotland PB, Lavau CP, Wechsler DS. A CALM-derived nuclear export signal is essential for CALM-AF10-mediated leukemogenesis. *Blood*. 2013;121(23):4758-4768.

41. Parikh I, Fardo DW, Estus S. Genetics of PICALM expression and Alzheimer's disease. *PLoS One*. 2014;9(3):e91242.

42. Narita M, Shimizu T, Hayashi Y, et al. Consistent detection of CALM-AF10 chimaeric transcripts in haematological malignancies with t(10;11)(p13;q14) and identification of novel transcripts. *Br J Haematol*. 1999;105(4):928-937.

43. Stoddart A, Tennant TR, Fernald AA, Anastasi J, Brodsky FM, Le Beau MM. The clathrin-binding domain of CALM-AF10 alters the phenotype of myeloid neoplasms in mice. *Oncogene*. 2012;31(4): 494-506.

44. Forgione MO, McClure BJ, Eadie LN, Yeung DT, White DL. KMT2A rearranged acute lymphoblastic leukaemia: Unravelling the genomic complexity and heterogeneity of this high-risk disease. *Cancer Lett*. 2020;469:410-418.

45. Zerkalenkova E, Lebedeva S, Kazakova A, et al. Acute myeloid leukemia with t(10;11)(p11-12;q23.3): results of Russian pediatric AML registration study. *Int J Lab Hematol*. 2019;41(2):287-292.

46. Deshpande AJ, Rouhi A, Lin Y, et al. The clathrin-binding domain of CALM and the OM-LZ domain of AF10 are sufficient to induce acute myeloid leukemia in mice. *Leukemia*. 2011;25(11):1718-1727.

47. Caudell D, Zhang Z, Chung YJ, Aplan PD. Expression of a CALM-AF10 fusion gene leads to Hoxa cluster overexpression and acute leukemia in transgenic mice. *Cancer Res*. 2007;67(17):8022-8031.

48. Deshpande AJ, Cusan M, Rawat VP, et al. Acute myeloid leukemia is propagated by a leukemic stem cell with lymphoid characteristics in a mouse model of CALM/AF10-positive leukemia. *Cancer Cell*. 2006; 10(5):363-374.

49. Grossmann V, Bacher U, Kohlmann A, et al. EZH2 mutations and their association with PICALM-MLLT10 positive acute leukaemia. *Br J Haematol*. 2012;157(3):387-390.

50. McCabe MT, Ott HM, Ganji G, et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature*. 2012;492(7427):108-112.

51. Lin YH, Kakadia PM, Chen Y, et al. Global reduction of the epigenetic H3K79 methylation mark and increased chromosomal instability in CALM-AF10-positive leukemias. *Blood*. 2009;114(3):651-658.

52. Valentin-Vega Y, Wang Y, Parker M, et al. Cancer-associated DDX3X mutations drive stress granule assembly and impair global translation. *Nat Sci Rep*. 2016;6:25996.

53. Epling LB, Grace CR, Lowe BR, Partridge JF, Enemark EJ. Cancer-associated mutants of RNA helicase DDX3X are defective in RNA-stimulated ATP hydrolysis. *J Mol Biol*. 2015;427(9):1779-1796.

54. Moriya K, Suzuki M, Watanabe Y, et al. Development of a multi-step leukemogenesis model of MLL-rearranged leukemia using humanized mice. *PLoS One*. 2012;7(6):e37892.

55. Andersson AK, Ma J, Wang J, et al. The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nat Genet*. 2015;47(4):330-337.

56. Emerenciano M, Barbosa TC, Almeida Lopes B, Meyer C, Marschalek R, Pombo-de-Oliveira MS. Subclonality and prenatal origin of RAS mutations in KMT2A (MLL)-rearranged infant acute lymphoblastic leukaemia. *Br J Haematol*. 2015;170(2):268-271.

57. Fedders H, Alsadeq A, Schmäh J, et al. The role of constitutive activation of FMS-related tyrosine kinase-3 and NRas/KRas mutational status in infants with KMT2A rearranged acute lymphoblastic leukemia. *Haematologica*. 2017;102:e438-e442.

58. Camus V, Miloudi H, Taly A, Sola B, Jardin F. XPO1 in B cell hematological malignancies: from recurrent somatic mutations to targeted therapy. *J Hematol Oncol*. 2017;10(1):47.

59. Sun YL, Liu F, Liu F, Zhao XH. Protein and gene expression characteristics of heterogeneous nuclear ribonucleoprotein H1 in esophageal squamous cell carcinoma. *World J Gastroenterol*. 2016;22(32): 7322-7331.

60. Qiao H, Li Y, Feng C, Duo S, Ji F, Jiao J. Nap1l1 controls embryonic neural progenitor cell proliferation and differentiation in the developing brain. *Cell Rep*. 2018;22(9):2279-2293.

61. Pan YR, Lee EY. UV-dependent interaction between Cep164 and XPA mediates localization of Cep164 at sites of DNA damage and UV sensitivity. *Cell Cycle*. 2009;8(4):655-664.

62. Rasila T, Saavalainen O, Attalla H, et al. Astroprincin (FAM171A1, C10orf38). *Am J Pathol*. 2019;189(1):177-189.

63. Okada Y, Feng Q, Lin Y, et al. hDOT1L links histone methylation to leukemogenesis. *Cell*. 2005;121(2):167-178.

64. Zhang H, Zhou B, Qin S, et al. Structural and functional analysis of the DOT1L-AF10 complex reveals mechanistic insights into MLL-AF10-associated leukemogenesis. *Genes Dev*. 2018;32(5–6):341-346.

65. Mohan M, Herz HM, Takahashi YH, et al. Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom). *Genes Dev*. 2010;24(6):574-589.

66. Bond J, Bergon A, Durand A, et al. Cryptic XPO1-MLLT10 translocation is associated with HOXA locus deregulation in T-ALL. *Blood*. 2014;124(19):3023-3025.

67. Collins CT, Hess JL. Role of HOXA9 in leukemia: dysregulation, cofactors and essential targets. *Oncogene*. 2016;35(9):1090-1098.

68. Kang H, Sharma ND, Nickl CK, et al. Dysregulated transcriptional networks in KMT2A- and MLLT10-rearranged T-ALL. *Biomark Res*. 2018;6:27.

69. Chen B, Deshpande A, Kleppe M, et al. Genomic and proteomic profiling of AF10-fusion Oncoproteins reveal mechanisms of Leukemogenesis and actionable targets. *Blood*. 2018;132(1):544.

70. Faber J, Krivtsov AV, Stubbs MC, et al. HOXA9 is required for survival in human MLL-rearranged acute leukemias. *Blood*. 2009;113 (11):2375-2385.

71. Ayton PM, Cleary ML. Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. *Genes Dev.* 2003;17(18):2298-2307.

72. Marschalek R. Another piece of the puzzle added to understand t(4; 11) leukemia better. *Haematologica.* 2019;104(6):1098-1100.

73. Dik WA, Brahim W, Braun C, et al. CALM-AF10+ T-ALL expression profiles are characterized by overexpression of HOXA and BMI1 oncogenes. *Leukemia.* 2005;19(11):1948-1957.

74. Umbreen S, Banday MM, Jamroze A, et al. COMMD3:BMI1 fusion and COMMD3 protein regulate C-MYC transcription: novel therapeutic target for metastatic prostate cancer. *Mol Cancer Ther.* 2019; 18(11):2111-2123.

75. Zheng D-F, Wang Q, Wang J-P, et al. The emerging role of sperm-associated antigen 6 gene in the microtubule function of cells and cancer. *Mol Therapy Oncol.* 2019;15:101-107.

76. Jiang M, Chen Y, Deng L, Luo X, Wang L, Liu L. Upregulation of SPAG6 in Myelodysplastic syndrome: knockdown inhibits cell proliferation via AKT/FOXO signaling pathway. *DNA Cell Biol.* 2019;38(5):476-484.

77. Micale MA. Pediatric T-cell acute lymphoblastic leukemia. *Atlas Genet Cytogenet Oncol Haematol.* 2018;22(11):451–464.

78. Litzow MR, Ferrando AA. How I treat T-cell acute lymphoblastic leukemia in adults. *Blood.* 2015;126(7):833-841.

79. Sarkaria SM, Christopher MJ, Klco JM, Ley TJ. Primary acute myeloid leukemia cells with IDH1 or IDH2 mutations respond to a DOT1L inhibitor in vitro. *Leukemia.* 2014;28(12):2403-2406.

80. Garrido Castro P, van Roon EHJ, Pinhancos SS, et al. The HDAC inhibitor panobinostat (LBH589) exerts in vivo anti-leukaemic activity against MLL-rearranged acute lymphoblastic leukaemia and involves the RNF20/RNF40/WAC-H2B ubiquitination axis. *Leukemia.* 2018;32(2):323-331.

81. Dafflon C, Craig VJ, Méreau H, et al. Complementary activities of DOT1L and Menin inhibitors in MLL-rearranged leukemia. *Leukemia.* 2016;31(6):1269-1277.

82. Chen L, Deshpande AJ, Banka D, et al. Abrogation of MLL-AF10 and CALM-AF10-mediated transformation through genetic inactivation or pharmacological inhibition of the H3K79 methyltransferase Dot1l. *Leukemia.* 2013;27(4):813-822.

83. Stein EM, Garcia-Manero G, Rizzieri DA, et al. The DOT1L inhibitor pinometostat reduces H3K79 methylation and has modest clinical activity in adult acute leukemia. *Blood.* 2018;131:2661-2669.

84. Alzrigat M, Párraga A, Majumder M, et al. The polycomb group protein BMI-1 inhibitor PTC-209 is a potent anti-myeloma agent alone or in combination with epigenetic inhibitors targeting EZH2 and the BET bromodomains. *Oncotarget.* 2017;8(61):103731-103743.

85. Cao L, Bombard J, Cintron K, Sheedy J, Weetall ML, Davis TW. BMI1 as a novel target for drug discovery in cancer. *J Cell Biochem.* 2011;112(10):2729-2741.

86. Bommi PV, Dimri M, Sahasrabuddhe AA, Khandekar J, Dimri GP. The polycomb group protein BMI1 is a transcriptional target of HDAC inhibitors. *Cell Cycle.* 2014;9(13):2663-2673.

87. Wang Q, Li Z, Wu Y, et al. Pharmacological inhibition of Bmi1 by PTC-209 impaired tumor growth in head neck squamous cell carcinoma. *Cancer Cell Int.* 2017;17:107.

88. Kong Y, Ai C, Dong F, et al. Targeting of BMI-1 with PTC-209 inhibits glioblastoma development. *Cell Cycle.* 2018;17(10):1199-1211.

89. Nishida Y, Maeda A, Kim MJ, et al. The novel BMI-1 inhibitor PTC596 downregulates MCL-1 and induces p53-independent mitochondrial apoptosis in acute myeloid leukemia progenitor cells. *Blood Cancer J.* 2017;7(2):e527.

90. Jeanne M, Lallemand-Breitenbach V, Ferhi O, et al. PML/RARA oxidation and arsenic binding initiate the antileukemia response of As2O3. *Cancer Cell.* 2010;18(1):88-98.

91. Phillips T, Forero-Torres A, Sher T, et al. Phase 1 study of the PI3Kδ inhibitor INCB040093 +/− JAK1 inhibitor itacitinib in relapsed/refractory B-cell lymphoma. *Blood.* 2018;132(3):293-306.

92. Eckschlager T, Plch J, Stiborova M, Hrabeta J. Histone Deacetylase inhibitors as anticancer drugs. *Int J Mol Sci.* 2017;18(7):1414.

93. Kim H, Bae S. Histone deacetylase inhibitors: molecular mechanisms of action and clinical trials as anti-cancer drugs. *Am J Transl Res.* 2011;3(2):166-179.

94. Dong Q, Sharma S, Liu H, et al. HDAC inhibitors reverse acquired radio resistance of KYSE-150R esophageal carcinoma cells by modulating Bmi-1 expression. *Toxicol Lett.* 2014;224(1):121-129.

95. Brown P, Pieters R, Biondi A. How I treat infant leukemia. *Blood.* 2019;133(3):205-214.

96. Chijimatsu I, Imanaka Y, Tomizawa D, et al. Azacitidine successfully maintained the second remission in an infant with KMT2A-rearranged acute lymphoblastic leukemia who relapsed after unrelated cord blood transplantation. *Pediatr Blood Cancer.* 2017;64(12):1-4.

97. Coudé M, Braun T, Berrou J, et al. BET inhibitor OTX015 targets BRD2 and BRD4 and decreases c-MYC in acute leukemia cells. *Oncotarget.* 2015;6(19):17698-17712.

98. Brzezinka K, Nevedomskaya E, Lesche R, et al. Characterization of the Menin-MLL interaction as therapeutic cancer target. *Cancers (Basel).* 2020;12(1):201.

99. Maifrede S, Martinez E, Nieborowska-Skorska M, et al. MLL-AF9 leukemias are sensitive to PARP1 inhibitors combined with cytotoxic drugs. *Blood Adv.* 2017;1(19):1467-1472.

100. Baker A, Gregory GP, Verbrugge I, et al. The CDK9 inhibitor Dinaciclib exerts potent apoptotic and antitumor effects in preclinical models of MLL-rearranged acute myeloid leukemia. *Cancer Res.* 2016;76(5):1158-1169.

101. Kerstjens M, Driessen E, Willekes M, et al. MEK inhibition is a promising therapeutic strategy for MLL-rearranged infant acute lymphoblastic leukemia patients carrying RAS mutations. *Oncotarget.* 2017;8(9):14835-14846.

102. Sun K, Atoyan R, Borek MA, et al. Dual HDAC and PI3K inhibitor CUDC-907 Downregulates MYC and suppresses growth of MYC-dependent cancers. *Mol Cancer Ther.* 2017;16(2):285-299.

103. Astorgues-Xerri L, Vazquez R, Odore E, et al. Insights into the cellular pharmacological properties of the BET-inhibitor OTX015/MK-8628 (birabresib), alone and in combination, in leukemia models. *Leuk Lymphoma.* 2019;60(12):3067-3070.

104. McCalmont H, Li KL, Jones L, et al. Efficacy of combined CDK9/BET inhibition in preclinical models of MLL-rearranged acute leukemia. *Blood Adv.* 2020;4(2):296-300.

1.5 The *HOXA* gene cluster is essential in *KMT2A*r and *MLLT10*r-mediated leukaemogenesis

1.5.1 The role of *HOXA* genes in development

As discussed in **sections 1.3** and **1.4**, fusion genes involving *KMT2A* or *MLLT10* induce dysregulation of H3K4me3 or H3K79me respectively. Truncation of *KMT2A* or *MLLT10* via gene fusions results in aberrant recruitment of associated protein complexes such as the super elongation complex (SEC) and DOT1L complex, and subsequently dysregulation of target gene expression, most notably the *HOXA* gene cluster.

HOX genes are highly conserved regulators of embryonic development. There are 39 human *HOX* genes localised to four clusters: *HOXA* (7p14-15), *HOXB* (17q21-22), *HOXC* (12q12-13) and *HOXD* (2q31-37)[16]. Additionally, numerous "orphan" *HOX* genes also exist, that occur outside of the four defined clusters but still contain the hallmark homeobox domain, such as *HOX11*[17]. In normal development, certain *HOX* genes function in a tissue-specific manner[18]. *HOXA-C* cluster genes are expressed in haematopoietic stem cells, and are gradually downregulated as cells differentiate[16]. In particular, *HOXA* cluster genes have a demonstrated role in leukaemogenesis (refer to section **1.5.2**)[16]. The structure and basic interactions of the *HOXA* clusters is described in **Figure 1.4**.
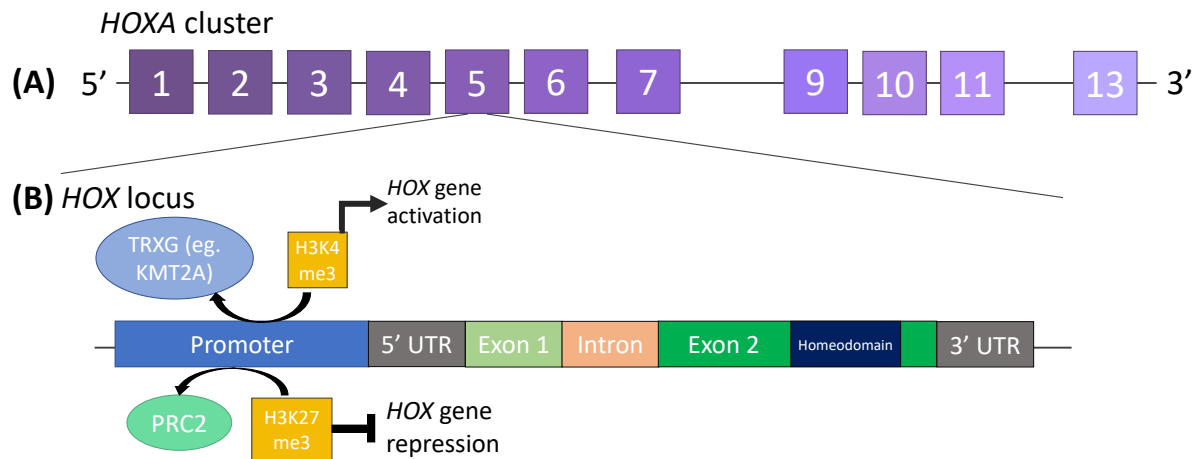
Figure 1.4 – *HOXA* gene cluster structure

Schematic of the *HOXA* gene cluster, and the basic DNA structure and promoter interactions of *HOXA* genes. **(A)** *HOXA* genes are ordered numerically 5' to 3', directly related to the order of *HOX* gene expression within a cluster throughout development. 3' *HOXA* genes are expressed anteriorly and earlier, and 5' genes are expressed posteriorly and later in development. 5' *HOX* genes exhibit a dominant phenotype over 3' genes (termed 'posterior prevalence'). **(B)** The basic DNA structure of a *HOX* gene is shown, including promoter region (blue), where histone modifiers such as TRXG (trithorax group) proteins such as KMT2A modify chromatin through H3K4me3 deposition to activate gene expression, and PRC2 (polycomb repressor complex 2) members can inactivate gene expression through H3K27me3 deposition. The characteristic domain of *HOX* genes is the DNA-binding homeodomain, located in exon 2 of all *HOX* members. Image adapted from Shah *et al.* 2010[18].

1.5.2 *HOXA* gene dysregulation in ALL

Aberrant *HOX* gene activation is central to many genomic subtypes of ALL. Fusion genes directly involving *HOX* genes include *TLX1*, *TLX3* (also known as *HOX11* and *HOX11L2* respectively) and gene rearrangements involving the *HOX* orphan gene *NKX2-1*, that together comprise approximately 30% of T-ALL diagnoses (**Figure 1.2**)[19,20]. *HOX* dysregulation can also occur indirectly, through mutations in regulators of *HOX* gene expression. Epigenetic regulation is integral to *HOX* gene expression control, and subsequently *HOX* dysregulation occurs as a result of mutations in transcriptional or epigenetic regulators including *SET-NUP214*, *MLLT10*r and *KMT2A*r, comprising 17% of new T-ALL diagnoses and 10% of B-ALL diagnoses[18,21].

Aberrant *HOX* expression influences tumour progression across a wide range of malignancies by dysregulating angiogenesis, proliferation, migration, metabolism and differentiation[18,19]. In leukaemia, aberrant *HOX* gene activation is thought to interfere with haematopoietic precursor cell differentiation, resulting in evasion of normal cellular regulatory mechanisms[18,19]. However, the precise mechanism by which aberrant *HOX* expression results in leukaemogenesis is not entirely clear, and further research in this field is required. As *HOX* dysregulation is a recurrently identified feature within T-ALL, its potential as a therapeutically targetable pathway should be evaluated.

1.6   Therapeutic targeting of *KMT2A*r and *MLLT10*r ALL

1.6.1   Chemotherapy in ALL

All subtypes of ALL are currently treated with multiagent chemotherapy regimens[5]. Treatment for ALL is generally comprised of three main stages: induction, intensification (or consolidation) and maintenance (or continuation) therapy[5,22]. Induction therapy generally includes a glucocorticoid (prednisone or dexamethasone), vincristine and asparaginase, and may also include anthracycline[5]. Consolidation therapy aims to eradicate residual leukaemic blasts, and typically comprises of methotrexate with mercaptopurine, asparaginase, vincristine and glucocorticoids[5]. Maintenance therapy typically lasts at least two years, and involves daily mercaptopurine with weekly methotrexate[5]. Vincristine and dexamethasone may also be administered. There are also selected cases where targeted therapies may be used in combination with chemotherapy, such as the use of tyrosine kinase inhibitors (TKIs) in *BCR-ABL1* positive or *BCR-ABL1*-like ALL[23].

Chemotherapy is effective at killing highly proliferative malignant cells, but lacks the precision efficacy of small molecule inhibitors that target specific cancer pathways. Consequently, chemotherapy also affects healthy rapidly-dividing cells, including hair follicles and mucosal linings such as the gastrointestinal tract, and so chemotherapy carries significant acute toxicities such as diarrhoea, hair loss, risk of infection due to pancytopenia, peripheral neuropathy and oral mucositis[24]. The chronic effects of chemotherapy are also significant, particularly for young children who survive ALL long-term, and include metabolic diseases, neurocognitive impairment, increased risk of secondary malignancies, microbiome dysbiosis[25], and impaired fertility[24,26,27].

1.6.2   Targeted therapies in *KMT2A*r and *MLLT10*r ALL

To date, no targeted therapies have been approved for the treatment of *KMT2A*r or *MLLT10*r ALL. This is of particular concern in *KMT2A*r acute leukaemia, where patients have poor outcomes despite intensified treatment protocols, and *KMT2A*r ALL is disproportionately common in infants who are adversely affected by the chronic toxicities discussed in section **1.6.1**. Consequently, treatment options that both improve long-term survival and reduce treatment-related morbidity are required. However, as discussed in section **1.3**, many classes of targeted therapies have been

trialled pre-clinically and/or clinically in *KMT2A*r B-ALL and AML, with varying degrees of efficacy. **Table 1.1** summarises pre-clinical data of novel targeted therapies in *KMT2A*r and *MLLT10*r acute leukaemia, including *in vitro* and *in vivo* models.

Many agents with pre-clinical efficacy have progressed to human clinical trials (**Table 1.2**). Birabresib (BET inhibitor)[28,29], pinometostat (DOT1L inhibitor)[30], and panobinostat[31] and vorinostat[32] (histone deacetylase inhibitors) have yielded underwhelming results as single agents, despite promising pre-clinical data in acute leukaemia models (**Tables 1.1-1.2**). M ost clinical trials involving these agents were broadly opened to relapsed or refractory (R/R) patients with ALL and/or AML, and participant genomic data is not published for most trials. It is therefore unclear whether these agents may have specific sensitivities in particular genomic subtypes such as *KMT2A*r.

Due to lack of efficacy as single agents, further trials are underway to combine targeted therapies with standard-of-care chemotherapy regimens, or demethylating agents such as azacitidine or decitabine, or with other targeted therapies such as venetoclax or TKIs (eg. dasatinib in *BCR-ABL1* ALL) (**Table 1.2**). A more recently established therapy is the curaxin CBL0137, that has exhibited promising pre-clinical efficacy in two PDX models of *KMT2A*r infant ALL[33,34]. It is currently not known whether CBL0137 is also efficacious in non-infant *KMT2A*r ALL, or other genomic subtypes of ALL, but CBL0137 represents a promising therapy that warrants further pre-clinical investigation. Clinical trials will provide further insight into the clinical feasibility of targeted therapies in ALL, to potentially supplement existing treatment protocols to improve patient outcomes and/or reduce treatment-related toxicities.

Table 1.1 – Pre-clinical data of novel therapeutic agents in *KMT2A*r acute leukaemia

Mechanism of action	Agent	*In vitro* results	*In vivo* results (murine models)
Anexelekto/MER tyrosine kinase inhibitor	ONO-7475	Induces growth arrest and apoptosis in *KMT2A*r and *FLT3*-ITD cell lines MOLM13 and MV4;11, but not *KMT2A*-WT and *FLT3*-WT cell lines HL60 and OCI-AML3[35]	Prolongs survival of MOLM13 and MV4;11 CDX models[35] (**Table 1.2**)
BCL-2 inhibitor	Venetoclax (ABT-199)	Induces apoptosis in a range of ALL cell lines and primary ALL samples, including both *KMT2A*r and *KMT2A*-WT[36]	Attenuates tumour growth of *KMT2A*r B-ALL PDX models[36] Currently in human clinical trials (**Table 1.2**)
Bromodomain and extra-terminal motif (BET) inhibitor	Birabresib (MK-8628/OTX015)	Induces apoptosis in a range of ALL cell lines, including both *KMT2A*r and *KMT2A*-WT[37] Inhibits proliferation and induces apoptosis in a range of ALL and AML cell lines; not *KMT2A*r-dependent[38]	In primary AML *ex vivo* assays: impairs leukaemia stem cell development in *CBF*r, *KMT2A*r or NPM1-MUT, but not in aneuploidy or *TP53*-MUT[39] (**Table 1.2**)
	FT-1101	Demonstrates anti-proliferative activity across a range of human leukaemia, lymphoma and multiple myeloma cell lines[40]	Significantly inhibits tumour growth of MV4;11 and THP-1 *KMT2A*r xenograft models[40] (**Table 1.2**)
Curaxin (p53 and NF-κB regulator)	CBL0137	Induces apoptosis at nanomolar concentrations across a range of *KMT2A*r and *KMT2A*-WT ALL and AML cell lines[33]	Induces objective clinical responses in PDX models of *KMT2A*r infant ALL[33,34], and non-infant paediatric PDX B- and T-ALL models of unknown genomic subtype[34]
DNA demethylating agents	Azacitidine (Vidaza), decitabine, zebularine	Inhibits cell proliferation of *KMT2A-AFF1* B-ALL cell lines SEM and RS4;11[41]	FDA-approved for treatment of MDS, and currently in human clinical trials as a combinatorial agent in leukaemia (**Table 1.2**)
DOT1L inhibitor	EPZ004777	Inhibits growth and promotes myeloid cell differentiation in primary adult AML samples (n=2/4 *KMT2A*r, 4/4 *KMT2A*-PTD, 1/1 *PICALM-MLLT10*, 7/7 *IDH1/2*-mutated, 0/2 normal karyotype)[42] Decreases proliferation and colony-forming potential, and induces differentiation, of murine bone marrow cells transformed with *KMT2A-MLLT10* or *PICALM-MLLT10*[43] Inhibits proliferation of various *KMT2A*r AML and ALL cell lines, but not *KMT2A*-WT cell lines[44]	Inhibits leukaemic initiation and progression of *KMT2A-MLLT10* and *PICALM-MLLT10*-transformed murine bone marrow cells[43] Increases median survival of MV4;11 xenograft mice[44] *Not further developed due to unfavourable pharmacokinetic (PK) profile – superseded by EPZ-5676*[45]

	Pinometostat (EPZ-5676)	Acts synergistically with chemotherapeutic agents in *KMT2A*r AML cell lines MOLM-13 & MV4;11, but not *KMT2A*-WT SKM-1[46]	Induces complete and sustained tumour regression in a MV4;11 PDX model[47] Currently in human clinical trials (**Table 1.2**)
HDAC inhibitor	Panobinostat (LBH-589)	Induces growth arrest and apoptosis in *KMT2A*r ALL cell lines SEM and KOPN8, but not *KMT2A*-WT REH and Jurkat cells[48]	Delays disease onset and significantly increases survival of *KMT2A-AFF1* PDX models[48] Currently in human clinical trials (**Table 1.2**)
	Vorinostat (MK0683)	Induces growth arrest and apoptosis in cell lines when applied in combination with numerous other targeted therapies and chemotherapeutic agents[49]	Numerous combinations with other targeted therapies and chemotherapeutic agents[49] Currently in human clinical trials (**Table 1.2**)
KMT2A/WDR5 inhibitor	MM-401	Induces cell cycle arrest, apoptosis and differentiation of *KMT2A*r murine bone marrow cells transformed with *KMT2A-MLLT3* and *KMT2A-MLLT1*, but not control cells[50]	Nil
	MM-589 (close analogue of MM-401)	Inhibits MOLM-13 and MV4;11 cell proliferation[51]	Nil
MDM2 inhibitor (p53 activator)	Idasanutlin (RG7388)	Inhibits viability in MV4;11 and MOLM-13 AML cell lines[52]	Induces growth arrest and apoptosis in xenograft models of p53-WT malignancies[53]
MEK inhibitor	MEK162	Reduces viability and induces apoptosis of *RAS*-MUT *KMT2A*r cell lines and primary infant leukaemic blasts[54] Synergistically enhances prednisolone sensitivity of *RAS*-WT and *RAS*-MUT *KMT2A-AFF1* cell lines[54]	Nil
	Selumetinib (AZD6244)	Reduces viability and induces apoptosis of *RAS*-MUT *KMT2A*r cell lines and primary infant leukaemic blasts[54] Synergistically enhances prednisolone sensitivity of *RAS*-WT and *RAS*-MUT *KMT2A-AFF1* cell lines[54] Selectively inhibits primary *RAS*-mutated ALL cell proliferation[55]	Reduces leukaemic burden and CNS involvement of *RAS*-mutated PDX models[55]
	Trametinib (GSK1120212)	Reduces viability and induces apoptosis of *RAS*-MUT *KMT2A*r cell lines and primary infant leukaemic blasts[54] Synergistically enhances prednisolone sensitivity of *RAS*-WT and *RAS*-MUT *KMT2A-AFF1* cell lines[54]	Reduces leukaemic burden in murine xenograft models of *RAS*-mutated *KMT2A-AFF1* SEM cells[56] Currently in human clinical trials (**Table 1.2**)
Menin/KMT2A inhibitor	KO-539	Induces apoptosis at nanomolar concentrations in a range of *KMT2A*r acute leukaemia cell lines, but not in *KMT2A*-WT cell lines[57]	Confers prolonged survival benefit in MV4;11 and MOLM13 PDX models, and *NPM1*-MUT/*DNMT3A*-MUT/*FLT3*-ITD PDX models, compared with FLT3 inhibitor quizartinib[57]

	MI-2 and MI-3	Induces growth arrest and inhibits transformation in murine bone marrow cells transduced with *KMT2A-MLLT3* and *KMT2A-MLLT1*, but not *E2A-HFL*. Also inhibits growth of human *KMT2A*r cell lines KOPN-8, ML-2 and MonoMac6[58]	*Superseded by subsequent analogues such as MI-503*
	MI-463 and MI-503	Inhibits growth and promotes myeloid differentiation of murine bone marrow cells transformed with *KMT2A-MLLT3*, and primary *KMT2A*r AML samples, but not primary *KMT2A*-WT samples[59]	Significantly inhibits tumour growth in MV4;11 PDX models and murine bone marrow cells transformed with *KMT2A-MLLT3*[59]
	VTP-50469	Induces apoptosis in human *KMT2A*r and *NPM1*-MUT acute leukaemia cell lines[60]	Abrogates self-renewal and proliferation of *NPM1*-MUT AML cells in mice[61] Significantly increases EFS of orthotopic infant *KMT2A*r PDX models compared to control mice[62] Prolongs survival of *NUP98*r AML PDX models[63]
Proteasome inhibitor	Carfilzomib (PR-171)	Induces apoptosis at nanomolar concentrations across six *KMT2A*r infant ALL cell lines, and synergises with several conventional chemotherapeutic agents[64]	Fails to provide survival benefit as a single agent or with induction chemotherapy in infant *KMT2A*r PDX models[64]
XPO1 inhibitor	Selinexor (KPT-330)	Nil	Inhibits growth of AML PDX models (normal karyotype, *FLT3*-ITD and two complex karyotype)[65] Currently in human clinical trials (**Table 1.2**)

Abbreviations: AML, acute myeloid leukaemia; CDX, cell line-derived xenografts; EFS, event-free survival; HDAC, histone deacetylase; MUT, mutated; MDS, myelodysplastic syndrome; Nil=no published data available PDX, patient-derived xenografts; PK, pharmacokinetic; PTD, partial tandem duplication; WT, wild-type

Table 1.2 – Current and recently-completed clinical trials for novel targeted therapies in acute leukaemia

Mechanism of action	Agent	Clinical trials in acute leukaemia (ClinicalTrials.gov reference numbers provided)	Number of patients who achieved an objective clinical response*×
AXL inhibitor	ONO-7475	**Phase I/II**: Single agent or plus venetoclax in adults with R/R AML or MDS (NCT03176277, recruiting)	No results available
BCL-2 inhibitor	Venetoclax (ABT-199)	**Phase I**: Paediatric and young adults (up to 21 years) with R/R malignancies (ALL, AML, NHL, neuroblastoma) (NCT03236857, recruiting[66])	No results available
		Phase I: Plus chemotherapy in adults with R/R ALL (NCT3319901, recruiting)	No results available
		Phase I: Plus navitoclax and chemotherapy in patients (≥4 years of age) with R/R ALL (NCT03181126, completed)	**B-ALL:** CR: 64% (16/25), PR: 12% (3/25)[67] **T-ALL:** CR: 53% (10/19)[67]
		Phase I: Plus chemotherapy, rituximab and dasatinib in R/R or newly-diagnosed B-ALL (NCT04872790, not yet recruiting)	No results available
		Phase Ib/II: Plus chemotherapy in adults with R/R ALL (NCT03808610, recruiting)	No results available
		Phase IB/II: Plus vincristine liposomal in adults with R/R ALL (NCT03504644, recruiting)	No results available
		Phase II: Plus azacitidine in adults with AML in the post-HSCT setting (NCT04128501, recruiting)	No results available
Bromodomain and extra-terminal motif (BET) inhibitor	Birabresib (MK-8628/ OTX015)	**Phase I**: Adults with advanced haematologic malignancies (ALL, AML, DLBCL, MM) (NCT01713582, completed)	5% (7/138)[28,29]
		Phase IB: Adults with R/R AML or DLBCL (NCT02698189, terminated)	0% (0/3)
BET inhibitor	FT-1101	**Phase I:** Single-agent or plus azacitidine in adults with R/R AML, MDS or NHL (NCT02543879, completed)	No results available
DOT1L inhibitor	Pinometostat (EPZ5676)	**Phase I**: Paediatric R/R *KMT2A*r leukaemia (NCT02141828, completed)	No responses observed[30]
		Phase I: Adults with R/R *KMT2A*r haematologic malignancies (AML, ALL, MLL, CMML, MPN, MDS) (NCT01684150, completed[68])	3.9% (2/51) + 3 cases had resolution of leukaemia cutis[68]
		Phase Ib/II: Plus azacitidine in adults with newly diagnosed or R/R *KMT2A-MLLT3* AML (NCT03701295, completed)	No results available
		Phase Ib/II: Plus chemotherapy in adolescents and adults (≥14 years of age) with newly diagnosed *KMT2A*r AML (NCT03724084, recruiting)	No results available

Histone deacetylase (HDAC) inhibitor	Panobinostat (LBH-589)	**Phase I:** Children with R/R haematologic malignancies (AML, ALL, NHL) (NCT01321346, completed)	0% (0/22)[31]
		Phase Ib: Plus chemotherapy in adults with R/R AML (NCT01055483, completed)	No results available
		Phase Ib: Plus chemotherapy in adults (≤65 years of age) with newly-diagnosed AML (NCT01242774, completed)	65% (30/46)[69]
		Phase IA/II: Adults with R/R advanced haematologic malignancies (NCT00621244, completed)	2.4% (2/83, dose escalation phase) 5.3% (1/19, expansion phase)
		Phase I/II: Plus chemotherapy in adults (≥65 years of age) with newly-diagnosed AML (NCT00840346, completed)	64% (24/38)[70]
		Phase II: Adults with R/R ALL or AML (NCT00723203, terminated due to lack of efficacy)	0% (0/12)
		Phase III: Panobinostat maintenance following HSCT in patients with high-risk AML or MDS (NCT04326764, active & not recruiting)	No results available
Histone deacetylase (HDAC) inhibitor	Vorinostat	**Pilot study:** Plus decitabine and chemotherapy in patients (age 1-21 years) with relapsed ALL (NCT01483690, terminated due to toxicity)	39% (9/23)[71]
		Phase I: Plus decitabine and cytarabine in adults with R/R AML (NCT01130506, completed)	35% (6/17, including 2/4 with *KMT2A*-PTD)[72]
		Phase I/II: Plus sorafenib and bortezomib in adults with newly-diagnosed or R/R AML (NCT01534260, completed)	40% (6/17)[73]
		Phase I/II: Plus bortezomib and chemotherapy in infant (up to 365 days old) ALL (NCT02553460, active & not recruiting)	No results available
		Phase II: Plus temozolomide in adults with R/R AML (NCT01550224, completed)	0% (0/23)
		Phase II: Plus azacitidine in adults with AML or MDS who are ineligible for other leukaemia protocols (NCT00948064, completed)	40% (12/30[74])
		Phase II: Adults with untreated high-risk or R/R good-risk AML (NCT00305773, completed)	1/37 (2.7%)[32]
		Phase II: Plus velcade in adults with high-risk MDS and R/R AML (NCT00818649, terminated due to toxicity)	8.3% (1/12)
		Phase II: Plus chemotherapy in patients (15-65 years of age) with high-risk MDS or AML (NCT00656617, completed)	76% (78/102)
		Phase II: Plus gemtuzumab in adults (≥60 years) with newly-diagnosed non-M3 AML (NCT00673153, terminated)	19% (6/31[75])

		Phase III: Plus chemotherapy in adults (18-60 years) with previously untreated AML (NCT01802333, completed)	61% (65/107)
MDM2 inhibitor (p53 activator)	Idasanutlin	**Phase IB**: Plus cobimetinib and venetoclax in adults with R/R AML (NCT02670044, completed)	15% CR (3/20), 5% PR (1/20)[76]
		Phase I/II: Plus chemotherapy or venetoclax in paediatric and young adults with R/R acute leukaemia or solid tumours (NCT04029688, recruiting)	No results available
		Phase III: Single agent or plus cytarabine in adults with R/R AML (NCT02545283, terminated due to lack of efficacy based on interim analysis[77])	No results available
MEK inhibitor	Selumetinib	**Phase I/II**: Plus dexamethasone in R/R *RAS*-pathway mutated ALL (NCT03705507, recruiting)	No results available
Menin-KMT2A inhibitor	KO-539	**Phase IA**: Adults with R/R AML (NCT04067336, recruiting)	No results available
Menin-KMT2A inhibitor	SNDX-5613 (close analog of VTP-50469)	**Phase I/II:** Infants, children & adults (aged $\geq$30 days) with R/R acute leukaemia with *KMT2A*r or *NPM1* mutation (NCT04065399, recruiting)	No results available
Menin-KMT2A inhibitor	JNJ-75276617	**Phase II:** Adults with R/R acute leukaemia with *KMT2A* or *NPM1* alterations (NCT04811560, recruiting)	No results available
XPO1 inhibitor	Selinexor (KPT-330)	**Phase I:** Plus venetoclax and chemotherapy in paediatric and young adults with R/R AML (NCT04898894, recruiting)	No results available
		Phase I: Plus chemotherapy in patients with newly diagnosed or R/R AML (NCT02573363, completed)	70% (14/20)[78]
		Phase I: Plus chemotherapy in paediatric and young adults (up to 24 years) with R/R acute leukaemia or MDS (NCT02212561, completed)	CR: 47% (7/15) PR 13% (2/15)[79]
		Phase I: Plus decitabine in adults with R/R AML (NCT02093403, completed)	40% (10/25)[80]
		Phase I/II: Plus sorafenib in adults with R/R *FLT3*-mutated AML (NCT02530476, completed)	(6/14)[81]
		Phase II: Single agent in older adults (aged $\geq$60 years) with R/R AML (NCT02088541, completed)	5.1% (6/118)

*Objective clinical response defined as partial response (PR) or complete response (CR) as assessed by investigator

[x]As per results posted on clinicaltrials.gov, unless otherwise cited

Note: "Adult" defined as $\geq$18 years of age, unless otherwise specified

Abbreviations: DLBCL, diffuse large B-cell lymphoma; HSCT, haematopoietic stem cell transplant; MLL, mixed lineage leukaemia; MPN, myeloproliferative neoplasms; MDS, myelodysplastic syndrome; NHL, Non-Hodgkin's lymphoma; MM, multiple myeloma.

1.7 Summary

Despite recent improvements in overall long-term outcomes for patients with ALL, patients with a *KMT2A*r continue to experience exceptionally poor outcomes. While the biology and clinical implications of *KMT2A*r B-ALL and AML have been well-studied, there is a lack of data regarding *KMT2A*r in a T-lineage specific setting. *In vitro* models of *KMT2A*r in HSCs and progenitor cells typically induce a mixed or myeloid phenotype, so the underlying molecular mechanisms of *KMT2A*r T-ALL aetiology remain unexplored. Further to this, B-ALL and AML cell lines and PDX models are typically utilised in the pre-clinical screening of targeted therapies, so it is also unclear whether *KMT2A*r T-ALL has a similar therapeutic sensitivity profile to B-ALL.

Similarly, the clinical and biological implications of *MLLT10*r in T-ALL are poorly characterised, but it is known that *KMT2A*r and *MLLT10*r share similar biological roles in epigenetic regulation of gene expression, including *HOXA* genes. This key similarity suggests that *MLLT10*r and *KMT2A*r leukaemia may share similarities in sensitivity to targeted therapies, but to date this has not been investigated.

Understanding the biology of *KMT2A*r and *MLLT10*r T-ALL will provide further insight into the lineage-specific mechanisms of leukaemogenic transformation by genomic aberrations that induce *HOXA* deregulation. Furthermore, screening targeted therapies for pre-clinical efficacy using T-ALL-specific *in vitro* models will identify optimal target populations for future clinical trials, both in terms of disease immunophenotype (T-ALL vs B-ALL vs AML) and genomic subtype.

1.7.1 Project aims

1. Generate *in vitro* models of *KMT2A*r B-ALL and T-ALL, and *MLLT10*r T-ALL, to characterise and compare changes in cellular proliferation, differentiation, and activation of downstream signalling pathways.

2. Generate *in vivo* murine models of *DDX3X-MLLT10* and *PICALM-MLLT10* T-ALL to assess the oncogenic driver capacity of *MLLT10*r in a T-ALL setting.

3. Interrogate the mutational landscape of leukaemic blasts from patients diagnosed with non-infant *KMT2A*r and *MLLT10*r ALL, to identify potentially

cooperative aberrations in epigenetic regulators and leukaemia-associated genes.

4. Explore an emerging targeted therapy (curaxin CBL0137) in *KMT2A*r and *KMT2A* wild-type human ALL cell lines.

1.7.2 Hypotheses

1. *KMT2A-AFF1, DDX3X-MLLT10* and *PICALM-MLLT10* expression will induce an aggressive phenotype in MOHITO T-cell line, and *KMT2A-AFF1* in Ba/F3 pro B-cell line, defined by increased proliferation and altered surface marker expression, but not activation of kinase-associated signalling pathways.

2. Immature T-cells (Arf$^{-/-}$ thymocytes) expressing either *DDX3X-MLLT10* or *PICALM-MLLT10* will engraft with accelerated onset in *NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/Sz* (NSG) mice, in comparison with control cells, and induce a distinctive immunophenotype.

3. Leukaemic blasts from non-infant patients with *KMT2A*r ALL will harbour very few cooperative mutations in leukaemia-associated genes and epigenetic regulators, whereas blasts from patients with *MLLT10*r ALL will have known pathogenic co-occurring genomic lesions.

4. CBL0137 will promote apoptosis across a broad range of human acute leukaemia cell lines, but will not induce specific efficacy in *KMT2A*r cell lines.

1.8 Chapter References

1   Pui, C.-H., Robison, L. L. & Look, A. T. Acute lymphoblastic leukaemia. *The Lancet* **371**, 1030-1043, doi:10.1016/s0140-6736(08)60457-2 (2008).

2   Mullighan, C. The molecular genetic makeup of acute lymphoblastic leukaemia. *Haematology* **1**, 389-396 (2012).

3   Pui, C. H. *et al.* Childhood Acute Lymphoblastic Leukemia: Progress Through Collaboration. *J Clin Oncol* **33**, 2938-2948, doi:10.1200/JCO.2014.59.1636 (2015).

4   Terwilliger, T. & Abdul-Hay, M. Acute lymphoblastic leukaemia: a comprehensive review and 2017 update. *Blood Cancer Journal* **7**, doi:doi:10.1038/bcj.2017.53 (2017).

5   Inaba, H., Greaves, M. & Mulligan, C. G. Acute lymphoblastic leukaemia. *The Lancet* **381**, 1943-1955, doi:10.1016/s0140-6736(12)62187-4 (2013).

6   Tasian, S. K., Loh, M. L. & Hunger, S. P. Childhood acute lymphoblastic leukemia: Integrating genomics into therapy. *Cancer* **121**, 3577-3590, doi:10.1002/cncr.29573 (2015).

7   Tasian, S. K. & Hunger, S. P. Genomic characterization of paediatric acute lymphoblastic leukaemia: an opportunity for precision medicine therapeutics. *Br J Haematol* **176**, 867-882, doi:10.1111/bjh.14474 (2017).

8   Roberts, K. & Mullighan, C. Genomics in acute lymphoblastic leukaemia: insights and treatment implications. *Nature Reviews* **12**, 344-356 (2015).

9   Inaba, H. & Mullighan, C. G. Pediatric acute lymphoblastic leukemia. *Haematologica* **105**, 2524-2539, doi:10.3324/haematol.2020.247031 (2020).

10   Tavakoli Shirazi, P. *et al.* The effect of co-occurring lesions on leukaemogenesis and drug response in T-ALL and ETP-ALL. *Br J Cancer* **122**, 455-464, doi:10.1038/s41416-019-0647-7 (2020).

11   Teachey, D. T. & Pui, C.-H. Comparative features and outcomes between paediatric T-cell and B-cell acute lymphoblastic leukaemia. *The Lancet Oncology* **20**, e142-e154, doi:10.1016/s1470-2045(19)30031-2 (2019).

12   Tanasi, I. *et al.* Efficacy of tyrosine kinase inhibitors in Ph-like acute lymphoblastic leukemia harboring ABL-class rearrangements. *Blood* **134**, 1351-1355, doi:10.1182/blood.2019001244 (2019).

13   Boer, J. & Boer, M. d. BCR-ABL1-like acute lymphoblastic leukaemia: From bench to bedside. *European Journal of Cancer* **82**, 203-218 (2017).

14   Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546-1558, doi:10.1126/science.1235122 (2013).

15   Greaves, M. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nature Reviews Cancer*, 471-484, doi:10.1038/s41568-018-0015-6 (2018).

16   Lappin, T.R.J, Grier, D. G., Thompson, A. & Halliday, H. L. HOX GENES: Seductive Science, Mysterious Mechanisms. *Ulster Medical Journal* **75**, 23-31 (2006).

17   Cavé, H. l. n. *et al.* Clinical significance of HOX11L2 expression linked to t(5;14)(q35;q32), of HOX11 expression, and of SIL-TAL fusion in childhood T-cell malignancies: results of EORTC studies 58881 and 58951. *Blood* **103**, 442-450, doi:10.1182/blood-2003-05-1495 (2004).

18   Shah, N. & Sukumar, S. The Hox genes and their roles in oncogenesis. *Nat Rev Cancer* **10**, 361-371, doi:10.1038/nrc2826 (2010).

19   Grabher, C., von Boehmer, H. & Look, A. T. Notch 1 activation in the molecular pathogenesis of T-cell acute lymphoblastic leukaemia. *Nat Rev Cancer* **6**, 347-359, doi:10.1038/nrc1880 (2006).

20   Van Vlierberghe, P., Pieters, R., Beverloo, H. B. & Meijerink, J. P. Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia. *Br J Haematol* **143**, 153-168, doi:10.1111/j.1365-2141.2008.07314.x (2008).

21   Li, B., Huang, Q. & Wei, G. H. The Role of HOX Transcription Factors in Cancer Predisposition and Progression. *Cancers (Basel)* **11**, doi:10.3390/cancers11040528 (2019).

22   Gokbuget, N. & Hoelzer, D. Treatment of adult acute lymphoblastic leukemia. *Semin Hematol* **46**, 64-75, doi:10.1053/j.seminhematol.2008.09.003 (2009).

23   Roberts, K. *et al.* Functional Analysis of Kinase-Activating Fusions in Ph-like Acute Lymphoblastic Leukemia. *Blood* **124**, 786 (2014).

24   Hunger, S. P. & Mullighan, C. G. Acute Lymphoblastic Leukemia in Children. *N Engl J Med* **373**, 1541-1552, doi:10.1056/NEJMra1400972 (2015).

25   Chua, L. L. *et al.* Temporal changes in gut microbiota profile in children with acute lymphoblastic leukemia prior to commencement-, during-, and post-cessation of chemotherapy. *BMC Cancer* **20**, 151, doi:10.1186/s12885-020-6654-5 (2020).

26   Essig, S. *et al.* Risk of late effects of treatment in children newly diagnosed with standard-risk acute lymphoblastic leukaemia: a report from the Childhood Cancer Survivor Study cohort. *The Lancet Oncology* **15**, 841-851 (2014).

27   Kanellopoulos, A. *et al.* Neurocognitive Outcome in Very Long-Term Survivors of Childhood Acute Lymphoblastic Leukemia After Treatment with Chemotherapy Only. *Pediatr Blood Cancer* **63**, 133-138, doi:10.1002/pbc.25690 (2016).

28   Odore, E. *et al.* Phase I Population Pharmacokinetic Assessment of the Oral Bromodomain Inhibitor OTX015 in Patients with Haematologic Malignancies. *Clin Pharmacokinet* **55**, 397-405, doi:10.1007/s40262-015-0327-6 (2016).

29   Amorim, S. *et al.* Bromodomain inhibitor OTX015 in patients with lymphoma or multiple myeloma: a dose-escalation, open-label, pharmacokinetic, phase 1 study. *The Lancet Haematology* **3**, e196-e204, doi:10.1016/s2352-3026(16)00021-1 (2016).

30   Shukla, N. *et al.* Final Report of Phase 1 Study of the DOT1L Inhibitor, Pinometostat (EPZ-5676), in Children with Relapsed or Refractory MLL-r Acute Leukemia. *Blood* **128**, 2780-2780 (2016).

31   Goldberg, J. *et al.* A phase I study of panobinostat in children with relapsed and refractory hematologic malignancies. *Pediatr Hematol Oncol* **37**, 465-474, doi:10.1080/08880018.2020.1752869 (2020).

32   Schaefer, E. W. *et al.* A phase 2 study of vorinostat in acute myeloid leukemia. *Haematologica* **94**, 1375-1382, doi:10.3324/haematol.2009.009217 (2009).

33   Somers, K. *et al.* Potent anti-leukemic activity of curaxin CBL0137 against MLL-rearranged leukemia. *Int J Cancer*, doi:10.1002/ijc.32582 (2019).

34   Lock, R. *et al.* Initial testing (stage 1) of the curaxin CBL0137 by the pediatricpreclinical testing program. *Pediatric Blood & Cancer* **64**, e26263, doi:10.1002/pbc.26263 (2016).

35   Ruvolo, P. P. *et al.* Anexelekto/MER tyrosine kinase inhibitor ONO-7475 arrests growth and kills FMS-like tyrosine kinase 3-internal tandem duplication mutant acute myeloid leukemia cells by diverse mechanisms. *Haematologica* **102**, 2048-2057, doi:10.3324/haematol.2017.168856 (2017).

36   Benito, J. M. *et al.* MLL-Rearranged Acute Lymphoblastic Leukemias Activate BCL-2 through H3K79 Methylation and Are Sensitive to the BCL-2-Specific Antagonist ABT-199. *Cell Rep* **13**, 2715-2727, doi:10.1016/j.celrep.2015.12.003 (2015).

37   Coudé, M. *et al.* BET inhibitor OTX015 targets BRD2 and BRD4 and decreases c-MYC in acute leukemia cells. *Oncotarget* **6**, 17698-17712, doi:10.18632/oncotarget.4131 (2015).

38   Brzezinka, K. *et al.* Functional diversity of inhibitors tackling the differentiation blockage of MLL-rearranged leukemia. *J Hematol Oncol* **12**, 66, doi:10.1186/s13045-019-0749-y (2019).

39   Masse, A. *et al.* BET inhibitors impair leukemic stem cell function only in defined oncogenic subgroups of acute myeloid leukaemias. *Leuk Res* **87**, 106269, doi:10.1016/j.leukres.2019.106269 (2019).

40   Millan, D. S. *et al.* FT-1101: A Structurally Distinct Pan-BET Bromodomain Inhibitor with Activity in Preclinical Models of Hematologic Malignancies. *Blood* **126**, 1367-1367, doi:10.1182/blood.V126.23.1367.1367 (2015).

41   Stumpel, D. J., Schneider, P., van Roon, E. H., Pieters, R. & Stam, R. W. Absence of global hypomethylation in promoter hypermethylated Mixed Lineage Leukaemia-rearranged infant

acute lymphoblastic leukaemia. *Eur J Cancer* **49**, 175-184, doi:10.1016/j.ejca.2012.07.013 (2013).

42 Sarkaria, S. M., Christopher, M. J., Klco, J. M. & Ley, T. J. Primary acute myeloid leukemia cells with IDH1 or IDH2 mutations respond to a DOT1L inhibitor in vitro. *Leukemia* **28**, 2403-2406, doi:10.1038/leu.2014.235 (2014).

43 Chen, L. *et al.* Abrogation of MLL-AF10 and CALM-AF10-mediated transformation through genetic inactivation or pharmacological inhibition of the H3K79 methyltransferase Dot1l. *Leukemia* **27**, 813-822, doi:10.1038/leu.2012.327 (2013).

44 Daigle, S. R. *et al.* Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. *Cancer Cell* **20**, 53-65, doi:10.1016/j.ccr.2011.06.009 (2011).

45 Marschalek, R. MLL Leukemia and Future Treatment Strategies. *Archiv der Pharmazie* **348**, 221-228, doi:10.1002/ardp.201400449 (2015).

46 Klaus, C. R. *et al.* DOT1L inhibitor EPZ-5676 displays synergistic antiproliferative activity in combination with standard of care drugs and hypomethylating agents in MLL-rearranged leukemia cells. *J Pharmacol Exp Ther* **350**, 646-656, doi:10.1124/jpet.114.214577 (2014).

47 Daigle, S. R. *et al.* Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood* **122**, 1017-1025 (2013).

48 Garrido Castro, P. *et al.* The HDAC inhibitor panobinostat (LBH589) exerts in vivo anti-leukaemic activity against MLL-rearranged acute lymphoblastic leukaemia and involves the RNF20/RNF40/WAC-H2B ubiquitination axis. *Leukemia* **32**, 323-331, doi:10.1038/leu.2017.216 (2018).

49 San Jose-Eneriz, E., Gimenez-Camino, N., Agirre, X. & Prosper, F. HDAC Inhibitors in Acute Myeloid Leukemia. *Cancers (Basel)* **11**, doi:10.3390/cancers11111794 (2019).

50 Cao, F. *et al.* Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. *Mol Cell* **53**, 247-261, doi:10.1016/j.molcel.2013.12.001 (2014).

51 Karatas, H. *et al.* Discovery of a Highly Potent, Cell-Permeable Macrocyclic Peptidomimetic (MM-589) Targeting the WD Repeat Domain 5 Protein (WDR5)-Mixed Lineage Leukemia (MLL) Protein-Protein Interaction. *J Med Chem* **60**, 4818-4839, doi:10.1021/acs.jmedchem.6b01796 (2017).

52 Lehmann, C., Friess, T., Birzele, F., Kiialainen, A. & Dangl, M. Superior anti-tumor activity of the MDM2 antagonist idasanutlin and the Bcl-2 inhibitor venetoclax in p53 wild-type acute myeloid leukemia models. *J Hematol Oncol* **9**, 50, doi:10.1186/s13045-016-0280-3 (2016).

53 Konopleva, M. *et al.* MDM2 inhibition: an important step forward in cancer therapy. *Leukemia* **34**, 2858-2874, doi:10.1038/s41375-020-0949-z (2020).

54 Kerstjens, M. *et al.* MEK inhibition is a promising therapeutic strategy for MLL-rearranged infant acute lymphoblastic leukemia patients carrying RAS mutations. *Oncotarget* **8**, 14835-14846, doi:10.18632/oncotarget.11730 (2017).

55 Irving, J. *et al.* Ras pathway mutations are prevalent in relapsed childhood acute lymphoblastic leukemia and confer sensitivity to MEK inhibition. *Blood* **124**, 3420-3430, doi:10.1182/blood-2014-04-531871 (2014).

56 Kerstjens, M. *et al.* Trametinib inhibits RAS mutant MLL-rearranged acute lymphoblastic leukemia at specific niche sites and reduces ERK phosphorylation in vivo. *Haematologica* **103**, e147-e150, doi:10.3324/haematol.2017.174060 (2018).

57 Burrows, F. *et al.* Abstract LB-A27: A novel small molecule menin-MLL inhibitor for potential treatment of MLL-rearranged leukemias and NPM1/DNMT3A-mutant AML. *Molecular Cancer Therapeutics* **17**, LB-A27-LB-A27, doi:10.1158/1535-7163.Targ-17-lb-a27 (2018).

58 Grembecka, J. *et al.* Menin-MLL inhibitors reverse oncogenic activity of MLL fusion proteins in leukemia. *Nat Chem Biol* **8**, 277-284, doi:10.1038/nchembio.773 (2012).

59 Borkin, D. *et al.* Pharmacologic inhibition of the Menin-MLL interaction blocks progression of MLL leukemia in vivo. *Cancer Cell* **27**, 589-602, doi:10.1016/j.ccell.2015.02.016 (2015).

60   Krivtsov, A. V. *et al.* Abstract 4958: VTP50469 is a novel, orally available menin-MLL1 inhibitor effective against MLL-rearranged and NPM1-mutant leukemia. *Cancer Research* **78**, 4958-4958, doi:10.1158/1538-7445.Am2018-4958 (2018).

61   Uckelmann, H. J. *et al.* Therapeutic targeting of preleukemia cells in a mouse model of NPM1 mutant acute myeloid leukemia. *Science* **367**, 586-590, doi:10.1126/science.aax5863 (2020).

62   Lock, R. B. *et al.* Abstract 3187: Pediatric Preclinical Testing Consortium evaluation of the menin inhibitor, VTP-50469, against xenograft models of MLL-rearranged infant acute lymphoblastic leukemia. *Cancer Research* **78**, 3187 (2018).

63   Emily B Heikamp, J. A. H., Florian Perner, Eric M Wong, Charles Hatton, Yanhe Wen, Sonali P Barwe, Anilkumar Gopalakrishnapillai, Haiming Xu, Hannah Julia Uckelmann, Sumiko Takao, Yaniv Kazansky, Yana Pikman, Gerard M. McGeehan, Edward A Kolb, Alex Kentsis, Scott A. Armstrong. The Menin-MLL1 interaction is a molecular dependency in NUP98-rearranged AML. *Blood*, doi:10.1182/blood.2021012806 (2021).

64   Cheung, L. C. *et al.* Preclinical Evaluation of Carfilzomib for Infant KMT2A-Rearranged Acute Lymphoblastic Leukemia. *Front Oncol* **11**, 631594, doi:10.3389/fonc.2021.631594 (2021).

65   Etchin, J. *et al.* Activity of a selective inhibitor of nuclear export, selinexor (KPT-330), against AML-initiating cells engrafted into immunosuppressed NSG mice. *Leukemia* **30**, 190-199, doi:10.1038/leu.2015.194 (2016).

66   Place, A. E. *et al.* Accelerating drug development in pediatric cancer: a novel Phase I study design of venetoclax in relapsed/refractory malignancies. *Future Oncology* **14**, 2115-2129, doi:10.2217/fon-2018-0121 (2018).

67   Pullarkat, V. A. *et al.* Venetoclax and Navitoclax in Combination with Chemotherapy in Patients with Relapsed or Refractory Acute Lymphoblastic Leukemia and Lymphoblastic Lymphoma. *Cancer Discov*, doi:10.1158/2159-8290.CD-20-1465 (2021).

68   Stein, E. M. *et al.* The DOT1L inhibitor pinometostat reduces H3K79 methylation and has modest clinical activity in adult acute leukemia. *Blood* **131**, 2661-2669 (2018).

69   DeAngelo, D. J. *et al.* Safety and efficacy of oral panobinostat plus chemotherapy in patients aged 65 years or younger with high-risk acute myeloid leukemia. *Leuk Res* **85**, 106197, doi:10.1016/j.leukres.2019.106197 (2019).

70   Ocio, E. M. *et al.* Panobinostat as part of induction and maintenance for elderly patients with newly diagnosed acute myeloid leukemia: phase Ib/II panobidara study. *Haematologica* **100**, 1294-1300, doi:10.3324/haematol.2015.129577 (2015).

71   Burke, M. J. *et al.* Decitabine and Vorinostat with Chemotherapy in Relapsed Pediatric Acute Lymphoblastic Leukemia: A TACL Pilot Study. *Clin Cancer Res* **26**, 2297-2307, doi:10.1158/1078-0432.CCR-19-1251 (2020).

72   Mims, A. S. *et al.* A novel regimen for relapsed/refractory adult acute myeloid leukemia using a KMT2A partial tandem duplication targeted therapy: results of phase 1 study NCI 8485. *Haematologica* **103**, 982-987, doi:10.3324/haematol.2017.186890 (2018).

73   Saliba, A. *et al.* Final Results of Phase I/II Study of Combination of Sorafenib, Vorinostat, and Bortezomib in Acute Myeloid Leukemia with FLT3-ITD Mutation or Poor-Risk Cytogenetics. *Blood* **130**, 3897, doi:10.1182/blood.V130.Suppl_1.3897.3897 (2017).

74   Montalban-Bravo, G. *et al.* A clinical trial for patients with acute myeloid leukemia or myelodysplastic syndromes not eligible for standard clinical trials. *Leukemia* **31**, 318-324, doi:10.1038/leu.2016.303 (2017).

75   Walter, R. B. *et al.* Phase II trial of vorinostat and gemtuzumab ozogamicin as induction and post-remission therapy in older adults with previously untreated acute myeloid leukemia. *Haematologica* **97**, 739-742, doi:10.3324/haematol.2011.055822 (2012).

76   Daver, N. *et al.* Preliminary Results from a Phase Ib Study Evaluating BCL-2 Inhibitor Venetoclax in Combination with MEK Inhibitor Cobimetinib or MDM2 Inhibitor Idasanutlin in Patients with Relapsed or Refractory (R/R) AML. *Blood* **130**, 813, doi:10.1182/blood.V130.Suppl_1.813.813 (2017).

77   Montesinos, P. *et al.* MIRROS: a randomized, placebo-controlled, Phase III trial of cytarabine ± idasanutlin in relapsed or refractory acute myeloid leukemia. *Future Oncology* **16**, 807-815, doi:10.2217/fon-2020-0044 (2020).

78   Wang, A. Y. *et al.* A phase I study of selinexor in combination with high-dose cytarabine and mitoxantrone for remission induction in patients with acute myeloid leukemia. *J Hematol Oncol* **11**, 4, doi:10.1186/s13045-017-0550-8 (2018).

79   Alexander, T. B. *et al.* Phase I Study of Selinexor, a Selective Inhibitor of Nuclear Export, in Combination With Fludarabine and Cytarabine, in Pediatric Relapsed or Refractory Acute Leukemia. *J Clin Oncol* **34**, 4094-4101, doi:10.1200/JCO.2016.67.5066 (2016).

80   Bhatnagar, B. *et al.* Selinexor in combination with decitabine in patients with acute myeloid leukemia: results from a phase 1 study. *Leuk Lymphoma* **61**, 387-396, doi:10.1080/10428194.2019.1665664 (2020).

81   Zhang, W. *et al.* Combinatorial targeting of XPO1 and FLT3 exerts synergistic anti-leukemia effects through induction of differentiation and apoptosis in FLT3-mutated acute myeloid leukemias: from concept to clinical trial. *Haematologica* **103**, 1642-1653, doi:10.3324/haematol.2017.185082 (2018).

Chapter 2 – Methods

2.1   General common reagents

Table 2.1 – Common reagents

Category	Reagent	Supplier	Catalogue No.
General Chemicals	Acetic acid	Sigma-Aldrich	#320099
	Ammonium chloride	Sigma-Aldrich	#213330
	Ammonium bicarbonate	Sigma-Aldrich	#09830
	CellTiter-Glo® Luminescent Cell Viability Assay	Promega	G7570
	Citric acid monohydrate	Sigma-Aldrich	C7129
	DC™ protein assay kit	Bio-Rad	#5000111
	Dulbecco's phosphate-buffered saline (PBS) (10X), no calcium, no magnesium	Gibco	#14190235
	EDTA	UNIVAR	AJA180
	Ethanol	Merck Millipore	#10098350000
	Formalin	Sigma-Aldrich	HT501128
	Glycerol	Sigma-Aldrich	SHBD0027V
	Isopropanol	Chem-Supply	#67-63-0
	Methanol	Chem-Supply	MA004-500M
	Nuclease-free water (NF-H_2O)	Sigma-Aldrich	W4502
	Sodium dodecyl sulphate (SDS), 10%	Sigma-Aldrich	L4509
	Sodium chloride (NaCl)	UNIVAR	AJA465
	Tris (hydroxymethyl) aminomethane	Sigma-Aldrich	#252859
	β-mercaptoethanol	Sigma-Aldrich	SHBC3203V
Cell Culture Reagents	Citrate dextrose solution (ACD-A)	Sigma-Aldrich	C3821
	Fetal bovine serum (FBS), triple 0.1 µm sterile filtered	CellSera	AU-FBS/SF
		HyClone	SH30084.03 Lot # DE27192266
	Dimethyl sulfoxide (DMSO)	Merck Millipore	#1029524000
	Doxycycline (10mg/mL)	Sigma-Aldrich	D9891
	Dulbecco's Modified Eagle's Medium (DMEM)	Sigma-Aldrich	D5671
	Murine Flt-3 ligand, carrier free	R&D Systems	#427-FL-025/CF

	L-glutamine 200 mM	Sigma-Aldrich	G7513
	Lymphoprep	STEMCELL Technologies	#07851
	Murine IL-2, recombinant	Peprotech	#212-12
	Murine IL-7, recombinant	Peprotech	#217-17
	Penicillin (500 U/mL) / Streptomycin (5000 µg/mL)	Sigma-Aldrich	P4333
	RPMI-1640 Medium	Sigma-Aldrich	R0883
	Trypan blue 0.4%	Gibco	#15250-061
	Trypsin (2.5%), no phenol red	Gibco	#15090046
Flow Cytometry Assay Reagents	7-Aminoactinomycin D (7AAD)	Invitrogen	A1310
	Annexin-V-PE	BD Pharmingen	#556421
	Bovine serum albumin (BSA)	Sigma-Aldrich	A9418
	Calcium Chloride (CaCl$_2$)	Sigma-Aldrich	C1016
	D-glucose	Sigma-Aldrich	G7528
	Hank's Balanced Salt Solution (HBSS)	Sigma-Aldrich	H9394
	HEPES solution, 1 M, pH 7.0-7.6	Sigma-Aldrich	H0887
	Paraformaldehyde powder, 95%	Fluka Chemica	#76240
Western Blotting Reagents	4-15% Criterion^{TM} Precast Midi Protein Gel, 18 well, 30 µL	Bio-Rad	#5671084
	Aprotinin, 1.7 mg/mL	Sigma-Aldrich	A1153
	β-glycerol phosphate	Sigma-Aldrich	G9422
	cOmplete^{TM} Mini EDTA-Free Protease Inhibitor Cocktail Tablets	Roche	#04693159001
	Glycine	Sigma-Aldrich	G8898
	Intercept Blocking Buffer	LI-COR	#197-70001
	Leupeptin, 1 mg/mL	Sigma-Aldrich	#11017128001
	NP40, 1% (Igepal^{TM})	Sigma-Aldrich	CA-630
	Precision Plus Protein^{TM} Kaleidoscope	Bio-Rad	#161-0375
	Phenylmethanesulfonyl fluoride, ≥98.5% (PMSF)	Sigma-Aldrich	P7626-250MG
	Immun-Blot® Low Fluorescence PVDF, 0.45 µm	Bio-Rad	#1620262

	Sodium fluoride	Sigma-Aldrich	L4509
	Sodium pyrophosphate (NaPPi)	Sigma-Aldrich	P8010
	Sodium orthovanadate (Na$_3$VO$_4$)	Sigma-Aldrich	S6508
	Trans-Blot® Turbo^{TM} RTA Midi LF PVDF Transfer Kit	Bio-Rad	#1704275
	Tween-20	Sigma-Aldrich	T8532
Genomics Reagents	10x Blue Juice (gel loading buffer)	Invitrogen	#10816-015
	1kb DNA ladder	New England Biolabs	N3232S
	Agarose	Sigma-Aldrich	A6013 – 500g
	BigDye^{TM} Terminator v3.1 Cycle Sequencing Kit	Thermo Fisher Scientific	#4337455
	BigDye XTERMINATOR^{TM} Purification Kit	Thermo Fisher Scientific	#4376486
	Chloroform	Merck Millipore	#100776B
	Deoxynucleotide set, 100 mM	Sigma-Aldrich	DNTP100-1KT
	DNA hydration solution	Qiagen	#158914
	Gel Loading Dye, Purple (6X), no SDS	New England Biolabs	B7024S
	GelRed^{TM} Nucleic Acid Stain	Biotium	#41002
	Glycogen, Molecular Biology Grade	Roche	#10 901 393 001
	UltraClean® 15 DNA Purification Kit	MO BIO Laboratories Inc.	#12100-300
	Nuclease-free water (NF-H$_2$O)	Qiagen	#129117
	Phusion High-Fidelity DNA polymerase	New England Biolabs	M0530
	Proteinase K	Roche	#03115887001
	Q5® High-Fidelity DNA polymerase	New England Biolabs	M0491
	QuantiTect Reverse Transcription Kit	Qiagen	#205313
	RNAse A	Qiagen	#19101
	RT^2 SYBR® Green qPCR Mastermix	Qiagen	#330503
	Silica Bead DNA Gel Extraction Kit	Thermo Scientific	K0513

	TRIzol® reagent	Thermo Fisher Scientific	#15596-018
	Truseq Stranded mRNA LT kit	Illumina	#20020595
	UltraPure™ Buffer Saturated Phenol	Invitrogen	#15593-031
	UltraPure™ Phenol:Chloroform:Isoamyl alcohol (25:24:1, v/v)	Invitrogen	#15593-031
Inhibitors	CBL0137	Selleckchem	S8483
Molecular Cloning Reagents	Ampicillin	Sigma-Aldrich	A5354
	BamHI-HF	New England Biolabs	R3136S
	BsmBI	New England Biolabs	R0580S
	CutSmart® Buffer (10x)	New England Biolabs	B7204S
	EcoRI-HF	New England Biolabs	R3101S
	Kanamycin Sulphate	Gibco	#15160-054
	Luria Broth Base	Invitrogen	#12795-084
	NEBuffer 2 (10x)	New England Biolabs	B7002S
	NEBuffer 3.1 (10x)	New England Biolabs	B7203S
	One Shot™ MAX Efficiency™ DH5α™-T1^R chemically competent *E. Coli*	Thermo Fisher Scientific	#12297016
	QIAprep Spin Miniprep Kit	Qiagen	#27106
	QIAGEN Plasmid Midi Kit	Qiagen	#12124
	Select Agar™	Invitrogen	#30391-023
	Shrimp Alkaline Phosphatase (rSAP)	New England Biolabs	M0371S
	S.O.C. Medium	Thermo Fisher Scientific	#15544034
	T4 DNA Ligase	New England Biolabs	M0202
	T4 DNA Ligase Reaction Buffer (10X)	New England Biolabs	B0202S
	T4 Polynucleotide Kinase	New England Biolabs	M0201S

	T7 Endonuclease I	New England Biolabs	M0302S
	XhoI	New England Biolabs	R0146S
Viral Transfection & Transduction Reagents	Lipofectamine 2000	Thermo Fisher Scientific	#11668500
	Opti-MEM™ Reduced Serum Medium, GlutaMAX™ Supplement	Thermo Fisher Scientific	#51985034
	Polybrene®	Santa Cruz Biotechnology	sc-134220
	RetroNectin® GMP grade	Takara Bio USA	T202

Table 2.2 – List of phospho-flow antibodies

Antibody	Isotype	Catalogue no.	Supplier
Akt (pS473)	IgG1, κ	#560378	BD Pharmingen
Erk (pT202/pY204)	IgG1	#612566	BD Pharmingen
IgG1-PE	IgG1, κ	#559320	BD Pharmingen
S6 (pS235/pS236)-PE	IgG$_1$, κ	#560433	BD Pharmingen
Stat5 (pY694)-PE	IgG$_1$, κ	#562077	BD Pharmingen
Syk (pY348)-PE	IgG$_1$, κ	#558529	BD Pharmingen
Zap70 (Y319)-PE	IgG$_1$, κ	#557881	BD Pharmingen

Note: 20 µL of stain was used per test.

Table 2.3 – List of murine surface marker flow cytometry antibodies

Antibody	Clone	Catalogue no.	Supplier
Alexa Fluor® 647 Rat Anti-Mouse CD5	53-7.3	#100613	Biolegend
Alexa Fluor® 647 Rat Anti-Mouse IgG2a, κ	RTK2758	#400526	Biolegend
Alexa Fluor® 700 Rat Anti-Mouse CD34	RAM34	#560518	BD Pharmingen
Alexa Fluor® 700 Rat Anti-Mouse CD8a	53-6.7	#100730	Biolegend
Alexa Fluor® 700 Rat Anti-Mouse IgG2a, κ	RTK2758	#400528	Biolegend
Alexa Fluor® 700 Rat IgG2a, κ Isotype Ctrl	RTK2758	#400528	Biolegend
APC Rat Anti-Mouse CD24	M1-69	#562349	BD Pharmingen
APC Rat Anti-Mouse CD25	PC61	#102012	Biolegend
APC Rat IgG1, λ Isotype Ctrl	G0114F7	#401904	Biolegend
APC/Cy7 Rat Anti-Mouse CD4	GK1.5	#100414	Biolegend
APC/Cy7 Rat IgG2b, κ Isotype Ctrl	RTK4530	#400624	Biolegend
BV650 Rat Anti-Mouse CD13	R3-242	#740488	BD Pharmingen
BV650 Rat Anti-Mouse IgG1, κ	R3-34	#563848	BD Pharmingen
BV786 Mouse Anti-Mouse CD249 (BP-1)	BP-1	#740882	BD Pharmingen
PE Rat Anti-Mouse B220	RA3-6B2	#103208	Biolegend
PE Rat Anti-Mouse CD2	RM2-5	#100108	Biolegend
PE Rat Anti-Mouse CD4	H129.19	#553653	BD Pharmingen
PE Rat Anti-Mouse CD8a	53-6.7	#553033	BD Pharmingen
PE Rat Anti-Mouse TER-119/Erythroid Cells	TER-119 (RUO)	#553673	BD Pharmingen
PE Rat IgG2b, λ Isotype Ctrl	G013B8	#403804	Biolegend
PE-CF594 Rat Anti-Mouse CD19	HIB19	#562291	BD Pharmingen
PE-CF594 Rat Anti-Mouse IgG2a, κ	R35-95	#562302	BD Pharmingen

PE-CF594 Rat Anti-Mouse Sca-1 (Ly-6a)	D7	#562730	BD Pharmingen
PE-Cy7 Rat Anti-Mouse CD117 (c-Kit)	2B8	#561681	BD Pharmingen
PE-Cy7 Rat Anti-Mouse CD43	S7	#562866	BD Pharmingen
PE-Cy7 Rat Anti-Mouse IgG2b, κ	R35-95	#552784	BD Pharmingen
PE/Cy7™ Rat Anti-Mouse CD44	IM7	#103030	Biolegend
PE/Cy7™ Rat IgG2b, κ Isotype Ctrl	RTK4530	#400618	Biolegend
PerCP/Cy5.5 Mouse Anti-Mouse CD45.2	104	#109828	Biolegend
PerCP/Cy5.5 Mouse IgG2a, κ Isotype Ctrl	MOPC-173	#400258	Biolegend
V450 Rat Anti-Mouse CD3 Molecular Complex	17A2	#561389	BD Pharmingen
V450 Rat Anti-Mouse IgG2b, κ	A95-1	#560457	BD Pharmingen

Table 2.4 – List of primary western blotting antibodies

Target	Clone	Molecular Weight (kDa)	Source	Dilution	Catalogue #	Supplier
GAPDH	14C10	37	Rabbit	1:1000	#2118	Cell Signaling Technology
p21	12D1	21	Rabbit	1:1000	#2947	Cell Signaling Technology
p53	1C12	53	Mouse	1:1000	#2524	Cell Signaling Technology

Note: All antibodies are monoclonal.

Table 2.5 – List of secondary western blotting antibodies

Target	Antibody	Source	Dilution	Cat #	Supplier
Goat IgG	IRDye® 800CW Donkey anti-Goat IgG	Donkey	1:15,000	#926-32214	LI-COR
Rabbit IgG	IRDye® 800CW Donkey anti-Rabbit IgG	Donkey	1:15,000	#925-32213	LI-COR

2.2 Cell culture media & solutions

All cell culture media were stored at 4°C and equilibrated at room temperature (RT) prior to use, unless otherwise stated.

Arf$^{-/-}$ thymocyte cell culture media

MEM α, medium, with nucleosides

1mM sodium pyruvate

20% Hyclone foetal bovine serum (FBS)

50 units/mL Penicillin

50 µg/mL Streptomycin

2mM L-glutamine

5 ng/mL Flt3 ligand (Flt3L)

5 ng/mL murine IL-7 (mIL-7)

Ba/F3 IL-3 independent, RS4;11 & WEHI-3B cell culture media

RPMI-1640 medium

2 mM L-glutamine

50 units/mL Penicillin

50 µg/mL Streptomycin

10% FBS

Ba/F3 parental wildtype cell culture media

RPMI-1640 medium

2 mM L-glutamine

50 units/mL Penicillin

50 µg/mL Streptomycin

10% FBS

5% WEHI-3B complete media (source of IL-3)

D10 media/HEK-293T cell culture media

Dulbecco's Modified Eagle's Medium (DMEM)

2 mM L-glutamine

50 units/mL Penicillin

50 µg/mL Streptomycin

10% FBS

MOHITO cell culture media

RPMI-1640 medium

2 mM L-glutamine

50 units/mL Penicillin

50 µg/mL Streptomycin

20% FBS

10 ng/mL murine IL-7

5 ng/mL murine IL-2

MOHITO cytokine-free culture media

RPMI-1640 medium

2 mM L-glutamine

50 units/mL Penicillin

50 µg/mL Streptomycin

20% FBS

OP9-DL1 cell culture media

MEM α medium, with nucleosides

1mM sodium pyruvate

10% FBS

50 units/mL Penicillin

50 µg/mL Streptomycin

2mM L-glutamine

Starve RPMI-1640 media

RPMI-1640 medium

2 mM L-glutamine

50 units/mL Penicillin

50 µg/mL Streptomycin

2% FBS

Cryoprotectant freezing solution

7:2:1 ratio of complete cell culture media:FBS:DMSO

2.3 Common buffers & solutions for flow cytometry analysis

All solutions stored at 4°C unless specified.

1x Binding buffer for Annexin-V / 7-AAD cell viability assay

500 mL HBSS

10 mM HEPES

5 mM $CaCl_2$

Solution was used ice-cold for each application.

1x Binding buffer for phospho-flow assay

1x PBS

1% BSA

Solution was prepared and filtered with a 0.2 µM sterile filter.

Methanol permeabilisation solution for phospho-flow assay

80% analytical grade methanol

20% Milli-Q® H_2O

Solution was stored at -20°C and used ice-cold for each application.

Flow cytometry fixative (FACS fix)

1x PBS

1% formaldehyde

111mM D-glucose

0.02% sodium azide

Cell blocking buffer

2% BSA

2% FCS

1x PBS

2.4   Common buffers & solutions for protein assays

4x Laemmli's reducing loading buffer

0.25 M Tris-HCl pH 6.8

8% SDS

40% glycerol

0.04% β-mercaptoethanol

0.05% bromophenol blue

RO H_2O

Aliquotted into screw-capped 1.5 mL tubes and stored at -20°C.

NP40 lysis buffer

10 mM Tris-HCl pH 7.4

137 mM NaCl

10% Glycerol

1% NP-40 (Igepal^{TM})

Sterile Milli-Q® H_2O

The above buffer was made in bulk (10 mL) and stored at 4°C. The following reagents were added to the concentrations indicated to a 1mL aliquot of the above immediately prior to each use.

10 mM β-glycerol phosphate

2 mM Sodium vanadate

2 mM Sodium fluoride

2 mM PMSF

10 mM Sodium pyrophosphate

1 µg/mL Leupeptin

5 µg/mL Aprotonin

1 x cOmplete^{TM} mini EDTA-free protease inhibitors cocktail tablet/10ml NP40 lysis buffer

1x Phosphate-buffered saline (1x PBS)

10x PBS stock

Diluted 1:10 in sterile Milli-Q® ultra-purified water

Solution was prepared and stored at 4°C.

10x SDS-PAGE gel running buffer

250 mM Tris

1.92 M Glycine

1% SDS

RO water

Solution was prepared and stored at RT.

1x solutions were prepared in RO water from 10x stocks and stored at RT.

10x Tris-buffered saline (10x TBS)

200 mM Tris-HCl (pH 7.5)

1.5 M NaCl

RO water

Solution was prepared and stored at RT.

1x solutions were prepared in RO water from 10x stocks and stored at RT.

1x Tris-buffered saline + 0.1% Tween-20 (1x TBST)

200 mM Tris-HCl (pH 7.5)

1.5 M NaCl

0.1% Tween-20

RO water

Solution was prepared and stored at RT.

10x Western blotting transfer buffer

250 mM Tris

1.92 mM Glycine

RO water

Solution was prepared and stored at 4°C.

1x solutions were prepared from 10x stocks in RO water with 20% methanol final and stored at 4°C.

2.5   Common molecular cloning buffers & reagents

1x TAE agarose gel for DNA electrophoresis

1-2% agarose powder

0.01% GelRed^{TM} nucleic acid stain

1x TAE buffer

Agarose powder was dissolved in 1x TAE by boiling in a microwave oven. Solution was cooled to approximately 40°C prior to the addition of GelRed^{TM} (Biotium, California, USA) nucleic acid stain and used for gel electrophoresis. Remaining gel was stored at RT for up to 1 month and was re-liquified prior to use in a microwave oven.

DNA lysis buffer

10 mM Tris

10 mM sodium chloride

10 mM EDTA

dNTP set (N=A, T, C, G)

25 mM (40 µL) of each dNTP

10 mM: 1:2.5 dilution of 25 mM stock in 80 µL DEPC water

5 mM: 1:5 dilution of 25 mM stock in 80 µL DEPC water

Solution was prepared and stored at -20°C.

Luria broth (LB)

7.5 g LB base

To 300 mL with Milli-Q® water

Solution was sterilised (autoclaved) and stored at RT.

LB agar

7.5 g LB base

4.5 g Select Agar^{TM}

To 300 mL with Milli-Q® water

Solution was sterilised (autoclaved) and stored at RT. Solidified agar stocks were heated to 100°C in a microwave oven and cooled to approximately 37°C prior to the

addition of antibiotics, depending on the antibiotic-resistance profile of plasmids (100 μg/μL ampicillin or 50 μg/μL kanamycin). LB agar solution was aseptically transferred into sterile petri dishes and allowed to solidify at RT. LB agar plates were transferred to a sterile biosafety cabinet and lids removed to enable condensation to evaporate. LB agar plates were then stored at 4°C for up to two months.

50x Tris-acetate EDTA (50x TAE) buffer

2 M Tris-HCl (pH 7.6)

1 M acetic acid

50 mM EDTA

Milli-Q® water

Solution was prepared and stored at RT.

1x TAE buffer

1x solutions were prepared from 50x stock in RO water and stored at RT.

2.6 Common buffers & solutions for *in vivo* and *ex vivo* murine experiments

Red cell lysing buffer (RCLB)

0.144 M NH_4Cl (ammonium chloride)

0.144 M NH_4HCO3 (ammonium bicarbonate)

Milli-Q® water

Solution was prepared and stored at RT.

Thymocyte thaw media

HBSS

1% Hepes

5% FBS

5% anticoagulant citrate dextrose solution (ACD-A)

Solution was prepared fresh prior to use and pre-warmed in a 37°C water bath.

White cell fluid (WCF)

2% glacial acetic acid

Milli-Q® water

Methyl violet crystals

Solution was prepared and stored at RT.

2.7   CBL0137 small molecule preparation and storage

CBL0137 was purchased as powdered stock from Selleckchem (Houston, USA). Powdered stock was dissolved in sterile diluent (DMSO) to 10 mM and stored long-term in 20 µL aliquots at -80°C. Immediately prior to use, serial dilutions were made from stored stock solutions in sterile diluent for use in cell death assays.

MW: 372.89 g/mol.

2.8   Ethics statement

Genomic material from patients diagnosed with ALL were used within this project. Participants were required to provide written informed consent for their samples to be used in accordance with the Declaration of Helsinki. The nature of the study was explained, where participants were informed of the intended use of samples for prospective medical research. The use of samples was approved by the Central Adelaide Local Health Network and the Royal Adelaide Hospital Human Research Ethics Committee (HREC/16/MH/251, 15/09/2017 and HREC/15/RAH/54, 25/02/2015).

2.9   Nucleic acid protocols

2.9.1   Mononuclear cell (MNC) isolation

Mononuclear cells (MNCs) were isolated from patient blood and bone marrow samples using Lymphoprep (STEMCELL Technologies, Vancouver, Canada) density gradient centrifugation according to manufacturer's instructions.

2.9.2   RNA extraction of TRIzol® samples

Approximately 5×10^6 cells were collected in 2 mL microcentrifuge tubes and pelleted at 550× g for 5 min. Supernatant was removed and cells were resuspended in 1 mL of TRIzol® reagent (Thermo Fisher Scientific, Massachusetts, USA) and incubated at RT for 5 min. If not immediately proceeding to RNA extraction, samples were stored at -80°C, and then thawed at RT immediately prior to RNA extraction. The suspension was then vigorously mixed with 200 µL of 100% chloroform (Sigma-Aldrich, Missouri, USA). All following centrifugation steps were performed at 4°C. The suspension was centrifuged at 12,000× g for 15 min at 4°C, and the upper aqueous phase retained in a fresh 1.7 mL microcentrifuge tube (Corning Inc., New York, USA). RNA was precipitated by mixing with 500 µL 100% isopropanol and 1 µL 20 µg/mL glycogen (Thermo Fisher Scientific). Samples were incubated for 10 min at RT, then centrifuged at 12,000× g for 10 min. The supernatant was discarded and RNA pellet washed with 1 mL 70% ethanol (Sigma-Aldrich), and centrifuged at 7,500× g for 5 min. Supernatant was removed and the pellet re-centrifuged to remove any residual ethanol. The pellet was air-dried for up to 10 min and resuspended in 30 µL of DEPC-treated NF-H_2O. Samples were incubated at 55°C for 10 min to ensure that RNA was completely dissolved, before quantification using a Nanodrop 8000 Spectrophotometer (Thermo Fisher Scientific). Samples were stored long-term at -80°C.

2.9.3   Complementary DNA (cDNA) synthesis

The QuantiTect Reverse Transcription Kit (Qiagen, Hilden, Germany) was used to synthesise cDNA from 1 µg of RNA. In a single reaction, 2 µL of 7x gDNA Wipeout Buffer was added with NF-H_2O and RNA to 14 µL final volume in a 0.2 mL PCR tube (Axygen®, California, USA). Samples were incubated for 2 min at 42°C to ensure gDNA elimination.

A volume of 6 µL master mix comprising 4 µL 5x Quantiscript buffer, 1 µL primer mix and 1 µL Quantiscript reverse transcriptase was added into each sample. Using an MJ Research DNA Engine thermal cycler, samples were incubated for 30 min at 42°C, then 3 min at 95°C. The resulting cDNA were stored at -20°C until required.

2.9.4   Polymerase chain reaction (PCR)

All PCR master mixes were set-up in 1.7 mL microcentrifuge tubes in a template-free PCR room before aliquoting into 0.2 µL PCR tubes or 8-well strips (Axygen®). depending on the number of samples, on an Isofreeze® tube cooler rack (Alpha Laboratories, Mount Wellington, New Zealand). Template was added in a designated UV cabinet. No template control (NTC) reactions with NF-H_2O were included to ensure absence of non-specific amplification products. PCR cycling reactions were performed on a T100 thermal cycler (Bio-Rad, Hercules, USA).

2.9.5   Quantitative reverse transcription PCR (qRT-PCR)

Set-up conditions are as described for other PCR reactions (refer to **2.9.4**), except using a MicroAmp^{TM} Optical 96-Well Reaction Plate (Thermo Fisher Scientific). Actin was used as an endogenous control for all reactions. qRT-PCR cycling reactions were performed on a QuantStudio 7 Real-Time PCR system (Thermo Fisher Scientific).

2.9.6   Sanger sequencing of plasmids and PCR products

Samples were either submitted for Sanger sequencing at the Australian Genome Research Facility (AGRF), or performed in-house using the BigDye^{TM} Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific) and the SeqStudio^{TM} Genetic Analyser System (Thermo Fisher Scientific). DNA chromatogram results were analysed using Benchling® online software.

Samples submitted to AGRF for Sanger sequencing were prepared in 1.7 mL microcentrifuge tubes, consisting of 2 µL of 5 µM stock per primer and 600-800 ng vector DNA per reaction, to a final volume of 14 µL with NF-H_2O.

Samples sequenced in-house using the SeqStudio^{TM} Genetic Analyser System were prepared according to the protocol specified in **Tables 2.6-2.7**. Samples were prepared in a MicroAmp^{TM} Optical 96-Well Reaction Plate (Thermo Fisher Scientific), or MicroAmp^{TM} Optical 8-Cap Strips (Thermo Fisher Scientific) depending on the number of samples.

Table 2.6 – Reagents and protocol for Sanger sequencing using BigDye^{TM} Terminator v3.1 Cycle Sequencing Kit

Purpose	Reagents	Volume/quantity	Total reaction volume (µL)
Perform cycle sequencing	BigDye^{TM} Terminator v3.1 Ready Reaction Mix	1 µL	20 µL Perform cycle sequencing in thermal cycler (**Table 2.7**)
	BigDye^{TM} Terminator v3.1 5x Sequencing Buffer	3.5 µL	
	Primer (5 µM)	0.65 µL (3.2 pmol)	
	DNA template	Varies: See **Table 2.8**	
	NF-H$_2$O	To 20 µL	
Purify sequencing reactions	SAM^{TM} Solution	90 µL per sequencing reaction	110 µL Apply reagents directly to cycle sequencing reaction
	BigDye XTerminator^{TM} bead solution	20 µL per sequencing reaction	

Table 2.7 – Reaction conditions for Sanger sequencing using BigDye^{TM} Terminator v3.1 Cycle Sequencing Kit

Purpose	Reaction conditions
Perform cycle sequencing	Cycle sequencing performed in a T100 thermal cycler (Bio-Rad) 96°C for 1 min **With a ramp rate of 1°C/sec:** 96°C for 10 sec, 50°C for 5 sec, 60°C for 4 min, repeat for total of 25 cycles 4°C for ∞
Purify sequencing reactions	Purification solution was prepared using the BigDye XTerminator^{TM} purification kit (SAM^{TM} Solution + bead solution, **Table 2.6**) was prepared, and 110 µL applied to each sequencing reaction. In a VWR® Advanced Microplate Vortex Mixer (VWR International, Pennsylvania, USA), samples were vortexed for 30 min at 2,000× g. In a swinging bucket centrifuge, samples were centrifuged at 1,000× g for 2 min.
Perform Sanger sequencing with SeqStudio^{TM} Genetic Analyser	Samples were inserted into the instrument and sequencing protocol carried out according to manufacturer's instructions[1]

Table 2.8 – Guide for template quantities used in Sanger sequencing using SeqStudio^{TM}

DNA template type and size	Quantity used for Sanger sequencing
Plasmid, double-stranded (any size)	750-850 ng
Plasmid, single-stranded (any size)	100-250 ng
PCR product: 100-200 bp	2-6 ng
PCR product: 200-400 bp	4-10 ng
PCR product: 400-600 bp	10- 16 ng
PCR product: 600-800 bp	20-30 ng
PCR product: >800 bp	30-75 ng

2.9.7 Genomic DNA (gDNA) extraction

For each sample, gDNA was extracted from 5 x 10^6 cells (MNCs) or 2 x 10^6 cells (cell lines). Cells were pelleted at 10,000× g for 5 min in a 1.7 mL microcentrifuge tube. The supernatant was aspirated, and the cell pellet frozen at -80°C as a 'whole-cell pellet (WCP)' until DNA extraction. WCPs were thawed and washed in 1 mL 1x PBS at 10,000× g for 5 min. Supernatant was aspirated and cells were resuspended in 480 µL DNA lysis buffer (**section 2.5**), then 12.5 µL 20% sodium dodecyl sulphate (SDS) (Thermo Fisher Scientific) and 10 µL 10 mg/mL proteinase K (Roche) was added, and the samples mixed by inversion. Samples were incubated overnight at 37°C or 2 h at 55°C.

Following incubation, 5 µL 100 mg/mL RNAse A (Qiagen, Hilden, Germany) was added to each tube, inverted approximately 50 times and incubated at 37°C for 10 min. Ten microlitres of 5 M sodium chloride and 1 µL 20 mg/mL glycogen carrier (Roche) were subsequently added to each tube. In a fume hood, 500 µL buffer saturated phenol (Invitrogen, California, USA) was added and tubes were mixed by vigorous shaking, achieving a homogenous milky solution. Samples were then centrifuged at 16,000× g for 5 min and the upper aqueous layer was retained in a fresh 1.7 mL microcentrifuge tube, to which 500 µL UltraPure™ phenol:chloroform:isoamyl alcohol (25:24:1, v/v) (Invitrogen) was added and mixed by vigorous shaking. Tubes were centrifuged 16,000× g for 5 min, and the aqueous phase retained in a fresh 1.7 mL tube. The DNA extract was washed with 1 mL of ice-cold 100% ethanol, mixed by inversion and DNA was then pelleted by centrifuging at 20,000× g for 10 min. The supernatant was discarded, and DNA pellet washed with 700-800 µL of 70% ethanol and then re-pelleted by centrifuging at 20,000× g for 5 min. The ethanol supernatant was discarded, and pellet centrifuged at 16,000× g for 1 min to remove residual ethanol. DNA pellets were air-dried for up to 10 min, resuspended in up to 200 µL DNA hydration solution (Qiagen), and incubated at 55°C for 2 h, and then 37°C overnight. DNA concentration was quantified using the NanoDrop 8000 Spectrophotometer (Thermo Fisher Scientific), and samples stored at -80°C.

2.9.8   DNA gel electrophoresis

Genomic products were visualised via DNA gel electrophoresis on a 1-2% agarose gel/1x TAE containing GelRed for 60 min at 100V, alongside 1 kb or 100 bp molecular weight ladders (New England Biolabs, Ipswich, USA) unless otherwise specified. Bands were visualised on the GelDoc™ XR+ (Bio-Rad) gel documentation system, ensuring correct sizes and purity of samples.

2.10   General cell culture and flow cytometry protocols

2.10.1   Sub-culturing of cell lines

Cells were assessed by Trypan blue exclusion (1:1 dilution) with a Neubauer haematocytometer (Marienfeld Superior GMBH & Co. KG, Lauda-Königshofen, Germany), and density and percentage viability quantified by microscopic visualisation under 10x magnification. In a biosafety class II cabinet, cells were sub-cultured in RT media (refer to **2.2**) into either T25, T75 or T150 culture flasks (Corning Inc.). Cells were sub-cultured at their recommended densities, depending on cell doubling time and desired final volume (**Table 2.9**).

Table 2.9 – Recommended seeding densities for *in vitro* cell lines

Cell line	Cell type	Species	Source	Culture properties	Recommended Live Cell Seeding Density for 3 days in culture
Ba/F3 IL-3 dependent	Pro-B cells	Mouse	DSMZ	Suspension	2×10^4 / mL
Ba/F3 IL-3 independent	Pro-B cells	Mouse	DSMZ	Suspension	1×10^4 / mL
HEK-293T	Embryonic kidney cells	Human	ATCC®	Adherent	1×10^4 / cm^2
OP9-DL1	Bone marrow stromal cells	Mouse	Sunnybrook Research Institute	Adherent	2-3 x 10^5 per 10 cm dish 1.5-2 x 10^4 per well in 12-well plate
MOHITO	CD4+CD8+ T-cells	Mouse	St Jude	Suspension	1-2 x 10^5 / mL

RS4;11	t(4;11) B lymphoblasts	Human	ATCC®	Suspension	2×10^5 / mL
NALM-19	B lymphoblasts	Human	DSMZ	Suspension	8×10^5 / mL
MV4;11	t(4;11) myeloblasts	Human	ATCC®	Suspension	1×10^5 / mL
RCH-ACV	t(1;19) B lymphoblasts	Human	DSMZ	Suspension	1×10^5 / mL
REH	t(12;21) B lymphoblasts	Human	DSMZ	Suspension	2×10^5 / mL
Jurkat	T lymphoblasts	Human	ATCC®	Suspension	6×10^4 / mL
U-937	t(10;11) myeloblasts	Human	ATCC®	Suspension	6×10^4 / mL
THP-1	t(9;11) myeloblasts	Human	ATCC®	Suspension	5×10^5 / mL

2.10.2 Cryopreservation of cells

Cells (approximately 5×10^6 to 2×10^7 cells per ampoule) were collected into a 50 mL centrifuge tube and pelleted at 1,400 rpm for 5 min at RT. For Ba/F3 and human cell lines, supernatant was discarded and cells resuspended in 700 µL of culture media per ampoule. A volume of 300 µL cryoprotectant freezing solution per ampoule (**2.2**) was added dropwise onto the cell pellet, with gentle agitation. For the MOHITO cell line and primary Arf$^{-/-}$ cells, supernatant was discarded and 1 mL of 90% FCS/10% DMSO per ampoule was added in a dropwise fashion, with gentle agitation.

The suspension was immediately transferred into 1.5 mL cryogenic ampoules (Nalgene Nunc International, New York, USA) and allowed to cool in a RT Mr. Frosty Freezing container (Nalgene Nunc International) overnight, at a rate of -1°C per minute in a -80°C ultracool freezer. Ampoules were then transferred for long-term storage into liquid nitrogen (-196°C).

2.10.1 Fluorescence-activated cell sorting of transduced cells

Fluorescence-activated cell sorting (FACS) was used to isolate transduced cell populations with stable fluorescent reporter gene expression (**Table 2.10**). Unsorted cells were centrifuged at 1,400 rpm for 5 min and resuspended in 1 mL RPMI-1640/1%

FCS to an approximate concentration of 5×10^6 cells/mL for sorting. GFP positive (GFP$^+$) gate was set based on the fluorescence of untransduced cells, and GFP$^+$ cells were collected in RPMI-1640 containing 20% FCS and sorted to >90% purity using a BD Biosciences (New Jersey, USA) FACSMelody^{TM} or FACSAria Fusion^{TM} flow cytometry cell sorter. Immediately following collection, cells were centrifuged for 5 min at 1,400 rpm. Supernatant was removed and cells were cultured at a suitable density (**Table 2.9**) in appropriate cell culture media.

Table 2.10 – Fluorescent reporter gene expression for expression vectors in transduced cell lines

Expression vector	Fluorescent reporter gene
pRUFiG2	Green fluorescent protein (GFP)
FgH1tUTG	GFP
FUCas9mCherry	mCherry

2.10.2 Immunophenotyping

Approximately 2×10^6 cells per sample were collected in a 5 mL FACS tube and centrifuged at 2,700 rpm for 2 min. Cells were washed with 1 mL RT 1x PBS and centrifuged at 2,700 rpm for 2 min. Supernatant was discarded and cells were blocked with 100-200 µL blocking buffer (**2.3**) on ice for 15 min. Meanwhile, antibody master mixes were prepared in 50 µL of 2% FCS/1x PBS per test. Antibody volume per test was prepared according to **Table 2.11** for MOHITO cell lines and primary Arf$^{-/-}$ cells, or **Table 2.12** for Ba/F3 cell lines. Antibodies were multiplexed where possible. Where multiplexing was performed, compensation was adjusted based on single stained and unstained controls. Following 15 min blocking, samples were washed in 1 mL RT 1x PBS and supernatant discarded. Samples were resuspended in 2% FCS/1x PBS (50 µL per test), and 50 µL aliquoted into each 5 mL FACS tube required, corresponding to the number of different antibody master mixes. Antibody master mixes (50 µL per sample) were applied to each sample, vortexed and incubated on ice in the dark for 30 min. Following incubation, cells were washed in 1 mL 1x PBS and resuspended in 100-200 µL FACS fix, and stored at 4°C until analysis on a BD FACSCanto^{TM}, FACSSymphony^{TM} or FACSFortessa^{TM} (BS Biosciences, USA).

Table 2.11 – Volume of antibodies for multicolour surface marker flow cytometry of MOHITO T-cell lines and Arf$^{-/-}$ thymocytes

Tube #1			
Antibody	**Fluorochrome**	**Isotype control**	**Volume per test**
CD3 molecular complex (sCD3)	V450	Rat IgG2b, κ	1 µL
CD4	APC/Cy7	Rat IgG2b, κ	5 µL
CD8a	Alexa Fluor® 700	Rat IgG2a, κ	1 µL
CD44	PE-Cy7	Rat IgG2b, κ	1 µL
CD25	APC	Rat IgG1, λ	0.5 µL
CD45.2	PerCP/Cy5.5	Mouse IgG2a, κ	1 µL
CD2	PE	Rat IgG2b, λ	1 µL
Tube #2			
Sca-1 (Ly-6A)	PE-CF594	Rat IgG2a, κ	1 µL
CD117 (c-Kit)	PE-Cy7	Rat IgG2b, κ	1 µL
CD34	Alexa Fluor® 700	Rat IgG2a, κ	1 µL
CD13	BV650	Rat IgG1, κ	1 µL
CD5	Alexa Fluor® 647	Rat IgG2a, κ	2 µL
All samples			
DAPI	DAPI	N/A	1 µL

Table 2.12 – Volume of antibodies for multicolour surface marker flow cytometry of Ba/F3 pro-B cell lines

Tube #1			
Antibody	**Fluorochrome**	**Isotype control**	**Volume per test (50 µL 2% FCS/1x PBS)**
CD45.2	PerCP/Cy5.5	Mouse IgG2a, κ	1 µL
CD19	PE-CF594	Mouse IgG1, κ	2 µL
CD43	PE-Cy7	Rat IgG2a, κ	2 µL
CD249	BV786	Mouse IgG1, κ	2 µL
CD25	APC	Rat IgG1, λ	0.5 µL
Tube #2			
Sca-1 (Ly-6A)	PE-CF594	Rat IgG2a, κ	2 µL
CD117 (c-Kit)	PE-Cy7	Rat IgG2b, κ	1 µL
CD34	Alexa Fluor® 700	Rat IgG2a, κ	1 µL
CD13	BV650	Rat IgG1, κ	1 µL
Tube #3			
B220	PE	Mouse IgG2a, κ	2 µL
CD24	APC	Rat IgG2b, κ	2 µL
All samples			
DAPI	DAPI	N/A	1 µL

2.11 CellTiter-Glo® Luminescent Cell Viability Assay

Cells were centrifuged at 1,400 rpm for 5 min and washed three times in 1x PBS. Cells were incubated in RPMI-1640 starve media (**2.2**) for 1 h at 37°C. Following starvation, cells were centrifuged at 1,400 rpm for 5 min, supernatant aspirated, and resuspended in either complete media or cytokine-free media for the respective cell line (Ba/F3 or MOHITO) at a concentration of 2×10^4 cells/mL (MOHITO) or 1×10^3 cells/mL (Ba/F3), and transferred into duplicate wells of a 12-well plate. Cells were incubated at 37°C/5% CO_2 for 5 days. Each day, including at day 0, each well was agitated with a p1000 pipette to ensure a homogenous cell suspension. Immediately following agitation, 25 µL from each well was transferred to a black bottom 96-well plate, and 20 µL of CellTiter-Glo® reagent was aliquoted into each well. The plate was vortexed for 1 min

and incubated in the dark for 15 min prior to quantifying luminescence with the PerkinElmer 2030 VICTOR X4 multilabel plate reader at 650-750 nm.

2.12 Phospho-flow Assay

Approximately 3×10^6 cells per sample were centrifuged at 1,400 rpm for 5 min and washed three times in 1x PBS. Washed cells were incubated in RPMI-1640 starve media (**2.2**) for 5 h at $37^{\circ}C/5\%$ CO_2. Cells were collected in a 15 mL centrifuge tube, and 100 µL of 16% paraformaldehyde (PFA) was added per millilitre of sample. Cells were incubated in a $37^{\circ}C$ humidified incubator (5% CO_2) for 10 min, to allow fixation of intracellular proteins. Cells were then centrifuged at 1,400 rpm for 5 min at RT. Supernatant was aspirated and cells were washed in 1 mL RT 1x PBS and pelleted as previously described. Supernatant was discarded and 80% ice-cold methanol was added in a drop-wise fashion, while agitating at 600 rpm to resuspend and permeabilise of cells. Samples were then stored at $-20^{\circ}C$ overnight.

The following day, cells were pelleted by centrifugation at 1,400 rpm for 5 min at RT, and the supernatant discarded. Cells were washed in 2 mL 1x PBS, then 1x phospho-flow buffer (**2.3**). Antibodies were prepared according to **Table 2.13**. Cells were briefly vortexed and samples incubated in the dark for 1 h at RT. Following incubation, cells were washed in 1 mL 1x PBS, pelleted and resuspended in 200 µL of 1x PBS.

Table 2.13 – Volume of antibodies for phospho-flow staining

Parameters	Antibody	Volume per 50 µL of 1x binding buffer, in each tube
Unstained control	NA	NA
Isotype control antibodies	IgG1-PE	2 µL
1° conjugated antibodies for phospho-proteins of interest	Stat5 (Y694)-PE Akt (S473)-PE Erk1/2 (pT202/pY204) S6 (pS235/pS236)-PE Syk (pY348)-PE Zap70 (Y319)-PE	20 µL

Note: To minimise pipetting error, stock solutions (n+1) of each antibody plus binding buffer were prepared, prior to delivery of staining solution into each tube.

2.13   Protein electrophoresis

2.13.1   Preparation of protein lysates

Approximately 1×10^7 cells were pelleted in a 10 mL centrifuge tube by centrifugation at 1,400 rpm for 5 min at RT. Cells were washed with 1 mL ice-cold 1x PBS and transferred to a pre-chilled 1.7 mL microcentrifuge tube, and centrifuged at 6,800 rpm for 5 min at 4°C. Supernatant was discarded and 60 µL NP40 lysis buffer with freshly added inhibitors (**2.4**) was added and mixed by gentle pipetting. Samples were incubated on ice for 10 min and centrifuged at 13,000 rpm for 10 min at 4°C. Lysate was transferred to a new pre-chilled 1.7 mL microcentrifuge tube, avoiding cell debris, and 1/3 volume of 4x Laemmli's loading buffer (4x LB) (**2.4**) was added to each lysate and mixed by gentle pipetting and pulse vortexing. Tubes containing remaining cell debris were retained for protein assay to determine lysate concentration (**2.13.2**). Lysates were boiled in a 100°C heat block for 10 min with weight applied to the lids. Boiled lysates were clarified by brief centrifugation prior to storage at -20°C.

2.13.2   Protein quantitation

Protein assays were performed using the Bio-Rad DC Assay kit according to manufacturer's instructions. Three microlitres of lysate was diluted 1:5 with water, and BSA protein standards at concentrations of 0, 0.2, 0.4, 0.8, 1.2, 1.6 and 2.0 µg/µL in water were analysed. Standards and diluted lysate samples were aliquoted into wells of a 96 well plate, in 5 µL duplicates. Twenty-five microlitres of Reagent A (stock solution of 20 µL Reagent S per 1 mL of Reagent A) and 200 µL of Reagent B were added into each well with a multi-channel pipette, and the plate was mixed by gentle vortexing. Samples were incubated in the dark for 15 min at RT, and protein concentration determined via the PerkinElmer 2030 VICTOR X4 multilabel plate reader at 650-750 nm. A standard curve was generated using the formula ($y = mx + c$) such that y is the optical density reading, m is the gradient generated in the formula, x is the concentration of BSA (µg/µL) and c refers to the constant. Ensuring that the R^2 value of this standard curve was as close to 1 as possible, the protein concentration in samples were then calculated according to this formula.

2.13.3   SDS-PAGE protein electrophoresis

Bio-Rad Criterion^{TM} pre-cast gels were assembled in Bio-Rad Criterion^{TM} protein electrophoresis apparatus. 1x SDS-PAGE running buffer was added into both inner and outer tanks, and wells of pre-cast gels were flushed with buffer. A volume of 10 µL of Bio-Rad Kaleidoscope protein ladder was added into the first lane, and 40-60 µg of lysate (thawed on ice, boiled to 100°C for 5 min and pulse centrifuged) was diluted to a final volume of 20-30 µL in 4x Laemmli's buffer and administered into each subsequent lane. An electrical force of 100V was applied for 20 min to stack proteins, and then voltage was increased to 200V for approximately 45 min.

2.13.4   Western blotting

Following separation of proteins via SDS-PAGE, proteins were transferred onto PDF membrane (Bio-Rad) using the Trans-Blot® Turbo^{TM} semi-dry western apparatus, using the 'Mixed setting' (1.3 A / 25 V for 7 min). Upon completion of protein transfer, membranes were cut enabling simultaneous incubation of multiple blots for differing proteins of interest. All following membrane washes and incubations were performed on a rocking platform (Ratek Instruments Pty Ltd., Australia) with gentle agitation unless otherwise specified. Membranes were blocked with Intercept blocking buffer (LI-COR, Lincoln, USA) for 60 min at RT. Following overnight incubation at 4°C with primary antibodies (**Table 2.4**), the membranes were washed 3 x with 1x TBST for 5 min per wash. Blots were then incubated in their respective secondary antibodies (**Table 2.5**) for 2 h at RT in the dark. Membranes were washed 3 x with 1x TBST for 5 min per wash and visualised on the LI-COR Odyssey® fluorescent scanner.

2.13.5   Densitometry analysis

Western blots were visualised and analysed using LI-COR Image Studio^{TM} Lite or Empiria Studio software. The intensity of each band was recorded, and the corresponding background value measured and subtracted. Intensities of phosphorylated proteins were compared to their total counterparts, and the ratio (phospho-protein:total-protein) obtained. Values for samples were graphed to quantify changes observed.

2.14   Annexin-V-PE / 7-AAD exclusion assay

2.14.1   Drug treatment of cells for Annexin-V-PE / 7-AAD cell-death assay

Serial dilutions of inhibitor stocks were freshly prepared in complete cell culture media, based on a pre-determined range of concentrations. For inhibitors requiring DMSO as a solvent, stocks were prepared such that the final concentration of DMSO in cell culture was not more than 0.3%. Following analysis of cell viability and density using Trypan blue exclusion, cells were prepared at a final concentration of the recommended seeding density for 3 days in culture (**Table 2.9**) in complete cell culture media. Cells were exposed to inhibitor concentrations in a round-bottom 96-well tissue culture plate (Corning Inc.) in technical duplicates. Cell culture plates were incubated at 37°C in a humidified incubator (5% CO_2) for 72 h prior to determination of viability by flow cytometry.

2.14.2   Annexin-V-PE / 7-AAD exclusion assay staining protocol

Following 72 h drug exposure, 96-well tissue culture plates were centrifuged at 1,400 rpm for 5 min at RT. Tissue culture media was removed by swift inversion, followed by gently blotting onto a multi-fold towel to completely remove supernatant. Cells were washed in ice-cold binding buffer (**2.3**), vortexed briefly and centrifuged at 1,400 rpm for 5 min at RT, and the supernatant discarded as previously described. Annexin-V-PE / 7-AAD staining solution was prepared in bulk and administered to samples according to **Table 2.14**. Annexin-V and 7-AAD stains were first separately prepared in bulk and added to control tubes, before single stains were combined to form the dual stain applied to experimental wells.

Table 2.14 – Volumes of reagents for Annexin-V-PE / 7-AAD staining solution

Parameter	Annexin-V-PE	7-AAD	1x Binding Buffer
Unstained control	-	-	20 µL
Annexin-V single-stained control	0.4 µL	-	20 µL
7-AAD single-stained control	-	0.04 µL	20 µL
Experimental sample (per well)	0.4 µL	0.04 µL	20 µL

Cells were incubated in staining solution for 20 min on ice in the dark, to allow binding of Annexin-V-PE and 7-AAD to apoptotic cells. Following staining, 200 µL of binding buffer was added into each well and analysed immediately. Control samples (unstained and single-stained) were analysed on the BD FACSCanto™ II flow cytometer to define forward and side-scatter parameters, and Annexin-V-PE / 7-AAD baseline fluorescence intensity gates. Experimental samples were acquired via the BD FACSCanto™ II high-throughput system (HTS), with compensation applied and data batch-analysed using FlowJo® software (FlowJo, LLC, USA).

2.15   Handling of patient specimens and bioinformatic data

2.15.1   Patient information

Leukaemic blasts from ALL patients with available mRNA seq data were assessed for the presence of fusion genes and single nucleotide variants (SNVs). Patient samples were included in analysis if they met all criteria described in **Table 2.15**.

Table 2.15 – Patient inclusion parameters for mRNA seq data analysis

Parameter	Criteria
Age	>1 year of age at time of sample collection
Disease status	Confirmed diagnosis of ALL Newly diagnosed
Genomic subtype	Detection of a fusion gene involving *KMT2A, MLLT10*, or *BCR-ABL1*

2.15.2 Preparation, MLPA and mRNA sequencing of patient samples

All protocols involving the handling of patient materials were performed by Dr. Barbara McClure, Dr. Susan Heatley and Ms. Caitlin Schutz. Peripheral blood and/or bone marrow samples were collected from patients with ALL at diagnosis or relapse. MNCs were isolated and RNA extracted according to sections **2.9.1**-**2.9.3**. MNCs were transcriptionally sequenced and analysed for single nucleotide variants (SNVs), insertion and deletion mutations (INDELs) and fusion genes. mRNA seq was performed using the Truseq Stranded mRNA LT kit (Illumina, California, USA) from 1 µg total RNA, and the Illumina NextSeq 500 platform. Paired-end reads of 75 bp length with an average read depth of 65 million reads was obtained. Copy number alterations (CNAs) were detected via Multiplex Ligation-dependent Probe Amplification (MLPA), Probemix product names P202, P335 and P383.

2.15.3 Filtering of mRNA-seq data

Fusion transcripts were identified from mRNA seq data with FusionCatcher (fusioncatcher.py 0.99.7c beta)[2], JAFFA (v1.09)[3], and SOAPfuse (v1.26)[4], then outputs combined. Only fusions identified by at least two callers were assessed further, and events were ranked by total number of supporting reads.

SNVs and INDELs were called using GATK HaplotypeCaller (v3.4.46)[5] and annotated by ANNOVAR software (2015-036-17)[6]. SNVs were further filtered using R Studio software. All codes are available on Github (https://github.com/michelleforgi/ALL_filtering.git). Previously published gene lists were used to interrogate for SNVs in leukaemia-associated genes[7,8] and epigenetic regulators[9]. SNVs were included in analysis if they met all of the criteria specified in **Table 2.16**.

Table 2.16 – mRNA-seq SNV filtering parameters

Parameter	Criteria
Deleterious prediction	Bioinformatically predicted as deleterious by both SIFT and PolyPhen2
Rarity of variant	Present in <5% of the normal 'healthy' population according to ExAC_all
Minimum reads	>20 total reads across reference and alternate allele

2.15.4 Manual interrogation of SNVs and INDELs within leukaemia-associated genes

Following filtering of data as per **Table 2.16**, variants were manually interrogated using the COSMIC[10] and VarSome[11] databases. Manual interrogation of INDELs was performed by Dr. Laura Eadie. SNVs were included in analysis if they were classified as pathogenic or likely pathogenic in VarSome[11] (ACMG score[12]) or COSMIC[10] (FATHMM[13] score). Variants of uncertain significance were also included in the analysis of SNVs affecting epigenetic regulators, to ensure the inclusion of potentially relevant novel variants. INDELs were included in analysis if they were classified as pathogenic or likely pathogenic in VarSome (ACMG score), or if there is published evidence of the INDEL in T-ALL.

2.16 Generation of Ba/F3 and MOHITO stable cell lines

The *KMT2A-AFF1* fusion gene was commercially synthesised in the pRUFiG2 mammalian retroviral expression vector backbone (**Figure 2.2B**)[14] by GenScript®. *DDX3X-MLLT10* and *PICALM-MLLT10* were identified from patient T-ALL bone marrow samples received at the Leukaemia Research Group, SAHMRI, and were amplified and cloned from patient cDNA into the pCR®-Blunt II-TOPO® construct (**Figure 2.2A**), and then sub-cloned into the pRUFiG2 expression vector (**Figure 2.2B**). HEK-293T cells were transfected with fusion retroviral expression constructs to generate retrovirus, followed by transduction into Ba/F3 parental cells (courtesy of Professor Andrew Zannettino and Dr. Stephen Fitter, The University of Adelaide) for *KMT2A-AFF1*, or MOHITO[15] parental cells (courtesy of Charles Mullighan, St Jude Children's Research Hospital) for *KMT2A-AFF1*, *DDX3X-MLLT10* and *PICALM-MLLT10*. Transduced GFP⁺ cells were selected by FACS, followed by assays to test cytokine independence and proliferation. A detailed description of these experimental processes are provided in sections **2.16.1**-**2.16.18** and a schematic diagram is shown in **Figure 2.1**.

1. PCR of *DDX3X-MLLT10* & *PICALM-MLLT10* from patient cDNA using high-fidelity polymerase

↓

2. Blunt-end cloning of fusions in pCR-Blunt II-TOPO® for amplification of DNA

↓

3. Amplification of construct by transformation in chemically competent DH5α *E. Coli*

↓

4. Insertion of fusion into pRUFiG2 mammalian expression vector

↓ → *KMT2A-AFF1*-pRUFiG2 commercially synthesised

5. Amplification of construct by transformation in chemically competent DH5α *E. Coli*

↓

6. Generation of viral particles by co-transfection of pRUFiG2 expression vector and packaging vector pEqECO in HEK-293T cells

↓

7. Transduction of Ba/F3 and MOHITO host cells with retrovirus

↓

8. Fluorescence-activated cell sorting of GFP-expressing cells to establish stable fusion-expressing cell lines

Figure 2.1 – Generation of Ba/F3 and MOHITO ALL stable cell lines via viral transduction

A. pCR™-Blunt II-TOPO®

B. pRufiG2 (pRufIRESGFP-loxP²)
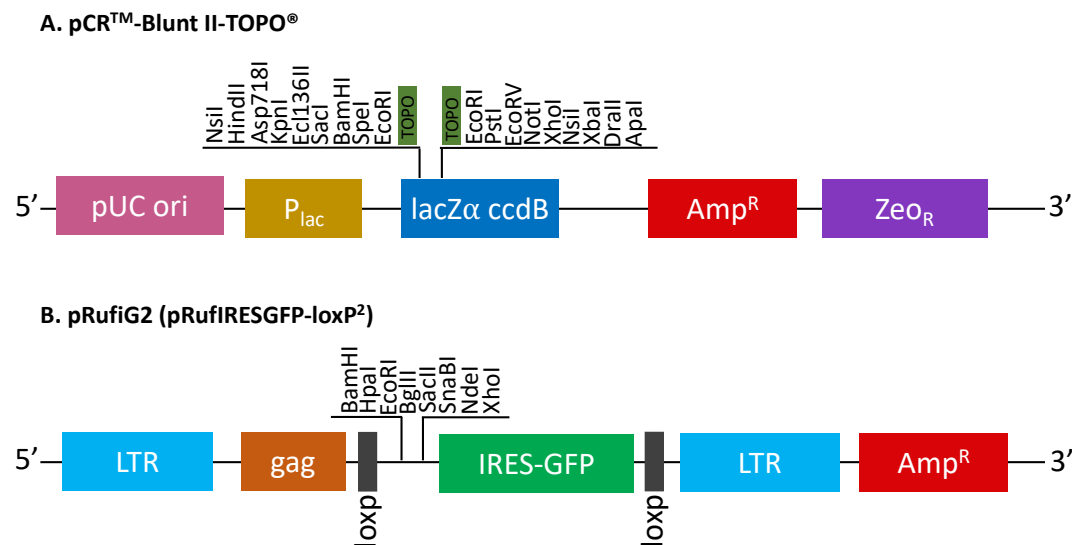
Figure 2.2 – Vector maps for (A) pCR®-Blunt II-TOPO and (B) pRUFiG2

2.16.1 Patient background

DDX3X-MLLT10 and *PICALM-MLLT10* were identified from independent T-ALL patient bone marrow samples received at the Leukaemia Research Group, SAHMRI, following transcriptomic sequencing and bioinformatic analysis using Fusion catcher, JAFFA and deFuse fusion prediction software. Full predicted fusion sequences from NCBI are provided in **Appendices 2.1-2.3**.

2.16.2 PCR amplification of fusions

All PCR were performed using 0.8 µL cDNA (**2.9.3**) diluted 1:2 with nuclease-free water, and performed in a Bio-Rad T100 thermal cycler unless otherwise specified. Primer sequences used for the amplification of *DDX3X-MLLT10* and *PICALM-MLLT10* are provided in **Table 2.17**, and PCR protocols are provided in **Tables 2.18-2.19.**

Table 2.17 – Primer sequences for the amplification of full-length *DDX3X-MLLT10* and *PICALM-MLLT10* from patient cDNA

Gene	Primer sequences (5' to 3')
DDX3X-MLLT10	Forward primer (*DDX3X* 5' UTR): TTAGCAGCGGAAGACTCCGA
	Reverse primer (*MLLT10* 3' UTR): CTGCAAAGGCAGCCAGATGA
PICALM-MLLT10	Forward primer (*PICALM* 5' UTR): CCAATGCAAAGAAGGTCTTG
	Reverse primer (*MLLT10* 3' UTR): TCCTATCATGCCTGTTGGTG

Table 2.18 – Volume of components for PCR amplification of *DDX3X-MLLT10*

Component	Final concentration	Volume for 1 sample (µL)
5X Phusion HF Buffer	1x	4 µL
10 mM dNTP	200 µM	0.4 µL
5 µM Forward primer	400 nM	0.8 µL
5 µM Reverse primer	400 nM	0.8 µL
Template DNA	Variable	0.8 µL
Phusion® High-Fidelity DNA polymerase	0.4 unit	0.2 µL

Nuclease-free water	-	To 20 µL
Final volume		20 µL

PCR cycling conditions for amplification of *DDX3X-MLLT10*

98ºC for 1 min

98ºC for 10 sec, 66ºC for 15 sec, 72ºC for 1 min: 34 cycles total

72ºC for 10 min

4ºC for ∞

Table 2.19 – Volume of components for PCR amplification of *PICALM-MLLT10*

Component	Final concentration	Volume for 1 sample (µL)
5 µM Forward primer	400 nM	0.8 µL
5 µM Reverse primer	400 nM	0.8 µL
Template DNA	Variable	0.8 µL
Platinum SuperFi Polymerase MasterMix	1x	12.5 µL
Nuclease-free water	-	To 20 µL
Final volume	-	**20 µL**

Cycling conditions for amplification of *PICALM-MLLT10*

98ºC for 30 sec

98ºC for 10 sec, 63ºC for 15 sec and 72ºC for 3 min: 34 cycles total

72ºC for 10 min

4ºC for ∞

2.16.3   DNA Gel Purification

Following amplification by PCR and confirmation of a single product size via DNA gel electrophoresis, the entire volume of each PCR were electrophoresed, and the band of the correct size excised using a scalpel and LED Transilluminator (Maestrogen, Hsinchu City, Taiwan). Gels were not imaged prior to excision to prevent UV-induced DNA damage. DNA was purified by gel extracted using the UltraClean® 15 DNA Purification Kit (MO BIO, San Diego, USA) according to manufacturer's instructions, with final elution volumes between 14 to 20 µL.

2.16.4 Molecular cloning of fusion genes in pCR-Blunt II-TOPO® vector

The Zero Blunt® TOPO® PCR Cloning Kit was used in the ligation of *DDX3X-MLLT10* and *PICALM-MLLT10* with the pCR-Blunt II-TOPO® vector (**Table 2.20**). Reactions were prepared on ice and incubated overnight at 4°C.

Table 2.20 – Volume of components for ligation of fusions with pCR-Blunt II-TOPO® vector

Component	Final volume
Purified fusion DNA (*DDX3X-MLLT10* or *PICALM-MLLT10*) or NF-H$_2$O (empty vector control)	4 µL
pCR-Blunt II-TOPO® vector	1 µL
TOPO® Salt Solution	1 µL

2.16.5 Transformation of competent cells for amplification of T-ALL fusion gene pCR-Blunt II-TOPO® plasmids

DDX3X-MLLT10 and *PICALM-MLLT10* constructs were transformed into 50 µL of One Shot^{TM} MAX Efficiency^{TM} DH5α^{TM}-T1^R chemically competent *E. Coli* via heat shock, as per manufacturer's instructions (Thermo Fisher Scientific). Briefly, competent cells were thawed on ice, followed by addition of the entire volume of ligation reactions (6 µL). Competent cells were gently agitated by flicking to mix cells with plasmid DNA, and incubated a further on ice for 30 min.

Following incubation, cells were then subjected to a 2 min 42°C 'heat shock' followed by immediate incubation on ice for 2 min. A volume of 250 µL RT S.O.C. medium (Thermo Fisher Scientific) added, and samples incubated at 37°C in a bacterial orbital shaker (Ratek Instruments Pty Ltd.) for 60 min. The entire reaction volume was then spread onto pre-warmed agar plates containing 50 µg/mL kanamycin and incubated overnight at 37°C. Twelve isolated single colonies each from the *DDX3X-MLLT10* and *PICALM-MLLT10* plates were sampled using a 200 µL sterile pipette tip and individually placed in 2 mL LB broth containing 50 µg/µL kanamycin, in 10 mL polypropylene tubes, and cultured overnight at 37°C with shaking in a dry bacterial incubator.

2.16.6   Extraction and purification of plasmid DNA

Plasmid DNA was extracted from liquid bacterial cultures using QIAprep® Spin Miniprep or Midiprep kits (Qiagen), according to manufacturer's instructions. Routinely, 2 mL cultures were used for Miniprep and 100 mL for Midiprep. Key steps included resuspending bacterial cells in an isotonic buffer containing 1 µg/µL RNAse A, followed by cell lysis using a solution of 0.2 M NaOH and 1% SDS, and neutralisation using a solution of 1 M potassium acetate and 10 M acetic acid. For Midipreps, plasmid DNA was precipitated using isopropanol (Merck Millipore, Burlington, USA), washed in 70% ethanol and rehydrated in nuclease-free water (NF-H_2O).

2.16.7   Screening of isolates for fusion gene ligation in pCR®-Blunt II-TOPO®

Following plasmid DNA extraction from bacterial cultures, restriction enzyme digest strategies were designed to screen for successful ligation of fusion genes in the pCR®-Blunt II-TOPO® vector. Fusion gene sequences were constructed from full-length wild-type gene sequences acquired from the National Center for Biotechnology Information (NCBI), and NEBCutter V2.0 software[16] was used to design a restriction digest strategy to screen plasmid DNA for directional cloning. NEBCutter identifies the location of restriction sites within DNA sequences, and fragment sizes for digested plasmids containing correctly and incorrectly orientated inserts were subsequently predicted. Restriction digest screening strategies are described in **Table 2.21**. All samples were incubated at 37°C for 90 min, then electrophoresed for analysis (**2.9.8**).

Table 2.21 – Strategies for screening of *DDX3X-MLLT10* and *PICALM-MLLT10* ligation in pCR®-Blunt II-TOPO® vector

Component	Volume
Purified plasmid DNA from bacterial isolates	1 µL
10x CutSmart® Buffer (NEB)	1 µL
Restriction enzyme XbaI (*DDX3X-MLLT10*) or ApaI (*PICALM-MLLT10*)	0.5 µL
NF-H$_2$O	7.5 µL
Final volume	**10 µL**

2.16.8   Sanger sequencing of plasmid DNA to confirm fusion gene sequence

Following identification of an isolate containing directionally cloned fusion gene within the pCR®-Blunt II-TOPO® vector, Sanger sequencing was employed to verify correct fusion gene sequence, using primers specific to the fusion and vector sequence (**Table 2.22**).

Table 2.22 – Primer sequences for the Sanger sequencing of *DDX3X-MLLT10*-pTOPO and *PICALM-MLLT10*-pTOPO

Fusion gene	Primers used for Sanger sequencing
DDX3X-MLLT10	Forward primer (M13F, pCR®-Blunt II-TOPO®): GTAAAACGACGGCCAG
	Forward primer (MLLT10 exon 15): TGGGCATTTGCAACAAGTAG
	Forward primer (MLLT10 exon 19): CAGCTTTCAGTGCCTTTTCC
	Reverse primer (M13R, pCR®-Blunt II-TOPO®): CAGGAAACAGCTATGAC
	Reverse primer (MLLT10 exon 24): CTGCAAAGGCAGCCAGATGA
PICALM-MLLT10	Forward primer (M13F, pCR®-Blunt II-TOPO®): GTAAAACGACGGCCAG
	Reverse primer (M13R, pCR®-Blunt II-TOPO®): CAGGAAACAGCTATGAC
	Forward primer (MLLT10 exon 19) CAGCTTTCAGTGCCTTTTCC
	Forward primer (MLLT10 exon 7): AAAGAGCAAACGGGGATCTAA
	Forward primer (MLLT10 exon 10): CAGACCCAAAGGAAACAAAAA
	Forward primer (MLLT10 exon 18): GAGAGGCAGTGGAGTGAAGG
	Forward primer (PICALM exon 6): CCAATGCAAAGAAGGTCTTG
	Reverse primer (PICALM exon 17): TCCTATCATGCCTGTTGGTG
	Reverse primer (MLLT10 exon 10): GGTGAAGCAGAAGAAACTGAGAG
	Reverse primer: (MLLT10 exon 17): GAGACGGACTAGGATTTGCTG
	Reverse primer: (MLLT10 exon 20): GAGGGCTCTTACTGCTGTTCA

2.16.9 Large-scale liquid culture preparation and plasmid DNA extraction

Following identification of an isolate confirmed to contain the desired fusion-vector construct, a small volume of remaining liquid culture from the corresponding correct isolate was streaked for single colonies onto agar containing 50 µg/mL kanamycin (for colonies transformed with pCR®-Blunt II-TOPO®) or 100 µg/mL ampicillin (for colonies transformed with pRUFiG2). After overnight incubation, liquid starter cultures were prepared by inoculating an individual colony from either plate in 5 mL LB containing the appropriate antibiotic, and rocking horizontally for 7 h at 37°C. The entire volume of this starter culture was then inoculated in 100 mL of sterile LB containing the appropriate antibiotic (50 µg/mL kanamycin or 100 µg/mL ampicillin), and cultured shaking overnight at 37°C. The Qiagen Plasmid Midi Kit was used to extract plasmid from these cultures and resulting DNA concentrations were quantified using a NanoDrop 8000 Spectrophotometer.

2.16.10 Long-term storage of transformed bacterial cells

Upon establishment of a confluent bacterial culture, glycerol stocks were made by diluting 700 μL of liquid culture with an equal volume of 100% sterile glycerol to a final volume of 1.4 mL, and samples were stored in 2 mL screw-capped vials (Nalgene Nunc International) at -80°C.

2.16.11 Cloning into pRUFiG2 retroviral expression vector

Once fusion genes were ligated in pCR®-Blunt II-TOPO® constructs, sub-cloning into a mammalian-compatible retroviral vector was required, as the pCR®-Blunt II-TOPO® constructs are non-expressing in mammalian cells. Fusions were sub-cloned into the pRUFiG2 vector, a mammalian-expressing retroviral vector courtesy of Professor Andrew Zannettino and Dr. Stephen Fitter (**Table 2.24, Figure 2.2**).

Table 2.23 – Strategies for sub-cloning of *MLLT10* fusion genes

Sub-cloning strategy	Fusion gene	
	DDX3X-MLLT10	*PICALM-MLLT10*
Preparation of pCR®-Blunt II-TOPO® donor construct	Perform restriction digest with EcoRI Gel purify with UltraClean® DNA purification kit (MoBio Inc., USA) and confirm purity of insert using agarose gel electrophoresis	Perform restriction digest with BamHI and XhoI Gel purify and confirm purity of insert using agarose gel electrophoresis
Preparation of pRUFiG2 recipient backbone	Perform restriction digest with EcoRI or BamHI and XhoI Dephosphorylate 5' and 3' ends with rSAP Gel purify and confirm purity using agarose gel electrophoresis	

The purified isolated fusion gene insert (from pCR®-Blunt II-TOPO® construct) and dephosphorylated pRUFiG2 vector backbone were ligated using T4 DNA ligase (New England Biolabs).

Ligation reactions consisted of 1 μL T4 DNA ligase (NEB), 1 μL 1x ligation buffer (NEB), 4 μL of purified *DDX3X-MLLT10* insert and 3 μL pRUFiG2, and were incubated overnight at 4°C. A dephosphorylated, digested empty vector control was also prepared.

Ligated products were transformed by heat shock as previously described (**2.16.5**) and selected on agar plates containing 100 µg/mL ampicillin (Sigma-Aldrich). Eight colonies from each gene were inoculated in 2 mL LB containing 100 µg/mL ampicillin and rocked horizontally overnight at 37°C.

Table 2.24 – Volume of reagents used for ligation of fusions into pRUFiG2 vector

Purpose	Materials	Volume/quantity	Total reaction volume (µL)
Restriction digest of pCR®-Blunt II-TOPO® construct	Purified plasmid DNA	1 µg	10 µL
	10x CutSmart® Buffer	1 µL	
	Restriction enzyme	0.5 µL	
	NF-H_2O	To 10 µL	
Dephosphorylation of pRUFiG2 with rSAP	Digested pRUFiG2	1 µg	20 µL
	Shrimp alkaline phosphatase (rSAP)	2 µL	
	10x CutSmart® Buffer	2 µL	
	NF-H_2O	To 20 µL	
T4 DNA ligation	T4 DNA ligase buffer (10x)	1 µL	10 µL
	T4 DNA ligase	1 µL	
	Purified insert DNA (or NF-H_2O for empty vector control reactions)	4 µL	
	Purified vector DNA	3 µL	
	NF-H_2O	1 µL	

2.16.12 Transformation of competent cells for amplification of *KMT2A-AFF1-pRUFiG2*

The *KMT2A-AFF1* fusion gene was commercially synthesised in the pRUFiG2 mammalian retroviral expression vector backbone[14] by GenScript®, as repeated attempts at PCR amplification of the entire fusion gene were unsuccessful, despite extensive troubleshooting. The construct was provided as a lyophilised powder, which was reconstituted in 20 µL NF-H_2O and then stored at -80°C. The fusion sequence synthesised is provided in **Appendix 2.3.**

One Shot^{TM} MAX Efficiency^{TM} DH5α^{TM}-T1^R chemically competent *E. Coli* (DH5α) were transformed with 1 µL of reconstituted *KMT2A-AFF1*-pRUFiG2 as previously described (**2.16.5**) and selected on agar plates containing 100 µg/mL ampicillin (Sigma-Aldrich). Six colonies were inoculated in 2 mL LB containing 100 µg/mL ampicillin and rocked horizontally overnight at 37°C.

2.16.13 Screening of isolates for *DDX3X-MLLT10, PICALM-MLLT10* and *KMT2A-AFF1* pRUFiG2

Selected isolates were screened for the presence of the *DDX3X-MLLT10*-pRUFiG2, *PICALM-MLLT10*-pRUFiG2 or *KMT2A-AFF1*-pRUFiG2 fusion gene constructs as previously described. Refer to **2.16.6**-**2.16.11** for detailed protocol on screening and storage protocols. Ampicillin (100 µg/µL) was used for selection of pRUFiG2 colonies.

2.16.14 Sanger sequencing to confirm ligation of fusions in pRUFiG2

Following identification of a correct isolate containing directionally cloned fusion gene within the pRUFiG2 vector, Sanger sequencing was used to verify the fusion gene sequence as previously described (**2.9.6**), using primers specific to the fusion and vector sequence (**Table 2.25**). Full coverage of each fusion gene was obtained, and adequate flanking sequence of the pRUFiG2 vector was covered to verify the orientation of the fusion within the vector.

Table 2.25 – Primer sequences for the Sanger sequencing of *DDX3X-MLLT10*, *KMT2A-AFF1* and PICALM-MLLT10 pRUFiG2 constructs

Fusion gene	Primers used for Sanger sequencing
DDX3X-MLLT10	Forward primer (RCF, pRUFiG2): GGGGACTCTGCTGACCAC
	Forward primer (*MLLT10* exon 19): CAGCTTTCAGTGCCTTTTCC
	Forward primer (*MLLT10* exon 15): TGGGCATTTGCAACAAGTAG
	Reverse primer (IRES, pRUFiG2): ACACCGGCCTTATTCCAAG
	Reverse primer (MLLT10 exon 24): CTGCAAAGGCAGCCAGATGA
KMT2A-AFF1	Forward primer (RCF, pRUFiG2): GGGGACTCTGCTGACCAC
	Forward primer (*KMT2A* exon 3): CTATTCGATCTGAACCAAGATCTCCTTCTC
	Forward primer (*KMT2A* exon 7): TCCACCACCAGAATCAGGTCC
	Forward primer (fusion breakpoint; *KMT2A* exon 9 to *AFF1* exon 5): GTGGACTTTAAGCAGACCTACTCCAATGAA
	Forward primer (*KMT2A* exon 9): AGGAGAATGCAGGCACTTTG
	Reverse primer (IRES, pRUFiG2): ACACCGGCCTTATTCCAAG
	Reverse primer (*KMT2A* exon 3): GAGAAGGAGATCTTGGTTCAGATCGAATAG
PICALM-MLLT10	Forward primer (RCF, pRUFiG2): GGGGACTCTGCTGACCAC
	Forward primer (*MLLT10* exon 19) CAGCTTTCAGTGCCTTTTCC
	Forward primer (MLLT10 exon 7): AAAGAGCAAACGGGGATCTAA
	Forward primer (*MLLT10* exon 10): CAGACCCAAAGGAAACAAAAA
	Forward primer (*MLLT10* exon 18): GAGAGGCAGTGGAGTGAAGG
	Forward primer (*PICALM* exon 6): CCAATGCAAAGAAGGTCTTG
	Reverse primer (IRES, pRUFiG2): ACACCGGCCTTATTCCAAG
	Reverse primer (*PICALM* exon 17): TCCTATCATGCCTGTTGGTG
	Reverse primer (*MLLT10* exon 10): GGTGAAGCAGAAGAAACTGAGAG
	Reverse primer: (*MLLT10* exon 17): GAGACGGACTAGGATTTGCTG
	Reverse primer: (*MLLT10* exon 20): GAGGGCTCTTACTGCTGTTCA

2.16.15 Generation of fusion gene retrovirus via transfection of HEK-293T cells

All work involving active virus was undertaken using appropriate containment and decontamination procedures. HEK-293T cells were seeded at a density of 9 x 10^5 in a T25 flask in 5 mL complete DMEM, and incubated overnight in a humidified incubator at 37°C and 5% CO_2. One flask per condition was prepared, including mock, pRUFiG2 empty vector control, and fusion-pRUFiG2.

The following day, pRUFiG2 vectors were co-transfected with packaging vector pEQEco[17] (courtesy of Dr. Stephen Fitter and Professor Andrew Zannettino, the University of Adelaide), in the presence of Lipofectamine™ 2000 (Thermo Fisher Scientific). A transfection solution consisting of 450 µL Opti-MEM™ Reduced Serum Media (Thermo Fisher Scientific) and 30 µL Lipofectamine™ 2000 per transfection was prepared in a 10 mL polypropylene tube. Retroviral pRUFiG2 expression vectors and packaging vector pEQEco (4-5 µg each) were prepared in 450 µL Opti-MEM™ in a 2 mL microcentrifuge tube, as well as a mock reaction containing no DNA. DNA and transfection solutions were incubated at RT for 10-20 min, and 500 µL of transfection solution was then added to each DNA solution and incubated for 2 h at RT. Following incubation, each transfection reaction was carefully added to a flask of HEK-293T cells and incubated overnight in a designated viral humidified incubator at 5% CO_2, 37°C. The following day, viral tissue culture media was discarded and replaced with 5 mL complete Ba/F3 media (for transduction of Ba/F3 cells), or DMEM (for transduction of MOHITO cells). Viral supernatant was then harvested 24 h later and replaced with 5 mL fresh media. Harvested supernatant was centrifuged at 1,400 rpm for 5 min, then filtered through a 0.45 µm sterile filter to obtain the viral supernatant. Viral supernatant was then either snap frozen at -80°C for later use or used for immediate transduction of cells (**2.16.16** or **2.16.17**). The viral supernatant harvesting process was repeated 24 h later, and HEK-293T cells were discarded.

2.16.16 Transfection of parental Ba/F3 pro-B cells with fusion retrovirus

Viral supernatant added to Ba/F3 pro-B parental cells (2 x 10^5 cells per well, in 3 mL final volume in a 6-well plate). Polybrene® was added to a final concentration of 8 µg/mL, and plates were centrifuged for 60 min at 1,800 rpm. At the end of the day, an additional 2 mL of complete Ba/F3 cell culture media was added to further dilute the Polybrene ®. Cells were incubated in a humidified incubator at 37°C and 5% CO_2 incubator for 3 d, with daily media replacement. Following incubation, GFP+ cells were collected by FACS (**2.10.1**).

2.16.17 Transfection of parental MOHITO T-cell line with fusion retrovirus

As transduction with retrovirus is less efficient in MOHITO cells than Ba/F3 cells, a modified RetroNectin® (Takara Bio, Kusatsu, Japan) transduction protocol was used. Non-tissue culture treated 6-well plates (Corning Inc.) were coated with 1 mL sterile 1x PBS/RetroNectin solution (30 µg/mL) and incubated at 4°C overnight. The following day, 1x PBS/RetroNectin solution was removed and stored long-term at 4°C for repeated use, and plate was blocked with 2 mL per well 2% BSA/1x PBS for 30 min at RT. Blocking solution was then discarded and wells washed with 1x PBS. Viral supernatant (1-1.5 mL per well) was then loaded onto RetroNectin coated plates by centrifugation at 2,000× g for 90 min. Viral supernatant was then discarded and 1.5 x 10^6 MOHITO cells suspended in 1 mL of complete MOHITO media were loaded into each well. Plates were centrifuged at 1,800 rpm for 60 min and incubated 37°C and 5% CO_2 in a humidified incubator. Media was changed every two days. Following three media changes, GFP$^+$ cells were collected by FACS (**2.10.1**). Cells were assessed for ability to transform MOHITO cells by culturing in MOHITO cytokine-free culture media (**2.2**), and viability was assessed using Trypan Blue exclusion and CellTiter-Glo® assays (**2.10.2**).

2.16.18 Confirmation of incorporation of fusion constructs in transduced cells using reverse transcription PCR (RT-PCR)

Following extraction of RNA from cell lines, cDNA was synthesised (**2.9.3**) and PCR was performed to verify fusion expression in transduced cell lines (**Table 2.26**).

Table 2.26 – Primer sequences for RT-PCR of fusion breakpoints

Fusion	Primers
DDX3X-MLLT10	Forward primer: GCAGTTTGCTGGCCTAGACCTGA
	Reverse primer: ATGAGATGGTGCCTGACTGAGAGAA
KMT2A-AFF1	Forward primer: TCCACCACCAGAATCAGGTCC
	Reverse primer: ATGCCACTGGTTCTGGAAGG
PICALM-MLLT10	Forward primer: CTCAGAACCAGAACCTTCCTGTTGC
	Reverse primer: ATGTCATGCAAGCACCAGTGGCT
ACTB (Beta-actin) (endogenous control)	Forward primer: GATCATTGCTCCTCCTGACC
	Reverse primer: GTCATAGTCCGCCTAGAAGCAT

Table 2.27– Volume of components for RT-PCR of fusion breakpoints

Component	Final concentration	Volume for 1 sample (µL)
5X Phusion HF Buffer	1x	4 µL
10 mM dNTP	200 µM	0.4 µL
5 µM Forward primer	400 nM	0.8 µL
5 µM Reverse primer	400 nM	0.8 µL
Template DNA	Variable	0.8 µL
Phusion® High-Fidelity DNA polymerase	0.4 unit	0.2 µL
Nuclease-free water	-	To 20 µL
Final volume	**-**	**20 µL**

PCR cycling conditions for amplification of fusion breakpoints:

1. 98°C for 1 min
2. 98°C for 10 sec
3. 68°C for 15 sec
4. 72°C for 1 min
5. Repeat steps 2-4, 34 cycles total
6. 72°C for 10 min
7. 4°C for ∞

2.16.19 qRT-PCR to assess differential expression of fusions and target genes in Ba/F3 and MOHITO cell lines

Refer to section **2.9.5** for qRT-PCR set-up conditions. The following primers were used to analyse expression of target genes in murine cell lines.

Table 2.28 – Primers for qRT-PCR of fusions and target genes in Ba/F3 and MOHITO cell lines

Gene	Primers
DDX3X-MLLT10	Forward primer: GCAGTTTGCTGGCCTAGACCTGA
	Reverse primer: ATGAGATGGTGCCTGACTGAGAGAA
KMT2A-AFF1	Forward primer: TCCACCACCAGAATCAGGTCC
	Reverse primer: ATGCCACTGGTTCTGGAAGG
PICALM-MLLT10	Forward primer: CTCAGAACCAGAACCTTCCTGTTGC
	Reverse primer: ATGTCATGCAAGCACCAGTGGCT
HOXA1	Forward primer: AGAAACCCTCCCAAAACAGG
	Reverse primer: TGAAGTGGAACTCCTTCTCCAG
HOXA3	Forward primer: ATGCAAAAAGCGACCTACTAC
	Reverse primer: CGCTCAGGGTGCGCAG
HOXA5	Forward primer: AGCCACAAATCAAGCACACA
	Reverse primer: GCTGAGATCCATGCCATTGT
HOXA7	Forward primer: CAGGTCAAAATTATGAGTTCTTCG
	Reverse primer: TCTGATAAAGGGGGCTGTTG
HOXA9	Forward primer: TCACACTTTGTCCCTGACTGA
	Reverse primer: TTCGCTGGGTTGTTTTTCTC
HOXA10	Forward primer: CCTTCAGAAAACAGTAAAGCTTCG
	Reverse primer: AAGGGCAGCGTTTCTTCC
MEIS1	Forward primer: TTGTAATGGACGGTCAGCAG
	Reverse primer: GCTACATACTCCCCTGGCATA
ACTB (endogenous control)	Forward primer: GATCATTGCTCCTCCTGACC
	Reverse primer: GTCATAGTCCGCCTAGAAGCAT

2.17 Generation of stable fusion-expressing Arf$^{-/-}$ thymocyte cells and *in vivo* murine xenograft

Thymocytes from female p19^{Arf} null C57BL/6J (Arf$^{-/-}$) mice[18] aged between 4-12 weeks of age were humanely euthanised, and immature thymocytes (CD4^{-}CD8^{-}TER119^{-}) were harvested by flow cytometry. Thymocytes were expanded by co-culture with the OP9-DL1 stromal cell line. *DDX3X-MLLT10* and *PICALM-MLLT10* pRUFiG2 expression constructs were generated as described in sections **2.16.1**-**2.16.14**. Retrovirus was generated and thymocytes were transduced (**2.17.4**) to generate stable fusion-expressing cells. GFP^{+} cells were expanded and immunophenotyped prior to freezing. *In vivo* murine xenografts were generated by thawing transduced thymocytes which were immediately injected into sublethally irradiated six-week-old *NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/Sz* (NSG) mice by tail vein injection (with the assistance of Dr. Laura Eadie). Engraftment was monitored by fortnightly teil vein bleeds and analysis of GFP^{+} populations by flow cytometry. Mice were monitored for clinical signs of engraftment, and were humanely euthanised when they appeared moribund. Cells were harvested for analysis from bone marrow (hind limbs), thymus, liver, spleen, and kidneys, and stored according to **Table 2.29**.

2.17.1 Harvesting of CD4^{-}CD8^{-}TER119^{-} Arf$^{-/-}$ thymocytes

Three female Arf$^{-/-}$ mice aged between 4-12 weeks were humanely euthanised. Thymi were harvested and stored in D10 media on ice prior to cell extraction. Single thymocytes were isolated by homogenising thymi through a 70 µM cell strainer, and cell count and viability was determined by WCF and Trypan blue cell counts (**2.17.11**). Cells were pelleted at 1,400 rpm for 5 min and resuspended in 90 µL 2% FCS/1x PBS per 10^{7} cells. A small volume was retained for immunophenotyping and unstained control samples. The remainder of cells were stained with 2.5 µL of each CD4-PE, CD8-PE and TER119-PE (BD Pharmingen, California, USA) per 10^{7} cells, to isolate immature thymocytes. TER119 staining was performed to exclude erythrocytes. Stained cells were mixed and incubated on ice for 30 min. Cells were then washed in 30 mL 2% FCS/1x PBS and centrifuged at 1,400 rpm for 5 min at 4°C. Supernatant was aspirated and cells resuspended in 2% FCS/1x PBS at a concentration of 5 x 10^{7} cells/mL, and PE negative (PE-) cells were sorted on a FACSFusion^{TM} flow cytometer.

PE-negative (PE-) cells were collected in Arf$^{-/-}$ thymocyte complete media (**2.2**). Following FACS, CD4$^-$CD8$^-$TER119$^-$ sorted cells were pelleted and resuspended in Arf$^{-/-}$ thymocyte complete media and co-cultured on OP9-DL1 stromal cells at a concentration of 0.5-1.0 x 10^6 cells per well in 2 mL of media (**2.17.2**).

2.17.2   Routine passage of OP9-DL1 stromal cells

OP9-DL1 cells were seeded at 3 x 10^5 cells in a 10 cm dish for routine passage. To split confluent cells, media was aspirated from the dish and cells were washed in 5 mL RT 1x PBS. The adherent OP9-DL1 cell monolayer was dissociated by incubation at 37°C for 5 min with 3 mL 0.25% trypsin. Trypsinised cells were harvested in a 15 mL tube containing 4 mL alpha-MEM media, and the plate washed with 7 mL 1x PBS to remove residual cells. Harvested cells were pelleted at 1,400 rpm for 5 min and resuspended in 2-5 mL of OP9-DL1 culture media (**2.2**). Cells were allowed to reach a maximum confluency of 80%, and were routinely passaged every 2-3 days. Cells were cultured up to a maximum of 20 passages.

For preparation of OP9-DL1 stromal layers for co-culture with primary thymocytes, OP9-DL1 cells were passaged in 12-well plates at a density of 2 x 10^4 cells/well, in 2 mL of OP9-DL1 culture media. In 2-3 days after seeding, confluent OP9-DL1 stromal layers were ready for co-culture with thymocytes.

2.17.3   Co-culture of primary thymocytes on OP9-DL1 stromal cells

Media was aspirated from confluent OP9-DL1 wells, and 0.5 x 10^6 to 1 x 10^6 thymocytes were applied to each well, in a final volume of 2 mL Arf$^{-/-}$ thymocyte complete media. Every 3-6 days, thymocytes were passaged onto fresh OP9-DL1 stromal cells. OP9-DL1 cells and thymocytes were disturbed with a sterile cell scraper, and a p1000 pipette was used to further disturb the scraped cell layer to isolate single thymocytes. The cell suspension was passed through the strainer into the 50 mL tube, and each well rinsed with 1 mL RT 1x PBS and transferred through the strainer. Cells were pelleted at 1,400 rpm for 5 min and resuspended in thymocyte complete media. Cells were counted using trypan blue dye exclusion, and passaged at 0.5 x 10^6 to 1 x 10^6 viable cells per well containing fresh confluent OP9-DL1 cells.

2.17.4   Generation of stable *MLLT10*r expressing Arf$^{-/-}$ thymocyte cell lines

Retrovirus expressing empty pRUFiG2, *DDX3X-MLLT10*-pRUFiG2 or *PICALM-MLLT10*-pRUFiG2 was generated as described in section **2.16.15**. Retroviral supernatant was applied to primary Arf$^{-/-}$ thymocytes by RetroNectin® transduction, as described in section **2.16.17**, with the following modifications.

Approximately 4 mL of retroviral supernatant was generated per condition (empty vector, *DDX3X-MLLT10* and *PICALM-MLLT10*), and 800 μL was applied to each well of a retronectin-coated 24-well plate, for a total of 5 wells per condition. Plates were centrifuged at 2,000× g for 90 min. During the centrifugation step, thymocytes were harvested from OP9-DL1 co-culture wells and counted by trypan blue dye exclusion. Following centrifugation of virus, retroviral supernatant was aspirated and 1 x 10^6 Arf$^{-/-}$ thymocytes were applied to each well. The plate containing cells was then centrifuged at 1,800 rpm for 60 min. Immediately following centrifugation, thymocytes were removed from the retronectin plate and transferred to fresh confluent OP9-DL1 wells in a 12-well plate.

Cells were sorted by FACS, gating on GFP$^+$ cells, and cultured on OP9-DL1 wells. GFP expression was checked after 5-10 days in culture, and re-sorted if required. Following confirmation of >95% GFP$^+$ cells in each condition, cells were cryopreserved for future use. Expression of the correct construct was confirmed by RT-PCR of either the pRUFiG2 empty vector multiple cloning site (MCS), or each fusion's specific breakpoint, as described in section **2.16.18**.

2.17.5   Preparation of cells for injection

One ampule of each cryopreserved Arf$^{-/-}$ thymocyte cell line (empty vector, *DDX3X-MLLT10* and *PICALM-MLLT10*) were thawed by incubation in a 37°C water bath for approximately 1 min. Thawed cells were transferred to a 50 mL tube and 12 mL thymocyte thaw media (**2.2**) was added dropwise. Cells were pelleted for 3 min at 450× g, supernatant aspirated, and washing repeated once with another 12 mL of thymocyte thaw media. Cells were resuspended in RT 1x PBS, WCF and Trypan blue cell counts performed, and cell suspension was adjusted to a concentration of 1.5 x 10^7 cells/mL of RT 1x PBS. Cell suspension was filtered through a 5 mL FACS tube with nylon

mesh strainer cap (STEMCELL Technologies) to remove clumps of cells, and promptly injected into mice.

2.17.6   Mouse injections

NSG mice between 6 and 8 weeks of age were used for *in vivo* experiments. Enrofloxacin (50 mg/mL) was diluted to a working concentration of 1 mg/mL in 1x PBS for injection as anti-bacterial prophylaxis. The NSG strain was utilised in this experiment as it is a well characterised and readily engraftable strain of mice. Four days prior to irradiation, mice were weighed and subcutaneously injected with 10 µL of 1 mg/mL enrofloxacin per g. Mice were sublethally irradiated with 200 cGy less than 24 hours prior to injection of cells via tail vein injection.

2.17.7   Mouse treatment

Mice were weighed and monitored daily to assess their general health. Mice were monitored for disease progression based on clinical signs and fortnightly tail vein bleeds (**2.17.8**). Mice were humanely killed when they appeared moribund.

2.17.8   Tail vein bleeds

Tail vein bleeds were performed fortnightly to monitor for disease progression. Mice were individually pre-warmed in a 37°C thermacage for 5 min, and 50-100 µL of blood was harvested by incision of the tail vein with a scalpel. Blood was collected in EDTA-coated microvettes and mixed thoroughly.

Red blood cells were lysed to enable white blood cell analysis. Approximately 0.5-1 mL of red cell lysis buffer (RCLB) (**2.6**) was applied to each blood sample, mixed and incubated at RT for 10 min. Samples were centrifuged at 500$\times$ g for 5 min and supernatant aspirated. Lysis was repeated, and remaining white cells resuspended in 100 µL FACS fix. Cell were stained with 1 µL DAPI (50 µg/mL) per sample and incubated on ice for 5 min prior to analysis on a BD FACSCanto™. DAPI negative single cells were assessed for percentage of GFP$^+$ cells.

2.17.9 *Ex vivo* sample collection and processing

Mice were humanely killed once they appeared moribund, by CO_2 euthanasia and cervical dislocation. Peripheral blood was harvested by either cardiac puncture or severing of the inferior vena cava. Harvested blood was collected in a 1.7 mL tube containing 10% of the total blood volume of 0.5 M EDTA. Approximately 30 µL of peripheral blood was stored separately for complete blood count (CBC) analysis (**2.17.15**). Mice were then dissected and spleen, liver and hind limbs from all mice, and the thymus and kidneys where indicated, were harvested and stored in 1x PBS in a petri dish. All organs were photographed prior to processing, and spleen and liver were weighed with a fine balance.

Organs were transferred to individual filters placed in a sterile 50 mL tube, and the plunger from a 3 mL syringe was used to puree organs. Single cells were harvested by homogenisation through a 70 µM filter. Filters were thoroughly washed with 2% FCS/1x PBS to ensure all cells were filtered. Bone marrow was harvested by flushing marrow cavities with 2% FCS/1x PBS using a 27 G needle.

Cells were pelleted by centrifugation at 500× g for 5 min. Supernatant was aspirated and RBCs lysed in 1-5 mL of RCLB for 10 min at RT. Cells were pelleted and RBC lysis was repeated if necessary, depending on the presence of residual RBCs in the cell pellet. Liver and kidney samples underwent gradient density centrifugation (**2.17.10**). Cells were resuspended in an appropriate volume (0.5-5 mL) of 2% FCS/1x PBS based on pellet size, and counted (**2.17.11**).

2.17.10 Gradient density centrifugation of low viability samples

Following RBC lysis, liver and kidney cell suspensions were resuspended in 6 mL 1x PBS and transferred to a 15 mL tube, and underlaid with 5 mL Lymphoprep. Tubes were centrifuged for 10 min at 1000× g with brake off. The live cell layer was transferred to a new 15 mL tube with a sterile pipette tip and washed to a maximal volume of 14 mL with 1x PBS to eliminate residual Lymphoprep. Cells were pelleted and resuspended in 2% FCS/1x PBS for counting.

2.17.11 Cell counting

Cell counts were performed using WCF to determine white cell number, and Trypan Blue to determine percentage of live cells. Percentage live white cells were calculated and cells were stored for various purposes according to **Table 2.29**.

Table 2.29 – Cell numbers required for storage of *ex vivo* mouse material

Purpose	Cell number required
Flow cytometric analysis (GFP only)	2×10^5
Immunophenotyping	1×10^6
Cryopreservation	Minimum 1×10^7
Whole cell pellet	1×10^7
TRIzol® (for RNA extraction)	4×10^6 to 1×10^7

2.17.12 Flow cytometry

An appropriate number of cells were transferred to 5 mL FACS tubes for GFP analysis and immunophenotyping. Some samples were thawed and immunophenotyped retrospectively. Refer to **2.10.2** for staining protocol.

2.17.13 Criteria for leukaemic engraftment

Mice were considered to have leukaemic engraftment if any harvested organ contained >20% GFP$^+$ live single cells or total WCC exceeded 25 K/µL.

2.17.14 Cryopreservation

Cells were cryopreserved in 90% FCS/10% DMSO. Refer to **2.10.2**.

2.17.15 Complete blood count (CBC)

Approximately 30 µL of peripheral blood was stored at 4°C for up to 1 week prior to CBC. Blood was analysed using the HEMAVET® 950 veterinary haematology system according to manufacturer's instructions.

2.17.16 Immunohistochemistry

Organ sections from each mouse were stored in 10% formalin for immunohistochemistry. A small section of the spleen and liver, and one femur were stored in for all mice, and sections of the thymus and kidney were also stored where indicated.

Samples were submitted to the Adelaide Health and Medical Science histology laboratory for parrafin embedding, sectioning, H&E staining, and GFP staining (antibodies specified in **Table 2.30**). On the day of sample submission, formaldehyde was removed and replaced with 70% ethanol.

Table 2.30 – Antibodies used for GFP immunohistochemistry

Primary antibody	Secondary antibody
GFP (D5.1) Rabbit mAb #2956S Supplier: Cell Signalling Technology Diluted 1:200 in 5% goat serum in PBST	Goat Anti-Rabbit IgG H&: (HRP) Supplier: Abcam # ab205718 Diluted 1:2,000-50,000 in 5% goat serum in PBST

2.17.17 qRT-PCR

RNA extraction, cDNA synthesis and qRT-PCR protocols are described in sections **2.9.2**-**2.9.5**. Primers for qRT-PCR targets are provided in **Table 2.31**.

Table 2.31 – Primers for qRT-PCR of targets in Arf$^{-/-}$ primary xenograft cells

Gene	Primers
DDX3X-MLLT10	Forward primer: GCAGTTTGCTGGCCTAGACCTGA
	Reverse primer: ATGAGATGGTGCCTGACTGAGAGAA
PICALM-MLLT10	Forward primer: CTCAGAACCAGAACCTTCCTGTTGC
	Reverse primer: ATGTCATGCAAGCACCAGTGGCT
HOXA1	Forward primer: AGAAACCCTCCCAAAACAGG
	Reverse primer: TGAAGTGGAACTCCTTCTCCAG
HOXA3	Forward primer: ATGCAAAAAGCGACCTACTAC
	Reverse primer: CGCTCAGGGTGCGCAG
HOXA5	Forward primer: AGCCACAAATCAAGCACACA
	Reverse primer: GCTGAGATCCATGCCATTGT
HOXA7	Forward primer: CAGGTCAAAATTATGAGTTCTTCG
	Reverse primer: TCTGATAAAGGGGGCTGTTG
HOXA9	Forward primer: TCACACTTTGTCCCTGACTGA
	Reverse primer: TTCGCTGGGTTGTTTTTCTC
HOXA10	Forward primer: CCTTCAGAAAACAGTAAAGCTTCG
	Reverse primer: AAGGGCAGCGTTTCTTCC
HOXB1	Forward primer: TCTACCCCTGGATGCGCAAA
	Reverse primer: AACCAGATCTTGATCTGGCGC
MEIS1	Forward primer: TTGTAATGGACGGTCAGCAG
	Reverse primer: GCTACATACTCCCCTGGCATA
BMI1	Forward primer: CTGGAAAGTGACTCTGGGAGTGACA
	Reverse primer: GGGCTGTTGCTGGTTCCATTCAT
ACTB (Beta-actin) (endogenous control)	Forward primer: GATCATTGCTCCTCCTGACC
	Reverse primer: GTCATAGTCCGCCTAGAAGCAT

2.18 Generation of *TP53* gene knockout in RS4;11 cells using CRISPR/Cas9 genome editing

TP53 was targeted for knockout in RS4;11 cells using a lentiviral doxycycline-inducible CRISPR/Cas9 genome editing system. *TP53* is recurrently mutated in ALL patient leukaemic blasts, and it was hypothesised that normal function of p53 is required for the action of the small molecule curaxin CBL0137. Details for the generation of *TP53* loss-of-function cell lines are described below. Further methodology can be found in the Methods and Supplementary Methods sections of **chapter 6.**

Process	Purpose
1. Generate FgH1tUTG sgRNA plasmid	Design of sgRNA for directed genome targeting
2. Generation of lentivirus by co-transfection of FuCas9mCherry and lentiviral packaging plasmids, and transduction in RS4;11 host cells	Generate RS4;11 cell lines with stable expression of Cas9 for genome editing
3. FACS of mCherry-expressing cells	Enrichment of Cas9-expressing cell population
4. Generation of lentivirus by co-transfection of FgH1tUTG sgRNA plasmid and lentiviral packaging plasmids, and transduction in RS4;11 FuCas9mCherry-expressing cells	Generate RS4;11 cell lines with stable Cas9 expression and sgRNA targeting exon for knockout of *TP53*
5. FACS of mCherry and GFP double-positive cells (if required)	Enrichment of cells containing both Cas9 and FgH1tUTG gRNA
6. Treatment of cells with doxycycline for 72 hours	Activation of FgH1tUTG sgRNA expression
7. gDNA extraction, and PCR amplification and Sanger sequencing of targeted exons	Screen cells for identification of mutations in target gene

Figure 2.3 – Workflow for the generation of gene knockout in RS4;11 cells using a lentiviral doxycycline-inducible CRISPR/Cas9 genome editing system

A. FgH1tUTG
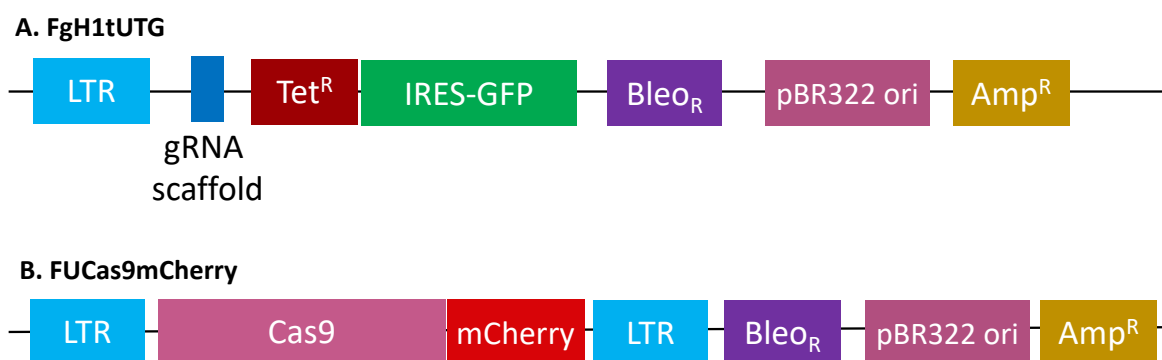
B. FUCas9mCherry

Figure 2.4 – Plasmid maps for (A) FgH1tUTG and (B) FUCas9mCherry

2.18.1 Generation of FgH1tUTG sgRNA plasmid

Benchling® software[19] was used to generate sgRNA oligonucleotides targeting an early exon of murine target gene *TP53* (**Table 2.32**). BsmBI restriction site overhangs were added to the 20 bp sequence to enable ligation with the FgH1tUTG plasmid backbone. sgRNA oligonucleotides were synthesised by Sigma-Aldrich and stored at -20°C when not in use. Protocols for the ligation of sgRNA and FgH1tUTG are provided in **Tables 2.33-2.34.**

Table 2.32 – sgRNA oligonucleotides designed for targeting of early *TP53* exons by CRISPR/Cas9

Target exon	sgRNA oligonucleotide (5' to 3')
TP53 exon 4, sgRNA 1	Forward: tcccACCAGCAGCTCCTACACCGG
	Reverse: aaacCCGGTGTAGGAGCTGCTGGT
TP53 exon 4, sgRNA 2	Forward: tcccCCATTGTTCAATATCGTCCG
	Reverse: aaacCGGACGATATTGAACAATGG

Note: 'tccc' and 'aaac' prefix sequences correspond to BsmBI restriction site overhangs required for ligation with FgH1tUTG

Table 2.33 – Protocol for phosphorylation of sgRNA oligonucleotides

Purpose	Materials	Volume/quantity	Total reaction volume (μL)
1. Phosphorylation of oligonucleotides	Oligonucleotide, forward (100 μM)	1 μL	10 μL
	Oligonucleotide, reverse (100 μM)	1 μL	
	T4 polynucleotide kinase (PNK)	1 μL	
	T4 ligase buffer	1 μL	
	NF-H_2O	6 μL	
2. Dilution (1:125) of phosphorylated oligonucleotides	Phosphorylated oligonucleotides from step 1	2 μL	250 μL
	NF-H_2O	248 μL	
3. Esp3I restriction digest of FgH1tUTG plasmid	Esp3I	1 μL	9 μL
	CutSmart buffer	1 μL	
	FgH1tUTG	1 μg	
	NF-H_2O	To 9 μL	
4. Dephosphorylation of Esp3I-digested FgH1tUTG with rSAP	Digested FgH1tUTG	9 μL	10 μL
	rSAP	1 μL	
5. Ligation of FgH1tUTG with phosphorylated oligonucleotides	FgH1tUTG (Esp3I-digested & rSAP-treated)	100 ng	20 μL
	1:125 diluted oligonucleotides	1 μL	
	T4 DNA ligase	1 μL	
	T4 DNA ligase buffer	1 μL	
	NF-H_2O	To 20 μL	

Table 2.34 – Reaction conditions for generation of FgH1tUTG sgRNAs

Purpose	Incubation conditions
Phosphorylation of oligonucleotides	**37°C for 30 min** **95°C for 5 min** **25°C for ∞, ramp rate of 0.1°C/sec**
Esp3I restriction digest of FgH1tUTG	37°C for 30 min *Addition of 1 µL rSAP* 37°C for a further 30 min 65°C for 5 min
Ligation of FgH1tUTG with phosphorylated oligonucleotides	4°C overnight

2.18.2 Transformation of competent cells for amplification of FgH1tUTG sgRNA plasmids

Competent cells were transformed with 10 µL of ligation reactions as previously described (**2.16.5**), except with the addition of 750 µL RT S.O.C. A 75 µL aliquot of transformation reactions were selected on agar plates containing 100 µg/mL ampicillin (Sigma-Aldrich). Four colonies were inoculated in 2 mL LB containing 100 µg/mL ampicillin and rocked horizontally overnight at 37°C.

2.18.3 Screening of isolates for FgH1tUTG-sgRNA ligation

Selected isolates were screened for the presence of desired sgRNA within the FgH1tUTG plasmid as previously described. Refer to **2.16.6-2.16.11** for detailed protocol on screening and storage protocols. Note that ampicillin (100 µg/µL) was used for selection of FgH1tUTG colonies.

2.18.4 Sanger sequencing to screen for ligation of oligonucleotides in FgH1tUTG

Sanger sequencing was employed to screen for ligation of sgRNAs with FgH1tUTG previously described (**refer to 2.9.6**), using the H1 primer targeted towards a region

of FgH1tUTG, 5' adjacent to sgRNA scaffold site (H1 primer sequence: TCGCTATGTGTTCTGGGAAA).

2.18.5   Generation of FUCas9mCherry and FgH1tUTG-sgRNA lentivirus via transfection of HEK-293T cells

Lentivirus containing either FUCas9mCherry or FgH1tUTG-sgRNA were generated as previously described (**2.16.15**), using lentiviral packaging vectors pMD2.G, pMDL/pRRE and pRSV-Rev. Quantities of expression plasmid and packaging vectors were prepared in a 1:1 molar ratio (**Table 2.35**).

Table 2.35 – Volume of reagents used for generation of lentivirus

Plasmid name	Plasmid size (bp)	Quantity for 1:1 molar ratio (µg)
pMD2.G	5,824	1.6
pMDL/pRRE	8,895	2.4
pRSV-Rev	4,174	1.1
FgH1tUTG-sgRNA	10,968	3.5
or		
FUCas9mCherry	10,948	3.5

2.18.6   Transfection of RS4;11 cells with FUCas9mCherry and FgH1tUTG-sgRNA lentiviruses

RS4;11 cells were transfected with FUCas9mCherry lentivirus as previously described (**2.18.6**). Following incubation, mCherry positive cells were collected by FACS (**2.10.1**).

Sorted cells were cultured for 1 week before undergoing transfection with FgH1tUTG-sgRNA as previously described (**2.18.6**). Due to high transfection efficiency of FgH1tUTG-sgRNA plasmids, mCherry and GFP double-positive cells were collected by FACS only if transfection efficiency was <75% (**2.10.1**). Cell populations generated are provided in **Table 2.36**.

Table 2.36 – Cell lines generated by transduction of FUCas9mCherry and FgH1tUTG sgRNA for genome editing by CRISPR/Cas9 system

Cell line name	Transduced plasmids	Purpose
RS4;11 Cas9	FUCas9mCherry	Parental RS4;11 control
RS4;11 TP53$^{+/-}$ #1	FUCas9mCherry FgH1tUTG-TP53-sgRNA-1	Heterozygous knockout of *TP53* using sgRNA 1
RS4;11 TP53$^{+/-}$ #2	FUCas9mCherry FgH1tUTG-TP53-sgRNA-2	Heterozygous knockout of *TP53* using sgRNA 2

2.18.7 Induction of FgH1tUTG sgRNA expression with doxycycline

RS4;11 cells transduced with FUCas9mCherry and FgH1tUTG sgRNA were treated with doxycycline (1 µg/mL) (Sigma-Aldrich) and cultured for 72 hours in complete cell culture media (**2.2**). Control cells transduced only with FUCas9mCherry were treated alongside experimental cells.

2.18.8 T7 endonuclease I mismatch cleavage assay to identify targeted mutagenesis in RS4;11 cell lines

Following doxycycline activation of FgH1tUTG, RS4;11 cell lines were assessed for targeted genome editing events using the T7 endonuclease I (T7E1) mismatch cleavage assay. In this assay, gDNA from edited cells is used as a direct PCR template for amplification with primers specific to the targeted region. The PCR product is then denatured and reannealed to produce heteroduplex mismatches where double-stranded breaks have occurred. T7E1 catalyses the cleavage of DNA mismatches, producing DNA fragments that can be visualised by agarose gel electrophoresis.

gDNA from doxycycline-treated cell lines was extracted as previously described (**2.9.7**) and PCR was performed to amplify targeted genomic regions with primers designed within intronic sequences flanking the target exon (**Table 2.37**). PCR protocol is provided in **Table 2.38**.

Table 2.37 – Primer sequences for PCR of *TP53* exon 4 targeted by CRISPR/Cas9

Target	Primers
TP53 exon 4	Forward primer: TCCTCTGACTGCTCTTTTCACCCAT
	Reverse primer: AATATTCAACTTTGGGACAGGAGTCAGAGA

Table 2.38 – Volume of components for PCR amplification of *TP53* exon 4

Component	Final concentration	Volume for 1 sample (µL)
5X Q5 Reaction Buffer (NEB)	1x	8 µL
10 mM dNTP	200 µM	0.8 µL
10 µM Forward primer	500 nM	2 µL
10 µM Reverse primer	500 nM	2 µL
Template DNA	100 ng	2 µL
Q5 High-Fidelity DNA polymerase (NEB)	0.4 unit	0.2 µL
Nuclease-free water	-	To 40 µL
Final volume	-	**40 µL**

PCR cycling conditions:

1. 98°C for 30 sec
2. 98°C for 5 sec
3. 69°C for 10 sec
4. 72°C for 20 sec
5. Repeat steps 2-4, 34 cycles total
6. 72°C for 2 min
7. 4°C for ∞

The entire volume (40 µL) of each PCR was electrophoresed with 15 µL Gel Loading Dye (NEB) on a 1% agarose gel containing GelRed (Biotium) for 70 min at 110 V. Bands corresponding to amplified exons were excised using a scalpel and LED Transilluminator (Maestrogen). Gels were not imaged prior to excision to prevent UV-induced DNA damage. DNA was gel extracted using the Silica Bead DNA Gel Extraction Kit (Thermo Scientific) according to manufacturer's instructions, with final elution volumes between 14 to 20 µL. Purified samples were then used for Sanger sequencing to determine mutations present, as previously described (**2.9.6**).

2.19 Chapter References

1   Applied Biosystems.   (ed Thermo Fisher Scientific Inc.) (United States of America, 2019).

2   Nicorici, D. *et al.* FusionCatcher – a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *BioRxiv*, doi:10.1101/011650 (2014).

3   Davidson, N. M., Majewski, I. J. & Oshlack, A. JAFFA: High sensitivity transcriptome-focused fusion gene detection. *Genome Med* **7**, 43, doi:10.1186/s13073-015-0167-x (2015).

4   Jia, W. *et al.* SOAPfuse: an algorithm for identifying fusion transcripts from paired-end RNA-Seq dat. *Genome Biology* **14**, 1-15, doi:10.1186/gb-2013-14-2-r12 (2013).

5   McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303, doi:10.1101/gr.107524.110 (2010).

6   Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164, doi:10.1093/nar/gkq603 (2010).

7   Chang, Y. H. *et al.* Targeted sequencing to identify genetic alterations and prognostic markers in pediatric T-cell acute lymphoblastic leukemia. *Sci Rep* **11**, 769, doi:10.1038/s41598-020-80613-6 (2021).

8   Liu, Y. *et al.* The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nat Genet* **49**, 1211-1218, doi:10.1038/ng.3909 (2017).

9   Huether, R. *et al.* The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nat Commun* **5**, 3630, doi:10.1038/ncomms4630 (2014).

10   Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res* **47**, D941-D947, doi:10.1093/nar/gky1015 (2019).

11   Kopanos, C. *et al.* VarSome: the human genomic variant search engine. *Bioinformatics* **35**, 1978-1980, doi:10.1093/bioinformatics/bty897 (2019).

12   Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* **17**, 405-424, doi:10.1038/gim.2015.30 (2015).

13   Shihab, H. A., Gough, J., Cooper, D. N., Day, I. N. & Gaunt, T. R. Predicting the functional consequences of cancer-associated amino acid substitutions. *Bioinformatics* **29**, 1504-1510, doi:10.1093/bioinformatics/btt182 (2013).

14   Noll, J. *et al.* SAMSN1 Is a Tumor Suppressor Gene in Multiple Myeloma. *Neoplasia* **16**, 572-585 (2014).

15   Kleppe, M., Mentens, N., Tousseyn, T., Wlodarska, I. & Cools, J. MOHITO, a novel mouse cytokine-dependent T-cell line, enables studies of oncogenic signaling in the T-cell context. *Haematologica* **96**, 779-783, doi:10.3324/haematol.2010.035931 (2011).

16   Vincze, T., Posfai, J. & Roberts, R. J. NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Research* **31**, 3688-3691 (2003).

17   Persons, D., Mehaffey, M., Kaleko, M., Nienhuis, A. & Vanin, E. An Improved Method for Generating Retroviral Producer Clones for Vectors Lacking a Selectable Marker Gene. *Blood, Cells, Molecules, and Diseases* **24**, 167-182 (1998).

18   Kamijo, T. *et al.* Tumor Suppression at the Mouse INK4a Locus Mediated by the Alternative Reading Frame Product p19ARF. *Cell* **91**, 649-659, doi:10.1016/S0092-8674(00)80452-3 (1997).

19   Benchling [Biology Software] (2019).

Chapter Appendices

Appendix 2.1 – *DDX3X-MLLT10* fusion sequence

TTAGCAGCGGAAGACTCCGAGTTCTCGGTACTCTTCAGGGATGAGTCATGTGGCAGTGGAAAATGCGCTCGGGCT
GGACCAGCAGTTTGCTGGCCTAGACCTGAACTCTTCAGATAATCAGAGTGGAGGAAGTACAGCCAGCAAAGGGCG
CTATATTCCTCCTCATTTAAGGAACCGAGAAGCTACTAAAGGTATTTATAACAGCAATGATGTAGCAGTATCGTT
TCCAAATGTAGTATCTGGCTCGGGATCTAGTACTCCTGTCTCCAGCTCTCACTTACCTCAGCAGTCTTCTGGGCA
TTTGCAACAAGTAGGAGCGCTCTCTCCCTCAGCTGTGTCATCTGCAGCCCCTGCTGTTGCTACAACTCAGGCAAA
TACTCTATCTGGATCTTCTCTCAGTCAGGCACCATCTCATATGTATGGCAATAGATCAAATTCATCAATGGCAGC
TCTTATAGCTCAGTCTGAAAACAATCAAACAGATCAAGATCTTGGAGACAATAGCCGCAACCTAGTTGGCAGAGG
AAGCTCACCCCGAGGAAGTCTCTCGCCACGATCCCCTGTAAGCAGCTTACAGATTCGCTATGATCAACCAGGCAA
CAGCAGTTTGGAAAATCTGCCTCCAGTAGCAGCCAGCATAGAACAGCTTTTGGAGAGGCAGTGGAGTGAAGGACA
GCAATTTTTACTAGAACAGGGTACTCCTAGTGACATTTTAGGAATGCTGAAGTCATTACACCAACTTCAAGTTGA
AAACCGAAGATTAGAGGAACAAATTAAAAACTTGACTGCCAAAAAGGAACGGCTTCAGTTATTGAATGCACAGCT
TTCAGTGCCTTTTCCAACAATAACAGCAAATCCTAGTCCGTCTCATCAAATACACACATTTTCAGCACAGACTGC
TCCTACTACTGATTCCTTGAACAGCAGTAAGAGCCCTCATATAGGAAACAGCTTTTTACCTGATAATTCTCTTCC
TGTATTAAATCAGGACTTAACCTCCAGTGGACAAAGTACCAGCAGCTCATCAGCTCTTTCTACCCCACCTCCTGC
TGGGCAGAGTCCGGCTCAACAAGGCTCAGGAGTGAGTGGAGTTCAGCAGGTCAATGGCGTGACAGTGGGGGCACT
AGCTAGTGGAATGCAGCCTGTAACTTCCACCATTCCTGCCGTGTCTGCAGTGGGTGGAATAATTGGAGCTTTGCC
AGGTAACCAACTGGCAATTAATGGCATTGTAGGAGCTTTAAATGGGGTTATGCAGACTCCTGTCACAATGTCCCA
GAACCCTACCCCTCTCACCCACACAACCGTACCACCTAATGCAACACATCCAATGCCAGCTACACTGACTAACAG
TGCCTCAGGACTAGGATTACTTTCTGACCAGCAACGACAAATACTTATTCATCAACAGCAGTTTCAGCAGTTGTT
AAATTCTCAACAGCTCACACCAGAACAACATCAAGCCTTTTTGTATCAGTTAATGCAACATCACCACCAGCAGCA
CCACCAACCTGAACTTCAGCAGCTGCAGATCCCTGGACCAACACAAATACCCATAAACAACCTTCTTGCAGGTAC
ACAGGCACCCCCACTTCACACAGCTACCACCAACCCATTTCTCACCATCCATGGAGATAATGCAAGTCAGAAAGT
AGCAAGACTTAGTGATAAAACTGGGCCTGTAGCTCAAGAGAAAAGTTGACACCTGAGAAACATCTAGAAATTGCC
TATCCTGCTGTTCTAGCACTTCATCTGGCTGCCTTTGCAG

Appendix 2.1 – *DDX3X-MLLT10* fusion sequence

The sequence of *DDX3X-MLLT10* is provided, with primers used to PCR-amplify the fusion from patient-derived cDNA highlighted in yellow (forward) and green (reverse). The entire coding sequence is 1,659 bp long, with the 5' *DDX3X* sequence in blue and the 3' *MLLT10* sequence in black. The entire PCR product size including primer sequences is 1,767 bp long. The start (ATG) and stop (TGA) codons are underlined.

Appendix 2.2 – *PICALM-MLLT10* fusion sequence

GACTCTAGAGCAGAGATGTCCGGCCAGAGCCTGACGGACCGAATCACTGCCGCCCAGCACAGTGTCACCGGCTCT
GCCGTATCCAAGACAGTATGCAAGGCCACGACCCACGAGATCATGGGGCCCAAGAAAAAGCACCTGGACTACTTA
ATTCAGTGCACAAATGAGATGAATGTGAACATCCCACAGTTGGCAGACAGTTTATTTGAAAGAACTACTAATAGT
AGTTGGGTGGTGGTCTTCAAATCTCTCATTACAACTCATCATTTGATGGTGTATGGAAATGAGCGTTTTATTCAG
TATTTGGCTTCAAGAAACACGTTGTTTAACTTAAGCAATTTTTTTGGATAAAAGTGGATTGCAAGGATATGACATG
TCTACATTTATTAGGCGGTATAGTAGATATTTAAATGAGAAAGCAGTTTCATACAGACAAGTTGCATTTGATTTC
ACAAAAGTGAAGAGAGGGGCTGATGGAGTTATGAGAACAATGAACACAGAAAAACTCCTAAAAACTGTACCAATT
ATTCAGAATCAGATGGATGCACTTCTTGATTTTAATGTTAATAGCAATGAACTTACAAATGGGGTAATAAATGCT
GCCTTCATGCTCCTGTTCAAAGATGCCATTAGACTGTTTGCAGCATACAATGAAGGAATTATTAATTTGTTGGAA
AAATATTTTGATATGAAAAAGAACCAATGCAAAGAAGGTCTTGACATCTATAAGAAGTTCCTAACTAGGATGACA
AGAATCTCAGAGTTCCTCAAAGTTGCAGAGCAAGTTGGAATTGACAGAGGTGATATACCAGACCTTTCACAGGCC
CCTAGCAGTCTTCTTGATGCTTTGGAACAACATTTAGCTTCCTTGGAAGGAAAGAAAATCAAAGATTCTACAGCT
GCAAGCAGGGCAACTACACTTTCCAATGCAGTGTCTTCCCTGGCAAGCACTGGTCTATCTCTGACCAAAGTGGAT
GAAAGGGAAAAGCAGGCAGCATTAGAGGAAGAACAGGCACGTTTGAAAGCTTTAAAGGAACAGCGCCTAAAAGAA
CTTGCAAAGAAACCTCATACCTCTTTAACAACTGCAGCCTCTCCTGTATCCACCTCAGCAGGAGGGATAATGACT
GCACCAGCCATTGACATATTTTCTACCCCTAGTTCTTCTAACAGCACATCAAAGCTGCCCAATGATCTGCTTGAT
TTGCAGCAGCCAACTTTTCACCCATCTGTACATCCTATGTCAACTGCTTCTCAGGTAGCAAGTACATGGGGAGAT
GCTGTTGATGATGCCATTCCAAGCTTAAATCCTTTCCTCACAAAAAGTAGTGGTGATGTTCACCTTTCCATTTCT
TCAGATGTATCTACTTTTACTACTAGGACACCTACTCATGAAATGTTTGTTGGATTCACTCCTTCTCCAGTTGCA
CAGCCACACCCTTCAGCTGGCCTTAATGTTGACTTTGAATCTGTGTTTGGAAATAAATCTACAAATGTTATTGTA
GATTCTGGGGGCTTTGATGAACTAGGTGGACTTCTCAAACCAACAGTGGCCTCTCAGAACCAGAACCTTCCTGTT
GCCAAACTCCCACCTAGCAAGTTAGTATCTGATGACTTGGATTCATCTTTAGCCAACCTTGTGGGCAATCTTGGC
ATCGGAAATGGAACCACTAAGAATGATGTAAATTGGAGTCAACCAGGTGAAAAGAAGTTAACTGGGGGATCTAAC
TGGCAACCAAAGGTTGCACCAACAACCGCTTGGAATGCTGCAACAATGGCACCCCCTGTAATGGCCTATCCTGCT
ACTACACCAACAGGCATGATAGGATATGGAATTCCTCCACAAATGGGAAGTGTTCCTGTAATGACGCAACCAACC
TTAATATACAGCCAGCCTGTCATGAGACCTCCAAACCCCTTTGGCCCTGTATCAGGAGCACAGAGATGTGAACTT
TGTCCCCATAAGGATGGAGCTTTAAAAAGAACAGATAATGGGGGTTGGGCCCATGTGGTTTGTGCCCTGTATATT
CCAGAGGTACAATTTGCCAATGTTTCCACAATGGAACCAATTGTTTTACAGTCTGTTCCGCATGATCGTTATAAT
AAGACTTGCTACATTTGTGATGAACAAGGAAGAGAAAGCAAAGCAGCCACTGGTGCTTGCxATGACATGTAATAA
ACATGGATGTCGACAGGCTTTCCATGTAACATGCGCTCAGTTTGCCGGACTGCTTTGTGAAGAAGAAGGTAATGG
TGCCGATAATGTCCAATACTGTGGCTACTGTAAATACCATTTTAGTAAGCTGAAAAAGAGCAAACGGGGATCTAA
TAGGTCATATGATCAAAGTTTAAGTGATTCTTCCTCTCACTCTCAGGATAAACATCATGAGAAAGAGAAAAAAA
ATATAAAGAGAAGGACAAACACAAACAGAAACACAAGAAGCAGCCAGAACCATCACCTGCATTGGTTCCATCCTT
GACTGTTACTACAGAAAAAACTTATACAAGCACTAGCAACAACTCTATATCTGGATCATTGAAGCGCTTGGAAGA
TACTACTGCACGATTTACAAATGCAAATTTCCAGGAAGTCTCTGCACACACCTCTAGTGGAAAAGATGTTTCAGA
GACTAGAGGGTCAGAGGGCAAAGGGAAGAAATCTTCAGCTCACAGCTCAGGTCAAAGGGGAAGAAAGCCTGGTGG
TGGAAGAAATCCAGGAACAACTGTGTCAGCAGCTAGCCCTTTTCCTCAAGGCAGTTTTTCAGGAACTCCAGGCAG
TGTAAAGTCATCTTCTGGAAGTTCAGTGCAGTCTCCCCAGGATTTCCTGAGCTTTACAGACTCAGATCTGCGTAA
TGACAGTTACTCTCACTCCCAACAGTCATCAGCAACCAAAGATGTACATAAAGGAGAGTCTGGAAGCCAGGAAGG

```
GGGGGTAAATAGTTTTAGTACCTTAATTGGCCTCCCTTCAACCTCAGCTGTTACTTCACAGCCTAAAAGCTTTGA
AAATTCACCTGGAGATTTGGGTAATTCCAGCCTTCCTACAGCAGGATATAAGCGGGCTCAAACTTCTGGCATAGA
AGAAGAAACTGTAAAGGAAAGAAAAGGAAAGGAAATAAACAAAGTAAGCATGGGCCTGGCAGACCCAAAGGAAA
CAAAAATCAAGAGAATGTTTCTCATCTCTCAGTTTCTTCTGCTTCACCAACATCATCTGTAGCATCAGCTGCAGG
AAGCATAACAAGCTCTAGTCTGCAGAAATCTCCTACATTGCTCAGGAATGGAAGTTTACAGAGCCTCAGTGTTGG
CTCATCTCCAGTTGGTTCAGAAATTTCCATGCAGTATCGGCATGATGGAGCTTGCCCAACAACTACGTTCTCAGA
GTTGCTGAATGCAATACACAACGGTATTTATAACAGCAATGATGTAGCAGTATCGTTTCCAAATGTAGTATCTGG
CTCGGGATCTAGTACTCCTGTCTCCAGCTCTCACTTACCTCAGCAGTCTTCTGGGCATTTGCAACAAGTAGGAGC
GCTCTCTCCCTCAGCTGTGTCATCTGCAGCCCCTGCTGTTGCTACAACTCAGGCAAATACTCTATCTGGATCTTC
TCTCAGTCAGGCACCATCTCATATGTATGGCAATAGATCAAATTCATCAATGGCAGCTCTTATAGCTCAGTCTGA
AAACAATCAAACAGATCAAGATCTTGGAGACAATAGCCGCAACCTAGTTGGCAGAGGAAGCTCACCCCGAGGAAG
TCTCTCGCCACGATCCCCTGTAAGCAGCTTACAGATTCGCTATGATCAACCAGGCAACAGCAGTTTGGAAAATCT
GCCTCCAGTAGCAGCCAGCATAGAACAGCTTTTGGAGAGGCAGTGGAGTGAAGGACAGCAATTTTTACTAGAACA
GGGTACTCCTAGTGACATTTTAGGAATGCTGAAGTCATTACACCAACTTCAAGTTGAAAACCGAAGATTAGAGGA
ACAAATTAAAAACTTGACTGCCAAAAAGGAACGGCTTCAGTTATTGAATGCACAGCTTTCAGTGCCTTTTCCAAC
AATAACAGCAAATCCTAGTCCGTCTCATCAAATACACACATTTTCAGCACAGACTGCTCCTACTACTGATTCCTT
GAACAGCAGTAAGAGCCCTCATATAGGAAACAGCTTTTTACCTGATAATTCTCTTCCTGTATTAAATCAGGACTT
AACCTCCAGTGGACAAAGTACCAGCAGCTCATCAGCTCTTTCTACCCCACCTCCTGCTGGGCAGAGTCCGGCTCA
ACAAGGCTCAGGAGTGAGTGGAGTTCAGCAGGTCAATGGCGTGACAGTGGGGGCACTAGCTAGTGGAATGCAGCC
TGTAACTTCCACCATTCCTGCCGTGTCTGCAGTGGGTGGAATAATTGGAGCTTTGCCAGGTAACCAACTGGCAAT
TAATGGCATTGTAGGAGCTTTAAATGGGGTTATGCAGACTCCTGTCACAATGTCCCAGAACCCTACCCCTCTCAC
CCACACAACCGTACCACCTAATGCAACACATCCAATGCCAGCTACACTGACTAACAGTGCCTCAGGACTAGGATT
ACTTTCTGACCAGCAACGACAAATACTTATTCATCAACAGCAGTTTCAGCAGTTGTTAAATTCTCAACAGCTCAC
ACCAGAACAACATCAAGCCTTTTTGTATCAGTTAATGCAACATCACCACCAGCAGCACCACCAACCTGAACTTCA
GCAGCTGCAGATCCCTGGACCAACACAAATACCCATAAACAACCTTCTTGCAGGTACACAGGCACCCCCACTTCA
CACAGCTACCACCAACCCATTTCTCACCATCCATGGAGATAATGCAAGTCAGAAAGTAGCAAGACTTAGTGATAA
AACTGGGCCTGTAGCTCAAGAGAAAAGTTGACACCTGAGAAACATCTAGAGGATCCAATTG
```

Appendix 2.2 – *PICALM-MLLT10* fusion sequence

The sequence of *PICALM-MLLT10* sequence is provided, with primers used to PCR-amplify the fusion from patient-derived cDNA highlighted in yellow (forward) and green (reverse). The entire coding sequence is 4,892 bp long. The truncated clathrin binding region of *PICALM* is shown in purple (corresponding to amino acid residues 414-641), and the full-length OM-LZ coding sequence of *MLLT10* is in green (residues 720-791). The entire PCR product size including primer sequences is 4,941 bp long. The start (ATG) and stop (TGA) codons are underlined.

Appendix 2.3 – *KMT2A-AFF1* fusion sequence

<u>ATG</u>GCGCACAGCTGTCGGTGGCGCTTCCCCGCCCGACCCGGGACCACCGGGGGCGGCGGCGGCGGGGGGCGCCGG
GGCCTAGGGGGCGCCCCGCGGCAACGCGTCCCGGCCCTGCTGCTTCCCCCCGGGCCCCGGTCGGCGGTGGCGGC
CCCGGGGCGCCCCCCTCCCCCCCGGCTGTGGCGGCCGCGGCGGCGGCGGCGGGAAGCAGCGGGGCTGGGGTTCCA
GGGGGAGCGGCCGCCGCCTCAGCAGCCTCCTCGTCGTCCGCCTCGTCTTCGTCTTCGTCATCGTCCTCAGCCTCT
TCAGGGCCGGCCCTGCTCCGGGTGGGCCCGGGCTTCGACGCGGCGCTGCAGGTCTCGGCCGCCATCGGCACCAAC
CTGCGCCGGTTCCGGGCCGTGTTTGGGGAGAGCGGCGGGGGAGGCGGCAGCGGAGAGGATGAGCAATTCTTAGGT
TTTGGCTCAGATGAAGAAGTCAGAGTGCGAAGTCCCACAAGGTCTCCTTCAGTTAAAACTAGTCCTCGAAAACCT
CGTGGGAGACCTAGAAGTGGCTCTGACCGAAATTCAGCTATCCTCTCAGATCCATCTGTGTTTTCCCCTCTAAAT
AAATCAGAGACCAAATCTGGAGATAAGATCAAGAAGAAAGATTCTAAAAGTATAGAAAAGAAGAGAGGAAGACCT
CCCACCTTCCCTGGAGTAAAAATCAAAATAACACATGGAAAGGACATTTCAGAGTTACCAAAGGGAAACAAAGAA
GATAGCCTGAAAAAAATTAAAAGGACACCTTCTGCTACGTTTCAGCAAGCCACAAAGATTAAAAAATTAAGAGCA
GGTAAACTCTCTCCTCTCAAGTCTAAGTTTAAGACAGGGAAGCTTCAAATAGGAAGGAAGGGGGTACAAATTGTA
CGACGGAGAGGAAGGCCTCCATCAACAGAAAGGATAAAGACCCCTTCGGGTCTCCTCATTAATTCTGAACTGGAA
AAGCCCCAGAAAGTCCGGAAAGACAAGGAAGGAACACCTCCACTTACAAAAGAAGATAAGACAGTTGTCAGACAA
AGCCCTCGAAGGATTAAGCCAGTTAGGATTATTCCTTCTTCAAAAAGGACAGATGCAACCATTGCTAAGCAACTC
TTACAGAGGGCAAAAAAGGGGGCTCAAAAGAAAATTGAAAAAGAAGCAGCTCAGCTGCAGGGAAGAAAGGTGAAG
ACACAGGTCAAAAATATTCGACAGTTCATCATGCCTGTTGTCAGTGCTATCTCCTCGCGGATCATTAAGACCCCT
CGGCGGTTTATAGAGGATGAGGATTATGACCCTCCAATTAAAATTGCCCGATTAGAGTCTACACCGAATAGTAGA
TTCAGTGCCCCGTCCTGTGGATCTTCTGAAAAATCAAGTGCAGCTTCTCAGCACTCCTCTCAAATGTCTTCAGAC
TCCTCTCGATCTAGTAGCCCCAGTGTTGATACCTCCACAGACTCTCAGGCTTCTGAGGAGATTCAGGTACTTCCT
GAGGAGCGGAGCGATACCCCTGAAGTTCATCCTCCACTGCCCATTTCCCAGTCCCCAGAAAATGAGAGTAATGAT
AGGAGAAGCAGAAGGTATTCAGTGTCGGAGAGAAGTTTTGGATCTAGAACGACGAAAAAATTATCAACTCTACAA
AGTGCCCCCCAGCAGCAGACCTCCTCGTCTCCACCTCCACCTCTGCTGACTCCACCGCCACCACTGCAGCCAGCC
TCCAGTATCTCTGACCACACACCTTGGCTTATGCCTCCAACAATCCCCTTAGCATCACCATTTTTGCCTGCTTCC
ACTGCTCCTATGCAAGGGAAGCGAAAATCTATTTTGCGAGAACCGACATTTAGGTGGACTTCTTTAAAGCATTCT
AGGTCAGAGCCACAATACTTTTCCTCAGCAAAGTATGCCAAAGAAGGTCTTATTCGCAAACCAATATTTGATAAT
TTCCGACCCCCTCCACTAACTCCCGAGGACGTTGGCTTTGCATCTGGTTTTTCTGCATCTGGTACCGCTGCTTCA
GCCCGATTGTTTTCGCCACTCCATTCTGGAACAAGGTTTGATATGCACAAAAGGAGCCCTCTTCTGAGAGCTCCA
AGATTTACTCCAAGTGAGGCTCACTCTAGAATATTTGAGTCTGTAACCTTGCCTAGTAATCGAACTTCTGCTGGA
ACATCTTCTTCAGGAGTATCCAATAGAAAAAGGAAAAGAAAAGTGTTTAGTCCTATTCGATCTGAACCAAGATCT
CCTTCTCACTCCATGAGGACAAGAAGTGGAAGGCTTAGTAGTTCTGAGCTCTCACCTCTCACCCCCCGTCTTCT
GTCTCTTCCTCGTTAAGCATTTCTGTTAGTCCTCTTGCCACTAGTGCCTTAAACCCAACTTTTACTTTTCCTTCT
CATTCCCTGACTCAGTCTGGGGAATCTGCAGAGAAAAATCAGAGACCAAGGAAGCAGACTAGTGCTCCGGCAGAG
CCATTTTCATCAAGTAGTCCTACTCCTCTCTTCCCTTGGTTTACCCCAGGCTCTCAGACTGAAAGAGGGAGAAAT
AAAGACAAGGCCCCCGAGGAGCTGTCCAAAGATCGAGATGCTGACAAGAGCGTGGAGAAGGACAAGAGTAGAGAG
AGAGACCGGGAGAGAGAAAAGGAGAATAAGCGGGAGTCAAGGAAAGAGAAAAGGAAAAAGGGATCAGAAATTCAG
AGTAGTTCTGCTTTGTATCCTGTGGGTAGGGTTTCCAAAGAGAAGGTTGTTGGTGAAGATGTTGCCACTTCATCT
TCTGCCAAAAAAGCAACAGGGCGGAAGAAGTCTTCATCACATGATTCTGGGACTGATATTACTTCTGTGACTCTT
GGGGATACAACAGCTGTCAAAACCAAAATACTTATAAAGAAAGGGAGAGGAAATCTGGAAAAAACCAACTTGGAC

CTCGGCCCAACTGCCCCATCCCTGGAGAAGGAGAAAACCCTCTGCCTTTCCACTCCTTCATCTAGCACTGTTAAA
CATTCCACTTCCTCCATAGGCTCCATGTTGGCTCAGGCAGACAAGCTTCCAATGACTGACAAGAGGGTTGCCAGC
CTCCTAAAAAAGGCCAAAGCTCAGCTCTGCAAGATTGAGAAGAGTAAGAGTCTTAAACAAACCGACCAGCCCAAA
GCACAGGGTCAAGAAAGTGACTCATCAGAGACCTCTGTGCGAGGACCCCGGATTAAACATGTCTGCAGAAGAGCA
GCTGTTGCCCTTGGCCGAAAACGAGCTGTGTTTCCTGATGACATGCCCACCCTGAGTGCCTTACCATGGGAAGAA
CGAGAAAAGATTTTGTCTTCCATGGGGAATGATGACAAGTCATCAATTGCTGGCTCAGAAGATGCTGAACCTCTT
GCTCCACCCATCAAACCAATTAAACCTGTCACTAGAAACAAGGCACCCCAGGAACCTCCAGTAAAGAAAGGACGT
CGATCGAGGCGGTGTGGGCAGTGTCCCGGCTGCCAGGTGCCTGAGGACTGTGGTGTTTGTACTAATTGCTTAGAT
AAGCCCAAGTTTGGTGGTCGCAATATAAAGAAGCAGTGCTGCAAGATGAGAAAATGTCAGAATCTACAATGGATG
CCTTCCAAAGCCTACCTGCAGAAGCAAGCTAAAGCTGTGAAAAAGAAAGAGAAAAAGTCTAAGACCAGTGAAAAG
AAAGACAGCAAAGAGAGCAGTGTTGTGAAGAACGTGGTGGACTCTAGTCAGAAACCTACCCCATCAGCAAGAGAG
GATCCTGCCCCAAAGAAAAGCAGTAGTGAGCCTCCTCCACGAAAGCCCGTCGAGGAAAAGAGTGAAGAAGGGAAT
GTCTCGGCCCCTGGGCCTGAATCCAAACAGGCCACCACTCCAGCTTCCAGGAAGTCAAGCAAGCAGGTCTCCCAG
CCAGCACTGGTCATCCCGCCTCAGCCACCTACTACAGGACCGCCAAGAAAAGAAGTTCCCAAAACCACTCCTAGT
GAGCCCAAGAAAAAGCAGCCTCCACCACCAGAATCAGGTCCAGAGCAGAGCAAACAGAAAAAAGTGGCTCCCCGC
CCAAGTATCCCTGTAAAACAAAAACCAAAAGAAAAGGAAAAACCACCTCCGGTCAATAAGCAGGAGAATGCAGGC
ACTTTGAACATCCTCAGCACTCTCTCCAATGGCAATAGTTCTAAGCAAAAAATTCCAGCAGATGGAGTCCACAGG
ATCAGAGTGGACTTTAAGCAGACCTACTCCAATGAAGTCCATTGTGTTGAAGAGATTCTGAAGGAAATGACCCAT
TCATGGCCGCCTCCTTTGACAGCAATACATACGCCTAGTACAGCTGAGCCATCCAAGTTTCCTTTCCCTACAAAG
GACTCTCAGCATGTCAGTTCTGTAACCCAAAACCAAAAACAATATGATACATCTTCAAAAACTCACTCAAATTCT
CAGCAAGGAACGTCATCCATGCTCGAAGACGACCTTCAGCTCAGTGACAGTGAGGACAGTGACAGTGAACAAACC
CCAGAGAAGCCTCCCTCCTCATCTGCACCTCCAAGTGCTCCACAGTCCCTTCCAGAACCAGTGGCATCAGCACAT
TCCAGCAGTGCAGAGTCAGAAAGCACCAGTGACTCAGACAGTTCCTCAGACTCAGAGAGCGAGAGCAGTTCAAGT
GACAGCGAAGAAAATGAGCCCCTAGAAACCCCAGCTCCGGAGCCTGAGCCTCCAACAACAAACAAATGGCAGCTG
GACAACTGGCTGACCAAAGTCAGCCAGCCAGCTGCGCCACCAGAGGGCCCCAGGAGCACAGAGCCCCCACGGCGG
CACCCAGAGAGTAAGGGCAGCAGCGACAGTGCCACGAGTCAGGAGCATTCTGAATCCAAAGATCCTCCCCCTAAA
AGCTCCAGCAAAGCCCCCCGGGCCCCACCCGAAGCCCCCCACCCCGGAAAGAGGAGCTGTCAGAAGTCTCCGGCA
CAGCAGGAGCCCCCACAAAGGCAAACCGTTGGAACCAAACAACCCAAAAAACCTGTCAAGGCCTCTGCCCGGGCA
GGTTCACGGACCAGCCTGCAGGGGGAAAGGGAGCCAGGGCTTCTTCCCTATGGCTCCCGAGACCAGACTTCCAAA
GACAAGCCCAAGGTGAAGACGAAAGGACGGCCCCGGGCCGCAGCAAGCAACGAACCCAAGCCAGCAGTGCCCCCC
TCCAGTGAGAAGAAGAAGCACAAGAGCTCCCTCCCTGCCCCCTCTAAGGCTCTCTCAGGCCCAGAACCCGCGAAG
GACAATGTGGAGGACAGGACCCCTGAGCACTTTGCTCTTGTTCCCCTGACTGAGAGCCAGGGCCCACCCCACAGT
GGCAGCGGCAGCAGGACTAGTGGCTGCCGCCAAGCCGTGGTGGTCCAGGAGGACAGCCGCAAAGACAGACTCCCA
TTGCCTTTGAGAGACACCAAGCTGCTCTCACCGCTCAGGGACACTCCTCCCCCACAAAGCTTGATGGTGAAGATC
ACCCTAGACCTGCTCTCTCGGATACCCCAGCCTCCCGGGAAGGGGAGCCGCCAGAGGAAAGCAGAAGATAAACAG
CCGCCCGCAGGGAAGAAGCACAGCTCTGAGAAGAGGAGCTCAGACAGCTCAAGCAAGTTGGCCAAAAAGAGAAAG
GGTGAAGCAGAAAGAGACTGTGATAACAAGAAATCAGACTGGAGAAGGAAATCAAATCACAGTCATCTTCATCT
TCATCCTCCCACAAAGAATCTTCTAAAACAAAGCCCTCCAGGCCCTCCTCACAGTCCTCAAAGAAGGAAATGCTC
CCCCCGCCACCCGTGTCCTCGTCCTCCCAGAAGCCAGCCAAGCCTGCACTTAAGAGGTCAAGGCGGGAAGCAGAC
ACCTGTGGCCAGGACCCTCCCAAAAGTGCCAGCAGTACCAAGAGCAACCACAAAGACTCTTCCATTCCCAAGCAG
AGAAGAGTAGAGGGGAAGGGCTCCAGAAGCTCCTCGGAGCACAAGGGTTCTTCCGGAGATACTGCAAATCCTTTT

CCAGTGCCTTCTTTGCCAAATGGTAACTCTAAACCAGGGAAGCCTCAAGTGAAGTTTGACAAACAACAAGCAGAC
CTTCACATGAGGGAGGCAAAAAAGATGAAGCAGAAAGCAGAGTTAATGACGGACAGGGTTGGAAAGGCTTTTAAG
TACCTGGAAGCCGTCTTGTCCTTCATTGAGTGCGGAATTGCCACAGAGTCTGAAAGCCAGTCATCCAAGTCAGCT
TACTCTGTCTACTCAGAAACTGTAGATCTCATTAAATTCATAATGTCATTAAAATCCTTCTCAGATGCCACAGCG
CCAACACAAGAGAAAATATTTGCTGTTTTATGCATGCGTTGCCAGTCCATTTTGAACATGGCGATGTTTCGTTGT
AAAAAAGACATAGCAATAAAGTATTCTCGTACTCTTAATAAACACTTCGAGAGTTCTTCCAAAGTCGCCCAGGCA
CCTTCTCCATGCATTGCAAGAAGCACAGGCACACCATCCCCTCTTTCCCCAATGCCTTCTCCTGCCAGCTCCGTA
GGGTCCCAGTCAAGTGCTGGCAGTGTGGGGAGCAGTGGGGTGGCTGCCACTATCAGCACCCCAGTCACCATCCAG
AATATGACATCTTCCTATGTCACCATCACATCCCATGTTCTTACCGCCTTTGACCTTTGGGAACAGGCCGAGGCC
CTCACGAGGAAGAATAAAGAATTCTTTGCTCGGCTCAGCACAAATGTGTGCACCTTGGCCCTCAACAGCAGTTTG
GTGGACCTGGTGCACTATACACGACAGGGTTTTCAGCAGCTACAAGAATTAACCAAAACACCTTAA

Appendix 2.3 – *KMT2A-AFF1* fusion sequence

The sequence of *KMT2A-AFF1* sequence is provided. The entire coding sequence is 6,816 bp long. The start (ATG) and stop (TAA) codons are underlined.

Chapter 3 – *DDX3X-MLLT10, PICALM-MLLT10* and *KMT2A-AFF1* confer distinct molecular characteristics *in vitro*

3.1 Introduction

Genome-wide comparisons of ALL leukaemic blasts have revealed that the presence of a *KMT2A*r is associated with distinct gene expression[1,2] and epigenetic[3] profiles. Most notably, upregulation of the *HOXA* cluster genes *HOXA7-10* and *HOX* cofactor *MEIS1* are well-characterised features of *KMT2A*r acute leukaemia[4]. However, the aetiology and precise molecular mechanisms underlying *KMT2A*r-mediated leukaemogenesis are still poorly understood. *KMT2A*r AML was recently successfully modelled in CD34[+] human cord blood cells[5], and *KMT2A*r B-ALL in CD34[+] human foetal liver cells[6], but no reliable transgenic models of *KMT2A*r T-ALL currently exist. A transgenic model of *KMT2A*r T-ALL would further our understanding of the biology and therapeutic sensitivities of *KMT2A*r T-ALL, as these models enable the study of genomic lesions in isolation.

In addition to *KMT2A*r, there are other ALL-associated fusion genes that have not yet been successfully modelled. *MLLT10*r T-ALL is a poorly characterised entity, despite accounting for approximately 10% of newly diagnosed T-ALL cases[7]. Wild-type *KMT2A* and *MLLT10* have closely related functions in epigenetic regulation of gene expression, and *MLLT10*r T-ALL cases share the *HOXA* dysregulated profile exhibited by *KMT2A*r[8,9]. Unlike *KMT2A*r, *MLLT10*r are never identified in B-ALL (except for *KMT2A-MLLT10*)[10], and appear to only induce T-ALL, AML, or on rare occasions MPAL[9]. As with *KMT2A*r, a pre-clinical model of *MLLT10*r T-ALL does not yet exist. A transgenic murine model of *PICALM-MLLT10* yielded disease with a predominantly myeloid phenotype and occasional aberrant T-cell marker expression[11]. A clear T-ALL phenotype was not observed in any mouse, despite expression of *PICALM-MLLT0* in the thymus as well as the bone marrow under the *vav* promoter. Further investigation into the aetiology of *MLLT10*r T-ALL would therefore improve the understanding of how genomic aberrations in epigenetic regulatory genes can induce acute leukaemia with different immunophenotypes.

3.2 Approach

This chapter characterises and compares the phenotypic changes induced by expression of *DDX3X-MLLT10*, *PICALM-MLLT10* and *KMT2A-AFF1* in the MOHITO cell line, and *KMT2A-AFF1* expression in the Ba/F3 cell line. The Ba/F3 and MOHITO cell lines are both murine model systems used to interrogate the oncogenic potential and associated downstream signalling pathways of cancer-associated genomic aberrations[12,13]. *KMT2A-AFF1* was studied as it is the most common *KMT2A*r in ALL, and *DDX3X-MLLT10* and *PICALM-MLLT10* were studied as they are the two most common *MLLT10*r identified in ALL.

DDX3X-MLLT10 and *PICALM-MLLT10* were identified in the leukaemic blasts of patients diagnosed with T-ALL. The fusion genes were PCR-amplified from patient-derived cDNA and cloned into the pCR®-Blunt II-TOPO® amplification vector, and then subcloned into the pRUFiG2 retroviral expression vector. *KMT2A-AFF1*-pRUFiG2 was commercially synthesised due to technical cloning difficulties. The resulting constructs were transduced into the CD4$^+$CD8$^+$ T-cell cell line MOHITO, and *KMT2A-AFF1* was also expressed in pro-B Ba/F3 cells. *DDX3X-MLLT10* and *PICALM-MLLT10* were not investigated in the Ba/F3 cell line, as these fusions do not occur in B-ALL.

In these cell lines, the molecular phenotype of *KMT2A*r and *MLLT10*r T-ALL were characterised and compared. The effect of host cell lineage on *KMT2A-AFF1*-mediated leukaemogenesis was also assessed, where B-ALL and T-ALL cell lines expressing *KMT2A-AFF1* were compared. Cell lines expressing each fusion were investigated by luminescence-based cell viability assays to assess cytokine-independent and dependent changes in cell proliferation. Differential expression of *HOXA* cluster genes was assessed by qRT-PCR, and surface marker staining was performed to determine cell immunophenotype and differentiation stage. Activation of kinase-associated signalling pathways was investigated by phospho-flow cytometry.

3.3 Results

3.3.1 *KMT2A-AFF1* expression promotes increased cytokine-dependent proliferation in MOHITO cells

Using molecular subcloning, the *DDX3X-MLLT10* and *PICALM-MLLT10* fusion genes were successfully PCR amplified from patient cDNA (**Figure 3.1**) and ligated into the pCR®-Blunt II-TOPO® amplification vector. The fusions were then isolated from the pCR®-Blunt II-TOPO® amplification vector and subcloned into the mammalian-expressing pRUFiG2 retroviral vector. *KMT2A-AFF1* could not be PCR-amplified due to its large size and G-C rich 5' region, so it was obtained as a commercially synthesised vector. Presence of the full-length fusion was confirmed by restriction enzyme digest (**Figure 3.2A**) and Sanger sequencing was performed across the entire length of each fusion sequence (breakpoints are shown in **Figure 3.2B**).

Vectors were transfected into HEK-293T cells with the pEQ-Eco packaging vector to generate retroviral particles. Wild-type MOHITO and Ba/F3 cells were exposed to viral supernatant, resulting in successful transduction based on GFP analysis by flow cytometry (**Figure 3.3**). GFP$^+$ cells were collected by FACS and cultured for ongoing experiments. The transduction efficiency of *KMT2A-AFF1*-pRUFiG2 was low due to the large size of the fusion construct, and consequently cells were re-sorted by FACS to obtain a predominantly GFP$^+$ population (**Figure 3.3E** and **G**). Following flow cytometric confirmation of >90% GFP$^+$ cells, RNA was extracted and RT-PCR performed across each fusion's breakpoint to confirm expression (**Figure 3.4**).

CellTiter-Glo® luminescent cell viability assays were performed on MOHITO (**Figure 3.5**) and Ba/F3 (**Figure 3.6**) cell lines to quantify ATP levels, as a measure of metabolic activity. Assays were performed both in the presence and absence of the cytokines required for each cell line's growth (IL-2 and IL-7 for MOHITO cells, and IL-3 for Ba/F3 cells). In the presence of cytokines, there were no significant differences in the proliferation rate of MOHITO cells expressing empty vector, *DDX3X-MLLT10* or *PICALM-MLLT10*, doubling 2.9 (±0.2), 2.8 (±0.1) and 2.9 (±0.2) times over four days respectively. MOHITO cells expressing *KMT2A-AFF1* exhibited a significant increase in proliferation in the presence of IL-2 and IL-7, doubling 4.0 (±0.4) times (p=0.004, compared to empty vector cells). Positive control MOHITO cells expressing the known pathogenic variant *NRAS*^{G12D} doubled 4.1 (±0.3) times (p=0.003, compared to empty

vector) (**Figure 3.5A**). However, expression of *KMT2A-AFF1* in Ba/F3 cells did not induce any cytokine-dependent changes in proliferation (**Figure 3.6A**). In the absence of cytokine, empty vector MOHITO and Ba/F3 cell lines lost all detectable metabolic activity within 48 hours (**Figures 3.5B** and **3.6B** respectively). Positive control MOHITO and Ba/F3 cell lines expressing the pathogenic *NRAS*^G12D variant doubled 3.7 (±0.2) and 1.8 (±0.4) times respectively in the absence of cytokines (p<0.01 for both MOHITO and Ba/F3 cell lines) (**Figures 3.5B** and **3.6B**). No fusion was capable of promoting cytokine-independent proliferation in either MOHITO or Ba/F3 cells.
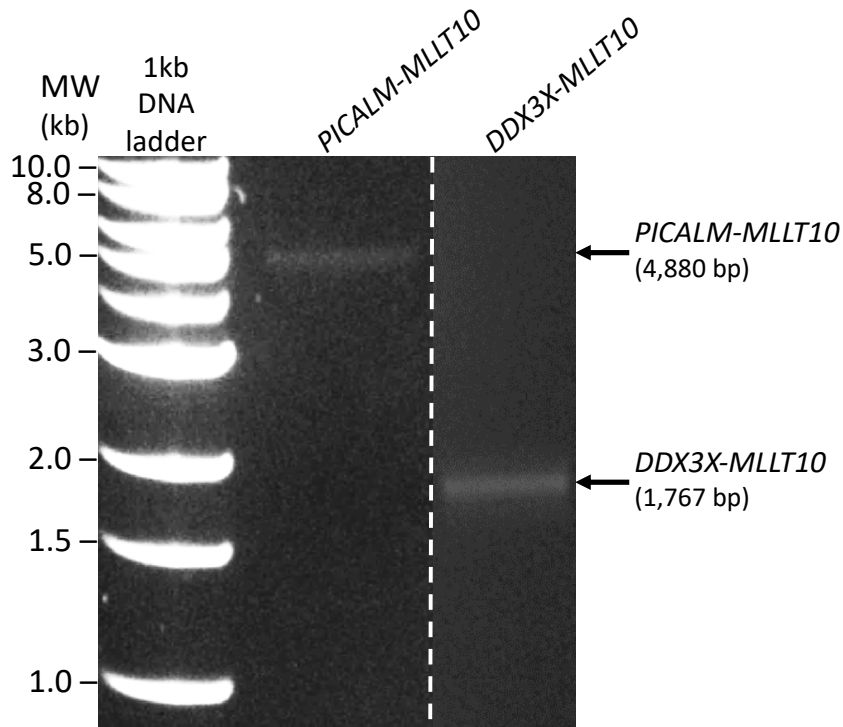
Figure 3.1 – Isolation of *PICALM-MLLT10* and *DDX3X-MLLT10* fusion genes from patient cDNA

DNA agarose gels of purified PCR products. *PICALM-MLLT10* (left, 4,880 bp) and *DDX3X-MLLT10* (right, 1,767 bp) were amplified from patient material and purified, then electrophoresed to verify successful purification prior to subcloning into the pCR®-Blunt II-TOPO® amplification vector. Abbreviations: MW, molecular weight marker.

(A)

(B)

(i) *DDX3X-MLLT10*

(ii) *PICALM-MLLT10*

(iii) *KMT2A-AFF1*
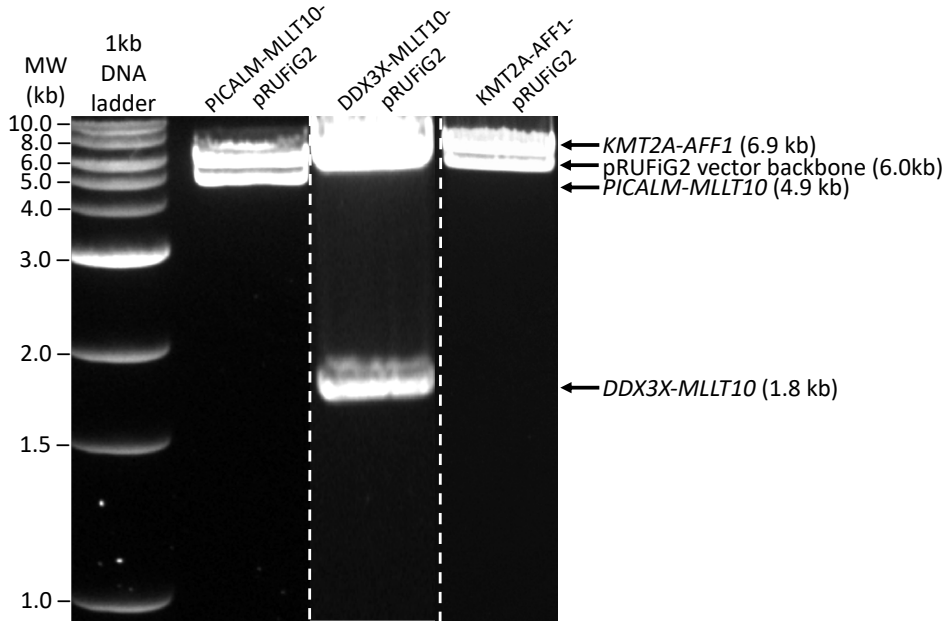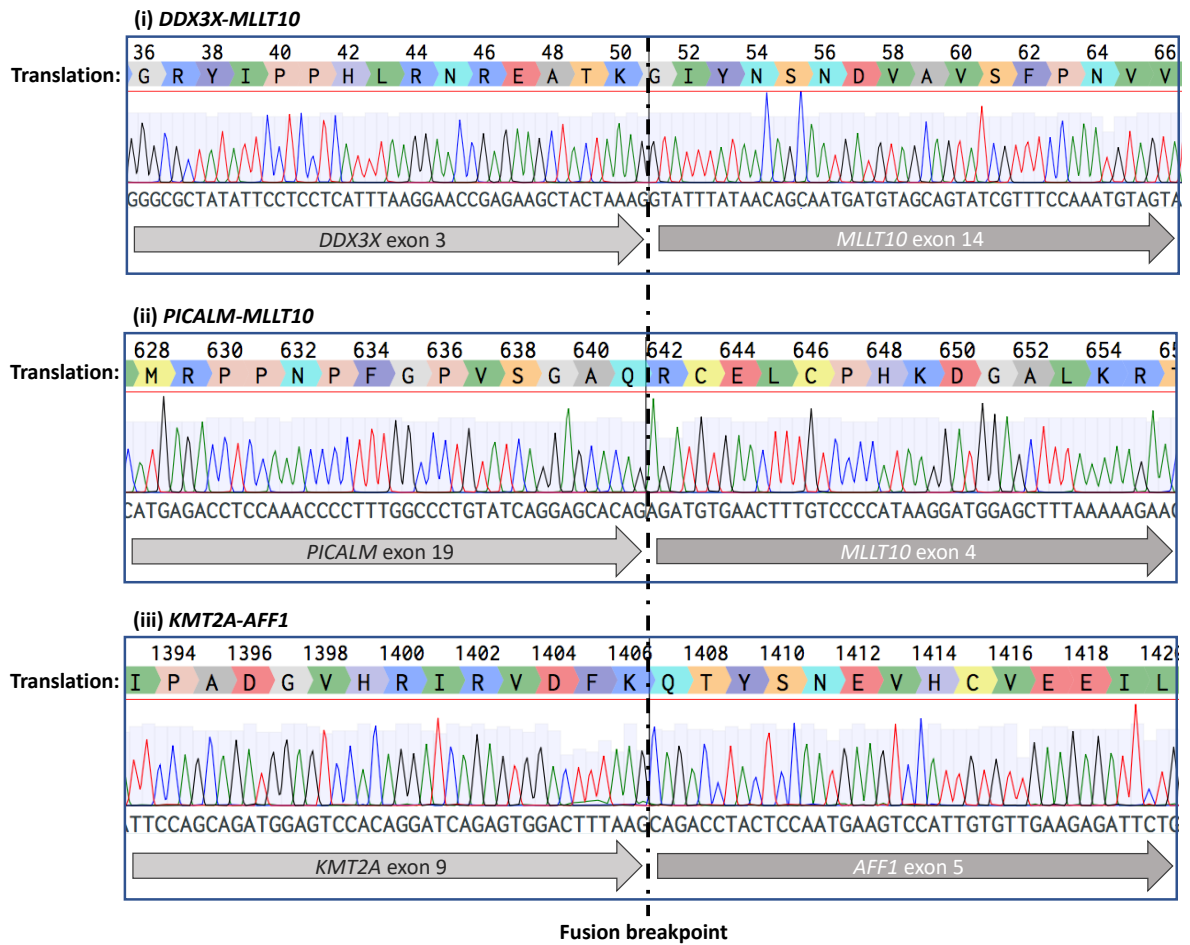
Fusion breakpoint

Figure 3.2 – Restriction digest and Sanger sequencing confirm presence of *DDX3X-MLLT10*, *PICALM-MLLT10* and *KMT2A-AFF1* fusion genes in pRUFiG2 retroviral expression vector

(**A**) DNA agarose gels and (**B**) Sanger sequencing of fusion breakpoints confirm the presence of the correct fusion sequence within the pRUFiG2 expression vector. Abbreviations: MW, molecular weight marker.

(A) **Flow cytometry gating strategy**
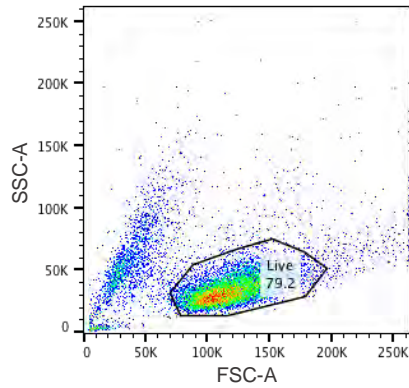
(B) **pRUFiG2 empty vector MOHITO**
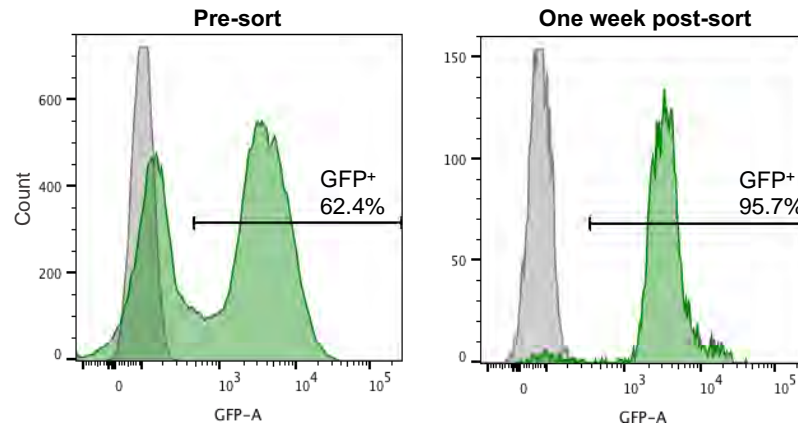
(C) ***DDX3X-MLLT10* MOHITO**
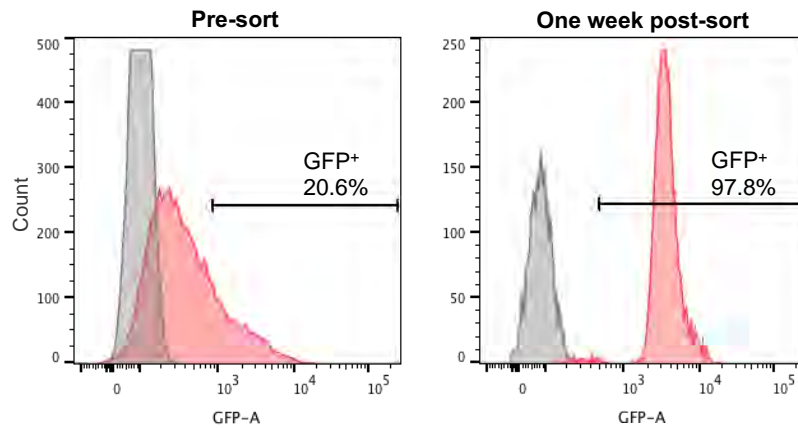
(D) ***PICALM-MLLT10* MOHITO**
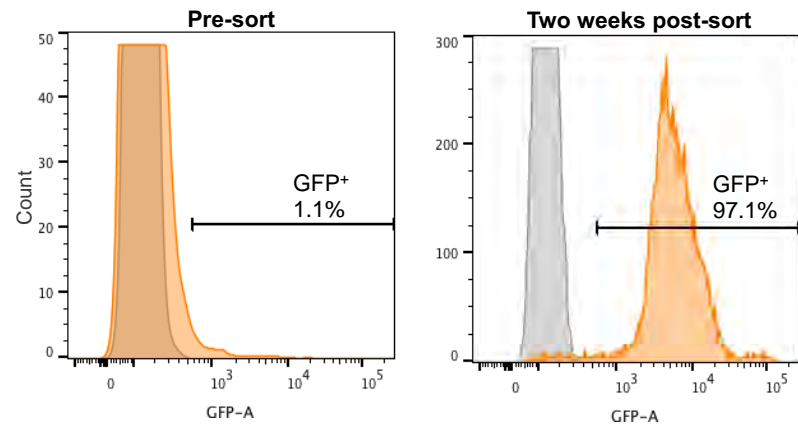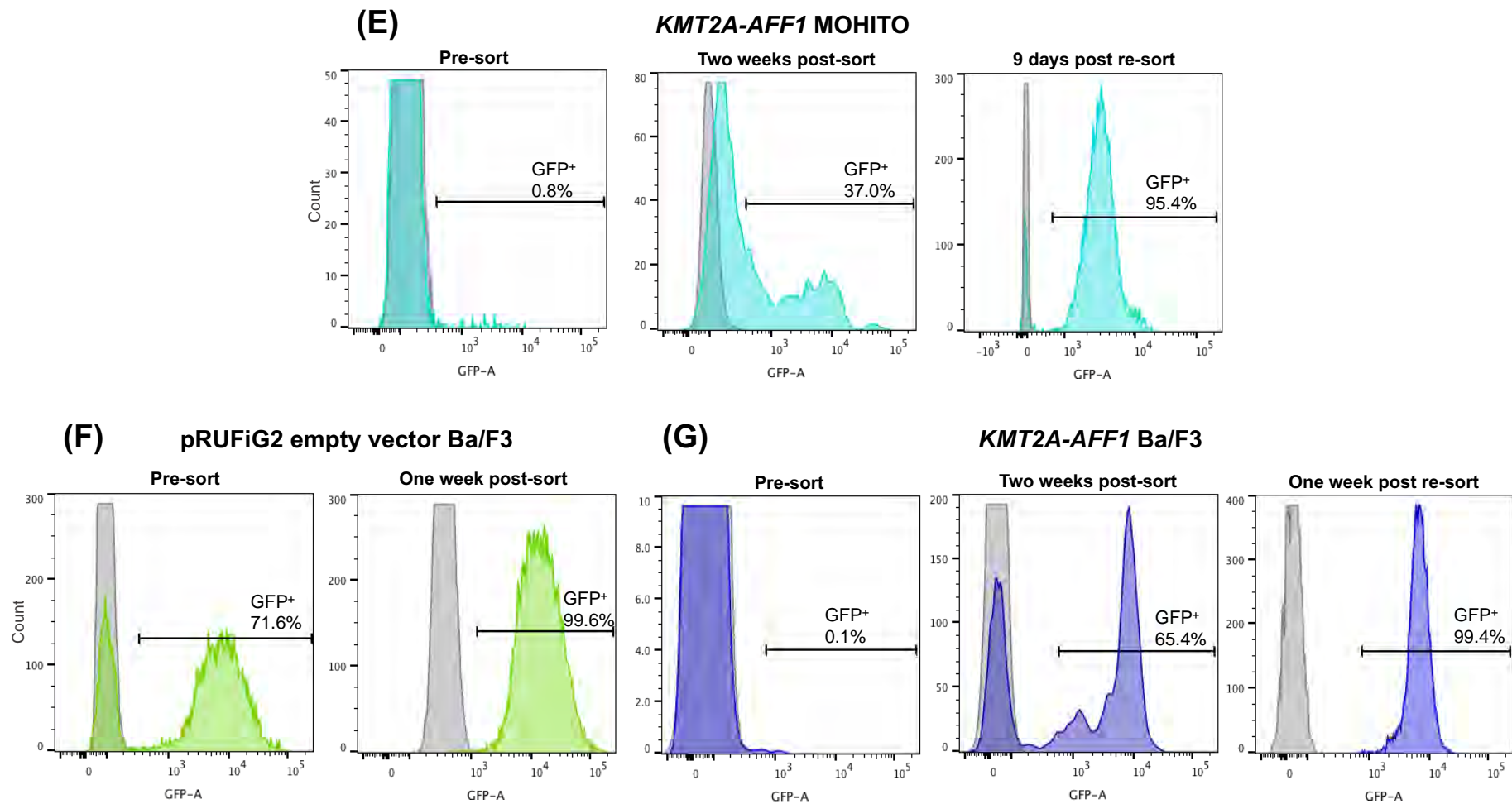
Figure 3.3 – MOHITO and Ba/F3 cells transduced with fusion-pRUFiG2 constructs express GFP

MOHITO cells were analysed and sorted based on GFP positivity using the BD FACSMelody™ approximately 1 week following retroviral transduction. GFP analysis was then repeated approximately 1-2 weeks post-sorting, to determine the efficiency of FACS.

Expression of GFP indicates successful integration and expression of retroviral constructs. (**A**) Gating strategy used for flow cytometry analysis and sorting of GFP⁺ cells. (**B-E**) GFP analysis of MOHITO cells transduced with (**B**) empty pRUFiG2 vector, (**C**) *DDX3X-MLLT10*-pRUFiG2, (**D**) *PICALM-MLLT10*-pRUFiG2 and (**E**) *KMT2A-AFF1*-pRUFiG2. (**F-G**) GFP analysis of Ba/F3 cells transduced with (**F**) empty pRUFiG2 vector and (**G**) *KMT2A-AFF1*-pRUFiG2.
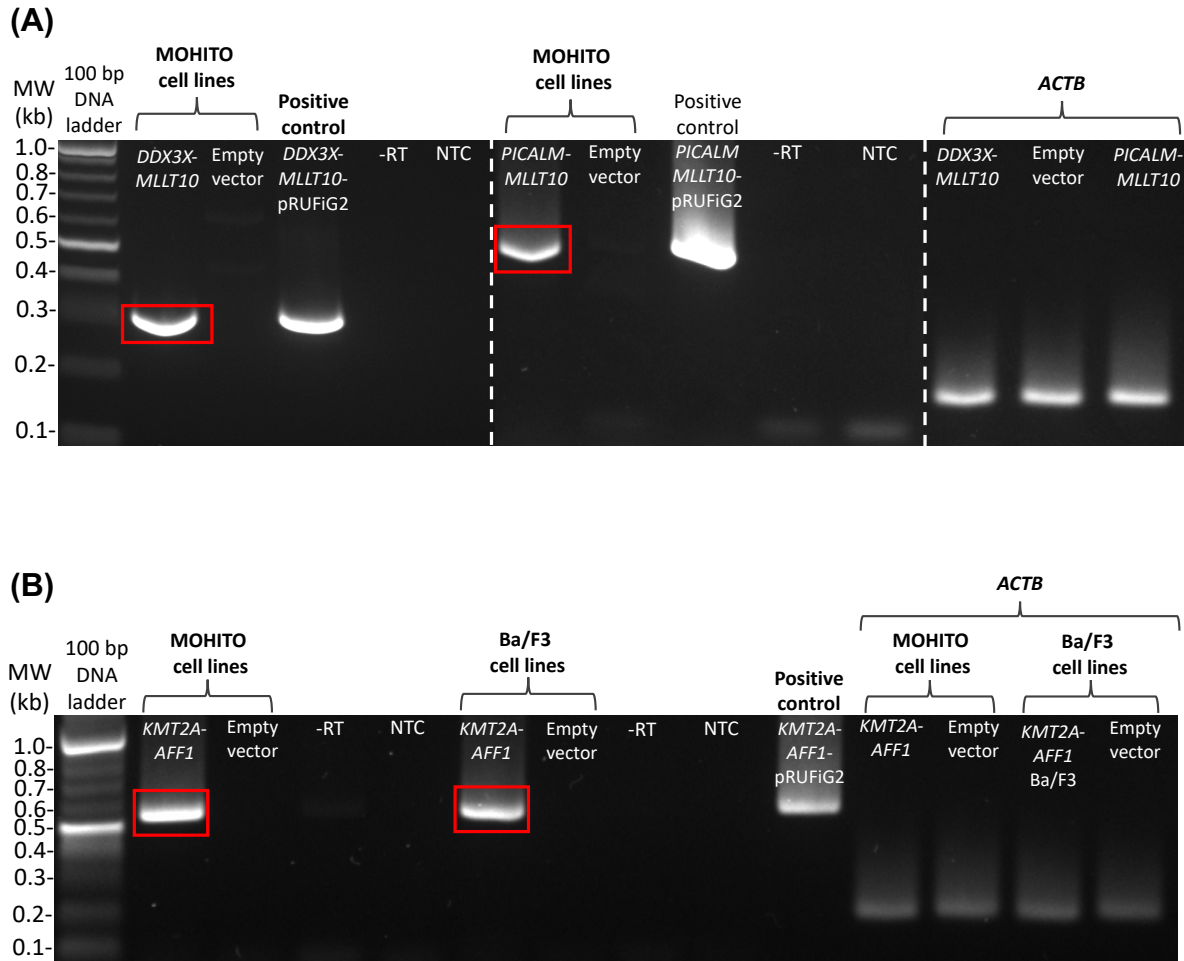
Figure 3.4 – MOHITO and Ba/F3 cells transduced with pRUFiG2 constructs express the correct fusion genes

Gel electrophoresis of RT-PCR products from (**A**) *PICALM-MLLT10* and *DDX3X-MLLT10* expressing MOHITO cells, and (**B**) *KMT2A-AFF1* expressing MOHITO and Ba/F3 cells, using primers specific for each fusion gene. Red rectangles highlight the presence of each fusion gene in their respective cell lines. *ACTB* was used as a positive control for PCR. Abbreviations: MW, molecular weight marker; -RT, no reverse transcriptase control; NTC, no template control; *ACTB*, actin.
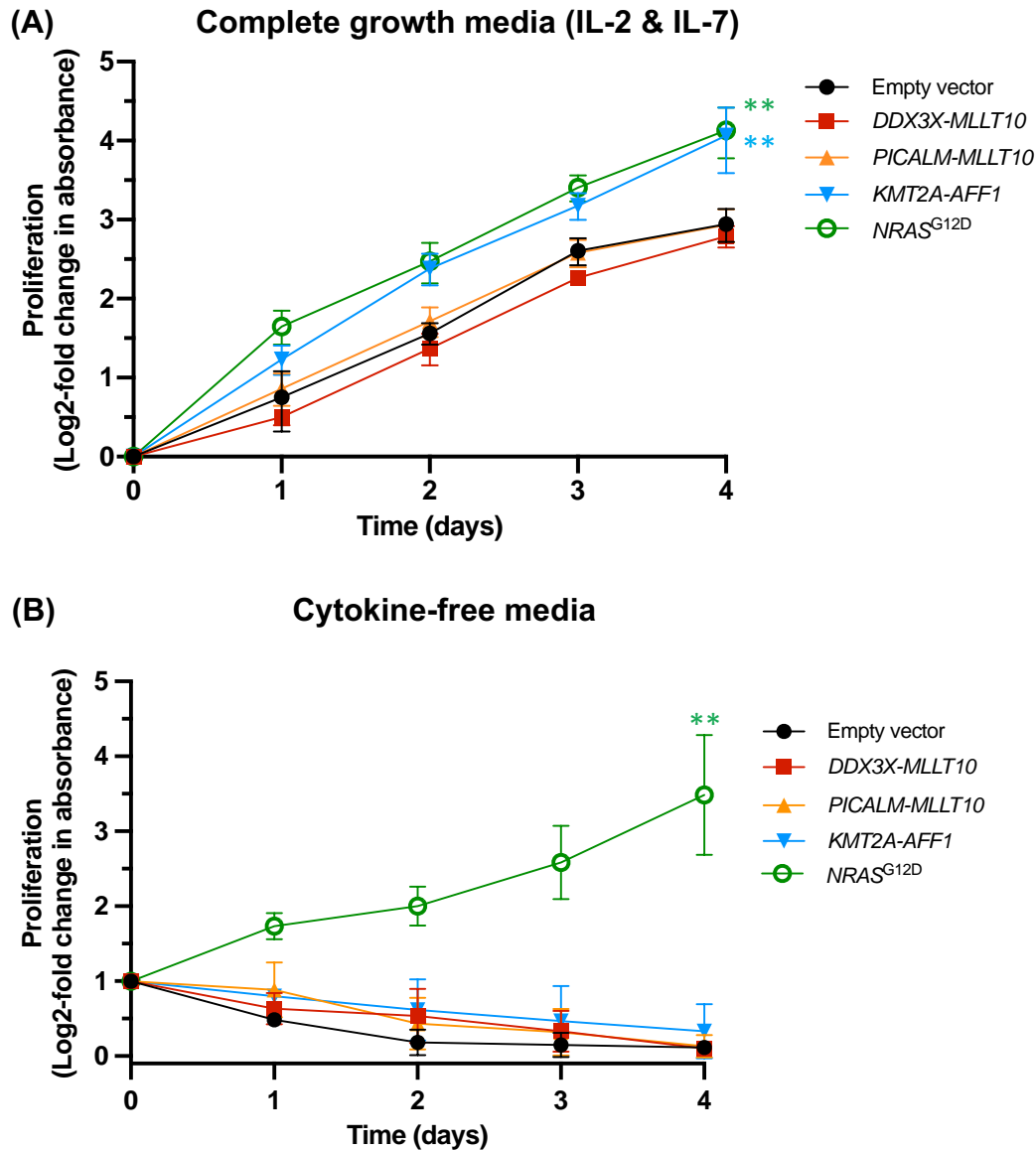
Figure 3.5 – Expression of *KMT2A-AFF1,* but not *MLLT10*r, promotes a cytokine-dependent increase in proliferation of MOHITO cells

MOHITO cell lines were cultured in technical duplicate for 5 days **(A)** in the presence of growth-supporting cytokines IL-2 and IL-7, and **(B)** in cytokine-free media. Cellular proliferation was quantified daily with the CellTiter-Glo® cell luminescent viability kit. Cytokine-independent MOHITO cells expressing *NRAS*[G12D] were used as a positive control. Change in proliferation is expressed as mean log2-fold change in absorbance, as measured by a PerkinElmer 2030 VICTOR X4 multilabel plate reader. All statistics represent one-way ANOVA, and statistically significant p-values are denoted by asterisks (**p<0.01, n.s.=not significant). Error bars represent mean ± SD of 3 biological replicates.
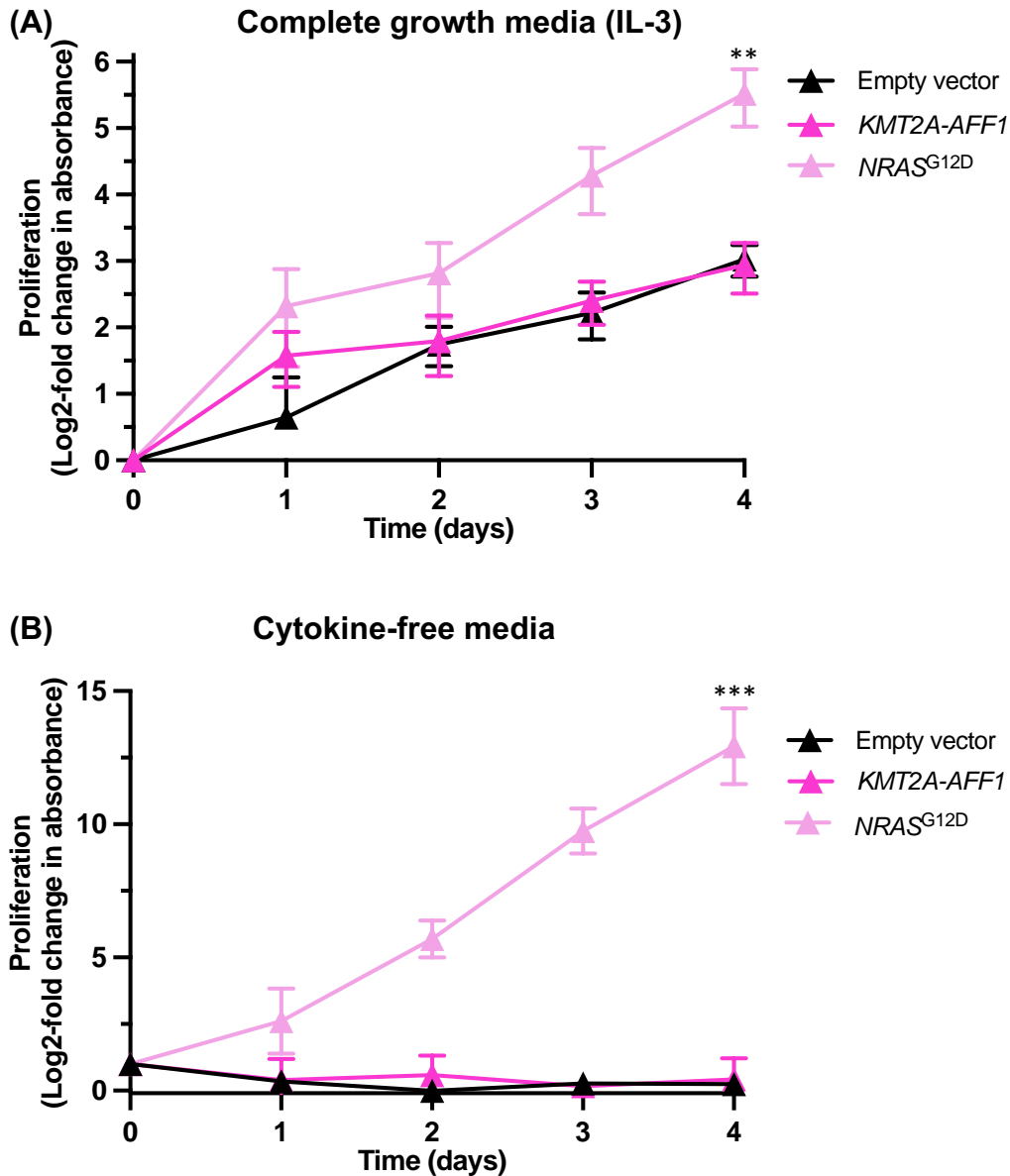
Figure 3.6 – Expression of *KMT2A-AFF1* does not alter proliferation of Ba/F3 cells

Ba/F3 cell lines were cultured for 4 days **(A)** in the presence of growth-supporting cytokine (IL-3) and **(B)** in cytokine-free media. Cellular proliferation was quantified daily using the CellTiter-Glo® cell luminescence viability kit. Change in proliferation is expressed as mean log2-fold change in absorbance, as measured by a PerkinElmer 2030 VICTOR X4 multilabel plate reader. Cytokine-independent Ba/F3 cells expressing *NRAS*G12D were used as a positive control. All statistics represent one-way ANOVA, and statistically significant p-values are denoted by asterisks (**p<0.01, ***p<0.001, n.s.=not significant). Error bars represent mean ± SD of 3 biological replicates.

3.3.2   The repertoire of *HOXA* gene cluster dysregulation induced by *KMT2A-AFF1* expression depends on parental cell lineage

HOXA gene cluster dysregulation is a hallmark feature of *KMT2A*r and *MLLT10*r B-ALL and AML[9,14], and subsequently this is a reliable validation tool for *KMT2A*r and *MLLT10*r models. Using SYBR® Green qRT-PCR analysis, differential expression levels of several *HOXA* gene cluster members were quantified.

No changes in *HOXA* gene expression were observed between wild-type MOHITO cells and cells expressing *DDX3X-MLLT10* (**Figure 3.7A**) or *PICALM-MLLT10* (**Figure 3.7B**). Increased expression of *HOXA9* was observed in Ba/F3 cells expressing *KMT2A-AFF1* (p=0.007) (**Figure 3.8A**). *KMT2A-AFF1* MOHITO cells exhibited significantly increased expression of *HOXA3, HOXA5* and *MEIS1* (p=0.003, <0.001 and 0.001 respectively) compared with wild-type (**Figure 3.8B**). Significant differences in expression levels of *HOXA3*, *HOXA5*, *HOXA9* and *MEIS1* (p=0.003, 0.01, 0.009 and 0.004 respectively) were observed between wildtype Ba/F3 and MOHITO cells and cells expressing *KMT2A-AFF1* (**Figure 3.8C**).
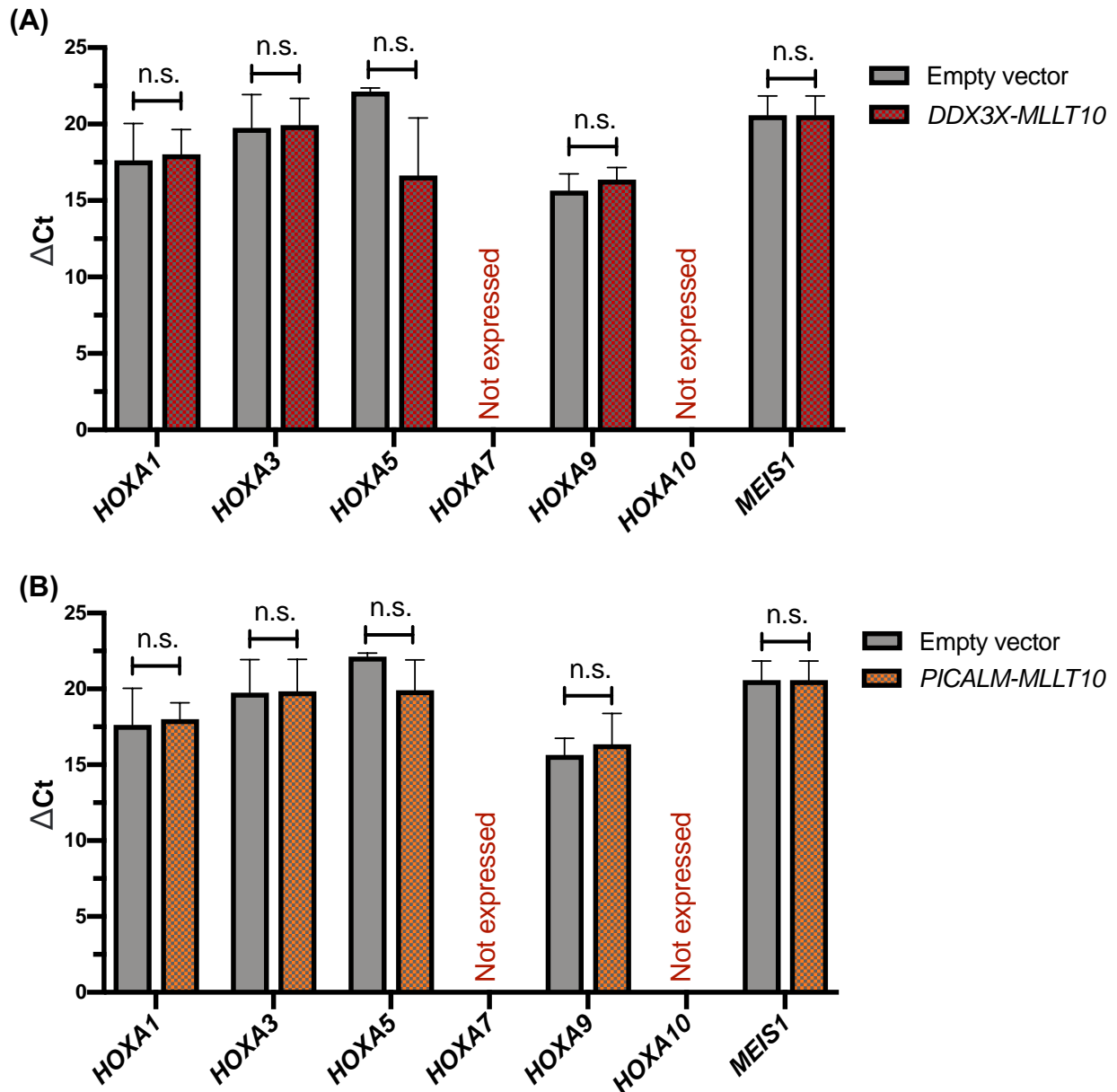
Figure 3.7 – Expression of *MLLT10*r in MOHITO cells does not induce changes in *HOXA* gene cluster expression

RNA was extracted from MOHITO cells (pRUFiG2 empty vector, *DDX3X-MLLT10* and *PICALM-MLLT10*), and cDNA synthesised for SYBR® Green qRT-PCR analysis of *HOXA* gene cluster members. Comparison of ΔCt values between empty vector and **(A)** *DDX3X-MLLT10* and **(B)** *PICALM-MLLT10* are shown*. ΔCt values were calculated by normalising Ct values to the *ACTB* housekeeping gene. A lower ΔCt value indicates higher expression. All statistics represent Student's t-tests (n.s.=not significant). Error bars represent mean ± SD of at least 3 biological replicates.
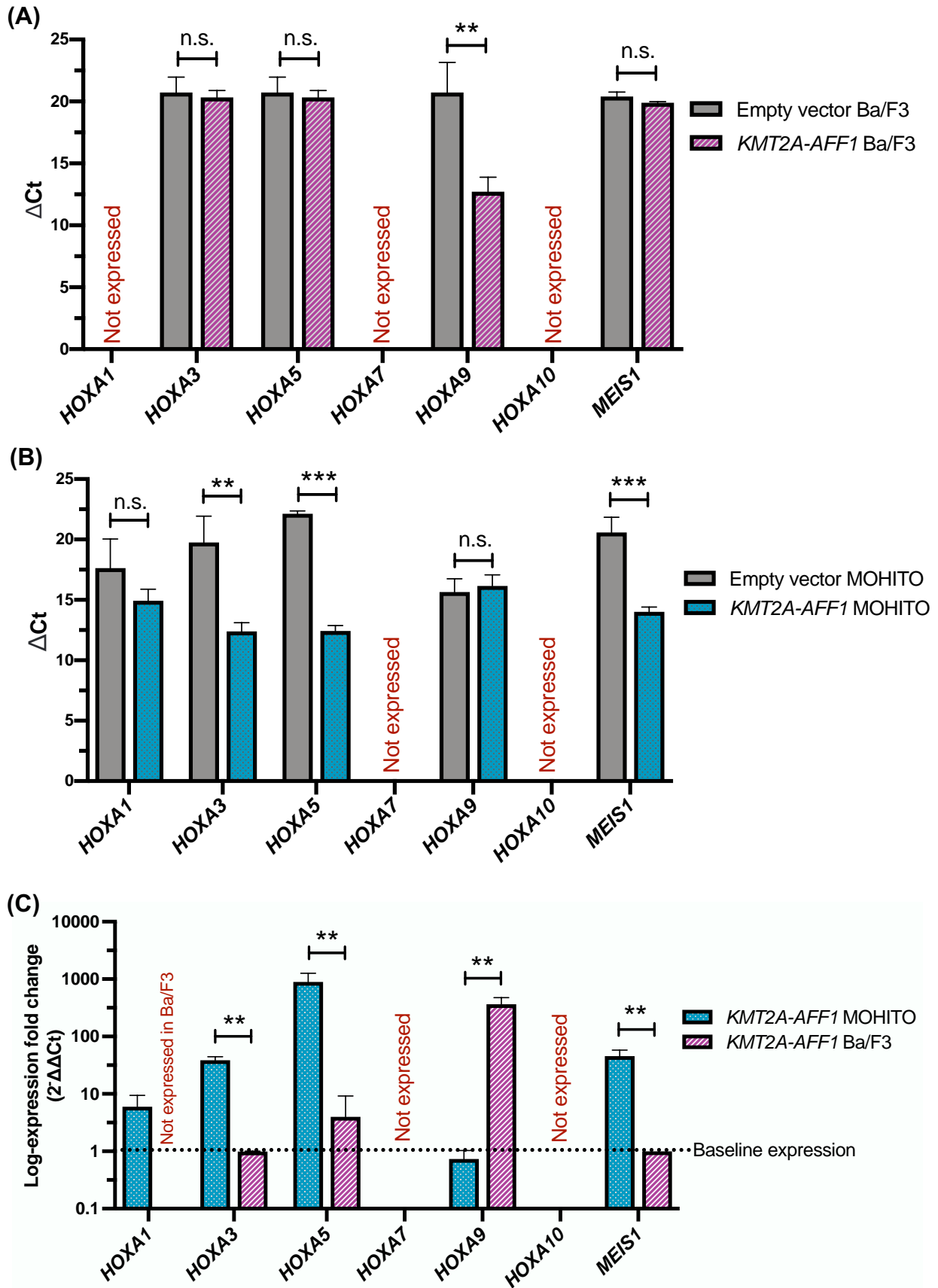
Figure 3.8 – Expression of *KMT2A-AFF1* confers a distinct *HOXA* gene expression profile in B-ALL and T-ALL

RNA was extracted from MOHITO and Ba/F3 cells (pRUFiG2 empty vector and *KMT2A-AFF1*), and cDNA synthesised for SYBR® Green qRT-PCR analysis of *HOXA* gene cluster members. **(A-B)** Comparison of ΔCt (cycle threshold) values between empty vector and *KMT2A-AFF1* expressing cells are shown for **(A)** MOHITO and **(B)** Ba/F3 cell lines, normalised to *ACTB* housekeeping gene. **(C)** Comparison of MOHITO and Ba/F3 *KMT2A-AFF1* cell lines are displayed as log-expression fold change ($2^{-\Delta\Delta Ct}$). All statistics represent Student's t-tests, and statistically significant p-values are denoted by asterisks (**$p<0.01$, ***$p<0.001$, n.s.=not significant). Error bars represent mean ± SD of at least 3 biological replicates.

3.3.3 Expression of *DDX3X-MLLT10*, *PICALM-MLLT10* or *KMT2A-AFF1* induces unique changes in surface marker expression

The role of *HOXA* dysregulation in *KMT2A*r and *MLLT10*r acute leukaemia suggests that cellular differentiation may play an important role in leukaemogenesis. Surface marker staining was performed on fusion-expressing Ba/F3 and MOHITO cell lines to explore cellular differentiation. All gating strategies were determined based on isotype control samples (**Supplementary Figure 3.1**). Wild-type and *KMT2A-AFF1* expressing Ba/F3 cells were analysed for expression of the pan-haematopoietic surface marker CD45, and B-cell associated markers CD43, CD19, B220, CD24 and BP-1. Wild-type Ba/F3 cells exhibited a pro/pre-B immunophenotype ($CD45^+CD43^+B220^{dim}CD19^-CD24^-BP-1^{dim}$)[15,16] (**Figure 3.9**). A significant reduction in B220 expression was noted in the *KMT2A-AFF1* Ba/F3 cell line (p=0.015). No expression of the myeloid marker CD13, or stem cell markers c-Kit, Sca-1 or CD34 was observed in either cell line (**Figure 3.10**).

Parental MOHITO cells exhibited a T-cell $CD45^+CD25^+CD44^{dim}CD2^+CD8a^+CD4^+CD5^-sCD3^-$ phenotype. c-Kit and Sca-1 were expressed, but not CD34 or CD13. The immunophenotype of wild-type MOHITO cells was compared to MOHITO cells expressing *DDX3X-MLLT10, PICALM-MLLT10* or *KMT2A-AFF1*. All MOHITO cell lines retained a $CD45^+CD2^+sCD3^-$ phenotype (**Figures 3.11-3.12**). Strong CD25 and weak CD44 expression was observed across all cell lines, but *KMT2A-AFF1* expressing cells had significantly greater CD44 expression compared with wild-type (p=0.04) (**Figure 3.13**). Marked loss of CD4 and CD8 was also observed in *KMT2A-AFF1* cells (p<0.0001, compared to wild-type), but not in cells expressing *MLLT10*r (**Figure 3.14**). CD5 expression was significantly reduced in *PICALM-MLLT10* cells compared with empty vector (p=0.04) (**Figure 3.15**). The haematopoietic stem cell markers c-Kit and Sca-1, but not CD34, were expressed in all cell lines, but Sca-1 expression was reduced in all three fusions compared to wild-type MOHITO cells (p=0.02, 0.005 and 0.007 respectively) (**Figures 3.15-3.16**).
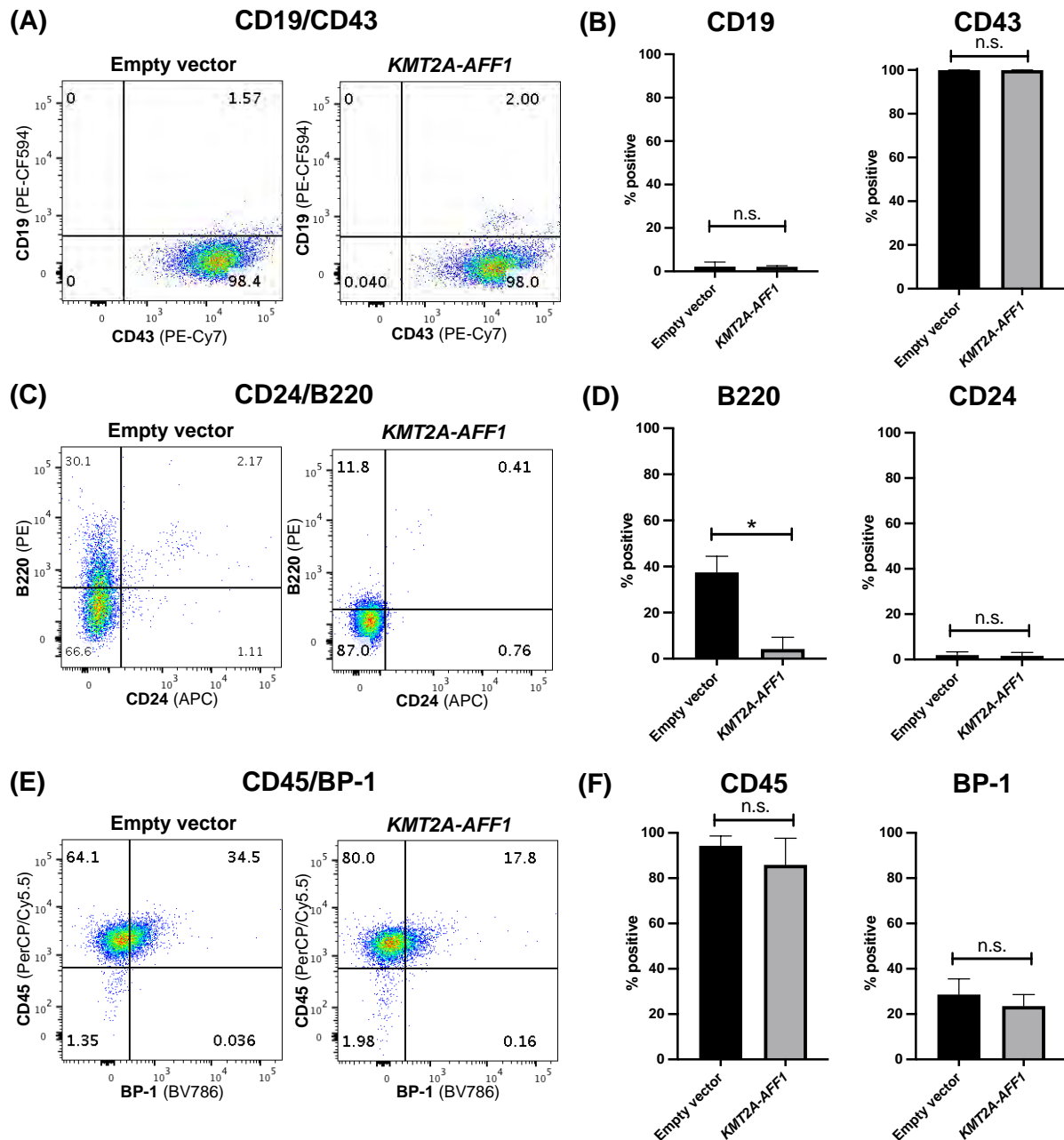
Figure 3.9 – *KMT2A-AFF1* expressing Ba/F3 cells retain a pro-B immunophenotype, with reduced B220 expression

Cell surface marker expression was analysed on Ba/F3 cells (pRUFiG2 empty vector and *KMT2A-AFF1*) on a BD LSRFortessa™. Scatter plots and mean percentage positive (% positive) values for **(A, B)** CD19/CD43, **(C, D)** CD24/B220, and **(E, F)** CD45/BP-1 dual stained samples. Percentage positive values for each cell line were calculated as a proportion of live GFP$^+$ cells and statistically compared by Student's t-test (n.s.=not significant). Scatter plots are representative of 3 biological replicates. Error bars represent mean ± SD of 3 biological replicates.
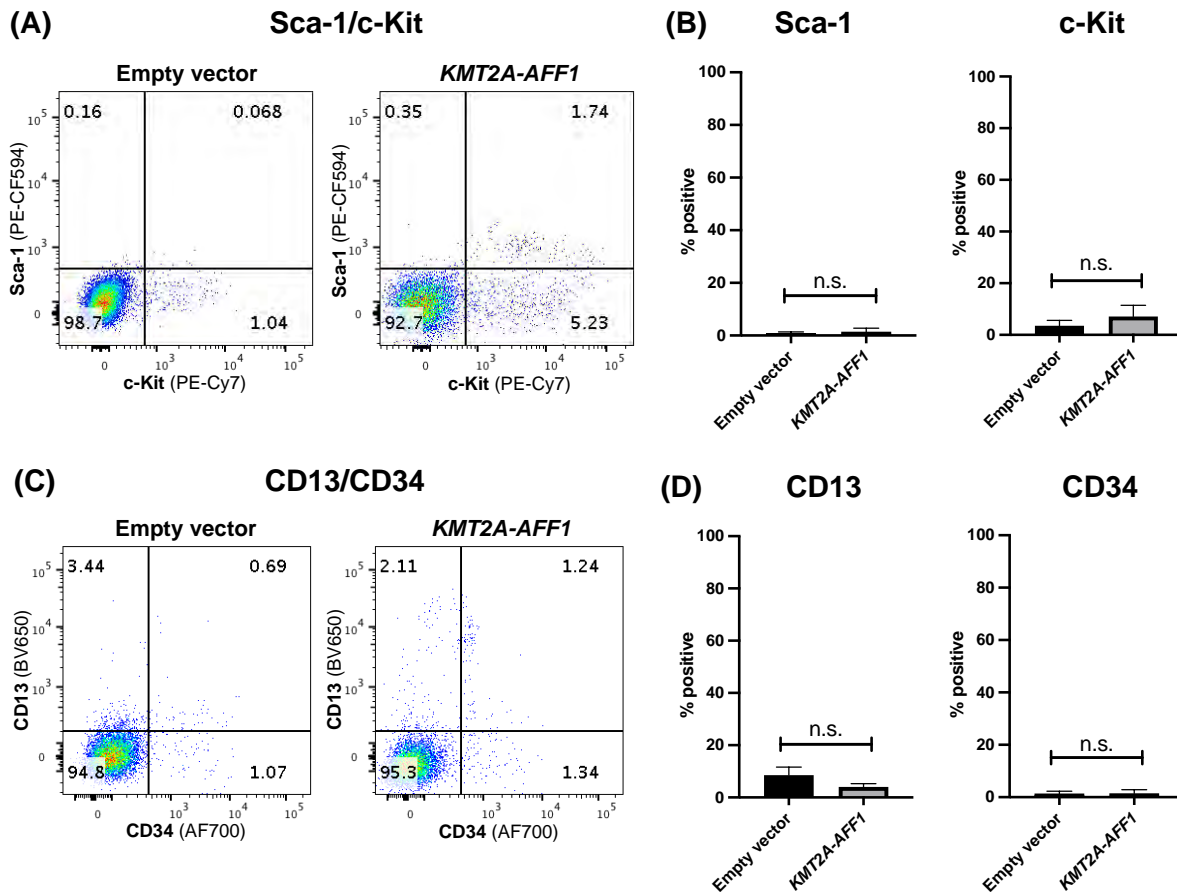
Figure 3.10 – Ba/F3 cells expressing empty vector or *KMT2A-AFF1* do not express the myeloid marker CD13, or stem cell markers c-Kit, Sca-1 or CD34

Cell surface marker expression was analysed on Ba/F3 cells (pRUFiG2 empty vector and *KMT2A-AFF1*) on a BD LSRFortessa™. **(A, C)** Scatter plots and mean percentage positive (% positive) values for **(A, B)** Sca-1/c-Kit and **(C, D)** CD13/CD34 dual stained samples. Percentage positive values for each cell line were calculated as a proportion of live GFP⁺ cells and statistically compared by Student's t-test (n.s.=not significant). Scatter plots are representative of 3 biological replicates. Error bars represent mean ± SD of 3 biological replicates.
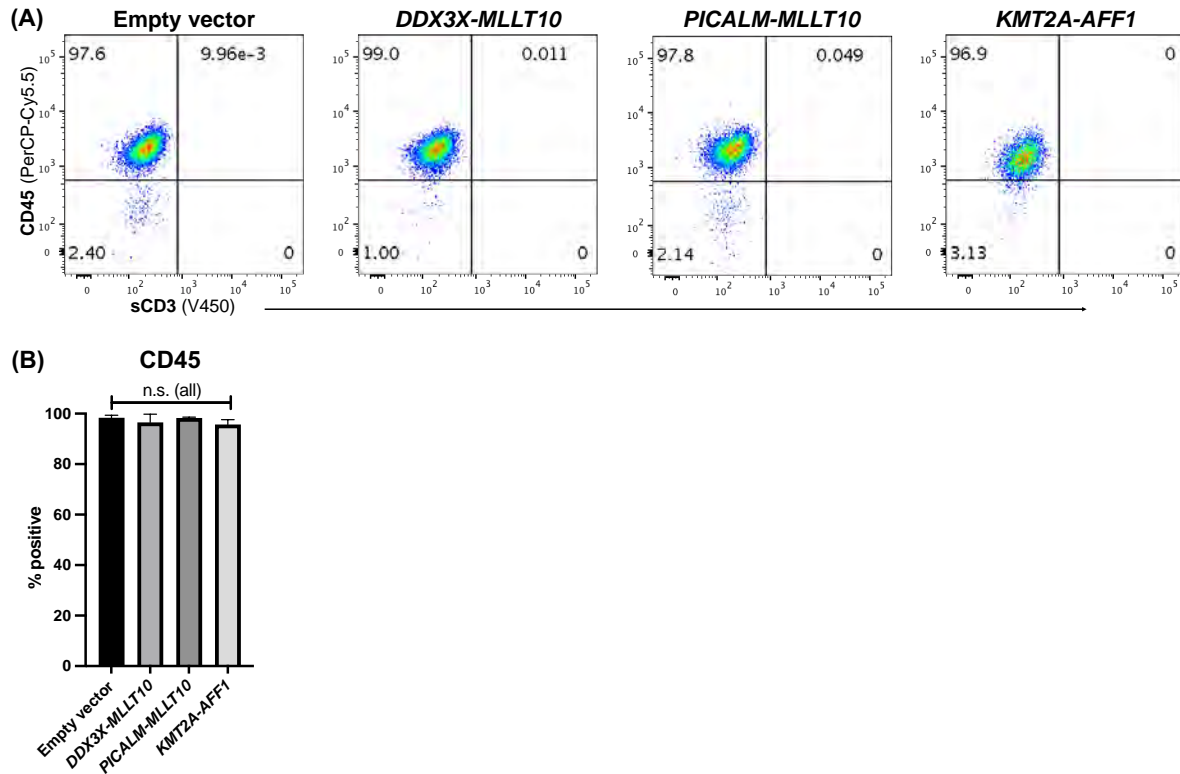
Figure 3.11 – MOHITO cell lines expressing empty vector, *DDX3X-MLLT10, PICALM-MLLT10* or *KMT2A-AFF1* exhibit a CD45+sCD3- immunophenotype

Cell surface marker staining was performed on MOHITO cells (pRUFiG2 empty vector, *DDX3X-MLLT10, PICALM-MLLT10* and *KMT2A-AFF1*) on a BD LSRFortessa™. **(A)** Scatter plots for CD45/sCD3 dual stained samples. **(B)** Percentage positive (% positive) values for each cell line were calculated as a proportion of live GFP+ cells and statistically compared by one-way ANOVA (n.s.=not significant). sCD3 graph not shown as all values were <1%. Scatter plots are representative of at least 3 biological replicates, and gates are set based on isotype controls. Error bars represent mean ± SD of 3 biological replicates.
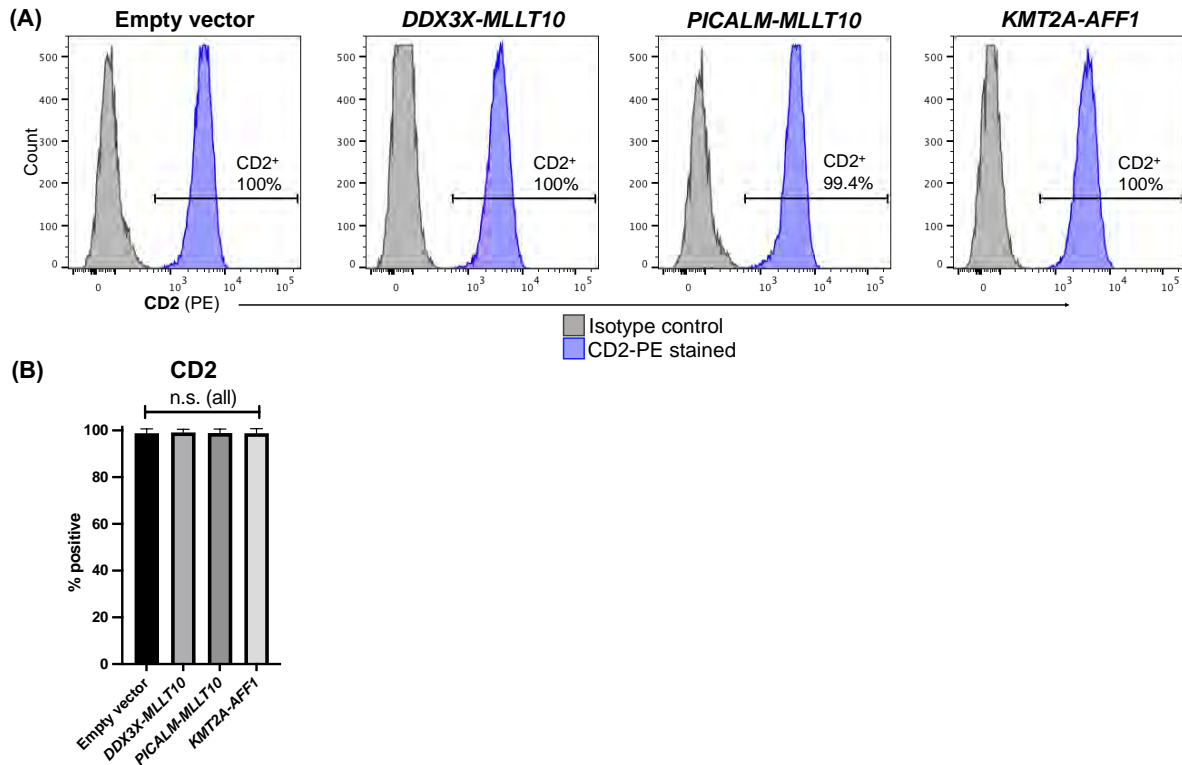
Figure 3.12 – MOHITO cell lines expressing empty vector, *DDX3X-MLLT10, PICALM-MLLT10* or *KMT2A-AFF1* express the T-cell marker CD2

Cell surface marker staining was performed on MOHITO cells (pRUFiG2 empty vector, *DDX3X-MLLT10, PICALM-MLLT10* and *KMT2A-AFF1*) on a BD LSRFortessa™. **(A)** Histograms for CD2 and isotype control single-stained samples. **(B)** Percentage positive (% positive) values for each cell line were calculated as a proportion of live GFP⁺ cells and statistically compared by one-way ANOVA (n.s.=not significant). Histograms are representative of at least 3 biological replicates, and gates are set based on isotype controls. Error bars represent mean ± SD of 3 biological replicates.
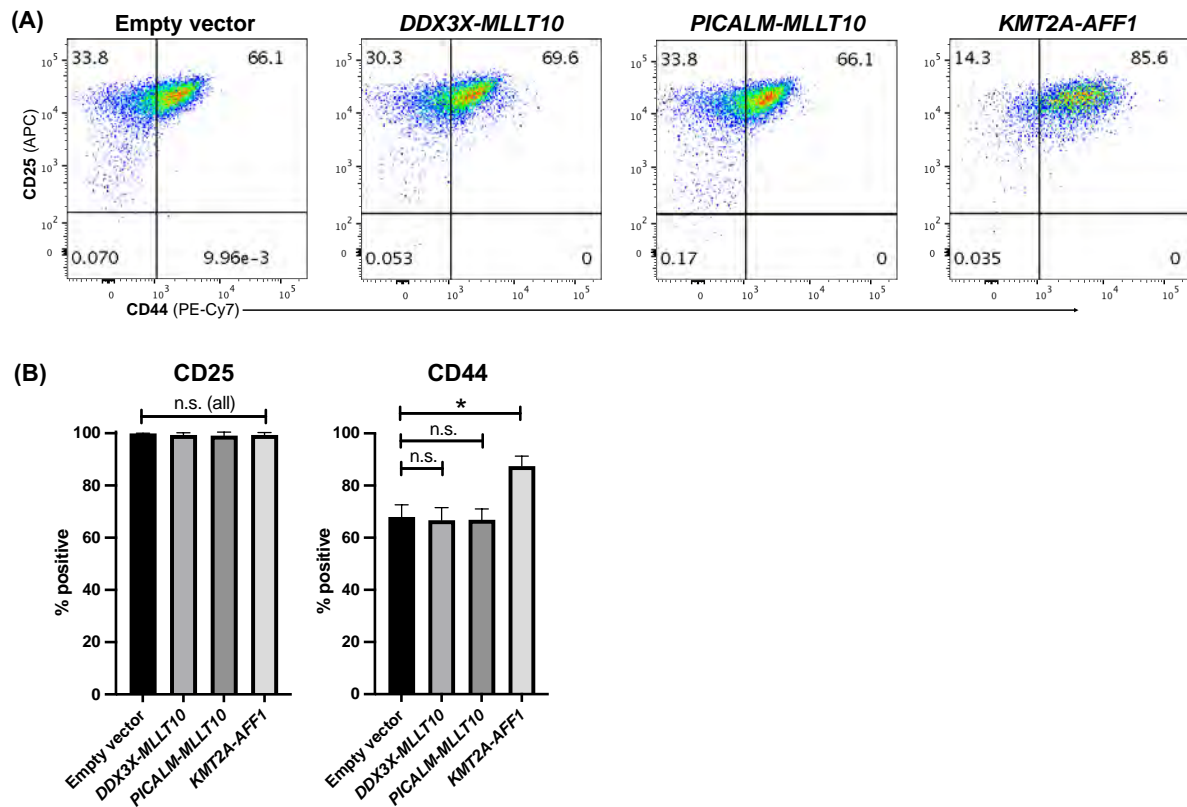
Figure 3.13 – Expression of *KMT2A-AFF1* in MOHITO cells promotes increased CD44 expression

Cell surface marker staining was performed on MOHITO cells (pRUFiG2 empty vector, *DDX3X-MLLT10, PICALM-MLLT10* and *KMT2A-AFF1*) on a BD LSRFortessa™. **(A)** Scatter plots for CD25/CD44 dual stained samples. **(B)** Percentage positive (% positive) values for each cell line were calculated as a proportion of live GFP⁺ cells and statistically compared by one-way ANOVA. Statistically significant values are denoted with asterisks (*p<0.05, n.s.=not significant). Scatter plots are representative of at least 3 biological replicates, and gates are set based on isotype controls. Error bars represent mean ± SD of 3 biological replicates.
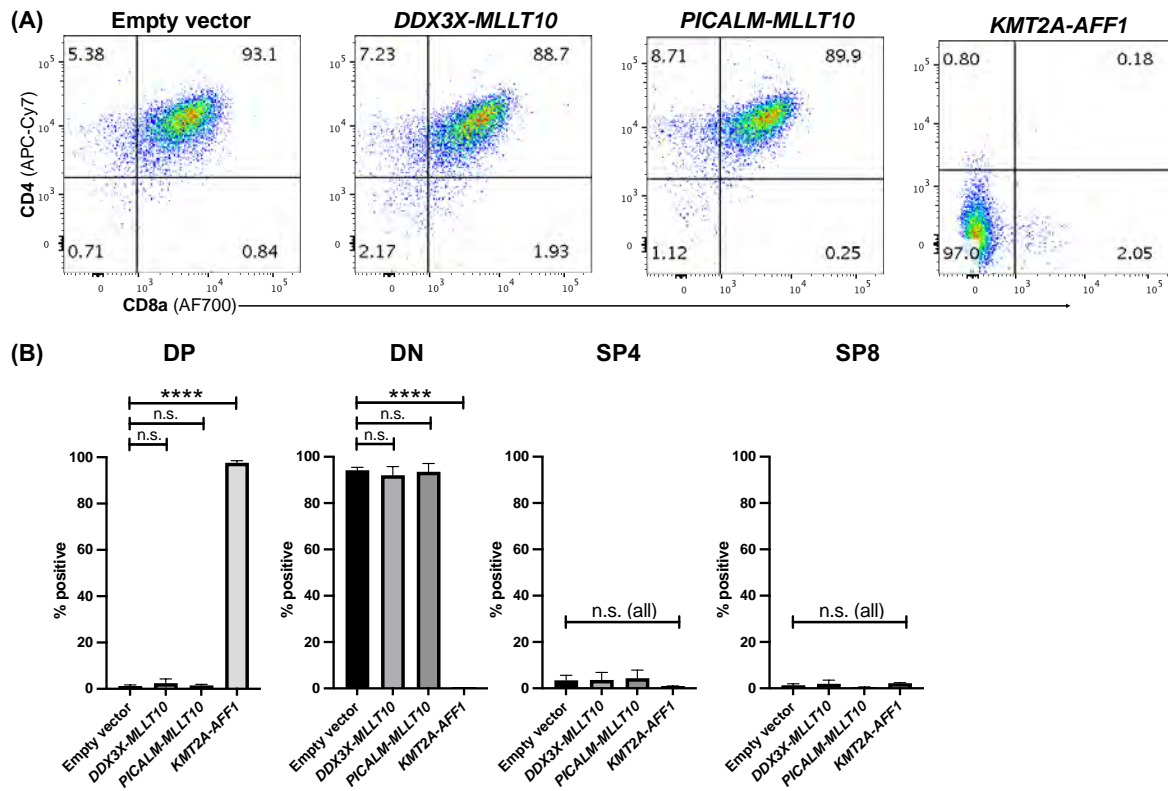
Figure 3.14 – MOHITO cells expressing *KMT2A-AFF1* exhibit an immature CD4⁻ CD8⁻ immunophenotype

Cell surface marker staining was performed on MOHITO cells (pRUFiG2 empty vector, *DDX3X-MLLT10, PICALM-MLLT10* and *KMT2A-AFF1*) on a BD LSRFortessa™. **(A)** Scatter plots for CD4/CD8 dual stained samples. **(B)** Percentage positive (% positive) for CD4⁺CD8⁺ (DP), CD4⁻CD8⁻ (DN), CD4⁺CD8⁻ (SP4) and CD4⁻CD8⁺ (SP8) values for each cell line were calculated as a proportion of live GFP⁺ cells and statistically compared by one-way ANOVA with Tukey's multiple comparisons. Statistically significant values are denoted with asterisks (***p<0.001, ***p<0.0001, n.s.=not significant). Scatter plots are representative of at least 3 biological replicates, and gates are set based on isotype controls. Error bars represent mean ± SD of 3 biological replicates.
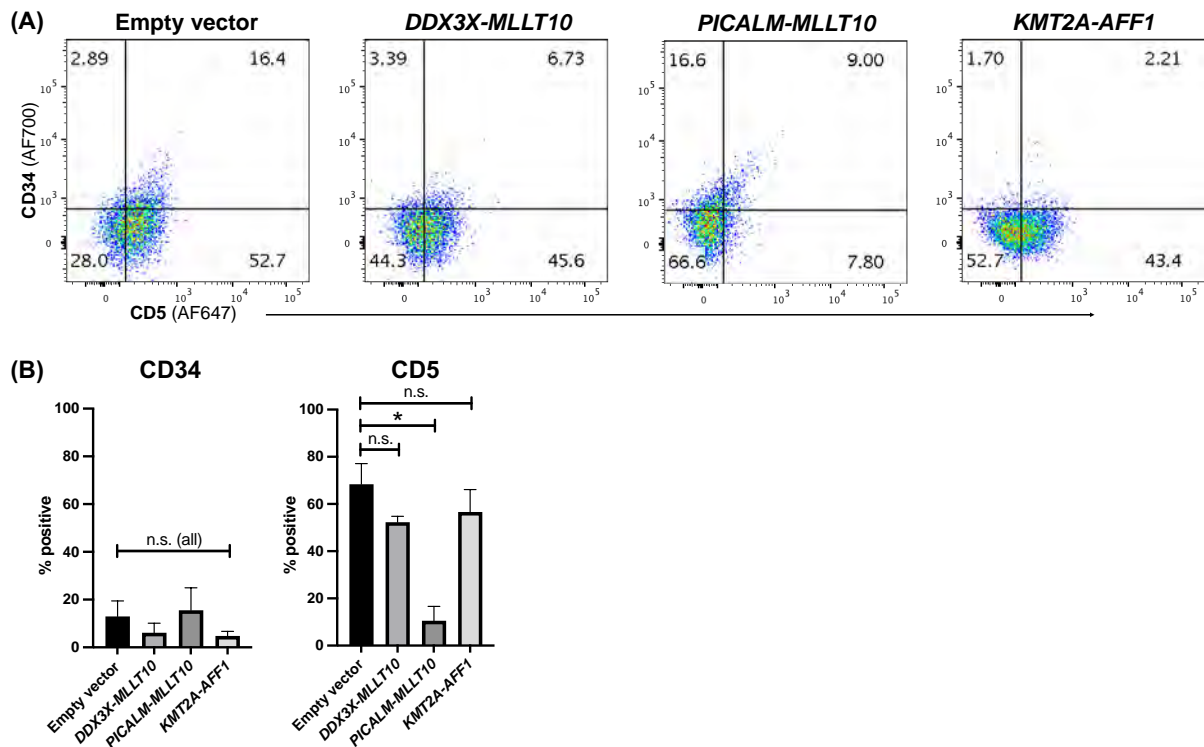
Figure 3.15 – MOHITO cell lines expressing *PICALM-MLLT10* exhibit reduced CD5 expression

Cell surface marker staining was performed on MOHITO cells (pRUFiG2 empty vector, *DDX3X-MLLT10, PICALM-MLLT10* and *KMT2A-AFF1*) by flow cytometry on a BD LSRFortessa™. **(A)** Scatter plots for CD34/CD5 dual stained samples. **(B)** Percentage positive (% positive) values for each cell line were calculated as a proportion of live GFP⁺ cells and statistically compared by one-way ANOVA. Statistically significant values are denoted with asterisks (*p<0.05, n.s.=not significant). Scatter plots are representative of at least 3 biological replicates, and gates are set based on isotype controls. Error bars represent mean ± SD of 3 biological replicates.
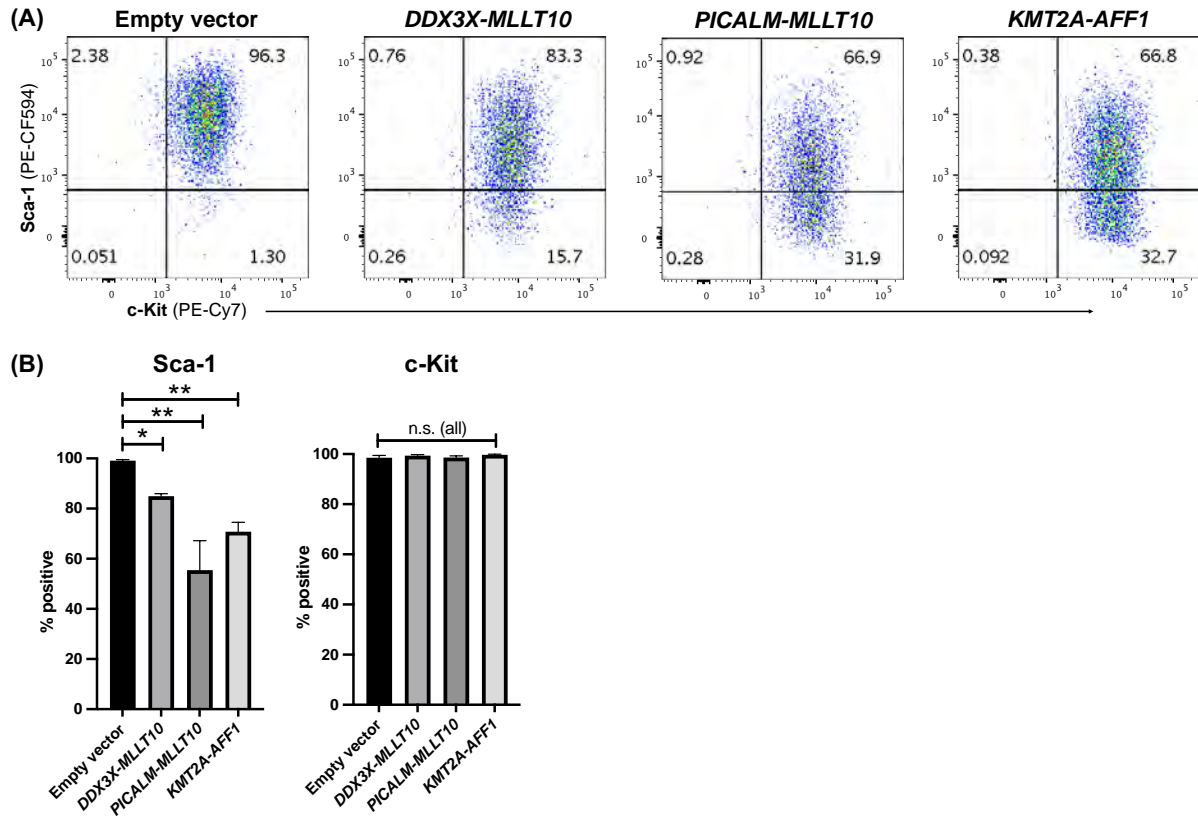
Figure 3.16 – MOHITO cell lines expressing *DDX3X-MLLT10, PICALM-MLLT10* and *KMT2A-AFF1* express c-Kit and reduced Sca-1

Cell surface marker staining was performed on MOHITO cells (pRUFiG2 empty vector, *DDX3X-MLLT10, PICALM-MLLT10* and *KMT2A-AFF1*) by flow cytometry on a BD LSRFortessa™. **(A)** Scatter plots for Sca-1/c-Kit dual stained samples. **(B)** Percentage positive (% positive) values for each cell line were calculated as a proportion of live GFP⁺ cells and statistically compared by one-way ANOVA. Statistically significant values are denoted with asterisks (*$p<0.05$, **$p<0.01$, n.s.=not significant). Scatter plots are representative of at least 3 biological replicates, and gates are set based on isotype controls. Error bars represent mean ± SD of 3 biological replicates.

3.3.4 Expression of *DDX3X-MLLT10*, *PICALM-MLLT10* or *KMT2A-AFF1* does not activate common leukaemia-associated signal transduction pathways

Activation of leukaemia-associated signal transduction pathways was investigated in cell lines by intracellular phospho-flow cytometry. The targets investigated and associated signalling pathways are summarised in **Table 3.1**.

Table 3.1 – Common leukaemia-associated signal transduction pathways, and targets investigated by phospho-flow cytometry

Signalling pathway	Target investigated
PI3K/AKT/mTOR	AKT[17], S6[17,18]
RAS/MAP-Kinase	ERK[19,20]
JAK/STAT	STAT5[21]
BCR	SYK[22,23], ZAP70[22,24]

Aberrant activation of signalling pathways typically results from genomic aberrations affecting members or regulators of a specific pathway, such as activating *JAK1-3* mutations that result in constitutive JAK/STAT activation[21]. While *KMT2A*r or *MLLT10*r are not direct members of the signalling pathways listed in **Table 3.1**, both *KMT2A* and *MLLT10* are epigenetic regulators that result in broadly aberrant transcriptional profiles, and therefore may potentially indirectly alter the expression of signalling pathway members. For example, a recent study identified that *MLLT10*r induces JAK/STAT activation in an AML model through increased expression of *JAK1*[25].

Signalling pathway activation was quantified in MOHITO (**Figure 3.17**) and Ba/F3 (**Figure 3.18**) cell lines by intracellular flow cytometry following a 5-hour cytokine starvation in 2% foetal calf serum (FCS). AKT, STAT5, SYK and S6 were significantly activated in all cell lines ($p < 0.05$), and non-significant activation of ERK and ZAP70 was observed. However, the presence of *KMT2A-AFF1, DDX3X-MLLT10* or *PICALM-MLLT10* did not lead to a statistically significant difference in activation levels compared to empty vector cells (**Figure 3.17**). No significant activation of any signalling pathway was observed in *KMT2A-AFF1* expressing Ba/F3 cells (**Figure 3.18**).
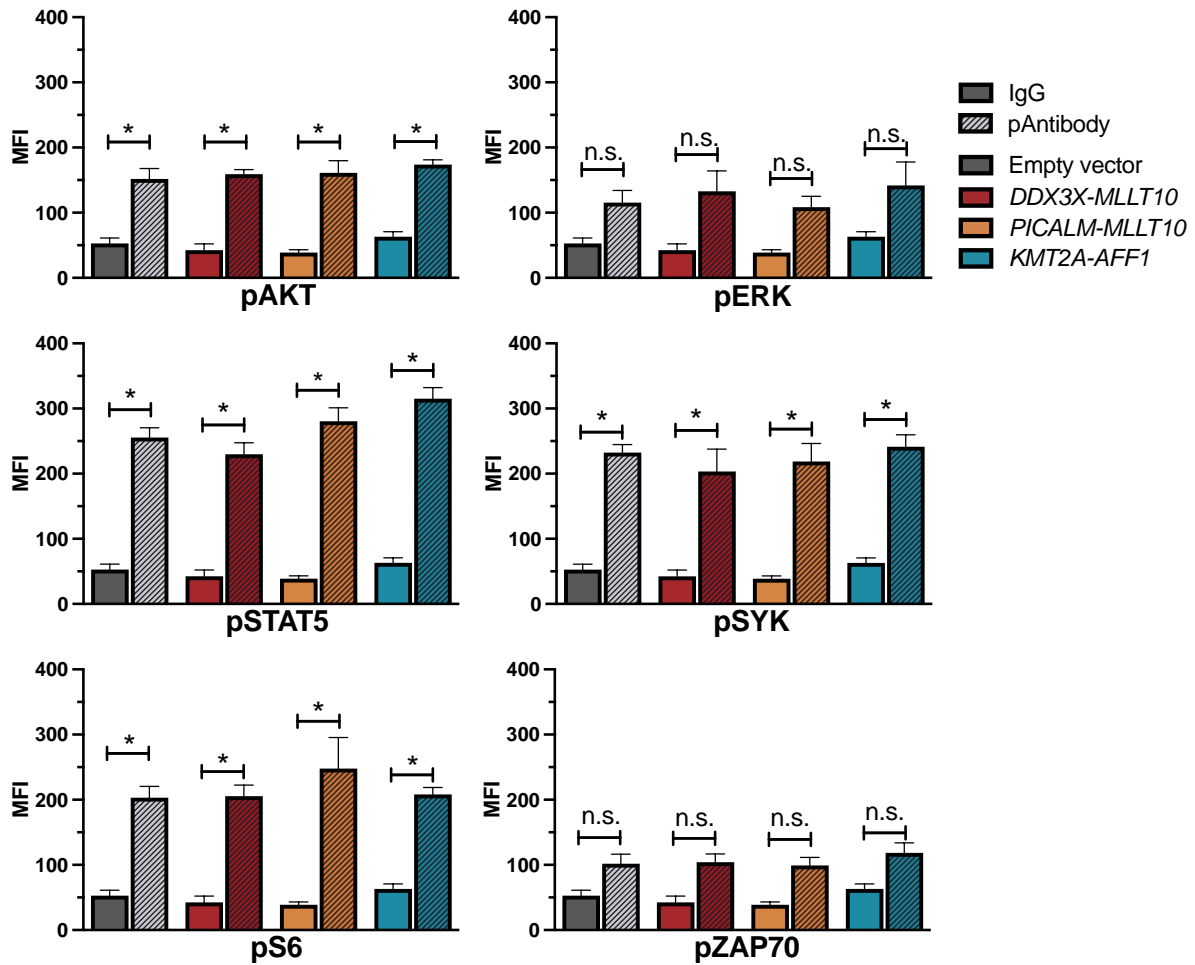
Figure 3.17 – *MLLT10*r or *KMT2A*r expression in MOHITO cells does not induce changes in the activation of leukaemia-associated signalling pathways

Signalling pathway activation was investigated by flow cytometric analysis of phosphorylated AKT, ERK, STAT5, SYK, S6 and ZAP70 in MOHITO cells expressing *DDX3X-MLLT10, PICALM-MLLT10, KMT2A-AFF1*, and control empty vector. Cells were incubated in 2% FCS for 5 hours prior to fixation. P-values were calculated by one-way ANOVA, comparing isotype and phospho-antibody (pAntibody) (*$p<0.05$, **$p<0.01$, n.s.=not significant). pAntibody MFI values for each fusion-expressing cell line were also compared to empty vector pAntibody by one-way ANOVA, and no statistically significant values were identified (annotations not shown on graph). Abbreviations: MFI, mean fluorescence intensity. Error bars represent mean ± SD of 3 biological replicates. Histograms are provided in **Supplementary Figure 3.2**.
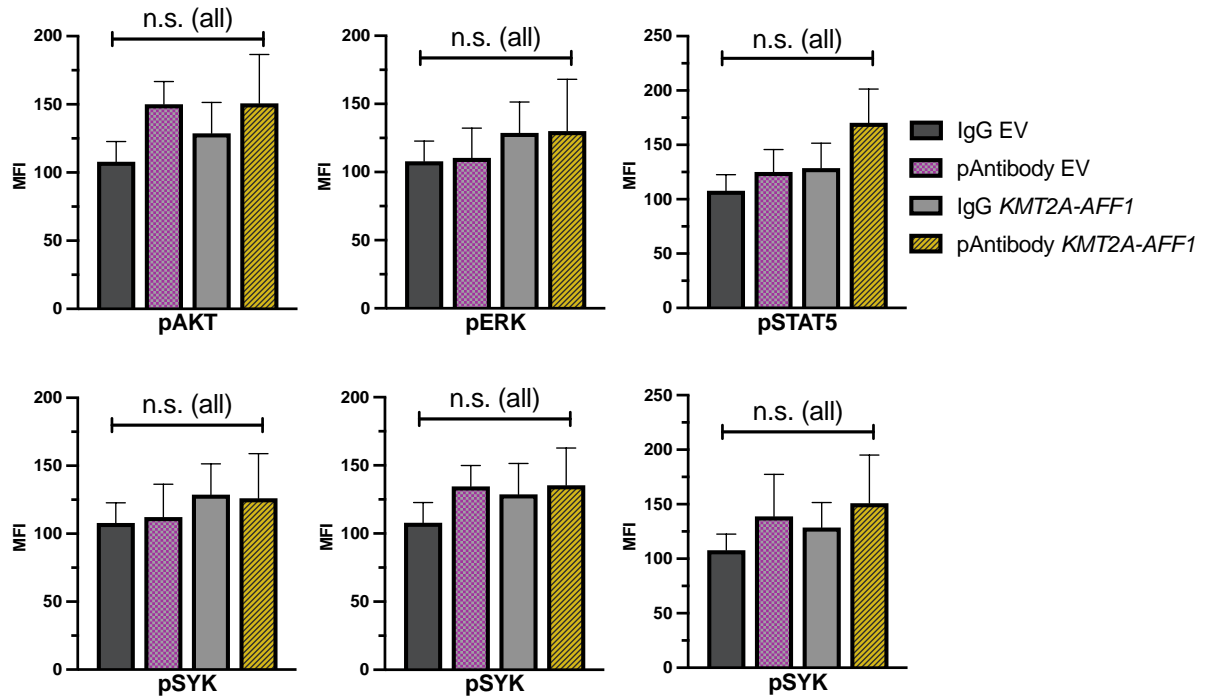
Figure 3.18 – Expression of *KMT2A-AFF1* in Ba/F3 cells does not induce changes in the activation of leukaemia-associated signalling pathways

Signalling pathway activation was investigated by flow cytometric analysis of phosphorylated AKT, ERK, STAT5, SYK, S6 and ZAP70 in Ba/F3 empty vector (EV) and *KMT2A-AFF1* expressing cells. Cells were incubated in 2% FCS for 5 hours prior to fixation. P-values were calculated using one-way ANOVA with Tukey's multiple comparisons. Abbreviations: MFI, mean fluorescence intensity, n.s.=not significant. Error bars represent mean ± SD of 3 biological replicates. Histograms are provided in **Supplementary Figure 3.3**.

3.4 Discussion

3.4.1 *KMT2A-AFF1* expression promotes increased cytokine-dependent MOHITO cell proliferation

KMT2A-AFF1 expression promoted a significant increase in cytokine-dependent proliferation of MOHITO cells (**Figure 3.5A**). However, expression of *KMT2A-AFF1* in Ba/F3 cells did not induce any changes in proliferation (**Figure 3.6A**). The ability of *KMT2A-AFF1* to confer a proliferative advantage in MOHITO cells indicates activation of oncogenic pathways. Positive control *NRAS*^{G12D} expressing cells were cytokine independent in both Ba/F3 and MOHITO cells, demonstrating assay validity (**Figure 3.5B** and **Figure 3.6B**).

KMT2A-AFF1, DDX3X-MLLT10 and *PICALM-MLLT10* are recurrently identified oncogenes. Therefore, the inability of these fusions to induce cytokine independence indicates that the oncogenic signalling pathways activated do not involve constitutive kinase activation. However, increased cytokine-dependent proliferation is induced by *KMT2A-AFF1* expression in MOHITO cells in the absence of constitutive activation of downstream effectors, demonstrates that classical kinase-associated pathways are not required to induce an aggressive ALL phenotype. Ba/F3 cells are transformed to cytokine independence by overexpression of kinase activating mutations such as *JAK2*^{V617F}, *EGFR*$^{G719S/L858R}$ and *MYB-TYK2*[26-28]. Similarly, MOHITO cells can be transformed to cytokine independence by the overexpression of *JAK1*^{A634D} or *BCR-ABL1*[13]. Importantly, all of these genomic aberrations result in constitutive activation of a kinase signalling pathway, resulting in cytokine independence. *KMT2A-AFF1* and *MLLT10*r do not have established associations with constitutive kinase activation, which explains the inability of these fusions to induce cytokine independence.

151

3.4.2 Expression of *KMT2A-AFF1*, but not *MLLT10*r, induces lineage-specific dysregulation of *HOXA* cluster genes

Dysregulation of the *HOXA* gene cluster is a characteristic feature of both *KMT2A*r and *MLLT10*r acute leukaemia. *HOX* genes are key regulators of haematopoietic stem cell differentiation[29]. It is likely that *HOXA* dysregulation induces *KMT2A*r and *MLLT10*r-mediated leukaemogenesis, by inducing a stem cell-like phenotype that promotes increased self-renewal, growth and survival advantages[4,30]. Knockdown of either *HOXA7, HOXA9, HOXA10* or *MEIS1* resulted in impaired engraftment of the *KMT2A-AFF1* B-ALL RS4;11 cell line in mice, demonstrating that these genes are both important and non-redundant[31]. Loss of *HOXA5* in *PICALM-MLLT10*-expressing murine bone marrow cells reduces serial replating capacity, suggesting an important role for *HOXA5* in *MLLT10*r leukaemia[30].

Here, the expression of *KMT2A-AFF1* induced specific changes in expression of individual *HOXA* genes between Ba/F3 and MOHITO cells. Only *HOXA9* expression was increased in *KMT2A-AFF1* Ba/F3 cells compared to empty vector Ba/F3 cells (**Figure 3.8A**), consistent with existing evidence that *HOXA9* is the most consistently upregulated gene in *KMT2A*r B-ALL patient leukaemic blasts[4,32,33]. Expression levels of *HOXA3, HOXA5*, and *HOX* cofactor *MEIS1,* but not *HOXA9*, were significantly increased in *KMT2A-AFF1* MOHITO cells compared with empty vector MOHITO cells (**Figure 3.8B**). *HOXA3* and *HOXA5* are upregulated in leukaemic blasts from patients with *KMT2A*r or *MLLT10*r T-ALL[32], whereas upregulation of these *HOX* genes is not typically reported in *KMT2A*r B-ALL cohorts[4,32,33]. This suggests a potential lineage-specific role for *HOXA3* or *HOXA5* in T-ALL, a notion confirmed by the results presented here.

It is well-established that leukaemic blasts from patients with *MLLT10*r T-ALL and AML exhibit *HOXA* dysregulation[32,34-37]. However, no changes in *HOXA* expression were observed in either *DDX3X-MLLT10* or *PICALM-MLLT10* expressing MOHITO cells compared with empty vector control cells (**Figure 3.7**). The lack of *HOXA* cluster upregulation observed here, in addition to lack of changes in proliferation and immunophenotype, indicates the MOHITO cell line is an incomplete model of *MLLT10*r T-ALL.

3.4.3 *KMT2A-AFF1* expressing Ba/F3 cells exhibit loss of pan-B-cell marker B220

Immunophenotyping of cell lines expressing *MLLT10*r or *KMT2A-AFF1* enabled exploration of changes in cell differentiation. Parental Ba/F3 cells exhibited a pro/pre-B immunophenotype (**Figures 3.8-3.9**). Haematopoietic stem cell markers c-Kit, Sca-1 and CD34, and myeloid marker CD13, were not expressed (**Figure 3.10**). Surface marker expression throughout immature B-cell development is summarised in **Figure 3.19**.

Ba/F3 cells expressing *KMT2A-AFF1* lost expression of the pan-B cell marker B220 (**Figure 3.9**). Lack of B220 indicates that *KMT2A-AFF1* alters Ba/F3 cells from a pro-B to a B-progenitor phenotype. A non-statistically significant reduction in BP-1 surface expression was also observed in *KMT2A-AFF1* Ba/F3 cells, consistent with loss of B-cell lineage identity (**Figure 3.9**).
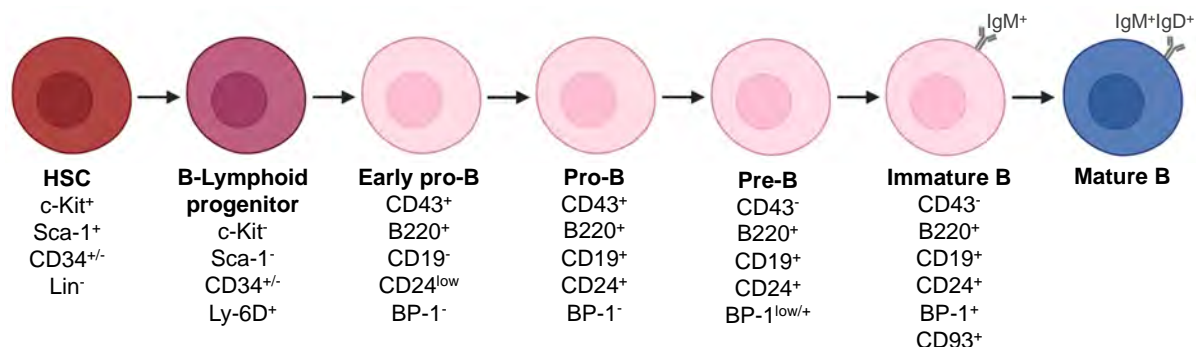
Figure 3.19 – Surface marker expression of murine B-lymphocytes at various stages of development

Lineage-negative (Lin⁻) haematopoietic stem cells (HSCs) mature into B-lymphocytes in the bone marrow. Various surface markers are expressed at different stages of development, enabling immunophenotyping of B-ALL leukaemic blasts to determine stage of differentiation arrest. Figure compiled based on information from references[16,38-40].

3.4.4 Immunophenotypic changes are induced by expression of *KMT2A-AFF1* or *MLLT10*r in MOHITO cells

Empty vector MOHITO cells possessed a mature CD4⁺CD8⁺ T-cell phenotype, consistent with the original published MOHITO immunophenotype[13]. The surface marker repertoire on each cell line is shown in **Table 3.2**. Expression of surface markers at various stages of thymic lymphocyte development are shown in **Figure 3.20**, and recognised EGIL T-ALL subsets are provided in **Table 3.3**.

Table 3.2 – Immunophenotype of MOHITO cell lines expressing *DDX3X-MLLT10*, *PICALM-MLLT10* or *KMT2A-AFF1*

Marker	Parental MOHITO[13]	Empty vector	*DDX3X-MLLT10*	*PICALM-MLLT10*	*KMT2A-AFF1*
CD45	+	+	+	+	+
sCD3	-	-	-	-	-
CD25	+	+	+	+	+
CD44	NA	Dim	Dim	Dim	+
CD4	+/-	+	+	+	-
CD8	+	+	+	+	-
CD2	NA	+	+	+	+
CD5	NA	+	+	-	+
CD34	-	-	-	-	-
c-Kit	+	+	+	+	+
Sca-1	+	+	Dim	Dim	Dim

Differences in surface marker expression level from empty vector are highlighted in green. Abbreviations: NA, data not available.
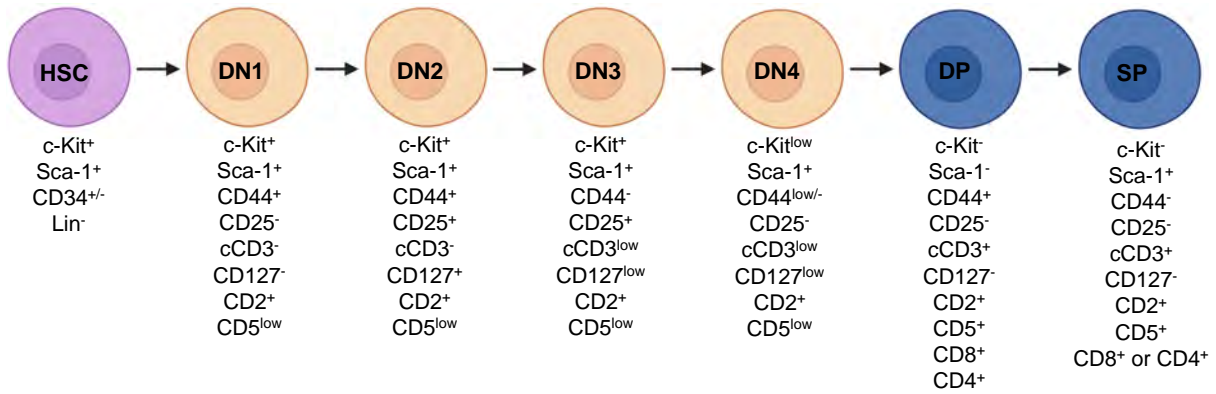
Figure 3.20 – Surface marker expression of murine lymphocytes at various stages of thymic development

Lin⁻ HSCs migrate to the thymus, where they undergo T-lymphocyte development. Figure was compiled based on information from references[41-45]. Abbreviations: Lin⁻, lineage negative; HSC, haematopoietic stem cells; DN, double negative; DP, double positive; SP, single positive.

Table 3.3 – ETP-ALL and recognised European Group for the Immunological Classification of Leukaemias (EGIL) T-ALL subsets, and associated T-ALL surface marker expression

T-ALL stage	Surface marker expression
ETP	cCD3⁺, CD7⁺, CD1a⁻, CD5$^{dim/-}$, sCD3⁻, CD4⁻, CD8⁻, and expression of at least one myeloid or stem cell antigen (c-Kit/CD117, HLA-DR, CD13, CD33, CD11b, CD65)
Immature	cCD3⁺, CD7⁺, CD1a⁻, CD2⁺ and/or CD5⁺, CD8⁺ or CD4⁺
Early cortical	cCD3⁺, CD7⁺, CD4⁺ and CD8⁺, CD5⁺, CD2⁺, CD1a⁻, sCD3⁻
Cortical	cCD3⁺, CD7⁺, CD1a⁺, sCD3⁻
Late cortical	cCD3⁺, CD7⁺, CD1a⁺, sCD3⁺
Mature	cCD3+, CD7⁺, CD1a⁻, sCD3⁺, and CD8a⁺ and/or CD4$^{+/-}$
Table compiled with information from references[38,46,47]	

All three fusion-expressing cell lines exhibited a significant reduction in Sca-1 expression compared with empty vector MOHITO cells (**Figure 3.16**). Sca-1 is a marker of murine HSCs in combination with c-Kit[48], but the human equivalent of Sca-1 is not known[48]. Sca-1 expression is variable in murine haematopoietic cells

155

depending on differentiation stage: HSCs exhibit strong expression of Sca-1, which is lost at the DN1 stage of thymocyte development. DP and SP thymocytes, and peripheral T-cells, then regain Sca-1 expression[48] (**Figure 3.20**). In one study, Sca-1 expression correlated with aggressive murine leukaemia, quantified by a shorter time to engraftment than low Sca-1 expressing disease[49]. However, further data is lacking, particularly since the human equivalent of Sca-1 is not known.

c-Kit expression was observed in all MOHITO cell lines (**Figure 3.16**). c-Kit is expressed in approximately 10% of T-ALLs overall[50,51], most frequently in immature ETP and pro-T cases (50% and 67% of cases respectively[50]). CD34 is a common marker of ALL, present in approximately 50% of T-ALL cases overall[52,53], but CD34 was not expressed by any MOHITO cell line. Despite being important markers of HSCs, CD34 and c-Kit are not reliable markers of immature disease alone. Indeed, other lineage-specific markers such as CD5, CD4 and CD8 are required to adequately immunophenotype T-ALL cases.

KMT2A-AFF1 expression in MOHITO cells induced a less mature immunophenotype, specifically loss of CD4 and CD8 (**Figure 3.14**), and increased CD44 expression (**Figure 3.13**). This immunophenotype is most similar to immature T-cells at the DN2 stage of development (**Figure 3.20**), and is consistent with an early thymic precursor (ETP) ALL phenotype due to strong c-Kit and dim CD5 expression (**Figure 3.20, Table 3.3**). Both *MLLT10*r cell lines retained CD4 and CD8 (**Figure 3.14**), but *PICALM-MLLT10* MOHITO cells lost CD5 expression, a hallmark feature of ETP-ALL. CD4 and CD8 are expressed in 53% and 60% of T-ALL cases overall[50], but there is currently no published data comparing the immunophenotype of T-ALL between different genomic subtypes.

CD44 expression is associated with poor outcomes across a range of malignancies, including adult T-ALL[54], AML[55], and various solid tumours[56,57]. DN1-2 immature thymocytes are CD44[+], then expression is lost in the DN3-4 stages, and later regained at the DP stage (**Figure 3.20**)[56]. CD44 is expressed in >90% of T-ALL cases, but intensity of expression is widely variable[50]. One study associated high CD44 expression with T-ALL cases that harboured mutations in *N/KRAS* or *FBXW7*[56]. However, this study only categorised cases by mutations in *NOTCH1, FBXW7, TLX3*

and *N/KRAS,* and the *STIL-TAL1* gene fusion*,* so the role of CD44 in *KMT2A*r T-ALL remains unclear.

Dim or absent CD5 is a defining feature of ETP-ALL. In normal thymic lymphocytes, CD5 expression is low throughout DN stages, and then increases in the DP and SP stages (**Figure 3.20**). In one study, inhibition of CD5 in patient MNCs was associated with increased IL-2 production and increased survival after 48 h in culture, suggesting that lack of CD5 expression may promote cancer cell survival[58]. While CD5 negative T-ALL is normally indicative of an immature phenotype, the *PICALM-MLLT10* MOHITO cells described here also express the mature T-cell markers CD4 and CD8, demonstrating an unusual T-ALL immunophenotype.

ETP-ALL was historically associated with poor outcomes when compared with non-ETP T-ALL cases[59,60], but studies of large T-ALL cohorts that describe immunophenotype and genomic subtype together are currently lacking. One published cohort included three infants with *KMT2A*r T-ALL, of which two had EGIL stage T-III and one had T-IV disease, indicating no association between *KMT2A*r and ETP-ALL in infants[61]. However, one study established that *HOXA* deregulated T-ALL is overrepresented within ETP-ALL in non-infants[62], and another identified that *PICALM-MLLT10* is overrepresented within the ETP-ALL immunophenotype[63].

To summarise, expression of *KMT2A-AFF1* and *MLLT10*r in the MOHITO cell line induced fusion-specific changes in immunophenotype that do not clearly correlate with established EGIL T-ALL subsets. Most notably, expression of *KMT2A-AFF1* induced loss of the mature T-cell markers CD4 and CD8, and increased expression of the immature T-cell marker CD44, overall indicating a less mature phenotype. *PICALM-MLLT10* and *DDX3X-MLLT10* cells retained CD4 and CD8 positivity, but cells expressing *PICALM-MLLT10* exhibited loss of CD5, a distinctive marker of ETP-ALL. Further immunophenotyping of leukaemic blasts from T-ALL patients, and identification of genomic subtype-specific patterns in surface marker expression, are required to validate these results. The immunophenotype of *MLLT10*r T-ALL is explored further in chapter 4.

3.4.5 Expression of *KMT2A-AFF1, DDX3X-MLLT10* or *PICALM-MLLT10* does not induce activation of common cancer-associated signalling pathways

To further investigate signalling pathway activation in *KMT2A-AFF1* and *MLLT10*r cell lines, activation of six common cancer-associated signalling pathways were explored through phospho-flow cytometry (**Figure 3.17** and **Figure 3.18**). All MOHITO cell lines exhibited significant activation of AKT, STAT5, SYK and S6 (**Figure 3.17**). STAT5 activation may occur due to the presence of the novel *JAK1*[S1042I] kinase domain mutation in the parental MOHITO cell line, although this mutation alone is incapable of inducing cytokine independence[13]. No significant changes in activation of ERK, STAT5, SYK, S6, AKT or ZAP70 were observed between empty vector and *KMT2A-AFF1* expressing Ba/F3 cells. These results demonstrate that expression of *KMT2A-AFF1* or *MLLT10*r does not induce changes in activation of signalling pathways involving ERK, STAT5 SYK, S6, AKT or ZAP70 kinase signalling.

3.4.6 Summary of chapter findings

This chapter demonstrates that expression of *KMT2A-AFF1* induces *HOXA* dysregulation and an immature T-ALL immunophenotype when expressed in the T-lineage MOHITO cell line. Taken together, these results suggest that phenotypic differentiation changes underpin leukaemogenesis driven by *KMT2A-AFF1*. The differentiation stage of leukaemic blasts is known to adversely affect prognosis, as ETP-ALL was historically associated with poor outcomes compared with other subtypes of T-ALL[64]. Subsequent intensification of ETP-ALL induction treatment protocols has resulted in improved outcomes[64,65]. This demonstrates the importance of identifying high-risk patients, based either on genomics or immunophenotype, to provide suitably intensified therapies to improve outcomes.

The Ba/F3 and MOHITO cell models are undoubtedly useful tools in characterising the biology of leukaemia-associated oncogenes. A notable benefit of these systems is they enable the study of genomic aberrations in isolation, ie. in the absence of each patient's complex and unique genomic landscape. However, cooperative mutations affect biology and therapeutic sensitivities, making it critical to identify and characterise recurrent co-occurring mutations. It is also important to note that direct

comparison of changes induced by a fusion's expression in Ba/F3 and MOHITO cells is not ideal, as they differ both in differentiation stage as well as lineage. MOHITO cells express the stem cell markers c-Kit and Sca-1, whereas Ba/F3 cells do not express c-Kit or Sca-1 (**Figure 3.10** and **Figure 3.16** respectively). It is therefore not clear whether the differences induced by *KMT2A-AFF1* expression – specifically, changes in proliferation and *HOXA* gene cluster expression – are specific to lineage (B vs T-lineage) or differentiation stage (cKit⁻Sca-1⁻ vs c-Kit⁺Sca-1⁺).

KMT2A-AFF1 expression in Ba/F3 cells induced hallmark upregulation of *HOXA9*, but did not induce any proliferative or immunophenotypic changes indicative of the aggressive pro-B ALL observed in patients. As *KMT2A-AFF1* is a well-characterised predictor of poor outcomes in B-ALL patients of all ages, it is clear that the Ba/F3 cell line is not an appropriate *in vitro* system for the modelling of *KMT2A*r. Modelling *KMT2A*r B-ALL has proven a major challenge in the field of ALL research for decades, and it is only very recently that *KMT2A-AFF1* pro-B ALL was successfully recapitulated in a transgenic model in human CD34⁺ foetal liver cells[6]. It is possible that this was not only a consequence of the parent cell's foetal gene expression signature, but also because the reciprocal *AFF1-KMT2A* fusion gene was expressed in this model. Further investigation of *KMT2A-AFF1* and *AFF1-KMT2A* transformation in haematopoietic progenitor cells may provide further progress towards understanding the aetiology of *KMT2A*r B- and T-ALL.

The expression of *DDX3X-MLLT10* or *PICALM-MLLT10* did not induce any changes in proliferation indicative of oncogenic transformation in MOHITO cells. *MLLT10*r are poorly characterised, and transgenic models of *MLLT10*r T-ALL are lacking. It is possible that, as demonstrated in *KMT2A*r, parental cell lineage is a critical factor in modelling the disease. However, *MLLT10*r other than *KMT2A-MLLT10* do not occur in B-ALL, indicating there must be key biological differences between *MLLT10*r and *KMT2A*r. An alternative model system to characterise the biology of *MLLT10*r T-ALL is discussed in chapter 4. The role of cooperative mutations should also be considered as a potential explanation for the inability of *MLLT10*r to transform MOHITO cells. A case series has described four patients with *DDX3X-MLLT10* T-ALL, where all four patients harboured a *NOTCH1* mutation, and two had *CDKN2A/B* deletions[37]. It is possible that cooperative events are required to induce *MLLT10*r ALL, a potential

explanation for the lack of transformation observed here. The mutational landscape of *MLLT10*r and *KMT2A*r ALL is discussed further in chapter 5.

The data within this chapter demonstrates that expression of a *KMT2A*r does not transform Ba/F3 cells, a common system used to model oncogenic fusion genes. This chapter also describes the first *in vitro* murine models of *KMT2A*r and *MLLT10*r T-ALL, using the MOHITO T-cell line, and establishes that expression of *KMT2A-AFF1* is sufficient to induce increased expression of specific *HOXA* genes and de-differentiation into an immature T-ALL immunophenotype. Moving forward, this model of *KMT2A*r T-ALL serves as a useful tool in pre-clinical drug discovery and screening, to establish more efficacious targeted therapies to treat this high-risk subtype.

3.5 Supplementary Figures

Supplementary Figure 3.1

(A)

(B)

(C)

(D)

(E)
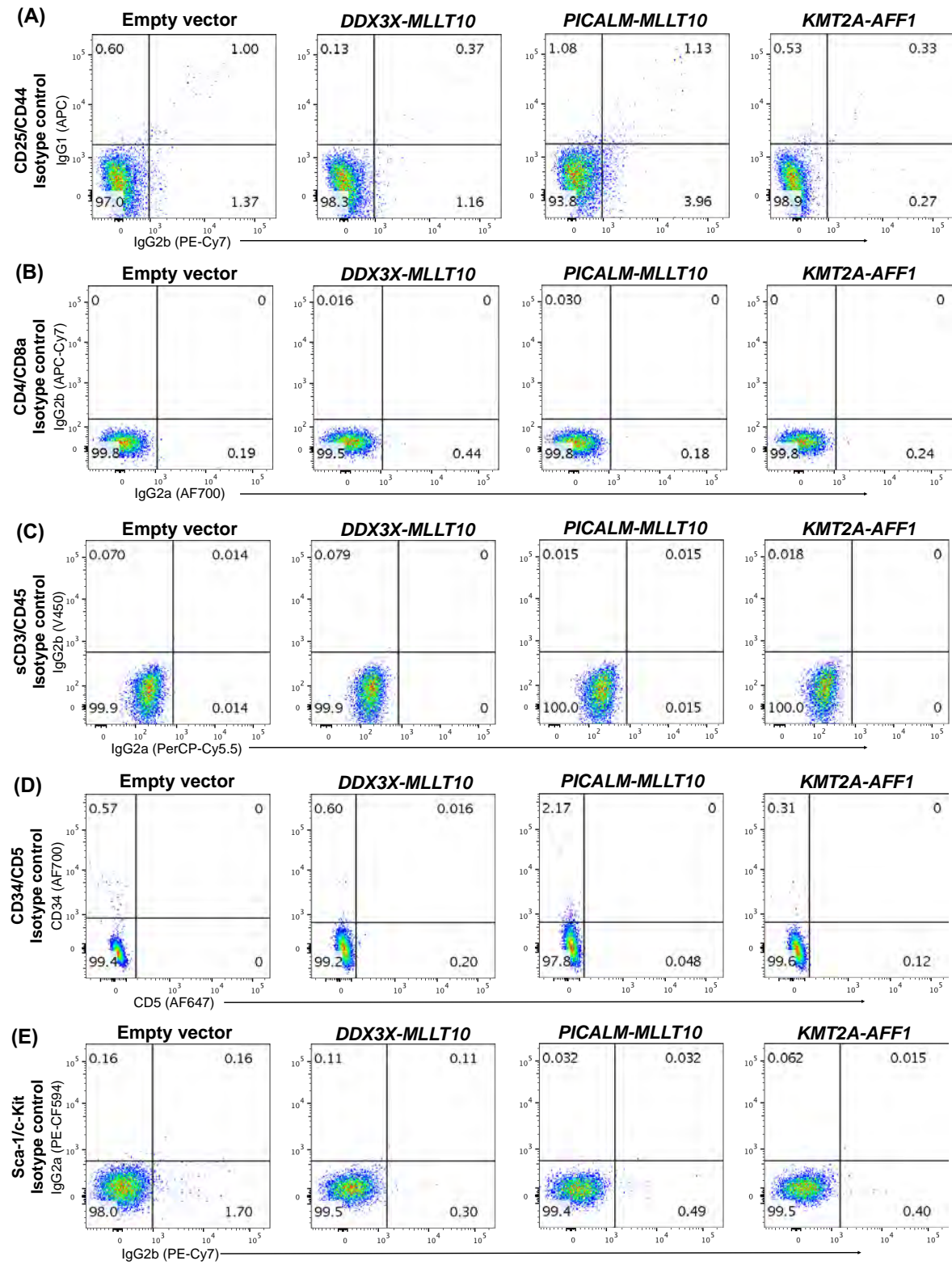
Supplementary Figure 3.1 – Isotype control scatter plots for MOHITO surface marker staining

All MOHITO surface marker staining gates were set according to fluorescence intensity of cells stained with relevant isotype control antibodies. All scatter plots are representative of 3 independent biological replicates.

Supplementary Figure 3.2
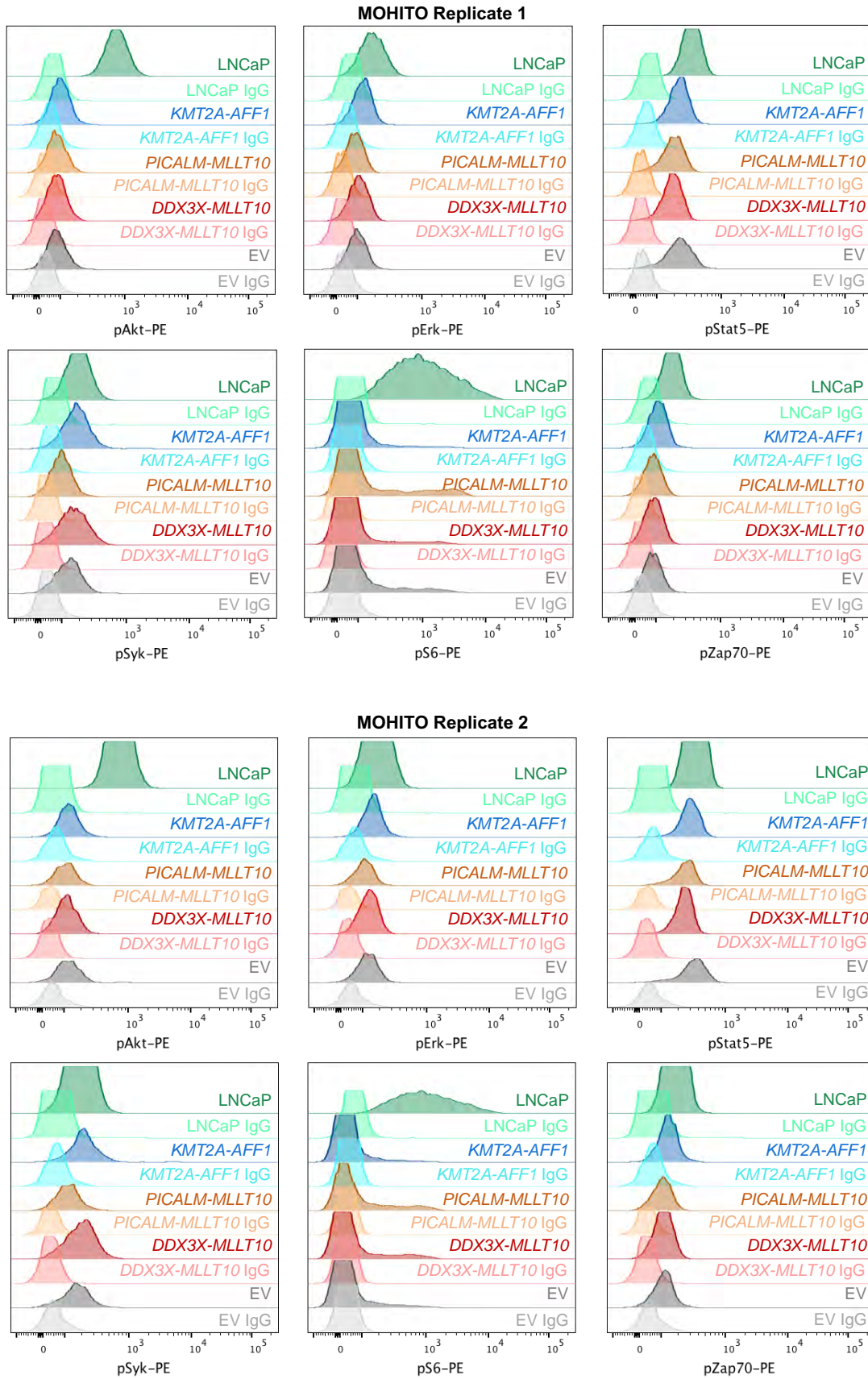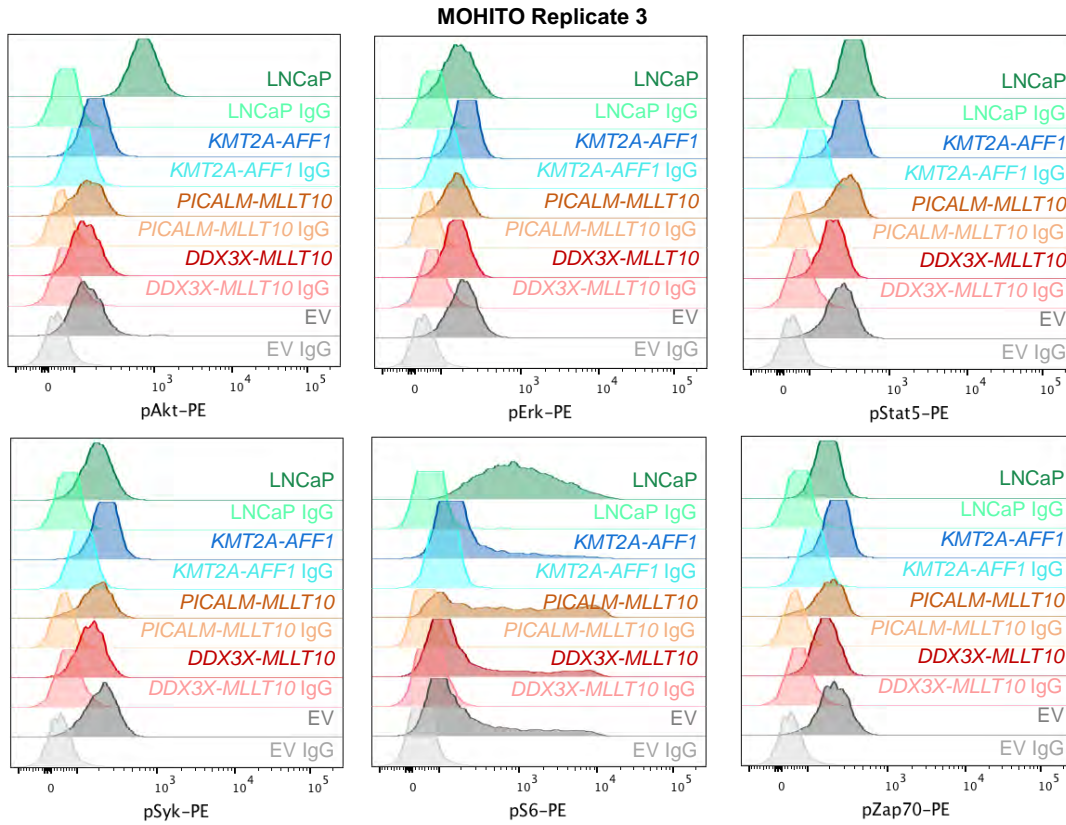
Supplementary Figure 3.2 – Phospho-flow analysis of signalling pathway activation in MOHITO cells

Signalling pathway activation was investigated by intracellular flow cytometric analysis of phosphorylated AKT, ERK, STAT5, SYK, S6 and ZAP70 in MOHITO cells expressing *DDX3X-MLLT10, PICALM-MLLT10, KMT2A-AFF1*, and control empty vector cells. All cells were incubated in 2% serum media for 5 hours prior to fixation. The prostate cancer cell line LNCaP was used as a positive control.

Supplementary Figure 3.3
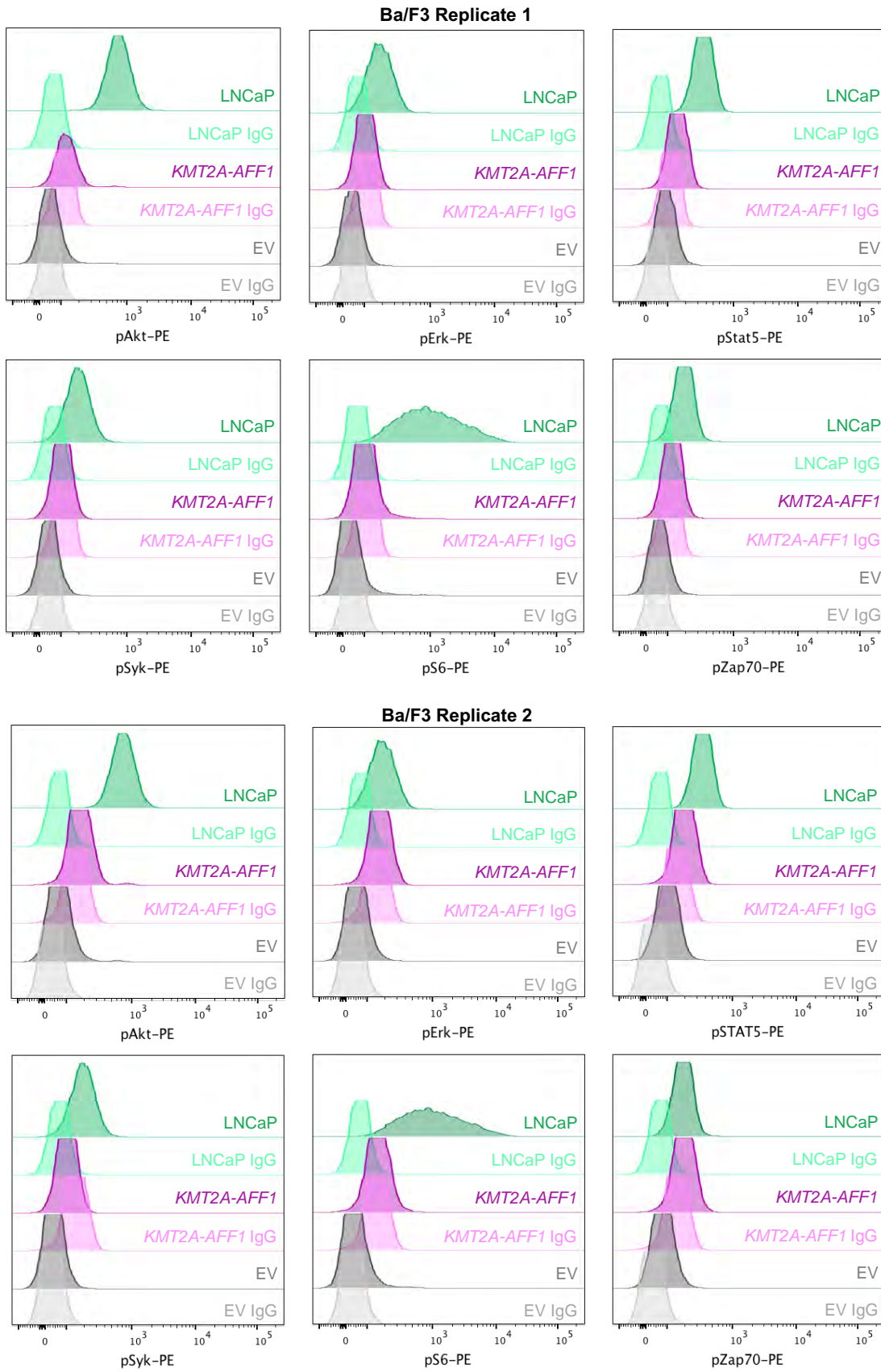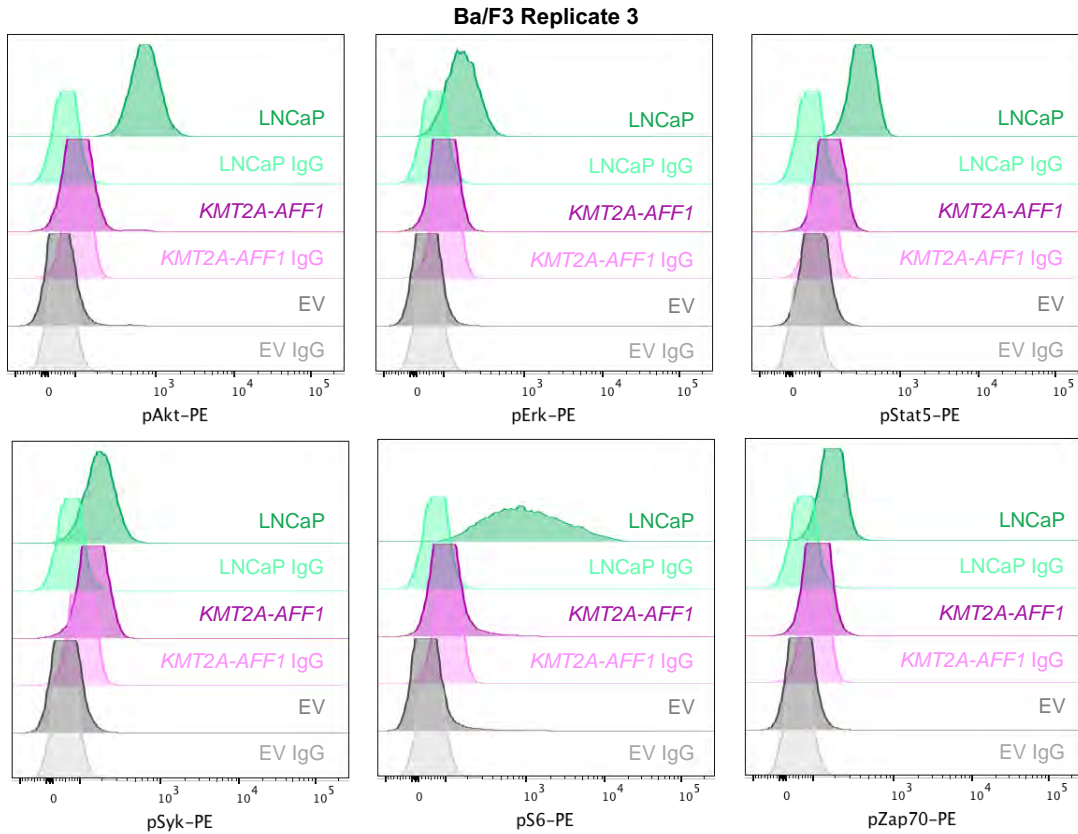
Supplementary Figure 3.3 – Phospho-flow analysis of signalling pathway activation in Ba/F3 cells

Signalling pathway activation was investigated by intracellular flow cytometric analysis of phosphorylated AKT, ERK, STAT5, SYK, S6 and ZAP70 in MOHITO cells expressing *DDX3X-MLLT10, PICALM-MLLT10, KMT2A-AFF1*, and control empty vector cells. All cells were incubated in low-serum media for 5 hours prior to fixation. The prostate cancer cell line LNCaP was used as a positive control.

3.6 Chapter References

1   Armstrong, S. A. *et al.* MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* **30**, 41-47, doi:10.1038/ng765 (2002).

2   Ferrando, A. *et al.* Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: dominance of HOX dysregulation. *Blood* **102**, 262-268, doi:10.1182/blood-2002- (2003).

3   Koldobskiy, M. A. *et al.* A Dysregulated DNA Methylation Landscape Linked to Gene Expression in MLL-Rearranged AML. *Epigenetics* **15**, 841-858, doi:10.1080/15592294.2020.1734149 (2020).

4   Winters, A. C. & Bernt, K. M. MLL-Rearranged Leukemias-An Update on Science and Clinical Approaches. *Front Pediatr* **5**, 1-21, doi:10.3389/fped.2017.00004 (2017).

5   Zeisig, B. B. *et al.* Functional reconstruction of human AML reveals stem cell origin and vulnerability of treatment-resistant MLL-rearranged leukemia. *Science Translational Medicine* **13**, eabc4822, doi:10.1126/scitranslmed.abc4822 (2021).

6   Rice, S. *et al.* A human fetal liver-derived infant MLL-AF4 acute lymphoblastic leukemia model reveals a distinct fetal gene expression program. *Nat Commun* **12**, 6905, doi:10.1038/s41467-021-27270-z (2021).

7   Belver, L. & Ferrando, A. The genetics and mechanisms of T cell acute lymphoblastic leukaemia. *Nat Rev Cancer* **16**, 494-507, doi:10.1038/nrc.2016.63 (2016).

8   Caudell, D. & Aplan, P. D. The role of CALM-AF10 gene fusion in acute leukemia. *Leukemia* **22**, 678-685, doi:10.1038/sj.leu.2405074 (2008).

9   Forgione, M. O., McClure, B. J., Yeung, D. T., Eadie, L. N. & White, D. L. MLLT10 rearranged acute leukemia: incidence, prognosis and possible therapeutic strategies. *Genes Chromosomes Cancer* **59**, 709-721, doi:10.1002/gcc.22887 (2020).

10   Iacobucci, I. & Mullighan, C. G. Genetic Basis of Acute Lymphoblastic Leukemia. *J Clin Oncol* **35**, 975-983, doi:10.1200/JCO.2016.70.7836 (2017).

11   Caudell, D., Zhang, Z., Chung, Y. J. & Aplan, P. D. Expression of a CALM-AF10 fusion gene leads to Hoxa cluster overexpression and acute leukemia in transgenic mice. *Cancer Res* **67**, 8022-8031, doi:10.1158/0008-5472.CAN-06-3749 (2007).

12   Warmuth, M., Sungjoon, K., Gu, X., Gang, X. & Francisco, A. Ba/F3 cells and their use in kinase drug discover. *Current Opinion in Oncology* **19**, 55-60, doi:10.1097/CCO.0B013E328011A25F (2007).

13   Kleppe, M., Mentens, N., Tousseyn, T., Wlodarska, I. & Cools, J. MOHITO, a novel mouse cytokine-dependent T-cell line, enables studies of oncogenic signaling in the T-cell context. *Haematologica* **96**, 779-783, doi:10.3324/haematol.2010.035931 (2011).

14   Forgione, M. O., McClure, B. J., Eadie, L. N., Yeung, D. T. & White, D. L. KMT2A rearranged acute lymphoblastic leukaemia: Unravelling the genomic complexity and heterogeneity of this high-risk disease. *Cancer Lett* **469**, 410-418, doi:10.1016/j.canlet.2019.11.005 (2020).

15   Sanz, E. *et al.* Ordering human CD34+CD10-CD19+ pre/pro-B-cell and CD19- common lymphoid progenitor stages in two pro-B-cell development pathways. *Proc Natl Acad Sci U S A* **107**, 5925-5930, doi:10.1073/pnas.0907942107 (2010).

16   Janke, L. J., Mullighan, C. G., Dang, J. & Rehg, J. E. Immunophenotyping of Murine Precursor B-Cell Leukemia/Lymphoma: A Comparison of Immunohistochemistry and Flow Cytometry. *Vet Pathol* **56**, 950-958, doi:10.1177/0300985819852138 (2019).

17   Teachey, D. T. & Pui, C.-H. Comparative features and outcomes between paediatric T-cell and B-cell acute lymphoblastic leukaemia. *The Lancet Oncology* **20**, e142-e154, doi:10.1016/s1470-2045(19)30031-2 (2019).

18   Wallington-Beddoe, C. *et al.* Identification of sphingosine kinase 1 as a therapeutic target in B-lineage acute lymphoblastic leukaemia. *British Journal of Haematology* **184**, 443-447, doi:10.1111/bjh.15097 (2018).

19   Moharram, S. A., Shah, K. & Kazi, J. U. T-cell Acute Lymphoblastic Leukemia Cells Display Activation of Different Survival Pathways. *J Cancer* **8**, 4124, doi:10.7150/jca.21725 (2017).

20   van der Zwet, J. *et al.* MAPK-ERK is a central pathway in T-cell acute lymphoblastic leukemia that drives steroid resistance. *Leukemia*, doi:doi.org/10.1038/s41375-021-01291-5 (2021).

21   Brooks, A. J. & Putoczki, T. JAK-STAT Signalling Pathway in Cancer. *Cancers (Basel)* **12**, doi:10.3390/cancers12071971 (2020).

22   Sadras, T. *et al.* Co-Expression of SYK and ZAP70 Subverts Negative B-Cell Selection and Enables Oncogenic Signaling in Multiple B-Cell Malignancies. *Blood* **134**, 295, doi:10.1182/blood-2019-128999 (2019).

23   Sender, S. *et al.* Precursor B-ALL Cell Lines Differentially Respond to SYK Inhibition by Entospletinib. *Int J Mol Sci* **22**, doi:10.3390/ijms22020592 (2021).

24   Alsadeq, A. *et al.* The role of ZAP70 kinase in acute lymphoblastic leukemia infiltration into the central nervous system. *haematologica* **102**, 346-355 (2017).

25   Chen, B.-R. *et al.* A JAK/STAT-mediated inflammatory signaling cascade drives oncogenesis in AF10-rearranged AML. *Blood* **137**, 3403-3415, doi:10.1182/blood.2020009023 (2021).

26   Tavakoli Shirazi, P. *et al.* Constitutive JAK/STAT signaling is the primary mechanism of resistance to JAKi in TYK2-rearranged acute lymphoblastic leukemia. *Cancer Lett* **512**, 28-37, doi:10.1016/j.canlet.2021.04.027 (2021).

27   Levine, R. L. *et al.* Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell* **7**, 387-397, doi:10.1016/j.ccr.2005.03.023 (2005).

28   Jiang, J. *et al.* Epidermal growth factor-independent transformation of Ba/F3 cells with cancer-derived epidermal growth factor receptor mutants induces gefitinib-sensitive cell cycle progression. *Cancer Res* **65**, 8968-8974, doi:10.1158/0008-5472.CAN-05-1829 (2005).

29   Lappin, T.R.J, Grier, D. G., Thompson, A. & Halliday, H. L. HOX GENES: Seductive Science, Mysterious Mechanisms. *Ulster Medical Journal* **75**, 23-31 (2006).

30   Okada, Y. *et al.* Leukaemic transformation by CALM-AF10 involves upregulation of Hoxa5 by hDOT1L. *Nat Cell Biol* **8**, 1017-1024, doi:10.1038/ncb1464 (2006).

31   Orlovsky, K. *et al.* Down-regulation of homeobox genes MEIS1 and HOXA in MLL-rearranged acute leukemia impairs engraftment and reduces proliferation. *Proc Natl Acad Sci U S A* **108**, 7956-7961, doi:10.1073/pnas.1103154108 (2011).

32   Kang, H. *et al.* Dysregulated transcriptional networks in KMT2A- and MLLT10-rearranged T-ALL. *Biomark Res* **6**, 27, doi:10.1186/s40364-018-0141-z (2018).

33   Chen, C. W. & Armstrong, S. A. Targeting DOT1L and HOX gene expression in MLL-rearranged leukemia and beyond. *Exp Hematol* **43**, 673-684, doi:10.1016/j.exphem.2015.05.012 (2015).

34   Dik, W. A. *et al.* CALM-AF10+ T-ALL expression profiles are characterized by overexpression of HOXA and BMI1 oncogenes. *Leukemia* **19**, 1948-1957, doi:10.1038/sj.leu.2403891 (2005).

35   Bond, J. *et al.* Cryptic XPO1-MLLT10 translocation is associated with HOXA locus deregulation in T-ALL. *Blood* **124**, 3023-3025, doi:10.1182/blood-2014-04-567636 (2014).

36   Bond, J. *et al.* NAP1L1-MLLT10 is a rare recurrent translocation that is associated with HOXA activation and poor treatment response in T-cell acute lymphoblastic leukaemia. *British Journal of Haematology* **174**, 470-473, doi:10.1111/bjh.13772 (2015).

37   Brandimarte, L. *et al.* DDX3X-MLLT10 fusion in adults with NOTCH1 positive T-cell acute lymphoblastic leukemia. *Haematologica* **99**, 64-66, doi:10.3324/haematol.2013.101725 (2014).

38   Chiaretti, S., Zini, G. & Bassan, R. Diagnosis and subclassification of acute lymphoblastic leukemia. *Mediterr J Hematol Infect Dis* **6**, e2014073, doi:10.4084/MJHID.2014.073 (2014).

39   Winkler, T. H. & Martensson, I. L. The Role of the Pre-B Cell Receptor in B Cell Development, Repertoire Selection, and Tolerance. *Front Immunol* **9**, 2423, doi:10.3389/fimmu.2018.02423 (2018).

40   Jensen, C. T., Lang, S., Somasundaram, R., Soneji, S. & Sigvardsson, M. Identification of Stage-Specific Surface Markers in Early B Cell Development Provides Novel Tools for Identification of Progenitor Populations. *J Immunol* **197**, 1937-1944, doi:10.4049/jimmunol.1600297 (2016).

41   Srivastava, A. & Makarenkova, H. P. Innate Immunity and Biological Therapies for the Treatment of Sjogren's Syndrome. *Int J Mol Sci* **21**, doi:10.3390/ijms21239172 (2020).

42   Whitmire, J. K., Eam, B. & Whitton, J. L. Mice deficient in stem cell antigen-1 (Sca1, Ly-6A/E) develop normal primary and memory CD4+ and CD8+ T-cell responses to virus infection. *Eur J Immunol* **39**, 1494-1504, doi:10.1002/eji.200838959 (2009).

43   Orkin, S. Diversification of Haematopoietic Stem Cells to Specific Lineages. *Nature Reviews Genetics* **1**, 57-64, doi:doi.org/10.1038/35049577 (2000).

44   Azzam, H. *et al.* CD5 Expression Is Developmentally Regulated By T Cell Receptor (TCR) Signals and TCR Avidity. *Journal of Experimental Medicine* **188**, 2301-2311, doi:10.1084/jem.188.12.2301 (1998).

45   Binder, C. *et al.* CD2 Immunobiology. *Front Immunol* **11**, 1090, doi:10.3389/fimmu.2020.01090 (2020).

46   Inukai, T. *et al.* Clinical significance of early T-cell precursor acute lymphoblastic leukaemia: results of the Tokyo Children's Cancer Study Group Study L99-15. *Br J Haematol* **156**, 358-365, doi:10.1111/j.1365-2141.2011.08955.x (2012).

47   Noronha, E. P. *et al.* The Profile of Immunophenotype and Genotype Aberrations in Subsets of Pediatric T-Cell Acute Lymphoblastic Leukemia. *Front Oncol* **9**, 316, doi:10.3389/fonc.2019.00316 (2019).

48   Upadhyay, G. Emerging Role of Lymphocyte Antigen-6 Family of Genes in Cancer and Immune Cells. *Front Immunol* **10**, 819, doi:10.3389/fimmu.2019.00819 (2019).

49   Hsu, Y. C., Mildenstein, K., Hunter, K., Tkachenko, O. & Mullen, C. A. Acute lymphoid leukemia cells with greater stem cell antigen-1 (Ly6a/Sca-1) expression exhibit higher levels of metalloproteinase activity and are more aggressive in vivo. *PLoS One* **9**, e88966, doi:10.1371/journal.pone.0088966 (2014).

50   Ohki, K. *et al.* Impact of immunophenotypic characteristics on genetic subgrouping in childhood acute lymphoblastic leukemia: Tokyo Children's Cancer Study Group (TCCSG) study L04-16. *Genes Chromosomes Cancer* **59**, 551-561, doi:10.1002/gcc.22858 (2020).

51   Noronha, E. P. *et al.* Immunophenotyping with CD135 and CD117 predicts the FLT3, IL-7R and TLX3 gene mutations in childhood T-cell acute leukemia. *Blood Cells Mol Dis* **57**, 74-80, doi:10.1016/j.bcmd.2015.12.003 (2016).

52   Tembhare, P. R. *et al.* Eleven-marker 10-color flow cytometric assessment of measurable residual disease for T-cell acute lymphoblastic leukemia using an approach of exclusion. *Cytometry B Clin Cytom* **100**, 421-433, doi:10.1002/cyto.b.21939 (2021).

53   Wang, Y. Z. *et al.* A seven-color panel including CD34 and TdT could be applied in >97% patients with T cell lymphoblastic leukemia for minimal residual disease detection independent of the initial phenotype. *Leuk Res* **72**, 12-19, doi:10.1016/j.leukres.2018.07.012 (2018).

54   Chagan-Yasutan, H. *et al.* Involvement of osteopontin and its signaling molecule CD44 in clinicopathological features of adult T cell leukemia. *Leuk Res* **35**, 1484-1490, doi:10.1016/j.leukres.2011.05.011 (2011).

55   Legras, S. *et al.* A Strong Expression of CD44-6v Correlates With Shorter Survival of Patients With Acute Myeloid Leukemia. *Blood* **91**, 3401-3413 (1998).

56   Marques, L. V. C. *et al.* CD44 Expression Profile Varies According to Maturational Subtypes and Molecular Profiles of Pediatric T-Cell Lymphoblastic Leukemia. *Front Oncol* **8**, 488, doi:10.3389/fonc.2018.00488 (2018).

57   Wang, Z., Zhao, K., Hackert, T. & Zoller, M. CD44/CD44v6 a Reliable Companion in Cancer-Initiating Cell Maintenance and Tumor Progression. *Front Cell Dev Biol* **6**, 97, doi:10.3389/fcell.2018.00097 (2018).

58   Rai, A. K. *et al.* Exonal switch down-regulates the expression of CD5 on blasts of acute T cell leukaemia. *Clin Exp Immunol* **190**, 340-350, doi:10.1111/cei.13019 (2017).

59   Jain, N. *et al.* Early T-cell precursor acute lymphoblastic leukemia/lymphoma (ETP-ALL/LBL) in adolescents and adults: a high-risk subtype. *Blood* **127**, 1863-1869, doi:10.1182/blood-2015-08-661702 (2016).

60   Vadillo, E., Dorantes-Acosta, E., Pelayo, R. & Schnoor, M. T cell acute lymphoblastic leukemia (T-ALL): New insights into the cellular origins and infiltration mechanisms common and unique among hematologic malignancies. *Blood Rev* **32**, 36-51, doi:10.1016/j.blre.2017.08.006 (2018).

61   Mansur, M. B. *et al.* Distinctive genotypes in infants with T-cell acute lymphoblastic leukaemia. *Br J Haematol* **171**, 574-584, doi:10.1111/bjh.13613 (2015).

62   Meijerink, J. P. Genetic rearrangements in relation to immunophenotype and outcome in T-cell acute lymphoblastic leukaemia. *Best Pract Res Clin Haematol* **23**, 307-318, doi:10.1016/j.beha.2010.08.002 (2010).

63   Abdelali, R. *et al.* The prognosis of CALM-AF10-positive adult T-cell acute lymphoblastic leukemias depends on the stage of maturation arrest. *Haematologica* **98**, 1711-1717, doi:10.3324/haematol.2013.086082 (2013).

64   Genescà, E. *et al.* Unique clinico-biological, genetic and prognostic features of adult early T-cell precursor acute lymphoblastic leukemia. *Haematologica* **105**, 1-4, doi:10.3324/haematol.2019.225078 (2020).

65   Burns, M. A. *et al.* Identification of prognostic factors in childhood T-cell acute lymphoblastic leukemia: Results from DFCI ALL Consortium Protocols 05-001 and 11-001. *Pediatr Blood Cancer* **68**, e28719, doi:10.1002/pbc.28719 (2021).

Chapter 4 – Generating an *in vivo* transplantation model of *DDX3X-MLLT10* and *PICALM-MLLT10* T-ALL

4.1 Introduction

DDX3X-MLLT10 and *PICALM-MLLT10* are poorly characterised recurrent genomic aberrations identified in approximately 10% of newly diagnosed T-ALL cases. To date, there are no published *in vitro* or *in vivo* models of *MLLT10*r T-ALL, and expression of *PICALM-MLLT10* in non-lineage committed haematopoietic precursors induces a myeloid or mixed phenotype[1]. Furthermore, no commercially available human T-ALL cell lines are known to express either *DDX3X-MLLT10* or *PICALM-MLLT10*. U937 cells express *PICALM-MLLT10*, exhibit an AML phenotype[2], and possess additional clinically relevant genomic aberrations including homozygous *TP53* loss-of-function, which is rare in *PICALM-MLLT10* T-ALL[3,4]. The AML phenotype and presence of additional genomic aberrations are likely to influence cancer biology and treatment response[5], making U937 a suboptimal system to study *MLLT10*r T-ALL. Generating a pre-clinical model of T-lymphoid-lineage *MLLT10*r leukaemia would enable the study of the biology and therapeutic sensitivities of *MLLT10*r in a T-ALL context.

In chapter 3, expression of *DDX3X-MLLT10* or *PICALM-MLLT10* in the murine CD4$^+$CD8$^+$ T-cell line MOHITO did not induce convincing evidence of oncogenicity. The occurrence of a *MLLT10*r may not be sufficient to induce leukaemogenesis alone; rather, certain cooperative mutations may be required to induce *MLLT10*r T-ALL. Alternatively, the MOHITO T-cell line may not serve as an appropriate model system for *MLLT10*r, potentially due to the CD4$^+$CD8$^+$ phenotype. CD4 and CD8 are expressed in 53% and 60% of T-ALL cases respectively[6], but there is currently no published data defining the expression of these markers specifically in *MLLT10*r T-ALL.

In this chapter, T-lineage cell models expressing *DDX3X-MLLT10* or *PICALM-MLLT10* were developed using less mature murine CD4$^-$CD8$^-$ thymocyte-derived cells harvested from p19^{Arf} null C57BL/6J (Arf$^{-/-}$) mice. Cells were retrovirally transduced with *DDX3X-MLLT10*-pRUFiG2 or *PICALM-MLLT10*-pRUFiG2 and transplanted into sub lethally irradiated 6-week-old *NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/Sz* (NSG) mice, to assess engraftment *in vivo*. The use of CD4$^-$CD8$^-$ thymocyte-derived cells enables exploration of the leukaemogenic potential of *MLLT10*r in an immature T-cell setting.

4.2 Approach

CD4$^-$CD8$^-$TER119$^-$ thymocytes were harvested from Arf$^{-/-}$ mice and retrovirally transduced with pRUFiG2 constructs using RetroNectin® to generate stable cell lines expressing *DDX3X-MLLT10* or *PICALM-MLLT10*, or control pRUFiG2 empty vector (refer to methods section **2.17**). Cells were expanded by sterile culture with OP9-DL1 stromal cells[7], and GFP$^+$ cells were collected by FACS using a BD FACSFusion^{TM}. The OP9-DL1 cell line is a bone-marrow-derived stromal cell line that ectopically expresses the Notch ligand, Delta-like 1 (DL1). With the addition of murine FLT3-ligand (mFLT3-L) and murine IL-7 (mIL-7), OP9-DL1 cells provide an appropriate microenvironment to promote lineage commitment, differentiation and proliferation of T-cells from murine haematopoietic progenitor cells[8,9].

To establish an *in vivo* model of *MLLT10*r T-ALL, NSG mice approximately 6 weeks of age were exposed to a sub lethal dose of radiation (200 cGy), and each was administered with 3x10^6 transduced cells by tail vein injection. Mice were monitored for engraftment by daily clinical assessment and fortnightly tail vein bleeds to monitor circulating GFP$^+$ cells in the peripheral blood. Mice were humanely killed when they appeared moribund, and haematopoietic organs were harvested for immunophenotyping, qRT-PCR of *HOX* genes, and histological analysis.

4.3 Results

4.3.1 Generation of stable-expressing *MLLT10*r Arf$^{-/-}$ murine T-lineage cells

Thymocytes were harvested from Arf$^{-/-}$ mice aged 4-12 weeks. Cells were co-stained for TER119 (erythrocyte lineage marker), CD4 and CD8a (mature T-cell lineage markers), and negative cells were collected by FACS (**Figure 4.1**). Cells were cultured with OP9-DL1 stromal cells for 3 days, prior to retroviral transduction with pRUFiG2 constructs. FACS was performed to determine transduction efficiency and collect GFP$^+$ cells, and FACS was repeated to enrich the proportion of GFP$^+$ cells to >90% (**Figure 4.2**).

RNA was extracted for qRT-PCR of fusion breakpoints, confirming that *DDX3X-MLLT10* and *PICALM-MLLT10* are expressed at similar levels in transduced cells (**Figure 4.3**). qRT-PCR revealed no significant difference in the expression of *HOX* cluster genes between empty vector, *DDX3X-MLLT10* and *PICALM-MLLT10* expressing cells (**Figure 4.4**). Immunophenotyping confirmed a heterogenous population of immature T-lineage cells and haematopoietic progenitors, prior to cryopreservation for later injection into mice. (**Figure 4.5**). The proportion of CD44$^+$ and CD5$^+$ cells was lower in *PICALM-MLLT10* expressing cells in comparison to *DDX3X-MLLT10* or empty vector, and expression of either fusion induced increased Sca-1 expression compared to empty vector cells. Statistical analysis was not performed, as there was only sufficient material for one replicate.
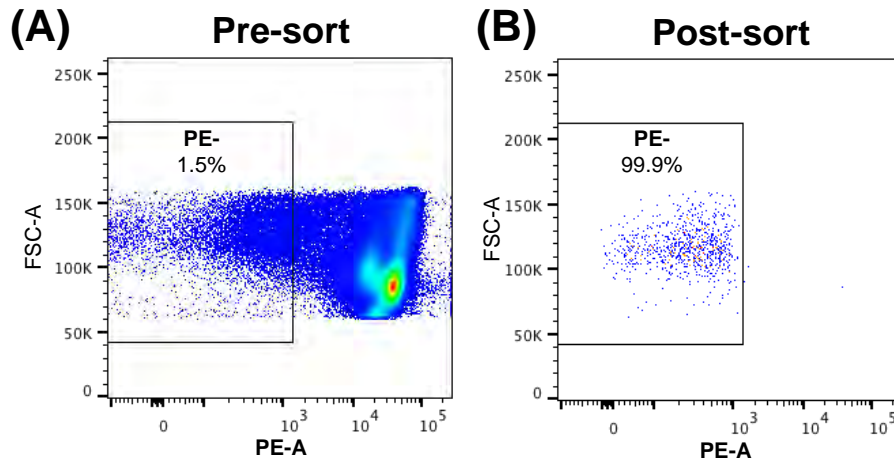
Figure 4.1 – FACS of CD4⁻CD8a⁻TER119⁻ thymocytes harvested from p19^Arf null C57BL/6J (Arf⁻/⁻) mice

Thymocytes were co-stained for CD4-PE, CD8a-PE and TER119-PE, and PE negative (PE-) cells were collected by FACS using a BD FACSFusion™. **(A)** Sorting gate was set around the PE- population as determined by unstained negative control cells, and **(B)** purity of sorted cells was confirmed immediately post-sort.
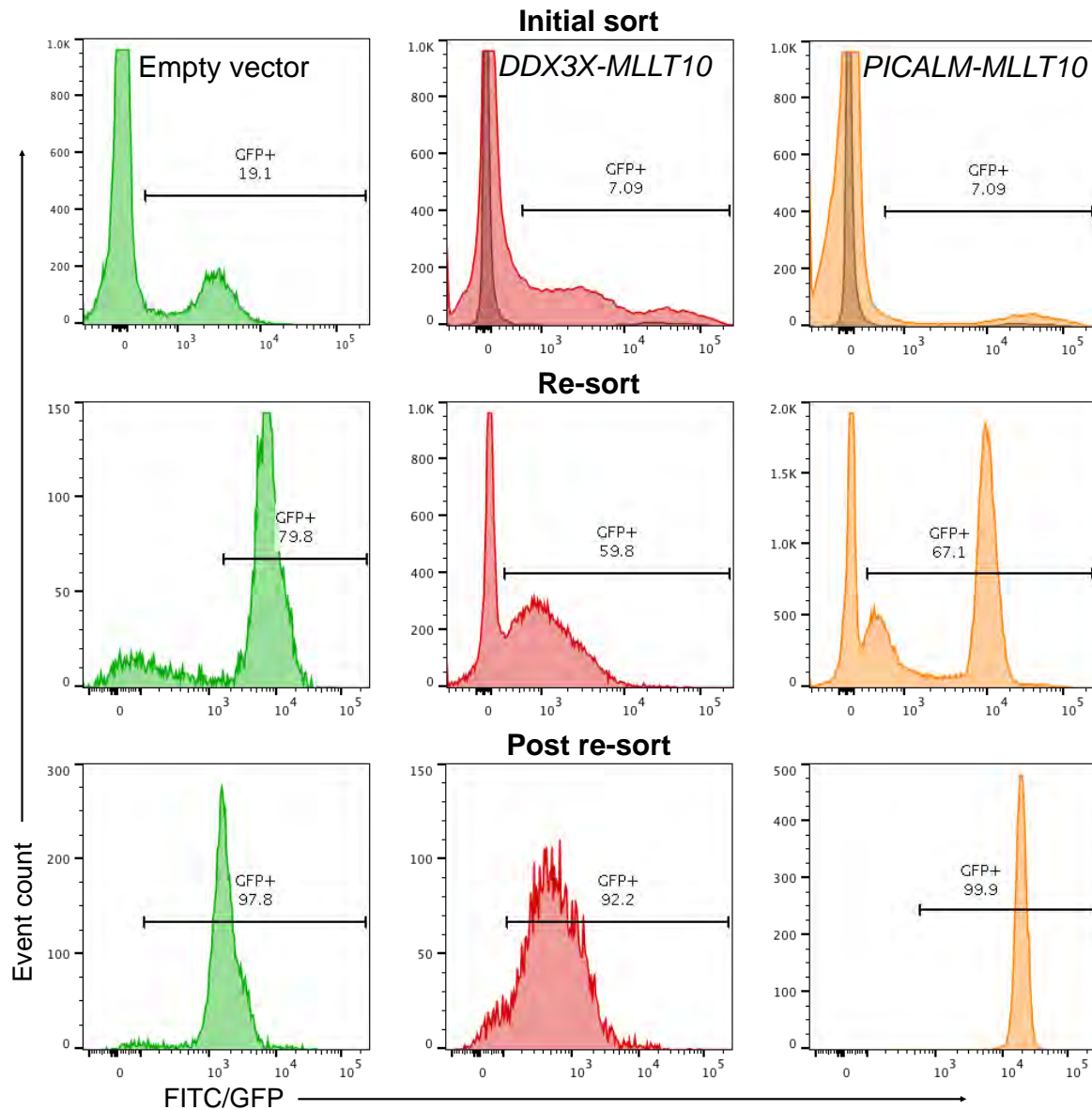
Figure 4.2 – CD4⁻CD8a⁻TER119⁻ Arf⁻ᐟ⁻ thymocytes transduced with fusion-pRUFiG2 constructs express GFP and are a homogenous population after sorting.

GFP⁺ cells were collected by FACS using either a BD FACSFusion™ or FACSMelody™ and further expanded by OP9-DL1 co-culture. Cells were expanded for one week, and FACS was repeated to obtain a population of >90% live GFP⁺ cells.
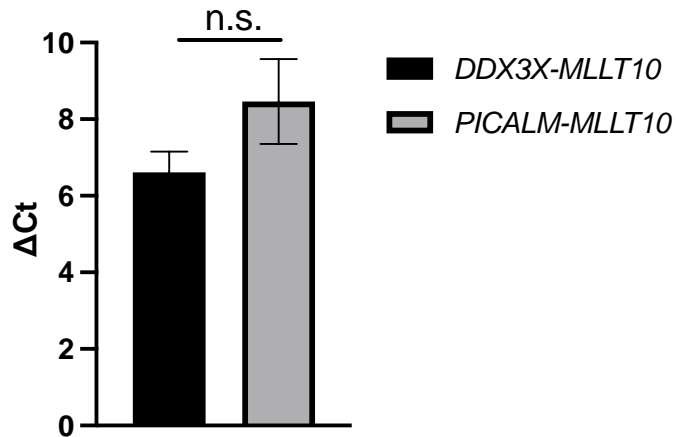
Figure 4.3 – *DDX3X-MLLT10* and *PICALM-MLLT10* pRUFiG2 constructs are expressed at similar levels in Arf$^{-/-}$ thymocytes

RNA was extracted from Arf$^{-/-}$ thymocytes (*DDX3X-MLLT10* and *PICALM-MLLT10*), and cDNA synthesised for SYBR® Green qRT-PCR analysis of fusion breakpoints. ΔCt values were calculated by normalising Ct values to *ACTB* housekeeping gene. A lower ΔCt value indicates higher expression. Neither fusion was expressed in empty vector cells. Statistical significance was determined by Student's t-test (n.s.=not significant). Error bars represent mean ± SD of 3 biological replicates.
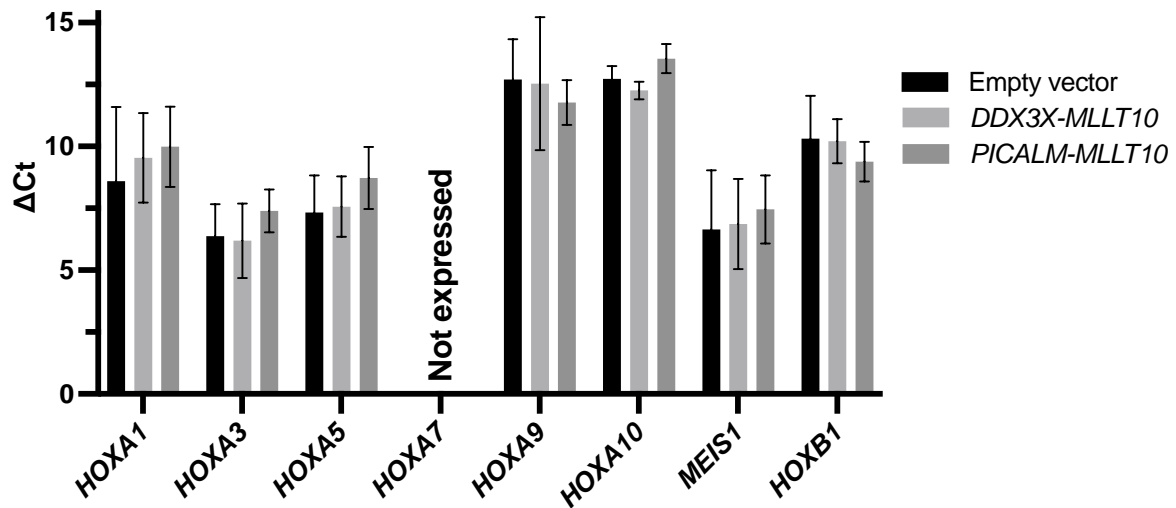
Figure 4.4 – Expression of *MLLT10*r does not induce changes in the expression of *HOX* genes in Arf^{-/-} thymocytes

RNA was extracted from Arf^{-/-} thymocytes (pRUFiG2 empty vector, *DDX3X-MLLT10* and *PICALM-MLLT10*), and cDNA synthesised for SYBR® Green qRT-PCR analysis of *HOX* genes. ΔCt values were calculated by normalising Ct values to *ACTB* housekeeping gene. A lower ΔCt value indicates higher expression. Differences in gene expression were statistically assessed by one-way ANOVA t-tests, and no significant differences were identified. Error bars represent mean ± SD of 3 biological replicates.
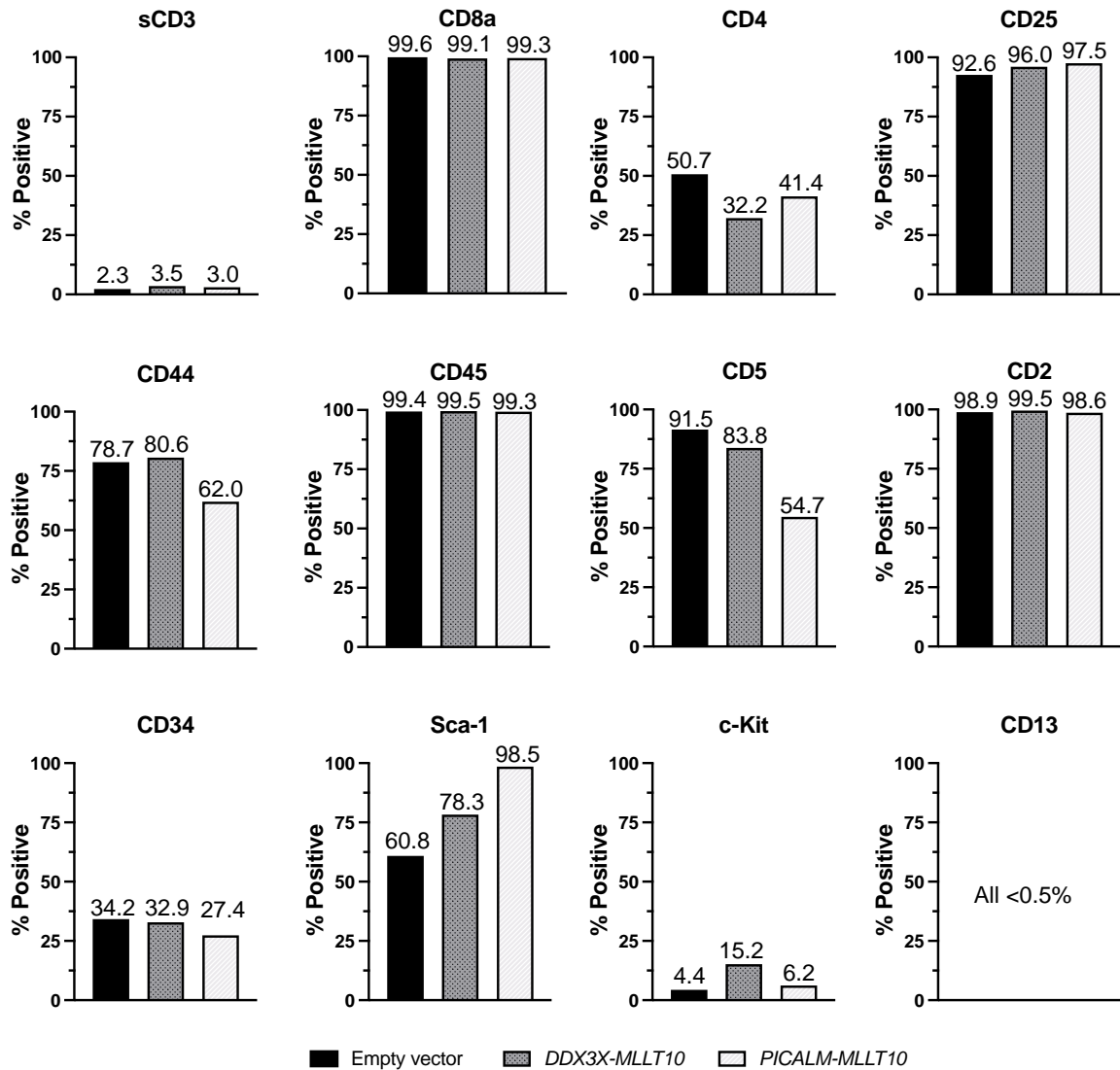
Figure 4.5 – Transduced Arf$^{-/-}$ thymocytes maintain a T-lineage immunophenotype

Cell surface marker staining was analysed on GFP$^+$ Arf$^{-/-}$ thymocytes on the day of cryopreservation using a BD FACSFortessa^{TM}. Gates were set based on isotype controls (**Supplementary Figure 4.1**), to determine proportion of positive single cells for each marker. Data represents one biological replicate only, as there was insufficient material for additional replicates.

4.3.2 Transduced Arf$^{-/-}$ thymocytes induce T-ALL in immunodeficient mice

A total of 3x10^6 live cells per mouse were injected into sub lethally irradiated NSG mice. Fifteen mice were included in the experiment (5 mice in each cohort: empty vector, *DDX3X-MLLT10* and *PICALM-MLLT10*). Mice were monitored for engraftment by daily clinical monitoring and fortnightly tail vein bleeds to assess for GFP$^+$ cells, and were humanely killed when they appeared moribund. Photographs documenting external organ morphology, FSC/SSC scatter plots and GFP expression histograms are provided for individual mice in **Appendices 4.1-4.3**. In all scatter plots, each mouse is depicted with a unique colour that is reported consistently throughout this chapter.

Empty vector, *DDX3X-MLLT10* and *PICALM-MLLT10* mice survived for a median of 117, 132 and 149 days post-injection respectively, indicating no significant difference in median survival (**Figure 4.6**). Three empty vector mice and one *DDX3X-MLLT10* mouse died without evidence of GFP$^+$ cell engraftment (discussed further in sections **4.4.9** and **4.4.10**). Peripheral blood, bone marrow from hind limbs and peripheral organs (spleen, liver and thymus if present) were harvested for analysis of GFP expression (**Figure 4.7**). Two empty vector mice demonstrated evidence of GFP$^+$ engraftment. EV-5 was culled due to clinical signs of illness on day 138, likely due to a grossly enlarged thymus containing 57.3% GFP$^+$ cells (**Appendix 4.1E**). EV-4 was culled at the experimental endpoint on day 234, and there was no evidence of GFP$^+$ cells in peripheral blood 1 week prior to humane killing.

Flow cytometric analyses revealed that all five *PICALM-MLLT10* mice exhibited >90% GFP$^+$ live single cells in all organs. In contrast, localisation of GFP$^+$ engraftment was variable within the *DDX3X-MLLT10* cohort. One mouse did not meet the criteria for leukaemic engraftment (DDX3X-1, **Appendix 4.2A**). The remaining mice exhibited thymus-predominant (DDX3X-2 and 5) or bone marrow-predominant (DDX3X-3 and 4) disease. DDX3X-2 and 5 had grossly enlarged thymi containing 97.0% and 41.0% GFP$^+$ cells, with minimal bone marrow involvement (1.5% and 9.7%) (**Appendix 4.2B & E**). In contrast, GFP$^+$ engraftment was observed in the bone marrow of DDX3X-3 and 4 (65.6% and 95.7%). GFP$^+$ cells were also identified in the thymi of DDX3X-3 and 4 (30.5% and 72.6%), but thymi were comparatively smaller, less fibrous and yielded few live cells (**Appendix 4.2C-D**).

No significant difference in liver or spleen weight (**Figure 4.8**), total white cell count (WCC), lymphocytes, neutrophils, red blood cells (RBC) or haemoglobin (Hb) (**Figure 4.9**) was observed between cohorts. Median organ weights, total WCC, lymphocyte count and neutrophil count were elevated, and RBC count and Hb were consistently low, relative to NSG reference values (**Supplementary Table 4.1**).

Sections of liver, spleen, thymus and femur were fixed in 10% formalin, prior to paraffin-embedding and sectioning for immunohistochemical staining with hematoxylin and eosin (H&E) and GFP antibody (**Figure 4.10**). All *PICALM-MLLT10* mice demonstrated widespread populations of densely packed small round cells with large nuclei in the bone marrow, liver, spleen and thymus, with diffuse GFP$^+$ staining evident. *DDX3X-MLLT10* mice with bone marrow predominant disease also demonstrated leukaemic infiltration in the liver. The thymi of *DDX3X-MLLT10* mice with thymus-predominant disease were infiltrated with GFP$^+$ lymphoblasts, whereas their bone marrow, liver and spleen retained normal structural features, with only minimal GFP staining evident in the bone marrow. IHC from empty vector mice with GFP$^+$ engraftment (mice EV-4 and EV-5) revealed scant GFP$^+$ cells in the bone marrow, liver and spleen (**Figure 4.11**).

RNA was extracted from liver specimens for quantification of *HOX* expression (**Figure 4.12**), revealing no statistically significant difference in *HOX* expression between cohorts. The liver was chosen as the comparator organ as this tissue demonstrated the most consistent leukaemic infiltration across cohorts. RT-PCR confirmed that the relevant fusion was expressed in each mouse except for DDX3X-1 (**Figure 4.13**). Neither fusion was detected in empty vector mice, confirming the absence of cross-contamination or endogenous fusion expression (**Figure 4.14**). The presence of empty pRUFiG2 multiple cloning site (MCS) was confirmed by RT-PCR and Sanger sequencing in mice transplanted with empty vector cells (**Supplementary Figure 4.2A-B**). The expected 245 bp sequence corresponding to the pRUFiG2 MCS was amplified in all mice except EV-3, with only faint bands observed in EV-1 and EV-2. An additional ~500 bp sequence was amplified in the positive control and EV-2, and Sanger sequencing confirmed non-specific amplification of a region from mouse chromosome 11 (**Supplementary Figure 4.2C**).
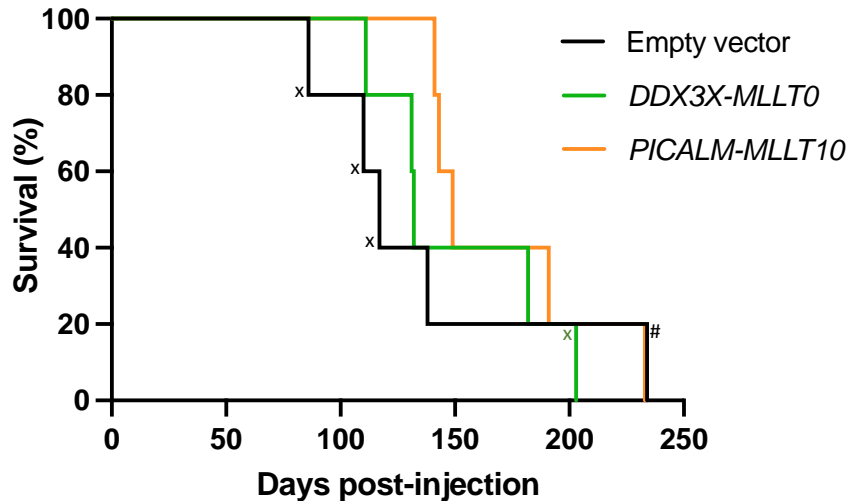
Figure 4.6 – Mice transplanted with Arf⁻ᐟ⁻ thymocytes expressing empty vector, *DDX3X-MLLT10* or *PICALM-MLLT10* survived for a median of 117, 132 and 149 days respectively

Kaplan-Meier curves depicting the survival of mice transplanted with cells expressing control empty vector, *DDX3X-MLLT10* or *PICALM-MLLT10*. No statistically significant difference in survival was identified (log-rank test). Data points marked with × represent mice who did not meet the criteria for leukaemic engraftment (criteria for leukaemic engraftment: total WCC >25 K/µL or ≥20% GFP⁺ live single cells in any organ). One empty vector mouse, marked by #, was culled at the experiments endpoint of 234 days post-injection.
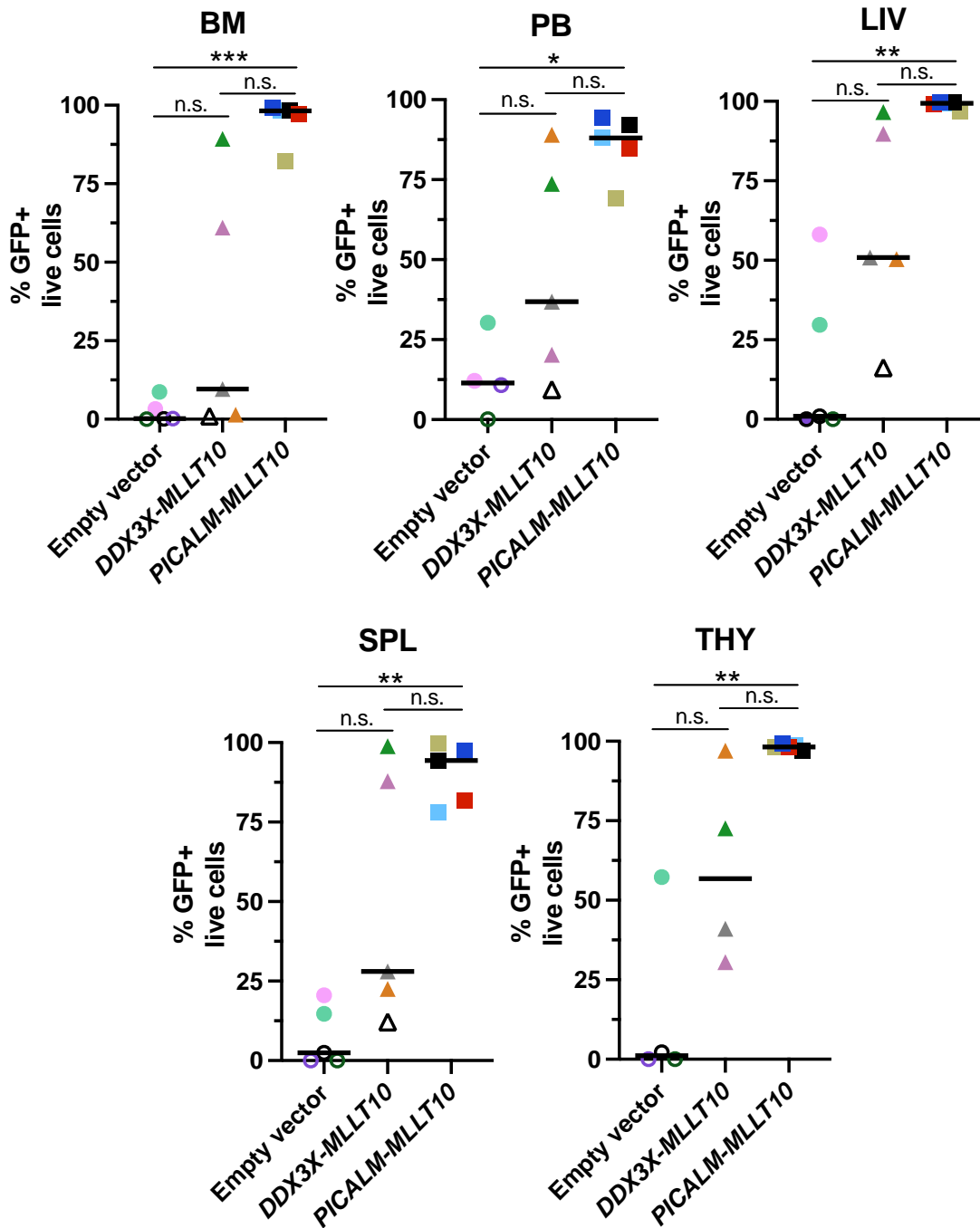
Figure 4.7 – Arf$^{-/-}$ thymocytes expressing *PICALM-MLLT10* induce greater engraftment in peripheral blood and haematopoietic organs, compared with cells expressing empty vector or *DDX3X-MLLT10*

Single cells were harvested from each organ and analysed for GFP expression on a BD FACSFortessa^{TM} or BD FACSCanto^{TM}. Percentage of GFP$^+$ cells was calculated as a proportion of total live cells. Each data point represents one animal, and the median of each cohort is represented by a horizontal bar. All statistics represent Kruskal-Wallis tests (*p<0.05, **p<0.01, ***p<0.001, n.s.=not significant). Hollow data

points ○ and △ represent mice who did not meet the criteria for leukaemic engraftment. Abbreviations: BM, bone marrow; PB, peripheral blood; LIV, liver; SPL, spleen; THY, thymus. Percentage of GFP$^+$ cells in each organ for individual mice are provided in **Supplementary Table 4.2**.
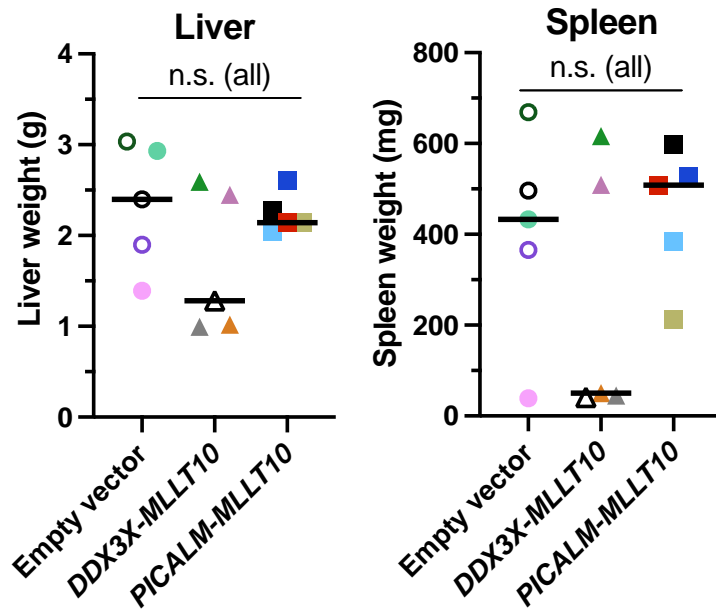
Figure 4.8 – Expression of *DDX3X-MLLT10* or *PICALM-MLLT10* does not induce significant changes in spleen or liver weight, compared with empty vector

Mouse spleens and livers were weighed following dissection. Statistical significance was assessed by Kruskal-Wallis tests (n.s.=not significant). Each data point represents one animal, and the median of each cohort is represented by a horizontal bar. Hollow data points ○ and △ represent mice who did not meet the criteria for leukaemic engraftment.
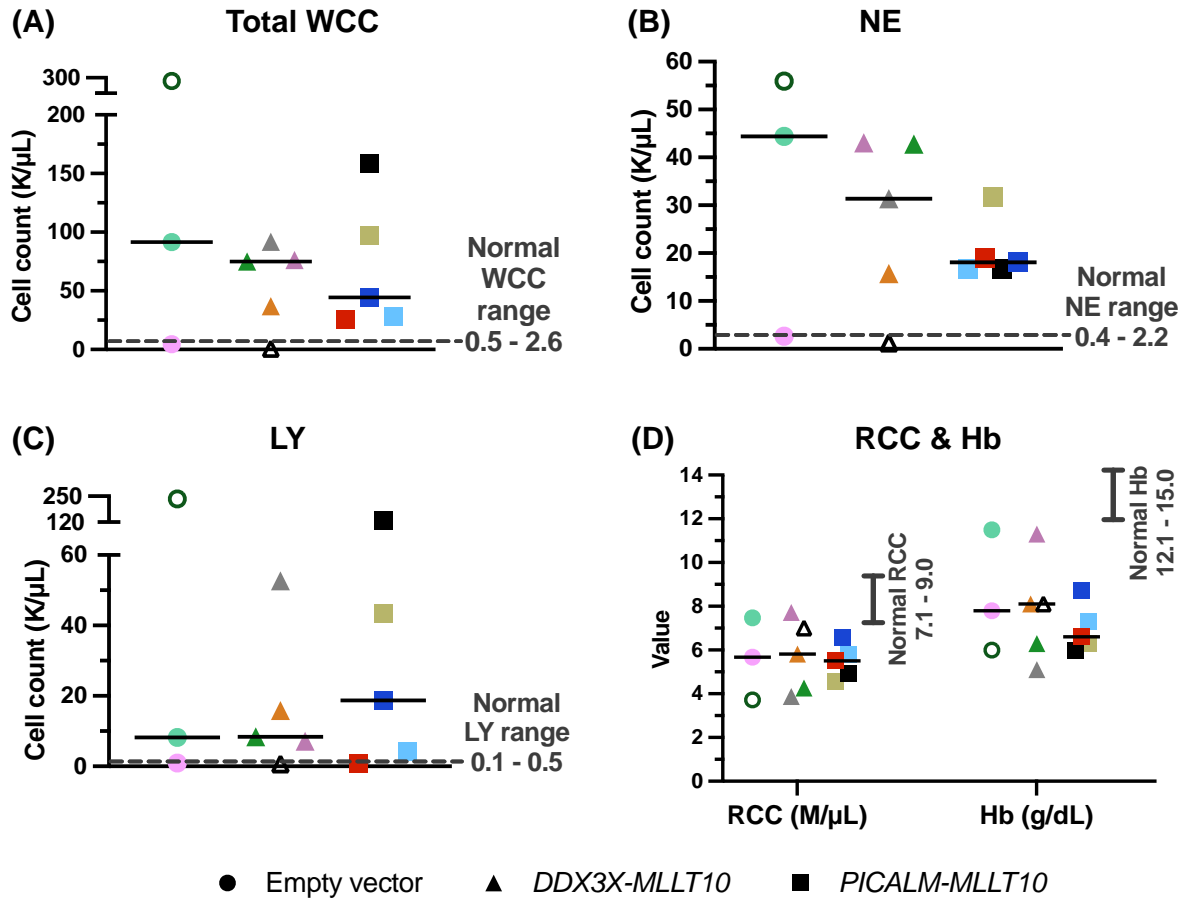
Figure 4.9 – Mice transplanted with Arf⁻/⁻ thymocytes expressing empty vector, *DDX3X-MLLT10* or *PICALM-MLLT10* exhibit abnormal blood profiles

Whole blood from each mouse was analysed using the HEMAVET® 950 veterinary haematology system to obtain values for **(A)** total white cell count, **(B)** neutrophil count, **(C)** lymphocyte count, **(D)** red cell count and haemoglobin concentration. Two empty vector mice did not have blood collected due to technical issues. NSG reference ranges are annotated on each graph[10]. No statistically significant difference between cohorts was observed, as determined by Kruskal-Wallis tests. Each data point is representative of one animal, and the median of each cohort is represented by a horizontal bar. Hollow data points ○ and △ represent mice who did not meet the criteria for successful GFP⁺ cell engraftment. Abbreviations: K/µL, thousand cells per microlitre; Hb, haemoglobin; LY, lymphocytes; M/µL, million cells per µL; NE, neutrophils; RCC, red cell count; WCC, white cell count. Blood profile values for each mouse are provided in **Supplementary Table 4.1**.

(A) H&E

	PICALM-MLLT10	DDX3X-MLLT10 Thymus predominant	DDX3X-MLLT10 Bone marrow predominant
Bone marrow (40x)			
Spleen (40x)			
Liver (40x)			
Thymus (40x)			N/A

(B)

GFP
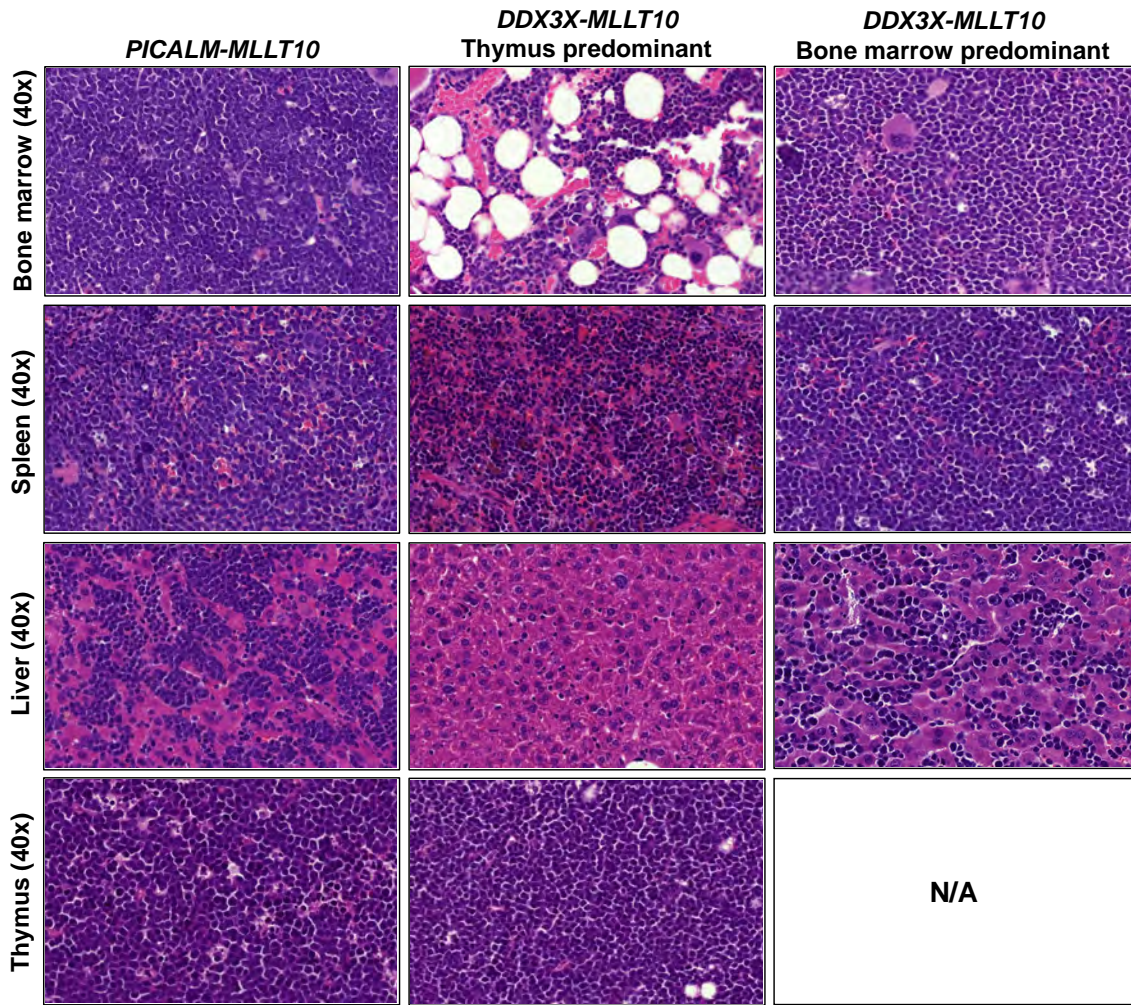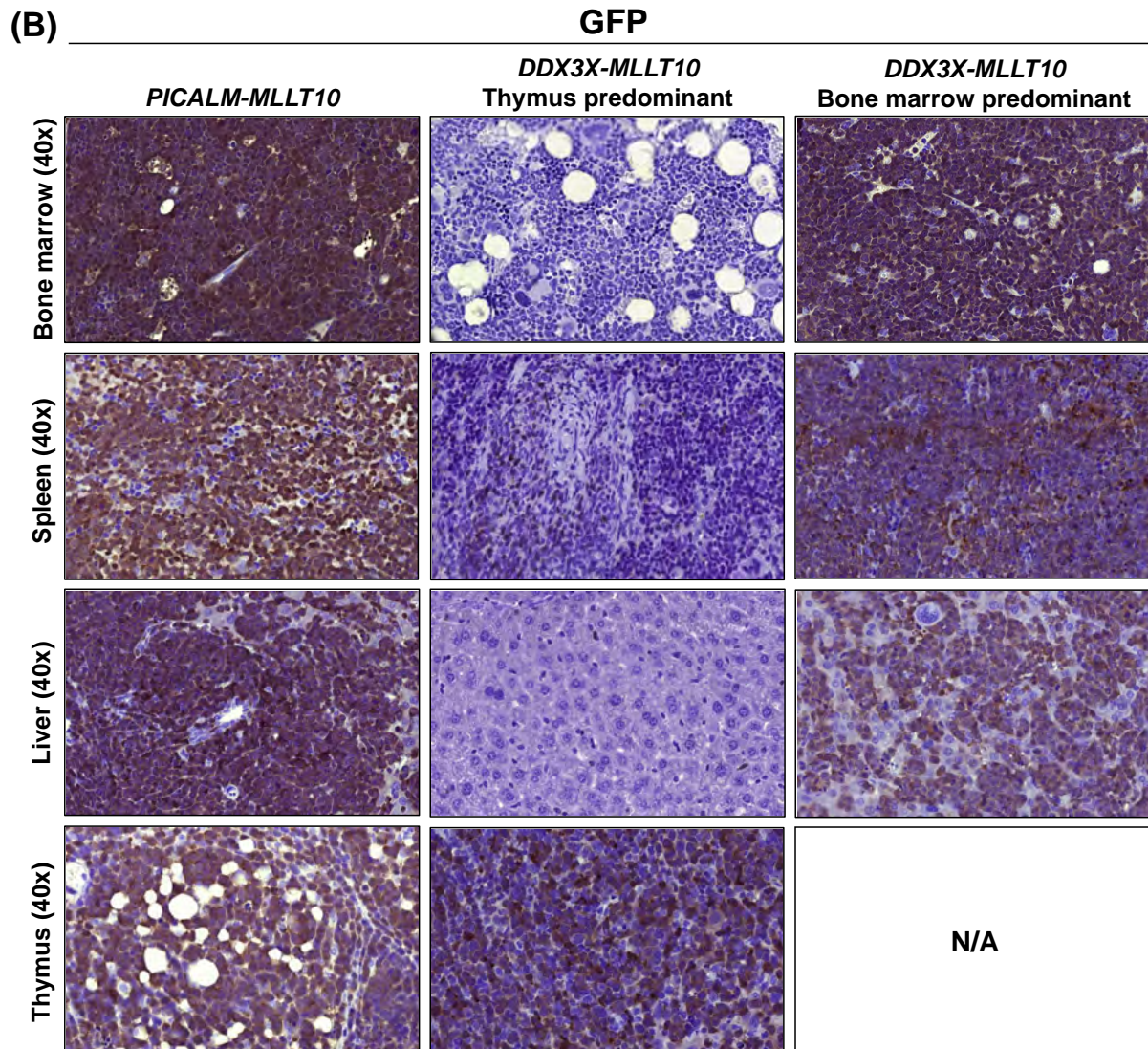
Figure 4.10 – Mice transplanted with *PICALM-MLLT10* or *DDX3X-MLLT10* expressing cells exhibit histological evidence of leukaemia

Representative **(A)** H&E and **(B)** GFP staining of bone marrow, spleen, liver and thymus sections from *PICALM-MLLT10* and *DDX3X-MLLT10* mice. Images analysed at 40x magnification with CaseViewer software. A representative no primary GFP antibody control is provided in **Supplementary Figure 4.3**. Abbreviations: N/A, not applicable.
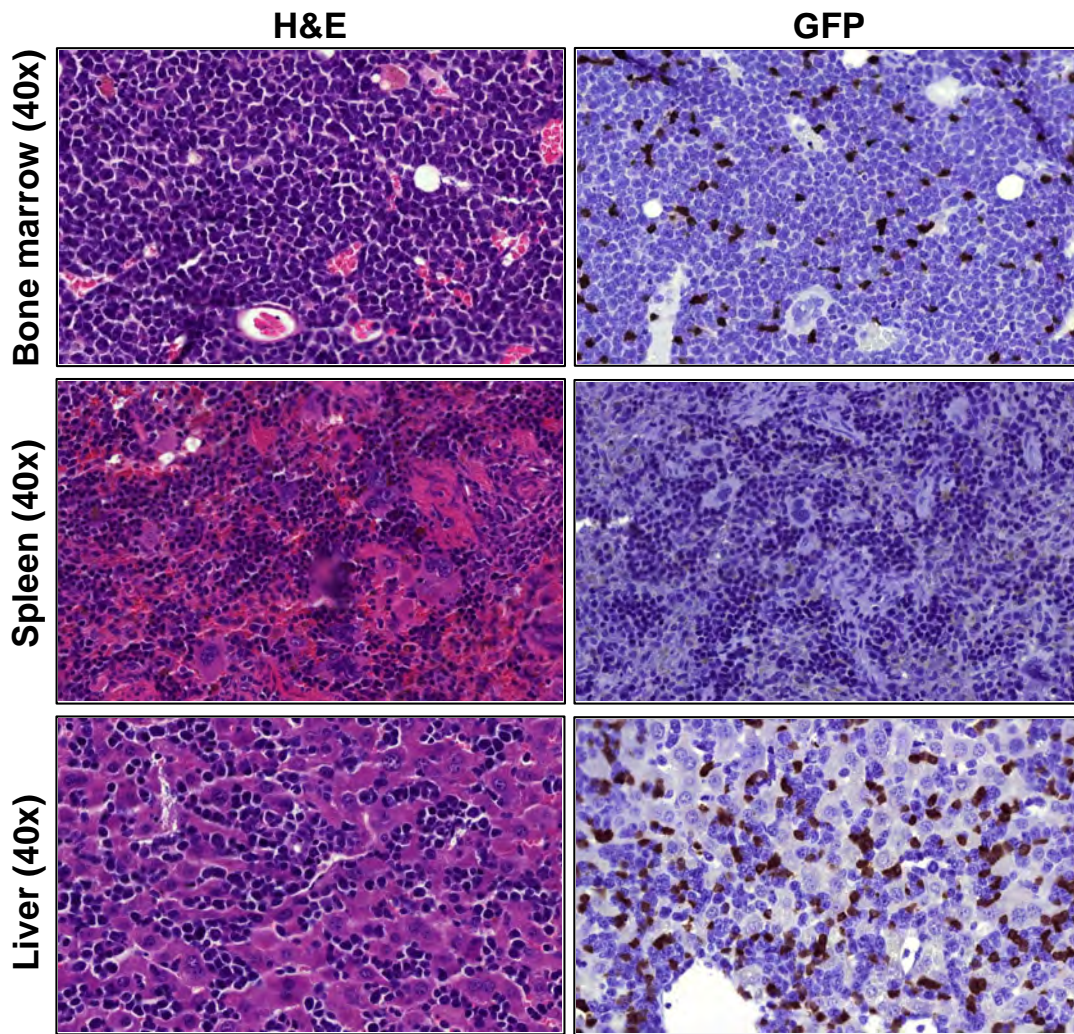
Figure 4.11 – Empty vector Arf$^{-/-}$ thymocytes engrafted in two mice

Representative H&E and GFP staining of sections from EV mice 4 and 5 bone marrow, spleen and liver. Images analysed at 40x magnification with CaseViewer software.
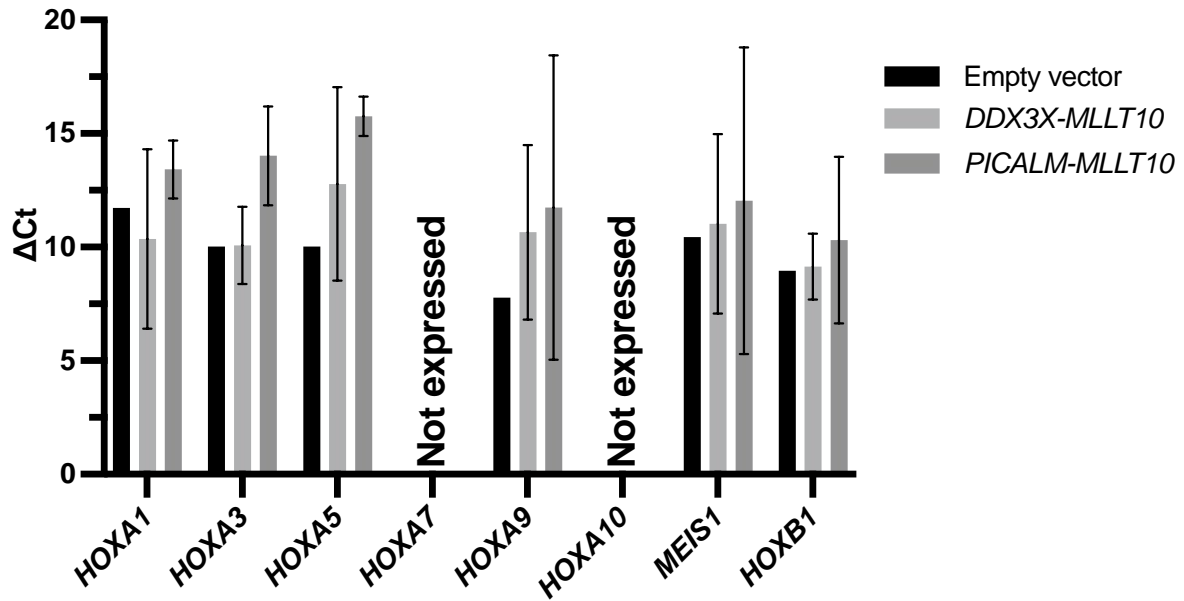
Figure 4.12 – *HOX* gene expression remains consistent in cells harvested from murine organs following transplantation

Liver RNA was extracted from each mouse, and cDNA synthesised for SYBR® Green qRT-PCR analysis of *HOX* genes. Comparison of ΔCt values between *DDX3X-MLLT10* and *PICALM-MLLT10* are shown, where Ct values were normalised to *ACTB* housekeeping gene. Differences in gene expression were statistically assessed by one-way ANOVA t-tests, and no significant differences were identified. Error bars represent mean ± SD for each cohort.
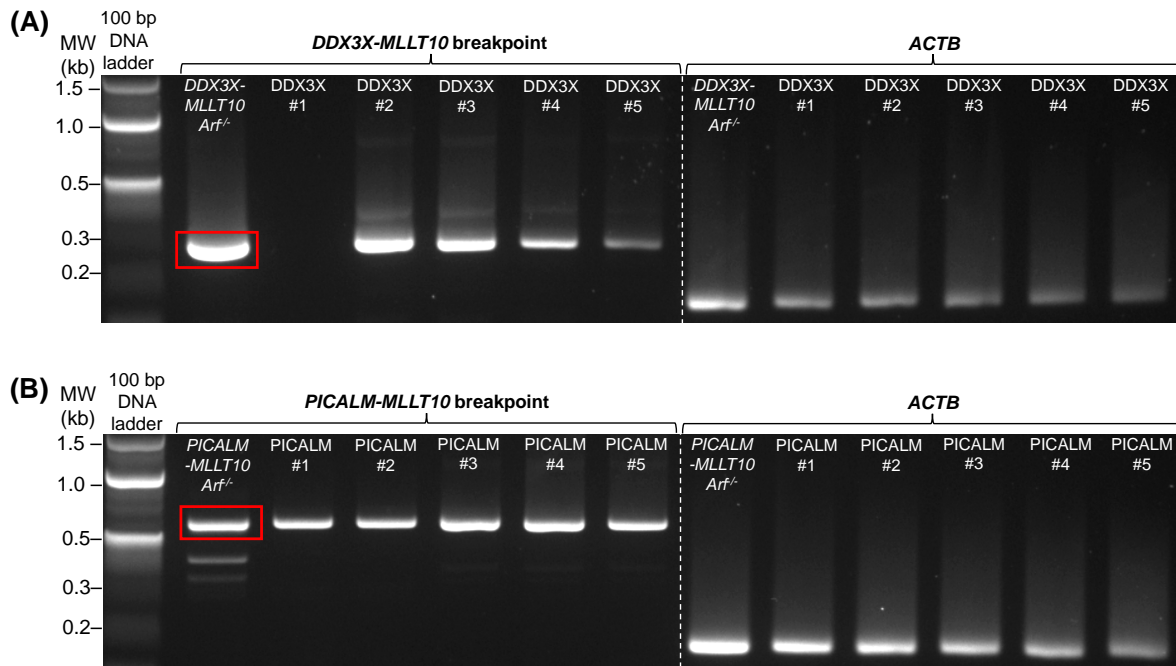
Figure 4.13 – *DDX3X-MLLT10* or *PICALM-MLLT10* are expressed in cells harvested from engrafted mice

Gel electrophoresis of RT-PCR fusion breakpoint products from mice transplanted with cells expressing **(A)** *DDX3X-MLLT10* (334 bp) and **(B)** *PICALM-MLLT10* (615 bp). Liver RNA was extracted from each mouse, and cDNA synthesised for PCR. Red rectangles highlight the presence of each fusion gene breakpoint in positive control samples. *ACTB* was used as a positive control for PCR. Abbreviations: MW, molecular weight marker; *ACTB*, actin.
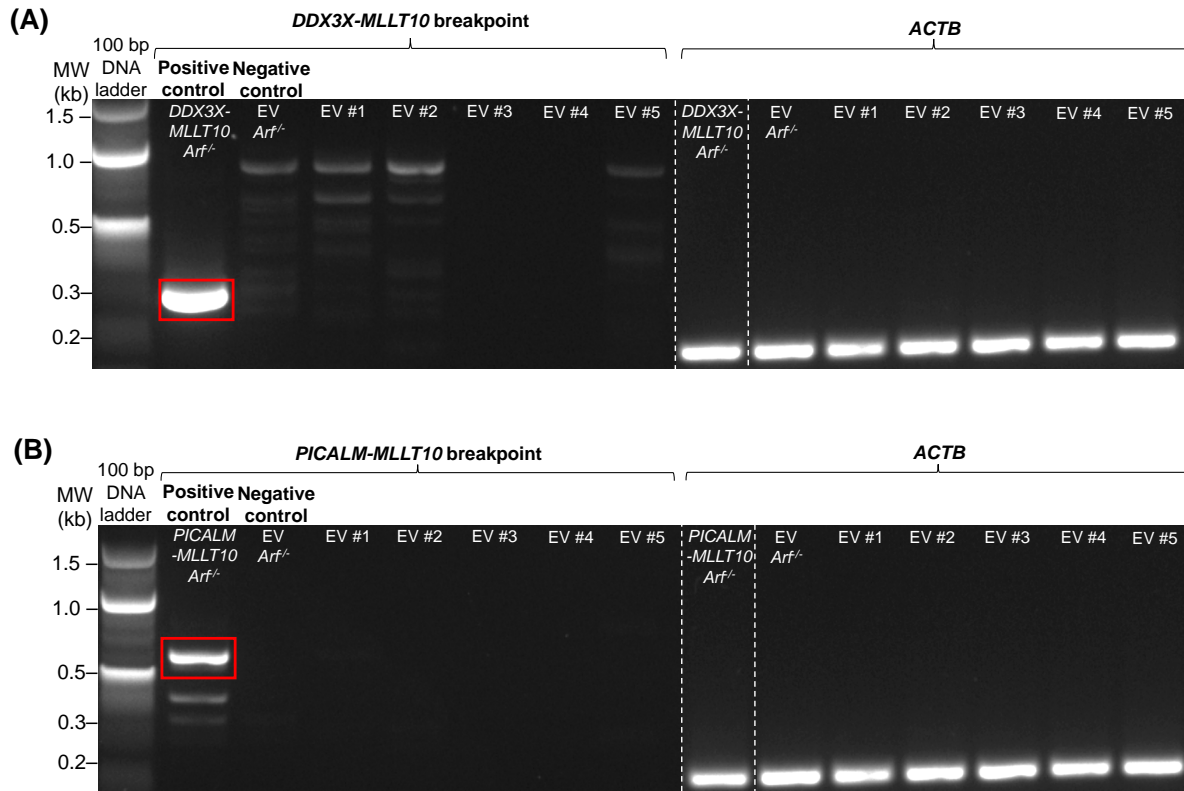
Figure 4.14 – Mice transplanted with pRUFiG2 empty vector expressing Arf⁻/⁻ cells do not express *DDX3X-MLLT10* or *PICALM-MLLT10*

Gel electrophoresis of RT-PCR products from empty vector mice. Liver RNA was extracted from each mouse, and cDNA synthesised for PCR of **(A)** *DDX3X-MLLT10* and **(B)** *PICALM-MLLT10* breakpoints. Red rectangles highlight the presence of each fusion gene breakpoint in positive control samples. *ACTB* was used as a positive control for PCR. Abbreviations: MW, molecular weight marker; *ACTB*, actin.

4.3.3 Cells expressing *PICALM-MLLT10* or *DDX3X-MLLT10* retain a CD4$^+$CD8$^+$ T-ALL immunophenotype with aberrant stem cell marker expression

GFP$^+$ cells were immunophenotyped as per the protocol described in **Table 2.11**. Surface marker staining was performed on cells harvested from bone marrow (or peripheral blood where bone marrow material was insufficient) (**Figure 4.15**), and thymi when sufficient material was present (**Figure 4.16**).

Cells harvested from both fusion cohorts retained an immunophenotype most consistent with early cortical T-ALL, based on the CD2$^+$CD4$^+$CD8$^+$CD5$^+$sCD3$^{-/dim}$ phenotype observed (**Table 3.3**). In humans, CD1a is a key marker that defines the early cortical subtype, but this was not examined here as mice do not express the CD1a isoform[11]. A striking loss of CD25 expression was observed in cells harvested from mice (median <10% CD25$^+$ cells for both cohorts), in comparison to pre-injection levels (>96% CD25$^+$). One exception was noted, a *PICALM-MLLT10* mouse that exhibited 75.7% (bone marrow) and 33.1% (thymus) CD25$^+$ cells. CD44 expression was variable and a trend towards lower CD44 expression in the thymi of *PICALM-MLLT10* mice was observed, but this was not statistically significant.

The stem cell markers CD34, c-Kit and Sca-1 were aberrantly expressed. *PICALM-MLLT10* cells exhibited a trend towards greater CD34 expression (bone marrow: 37.9% vs 98.8% and thymus: 72.9% vs 99.7%, *DDX3X-MLLT10* vs *PICALM-MLLT10*, not statistically significant). *DDX3X-MLLT10* cells exhibited a trend towards greater expression of c-Kit (bone marrow: 48.9% vs 25.0% and thymus: 52.4% vs 15.8%, *DDX3X-MLLT10* vs *PICALM-MLLT10*, not statistically significant). Sca-1 expression was significantly greater in *DDX3X-MLLT10* thymi than *PICALM-MLLT10* thymi (92.6% vs 56.8%, p=0.04) (**Figure 4.16**).

EV-5 was the only empty vector mouse with sufficient engraftment of GFP$^+$ cells for immunophenotyping, where >20% GFP$^+$ cells were identified in the peripheral blood, thymus and liver (**Supplementary Table 4.2**). GFP$^+$ cells from the peripheral blood of EV-5 exhibited an unusual mixed/ETP phenotype, including expression of myeloid marker CD13, mature T-cell marker sCD3 and stem cell markers Sca-1 and c-Kit, and heterogenous CD34 and CD2 (**Figure 4.18**).
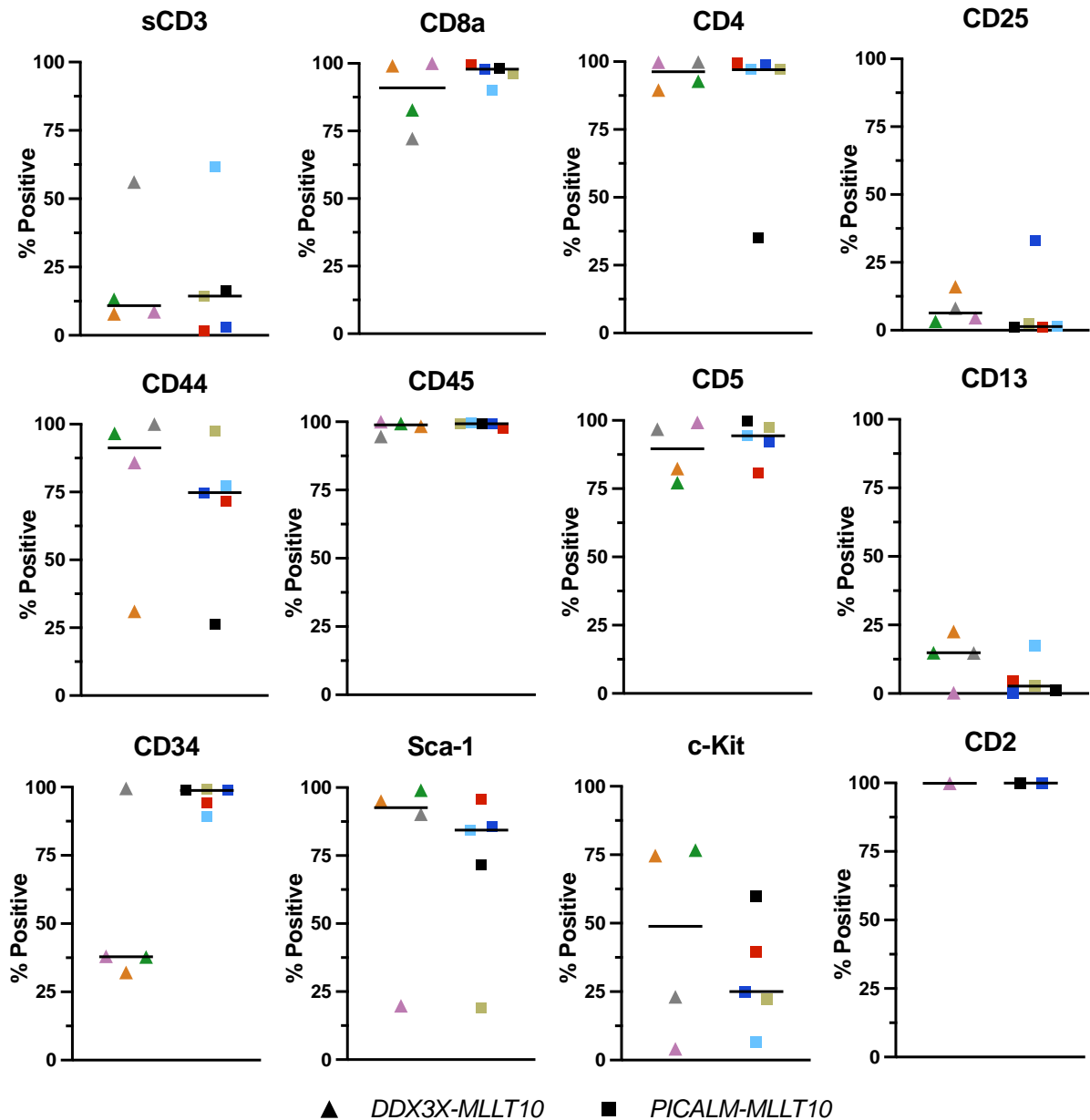
Figure 4.15 – GFP⁺ cells expressing *DDX3X-MLLT10* or *PICALM-MLLT10* harvested from murine bone marrow or peripheral blood exhibit a T-ALL immunophenotype

Cell surface marker staining was analysed on GFP⁺ cells harvested from the bone marrow or peripheral blood of mice, using a BD FACSFortessa™. Gating was set based on isotype controls to determine the proportion of positive cells for each marker (**Supplementary Figure 4.1**). Each data point represents one mouse, and the median of each cohort is represented by a horizontal bar. No statistically significant difference between cohorts was identified by Mann Whitney U tests. Immunophenotyping data for each mouse is provided in **Supplementary Tables 4.3-4.4.**
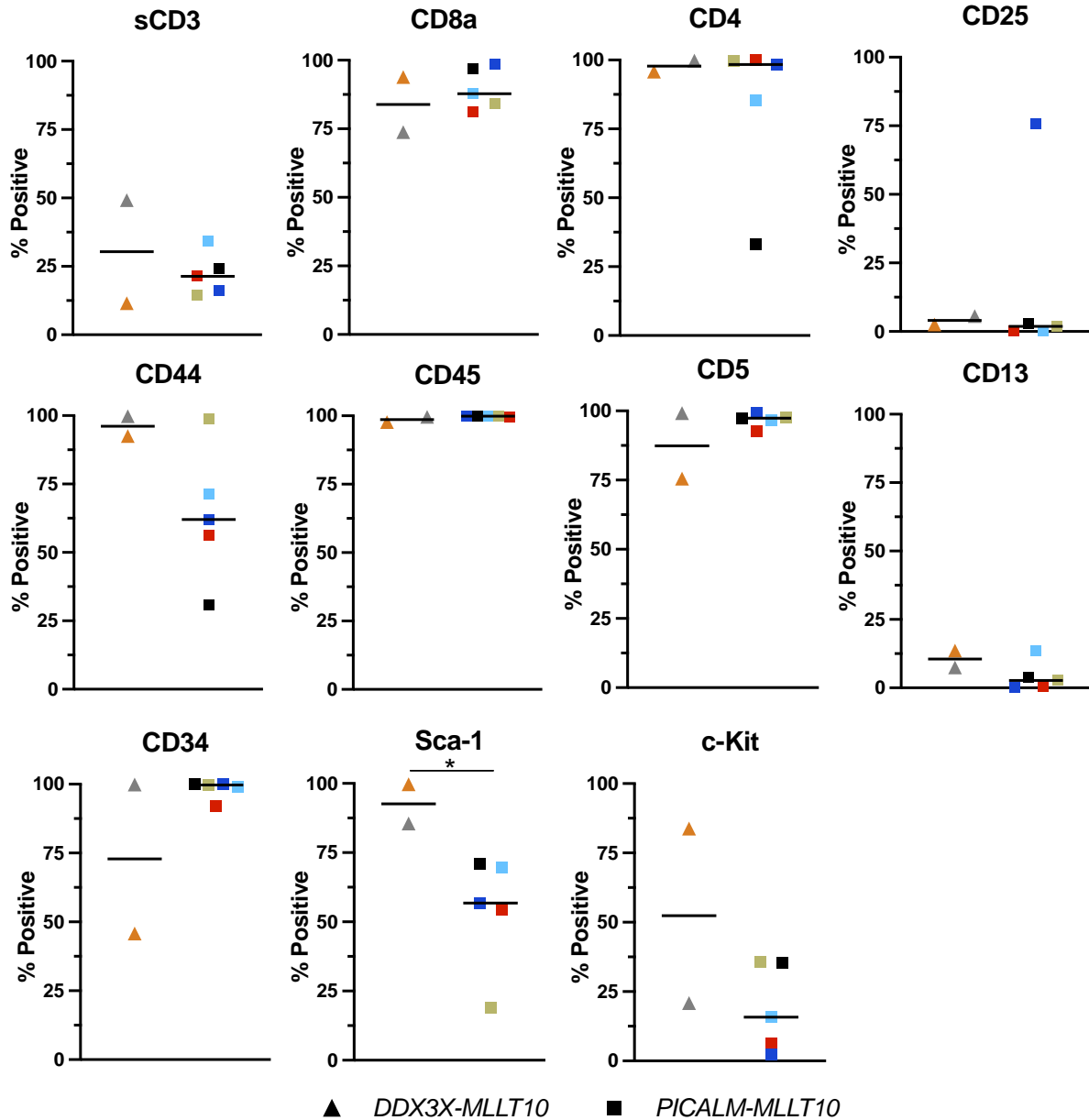
Figure 4.16 – GFP⁺ cells expressing *DDX3X-MLLT10* or *PICALM-MLLT10* harvested from murine thymi exhibit a T-ALL immunophenotype

Cell surface marker staining was analysed on GFP⁺ single cells using a BD FACSFortessa™. Gating was set based on isotype controls, to determine percentage of positive cells for each marker. Each data point represents one animal, and the median of each cohort is represented by a horizontal bar. CD2 was not investigated due to external limitations relating to antibody availability at the time of assay. Statistical significance was determined by Mann Whitney U tests (*p<0.05). On plots where statistical annotation is not present, the results were not significant. Immunophenotyping data for each mouse is provided in **Supplementary Tables 4.3-4.4**.
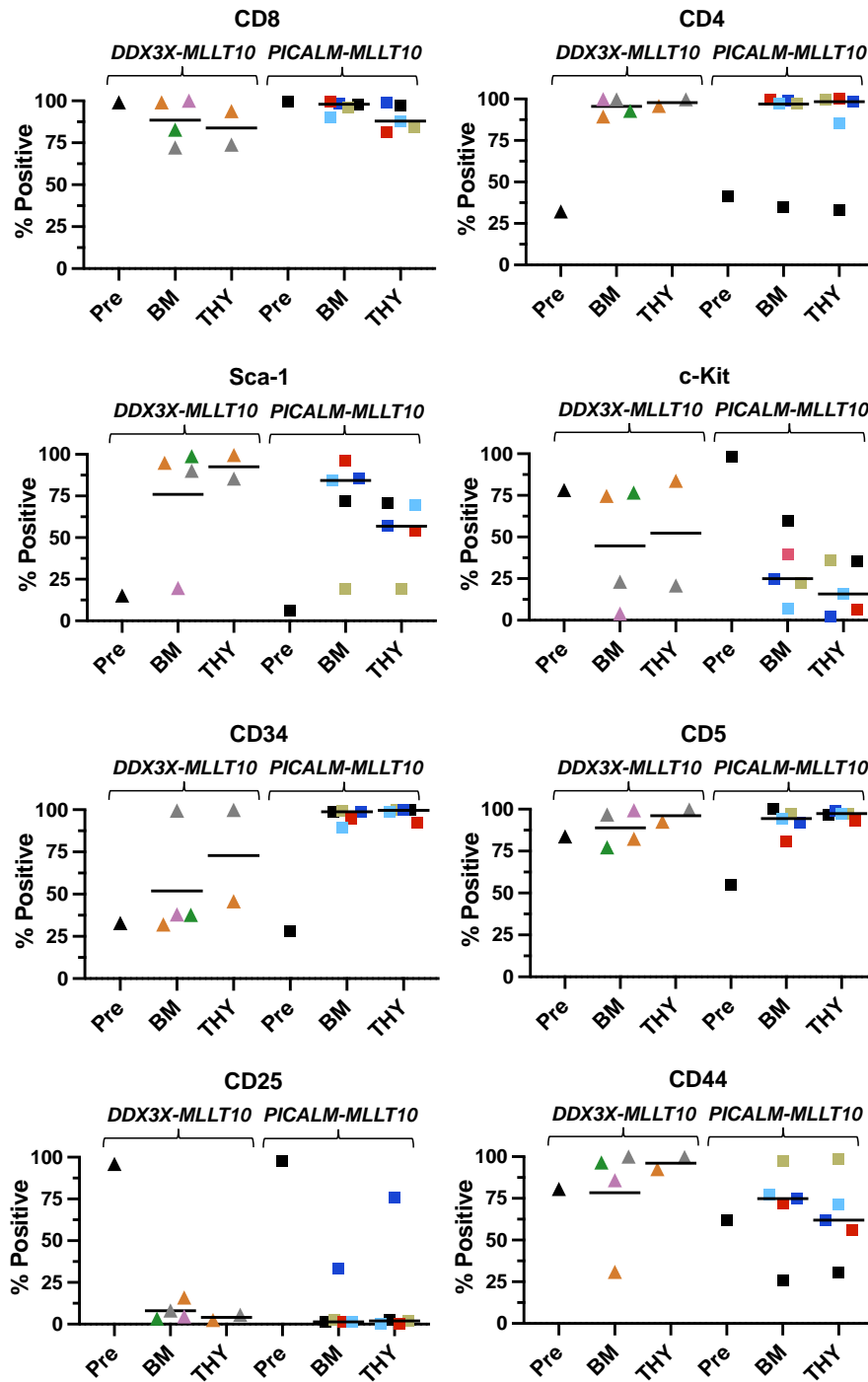
Figure 4.17 – Summary of *DDX3X-MLLT10* and *PICALM-MLLT10* Arf$^{-/-}$ thymocyte surface marker expression

A summary of CD4, CD8, Sca-1, c-Kit, CD34, CD5, CD25 and CD44 expression is provided for Arf$^{-/-}$ thymocytes expressing *DDX3X-MLLT10* or *PICALM-MLLT10*. Each data point represents one animal, and the median of each cohort is represented by a horizontal bar. Abbreviations: BM, cells harvested from murine bone marrow; Pre, immunophenotype of cells pre-injection; THY, cells harvested from murine thymi.
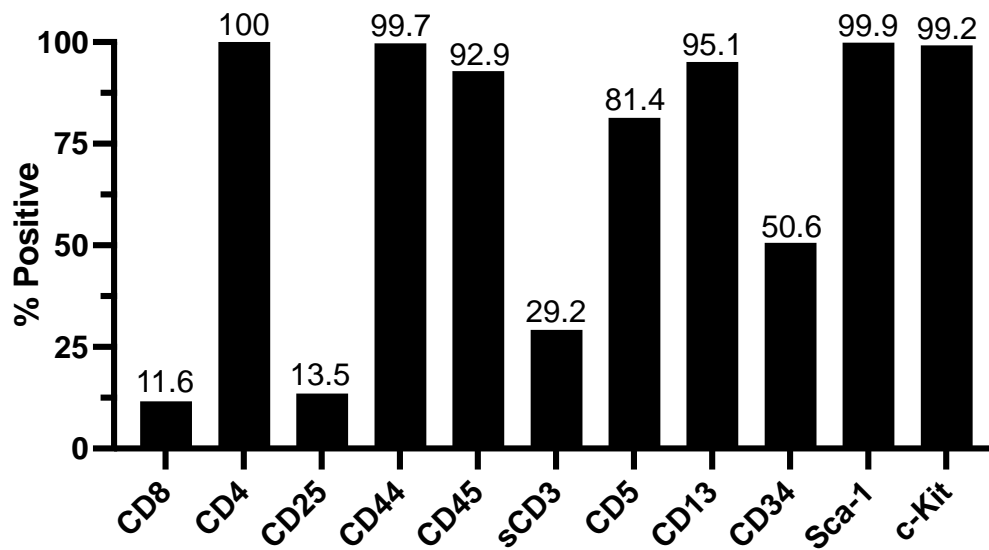
Figure 4.18 – GFP⁺ cells harvested from empty vector mouse EV-5 exhibit a mixed myeloid/T immunophenotype

Cell surface marker staining was analysed on GFP⁺ cells harvested from the peripheral blood of empty vector mouse EV-5 using a BD FACSFortessa™. Gating was set based on isotype controls to determine the proportion of positive cells for each marker (**Supplementary Figure 4.1**). Percentage of positive cells for each marker are annotated above each bar.

4.3.4 GFP-negative T-lineage malignancies developed in three control empty vector mice

Three control mice (EV-1, 2 and 3) developed clinical signs of illness at 20-22 weeks of age and were humanely killed on days 86, 110 and 117 post-transplantation. Signs of leukaemic infiltration were present in haematopoietic organs, including an enlarged thymus, liver and spleen, and pale bone marrow indicative of anaemia (**Supplementary Figure 4.1A-C**). Flow cytometry revealed FSC/SSC properties indicative of lymphocyte populations in the peripheral blood, spleen, liver, thymus and bone marrow (**Supplementary Figure 4.1A-C**). However, no evidence of GFP$^+$ cell engraftment was identified in any haematopoietic site by flow cytometry (**Figure 4.7**). A complete blood count (CBC) was performed for one mouse, revealing an extremely high lymphocyte count of 236 K/µL, 98-fold greater than the upper limit of normal for NSG mice (2.6 K/µL) (**Figure 4.9** & **Supplementary Table 4.1**). H&E staining revealed dense infiltration of leukaemic blasts, and GFP staining was negative (**Figure 4.19**). Immunophenotyping was relatively consistent between the three affected mice, demonstrating expression of T-lineage markers CD8, CD4 and CD5, and stem cell marker CD34 (**Figure 4.20**).
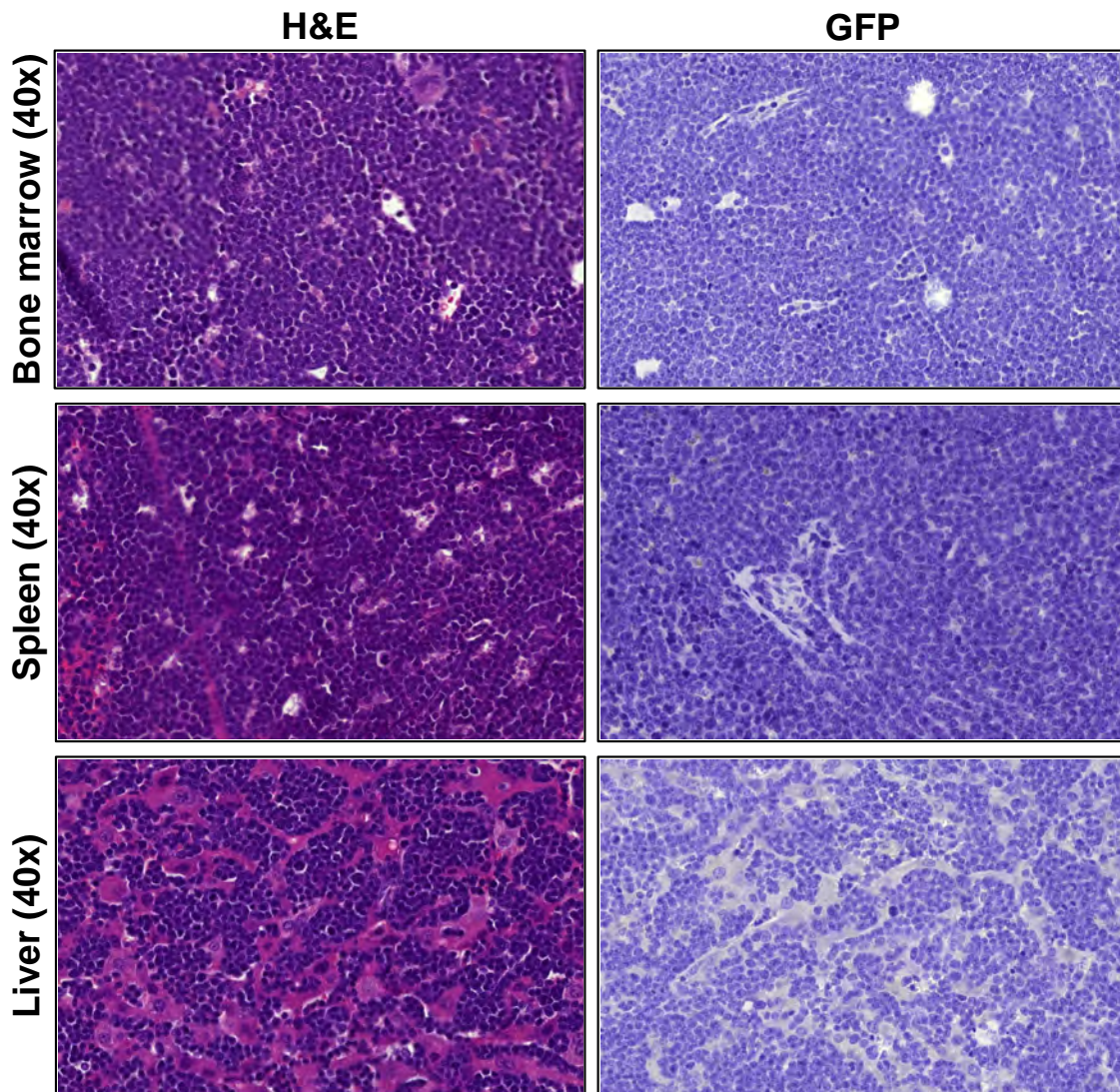
Figure 4.19 – Three empty vector mice exhibit histological evidence of a GFP-negative haematopoietic malignancy

Representative H&E and GFP staining of sections from an empty vector mouse who developed a GFP-negative T-cell malignancy. Images analysed at 40x magnification with CaseViewer software.

(A)
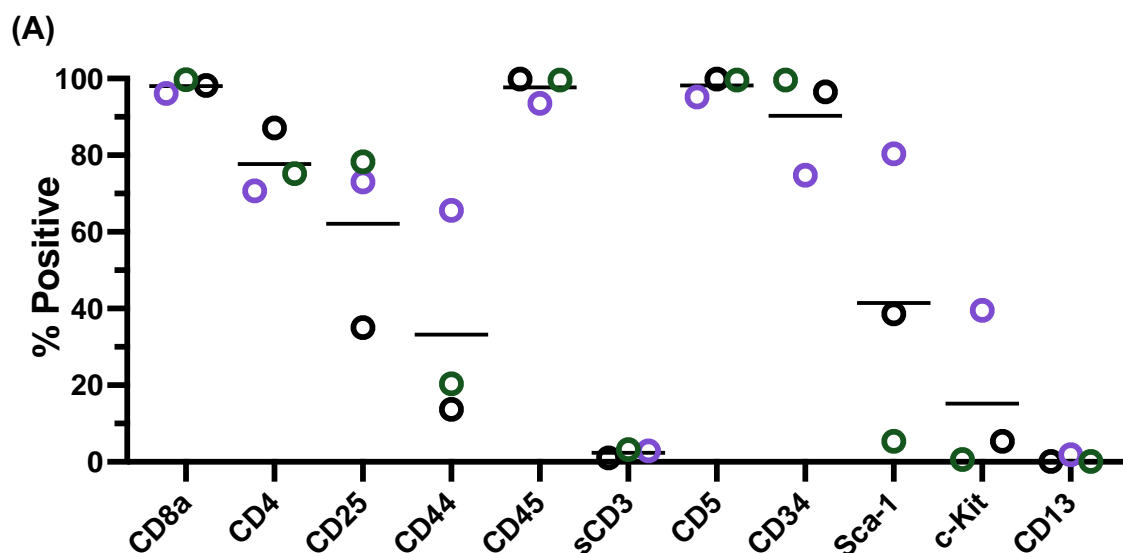
(B)

Marker	EV-1	EV-2	EV-3	Published phenotype[12]
CD8	+	+	+	+ or -
CD4	+	+	+	+
CD25	Heterogenous[x]	+	Heterogenous	+ or -
CD44	-	Dim	Heterogenous	-
CD45	+	+	+	+
sCD3	-	-	-	ND
CD5	+	+	+	+ or -
CD34	+	+	+	ND
Sca-1	Dim*	+	-	ND
c-Kit	-	Dim	-	-
CD13	-	-	-	ND

Figure 4.20 – Immunophenotype of white blood cells harvested from control mice is consistent with T-cell leukaemia/lymphoma

(A) Cell surface marker staining was analysed on all live single cells from the livers of empty vector mice EV-1, 2 and 3, using a BD FACSFortessa[TM]. Gating was set based on isotype controls, to determine percentage of positive cells for each marker. Each data point represents one animal, and the median is represented by a horizontal bar. **(B)** A table summarising the immunophenotype of the three affected mice, compared with a published case series of sporadic T-lineage leukaemias/lymphomas in NSG mice[12]. [x]Heterogenous expression is defined when there are two clearly discernible populations of positive and negatively stained cells. *Dim expression is defined when

a population of cells has weak expression of a marker, but the population is not clearly distinguished from the negative population (**Supplementary Figure 4.1**). Abbreviations: ND, no data.

4.4 Discussion

In this chapter, Arf$^{-/-}$ thymocytes expressing *DDX3X-MLLT10, PICALM-MLLT10* or control empty vector were transplanted into NSG mice to explore the oncogenic potential of *MLLT10*r *in vivo*. Haematopoietic material was harvested from mice and interrogated by surface marker staining, immunohistochemistry (IHC) and qRT-PCR.

4.4.1 Expression of *MLLT10*r does not change *HOX* gene expression in Arf$^{-/-}$ thymocytes

No statistically significant difference in *HOX* expression was identified between cohorts for any assessed gene *in vitro* (**Figure 4.4**) or *ex vivo* (**Figure 4.12**), indicating that *MLLT10*r expression does not induce upregulation of *HOX* cluster genes in Arf$^{-/-}$ thymocytes. The lack of *HOX* gene upregulation has important implications for the validity of this model system, as *HOX* upregulation is a well-established characteristic of *MLLT10*r leukaemia[13-17]. Expression of *HOXA1, 3, 5, 10* and *MEIS1* decreased relative to pre-injection cells, whereas *HOXA9* and *HOXB1* expression levels remained constant (**Supplementary Table 4.5**). *HOX* expression was variable between mice, likely due to the differences in proportion of GFP$^+$ cells present in each sample. The percentage of GFP$^+$ cells in murine livers varied between 50-99.6%, but the quantity of cells harvested from organs with low GFP expression was too small for FACS collection of a pure GFP$^+$ population. It is therefore difficult to accurately compare *HOX* expression between mice and pre-injection values, as is likely that these results are skewed by the varying proportion of GFP-negative endogenous murine cells in each sample.

All *HOXA* genes except *HOXA7* were expressed, indicating a relatively immature parent cell differentiation stage. It is possible that since most *HOXA* genes are sufficiently expressed in Arf$^{-/-}$ thymocytes, a further increase in expression is not necessary to induce leukaemogenesis *in vivo*. In both *MLLT10*r and empty vector thymocytes, *HOXA1, HOXA3, HOXA5* and *MEIS1* have higher expression (ΔCt values of <10) than *HOXA7-10* (ΔCt values of 10-15). In chapter 3, *HOXA1, HOXA3, HOXA5* and *MEIS1* expression was significantly increased in *KMT2A-AFF1* MOHITO cells relative to empty vector MOHITO cells (**Figure 3.8B**), whereas *HOXA7-10* expression did not change. *HOXA3* and *HOXA5* are upregulated in leukaemic blasts from patients with *KMT2A*r or

*MLLT10*r T-ALL[14], and also in two published models of murine *PICALM-MLLT10* AML[18,19]. Taken together, these data support the suggestion that *HOXA1, HOXA3* and *HOXA5* play important lineage-specific roles in myeloid and T-lineage haematopoiesis and malignancies.

4.4.2 Murine-derived CD4$^-$CD8$^-$TER119$^-$ thymocytes retain a T-cell immunophenotype but regain CD8 and CD4 expression in culture

Surface marker staining of Arf$^{-/-}$ thymocytes prior to transplantation into mice confirmed an immunophenotype consistent with T-ALL (**Figure 4.5**). Cells were >90% CD8$^+$ and 32-50% CD4$^+$ prior to injection, despite the initial collection of CD4$^-$CD8$^-$ (PE-negative) cells (**Figure 4.1**). It is likely that cells were initially CD4$^-$CD8$^-$, and gradually differentiated over the 37 days spent in culture. This differentiation is supported by the OP9-DL1 co-culture protocol published by Holmes *et al.*[8], where it was observed that cultured murine bone marrow-derived CD44$^-$CD25$^-$CD4$^-$CD8$^-$ HSCs differentiated over time, with 39% CD4$^+$CD8$^+$ cells present after 16 days in culture. The cells used in this study were cultured for 37 days due to the time required to transduce and expand cell populations. The high percentage of further differentiated, but still primitive, CD4$^+$CD8$^+$ T-cells is therefore expected after extended time in culture. Holmes *et al.* also demonstrate that differentiation kinetics are accelerated in lymphocyte progenitors, in contrast to less mature HSCs. As CD4$^-$CD8$^-$ cells were collected, it is likely that a mixed population of CD4$^-$CD8$^-$CD44$^{+/-}$CD25$^{+/-}$ progenitors were present in culture, resulting in more accelerated differentiation than if a pure population of CD44$^-$CD25$^-$CD4$^-$CD8$^-$ cells were collected.

It was hypothesised that transplantation of the mixed CD4 and CD8 expressing population would result in competitive outgrowth of the cell differentiation stage that is most susceptible to leukaemic transformation by *MLLT10*r. Immunophenotyping on harvested material from engrafted mice would determine whether CD4 and CD8 positive or negative cells induced disease. Sorting of CD4$^-$CD8$^-$ cells was therefore not performed prior to transplantation, and the bulk population of cells were transplanted into sub lethally irradiated NSG mice.

4.4.3 Expression of *PICALM-MLLT10* induces leukaemia with widespread organ penetrance

All *PICALM-MLLT10* mice exhibited >90% GFP+ live single cells in all organs analysed, indicating a malignancy with aggressive organ penetrance (**Figure 4.7**). PICALM-5 also presented with enlarged, pale kidneys with white patches, which was not observed in any other cohort. IHC revealed infiltration of GFP+ cells in the kidney tissue of PICALM-5, but not PICALM-3 (**Supplementary Figure 4.4**), demonstrating this was an additional site of leukaemic involvement in PICALM-5 only, indicative of a particularly aggressive malignancy in this mouse. Renal involvement of ALL is a rare occurrence that typically presents as bilateral kidney enlargement and acute renal failure[20-23]. Enlarged kidneys at diagnosis was historically associated with poor outcomes in childhood ALL[24,25], but it is not clear whether this remains the case with modern intensified treatment protocols.

Mice engrafted with cells expressing *DDX3X-MLLT10* exhibited engraftment of predominantly either the thymus or bone marrow (**Figure 4.10** and **Appendix 4.2**), indicating a reduced ability of *DDX3X-MLLT10* expressing cells to penetrate secondary organs in comparison to *PICALM-MLLT10* cells. Thymus size did not correlate with percentage of GFP+ cells or onset of disease. DDX3X-4 had a small thymus that yielded very few cells, of which 72.6% were GFP+, whereas DDX3X-5 had a comparatively larger thymus that was only 41.0% GFP+. The low proportion of GFP+ in DDX3X-5 may be the result of non-malignant cells involved in the thymus structure and microenvironment, required to support the organ's large size.

It is not clear why *DDX3X-MLLT10* expressing cells exhibited variable homing patterns. It is possible that different clones emerged in each mouse, potentially driven by the acquisition of distinct cooperative alterations. This could be further explored in cells harvested from individual mice by mRNA-seq. Alternatively, there may be a fusion-specific functional mechanism underlying this phenomenon, where the 5' fusion partner, *DDX3X* or *PICALM*, may promote cell homing and dissemination by distinct mechanisms.

It was expected that transplanted thymocytes would initially home to the thymus. NSG mice undergo thymic atrophy, where a thymus is typically only visible until approximately 12 weeks of age[26]. Thymic atrophy also normally occurs in humans,

where functional lymphopoietic tissue is replaced by adipose tissue over time as a normal part of the ageing process[27]. Despite thymic atrophy, NSGs are capable of supporting thymopoiesis in response to engraftment of human or murine thymocytes[26]. This is supported by the presence of GFP+ cells in the thymi of all mice engrafted with fusion-expressing cells, despite some thymi being small (particularly DDX3X-3 and 4). However, a thymus could not be identified in EV-4, despite >20% GFP+ cells in the liver and spleen but only 3.3% in the bone marrow, suggesting the liver and spleen may be involved in the initial homing of transplanted cells. A bioluminescent imaging system, such as luciferase imaging, would assist in identifying homing patterns of transplanted cells throughout engraftment, particularly establishing whether the thymus is a site of initial leukaemic engraftment, and whether cell homing differs between fusions (discussed further in **4.4.11**).

Despite a high percentage of GFP+ cells in *PICALM-MLLT10* mice, no significant difference in liver or spleen weight (**Figure 4.8**) or blood profile (**Figure 4.9**) was observed between cohorts. Abnormal blood results were observed across all three cohorts, with high WCC indicative of leukaemic involvement. RBC count and Hb were consistently low relative to NSG reference values, indicating anaemia potentially induced by RBC depletion due to leukaemic infiltration.

4.4.4 Two empty vector mice engrafted with GFP+ disease

Empty vector mice EV-4 and 5 both demonstrated evidence of GFP+ engraftment by flow cytometry (**Appendix 4.1D-E**) and IHC (**Figure 4.11**). EV-5 was humanely killed due to clinical signs of illness on day 138, whereas EV-4 was humanely killed at the experimental endpoint on day 234, based on absence of GFP+ cells in peripheral blood 1 week prior to termination. It is likely that EV-4 would have eventually developed clinical signs of leukaemia, as 58.1% and 20.5% GFP+ cells were evident in the liver and spleen respectively (**Supplementary Table 4.2**).

As analysis of only two GFP+ empty vector mice was possible, who died at vastly different time points, it is not clear whether control Arf[-/-] thymocytes engraft with a similar latency and localisation in comparison to cells expressing *MLLT10*r.

4.4.5 An unusual mixed/ETP phenotype developed in one empty vector mouse

Curiously, EV-5 was the only mouse where high expression of the myeloid marker CD13 occurred (95% CD13$^+$, **Figure 4.18**), indicating a mixed myeloid/T-ALL phenotype. CD13 is a myeloid antigen identified in >50% of ETP-ALL cases[28,29]. As this experiment involved the transduction of a pool of CD4$^-$CD8$^-$ precursor cells, it is possible that expansion of a primitive precursor with myeloid differentiation potential occurred. GFP$^+$ cells from this mouse also exhibited expression of mature T-cell marker sCD3 and stem cell markers Sca-1 and c-Kit, and heterogenous CD34 and CD2 expression, which is different from the immunophenotype observed in the *MLLT10*r cohorts. However, this finding is difficult to interpret as only one mouse was affected, and further replicates are required to determine whether this unusual phenotype is consistently recapitulated by empty vector Arf$^{-/-}$ thymocytes, or if it was an isolated incident.

4.4.6 Cells expressing *PICALM-MLLT10* or *DDX3X-MLLT10* are CD4$^+$CD8$^+$ following transplantation

Cells harvested from *MLLT10*r mice were consistently CD4$^+$CD8$^+$. The proportion of CD8 and CD4 positive cells pre- and post-injection is summarised in **Figure 4.5**. CD8 expression was maintained *in vivo*, whereas the proportion of CD4$^+$ cells increased in the bone marrow and thymus. The CD4$^+$CD8$^+$ immunophenotype was specific to fusion-expressing cells, whereas GFP$^+$ cells from the peripheral blood of EV-5 were CD4$^+$CD8$^-$ (**Figure 4.18**). CD4 and CD8 expression is variable in human T-ALL, positive in 53% and 60% of cases respectively[6], and CD4$^+$CD8$^+$ double-positive disease is identified in approximately 40% of cases[30,31], including in *PICALM-MLLT10* T-ALL[32]. As CD4 and CD8 expression are variable within human T-ALL, additional factors external of fusion expression are likely involved in determining T-ALL immunophenotype.

4.4.7 Cells expressing *PICALM-MLLT10* or *DDX3X-MLLT10* exhibit aberrant stem cell marker expression

The stem cell markers CD34, Sca-1 and c-Kit were aberrantly expressed in cells expressing *DDX3X-MLLT10* and *PICALM-MLLT10*. Differences in stem cell marker expression are summarised in **Figure 4.5**. *DDX3X-MLLT10* cells harvested from thymi exhibited a higher percentage of Sca-1$^+$ cells compared to *PICALM-MLLT10*, suggesting a potential role for this murine stem cell marker in *DDX3X-MLLT10* T-ALL pathogenesis. Sca-1 expression correlates with more aggressive murine leukaemia, quantified by a shorter time to engraftment than Sca-1$^{low/-}$ disease in one study[33], but no significant difference in time to engraftment was observed here. It is possible that Sca-1$^+$ precursors may represent a primitive cell type that permits increased growth within the thymus but not extra-thymic dissemination of engrafted cells. The clinical relevance of Sca-1 expression in human leukaemia is unclear, as the human equivalent of Sca-1 is unknown.

There were trends toward increased expression of c-Kit in *DDX3X-MLLT10* cells and CD34 in *PICALM-MLLT10* cells. CD34 is the most frequently identified stem cell marker expressed on ALL blasts, present in >50% of T-ALL cases[34,35], including 75% of ETP-ALL, 83% of pro-T, and 20-50% of pre-T, cortical and mature T-ALL cases[6]. c-Kit is expressed in approximately 10-20% of T-ALLs overall[6,36], predominantly in immature ETP and pro-T cases (50% and 67% of new diagnoses respectively[6]). Overall, these results indicate that stem cell marker expression is variable in *MLLT10*r pathogenesis. Further comprehensive immunophenotyping of *MLLT10*r patient cohorts would provide further insight into the functional significance of stem cell marker expression in *MLLT10*r T-ALL.

4.4.8 Changes in expression of T-cell markers CD25, CD44 and CD5 occurred in *MLLT10*r cells

Expression of CD25, CD44 and CD5 in *DDX3X-MLLT10* and *PICALM-MLLT10* cells are summarised in **Figure 4.5**. One mouse from either fusion cohort exhibited a CD25[+] (PICALM-1) or CD44[-] (DDX3X-4 and PICALM-2) immunophenotype, demonstrating variable expression of these markers. CD25 and CD44 are immature T-cell markers that define the stage of double negative (DN) thymocyte differentiation. CD25 is expressed in 13% of AML[37] and 43-54% of adult B-ALL diagnoses[38,39] and is disproportionately associated with *BCR-ABL1* positive B-ALL[39], but incidence in T-ALL is not known. Investigating the expression of CD44 and CD25 in T-ALL patients would establish whether these markers have an association with genomic subtype or outcomes.

CD5 expression was notably greater in *DDX3X-MLLT10* cells pre-injection (median % CD5[+] cells = 83.8% vs 54.7%, *DDX3X-MLLT10* vs *PICALM-MLLT10*), suggesting that *PICALM-MLLT10* is capable of inducing a less mature immunophenotype indicative of ETP-ALL[40]. However, engrafted cells were consistently CD5[+] in the bone marrow and thymus, suggesting that lack of CD5 expression is not an important contributor to leukaemic engraftment in this context.

4.4.9 Likely familial T-lineage malignancies developed in three control empty vector mice

Three control mice (EV-1, 2 and 3) developed clinical signs of illness at 20-22 weeks of age and were humanely killed on days 86, 110 and 117 post-injection. CBC analysis (**Figure 4.9**), IHC (**Figure 4.19**) and flow cytometry (**Figure 4.20**) revealed a GFP-negative T-lymphoid malignancy that infiltrated the liver, spleen, thymus, bone marrow and peripheral blood.

As all three affected mice were littermates, the occurrence of a familial genetic abnormality is possible. One published case series describes sporadic haematopoietic malignancies in 12 NSG mice aged 20 to 38 weeks[12], that presented with grossly enlarged spleen, liver, thymus, bone marrow and/or lymph nodes. Immunophenotyping revealed that all mice had acquired T-cell lymphoblastic

leukaemia/lymphomas (T-LBL). Consistent with this evidence, the immunophenotype of the three mice affected in this experiment are indicative of a T-cell malignancy (**Figure 4.20**).

It is not clear why sporadic lymphomas/leukaemias develop in NSG mice. Tillman *et al.* propose that malignant lymphocytosis may evolve as an alternative maturation mechanism in mice where normal physiologic processes are perturbed. They also report concurrent viral infections in some of the affected mice, which may serve as a trigger for T-cell malignancies[12]. As the three affected mice were housed in the same cage, it is possible that a viral infection was sustained. The mice developed ALL within a close time frame (within 31 days of each other), potentially suggesting the involvement of a contagious entity. However, considering that NSGs are severely immunocompromised and therefore highly susceptible to infection, it is unlikely that the other two mice housed in the same cage would remain unaffected. It is therefore unlikely that viral infection was the solitary cause of these malignancies, but the combination of a genetic predisposition and a viral infection cannot be excluded as potential causes.

Ultimately, it is likely that these three mice developed leukaemia/lymphoma with a familial genetic basis. The development of an aggressive disease in short succession in three littermates, but not the other two unrelated (not littermate) mice who were housed in the same cage and transplanted with the same empty vector cells, is strong evidence to suggest a familial basis for this phenomenon.

4.4.10 One *DDX3X-MLLT10* mouse died of likely unrelated causes

DDX3X-1 is the only mouse in either of the fusion cohorts who did not develop leukaemia. This mouse exhibited intermittent eye redness and swelling for three months, but no other signs of clinical deterioration were evident until 2 weeks prior to termination, when gradual weight loss and declining activity occurred. Some GFP⁺ cell engraftment was noted in the liver and spleen (LIV 16.2% and SPL 12.1%, **Supplementary Table 4.2**), but organs were of normal size and morphology (**Appendix 4.2A**). RT-PCR of RNA harvested from the liver failed to detect the *DDX3X-MLLT10* breakpoint (**Figure 4.13**). The relatively low proportion of GFP⁺ cell engraftment is sufficient to establish that systemic leukaemia was not the cause of

illness. The intermittent eye swelling observed raised the possibility of ocular infiltration of transplanted cells. Optic nerve infiltration is a rarely reported occurrence in ALL[41,42], but IHC of the optic nerve from both the affected and unaffected eye did not reveal any evidence of leukaemic infiltration (**Supplementary Figure 4.5**). Unlike the three GFP⁻ EV mice who developed a GFP-negative T-cell malignancy, this mouse did not exhibit any other signs of haematopoietic illness. A normal blood profile (**Figure 4.8**) and normal organ morphology (**Appendix 4.2A**) were observed. IHC reveals absence of GFP⁺ lymphoblasts, except for scant staining in the spleen (**Supplementary Figure 4.6**). The cause of deterioration in this mouse is therefore unknown, and while no evidence of a systemic infection was identified (such as elevated WCC), the ongoing eye problems indicate a potential eye infection that may have induced systemic deterioration in this immunocompromised mouse.

4.4.11 Limitations and future directions

Some limitations and technical difficulties were encountered in developing this model. For further studies, the following changes would be implemented.

1. Addressing the issue of CD4⁺CD8⁺ differentiation in culture

In future experiments, optimisation of transduction protocols to limit the time that cells spend in culture (ideally to <2 weeks) to minimise CD4⁺CD8⁺ differentiation is warranted. Immature cells could also be re-selected immediately prior to injection, to enrich the population of immature lymphoid precursors. This methodology was utilised in one published study, where CD4⁻CD8⁻ thymocytes were collected, and then CD44⁻CD25⁻CD4⁻CD8⁻ cells were re-selected by FACS immediately prior to injection into mice[7].

In the present experiment, cells were passaged and media replenished every 3 days, with one replenishment after 6 days at 22 days post-transduction (due to external limitations). The 6 day incubation without media change may have affected cellular differentiation, as low IL-7 concentration accelerates CD4⁺CD8⁺ differentiation of OP9-DL1 cultured cells[8]. More frequent and regular media replenishment may delay cell differentiation, by ensuring exposure to consistent quantities of IL-7.

2. Introducing a bioluminescence imaging system to track cell homing and engraftment in organs

Leukaemic engraftment was monitored in mice by fortnightly tail vein bleeds. Mice exhibited very low levels of circulating GFP$^+$ cells in the peripheral blood throughout the experiment. In many cases, <1% circulating GFP$^+$ cells were identified by tail vein bleed and flow cytometric analysis 1 week prior to a mouse being culled due to clinical deterioration. Most mice exhibited >60% GFP$^+$ cells in the peripheral blood at the time of culling, suggesting that circulation of GFP$^+$ cells in the peripheral blood occurred at a late stage in disease development. As a result, tail vein bleeds were not an informative means of monitoring disease progression throughout the experiment. Introducing a bioluminescent imaging system, such as a luciferase vector, would enable live cell imaging to track cell homing and engraftment within organs such as the bone marrow, thymus, liver and spleen. This would allow tracking of overall disease burden within each mouse and provide an understanding of the homing patterns of transplanted cells.

3. Understanding the engraftment of empty vector expressing cells

In this experiment, 3/5 empty vector control mice were affected by a likely familial T-cell malignancy, a rare but recognised phenomenon in NSG mice. The remaining two mice developed GFP$^+$ cell engraftment, but at vastly different time points (138 and 234 days post-injection). Consequently, it is not clear whether control Arf$^{-/-}$ thymocytes consistently engraft with a similar latency and immunophenotype of *MLLT10*r expressing cells. It is therefore necessary to repeat the experiment, to confirm that the GFP-negative disease developed by three control mice was indeed a sporadic event, and to clarify the engraftment potential and immunophenotype of empty vector-expressing Arf$^{-/-}$ thymocytes.

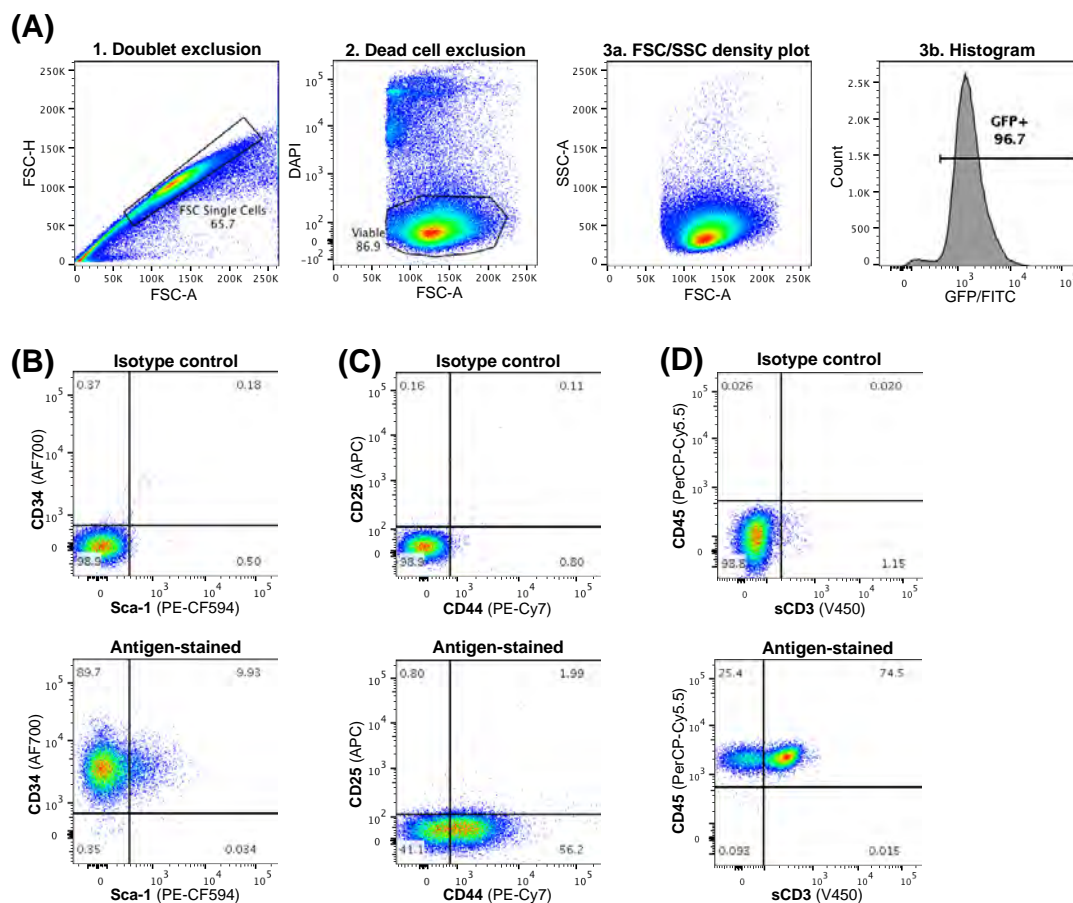## 4. Generating a secondary xenograft from harvested cells

Generating a secondary xenograft mouse model using GFP$^+$ cells harvested from primary xenografts would provide further insight into the role of *MLLT10*r in T-ALL

development and clonal selection[43]. Harvested cells expressing *DDX3X-MLLT10,* *PICALM-MLLT10* or empty vector would be transplanted into NSG mice to assess time to engraftment, organ penetrance and immunophenotype. mRNA-seq would also provide insight into clonal selection by characterising the landscape of acquired secondary genomic alterations in engrafted cells.

4.4.12 Summary of chapter findings

This chapter describes, for the first time, an *in vivo* model of *MLLT10*r T-ALL. Mice transplanted with Arf$^{-/-}$ thymocytes expressing *PICALM-MLLT10* developed a widespread haematopoietic malignancy, affecting the thymus, bone marrow, liver, spleen and peripheral blood. In contrast, *DDX3X-MLLT10* expressing cells induced disease that resided predominantly in either the bone marrow or thymus. Overall, this indicates that while no significant difference in survival were observed, *PICALM-MLLT10* expression induces a comparatively aggressive T-ALL, characterised by widespread disease dissemination, and a trend towards increased CD34 expression. The results in this chapter demonstrate that *MLLT10*r fusion gene retroviral constructs are able to be stably expressed in murine Arf$^{-/-}$ thymocytes, that induce leukaemia when transplanted into NSG mice. This represents a promising system to investigate the biology of genomic aberrations associated with T-ALL, and pre-clinically test targeted therapies.
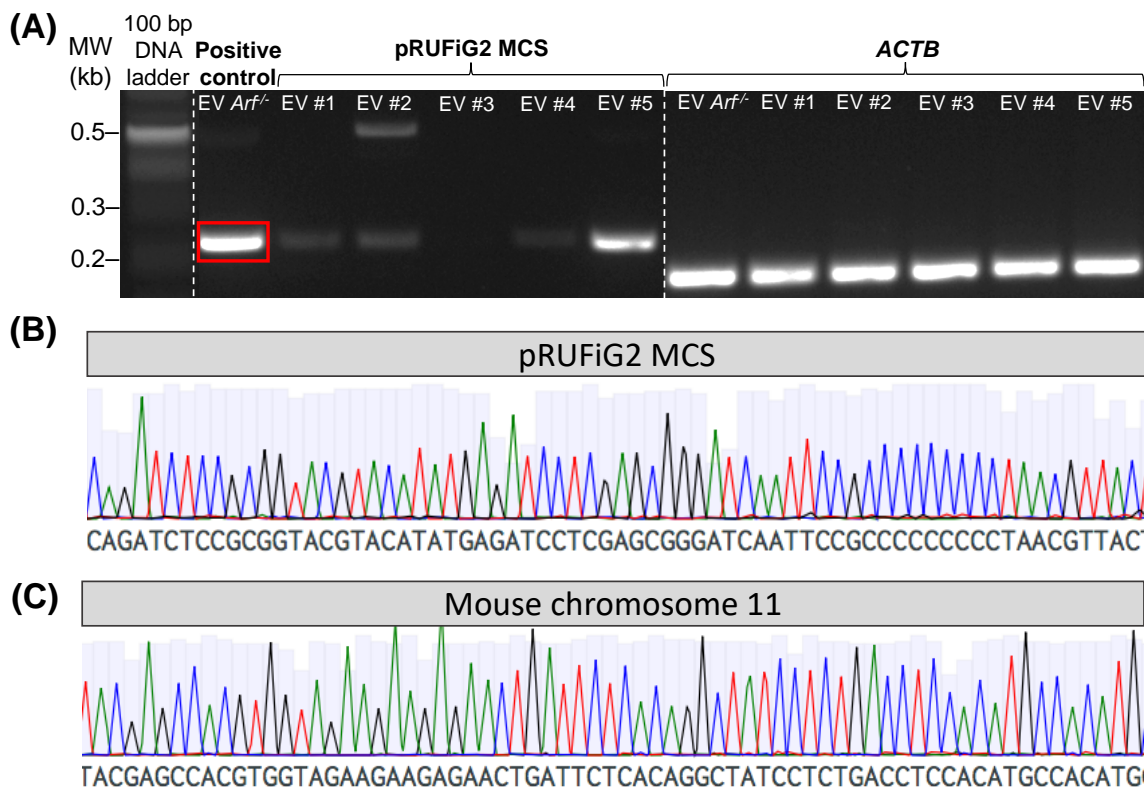
4.5 Supplementary Figures

(A)
1. Doublet exclusion
2. Dead cell exclusion
3a. FSC/SSC density plot
3b. Histogram

(B) Isotype control / Antigen-stained — CD34 (AF700) vs Sca-1 (PE-CF594)
(C) Isotype control / Antigen-stained — CD25 (APC) vs CD44 (PE-Cy7)
(D) Isotype control / Antigen-stained — CD45 (PerCP-Cy5.5) vs sCD3 (V450)

Supplementary Figure 4.1 – Flow cytometry gating strategies

(A) (1) Singlets were isolated on FSC-H vs FSC-A density plot and **(2)** live cells isolated by excluding cells positive for the dead cell stain DAPI. **(3a)** Granularity and size of live single cells was assessed on SSC-A vs FSC-A density plots and **(3b)** proportion of GFP$^+$ cells within the total live cell population was determined. **(B-C)** Immunophenotyping was performed by antibody staining for surface markers. Gating on live GFP$^+$ cells was first performed as described in **(A)**, except for EV mice 1-3, where all live cells were included in analysis. Negative gates were set for each antibody analysed based on isotype control-stained samples. A population was considered positive for a marker if ≥80% of cells stained positive (CD34, panel **B**), and cells were negative for a marker if <10% of cells stained positive (Sca-1, panel **B** and CD25, panel **C**). **(C)** Samples were categorised as having dim expression if 10-79.9% of cells were positive for an antigen, but the positive and negative populations were not clearly discernible (CD44). **(D)** Samples were categorised as having heterogenous

expression if 10-79.9% of cells were positive for an antigen, and the positive and negative populations were clearly discernible (sCD3).

(A)

(B) pRUFiG2 MCS

CAGATCTCCGCGGTACGTACATATGAGATCCTCGAGCGGGATCAATTCCGCCCCCCCCCTAACGTTACT

(C) Mouse chromosome 11

TACGAGCCACGTGGTAGAAGAAGAGAACTGATTCTCACAGGCTATCCTCTGACCTCCACATGCCACATG

Supplementary Figure 4.2 – Validation of the pRUFiG2 multiple cloning site from empty vector expressing Arf$^{-/-}$ cells

(A) Gel electrophoresis of RT-PCR products from empty vector mice. Liver RNA was extracted from each mouse, and cDNA synthesised for PCR of the empty vector MCS. Red rectangles highlight the 245 bp fragment corresponding to the size of the MCS. *ACTB* was used as a positive control for PCR. **(B)** The 245 bp fragment corresponding to the empty vector MCS was purified and Sanger sequenced to confirm the intact MCS. **(C)** The ~500 bp fragment amplified in EV-2 was purified and Sanger sequenced, revealing 100% homology with a region of mouse chromosome 11. Abbreviations: *ACTB*, actin; MCS, multiple cloning site; MW, molecular weight marker.

Without primary antibody   **With primary antibody**
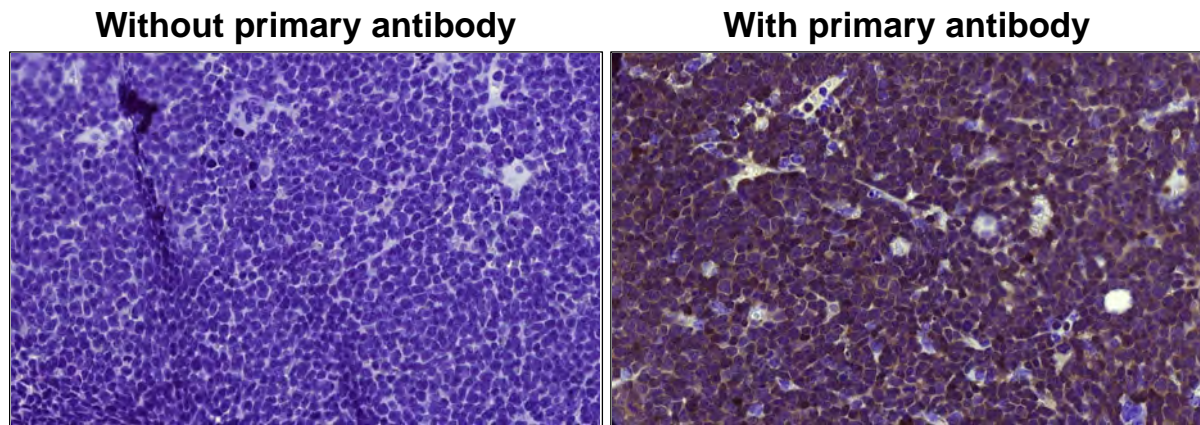
Supplementary Figure 4.3 – Representative no primary GFP antibody control for IHC of formalin-fixed mouse tissue

Murine tissues were fixed in 10% formalin, prior to paraffin-embedding and sectioning for immunohistochemical staining with a GFP primary monoclonal antibody. For each stained section, a control section was performed alongside, which was treated with the antigen retrieval and secondary antibody agents, but not stained with the primary GFP antibody. This control confirms the absence of background GFP staining. A representative section of murine bone marrow is provided to demonstrate the difference between the non-primary antibody stained (left) and primary antibody stained (right) samples.
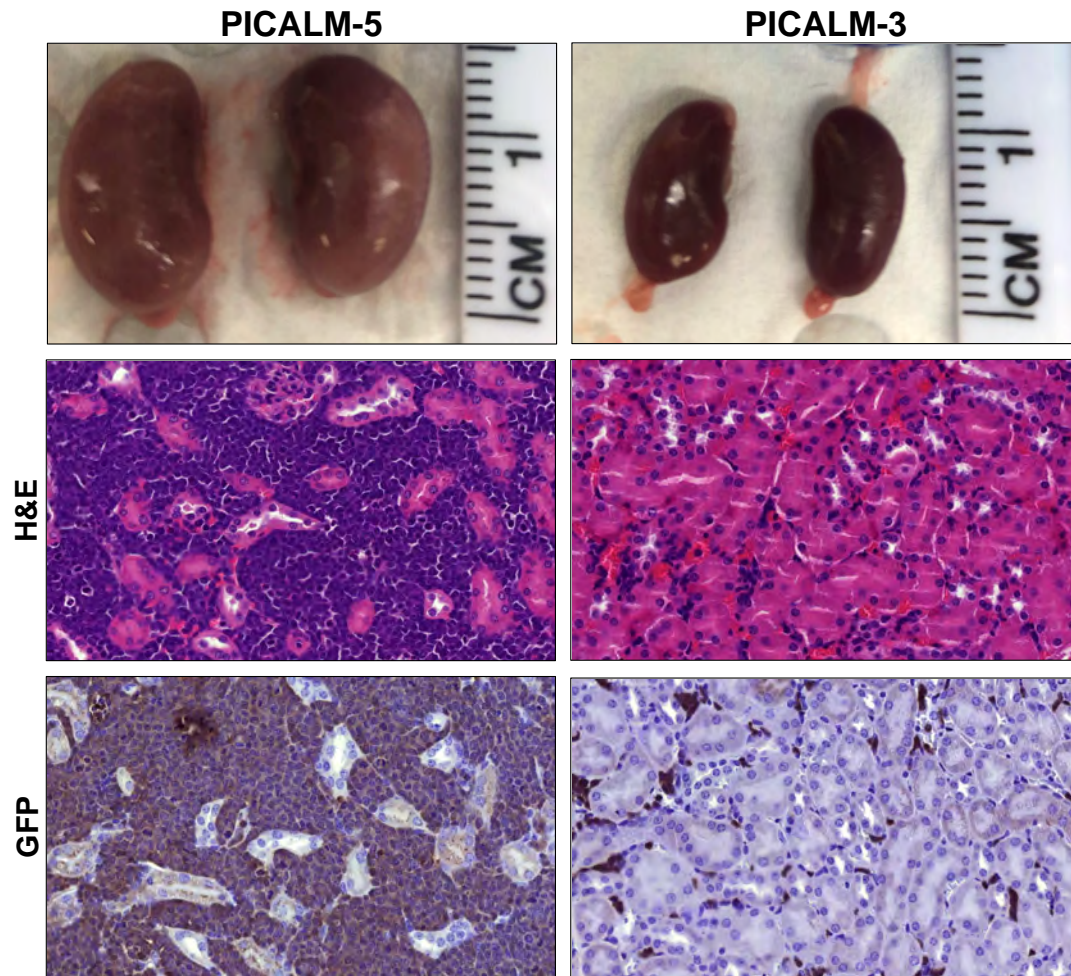
Supplementary Figure 4.4 – Evidence of leukaemic engraftment is present in the kidneys of PICALM-5

Photographs and IHC staining of kidney sections from *PICALM-MLLT10* mice PICALM-5 (left) and PICALM-3 (right). Images analysed at 40x magnification with CaseViewer software.

	Affected eye	Unaffected eye
H&E		
GFP		
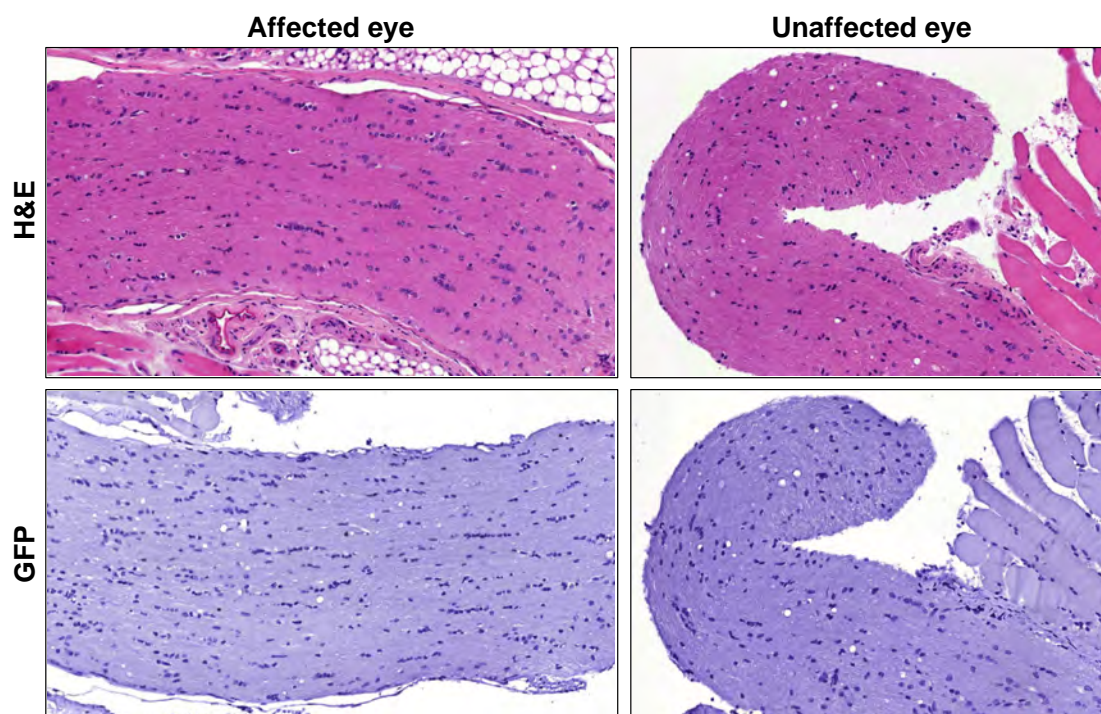
Supplementary Figure 4.5 – H&E and GFP staining of eye sections harvested from DDX3X-1 do not reveal evidence of leukaemic engraftment

H&E and GFP staining of optic nerve sections from DDX3X-1. Images analysed at 20x magnification with CaseViewer software.
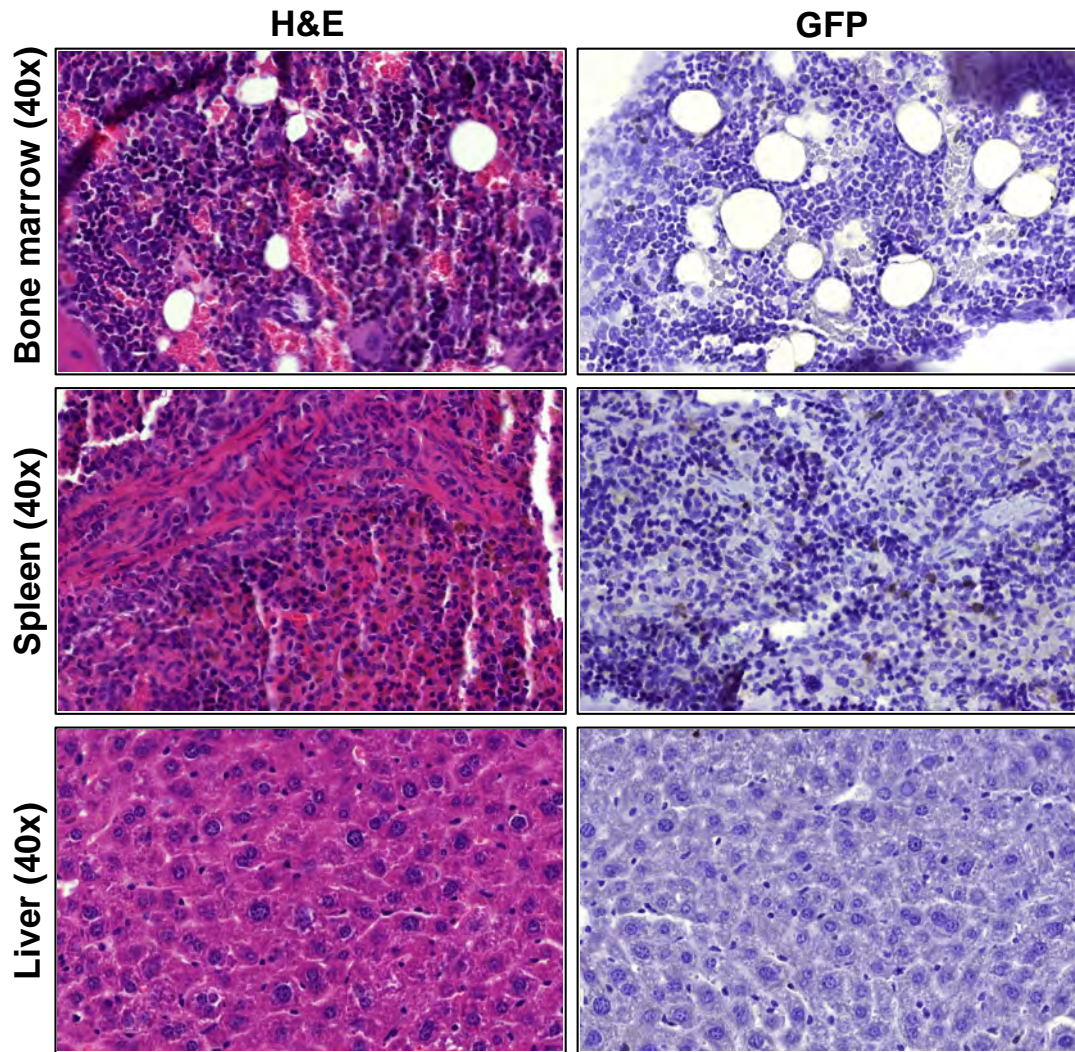
Supplementary Figure 4.6 – One mouse transplanted with *DDX3X-MLLT10* cells did not demonstrate histological evidence of leukaemic engraftment

H&E and GFP staining of organ sections from DDX3X-1. Images analysed at 40x magnification with CaseViewer software.

Supplementary Table 4.1 – CBC values for mice on the day of organ harvest

	Empty vector	*DDX3X-MLLT10*	*PICALM-MLLT10*	Reference ranges (LLN - ULN)
Total white cell count (WCC) K/µL	EV-1: ND EV-2: ND EV-3: 294.8 EV-4: 4.5 EV-5: 91.6	DDX3X-1: 0.6 DDX3X-2: 36.8 DDX3X-3: 76.4 DDX3X-4: 75.0 DDX3X-5: 91.9	PICALM-1: 44.5 PICALM-2: 25.2 PICALM-3: 28.2 PICALM-4: 97.4 PICALM-5: 158.0	0.5 - 2.6
Median (range)	**91.6 (4.5-294.8)**	**75.0 (0.6-91.9)**	**44.5 (25.2-158.0)**	
Total neutrophil count (NE) K/µL	1: ND 2: ND 3: 44.4 4: 2.7 5: 55.9	1: 1.2 2: 15.8 3: 43.1 4: 42.8 5: 31.4	1: 18.1 2: 18.9 3: 16.6 4: 31.7 5: 16.6	0.4 - 2.2
Median (range)	**44.4 (2.7-55.9)**	**31.4 (1.2-43.1)**	**18.1 (16.5-31.7)**	
Total lymphocyte count (LY) K/µL	1: ND 2: ND 3: 236.0 4: 1.0 5: 8.2	1: 0.6 2: 15.8 3: 7.2 4: 8.4 5: 52.7	1: 18.7 2: 0.8 3: 4.2 4: 43.6 5: 128.0	0.1 - 0.5
Median (range)	**8.2 (1.0-236.0)**	**8.4 (0.6-52.7)**	**18.7 (0.8-128.0)**	
Total monocyte count (MO) K/µL	1: ND 2: ND 3: 12.3 4: 0.2 5: 6.8	1: 0.1 2: 3.7 3: 2.6 4: 13.1 5: 5.7	1: 2.0 2: 0.0 3: 1.3 4: 14.6 5: 10.4	0.0 - 0.6
Median (range)	**6.8 (0.2-12.3)**	**3.7 (0.1-13.1)**	**2.0 (0.0-14.6)**	
Total eosinophil count (EO) K/µL	1: ND 2: ND 3: 2.1 4: 0.6 5: 15.1	1: 0.2 2: 1.4 3: 16.2 4: 9.3 5: 2.1	1: 4.0 2: 5.7 3: 4.9 4: 6.8 5: 2.2	0.0 - 0.4
Median (range)	**2.1 (0.6-15.1)**	**2.1 (0.2-16.2)**	**4.9 (2.2-6.8)**	
Total basophil count (BA) K/µL	1: ND 2: ND 3: 0.2 4: 0.6 5: 5.8	1: 0.1 2: 0.3 3: 8.5 4: 1.9 5: 0.3	1: 1.3 2: 1.3 3: 1.2 4: 0.9 5: 0.8	Unknown
Median (range)	**0.6 (0.2-5.8)**	**0.3 (0.1-8.5)**	**1.2 (0.8-1.3)**	
Total red cell count (RCC) M/µL	1: ND 2: ND 3: 3.7 4: 5.7 5: 7.5	1: 7.0 2: 5.8 3: 7.7 4: 4.3 5: 3.9	1: 6.6 2: 5.5 3: 5.8 4: 4.6 5: 4.9	7.0 - 9.1

Median (range)	**5.7 (3.7-7.5)**	**5.8 (3.9-7.7)**	**5.5 (4.6-6.6)**	
Haemoglobin (Hb) g/dL	1: ND	1: 8.1	1: 8.7	12.1 - 15.0
	2: ND	2: 8.1	2: 6.6	
	3: 6.0	3: 11.3	3: 7.3	
	4: 7.8	4: 6.3	4: 6.3	
	5: 11.5	5: 5.1	5: 6.0	
Median (range)	**7.8 (6.0-11.5)**	**8.1 (5.1-11.3)**	**6.6 (6.0-8.7)**	
Platelets (PLT) K/µL	1: ND	1: 290	1: ND	782 - 1166
	2: ND	2: 239	2: 225	
	3: 816	3: 513	3: 204	
	4: ND	4: 266	4: 189	
	5: 382	5: 274	5: 522	
Median (range)	**599.0 (382-816)**	**274 (239-513)**	**215 (189-522)**	

Abbreviations: g/dL, grams per decilitre; K/µL, thousand cells per microlitre; LLN, lower limit of normal; M/µL, million cells per microlitre; ND, no data; ULN, upper limit of normal

Supplementary Table 4.2 – GFP expression in mouse organs

Organ	% positive		
	Empty vector	***DDX3X-MLLT10***	***PICALM-MLLT10***
PB	EV-1: ND	DDX3X-1: 9.4	PICALM-1: 94.3
	EV-2: 10.8	DDX3X-2: 89.0	PICALM-2: 87.0
	EV-3: 0.1	DDX3X-3: 20.3	PICALM-3: 90.9
	EV-4: 12.1	DDX3X-4: 73.7	PICALM-4: 69.2
	EV-5: 30.3	DDX3X-5: 36.9	PICALM-5: 92.1
Median (range)	**11.5 (0.1-30.3)**	**36.9 (9.4-89.0)**	**88.0 (69.2-94.3)**
BM	1: 0.2	1: 1.0	1: 99.1
	2: 0.2	2: 1.5	2: 97.2
	3: 0.0	3: 61.0	3: 98.2
	4: 3.3	4: 89.3	4: 82.1
	5: 8.7	5: 9.7	5: 98.4
Median (range)	**0.2 (0.0-8.7)**	**9.7 (1.0-89.3)**	**98.2 (82.1-99.1)**
THY	1: 2.2	1: ND	1: 99.4
	2: 0.0	2: 97.0	2: 98.2
	3: 0.1	3: 30.5	3: 98.7
	4: ND	4: 72.6	4: 98.2
	5: 57.3	5: 41.0	5: 97.0
Median (range)	**1.1 (0.0-57.3)**	**37.6 (29.0-97)**	**98.2 (97.0-99.4)**
SPL	1: 2.5	1: 12.1	1: 81.8
	2: 0.0	2: 22.5	2: 81.8
	3: 0.0	3: 87.9	3: 78.1
	4: 20.5	4: 98.9	4: 99.8
	5: 14.7	5: 28.0	5: 94.4
Median (range)	**2.5 (0.0-20.5)**	**28.0 (12.1-98.9)**	**94.4 (78.1-99.8)**
LIV	1: 0.9	1: 16.2	1: 99.6
	2: 0.0	2: 50.4	2: 99.2
	3: 0.0	3: 89.8	3: 99.4
	4: 58.1	4: 96.6	4: 96.8
	5: 29.7	5: 50.9	5: 99.6
Median (range)	**0.9 (0.0-58.1)**	**50.9 (16.2-96.6)**	**99.4 (96.8-99.6)**

Abbreviations: ND, no data

Supplementary Table 4.3 – Immunophenotyping of mouse BM or PB

Marker	% positive		
	Empty vector	***DDX3X-MLLT10***	***PICALM-MLLT10***
CD8 (AF700)	EV-1: ND	DDX3X-1: ND	PICALM-1 (BM): 97.9
	EV-2: ND	DDX3X-2 (PB): 99.1	PICALM-2 (BM): 98.1
	EV-3: ND	DDX3X-3 (BM):100	PICALM-3 (BM): 96.0
	EV-4: ND	DDX3X-4 (PB): 82.8	PICALM-4 (BM): 99.6
	EV-5 (PB): 11.6	DDX3X-5 (PB): 72.2	PICALM-5 (BM): 90.2
Median (range)	**11.6**	**90.3 (72.2-100.0)**	**97.9 (90.2-99.6)**
CD4 (APC-Cy7)	1: ND	1: ND	1: 98.9
	2: ND	2: 92.8	2: 35.0
	3: ND	3: 99.8	3: 97.0
	4: ND	4: 89.5	4: 99.6
	5: 100	5: 99.9	5: 97.0
Median (range)	**100**	**96.3 (89.5-99.9)**	**97.0 (35.0-99.6)**
CD25 (APC)	1: ND	1: ND	1: 33.1
	2: ND	2: 3.3	2: 1.2
	3: ND	3: 4.6	3: 2.4
	4: ND	4: 16.0	4: 1.3
	5: 13.5	5: 8.1	5: 1.4
Median (range)	**13.5**	**6.4 (3.3-16.0)**	**1.4 (1.2-33.1)**
CD44 (PE-Cy7)	1: ND	1: ND	1: 74.8
	2: ND	2: 96.6	2: 26.2
	3: ND	3: 85.9	3: 97.6
	4: ND	4: 31.0	4: 71.7
	5: 99.7	5: 100.0	5: 77.2
Median (range)	**99.7**	**91.3 (31.0-100.0)**	**74.8 (26.2-97.6)**
CD45 (PerCP-Cy5.5)	1: ND	1: ND	1: 99.4
	2: ND	2: 99.4	2: 99.3
	3: ND	3: 100	3: 99.2
	4: ND	4: 98.4	4: 97.6
	5: 92.9	5: 94.6	5: 99.6
Median (range)	**92.9**	**98.9 (94.6-100.0)**	**99.3 (97.6-99.6)**
sCD3 (V450)	1: ND	1: ND	1: 3.0
	2: ND	2: 13.2	2: 16.4
	3: ND	3: 8.6	3: 14.4
	4: ND	4: 7.9	4: 1.8
	5: 29.2	5: 56.1	5: 61.6
Median (range)	**29.2**	**10.9 (7.9-56.1)**	**14.4 (1.8-61.6)**
CD5 (APC)	1: ND	1: ND	1: 92.0
	2: ND	2: 77.2	2: 99.9
	3: ND	3: 99.3	3: 97.5
	4: ND	4: 82.4	4: 80.8
	5: 81.4	5: 96.8	5: 94.4
Median (range)	**81.4**	**89.6 (77.2-93.3)**	**94.4 (80.8-99.9)**

c-Kit (PE-Cy7)	1: ND	1: ND	1: 25.0
	2: ND	2: 76.7	2: 59.8
	3: ND	3: 4.1	3: 22.4
	4: ND	4: 74.7	4: 39.7
	5: 99.2	5: 23.1	5: 59.8
Median (range)	**99.2**	**48.9 (4.1-76.7)**	**25.0 (6.7-59.8)**
CD34 (AF700)	1: ND	1: ND	1: 98.8
	2: ND	2: 37.8	2: 99.0
	3: ND	3: 38.0	3: 99.3
	4: ND	4: 32.1	4: 94.4
	5: 50.6	5: 99.5	5: 89.2
Median (range)	**50.6**	**37.9 (32.1-99.5)**	**98.8 (89.2-99.3)**
Sca-1 (PE-CF594)	1: ND	1: ND	1: 85.7
	2: ND	2: 99.0	2: 71.7
	3: ND	3: 19.8	3: 19.1
	4: ND	4: 95.0	4: 95.9
	5: 99.9	5: 90.2	5: 84.4
Median (range)	**99.9**	**92.6 (19.8-99.0)**	**84.4 (19.1-95.9)**
CD13 (BV650)	1: ND	1: ND	1: 0.2
	2: ND	2: 14.9	2: 1.2
	3: ND	3: 0.3	3: 2.7
	4: ND	4: 22.6	4: 4.4
	5: 95.1	5: 14.8	5: 17.4
Median (range)	**95.1**	**14.9 (0.3-22.6)**	**2.7 (0.2-17.4)**
CD2 (PE)	1: ND	1: ND	1: 100
	2: ND	2: ND	2: 99.9
	3: ND	3: 99.9	3: ND
	4: ND	4: ND	4: ND
	5: 47.6	5: ND	5: ND
Median (range)	**47.6**	**99.9**	**100.0**

Abbreviations: BM, bone marrow; ND, no data; PB, peripheral blood

Supplementary Table 4.4 – Immunophenotyping of mouse THY

Marker	% positive	
	DDX3X-MLLT10	*PICALM-MLLT10*
CD8 (AF700)	DDX3X-1: ND	PICALM-1: 98.7
	DDX3X-2: 93.9	PICALM-2: 97.0
	DDX3X-3: ND	PICALM-3: 84.3
	DDX3X-4: ND	PICALM-4: 81.3
	DDX3X-5: 73.8	PICALM-5: 87.8
Median (range)	**83.9 (73.8-93.9)**	**87.8 (81.3-98.7)**
CD4 (APC-Cy7)	1: ND	1: 98.4
	2: 95.8	2: 33.0
	3: ND	3: 99.5
	4: ND	4: 99.9
	5: 99.8	5: 85.2
Median (range)	**97.8 (95.8-99.8)**	**98.4 (33.0-99.9)**
CD25 (APC)	95.1c: ND	1: 75.7
	95.1d: 2.5	2: 2.8
	98.1a: ND	3: 1.9
	98.1b: ND	4: 0.3
	98.1c: 5.6	5: 0.4
Median (range)	**4.1 (2.5-5.6)**	**1.9 (0.3-75.7)**
CD44 (PE-Cy7)	1: ND	1: 62.0
	2: 92.5	2: 30.8
	3: ND	3: 98.7
	4: ND	4: 56.2
	5: 99.8	5: 71.3
Median (range)	**96.2 (92.5-99.8)**	**62.0 (30.8-98.7)**
CD45 (PerCP-Cy5.5)	1: ND	1: 100.0
	2: 97.7	2: 100.0
	3: ND	3: 100.0
	4: ND	4: 100.0
	5: 99.6	5: 100.0
Median (range)	**98.7 (97.7-99.6)**	**100.0**
sCD3 (V450)	1: ND	1: 16.0
	2: 11.5	2: 24.1
	3: ND	3: 14.4
	4: ND	4: 21.4
	5: 49.2	5: 34.3
Median (range)	**30.4 (11.5-49.2)**	**21.4 (14.4-34.3)**
CD5 (APC)	1: ND	1: 99.1
	2: 75.5	2: 97.4
	3: ND	3: 97.5
	4: ND	4: 92.8
	5: 99.2	5: 96.7
Median (range)	**87.4 (75.5-99.2)**	**97.4 (92.8-99.1)**

c-Kit (PE-Cy7)	1: ND	1: 2.2
	2: 83.8	2: 35.3
	3: ND	3: 22.4
	4: ND	4: 6.4
	5: 20.9	5: 25.8
Median (range)	**52.4 (20.9-83.8)**	**15.8 (2.2-35.7)**
CD34 (AF700)	1: ND	1: 99.9
	2: 45.8	2: 99.9
	3: ND	3: 99.7
	4: ND	4: 92.1
	5: 99.9	5: 98.9
Median (range)	**72.9 (45.8-99.9)**	**99.7 (92.1-99.9)**
Sca-1 (PE-CF594)	1: ND	1: 56.8
	2: 99.7	2: 70.9
	3: ND	3: 19.1
	4: ND	4: 54.4
	5: 85.5	5: 69.7
Median (range)	**92.6 (85.5-99.7)**	**56.8 (19.1-70.9)**
CD13 (BV650)	1: ND	1: 0.2
	2: 13.6	2: 3.8
	3: ND	3: 2.7
	4: ND	4: 0.6
	5: 7.4	5: 13.5
Median (range)	**10.5 (7.4-13.6)**	**2.7 (0.2-13.5)**
CD2 (PE)	1: ND	1: 100.0
	2: ND	2: 98.5
	3: ND	3: ND
	4: ND	4: ND
	5: ND	5: ND
Median (range)	**ND**	**99.3 (98.5-100.0)**

Abbreviations: ND, no data

Supplementary Table 4.5 – Comparison of *HOX* gene expression pre-injection vs post-injection

	ΔCt, pre-injection cells +SD			ΔCt, cells from murine liver* +SD		
	EV	*DDX3X*	*PICALM*	**EV**[x]	*DDX3X*	*PICALM*
HOXA1	8.6 (3.0)	9.5 (1.8)	10.0 (1.6)	11.7	10.4 (4.0)	13.4 (1.3)
HOXA3	6.4 (1.3)	6.2 (1.5)	7.4 (0.9)	10.0	10.1 (1.7)	14.0 (2.2)
HOXA5	7.3 (1.5)	7.6 (1.2)	8.7 (1.3)	10.0	12.8 (4.3)	15.8 (0.9)
HOXA7	NE	NE	NE	NE	NE	NE
HOXA9	12.7 (1.6)	12.5 (2.7)	11.8 (0.9)	7.8	10.7 (3.8)	11.7 (6.7)
HOXA10	12.7 (0.5)	12.3 (0.4)	13.5 (0.6)	NE	NE	NE
MEIS1	6.6 (2.4)	6.9 (1.8)	7.5 (1.4)	10.4	11.0 (4.0)	12.0 (6.8)
HOXB1	10.3 (1.7)	10.2 (0.9)	9.4 (0.8)	9.0	9.1 (1.5)	10.3 (3.7)

Data presented as mean ΔCt values (+SD), where a lower ΔCt indicates higher expression.

*Except for EV-5, where spleen RNA was used.

[x]SD not provided as only one EV mouse was analysed.

Abbreviations: EV, empty vector; NE, not expressed; SD, standard deviation.

4.6 Chapter References

1   Stoddart, A. *et al.* The clathrin-binding domain of CALM-AF10 alters the phenotype of myeloid neoplasms in mice. *Oncogene* **31**, 494-506, doi:10.1038/onc.2011.251 (2012).

2   Dreyling, M. *et al.* The t(10;11)(p13;q14). in the U937 cell line results in the fusion of the AF10 gene and CALM, encoding a new member of the AP-3 clathrin assembly protein family. *Medical Sciences* **93**, 4804-4809, doi:10.1073/pnas.93.10.4804 (1996).

3   Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603-607, doi:10.1038/nature11003 (2012).

4   Liu, Y. *et al.* The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nat Genet* **49**, 1211-1218, doi:10.1038/ng.3909 (2017).

5   Zhu, G. *et al.* Mutant p53 in Cancer Progression and Targeted Therapies. *Front Oncol* **10**, 595187, doi:10.3389/fonc.2020.595187 (2020).

6   Ohki, K. *et al.* Impact of immunophenotypic characteristics on genetic subgrouping in childhood acute lymphoblastic leukemia: Tokyo Children's Cancer Study Group (TCCSG) study L04-16. *Genes Chromosomes Cancer* **59**, 551-561, doi:10.1002/gcc.22858 (2020).

7   Treanor, L. M. *et al.* Interleukin-7 receptor mutants initiate early T cell precursor leukemia in murine thymocyte progenitors with multipotent potential. *J Exp Med* **211**, 701-713, doi:10.1084/jem.20122727 (2014).

8   Holmes, R. & Zuniga-Pflucker, J. C. The OP9-DL1 system: generation of T-lymphocytes from embryonic or hematopoietic stem cells in vitro. *Cold Spring Harb Protoc* **2009**, pdb prot5156, doi:10.1101/pdb.prot5156 (2009).

9   de Pooter, R. & Zuniga-Pflucker, J. C. T-cell potential and development in vitro: the OP9-DL1 approach. *Curr Opin Immunol* **19**, 163-168, doi:10.1016/j.coi.2007.02.011 (2007).

10   Layssol-Lamour, C. J., Sarry, J. E., Braun, J. D., Trumel, C. & Bourges-Abella, N. H. Reference Values for Hematology, Plasma Biochemistry, Bone Marrow Cytology and Bone Histology of NOD.Cg-Prkdc(scid) Il2rg(tm1Wjl)/ SzJ Immunodeficient Mice. *J Am Assoc Lab Anim Sci* **60**, 4-17, doi:10.30802/AALAS-JAALAS-20-000020 (2021).

11   Lockridge, J. L. *et al.* Analysis of the CD1 antigen presenting system in humanized SCID mice. *PLoS One* **6**, e21701, doi:10.1371/journal.pone.0021701 (2011).

12   Tillman, H., Janke, L. J., Funk, A., Vogel, P. & Rehg, J. E. Morphologic and Immunohistochemical Characterization of Spontaneous Lymphoma/Leukemia in NSG Mice. *Vet Pathol* **57**, 160-171, doi:10.1177/0300985819882631 (2020).

13   Dik, W. A. *et al.* CALM-AF10+ T-ALL expression profiles are characterized by overexpression of HOXA and BMI1 oncogenes. *Leukemia* **19**, 1948-1957, doi:10.1038/sj.leu.2403891 (2005).

14   Kang, H. *et al.* Dysregulated transcriptional networks in KMT2A- and MLLT10-rearranged T-ALL. *Biomark Res* **6**, 27, doi:10.1186/s40364-018-0141-z (2018).

15   Bond, J. *et al.* Cryptic XPO1-MLLT10 translocation is associated with HOXA locus deregulation in T-ALL. *Blood* **124**, 3023-3025, doi:10.1182/blood-2014-04-567636 (2014).

16   Bond, J. *et al.* NAP1L1-MLLT10 is a rare recurrent translocation that is associated with HOXA activation and poor treatment response in T-cell acute lymphoblastic leukaemia. *British Journal of Haematology* **174**, 470-473, doi:10.1111/bjh.13772 (2015).

17   Brandimarte, L. *et al.* DDX3X-MLLT10 fusion in adults with NOTCH1 positive T-cell acute lymphoblastic leukemia. *Haematologica* **99**, 64-66, doi:10.3324/haematol.2013.101725 (2014).

18   Caudell, D., Zhang, Z., Chung, Y. J. & Aplan, P. D. Expression of a CALM-AF10 fusion gene leads to Hoxa cluster overexpression and acute leukemia in transgenic mice. *Cancer Res* **67**, 8022-8031, doi:10.1158/0008-5472.CAN-06-3749 (2007).

19   Okada, Y. *et al.* Leukaemic transformation by CALM-AF10 involves upregulation of Hoxa5 by hDOT1L. *Nat Cell Biol* **8**, 1017-1024, doi:10.1038/ncb1464 (2006).

20   Escobar, H. *et al.* Acute renal failure associated with bilateral enlargement of the kidneys: a rare manifestation of acute lymphoblastic leukemia (ALL). *Klin Padiatr* **221**, 176-178, doi:10.1055/s-0029-1216365 (2009).

21   Skeith, L., Lazo-Langner, A. & Mangel, J. Kidney and pancreatic extramedullary relapse in adult acute lymphoblastic leukemia: a case report and review of the literature. *Case Rep Hematol* **2013**, 637264, doi:10.1155/2013/637264 (2013).

22   Asdahl, P. H., Warner, L. F., Bendix, K. & Hasle, H. Acute renal failure and normal blood count: A rare presentation of T-cell acute lymphoblastic leukemia. *Leuk Res Rep* **3**, 14-16, doi:10.1016/j.lrr.2013.11.002 (2014).

23   Vijayasekharan, K. *et al.* Bilateral massive nephromegaly-A rare presentation of t-cell acute lymphoblastic leukemia. *Leuk Res Rep* **15**, 100246, doi:10.1016/j.lrr.2021.100246 (2021).

24   D'Angelo, P. *et al.* Prognostic Value of Nephromegaly at Diagnosis of Childhood Acute Lymphoblastic Leukemia. *94* **2**, 84-85, doi:10.1159/000203979 (1995).

25   Hann, I., Lees, PD, Palmer, MK, Gupta, S, Morris-Jones, PH. Renal Size as a Prognostic Factor in Childhood Acute Lymphoblastic Leukemia. *Cancer* **48**, 207-209, doi:10.1002/1097-0142 (1981).

26   Khosravi-Maharlooei, M. *et al.* Rapid thymectomy of NSG mice to analyze the role of native and grafted thymi in humanized mice. *Eur J Immunol* **50**, 138-141, doi:10.1002/eji.201948205 (2020).

27   Cepeda, S. & Griffith, A. V. Thymic stromal cells: Roles in atrophy and age-associated dysfunction of the thymus. *Exp Gerontol* **105**, 113-117, doi:10.1016/j.exger.2017.12.022 (2018).

28   Van Vlierberghe, P. *et al.* Prognostic relevance of integrated genetic profiling in adult T-cell acute lymphoblastic leukemia. *Blood* **122**, 74-82, doi:10.1182/blood-2013-03-491092 (2013).

29   Noronha, E. P. *et al.* T-lymphoid/myeloid mixed phenotype acute leukemia and early T-cell precursor lymphoblastic leukemia similarities with NOTCH1 mutation as a good prognostic factor. *Cancer Manag Res* **11**, 3933-3943, doi:10.2147/CMAR.S196574 (2019).

30   Sayed, D. M. *et al.* Outcome and Clinical Significance of Immunophenotypic Markers Expressed in Different Treatment Protocols of Pediatric Patients With T-ALL in Developing Countries. *Clin Lymphoma Myeloma Leuk* **17**, 443-449, doi:10.1016/j.clml.2017.05.012 (2017).

31   Noronha, E. P. *et al.* The Profile of Immunophenotype and Genotype Aberrations in Subsets of Pediatric T-Cell Acute Lymphoblastic Leukemia. *Front Oncol* **9**, 316, doi:10.3389/fonc.2019.00316 (2019).

32   Abdelali, R. *et al.* The prognosis of CALM-AF10-positive adult T-cell acute lymphoblastic leukemias depends on the stage of maturation arrest. *Haematologica* **98**, 1711-1717, doi:10.3324/haematol.2013.086082 (2013).

33   Hsu, Y. C., Mildenstein, K., Hunter, K., Tkachenko, O. & Mullen, C. A. Acute lymphoid leukemia cells with greater stem cell antigen-1 (Ly6a/Sca-1) expression exhibit higher levels of metalloproteinase activity and are more aggressive in vivo. *PLoS One* **9**, e88966, doi:10.1371/journal.pone.0088966 (2014).

34   Tembhare, P. R. *et al.* Eleven-marker 10-color flow cytometric assessment of measurable residual disease for T-cell acute lymphoblastic leukemia using an approach of exclusion. *Cytometry B Clin Cytom* **100**, 421-433, doi:10.1002/cyto.b.21939 (2021).

35   Wang, Y. Z. *et al.* A seven-color panel including CD34 and TdT could be applied in >97% patients with T cell lymphoblastic leukemia for minimal residual disease detection independent of the initial phenotype. *Leuk Res* **72**, 12-19, doi:10.1016/j.leukres.2018.07.012 (2018).

36   Noronha, E. P. *et al.* Immunophenotyping with CD135 and CD117 predicts the FLT3, IL-7R and TLX3 gene mutations in childhood T-cell acute leukemia. *Blood Cells Mol Dis* **57**, 74-80, doi:10.1016/j.bcmd.2015.12.003 (2016).

37   Gonen, M. *et al.* CD25 expression status improves prognostic risk classification in AML independent of established biomarkers: ECOG phase 3 trial, E1900. *Blood* **120**, 2297-2306, doi:10.1182/blood-2012-02-414425 (2012).

38   Mohammed Said, R. M., Asfour, I. A., Shams El Din, G. M. K. & Juadam, E. M. Evaluation of CD 25 (IL2 Receptor Alpha) Expression in Adult Acute Lymphoblastic Leukemia Patients. *Clinical Lymphoma Myeloma and Leukemia* **17**, doi:10.1016/j.clml.2017.07.041 (2017).

39   Aref, S., El Agdar, M., Khaled, N., Ibrahim, L. & El-Ghonemy, M. S. Clinical Impact of CD25/CD123 Coexpression in Adult B-Cell Acute Lymphoblastic Leukemia Patients. *Adv Hematol* **2020**, 9545717, doi:10.1155/2020/9545717 (2020).

40   Chiaretti, S., Zini, G. & Bassan, R. Diagnosis and subclassification of acute lymphoblastic leukemia. *Mediterr J Hematol Infect Dis* **6**, e2014073, doi:10.4084/MJHID.2014.073 (2014).

41   Bandyopadhyay, S., Das, D., Das, G. & Gayen, S. Unilateral optic nerve infiltration as an initial site of relapse of acute lymphoblastic leukemia in remission. *Oman J Ophthalmol* **3**, 153-154, doi:10.4103/0974-620X.71902 (2010).

42   Pflugrath, A. E. & Brar, V. S. Bilateral optic nerve and retinal infiltration as an initial site of relapse in a child with T-cell acute lymphoblastic leukemia. *Am J Ophthalmol Case Rep* **18**, 100695, doi:10.1016/j.ajoc.2020.100695 (2020).

43   Zhou, S. *et al.* Mouse transplant models for evaluating the oncogenic risk of a self-inactivating XSCID lentiviral vector. *PLoS One* **8**, e62333, doi:10.1371/journal.pone.0062333 (2013).

Appendix 4.1

(A) EV-1

(B) EV-2

BM

SPL

LIV

THY

GFP+
0.24

GFP+
0

GFP+
1.69

GFP+
0.15
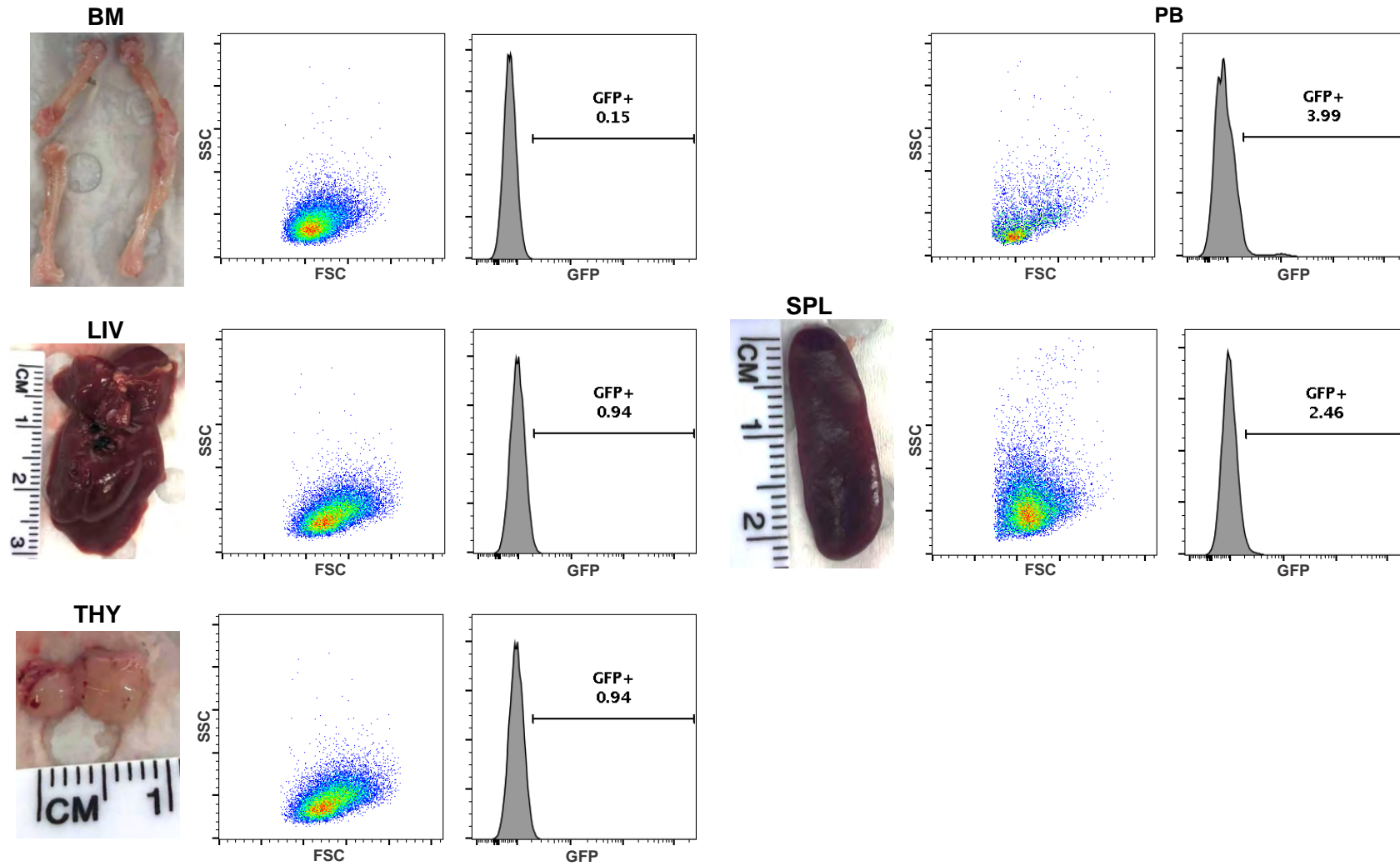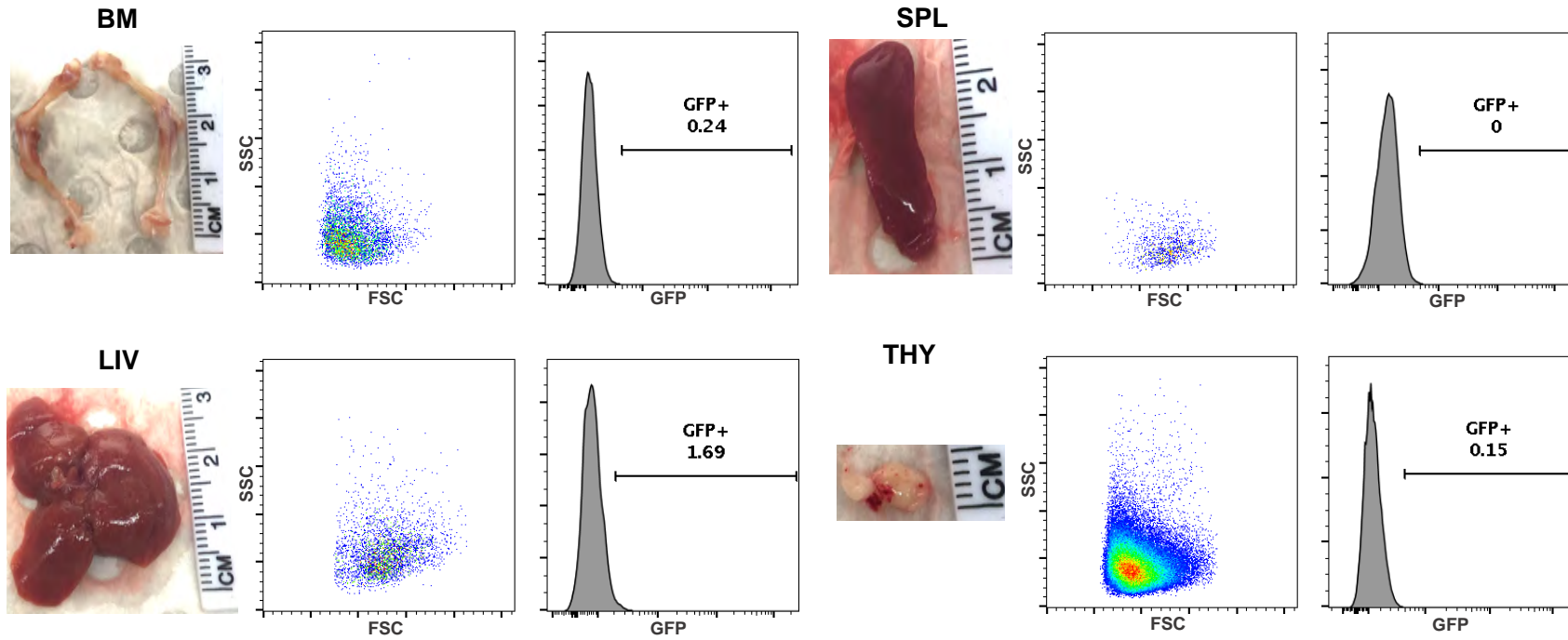
(C) EV-3
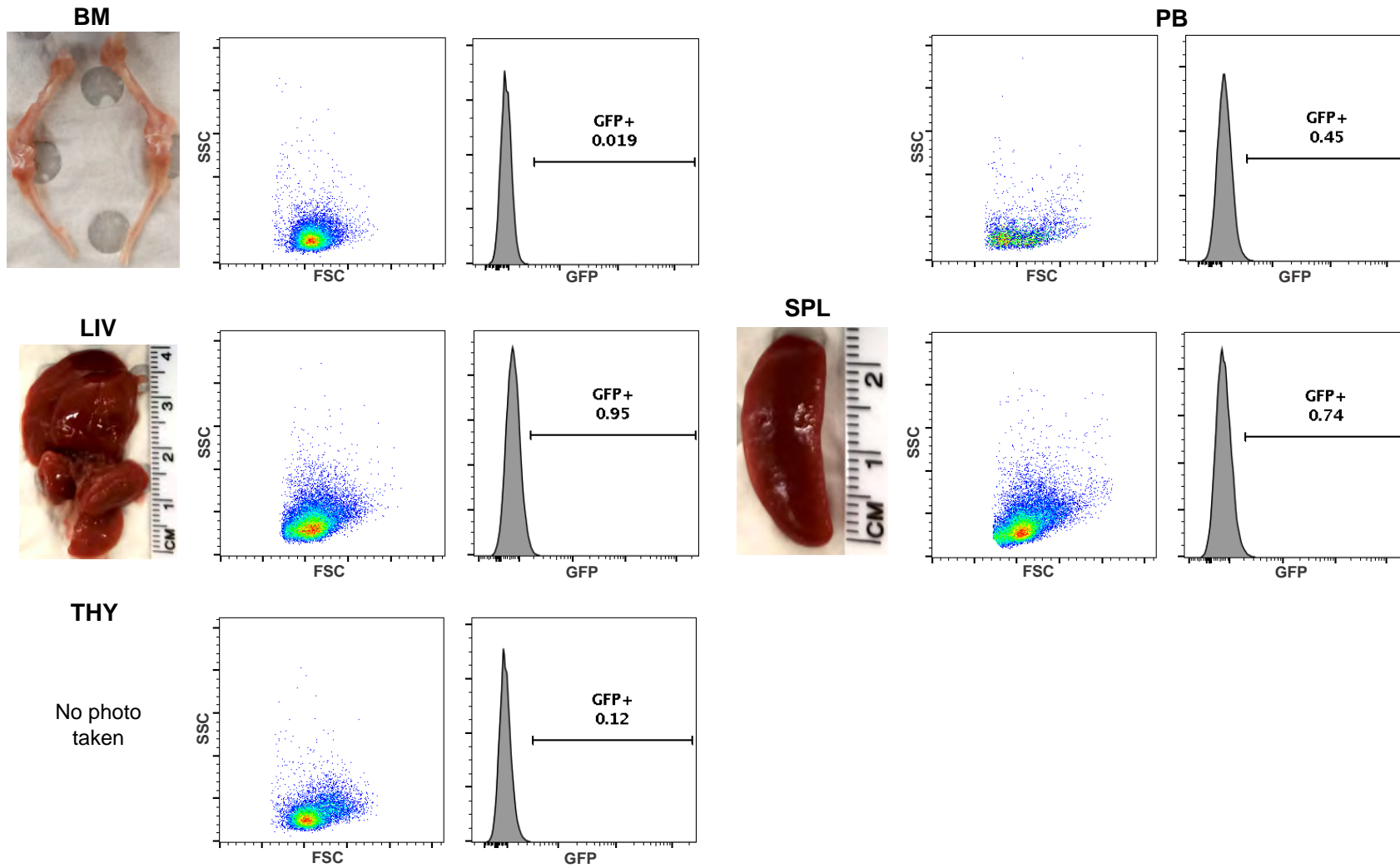
(D) EV-4
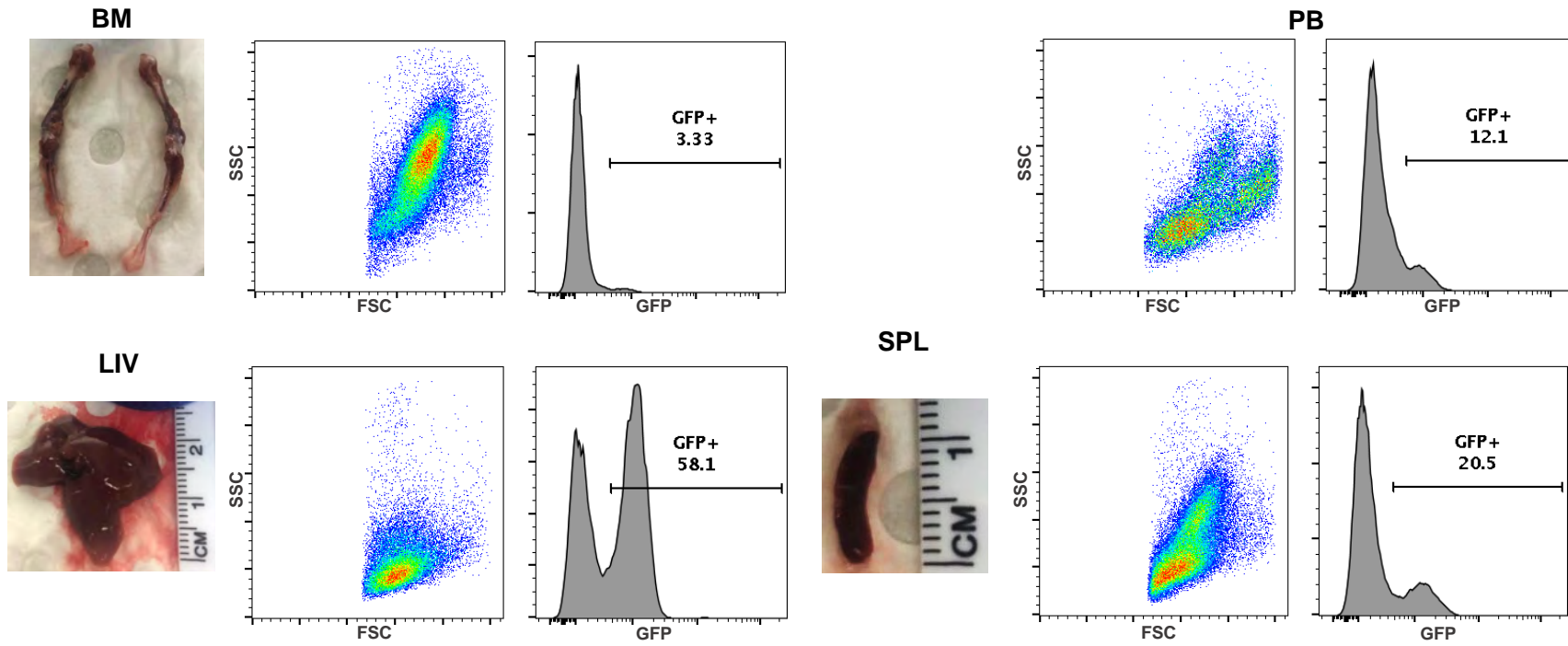
(E) EV-5
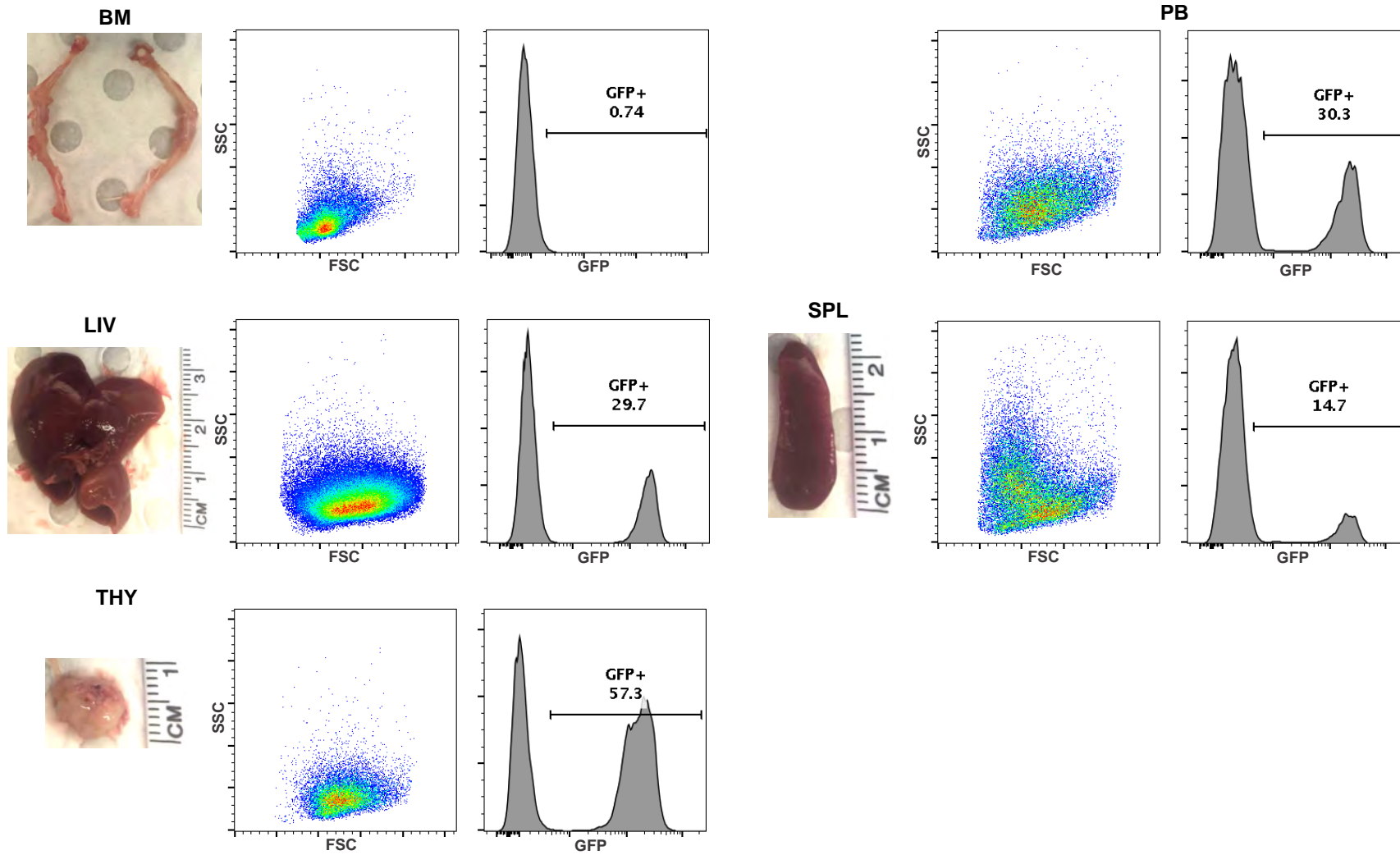
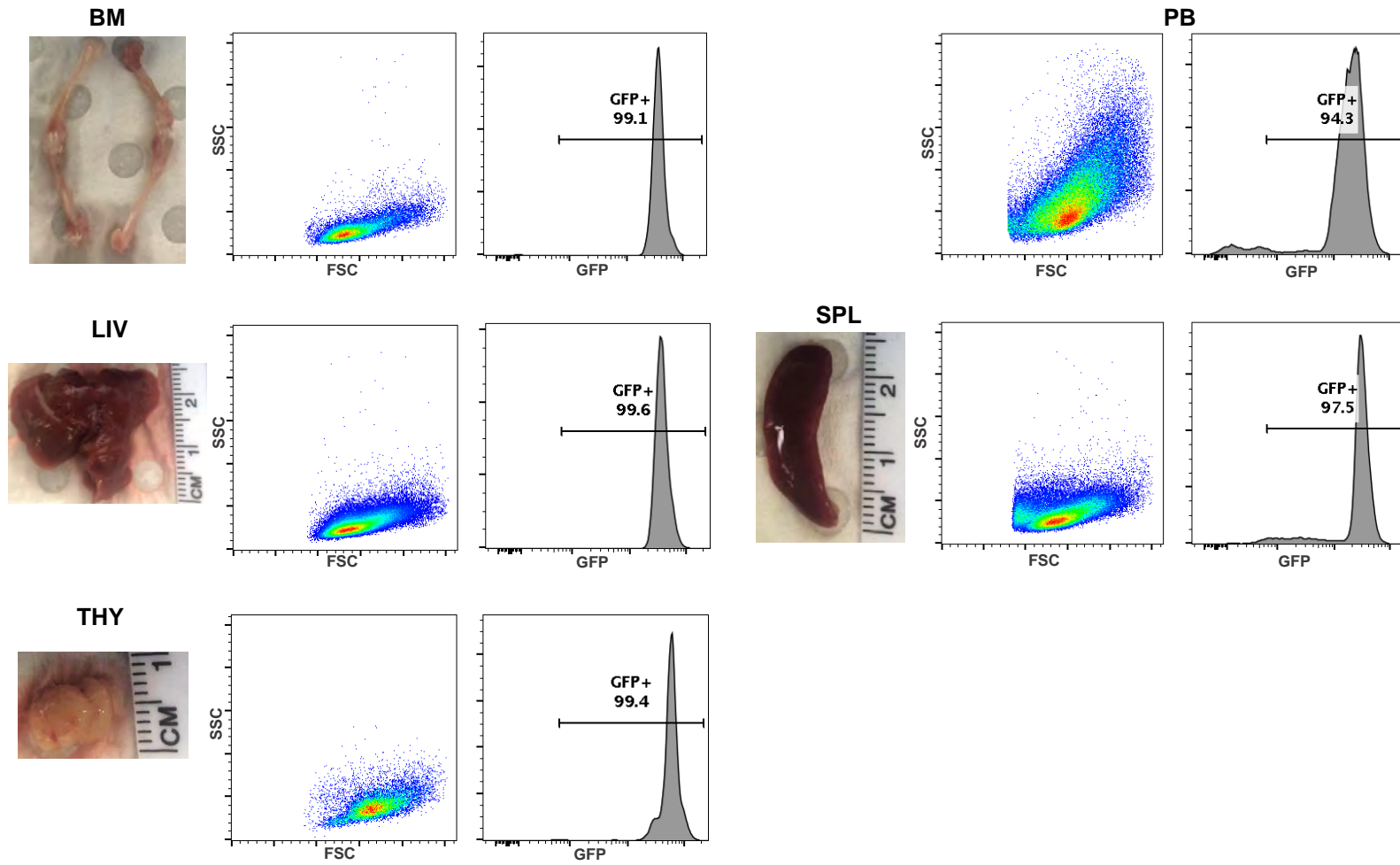**Appendix 4.1 – External haematopoietic organ morphology and properties of cells harvested from mice injected with Arf$^{-/-}$ thymocytes expressing empty vector**
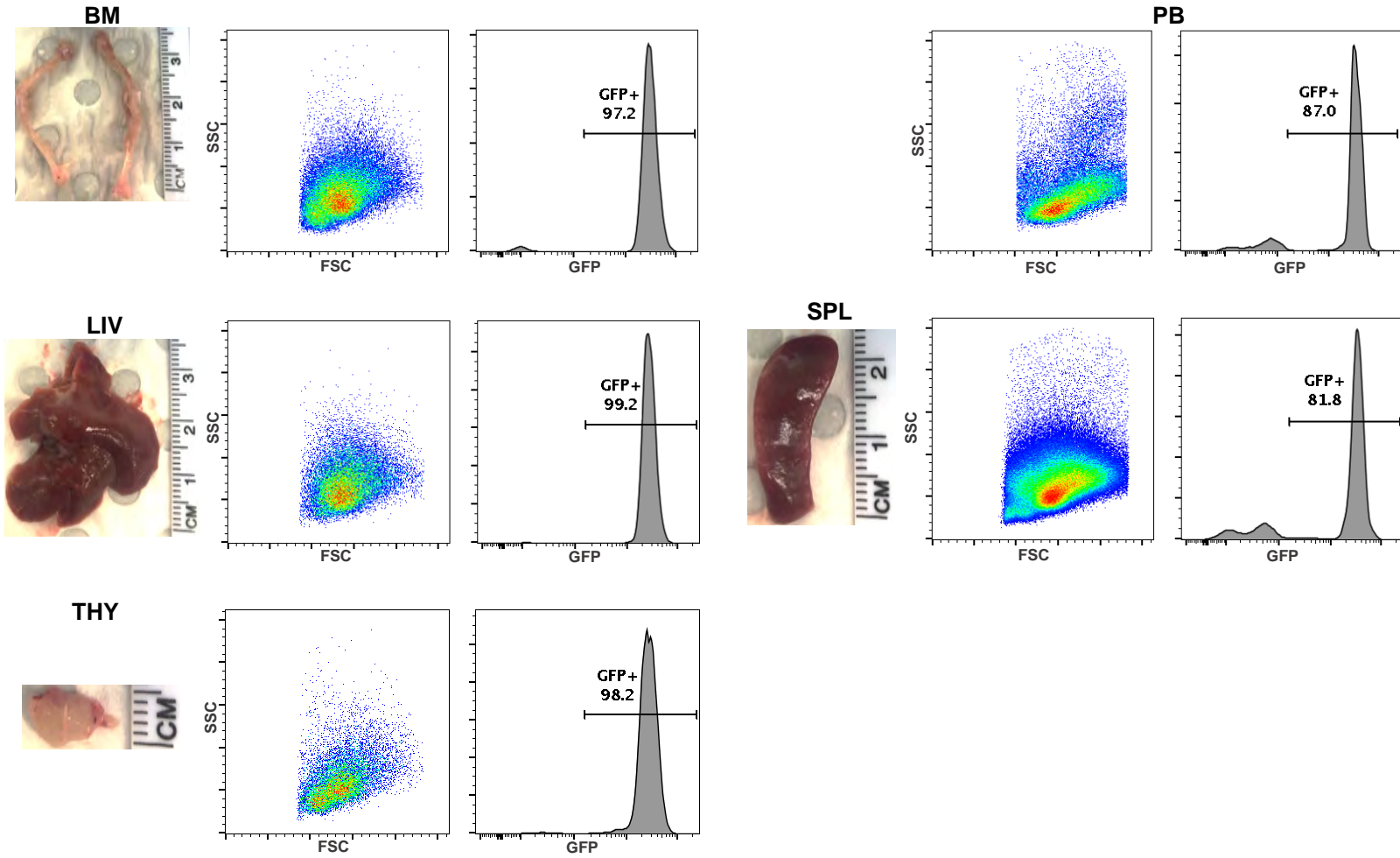
Peripheral blood (PB), liver (LIV), spleen (SPL), thymi (THY) and bone marrow from hind limbs (BM) were harvested from mice for analysis. PB was not harvested from EV-2, and no THY was identified in EV-4. Organs were documented with photographs, and live cells were assessed for size and granularity (FSC vs SSC) and GFP expression (histogram) by flow cytometry on a BD FACSCanto^{TM} or BD FACSFortessa^{TM}.

Appendix 4.2

(A) DDX3X-1
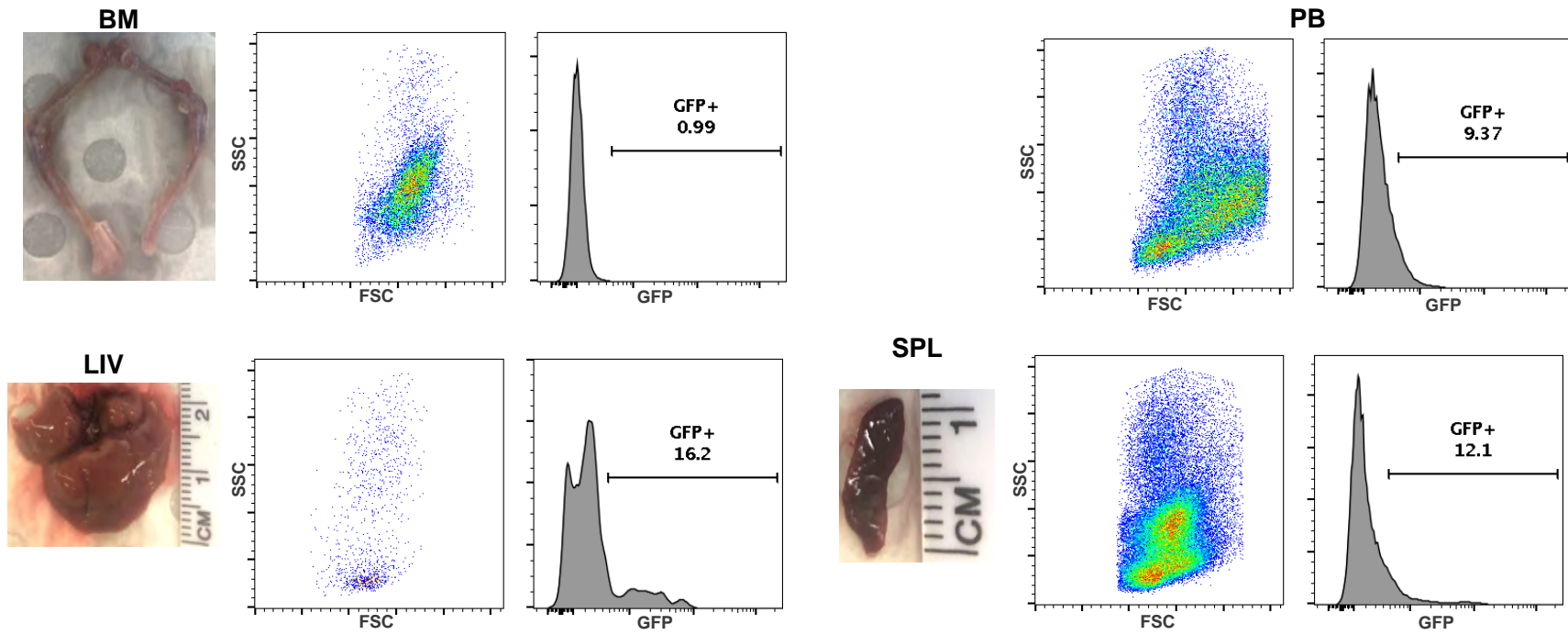
(B) DDX3X-2
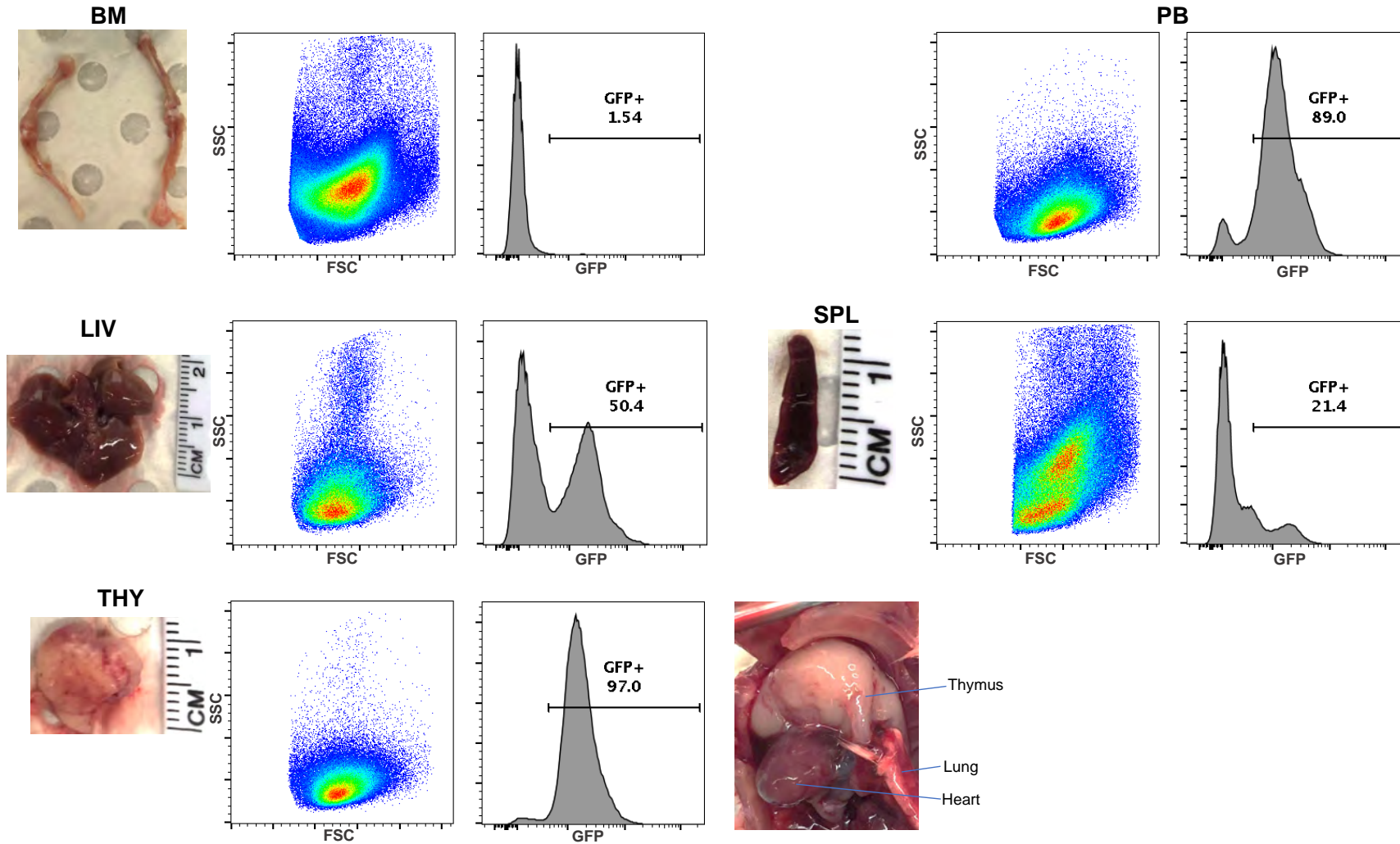
(C) DDX3X-3
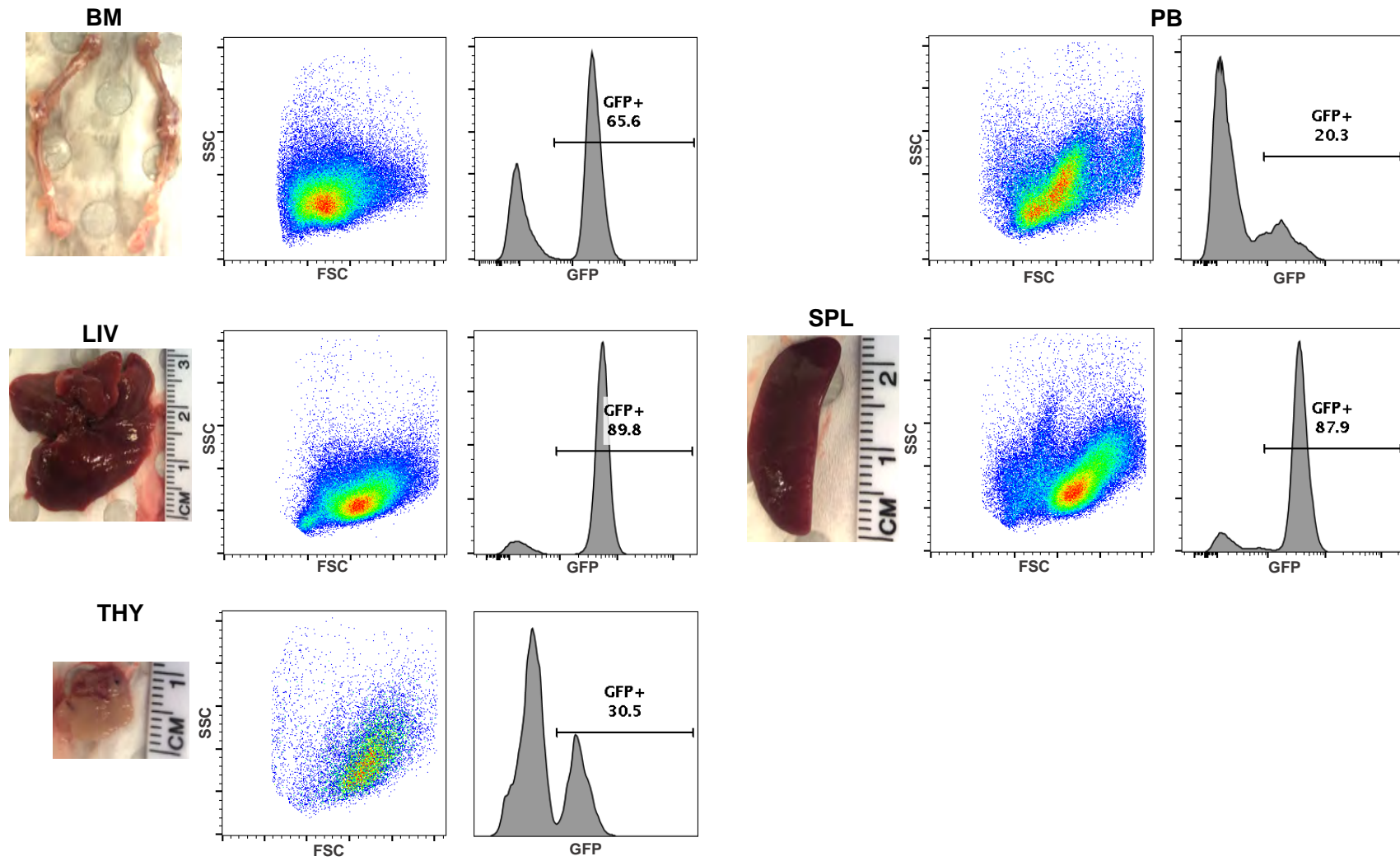
(D) DDX3X-4
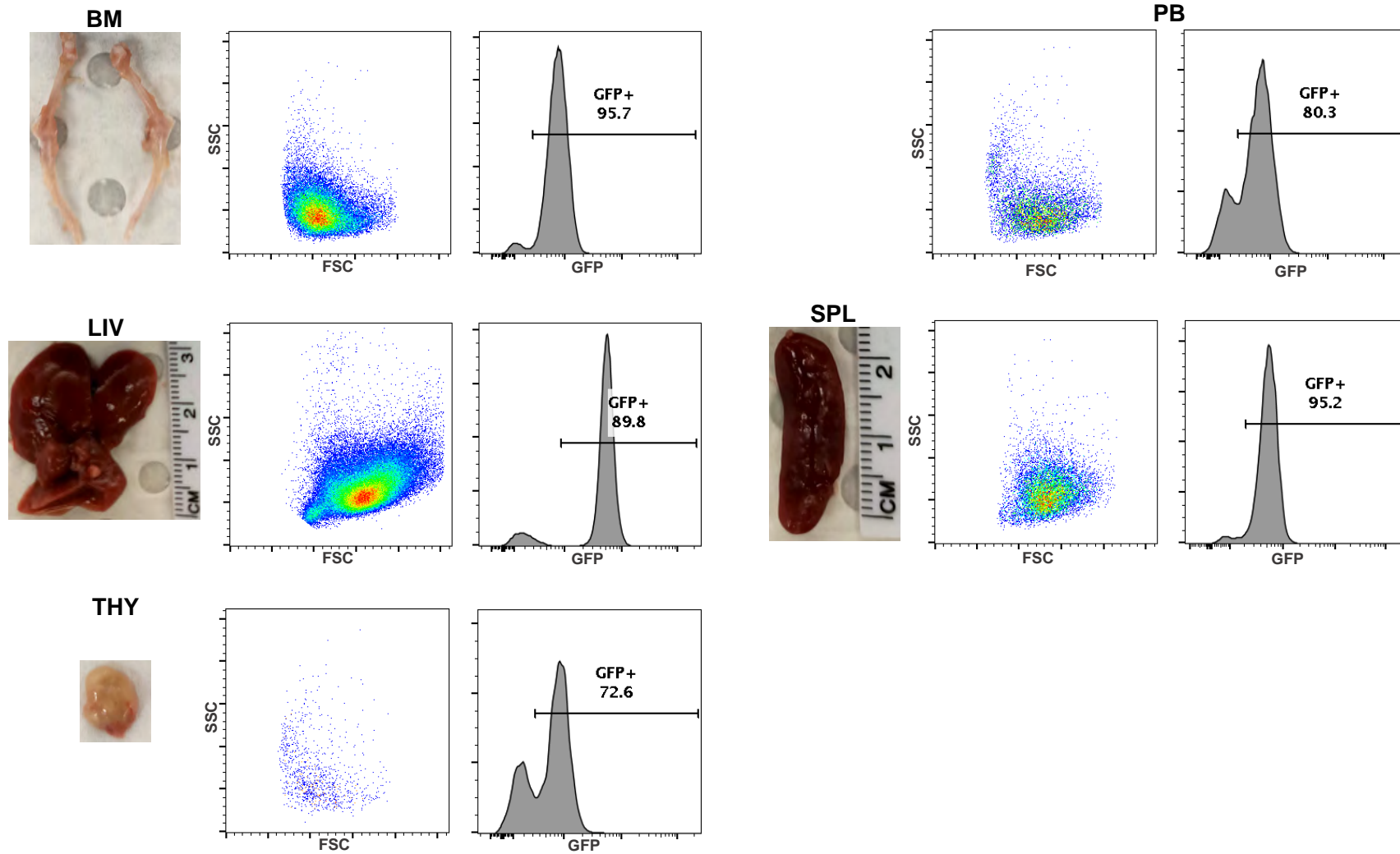
(E) DDX3X-5
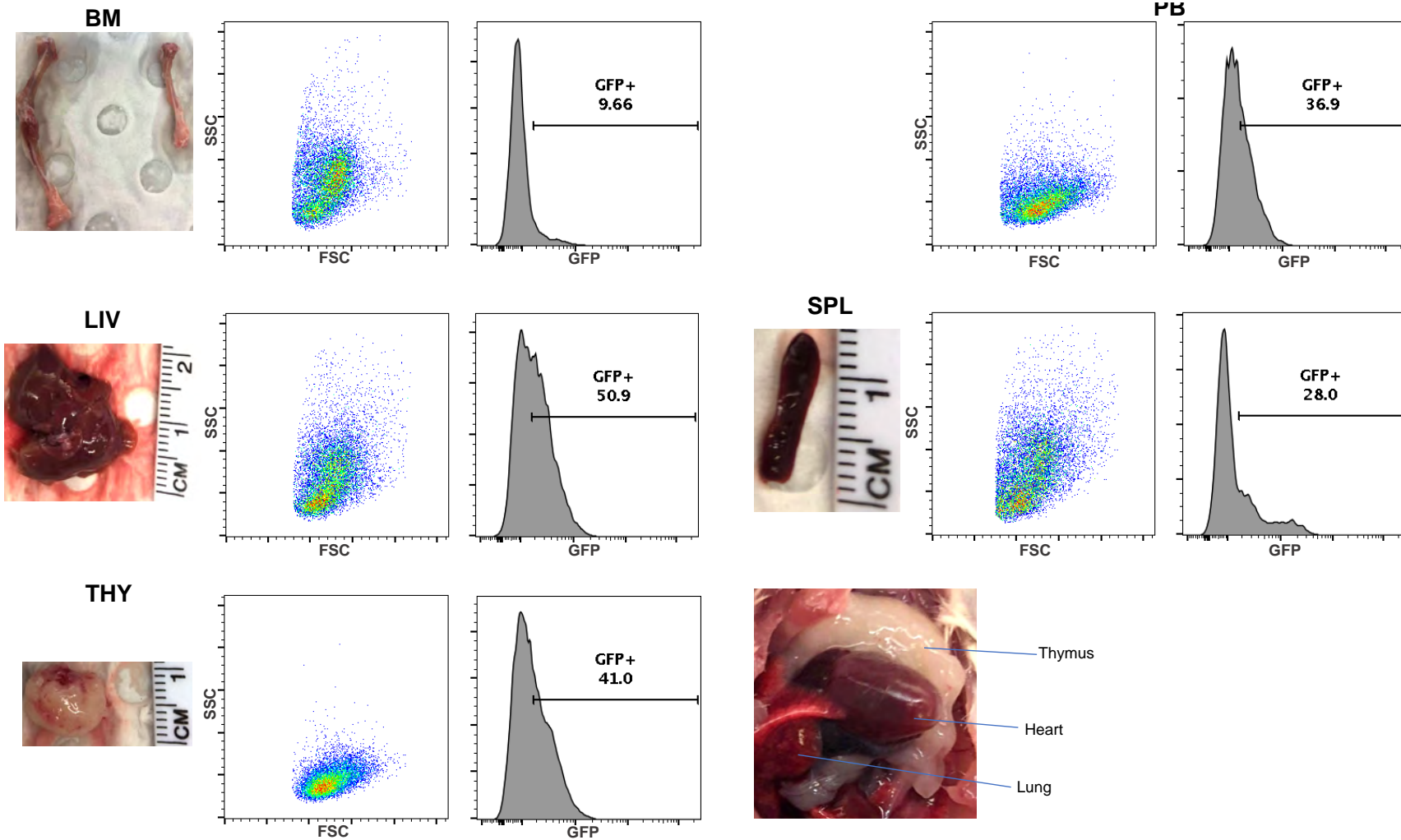
Appendix 4.2 – Documentation of external haematopoietic organ morphology and properties of cells harvested from mice injected with Arf$^{-/-}$ thymocytes expressing *DDX3X-MLLT10*

Peripheral blood (PB), liver (LIV), spleen (SPL), thymi (THY) and bone marrow from hind limbs (BM) were harvested from euthanised mice for analysis. Organs were documented with photographs, and live cells were assessed for size and granularity (FSC vs SSC) and GFP expression (histogram) by flow cytometry on a BD FACSCanto™ or BD FACSFortessa™.

Appendix 4.3

(A) PICALM-1

(B) PICALM-2

(C) PICALM-3
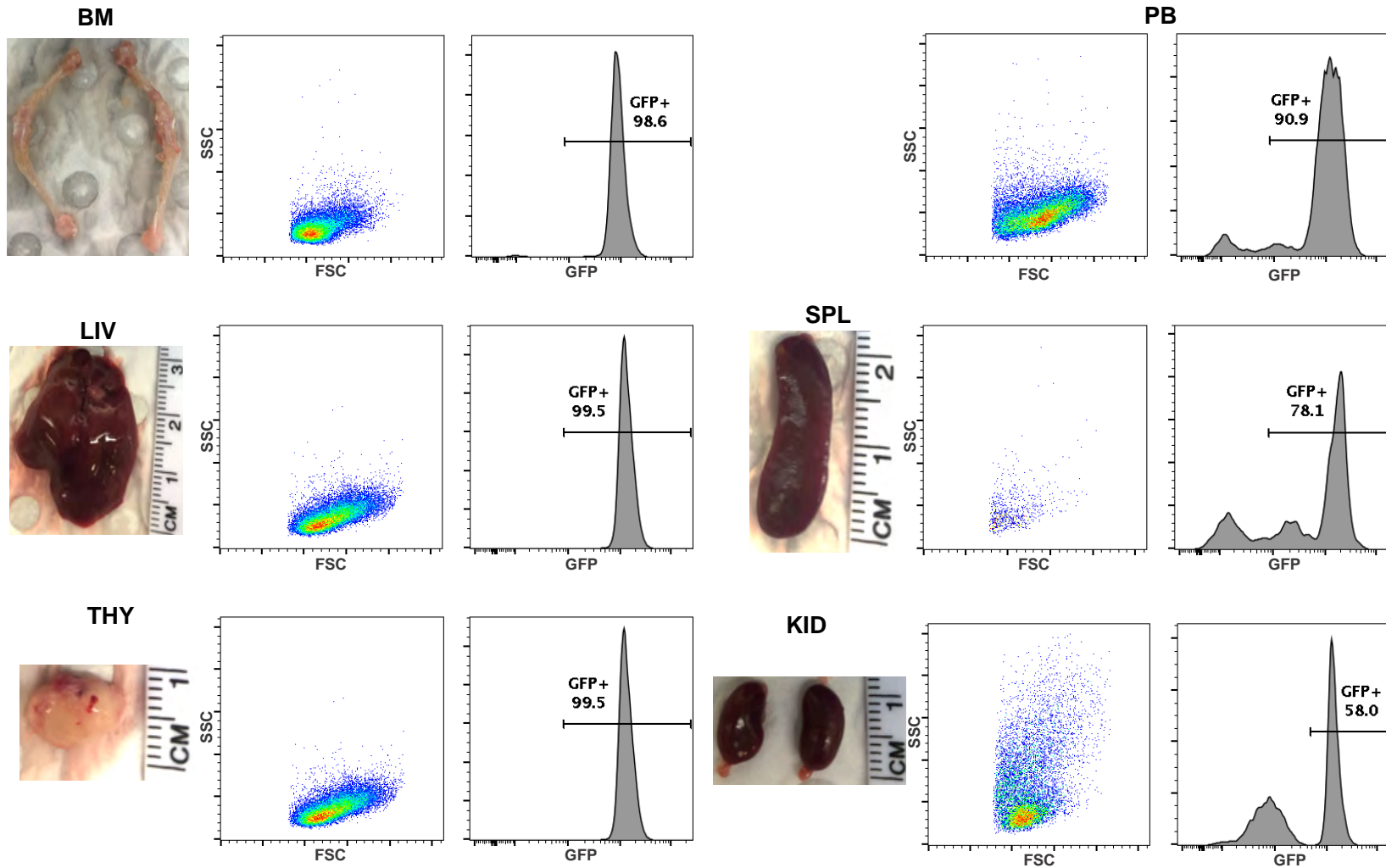
(D) PICALM-4
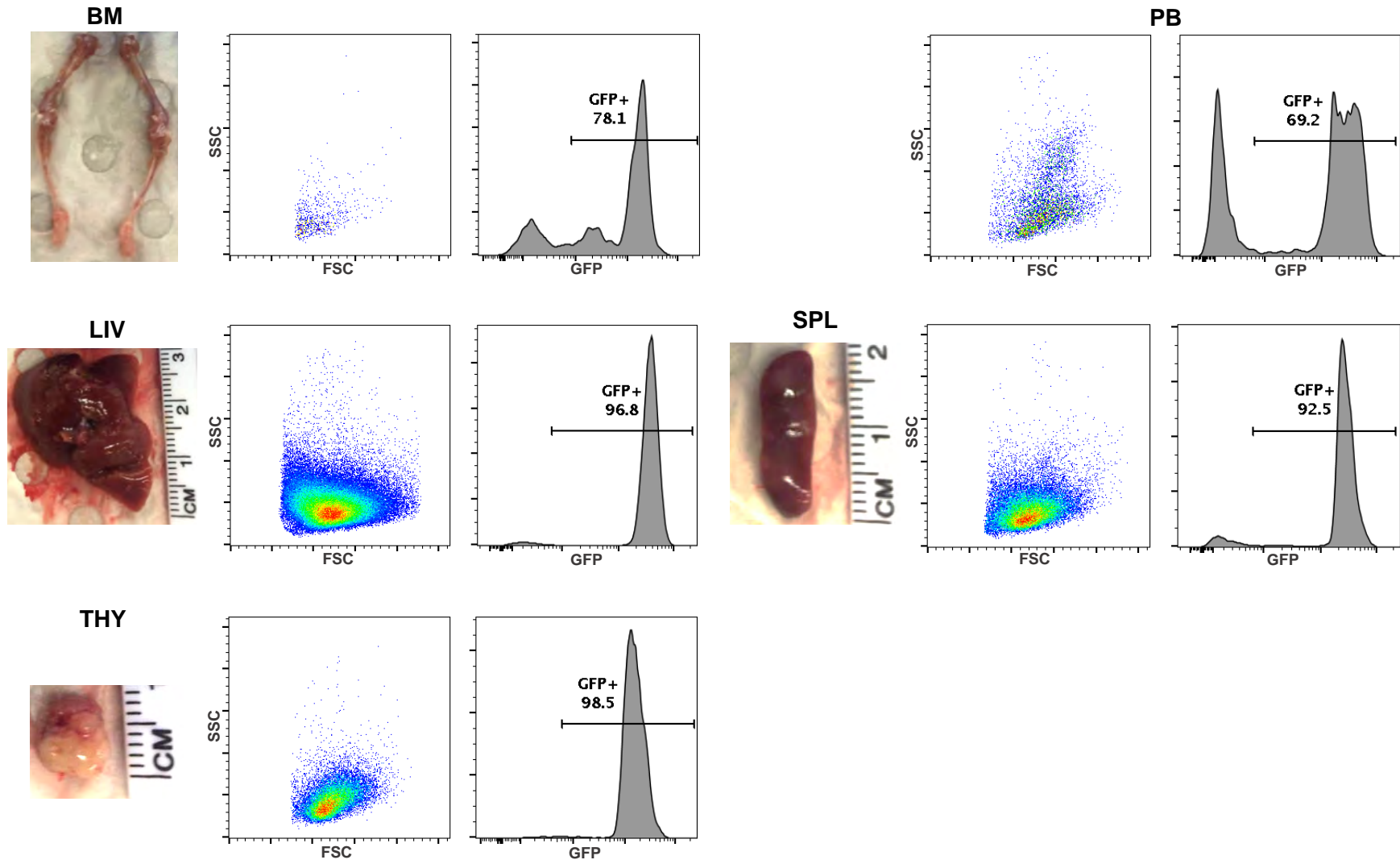
(E) PICALM-5
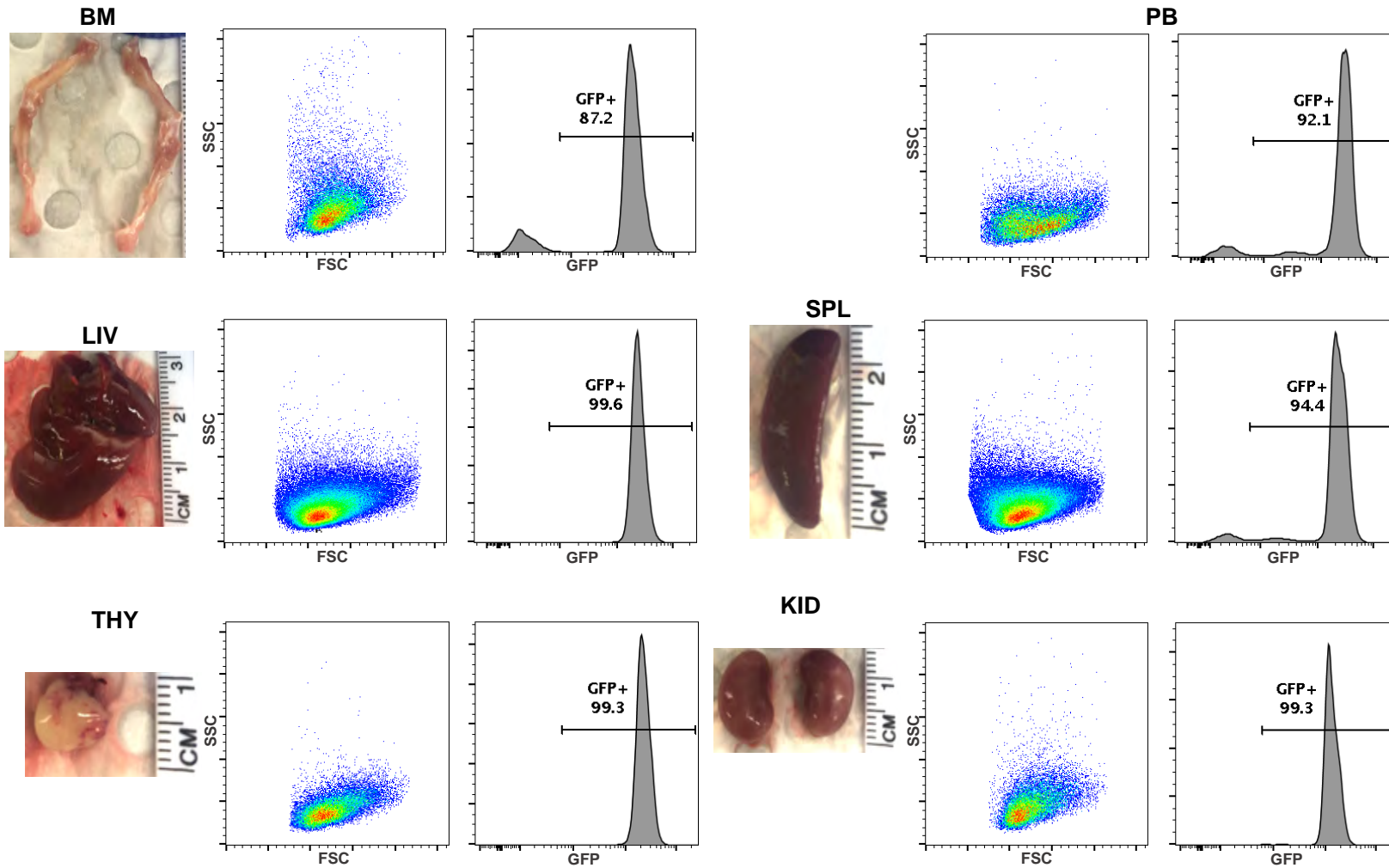
Appendix 4.3 – Documentation of external haematopoietic organ morphology and properties of single cells harvested from mice injected with Arf$^{-/-}$ thymocytes expressing *PICALM-MLLT10*

Peripheral blood (PB), liver (LIV), spleen (SPL), thymi (THY) and bone marrow from hind limbs (BM) were harvested from euthanised mice for analysis. Kidneys (KID) were harvested from PICALM-3 and PICALM-5 only. Organs were documented with photographs, and live single cells were assessed for size and granularity (FSC vs SSC) and GFP expression (histogram) by flow cytometry on a BD FACSCanto^{TM} or BD FACSFortessa^{TM}.

Chapter 5 – Characterising the genomic landscape of non-infant *KMT2A*r and *MLLT10*r ALL

5.1 Introduction

Co-operative genomic alterations are often present in ALL at diagnosis. These alterations include structural aberrations (insertions, deletions and inversions), copy number alterations (CNAs) and single nucleotide variants (SNVs) affecting oncogenes and tumour suppressors[1], that co-operate with initiating alterations (typically fusion genes) to induce ALL[2].

The high incidence of *KMT2A*r in infant ALL (approximately 70% of newly-diagnosed infant ALL cases[3]) suggests a unique aetiology that is distinct from other recurrent genomic subtypes. Infant *KMT2A*r ALL carries a very low mutational burden, and in many cases no co-operative genomic events are identified[4-7]. The absence of co-operative alterations in infants indicates that *KMT2A*r are highly pathogenic lesions, as additional genomic events are not required to induce leukaemia. It was recently demonstrated that human foetal liver haematopoietic stem and progenitor cells (HSPCs) provide a molecular environment that is integral to the distinct gene expression profile of *KMT2A-AFF1* infant B-ALL, which differs from the expression profile of non-infant paediatric *KMT2A-AFF1* B-ALL[8]. The mutational landscape of non-infant *KMT2A*r ALL is poorly characterised, so it is unknown whether *KMT2A*r co-occur with other genomic events to induce ALL in non-infant patients. Detailing the genomic landscape of non-infant *KMT2A*r ALL is critical for understanding the aetiology and biology of *KMT2A*r ALL.

*MLLT10*r T-ALL is another high-risk ALL subtype where co-operative mutations are currently poorly characterised. In chapters 3 and 4, two different murine models expressing *MLLT10*r did not induce the *HOXA* cluster dysregulation observed in the leukaemic blasts of *MLLT10*r positive ALL patients, raising the possibility that additional genomic events are required. Activating *NOTCH1* (60-80% of T-ALL diagnoses)[9,10], frameshift and non-frameshift insertions affecting *PTEN* (14-23%)[9,11], deletion of cell cycle regulators *CDKN2A/B* (70-78%)[9,10] and *JAK-STAT* signalling pathway variants (25%)[9] are common alterations identified in T-ALL[12], including in *HOXA*-deregulated cases. It is difficult to establish the mutational landscape of *MLLT10*r T-ALL from existing published studies, as the *HOXA*-deregulated T-ALL subtype also includes other fusions such as *KMT2A*r and *SET-NUP214*. Further interrogation of the co-operative lesions present in *MLLT10*r T-ALL may assist in

developing a model of *MLLT10*r T-ALL for pre-clinical studies of disease biology and therapeutic sensitivities.

The role of aberrant epigenetic profiles has emerged as an important consideration in ALL[13]. DNA methylation profiling suggests that different genomic ALL subtypes possess overlapping but distinct DNA methylation signatures[14]. Therapies targeted towards the epigenome have subsequently emerged, including inhibitors of DOT1L[15], BET[16,17] and histone deacetylases (HDACs)[18], and DNA demethylating agents such as azacitidine. The identification of mutations in genes associated with epigenetic regulation are of interest, to understand the causal mechanisms of epigenetic dysregulation in ALL and predict sensitivity to epigenome-targeted therapies.

This chapter explores the genomic landscape of non-infant *KMT2A*r B-ALL and *MLLT10*r T-ALL at diagnosis. Two primary aims are addressed:

1. Identify clinically relevant SNVs and CNAs in genes associated with leukaemia in *KMT2A*r B-ALL and *MLLT10*r T-ALL, to identify potentially co-operative genomic variants that may influence leukaemogenesis.
2. Identify novel or known SNVs within genes that have a known epigenetic regulatory function, that may be associated with epigenetic dysregulation or response to therapies targeted towards the epigenome.

5.2 Approach

All methods are described in detail in section **2.15.** Patient MNCs were collected from peripheral blood or bone marrow and RNA was extracted. Multiplex-ligation dependent probe amplification (MLPA) was performed to screen for CNAs in a panel of ALL-associated genes, and mRNA-seq was performed and bioinformatically analysed for fusion transcripts and SNVs. Insertion/deletion mutations (INDELs) were also investigated for the T-ALL cohort using mRNA-seq, as this data was readily available. INDELs were not investigated for the B-ALL cohort due to extraneous limitations in data availability, but this was not considered a major limitation as clinically relevant small gene insertions or deletions are relatively rare in B-ALL at diagnosis, relative to SNVs and whole gene CNVs[19].

Selected genes were interrogated based on previously published studies, to identify SNVs (B-ALL and T-ALL) and INDELs (T-ALL only) in leukaemia-associated genes[9,11] (**Supplementary Table 5.1**) and SNVs within epigenetic regulators[20] (**Supplementary Table 5.2**). Bioinformatically filtered SNVs and INDELs were manually interrogated using the COSMIC[21] database to obtain the FATHMM[22] score (SNVs only) and establish whether alterations had been previously identified in cancer, and VarSome[23] to determine the American College of Medical Genetics (ACMG) classification of variant pathogenicity[24]. SNVs occurring in leukaemia-associated genes were included in analysis if they were classified as pathogenic or likely pathogenic by FATHMM or ACMG. Variants of uncertain significance were also included in the analysis of SNVs affecting epigenetic regulators, as these may be potentially relevant novel variants. FATHMM scores are not applicable for INDELs, so a pathogenic/likely pathogenic ACMG classification or previously published reports of the INDEL in T-ALL was required for inclusion in analysis.

An age-matched cohort of *BCR-ABL1* positive B-ALL was utilised for comparison with the *KMT2A*r B-ALL cohort, to identify trends in genomic alterations. *BCR-ABL1* was selected as the comparator genomic subtype as it is another high-risk B-ALL subtype that occurs in both children and adults.

5.3 Results

Cohort demographics are specified in **Table 5.1**. ALL-associated SNVs and CNAs were interrogated in *KMT2A*r and *BCR-ABL1* B-ALL (**Figures 5.1-5.2**), and the alterations present in each case are specified in **Supplementary Tables 5.3** (*KMT2A*r) and **5.4** (*BCR-ABL1*). At least one alteration was identified in 34.7% (8/23) of *KMT2A*r ALL cases, compared with 69.6% (16/23) of *BCR-ABL1* cases (p=0.03, **Figure 5.2**). The most common CNA in the *KMT2A*r cohort was *PAR1* deletion (15.0% or 3/20 for whom MLPA data were available). Heterozygous deletion of *ETV6, IKZF1* or *PAX5* occurred in 2 cases each (10.0%, 2/20). In the *BCR-ABL1* cohort, deletions of *CDKN2A/B* and *IKZF1* were common (50%, 11/22 and 55%, 12/22 respectively) but were rarely observed in *KMT2A*r patients (5.0%, 1/20 and 10.0%, 2/20 respectively).

Alterations identified in *MLLT10*r ALL are summarised in **Figure 5.3**, and the alterations present in each case are specified in **Supplementary Table 5.5**. At least one clinically relevant alteration was identified in all *MLLT10*r cases. *NOTCH1* alterations (71.4%, 5/7), frameshift and non-frameshift alterations of *PTEN* (42.9%, 3/7), *CDKN2A/B* deletion (33.3%, 2/6 with MLPA available) and *RAS* pathway alterations (42.9%, 2/7) were recurrent events. Comprising the five *NOTCH1* alterations were three SNVs, and one frameshift insertion and one non-frameshift insertion, that all result in *NOTCH1* activation. Four *PTEN* alterations were identified across three patients, where one patient (ID: AYAI0082) harboured both *PTEN* p.A418fs (frameshift insertion) and p.P421delinsLS (non-frameshift insertion).

Table 5.1 – Cohort demographics

Characteristic	*KMT2A*r B-ALL	*MLLT10r* ALL	*BCR-ABL1* B-ALL
N =	23	7	23
% female	61%	28.6%[^]	52%
Age Mean (range)	33.6 (11-74)*	20 (15-75)[x]	35.0 (8-71)

*Age unknown for 1 patient

[x]Age unknown for 4 patients

[^]T-ALL affects males more than females, so this percentage is expected[25]
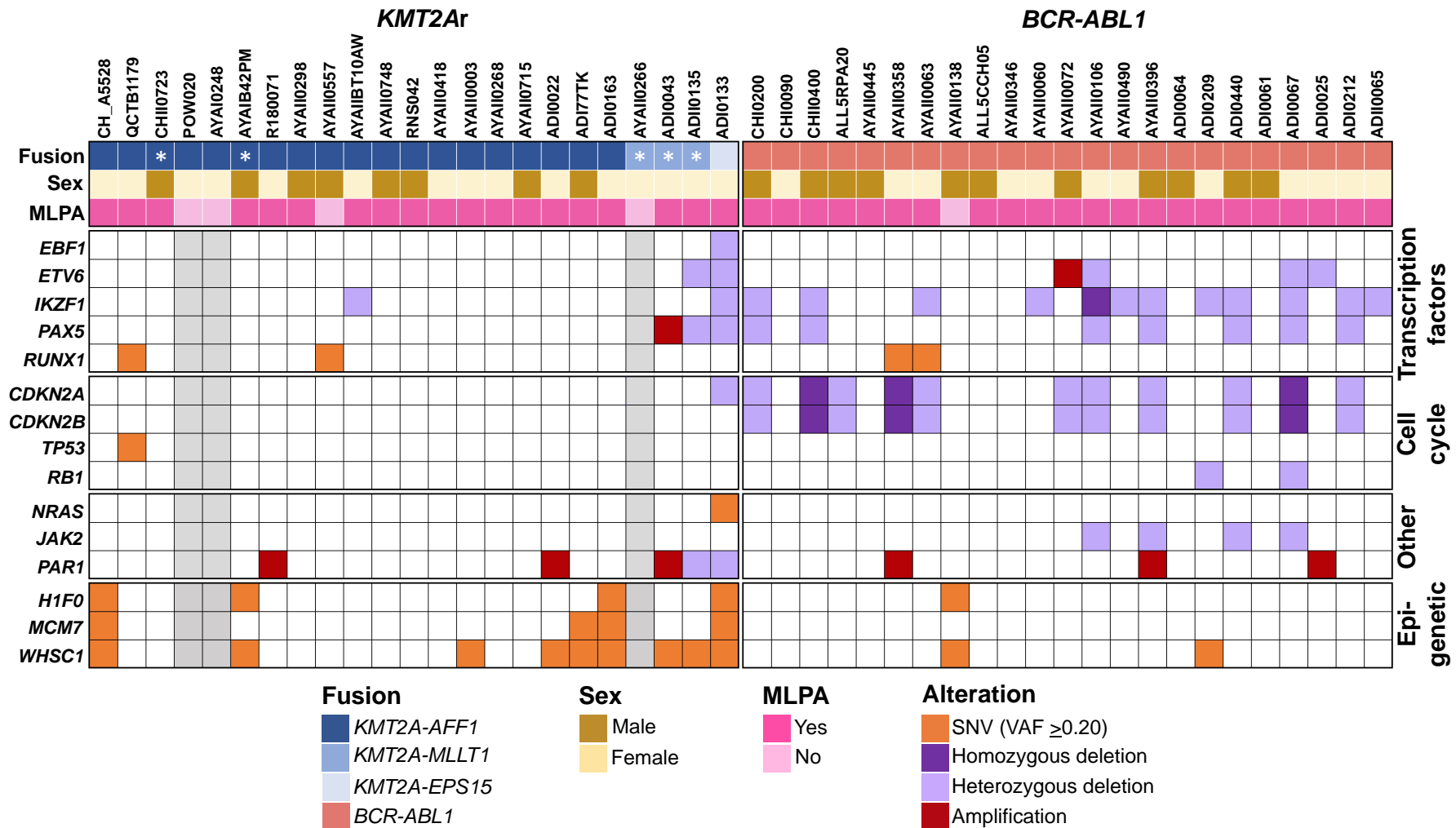
Figure 5.1 – Non-infant *KMT2A*r B-ALL harbours few leukaemia-associated SNVs and CNAs at diagnosis, in comparison to *BCR-ABL1* B-ALL

SNVs and CNAs were interrogated in the leukaemic blasts of patients diagnosed with *KMT2A*r or *BCR-ABL1* B-ALL. RNA was extracted from patient MNCs. Fusion genes and SNVs were identified within a panel of leukaemia-associated genes by mRNA-seq,

and MLPA was performed to detect CNAs. SNVs with a variant allele frequency (VAF) of ≥0.2 were included in analysis. Grey shading indicates samples that do not have MLPA data and no SNVs or INDELs were identified. *No in-frame reciprocal fusion (for example *AFF1-KMT2A*) was identified in these cases. A summary of variants identified in each case are provided in **Supplementary Tables 5.3-5.4**.
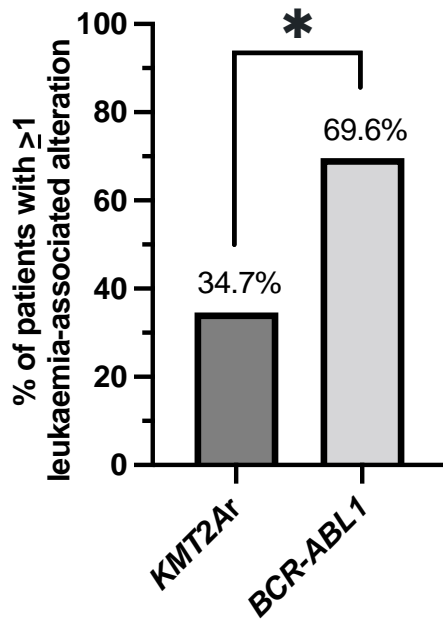
Figure 5.2 – Leukaemia-associated genomic alterations (CNAs and SNVs) are rare in *KMT2A*r B-ALL, relative to *BCR-ABL1* B-ALL leukaemic blasts

SNVs and CNAs were identified in the leukaemic blasts of patients diagnosed with *KMT2A*r or *BCR-ABL1* B-ALL by mRNA-seq (SNVs) and MLPA (CNAs). The number of patients with at least one pathogenic or likely pathogenic alteration was documented. Statistical significance was determined by Mann-Whitney U test (*p<0.05).
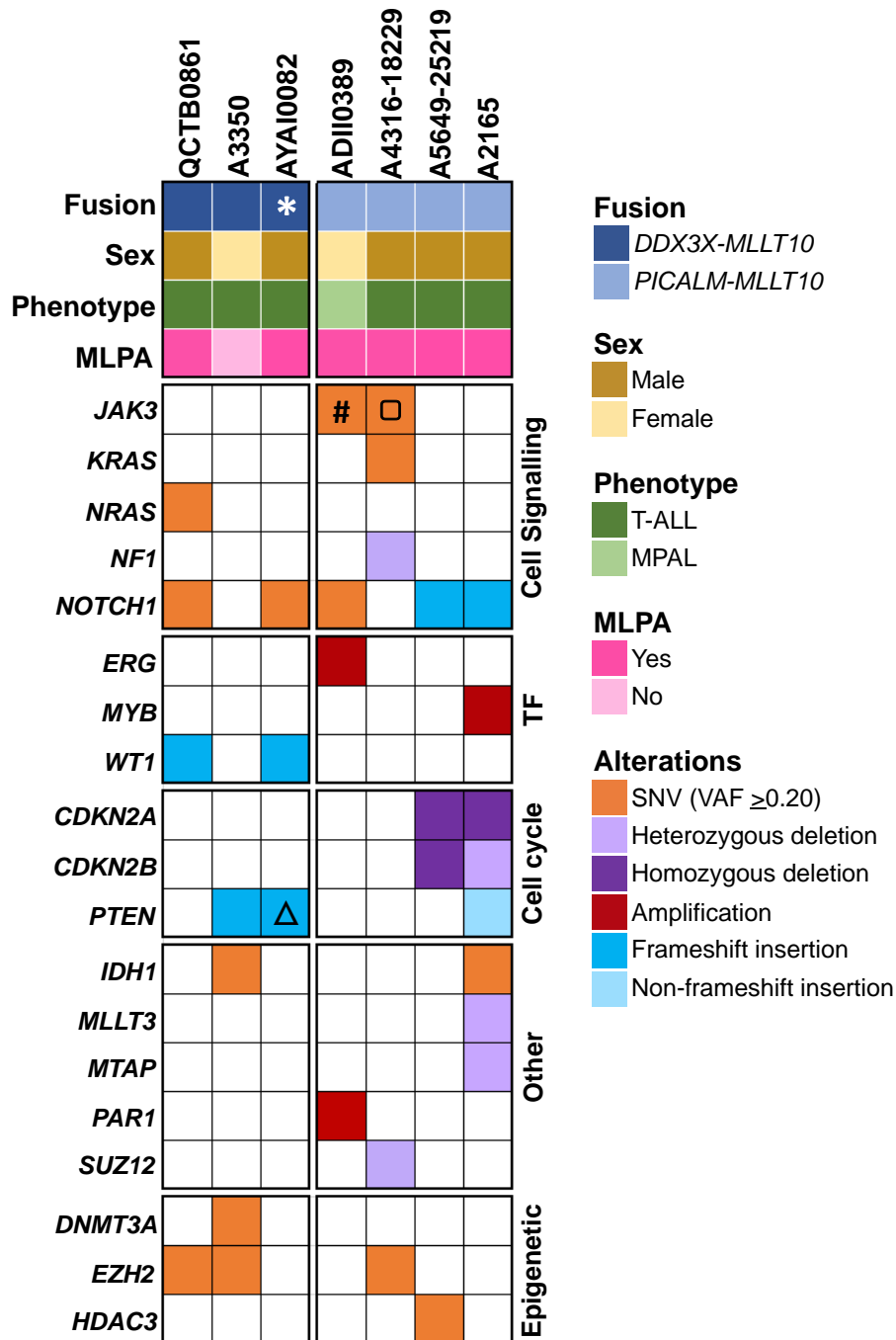
Figure 5.3 – Leukaemic blasts from patients diagnosed with *MLLT10*r acute leukaemia harbour co-operative genomic alterations at diagnosis

SNVs, INDELs and CNAs were identified in the leukaemic blasts of patients diagnosed with *MLLT10*r ALL. RNA was extracted from patient MNCs. Fusion genes, SNVs and INDELs were identified by mRNA-seq and MLPA was performed to detect CNAs. SNVs and INDELs with a VAF of ≥0.2 were included in analysis. A summary of variants identified in each case is provided in **Supplementary Table 5.5**. Abbreviations: MPAL, mixed phenotype acute leukaemia; TF, transcription factors. *Reciprocal fusion

(*MLLT10-DDX3X*) was not detected by mRNA-seq in this case. #VAF=0.16, but this variant was reported due to being a highly pathogenic SNV. □Two concurrent SNVs within *JAK3* were identified in this sample. △Both a frameshift and non-frameshift *PTEN* insertion were identified in this sample.

SNVs within epigenetic regulators were interrogated in the leukaemic blasts of patients with *KMT2A*r or *BCR-ABL1* B-ALL (**Figure 5.4** & **Supplementary Tables 5.3-5.4**). The FATHMM and ACMG scores for each identified variant are provided in **Supplementary Table 5.7**. *WHSC1* (also known as *NSD2* or *MMSET*) p.E1099K was identified in 39.1% (9/23) of *KMT2A*r and 8.7% (2/23) of *BCR-ABL1* cases (p=0.035). *H1F0* p.A89P and *MCM7* p.A433P were each identified in 17.4% (4/23) *KMT2A*r B-ALL cases, whereas only *H1F0* p.A89P was identified in one (4.3%) *BCR-ABL1* case (p=not significant). In the *MLLT10*r cohort, *DNMT3A* p.G298E and *HDAC3* p.H135Q were both identified once. Three patients each harboured a different *EZH2* variant (p.V679M, p.A255T and p.D185H) (**Figure 5.5**).
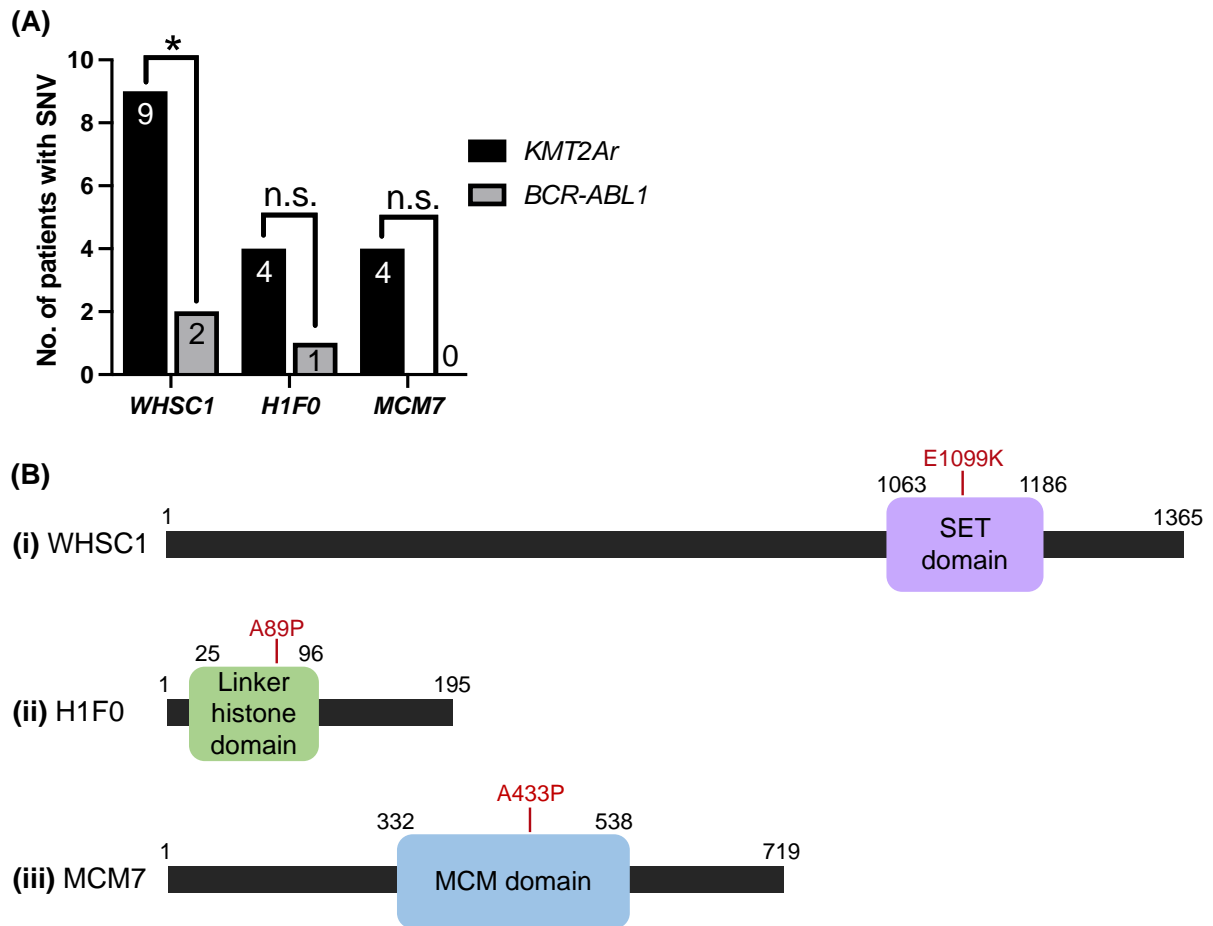
(A)

(B)

Figure 5.4 – Recurrent SNVs within epigenetic regulatory genes occur in leukaemic blasts from patients diagnosed with *KMT2A*r B-ALL

*KMT2A*r or *BCR-ABL1* B-ALL blasts were interrogated for SNVs within selected epigenetic regulatory genes. SNVs with a FATHMM or ACMG prediction of pathogenic, likely pathogenic or uncertain significance were included in analysis. **(A)** The frequency of SNV identification in each gene is provided, with the number of patients from either genomic subtype annotated above each bar. **(B)** Protein domain maps demonstrating the location of variants identified in **(i)** WHSC1, **(ii)** H1F0 and **(iii)** MCM7. Protein domains were mapped with amino acid coordinates provided by St Jude ProteinPaint software[26]. Statistical significance was determined by Mann-Whitney U test (*$p<0.05$). Abbreviations: MCM, minichromosome maintenance complex domain; SET, Su(var)3-9, Enhancer-of-zeste, Trithorax; *WHSC1*, Wolf-Hirschhorn Syndrome Candidate 1.
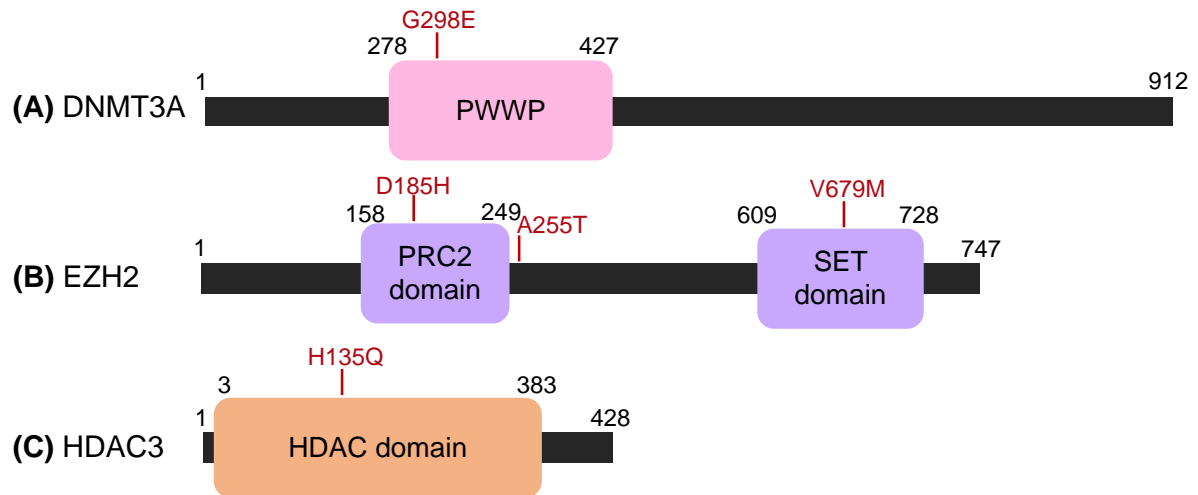
Figure 5.5 – Leukaemic blasts from patients diagnosed with *MLLT10*r T-ALL harbour SNVs within epigenetic regulatory genes

*MLLT10*r ALL leukaemic blasts were interrogated for SNVs within a panel of epigenetic regulatory genes. SNVs with an ACMG or FATHMM prediction of pathogenic, likely pathogenic or uncertain significance were included in analysis. SNVs were identified in **(A)** DNMT3A (n=1 patient), **(B)** EZH2 (n=3) and **(C)** HDAC3 (n=1). Protein domains were mapped with amino acid coordinates provided by St Jude ProteinPaint software[26]. Abbreviations: PWWP, proline-tryptophan-tryptophan-proline; SET, Su(var)3-9, Enhancer-of-zeste, Trithorax; HDAC, histone deacetylase; MCM, minichromosome maintenance complex; PRC2 domain, Polycomb repressive complex 2 tri-helical domain.

5.4 Discussion

This chapter interrogates the genomic landscape of genomic alterations present at diagnosis in non-infant *KMT2A*r B-ALL and *MLLT10*r T-ALL, two subgroups of ALL patients that are poorly characterised. The *KMT2A*r B-ALL genomic landscape was compared to an age-matched cohort of *BCR-ABL1* B-ALL cases. INDELs (T-ALL only), SNVs and CNAs within leukaemia-associated genes were identified, with each case summarised in **Supplementary Tables 5.3-5.5**. Evidence supporting the pathogenic or likely pathogenic status of each reported SNV is provided in **Supplementary Table 5.6**. SNVs within epigenetic regulatory genes were identified, with variant details provided in **Supplementary Table 5.7**.

5.4.1 Non-infant *KMT2A*r B-ALL blasts harbour few clinically relevant genomic aberrations at diagnosis

The low incidence of mutations in the *KMT2A*r B-ALL cohort supports the notion that *KMT2A*r are sufficiently pathogenic to induce leukaemia in the absence of co-operative mutations, as observed in published infant cohorts[4-7]. It was recently identified that a foetal-specific gene expression signature underlies infant B-ALL, which is distinct from non-infant, paediatric *KMT2A-AFF1* ALL[8]. This distinct foetal gene expression program is proposed to explain the short latency and absence of co-operative events observed in infant ALL[8]. The data presented in **Figure 5.1** establishes that non-infant *KMT2A*r B-ALL lacks recurrent leukaemia-associated CNAs and SNVs that are common in other subtypes of B-ALL, indicating a highly aggressive aetiology in *KMT2A*r patients of all ages. The *BCR-ABL1* comparator cohort more frequently harbours co-operative genomic events than *KMT2A*r, and in the majority of cases these alterations are CNAs. *BCR-ABL1* is a pathogenic fusion gene that is capable of transforming Ba/F3 cells to cytokine independence, suggesting that *BCR-ABL1* expression is sufficient to induce ALL alone[27]. Indeed, no co-operative genomic events were identified in 30.4% (7/23) of *BCR-ABL1* cases (**Figure 5.1**). While *KMT2A-AFF1* does not induce Ba/F3 cytokine independence, its association with high-risk ALL is sufficient evidence to demonstrate its pathogenicity. The finding that *KMT2A*r cases harbour cooperative genomic events even less frequently than *BCR-ABL1* cases demonstrates that *KMT2A*r B-ALL is a uniquely aggressive entity.

Three or more leukaemia-associated aberrations were identified in three *KMT2A*r cases. One case (ID: ADI0133) harboured the rare *KMT2A-EPS15* fusion in addition to multiple alterations: *NRAS* p.Q61K (VAF=0.30) and heterozygous deletion of *EBF1, ETV6, IKZF1, PAX5, CDKN2A* and *PAR1*. Two *KMT2A-MLLT1* cases also possessed multiple alterations: ADI0043 harboured amplification of *PAX5* and the *PAR1* region, and ADII00135 harboured deletions of *ETV6, PAX5* and *PAR1*. Multiple alterations were not identified in any *KMT2A-AFF1* case, suggesting that rare *KMT2A*r such as *KMT2A-EPS15* and *KMT2A-MLLT1* require additional genomic events to induce leukaemia, whereas *KMT2A-AFF1* does not, but further studies on larger cohorts are required. This theory is supported by one T-ALL study, where 2/4 *KMT2A-MLLT4* and 2/2 *KMT2A-MLLT1* T-ALL patients had additional alterations of clinical relevance, most commonly *NOTCH1* and *EZH2* variants[12].

Other than case ADI0133, no *RAS* pathway variants were identified in the B-ALL cohort (total incidence in *KMT2A*r B-ALL cohort = 4.3%). This is consistent with one study that identified *K/NRAS* SNVs in 8% of adult *KMT2A*r ALL cases[28], rarer than the incidence in infant *KMT2A*r ALL (approximately 50%)[29-32]. *NRAS* p.Q61K is a clinically relevant, pathogenic SNV[33] that confers sensitivity of *KMT2A*r ALL to MEK inhibitors such as selumetinib in pre-clinical studies[34]. A clinical trial is currently recruiting patients to evaluate the role of selumetinib in relapsed/refractory *RAS*-mutated paediatric and adult ALL (NCT03705507, refer to **Tables 1.1-1.2**).

The *RUNX1* p.L56S single nucleotide variant (SNV) was identified in two *KMT2A*r patients. The p.L56S variant is located within the *KMT2A*-interacting domain[35], but it is currently unclear whether L56S is a benign single nucleotide polymorphism (SNP) or leukaemia-predisposing variant. One study reports an increased incidence of the L56S variant in MDS/AML (incidence of 4.9% in MDS/AML patients, compared with 1.7% the general European population, not statistically analysed)[35]. However, an independent study did not establish any significant differences in p.L56S incidence between AML patients and healthy controls[36]. The p.L56S variant impairs recruitment of KMT2A by RUNX1 to target gene promoter regions in a *KMT2A* wild-type human erythroleukaemia cell line[37] but it is not certain how this process is affected when a *KMT2A*r is present. It would be interesting to co-express *KMT2A*r and *RUNX1* p.L56S to explore the functional significance of this variant in a *KMT2A*r context but this examination is outside the scope of this study.

Overall, this study establishes that ALL-associated SNVs and CNAs are uncommon in non-infant *KMT2A*r B-ALL, indicating a low mutational burden similar to that observed in infant *KMT2A*r ALL. This supports the notion that *KMT2A*r are sufficiently pathogenic to induce B-ALL without the requirement of co-operative genomic alterations, demonstrating an aggressive aetiology regardless of patient age. A clinically actionable variant was identified in one *KMT2A*r patient, the *NRAS* p.Q61K variant that confers sensitivity to MEK inhibition.

5.4.2   Clinically relevant co-operative mutations are present in *MLLT10*r T-ALL blasts at diagnosis

NOTCH1, PTEN and *CDKN2A/B* alterations were identified in 71.4%, 42.9% and 33.3% of patients respectively (**Figure 5.3**). The incidence of *NOTCH1* alterations in the present study is consistent with other publications describing T-ALL cohorts. Activating *NOTCH1* signalling is one of the most prevalent genomic events in T-ALL, reported in approximately 70% of new T-ALL diagnoses overall. *PTEN* is a tumour suppressor gene that is inactivated in various malignancies, including in 14-23% of T-ALL cases[9,11]. *CDKN2A/B* deletions appear to be underrepresented in the present study, previously reported to occur in >70% of paediatric and adolescent T-ALLs[9,38] and 30% of adult T-ALLs[39]. The age of four *MLLT10*r patients described in this chapter are unknown, but it is possible that the low incidence of *CDKN2A/B* deletion could be explained if the patients are adults. Interestingly, existing data suggests that *CDKN2A* deletions are very rare in *MLLT10*r T-ALL. In one study, no *CDKN2A* deletions were identified in 23 *PICALM-MLLT10* T-ALL patients by Southern blot analysis[40]. In an adult T-ALL cohort, 11.1% (1/9) of *HOXA* deregulated cases harboured a *CDKN2A/B* deletion, compared with 28.3% (15/53) of the cohort overall[39].

RAS pathway alterations were identified in 42.9% of *MLLT10*r cases. *NRAS* p.G12V and *KRAS* p.G12A were identified in one *DDX3X-MLLT10* and *PICALM-MLLT10* case respectively. *K/NRAS* G12V and G12A variants are canonical hotspot variants that occur in a GTP-binding region, that result in K/NRAS inhibition and subsequent downstream MEK pathway activation. *K/NRAS* p.G12 variants confer pre-clinical sensitivity to MEK inhibition in B-ALL, indicating this as a potential therapeutic avenue for these patients[34,41]. An additional patient had heterozygous inactivation of *NF1*,

which also confers sensitivity to MEK inhibitors[42,43]. T-ALL patients are eligible for participation in a currently recruiting trial assessing selumetinib in relapsed/refractory *RAS*-mutated ALL (NCT03705507, refer to **Tables 1.1-1.2**), which will provide insight into the clinical benefit of MEK inhibition in ALL.

The *JAK3* p.M511I and p.R657Q activating SNVs were identified in one *PICALM-MLLT10* case (ID: A4316-18229), and ADI0389 had p.M511I with a low VAF (0.16). *JAK3* mutations are present in 15-20% of T-ALL diagnoses[9,44], typically the pathogenic p.M511I and p.A573V variants[45]. R657Q is comparatively rare, previously reported in isolated T-ALL incidences[44,46], and is an activating variant as it confers STAT5 activation and cytokine independence in Ba/F3 cells[47]. Ectopic co-expression of *HOXA9* and *JAK3* p.M511I induces leukaemia *in vivo* with a shorter latency than either *HOXA9* or *JAK3* p.M511I expression alone[48]. As *MLLT10*r induces *HOXA* dysregulation, there may be a co-operative advantage conferred by *MLLT10*r and *JAK3* activation. The therapeutic relevance of *JAK3* inhibitors in ALL should be further investigated, but it is important to note the presence of activating mutations does not necessarily confer clinically significant inhibitor sensitivity, particularly in cases where VAF is low (<0.2). One case report describes a T-ALL patient with *DDX3X-MLLT10* and two *JAK3* variants (p.R657W and p.M511I), who failed to respond to the *JAK1/JAK3* inhibitor tofacitinib[44].

In addition to *JAK3* p.R657Q and p.M511I, patient A4316-18229 also harbours heterozygous deletion of *SUZ12*, a critical component of the Polycomb Repressor Complex 2 (PRC2) that epigenetically regulates transcriptional repression[49,50]. Inactivation of *SUZ12* cooperates with *JAK3* p.M511I to induce T-ALL through activation of PI3K/mTOR, VEGF and WNT signalling pathways, inducing an aggressive leukaemia *in vivo*[49]. This leukaemia is sensitive to inhibitors of PI3K/mTOR, VEGF receptor, HDAC and HSP90, indicating novel pre-clinical therapeutic targets for patients with co-occurring *SUZ12* inactivating and *JAK3* activating alterations[49]. This study also reported that PRC2 mutations, including inactivating mutations of *SUZ12*, were disproportionately common in *HOXA* mutated cases, relative to T-ALL subtypes[49]. Together, these data suggest that *JAK3* activation and/or PRC2 complex dysregulation may cooperate with *HOXA* dysregulation to induce T-ALL in some cases, identifying a potential opportunity for novel therapeutic approaches in these patients.

In summary, clinically relevant alterations were identified in all *MLLT10*r cases, including five *NOTCH1* alterations, four *PTEN* frameshift and non-frameshift insertions, three activating *RAS* pathway alterations, two *CDKN2A/B* deletions, and two *JAK3* activating variants. The *RAS* and *JAK3* variants may be actionable with clinically available inhibitors such as selumetinib (MEK inhibitor) or tofacitinib (JAK3 inhibitor), but further pre-clinical and clinical trial data of inhibitor efficacy is required in a T-ALL setting. Overall, this suggests that alterations affecting cell cycle and cell signalling contribute to *MLLT10*r-mediated leukaemogenesis, which should be considered when generating pre-clinical models of *MLLT10*r T-ALL.

5.4.3 Known and novel SNVs in epigenetic regulatory genes were identified in non-infant *KMT2A*r B-ALL

Recurrent SNVs within epigenetic regulatory genes were identified in *KMT2A*r B-ALL, that were uncommon or absent in the *BCR-ABL1* cohort. *WHSC1* p.E1099K was identified in nine *KMT2A*r B-ALL cases, compared to two *BCR-ABL1* cases (p=0.035, **Figure 5.4**). *WHSC1* is a histone lysine methyltransferase[51], and the E1099K variant is frequently reported in lymphoid malignancies including 5-10% of paediatric B-ALL cases[51-53]. The p.E1099K variant functionally disrupts the SET domain of WHSC1[53] (**Figure 5.4A**), and the presence of p.E1099K in human ALL cell lines is associated with enhanced H3K36me2 activity and aberrant global chromatin methylation profiles[54]. Induction of heterozygous *WHSC1* p.E1099K by CRISPR/Cas9 into SEM (*KMT2A*r) and RCH-ACV (*KMT2A*-WT) B-ALL cell lines resulted in shortened disease latency and increased incidence of brain infiltration in murine xenografts, compared with wild-type cells[52]. These findings indicate that *WHSC1* may serve as an important genomic event in some cases of B-ALL. No direct inhibitors of *WHSC1* are clinically available, but sinefungin derivatives have demonstrated preliminary promise as inhibitors of *WHSC1* function[55]. As the p.E1099K variant is disproportionately common in the *KMT2A*r cohort compared with the *BCR-ABL1* cohort, the co-operative potential of *KMT2A*r and *WHSC1* p.E1099K should be investigated, particularly as they are both lysine methyltransferases that induce widespread changes in gene expression profiles[51,56].

H1F0 p.A89P was more common in *KMT2A*r than *BCR-ABL1* cases (17.4%, 4/23 vs 4.3%, 1/23 patients). *H1F0* encodes the H1.0 linker histone that is essential in nucleosome-DNA stabilisation[57]. H1.0 deletion results in transcriptional upregulation of gene programs associated with enhanced cell self-renewal, such as PRC2 components[57]. The functional effect of the p.A89P variant on H1.0 function is currently unknown but is bioinformatically classified as deleterious. A89 is a highly conserved residue that occurs within the histone linker domain of H1.0 (**Figure 5.4B**), and the amino acid substitution of non-polar alanine to polar proline likely induces significant alteration in protein structure and function.

MCM7 p.A433P was identified in 17.4% (4/23 patients) *KMT2A*r, but not in any *BCR-ABL1* cases. *MCM7* is a critical component of the mini-chromosome maintenance (MCM) complex, the putative replicative helicase complex that ensures 'once per cycle' DNA replication[58]. MCM proteins 2-7 are highly conserved, and ectopic upregulation results in enhanced tumour development in mice[58]. The p.A433P variant is located within the functional MCM domain of MCM7 (**Figure 5.4C**), and the amino acid change of non-polar alanine to polar proline likely induces a change in protein folding and subsequently impact protein function. However, no further functional understanding of this variant is currently available.

In summary, potentially clinically relevant SNVs were recurrently identified in the *KMT2A*r B-ALL cohort. The pathogenic *WHSC1* p.E1099K variant, and novel variants *MCM7* p.A433P and *H1F0* p.A89P were identified more frequently in *KMT2A*r than in *BCR-ABL1* B-ALL. Further investigation of these variants in *KMT2A*r ALL is required to establish their role in leukaemic development and therapeutic sensitivity.

5.4.4 SNVs in epigenetic regulatory genes were identified in *MLLT10*r ALL

Five potentially relevant SNVs within epigenetic regulators were identified in the leukaemic blasts of four *MLLT10*r cases. *DNMT3A* is a DNA methyltransferase with important roles in haematopoietic cell differentiation. The p.G298E variant occurred in one *MLLT10*r patient and is likely pathogenic, as it resides in the regulatory PWWP/catalytic methyltransferase domain of DNMT3A[59] (**Figure 5.5A**). *DNMT3A* SNVs are identified in adults with AML, T-ALL and MPAL with an incidence of approximately 20-30%. The presence of a variant at *DNMT3A* p.R882 is associated

with significantly shorter overall survival in AML[60]. *DNMT3A* p.R882 variants also sensitise AML cells to the DNA methyltransferase inhibitor azacitidine[61], which is currently in ALL clinical trials (**Table 1.2**). However, the prognostic or functional relevance of *DNMT3A* p.G298E is uncertain[59,62].

EZH2 encodes the PRC2 protein EZH2[63]. The p.V679M variant is considered a pathogenic variant as it occurs within the highly conserved SET domain, responsible for histone-lysine methyltransferase activity[64] (**Figure 5.5B**). SET domain gain-of-function variants occur in 7-22% of B-cell lymphomas, including one report of a sub clonal V679M variant[64]. The p.D185H variant occurs within the PRC2 domain that is associated with interaction with members of the PRC2 complex and subsequent epigenetic silencing[65] (**Figure 5.5B**). The p.D185H variant is often observed as a germline variant and has been associated with an increased risk of myeloid malignancies[66]. The p.A255T variant resides close to the PRC2 domain (**Figure 5.5B**). No functional studies or published reports of p.A255T in cancer have been identified, warranting further investigation. *EZH2* alterations have been identified in isolated instances of T-ALL[45], including in one study that reports *EZH2* variants in 4/15 (27%) of *PICALM-MLLT10* T-ALL/AML cases, compared with 0/12 non-*PICALM-MLLT10* cases[67]. *EZH2* inhibitors such as tazometostat are currently in clinical trials for a range of malignancies including B-cell lymphomas (NCT03456726), and efficacy of such inhibitors in ALL should be tested[68].

HDAC3 p.H135Q was identified in one *MLLT10*r case (**Figure 5.5C**). *HDAC3* encodes a histone deacetylase that is overexpressed in a range of cancers[69]. *HDAC3* p.H135 is a highly conserved residue that resides near the Zn ion binding site, and the H135A substitution completely inactivates HDAC3 function in HEK283T cells[70]. While no functional studies have been published on the H135Q variant, it is likely that the substitution of histidine (polar charged/basic) with glutamine (polar neutral) would induce a change in protein structure and/or Zn ion binding affinity. HDAC inhibitors such as panobinostat and vorinostat have pre-clinical efficacy across a range of solid tumours and haematological malignancies[69,71], and functional studies investigating the role of HDAC p.H135 variants in HDAC activity and sensitivity to HDAC inhibitors would be clinically useful.

Interestingly, *EZH2* p.V679M and *DNMT3A* p.G298E occurred in the same patient (ID: A3350) (**Figure 5.3**). *EZH2* interacts with *DNMT3A* to regulate promoter methylation of *EZH2* target genes (such as *MYT1, WNT1* and *CRN1*)[72]. The known interaction of *EZH2* and *DNMT3A* raises the possibility that dysregulation of both these targets may induce a co-operative leukaemic effect. Further functional studies on *EZH2* p.V679M and *DNMT3A* p.G298E should therefore be performed. It would also be insightful to examine *EZH2* target gene expression levels in the leukaemic blasts of patients that possess *EZH2* p.V679M or *DNMT3A* p.G298E, to establish whether these variants have functional consequences on target gene expression.

In summary, 57.1% (4/7) *MLLT10*r cases harboured an SNV of interest within an epigenetic regulatory gene. The presence of a mutation in *DNMT3A, SUZ12* or *EZH2* has been previously shown to confer inferior overall survival in ETP-ALL[73], indicating that the role of these variants in T-ALL leukaemogenesis, and the viability of therapeutic targeting, should be further investigated.

5.4.5 Limitations

1. INDELs were not investigated in the B-ALL cohort

INDELs were not analysed in the B-ALL cohort due to extraneous limitations in data availability. This was not considered a major limitation as clinically relevant small gene insertions or deletions are relatively rare in B-ALL at diagnosis, relative to SNVs and whole gene CNVs[19]. However, it is likely that some rare instances of INDELs occurred in the B-ALL cohort, such as in the tumour suppressor gene *TP53*. In one study of *TP53* alterations in ALL, 18.2% of identified *TP53* mutations were alterations other than missense SNVs, including frameshift INDELs (7.2%), non-frameshift INDELs (4.5%), splice-site alterations (2.7%), and nonsense mutations (3.6%)[74]. It is therefore likely that the frequency of clinically relevant variants in tumour suppressor genes such as *TP53* is underreported here.

2. Unable to differentiate germline and somatic variants

A matched remission or otherwise non-leukaemic sample was not available for differentiation of germline and somatic variants. This issue was overcome in one study

by using a "likely germline" approach, where germline lesions were predicted based on a VAF of $\geq$0.44, after adjusting for chromosome copy number[75]. A similar analysis could be conducted here to allow for identification of germline variants.

5.4.6 Concluding remarks

The genomic landscapes of non-infant *KMT2A*r B-ALL and *MLLT10*r T-ALL were interrogated for known leukaemia-associated variants and SNVs within epigenetic regulatory genes. In chapters 3 and 4, expression of *DDX3X-MLLT10* or *PICALM-MLLT10* in the MOHITO cell line (**chapter 3**) or primary Arf[-/-] thymocytes (**chapter 4**) did not induce *HOXA* deregulation, which is a characteristic feature of *MLLT10*r ALL. Furthermore, expression of *KMT2A-AFF1* in Ba/F3 cells (**chapter 3**) did not induce a phenotype reflective of the aggressive B-ALL observed in patients. These findings raised the possibility that co-operative lesions may be involved in *MLLT10*r or *KMT2A*r leukaemogenesis, and investigation of co-operative lesions in these subtypes was conducted to inform disease aetiology and pre-clinical modelling design and cell type.

Leukaemia-associated co-operative mutations were identified in significantly fewer cases of *KMT2A*r B-ALL when compared with *BCR-ABL1*, another high-risk genomic subtype of B-ALL (34.7%, 8/23 vs 69.6%, 16/23 **Figures 5.1-5.2**). Despite a low incidence of leukaemia-associated alterations, recurrent SNVs within epigenetic regulators were identified in *KMT2A*r patients (**Figure 5.4**). The known pathogenic *WHSC1* p.E1099K SNV occurred in 39.1% of *KMT2A*r patients (vs 4.3% of *BCR-ABL1*). Two potentially relevant novel variants were also identified: *H1F0* p.A89P (17.4% *KMT2A*r vs 4.3% *BCR-ABL1*) and *MCM7* p.A433P (17.4% *KMT2A*r vs 0% *BCR-ABL1*). Further functional studies are required to assess the functional significance of these variants in cooperation with *KMT2A*r. Induction of heterozygous SNVs by CRISPR/Cas9 into human *KMT2A*r cell lines such as SEM and RS4;11 and subsequent *in vivo* transplantation would provide insight into the influence of these variants on disease latency and dissemination, to establish the role of these variants in leukaemogenesis[52].

Co-operative leukaemia-associated alterations were identified in all seven cases of *MLLT10*r ALL (**Figure 5.3**), which is a substantially higher frequency than in the *KMT2A*r and *BCR-ABL1* B-ALL cohorts investigated. *NOTCH1* alterations (71.4% of

patients), frameshift and non-frameshift insertions affecting *PTEN* (42.9%), *RAS* pathway activating alterations (42.9%) and *CDKN2A/B* deletion (33.3%) were recurrent events in the *MLLT10*r cohort. These results suggest that co-operative variants, particularly those involved in cell signalling and cell cycle regulation, are relevant genomic events in *MLLT10*r ALL, and this must be considered when generating pre-clinical models of *MLLT10*r ALL. In addition to these known leukaemia-associated co-operative events, 4/7 *MLLT10*r cases harboured at least one potentially relevant SNV within an epigenetic regulatory gene, including three patients with *EZH2* variants. Further investigation on the incidence of identified variants in other subtypes of T-ALL, and their role in T-ALL leukaemogenesis, is warranted.

These results demonstrate that:

a. Similar to infant *KMT2A*r B-ALL, non-infant *KMT2A*r B-ALL has a low incidence of leukaemia-associated co-operative variants, when compared to *BCR-ABL1* B-ALL cases

b. Variants within epigenetic regulatory genes *WHSC1*, *H1F0* and *MCM7* are recurrent events in non-infant *KMT2A*r B-ALL

c. Co-operative genomic events are common in *MLLT10*r T-ALL, particularly *NOTCH1* and *RAS* pathway activating alterations, *PTEN* frameshift and non-frameshift insertions, and *CDKN2A/B* deletions

d. Variants within epigenetic regulatory genes occur in *MLLT10*r T-ALL, including *DNMT3A*, *EZH2* and *HDAC3*, but the clinical relevance of these variants is currently unknown

5.5 Supplementary Materials

Supplementary Table 5.1 – List of genes commonly altered in ALL[9,11]

ABL1	*ERG*	*KRAS*	*PTPN11*
ABL2	*ETV6*	*LEF1*	*RB1*
BCL11B	*FBXW7*	*LMO1*	*RPL10*
BCOR	*FLT3*	*LMO2*	*RUNX1*
BRAF	*GATA2*	*MTOR*	*STAT5B*
CBL	*GATA3*	*NF1*	*SH2B3*
CDKN2A	*IDH1*	*NOTCH1*	*STIL*
CDKN2B	*IDH2*	*NOTCH2*	*TP53*
CNOT3	*IKZF1*	*NRAS*	*WT1*
CRLF2	*IL7R*	*PAR1*	
CSF1R	*JAK1*	*PAX5*	
EBF1	*JAK2*	*PHF6*	
EPOR	*JAK3*	*PTEN*	

Supplementary Table 5.2 – List of genes used to interrogate variants within epigenetic regulators[20]

Gene	Function	Gene	Function	Gene	Function
ACTA1	remodel	*BMI1*	writer	*CECR2*	reader
ACTB	remodel	*BPTF*	bind_eraser	*CENPA*	histone
ACTL6A	remodel	*BRCA1*	bind_writer	*CENPB*	remodel
ACTL6B	remodel	*BRCA2*	bind_writer	*CENPC1*	remodel
ACTR3B	remodel	*BRCC3*	writer	*CENPE*	remodel
ACTR5	remodel	*BRD1*	reader	*CENPF*	remodel
ACTR6	reader	*BRD2*	reader	*CENPH*	remodel
ACTR8	remodel	*BRD3*	reader	*CENPI*	remodel
AEBP2	bind_writer	*BRD4*	reader	*CENPJ*	remodel
AIRE	reader	*BRD7*	reader	*CENPK*	remodel
AP2A1	bind_writer	*BRD8*	reader	*CENPL*	remodel
APBB1	reader	*BRD9*	reader	*CENPM*	remodel
ARID1A	remodel	*BRDT*	reader	*CENPN*	remodel
ARID1B	remodel	*BRE*	writer	*CENPO*	remodel
ARID2	remodel	*BRPF1*	reader	*CENPP*	remodel
ASCL2	reader	*BRPF3*	reader	*CENPQ*	remodel
ASF1A	reader	*BRWD1*	reader	*CENPT*	remodel
ASF1B	reader	*BRWD3*	reader	*CENPV*	remodel
ASH1L	writer	*C14orf169*	eraser	*CENPW*	remodel
ASH2L	bind_writer	*C16orf53*	bind_writer	*CHAF1A*	dna_mod
ASXL1	bind_eraser	*C20orf20*	reader	*CHD1*	reader
ASXL2	bind_eraser	*CARM1*	writer	*CHD1L*	reader
ASXL3	bind_eraser	*CBX1*	reader	*CHD2*	reader
ATAD2	reader	*CBX2*	bind_writer	*CHD3*	bind_eraser
ATAD2B	reader	*CBX3*	reader	*CHD4*	bind_eraser
ATAT1	writer	*CBX4*	bind_writer	*CHD5*	reader
ATF2	writer	*CBX5*	reader	*CHD6*	reader
ATF7IP	bind_writer	*CBX6*	bind_writer	*CHD7*	remodel
ATF7IP2	bind_writer	*CBX7*	bind_writer	*CHD8*	reader
ATM	writer	*CBX8*	bind_writer	*CHRAC1*	remodel
ATR	writer	*CCDC101*	reader	*CHUK*	writer
ATRX	reader	*CDC6*	dna_mod	*CLOCK*	writer
ATXN7	bind_writer	*CDK17*	writer	*CREBBP*	writer
ATXN7L3	bind_writer	*CDK3*	writer	*CSNK2A1*	writer
AURKB	writer	*CDK5*	writer	*CXXC1*	bind_writer
BABAM1	reader	*CDY1*	writer	*DAPK3*	writer
BAP1	bind_writer	*CDY1B*	writer	*DAXX*	bind_writer
BAZ1A	reader	*CDY2A*	writer	*DMAP1*	dna_mod
BAZ1B	writer	*CDY2B*	writer	*DNA2*	dna_mod
BAZ2A	reader	*CDYL*	writer	*DNMT1*	dna_mod
BAZ2B	reader	*CDYL2*	writer	*DNMT3A*	dna_mod

Gene	Function	Gene	Function	Gene	Function
DNMT3B	dna_mod	*GFI1B*	bind_eraser	*HIST1H1B*	histone
DNMT3L	dna_mod	*GLYR1*	reader	*HIST1H1C*	histone
DOT1L	writer	*GSG2*	writer	*HIST1H1D*	histone
DPF1	remodel	*GTF3C4*	writer	*HIST1H1E*	histone
DPF2	remodel	*H1F0*	histone	*HIST1H1T*	histone
DPF3	remodel	*H1FNT*	histone	*HIST1H2AA*	histone
DPY30	bind_writer	*H1FOO*	histone	*HIST1H2AB*	histone
DSN1	remodel	*H1FX*	histone	*HIST1H2AC*	histone
DTX3L	writer	*H2AFB1*	histone	*HIST1H2AD*	histone
EED	bind_writer	*H2AFB2*	histone	*HIST1H2AE*	histone
EHMT1	writer	*H2AFB3*	histone	*HIST1H2AG*	histone
EHMT2	writer	*H2AFJ*	histone	*HIST1H2AH*	histone
ELP3	writer	*H2AFV*	histone	*HIST1H2AI*	histone
ENY2	bind_writer	*H2AFX*	histone	*HIST1H2AJ*	histone
EP300	writer	*H2AFY*	histone	*HIST1H2AK*	histone
EP400	bind_writer	*H2AFY2*	histone	*HIST1H2AL*	histone
EPC1	bind_writer	*H2AFZ*	histone	*HIST1H2AM*	histone
EPC2	remodel	*H2BFM*	histone	*HIST1H2BA*	histone
ERCC6	dna_mod	*H2BFWT*	histone	*HIST1H2BB*	histone
ESCO1	dna_mod	*H3F3A*	histone	*HIST1H2BC*	histone
ESCO2	dna_mod	*H3F3B*	histone	*HIST1H2BD*	histone
EYA1	eraser	*H3F3C*	histone	*HIST1H2BE*	histone
EYA2	eraser	*HAT1*	writer	*HIST1H2BF*	histone
EZH1	writer	*HCFC1*	bind_writer	*HIST1H2BG*	histone
EZH2	writer	*HDAC1*	eraser	*HIST1H2BH*	histone
FAM175A	reader	*HDAC10*	eraser	*HIST1H2BI*	histone
FAM175B	bind_writer	*HDAC11*	eraser	*HIST1H2BJ*	histone
FBXL19	bind_writer	*HDAC2*	eraser	*HIST1H2BK*	histone
FEN1	dna_mod	*HDAC3*	eraser	*HIST1H2BL*	histone
FOXA1	dna_mod	*HDAC4*	eraser	*HIST1H2BM*	histone
FOXA2	dna_mod	*HDAC5*	eraser	*HIST1H2BN*	histone
FOXA3	dna_mod	*HDAC6*	eraser	*HIST1H2BO*	histone
FOXO1	dna_mod	*HDAC7*	eraser	*HIST1H3A*	histone
FOXO3	dna_mod	*HDAC8*	eraser	*HIST1H3B*	histone
FOXO4	dna_mod	*HDAC9*	eraser	*HIST1H3C*	histone
FOXP1	dna_mod	*HDGF*	dna_mod	*HIST1H3D*	histone
FOXP2	dna_mod	*HDGFL1*	dna_mod	*HIST1H3E*	histone
FOXP3	dna_mod	*HDGFRP2*	dna_mod	*HIST1H3F*	histone
FOXP4	dna_mod	*HDGFRP3*	dna_mod	*HIST1H3G*	histone
GATAD2A	reader	*HIRA*	reader	*HIST1H3H*	histone
GATAD2B	reader	*HIST1H1A*	histone	*HIST1H3I*	histone

Gene	Function	Gene	Function	Gene	Function
HIST1H3J	histone	*ING4*	reader	*KMT2C*	writer
HIST1H4A	histone	*ING5*	reader	*KMT2D*	writer
HIST1H4B	histone	*INO80*	remodel	*KMT2E*	writer
HIST1H4C	histone	*INO80B*	remodel	*KMT2F*	writer
HIST1H4D	histone	*INO80C*	remodel	*KMT2G*	writer
HIST1H4E	histone	*INO80D*	remodel	*KNTC1*	remodel
HIST1H4F	histone	*INO80E*	remodel	*L3MBTL1*	bind_writer
HIST1H4G	histone	*INTS12*	dna_mod	*L3MBTL2*	bind_writer
HIST1H4H	histone	*JAK2*	writer	*L3MBTL3*	bind_writer
HIST1H4I	histone	*JARID2*	bind_writer	*L3MBTL4*	bind_writer
HIST1H4J	histone	*JHDM1D*	eraser	*LMNA*	dna_mod
HIST1H4K	histone	*JMJD1C*	eraser	*LMNB1*	dna_mod
HIST1H4L	histone	*JMJD5*	eraser	*LMNB2*	dna_mod
HIST2H2AA3	histone	*JMJD6*	eraser	*LRWD1*	reader
HIST2H2AA4	histone	*KAT2A*	writer	*MAP3K12*	writer
HIST2H2AB	histone	*KAT2B*	writer	*MAZ*	bind_eraser
HIST2H2AC	histone	*KAT5*	writer	*MBD1*	bind_writer
HIST2H2BE	histone	*KAT6A*	writer	*MBD2*	reader
HIST2H3A	histone	*KAT6B*	writer	*MBD3*	dna_mod
HIST2H3C	histone	*KAT7*	writer	*MBD4*	dna_mod
HIST2H3D	histone	*KAT8*	writer	*MBD5*	dna_mod
HIST2H4A	histone	*KDM1A*	eraser	*MBD6*	dna_mod
HIST2H4B	histone	*KDM1B*	eraser	*MBTD1*	reader
HIST3H2A	histone	*KDM2A*	eraser	*MCM2*	dna_mod
HIST3H2BB	histone	*KDM2B*	eraser	*MCM3*	dna_mod
HIST3H3	histone	*KDM3A*	eraser	*MCM4*	dna_mod
HIST4H4	histone	*KDM3B*	eraser	*MCM5*	dna_mod
HJURP	dna_mod	*KDM4A*	eraser	*MCM6*	dna_mod
HLCS	writer	*KDM4B*	eraser	*MCM7*	dna_mod
HLTF	remodel	*KDM4C*	eraser	*MCRS1*	bind_writer
HMG20B	bind_eraser	*KDM4D*	eraser	*MDC1*	reader
HMGA1	dna_mod	*KDM4DL*	eraser	*MEAF6*	bind_writer
HMGN1	reader	*KDM5A*	eraser	*MECOM*	reader
HMGN2	reader	*KDM5B*	eraser	*MeCP2*	dna_mod
HMGN3	reader	*KDM5C*	eraser	*MEN1*	bind_writer
HMGN4	dna_mod	*KDM5D*	eraser	*MGEA5*	writer
HMGN5	reader	*KDM6A*	eraser	*MINA*	eraser
HUWE1	writer	*KDM6B*	eraser	*MIS12*	remodel
ING1	reader	*KEAP1*	reader	*MIS18A*	remodel
ING2	reader	*KMT2A*	writer	*MIS18BP1*	remodel
ING3	reader	*KMT2B*	writer	*MLL*	writer

Gene	Function	Gene	Function	Gene	Function
MLL2	writer	*ORC1*	bind_writer	*PHF5A*	reader
MLL3	writer	*PADI4*	eraser	*PHF6*	reader
MLL4	writer	*PAF1*	reader	*PHF7*	reader
MLL5	writer	*PAK2*	writer	*PHF8*	reader
MLLT6	reader	*PARP1*	writer	*PHIP*	reader
MORF4L1	reader	*PARP10*	writer	*PKN1*	writer
MORF4L2	reader	*PARP11*	writer	*POLE3*	remodel
MSH6	dna_mod	*PARP12*	writer	*PPP4C*	eraser
MSL2	bind_writer	*PARP14*	writer	*PPP5C*	eraser
MSL3	reader	*PARP15*	writer	*PRDM1*	writer
MST1	writer	*PARP16*	writer	*PRDM10*	writer
MTA1	bind_eraser	*PARP2*	writer	*PRDM11*	writer
MTA2	bind_eraser	*PARP3*	writer	*PRDM12*	writer
MTA3	bind_eraser	*PARP4*	writer	*PRDM13*	writer
MTF1	bind_writer	*PARP6*	writer	*PRDM14*	writer
MTF2	bind_writer	*PARP8*	writer	*PRDM15*	writer
MUM1	reader	*PARP9*	writer	*PRDM16*	writer
MUM1L1	reader	*PBRM1*	remodel	*PRDM2*	writer
MYC	bind_writer	*PCGF1*	bind_writer	*PRDM4*	writer
MYSM1	eraser	*PCGF2*	bind_writer	*PRDM5*	writer
NAT10	writer	*PCNA*	reader	*PRDM6*	writer
NBN	reader	*PHC1*	bind_writer	*PRDM7*	writer
NCOA1	writer	*PHC2*	bind_writer	*PRDM8*	writer
NCOA3	writer	*PHC3*	bind_writer	*PRDM9*	writer
NCOA6	writer	*PHF1*	reader	*PRKCB*	writer
NCOR1	bind_eraser	*PHF10*	remodel	*PRKCD*	writer
NCOR2	bind_eraser	*PHF11*	remodel	*PRKDC*	writer
NDC80	remodel	*PHF12*	bind_eraser	*PRMT1*	writer
Nek6	writer	*PHF13*	remodel	*PRMT2*	writer
Nek9	writer	*PHF14*	reader	*PRMT3*	writer
NFE2	bind_writer	*PHF15*	reader	*PRMT5*	writer
NFRKB	remodel	*PHF16*	reader	*PRMT6*	writer
NFYA	bind_writer	*PHF17*	reader	*PRMT7*	writer
NFYB	bind_writer	*PHF19*	reader	*RAD18*	reader
NFYC	bind_writer	*PHF2*	reader	*RAD51*	remodel
NSD1	writer	*PHF20*	reader	*RAD54B*	remodel
NSD2	writer	*PHF20L1*	reader	*RAD54L*	remodel
NSL1	remodel	*PHF21A*	reader	*RAG1*	reader
NUF2	remodel	*PHF21B*	reader	*RAG2*	reader
OGT	writer	*PHF23*	reader	*RAI1*	remodel
OIP5	remodel	*PHF3*	reader	*RBBP4*	bind_writer

Gene	Function	Gene	Function	Gene	Function
RBBP5	bind_writer	*SIRT5*	eraser	*SUV420H2*	writer
RBBP7	bind_writer	*SIRT6*	eraser	*SUZ12*	bind_writer
RCOR1	bind_eraser	*SIRT7*	eraser	*TADA1*	bind_writer
REST	bind_eraser	*SKA1*	remodel	*TADA2B*	bind_writer
RING1	writer	*SKA2*	remodel	*TADA3*	bind_writer
RNF168	writer	*SKA3*	remodel	*TAF1*	writer
RNF17	reader	*SLK*	writer	*TAF10*	writer
RNF2	writer	*SMARCA1*	reader	*TAF12*	writer
RNF20	writer	*SMARCA2*	reader	*TAF1L*	writer
RNF40	writer	*SMARCA4*	reader	*TAF3*	writer
RNF8	writer	*SMARCA5*	reader	*TAF5L*	writer
RPS6KA4	writer	*SMARCAD1*	reader	*TAF6L*	writer
RPS6KA5	writer	*SMARCAL1*	reader	*TAF9*	writer
RSF1	dna_mod	*SMARCB1*	reader	*TDRD3*	reader
RUVBL1	reader	*SMARCC1*	reader	*TERF2*	dna_mod
RUVBL2	reader	*SMARCC2*	remodel	*TET1*	dna_mod
SAP30	bind_eraser	*SMARCD1*	remodel	*TET2*	dna_mod
SCMH1	bind_writer	*SMARCD2*	remodel	*TET3*	dna_mod
SCML2	bind_writer	*SMARCD3*	remodel	*TFPT*	remodel
SERPINE1	reader	*SMARCE1*	remodel	*TP53BP1*	reader
SET	writer	*SMC1A*	reader	*TRIM24*	reader
SETD1A	writer	*SMC1B*	remodel	*TRIM28*	reader
SETD1B	writer	*SMYD1*	writer	*TRIM33*	writer
SETD2	writer	*SMYD2*	writer	*TRIM66*	reader
SETD3	writer	*SMYD3*	writer	*TRRAP*	bind_writer
SETD4	writer	*SMYD4*	writer	*UBE2A*	writer
SETD5	writer	*SMYD5*	writer	*UBE2B*	writer
SETD6	writer	*SP100*	dna_mod	*UBR1*	writer
SETD7	writer	*SP140*	reader	*UBR2*	writer
SETD8	writer	*SP140L*	reader	*UBR3*	writer
SETDB1	writer	*SPC24*	remodel	*UBR4*	writer
SETDB2	writer	*SPC25*	remodel	*UBR5*	writer
SETMAR	writer	*SRC*	writer	*UBR7*	writer
SFMBT1	reader	*SRCAP*	reader	*UCHL5*	writer
SFMBT2	reader	*SSRP1*	dna_mod	*UHRF1*	writer
SHPRH	writer	*STK4*	writer	*UHRF2*	writer
SIN3A	dna_mod	*SUPT3H*	bind_writer	*UIMC1*	bind_writer
SIRT1	eraser	*SUPT7L*	bind_writer	*UNG*	dna_mod
SIRT2	eraser	*SUV39H1*	writer	*USP12*	eraser
SIRT3	eraser	*SUV39H2*	writer	*USP16*	eraser
SIRT4	eraser	*SUV420H1*	writer	*USP21*	eraser

Gene	Function
USP22	eraser
USP3	eraser
USP46	eraser
USP7	eraser
UTY	eraser
VPS72	reader
VRK1	writer
WDR5	bind_writer
WDR82	bind_writer
WHSC1	writer
WHSC1L1	writer
YEATS4	bind_writer
YY1	dna_mod
ZAK	writer
ZBTB33	dna_mod
ZBTB38	dna_mod
ZBTB4	dna_mod
ZMYM1	reader
ZMYM2	bind_eraser
ZMYM3	bind_eraser
ZMYM4	reader
ZMYM5	reader
ZMYM6	reader
ZMYND11	reader
ZMYND8	reader
ZW10	remodel
ZWILCH	remodel

Note: In genes where multiple names are commonly used, all gene names were listed where possible (eg. *MLL*/*KMT2A*, *MLL2/MLL4/KMT2D, NSD2/WHSC1*).

Supplementary Table 5.3 – Variants present at diagnosis in *KMT2A*r B-ALL patient leukaemic blasts

Patient number	Sex	Age	Fusion	Reciprocal*	Variants & copy number alterations[x]
CH_A5528	F	UNK	*KMT2A-AFF1*	Y	*H1F0* p.A89P (VAF=0.32) *MCM7* p.A433P (0.32) *WHSC1* p.E1099K (0.51)
QCTB1179	F	11	*KMT2A-AFF1*	Y	*TP53* p.R248Q (0.48) *RUNX1* p.L56S (0.41)
CHII0723	M	14	*KMT2A-AFF1*	N	Nil
POW020	F	17	*KMT2A-AFF1*	Y	Nil
AYAI0248	F	18	*KMT2A-AFF1*	Y	Nil
AYAIB42PM	M	20	*KMT2A-AFF1*	N	*H1F0* p.A89P (0.32) *WHSC1* p.E1099K (0.45)
R180071	F	27	*KMT2A-AFF1*	Y	*PAR1* region amplification
AYAII0298	M	28	*KMT2A-AFF1*	Y	Nil
AYAII0557	M	29	*KMT2A-AFF1*	Y	*RUNX1* p.L56S (0.51)
AYAIIBT-10AW	F	29	*KMT2A-AFF1*	Y	*IKZF1* ex 1-10 het. deletion
AYAII0748	M	29	*KMT2A-AFF1*	Y	Nil
RNS042	M	34	*KMT2A-AFF1*	Y	Nil
AYAII0418	F	35	*KMT2A-AFF1*	Y	Nil
AYAII0003	F	37	*KMT2A-AFF1*	Y	*WHSC1* p.E1099K (0.43)
AYAII0268	F	37	*KMT2A-AFF1*	Y	Nil
AYAII0715	M	38	*KMT2A-AFF1*	Y	Nil
ADI0022	F	40	*KMT2A-AFF1*	Y	*PAR1* region amplification *WHSC1* p.E1099K (0.43)
ADI77TK	M	42	*KMT2A-AFF1*	Y	*MCM7* p.A433P (0.34) *WHSC1* p.E1099K (0.44)
ADI0163	F	51	*KMT2A-AFF1*	Y	*H1F0* p.A89P (0.31) *MCM7* p.A433P (0.35) *WHSC1* p.E1099K (0.40)
AYAII0266	F	22	*KMT2A-MLLT1*	N	Nil
ADI0043	F	60	*KMT2A-MLLT1*	N	*PAR1* region amplification *PAX5* ex 1-10 amplification *WHSC1* p.E1099K (0.40)

ADII0135	F	74	*KMT2A-MLLT1*	N	*ETV6* ex 1, *PAX5* ex 1-10 and *PAR1* region het. deletion *WHSC1* p.E1099K (0.35)
ADI0133	F	47	*KMT2A-EPS15*	Y	*NRAS* p.Q61K (0.31) *EBF1* ex 1*, ETV6* ex 1-8*, IKZF1* ex 1-8*, PAX5* ex 1-10*, CDKN2A* ex 2-4 and *PAR1* region het. deletion *H1F0* p.A89P (0.36) *MCM7* p.A433P (0.35) *WHSC1* p.E1099K (0.44)

*Refers to whether in-frame reciprocal *KMT2A*r (ie. *AFF1-KMT2A*) was detected by mRNA-seq

ˣVAF is reported in brackets for each SNV

Abbreviations: ex, exon; het., heterozygous; UNK, unknown

Supplementary Table 5.4 – Variants present at diagnosis in *BCR-ABL1* B-ALL patient leukaemic blasts

Patient	Sex	Age	Variants & copy number alterations[x]
CHI0200	M	8	*IKZF1* ex. 4-7, *PAX5* ex. 2-10 & *CDKN2A/B* het. deletion
CHII0090	F	15	Nil
CHII0400	M	15	*IKZF1* ex 1 & 4-7, *PAX5* ex 1-10 het. deletion *CDKN2A/B* hom. deletion
ALL5RPA20	M	18	*CDKN2A/B* het. deletion
AYAII0445	M	22	Nil
AYAII0358	F	22	*CDKN2A/B* hom. deletion *PAR1* duplication
AYAII0063	F	27	*RUNX1* p.L59S (0.48) *IKZF1* ex 1-8, *PAX5* ex. 2-5 & *CDKN2A/B* het. deletion
AYAII0138	M	28	*H1F0* p.A89P (0.34) *WHSC1* p.E1099K (0.43)
ALL5CCH05	M	29	Nil
AYAII0346	F	34	Nil
AYAII0060	F	36	*IKZF1* ex. 1-8 het. del
AYAII0072	M	36	*CDKN2A/B* het. deletion & *ETV6* ex 1-8 duplication
AYAII0106	F	37	*IKZF1* ex 1-3, *PAX5* ex 1-10, *ETV6* ex 1-8, *JAK2* ex 23 & *CDKN2A/B* het. del *IKZF1* ex 4-7 hom. del
AYAII0490	F	38	*IKZF1* ex 4-7 het. deletion
AYAII0396	M	39	*IKZF1* ex 1-8, *PAX5* ex 1-10, *JAK2* ex 23 & *CDKN2A/B* het. deletion *PAR1* duplication
ADI0064	M	41	Nil
ADI0209	F	42	*IKZF1* ex 2-8 & *RB1* ex 19-26 het. deletion *WHSC1* p.E1099K (0.42)
ADI0440	M	43	*IKZF1* ex 1-8, *PAX5* ex 1-10, *JAK2* ex 23 & *CDKN2A/B* het. deletion
ADI0061	M	45	Nil
ADI0067	F	47	*IKZF1* ex 4-7, *ETV6* ex 1-8, *JAK2* ex 23, *PAX5* ex 1-10 & *RB1* ex 19-26 het. deletion *CDKN2A/B* hom. deletion
ADI0025	F	51	*ETV6* ex 1-8 het. deletion *PAR1* duplication
ADI0212	F	60	*IKZF1* ex 1-8, *PAX5* ex 2-10 & *CDKN2A/B* het. deletion

| ADII0065 | F | 71 | *IKZF1* ex 1-8 het. deletion |

Note: "*CDKN2A/B* deletion" (het. or hom.) refers to deletion of *CDKN2A* ex. 2 & 4, and *CDKN2B* ex. 2

[x]VAF is reported in brackets for each SNV

Abbreviations: ex, exon; het., heterozygous; UNK, unknown

Supplementary Table 5.5 – Variants present at diagnosis in *MLLT10*r patient leukaemic blasts

Patient	Sex	Age	Fusion	Reciprocal*	Variants & copy number alterations^x
QCTB0861	M	15	*DDX3X-MLLT10*	Y	*EZH2* p.A255T (VAF=0.46) *NOTCH1* p.I1680N (0.46) *NRAS* p.G12V (0.58) *WT1* p.R353fs/R370fs (0.48)
A3350	F	UNK	*DDX3X-MLLT10*	Y	*DNMT3A* p.G298E (0.52) *EZH2* p.V679M (0.49) *IDH1* p.V178I (0.62) *PTEN* p.F241fs (0.51)
AYAI0082	M	20	*DDX3X-MLLT10*	N*	*NOTCH1* p.L1678Q (0.51) *PTEN* p.A418fs (0.23) *PTEN* p.P421delinsLS (0.64) *WT1* p.R369fs (0.93)
ADI0389	F	75	*PICALM-MLLT10*	Y	*JAK3* p.M511I (VAF=0.16)^# *NOTCH1* p.P401L (0.30) Duplication of *ERG* and *PAR1* region
A4316-18229	M	UNK	*PICALM-MLLT10*	Y	*EZH2* p.D185H (0.57) *JAK3* p.R657Q (0.47) *JAK3* p.M511I (0.31) *KRAS* p.G12A (0.51) *NF1* ex 26 & 58 het. del *SUZ12* ex 10 & 15 het. del
A5649-25219	M	UNK	*PICALM-MLLT10*	Y	*CDKN2A/B* hom. del^ *HDAC3* p.H135Q (0.58) *NOTCH1* p.I2456fs (0.36)
A2165	M	UNK	*PICALM-MLLT10*	Y	Duplication of *MYB* *CDKN2A* hom. del *CDKN2B, MLLT3* ex 1 & 7, and *MTAP* ex 1 het. del *IDH1* p.V178I (0.88) *NOTCH1* p.L2468fs (0.57) *PTEN* p.G129delinsGRL (0.38)

Note: "*CDKN2A/B* deletion" (het. or hom.) refers to deletion of *CDKN2A* ex. 2 & 4 and *CDKN2B* ex. 2

*Refers to whether in-frame reciprocal *MLLT10*r (ie. *MLLT10-DDX3X/PICALM*) was detected by mRNA-seq. For AYAI0082, *DDX3X-MLLT10* (chrX-chr10) was identified by mRNA-Seq. The LACTB-DDX3X (chr15-chrX) gene fusion was also identified. Cytogenetics indicated there was a 4-way translocation between chromosomes 10, 15, 17, X. However, mRNA-Seq did not identify a clinically relevant gene fusion involving chr17

ˣVAF is reported in brackets for each SNV and INDEL

#Variant is below VAF cut-off (0.2), but was reported as it is a highly pathogenic and recurrent variant identified in T-ALL

Abbreviations: ex, exon; het., heterozygous; UNK, unknown

Supplementary Table 5.6 – Evidence of pathogenicity for cancer-associated SNVs and indels included in analysis

Variant	Frequency (%)[x]			Evidence of pathogenicity*
	K	B	M	
IDH1 p.V178I	0	0	28.6	ACMG: Benign FATHMM: Pathogenic (0.99) 19 counts in COSMIC, of which 5 occurred in haematopoietic and lymphoid malignancies
JAK3 p.M511I	0	0	14.3	ACMG: Likely pathogenic FATHMM: Pathogenic (0.95) 76 counts in COSMIC, all of which occurred in haematopoietic and lymphoid malignancies
JAK3 p.R657Q	0	0	14.3	ACMG: Uncertain significance/likely pathogenic FATHMM: Pathogenic (0.99) 41 counts in COSMIC, of which 36 occurred in haematopoietic and lymphoid malignancies
KRAS p.G12A	0	0	14.3	ACMG: Pathogenic FATHMM: Pathogenic (0.98) 2520 counts in COSMIC, of which 73 occurred in haematopoietic and lymphoid malignancies
NOTCH1 p.I1680N	0	0	14.3	ACMG: Likely pathogenic FATHMM: Pathogenic (0.96) 21 counts in COSMIC, of which 20 occurred in haematopoietic and lymphoid malignancies
NOTCH1 p.L1678Q	0	0	14.3	ACMG: Likely pathogenic FATHMM: N/R 8 counts in COSMIC, of which all occurred in haematopoietic and lymphoid malignancies
NOTCH1 p.P401L	0	0	14.3	ACMG: Likely pathogenic FATHMM: Pathogenic (0.97) 3 counts in COSMIC, of which 1 occurred in a haematopoietic and lymphoid malignancy
NOTCH1 p.I2456fs	0	0	14.3	ACMG: Pathogenic Non-frameshift insertion that occurs in a mutational hotspot. Alternate frameshifts at this location previously reported in T-ALL[9,76]
NOTCH1 p.L2468fs	0	0	14.3	ACMG: Pathogenic Frameshift insertion, previously identified in T-ALL[9]
NRAS p.G12V	0	0	14.3	ACMG: Pathogenic FATHMM: Pathogenic (0.92)

			Widely recognised oncogenic variant; 135 counts in COSMIC, of which 102 occurred in haematopoietic and lymphoid malignancies
NRAS p.Q61K	4.3	0	0	ACMG: Pathogenic FATHMM: Pathogenic (0.99) 1423 counts in COSMIC, of which 218 occurred in haematopoietic and lymphoid malignancies
PTEN p.G129delinsGRL	0	0	14.3	ACMG: Likely pathogenic Non-frameshift insertion that introduces two electrically charged residues. Occurs within mutation hotspot
PTEN p.F241fs	0	0	14.3	ACMG: Pathogenic Frameshift insertion, occurs within mutation hotspot[77]
PTEN p.A418fs	0	0	14.3	ACMG: Pathogenic Frameshift insertion, residue lies within conserved PTEN domain
PTEN p.P421delinsLS	0	0	14.3	ACMG: Likely pathogenic Non-frameshift insertion, residue lies within conserved PTEN domain
RUNX1 p.L56S	8.7	8.7	0	ACMG: Uncertain significance FATHMM: Pathogenic (0.90) 18 counts in COSMIC, of which 17 occurred in haematopoietic and lymphoid malignancies
TP53 p.R248Q	4.3	0	0	ACMG: Pathogenic FATHMM: Pathogenic (0.98) >1400 counts in COSMIC, of which 134 occurred in haematopoietic and lymphoid malignancies
WT1 p.R353fs/R370fs	0	0	14.3	ACMG: Pathogenic Frameshift insertion, occurs in hotspot within conserved WT1 domain. Previously identified in T-ALL[78]
WT1 p.R369fs	0	0	14.3	ACMG: Pathogenic Frameshift insertion, previously reported in T-ALL[79]

[x]Refers to the percentage of patients in which each variant was identified (K, *KMT2A*r B-ALL; B, *BCR-ABL1* B-ALL; M, *MLLT10*r T-ALL). Abbreviations: N/R, not reported.

*SNVs with a FATHMM score of deleterious or pathogenic (>0.5), or an ACMG classification of likely pathogenic or pathogenic according to VarSome, were included in analysis. For INDELs, a pathogenic/likely pathogenic ACMG classification or previously published reports of the INDEL in T-ALL was required for inclusion in analysis.

Supplementary Table 5.7 – SNVs identified within epigenetic regulators

Variant	Frequency (%)[x]			Evidence of pathogenicity*
	K	**B**	**M**	
DNMT3A p.G298E	0	0	14.3	ACMG: Likely pathogenic FATHMM: Pathogenic (0.99) 2 counts in COSMIC, including 1 in a haematopoietic malignancy Resides in the regulatory region (PWWP/catalytic methyltransferase domain)[59]
EZH2 p.D185H	0	0	14.3	ACMG: Pathogenic FATHMM: Pathogenic (1.0) 98 counts in COSMIC, including 9 in haematopoietic and lymphoid malignancies. Known cancer predisposition variant, previously identified in T-ALL[80]
EZH2 p.A255T	0	0	14.3	ACMG: Likely pathogenic FATHMM: Pathogenic (0.92) 5 counts in COSMIC, including 3 in haematopoietic and lymphoid malignancies
EZH2 p.V679M	0	0	14.3	ACMG: Likely pathogenic FATHMM: Pathogenic (0.99) 11 counts in COSMIC, all in haematopoietic malignancies Occurs within catalytic SET domain[64]
H1F0 p.A89P	17.4	4.3	0	ACMG: Variant of uncertain significance FATHMM: N/R Highly conserved position, change of A to P is likely to structurally change protein due to proline secondary structure
HDAC3 p.H135Q	0	0	14.3	ACMG: Pathogenic FATHMM: N/R Occurs at a highly conserved, catalytically important residue involved in Zn binding. Substitution with alanine completely inactivates HDAC3 enzyme activity[70]
MCM7 p.A433P	17.4	0	0	ACMG: Variant of uncertain significance/likely pathogenic FATHMM: N/R Bioinformatically predicted as pathogenic by 9/12 callers, but no functional data available
WHSC1 p.E1099K	39.1	8.7	0	ACMG: Likely pathogenic FATHMM: Pathogenic (1.00)

				50 counts in COSMIC, including 34 in haematopoietic malignancies
				Occurs within a highly conserved residue within the SET domain[53]

ˣRefers to the percentage of patients in which each variant was identified (K, *KMT2A*r B-ALL; B, *BCR-ABL1* B-ALL; M, *MLLT10*r T-ALL).

*SNVs with a FATHMM score of deleterious or pathogenic (>0.5), or an ACMG classification of uncertain significance/likely pathogenic or pathogenic according to VarSome, were included in analysis.

Abbreviations: N/R, not reported.

5.6 Chapter References

1   Pui, C. H., Nichols, K. E. & Yang, J. J. Somatic and germline genomics in paediatric acute lymphoblastic leukaemia. *Nat Rev Clin Oncol* **16**, 227-240, doi:10.1038/s41571-018-0136-6 (2019).

2   Greaves, M. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nature Reviews Cancer*, 471-484, doi:10.1038/s41568-018-0015-6 (2018).

3   Winters, A. C. & Bernt, K. M. MLL-Rearranged Leukemias-An Update on Science and Clinical Approaches. *Front Pediatr* **5**, 1-21, doi:10.3389/fped.2017.00004 (2017).

4   Andersson, A. K. *et al.* The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nature genetics* **47**, 330 (2015).

5   Agraz-Doblas, A. *et al.* Unraveling the cellular origin and clinical prognostic markers of infant B-cell acute lymphoblastic leukemia using genome-wide analysis. *Haematologica* **104**, 1176-1188, doi:10.3324/haematol.2018.206375 (2019).

6   Inaba, H., Greaves, M. & Mullighan, C. G. Acute lymphoblastic leukaemia. *The Lancet* **381**, 1943-1955, doi:10.1016/s0140-6736(12)62187-4 (2013).

7   Burns, M., Armstrong, S. A. & Gutierrez, A. in *Hematology*   1005-1019.e1011 (2018).

8   Rice, S. *et al.* A human fetal liver-derived infant MLL-AF4 acute lymphoblastic leukemia model reveals a distinct fetal gene expression program. *Nat Commun* **12**, 6905, doi:10.1038/s41467-021-27270-z (2021).

9   Liu, Y. *et al.* The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nat Genet* **49**, 1211-1218, doi:10.1038/ng.3909 (2017).

10   Mansur, M. B. *et al.* The genomic landscape of teenage and young adult T-cell acute lymphoblastic leukemia. *Cancer Med* **10**, 4864-4873, doi:10.1002/cam4.4024 (2021).

11   Chang, Y. H. *et al.* Targeted sequencing to identify genetic alterations and prognostic markers in pediatric T-cell acute lymphoblastic leukemia. *Sci Rep* **11**, 769, doi:10.1038/s41598-020-80613-6 (2021).

12   Eadie, L. *et al.* Next Generation Genomic Analyses in T-ALL Patients Identify Recurrent and Novel Genomic Abnormalities. *Blood* **136**, 13-14, doi:10.1182/blood-2020-138995 (2020).

13   Nordlund, J. & Syvanen, A. C. Epigenetics in pediatric acute lymphoblastic leukemia. *Semin Cancer Biol*, doi:10.1016/j.semcancer.2017.09.001 (2017).

14   Figueroa, M. E. *et al.* Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia. *J Clin Invest* **123**, 3099-3111, doi:10.1172/JCI66203 (2013).

15   Shukla, N. *et al.* Final Report of Phase 1 Study of the DOT1L Inhibitor, Pinometostat (EPZ-5676), in Children with Relapsed or Refractory MLL-r Acute Leukemia. *Blood* **128**, 2780-2780 (2016).

16   Odore, E. *et al.* Phase I Population Pharmacokinetic Assessment of the Oral Bromodomain Inhibitor OTX015 in Patients with Haematologic Malignancies. *Clin Pharmacokinet* **55**, 397-405, doi:10.1007/s40262-015-0327-6 (2016).

17   Amorim, S. *et al.* Bromodomain inhibitor OTX015 in patients with lymphoma or multiple myeloma: a dose-escalation, open-label, pharmacokinetic, phase 1 study. *The Lancet Haematology* **3**, e196-e204, doi:10.1016/s2352-3026(16)00021-1 (2016).

18   Goldberg, J. *et al.* A phase I study of panobinostat in children with relapsed and refractory hematologic malignancies. *Pediatr Hematol Oncol* **37**, 465-474, doi:10.1080/08880018.2020.1752869 (2020).

19   Ma, X. *et al.* Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. *Nat Commun* **6**, 6604, doi:10.1038/ncomms7604 (2015).

20   Huether, R. *et al.* The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nat Commun* **5**, 3630, doi:10.1038/ncomms4630 (2014).

21   Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res* **47**, D941-D947, doi:10.1093/nar/gky1015 (2019).

22   Shihab, H. A., Gough, J., Cooper, D. N., Day, I. N. & Gaunt, T. R. Predicting the functional consequences of cancer-associated amino acid substitutions. *Bioinformatics* **29**, 1504-1510, doi:10.1093/bioinformatics/btt182 (2013).

23   Kopanos, C. *et al.* VarSome: the human genomic variant search engine. *Bioinformatics* **35**, 1978-1980, doi:10.1093/bioinformatics/bty897 (2019).

24   Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* **17**, 405-424, doi:10.1038/gim.2015.30 (2015).

25   Sayed, D. M. *et al.* Outcome and Clinical Significance of Immunophenotypic Markers Expressed in Different Treatment Protocols of Pediatric Patients With T-ALL in Developing Countries. *Clin Lymphoma Myeloma Leuk* **17**, 443-449, doi:10.1016/j.clml.2017.05.012 (2017).

26   Zhou, X. *et al.* Exploring genomic alteration in pediatric cancer using ProteinPaint. *Nat Genet* **48**, 4-6, doi:10.1038/ng.3466 (2016).

27   Massimino, M. *et al.* Impact of the Breakpoint Region on the Leukemogenic Potential and the TKI Responsiveness of Atypical BCR-ABL1 Transcripts. *Front Pharmacol* **12**, 669469, doi:10.3389/fphar.2021.669469 (2021).

28   Prelle, C., Bursen, A., Dingermann, T. & Marschalek, R. Secondary mutations in t(4;11) leukemia patients. *Leukemia* **27**, 1425-1427, doi:10.1038/leu.2012.365 (2013).

29   Emerenciano, M. *et al.* Subclonality and prenatal origin of RAS mutations in KMT2A (MLL)-rearranged infant acute lymphoblastic leukaemia. *British journal of haematology* **170**, 268-271 (2015).

30   Fedders, H. *et al.* The role of constitutive activation of FMS-related tyrosine kinase-3 and NRas/KRas mutational status in infants with KMT2A rearranged acute lymphoblastic leukemia. *Haematologica* **102**, e438-e442 (2017).

31   Driessen, E. M. *et al.* Frequencies and prognostic impact of RAS mutations in MLL-rearranged acute lymphoblastic leukemia in infants. *Haematologica* **98**, 937-944, doi:10.3324/haematol.2012.067983 (2013).

32   Liang, D. C. *et al.* Mutational status of NRAS, KRAS, and PTPN11 genes is associated with genetic/cytogenetic features in children with B-precursor acute lymphoblastic leukemia. *Pediatr Blood Cancer* **65**, e26786, doi:10.1002/pbc.26786 (2018).

33   Yu, R. T. D. & Garcia, R. L. NRAS mutant E132K identified in young-onset sporadic colorectal cancer and the canonical mutants G12D and Q61K affect distinct oncogenic phenotypes. *Sci Rep* **10**, 11028, doi:10.1038/s41598-020-67796-8 (2020).

34   Kerstjens, M. *et al.* Trametinib inhibits RAS mutant MLL-rearranged acute lymphoblastic leukemia at specific niche sites and reduces ERK phosphorylation in vivo. *Haematologica* **103**, e147-e150, doi:10.3324/haematol.2017.174060 (2018).

35   Prieto-Conde, M. I. *et al.* Genomic analysis of a familial myelodysplasia/acute myeloid leukemia and inherited RUNX1 mutations without a pre-existing platelet disorder. *Leuk Lymphoma* **61**, 181-184, doi:10.1080/10428194.2019.1648801 (2020).

36   Palmer, D. & Nacheva, E. An analysis of the RUNX1p.(Leu56Ser) variant in a cohort of individuals with myeloid neoplasms; suggests it is a benign germline variant. *Leuk Lymphoma* **62**, 1255-1258, doi:10.1080/10428194.2020.1861272 (2021).

37   Huang, G. *et al.* The ability of MLL to bind RUNX1 and methylate H3K4 at PU.1 regulatory regions is impaired by MDS/AML-associated RUNX1/AML1 mutations. *Blood* **118**, 6544-6552, doi:10.1182/blood-2010-11-317909 (2011).

38   Noronha, E. P. *et al.* The Profile of Immunophenotype and Genotype Aberrations in Subsets of Pediatric T-Cell Acute Lymphoblastic Leukemia. *Front Oncol* **9**, 316, doi:10.3389/fonc.2019.00316 (2019).

39   Van Vlierberghe, P. *et al.* Prognostic relevance of integrated genetic profiling in adult T-cell acute lymphoblastic leukemia. *Blood* **122**, 74-82, doi:10.1182/blood-2013-03-491092 (2013).

40   Dik, W. A. *et al.* CALM-AF10+ T-ALL expression profiles are characterized by overexpression of HOXA and BMI1 oncogenes. *Leukemia* **19**, 1948-1957, doi:10.1038/sj.leu.2403891 (2005).

41   Kerstjens, M. *et al.* MEK inhibition is a promising therapeutic strategy for MLL-rearranged infant acute lymphoblastic leukemia patients carrying RAS mutations. *Oncotarget* **8**, 14835-14846, doi:10.18632/oncotarget.11730 (2017).

42   Knight, T. & Irving, J. A. Ras/Raf/MEK/ERK Pathway Activation in Childhood Acute Lymphoblastic Leukemia and Its Therapeutic Targeting. *Front Oncol* **4**, 160, doi:10.3389/fonc.2014.00160 (2014).

43   Heatley, S. L. *et al.* Modeling Relapsed, Refractory Acute Lymphoblastic Leukemia from a Child with Neurofibromatosis. *Blood* **138**, 1317, doi:10.1182/blood-2021-150086 (2021).

44   Wong, J. *et al.* Failure of tofacitinib to achieve an objective response in a DDX3X-MLLT10 T-lymphoblastic leukemia with activating JAK3 mutations. *Cold Spring Harb Mol Case Stud* **6**, doi:10.1101/mcs.a004994 (2020).

45   Zhang, J. *et al.* The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* **481**, 157-163, doi:10.1038/nature10725 (2012).

46   Bains, T. *et al.* Newly described activating JAK3 mutations in T-cell acute lymphoblastic leukemia. *Leukemia* **26**, 2144-2146, doi:10.1038/leu.2012.74 (2012).

47   Sato, T. *et al.* Functional analysis of JAK3 mutations in transient myeloproliferative disorder and acute megakaryoblastic leukaemia accompanying Down syndrome. *Br J Haematol* **141**, 681-688, doi:10.1111/j.1365-2141.2008.07081.x (2008).

48   de Bock, C. E. & Cools, J. JAK3 mutations and HOXA9 expression are important cooperating events in T-cell acute lymphoblastic leukemia. *Mol Cell Oncol* **5**, e1458014, doi:10.1080/23723556.2018.1458014 (2018).

49   Broux, M. *et al.* Suz12 inactivation cooperates with JAK3 mutant signaling in the development of T-cell acute lymphoblastic leukemia. *Blood* **134**, 1323-1336, doi:10.1182/blood.2019000015 (2019).

50   Ntziachristos, P. *et al.* Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. *Nat Med* **18**, 298-301, doi:10.1038/nm.2651 (2012).

51   Jaffe, J. D. *et al.* Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. *Nat Genet* **45**, 1386-1391, doi:10.1038/ng.2777 (2013).

52   Swaroop, A. *et al.* An activating mutation of the NSD2 histone methyltransferase drives oncogenic reprogramming in acute lymphocytic leukemia. *Oncogene*, doi:10.1038/s41388-018-0474-y (2018).

53   Ding, L. W. *et al.* Mutational Landscape of Pediatric Acute Lymphoblastic Leukemia. *Cancer Res* **77**, 390-400, doi:10.1158/0008-5472.CAN-16-1303 (2017).

54   Oyer, J. *et al.* Point mutation E1099K in MMSET/NSD2 enhances its methyltranferase activity and leads to altered global chromatin methylation in lymphoid malignancies. *Leukemia* **28**, 198-201 (2014).

55   Tisi, D. *et al.* Structure of the Epigenetic Oncogene MMSET and Inhibition by N-Alkyl Sinefungin Derivatives. *ACS Chem Biol* **11**, 3093-3105, doi:10.1021/acschembio.6b00308 (2016).

56   Rao, R. C. & Dou, Y. Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat Rev Cancer* **15**, 334-346, doi:10.1038/nrc3929 (2015).

57   Torres, C. M. *et al.* The linker histone H1.0 generates epigenetic and functional intratumor heterogeneity. *Science* **353**, doi:10.1126/science.aaf1644 (2016).

58   Honeycutt, K. A. *et al.* Deregulated minichromosomal maintenance protein MCM7 contributes to oncogene driven tumorigenesis. *Oncogene* **25**, 4027-4032, doi:10.1038/sj.onc.1209435 (2006).

59   Yang, L., Rau, R. & Goodell, M. A. DNMT3A in haematological malignancies. *Nat Rev Cancer* **15**, 152-165, doi:10.1038/nrc3895 (2015).

60   Wong, K., Lawrie, CH, Green, TM. Oncogenic Roles and Inhibitors of DNMT1, DNMT3A, and DNMT3B in Acute Myeloid Leukaemia. *Biomarker Insights* **14**, 1-12, doi:10.1177/1177271919846454 (2019).

61   Scheller, M. *et al.* Hotspot DNMT3A mutations in clonal hematopoiesis and acute myeloid leukemia sensitize cells to azacytidine via viral mimicry response. *Nature Cancer* **2**, 527-544, doi:10.1038/s43018-021-00213-9 (2021).

62   Steca, I., Naglb, SB, van Ommena, G, den Dunnen, JT. The PWWP domain: a potential protein^protein interaction domain in nuclear proteins influencing differentiation? *FBBS Letters* **473**, 1-5, doi:10.1016/S0014-5793(00)01449-6 (2000).

63   Kim, K. H. & Roberts, C. W. Targeting EZH2 in cancer. *Nat Med* **22**, 128-134, doi:10.1038/nm.4036 (2016).

64   Bodor, C. *et al.* EZH2 mutations are frequent and represent an early event in follicular lymphoma. *Blood* **122**, 3165-3168, doi:10.1182/blood-2013-04-496893 (2013).

65   Yan, K. S. *et al.* EZH2 in Cancer Progression and Potential Application in Cancer Therapy: A Friend or Foe? *Int J Mol Sci* **18**, doi:10.3390/ijms18061172 (2017).

66   Burgos, S. *et al.* Novel EZH2 mutation in a patient with secondary B-cell acute lymphocytic leukemia after deletion 5q myelodysplastic syndrome treated with lenalidomide: A case report. *Medicine (Baltimore)* **98**, e14011, doi:10.1097/MD.0000000000014011 (2019).

67   Grossmann, V. *et al.* EZH2 mutations and their association with PICALM-MLLT10 positive acute leukaemia. *British Journal of Haematology* **157**, 387-390, doi:10.1111/j.1365-2141.2011.08986.x (2012).

68   Italiano, A. *et al.* Tazemetostat, an EZH2 inhibitor, in relapsed or refractory B-cell non-Hodgkin lymphoma and advanced solid tumours: a first-in-human, open-label, phase 1 study. *The Lancet Oncology* **19**, 649-659, doi:10.1016/s1470-2045(18)30145-1 (2018).

69   Adhikari, N., Amin, S. A., Trivedi, P., Jha, T. & Ghosh, B. HDAC3 is a potential validated target for cancer: An overview on the benzamide-based selective HDAC3 inhibitors through comparative SAR/QSAR/QAAR approaches. *Eur J Med Chem* **157**, 1127-1142, doi:10.1016/j.ejmech.2018.08.081 (2018).

70   Sun, Z. *et al.* Deacetylase-independent function of HDAC3 in transcription and metabolism requires nuclear receptor corepressor. *Mol Cell* **52**, 769-782, doi:10.1016/j.molcel.2013.10.022 (2013).

71   Hsieh, H. Y. *et al.* Targeting breast cancer stem cells by novel HDAC3-selective inhibitors. *Eur J Med Chem* **140**, 42-51, doi:10.1016/j.ejmech.2017.08.069 (2017).

72   Yang, Y., Dai, Y., Yang, X., Wu, S. & Wang, Y. DNMT3A Mutation-Induced CDK1 Overexpression Promotes Leukemogenesis by Modulating the Interaction between EZH2 and DNMT3A. *Biomolecules* **11**, doi:10.3390/biom11060781 (2021).

73   Neumann, M. *et al.* Whole-exome sequencing in adult ETP-ALL reveals a high rate of DNMT3A mutations. *Blood* **121**, 4749-4752, doi:10.1182/blood-2012-11-465138 (2013).

74   Stengel, A. *et al.* TP53 mutations occur in 15.7% of ALL and are associated with MYC-rearrangement, low hypodiploidy, and a poor prognosis. *Blood* **124**, 251-258, doi:10.1182/blood-2014-02-558833 (2014).

75   de Smith, A. J. *et al.* Predisposing germline mutations in high hyperdiploid acute lymphoblastic leukemia in children. *Genes Chromosomes Cancer* **58**, 723-730, doi:10.1002/gcc.22765 (2019).

76   Ma, X. *et al.* Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature* **555**, 371-376, doi:10.1038/nature25795 (2018).

77   Li, B. *et al.* Therapy-induced mutations drive the genomic landscape of relapsed acute lymphoblastic leukemia. *Blood* **135**, 41-55, doi:10.1182/blood.2019002220 (2020).

78   Grobner, S. N. *et al.* The landscape of genomic alterations across childhood cancers. *Nature* **555**, 321-327, doi:10.1038/nature25480 (2018).

79   Vicente, C. *et al.* Targeted sequencing identifies associations between IL7R-JAK mutations and epigenetic modulators in T-cell acute lymphoblastic leukemia. *Haematologica* **100**, 1301-1310, doi:10.3324/haematol.2015.130179 (2015).

80   Schafer, V. *et al.* EZH2 mutations and promoter hypermethylation in childhood acute lymphoblastic leukemia. *J Cancer Res Clin Oncol* **142**, 1641-1650, doi:10.1007/s00432-016-2174-8 (2016).

Chapter 6 – *TP53* loss-of-function mutations reduce sensitivity of acute leukaemia to the curaxin CBL0137

6.1 Statement of Authorship

Title of Paper	*TP53* loss-of-function mutations reduce sensitivity of acute leukaemia to the curaxin CBL0137
Publication Status	☒ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Authors: Michelle O. Forgione, Barbara J. McClure, Elyse C. Page, David T. Yeung, Laura N. Eadie, Deborah L. White Published in *Oncology Reports.* DOI: 10.3892/or.2022.8310

Principal Author

Name of Principal Author (Candidate)	Michelle Forgione		
Contribution to the Paper	Conceptualised idea, conducted research, constructed manuscript.		
Overall percentage (%)	80%		
Certification:	This paper is original research that I conducted during the period of my Higher Degree by Research candidature, and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	10/1/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Barbara J. McClure		
Contribution to the Paper	Conceptualised idea, critical review of manuscript.		
Signature		Date	10/1/2022

Name of Co-Author	Elyse C. Page

Contribution to the Paper	Designed experiments, critical review of manuscript		
	296		
Signature		Date	10/1/2022

Name of Co-Author	David T. Yeung		
Contribution to the Paper	Critical review of manuscript.		
Signature		Date	10/1/2022

Name of Co-Author	Laura N. Eadie		
Contribution to the Paper	Critical review of manuscript.		
Signature		Date	10/1/2022

Name of Co-Author	Deborah L. White		
Contribution to the Paper	Conceptualised idea, critical review of manuscript.		
Signature		Date	10/1/2022

TP53 loss-of-function mutations reduce sensitivity of acute leukaemia to the curaxin CBL0137

MICHELLE O. FORGIONE[1,2], BARBARA J. McCLURE[1,3], ELYSE C. PAGE[1,2],
DAVID T. YEUNG[1,3,4], LAURA N. EADIE[1,3] and DEBORAH L. WHITE[1-3,5,6]

[1]Cancer Program, Precision Medicine Theme, South Australian Health and Medical Research Institute (SAHMRI);
Faculties of [2]Science and [3]Health and Medical Science, University of Adelaide; [4]Department of Haematology,
Royal Adelaide Hospital, Adelaide, South Australia 5000; [5]Australian and New Zealand Children's Haematology/Oncology
Group (ANZCHOG), Melbourne, Victoria 3168; [6]Australian Genomics Health Alliance (AGHA),
The Murdoch Children's Research Institute, Melbourne, Victoria 3052, Australia

Abstract. The presence of a *TP53* mutation is a predictor of poor outcome in leukaemia, and efficacious targeted therapies for these patients are lacking. The curaxin CBL0137 has demonstrated promising antitumour activities in multiple cancers such as glioblastoma, acting through p53 activation, NF-κB inhibition and chromatin remodelling. In the present study, it was revealed using Annexin-V/7-AAD apoptosis assays that CBL0137 has efficacy across several human acute leukaemia cell lines with wild-type *TP53*, but sensitivity is reduced in *TP53*-mutated subtypes. A heterozygous *TP53* loss-of-function mutation in the *KMT2A-AFF1* human RS4;11 cell line was generated, and it was demonstrated that heterozygous *TP53* loss-of-function is sufficient to cause a significant reduction in CBL0137 sensitivity. To the best of our knowledge, this is the first evidence to suggest a clinically significant role for functional p53 in the efficacy of CBL0137 in acute leukaemia. Future CBL0137 clinical trials should include *TP53* mutation screening, to establish the clinical relevance of *TP53* mutations in CBL0137 efficacy.

Correspondence to: Professor Deborah L. White, Cancer Program, Precision Medicine Theme, South Australian Health and Medical Research Institute (SAHMRI), North Terrace, Adelaide, South Australia 5000, Australia
E-mail: deborah.white@sahmri.com

Abbreviations: ALL, acute lymphoblastic leukaemia; AML, acute myeloid leukaemia; *KMT2A*r, *KMT2A*-rearranged; EFS, event-free survival; PDX, patient-derived xenograft; B-ALL, B-cell ALL; T-ALL, T-cell ALL; *TP53*[WT] *TP53*-wild type; *TP53*[MUT], *TP53*-mutated

Key words: p53, TP53, curaxin, acute leukaemia, ALL, CBL0137

Introduction

Despite recent improvements in long-term survival rates for acute lymphoblastic leukaemia (ALL) (1) and acute myeloid leukaemia (AML) (2), subgroups of patients continue to experience poor long-term outcomes. *TP53* alterations are a high-risk genomic feature present in 16% of newly diagnosed ALL (3,4) and 13% of newly diagnosed AML cases (4). The presence of a *TP53* alteration is associated with poor long-term overall survival compared with *TP53* wild-type (*TP53*[WT]) cases in both ALL (24 months vs. not reached, P=0.001) and AML (6 vs. 26 months, P<0.001) (3). To date, no targeted therapies are currently available to treat *TP53*-mutated (*TP53*[MUT]) acute leukaemia outside of clinical trials, and more efficacious therapeutic options are required to improve long-term outcomes.

CBL0137 is a small molecule curaxin that epigenetically modulates multiple cancer-related signalling pathways (5,6), including casein kinase 2 (CK2)-mediated activation of the *TP53*-encoded p53 protein through the Facilitates Chromatin Transcription (FACT) complex (7). CBL0137 is effective in pre-clinical ALL patient-derived xenograft (PDX) models, most notably in infant *KMT2A*-rearranged (*KMT2A*r) ALL (8,9). *KMT2A*r is present in 5-10% of newly-diagnosed acute leukaemia cases overall (10), including 70-80% of infant ALL and 35-50% of infant AML diagnoses (11), and *TP53* mutations occur in ~16% of *KMT2A*r ALL cases (4). Outcomes for *KMT2A*r leukaemia are exceptionally poor across all age groups, with five-year event-free survival (EFS) rates of 30-50% (12). Evidently, advances in treatment options are required to improve patient outcomes.

In a study by Lock *et al* (8), CBL0137 induced complete remission in a *KMT2A*r infant ALL PDX model, and partial responses were observed in several B-cell ALL (B-ALL) and T-cell ALL (T-ALL) models of unknown genomic subtype. Somers *et al* (9) similarly reported CBL0137 efficacy in infant *KMT2A*r B-ALL PDX models, with responses ranging from delayed cancer progression to maintained complete remission. These early data suggested that the role of CBL0137 deserves further exploration. For instance, the presence of

loss-of-function *TP53* alterations, a reasonably common event in *KMT2A*r B-ALL, may reasonably be expected to confer resistance. Furthermore, given its *in vitro* efficacy against *KMT2A*r B-ALL, CBL0137 is also likely to be active against other subtypes of B-ALL. In the present study, it was confirmed that CBL0137 induces leukaemic cell apoptosis in a number of cell lines with various driver alterations, including those with *KMT2A* rearrangements, with LC_{50} concentrations in the range of 166 to 676 nM. Notably, it was also demonstrated that the potency of CBL0137 is attenuated in the presence of loss of function *TP53* alterations.

Materials and methods

Cell line maintenance. U-937 (CRL-1593.2), MV4;11 (CRL-9591), Jurkat (TIB-152), THP-1 (TIB-202) and RS4;11 (CRL-1873) cell lines were purchased from the American Type Culture Collection (ATCC). RCH-ACV (ACC 548), REH (ACC 22) and NALM-19 (ACC 522) cell lines were purchased from DSMZ. All cell lines were maintained in culture at 37°C in RPMI-1640 media (cat. no. R0883; Sigma-Aldrich; Merck) containing foetal calf serum (FCS) (AU-FBS/SF; CellSera), 5 mM L-glutamine, 50 U/ml penicillin and 50 µg/ml streptomycin. THP-1 cells were cultured in 20% FCS, and all other cell lines were cultured in 10% FCS. Cells up to 20 passages from the original stock were used for experiments.

Inhibitor storage. CBL0137 (cat. no. S0507; Selleck Chemicals) was stored long-term at 10 mM in DMSO at -80°C and diluted in DMSO immediately prior to use.

Generation of CRISPR/Cas9 TP53 knock-out RS4;11 cell lines. Vectors FUCas9mCherry and FgH1tUTG were a gift from Professor Marco Herald (Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia) (13). Two independent single guide RNA (sgRNA) targets were designed, to account for off-target effects. sgRNA target sequences were designed to generate random indel and frameshift mutations in exon 4 of *TP53*, to disrupt the p53 protein prior to the critical DNA-binding domain, where deleterious *TP53* mutations identified in human leukaemia cell lines are located (Fig. 1A). sgRNA oligonucleotides were designed with online Benchling® Software (Benchling) and purchased from Sigma-Aldrich; Merck KGaA: TP53 Oligo 1 forward, 5'-TCCCACCAGCAG CTCCTACACCGG-3' and reverse, 5'-AAACCCGGTGTA GGAGCTGCTGGT-3'; and TP53 Oligo 2 forward, 5'-TCC CCCATTGTTCAATATCGTCCG-3' and reverse, 5'AAACCG GACGATATTGAACAATGG-3'.

TP53 FgH1tUTG sgRNA vectors were generated by *Bsm*BI digestion and T4 DNA ligation, incubated overnight at 4°C and transformed by 42°C heat shock for 2 min in DH5α™-T1ᴿ chemically competent *E. coli* (Thermo Fisher Scientific, Inc.), and selected based on ampicillin resistance (50 µg/ml) conferred by transformation of the FgH1tUTG plasmid. Successful ligation of inserts was confirmed by Sanger sequencing with the BigDye™ Terminator v3.1 Cycle Sequencing Kit (used according to manufacturer's protocol) and SeqStudio™ Genetic Analyser System (both from Thermo Fisher Scientific, Inc.). DNA chromatogram results were analysed using online Benchling® Software.

293T cells (CRL-3216; purchased from ATCC) were co-transfected with FUCas9mCherry (4.2 µg) or FgH1tUTG sgRNA (4.2 µg) vectors, and 2nd generation lentiviral packaging vectors pMD2.G (1.6 µg), pMDLg/pRRE (2.4 µg) and pRSV-Rev (1.1 µg) (Addgene, Inc.). Each transfection was prepared in 450 µl Opti-MEM™ Reduced Serum Media and 30 µl Lipofectamine™ 2000 transfection reagent (both from Thermo Fisher Scientific, Inc.), and incubated for 1 h at room temperature prior to application to 293T cells. Viral supernatant was harvested from 293T cultures 48 h post-transfection (MOI not quantified), and RS4;11 cells were transduced by spinfection in the presence of 4 mg/ml Polybrene® (Santa Cruz Biotechnology, Inc.) at 220 x g for 1 h at room temperature in six-well tissue culture plates. One week post-transduction, GFP and mCherry double-positive populations were sorted with a BD FACSFusion flow cytometer (BD Biosciences). Sorted populations were activated with doxycycline (1 µg/ml) for 3 days, and genomic DNA was extracted by phenol-chloroform to confirm induction of frameshift mutations by PCR amplification of TP53 exon 4 using Q5® High-Fidelity DNA Polymerase (cat. no. M0491; New England Biolabs, Inc.) and the following primers: TP53 intron 4 forward, 5'-TCCTCT GACTGCTCTTTTCACCCAT-3' and reverse, 5'-AATATT CAACTTTGGGACAGGAGTCAGAGA-3'. Thermocycling conditions were as follows: 98°C for 1 min, followed by 33 cycles at 98°C for 10 sec, 64°C for 15 sec and 72°C for 2 min, followed by 1 cycle at 72°C for 10 min. Samples were then stored at 4°C until they were visualised in 2% agarose gel containing 1:10,000 GelRed (Biotium, Inc.).

Sanger sequencing was utilised to identify the profile of mutations present (Fig. S1). Sanger sequencing was performed using the BigDye™ Terminator v3.1 Cycle Sequencing Kit according to manufacturer's instructions (Thermo Fisher Scientific, Inc.) and SeqStudio™ Genetic Analyser System (Thermo Fisher Scientific, Inc.). DNA chromatogram results were analysed using online Benchling® Software. RNA was harvested after 7 and 14 days in culture, to quantify *TP53* expression by quantitative reverse transcription-quantitative PCR (RT-qPCR). As a comparator, RS4;11 cells expressing Cas9 vector only were used as a *TP53*ᵂᵀ control in all relevant experiments. RT-qPCR thermocycling conditions were as follows: 10 min at 95°C for one cycle, followed by 40 cycles at 95°C for 15 sec and 60°C for 60 sec. RT-qPCR cycling reactions were performed on a QuantStudio 7 Real-Time PCR system (Thermo Fisher Scientific, Inc.).

Apoptosis detection via Annexin V/7-Aminoactinomycin D staining. Cells were seeded at $2x10^5$ cells/ml and treated for 72 h with a range of CBL0137 doses in 0.3% DMSO, and incubated in 96-well tissue culture plates at 37°C/5% CO_2 for 72 h. Treated cells were harvested at room temperature by centrifugation at 220 x g for 5 min, washed twice in flow cytometry binding buffer (Hank's Balanced Salt Solution (cat. no. H9394; Sigma-Aldrich; Merck KGaA), 10 mM HEPES, 5 mM $CaCl_2$), and cells in each well stained with Annexin V (cat. no. 556421; BD Biosciences) and 7-Aminoactinomycin (7-AAD) (cat. no. A1310; Thermo Fisher Scientific, Inc.) according to the supplier's protocol (BD Biosciences), and incubated on ice for 30 min. Assays were analysed on a BD FACSCanto flow cytometer (BD Biosciences) and data were
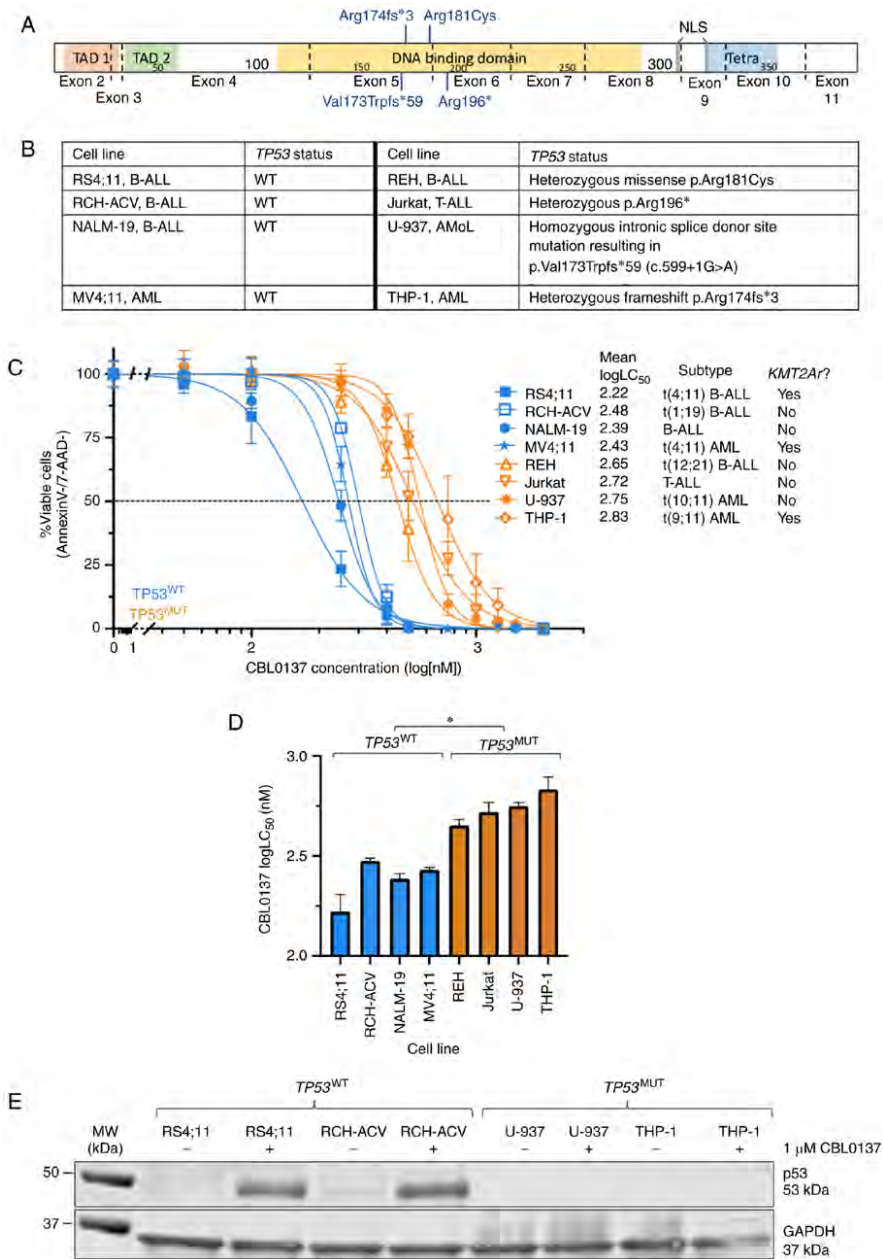
Figure 1. CBL0137 induces apoptosis in acute leukaemia cell lines, and LC_{50} are higher in *TP53*-mutated cell lines. (A) TP53 mutations identified in leukaemia cell lines. Data was obtained from the Broad Institute Cancer Cell Line Encyclopedia portal. (B) Protein domain schematic of p53, with mutations identified in cell lines annotated. Exons are denoted by dotted lines. (C) Human acute leukaemia cell lines were exposed to increasing doses of CBL0137 in duplicate and apoptosis was measured by Annexin-V/7-AAD staining after 72 h. Concentration-response curves and $logLC_{50}$ values were extrapolated using non-linear regression analysis in GraphPad Prism. Data are mean + SD for 3 biological replicates. (D) Mean CBL0137 $logLC_{50}$ values for each cell line are shown. LC_{50} values are provided in Table SII. For statistical analysis, $^{*}P<0.05$ for comparison of TP53 wild-type (TP53WT) and mutated (TP53MUT) mean $logLC_{50}$, Kruskal-Wallis test with multiple comparisons. (E) Representative immunoblot of total p53 protein in DMSO vehicle-treated (-) and 1 μM CBL0137-treated (+) cell lines. AML, acute myeloid leukaemia; AMoL, acute monocytic leukaemia; B-ALL, B-lineage acute lymphoblastic leukaemia; MUT, mutated; WT, wild-type; T-ALL, T-lineage acute lymphoblastic leukaemia; TAD, transactivation domain; NLS, nuclear localisation signal; Tetra, Tetramerisation motif.

analysed using FlowJo version 10 software (FlowJo LLC) to determine apoptotic (Annexin V and/or 7-AAD positive) and non-apoptotic (Annexin V and 7-AAD negative) populations.

The *in vitro* efficacy of CBL0137 to induce apoptosis was assessed on four $TP53^{WT}$ (RS4;11, RCH-ACV, NALM-19, MV4;11) and four $TP53^{MUT}$ (REH, Jurkat, U-937 and THP-1) human acute leukaemia cell lines using Annexin-V/7-AAD. The U-937 cell line was initially classified as histiocytic lymphoma (14) but is now classified as acute monocytic leukaemia (AMoL) and is therefore referred to as such throughout the manuscript (15). *TP53* mutations present in each cell line investigated are provided in Fig. 1A, obtained from the Broad Institute Cancer Cell Line Encyclopedia portal (16) (https://sites.broadinstitute.org/ccle/). All identified *TP53* mutations within the cell lines utilised are predicted to be deleterious according to COSMIC (17) (https://cancer.sanger.ac.uk) or ClinVar (18) (https://www.ncbi.nlm.nih.gov/clinvar/) databases. Later, the *in vitro* efficacy of CBL0137 to induce apoptosis was assessed on CRISPR/Cas9 *TP53* knock-out RS4;11 cell lines.

RNA analysis. RNA was extracted from $5x10^6$-$1x10^7$ cells using TRIzol® reagent (Thermo Fisher Scientific, Inc.). RNA concentration and purity were quantified with a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Inc.). The QuantiTect reverse transcription kit (Qiagen GmbH) was used to synthesise cDNA from 1 μg of RNA as per manufacturer's instructions. RT-qPCR reactions were prepared in duplicate as follows: 12.5 μl RT² SYBR® Green ROX™ qPCR Mastermix (Qiagen GmbH), 400 nM each of forward and reverse primers, 1 μl cDNA, nuclease-free H_2O to 25 μl final volume. Cycling reactions were performed on a QuantStudio 7 Real-Time PCR system (Thermo Fisher Scientific, Inc.) with the following primers: TP53 qPCR forward, 5'-GAAGGAAATTTGCGTGTGG-3' and reverse, 5'-TGTTACACATGTAGTTGTAGTGG-3'. Values were normalised against housekeeping gene ACTB (forward, 5'-GATCATTGCTCCTCCTGAGC-3' and reverse, 5'-TCTGCGCAAGTTAGGTTTTGTC-3'). The thermocycling conditions were as aforementioned. Relative gene expression was calculated by the $\Delta\Delta$Cq method and fold-change expression ($2^{-\Delta\Delta Cq}$) was calculated relative to Cas9 (TP53-WT) control (19).

Protein analysis. Cells were treated for 6 h in either 0.3% DMSO or 1 μM CBL0137 (0.3% DMSO) final concentration in RPMI-1640 + 2% FCS + 50 U/ml penicillin + 50 μg/ml streptomycin. Cells were pelleted by centrifugation at 10,000 x g for 10 min at 4°C and lysed in 60 μl 1% NP-40 (IGEPAL®) buffer (Sigma-Aldrich; Merck KGaA). Protein concentration was determined by generating a standard curve using a bicinchoninic acid (BCA) assay. Lysates were denatured and 60 μg protein was electrophoresed on 4-15% pre-cast gels (Bio-Rad Laboratories, Inc.). Proteins were transferred to polyvinylidene difluoride membranes (Bio-Rad Laboratories, Inc.) and blocked with 1X Intercept blocking buffer (LI-COR Biosciences) for 1 h at room temperature before incubation with primary antibodies (mouse anti-p53, product no. 2524; rabbit anti-p21, product no. 2947; and rabbit anti-GAPDH, product no. 2118) at 4°C for 17-48 h. All primary antibodies

were purchased from Cell Signaling Technology, Inc. and diluted at 1:1,000 in 1X Intercept blocking buffer. After washing with TBS-T (containing 0.1% Tween-20) and TBS, membranes were incubated with secondary antibodies [IRDye® 800CW anti-mouse (1:10,000) or IRDye® 800CW anti-rabbit (1:10,000)] (cat nos. 926-32212 and 925-32213, respectively; LI-COR Biosciences) for 2 h in the dark at room temperature and visualised on the LI-COR Odyssey® fluorescent scanner. For calculation of p21 expression fold change, Empiria Studio® Software version 2.0 (https://www.licor.com/bio/empiria-studio/) was used to normalise protein expression to housekeeping GAPDH protein. Fold expression change was then calculated for CBL0137-treated samples, relative to untreated samples for each respective cell line.

Statistical analysis. All calculations and statistical analysis of data were performed using GraphPad Prism Version 9.2.0 for Mac OS (GraphPad Software, Inc.). Kruskal-Wallis test with Dunn's multiple comparison post hoc test or Mann-Whitney U tests were performed to compare mean $logLC_{50}$ values between each cell line tested. $LogLC_{50}$ and LC_{50} values were extrapolated using non-linear regression analysis. All data was generated from 3 independent biological replicates. $P<0.05$ was considered to indicate a statistically significant difference.

Results

All identified *TP53* mutations within the cell lines REH, Jurkat, U-937 and THP-1 occur within a critical region of the p53 DNA binding domain (Fig. 1A and B). The presence of *TP53* loss-of-function genomic alterations significantly reduced the sensitivity of cells to CBL0137, independent of other genomic lesions present in each cell line (Fig. 1C and D). The aggregate mean $logLC_{50}$ was greater for $TP53^{MUT}$ cell lines compared with $TP53^{WT}$ cell lines (mean $logLC_{50}$=2.38 nM vs. 2.75 nM, P=0.046; Fig. 1D), with a more than two-fold increase in LC_{50} (Table SI). Notably, the *KMT2A* WT B-ALL cell lines NALM-19 and RCH-ACV exhibited low $logLC_{50}$ values ($logLC_{50}$=2.39 and 2.48 nM respectively; Fig. 1C). The *KMT2A*r cell line THP-1 was the most resistant cell line investigated ($logLC_{50}$=2.83 nM; Fig. 1C and D). A total of four cell lines, including two $TP53^{WT}$ and two $TP53^{MUT}$ were probed for the presence of total p53 protein, demonstrating increased levels of p53 following CBL0137 treatment in $TP53^{WT}$ cell lines only (Fig. 1E).

To further investigate the role of p53 in CBL0137 efficacy, heterozygous *TP53* loss-of-function cell lines ($TP53^{+/-}$) were generated using CRISPR/Cas9 in the human *KMT2A-AFF1* ALL cell line RS4;11. The RS4;11 cell line was selected as a representative of *KMT2A*r cell lines as it is highly sensitive to CBL0137 and expresses *KMT2A-AFF1,* the most common *KMT2A*r identified in B-ALL (10). Heterozygous *TP53* loss-of-function was confirmed in cell lines by RT-qPCR and immunoblot analysis (Fig. 2A and B). Both $TP53^{+/-}$ cell lines exhibited a significant reduction in *TP53* expression by RT-qPCR, compared with control $TP53^{WT}$ RS4;11 cells (Fig. 2A). Immunoblot analysis demonstrated that treatment with 1 μM CBL0137 stimulated p53 expression in WT RS4;11 and positive control NALM-19 cells, but this effect was abrogated in both $TP53^{+/-}$ RS4;11 cell lines
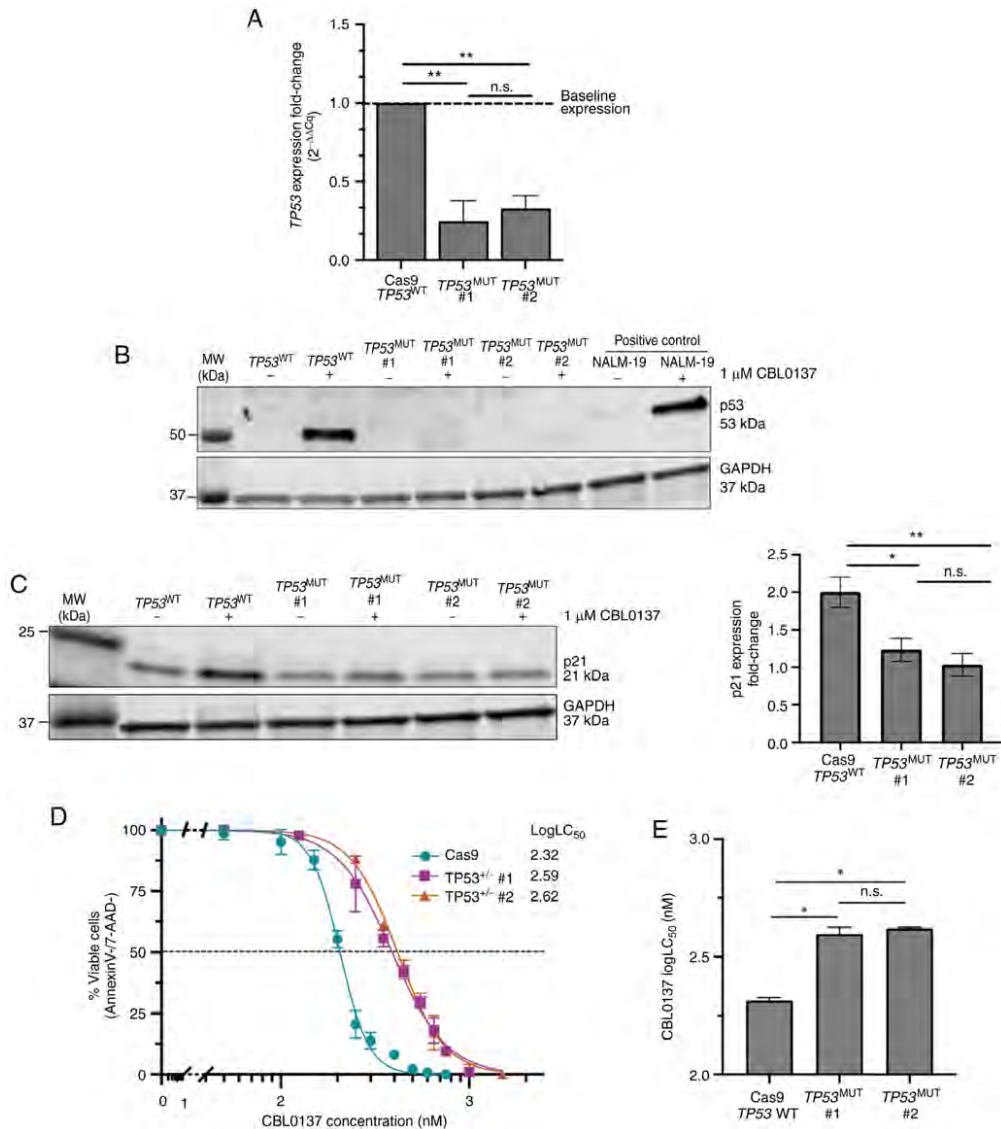
Figure 2. Heterozygous loss of *TP53* (*TP53*[+/-]) promotes CBL0137 resistance in *KMT2A*r ALL. (A) Reverse transcription-quantitative PCR analysis of TP53 expression in RS4;11 cell lines validated heterozygous loss of TP53 expression. Data are mean $2^{-\Delta\Delta Cq}$ + SD, relative to Cas9 control RS4;11 cells. Baseline expression refers to TP53 expression in *TP53*[WT] RS4;11 cells. (B) Representative immunoblot of total p53 protein in DMSO vehicle-treated (-) and 1 μM CBL0137-treated (+) RS4;11 cell lines. (C) Representative immunoblot of total p21 protein in DMSO vehicle-treated (-) and 1 μM CBL0137-treated (+) RS4;11 cell lines. Fold change in p21 expression following CBL0137 treatment was calculated relative to untreated samples, after normalizing to GAPDH. (D and E) CBL0137 dose-response and logLC$_{50}$ for *TP53*[WT] and *TP53*[MUT] RS4;11 cell lines, measured by Annexin-V/7-AAD apoptosis assay. LC$_{50}$ values are provided in Table SII. Data are presented as the mean + SD. For all statistical analysis, [*]P<0.05 and [**]P<0.01, for comparison of cell lines by Kruskal-Wallis test with multiple comparisons. All data is representative of 3 independent biological replicates. *KMT2A*r, KMT2A-rearranged; ALL, acute lymphoblastic leukaemia; *TP53*[WT] *TP53*-wild type; *TP53*[MUT], *TP53*-mutated; n.s., non-significant.

(Fig. 2B). The p53 effector protein p21 was upregulated ~two-fold in CBL0137-treated *TP53*[WT] RS4;11 cells but not *TP53*[MUT] cells, demonstrating that heterozygous *TP53*

loss-of-function is sufficient to abrogate activation of downstream p53 pathways (Fig. 2C). The cell senescence effector protein p21 is a well-characterised p53 effector protein, where

p21 is rapidly activated following p53 activation, to induce cellular senescence and apoptosis (20,21). Annexin-V/7-AAD staining of $TP53^{+/-}$ cell lines revealed a ~two-fold increase in CBL0137 $logLC_{50}$ (mean $logLC_{50}$ WT=2.32 nM vs. $TP53^{+/-}$ #1=2.59 nM, P=0.043; WT vs. $TP53^{+/-}$ #2=2.62 nM, P=0.021; Fig. 2D and E), indicating that heterozygous *TP53* loss-of-function is sufficient to cause a significant reduction in the sensitivity of cell lines to CBL0137.

Discussion

In the present study, it was demonstrated that CBL0137 has similar potency in all tested $TP53^{WT}$ acute leukaemia cell lines, regardless of the presence of any additional genomic lesions, including *KMT2A*r. However, the presence of a *TP53* loss-of-function mutation significantly increased CBL0137 LC_{50}, with a ~2-fold increase (range 1.7-4.0-fold) compared with $TP53^{WT}$ cell lines, regardless of genotype. It was also demonstrated, for the first time to the best of our knowledge, that heterozygous *TP53* loss-of-function is alone sufficient to cause a significant increase in the LC_{50} of CBL0137 in the *KMT2A*r B-ALL cell line RS4;11.

These results suggested that CBL0137 is indeed a promising therapy which may be broadly applicable in acute leukaemia. These data also indicated that cytotoxicity of CBL0137 is not specific to *KMT2A*r, but rather the target depends on the presence or absence of functional *TP53*. Reduced sensitivity is expected in the context of *TP53* mutated malignancies, and further testing is warranted to understand the clinical significance of *TP53* mutation status in CBL0137 efficacy. It is important to note that thus far, all ALL xenograft models tested by the Pediatric Preclinical Testing Program were $TP53^{WT}$ (8). An additional 6/31 paediatric solid tumour models exhibited tumour growth delay, and 5/6 of these were $TP53^{WT}$ (8). This should also be investigated through further pre-clinical testing, and anticipated in its clinical development program. For instance, two recently completed Phase I studies of CBL0137 reported acceptable tolerability and some clinical activity against cancers of the liver, prostate, uterus, breast and ovary (22,23). The *TP53* mutational status would optimise patient selection and target those most likely to benefit from this drug for further clinical testing. It is possible that mutations in p53 downstream mediators may also influence CBL0137 potency. However, alterations within p53 downstream mediators such as *CDKN1A* (p21) and *BCL2* (BCL-2) are rarely reported in acute leukaemia. It is also possible that alterations within CK2 or the FACT complex may influence CBL0137 potency, as CBL0137-mediated activation of p53 occurs via these pathways (7), but data on the occurrence of mutations in these pathways in acute leukaemia is currently lacking (24).

The present findings highlighted the importance of p53 activation in CBL0137 efficacy in acute leukaemia, indicating that *TP53* mutation status is an important factor in the clinical application of CBL0137. There is an ongoing need to identify therapeutic strategies for patients with high-risk acute leukaemia, such as those with *TP53* mutations, to improve outcomes for these patients. The data of the present study indicated that CBL0137 is a promising anticancer therapy that depends on WT p53 activity, which thus far has not been therapeutically targetable. The clinical feasibility of CBL0137 across both $TP53^{MUT}$ and $TP53^{WT}$ malignancies will become evident as further clinical trials are performed. In the present study, it was clearly demonstrated that the potency profile of CBL0137 is tightly linked to *TP53* mutation status, and it was revealed for the first time that heterozygous *TP53* loss-of-function alone significantly affects response to CBL0137 *in vitro*. These results support the need for accurate determination of *TP53* mutation status for patients enrolling in future CBL0137 clinical trials.

Availability of data and materials

The datasets used and/or analysed in the present study are available from the corresponding author on reasonable request.

Authors' contributions

MOF conceptualised the presented idea. MOF performed experiments and constructed the manuscript in consultation with BJM, ECP, DTY, LNE and DLW. MOF and BJM confirm the authenticity of all the raw data. All authors provided critical feedback and helped shape the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The present study was performed under the International Bioethics Committee (IBC) certification number BC02/2018, and all work contained within the present study was approved by the Royal Adelaide Hospital HREC committee (approval no. HREC/15/RAH/54).

Patient consent for publication

Not applicable.

Competing interests

DLW receives research support from Bristol-Meyers Squibb, and Honoraria from Bristol-Meyers Squibb and Amgen. DTY receives research support from Bristol-Meyers Squibb and Novartis, and Honoraria from Bristol-Meyers Squibb, Novartis, Pfizer and Amgen. None of these agencies have had a role in the preparation of this manuscript. All other authors declare that they have no competing interests.

References

1. Inaba H and Mullighan CG: Pediatric acute lymphoblastic leukemia. Haematologica 105: 2524-2539, 2020.

2. Shallis RM, Wang R, Davidoff A, Ma X and Zeidan AM: Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges. Blood Rev 36: 70-87, 2019.

3. Stengel A, Kern W, Haferlach T, Meggendorfer M, Fasan A and Haferlach C: The impact of TP53 mutations and TP53 deletions on survival varies between AML, ALL, MDS and CLL: An analysis of 3307 cases. Leukemia 31: 705-711, 2017.

4. Stengel A, Schnittger S, Weissmann S, Kuznia S, Kern W, Kohlmann A, Haferlach T and Haferlach C: TP53 mutations occur in 15.7% of ALL and are associated with MYC-rearrangement, low hypodiploidy, and a poor prognosis. Blood 124: 251-258, 2014.

5. Kantidze OL, Luzhin AV, Nizovtseva EV, Safina A, Valieva ME, Golov AK, Velichko AK, Lyubitelev AV, Feofanov AV, Gurova KV, et al: The anti-cancer drugs curaxins target spatial genome organization. Nat Commun 10: 1441, 2019.

6. Dallavalle S, Mattio LM, Artali R, Musso L, Aviñó A, Fàbrega C, Eritja R, Gargallo R and Mazzini S: Exploring the interaction of curaxin CBL0137 with G-quadruplex DNA oligomers. Int J Mol Sci 22: 6476, 2021.

7. Gasparian AV, Burkhart CA, Purmal AA, Brodsky L, Pal M, Saranadasa M, Bosykh DA, Commane M, Guryanova OA, Pal S, et al: Curaxins: Anticancer compounds that simultaneously suppress NF-κB and activate p53 by targeting FACT. Sci Transl Med 3: 95ra74, 2011.

8. Lock R, Carol H, Maris JM, Kolb EA, Gorlick R, Reynolds CP, Kang MH, Keir ST, Wu J, Purmal A, et al: Initial testing (stage 1) of the curaxin CBL0137 by the pediatric preclinical testing program. Pediatr Blood Cancer 64: e26263, 2017.

9. Somers K, Kosciolek A, Bongers A, El-Ayoubi A, Karsa M, Mayoh C, Wadham C, Middlemiss S, Neznanov N, Kees UR, et al: Potent antileukemic activity of curaxin CBL0137 against MLL-rearranged leukemia. Int J Cancer 146: 1902-1916, 2020.

10. Forgione MO, McClure BJ, Eadie LN, Yeung DT and White DL: KMT2A rearranged acute lymphoblastic leukaemia: Unravelling the genomic complexity and heterogeneity of this high-risk disease. Cancer Lett 469: 410-418, 2020.

11. Muntean AG and Hess JL: The pathogenesis of mixed-lineage leukemia. Annu Rev Pathol 7: 283-301, 2012.

12. Marks DI, Moorman AV, Chilton L, Paietta E, Enshaie A, DeWald G, Harrison CJ, Fielding AK, Foroni L, Goldstone AH, et al: The clinical characteristics, therapy and outcome of 85 adults with acute lymphoblastic leukemia and t(4;11)(q21;q23)/MLL-AFF1 prospectively treated in the UKALLXII/ECOG2993 trial. Haematologica 98: 945-952, 2013.

13. Aubrey BJ, Kelly GL, Kueh AJ, Brennan MS, O'Connor L, Milla L, Wilcox S, Tai L, Strasser A and Herold MJ: An inducible lentiviral guide RNA platform enables the identification of tumor-essential genes and tumor-promoting mutations in vivo. Cell Rep 10: 1422-1432, 2015.

14. Ralph P, Moore M and Nilsson K: Lysozyme synthesis by established human and murine histiocytic lymphoma cell lines. J Exp Med 143: 1528-1533, 1976.

15. Chanput W, Peters V and Wichers H: THP-1 and U937 cells. In: The Impact of Food Bioactives on Health. Verhoeckx K, Cotter P, López-Expósito I, Kleiveland C, Lea T, Mackie A, Requena T, Swiatecka D and Wichers H (eds). Springer, Cham, pp147-159, 2015.

16. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehár J, Kryukov GV, Sonkin D, et al: The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 483: 603-607, 2012.

17. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, et al: COSMIC: The catalogue of somatic mutations in cancer. Nucleic Acids Res 47 (D1): D941-D947, 2019.

18. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W, et al: ClinVar: Improving access to variant interpretations and supporting evidence. Nucleic Acids Res 46 (D1): D1062-D1067, 2018.

19. Zhang JD, Ruschhaupt M and Biczok R: ddCt method for qRT-PCR data analysis. Bioconductor, 2013. http://www.bioconductor.org/packages/release/bioc/vignettes/ddCt/inst/doc/rtPCR.pdf. Accessed October 26, 2021.

20. He G, Siddik ZH, Huang Z, Wang R, Koomen J, Kobayashi R, Khokhar AR and Kuang J: Induction of p21 by p53 following DNA damage inhibits both Cdk4 and Cdk2 activities. Oncogene 24: 2929-2943, 2005.

21. Xia M, Knezevic D and Vassilev LT: p21 does not protect cancer cells from apoptosis induced by nongenotoxic p53 activation. Oncogene 30: 346-355, 2011.

22. Sarantopoulos J, Mahalingam D, Sharma N, Iyer RV, Ma WW, Ahluwalia MS, Johnson S, Purmal A, Shpigotskaya P, Hards A, et al: Results of a completed phase I trial of CBL0137 administered intravenously (IV) to patients (Pts) with advanced solid tumors. J Clin Oncol 38 (Suppl 15): S3583, 2020.

23. Fedyanin M, Tryakin A, Lisyanskaya AS, Solovyeva E, Fadeeva N, Gladkov O, Moiseyenko V, Cheporov SV, Shpigotskaya P, Purmal A, et al: Results of a completed first-in-human phase Ib dose-escalation study of oral CBL0137 in patients with advanced solid tumors. J Clin Oncol 38 (Suppl 15): S3607, 2020.

24. Chua MMJ, Lee M and Dominguez I: Cancer-type dependent expression of CK2 transcripts. PLoS One 12: e0188854, 2017.

Figure S1. Sanger sequencing reveals heterozygous *TP53* deletion and frameshift mutations induced in CRISPR/Cas9-edited RS4;11 cells. Genomic DNA was extracted from RS4;11 cell lines transduced with *TP53* sgRNA constructs, following doxycycline activation. *TP53* exon 4 was PCR-amplified and purified, and Sanger sequencing of the purified PCR product verified the presence of mutations in cells transduced with (A) *TP53* sgRNA #1 and (B) *TP53* sgRNA #2. A heterogenous population of alterations were present in both cell lines, where an indel size of 0 denotes wild-type sequence. In *TP53* sgRNA #1 cells, the dominant isoform is p.Ala78_Pro87del (40% of cells). In *TP53* sgRNA #2 cells, the dominant isoform is p.Lys101?fs*192 (33% of cells). sgRNA, single guide RNA.
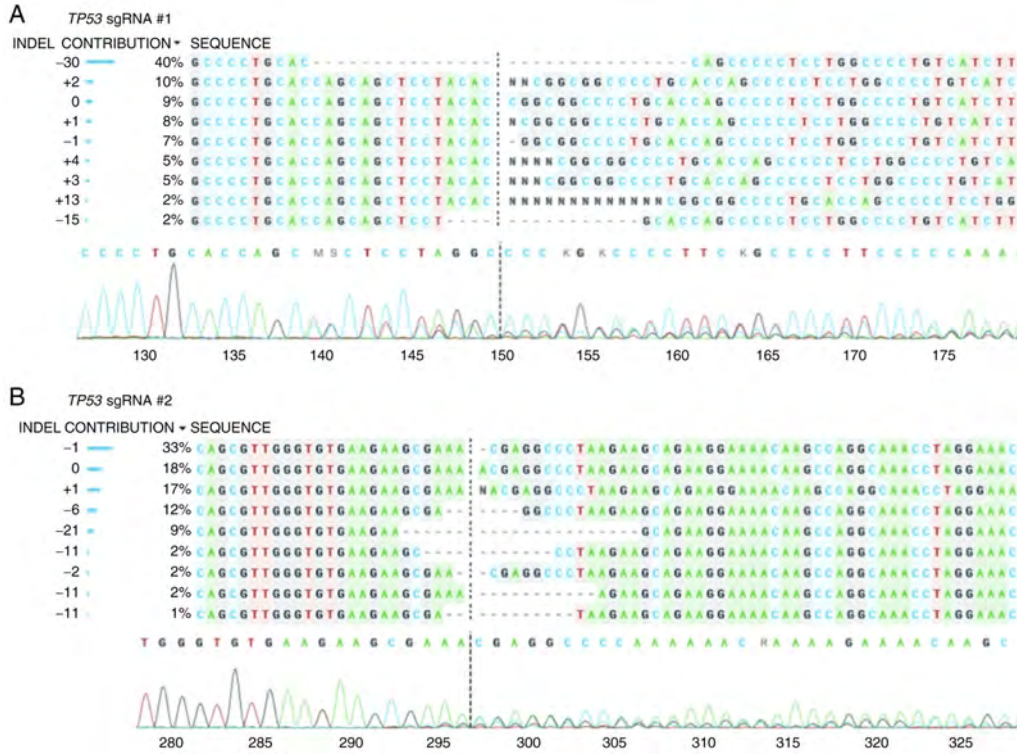
Table SI. One-way ANOVA reveals significant differences between TP53-mutated and TP53 wild-type human leukaemia cell lines.

Cell lines (A) vs. (B)	logLC$_{50}$ (A) (nM)	logLC$_{50}$ (B) (nM)	Mean difference	Significance	P-value
RS4;11 vs. NALM-19	2.222	2.386	-0.1642	No	0.4579
RS4;11 vs. MV4;11	2.222	2.430	-0.2078	No	0.1376
RS4;11 vs. RCH-ACV	2.222	2.476	-0.2535	Yes	0.0260
RS4;11 vs. REH	2.222	2.652	-0.4298	Yes	0.0023
RS4;11 vs. Jurkat	2.222	2.719	-0.4968	Yes	<0.0001
RS4;11 vs. THP-1	2.222	2.833	-0.6107	Yes	<0.0001
RS4;11 vs. U-937	2.222	2.749	-0.5269	Yes	<0.0001
NALM-19 vs. MV4;11	2.386	2.430	-0.04359	No	0.4579
NALM-19 vs. RCH-ACV	2.386	2.476	-0.08931	No	0.1376
NALM-19 vs. REH	2.386	2.652	-0.2656	Yes	0.0209
NALM-19 vs. Jurkat	2.386	2.719	-0.3326	Yes	0.0015
NALM-19 vs. THP-1	2.386	2.833	-0.4465	Yes	<0.0001
NALM-19 vs. U-937	2.386	2.749	-0.3627	Yes	0.0004
MV4;11 vs. RCH-ACV	2.430	2.476	-0.04572	No	0.4579
MV4;11 vs. REH	2.430	2.652	-0.2220	No	0.1171
MV4;11 vs. Jurkat	2.430	2.719	-0.2890	Yes	0.0150
MV4;11 vs. THP-1	2.430	2.833	-0.4029	Yes	0.0004
MV4;11 vs. U-937	2.430	2.749	-0.3191	Yes	0.0047
RCH-ACV vs. REH	2.476	2.652	-0.1763	No	0.4095
RCH-ACV vs. Jurkat	2.476	2.719	-0.2433	No	0.0909
RCH-ACV vs. THP-1	2.476	2.833	-0.3572	Yes	0.0047
RCH-ACV vs. U-937	2.476	2.749	-0.2734	Yes	0.0373
REH vs. Jurkat	2.652	2.719	-0.06705	No	0.3865
REH vs. THP-1	2.652	2.833	-0.1809	Yes	0.0455
REH vs. U-937	2.652	2.749	-0.09715	No	0.2085
Jurkat vs. THP-1	2.719	2.833	-0.1139	No	0.2568
Jurkat vs. U-937	2.719	2.749	-0.03010	No	0.6952
THP-1 vs. U-	2.833	2.749	0.08377	No	0.4579

937					

Human acute leukaemia cell lines were exposed to increasing doses of CBL0137 in duplicate and apoptosis was measured by Annexin-V/7-AAD staining after 72 h. Mean CBL0137 $logLC_{50}$ values were calculated from 3 independent biological replicates. Kruskal-Wallis tests with multiple comparisons tests were performed to compare mean $logLC_{50}$ values between each cell line tested. Comparisons between TP53-wild-type and TP53-mutated cell lines are highlighted in yellow.

2

Table SII. LogLC$_{50}$ and LC$_{50}$ for CBL0137 in acute leukaemia cell lines.

Cell line	logLC$_{50}$ (nM)	LC$_{50}$ (nM)
RS4;11	2.22	166.0
RCH-ACV	2.48	302.0
NALM-19	2.39	245.5
MV4;11	2.43	269.2
REH	2.65	446.7
Jurkat	2.72	524.8
U-937	2.75	562.3
THP-1	2.83	676.1
RS4;11 Cas9	2.32	208.9
RS4;11 $TP53^{MUT}$ #1	2.59	389.1
RS4;11 $TP53^{MUT}$ #2	2.62	416.9

Human acute leukaemia cell lines were exposed to increasing doses of CBL0137 in duplicate and apoptosis was measured by Annexin-V/7-AAD staining after 72 h. Mean CBL0137 logLC$_{50}$ values were calculated from 3 independent biological replicates. LogLC$_{50}$ and LC$_{50}$ values were extrapolated using non-linear regression analysis.

3

Chapter 7 – Discussion & Concluding Remarks

7.1 Introduction

Rearrangements of the *KMT2A* gene (*KMT2A*r) induce highly aggressive ALL associated with poor outcomes, yet little is known about the biology of *KMT2A*r ALL. *MLLT10* rearrangements (*MLLT10*r) are another subtype of ALL with poorly characterised biology, despite being one of the most common fusion genes identified in T-ALL (approximately 10% of new T-ALL diagnoses). *KMT2A*r and *MLLT10*r leukaemic blasts both feature transcriptional dysregulation of the *HOXA* cluster, suggesting a similar aetiology. An improved understanding of the biology of *KMT2A*r and *MLLT10*r ALL is required to develop precision medicine approaches to improve outcomes and limit toxicities for these subtypes.

The research presented in this thesis:

1. Generates and compares *in vitro* models of *KMT2A*r B-ALL and T-ALL
2. Generates and characterises *in vitro* and *in vivo* models of *MLLT10*r T-ALL
3. Interrogates the genomic landscape of non-infant *KMT2A*r B-ALL and *MLLT10*r ALL
4. Establishes the clinical scope of the curaxin CBL0137, an inhibitor with pre-clinical promise in *KMT2A*r infant ALL

7.2 Major findings

7.2.1 Parent cell identity is critical in modelling *KMT2A-AFF1 in vitro*

The data presented in chapter 3 support the notion that parent cell identity and maturity are critical in the transformative capacity of *KMT2A-AFF1*. *KMT2A-AFF1* expression in B-lineage (Ba/F3) and T-lineage (MOHITO) cell lines induced differential changes in proliferation, *HOXA* gene expression and surface marker expression. The *in vitro* models generated and characterised in this thesis are summarised in **Figure 7.1**.
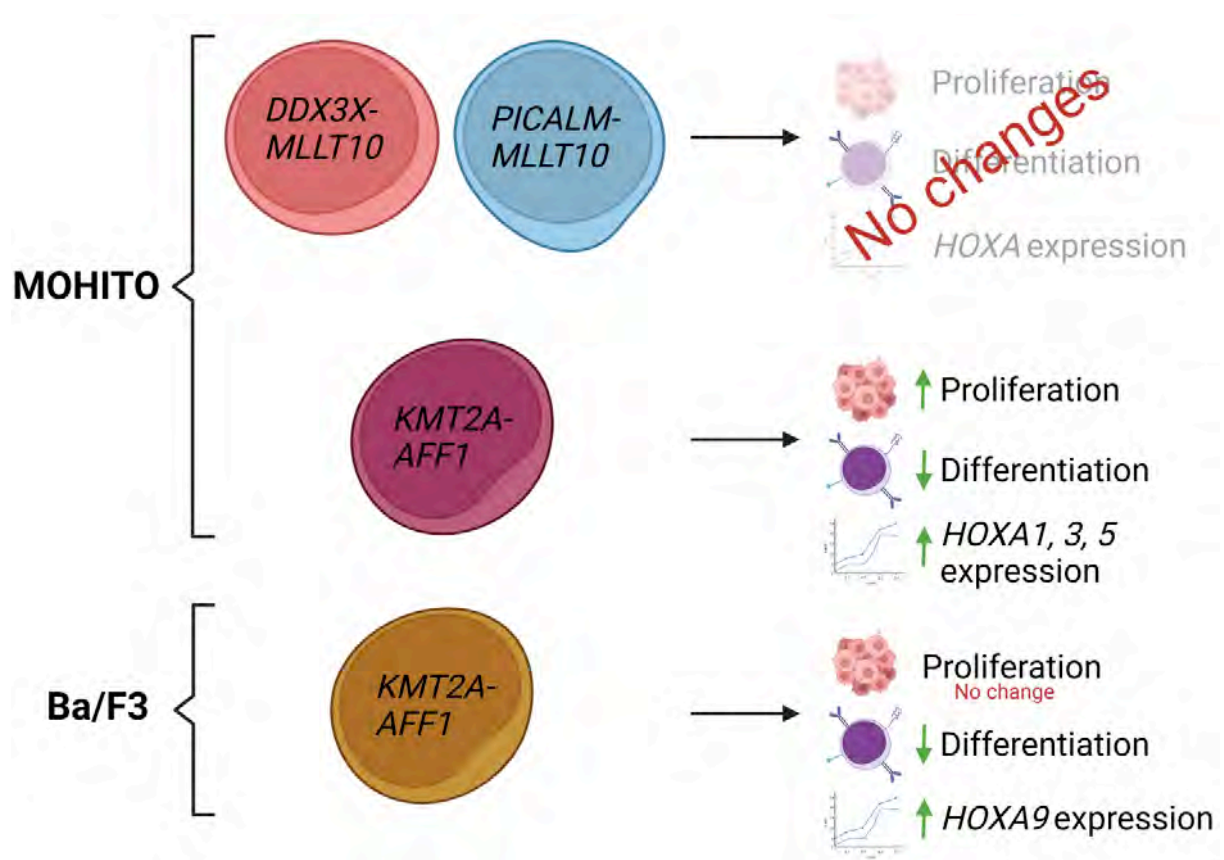

Figure 7.1 – Summary of findings from generation and interrogation of *in vitro* *MLLT10*r and *KMT2A-AFF1* models

The MOHITO and Ba/F3 murine cell lines were utilised to generate *in vitro* models of *DDX3X-MLLT10, PICALM-MLLT10* and *KMT2A-AFF1* B-ALL and T-ALL. The ability of fusion expression to induce changes in cellular proliferation, differentiation (surface marker expression) and *HOXA* expression are summarised. Image created with BioRender.com.

Extensive clinical data has clearly established *KMT2A-AFF1* as an aggressive driver of pro-B ALL, yet no convincing evidence of oncogenic transformation was observed in pro-B Ba/F3 cells, other than increased expression of *HOXA9.* While Ba/F3 cells are a useful and widely accepted system used for *in vitro* modelling of kinase-associated genomic alterations, they are not appropriate for studying *KMT2A*r, where leukaemic transformation likely occurs through influencing cellular differentiation rather than kinase signalling pathways. This is a critical point, as it emphasises an important limitation of the Ba/F3 system in testing the oncogenicity of genomic aberrations.

KMT2A-AFF1 expression induced hallmark indications of transformation in the murine CD4[+]CD8[+] T-ALL MOHITO cell line, most notably increased cytokine-dependent cellular proliferation and upregulation of the *HOXA* genes *HOXA1, HOXA3* and *HOXA5*, and *HOX* cofactor *MEIS1*. *KMT2A-AFF1* expression induced an immature T-lineage immunophenotype, characterised by loss of mature T-cell surface markers, CD4 and CD8, and increased expression of the immature T-cell marker CD44. High CD44 expression is associated with poor outcomes in T-ALL[1] as well as other haematopoietic and solid tumours[2-4], while the prognostic significance of CD4 and CD8 expression in T-ALL patients is unknown. Importantly, this is the first described *in vitro* model of *KMT2A-AFF1* T-ALL, which will be particularly useful in the pre-clinical screening of targeted therapies, to improve therapeutic options for patients with *KMT2A*r T-ALL.

The different phenotypic outcomes induced by *KMT2A-AFF1* expression in MOHITO and Ba/F3 cells is likely a consequence of parent cell differentiation stage and background genetic alterations, rather than cell lineage. It is clear from clinical data that *KMT2A-AFF1* is a recurrent lesion that induces aggressive B-ALL. It is likely that the different maturity stages of Ba/F3 and MOHITO cells is the underlying cause of the transformative discrepancies observed, where expression of stem cell markers such as Sca-1, c-Kit and/or CD34 are likely to be important in the initiation of *KMT2A*r ALL[5]. Indeed, expression of the haematopoietic stem cell markers c-Kit and Sca-1 in MOHITO cells but not Ba/F3 cells may be a contributing factor to the differences in leukaemic transformation induced by *KMT2A-AFF1* expression. It is also possible that cell-type specific collaborative genomic events may influence the phenotype induced by *KMT2A-AFF1* expression. For example, *NOTCH1-3* and *STAT5* genes are involved in maintaining quiescence of HSCs, but promote survival and proliferation in other

haematopoietic cell types[6,7], suggesting that some genomic alterations identified in ALL are highly cell type-specific.

Another important consideration when generating pre-clinical models of *KMT2A*r ALL is the role of the reciprocal *AFF1-KMT2A* fusion. An in-frame reciprocal fusion is identified in most *KMT2A-AFF1* B-ALL patients[8], including 89.4% (17/19) of *KMT2A-AFF1* cases described in chapter 5. In instances where reciprocal expression was not detected, it is possible that fusions were complex rearrangements (for example, 3-way translocations that are common with *KMT2A-MLLT1*[9]) rather than direct reciprocal translocations. Other groups have demonstrated that *AFF1-KMT2A* expression is critical in the transformation of *KMT2A*r B-ALL[8]. Transplantation of Lin⁻Sca-1⁺ murine stem cells expressing *AFF1-KMT2A* alone induces pro-B ALL when transplanted into mice, albeit with a low penetrance (35%) and long latency (median = 233 days)[10]. In this study, expression of both reciprocal fusions together induced a less mature B/T precursor or mixed lineage leukaemia with similarly low penetrance and delayed latency, whereas *KMT2A-AFF1* expression alone did not induce disease at all[10]. This study also emphasises the importance of parent cell maturity when modelling *KMT2A*r ALL, as Lin⁻Sca-1⁺ cells were utilised, whereas Ba/F3 cells are Lin⁺Sca-1⁻. Recent work by Rice *et al.* describes the first published pre-clinical model of *KMT2A-AFF1* that recapitulates aggressive B-ALL[11]. In this study, t(4;11) was induced in CD34⁺ human foetal liver cells using CRISPR/Cas9, resulting in expression of both *KMT2A-AFF1* and *AFF1-KMT2A* and inducing pro-B ALL when transplanted into mice, with a median latency of 18 weeks (approximately 126 days)[11]. The reciprocal fusion was not expressed in the *in vitro* models described in this thesis, which may explain the limited transformation observed in Ba/F3 cells. In light of these recent studies, it would be essential to establish whether *AFF1-KMT2A* expression in MOHITO cells induces further leukaemic changes, such as inducing cytokine-independent proliferation, *HOXA9* upregulation or a shortening of disease latency in an *in vivo* transplantation model.

The Rice *et al.* model also addressed the *in utero* origin of *KMT2A-AFF1* in infants[12], by modelling the fusion in human foetal liver cells, a source of primitive progenitors. This model supports the notion that a foetal-specific gene expression signature is critical in the development of aggressive and therapy-resistant infant *KMT2A*r ALL, which is distinct from the gene expression signature of childhood *KMT2A*r ALL[10]. This

is supported by a study by Barrett *et al.*, where conditional *KMT2A-AFF1* expression enhanced murine B lymphoid repopulation and self-renewal capacity most prominently in the primitive lymphoid-primed multi-potential progenitor (LMPP) population[13]. *KMT2A-AFF1* expression in human foetal liver derived LMPPs induces a distinct gene expression signature indicative of highly proliferative and oncogenic nature, including upregulation of cyclins and cyclin-dependent kinases, MTORC1 signalling pathway, and the proto-oncogenes *MYC* and *MYCN*[14]. However, *KMT2A*r ALL is also an aggressive disease in non-infant patients, so there is a clear need to continue to seek alternate pre-clinical models to investigate this entity.

The results described in this thesis have contributed to the complex challenge of pre-clinically modelling *KMT2A*r ALL. Critically, it is demonstrated that *KMT2A-AFF1* does not transform Ba/F3 cells but does induce leukaemic changes in the MOHITO T-cell line, providing a potential model to investigate inhibitor efficacy in a *KMT2A*r T-ALL setting. Further studies addressing the role of reciprocal fusion expression in B-lineage, T-lineage and non-lineage committed parent cells is required elucidate the respective importance of reciprocal fusion expression and parent cell maturity.

7.2.2   Arf⁻/⁻ thymocytes expressing *PICALM-MLLT10* or *DDX3X-MLLT10* induce T-ALL *in vivo*

The two most common *MLLT10*r identified in T-ALL, *PICALM-MLLT10* and *DDX3X-MLLT10*, were expressed in MOHITO cells to establish a model that would enable the exploration of biology and therapeutic sensitivity in a T-ALL context. No changes were observed in proliferation or *HOXA* dysregulation due to expression of either *MLLT10*r. Cells expressing *PICALM-MLLT10* exhibited reduced expression of CD5 but retained expression of CD4 and CD8 in comparison to cells expressing *DDX3X-MLLT10*, which was more indicative of an ETP-like immunophenotype. Overall, no clear evidence of *in vitro* leukaemic transformation was induced in MOHITO cells by *MLLT10*r expression, indicating a need to explore alternate cellular model systems or co-expression of additional genomic aberrations (refer to section **7.2.3**).

DDX3X-MLLT10 and *PICALM-MLLT10* were expressed in primary Arf⁻/⁻ thymocytes and transplanted into NSG mice to assess *in vivo* leukaemic potential. The Arf⁻/⁻ thymocyte system was utilised as it enabled the use of primary thymus-derived cells,

as opposed to the spleen-derived MOHITO cell line[15]. The Arf[-/-] system also enabled the collection of CD4[-]CD8[-] cells, that are less mature than the CD4[+]CD8[+] MOHITOs, although some CD8 and CD4 expression was re-gained prior to injection of transduced cells. A summary of the findings obtained from *in vivo* experiments is summarised in **Figure 4.2**.
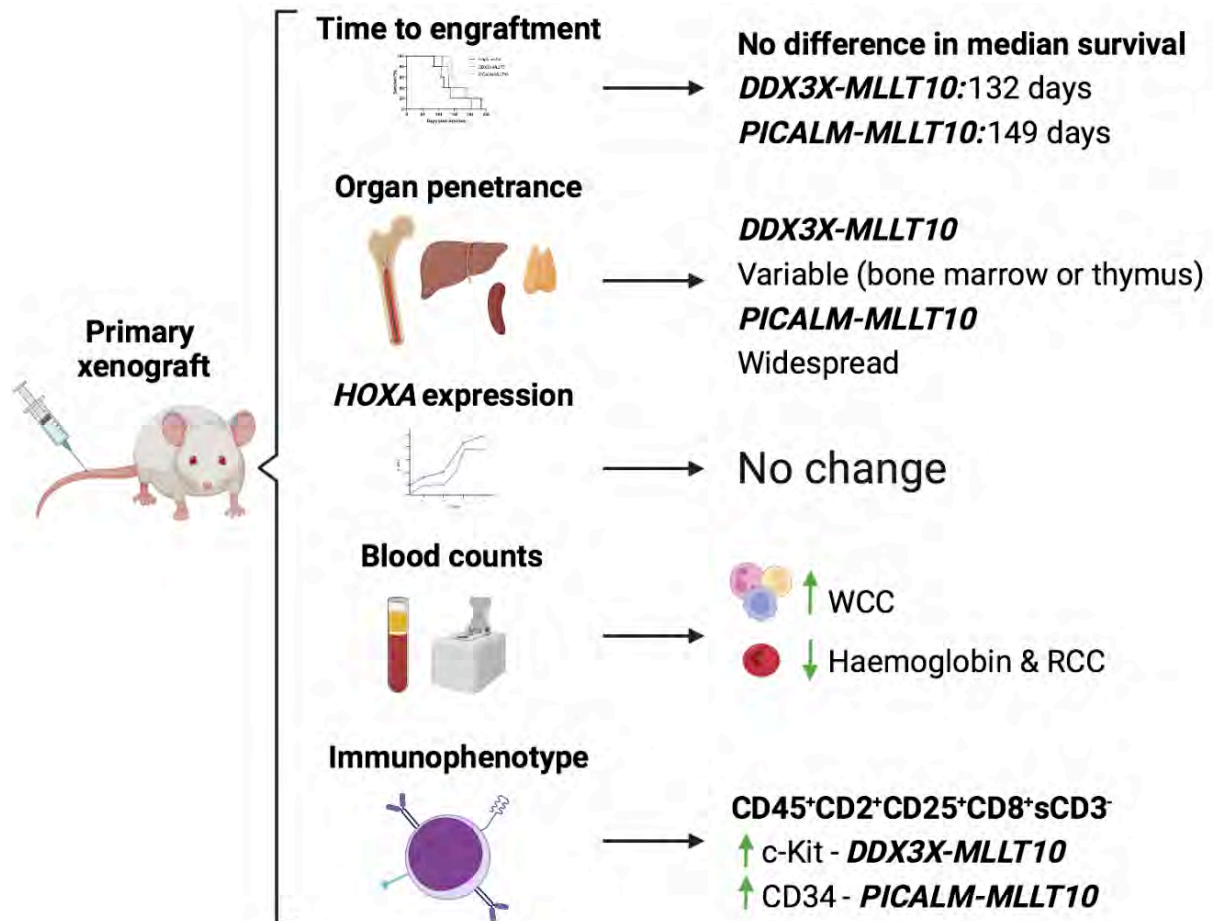

Figure 4.2 – Summary of findings from generation and interrogation of *in vitro* MLLT10r and *KMT2A-AFF1* models

DDX3X-MLLT10 and *PICALM-MLLT10* were expressed in primary Arf[-/-] thymocytes and transplanted into NSG mice to assess *in vivo* leukaemic potential. The ability of fusion expression to induce changes in time to engraftment, organ penetrance, *HOXA* expression, blood counts and immunophenotype are summarised. Image created with BioRender.com.

Transplantation of *PICALM-MLLT10* expressing cells induced a widely disseminated malignancy with involvement of the spleen, liver, bone marrow and thymus, whereas mice transplanted with *DDX3X-MLLT10* expressing cells induced disease that infiltrated predominantly either the bone marrow or thymus. It is interesting that *DDX3X-MLLT10* and *PICALM-MLLT10* expression induced distinct patterns in cell engraftment, since all *MLLT10*r are classified as the same entity in T-ALL patients (typically under the subtype of '*HOXA*-deregulated T-ALL'), regardless of 5' fusion partner. A recent publication demonstrates that *DDX3X-MLLT10* and *PICALM-MLLT10* lineage-ambiguous leukaemias cluster distinctly based on t-SNE analysis of gene expression data[16]. Taken together, these data suggest that the 5' fusion partner may play a role in the engraftment of *MLLT10*r leukaemia, and further investigation is warranted to understand the influence of the 5' *MLLT10*r fusion partner on biology and patient outcomes. This newly-developed *in vivo* model of *DDX3X-MLLT10* and *PICALM-MLLT10* will aid in further understanding the role of the 5' fusion partner. Specifically, other *MLLT10*r can be modelled, such as *HNRNPH1-MLLT10* and *KMT2A-MLLT10*, and ChIP-seq or mRNA-seq can be utilised to establish the differential gene expression profiles induced by *MLLT10*r in a T-ALL context, and the differences between thymic and bone marrow disease in the *DDX3X-MLLT10* model.

No changes in *HOXA* gene expression were observed in MOHITOs or Arf$^{-/-}$ thymocytes expressing *MLLT10*r, in comparison to empty vector cells. An important difference between the MOHITO and Arf$^{-/-}$ cells is that MOHITO cells have very low expression of *HOXA* genes, whereas primary Arf$^{-/-}$ thymocytes have intermediate expression levels. This may be a contributing factor in the ability of Arf$^{-/-}$ thymocytes expressing *MLLT10*r to engraft and induce leukaemia without requiring a further increase in *HOXA* expression. Alternatively, perhaps a less mature cell of origin is required to induce *HOXA*-dysregulated *MLLT10*r T-ALL. Expression of *MLLT10*r in CD34$^+$ lymphoid lineage-primed cells or haematopoietic stem cells (such as cord blood or foetal liver cells) may provide further insight into the role of *HOXA* dysregulation in *MLLT10*r T-ALL, as haematopoietic stem cells have higher expression of *HOXA* genes than lymphoid precursors[17].

There were some similarities and differences in surface marker expression between the MOHITO and Arf$^{-/-}$ *MLLT10*r *in vitro* models. *In vitro*, both models were CD45$^+$CD2$^+$CD25$^+$CD8$^+$sCD3$^-$, indicative of a T-lineage immunophenotype. *PICALM-*

MLLT10 expressing cells exhibited reduced CD5 expression relative to *DDX3X-MLLT10* cells *in vitro* but not *in vivo*, suggesting loss of CD5 expression is not important in the engraftment of *PICALM-MLLT10* disease. Stem cell marker expression was widely variable across both models. Most prominently, CD34 expression was increased in *PICALM-MLLT10* transplanted cells, indicating an important role for this marker in T-ALL engraftment. CD34 is the most frequently identified stem cell marker expressed on ALL blasts, present in >50% of T-ALL cases[18,19]. One study concludes that CD34⁺CD7⁺ T-ALL cells are enriched in leukemia-initiating activity compared with CD34⁻CD7⁺. In this study, CD34⁺CD7⁺ leukaemic blasts engrafted more readily into NSG mice than CD34⁻CD7⁺ cells, despite harbouring the same genomic abnormality (*SIL-TAL1*)[20]. This indicates that CD34 expression results in a more aggressive T-ALL phenotype, which is supported by a clinical correlation between CD34 expression and poor 5-year overall survival in paediatric T-ALL[21]. However, no clinical data has yet established whether CD34 expression varies between genomic subtypes of T-ALL. It is also important to note that the NSG mouse strain is by necessity a severely immunocompromised host, which is very different to the immune system of human hosts.

It is evident that parent cell differentiation stage is important in *MLLT10*r leukaemogenesis, similar to *KMT2A*r ALL. *MLLT10* and *KMT2A* are functionally linked, as they are both epigenetic regulators of cellular differentiation associated with regulation of *HOXA* expression[22]. However, *MLLT10*r do not occur in B-ALL or infant leukaemia (other than *KMT2A-MLLT10*), indicating there must be some difference in transformation biology[23]. The *in vivo* model described in this thesis is the first known *in vivo* model of *MLLT10*r T-ALL, that has provided insight into the distinctly aggressive nature of *PICALM-MLLT10* driven T-ALL, suggesting that further studies comparing *MLLT10*r with different 5' fusion partners is worthwhile.

7.2.3 Co-operative alterations are present at diagnosis in *MLLT10*r ALL

At least one pathogenic leukaemia-associated cooperative alteration was identified in all *MLLT10*r cases where MLPA analysis was available. Most frequently these were *NOTCH1* activating mutations (71.4% or 5/7 patients), *RAS/JAK3* activating mutations (42.9% or 3/7), *PTEN* frameshift or non-frameshift INDELs (42.9% or 3/7), and/or

CDKN2A/B deletions (33.3% or 2/6). This suggests that a range of genomic alterations are important to consider when generating pre-clinical models of *MLLT10*r ALL. Five additional pathogenic variants within epigenetic regulatory genes were identified across four *MLLT10*r patients (affecting the genes *DNMT3A*, *EZH2* and *HDAC3*). These alterations are sensitive to clinically available inhibitors, but the role of these variants is presently unknown in the aetiology of T-ALL.

MOHITO cells and Arf$^{-/-}$ thymocytes both harbour clinically relevant genomic alterations. MOHITO cells are homozygous for the *JAK1* p.S1042I variant that transforms Ba/F3 but not MOHITO cells to cytokine independence[15]. MOHITO cells also possess the pathogenic *NOTCH1* p.I1708S variant within the heterodimerisation (HD) domain, and a truncating frameshift mutation (7463-64insC) in the PEST domain, that both result in increased NOTCH1 activity[15,24]. The data presented in this thesis indicates that *JAK1* and *NOTCH1* activation does cooperate with *MLLT10*r to induce T-ALL in MOHITO cells *in vitro*.

Arf$^{-/-}$ thymocytes are homozygous for loss of murine p19^{Arf}, a tumour suppressor encoded by *CDKN2A*[25]. Two *MLLT10*r patients (33.3%) in our cohort presented with whole gene deletions of *CDKN2A*, and consequently lack expression of p14^{Arf} (the human equivalent of p19^{Arf}) and p16^{Ink4a}. The successful *in vivo* engraftment of *PICALM-MLLT10* and *DDX3X-MLLT10* expressing cells demonstrates that this combination of genomic events is sufficient to induce T-ALL in this system. However, it is necessary to repeat the *in vivo* experiment to establish the ability of empty vector Arf$^{-/-}$ cells to engraft, as 3/5 empty vector mice died due to a likely familial abnormality. A similar model should also be constructed to determine the cooperative potential of other alterations such as *NOTCH1* or *RAS* pathway activation in a *MLLT10*r context, as not all *MLLT10*r patients present with *CDKN2A* deletion. *CDKN2B* was also deleted in both *MLLT10*r patients with *CDKN2A* deletion, where *CDKN2B* encodes the tumour suppressor p15 that also promotes leukaemogenesis[26], and the role of *CDKN2B* deletion in *MLLT10*r ALL should therefore also be investigated.

7.2.4   Cooperative leukaemia-associated genomic alterations are rare at diagnosis in non-infant *KMT2A*r B-ALL

The results presented in chapter 5 demonstrate that the genomic landscape of non-infant *KMT2A*r B-ALL is largely devoid of known cooperative leukaemia-associated CNAs and SNVs, in comparison to an age-matched *BCR-ABL1* B-ALL cohort. This demonstrated that non-infant *KMT2A*r ALL has mutational landscape that is similar to that of infant *KMT2A*r B-ALL, despite the different gene expression signatures identified in infant and paediatric *KMT2A*r ALL[11].

Interrogation of epigenetic regulatory genes revealed three recurrent SNVs in *KMT2A*r B-ALL, that were either uncommon or absent from the *BCR-ABL1* cohort (*H1F0* p.A89P, *MCM7* p.A433P and *WHSC1* p.E1099K). *KMT2A*r are known to induce widespread epigenetic dysregulation, which has generally been attributed to the direct dysregulation of WT *KMT2A*[27,28]. *WHSC1* p.E1099K is known to shorten *in vivo* ALL latency in both *KMT2A*r and *KMT2A*-WT cell lines[29], indicating it is significant in the initiation of ALL. These data demonstrate that while *KMT2A*r B-ALL lacks genomic alterations that are common in other subtypes of B-ALL, variants within epigenetic regulators are disproportionately common, that may contribute to the unique biology underlying *KMT2A*r ALL.

7.2.5   CBL0137 is a promising ALL therapy in *TP53*^WT disease, independent of *KMT2A*r status

In response to two recent publications that cite particular sensitivity of *KMT2A*r infant ALL PDX models to the curaxin CBL0137[30,31], the role of this promising targeted therapy in acute leukaemia was characterised. Findings presented here demonstrate that CBL0137 does not specifically confer therapeutic sensitivity to *KMT2A*r ALL, relative to other genomic subtypes of ALL. Rather, using *in vitro* sensitivity assays, it was confirmed that wild-type p53 is required for optimal activity of CBL0137, demonstrating that CBL0137 may serve as a useful drug in *TP53*^WT ALL overall.

TP53 mutations are rare in infants with *KMT2A*r ALL, identified in 4.3% (2/47) of infant *KMT2A*r ALL cases at diagnosis in one study[32]. In non-infants, *TP53* mutations occur in 15.7% of new ALL diagnoses overall, but variants are often subclonal at diagnosis

and emerge in the dominant clone at relapse[33]. In the non-infant *KMT2A*r B-ALL cohort described in this thesis, only one *TP53* variant (p.R248Q) was identified in one case of non-infant *KMT2Ar* ALL (overall incidence = 4.4%), noting indels were not investigated. The fact that most *KMT2A*r patients are *TP53*^WT indicates that CBL0137 is a promising therapeutic strategy. These findings emphasise the importance of a genome-wide approach to understanding the complete mutational landscape of each ALL patient, to optimise precision medicine approaches.

7.2.6 Concluding remarks

In this thesis, the first *in vitro* model of *KMT2A*r T-ALL was established. This demonstrated that *KMT2A-AFF1* expression is sufficient to induce oncogenic changes in MOHITO but not Ba/F3 cells, emphasising the importance of parent cell maturity in initiation of *KMT2A*r ALL. The first *in vivo* model of *MLLT10*r T-ALL was also generated, demonstrating that *DDX3X-MLLT10* and *PICALM-MLLT10* induce T-ALL with distinct immunophenotypes and organ penetrance. The landscape of genomic alterations within non-infant *KMT2A*r and *MLLT10*r ALL was interrogated, and it was established that while *KMT2A*r ALL carries a very low burden of typical leukaemia-associated alterations, known and novel variants within epigenetic regulators were identified in *KMT2A*r B-ALL and *MLLT10*r T-ALL, some of which are therapeutically targetable by clinically available inhibitors. In *TP53*^WT infant and non-infant *KMT2A*r ALL , therapeutic sensitivity to the curaxin and p53 activator CBL0137 was noted, highlighting the importance a complete genomic profile for precision medicine approaches. Novel approaches are required to accurately model genomic subtypes of ALL that rely on dysregulation of cellular differentiation, unlike classical kinase-associated alterations. Specifically, models should use immature/foetal-derived haematopoetic stem and progenitor cells, and address the importance of reciprocal fusion expression and co-occurring genomic alterations. Ultimately, these studies give novel insight into the development of disease-representative *in vitro* and *in vivo* models of *KMT2A*r and *MLLT10*r ALL to aid in the screening of targeted therapies, which is essential in improving outcomes and reducing treatment-related toxicities in ALL.

7.3 Chapter References

1   Chagan-Yasutan, H. *et al.* Involvement of osteopontin and its signaling molecule CD44 in clinicopathological features of adult T cell leukemia. *Leuk Res* **35**, 1484-1490, doi:10.1016/j.leukres.2011.05.011 (2011).

2   Marques, L. V. C. *et al.* CD44 Expression Profile Varies According to Maturational Subtypes and Molecular Profiles of Pediatric T-Cell Lymphoblastic Leukemia. *Front Oncol* **8**, 488, doi:10.3389/fonc.2018.00488 (2018).

3   Wang, Z., Zhao, K., Hackert, T. & Zoller, M. CD44/CD44v6 a Reliable Companion in Cancer-Initiating Cell Maintenance and Tumor Progression. *Front Cell Dev Biol* **6**, 97, doi:10.3389/fcell.2018.00097 (2018).

4   Legras, S. *et al.* A Strong Expression of CD44-6v Correlates With Shorter Survival of Patients With Acute Myeloid Leukemia. *Blood* **91**, 3401-3413 (1998).

5   Marschalek, R. Mixed lineage leukemia: roles in human malignancies and potential therapy. *FEBS J* **277**, 1822-1831, doi:10.1111/j.1742-4658.2010.07608.x (2010).

6   Sottoriva, K. & Pajcini, K. V. Notch Signaling in the Bone Marrow Lymphopoietic Niche. *Front Immunol* **12**, 723055, doi:10.3389/fimmu.2021.723055 (2021).

7   Schepers, H., Wierenga, A. T., Vellenga, E. & Schuringa, J. J. STAT5-mediated self-renewal of normal hematopoietic and leukemic stem cells. *JAKSTAT* **1**, 13-22, doi:10.4161/jkst.19316 (2012).

8   Marschalek, R. The reciprocal world of MLL fusions: A personal view. *Biochim Biophys Acta Gene Regul Mech* **1863**, 194547, doi:10.1016/j.bbagrm.2020.194547 (2020).

9   de Matos, R. *et al.* KMT2A-MLLT1 and the Novel SEC16A-KMT2A in a Cryptic 3-Way Translocation t(9;11;19) Present in an Infant With Acute Lymphoblastic Leukemia. *Pediatric Hematology/Oncology* **-**, doi:10.1097/MPH.0000000000002386 (2021).

10   Bursen, A. *et al.* The AF4.MLL fusion protein is capable of inducing ALL in mice without requirement of MLL.AF4. *Blood* **115**, 3570-3579, doi:10.1182/blood-2009-06-229542 (2010).

11   Rice, S. *et al.* A human fetal liver-derived infant MLL-AF4 acute lymphoblastic leukemia model reveals a distinct fetal gene expression program. *Nat Commun* **12**, 6905, doi:10.1038/s41467-021-27270-z (2021).

12   Gale, K. B. *et al.* Backtracking leukemia to birth: Identification of clonotypic gene fusion sequences in neonatal blood spots. *Proceedings of the National Academy of Sciences* **94**, 13950-13954, doi:10.1073/pnas.94.25.13950 (1997).

13   Barrett, N. A. *et al.* Mll-AF4 Confers Enhanced Self-Renewal and Lymphoid Potential during a Restricted Window in Development. *Cell Rep* **16**, 1039-1054, doi:10.1016/j.celrep.2016.06.046 (2016).

14   Symeonidou, V. *et al.* Defining the fetal origin of MLL-AF4 infant leukemia highlights specific fatty acid requirements. *Cell Reports* **37**, 109900, doi:10.1016/j.celrep.2021.109900 (2021).

15   Kleppe, M., Mentens, N., Tousseyn, T., Wlodarska, I. & Cools, J. MOHITO, a novel mouse cytokine-dependent T-cell line, enables studies of oncogenic signaling in the T-cell context. *Haematologica* **96**, 779-783, doi:10.3324/haematol.2010.035931 (2011).

16   Montefiori, L. E. *et al.* Enhancer Hijacking Drives Oncogenic BCL11B Expression in Lineage-Ambiguous Stem Cell Leukemia. *Cancer Discov* **11**, 2846-2867, doi:10.1158/2159-8290.CD-21-0145 (2021).

17   Lebert-Ghali, C. E. *et al.* Hoxa cluster genes determine the proliferative activity of adult mouse hematopoietic stem and progenitor cells. *Blood* **127**, 87-90, doi:10.1182/blood-2015-02-626390 (2016).

18   Tembhare, P. R. *et al.* Eleven-marker 10-color flow cytometric assessment of measurable residual disease for T-cell acute lymphoblastic leukemia using an approach of exclusion. *Cytometry B Clin Cytom* **100**, 421-433, doi:10.1002/cyto.b.21939 (2021).

19   Wang, Y. Z. *et al.* A seven-color panel including CD34 and TdT could be applied in >97% patients with T cell lymphoblastic leukemia for minimal residual disease detection independent of the initial phenotype. *Leuk Res* **72**, 12-19, doi:10.1016/j.leukres.2018.07.012 (2018).

20   Gerby, B. *et al.* Expression of CD34 and CD7 on human T-cell acute lymphoblastic leukemia discriminates functionally heterogeneous cell populations. *Leukemia* **25**, 1249-1258, doi:10.1038/leu.2011.93 (2011).

21   van Grotel, M. *et al.* CD34 expression is associated with poor survival in pediatric T-cell acute lymphoblastic leukemia. *Pediatr Blood Cancer* **51**, 737-740, doi:10.1002/pbc.21707 (2008).

22   Chen, L. *et al.* Abrogation of MLL-AF10 and CALM-AF10-mediated transformation through genetic inactivation or pharmacological inhibition of the H3K79 methyltransferase Dot1l. *Leukemia* **27**, 813-822, doi:10.1038/leu.2012.327 (2013).

23   Forgione, M. O., McClure, B. J., Yeung, D. T., Eadie, L. N. & White, D. L. MLLT10 rearranged acute leukemia: incidence, prognosis and possible therapeutic strategies. *Genes Chromosomes Cancer* **59**, 709-721, doi:10.1002/gcc.22887 (2020).

24   Ferrando, A. A. The role of NOTCH1 signaling in T-ALL. *Hematology Am Soc Hematol Educ Program*, 353-361, doi:10.1182/asheducation-2009.1.353 (2009).

25   Stott, F., Bates, S, James, MC. McConnell, BB, Starborg, M, Brookes, S, Palmero, I, Ryan, K, Hara, E, Vousden, KH, Peters, G. The alternative product from the human CDKN2A locus, p14ARF, participates in a regulatory feedback loop with p53 and MDM2. *The EMBO Journal* **17**, 5011-5014, doi:10.1093/emboj/17.17.5001 (1998).

26   Van Vlierberghe, P. *et al.* Prognostic relevance of integrated genetic profiling in adult T-cell acute lymphoblastic leukemia. *Blood* **122**, 74-82, doi:10.1182/blood-2013-03-491092 (2013).

27   Nordlund, J. & Syvanen, A. C. Epigenetics in pediatric acute lymphoblastic leukemia. *Semin Cancer Biol*, doi:10.1016/j.semcancer.2017.09.001 (2017).

28   Zeisig, B. B. & So, C. W. E. in *Chromosomal Translocations and Genome Rearrangements in Cancer* (eds Janet D. Rowley, Michelle M. Le Beau, & Terence H. Rabbitts) 223-250 (Springer International Publishing, 2015).

29   Swaroop, A. *et al.* An activating mutation of the NSD2 histone methyltransferase drives oncogenic reprogramming in acute lymphocytic leukemia. *Oncogene*, doi:10.1038/s41388-018-0474-y (2018).

30   Somers, K. *et al.* Potent anti-leukemic activity of curaxin CBL0137 against MLL-rearranged leukemia. *Int J Cancer*, doi:10.1002/ijc.32582 (2019).

31   Lock, R. *et al.* Initial testing (stage 1) of the curaxin CBL0137 by the pediatricpreclinical testing program. *Pediatric Blood & Cancer* **64**, e26263, doi:10.1002/pbc.26263 (2016).

32   Andersson, A. K. *et al.* The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nature genetics* **47**, 330 (2015).

33   Stengel, A. *et al.* TP53 mutations occur in 15.7% of ALL and are associated with MYC-rearrangement, low hypodiploidy, and a poor prognosis. *Blood* **124**, 251-258, doi:10.1182/blood-2014-02-558833 (2014).