Tampere University

LAURA OKSA

Novel Genomic Biomarkers for Childhood Acute Lymphoblastic Leukemia

Tampere University Dissertations 898

Tampere University Dissertations 898

LAURA OKSA

Novel Genomic Biomarkers for Childhood Acute Lymphoblastic Leukemia

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Medicine and Health Technology
of Tampere University,
for public discussion in the auditorium
of the Finn-Medi 5, Biokatu 12, Tampere,
on 15 December 2023, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Medicine and Health Technology
Finland

Responsible supervisor and Custos	Docent Olli Lohi Tampere University Finland	
Supervisor	Professor Merja Heinäniemi University of Eastern Finland Finland	
Pre-examiners	Professor Eeva-Riitta Savolainen University of Oulu Finland	Docent Leena Latonen Tampere University Finland
Opponent	Professor Anindita Roy University of Oxford United Kingdom	

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Cover design: Roihu Inc.

ISBN 978-952-03-3143-6 (print)
ISBN 978-952-03-3144-3 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-3144-3

Carbon dioxide emissions from printing Tampere University dissertations have been compensated.

PunaMusta Oy – Yliopistopaino
Joensuu 2023

To my family.

ACKNOWLEDGEMENTS

The research for this thesis was carried out at the Tampere Center for Child, Adolescent, and Maternal Health Research (TamCAM), Faculty of Medicine and Health Technology, Tampere University, during 2016-2023. First, I would like to express my deepest gratitude to all the people who advised and helped me with my work during these years. I also want to thank all the patients and their families who were willing to participate in the research. I am thankful for the financial support from the Blood Disease Research Foundation, ERAPerMed, Cancer Foundation Finland, Jane and Aatos Erkko Foundation, Sigrid Juselius Foundation, and the Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital, which enabled me to concentrate on my work and participate in world-class meetings.

I would like to thank one of the founders of the Tampere Center of Child Health Research, Emer. Prof. Markku Mäki, and the current director of TamCAM Prof. Per Ashorn for the good and interesting working environment in the field of children's health.

I would like to whole-heartedly thank my supervisors, Docent Olli Lohi, MD, PhD and Prof. Merja Heinäniemi. Olli, thank you for the opportunity to join your research group and work with exciting research projects. I am greatly thankful for the trust and opportunities you have given to me, and the knowledge you have shared with me. I am honored to be a part of the HemoRes research group. Merja, this thesis and I as the scientist would be something totally different without your contribution and help. Thank you for believing in and challenging me to learn new. I want to warmly thank you, Olli and Merja, for your support over the years. If I have sometimes lost my faith, you have always encouraged and trusted me. I admire the passion for the science what you both have, your can-do attitude, and your never-ending flow of new ideas and research questions.

I would like to thank the pre-examiners of this thesis, Prof. Eeva-Riitta Savolainen and Docent Leena Latonen, PhD, for their valuable comments that improved this thesis. I also thank the members of my thesis supervision committee, Docent Kirsi Rautajoki, PhD, and Docent Keijo Viiri, PhD, for their advice and kind comments regarding my research projects. I would also like to warmly thank all my

co-authors, especially Valtteri Wirta, PhD, Assoc. Prof. Jessica Nordlund, and Otto Kauko, MD, PhD; your help and contribution made these publications possible.

I wish to thank Prof. Markus Müschen for the opportunity to join his research lab for the research visit. The time in Müschen Lab gave me an opportunity to challenge myself, expand my knowledge, and gave me a viewpoint on the world of top-class science. I also want to thank all the members of the Müschen Lab for their support, as well as for great memories in and out of the lab.

From the bottom of my heart, I want to thank the past and present members of HemoRes group, especially Saara Mathlein, PhD, Susanna Teppo, PhD, Toni Grönroos, PhD, Artturi Mäkinen, PhD, Atte Nikkilä, PhD, Veronika Zapilko, PhD, Julia Ventelä, MD, Noora Hyvärinen, MSc, and Miikka Vouttilainen MSc. Thank you for being such a wonderful companion these past years, I am privileged to have worked with you. I would like to warmly thank Jorma Kulmala and Eini Eskola for your help in the lab. I also wish to thank all the other colleagues with whom I have had great pleasure working with. Members of the Heinäniemi group thank you for all the collaboration and memorable time on and off work. Especially, I want to thank Mari Lahnalampi, MSc, Juha Mehtonen, MSc, Petri Pölönen, PhD, and Sanni Moisio, MSc. ISE group members, the present and past, especially Laura Martin Diaz, MSc, Valeriia Dotsenko, MSc, Hana Le, PhD, Mikko Oittinen, MSc and Joel Johnsson, PhD, thank you for your great company at the office, your support, and for making the working days more fun. I would also like to thank the past and present members of the CeliRes group for their peer support.

I have been lucky to have so many friends supporting me, thank you. Anna, Pauski, Mikke, and Martti thank you for your life-long friendship. Elli, thank you for cheering me on over the years, and reminding me that there is also much else in the world than work. Birgitta and Maria, thank you for staying by my side no matter where we live, special thanks for the help with coding. Laura, Miina and Heidi, thank you for your peer support and friendship. Big thanks also to my hockey family, I am lucky to have you in my life.

Finally, I am deeply grateful to my family. Mom, Dad, Lauri and Lasse thank you for your never-ending support and encouragement. Lauri, on the times I have had a hard time believing myself, your "you can do anything"–attitude has meant the world to me, kiitos.

Tampere, October 2023

Laura Oksa

ABSTRACT

Acute lymphoblastic leukemia (ALL) is the most common form of childhood cancer. It can be classified into T and B cell lineage leukemias, with B-ALL being more common, and accounting for approximately 85% of cases. The survival rate of ALL patients has exceeded 90% due to significant improvements in treatment over the past decades. However, the prognosis of relapsed disease is much worse, and relapsed T-ALL is particularly difficult to cure.

B-ALL can be divided into several subtypes based on the presence of specific genetic alterations. One of these is the *ETV6::RUNX1* subtype, which is found in one-quarter of pediatric ALL cases. In contrast to B-ALL, T-ALL lacks a similar prognostic classification system. Therefore, studies seeking to identify clinically applicable biomarkers for T-ALL are highly relevant.

In this thesis, we used a large gene expression data set of hematological malignancies and healthy samples to screen for potential candidate genes that are biologically relevant to T-ALL pathogenesis and could serve as prognostic biomarkers. Two genes, protein arginine methyltransferase 7 (*PRMT7*) and SIX homeobox 6 (*SIX6*), were highly expressed in the T-ALL samples. *PRMT7* promoted the growth of T-ALL cells in colony formation assays, regulated the expression of various genes associated with T-ALL pathogenesis, and altered the arginine monomethylation of the *RUNX1* transcription factor, leading to the altered expression of *RUNX1* target genes. High *PRMT7* expression was associated with a trend toward inferior event-free survival and overall survival. Similarly, high levels of *SIX6* messenger RNA were associated with inferior event-free survival and overall survival. The expression of *SIX6* was directly upregulated by *TAL1*, a master regulator of transcription in T-ALL.

Patients with the *ETV6::RUNX1* fusion typically have excellent outcomes. However, a subset of patients has a slow early response to therapy, which is associated with an increased risk of relapse. We hypothesized that secondary genetic alterations would determine the treatment response in *ETV6::RUNX1*-positive leukemia. In leukemia samples from a cohort of 175 pediatric *ETV6::RUNX1* patients, we identified large genomic deletions and amplifications associated with early therapy response. The fast-responding leukemic

blasts frequently exhibited the APOBEC mutational signature, had higher cell cycle activity, and were often missing a region of chromosome 12p. In contrast, blasts from slow-responding leukemias had frequent rearrangements in the immunoglobulin kappa gene and kappa-deleting elements. Upon treatment, the blasts with productive rearrangements were enriched, suggesting the activation of BCR signaling.

In summary, we discovered two novel prognostic biomarkers for T-ALL and the response-associated genetic features of B-ALL. These findings may pave the way for more individualized therapy in pediatric ALL in the future.

TIIVISTELMÄ

Akuutti lymfoblastileukemia (ALL) on lapsuusiän yleisin syöpä. Se voidaan luokitella B- ja T-solulinjojen taudeiksi taudin alkuperän mukaan. Näistä kahdesta leukemiatyypistä B-ALL on selkeästi yleisempi, ja kaikista lasten ALL:ista noin 85 % on B-solulinjan tauteja. ALL:n ennuste on nykyään erinomainen vuosikymmenien saatossa asteittain kehitettyjen lääkehoitojen ansiosta, ja niinpä B-ALL potilaista noin 90 % paranee pysyvästi. Taudin uusiutuessa ennuste on selvästi heikompi, ja erityisesti T-ALL on harvemmin parannettavissa uusiutumisen jälkeen.

B-ALL voidaan luokitella useaan eri alatyyppiin erityisten geneettisten muutosten perusteella, ja tyypillisen tällaisen alaryhmän muodostavat potilaat, joiden leukemiasoluissa todetaan *ETV6::RUNX1*-fuusiogeeni. T-ALL:n luokittelua ei ole samalla tavalla vakiinnutettu ja niinpä tutkimukset, joissa haetaan ennusteellisia merkkiaineita, ovat hyvin ajankohtaisia.

Tässä väitöskirjassa löydettiin kaksi uutta merkkiainetta T-ALL:aan hyödyntämällä aikaisemmin kerättyä laajaa verisairauksien geeni-ilmentymisen tietopankkia. *PRMT7*- ja *SIX6*-geenit ilmentyivät vahvasti T-ALL:ssa verrattuna muihin verisyöpiin ja terveisiin T-soluihin. Solukokeissa havainnoimme, että arginiinimetyylitransferaasi PRMT7 lisäsi T-ALL-solujen pesäkekasvua, muutti T-ALL:n patogeneesiin osallistuvien geenien ilmentymistä ja häiritsi RUNX1-proteiinin arginiinijäänteiden monometylaatiota ja sen toimintaa. Korkea *PRMT7*- ja *SIX6*-geenien ilmentyminen liittyi huonompaan ennusteeseen. T-ALL:ssa hyvin tunnettu transkriptiotekijä TAL1 sääteli suoraan *SIX6*:n ilmentymistä.

Noin neljännes lapsuusiän ALL-tapauksista luokitellaan *ETV6::RUNX1*-alatyypin leukemiaksi. Tämän leukemiatyypin ennuste on yleensä erittäin hyvä, mutta tuntemattomasta syystä pieni osa potilaista reagoi huonosti alkuvaiheen lääkehoidolle, ja heillä on lisääntynyt taudin uusiutumisen riski. Tutkimushypoteesinamme oli, että hoitovasteen puutteen takana ovat geneettiset tekijät, ja tutkimme asiaa 175 potilaan näytteistä. Tunnistimme isoja geenialueiden puutoksia ja monistumia, jotka liittyivät hoitovasteeseen. Nopeasti hoitoon vastaavien leukemiasolujen geeneissä oli usein havaittavissa APOBEC-mutaatiojälki ja kromosomialue 12p puutos. Solut olivat myös usein aktiivisesti solusyklin jakautumisvaiheessa. Hitaasti hoitoon vastaavilla leukemiasoluilla

havaittiin immunoglobuliinigeenin (kappa) uudelleenjärjestymiä ja kappa-deleetio-alueen muutoksia. Tästä johtuen solujen fenotyyppi muuttui, mikä viittaa B-solureseptorin signaloinnin aktivoitumiseen.

Yhteenvetona voidaan todeta, että tässä väitöskirjatutkimuksessa tunnistettiin kaksi uutta merkkiainetta T-soluiseen ALL:aan sekä geneettisiä muutoksia, jotka selittävät hoitovastetta *ETV6::RUNX1*-tyypin leukemiassa. Tämän kaltaiset tutkimukset mahdollistavat tulevaisuudessa paremmin kohdennettujen hoitojen suunnittelun leukemiaan.

CONTENTS

List of Figures

List of Tables

ABBREVIATIONS

ABL1	ABL Proto-Oncogene 1, Non-Receptor Tyrosine Kinase
ADMA	ω-NG, NG-asymmetric dimethylarginine
AICDA, AID	Activation-induced cytidine deaminase
ALL	Acute lymphoblastic leukemia
AML	Acute myeloid leukemia
ANGPT1	Angiopoietin 1
APOBEC	Apolipoprotein B mRNA editing enzyme
ARPP21	CAMP Regulated Phosphoprotein 21
ATF7IP	Activating Transcription Factor 7 Interacting Protein
B-ALL	B cell acute lymphoblastic leukemia
BCL11B	BCL11 Transcription Factor B
BCR	B cell reseptor
bHLH	The basic-helix-loop-helix
BM	Bone marrow
BMI1	BMI1 Proto-Oncogene, Polycomb Ring Finger
CCL19	C-C Motif Chemokine Ligand 19
CCL25	C-C Motif Chemokine Ligand 25
CCR7	C-C Motif Chemokine Receptor 7
CCR9	C-C Motif Chemokine Receptor 9
CD28	CD28 Molecule
CD3ζ	CD3 Epsilon Subunit Of T-Cell Receptor Complex
CD4	CD4 Molecule
CD8	CD8 Molecule
CDKN2A	Cyclin Dependent Kinase Inhibitor 2 A
CDKN2B	Cyclin Dependent Kinase Inhibitor 2 B
CFU	Colony formation unit
CGH array	Comparative genomic hybridization
CLL	Chronic lymphocytic leukemia
CML	Chronic myeloid leukemia
CNV	Copy number variation

CLP	Common lymphoid progenitor
CMP	Common myeloid progenitor
CORUM	Mammalian protein complexes database
COSMIC	Catalogue of Somatic Mutations in Cancer
CXCL12	C-X-C motif chemokine ligand 12
DDR	DNA damage response
DC	Dentric cell
DLBCL	Diffuse large B cell lymphoma
DN	Double negative
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DP	Double positive
DUX4	Double Homeobox 4
E2A	E2A immunoglobulin enhancer-binding factors E12/E47
E3F1	EBF Transcription Factor 3
EBF1	EBF Transcription Factor 1
EFS	Event-free survival
EOC	End-of-consolidation
EOI	End-of-induction
ETP	Early thymic precursor
ETS	E26 transformation-specific family
ETV6	ETS Variant Transcription Factor 6
ETV6::RUNX1	Fusion of ETV6 and RUNX1 genes
FBXW7	F-Box And WD Repeat Domain Containing 7
FDR	False discovery rate
FISH	Fluorescence in-situ hybridization
FP	False positive
GATA1	GATA binding protein 1
GATA2	GATA binding protein 2
GATA3	GATA binding protein 3
GC	Germinal center
GFI1	Growth factor independence 1
GMP	Granulocyte/macrophage progenitor
GnRH	Gonadotropin-releasing hormone
gRNA	Guide RNA
H3K27ac	Acetylation of histone 3 lysine 27

H3K4me1	Monomethylation of histone 3 lysine 4
H3K4me3	Trimethylation of histone 3 lysine 4
H4K4me3	Trimethylation of histone 3 lysine 4
HAT	Histone acetyl transferase
HDAC	Histone deacetylase
HDACi	HDAC inhibitors
HeH	High hyperdiploid
HSC	Hematopoietic stem cell
Hi-C	Chromosome conformation capture method
HOX	Homeobox gene
HOXA	Homeobox A
HR	High risk
iAMP21	Intrachromosomal amplification of chromosome 21
IG	Immunoglobulin
IGH	Immunoglobulin heavy chain
IGL	Immunoglobulin light chain
IKZF1	IKAROS Family Zinc Finger 1
IL-7R	Interleukin-7 receptor
im	Intramuscular
InDel	Small indel or deletion
IR	Intermediate risk
it	Intrathecal
ITAM	Immunoreceptor tyrosine-based activation motifs
iv	Intravenous
JAK	Janus Kinase
KMT2A	Lysine Methyltransferase 2A
KRAS	Kirsten rar sarcoma proto-oncogene
LAT	Linker for activation of T cells
LCK	LCK proto-oncogene
LEF1	Lymphoid Enhancer Binding Factor 1
Lin28B	Lin-28 Homolog B
LMO1	LIM domain only 1
LMO2	LIM domain only 2
logFC	Log fold chance
LOH	Loss-of-heterozygosity
LT-HSC	Long-term HSC

MAPK	Mitogen-activated protein kinases
Mb	Megabase
MEP	Megakaryocyte/erythroid progenitor
MEP50	Methylosome protein 50
MHC	Major histocompatibility complex
MLL	Mixed lineage-leukemia
MMA	ω-NG-monomethylarginine
MMP9	Matrix metallopeptidase 9
MNC	Mononuclear cell
MRD	Minimal residual disease
MS	Mass spectrometry
MPP	Multipotent progenitor cell
MYC	MYC proto-oncogene
Ly-Bi	Lymphoid-biased
My-Bi	Myeloid-biased
NF-ϰB	Nuclear factor-ϰB
NGS	Next-generation sequencing
NK cells	Natural killer cell
NKP	NK cell precursor
NKX2-1	NK2 homeobox 1
NOPHO	Nordic society of peadiatric hematology and oncology
NOTCH1	Notch receptor 1
NR3C1	Nuclear Receptor Subfamily 3 Group C Member 1
OS	Overall survival
PAX5	Paired Box 5
PEG-ASP	Pegylated Asparaginase
Ph	Philadelphia chromosome
PHF6	Plant homeodomain finger 6
PLCy1	Phospholipase Cγ1
PRC2	Polycomp repressive complex 2
pre-B	Precursor B
pre-BCR	pre B cell receptor
pre-Tα	Surrogate α receptor
PRMT	Protein Arginine Methyltransferases
PRMT7	Protein Arginine Methyltransferase 7
pro-B	Progenitor B

PU.1	Spi-1 Proto-Oncogene
RAG	Recombination-activating gene
RAS	Rat sarcoma
RB1	RB transcriptional corepressor 1
RNA	Ribonucleic acid
RPKM	Normalized per million mapped
RPL10	Ribosomal protein L10
RPL11	Ribosomal protein L11
RPL22	Ribosomal protein L22
RPL5	Ribosomal protein L5
RT-PCR	Real time quantitative polymerase chain reaction
RUNX1	RUNX Family Transcription Factor 1
SAM/AdoMet	S-adenosyl-L-methionine
sc	Single cell
scRNA-seq	Single-cell RNA sequencing
SDMA	ω-NG, N'G-symmetric dimethylarginine
SELE	Endothelial-cell (E)-selectin
SET	SET Nuclear Proto-Oncogene
ST-HSC	Short-term HSC
SIX	The Sine Oculis homeobox
SIX6	SIX Homeobox 6
SLC	Surrogate light chains
SNP	Single nucleotide polymorphisms
SNV	Singel nucleotide variant
SP	Single positive
SR	Standard risk
STAT	Signal transducer and activator of transcription
SV	Structural variant
T-ALL	T cell acute lymphoblastic leukemia
TAL1	TAL BHLH Transcription Factor 1
TBL1XR1	Transducin Beta Like 1 X-Linked Receptor 1
TCL	T cell lymphoma
TCR	T cell receptor
TF	Transcription factor
Th cell	T-helper cell
TIGAR	TP53 Induced Glycolysis Regulatory Phosphatase

TLX1	T Cell Leukemia Homeobox 1
TLX3	T Cell Leukemia Homeobox 3
TP	True positive
TP53	Tumor Protein 53
Treg cell	Regulatory T cell
TSP	Thymic seeding progenitors
USP7	Ubiquitin-specific-processing protease 7
VCAM1	Vascular cell adhesion molecule 1
VrepB	V-Set Pre-B Cell Surrogate Light Chain 1
WES	Whole exome sequencing
WGS	Whole genome sequencing
WHO	World Health Organization
WT1	WT1 Transcription Factor
ZAP70	Zeta-chain-associated protein kinase 70

ORIGINAL PUBLICATIONS

Publication I Laukkanen, S., **Oksa, L.,** Nikkilä, A., Lahnalampi, M., Parikka, M., Seki, M., Takita, J., Degerman, S., de Bock, C. E., Heinäniemi, M., & Lohi, O. (2020). SIX6 is a TAL1-regulated transcription factor in T-ALL and associated with inferior outcome. *Leukemia & lymphoma*, *61*(13), 3089–3100. https://doi.org/10.1080/10428194.2020.1804560

Publication II **Oksa, L.**, Mäkinen, A., Nikkilä, A., Hyvärinen, N., Laukkanen, S., Rokka, A., Haapaniemi, P., Seki, M., Takita, J., Kauko, O., Heinäniemi, M., & Lohi, O. (2022). Arginine Methyltransferase PRMT7 Deregulates Expression of RUNX1 Target Genes in T cell Acute Lymphoblastic Leukemia. *Cancers*, *14*(9), 2169. https://doi.org/10.3390/cancers14092169.

Publication III **Oksa, L.**, Moisio S., Maqbool, K., Kramer, R. Nikkilä, A., Foroughi-Asl H., Suhonen, Krali, O., Voutilainen, M., Lahnalampi, M., J., Duque-Afonso, J., Hauer J., Vepsäläinen, K., Nordlund, J., Wirta, V., Lohi O.* & Heinäniemi M.* Genomic determinants of therapy response in *ETV6::RUNX1* leukemia. Submitted.

*co-senior authors

AUTHOR'S CONTRIBUTION

Publication I: The author designed, and performed laboratory works and data analyses, such as generation of CRISPR-Cas9 knockout cell lines and their phenotypic characterization, *in vitro* experiments and RNA-seq analysis. The author actively participated in preparation and writing of the manuscript.

Publication II: The author participated in the study design with the supervisors. The author conducted a literature search and collected the data for the study. The author designed and performed most of the laboratory work and data analysis. The author was responsible for writing the first draft of the manuscript and finalizing the manuscript for submission.

Manuscript III: The author participated in the study design with the supervisors. The author conducted a literature search and coordinated the clinical data collection for the study. The author participated actively in sequencing data analysis (WGS, RNAseq), and performed the laboratory works. The author was responsible for writing the first draft of the manuscript and finalizing the manuscript for submission.

1 INTRODUCTION

Acute leukemia is a hematological malignancy characterized by an increased number of immature blast cells in the bone marrow and blood. Leukemia is the most common childhood cancer, and approximately 4,000 new cases are diagnosed annually in Europe (Suomen Syöpärekisterin tilastot.; Steliarova-Foucher et al., 2017). The incidence peak of acute lymphoblastic leukemia (ALL) is between 2 and 5 years. During the past few decades, the prognosis of ALL has improved remarkably, with the current cure rate exceeding 90 % (Norén-Nyström et al., 2023; Quist-Paulsen et al., 2020; Toft et al., 2018). Success was built on stepwise improvement of the therapy. The major steps were the introduction of combination chemotherapies in the 1960s and, more recently, improved risk grouping based on early therapy response as measured by minimal residual disease (MRD).

ALL is divided into B- and T-lineage diseases according to the cell of origin. Around 15%–20% of childhood ALL cases represent T-ALL, while the rest are B-ALL. B-ALL cases can be further divided into more than 20 subtypes based on the genetic alterations that leukemia cells harbor (Inaba and Pui, 2021). Subtyping of B-ALL informs clinicians about treatment decisions, as the outcome is closely related to the subtype. However, this is not always the case, as even patients belonging to low-risk genetic subtypes may encounter disease recurrence. The reasons for treatment failure are unknown, and no consistent predictive genetic biomarkers have been identified. In the case of T-ALL, no similar subtyping exists, and there is a quest for novel biomarkers and improved prognosticators for this disease.

This thesis aimed to discover abnormally regulated genes in T-ALL, study their function as possible novel biomarkers, and identify genetic alterations that could predict treatment response in *ETV6::RUNX1*-positive leukemia.

2 REVIEW OF THE LITERATURE

2.1 Hematopoiesis

Hematopoiesis refers to a process in which blood cells are formed through a differentiation cascade from a multipotent self-renewing stem cell into highly specialized cells responsible for tissue oxygenation, clot formation, and immunological defense of the body. Hematopoiesis occurs from early embryonic development to adulthood, and its main function is to constantly replenish various blood and immune cells over the lifetime. Many blood cell diseases, including leukemia, arise from disruptions in this process. Therefore, understanding the exact molecular mechanisms underlying hematopoiesis is crucial.

In vertebrates, hematopoiesis occurs in different anatomic locations in two waves called primitive and definitive waves (Galloway and Zon, 2003). The primitive wave begins in the early embryonic stage at the third week of gestation (Palis and Yoder, 2001), occurs in the yolk sac, and produces erythroblasts to ensure tissue oxygenation during the rapid growth of the embryo (Galloway and Zon, 2003; Orkin and Zon, 2008; Palis and Yoder, 2001). Macrophages and megakaryocytes are also produced during primitive waves in the yolk sac. Primitive wave is transient and does not produce pluripotent cells with self-renewal capacity, and it is rapidly replaced by adult-type hematopoiesis, which is called a definitive wave of hematopoiesis (Jagannathan-Bogdan and Zon, 2013; Orkin and Zon, 2008). A definite wave occurs later in embryonic development. The definitive wave starts in the yolk sac, when the first pluripotent cells, hematopoietic stem cells (HSC), are generated from the hemogenic endothelium in the aorta-gonad-mesonephros (AGM) region of the human embryo between days 27 and 40 of human embryonic development (Galloway and Zon, 2003; Julien et al., 2016; Palis and Yoder, 2001). AGM plays a crucial role during the definitive wave and has an autonomous capacity to generate hematopoietic progenitors (Julien et al., 2016). Later in a definitive wave, approximately during weeks 4–5 after gestation, HSCs start moving into the fetal liver, where they begin to differentiate into various hematopoietic cell lineages, and by weeks 8–9, they also circulate into the thymus (Baron et al., 2012; Galloway and Zon, 2003; Haddad et al., 2006; Haynes and Heinly, 1995; Tavian and Péault, 2005).

After week 10.5 of gestation, HSCs colonize the spleen and bone marrow (BM) (Baron et al., 2012; Galloway and Zon, 2003; Rieger and Schroeder, 2012; Tavian and Péault, 2005). Thereafter, the thymus and BM are humans' primary sites of hematopoiesis (Galloway and Zon, 2003; Julien et al., 2016; Rieger and Schroeder, 2012). HSCs were the first tissue-specific stem cells identified by Till and McCulloch in the 1960 century (Till and McCulloch, 1961).

In 1957, Conrad Waddington presented a theory according to which mammalian development is unidirectional, where embryonic stem cells develop into more mature differentiated states. This is now known as Waddington's epigenetic model (Bhattacharya et al., 2011; Brown and Ceredig, 2019; Waddington, 1957). Waddington suggested that genes determine which path cells follow during development. This process is often illustrated as a ball rolling down from the top of a mountain along branching valleys and ridges; the further down, the more specialized or stable the cell state is (Bhattacharya et al., 2011; Brown and Ceredig, 2019; Waddington, 1957). This model has been widely incorporated into models of hematopoiesis.

The study of hematopoiesis started in the 1960s, when it was shown that bone marrow cells were capable to form spleen colonies in irradiated mice, and later, bone marrow cells were shown to form heterogeneous colonies in semi-solid media (Bradley and Metcalf, 1966; Till and McCulloch, 1961). In the classical model of hematopoiesis, HSCs lay on top of the hematopoietic hierarchy, from which different progenitor cells with increasingly restricted lineage potential give rise to all cell lineages (Figure 1.) (Rieger and Schroeder, 2012). HSCs are the only hematopoietic cell type with self-renewal and differentiation potential (Rieger and Schroeder, 2012). In this model, long-term HSCs (LT-HSCs) differentiate into short-term HSCs (ST-HSCs), which exhibit decreased self-renewal capacity. ST-HSCs give rise to multipotent progenitor cells (MPP), which branch into common myeloid progenitors (CMPs) and common lymphoid progenitors (CLPs). CMPs can differentiate into myeloid, erythroid, and megakaryocytic lineages, whereas CLPs can produce T cells, B cells, natural killer cells (NK), and dendritic cells. During myelopoiesis, CMPs give rise to granulocyte/macrophage progenitor cells, which further differentiate into granulocytes, monocytes, macrophages, and dendritic cells (DCs) (Orkin and Zon, 2008). These cells have an essential role in the regulation of innate and adaptive immune responses, having roles in inflammation, infection, cellular immunity, and tissue homeostasis (Amodio et al., 2019; De Kleer et al., 2014). The two major functions of myeloid cells are phagocytosis of foreign material or pathogens and the secretion of inflammatory cytokines and chemokines to enhance

the inflammatory response. CMPs also give rise to megakaryocyte-erythrocyte progenitors that differentiate into megakaryocytes and erythrocytes (Orkin and Zon, 2008; Patel et al., 2005). Megakaryocytes are primarily located in the BM, and their main function is to produce platelets by tailoring their own cytoplasm and membrane synthesis in a process called thrombopoiesis (Eldor et al., 1989; Patel et al., 2005). The primary function of platelets is to regulate hemostasis, and it has been shown that platelets also play a role in the regulation of immunity (Holinstat, 2017; Patel et al., 2005). Erythrocytes are responsible for carrying oxygen from the lungs to the tissues and carbon dioxide from the tissues back to the lungs and for maintaining systemic acid/base equilibria (Kuhn et al., 2017). All steps in this tree-like hierarchy model of hematopoiesis are tightly regulated by transcription factors and cytokines (Cheng et al., 2020), as discussed later in this thesis.

Although the classical model of hematopoiesis has offered a basic understanding of the differentiation process, it was established by using immunophenotyping of bulk cell populations and functional assays of bulk cells. New single-cell omics, such as single-cell transcriptomics, proteomics, and multi-omics methods, have improved the resolution of studies and revealed that differentiation is more complex (Brown and Ceredig, 2019; Cheng et al., 2020; Zhang et al., 2022). Accordingly, new subpopulations of HSCs have been identified, and it seems that HSCs and MPPs are more heterogeneous than previously thought, and cell fate commitment already occurs at the HSC level (Belluschi et al., 2018; Brown and Ceredig, 2019; Karamitros et al., 2018; Zhang et al., 2022). One of the revised models shows that HSCs harbor lineage specificity at an early stage, where HSCs have been defined as myeloid-biased (My-Bi), balanced (Ba), and lymphoid-biased (Ly-Bi) HSCs (Müller-Sieburg et al., 2004, 2002). Research groups led by Trumpp and Passegue have further revised the model by dividing the MPP population into four groups, MPP1, MPP2, MPP3, and MPP4, according to their immunophenotype, cell cycle status, lineage bias, drug resistance, and bone marrow abundance (Pietras et al., 2015; Wilson et al., 2008). In this model, MPP1 cells give rise to all lineages, whereas MPP2 cells are myeloid- and platelet-biased, MPP3 cells are myeloid-biased, and MPP4 cells are lymphoid-biased.

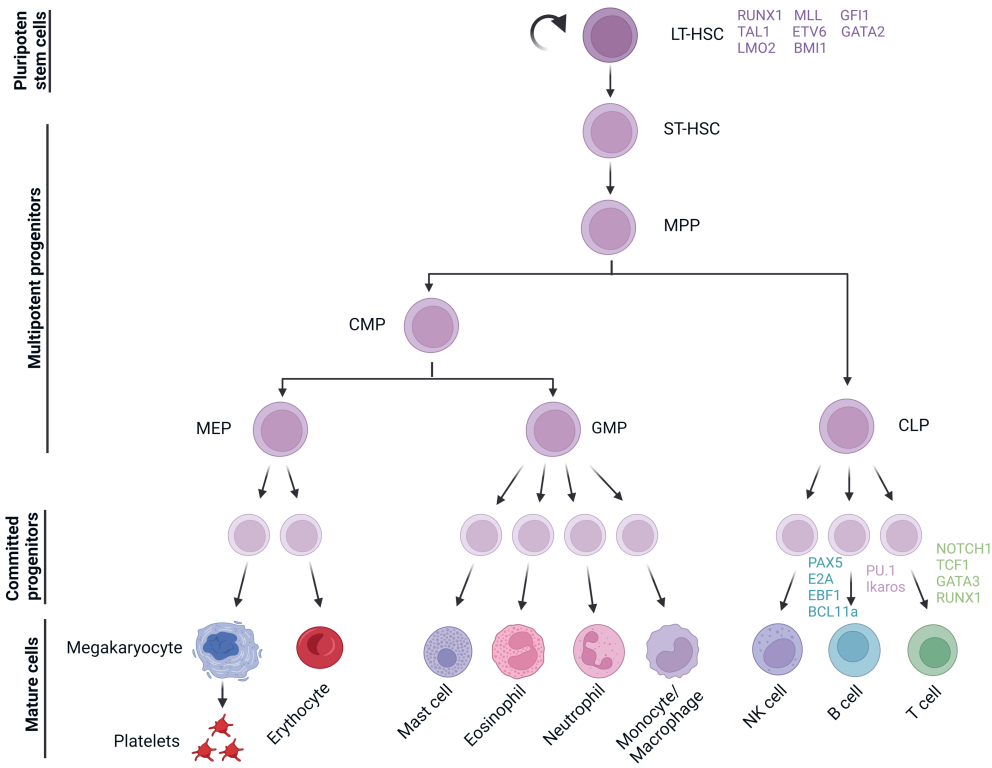


Figure 1. The classical model of hematopoiesis. The most central transcription factors related to the HSC self-renewal and differentiation, and lymphocyte differentiation are indicated. LT-HSC = Long-term hematopoietic stem cells, ST-HSC = Short-term hematopoietic stem cells, MPP = Multipotent progenitor, CMP = Common myeloid progenitor, CLP = Common lymphoid progenitor, MEP = Megakaryocyte/erythroid progenitor, GMP = Granulocyte/macrophage progenitor. Adapted from Orkin and Zon 2008, and Rieger and Schroeder 2012. Created with BioRender.com

2.1.1 Lymphoid cells and their differentiation

CLPs give rise to mature lymphoid cells, T lymphocytes (thymus-dependent lymphocytes), B lymphocytes (bursa/bone marrow –dependent lymphocytes), and NK cells (Cooper, 2015; Orkin and Zon, 2008). In the healthy stage, these cells play an essential role in antibody production, killing infected and tumor cells by direct cell-mediated killing, and regulation of the immune response. T and B cells are effectors of adaptive immunity, whereas NK cells function as cytotoxic lymphocytes

in the innate immune system. NK cells lack the expression of recombinant antigen receptors, and for this reason, their maturation process differs from that of other lymphoid cells (LaRosa and Orange, 2008). NK cell development occurs mostly in the bone marrow, but also in other tissues such as the thymus, lymph nodes, spleen, and liver from committed NK cell precursors (NKPs). During the differentiation process, NKPs form immature NK cells when they begin to express activating and inhibitory receptors for specific MHC class I molecules and engage self-MHC molecules, this process is called "NK cell education" (Huntington et al., 2007). Educated NK cells acquire a mature phenotype, which is tolerant to self and is capable of recognizing MHC class I expression levels in target cells (Huntington et al., 2007).

T cell development

T cells are the key effectors of cell-mediated immunity, defending against infections and mediating immune surveillance against tumor cells and foreign tissues, by the activation of T cells in response to antigens from pathogens, tumors and the environment (Koch and Radtke, 2011; Kumar et al., 2018). Mature T cells develop in the thymus from progenitor cells through the multistage and coordinated developmental process (Figure 2.) (LaRosa and Orange, 2008; Rothenberg et al., 2008). The thymus is a small gland in the lymphatic system between the lungs and behind the sternum, which is specialized for the development and production of T cells. It can be subdivided into four major compartments: the cortex, subcapsular zone, medulla, and corticomedullary junction, all of which have specific microenvironments and host different stages of T cell development (Koch and Radtke, 2011; Rothenberg et al., 2008). Thymic T cell development starts when progenitor cells, called as thymic seeding progenitors (TSP), enter the thymus, and face the thymic epithelium. TSPs can be of several different types, such as MMPs, CLPs, or committed T cell progenitors, but the major progenitor type in mice studies is lymphoid-primed MMPs (Rothenberg et al., 2008; Zlotoff and Bhandoola, 2011). TSPs settle in the thymus guided by CCL19 and CCL25 chemokines secreted by thymic endothelial cells, which bind to CCR9 and CCR7 receptors on the surface of progenitor cells (Rothenberg et al., 2008; Zlotoff and Bhandoola, 2011).

TSPs first differentiate into early thymic precursor (ETP) cells and CD4-/CD8-double-negative (DN) DN1 cell subsets at the corticomedullary junction. These cells retain the capacity to differentiate into the DCs, NK cells, and macrophages. However, at the T cell DN1 checkpoint, Notch1 signaling inhibits the alternative

cell development potential and cells commit to T cell fate. Once DN1 cells migrate into the cortex, they differentiate into the DN2a and further DN2b stages. Here, recombination-activating gene (RAG) 1- and RAG2-mediated gene rearrangement at the *TCRγ*, *TCRδ*, and *TCRβ* gene loci is initiated (Ciofani and Zúñiga-Pflücker, 2007; Rothenberg et al., 2008). When cells reach the DN3 stage, they stop proliferating and complete TCR rearrangement, resulting in the first fully rearranged TCR loci. At the DN3 stage, the second checkpoint called β-selection occurs, where cells with successfully rearranged β-chains pair with the surrogate α receptor (pre-Tα), forming a functional pre-TCR complex, and continue differentiation into αβ T cells through DN3a and DN3b stages (Ciofani and Zúñiga-Pflücker, 2007; Koch and Radtke, 2011). Cells with successfully rearranged γ- or δ-chains develop into γδ-T cells via γδ-selection and remain as DN cells, migrating from the thymus into lymphoid tissues and epithelia (Ciofani and Zúñiga-Pflücker, 2007; LaRosa and Orange, 2008; Rothenberg et al., 2008). Finally, at the DN4 stage, cells move from the subcapsular zone to the medulla, and after active pre-TCR signaling, they begin expressing CD4 and CD8 molecules (double-positive DP cells) (Germain, 2002; Rothenberg et al., 2008). Here, TCRα recombination is initiated, resulting in DP TCRαβ T cells, which undergo MHC-mediated selection to yield mature CD4 or CD8 single positive (SP) T cells (Koch and Radtke, 2011; Spits, 2002). In the selection process, DP cells that bind to MHC with sufficient affinity are positively selected to survive: DP cells with TCRs that bind to the MHC-class-I-complexes become CD8+ T cells, whereas cells that bind to the MHC-class-II become CD4+ T cells (Ciofani and Zúñiga-Pflücker, 2007; Spits, 2002; Starr et al., 2003).

In the medulla, SP cells undergo negative selection, and T cells with high-affinity TCRs for self-antigens are eliminated through apoptosis. CD4+ cells specialize in T-helper (Th) cells and regulatory T-(Treg), whereas CD8+ cells become cytotoxic T cells. After the differentiation process is complete, the cells leave the thymus and move into the lymphoid tissues throughout the body (Germain, 2002; LaRosa and Orange, 2008).

TCR signaling is essential for the normal development of T cells. T cell receptor is formed through multiple rearrangements of *TCRa*, *TCRβ*, *TCRγ*, and *TCRδ* genes, resulting in a TCR heterodimer complex, with unique peptide-MHC antigen specificity (Courtney et al., 2018; Mariuzza et al., 2020; Shah et al., 2021). In brief, TCR signaling is initiated when TCR binds to an antigen, and the signal is forwarded to the cytosol by invariant C3 dimers, which are non-covalently associated with TCR. Next, LCK is recruited to the TCR complex by the colocalization of CD4 or CD8 coreceptors, where LCK phosphorylates ITAM signaling motifs on CD3ζ chains.

Phosphorylated ITAMs create binding sites for ZAP70 kinase. ZAP70 is activated by LCK phosphorylation, and activated ZAP70 phosphorylates adaptor protein LAT (Courtney et al., 2018; Mariuzza et al., 2020; Shah et al., 2021). LAT recruits multiple additional signaling effectors that become activated; for example, activation of LAT recruits PLCγ1 to provide calcium and activate RAS-MAPK signaling (Courtney et al., 2018). LAT contains four major phosphorylation sites and different sites recruit different effectors. LCK can also interact with CD28, leading to binding of the regulatory subunit of PI3K and activation of PI3K-AKT signaling (Courtney et al., 2018). The main function of TCR signaling activation is the initiation of T cell response, including proliferation, migration, survival, and cytokine production (Shah et al., 2021).

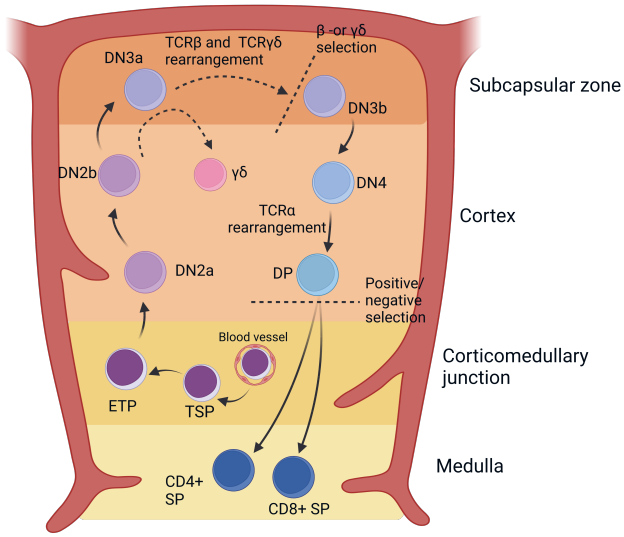


Figure 2. Schematic presentation of early T cell development in thymus. Abbreviations: TSP, thymus progenitor; ETP, early T cell precursor; DN, double negative; DP, double positive; SP, single positive; TCR, T cell receptor. Adapted from Rothenberg et al. 2008, and created using BioRender.com.

B cell development

B cells are responsible for humoral immunity by producing antibodies against various pathogens and foreign antigens. Moreover, they produce cytokines, present antigens, and regulate innate immunity (LaRosa and Orange, 2008). B cells develop

from HSCs in the bone marrow through a multistage process into naïve B cells and achieve maturity in secondary lymphoid organs such as lymph nodes and spleen (Figure 3.) (LeBien and Tedder, 2008). At the beginning of B cell development, CLPs can be subdivided into CD19- pre-pro-B cells and CD19+ pro-B cells (Fischer et al., 2020; Hardy and Hayakawa, 2001). The proliferation of these cells is highly dependent on interleukin-7 receptor (IL-7R) signaling, as shown in mouse models, where a defect in IL7-R signaling is arrested at the pro-B cell stage (Buchner et al., 2015; Clark et al., 2013). RAG-mediated immunoglobulin (IG) gene arrangement is first initiated at the CLP or pre-pro-B cell stage in the immunoglobulin heavy chain loci, IGH, with D_H to J_H rearrangements (Herzog et al., 2009; Igarashi et al., 2002). Later, during the pro-B cell stage, V_H to DJ_H rearrangements occur, and if V(D)J recombination is productive, pro-B cells progress to the pre-B cell stage (Herzog et al., 2009). This initiates the expression of the Igμ chain in pre-B cells, which pairs with surrogate light chains (SLC), VrepB, and λ5, resulting the pre-B cell receptor (pre-BCR) complex (Clark et al., 2013; Herzog et al., 2009; LeBien and Tedder, 2008; Müschen, 2018). The pre-BCR complex associates with the signaling subunits Igα and Igβ and is capable of initiating autonomous signaling that functions as an important checkpoint in B cell development (Herzog et al., 2009). If pre-B signaling is functional, the cells survive and maturation to large pre-B cell continues. However, if V(D)J recombination is non-productive, pro-B cells can either rearrange the second allele of the IgH locus or undergo cell death. Large pre-B cells expressing an autoreactive Igμ, which is recognized by SLC, will undergo apoptosis through negative selection. Positively selected cells continue maturation towards small pre-B cells, which triggers immunoglobulin light chain (IgL) gene rearrangement. Productive IgL recombination leads to the expression of a mature BCR complex, which is formed by pairing two Igκ (the κ-chain of IgL) or Igλ (the λ-chain of IgL) chains with existing Igμ chains on immature B cells (Clark et al., 2013; Herzog et al., 2009; Müschen, 2018). Active BCR signaling functions as a survival signal for immature B cells and the cells exit the BM. If BCR signaling is autoreactive, cells remain in the BM and undergo multiple rounds of receptor editing, resuming back to the small pre-B cell stage. These cells have a lifespan of approximately 3 days, and unless they can form a tolerated BCR, they undergo negative selection (Collins and Watson, 2018; Herzog et al., 2009; Müschen, 2018). Immature B cells express IgM, CD21, and CD22 and differentiate to mature B cells, mostly residing in lymphoid follicles of the spleen and lymph nodes, where they further differentiate into plasma cells or enter to germinal center (GC) where they undergo somatic hypermutation

and class-switch of immunoglobulin genes (IgD) and develop into memory B cells (Fischer et al., 2020; LeBien and Tedder, 2008).

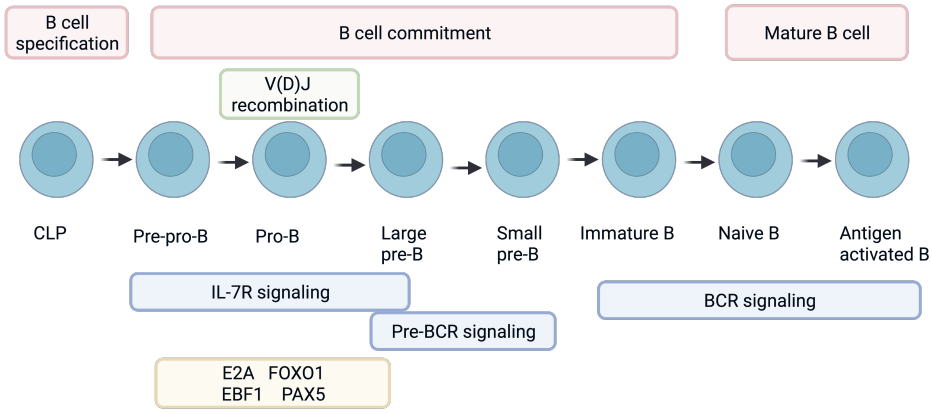


Figure 3. Schematic presentation of B cell development with key differentiation stages and core transcription factors (yellow). Abbreviations: BCR, B cell receptor; CLP = Common lymphoid progenitor; pro-B, progenitor B cell; pre-B, precursor B cell. Adapted from Fischer et al. 2020 and Buncher et al. 2015. Created with BioRender.com

2.1.2 Bone marrow microenvironment

The BM is the primary site of postnatal and adult hematopoiesis, and only a minority of hematopoietic activity and hematopoietic cell maturation occur in other tissues, such as the liver, spleen, and thymus, which is the main maturation site for T cells. The bone marrow is an architecturally complex tissue, consisting of hematopoietic, mesenchymal, endothelial, and neuronal cells, which are connected by a vascular and innervated network within the cavities of long and axial bones (Baccin et al., 2019; Ho et al., 2015; Pinho and Frenette, 2019). The maintenance and regulation of HSCs are highly dependent on their microenvironment (Wilson and Trumpp, 2006). In 1978, Schofield et al. proposed that the contact of HSCs with the bone was critical for maintaining their self-renewal capacity and inhibiting their maturation (Schofield, 1978; Wilson and Trumpp, 2006). Later studies have shown that the endosteal bone marrow microenvironment is mainly formed by osteoblasts and osteoclasts. Another niche for HSCs, called as the perivascular niche, is mostly composed of endothelial cells (Ho et al., 2015; Wilson and Trumpp, 2006).

Quiescent and slowly dividing LT-HSCs are localized in the endosteal niche. Many studies have shown that osteoblasts play a major role in the maintenance of self-renewal and regulation of LT-HSCs, and an increased number of osteoblasts is directly associated with an increased number of HSCs (Ho et al., 2015; Mikkola and Orkin, 2006; Wilson and Trumpp, 2006). Osteoblasts secrete many signaling molecules that promote LT-HSC adherence, quiescence, and survival, such as Angiopoietin 1 (ANGPT1) and C-X-C motif chemokine ligand 12 (CXCL12) (Arai et al., 2004; Ho et al., 2015; Katayama et al., 2006; Mikkola and Orkin, 2006). Osteoblasts also support the differentiation and survival of HSCs via the NOTCH and WNT signaling, thus expanding the HSC pool in the niche (Ho et al., 2015; Reya et al., 2003, 2000). Osteoclasts, which resorb bone, regulate HSCs homeostasis in the endosteal niche by inducing the mobilization of hematopoietic precursors and LT-HSCs from the bone marrow to the bloodstream (Ho et al., 2015). When osteoclasts are active and resorb the bone, they secrete proteolytic enzymes, such as matrix metallopeptidase 9 (MMP9) and cathepsin K and G, which cleave the CXCL12 and Kit ligands that are responsible for anchoring HSCs to the niche, thereby releasing LT-HSCs (Ho et al., 2015; Kollet et al., 2006). In the endosteal niche, osteoblasts and osteoclasts are key regulators that maintain the number of HSCs.

The perivascular niche, which consists largely of arterial and sinusoidal cells, is a site for proliferative and fast-cycling ST-HSCs (Baccin et al., 2019; Calvi et al., 2003; Wilson and Trumpp, 2006). Compared to arterial cells, sinusoidal cells have the functional capacity to maintain HSCs by expressing important molecules for HSCs mobilization, homing, and engraftment, such as CXCL12, endothelial-cell (E)-selectin (SELE) and vascular cell adhesion molecule 1 (VCAM1). Recently, a single-cell study by Baccin et al. (2019) revealed two novel subsets of mesenchymal CXCL12-abundant (CAR) cells, Adipo-CAR and Osteo-CAR cells, which produced the highest quantities of cytokines mediating lymphoid and myeloid differentiation (Baccin et al., 2019). There is evidence that the perivascular niche is not able to maintain the LT-HSCs, which indicates that it acts as a "secondary niche", and its primary purpose is to serve as a niche for MPPs to differentiate into B cells, megakaryocytes, erythrocytes, and myeloid cells or to migrate to the thymus, where cells differentiate into T cells (Wilson and Trumpp, 2006).

2.1.3 Transcriptional regulation in hematopoiesis

Hematopoiesis is tightly regulated by multiple transcription factors (TFs). Accordingly, alterations in these TFs are often observed in hematological diseases and malignancies. Transcription factors are proteins that bind to DNA sequences in gene regulatory regions and control gene transcription by either activating or repressing them (Latchman, 1993). Primitive hematopoiesis is mostly regulated by GATA1 and Pu.1. transcription factors, where GATA1 is the regulator of primitive erythropoiesis and Pu.1 of myelopoiesis. These two transcription factors have a cross-inhibitory relationship in the regulation of primitive erythroid and myeloid cell fates. Knockout of both TFs is lethal to embryonic development in mice (Ferreira et al., 2005; Fisher and Scott, 1998; Iwasaki et al., 2005; Jagannathan-Bogdan and Zon, 2013; Scott et al., 1994; Suwabe et al., 1998). In definitive hematopoiesis, multiple TFs, such as TAL1, LMO2, RUNX1, MLL, ETV6, GATA2, BMI1, and GFI1, are critically important for the regulation of early HSC development, self-renewal, and later differentiation into different lineages (Figure 1.) (Orkin and Zon, 2008).

The basic-helix-loop-helix (bHLH) factor TAL1 (also known as SLC) is essential for HSCs erythroid and megakaryocytic lineage development in both primitive and definite phases of hematopoiesis (Robb et al., 1995). The absence of *TAL1* leads to lethality in mice due to the complete absence of primitive hematopoiesis (Robb et al., 1995). *TAL1* homozygous null mice cannot develop any lineage blood cells and HSCs are unable to form colonies (Orkin and Zon, 2008; Robb et al., 1995; Robertson et al., 2000). Interestingly, in murine models, multiple cis-regulatory elements of *TAL1* have been recognized, and three of them have been identified as hematopoietic enhancers, each driving a different differentiation stage. The -4 *TAL1* enhancer drives the expression of endothelium and fetal blood progenitors, while the +19 *TAL1* enhancer activates *TAL1* expression in HSCs, hematopoietic progenitors, and endothelial cells, and the +40 *TAL1* enhancer displays activity in erythroblasts.

Lim-only protein 2 (LMO2) is essential for primitive and definitive hematopoiesis, and is widely expressed during hematopoiesis in all states, except for mature T cells. LMO2 does not directly bind to DNA. Instead, it forms a DNA-binding complex with other TFs such as LDB1, TAL1, E2A, and GATA1/2. LMO2 is required for primitive erythropoiesis, and its absence leads to an identical phenotype as the absence of *TAL1*. Aberrant expression of *LMO2* is known to result T cell leukemias (Wilson et al., 2010).

RUNX1 (also known as AML1) is a member of the runt transcription factor family. It participates in the generation of HSCs in the AGM, and later in erythroid, myeloid, and lymphoid cell differentiation (Jagannathan-Bogdan and Zon, 2013; Wang et al., 1996). To regulate multiple hematopoietic lineage-specific genes, RUNX1 dimerizes with the common β-subunit (CBFβ) to bind to its target sequence (Figure 4.)(Ichikawa et al., 2004b). In addition to hematopoietic regulation, RUNX1 regulates variety of other biological functions shown in Figure 4. In *RUNX1* knockout mice, primitive yolk sac hematopoiesis was unaffected, but definitive hematopoiesis in the fetal liver was absent, leading to the death of the embryos around embryonic day 12.5 (Okuda et al., 1996; Zhu and Emerson, 2002). *RUNX1* knockout led to impaired maturation of megakaryocytes, thus reducing the number of circulating platelets. In addition, *RUNX1* is expressed in developing thymocytes and B cells and regulates the expression of T and B cell-specific genes, and for that reason, *RUNX1* knockout impairs the production of mature lymphocytes (Ichikawa et al., 2004b). However, the absence of *RUNX1* does not affect the long-term maintenance of HSCs (Ichikawa et al., 2004a). Genetic aberrations in the *RUNX1* gene are common in leukemogenesis, thus indicating its role as a critical regulator of definitive hematopoiesis (Golub et al., 1995; Miyoshi et al., 1991; Sood et al., 2017).

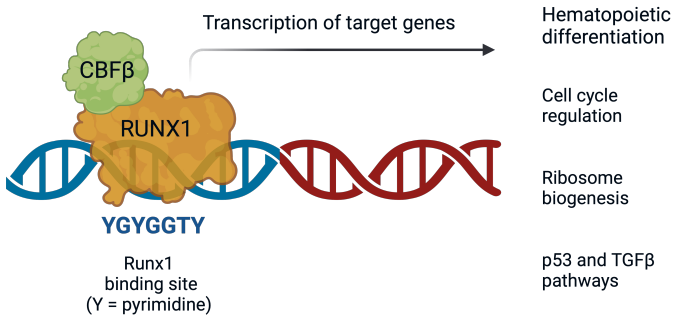


Figure 4. Schematic presentation of RUNX1 function. RUNX1 heterodimerizes with its binding partner, CBFβ. The complex binds to specific binding sites at promoters of the target genes, regulating their transcription. Adapted from Sood et al. 2017. Created with BioRender.com.

Like RUNX1, the SET-domain containing histone methyltransferase transcription factor KMT2A (also known as MLL), encoded by the Mixed-Linage Leukemia gene, is essential for the generation of HSCs in the AGM during the

embryonic development and differentiation of HSCs to multipotent progenitors (Ernst et al., 2004; Orkin and Zon, 2008). During embryonic development, KMT2A maintains the expression of *HOX* genes that are essential for normal embryonic development. *KMT2A* is expressed in both fetal and adult hematopoietic cells throughout development (Ernst et al., 2004; Kawagoe et al., 1999). Knockout studies have shown that *KMT2A* deficient cells fail to contribute to fetal liver HSC and progenitor populations by failing to form lymphocytes. In *in vitro* studies, the absence of *KMT2A* led to a deficiency in B cell differentiation (Ernst et al., 2004). Translocations of *KMT2A* with genes on chromosome (chr) segment 11q23 are found in acute lymphoblastic leukemias, especially among infants or in chemotherapy-induced leukemias (Armstrong et al., 2001).

The ETS family transcription factor ETV6 (also known as TEL) is a transcriptional repressor that is essential for HSC maintenance (Hock and Shimamura, 2017; Wang et al., 1998). ETV6 participates in the transition of hematopoietic activity from the fetal liver to the bone marrow. In the BM, *ETV6* expression is essential for the differentiation of all hematopoietic lineages. The absence of *ETV6* is lethal to mouse embryos due to the defects in yolk sac angiogenesis (Wang et al., 1998). In leukemia, translocations between *ETV6* and other hematopoietic TF genes are common; for example, the most common fusion found in childhood ALL is between *ETV6* and *RUNX1* genes (Golub et al., 1995).

The growth factor independence 1 (*GFI1*) gene is expressed in HSCs, granulocytes, monocytes, and activated macrophages. The absence of *GFI1* fails to maintain long-term hematopoiesis due to the depletion of the HSC pool and lack of neutrophils. GFI1 has also been shown to be part of the wider HSC regulatory network because it is known to bind to TAL1, PU.1, RUNX1, and GATA2 (Wilson et al., 2010). Like GFI1, also the Polycomb complex protein BMI1 is a master regulator of HSCs and is required for the self-renewal and long-term maintenance of HSCs (Hidalgo et al., 2022; Park et al., 2003).

GATA-2 is required for the maintenance of primitive and definitive hematopoiesis during the fetal period, as well as in adulthood (Vicente et al., 2012; Zhu and Emerson, 2002). GATA-2 is an important TF for the proliferation and maintenance of HSCs and multipotent progenitors (Vicente et al., 2012). *GATA-2* knockout leads to defects in primitive erythropoiesis, leading to lethality in mouse embryos. In definitive hematopoiesis, the absence of *GATA-2* prevents the development of hematopoietic lineages in both the fetal liver and adult bone marrow (Zhu and Emerson, 2002).

Two transcription factors PU.1 and Ikaros, are essential TFs in the regulation of lymphopoiesis. PU.1 is required for the generation of CLPs from HSCs and activates lymphoid genes directing cell differentiation toward LMPPs (Pang et al., 2018). PU.1 concentration also regulates the differentiation of B cells and macrophages: low concentration of PU.1 induces B cell development, whereas high concentration promotes macrophage differentiation and blocks B cell development, respectively (DeKoter and Singh, 2000). It has been shown that progenitors lacking *PU.1* expression do not express the *IL-7Ra* gene, thus preventing IL-7R mediated B cell differentiation (DeKoter et al., 2002). Ikaros promotes the transcriptional priming of lymphocyte differentiation-promoting genes in MMPs (Ng et al., 2009; Yoshida and Georgopoulos, 2014). After commitment to the lymphoid lineage, Ikaros expression increases in B and T cell precursors, which is necessary to support cell maturation and prevent leukemogenesis (Yoshida and Georgopoulos, 2014). It is known that loss of Ikaros promotes self-renewal and proliferation in both T- and B cells and blocks B cell differentiation leading to leukemogenesis (Schwickert et al., 2014; Wang et al., 1996; Yoshida and Georgopoulos, 2014).

In B cell fate initiation, E2A, EBF1, and PAX5 are upregulated. PAX5 is a master regulator of B cell fate, whereas EBF1 represses genes necessary for NK, dendritic, and T cell lineage differentiation, and together with E2A, induces B cell lineage-specific genes (Kwon et al., 2008). The T cell fate initiation requires NOTHC1 signaling, which activates the expression of a variety of T cell lineage-specific genes and downregulates genes required for B and NK cell development (Li and von Boehmer, 2011). *TCF1* is one of the NOTCH1 target genes, which in turn upregulates *GATA3* expression (Weber et al., 2011). A balance in *GATA3* expression is required for T progenitor cells to develop toward the DN2 stage (Ho et al., 2009). Too low *GATA3* expression impairs differentiation, and excessively high expression blocks the T cell development and drives cells towards the megakaryocyte, erythrocyte, and mast cell differentiation (Ho et al., 2009).

2.2 Leukemia

Leukemias are a heterogeneous group of hematological malignancies characterized by an increased number of immature blast cells in the BM and blood (Figure 5.). (Juliusson and Hough, 2016). According to World Health Organization (WHO) almost 500 000 new leukemia cases were diagnosed worldwide during year 2020,

making it the most common cancer in children, and among adults, it is included in the top 15 most common cancers (Kampen, 2012; https://www.wcrf.org/cancer-trends/worldwide-cancer-data/, 7.12.2022). In the early 19th century first description of leukemia as a "milky blood" was made by Peter Cullen, and later in 1845 the disease was named as leucocythemia based on morphological studies by John Bennet (Kampen, 2012; Piller, 2001). Leukemias are divided into chronic and acute forms based on the speed of disease progression, and lymphoblastic or myeloid lineage leukemias according to the cell phenotype affected by malignancy (Bain, 2017; Bullinger and Fröhling, 2013).

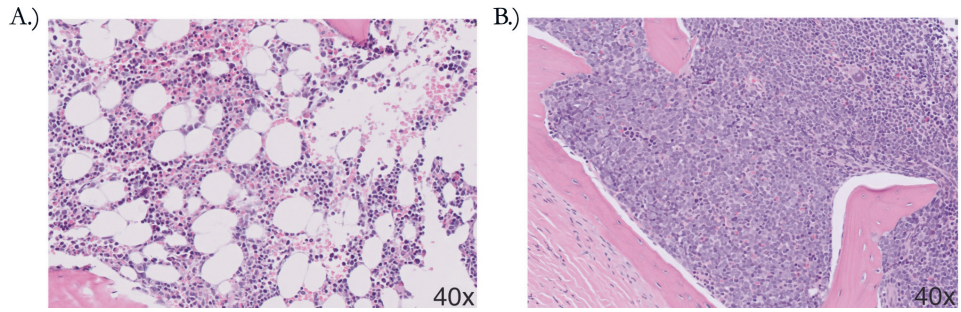


Figure 5. Healthy and leukemic bone marrow. Hematoxylin&eosin staining of (A) healthy bone marrow, which contains multiple hematopoietic cell linages in various maturation stages, and (B) bone marrow of a patient with ALL, which is filled leukemic blast cells.

2.2.1 Acute lymphoblastic leukemia

Acute lymphoblastic leukemia (ALL) can occur in both children and adults, but it is most common in children and adolescents younger than 20 years of age. Approximately 60% of ALL cases occur in this age group, making ALL the most common pediatric cancer (Iacobucci et al., 2021; Inaba et al., 2013; Inaba and Pui, 2021). Yearly, approximately 4000 children in Europe and 50 in Finland are diagnosed with ALL (Lohi et al., 2013; Pitkäniemi, 2020; Steliarova-Foucher et al., 2017). ALL is more frequent in boys than in girls, and also race and ethnic group affect the ALL incidence; majority of the new ALL cases are diagnosed among the Hispanic population (Inaba and Pui, 2021). ALL can be classified based on cell immunophenotyping into B-lymphoid lineage ALL (B-ALL), comprising approximately 85% of the cases, and T-lymphoid lineage ALL (T-ALL). Nowadays, ALL can be subdivided into over 20 B-ALL and more than 10 T-ALL subtypes

(Figure 6.) (Inaba and Pui, 2021). These subtypes are characterized by chromosomal gains and losses (aneuploidy) and chromosomal rearrangements, which often deregulate oncogenes or encode oncogenic fusion proteins, single point mutations, enhancer hijacking events, and aberrant gene expression profiles. Secondary mutations and DNA copy number alterations are critical for driving leukemogenesis in most ALL subtypes (Iacobucci et al., 2021; Inaba et al., 2013; Inaba and Pui, 2021; Pui and Evans, 2006).

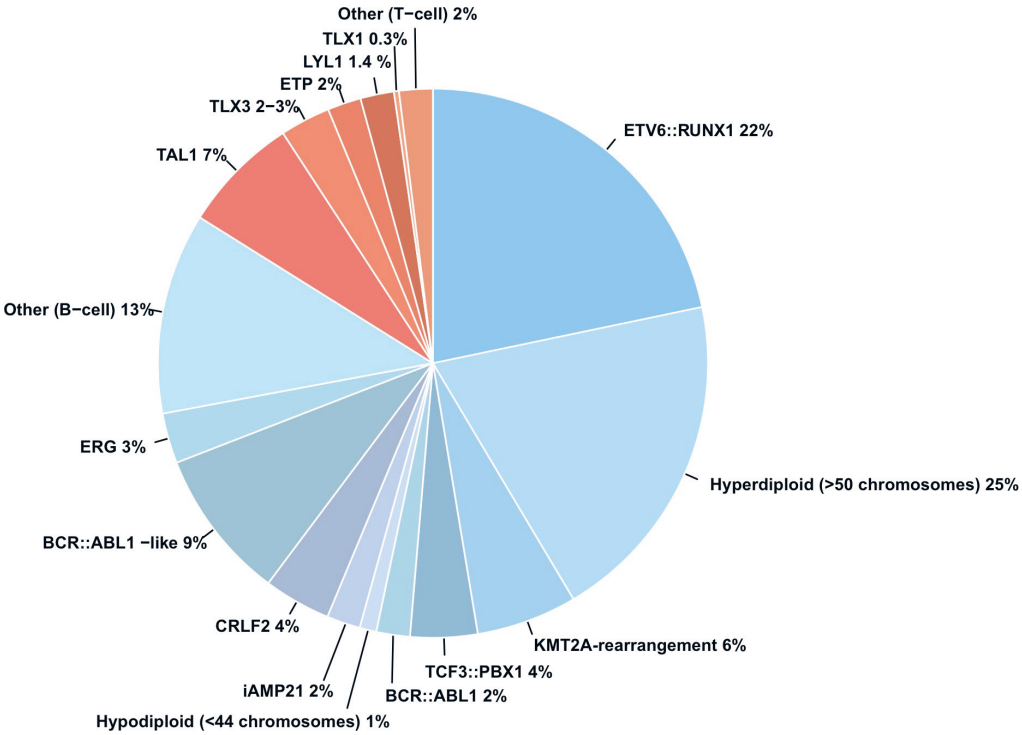


Figure 6. Estimated frequency of childhood ALL subtypes. Blue and orange segments indicate B-ALL and T-ALL subtypes, respectively. Modified from Inaba et al. 2013 and Inaba & Pui 2021.

In high-income countries, the cure rate for childhood ALL exceeds 90%. However, the outcome in adults has remained poor, and less than 40% of adults over 40 years of age can be cured (Inaba and Pui, 2021; Pui et al., 2008, 2004). In the last decades, MRD-based treatment stratification has further improved the outcome of the ALL patients, and new targeted therapies and immunotherapy approaches

have shown promising results (Inaba and Pui, 2021; Jeha et al., 2019; Pui et al., 2009; van Dongen et al., 2015).

Cancer in general arises from interactions between exogenous (e.g., lifestyle-related risk factors and infection) or endogenous (e.g., oxidative stress) exposures and innate genetic susceptibility (Greaves, 2018; Inaba et al., 2013; Schmidt et al., 2021). In adults, lifestyle-related risk factors, such as smoking, obesity, and abundant alcohol consumption, play major roles in the onset of many cancer types. In contrast, lifestyle-related risk factors are not significantly associated with cancer development in children (Schmidt et al., 2021). The association between maternal exposures during pregnancy (e.g., consumption of alcohol, tea, coffee, or smoking) and leukemia development has been studied, but no clear association has been found (Greaves, 2018). Likewise, epidemiological studies have suggested that many environmental exposures contribute to childhood ALL; however, only ionizing radiation has been accepted as a causal environmental agent in leukemia development (Doll and Wakeford, 1997; Greaves, 2018; Hsu et al., 2013; Inaba et al., 2013). In addition, the relevance of background and natural radiation in ALL development has been studied, and has been shown to be associated with a slightly increased risk of ALL (Kendall et al., 2011; Nikkilä et al., 2016). In some ALL subtypes, predisposing "first hit" can already be found at the time of birth. For example, *ETV6::RUNX1* gene fusion can be found in 1:100 newborns at birth; however, only 1:10 000 develop leukemia.

Exposure to infections as a causal factor in the development of childhood ALL was suggested over 100 years ago (Greaves, 2018; Ward, 1917). Later in 1988, two revised models were suggested by Leo Kinlen and Mel Greaves, where leukemia arose as a consequence of an abnormal immune response to common infection(s) (Greaves, 2018, 2006; Kinlen, 1988). The Kinlen hypothesis, also known as "the population-mixing hypothesis," is based on the model that childhood leukemia is caused by an abnormal reaction to a common infection of low pathogenicity in a population lacking herd immunity (Greaves, 2018; Kinlen, 2011, 1988). Greave's "delayed infection" hypothesis is based on more biological origin than epidemiological and proposes the role of infection in the context of a "two-hit" model (Greaves, 1988). In this model, an underexposed immune system during infancy leads to dysregulated responses to common infections later in childhood that could overwhelm the lymphocyte regeneration capacity and trigger leukemia development by causing deleterious secondary mutations in susceptible individuals (e.g., carriers of *ETV6::RUNX1* fusion) (Greaves, 2018, 2006, 1988).

Some evidence of a genetic predisposition to ALL has been reported. Rare constitutional syndromes (e.g., Down syndrome), familial cancer syndromes (e.g., Li-Fraumeni syndrome), non-coding polymorphisms, and alterations in specific genes (e.g., *ARID5B*, *ETV6*, *TP53*, *PAX5*, *IKZF1*, *RUNX1, and CDKN2A/2B)* harboring germline non-silent variants have been shown to increase the risk of ALL (Inaba and Mullighan, 2020; Treviño et al., 2009).

Methods for disease classification

Leukemia is classified at the time of diagnosis based on integration of cell morphology, immunophenotyping, cytogenetics, genetics, and clinical data (Alaggio et al., 2022; Arber et al., 2022; Swerdlow et al., 2016). Morphological and immunophenotyping studies provide the first glimpse of the disease, providing information on lymphoblast involvement in blood and BM, cell lineage, and markers, which are helpful for later assessment of MRD (Alaggio et al., 2022; Béné et al., 2011; Malard and Mohty, 2020). For chromosomal analysis and detection of smaller genomic abnormalities, fluorescence in-situ hybridization (FISH), SNP or CGH arrays, and RT-PCR are used (Alaggio et al., 2022; Malard and Mohty, 2020). Using these details, leukemias can be classified into subtypes for example, according to WHO classification guidelines. A recent update of the 2016 WHO classification of leukemias introduced several new ALL subtypes (Alaggio et al., 2022; Arber et al., 2022). Detection and diagnosis of novel subtypes often require advanced next-generation sequencing (NGS) methods, such as whole transcriptome sequencing, exome sequencing, and/or panel sequencing (Arber et al., 2022). These new NGS methods will be more widely used in the future as they become routinely available and more affordable (Malard and Mohty, 2020).

High-throughput DNA sequencing methods, such as whole genome sequencing (WGS), whole exome sequencing (WES), and panel sequencing, provide information on genetic alterations in tumor samples compared to normal (healthy) samples (Arber et al., 2022; Buzdin et al., 2021; Inaba et al., 2013; Meldrum et al., 2011). This data can provide information from one single base pair mutations to large DNA copy number alterations, which could be used to replace some traditional cytogenetic methods, predict the outcome, and design targeted treatments for the patients in the clinic (Meldrum et al., 2011; Pleasance et al., 2010). RNA sequencing methods provide crucial information on gene expression profiles in bulk cells or at the single-cell level (Buzdin et al., 2021; Meldrum et al., 2011). For example, two newer ALL subtypes *BCR::ABL1*-like and *ETV6::RUNX1* fusion-like have similar gene

expression profiles when compared to *BCR::ABL1* and *ETV6::RUNX1* subtypes, even though they do not have the translocation causing the fusion protein expression (Arber et al., 2022; Iacobucci et al., 2021; Lilljebjörn et al., 2016). To identify these subtypes, the use of RNA sequencing in clinical settings will help to specify the diagnosis. In addition, single-cell genomics and transcriptomics could provide the possibility of identifying different cell populations in tumor samples and help to identify drug-resistant cell populations in the early phase, which might affect the possibility of relapse (Buzdin et al., 2021; Mehtonen et al., 2020). Combining these methods, multiplatform genomic analysis can for example help identify cancer-driving genes, gene expression profiles, and tumor clonality to obtain diagnostic and more detailed prognostic information (Brady et al., 2022).

2.2.2 T cell acute lymphoblastic leukemia

T cell acute lymphoblastic leukemia (T-ALL) is an aggressive hematological cancer accounting for approximately 15 – 25 % of pediatric and adult ALLs (Girardi et al., 2017). Despite the significant improvements during the last decade, the prognosis of T-ALL is less favorable compared to B-ALL (Roti G and Stegmaier K, 2014). The five-year-event-free survival (EFS) and overall survival (OS) rates for T-ALL vary depending on the treatment protocol used, but they are generally around 76–82% and 80–88%, respectively (Patrick et al., 2014; Quist-Paulsen et al., 2020; Schrappe et al., 2011; Winter et al., 2018). T-ALL patients often suffer from failure of induction therapy or early relapse, which is associated with a dismal prognosis. According to the NOPHO ALL2008 study, approximately 14% of T-ALL patients experience relapse, and only about 8% of relapsed patients were alive at the time of the last follow-up (Quist-Paulsen et al., 2020).

Genetic basis of T-ALL

T-ALL is caused by the accumulation of genomic alterations, which disrupt the key developmental pathways responsible for normal control of T cell differentiation, proliferation, and survival (Girardi et al., 2017; Sanchez-Martin and Ferrando, 2017). T-ALL can be subclassified according to the stage of cell maturation to pro-, pre-, cortical, or mature T cell leukemias (Liu Y et al., 2017). Molecular subtype classification of T-ALL can be done by the detection of a specific transcriptional profiles or ectopic expression of a particular oncogenic transcription factor (Girardi

et al., 2017). There are six major T-ALL subgroups, the most common of which is defined by the expression of *TAL1* (sometimes together with *LMO1/LMO2*). Other major subgroups are defined by the expression of *TLX1*, *TLX3*, *HOXA*, *LMO2,* or *NKX2-1* (Girardi et al., 2017; Liu Y et al., 2017). The early T cell precursor subgroup can be identified by the expression of *LYL1* (Girardi et al., 2017; Zhang et al., 2012). The classification of T-ALL subgroups is not fully clear, and the number of subgroups varies in the literature. For example, Liu et al. (2017) presented two rare subtypes defined by the expression of LMO2/LMO1 (10/264 T-ALL cases) and TAL2 (8/264) (Liu Y et al., 2017).

Genetic alterations leading to T-ALL pathogenesis are most frequently activating or inactivating mutations leading to the aberrant function of signaling pathways, transcription factors, epigenetic factors, or translation and RNA stability (Girardi et al., 2017; Inaba and Mullighan, 2020). Rearrangements leading to aberrant expression of target genes or formation of chimeric fusions are well known. Most commonly, T-ALL can be characterized by either the deregulated expression of T-linage transcription factors, altered NOTCH1/MYC signaling, or loss of cell cycle control (Inaba and Mullighan, 2020; Sanchez-Martin and Ferrando, 2017).

Class I transmembrane protein NOTCH1 is a ligand-activated TF that transduces extracellular signals, and with a complex of other TFs and coactivators, leads to changes in gene expression. NOTCH1 signaling is essential for T cell development and differentiation, and its oncogenic activation is involved in almost 80% of T-ALL cases (Liu Y et al., 2017; Palomero et al., 2006; Sanchez-Martin and Ferrando, 2017). The NOTCH1 activation is mostly caused by activating *NOTCH1* mutation (in >75% of cases), or loss-of-function of negative regulators of NOTHC1, such as a mutation in *FBXW7* (in >25% of cases) leading to increased NOTCH1 protein stability (Girardi et al., 2017; Inaba and Mullighan, 2020; Sanchez-Martin and Ferrando, 2017). In 1% of T-ALL patients, oncogenic activation of *NOTCH1* is caused by chromosomal rearrangement that places N-terminal truncated NOTCH1 under the control of strong T cell-specific enhancer close to *TCRB* locus, leading to the aberrant expression of an intracellular constitutively active form of NOTCH1 (Ellisen et al., 1991; Palomero et al., 2006; Sanchez-Martin and Ferrando, 2017). The frequency of *NOTCH1* mutation is the highest in early cortical T-ALL subtypes (*TLX1* and *TLX3*) compared to the early T cell precursor (*LMO2/LYL1* and *HOXA*) and late cortical (*NKX2-1* and *TAL1*) subtypes (Liu Y et al., 2017; Sanchez-Martin and Ferrando, 2017). NOTCH1 can also directly activate *MYC* expression via the interaction of long-range NOTCH-MYC enhancer (NMe), in which focal duplication is seen in about 5% of T-ALL patients (Herranz et al., 2014; Sanchez-

Martin and Ferrando, 2017). During normal hematopoiesis, MYC maintains the balance of HSC differentiation and plays a role in lymphoid cell development and maturation. Overexpression of *MYC* leads to uncontrolled growth, proliferation, and leukemia-initiating cell activity in T-ALL (Herranz et al., 2014; Sanchez-Martin and Ferrando, 2017).

Interleukin 7 (IL-7) signaling is essential for normal T cell development and homeostasis in mature T cells. This signaling pathway is triggered when IL7 binds to its heterodimeric receptor IL-7R (formed by subunits α and γc) (Girardi et al., 2017; Ribeiro et al., 2013; Zenatti et al., 2011). Around 10% of T-ALL patients have activating mutations in *IL-7R*, and in most cases it leads to the formation of disulfide bonds between two IL-7Rα subunits, which constitutively activate the downstream signaling pathways, most commonly PI3K-AKT, JAK-STAT and RAS-MAPK pathways, driving cells towards tumorigenesis (Ribeiro et al., 2013; Zenatti et al., 2011). IL-7 signaling can be indirectly enhanced by *NOTCH1* expression, which upregulates the expression of IL-7R on the cell surface of thymocytes (Girardi et al., 2017; González-García et al., 2012).

Mutations activating kinase signaling pathways in T-ALL are common (65% of cases), and the most frequently mutated signaling pathways are PI3K-AKT (~29% of cases), JAK-STAT (~25% of cases) and Ras (~14% of cases) signaling pathways (Liu Y et al., 2017). Also, rearrangement of *ABL1* kinase is observed in 6% of T-ALL cases (Graux et al., 2009; Inaba and Mullighan, 2020). PI3K-AKT activation in T-ALL is mostly caused by inactivating mutation or deletion of phosphatase and tensin homolog (*PTEN*), which is the main negative regulator of this pathway. In some T-ALL cases, the PI3K-AKT pathway is also activated by gain-of-function mutations in regulatory and catalytic subunits of PI3K or downstream effectors of the signaling cascade such as *AKT* or *mTOR* (Girardi et al., 2017). Also, IL7 stimulation or RAS-MAPK pathway activation can lead to enhanced PI3K-AKT signaling (Cully et al., 2006; Girardi et al., 2017). The PI3K-AKT cascade is a crucial signal transduction pathway for normal cell growth and survival, and its aberrant activity leads to enhanced cell metabolism, proliferation, and impaired apoptosis (Bressanin et al., 2012; Cully et al., 2006). PI3K-AKT mutations are more frequent in more mature TAL1 subtype T-ALL compared to other subtypes (Liu Y et al., 2017). Activation of the JAK-STAT pathway occurs due to activating mutations in *IL7R*, *JAK1*, *JAK3,* or *STAT5B*, (20-30% of cases) or loss-of-function in JAK-STAT negative regulators *PTPN2* and *SH2B3* (less than 5% of the cases) (Inaba and Mullighan, 2020; Kontro et al., 2014; Liu Y et al., 2017; Zenatti et al., 2011). Activation of the pathway by phosphorylation of JAK1 and JAK2 leads to STAT5

dimerization, after which it translocates to the nucleus and regulates multiple target genes, such as, antiapoptotic B-cell lymphoma 2 (BCL-2) family of proteins (Girardi et al., 2017). Mutations in the JAK-STAT pathway are most commonly present in more immature TLX1, TLX3, and HOXA T-ALL subtypes (Liu Y et al., 2017). Oncogenic activation of RAS-MAPK signaling is caused by point mutations, deletions, or chromosomal translocations in multiple genes, including integral components of the pathway (such as *RAS*, *BRAF*, *PI3K*, and *PTEN*), upstream activators (RTKs, such as *FLT3*), and regulatory proteins (Knight and Irving, 2014; Steelman et al., 2011). Activating mutations of *NRAS* and *KRAS* are common in ALL, and activated RAS has been reported in approximately 50% of all T-ALL cases (Kindler et al., 2008; von Lintig et al., 2000). Compared to healthy T cells, in T-ALL RAS-MAPK pathway can be also activated by IL7 (Girardi et al., 2017). The RAS-MAPK signaling pathway has a key role in regulating diverse cellular functions, including cell proliferation, differentiation, survival, and apoptosis, and aberrantly activated RAS-MAPK signaling pathway leads to uncontrolled proliferation and survival of ALL cells (Chung and Kondo, 2011; Knight and Irving, 2014). Mutations in RAS signaling are most commonly present in the immature HOXA and ETP-ALL subtypes (Liu Y et al., 2017). Activating *RAS* mutations have also been associated with relapse risk, poorer overall outcome, and resistance to methotrexate (Irving et al., 2014; Oshima et al., 2016).

Activation of ABL1 kinase is observed in approximately 8% of T-ALL cases. In 6% of the cases, chimeric fusion gene *NUP214::ABL1* is present, whereas *BCR::ABL1* and *ETV6::ABL1* fusions are present in less than 1% of the cases (Girardi et al., 2017; Graux et al., 2009). *NUP214::ABL1* forms a constitutively active tyrosine kinase, activating STAT5 and the RAS-MAPK pathways; however, its kinase activity is weaker compared to BCR::ABL1 and ETV6::ABL1 protein fusions (Graux et al., 2009). It is shown that NUP214::ABL1 fusion kinase functions as a weak oncogene, which cooperates with secondary oncogenic events to drive T-ALL development (Graux et al., 2009). In a study by Liu et al. 2017, *NUP214::ABL1* was not associated with a specific T-ALL subtype, suggesting it may drive signaling pathways in multiple stages of T cell maturation (Liu Y et al., 2017), even though earlier studies have shown that *NUP214::ABL1* fusion was often found with the expression of TLX1 or TLX3 (Graux et al., 2009, 2004).

The oncogenesis of TCR signaling in T-ALL is not well understood, and no mutations in the central signaling components have been found in T-ALL. However, translocations involving TCR loci and several different partner genes are found in ~35% of T-ALL cases (Kang et al., 2012; Larmonie et al., 2013). Usually, these

translocations lead to the deregulation of partner genes located near the TCR regulatory element (Kang et al., 2012; Larmonie et al., 2013). A hallmark of TCR signaling is its sensitivity and specificity; both TCR signaling that is too weak and too strong may disrupt normal T cell function. Mouse studies have shown that basal or weak signaling favors leukemogenesis over normal T cell development, and a high level of TCR signaling can be detrimental to leukemia cells (Courtney et al., 2018; dos Santos et al., 2019). However, TCR signaling resides upstream of pathways such as RAS-MAPK and PI3K-AKT, which have a well-known role in T-ALL development (Courtney et al., 2018). Instead of T-ALL, aberrant expressions of TCR components, such as *LCK* and *ZAP70*, have been observed in chronic lymphocytic leukemia (CLL), acute myeloid leukemia (AML), and B-ALL (Bommhardt et al., 2019; Sadras et al., 2021).

Cell cycle regulation and tumor suppression pathways are commonly mutated in T-ALL (~80% of cases) (Liu Y et al., 2017). Deletion of tumor suppressor *CDKN2A/CDKN2B* is observed in over 70% of T-ALL cases, and deletions of *RB1* and *CDKN1B*, encoding cell cycle regulators, are observed approximately in 15% of cases respectively (Liu Y et al., 2017).

Mutations in genes affecting translation machinery have been described in ~20% of T-ALL cases (Girardi et al., 2018; Liu Y et al., 2017). Mutations affecting the function of ribosomal proteins RPL5 and RPL10 have been described in 10% of pediatric T-ALLs, and in addition, *RPL22* is inactivated in another 10% of T-ALLs (Girardi et al., 2018). Rare mutations in *RPL11* have been found in relapsed T-ALL patients (Girardi et al., 2018). *RPL10* is an oncogenic driver gene in T-ALL (Liu Y et al., 2017). The *RPL10* R98S missense mutational hotspot is known to impair ribosome formation and proliferation in mouse lymphoid cells (Girardi et al., 2018; Kampen et al., 2019). Cells expressing *RPL10* R98S mutation have shown overexpression of several JAK-STAT signaling pathway proteins, and it is suggested that *RPL10* R98S mimic the JAK-STAT activation T-ALL (Girardi et al., 2018). RPL22 functions as a tumor suppressor, and the loss of one copy of *RPL22* is sufficient to enhance T cell leukemogenesis by activating NF-κB and inducing the stem cell factor Lin28B (Rao et al., 2016).

Transcriptional Regulation in T-ALL

Deregulation of TFs by activating or inactivating mutations is a common mechanism by which progenitor T cells gain oncogenic properties and progress towards a malignant cell state (Roy and Raghavan, 2021).

TAL1 is the most commonly deregulated TF in T-ALL as it is overexpressed in over 30% of pediatric T-ALLs (Girardi et al., 2017; Liu Y et al., 2017). It has a key role in normal hematopoiesis, where its expression is decreased during T cell development (Tan et al., 2019). In T-ALL, all mechanisms deregulating *TAL1* expression lead to the high expression of the *TAL1* gene (Tan et al., 2019). *TAL1* overexpression is caused by chromosomal translocations, intrachromosomal rearrangements, or somatic mutations in intergenic elements (Tan et al., 2019). One hallmark of T-ALL is the *TCR* gene translocations with genes, which locates the partner gene under the control of the TCR regulatory element and leads to abnormal expression of the partner gene (e.g. *TAL1*) (Begley et al., 1989; Look, 1997; Tan et al., 2019). The most common mechanism driving *TAL1* expression in T-ALL is the SIL-TAL1 fusion, which is caused by intrachromosomal deletion between the *TAL1* and *SIL* genes that switches *TAL1* under the control of *SIL* gene regulatory element (Janssen et al., 1993; Liu Y et al., 2017). Also, an intergenic somatic mutation upstream of *TAL1* can introduce a MYB-binding site and an enhancer (*MuTe*), driving high *TAL1* expression (Mansour et al., 2014; Navarro et al., 2015). TAL1 can mediate the oncogenic effect by two main mechanisms: disrupting the tumor suppressor E2A function by forming TAL1-E2A heterodimers, or through its participation in the core regulatory circuit of T-ALL (Palii et al., 2011; Sanda et al., 2012). In core regulatory circuits, TFs form interconnected feed-forward transcriptional loops to drive cells to specific expression programs that maintain the state of normal or malignant cells (Boyer et al., 2005; Saint-André et al., 2016). In T-ALL, TAL1 regulates the expression of the *ETS* family, *GATA3*, *MYB*, and *RUNX1* genes (Sanda et al., 2012; Tan et al., 2019). These TFs form a complex that binds to several regulatory elements, including their own enhancers, and drives oncogenic gene expression programs (Palii et al., 2011; Sanda et al., 2012). Even though *TAL1* acts as an oncogene in T-ALL, multiple studies have shown, that *TAL1* alone is not capable of initiating the full T-ALL development, but needs the secondary hits, such as co-expression of *LMO1/LMO2*, for full oncogenic capability (Girardi et al., 2017; Tan et al., 2019).

In addition to being part of the core regulatory circuit, *GATA3*, *MYB*, *RUNX1*, and *ETS* family member *ETV6* are mutated in T-ALL. *GATA3* inactivating mutations or deletions are found in ~5% of pediatric T-ALLs. Low *GATA3* expression is especially associated with ETP-ALL. However, the *GATA3* expression level has not been noted to affect survival (Fransecky et al., 2016). *MYB* overexpression in T-ALL (~7% of the cases) is caused commonly due to rearrangements, for example with the *TCR* gene, duplications, or activating

mutations (Liu Y et al., 2017). *MYB* functions as an oncogene by targeting genes that regulate cell proliferation, cell cycle progression, and cell survival. *RUNX1* mutations and deletions are observed in 8% of pediatric T-ALL, with a trend of higher incidence in ETP-ALL cases (Grossmann et al., 2011; Liu Y et al., 2017). *RUNX1* mutations are associated with poor overall survival in T-ALL (Grossmann et al., 2011). Loss-of-function alterations of the *ETS* family member *ETV6* gene is observed in ~8% of pediatric T-ALL cases, mostly in more immature T-ALL subtypes (Girardi et al., 2017; Liu Y et al., 2017; Nishii et al., 2021; Roy and Raghavan, 2021; Van Vlierberghe et al., 2011). In immature T-ALLs, in 80% of the cases, *ETV6* mutations co-occur with *NOTHC1* mutations (Roy and Raghavan, 2021). Loss of *ETV6* expression leads to loss of transcriptional repression activity, transforming the cell in cooperation with other TFs. Also, loss-of-function alterations in *BCL11B*, *LEF1*, and *WT1* have been observed in pediatric T-ALL with frequencies of 10-19% (Girardi et al., 2017; Liu Y et al., 2017). BCL11B has a key role in maintaining T cell identity and differentiation. Its loss-of-function alterations in T-ALL confer NK cell-like properties to T cells, resulting in enhanced proliferation and survival of immature T cells (Roy and Raghavan, 2021). LEF1 is a key mediator of Wnt signaling and it has a crucial role in early lymphocyte development (Guo et al., 2015). In pediatric T-ALL, *LEF1* inactivating alterations have been associated with activating *NOTCH1* and PI3K/AKT mutations and loss of *PTEN* (Gutierrez et al., 2010). Low *LEF1* expression functions as a biomarker for the ETP T-ALL subtype, while higher *LEF1* expression indicates a favorable prognosis in children (Jia et al., 2015; Wang and Zhang, 2020). *WT1* is considered a tumor suppressor; however, depending on cell context, it might also have oncogenic properties (Loeb and Sukumar, 2002; Tosello et al., 2009). In T-ALL, *WT1* inactivating mutations and deletions are most frequently associated with the TLX3 subtype, but it is also mutated in ETP ALL, especially in LMO2/LYL1 subtype leukemias (Liu Y et al., 2017). It has also been shown that higher *WT1* expression is associated with higher age and reduced expression, with a higher risk of relapse in children (Boublikova et al., 2006).

Epigenomics in T-ALL

Disruption of epigenomic homeostasis in T cells plays a significant role in T-ALL progression. Pediatric T-ALL has a high incidence of epigenetic lesions, and around 56% of the patients have mutations in genes encoding epigenetic regulators (Girardi et al., 2017; Roy and Raghavan, 2021). Epigenetic modifications include DNA

methylation, histone modification, and nucleosome remodeling (Van der Meulen et al., 2014). DNA methylation is executed by DNA methyltransferases DNMT2, DNMT3A, and DNMT3B. Compared to adult T-ALL, where mutations in *DNMT3A* are observed in 4-18% of the cases, DNA methyltransferase mutations in children T-ALL are rare and seen only in ~1% of the cases (Roy and Raghavan, 2021). Histones undergo multiple different post-translational modifications, including acetylation, methylation, phosphorylation, ubiquitinylation, and SUMOylation, mediated by various epigenetic regulators. Mutations in histone acetyltransferases (HATs) and deacetylases (HDACs) have been observed in both pediatric and adult T-ALL. Typically, they regulate the balance of acetylation and deacetylation of histones and non-histone proteins, such as MEB, E3F1, TP53, and RB1, leading to tightly regulated gene transcription. Disruption of this balance by somatic mutations affects the initiation of cancer. Also, inactivating mutations in the core components (*EZH2*, *SUZ12*, and *EED*) of PRC2 (Polycomp repressive complex 2) have been observed in almost half of pediatric ETP ALL cases (Zhang et al., 2012). This complex mediates transcriptional gene repression by regulating the trimethylation of histone 3 lysine 23 (H3K27), and the loss of function of PRC2 components has been shown to lead to the initiation of T cell malignancy (Girardi et al., 2017; Roy and Raghavan, 2021; Zhang et al., 2012). Alterations in *USP7*, which is responsible for deubiquitylation, are observed exclusively in TAL1 subtype T-ALL cases. Inactivating mutations of this gene are known to destabilize tumor suppressors TP53 and PTEN (Liu Y et al., 2017; Roy and Raghavan, 2021). Tumor suppressor PHF6 (plant homeodomain finger 6) is also recognized as an epigenetic modifier because it contains two PHD-like domains, which are known to interact with DNA and transcriptional regulators. Inactivating alterations of this gene are observed in approximately 20% of pediatric T-ALLs, especially in TLX1 and TLX3 subtypes, and lead to aberrant expression of oncogenic TFs (Liu Y et al., 2017; Roy and Raghavan, 2021; Van der Meulen et al., 2014; Van Vlierberghe et al., 2010).

SIX6 homeobox transcription factor

In addition to key regulators in leukemia, multiple TFs and transcriptional regulators have a role in leukemia. In this thesis the role of The Sine Oculis homeobox (SIX) was studied more closely in the context of T-ALL. The homeobox (HOX) gene superfamily encodes transcription factors that regulate development by activating or repressing their target genes. The expression of these TFs is essential for

embryogenesis and development, but their aberrant expression plays a significant role in tumorigenesis (Abate-Shen, 2002; Feng et al., 2021). One subfamily of the HOX family is The Sine Oculis homeobox (SIX) protein family, which usually participates in the development of the visual system (Kawakami et al., 2000; Lee et al., 2012).

In mammals, the SIX homeobox family comprises six genes (*SIX1-6*) encoding TFs. SIX family members can be characterized by two evolutionarily conserved domains, the homeobox nucleic acid recognition domain (HD) and SIX domain (SD), both of which are essential for DNA binding and cooperative interactions with target proteins (Kawakami et al., 2000). SIX family members are further subdivided into three subgroups based on amino acid sequence similarities and structures of the N- and C-terminal regions (Figure 7.) (Kawakami et al., 2000; Kumar, 2009). SIX proteins play a critical role in tissue formation and organogenesis in many tissues (e.g., retina, brain, skeletal muscle, and kidney), and in regulating cell proliferation, differentiation, and survival (Jean et al., 1999; Kobayashi et al., 1998; Oliver et al., 1995; Self et al., 2006; Xu et al., 2003). Depending on their interactions with other regulatory proteins, these TFs can act as transcriptional activators or repressors (Kawakami et al., 2000; Kumar, 2009; Meurer et al., 2021). The most studied transcriptional activation complexes are formed by Eyes absent (EYA) and SIX proteins (Kumar, 2009; Xu, 2013). SIX proteins, especially SIX1-3 and SIX6, are known to bind to the Groucho/TLE and DACH family proteins to form transcriptional repressor complexes (Kumar, 2009; Meurer et al., 2021). For example, the SIX6-DACH1 complex controls cell proliferation by repressing CDKN1b, an inhibitor of the cell cycle G1/S transition (Li et al., 2002; Wu et al., 2003). In mice, SIX6 has also been shown to positively regulate the transcription of gonadotropin-releasing hormone (GnRH) by directly binding to its promoter, thereby affecting fertility (Larder et al., 2011).

SIX6 has crucial and conserved functions in forebrain and retina development. SIX3 and SIX6 are major SIX family proteins expressed in the hypothalamus and retina (Kumar, 2009; Lee et al., 2012). During eye development, these genes are first expressed in the optic vesicle and optic stalk before expanding to the neural retina (Diacou et al., 2022; Jean et al., 1999; Kumar, 2009). Compared to SIX3, SIX6 expression is not found in the lens (Jean et al., 1999). Gain- and loss-of-function studies in vertebrate embryos have revealed multiple sites of differentiation, where *SIX6* expression is essential for normal eye development (Lee et al., 2012). Expression of *SIX6,* for example promotes the transdifferentiation of pigment epithelial cells into retinal neurons and photoreceptors. However, overexpression

has been shown to cause eye enlargement, ectopic eye-like structures, and retinal tissue in the developing brain. Loss of expression leads to impaired retinal development (Bernier et al., 2000; Lee et al., 2012; Li et al., 2002; Toy and Sundin, 1999; Zuber et al., 1999). In brain development, SIX6 is first expressed in the neural plate and subsequently in the ventral forebrain, where its expression is primarily localized to the hypothalamus and pituitary gland (Jean et al., 1999; López-Ríos et al., 1999). SIX3 and SIX6 both belong to the Optix subgroup, sharing many similarities in protein binding and function; perhaps for this reason, it has been shown that they compensate for each other, both *in vivo* and *in vitro* (Xie et al., 2015).

Cell proliferation or migration promoted by the ectopic expression of SIX family members is associated with the formation, survival, and metastasis of several cancer types (Meurer et al., 2021). High expression of *SIX6* is associated with T-ALL, more specifically, with the TAL1 subtype of T-ALL (Chen et al., 2018; Soulier et al., 2005). In breast and non-small cell lung carcinomas, increased *SIX6* expression is also associated with poor overall survival (Q. Liu et al., 2016; Xu et al., 2016). Also, the *SIX6* methylation level has been shown to be significantly higher in cancer tissues compared to normal samples (Dong et al., 2022). Dong et al. (2022) showed that *SIX6* was universally hypermethylated in cancer samples compared to normal samples when studied across 15 different cancer types from the TCGA data set and 678 clinical tumor samples, including 73 leukemia samples (Dong et al., 2022). Their data also showed that *SIX6* hypermethylation is an early marker of tumor progression, especially when healthy cells transform into precancerous cells.

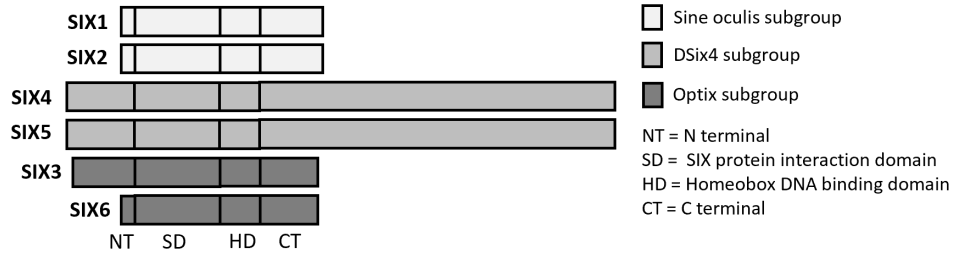


Figure 7. Subgroups and structure of the SIX homeobox transcription factor members. Adapted from the study by Kumar et al. 2009.

Overall, around 0.8% of protein coding genes in human encodes methyltransferases, making the methyltransferases one of the largest enzyme classes (Horning et al., 2016; Katz et al., 2003; Petrossian and Clarke, 2011). Methyltransferases catalyze the addition of methyl groups to the nitrogen, carbon, sulfur, and oxygen atoms of small molecules, lipids, and nucleic and amino acids (Bedford and Richard, 2005; Paik and Kim, 1971). One group of methyltransferases is Protein Arginine Methyltransferases (PRMTs). PRMTs are responsible for arginine methylation. Arginine methylation is a common post-translational modification that occurs on histones, RNA-binding proteins, and many other cellular proteins, affecting their function by altering the protein-protein and protein-nucleic acid interactions (Xu and Richard, 2021). They function as epigenetic transcriptional regulators and play important roles in pre-mRNA processing, DNA damage, cell signaling, immune responses, and cell fate decision (Blanc and Richard, 2017).

The mammalian genome encodes nine sequence-related PRMTs (Bedford and Clarke, 2009; Feng Y et al., 2013; Miranda et al., 2004). In eukaryotic cells, they catalyze reactions in which a methyl group from S-adenosyl-L-methionine (SAM/AdoMet) is transferred to an arginine residue on protein substrates (Miranda et al., 2004; Paik and Kim, 1971). PRMTs catalyze three types of methylation: monomethylation, asymmetric dimethylation, and symmetric dimethylation. PRMTs are divided into three groups based on their methylation capabilities. PRMT1, PRMT2, PRMT3, Coactivator-Associated Arginine Methyltransferase 1 (CARM1, also known as PRMT4), PRMT6 and PRMT8, are type I PRMTs. Type I PRMTs catalyze the formation of ω-NG-monomethylarginine (MMA) and ω-N^G, N^G-asymmetric dimethylarginine residues (ADMA). PRMT5 and PRMT9 are type II PRMTs, and they catalyze the formation of ω-N^G, N'^G-symmetric dimethylarginine (SDMA). PRMT7 is classified as a type III arginine methyltransferase, and it only generates ω-monomethyl arginine (MMA) residues on its substrates, typically at basic RXR sequences in peptides and histones (Miranda et al., 2004). The reactions catalyzed by different PRMT types are illustrated in (Figure 8.).

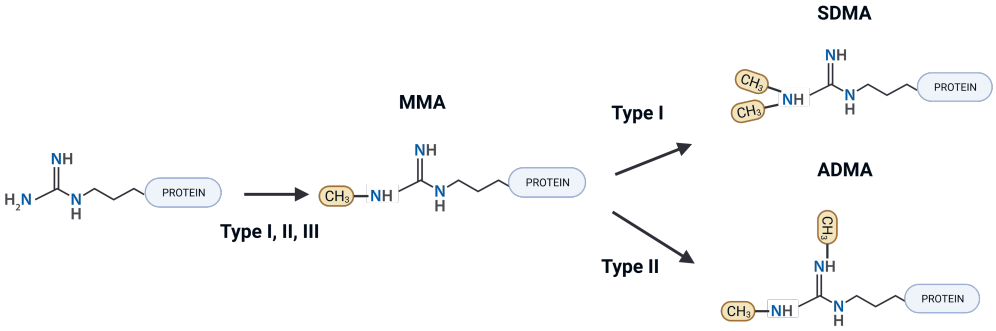


Figure 8. Types of protein arginine methylation catalyzed by different PRMT enzymes. Abbreviations: ADMA, asymmetric dimethylarginine; SDMA, symmetric dimethylarginine; MMA, monomethylarginine. Adapted from Yang and Bedford 2012, and Hwang et al. 2021.

The basic structure of PRMT proteins is similar and consists of two domains: an N-terminal Rossman fold, including the catalytic AdoMet binding domain, and a C-terminal B-barrel domain. Compared to other PRMTs, the activity of PRMT5 requires an additional protein cofactor, for example, MEP50 (methylosome protein 50) (Antonysamy et al., 2012; Ho et al., 2013; Tewary et al., 2019; J. Xu & Richard, 2021). Methylation of the guanidinium moiety increases the hydrophobicity of the arginine residue, thereby increasing the bulkiness of the side chain of the protein. The presence of methylated side chains decreases the hydrogen bonding potential and alters the binding of methylarginine to protein modules that read the arginine methylation marks, such as the PHD finger, Tudor domain, and SH3 domain (Bedford et al., 2000; Côté and Richards, 2005; Li et al., 2021; Liu H et al., 2010; Liu R et al., 2018). Especially, arginine methylation is known to enhance the interactions with Tudor domains (Selenko et al., 2001; Tripsianes et al., 2011).

In this thesis, the role of PRMT7 was studied more closely in the context of T-ALL development. PRMT7 has some unique features compared to other PRMT family members: it catalyzes only the formation of MMA and has two AdoMet-binding motifs. Two AdoMet-binding motifs support the idea that PRMT7 results from gene duplication (Miranda et al., 2004). Although PRMT7 is the least well-known member of the PRMT family, studies conducted over the last 20 years have started to reveal its function. PRMT7 participates in many cellular processes, including transcriptional regulation, DNA damage repair, cellular stress response, RNA splicing, cell differentiation and proliferation, antiviral responses, and the ability of cancer to metastasize (Feng Y et al., 2013; Haghandish et al., 2019; Karkhanis et al., 2012; Liu F et al., 2020; Szewczyk et al., 2020; Zhu et al., 2021). *In vivo* mouse studies have shown that the loss of *PRMT7* leads to developmental delays,

premature aging, and loss of proliferation and differentiation of muscle stem cells, leading to defects in skeletal muscles (Akawi et al., 2015; Blanc et al., 2016; Jeong et al., 2016). Arginine regulation plays a critical role in the maintenance of lymphoid and myeloid lineage cells, especially when catalyzed by PRMT1, CRAM1, and PRMT5 (Blanc and Richard, 2017). Knockout of *PRMT1* and *PRMT5* is known to lead to impaired B cell differentiation, whereas knockout of *PRMT1* and *CARM1* leads to the loss of erythroid differentiation, growth, and survival (Hata et al., 2016; Hua et al., 2013; Liu et al., 2015; Streubel et al., 2013). PRMT5 is essential for promoting hematopoietic cell commitment and the maintenance of progenitor cell expansion in mice (Liu et al., 2015). PRMT7 has been shown to play a role in germinal center formation by regulating *BCL6* expression. In mouse models, loss of *PRMT7* in B cells led to a decreased number of mature marginal zone B cells, an increased number of follicular B cells, and promoted germinal center formation (Ying et al., 2015).

The deregulation of PRMTs has been observed in many cancers (Blanc and Richard, 2017; Yang and Bedford, 2013). Table 1. summarizes the current literature and shows the roles of individual PRMT proteins in tumorigenesis.

Table 1. Deregulation of PRMTs in cancer.

PRMT	Deregulation in Cancer	References
PRMT1	Overexpressed in breast, prostate, lung, colon, and bladder cancer, and AML, ALL, Hodgkin's lymphoma.	Leonard 2012 Zhu 2019 Zou 2012 Shia 2012
PRMT2	Overexpressed in breast cancer. Overexpression in glioblastoma correlates with poor prognosis.	Zhong 2011 Dong 2018
PRMT3	Overexpressed in hepatocellular carcinoma, pancreatic and colorectal cancer, and promotes cell invasion and metastasis in breast cancer.	Hsu 2019 Zhang 2021 Lei 2022 Zhi 2023
CARM1	Overexpressed in the breast, prostate and colorectal cancer, and Hodgkin's lymphoma and AML.	Leonard 2012 Kim 2010
PRMT5	Overexpressed or increased enzymatic activity in gastric, colorectal and lung cancer, and mantel cell lymphoma and AML	Kim 2005 Pal 2008 Wang 2008 Wei 2012
PRMT6	Overexpressed in the bladder and lung cancer, and lymphoma.	Yoshimatsu 2011 Tang 2020 Chen 2022
PRMT7	Overexpression promotes cell invasion and metastasis of breast cancer and the metastatic phenotype in human non-small-cell lung cancer.	Baldwin 2014 Yao 2014 Cheng 2018
PRMT8	Overexpressed in breast, ovarian, and cervical cancer. Overexpression decreases survival in gastric cancer.	Hernandez 2017
PRMT9	Overexpression promotes hepatocellular carcinoma invasion and metastasis.	Jiang 2018

2.2.3 B cell acute lymphoblastic leukemia

B-ALL originates from progenitor B cells, causing the clonal proliferation mainly in BM (Mullighan, 2012). The incidence of B-ALL peaks at the age of 2–6 years (Hauer et al., 2021). Compared to T-ALL, the prognosis of B-ALL is more favorable, with survival rates exceeding 90%. However, the outcome of patients with early relapse is poor (5-year OS 38.0±10.6%) (Oskarsson et al., 2016). Treatment intensity depends on risk factors, and short- and long-term side effects vary considerably among patients (Fulbright et al., 2011; Malard and Mohty, 2020).

Genetics of B-ALL

Most of the B-ALL cases can be categorized into distinct subtypes based on the presence of genetic alterations, such as aneuploidy or chromosomal rearrangements (Mullighan, 2012). These alterations are crucial events in the onset of leukemia; however, they are often insufficient to initiate leukemia on their own (Mullighan, 2012). Novel NGS technologies have revealed numerous new lesions involving inactivating mutations in tumor suppressor and cell cycle genes that regulate lymphoid development and oncogenic signaling pathways (Brady et al., 2022; Gu et al., 2019; Lilljebjörn et al., 2016). Classical B-ALL subtypes are high hyperdiploidy (HeH), hypodiploidy, *ETV6::RUNX1*, *TCF3::PBX1,* and *BCR::ABL1* translocations, and various *KMT2A*-rearrangements.

ETV6::RUNX1 subtype

ETV6::RUNX1 (also known as *TEL::AML1*) fusion is found in approximately 25% of children diagnosed with B-ALL (Bhojwani et al., 2012; Piette et al., 2018; Shurtleff et al., 1995). The peak incidence of *ETV6::RUNX1* leukemia occurs in children in age 1–9 years, and it is rarely seen in adults (Bhojwani et al., 2012; Madzo et al., 2003; Shurtleff et al., 1995). This subgroup B-ALL is generally associated with a favorable prognosis with a good response to chemotherapy and low-risk features.

The translocation t(12;21)(p13;q22) between *ETV6* and *RUNX1* genes, was found in lymphoid leukemias in the 1990s (Kobayashi and Rowley, 1995; Romana et al., 1994). *ETV6* breakpoints are scattered within the 14 kb of the intronic region between exons 5 and 6 and in *RUNX1* within two putative regions in intron 1 (Figure 9.) (Wiemels et al., 1999). In this chromosomal rearrangement two genes located in different chromosomes fuses together, forming a new gene, called fusion gene with altered function compared to wild type genes. ETV6 and RUNX1 are both transcription factors required for normal definitive hematopoiesis. Even though wild-type RUNX1 acts as a transcriptional activator, the ETV6::RUNX1 fusion protein functions as a transcriptional repressor that downregulates RUNX1 target genes (Golub et al., 1995; Guidez et al., 2000; Teppo et al., 2016). *ETV6::RUNX1* also causes overexpression of the erythropoietin receptor (*EPOR*), activation of the JAK-STAT signaling pathway, and upregulation of the anti-apoptotic *BCL-XL* gene, thereby promoting cell survival, progenitor self-renewal, and proliferation (Inthal et al., 2008; Torrano et al., 2011). *ETV6::RUNX1* fusion is not able to initiate leukemia

on its own, but requires secondary genetic events even years before leukemia initiation (Andreasson et al., 2001; Morrow et al., 2004; Wiemels et al., 1999).

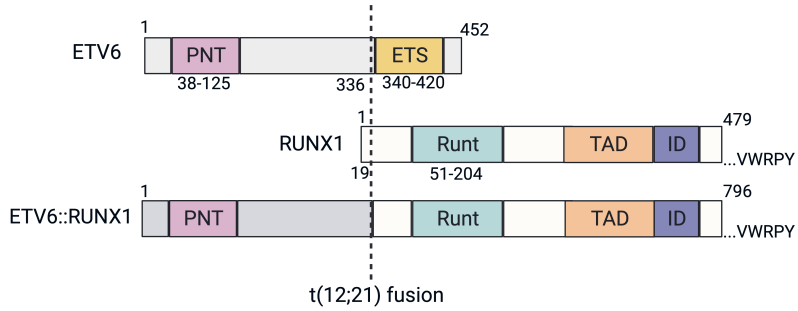


Figure 9. Schematic presentation of full-length ETV6, RUNX1, and ETV6::RUNX1 fusion proteins. The ETV6::RUNX1 fusion protein is a product of the t(12;21) chromosomal translocation (indicated by dashed line), a common genetic abnormality in acute lymphoblastic leukemia. The fusion protein includes the pointed (PNT) domain of ETV6, and the Runt DNA-binding domain (DBD) of RUNX1. However, fusion protein lacks the ETS DBD of ETV6, which is essential to function of wild-type ETV6. The carboxy-terminal VWRPY motif of RUNX1, which can bind Groucho-related corepressors, is also indicated. TAD = transcriptional activation domain, ID = Runx inhibitory domain. Adapted from Teppo et al. 2017 and Zelent et al. 2004. Created with BioRender.com.

ETV6::RUNX1 fusion is commonly detectable already at birth, suggesting that it commonly occurs in fetal hematopoietic progenitors in utero. In cord blood studies, *ETV6::RUNX1* has been found in 2–5% of healthy newborns (Mori et al., 2002; Zuna et al., 2011). Other evidence supporting the prenatal origin of *ETV6::RUNX1* comes from monochronic twins with the same translocation breakpoints when developing *ETV6::RUNX1*-positive leukemia (Ford et al., 1998; Ford and Greaves, 2017). In cases where only one of the twins developed leukemia, pre-leukemic *ETV6::RUNX1* cells were also found in healthy twin (Ford et al., 1998; Wiemels et al., 1999).

Aberrant RAG recombinase activity drives the formation of critical secondary events in *ETV6::RUNX1* leukemia (Papaemmanuil et al., 2014). Secondary genetic abnormalities typically include focal deletions that encompass only one to two genes but also the gains of chromosome arms and extra chromosomes (Brady et al., 2022; Lilljebjörn et al., 2010; Mullighan et al., 2007; Papaemmanuil et al., 2014; Tsuzuki et al., 2007). The most frequently deleted genes co-occurring with *ETV6::RUNX1* are *ETV6*, *CDKN2A/B*, *PAX5*, *ATF7IP*, *TCF4*, *TL1XR1*, *EBF1*, *CD200*, *BTLA*,

TOX, *NR3C1,* and *BTG1* (Brady et al., 2022; Lilljebjörn et al., 2010; Mullighan et al., 2007; Papaemmanuil et al., 2014).

The deletion of wild-type or non-rearranged *ETV6* is the most common secondary event in *ETV6::RUNX1* leukemia. It is present in approximately 70% of all *ETV6::RUNX1* cases, with a variation of 11–90% depending on the study (Brady et al., 2022; Cavé et al., 1997; Lilljebjörn et al., 2010; Mullighan, 2012; Papaemmanuil et al., 2014; Stams et al., 2006; Zakaria et al., 2012)(Raynaud et al. 1996, Cave 1997, Stams 2006, Mullighan 2008, Lilljebjörn 2010, Zakaria 2012, Papaemmanuil 2014, Brady 2022). The exact mechanism of how *ETV6* deletion drives the leukemia onset is unclear, but ETV6 can function as a tumor suppressor, inducing the cell cycle arrest or changing epigenetic cell state (Lopez et al., 1999; Neveu et al., 2018; Rompaey et al., 2000). Another common deletion, seen in up to 25% of *ETV6::RUNX1*-positive leukemias, is a 9p deletion, which often affects the tumor suppressor *CDKN2A/CDKN2B* and B cell differentiation regulator *PAX5* (Brady et al., 2022; Lilljebjörn et al., 2010; Stams et al., 2006). Typical gains are an extra *RUNX1* gene (23% of *ETV6::RUNX1* cases) and/or an extra der(21)t(12;21) chromosome (encoding *RUNX1::ETV6*) (10% of *ETV6::RUNX1*cases) (Stams et al., 2006). The extra copy of der(21)t(12;21) is associated with a more aggressive disease and increased relapse risk (Hernandes et al., 2012; Kobayashi et al., 1996; Peter et al., 2009; Stams et al., 2006).

Despite the good prognosis, some *ETV6::RUNX1* patients have a slow response to induction treatment and carry an increased risk of relapse (Coustan-Smith et al., 2000; Madzo et al., 2003; O'Connor et al., 2018; Wang et al., 2018). Despite the good prognosis, out of all relapsed B-ALL cases, around 20% harbor the *ETV6::RUNX1* fusion (Kuster et al., 2011; Loh et al., 2006; Norén-Nyström et al., 2023). Usually, relapse occurs in patients after cessation of treatment, and it may occur even 10–20 years after the primary diagnosis (Forestier et al., 2008; Konrad et al., 2003; Seeger et al., 2001; van Delft et al., 2011). It has been shown that relapsed *ETV6::RUNX1* leukemia remains chemosensitive, suggesting that its initial survival and reemergence are related to mechanisms other than mutation-based drug resistance (Lawson et al., 2000; van Delft et al., 2011). Nevertheless, a widely recognized genetic alteration that affects drug response and increases the risk of relapse is the deletion of the *NR3C1* gene. *NR3C1* encodes the glucocorticoid receptor, which is necessary for treatment response when using glucocorticoids, which form the backbone of leukemia treatment. *NR3C1* loss-of-function alterations are associated with relapsed *ETV6::RUNX1* leukemias, and they are not found at the time of the diagnosis from the same patient (Kuster et al., 2011). Studies of the clonal origin of relapse in

ETV6::RUNX1 leukemia have revealed that relapse may occur from a minor clone present already at the time of diagnosis. The nature of relapse-associated minor clones is still unclear; however, one suggestion is that these minor clones might represent preleukemic cells at the time of diagnosis, and when they acquire secondary genetic alterations, relapse evolves (Konrad et al., 2003; van Delft et al., 2011; Zuna et al., 2004).

Other B-ALL subtypes

The HeH subtype with 51–67 chromosomes is the most common subtype in childhood, presented in 25–30% of the cases, and is associated with excellent prognosis (Mullighan, 2012; Paulsson and Johansson, 2009; Safavi and Paulsson, 2017). The HeH subtype harbors relatively few additional genetic alterations, suggesting that extra chromosomes are the driving event for leukemia onset. Secondary hits usually affect the RAS signaling pathway and histone modifiers (Inthal et al., 2012; Paulsson et al., 2008). Hypodiploidy, another aneuploidy with 44 chromosomes or fewer, is a relatively rare subtype in children (~2% of cases) and is associated with poor prognosis (Mullighan et al., 2015; Nachman et al., 2007; Safavi and Paulsson, 2017).

The t(1;19)(q23;p13) translocation encoding fusion protein *TCF3::PBX1* is present in 3–6% of childhood B-ALL cases (Andersen et al., 2011; Mullighan, 2012). Typically, translocation is unbalanced, with duplication of 1q distal to *PBX1* gene (Andersen et al., 2011; Mullighan, 2012). *TCF3::PBX1* interferes with hematopoietic differentiation by binding to HOX proteins and disturbing their downstream target gene expression (Mullighan, 2012). Although *TCF3::PBX1* fusion-positive patients have a favorable prognosis with intensified therapy, they have an increased risk for central nervous system (CNS) relapse (Andersen et al., 2011; Jeha et al., 2009; Rivera et al., 1991).

The *BCR::ABL1* fusion is the result of t(9;22)(q34;11.2) translocation, also known as the Philadelphia chromosome (Mullighan, 2012; Ribeiro et al., 1987). In children, *BCR::ABL1* leukemia is rare and is present in approximately 3% of cases (Cazzaniga et al., 2018; Pui and Evans, 2006). *BCR::ABL1* activates multiple signaling pathways that increase cell proliferation (Duy et al., 2011). This subtype has been associated previously with extremely poor outcome, but the addition of tyrosine kinase inhibitors to the treatment has improved the outcome of patients (Cazzaniga et al., 2018; Schultz et al., 2014).

KTM2A-rearrangements are rare, occurring in approximately 1% of B-ALL cases, mainly affecting infants. Patients with this subtype of leukemia typically respond poorly to therapies and have less favorable prognosis (Chessells et al., 2002; Iacobucci and Mullighan, 2022). *KTM2A* can rearrange with multiple genes, but the most common rearrangement is *KMT2A::AF4*, which is found in approximately 70–80% of cases (Iacobucci and Mullighan, 2022). Many of the rearranged partner genes are normally involved in transcriptional regulation, and their alterations may lead to transcriptional dysregulation and drive leukemogenesis (Mullighan, 2012). *KMT2A*-rearranged leukemias may exhibit lineage-switched relapse, where relapsed leukemia cells express myeloid markers (Iacobucci and Mullighan, 2022).

For a long time, approximately 70% of childhood B-ALL cases could be classified routinely, and the remaining 30% were classified into the "B-other" group lacking the known genetic features (Lilljebjörn et al., 2016; Ryan et al., 2023). However, more recently, new genomic tools have uncovered new subtypes among the B-other groups. Approximately 10% of B-ALL cases lack the recurrent genetic abnormalities, and are referred as "B-other NOS" (not otherwise specified). They are typically more heterogeneous than other subtypes of B-ALL, and their prognoses are more variable. In 2016, the WHO revised the classification of lymphoid neoplasms by adding two new subtypes into the classification, namely the *BCR:ABL1*-like and iAMP21 subtypes (Arber et al., 2016). The *BCR::ABL1*-like subtype (also known as Ph-like ALL) is characterized by a gene expression profile similar to *BCR::ABL1*-positive ALL, but lacks the *BCR::ABL1* fusion (Boer et al., 2009; Mullighan et al., 2009). Many *BCR::ABL1*-like cases harbor translocations involving other tyrosine kinases such as ABL2 or PDGFR (Boer et al., 2009; Mullighan et al., 2009). Deletions or mutations in lymphoid transcription factor genes are the hallmark of *BCR::ABL1*-like ALL (Mullighan et al., 2009; Roberts et al., 2014). This subtype is present in approximately 10% of pediatric B-ALL cases and is associated with poor outcome (Roberts et al., 2014).

The iAMP21 subtype is characterized by extra copies of RUNX1 with a complex structure comprising multiple gains, amplifications, inversions, and deletions (Harrison et al., 2014; Heerema et al., 2013). The incidence of iAMP21 in all pediatric B-ALL cases is approximately 2% and is associated with inferior outcome. An improvement in outcome has been observed when patients are treated with high-risk therapy (Harrison et al., 2014).

In a more recent update of the B-ALL classification, 27 subtypes were identified (Arber et al., 2022). The identification of some of these subtypes is dependent on NGS methods, such as RNA sequencing, which are not routinely used in all clinical

laboratories. For example, *ETV6::RUNX1*-like and *DUX4*-rearranged subtypes have been identified using RNA sequencing (Lilljebjörn et al., 2016). *ETV6::RUNX1*-like subtype is found in approximately 3–4% of all B-ALL cases and is characterized by a similar gene expression profile to *ETV6::RUNX1*-positive B-ALL but lacks the translocation (Gu et al., 2019; Lilljebjörn et al., 2016; Zaliova et al., 2017). In both RNA sequencing and methylation studies, *ETV6::RUNX1*-like cases have been shown to cluster with *ETV6::RUNX1* cases (Gu et al., 2019; Lilljebjörn et al., 2016; Nordlund et al., 2015). This subtype often has secondary alterations in the *ETV6*, *IKZF1*, and *ARPP2* genes (Z. Gu et al., 2016; Lilljebjörn et al., 2016; Zaliova et al., 2017).

DUX4-rearranged subtype comprises approximately 4–7% of B-ALL cases (Lilljebjörn et al., 2016; Zhang et al., 2016). It displays *IGH::DUX4* or *ERG::DUX4* fusions, which lead to overexpression of *DUX4*. Intergenic deletions of *ERG* are frequently associated with this subtype (Lilljebjörn et al., 2016). This subtype of leukemia has a favorable prognosis, even with *IKFZ1* deletion, which is commonly associated with poor prognosis (Clappier et al., 2014; Lilljebjörn and Fioretos, 2017; Zhang et al., 2016).

MEFD2 and *ZNF384* rearrangements are both accounting for approximately 3% of pediatric B-ALL patients (Gu et al., 2016; Y.-F. Liu et al., 2016; Yasuda et al., 2016). *MEFD2* can rearrange with multiple genes but is most commonly seen to fuse with *BCL9* (Gu 2016). These rearrangements result in enhanced MEFD2 transcriptional activity. *MEFD2*-rearranged leukemias are sensitive to HDAC inhibitors (HDACi) due to the deregulation of *HDAC9*, which is one of the MEFD2 target genes (Zhaohui Gu et al., 2016). *ZNF384* frequently rearranges with a transcriptional regulator or chromatin modifier, and this subtype of leukemia often expresses myeloid markers (Yasuda et al., 2016). *ZNF384*-rearranged B-ALL has an intermediate prognosis. Upregulation of JAK-STAT is often seen in this subtype and may provide an opportunity for targeted treatment with JAK inhibitors (Yasuda et al., 2016). Together, E/R-like, *DUX4-, MEFD2-* , and *ZNF384* rearranged subtypes of B-ALL account for up to 50% of the historical "B-other" subtype (Lilljebjörn and Fioretos, 2017).

2.2.4 Current treatment of acute lymphoblastic leukemia

ALL treatment is intense and lasts for over two years from diagnosis (Toft et al., 2018). In Nordic countries, pediatric ALL patients have been treated according to NOPHO (The Nordic Society of Paediatric Haematology and Oncology) treatment protocols (Toft et al., 2018; ClinicalTrials.gov: NCT03911128& NCT00819351). There are separate protocols for infants with ALL (Interfant protocols) and Ph+ALL (EsPhALL protocols) (Biondi et al., 2017; Pieters et al., 2019). Likewise, relapsed leukemias have their own designated treatment protocols (IntreALL protocols) (Locatelli et al., 2021). ALL treatment protocols are regularly revised to improve the efficacy and safety of treatments. Currently, patients in Nordic countries are treated according to the semi-European ALLTogether protocol, which was opened in 2021 (ClinicalTrials.gov: NCT03911128). However, the patients in this thesis were treated mainly according to the NOHPO ALL2008 protocol, and for this reason, an overview of the therapy is made using this protocol.

Therapy consisted of induction therapy, consolidation(s), intensification(s), and maintenance phases (Eiser et al., 2017; Inaba and Pui, 2021; Jeha et al., 2019; Toft et al., 2018). Treatment started with the four-week induction therapy, followed by the consolidation phase, as described in Figure 7. Considering this thesis, induction and consolidation therapies are the most relevant parts of the treatment, because in Publication III, post-induction and post-consolidation MRDs were used to categorize patients into the treatment response groups.

In the NOPHO ALL2008 protocol, patients were treated with induction therapy for the first 29 days (Figure 10.) (Toft et al., 2018; ClinicalTrials.gov: NCT00819351). Prednisolone is mainly used as a glucocorticosteroid, but for patients with T-ALL or high WBC ($\geq 100 \times 10^9$/L) at diagnosis, prednisolone was replaced with dexamethasone (Toft et al., 2018). One month after therapy, treatment was administered according to the stratified risk classification system, which was determined by clinical factors (such as age, white blood cell count (WBC), cytogenetics, targeted genetic analyses, and treatment response evaluated by measuring the minimal residual disease (MRD). Roughly 50–60% of the patients were classified into the standard risk (SR) group, which included criteria such as B-lineage leukemia and post-induction MRD less than 0.1%. The intermediate risk (IR) group consisted of 30% of all patients, and the inclusion criteria for this group were intermediate response to induction therapy (end-of-induction (EOI) MRD $>10^{-3}$ and <5%) or specific cytogenetic aberrations such as iAMP21 or *TCF3::PBX1* translocation (Toft et al., 2018). *ETV6::RUNX1* patients were mainly treated

according to the SR or IR arms. Overall, approximately 15% of the patients were categorized into the high-risk (HR) risk group. Criteria determining HR are, for example, poorly responding T-ALL (EOI MRD > 0.1%), high WBC and/or poor response to induction therapy (EOI MRD >5%), or high-risk cytogenetics such as hypodiploidy or *KTM2A*-rearrangement (Quist-Paulsen et al., 2020).

Patients in the HR group received more intense therapy regimens that included a greater number of drugs than those in the SR/IR groups did. For HR patients, allogeneic hematopoietic stem cell transplantation was also an option in cases where the treatment response was very poor at the end of the induction or consolidation phases (Quist-Paulsen et al., 2020; Toft et al., 2018).

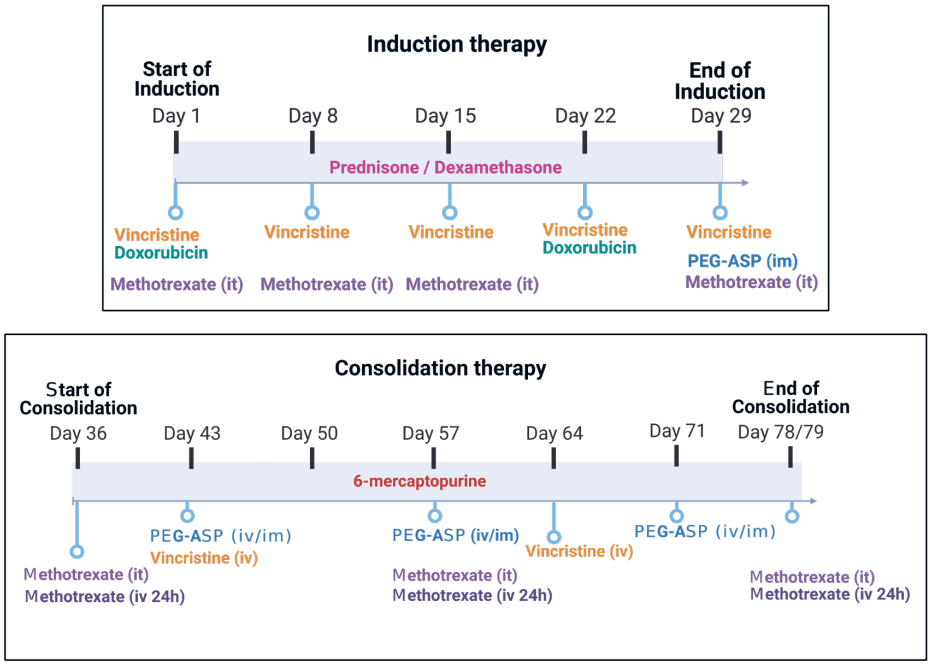


Figure 10. Overview of induction and consolidation therapies for SR and IR patients in the NOPHPO ALL2008 protocol. Abbreviations: im = intramuscular, intramuscular; IV, intravenous; PEG-ASP, pegylated asparaginase.

3 AIMS OF THE STUDY

Although the prognosis of ALL has improved significantly, the outcome for certain subgroups of patients with T-ALL and B-ALL, as well as for patients with relapsed disease, remain poor. The development of new therapies requires a better understanding of disease genetics and an improved classification. Here, we sought to identify novel prognostic biomarkers for T-ALL, which lacks such markers, and to deeply characterize the genomics of *ETV6::RUNX1*-positive leukemias to better understand the genetic factors that determine therapy response. We used the Hemap data set, which was compiled in our study consortia, to search for novel biomarkers in T-ALL. We investigated the genomics of the *ETV6::RUNX1* subtype using several patient cohorts that had either samples or sequencing data available along with detailed therapy response data.

The specific aims of this study were:

1.) to study the prognostic impact of *SIX6* expression in T-ALL and its role in disease development and maintenance (Study I);

2.) to investigate the role of high *PRMT7* expression in the development of T-ALL (Study II);

3.) to study the genomic determinants of therapy response in *ETV6::RUNX1* leukemia (Study III).

4 MATERIALS AND METHODS

More detailed information of the methods can be found in the online versions of the original publications, which are referred to by Roman numerals (I-III).

4.1 Cell culture (I-II)

Jurkat, and Jurkat Cas9 cells were cultured in RPMI media (Gibco, Thermo Scientific, Waltham, MA, USA), supplemented with 2 mM L-glutamine (Gibco), 1 % antibiotics (0.5 U/ml penicillin and 0.5 μg/ml streptomycin; Lonza, Basel, Switzerland), and 10 % FBS (Gibco) at the standard cell culture conditions, 37°C in 5% CO2 (referred as a standard cell culture condition), and split every 2–3 days. Molt-4 cells were cultured in RPMI media (Gibco), supplemented with 2 mM L-glutamine (Gibco), 1 % antibiotics (0.5 U/ml penicillin and 0.5 μg/ml streptomycin; Lonza, Basel, Switzerland), and 20 % FBS (Gibco) under the standard cell culture conditions, 37°C in 5% CO2, and split every 3 days. The Jurkat cell line was obtained from DSMZ (Braunschweig, Germany) and the Molt-4 cell lines were obtained from Synthego (Synthego Corporation, Menlo Park, CA, USA). As a courtesy, the Jurkat Cas9 cell line was obtained from the laboratory of Jan Cools (VIB-KU, Leuven, Belgium). All cell lines have been tested and proven regularly negative for mycoplasma. Jurkat and Jurkat Cas9 cell lines have been authenticated using STR genotyping services provided by Eurofins Genomics (Ebersger, Germany).

4.2 Patient samples for WGS and RNA-seq characterization (I-III)

Bone marrow and blood samples of 35 B-ALL patients were collected before the initiation of therapy and at remission when no blast cells were detectable by morphology and flow cytometry. Patients/caretakers were consented to the study before the sample collection. Patients were treated according to the NOPHO ALL2008 protocol. The study was approved by the regional ethical committee

(Pirkanmaa Hospital District Ethical Committee, ETL code R13109) and was conducted according to the principles of the Declaration of Helsinki.

4.3 CRISPR-Cas9 mediated genome editing (I-II)

CRISPR-Cas9 mediated knockout cell lines used in this study were generated either in-house, or a knockout bulk cell line pool was obtained from a commercial provider.

CRISPR guide RNAs against *PRMT7* and *SIX6* genes were designed using the Desktop genetics online tool (London, UK). Guide RNAs were cloned into a px321-GFP vector containing a GFP fluorochrome marker. The px321 vectors were courtesy of the Jan Cools laboratory (VIB-KU). Jurkat Cas9 cells, stably expressing Cas9 enzyme, were transfected with 2 μg of the plasmids using 4D Nucleofector electroporation equipment (Lonza) and electroporation program CL-120 with the SE solution. After 24 h, GFP-positive cells were single-cell sorted into 96-well plates in conditioned media, containing 20–40 % of media harvested from cultured Jurkat and Jurkat-Cas9 cells and 60–80 % RPMI 1640 media.

Molt-4 derived *PRMT7* knockout cell pool was obtained from Synthego (Synthego Corporation). The bulk cells were single-cell sorted into 96-well plates in conditioned media containing 40% media harvested from unmodified and modified Molt-4 cells and 60% RPMI-1640 media.

Single-cell clones were screened for mutations using T7 endonuclease (New England Biolabs, Ipswich, MA, USA) and confirmed by Sanger sequencing (DNA sequencing and genomics service, University of Helsinki, Helsinki, Finland). MutationTaster2 (Schwarz et al. 2014) was used to study the functional effects of the mutations and changes in amino acid sequences and to predict the disease-causing potential.

Silencing of the protein level of PRMT7 was confirmed using Immunohistochemistry staining and Western blot; for SIX6 only Western blot was used.

4.4 Characterization of genetically modified cell lines (I-II)

After genetic modifications, the cell lines were analyzed for viability, proliferation, cell cycle, apoptosis, and clonogenic capacity.

Viability was measured using AlamarBlue (Invitrogen, Carlsbad, CA, USA) or CellTiter-Glo (Promega, Madison, WI, USA) luminescent assay. For viability studies with AlamarBlue, a reagent was added at a ratio of 1:10 into cells (10, 000 cells in 100µl). The cells were then incubated for 2 h under standard cell culture conditions. CellTiter-Glo reagent was added 1:1 to the cells. Cells were incubated for 15 min on a shaker, and luminescence was measured using a Wallac 1420 Victor2 Microplate Reader (Perkin Elmer, Waltham, MA, USA) with a 500 ms integration time. Viability was measured at 0, 24, 48, and 72 hours. The results were normalized using an unmodified control cell line or DMSO-treated unmodified control cells to obtain relative viability.

For *SIX6* and *PRMT7* knockout cell proliferation studies, 100 000 cells per cell line were seeded into a 24-well plate in 1 ml of cell culture media, as described earlier. The cells were incubated for 0, 24, 48, and 72 h under standard cell culture conditions to perform a time-series study. The cell count was calculated using a Bürker chamber at the chosen time points. Control cells were used to normalize to results to obtain a relative proliferation rate.

The Click-iT™ EdU Alexa Fluor™ 647 Flow Cytometry Assay (Invitrogen) was used to measure the cell cycle and the number of cells in different cell cycle phases (G1, S, G2/M phases) of the *SIX6*, and *PRMT7* knockout cells compared to control cells. Assay was performed according to the manufacturer's protocol, shortly by adding 1µl of EdU to 1 x 10^6 cells in 1 ml, after 24 h of seeding the cells and incubating at 37°C for 1 hour. Apoptosis was measured using the Annexin V Apoptosis Detection Kit-APC (eBioscience, Thermo Fisher Scientific) according to the manufacturer's instructions. EdU and apoptosis assay results were recorded using BD FACS Canto II (BD Biosciences, San Jose, CA, USA). The results were analyzed using the FlowJo software (BD Biosciences).

The clonogenic capacity of *PRMT7*-knockout cells was studied using colony formation analysis. Methylcellulose-based MethoCult medium (H4230, Stemcell Technology) was prepared according to the manufacturer's protocol; the mixture was added to 6-well plates at a density of 1000 cells/well. The cells were incubated under standard culture conditions for 14 days. Colonies were imaged on days 10 or 14 using a Nikon AZ100 Fluorescence Macroscope (Nikon, Minato, Tokyo, Japan).

4.5 Protein expression studies (I-II)

4.5.1 Western Blot (I-II)

For protein extraction, cells were lysed with CelLytic M reagent (Sigma Aldrich) according to the manufacturer's protocol.

Protein samples were loaded on Mini-PROTEAN® TGX Stain-Free™ Precast 10% or 12% gels (Bio-Rad) and transferred to a 0.2μM PVDF membrane using Trans-Blot Turbo Transfer Pack and Trans-Blot Turbo Transfer System (Bio-Rad). Membranes were processed using standard procedures, with PRMT7 (1:1000 in 3% BSA, #14762, Cell Signaling Technology, Danvers, MA, USA), SIX6 (1:2000 in 5% BSA, #ab251558, Abcam), and Histone H3 (1:10,000 in 3% BSA, #9715S, Cell Signaling Technology) primary antibodies, and Horseradish peroxidase conjugated anti-rabbit (1:5000 in 3% BSA, #7074S, Cell Signaling Technology) secondary antibody. Amersham ECL Reagent (GE Healthcare) was used for chemiluminescence reaction, and chemiluminescence was detected with ChemiDocTM XRS+ using Image LabTM Software (Bio-Rad). A PageRuler Plus pre-stained protein ladder (Thermo Fisher Scientific) was used as a reference for protein size. ImageJ software was used to quantify protein expression (Schneider et al., 2012).

4.6 Monomethyl arginine analysis (II)

4.6.1 Mass spectrometry-based arginine monomethylation analysis

For protein extraction, $1x10^8$ Jurkat *PRMT7* EV, Jurkat *PRMT7* KO3, and KO5 cells were collected. Proteins were extracted using a urea lysis buffer (200 mM HEPES pH 8.0 (Sigma-Aldrich), 9 M urea (Thermo Fisher Scientific), 2.5 mM sodium orthovanadate (Sigma-Aldrich), 1 mM sodium pyrophosphate (Sigma-Aldrich), 1 mM β-glycerophosphate (Cayman Chemical, Ann Arbor, MI, USA), and lysates were sonicated (1 min, pulse 1 s on/off with 40 % amplitude, three times) using a Vibracell 500-Watt ultrasonic processor (VC 505) sonicator (3 mm microtip) (Sonics & Materials Inc., Newtown, CT, USA) to break apart DNA. Samples were processed using the PTMScan® Mono-Methyl Arginine Mo-tif [mme-RG] kit (Cell

Signaling Technology). Briefly, proteins were digested with proteases, and the resulting peptides were purified by reversed-phase extraction. The peptides then were subjected to immunoaffinity purification. Unbound peptides were removed by washing and the captured MMA-containing peptides were eluted with dilute acid. Enriched peptides were analyzed using a Q Exactive HF mass spectrometer software (Thermo Fisher Scientific). Label-free quantification was performed at the peptide isoform level. Abundance values were calculated based on the intensities of peptide precursor ions.

4.6.2 LC-MS/MS analysis

LC-ESI-MS/MS analyses were performed on a nanoflow HPLC system (Easy-nLC1200, Thermo Fisher Scientific) coupled to Q Exactive HF mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nano-electrospray ionization source. Peptides were first loaded onto a trapping column and subsequently separated on a 15 cm C18 column (75 μm × 15 cm, ReproSil-Pur 3 μm 120 Å C18-AQ, Dr. Maisch HPLC GmbH, Ammerbuch-Entringen, Germany). The mobile phase consisted of water with 0.1% formic acid (solvent A), and acetonitrile/water (80:20 (v/v)) with 0.1% formic acid (solvent B). Peptides were eluted with a 120 min gradient: from 2% to 39% solvent B in 105 min, from 39% to 100% solvent B in 10 min, followed by a wash for 5 min in 100% solvent B. MS data were acquired automatically using Thermo Xcalibur 4.1. (Thermo Fisher Scientific). An information-dependent acquisition method consisted of an Orbitrap MS survey scan with a mass range of m/z 300–1750, followed by HCD fragmentation for the most intense peptide ions.

Database searches were performed using Proteome Discoverer 2.4 software (Thermo Fisher Scientific) connected to in-house Mascot 2.7.0 search engine software (Matrix Science) against UniProt Swiss-Prot (version 2021_02) Homo sapiens protein database. The search criteria were as follows: trypsin as an enzyme; cysteine carbamidomethylation as fixed modification; arginine methylation, lysine methylation, methionine oxidation, and N-terminal acetylation as variable modifications; peptide mass tolerance 10ppm; and MS/MS ion tolerance 0.02 Da. Label-free quantification, validation, and filtering of the results were performed using the Proteome Discoverer 2.4 software.

4.7 Sequencing (I-III)

4.7.1 RNA sequencing (I-III)

Total RNAs was extracted from cell lines using the PureLink™ RNA Mini Kit with the On-Column PureLink® DNase Treatment Protocol (Ambion®, Life Technologies, Thermo Fisher Scientific). Library preparation and RNA sequencing of cell lines from unmodified Jurkat and Molt-4 cells and modified *SIX6* and *PRMT7* knockout lines were performed by Novogene (Cambridge, UK). The quality of the raw sequencing reads was ensured with FastQC (v0.11.8). The reads were mapped to the human reference genome version hg19 using the STAR aligner software (2.5.3a modified); reads aligned to more than two locations were discarded. The alignment file was turned into tag directories and read counts were calculated using the HOMER toolkit (v.4.11). Differential gene expression was analyzed by the quasi-likelihood F-test using the edgeR R package (Robinson 2019). Top genes from differential gene expression analysis were selected using threshold log fold change (logFC) $\geq \pm 2$ and false discovery rate (FDR)<0.01, for Jurkat *SIX6* WT and KO cells, and logFC $\geq \pm 1.5$, and FDR < 0.05, for Jurkat *PRMT7* EV and KO and Molt-4 WT and KO cells. Volcano plots were drawn using the R-package ggplot2 Volcano plot function, combining the three biological replicates, and using log2 fold-change values and FDR values. Heatmaps presenting all biological replicates were drawn using a row-scaled count matrix of the normalized per million mapped (RPKM) reads and the ComplexHeatmap R-package (Z. Gu et al., 2016).

RNA of *ETV6:RUNX1* (n=35) samples was extracted from diagnostic bone marrow or peripheral blood samples collected in PAXgene blood RNA tubes using the PAXgene Blood RNN kit (cat. #762174, Qiagen GmbH, Hilden, Germany) following the v2 instructions for manual purification or using or from isolated MNC cells with PureLink™ RNA Mini kit (Invitrogen) according to the manufacturer's protocol. RNA sequencing and library preparation were performed in three batches (Table 2). Sequencing of all samples were performed by Novogene, using Illumina Novaseq 6000 instrument (Illumina Inc.) Raw sequencing data were processed using nextflow-based (v21.02.0. edge) nf-core/rnaseq (v3.0) pipeline, including trimming of paired-end reads and alignment to GRCh38.103 with STAR (v2.6.1d). Quantification of aligned reads at the transcript level were made using Salmon (v1.4.0), and normalization used limma (v3.52.2) (Patro et al., 2017; Ritchie et al., 2015). Expression of the genes involved in the S and G2/M phases during the cell

cycle was studied as individual genes, but gene set scores were calculated from the normalized count matrix using GSVA ssGSEA (v1.48.0) (Hänzelmann et al., 2013). The regulons analyzed were chosen from our earlier publication by Mehtonen et al. 2020. The gene, gene set, and regulon expressions were drawn using the ComplexHeatmap R-package (Z. Gu et al., 2016). Correlation tests were performed using Spearman's correlation by comparing gene set expression or regulon activity and MRDs (mid-induction and EOI).

Table 2. RNA-sequencing sample preparation.

Batch	Pre-processing kit	Library preparation
C101HW18062568	Globin-Zero Gold rRNA Removal Kit	NEBNext Ultra Directional RNA Library Prep Kit
X201SC19081411-Z01	Ribo-Zero rRNA Removal kit	NEBNext Ultra Directional RNA Library Prep Kit
X204SC22010852-Z01	Ribo-Zero rRNA Removal kit	NEBNext Ultra Directional RNA Library Prep Kit

4.7.2 Singe cell RNA-seq (III)

BM or blood mononuclear cells collected at diagnosis and mid-induction (day 15) were analyzed from viable cryopreserved samples. Cells were thawed and counted, followed by blocking with Human TruStain FcX Blocking Solution (BioLegend, San Diego, CA, USA) 10 min at 4°C. Antibody pool (0.25 μg per million cells per specific antibody, except 0.125 μg for CD45 and CD8; 0.1 μg per hashtag antibody) was added to the cells and incubated for 30 min at 4°C. After staining, cells were washed, counted and sample pools were prepared, and samples were loaded onto Chromium lanes (10x Genomics). The 5' gene expression, ADT- and VDJ (BCR and TCR) libraries were prepared following the manufacturer's instructions (10x Genomics, Pleasanton, CA, USA). Libraries were index barcoded and sequenced using Illumina Novaseq (S1/S2).

Primary ALL scRNA-data from VDJ sequencing was aligned with Cell Ranger 6.0 version to human reference (hg20) with default parameters. Donor (and singlet/doublet) assignments were performed using DSB-normalized hashtag signals, and downstream analysis was performed using Seurat (Hao et al., 2021; Satija et al., 2015).

4.7.3 DNA sequencing (III)

DNA of *ETV6::RUNX1* samples (n=35) was extracted from primary ALL bone marrow or blood cells after Ficoll/Lymphoprep™ (GE Healthcare, IL, USA / Stemcell Technologies, Vancouver, Canada) gradient separation or PAXGene Blood RNA Tubes (cat. #762165, PreAnalytiX, Qiagen) using the QIAamp DNA Blood Midi kit (cat. #51183, Qiagen GmbH, Hilden, Germany), GeneJET Genomic DNA Purification kit (Thermo Fischer), or AllPrep DNA/RNA/miRNA Universal Kit (Qiagen).

Whole genome sequencing (III)

WGS library preparation and DNA sequencing were performed in five different batches (see details in Materials and Methods of Publication III). Shortly, libraries were prepared using the library preparation kits in Table 3. All samples were fragmented aiming at 350-bp fragments, DNA fragments were end-repaired, A-tailed, and ligated. Before sequencing, library quantification was performed. Paired-end sequencing (150 bp) was performed for all sample libraries: diagnostic samples were sequenced at the depth of 60x or 90x and remission samples at 30x. Sequencing was performed using the Hiseq X or Novaseq 6000 sequencing systems (Illumina Inc.).

Table 3. WGS sequencing library preparation.

Batch	Library preparation kit
1	TruSeq DNA PCR-free library preparation kit
2-4	NEBNext® Ultra™ DNA Library Prep Kit
5	NexSeq AmpFREE Low DNA Fragment Library kit

Targeted genome sequencing (III)

Library preparation for targeted sequencing was performed using KAPA HyperPlus (Roche, Basel, Switzerland) with the following modifications: fragmentation with 12.5 min incubation, xGen Duplex Seq adapters (3–4 nt unique molecular identifiers (UMI), 0.6 mM, Integrated DNA Technologies) were used for ligation, and xGen Indexing primers (2 mM, with unique dual indices, Integrated DNA Technologies)

were used for PCR amplification (5–13 cycles depending on the amount of DNA input). Target enrichment was performed in a multiplex fashion with a library amount of 187.5 ng (8-plex). The libraries were hybridized with a custom-developed 243 gene panel GMS Lymphoid panel v1 (Twist Bioscience) with the addition of Twist Universal Blockers and Blocking solution for 16 h. The panel design targeted the full protein-coding sequences of the 243 genes. Post-capture PCR was performed with xGen Library Amp Primer (0.5 mM, Integrated DNA Technologies) for 10 cycles. Quality control was performed using the Quant-iT dsDNA HS assay (Invitrogen) and TapeStation HS D1000 assay (Agilent). Sequencing was performed on NovaSeq 6000 (Illumina) using paired-end 150 nt readout, aiming at 40 M (tumor) or 20 M (normal) read pairs per sample. Demultiplexing was performed using Illumina bcl2fastq Conversion Software v2.20.

4.7.4 DNA methylation array and CNV detection (III)

Copy number variations (CNVs) can be detected using array-based DNA methylation data (Kilaru 2021). In this study, DNA methylation data was used to increase the number of samples in the CNV analysis, and genome-wide DNA methylation levels were determined using the Infinium HumMeth450K BeadChip assay (450k array, Illumina). Methylation data was analyzed in R version 4.1.2 using Bioconductor (v3.14) packages methylumi (v2.40.1) and conumee (v1.28.0). Background correction and normalization of using Illumina controls was made using methylumi. Conumee was used to detect copy number variations after fitting the data to a set of normal reference samples, for which 86 normal samples from the GEO data set GSE49031 were used. Script defaults were used for all processing. Conumee output includes genome plots and segmentation files containing statistically interesting intervals of the genome. In addition, the segments with minimal deviation from the baseline signal were included in the tool output, as regions with a median signal between -0.05 and 0.1 were accounted for most regions as copy number change events. Therefore, segments with a median signal of -0.05 or below, were considered deletions, and segments with a median signal of 0.1 or above were classified as amplifications. In RNA-seq expression data that was available for a subset of samples, we found that putatively deleted regions were enriched in mono- and nearly mono-allelic expression as would be expected, which supported our choice of thresholds. Additionally, events overlapping with Encode Blacklist regions (Amemiya et al., 2019) by 50% or more were excluded from further

analysis. Events larger than 1 Mb in size were considered CNVs, and events smaller than 1 Mb were classified as SVs in downstream analyses. Two samples, ALL_309 and ALL_977, were excluded from downstream analyses because of unsuccessful normalization.

4.7.5 Alignment, annotation, variant calling, and filtering (III)

Bioinformatics analysis of WGS was done using the BALSAMIC workflow (Foroughi-Asl et al., 2019)(v8.2.10, v8.1.0, and v9.0.1). The Sentieon was used for alignment to reference genome hg19 and Sentieon TNscope (Sentieon version 201911) for calling of SNVs and InDels. Calling of CNA was done using ascatNGS (Raine et al., 2016) and calling of SVs was done using a combination of Manta v1.6 (Illumina), Delly v0.8.7 (Rausch et al., 2012), and BRASS (https://github.com/cancerit/BRASS, refered 21.3.2023). Targeted sequencing analysis was also performed using the BALSAMIC workflow (v7.2.2), but alignment to the reference genome hg19 was performed using BWA MEM (v0.7.15) (Li and Durbin, 2009). Somatic mutations were called using VarDict (v2019.06.04) (Lai et al., 2016) in tumor/normal mode and were annotated using Ensembl VEP (v94.5).

The SNVs and InDels of WGS calls were intersected with the variants of targeted sequencing using Bedtools (Quinlan and Hall, 2010) to identify common variants. These variants were reviewed manually in Integrative Genomic Viewer (IGV) (Robinson et al., 2011) and classified into three categories: true positive, false positive somatic variant (evidence in the normal sample in the targeted sequencing data), and false positive variant with no evidence in the targeted panel. The TNscope parameters between these categories were then compared and used to determine further filtering criteria for the SNVs and indels. In addition to the default TNscope filter, the following criteria were used to select the variants for further downstream analysis: sum of AD values (allelic depths for the ref and alt alleles) ≥ 10 and $\geq$ AFDP (read depth to calculate allele fraction), alternative allele AD ≥ 3, and AF (allele fraction) ≥ 0.05 and < 1, QUAL ≥ 40 and/or PV2 (Fisher's exact test p-value for high-confidence reads) ≤ 0.05, SOR (symmetric Odds Ratio of 2x2 contingency table to detect strand bias) ≤ 3, and variant locating outside Encode Blacklist regions (Amemiya et al., 2019). Variants filtered using these criteria were called moderately filtered variants, including subclonal alterations.

In addition to the default Manta and Delly filters, the SVs were further filtered to include only variants with at least 5 reads supporting the alternative allele in the

tumor samples, and exclude the variants overlapping more than 50% with the Encode Blacklist region (Amemiya et al., 2019). Variants by size between 150 bp and 1Mb were classified as SVs. Collective SV list were constructed by combining the filtered results from both callers, with following criteria: deletions, insertions, and duplications with 80% or more overlap were included and translocations within less than 5000 bp from each other, prioritizing the breakpoints reported. AnnotSV (v3.1.1) was used to perform SV annotation, and further consequence annotation for the coding variants is presented in Table 4. (Geoffroy et al., 2018). Variants that did not fulfill the criteria listed in Table 4. were annotated as non-coding variants. Variants without gene annotations were categorized as intergenic variants, and the closest upstream and downstream genes were determined.

The CNVs detected using ASCAT were filtered by size, accounting for events larger than 1 Mb as CNVs. Events overlapping 50% or more with the Encode Blacklist regions were excluded. CNVs detected by the ASCAT and over 1 Mb deletions and duplications detected by Manta and Delly were combined to the one collective CNV tool: including the events which overlap 80%, prioritizing Manta breakpoint over Delly, and Delly over ASCAT. The merged CNV results were annotated using AnnotSV, and consequence filtering was performed using the same criteria as those for SV deletions and duplications (Table 4). CNVs reported with a copy number of 2:0 (total copy number:minor copy number) by ASCAT were classified as CNV-neutral loss-of-heterozygosity (LOH).

For detection of subclonality and CNVs the cgpBattenberg algorithm was used (https://github.com/cancerit/cgpBattenberg, v3.5.3). The CNVs were compared with ASCAT CNVs to confirm the detection accuracy.

Immunoglobulin gene rearrangements were analyzed by utilizing the IgCaller tool (v1.2.1) (Nadeu et al., 2020), providing tumor and normal bam files as input and setting the tumorPurity parameter based on the consensus estimates of blast percentage. The filtered output file from the tool was included in further analysis, and only events with a minimum of five split reads and/or abnormal insert size reads supporting the rearrangement were included.

Table 4. Criteria for consequence filtering of coding SVs.

Consequence	SV type	Criteria
Loss of function	Deletion	100% overlap with the coding region of annotated gene
Gain of function	Duplication	100% overlap with the coding region of annotated gene
Coding sequence variant	Deletion, duplication	1-99% overlap with the coding region
Coding sequence variant	Inversion, translocation	Breakends locating between the transcript start and end coordinates

4.7.6 Evaluation of CNV results to karyotype (III)

The CNV results from the WGS analysis were compared with the clinical karyotypes to evaluate the findings. The ASCAT CNV profiles were compared with the clinical karyotypes of each patient. Systematically, the most common alterations, such as derivative chromosome 21, amplification of chromosome 21, and deletion of p12, were checked to confirm that ASCAT reliably detected CNVs. The clinical karyotype was also revised using ASCAT CNVs and subclone CNVs detected by the cgpBattenberg algorithm.

4.7.7 Mutation Signature Analysis (III)

Mutational signature analysis was performed using MuSiCa-Mutational Signatures in Cancer (version 1.2, released September 24, 2020) (Díaz-Gay et al., 2018). For mutational signature analysis, only the strictly filtered SNVs were used. The expression of COSMIC signatures per patient were plotted using the ComplexHeatmap (Z. Gu et al., 2016). Correlation tests were performed using Spearman's correlation by comparing the signature contributions and MRDs (mid-induction and EOI).

4.8 *In vitro* drug treatments (I)

To study drug responses *in vitro*, 10 000 Jurkat *SIX6* wildtype and Jurkat *SIX6* KO6 cells were plated per 96-well plate. Cells were treated with a 2–5 fold dilution series of dexamethasone (0.1–5 μM; D8893; Sigma-Aldrich), prednisolone (2.5–10 μM; P6004; Sigma-Aldrich), vincristine (0.5–2 nM; V8879; Sigma-Aldrich), L-asparaginase (0.5–2 U/ml; A3809; Sigma-Aldrich), doxorubicin (5–20 nM; D1515; Sigma-Aldrich), dasatinib (10–10000 nM; 9052S; Cell Signaling Technology) or temsirolimus (4-fold serie; 1-4096nM; PZ0020; Sigma-Aldrich) and incubated for 72 hours at 37°C in 5% CO2. After incubation, cell viability was measured using AlamarBlue assay (Invitrogen) and Tecan fluorescence reader (Tecan, Männedorf, Switzerland).

4.9 Data sets (I-III)

In this study, multiple patient sample cohorts and data sets were generated, used, and reanalyzed. All used data sets, their descriptions and availability are summarized in Table 5.

Table 5. Summary of data sets produced and re-analyzed in studies I-III. In accession codes EGA refers to European Genome-Phenome Archive, GSE to NCBI Gene Expression Omnibus database and PXD ProteomeXchange Consortium

Patient sample data sets generated in this study				
Cohort	**Number of samples**	**Data type**	**Accession code / Availability**	**Publication**
GEPARD WGS	35	WGS	EGAD00001010164	III
GEPARD RNA	35	lncRNA-seq	EGAD00001010128, GSE228632	III
GEPARD panel	35	Targeted panel sequencing	NA	III
Patient sample data sets re-analyzed				
Cohort	**Number of samples**	**Data type**	**Availability**	**Publication**
DNA Methylation	143	DNA methylation levels of 435,941 CpG sites	GSE49031	III

Nordic RNA	19	lncRNA-seq	GSE227832	III
TARGET	264	RNA-seq and survival	https://portal.gdc.cancer.gov/proj ects (Liu et al., 2017; Roberts et al., 2014)	I, II
Hemap	9544	microarray	http://hemap.uta.fi/ (Mehtonen et al., 2019; Pölönen et al., 2019)	I, II
sc-Hemap	20,753 healthy B cells from BM	scRNA-seq 20 753 healthy B cells in	https://bioinformatics.uef.fi/cells/public/ HCA_Blineage/?ds=HCA_Blineage	III
scRNA-seq+VDJ-seq	4	CITE-seq (scRNA-seq) and scVDJ-seq from *ETV6::RUNX1* samples	GSE230295	III
Nordic	30	RNA-seq and survival	See Haider et al. 2019	I, II
Japanese	120	RNA-seq and survival	See Seki et al. 2017	I, II
Human thymus	1	Sc-RNAseq	https://developmentcellatlas.ncl.a c.uk (Park et al., 2020)	I, II
Human cell line data sets generated in this study				
Cohort	**Number of replicates**	**Data type**	**Availability**	**Publication**
Jurkat	3	RNA-seq	GSE186238	I, II
Molt-4	3	RNA-seq	GSE186238	I
Jurkat PRMT7-KO	3x4	RNA-seq	GSE186238	I
Molt-4 PRMT7-KO	3x2	RNA-seq	GSE186238	I
Jurkat SIX6-KO	3x2	RNA-seq	GSE148658	II
Jurkat, Molt-4, CCRF-CEM, P12-Ichikawa	1	GRO-seq	NA	II
Jurkat, Jurkat PRMT7-KO	3x3	MS-based arginine mono-methylation	PXD032819	I
Human cell line data sets re-analyzed				
Cohort	**Number of replicates**	**Data type**	**Availability**	**Publication**
Jurkat	1	H3K27ac ChIP-seq	GSE68976 (Hnisz et al., 2016)	II

Molt-4, CCRF-CEM, P12-Ichikawa	1	H3K27ac ChIP-seq	GSE76783 (Abraham et al., 2017)	II
Jurkat	1	H3K4me1 ChIP-seq	GSE97514 (Leong et al., 2017)	II
Jurkat	2	H3K4me3 ChIP-seq	GSE60104 (Orlando et al., 2014)	II
Jurkat	1	TAL1 & GATA3 ChIP-seq	GSE29180 (Sanda et al., 2012)	II
Jurkat	1	Hi-C	GSE122958 (Lucic et al., 2019)	II
Jurkat shRNA	4	RNA-seq	GSE29179 (Sanda et al., 2012)	II

4.10 Survival analysis (I-II)

The gene expression of *PRMT7* and *SIX6* was explored in the context of survival in pediatric T-ALL patients in two different cohorts (TARGET, n = 264; Japanese leukemia group, n = 119) (Liu Y et al., 2017; Roberts et al., 2014; Seki et al., 2017; Sood et al., 2017). For survival analysis associated to *SIX6* expression also the third Nordic cohort was used (Haider et al., 2019). The Kaplan-Meier method was used to estimate survival, and group differences were tested using the log-rank method. The events included induction failure, disease progression, relapse, secondary malignant disease, and death by any cause. The Cox proportional hazards model was used to examine the association between the individual risk factors and patient survival. The proportionality assumption was tested with Schoenfeld residuals. The R-packages *survminer* and *survival* were used for data visualization and ordered expression graphs were plotted. The R (v. 3.6.2) software environment (R Foundation for Statistical Computing, Vienna, Austria) was used for statistical analyses.

4.11 Statistical tests

All in vitro studies had three or more biological replicates. Dose-response, growth curves, cell cycle apoptosis, CFU, and RT-qPCR results are presented as the mean and standard error of the mean or all data points with median values. Statistical tests were performed mostly using Student's two-tailed t-test unless otherwise mentioned

in the text. A p-value < 0.05 was considered statistically significant. The Wilcoxon test was used to compare the statistical significance of the differences in gene expression between different leukemia types. In mass spectrometry-based arginine monomethylation analysis, statistical testing was performed with a two-tailed Student's t-test, and p-values were adjusted using the Benjamini-Hochberg procedure to account for multiple testing. Unsupervised clustering was performed using Pearson's correlation as the distance metric and Ward's linkage method. The statistical tests used in Publication III are described in the text for each method.

The R (v4.0.2, v4.3.0) software environment (R Foundation for Statistical Computing, Vienna, Austria), Microsoft Excel (Redmond, WA, USA), and GraphPad Prism 8 software (GraphPad, San Diego, CA, USA) were used for statistical analyses unless otherwise mentioned in the text.

5 RESULTS

5.1 High expression of TAL1-regulated *SIX6* as a marker for inferior outcome in T-ALL (I)

5.1.1 High expression of *SIX6* in the TAL1 subgroup of T-ALL

Using the Hemap data set, we identified *SIX6* as a highly expressed gene in T-ALL. The expression of *SIX6* was approximately 1.3-fold higher in T-ALL compared to healthy T-lymphocytes (p-value < 0.01) (Figure 11). In different T-ALL subtypes, *SIX6* expression was especially associated with the *TAL1* and *TAL1/LMO1* subtypes, indicating that *SIX6* is expressed in more mature T-ALL, as *TAL1* subtypes arise from differentiating cortical T cells (Liu Y et al., 2017).

SIX6 expression was not detected in any of the healthy DN, DP or SP stage cells in the scRNA-seq data (see study I Figure 1e), suggesting that strong expression of *SIX6* in T-ALL is aberrant and associated with the leukemogenic process.

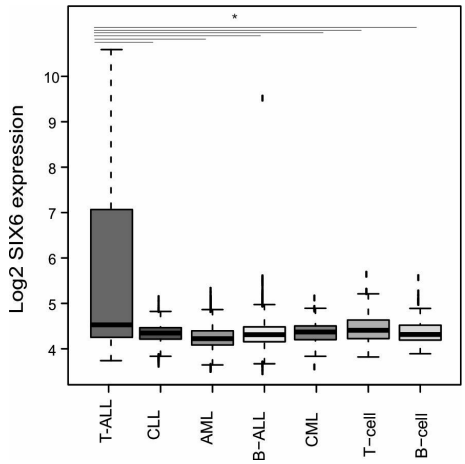


Figure 11. Expression of *SIX6* in T-ALL compared to other hematological malignancies and healthy T and B cells. The asterisk indicates a p-value < 0.01. Number of samples in each sample group: AML 1,713, B cells 75, CLL 801, CML 215, B-ALL 1,304, T-ALL 385, and T cells 247. Adapted from Laukkanen et al., 2020 *Leukemia & Lymphoma*.

5.1.2 *SIX6* is upregulated by TAL1 and GATA3

Expression of *SIX6* in the *TAL1* subtype suggests a possible regulatory relationship between *SIX6* and *TAL1*. To explore this, we utilized gene silencing data from Sanda et al. (2017). *SIX6* downregulation was triggered by the silencing of either *TAL1* or *GATA3* in Jurkat cells (see Study I, Figure 1f). However, chromatin immunoprecipitation sequencing (ChiP-seq) data showed that neither TAL1 nor GATA3 bound to the proximal promoter region of the *SIX6* gene, suggesting that the regulation of *SIX6* may be caused by longer-range enhancer elements. Accordingly, analysis of nascent RNA expression by Global Run-on Sequencing (GRO-seq) showed an active enhancer element 165 kb upstream from the *SIX6* gene. Three cell lines with high *SIX6* expression showed co-expression of *SIX6* and the enhancer RNA (eRNA), while cell lines lacking *SIX6* expression did not express the studied eRNA. Interestingly, ChiP-seq data revealed that TAL1 and *GATA3* TFs were also bound to this enhancer locus (Figure 12).

The activity of the enhancer in *SIX6*-expressed cell lines was confirmed by using histone modification ChiP-seq data, where acetylation of histone 3 lysine 27 (H3K27ac), monomethylation of histone 3 lysine 4 (H3K4me1), and trimethylation of histone 3 lysine 4 (H4K4me3) all pointed toward transcriptional activity of the locus. Finally, high-throughput chromosome conformation capture (Hi-C) data (Lucic et al., 2019) demonstrated physical looping and interaction between the distal enhancer element and the promoter of *SIX6*.

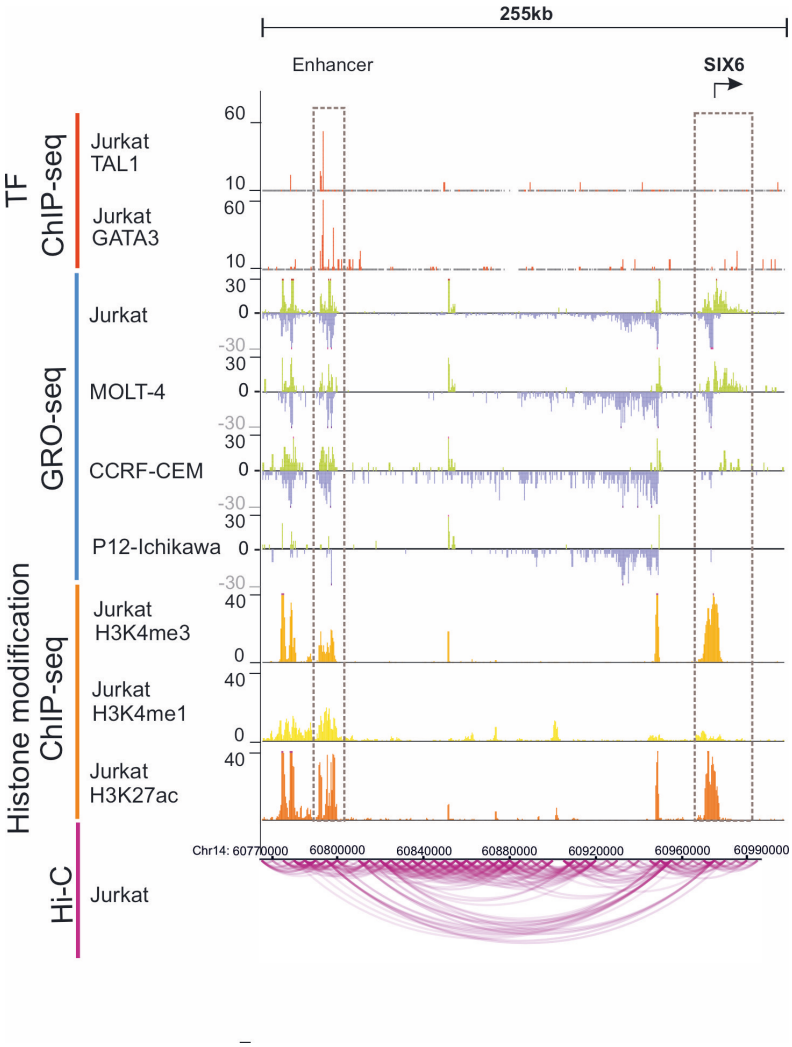


Figure 12. Regulation of *SIX6* expression. ChIP-seq (red and orange), GRO-seq (green and blue), and Hi-C (purple) analyses provide valuable insights into the activity and interaction of enhancers with the *SIX6* promoter. ChIP-seq data reveals that the TAL1 and GATA3 TFs bind close proximity to the *SIX6* gene. ChIP-seq data for histone modifications indicates enhancer activity, and GRO-seq results demonstrate the expression levels of both the enhancer and the *SIX6* gene. Hi-C data, shows long-range interactions within the depicted chromatin region. The y-axis of the sequencing data tracks represents the normalized read density. The dashed boxes highlight both the SIX6 gene locus and the distal enhancer region, drawing attention to their respective roles in gene regulation. Adapted from Laukkanen et al. 2020, Leukemia & Lymphoma.

5.1.3 High expression of *SIX6* defines a subgroup of T-ALL with an inferior outcome

The association of *SIX6* expression with patient survival was explored in three independent patient cohorts. In addition to the TARGET and Japanese cohorts, we obtained data from a Nordic patient cohort. Based on the distribution of *SIX6* expression data in the TARGET data set, patients were categorized into three groups: *SIX6*^{high}, *SIX6*^{medium}, and *SIX6*^{low}. *SIX6*^{high} cases had a 5-y EFS of 80% (95% CI, 70%–92%), *SIX6*^{medium} cases were 90% (95% CI, 84%–96%), and *SIX6*^{low} cases were 94% (95% CI, 89%–99%), indicating that the level of *SIX6* expression was associated with patient outcomes (see Study I, Figure 3). The difference between the *SIX6*^{low} and *SIX6*^{high} groups reached statistical significance (log-rank test, p-value = 0.011). In the Cox proportional hazard model, high expression of *SIX6* (*SIX6*^{high}) was associated with inferior outcomes in multivariate (HR 3.46, 95% CI 1.10–10.86, p-value = 0.03) and univariate models (HR 3.15, 95% CI, 1.22–8.14, p-value = 0.02) (see Study I, Table 1.). Interestingly, high *SIX6* messenger RNA (mRNA) had a higher hazard ratio than EOI MRD (HR 1.66, 95% CI 0.8–3.47, p-value = 0.18). Notably, the TARGET data set does not include cases with more than 5% residual disease at the EOI, thus excluding cases with the highest risk of therapy failure.

Patients from the Japanese and Nordic cohorts were categorized into *SIX6*^{high} and *SIX6*^{low} groups, respectively (Haider et al., 2019; Seki et al., 2017). In agreement with the TARGET data, higher than median *SIX6* mRNA level associated with inferior EFS as well as lower OS in both cohorts. However, the results did not reach statistical significance (see Study I, Figure 3b-c). Despite the direct regulation of *SIX6* by *TAL1*, *SIX6* remained an independent risk factor, even in the presence of a *SIX6*/*TAL1* interaction term.

5.1.4 SIX6 does not have oncogenic driver potential in T-ALL

To study the specific role of *SIX6* in T-ALL pathogenesis, the *SIX6* gene was deleted in the *TAL1*-positive Jurkat-Cas9 cell line using CRISPR-Cas9 technology. In RNA-seq analysis, 113 significantly up- or downregulated genes were found in *SIX6* knockout (KO) cells compared to wild-type Jurkat cells (log2FC $\geq \pm 2$, adjusted p-value < 0.01). Two well-known T-ALL-associated genes, *ILR7* and *JAG1,* were downregulated in the *SIX6* KO cells compared to the wild-type cells (see Study I, Figure 4b). Even though *JAG1* is an activating ligand of NOTCH1, we did not see any differences in NOTHC1 cleavage between *SIX6* KO and wild-type cells.

Next, the affected cellular pathways were explored by applying gene set enrichment analysis (GSEA). Enrichment was noted in mTOR-, KRAS-, and TNFα-related molecular signatures (see Study I, Figures 4E and S6A–B). Several genes with altered expressions belonging to the mTOR pathway were found, but they did not contribute to altered sensitivity to the mTOR inhibitor temsirolimus in *SIX6* KO cells compared to wild-type cells (see Study I, S5b).

The role of *SIX6* expression in the maintenance of tumor growth and chemotherapeutic resistance was investigated in tissue cultures and in a zebrafish model of T-ALL. Ablation of *SIX6* did not significantly affect the cell cycle, apoptosis, cell viability, or drug response when treated with dexamethasone, prednisolone, vincristine, l-asparaginase, doxorubicin, or dasatinib in tissue cultures. T-ALL maintenance and progression related to *SIX6* expression were probed in the T-ALL zebrafish model (Rag2:mMYC), in which endogenous zebrafish *SIX6* gene expression was low. Zebrafish were coinjected with Rag2:mMYC and Rag2:h*SIX6*-TdTomato constructs and compared to Rag2:mMYC-TdTomato-injected animals by analyzing the spread of a fluorochrome marker. All animals developed T lineage leukemia, but no statistically significant differences were noted in latency between the groups (see Study I, Figures 4H and S7B).

5.2 PRMT7 regulates RUNX1 target gene expression in T-ALL (II)

5.2.1 *PRMT7* is highly expressed in mature T-ALL subgroups

Hemap is a microarray based data set of global gene expression data of various hematologic malignancies, healthy tissues, and cell lines, consisting of 9,544 expression profiles, including 4,430 leukemias (such as 385 T-ALL, 1,304 pre-B-ALL, 1,713 AML, and 801 CLL, 1,306 lymphomas (208 T cell lymphoma (TCL) and 743 diffuse large B cell lymphoma (DLBCL)), and 428 healthy samples (such as 247 T-lymphocytes and 75 B-lymphocytes) (Mehtonen J et al., 2019; Pölönen P et al., 2019; Hemap resource: http://hemap.uta.fi/hemap/index.html). Hemap includes gene expression profiles collected from multiple previously published studies from the NCBI GEO database, and malignant samples broadly represent hematological disease types. This resource allows for the easy comparison of molecular phenotypes across hematologic malignancies and healthy cells.

Using the Hemap database, we sought to find genes that showed aberrant and disease-specific expression patterns in ALL. To this end, we identified *PRMT7* as highly expressed in T-ALL compared to other hematological malignancies or healthy cells (Figure 13a). Expression was approximately 2.6-fold higher in T-ALL compared to healthy T-lymphocytes (p-value < 0.01). When T-ALL cases were classified based on their expression profiles (Liu Y et al., 2017), *PRMT7* showed the strongest expression in mature subtypes of T-ALL (*TLX3, TLX1,* and *NKX2-1*) compared to less differentiated subtypes such as *LYL1/LMO2,* and *HOXA* (see Study II, Figure 1c).

We utilized a publicly available scRNA-seq data set to study whether the expression of *PRMT7* was evident during the development and maturation of healthy T cells (Park et al., 2020). Only weak expression of *PRMT7* was observed at later stages of differentiation or maturation of T cells (see Study II, Figures 1e), suggesting that strong expression of *PRMT7* is associated with the leukemogenic process.

Messenger RNA expression does not always reflect protein expression (Latonen et al., 2018). Therefore, we examined the expression of the PRMT7 protein in 84 diagnostic bone marrow biopsies that included 6 T-ALL and 78 B-ALL cases. Strong staining of the PRMT7 protein was seen in two of the six T-ALL cases, with 49% and 57% of the leukemic blasts positive, respectively, while two cases showed weak expression, and two stained negative (see Study II, Figures 1d and S1b). Among the B-ALL cases, strong PRMT7 staining was observed in 7 cases (range 21%–36%), weak staining was observed in 57 cases (range 1%–20%), and 10 of the cases were PRMT7 negative.

5.2.2 High *PRMT7* expression suggests trend towards inferior outcome

Two independent patient cohorts were used to study the association of *PRMT7* mRNA levels with patient survival. The patients were categorized into two groups based on their median *PRMT7* expression values ($PRMT7^{high}$ and $PRMT7^{low}$). In the TARGET data set, which includes gene expression and clinical data from 264 T-ALL patients (Liu Y et al., 2017; Roberts et al., 2014), $PRMT7^{high}$ cases had a 5-year EFS of 86% (95% CI, 81%–93%), while $PRMT7^{low}$ cases had an EFS of 92% (95% CI, 88%–97%) (p = 0.098) (Figure 13b). To study the prognostic effect of *PRMT7* expression in conjunction with known risk factors (e.g., age, white blood cell count, and MRD), the Cox proportional hazard model was applied. $PRMT7^{high}$ showed a

(statistically non-significant) trend toward an increased hazard for an event in both the univariate HR = 1.81, 95% CI, 0.89–3.71, p = 0.1) and multivariate models (HR = 1.77, 95% CI, 0.85–3.68, p = 0.13) (see Study II, Table S1). A Japanese patient cohort included expression and survival data for 119 T-ALL cases (Seki et al. 2017). *PRMT7*^{high} cases had worse EFS (Figure 8b) and an increased hazard for an event (HR = 1.41, 95% CI 0.76–2.6) (p = 0.272), although statistical significance was not reached.

5.2.3 *PRMT7* knockdown deregulates the expression of RUNX1 target genes

To study the functional role of *PRMT7* in T-ALL, we performed a genetic knockout of *PRMT7* in two T-ALL cell lines, Jurkat and Molt-4, using CRISPR-Cas9 technology. Guide RNAs targeting the first exon of the *PRMT7* gene led to small indels and deletions, which introduced an early stop codon to the coding sequence of *PRMT7*. In total, six single-cell sorted *PRMT7* KO cell lines were generated, with each harboring a different mutation (see Study II, Figure S2). Ablation of PRMT7 protein expression was confirmed by IHC staining and western blotting. The KO of *PRMT7* significantly decreased colony formation capability when compared to the unmodified cell lines (Figure 13c–d). In the cell cycle, apoptosis and cell viability assays showed no consistent changes in either direction, suggesting that *PRMT7* alone is not critical for these cellular processes.

As PRMT7 methylates arginine residues in histone and non-histone proteins in a monomethyl manner, we investigated arginine monomethylation patterns in KO and unmodified Jurkat cells. Global alterations in arginine monomethylation patterns in protein complexes were examined by utilizing the Comprehensive Resource of Mammalian protein complexes (CORUM) database (Giurgiu et al., 2019). Out of 243 protein complexes, the average peptide methylation was statistically significantly changed in 37 (Benjamini–Hochberg adjusted p-value < 0.05). Most of the differentially regulated peptides belonged to protein complexes affecting RNA and DNA processing (see Study II, Figure 4b). Protein complexes affecting RNA binding had increased average arginine monomethylation compared to the control cells. Complexes that play a role in DNA processing had mostly decreased average arginine monomethylation levels. We also found decreased average arginine monomethylation levels in two complexes that play an essential role in tumor development (H2AX complex II and PIDDsome).

To complement the methylation analysis, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Tenenbaum & Maintainer, 2021; version 1.35.0 Release 3.14) was used to explore the pathways associated with *PRMT7* alterations. The average peptide methylation level was significantly altered in 12 pathways (Benjamini–Hochberg adjusted p-value < 0.05) in the *PRMT7* KO cells. This analysis also showed increased methylation levels in pathways associated with RNA processing (see Study II, Figure 4c).

Finally, we studied changes in arginine monomethylation patterns in proteins that have a well-established role in the pathogenesis of T-ALL (Hnisz et al., 2016). Changes in arginine monomethylation levels were observed in peptides belonging to LEF1, LYL1, BCL11B, MYB, and NRAS. Moreover, KO of *PRMT7* affected the arginine monomethylation of several RUNX1 peptides (see Study II, Figure 4d). Interestingly, RNA-seq and Enrichr analysis revealed that five different RUNX1 target gene sets were statistically significantly altered in *PRMT7* KO cells when compared to the control cells ($\log2FC \geq \pm 1.5$, adjusted p-value < 0.05). *PRMT7* KO also led to the downregulation of many genes associated with the pathogenesis of T-ALL (*ILR7, CCDN2*), T cell differentiation (*DNTT*), tumorigenesis (*BCL11A*), and oncogenic signaling (*NEDD9*) (see Study II, Figure 5).

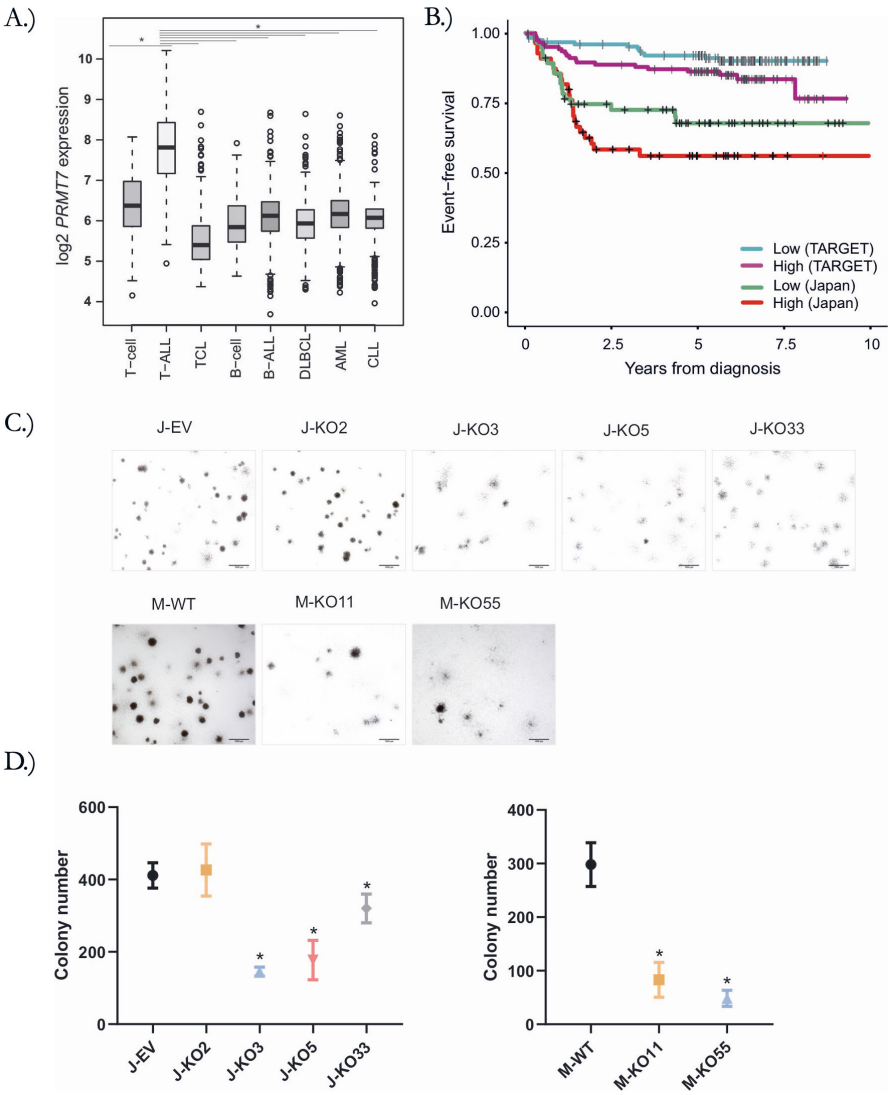


Figure 13. *PRMT7* expression in T-ALL. (A) *PRMT7* expression is significantly higher in T-ALL than in other hematological cancers and healthy T and B cells. Number of samples in each group: AML 1,713, B cells 75, CLL 801, CML 215, B-ALL 1,304, DLBCL 743, T-ALL 385, TCL 208, and T cells 247. An asterisk indicates a p-value < 0.01. (B) Association of *PRMT7* with patient outcomes in the TARGET and Japanese patient cohorts. (C and D) Effects of the *PRMT7* knockout on colony formation capability of unmodified (J-EV and M-WT) and knockout cell lines (an asterisk indicates a p-value < 0.05). Jurkat KO cells marked as "J-KO", and Molt-4 KO cell lines marked as "M-KO", numbers indicate the different KO cell lines. Adapted from Oksa et al. 2022, *Cancers*.

5.3 Genomic determinants of treatment response in *ETV6::RUNX1* positive B-ALL (III)

5.3.1 12p deletion associates with the faster treatment response

Although the prognosis of the *ETV6::RUNX1* subtype is generally good, a subset of patients had an inadequate early therapy response associated with a significantly increased risk of relapse. To date, no consistent genomic biomarkers have been identified that predict the therapy response in this subtype. We hypothesized that genetic differences in leukemic cells might contribute to drug sensitivity, and used multiomic data from 175 children with newly diagnosed *ETV6::RUNX1* leukemias treated according to the NOPHO ALL2000 and ALL2008 protocols to identify genomic features that predict the therapy response in this leukemia subtype (see Study III, Figure 1b). Our multiomic data comprised diagnostic and remission DNA analyzed using array-based copy number data (n = 142) or WGS data (n = 33). Diagnostic bulk RNA-seq (n = 51) data was complemented with single-cell multiomics (RNA, surface protein, and VDJ-seq) data from leukemic bone marrow cells upon chemotherapy (n = 2). The patients were categorized into three groups based on MRD at EOI therapy: fast responders with negative MRD, intermediate responders with positive but below 0.1% MRD, and slow responders with equal to or above 0.1% MRD. Response data was complemented with mid-induction (day 15) and end-of-consolidation (EOC) data, when available.

The number of SNVs, InDels, SVs, and CNVs did not differ by responder group (see Study III, Figure S1 a–d). Likewise, the distribution of breakpoints in the *ETV6::RUNX1* translocation locus was similar across the response groups (see Study III, Figure 1c). In contrast, copy number analysis revealed a more frequent deletion of p12 in the fast and intermediate response groups than in the slow response group when classified according to EOI and EOC MRD (Figure 14). This locus includes genes such as *KRAS* and *USP5,* which may play roles in the cell cycle, proliferation, and drug response. Previously, the deletion of the *TBL1XR1* gene was linked to poor outcomes (Brady et al., 2022). We found that chr3 deletions were more frequent among slow responders than in fast responders, as determined by EOC classification (see Study III, Figure S5h).

Our analysis also revealed that slow responders had more frequent amplification of the first 12 Mb of chr12 (12p), corresponding to a der21(t12;21) gain (see Study III, Figures 4b and S5e–f). Amplification includes genes such as *ETV6*, *TIGAR*,

SLC2A3, and *USP5*, which showed elevated RNA expression in the MRD-positive samples. In earlier studies, these genes were associated with tumor progression, cancer cell growth, and chemoresistance (Liu et al., 2019; Ma et al., 2018; Ziegler et al., 2020). Similarly, a higher frequency of partial amplification of chr (der21) was also observed in slow responders (see Study III, Figure S5g).

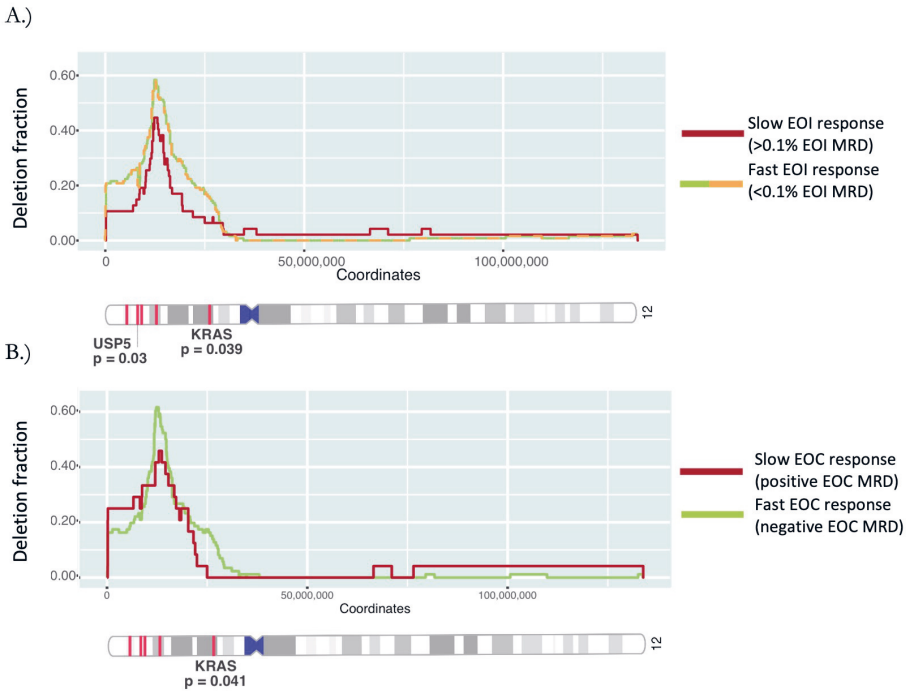


Figure 14. 12p deletion fractions by treatment response. A.) Represents 12p deletion fractions by EOI treatment stratification and B.) by EOC, respectively. The x-axis represents the coordinates of chr12, and the y-axis represents the 12p deletion fraction in each responder group. Relevant genes in 12p are indicated on the chr12 illustration. Red curves represent slow responders; dashed green–orange curves represent fast responders by EOI classification; and green curves represent fast responders by EOC classification. Adapted from Study III.

5.3.2 *ETV6::RUNX1* diagnostic cells differ in light chain recombination status

In healthy cells, RAG-mediated V(D)J recombination leads to antigen receptor gene rearrangements during early lymphocyte development (Gellert, 2002). First, at pro-B cell stage, rearrangement of the immunoglobulin heavy chain (IGH) occurs, followed by light chain rearrangements at pre-B cell stage (Herzog et al., 2009; Schatz and Swanson, 2011). Usually, diagnostic *ETV6::RUNX1*-positive cells are thought to be arrested at the pro-B cell stage based on their immunophenotype and transcriptome profiles (Mehtonen et al., 2020). Aberrant RAG recombinase activity is observed in *ETV6::RUNX1* leukemias, and it is known to drive secondary genomic alterations (Papaemmanuil et al., 2014). Strikingly, analysis of *IG* gene rearrangements revealed that *ETV6::RUNX1* patients with a slow treatment response had more *IG kappa (IGK)* gene rearrangements than fast responders (p-value = 0.03) (see Study III, Figure 3a). To study this further, we examined the V(D)J repertoire at the single-cell level and performed scRNA-seq coupled with scVDJ-seq at diagnosis, day 2, and mid-induction. The case selected for this analysis, slow responder (*IGK*+), had a typical pro-B-like expression status at diagnosis, and only the *IGHM* gene was expressed (Figure 15a). Upon chemotherapy, light chain expression was highly elevated in the remaining blast cells at mid-induction (Figure 12a and see Study III, Figures 3b–c). At this point, the blast status was confirmed by comparing the *IGH* sequences obtained from the WGS to scVDJ-seq data (see Study III Figure S4a). Surprisingly, in the slow-responding case, the productive rearrangements were enriched upon the treatment (Figure 15b, see Study III Figure 3b and d). In healthy cells, this phenotype is acquired only upon successful heavy and light chain recombination in precursor cells, suggesting that the more frequent IG light chain rearrangements in slow-responding *ETV6::RUNX1* blast cells may indicate acquisition of survival advantage through active BCR signaling.

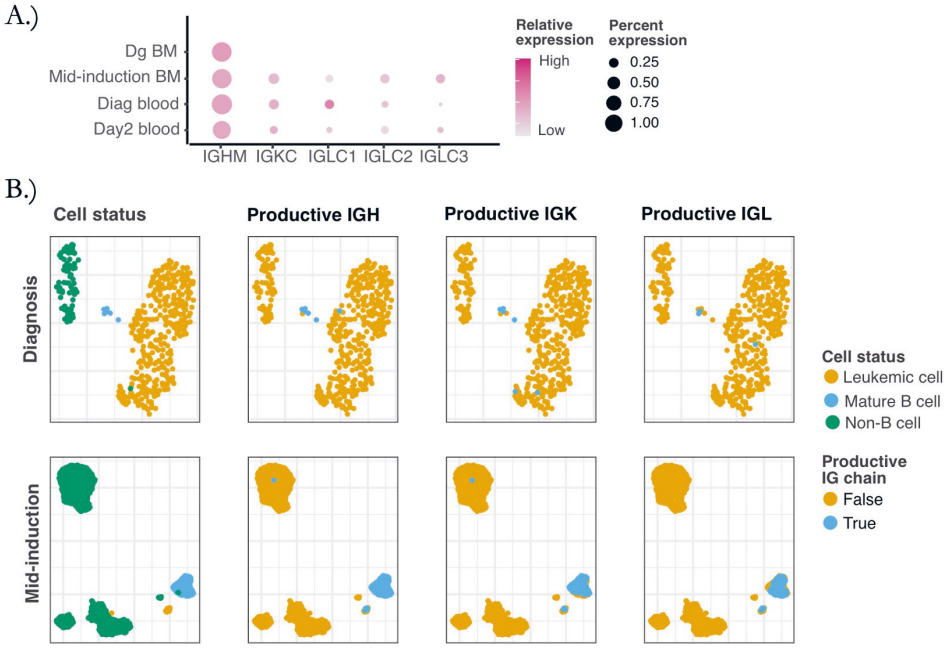


Figure 15. Immunoglobulin heavy- and light chain expression upon the treatment. A.) Dot plot heatmap presenting expression status of *IG* chain genes at diagnosis, day 2 and at mid-induction. B.) Cell status and expression of *IGH*, *IGK*, and *IGL* chains in scRNA-seq with integrated V(D)J sequencing data presented as UMAPs at diagnosis and mid-induction. Adapted from Study III.

5.3.3 APOBEC mutational signature is more frequent in fast responders

Mutational signature status is widely used in different cancers to study the mutagenic process that leads to DNA damage in the cancer genome (Koh et al., 2021). In our study, mutational signature analyses revealed significant COSMIC Signature 2 (APOBEC) enrichment in the *ETV6::RUNX1* samples, similar to a recent report (Brady et al., 2022). In addition, we found a significant negative correlation between the APOBEC signature and MRD measured at mid-induction; thus, patients with the APOBEC signature more frequently had fast early treatment responses (see Study III, Figure 2b). Accordingly, Signature 1 (a clock-like cell-intrinsic process) and Signature 8 (unknown) were found more frequently from slow responders (mid-induction time point), whereas slow responders at EOI exhibited Signature 3 more frequently, which stems from the failure of the DNA double-strand break repair

(DSBR) system, and may be related to high RAG activity in this subtype (Alexandrov et al., 2015, 2013) (see Study III, Figure 2a–b).

The overexpression of *APOBEC3A* (*A3A*) and *APOBEC3B* (*A3B*) genes is known to correlate with the genes associated with the cell cycle and DNA damage response (DDR), and *A3B* expression has been reported to be enriched in the G2/M phase of the cell cycle (Ng et al., 2019). To study this more closely, we analyzed the expression of the APOBEC family genes in the scRNA-seq data of healthy bone marrow cells and observed that *A3B* is expressed primarily on cycling HSPC, pro- and pre-B cells. Gene set scores from genes involved in the S and G2/M cell cycle phases, as well as the expression of the S phase-specific gene *PCNA*, the G2/M-phase specific gene *TOP2A*, and the early B cell marker gene *DNTT*, confirmed that the majority of the *A3B* expressing cells were early B cells that were actively cycling (see Study III, Figure 2c). Therefore, as actively cycling cells are likelier to be exposed to APOBEC enzymes, they may accumulate more APOBEC-related mutations.

5.3.4 *ETV6::RUNX1* cells differ in their cell cycle activity

Regarding the APOBEC signature, *A3B* expression, and cell cycle, we used bulk RNA-seq data (n = 51) to study this association further. We found that the gene set score from the S phase genes had a significant negative correlation to mid-induction ($r = -0.30$, $p < 0.05$) and EOI MRD ($r = -0.28$, $p < 0.05$), so that patients with a faster treatment response had more leukemic cells in the S phase (Figure 16.). A trend toward a negative correlation was seen among genes involved in the G2/M phase ($r = -0.21$, p-value = 0.15).

A regulon comprises a group of target genes that are regulated by a specific transcription factor. Regulons have distinct biological functions, which are determined by the genes that are regulated by the TF (Xia, 2020). Therefore, we studied E2F regulon activities and found that at EOI, fast responders had higher activity of E2F3 and E2F8 regulons (Figure 16). The E2F family of TFs regulates various genes involved in the cell cycle and apoptosis (Dyson, 1998), and E2F3 and E2F8 especially play a role in controlling the G1/S transition (Cuitiño et al., 2019; Leone et al., 1998). Our results support the observation from our previous scRNA-seq study (Mehtonen et al., 2020), where the diagnostic *ETV6::RUNX1* cells differed in their cell cycle activity, so that the faster responding cases had a higher proportion of cells in the active phase of the cell cycle.

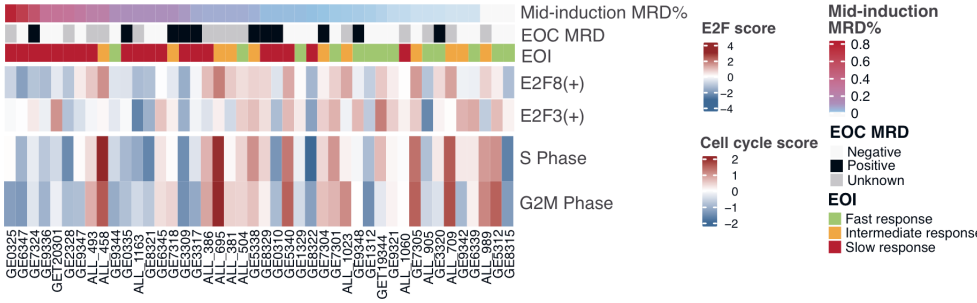


Figure 16. Changes in genes related to various cell cycle phases. Heatmaps show the expression scores of E2F2 and E2F8 regulons and the S- and G2/M-phase gene set scores by mid-induction treatment response. On the heatmaps, blue indicates downregulated gene expression, and red indicates upregulated gene expression. Adapted from Study III.

6 DISCUSSION

6.1 Novel biomarkers for T-ALL

T-ALL is an aggressive leukemia that accounts for approximately 15–25 % of pediatric and adult ALL cases and lacks well-established prognostic biomarkers. Using the Hemap data set, we established high expression of *SIX6* and *PRMT7* genes in T-ALL compared to healthy T cells. Alterations in both genes have previously been reported in other cancers, but their role in T-ALL is unknown. We have shown that SIX6 and PRMT7 are dysregulated in T-ALL and have further examined their utility as prognostic biomarkers.

6.1.1 Transcriptional regulator SIX6 affects T-ALL outcomes

SIX6 is a TF that functions normally in both forebrain and retinal development (Kumar, 2009). In breast and non-small cell cancers, high *SIX6* expression is associated with poor OS of patients(Q. Liu et al., 2016; Xu et al., 2016). In T-ALL, high *SIX6* expression is recognized, especially in the *TAL1* subtype, but its detailed role remains known (Chen et al., 2018; Soulier et al., 2005).

Using three independent pediatric T-ALL data cohorts, we showed that high *SIX6* expression was associated with inferior outcomes in T-ALL. Our finding that *SIX6* could be used as a predictive biomarker for patient outcomes is significant, as there are currently no adopted biomarkers for diagnosing T-ALL (Bardelli et al., 2021).

High expression of *SIX6* was detected, especially in the *TAL1* subtype, when compared to healthy T cells using the Hemap dataset. scRNA-seq data from healthy thymic cells revealed that *SIX6* is not expressed in healthy T cells during T cell differentiation (Park et al., 2020), confirming that its expression in T-ALL is not related with immature cell phenotype. A clear correlation between the expression of *SIX6* and *TAL1* was found, suggesting a regulatory relationship between them. Our analysis of protein–DNA interactions (ChiP-seq), chromatin interactions (Hi-C), and nascent transcription (GRO-seq) revealed that TAL1 directly regulated the

expression of *SIX6* by binding to its distal upstream enhancer together with GATA3. Expression of *TAL1* and *GATA3* was also correlated with expression of *SIX6* in a silencing experiment, supporting the regulatory relationship (Sanda et al., 2012). As TAL1 and GATA3 are part of the T-ALL core regulatory circuit, our findings suggest that *SIX6* plays a role in central regulatory events that controls T-ALL development (Sanda et al., 2012).

To better understand the molecular mechanisms and relevance of high *SIX6* expression in T-ALL, we generated SIX6-deficient Jurkat cell lines and analyzed their gene expression using RNA-seq data. Silencing of *SIX6* expression led to changes in the expression of genes belonging to the mTOR-, KRAS-, and TNFα-related pathways. In T-ALL, these signaling pathways play a role in proliferation, cell survival, and apoptosis (Evangelisti et al., 2018; Prior et al., 2012; Zhou et al., 2017). For example, mTOR activity is frequently associated with oncogenicity, and especially in, T-ALL, it is known to regulate cell proliferation, survival, metabolism, and drug resistance (Evangelisti et al., 2018; Khanna et al., 2018). We hypothesized that changes in genes related to mTOR pathway mediated by *SIX6* deletion played a role in T-ALL maintenance and drug resistance; however, cellular functional studies did not reveal any major differences between the *SIX6* KO and WT cells. As the mTOR pathway can be targeted by small molecule inhibitors (Evangelisti et al., 2018; Zanardi et al., 2015), and for that reason we treated cells with the mTOR-specific inhibitor temsirolimus, but found no significant changes when comparing the *SIX6* KO and WT cells. Furthermore, overexpression of *SIX6* in the zebrafish T-ALL model did not show significant impact on disease progression.

Our *in vitro* results indicate that SIX6 does not play a dominant oncogenic role in T-ALL. However, SIX6 seems to be part of the larger regulatory complex, and changes in *SIX6* expression might affect the larger entity. One possibility that must be considered is that other members of the *SIX* family might replace the function of SIX6 when its expression is absent. Accordingly, we noticed a slight increase in *SIX3* expression in the *SIX6* KO cells. This might be an explanation why no major changes were observed in the tested cellular processes. In future, interaction partners for SIX6 should be investigated more thoroughly. In survival analyses, the Cox proportional hazard modelling in the TARGET cohort showed that *SIX6* is an independent risk factor in T-ALL, indicating its potential utility in patient risk classification.

6.1.2 PRMT7 has an active role in T-ALL pathogenesis

The PRMT proteins methylate arginine residues affecting various cellular functions such as gene transcription, RNA splicing, and DNA damage response (Feng Y et al., 2013; Liu F et al., 2020; Szewczyk et al., 2020). Histone proteins are a major substrate for PRMTs, which use them to fine-tune gene expression levels. It is therefore unsurprising that they play important roles also in cancer. PRMT7 is a unique protein among PRMTs, as it only generates ω-monomethyl arginine residues in its substrates. Compared to other PRMT family members, role of PRMT7 has not been as widely studied. High expression of *PRMT7* has been associated with metastasis and decreased survival in breast cancer, renal cell carcinoma, and non-small cell lung cancer (Baldwin et al., 2014; Cheng D et al., 2018; Yao R et al., 2014). However, no comprehensive studies have been conducted on the role of PRMT7 in hematological cancers. We studied the expression of *PRMT* family genes across the hematological cancers and observed increased expression of *PRMT5* and *PRMT7* in T-ALL compared to healthy T cells. Increased expression of *PRMT5* was observed across all studied hematological cancers, but high expression of *PRMT7* was more unique to T-ALL, and was thus the reason for focusing on its role in this thesis. When the expression of *PRMT7* was studied between different T-ALL subtypes, we noticed significantly higher levels in mature T-ALL subtypes compared to thymic precursor cells. In two independent patient cohorts, high levels of *PRMT7* mRNA were associated with a trend toward decreased survival. The number of patients in both cohorts was limited, impacting the statistical power of comparison, and patients with more than 5% of the residual disease were excluded from the TARGET cohort. In the future, if our findings are replicated in a larger study, the expression of *PRMT7* may be a useful prognostic marker for T-ALL.

In total, seven *PRMT7* KO cell lines were generated to study the oncogenic potential of *PRMT7* and its role in T-ALL. Various cellular functional assays were performed, including the colony formation assay, which showed decreased oncogenic potential in *PRMT7* KO cells. This suggests that high level of *PRMT7* mRNA may promote aggressiveness in T-ALL.

In our analysis of PRMT7, we focused on well-known T-ALL-related genes. Interestingly, alterations in the three arginine residues of the RUNX1 protein were observed in the PRMT7 KO cells. Simultaneously, the expression of the RUNX1 target genes was significantly downregulated, suggesting that PRMT7 activity fine-tunes the transcriptional activity of RUNX1 by altering the methylation of its arginine residues. RUNX1 itself plays an essential role in normal lymphocyte

differentiation, and in T-ALL it has shown to drive oncogenesis by cooperating with TAL1 and NOTCH1 (Islam et al., 2023; Sanda, 2017), so alterations in RUNX1 function might contribute to T-ALL. RUNX1 target genes play many critical roles in cellular processes, especially in hematopoietic cells (Okuda et al., 1996; Wang et al., 1996; Zhu et al., 2021)(e.g., they are involved in differentiation and cell cycle regulation). Previous studies have shown that changes in RUNX1 arginine monomethylation via PRMT1 and PRMT4 promote cell survival, resistance to apoptosis, and the maintenance of undifferentiated phenotypes in hematopoietic cells (Matsumura T et al., 2020; Mizutani et al., 2015; Zhao et al., 2008).

One popular way to study the role of individual genes is to perform RNA-seq analysis. It can be used to identify differences in gene expression between modified and unmodified cells, which can reveal potential changes in cell state. Hence, we performed RNA-seq on the *PRMT7* KO and WT cells and showed that the deletion of *PRMT7* altered the expression of many genes involved in drug resistance, tumor progression, and T-ALL pathogenesis. Among the downregulated genes in *PRMT7* KO cells, three genes are directly involved in drug resistance and relapse, namely *NR3C2*, *ST8SIA6*, and *ARHGEF26* (Li et al., 2020; Ma et al., 2015; Zhang et al., 2015). This suggests that high PRMT7 activity leads to alterations in arginine methylation of proteins that associate with drug sensitivity, and thereby may impact therapy response. *PRMT7* KO led also to the downregulation of the *BCL11A* gene, which has a well-known function in many hematological cancers in as it can block cell differentiation, inhibit apoptosis, and promote cell proliferation (Yu et al., 2012). High PRMT7 could also mediate its effect on T-ALL via *IL7R*, *CCDN2*, and *NKX2-2*, whose expression was altered in KO cells, as they participate in cell cycle regulation, cellular growth, and cellular transformation (Clappier et al., 2006; Nagel et al., 2018, 2017; Ribeiro et al., 1987). Finally, alterations in the *RUNX1* target genes were noted in the *PRMT7* KO cells, indicating that PRMT7 is one of the regulators affecting to the oncogenic role of *RUNX1* in T-ALL.

Taken together, our results suggest that PRMT7 is involved in the pathogenesis of T-ALL, and may promote the aggressiveness of the disease. To determine a more detailed mechanistic function of PRMT7 in T-ALL in future, more extensive proteomic studies should be conducted. The study of arginine monomethylation alterations requires immunoaffinity purification and capturing of the desired peptides. An enhanced peptide capture technology, and larger sample sizes would provide additional information about the mechanisms in the future.

6.2 The effect of genomic alterations in *ETV6::RUNX1* leukemia treatment response

The *ETV6::RUNX1* subgroup is the second most common subtype of B-ALL. Patients belonging to this subgroup respond well to chemotherapy and have excellent outcomes. However, a subset of patients has insufficient response to induction therapy and increased risk of relapse (Coustan-Smith et al., 2000; Madzo et al., 2003; O'Connor et al., 2018; Wang et al., 2018). In total, around 5–6% of *ETV6::RUNX1* cases relapse. Despite the low overall relapse rate, these relapses accounts for approximately 20% of all ALL relapses, due to the high prevalence of *ETV6::RUNX1* subtype (Kuster et al., 2011; Loh et al., 2006; Norén-Nyström et al., 2023). This highlights the need to understand the underlying causes and develop better treatment strategies for cases with increased risk of relapse. Additionally, the short- and long-term side effects of leukemia treatment can be severe, making it important to learn more about the biology of good responders in order to reduce the burden of current treatment and minimize possible side effects.

In the past, vast majority of genomic studies have focused on the high-risk subtypes. Genetic risk factors in *ETV6::RUNX1* leukemia have also been investigated, with a focus on the relapse rate and OS (Bokemeyer et al., 2014; Brady et al., 2022; Grausenburger et al., 2016; Stams et al., 2006; van Delft et al., 2011). Our focus was on the genomic features that underlie early therapy response, a phenomenon that is closely related to patient outcome.

In CNV analysis, we found several alterations, which associated with treatment response. The most prominent finding was the deletion of 12p, which was more frequent in patients with the fast treatment response. This deletion includes multiple genes, such as *KRAS*. While activating *RAS* mutations are among the most frequent mutations in cancer and occur in about 15% of pediatric B-ALL cases (Irving et al., 2014; Jerchel et al., 2018; Ward et al., 2012), deletion of *KRAS* has not been recognized earlier as a relevant finding (Irving et al., 2014; Messina et al., 2016). Our findings also strengthen the findings of a recent study in which a nominal association between 12p deletion and a fast EOI response was found (Brady et al., 2022). In clinical laboratory, 12p deletion is routinely detected, thereby potentially facilitating integration of our findings into clinical use. Extra copies of the fusion der(21)t(12;21) have been associated with more aggressive disease and a higher relapse risk (Fears et al., 1996; Lilljebjörn et al., 2010; Loncarevic et al., 1999; McLean et al., 1996). Consistent with earlier studies, amplification of der(21) was more frequent in slow responders in our patient cohort. Additionally, we identified a deletion on chr3,

including the *TBL1XR1* gene, which exhibited a higher frequency among slow responders (MRD at EOC). The *TBL1XR1* deletion was recently recognized as the most significant genetic change in predicting EFS and OS in *ETV6::RUNX1* leukemia (Brady et al., 2022).

In the mutational signature analysis, we observed that the COSMIC Signature 2, also known as the APOBEC signature, correlated with fast early treatment responses. Overexpression of genes that drive APOBEC signature, namely *A3A* and *A3B*, associate with genes regulating the cell cycle, and *A3B* expression is known to be enriched in the G2/M phase of the cell cycle (Ng et al., 2019). Our results also indicated expression of *A3B* primarly in cycling cells. In addition, we found that Signature 1 (clock-like cell-intrinsic process) and Signature 8 (unknown) were more frequent in slow responders. Signature 1 is detected in more than 98% of pediatric cancers, and previous studies have linked it to the patient's age at the diagnosis. To our knowledge, this is the first time that this signature has been associated with treatment response. Signature 8 is present in ~20% of all pediatric cancers, and it is known to arise often due to late replication errors and mutations in HR genes (Gröbner et al., 2018; Singh et al., 2020; Thatikonda et al., 2023). This signature is often associated with mutations of *BRCA1* and *BRCA2* in breast and ovarian cancers. However, this signature marks an overall failure in DNA DSBR, which primarily functions during the late S and G2 phases (Polak et al., 2017).

In 2020, Mehtonen et al. showed that two common features of chemoresistance in *ETV6::RUNX1* leukemia are the lack of a resting G1 cell cycle state in diagnostic leukemia cells and the lack of differentiation-associated regulatory network changes during induction therapy. These observations were made using a small sample size, and mostly with scRNA-seq data, so we wanted to explore this more using a larger sample size. The mutational signature results indicate changes in the cell cycle, and the earlier findings led us to investigate cell cycle activity at the RNA expression level. We observed that fast-responding *ETV6::RUNX1* cases have a higher proportion of cells in the active cell cycle based on marker gene expression and regulon activity data. By combining the gene expression findings with evidence of E2F regulon activity, we reliably linked the active cell cycle to a faster treatment response in *ETV6::RUNX1* ALL.

When analyzing IG gene rearrangements, we observed that the slow responders had more frequently rearrangements in the *IGK* genes and kappa-deleting elements than fast responders. This is a surprise, since *ETV6::RUNX1* -positive leukemia cells are thought to be arrested at the pro-B cell stage, where *IGK* light chains are not yet rearranged (Pine et al., 2003; Sun et al., 2017; van Zelm et al., 2005). Using

scRNA-seq combined with V(D)J profiling, we confirmed that *ETV6::RUNX1* cells acquired enhanced expression of IG light chains during treatment. Moreover, cells carrying a productive light chain rearrangement were enriched upon the therapy. Because this phenotype is acquired in healthy BM precursors only upon successful heavy and light chain recombination (Clark et al., 2013; Herzog et al., 2009; Müschen, 2018), we suggest that slow-responding *ETV6::RUNX1* cells may have acquired a survival advantage through the acquisition of active BCR signaling. BCR signaling can be targeted by BCR-related kinase inhibitors, such as Burton tyrosine kinase (BTK) inhibitor ibrutinib, which is approved for treatment of some lymphoma types (Burger and Wiestner, 2018). Preclinical potential of ibrutinib in ALL has been proven, suggesting that it could be a potential targeted agent for specific ALL cases with elevated pre-BCR and BCR signaling activity (Burger and Wiestner, 2018; Kim et al., 2017).

In conclusion, our approach to combining multiomic and treatment response data revealed genetic alterations and biological processes associated with early therapy responses in pediatric *ETV6::RUNX1* leukemia. A limitation of our study is the small number of samples in the WGS analysis, which limited the study power. We are planning to perform targeted sequencing in a cohort of over 250 *ETV6::RUNX1* cases to validate the current findings.

7 SUMMARY AND CONCLUSION

This thesis investigated novel prognostic biomarkers for T-ALL and the genomic alterations affecting the treatment response of the *ETV6::RUNX1* subtype of B-ALL.

T-ALL is an aggressive leukemia that lacks well-established prognostic biomarkers. Aberrant transcriptional regulation and post-translational modifications are central mechanisms that underlie leukemia development and maintenance. In this thesis, we discovered two novel genes, *PRMT7* and *SIX6,* as highly expressed in T-ALL when compared to other leukemia types and healthy T cells. Expression of *PRMT7* was associated with increased oncogenicity, while *SIX6* was directly upregulated by the *TAL1* gene, a well-established oncogene in T-ALL that defines the largest T-ALL subgroup. Using independent patient cohorts, we showed a significant association between high *SIX6* expression and inferior EFS, whereas high *PRMT7* expression showed a trend toward inferior EFS and OS. Further studies are needed to elucidate the exact role of these genes in T-ALL pathogenesis and the clinical relevance of prognostic findings.

ETV6::RUNX1 -positive leukemia is a common B-ALL subtype and usually associated with good outcome. However, suboptimal early therapy response is associated with increased risk of relapse. Here, we uncovered genetic alterations and mutational processes that were associated with treatment response. Our study revealed the genetic and transcriptomic heterogeneity of *ETV6::RUNX1* leukemia, and pave the way for identification of robust biomarkers in future.

8 REFERENCES

Abate-Shen, C. (2002). Deregulated homeobox gene expression in cancer: Cause or consequence? Nature Reviews Cancer, 2(10), 777–785. https://doi.org/10.1038/nrc907

Akawi, N., McRae, J., Ansari, M., Balasubramanian, M., Blyth, M., Brady, A. F., … Hurles, M. E. (2015). Discovery of four recessive developmental disorders using probabilistic genotype and phenotype matching among 4,125 families. Nature Genetics, 47(11), 1363–1369. https://doi.org/10.1038/ng.3410

Alaggio, R., Amador, C., Anagnostopoulos, I., Attygalle, A. D., Araujo, I. B. de O., Berti, E., … Xiao, W. (2022). The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. Leukemia, 36(7), 1720–1748. https://doi.org/10.1038/s41375-022-01620-2

Alexandrov, L. B., Jones, P. H., Wedge, D. C., Sale, J. E., Campbell, P. J., Nik-Zainal, S., & Stratton, M. R. (2015). Clock-like mutational processes in human somatic cells. Nature Genetics, 47(12), 1402–1407. https://doi.org/10.1038/ng.3441

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R., Behjati, S., Biankin, A. V., … Stratton, M. R. (2013). Signatures of mutational processes in human cancer. Nature, 500(7463), 415–421. https://doi.org/10.1038/nature12477

Amemiya, H. M., Kundaje, A., & Boyle, A. P. (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. Scientific Reports, 9(1), 9354. https://doi.org/10.1038/s41598-019-45839-z

Amodio, G., Cichy, J., Conde, P., Matteoli, G., Moreau, A., Ochando, J., … Gregori, S. (2019). Role of myeloid regulatory cells (MRCs) in maintaining tissue homeostasis and promoting tolerance in autoimmunity, inflammatory disease and transplantation. *Cancer Immunology, Immunotherapy*, *68*(4), 661–672. https://doi.org/10.1007/S00262-018-2264-3/FIGURES/2

Andersen, M. K., Autio, K., Barbany, G., Borgström, G., Cavelier, L., Golovleva, I., … Forestier, E. (2011). Paediatric B-cell precursor acute lymphoblastic leukaemia with t(1;19)(q23;p13): Clinical and cytogenetic characteristics of 47 cases from the Nordic countries treated according to NOPHO protocols. *British Journal of Haematology*, *155*(2), 235–243. https://doi.org/10.1111/j.1365-2141.2011.08824.x

Andreasson, P., Schwaller, J., Anastasiadou, E., Aster, J., & Gilliland, D. G. (2001). The expression of ETV6/CBFA2 (TEL/AML1) is not sufficient for the transformation of hematopoietic cell lines in vitro or the induction of hematologic disease in vivo. *Cancer Genetics and Cytogenetics*, *130*(2), 93–104. https://doi.org/10.1016/s0165-4608(01)00518-0

Antonysamy, S., Bonday, Z., Campbell, R. M., Doyle, B., Druzina, Z., Gheyi, T., … Emtage, S. (2012). Crystal structure of the human PRMT5:MEP50 complex. *Proceedings of the National Academy of Sciences*, *109*(44), 17960–17965. https://doi.org/10.1073/pnas.1209814109

Arai, F., Hirao, A., Ohmura, M., Sato, H., Matsuoka, S., Takubo, K., … Suda, T. (2004). Tie2/Angiopoietin-1 Signaling Regulates Hematopoietic Stem Cell Quiescence in the Bone Marrow Niche. *Cell*, *118*(2), 149–161. https://doi.org/10.1016/J.CELL.2004.07.004

Arber, D. A., Orazi, A., Hasserjian, R. P., Borowitz, M. J., Calvo, K. R., Kvasnicka, H.-M., … Tefferi, A. (2022). International Consensus Classification of Myeloid Neoplasms and Acute Leukemias: Integrating morphologic, clinical, and genomic data. *Blood*, *140*(11), 1200–1228. https://doi.org/10.1182/blood.2022015850

Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., … Vardiman, J. W. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*, *127*(20), 2391–2405. https://doi.org/10.1182/blood-2016-03-643544

Armstrong, S. A., Staunton, J. E., Silverman, L. B., Pieters, R., Den Boer, M. L., Minden, M. D., … Korsmeyer, S. J. (2001). MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature Genetics 2001 30:1*, *30*(1), 41–47. https://doi.org/10.1038/ng765

Baccin, C., Al-Sabah, J., Velten, L., Helbling, P. M., Grünschläger, F., Hernández-Malmierca, P., … Haas, S. (2019). Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. *Nature Cell Biology 2019 22:1*, *22*(1), 38–48. https://doi.org/10.1038/s41556-019-0439-6

Bain, B. J. (2017). WHO Classification of Leukemia☆. In *Reference Module in Life Sciences*. Elsevier. https://doi.org/10.1016/B978-0-12-809633-8.07356-8

Baldwin, B., Morettin A, & Côté J. (2014). Role of PRMTs in cancer: Could minor isoforms be leaving a mark? *World Journal of Biological Chemistry*, *5*(2), 115–129. https://doi.org/10.4331/WJBC.V5.I2.115

Bardelli, V., Arniani, S., Pierini, V., Di Giacomo, D., Pierini, T., Gorello, P., … La Starza, R. (2021). T-Cell Acute Lymphoblastic Leukemia: Biomarkers and Their Clinical Usefulness. *Genes*, *12*(8), 1118. https://doi.org/10.3390/genes12081118

Baron, M. H., Isern, J., & Fraser, S. T. (2012). The embryonic origins of erythropoiesis in mammals. *Blood*, *119*(21), 4828. https://doi.org/10.1182/BLOOD-2012-01-153486

Bedford, M., Frankel A, Yaffe MB, Clarke S, Leder P, & Richard S. (2000). Arginine methylation inhibits the binding of proline-rich ligands to Src homology 3, but not WW, domains. *The Journal of Biological Chemistry*, *275*(21), 16030–16036. https://doi.org/10.1074/JBC.M909368199

Bedford, M. T., & Clarke, S. G. (2009). Protein Arginine Methylation in Mammals: Who, What, and Why. *Molecular Cell*, *33*(1), 1. https://doi.org/10.1016/J.MOLCEL.2008.12.013

Bedford, M. T., & Richard, S. (2005). Arginine Methylation: An Emerging Regulatorof Protein Function. *Molecular Cell*, *18*(3), 263–272. https://doi.org/10.1016/J.MOLCEL.2005.04.003

Begley, C. G., Aplan, P. D., Davey, M. P., Nakahara, K., Tchorz, K., Kurtzberg, J., … Waldmann, T. A. (1989). Chromosomal translocation in a human leukemic stem-cell line disrupts the T-cell antigen receptor delta-chain diversity region and results in a previously unreported fusion transcript. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(6), 2031–2035. https://doi.org/10.1073/pnas.86.6.2031

Belluschi, S., Calderbank, E. F., Ciaurro, V., Pijuan-Sala, B., Santoro, A., Mende, N., … Laurenti, E. (2018). Myelo-lymphoid lineage restriction occurs in the human

haematopoietic stem cell compartment before lymphoid-primed multipotent progenitors. *Nature Communications*, *9*(1). https://doi.org/10.1038/S41467-018-06442-4

Béné, M. C., Nebe, T., Bettelheim, P., Buldini, B., Bumbea, H., Kern, W., … Porwit, A. (2011). Immunophenotyping of acute leukemia and lymphoproliferative disorders: A consensus proposal of the European LeukemiaNet Work Package 10. *Leukemia*, *25*(4), 567–574. https://doi.org/10.1038/leu.2010.312

Bernier, G., Panitz, F., Zhou, X., Hollemann, T., Gruss, P., & Pieler, T. (2000). Expanded retina territory by midbrain transformation upon overexpression of Six6 (Optx2) in Xenopus embryos. *Mechanisms of Development*, *93*(1–2), 59–69. https://doi.org/10.1016/s0925-4773(00)00271-9

Bhattacharya, S., Zhang, Q., & Andersen, M. E. (2011). A deterministic map of Waddington's epigenetic landscape for cell fate specification. *BMC Systems Biology*, *5*(1), 1–12. https://doi.org/10.1186/1752-0509-5-85/FIGURES/6

Bhojwani, D., Pei, D., Sandlund, J. T., Jeha, S., Ribeiro, R. C., Rubnitz, J. E., … Pui, C.-H. (2012). ETV6-RUNX1-positive childhood acute lymphoblastic leukemia: Improved outcome with contemporary therapy. *Leukemia*, *26*(2), 265–270. https://doi.org/10.1038/leu.2011.227

Biondi, A., Gandemer, V., Campbell, M., Castor, A., Pieters, R., De Lorenzo, P., … Valsecchi, M. G. (2017). A Treatment Protocol with Imatinib and Intensive Chemotherapy for Pediatric Philadelphia Positive Acute Lymphoblastic Leukemia Patients: A Single-Arm, Intergroup Study (EsPhALL 2010-2014). *Blood*, *130*(Supplement 1), 97. https://doi.org/10.1182/blood.V130.Suppl_1.97.97

Blanc, Roméo S., & Richard, S. (2017). Arginine Methylation: The Coming of Age. *Molecular Cell*, *65*(1), 8–24. https://doi.org/10.1016/j.molcel.2016.11.003

Blanc, Roméo Sébastien, Vogel, G., Chen, T., Crist, C., & Richard, S. (2016). PRMT7 Preserves Satellite Cell Regenerative Capacity. *Cell Reports*, *14*(6), 1528–1539. https://doi.org/10.1016/j.celrep.2016.01.022

Boer, M. L. D., Slegtenhorst, M. van, Menezes, R. X. D., Cheok, M. H., Buijs-Gladdines, J. G., Peters, S. T., … Pieters, R. (2009). A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: A genome-wide classification study. *The Lancet Oncology*, *10*(2), 125–134. https://doi.org/10.1016/S1470-2045(08)70339-5

Bokemeyer, A., Eckert, C., Meyr, F., Koerner, G., von Stackelberg, A., Ullmann, R., … Seeger, K. (2014). Copy number genome alterations are associated with treatment response and outcome in relapsed childhood ETV6/RUNX1-positive acute lymphoblastic leukemia. *Haematologica*, *99*(4), 706–714. https://doi.org/10.3324/haematol.2012.072470

Bommhardt, U., Schraven, B., & Simeoni, L. (2019). Beyond TCR Signaling: Emerging Functions of Lck in Cancer and Immunotherapy. *International Journal of Molecular Sciences*, *20*(14), 3500. https://doi.org/10.3390/ijms20143500

Boublikova, L., Kalinova, M., Ryan, J., Quinn, F., O'Marcaigh, A., Smith, O., … Lawler, M. (2006). Wilms' tumor gene 1 (WT1) expression in childhood acute lymphoblastic leukemia: A wide range of WT1 expression levels, its impact on prognosis and minimal residual disease monitoring. *Leukemia*, *20*(2), 254–263. https://doi.org/10.1038/sj.leu.2404047

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., … Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, *122*(6), 947–956. https://doi.org/10.1016/j.cell.2005.08.020

Bradley, T. R., & Metcalf, D. (1966). The growth of mouse bone marrow cells in vitro. *The Australian Journal of Experimental Biology and Medical Science*, *44*(3), 287–299. https://doi.org/10.1038/ICB.1966.28

Brady, S. W., Roberts, K. G., Gu, Z., Shi, L., Pounds, S., Pei, D., … Mullighan, C. G. (2022). The genomic landscape of pediatric acute lymphoblastic leukemia. *Nature Genetics*, *54*(9), 1376–1389. https://doi.org/10.1038/s41588-022-01159-z

Bressanin, D., Evangelisti, C., Ricci, F., Tabellini, G., Chiarini, F., Tazzari, P. L., … Martelli, A. M. (2012). Harnessing the PI3K/Akt/mTOR pathway in T-cell acute lymphoblastic leukemia: Eliminating activity by targeting at different levels. *Oncotarget*, *3*(8), 811–823.

Brown, G., & Ceredig, R. (2019). Modeling the hematopoietic landscape. *Frontiers in Cell and Developmental Biology*, *7*(JUN), 104. https://doi.org/10.3389/FCELL.2019.00104/BIBTEX

Buchner, M., Swaminathan, S., Chen, Z., & Müschen, M. (2015). Mechanisms of pre-B-cell receptor checkpoint control and its oncogenic subversion in acute lymphoblastic leukemia. *Immunological Reviews*, *263*(1), 192–209. https://doi.org/10.1111/IMR.12235

Bullinger, L., & Fröhling, S. (2013). Chapter 58—Leukemias. In G. S. Ginsburg & H. F. Willard (Eds.), *Genomic and Personalized Medicine (Second Edition)* (pp. 675–690). Academic Press. https://doi.org/10.1016/B978-0-12-382227-7.00058-6

Burger, J. A., & Wiestner, A. (2018). Targeting B cell receptor signalling in cancer: Preclinical and clinical advances. *Nature Reviews Cancer*, *18*(3), 148–167. https://doi.org/10.1038/nrc.2017.121

Buzdin, A., Skvortsova, I. I., Li, X., & Wang, Y. (2021). Editorial: Next Generation Sequencing Based Diagnostic Approaches in Clinical Oncology. *Frontiers in Oncology*, *10*. Retrieved from https://www.frontiersin.org/articles/10.3389/fonc.2020.635555

Calvi, L. M., Adams, G. B., Weibrecht, K. W., Weber, J. M., Olson, D. P., Knight, M. C., … Scadden, D. T. (2003). Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature 2003 425:6960*, *425*(6960), 841–846. https://doi.org/10.1038/nature02040

Cavé, H., Cacheux, V., Raynaud, S., Brunie, G., Bakkus, M., Cochaux, P., … Grandchamp, B. (1997). ETV6 is the target of chromosome 12p deletions in t(12;21) childhood acute lymphocytic leukemia. *Leukemia*, *11*(9), 1459–1464. https://doi.org/10.1038/sj.leu.2400798

Cazzaniga, G., Lorenzo, P. D., Alten, J., Röttgers, S., Hancock, J., Saha, V., … Biondi, A. (2018). Predictive value of minimal residual disease in Philadelphia-chromosome-positive acute lymphoblastic leukemia treated with imatinib in the European intergroup study of post-induction treatment of Philadelphia-chromosome-positive acute lymphoblastic leukemia, based on immunoglobulin/T-cell receptor and BCR/ABL1 methodologies. *Haematologica*, *103*(1), 107–115. https://doi.org/10.3324/haematol.2017.176917

Chen, B., Jiang, L., Zhong, M.-L., Li, J.-F., Li, B.-S., Peng, L.-J., … Chen, S.-J. (2018). Identification of fusion genes and characterization of transcriptome features in T-cell acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences*, *115*(2), 373–378. https://doi.org/10.1073/pnas.1717125115

Cheng D, Z, H., L, Z., D, X., S, D., & P, Z. (2018). PRMT7 contributes to the metastasis phenotype in human non-small-cell lung cancer cells possibly through the interaction with HSPA5 and EEF2. *OncoTargets and Therapy*, *11*, 4869–4876. https://doi.org/10.2147/OTT.S166412

Cheng, H., Zheng, Z., & Cheng, T. (2020). New paradigms on hematopoietic stem cell differentiation. *Protein & Cell*, *11*(1), 34. https://doi.org/10.1007/S13238-019-0633-0

Chessells, J. M., Harrison, C. J., Kempski, H., Webb, D. K. H., Wheatley, K., Hann, I. M., … MRC Childhood Leukaemia working party. (2002). Clinical features, cytogenetics and outcome in acute lymphoblastic and myeloid leukaemia of infancy: Report from the MRC Childhood Leukaemia working party. *Leukemia*, *16*(5), 776–784. https://doi.org/10.1038/sj.leu.2402468

Chung, E., & Kondo, M. (2011). Role of Ras/Raf/MEK/ERK signaling in physiological hematopoiesis and leukemia development. *Immunologic Research*, *49*(1–3), 248–268. https://doi.org/10.1007/s12026-010-8187-5

Ciofani, M., & Zúñiga-Pflücker, J. C. (2007). The Thymus as an Inductive Site for T Lymphopoiesis. *Https://Doi.Org/10.1146/Annurev.Cellbio.23.090506.123547*, *23*, 463–493. https://doi.org/10.1146/ANNUREV.CELLBIO.23.090506.123547

Clappier, E., Auclerc, M. F., Rapion, J., Bakkus, M., Caye, A., Khemiri, A., … Cavé, H. (2014). An intragenic ERG deletion is a marker of an oncogenic subtype of B-cell precursor acute lymphoblastic leukemia with a favorable outcome despite frequent IKZF1 deletions. *Leukemia*, *28*(1), 70–77. https://doi.org/10.1038/leu.2013.277

Clappier, E., Cuccuini, W., Cayuela, J. M., Vecchione, D., Baruchel, A., Dombret, H., … Soulier, J. (2006). Cyclin D2 dysregulation by chromosomal translocations to TCR loci in T-cell acute lymphoblastic leukemias. *Leukemia*, *20*(1), 82–86. https://doi.org/10.1038/SJ.LEU.2404008

Clark, M. R., Mandal, M., Ochiai, K., & Singh, H. (2013). Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nature Reviews Immunology 2013 14:2*, *14*(2), 69–80. https://doi.org/10.1038/nri3570

Collins, A. M., & Watson, C. T. (2018). Immunoglobulin light chain gene rearrangements, receptor editing and the development of a self-tolerant antibody repertoire. *Frontiers in Immunology*, *9*(OCT), 2249. https://doi.org/10.3389/FIMMU.2018.02249/BIBTEX

Cooper, M. D. (2015). The early history of B cells. *Nature Reviews. Immunology*, *15*(3), 191–197. https://doi.org/10.1038/nri3801

Côté, J., & Richards, S. (2005). Tudor domains bind symmetrical dimethylated arginines. *The Journal of Biological Chemistry*, *280*(31), 28476–28483. https://doi.org/10.1074/JBC.M414328200

Courtney, A. H., Lo, W.-L., & Weiss, A. (2018). TCR Signaling: Mechanisms of Initiation and Propagation. *Trends in Biochemical Sciences*, *43*(2), 108–123. https://doi.org/10.1016/j.tibs.2017.11.008

Coustan-Smith, E., Sancho, J., Hancock, M. L., Boyett, J. M., Behm, F. G., Raimondi, S. C., … Campana, D. (2000). Clinical importance of minimal residual disease in childhood acute lymphoblastic leukemia. *Blood*, *96*(8), 2691–2696. https://doi.org/10.1182/blood.V96.8.2691

Cuitiño, M. C., Pécot, T., Sun, D., Kladney, R., Okano-Uchida, T., Shinde, N., … Leone, G. (2019). Two Distinct E2F Transcriptional Modules Drive Cell Cycles and Differentiation. *Cell Reports*, *27*(12), 3547-3560.e5. https://doi.org/10.1016/j.celrep.2019.05.004

Cully, M., You, H., Levine, A. J., & Mak, T. W. (2006). Beyond PTEN mutations: The PI3K pathway as an integrator of multiple inputs during tumorigenesis. *Nature Reviews Cancer*, *6*(3), 184–192. https://doi.org/10.1038/nrc1819

De Kleer, I., Willems, F., Lambrecht, B., & Goriely, S. (2014). Ontogeny of myeloid cells. *Frontiers in Immunology*, *5*(AUG), 423. https://doi.org/10.3389/FIMMU.2014.00423/BIBTEX

DeKoter, R. P., Lee, H. J., & Singh, H. (2002). PU.1 regulates expression of the interleukin-7 receptor in lymphoid progenitors. *Immunity*, *16*(2), 297–309. https://doi.org/10.1016/S1074-7613(02)00269-8

DeKoter, R. P., & Singh, H. (2000). Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science (New York, N.Y.)*, *288*(5470), 1439–1442. https://doi.org/10.1126/SCIENCE.288.5470.1439

Diacou, R., Nandigrami, P., Fiser, A., Liu, W., Ashery-Padan, R., & Cvekl, A. (2022). Cell fate decisions, transcription factors and signaling during early retinal development. *Progress in Retinal and Eye Research*, *91*, 101093. https://doi.org/10.1016/j.preteyeres.2022.101093

Díaz-Gay, M., Vila-Casadesús, M., Franch-Expósito, S., Hernández-Illán, E., Lozano, J. J., & Castellví-Bel, S. (2018). Mutational Signatures in Cancer (MuSiCa): A web application to implement mutational signatures analysis in cancer samples. *BMC Bioinformatics*, *19*(1), 224. https://doi.org/10.1186/s12859-018-2234-y

Doll, R., & Wakeford, R. (1997). Risk of childhood cancer from fetal irradiation. *The British Journal of Radiology*, *70*, 130–139. https://doi.org/10.1259/bjr.70.830.9135438

Dong, S., Yang, Z., Xu, P., Zheng, W., Zhang, B., Fu, F., … Yu, W. (2022). Mutually exclusive epigenetic modification on SIX6 with hypermethylation for precancerous stage and metastasis emergence tracing. *Signal Transduction and Targeted Therapy*, *7*(1), 1–4. https://doi.org/10.1038/s41392-022-01026-7

dos Santos, N. R., Ghysdael, J., & Tran Quang, C. (2019). The TCR/CD3 complex in leukemogenesis and as a therapeutic target in T-cell acute lymphoblastic leukemia. *Advances in Biological Regulation*, *74*, 100638. https://doi.org/10.1016/j.jbior.2019.100638

Duy, C., Hurtz, C., Shojaee, S., Cerchietti, L., Geng, H., Swaminathan, S., … Müschen, M. (2011). BCL6 enables Ph+ acute lymphoblastic leukaemia cells to survive BCR-ABL1 kinase inhibition. *Nature*, *473*(7347), 384–388. https://doi.org/10.1038/nature09883

Dyson, N. (1998). The regulation of E2F by pRB-family proteins. *Genes & Development*, *12*(15), 2245–2262. https://doi.org/10.1101/gad.12.15.2245

Eiser, C., Stride, C. b., Vora, A., Goulden, N., Mitchell, C., Buck, G., … the National Cancer Research Institute Childhood Leukaemia Sub-Group and UK Childhood Leukaemia Clinicians Network. (2017). Prospective evaluation of quality of life in children treated in UKALL 2003 for acute lymphoblastic leukaemia: A cohort study. *Pediatric Blood & Cancer*, *64*(11), e26615. https://doi.org/10.1002/pbc.26615

Eldor, A., Vlodavsky, I., Deutsch, V., & Levine, R. F. (1989). Megakaryocyte function and dysfunction. *Bailliere's Clinical Haematology*, *2*(3), 543–568. https://doi.org/10.1016/S0950-3536(89)80033-2

Ellisen, L. W., Bird, J., West, D. C., Soreng, A. L., Reynolds, T. C., Smith, S. D., & Sklar, J. (1991). TAN-1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell*, *66*(4), 649–661. https://doi.org/10.1016/0092-8674(91)90111-b

Ernst, P., Fisher, J. K., Avery, W., Wade, S., Foy, D., & Korsmeyer, S. J. (2004). Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Developmental Cell*, *6*(3), 437–443. https://doi.org/10.1016/S1534-5807(04)00061-9

Evangelisti, C., Chiarini, F., McCubrey, J. A., & Martelli, A. M. (2018). Therapeutic Targeting of mTOR in T-Cell Acute Lymphoblastic Leukemia: An Update. *International Journal of Molecular Sciences*, *19*(7), 1878. https://doi.org/10.3390/ijms19071878

Fears, S., Vignon, C., Bohlander, S. K., Smith, S., Rowley, J. D., & Nucifora, G. (1996). Correlation between the ETV6/CBFA2 (TEL/AML1) fusion gene and karyotypic abnormalities in children with B-cell precursor acute lymphoblastic leukemia. *Genes, Chromosomes & Cancer*, *17*(2), 127–135. https://doi.org/10.1002/(SICI)1098-2264(199610)17:2<127::AID-GCC8>3.0.CO;2-7

Feng Y, Maity R, Whitelegge JP, Hadjikyriacou A, Li Z, C, Z.-L., … SG, C. (2013). Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions. *The Journal of Biological Chemistry*, *288*(52), 37010–37025. https://doi.org/10.1074/JBC.M113.525345

Feng, Y., Zhang, T., Wang, Y., Xie, M., Ji, X., Luo, X., … Xia, L. (2021). Homeobox Genes in Cancers: From Carcinogenesis to Recent Therapeutic Intervention. *Frontiers in Oncology*, *11*. Retrieved from https://www.frontiersin.org/articles/10.3389/fonc.2021.770428

Ferreira, R., Ohneda, K., Yamamoto, M., & Philipsen, S. (2005). GATA1 Function, a Paradigm for Transcription Factors in Hematopoiesis. *Molecular and Cellular Biology*, *25*(4), 1215. https://doi.org/10.1128/MCB.25.4.1215-1227.2005

Fischer, U., Yang, J. J., Ikawa, T., Hein, D., Vicente-Dueñas, C., Borkhardt, A., & Sánchez-García, I. (2020). Cell Fate Decisions: The Role of Transcription Factors in Early B-cell Development and Leukemia. *Blood Cancer Discovery*, *1*(3), 224–233. https://doi.org/10.1158/2643-3230.BCD-20-0011

Fisher, R. C., & Scott, E. W. (1998). Role of PU.1 in Hematopoiesis. *STEM CELLS*, *16*(1), 25–37. https://doi.org/10.1002/stem.160025

Ford, A. M., Bennett, C. A., Price, C. M., Bruin, M. C. A., Van Wering, E. R., & Greaves, M. (1998). Fetal origins of the TEL-AML1 fusion gene in identical twins with leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(8), 4584–4588.

Ford, A. M., & Greaves, M. (2017). ETV6-RUNX1+Acute Lymphoblastic Leukaemia in Identical Twins. In Y. Groner, Y. Ito, P. Liu, J. C. Neil, N. A. Speck, & A. van Wijnen (Eds.), *RUNX Proteins in Development and Cancer* (pp. 217–228). Singapore: Springer. https://doi.org/10.1007/978-981-10-3233-2_14

Forestier, E., Heyman, M., Andersen, M. K., Autio, K., Blennow, E., Borgström, G., … NOPHO Leukaemia Cytogenetic Study Group (NLCSG). (2008). Outcome of ETV6/RUNX1-positive childhood acute lymphoblastic leukaemia in the NOPHO-ALL-1992 protocol: Frequent late relapses but good overall survival. *British Journal of Haematology*, *140*(6), 665–672. https://doi.org/10.1111/j.1365-2141.2008.06980.x

Foroughi-Asl, H., Jeggari, A., Maqbool, K., Ivanchuk, V., Elhami, K., & Wirta, V. (2019). *BALSAMIC: Bioinformatic Analysis pipeLine for SomAtic MutatIons in Cancer* [Python]. Retrieved from https://github.com/Clinical-Genomics/BALSAMIC (Original work published 2018)

Fransecky, L., Neumann, M., Heesch, S., Schlee, C., Ortiz-Tanchez, J., Heller, S., … Baldus, C. D. (2016). Silencing of GATA3 defines a novel stem cell-like subgroup of ETP-ALL. *Journal of Hematology & Oncology*, *9*(1), 95. https://doi.org/10.1186/s13045-016-0324-8

Fulbright, J. M., Raman, S., McClellan, W. S., & August, K. J. (2011). Late effects of childhood leukemia therapy. *Current Hematologic Malignancy Reports*, *6*(3), 195–205. https://doi.org/10.1007/s11899-011-0094-x

Galloway, J. L., & Zon, L. I. (2003). 3 Ontogeny of hematopoiesis: Examining the emergence of hematopoietic cells in the vertebrate embryo. In *Current Topics in Developmental Biology* (Vol. 53, pp. 139–158). Academic Press. https://doi.org/10.1016/S0070-2153(03)53004-6

Gellert, M. (2002). V(D)J recombination: RAG proteins, repair factors, and regulation. *Annual Review of Biochemistry*, *71*, 101–132. https://doi.org/10.1146/annurev.biochem.71.090501.150203

Geoffroy, V., Herenger, Y., Kress, A., Stoetzel, C., Piton, A., Dollfus, H., & Muller, J. (2018). AnnotSV: An integrated tool for structural variations annotation. *Bioinformatics*, *34*(20), 3572–3574. https://doi.org/10.1093/bioinformatics/bty304

Germain, R. N. (2002). T-cell development and the CD4–CD8 lineage decision. *Nature Reviews Immunology 2002 2:5*, *2*(5), 309–322. https://doi.org/10.1038/nri798

Girardi, T., Vereecke, S., Sulima, S. O., Khan, Y., Fancello, L., Briggs, J. W., … De Keersmaecker, K. (2018). The T-cell leukemia-associated ribosomal RPL10 R98S mutation enhances JAK-STAT signaling. *Leukemia*, *32*(3), 809–819. https://doi.org/10.1038/leu.2017.225

Girardi, Tiziana, Vicente, C., Cools, J., & De Keersmaecker, K. (2017). The genetics and molecular biology of T-ALL. *Blood*, *129*(9), 1113–1123. https://doi.org/10.1182/blood-2016-10-706465

Giurgiu, M., Reinhard, J., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., … Ruepp, A. (2019). CORUM: the comprehensive resource of mammalian protein complexes—2019. *Nucleic Acids Research*, *47*(Database issue), D559. https://doi.org/10.1093/NAR/GKY973

Golub, T. R., Barker, G. F., Bohlander, S. K., Hiebert, S. W., Ward, D. C., Bray-Ward, P., … Gilliland, D. G. (1995). Fusion of the TEL gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(11), 4917–4921. https://doi.org/10.1073/PNAS.92.11.4917

González-García, S., García-Peydró, M., Alcain, J., & Toribio, M. L. (2012). Notch1 and IL-7 Receptor Signalling in Early T-cell Development and Leukaemia. In F. Radtke (Ed.), *Notch Regulation of the Immune System* (pp. 47–73). Berlin, Heidelberg: Springer. https://doi.org/10.1007/82_2012_231

Grausenburger, R., Bastelberger, S., Eckert, C., Kauer, M., Stanulla, M., Frech, C., … Panzer-Grümayer, R. (2016). Genetic alterations in glucocorticoid signaling pathway components are associated with adverse prognosis in children with relapsed ETV6/RUNX1-positive acute lymphoblastic leukemia. *Leukemia & Lymphoma*, *57*(5), 1163–1173. https://doi.org/10.3109/10428194.2015.1088650

Graux, C., Cools, J., Melotte, C., Quentmeier, H., Ferrando, A., Levine, R., … Hagemeijer, A. (2004). Fusion of NUP214 to ABL1 on amplified episomes in T-cell acute lymphoblastic leukemia. *Nature Genetics*, *36*(10), 1084–1089. https://doi.org/10.1038/ng1425

Graux, C., Stevens-Kroef, M., Lafage, M., Dastugue, N., Harrison, C. J., Mugneret, F., … Belgian Cytogenetic Group for Hematology and Oncology. (2009). Heterogeneous patterns of amplification of the NUP214-ABL1 fusion gene in T-cell acute

lymphoblastic leukemia. *Leukemia*, *23*(1), 125–133. https://doi.org/10.1038/leu.2008.278

Greaves, M. (1988). Speculations on the cause of childhood acute lymphoblastic leukemia. *Leukemia*, *2*(2), 120–125.

Greaves, M. (2006). Infection, immune responses and the aetiology of childhood leukaemia. *Nature Reviews Cancer*, *6*(3), 193–203. https://doi.org/10.1038/nrc1816

Greaves, M. (2018). A causal mechanism for childhood acute lymphoblastic leukaemia. *Nature Reviews Cancer*, *18*(8), 471–484. https://doi.org/10.1038/s41568-018-0015-6

Gröbner, S. N., Worst, B. C., Weischenfeldt, J., Buchhalter, I., Kleinheinz, K., Rudneva, V. A., … Pfister, S. M. (2018). The landscape of genomic alterations across childhood cancers. *Nature*, *555*(7696), 321–327. https://doi.org/10.1038/nature25480

Grossmann, V., Kern, W., Harbich, S., Alpermann, T., Jeromin, S., Schnittger, S., … Kohlmann, A. (2011). Prognostic relevance of RUNX1 mutations in T-cell acute lymphoblastic leukemia. *Haematologica*, *96*(12), 1874–1877. https://doi.org/10.3324/haematol.2011.043919

Gu, Z., Eils R, & Schlesner M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics (Oxford, England)*, *32*(18), 2847–2849. https://doi.org/10.1093/BIOINFORMATICS/BTW313

Gu, Zhaohui, Churchman, M. L., Roberts, K. G., Moore, I., Zhou, X., Nakitandwe, J., … Mullighan, C. G. (2019). PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nature Genetics*, *51*(2), 296–307. https://doi.org/10.1038/s41588-018-0315-5

Gu, Zhaohui, Churchman, M., Roberts, K., Li, Y., Liu, Y., Harvey, R. C., … Mullighan, C. G. (2016). Genomic analyses identify recurrent MEF2D fusions in acute lymphoblastic leukaemia. *Nature Communications*, *7*, 13331. https://doi.org/10.1038/ncomms13331

Guidez, F., Petrie, K., Ford, A. M., Lu, H., Bennett, C. A., MacGregor, A., … Zelent, A. (2000). Recruitment of the nuclear receptor corepressor N-CoR by the TEL moiety of the childhood leukemia-associated TEL-AML1 oncoprotein. *Blood*, *96*(7), 2557–2561.

Guo, X., Zhang, R., Liu, J., Li, M., Song, C., Dovat, S., … Ge, Z. (2015). Characterization of LEF1 High Expression and Novel Mutations in Adult Acute Lymphoblastic Leukemia. *PLoS ONE*, *10*(5), e0125429. https://doi.org/10.1371/journal.pone.0125429

Gutierrez, A., Sanda, T., Ma, W., Zhang, J., Grebliunaite, R., Dahlberg, S., … Look, A. T. (2010). Inactivation of LEF1 in T-cell acute lymphoblastic leukemia. *Blood*, *115*(14), 2845–2851. https://doi.org/10.1182/blood-2009-07-234377

Haddad, R., Guimiot, F., Six, E., Jourquin, F., Setterblad, N., Kahn, E., … Canque, B. (2006). Dynamics of Thymus-Colonizing Cells during Human Development. *Immunity*, *24*(2), 217–230. https://doi.org/10.1016/J.IMMUNI.2006.01.008

Haghandish, N., Baldwin, R. M., Morettin, A., Dawit, H. T., Adhikary, H., Masson, J.-Y., … Côté, J. (2019). PRMT7 methylates eukaryotic translation initiation factor 2α and regulates its role in stress granule formation. *Molecular Biology of the Cell*, *30*(6), 778–793. https://doi.org/10.1091/mbc.E18-05-0330

Haider, Z., Larsson, P., Landfors, M., Köhn, L., Schmiegelow, K., Flægstad, T., … Degerman, S. (2019). An integrated transcriptome analysis in T-cell acute lymphoblastic leukemia links DNA methylation subgroups to dysregulated TAL1 and

ANTP homeobox gene expression. *Cancer Medicine*, *8*(1), 311–324. https://doi.org/10.1002/cam4.1917

Hänzelmann, S., Castelo, R., & Guinney, J. (2013). GSVA: Gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*, *14*, 7. https://doi.org/10.1186/1471-2105-14-7

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S., Butler, A., … Satija, R. (2021). Integrated analysis of multimodal single-cell data. *Cell*, *184*(13), 3573-3587.e29. https://doi.org/10.1016/j.cell.2021.04.048

Hardy, R. R., & Hayakawa, K. (2001). B cell development pathways. *Annual Review of Immunology*, *19*, 595–621. https://doi.org/10.1146/ANNUREV.IMMUNOL.19.1.595

Harrison, C. J., Moorman, A. V., Schwab, C., Carroll, A. J., Raetz, E. A., Devidas, M., … Haas, O. A. (2014). An international study of intrachromosomal amplification of chromosome 21 (iAMP21): Cytogenetic characterization and outcome. *Leukemia*, *28*(5), 1015–1021. https://doi.org/10.1038/leu.2013.317

Hata, K., Yanase, N., Sudo, K., Kiyonari, H., Mukumoto, Y., Mizuguchi, J., & Yokosuka, T. (2016). Differential regulation of T-cell dependent and T-cell independent antibody responses through arginine methyltransferase PRMT1 in vivo. *FEBS Letters*, *590*(8), 1200–1210. https://doi.org/10.1002/1873-3468.12161

Hauer, J., Fischer, U., & Borkhardt, A. (2021). Toward prevention of childhood ALL by early-life immune training. *Blood*, *138*(16), 1412–1428. https://doi.org/10.1182/blood.2020009895

Haynes, B. F., & Heinly, C. S. (1995). Early human T cell development: Analysis of the human thymus at the time of initial entry of hematopoietic stem cells into the fetal thymic microenvironment. *The Journal of Experimental Medicine*, *181*(4), 1445. https://doi.org/10.1084/JEM.181.4.1445

Heerema, N. A., Carroll, A. J., Devidas, M., Loh, M. L., Borowitz, M. J., Gastier-Foster, J. M., … Raetz, E. A. (2013). Intrachromosomal Amplification of Chromosome 21 Is Associated With Inferior Outcomes in Children With Acute Lymphoblastic Leukemia Treated in Contemporary Standard-Risk Children's Oncology Group Studies: A Report From the Children's Oncology Group. *Journal of Clinical Oncology*, *31*(27), 3397–3402. https://doi.org/10.1200/JCO.2013.49.1308

Hernandes, M. A. F., Marques-Salles, T. de J., Mkrtchyan, H., Soares-Ventura, E. M., Leite, E. P., Muniz, M. T. C., … Silva, M. L. M. (2012). Extra Copies of der(21)t(12;21) plus Deletion of ETV6 Gene due to dic(12;18) in B-Cell Precursor ALL with Poor Outcome. *Case Reports in Genetics*, *2012*, 186532. https://doi.org/10.1155/2012/186532

Herranz, D., Ambesi-Impiombato, A., Palomero, T., Schnell, S. A., Belver, L., Wendorff, A. A., … Ferrando, A. A. (2014). A NOTCH1-driven MYC enhancer promotes T cell development, transformation and acute lymphoblastic leukemia. *Nature Medicine*, *20*(10), 1130–1137. https://doi.org/10.1038/nm.3665

Herzog, S., Reth, M., & Jumaa, H. (2009). Regulation of B-cell proliferation and differentiation by pre-B-cell receptor signalling. *Nature Reviews. Immunology*, *9*(3), 195–205. https://doi.org/10.1038/nri2491

Hidalgo, I., Wahlestedt, M., Yuan, O., Zhang, Q., Bryder, D., & Pronk, C. J. (2022). Bmi1 induction protects hematopoietic stem cells against pronounced long-term hematopoietic stress. *Experimental Hematology*, *109*, 35–44. https://doi.org/10.1016/J.EXPHEM.2022.02.004

Hnisz, D., Weintrau, A. S., Day, D. S., Valton, A. L., Bak, R. O., Li, C. H., … Young, R. A. (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science (New York, N.Y.)*, *351*(6280), 1454–1458. https://doi.org/10.1126/SCIENCE.AAD9024

Ho, I.-C., Tai, T.-S., & Pai, S.-Y. (2009). GATA3 and the T-cell lineage: Essential functions before and after T-helper-2-cell differentiation. *Nature Reviews Immunology*, *9*(2), 125–135. https://doi.org/10.1038/nri2476

Ho, M. S. H., Medcalf, R. L., Livesey, S. A., & Traianedes, K. (2015). The dynamics of adult haematopoiesis in the bone and bone marrow environment. *British Journal of Haematology*, *170*(4), 472–486. https://doi.org/10.1111/BJH.13445

Ho, M.-C., Wilczek, C., Bonanno, J. B., Xing, L., Seznec, J., Matsui, T., … Shechter, D. (2013). Structure of the Arginine Methyltransferase PRMT5-MEP50 Reveals a Mechanism for Substrate Specificity. *PLOS ONE*, *8*(2), e57008. https://doi.org/10.1371/journal.pone.0057008

Hock, H., & Shimamura, A. (2017). ETV6 in hematopoiesis and leukemia predisposition. *Seminars in Hematology*, *54*(2), 98–104. https://doi.org/10.1053/j.seminhematol.2017.04.005

Holinstat, M. (2017). Normal platelet function. *Cancer Metastasis Reviews*, *36*(2), 195. https://doi.org/10.1007/S10555-017-9677-X

Horning, B. D., Suciu, R. M., Ghadiri, D. A., Ulanovskaya, O. A., Matthews, M. L., Lum, K. M., … Cravatt, B. F. (2016). Chemical Proteomic Profiling of Human Methyltransferases. *Journal of the American Chemical Society*, *138*(40), 13335–13343. https://doi.org/10.1021/jacs.6b07830

Hsu, W.-L., Preston, D. L., Soda, M., Sugiyama, H., Funamoto, S., Kodama, K., … Mabuchi, K. (2013). The Incidence of Leukemia, Lymphoma and Multiple Myeloma among Atomic Bomb Survivors: 1950–2001. *Radiation Research*, *179*(3), 361–382. https://doi.org/10.1667/RR2892.1

Hua, W.-K., Chang, Y.-I., Yao, C.-L., Hwang, S.-M., Chang, C.-Y., & Lin, W.-J. (2013). Protein Arginine Methyltransferase 1 Interacts with and Activates p38α to Facilitate Erythroid Differentiation. *PLoS ONE*, *8*(3), e56715. https://doi.org/10.1371/journal.pone.0056715

Huntington, N. D., Vosshenrich, C. A. J., & Di Santo, J. P. (2007). Developmental pathways that generate natural-killer-cell diversity in mice and humans. *Nature Reviews Immunology 2007 7:9*, *7*(9), 703–714. https://doi.org/10.1038/nri2154

Iacobucci, I., Kimura, S., & Mullighan, C. G. (2021). Biologic and Therapeutic Implications of Genomic Alterations in Acute Lymphoblastic Leukemia. *Journal of Clinical Medicine*, *10*(17), 3792. https://doi.org/10.3390/jcm10173792

Iacobucci, I., & Mullighan, C. G. (2022). KMT2A-rearranged leukemia: The shapeshifter. *Blood*, *140*(17), 1833–1835. https://doi.org/10.1182/blood.2022017645

Ichikawa, M., Asai, T., Chiba, S., Kurokawa, M., & Seishi Ogawa. (2004). Runx1/AML-1 ranks as a master regulator of adult hematopoiesis. *Cell Cycle (Georgetown, Tex.)*, *3*(6), 722–724.

Ichikawa, M., Asai, T., Saito, T., Yamamoto, G., Seo, S., Yamazaki, I., … Kurokawa, M. (2004). AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. *Nature Medicine*, *10*(3), 299–304. https://doi.org/10.1038/NM997

Igarashi, H., Gregory, S. C., Yokota, T., Sakaguchi, N., & Kincade, P. W. (2002). Transcription from the RAG1 locus marks the earliest lymphocyte progenitors in

bone marrow. *Immunity*, *17*(2), 117–130. https://doi.org/10.1016/S1074-7613(02)00366-7

Inaba, H., Greaves, M., & Mullighan, C. G. (2013). Acute lymphoblastic leukaemia. *The Lancet*, *381*(9881), 1943–1955. https://doi.org/10.1016/S0140-6736(12)62187-4

Inaba, H., & Mullighan, C. G. (2020). Pediatric acute lymphoblastic leukemia. *Haematologica*, *105*(11), 2524–2539. https://doi.org/10.3324/haematol.2020.247031

Inaba, H., & Pui, C.-H. (2021). Advances in the Diagnosis and Treatment of Pediatric Acute Lymphoblastic Leukemia. *Journal of Clinical Medicine*, *10*(9), 1926. https://doi.org/10.3390/jcm10091926

Inthal, A., Zeitlhofer, P., Zeginigg, M., Morak, M., Grausenburger, R., Fronkova, E., … Panzer-Grümayer, R. (2012). CREBBP HAT domain mutations prevail in relapse cases of high hyperdiploid childhood acute lymphoblastic leukemia. *Leukemia*, *26*(8), 1797–1803. https://doi.org/10.1038/leu.2012.60

Inthal, Andrea, Krapf, G., Beck, D., Joas, R., Kauer, M. O., Orel, L., … Panzer-Grümayer, E. R. (2008). Role of the Erythropoietin Receptor in ETV6/RUNX1-Positive Acute Lymphoblastic Leukemia. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *14*(22), 7196–7204. https://doi.org/10.1158/1078-0432.CCR-07-5051

Irving, J., Matheson, E., Minto, L., Blair, H., Case, M., Halsey, C., … Eckert, C. (2014). Ras pathway mutations are prevalent in relapsed childhood acute lymphoblastic leukemia and confer sensitivity to MEK inhibition. *Blood*, *124*(23), 3420–3430. https://doi.org/10.1182/blood-2014-04-531871

Islam, R., Jenkins, C. E., Cao, Q., Wong, J., Bilenky, M., Carles, A., … Hirst, M. (2023). RUNX1 colludes with NOTCH1 to reprogram chromatin in T cell acute lymphoblastic leukemia. *IScience*, *26*(6), 106795. https://doi.org/10.1016/j.isci.2023.106795

Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E. A., Iwasaki-Arai, J., Mizuno, S. I., … Akashi, K. (2005). Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood*, *106*(5), 1590. https://doi.org/10.1182/BLOOD-2005-03-0860

Jagannathan-Bogdan, M., & Zon, L. I. (2013). Hematopoiesis. *Development (Cambridge, England)*, *140*(12), 2463. https://doi.org/10.1242/DEV.083147

Janssen, J. W., Ludwig, W. D., Sterry, W., & Bartram, C. R. (1993). SIL-TAL1 deletion in T-cell acute lymphoblastic leukemia. *Leukemia*, *7*(8), 1204–1210.

Jean, D., Bernier, G., & Gruss, P. (1999). Six6 (Optx2) is a novel murine Six3-related homeobox gene that demarcates the presumptive pituitary/hypothalamic axis and the ventral optic stalk. *Mechanisms of Development*, *84*(1–2), 31–40. https://doi.org/10.1016/s0925-4773(99)00068-4

Jeha, S., Pei, D., Choi, J., Cheng, C., Sandlund, J. T., Coustan-Smith, E., … Pui, C.-H. (2019). Improved CNS Control of Childhood Acute Lymphoblastic Leukemia Without Cranial Irradiation: St Jude Total Therapy Study 16. *Journal of Clinical Oncology*, *37*(35), 3377–3391. https://doi.org/10.1200/JCO.19.01692

Jeha, S., Pei, D., Raimondi, S. C., Onciu, M., Campana, D., Cheng, C., … Pui, C.-H. (2009). Increased Risk for CNS Relapse in Pre-B Cell Leukemia with the t(1;19)/TCF3-PBX1. *Leukemia : Official Journal of the Leukemia Society of America, Leukemia Research Fund, U.K*, *23*(8), 1406–1409. https://doi.org/10.1038/leu.2009.42

Jeong, H.-J., Lee, H.-J., Vuong, T. A., Choi, K.-S., Choi, D., Koo, S.-H., … Kang, J.-S. (2016). Prmt7 Deficiency Causes Reduced Skeletal Muscle Oxidative Metabolism and Age-Related Obesity. *Diabetes*, *65*(7), 1868–1882. https://doi.org/10.2337/db15-1500

Jerchel, I. S., Hoogkamer, A. Q., Ariës, I. M., Steeghs, E. M. P., Boer, J. M., Besselink, N. J. M., … den Boer, M. L. (2018). RAS pathway mutations as a predictive biomarker for treatment adaptation in pediatric B-cell precursor acute lymphoblastic leukemia. *Leukemia*, *32*(4), 931–940. https://doi.org/10.1038/leu.2017.303

Jia, M., Zhao, H.-Z., Shen, H.-P., Cheng, Y.-P., Luo, Z.-B., Li, S.-S., … Tang, Y.-M. (2015). Overexpression of lymphoid enhancer-binding factor-1 (LEF1) is a novel favorable prognostic factor in childhood acute lymphoblastic leukemia. *International Journal of Laboratory Hematology*, *37*(5), 631–640. https://doi.org/10.1111/ijlh.12375

Julien, E., El Omar, R., & Tavian, M. (2016). Origin of the hematopoietic system in the human embryo. *FEBS Letters*, *590*(22), 3987–4001. https://doi.org/10.1002/1873-3468.12389

Juliusson, G., & Hough, R. (2016). Leukemia. *Tumors in Adolescents and Young Adults*, *43*, 87–100. https://doi.org/10.1159/000447076

Kampen, K. R. (2012). The discovery and early understanding of leukemia. *Leukemia Research*, *36*(1), 6–13. https://doi.org/10.1016/j.leukres.2011.09.028

Kampen, K. R., Sulima, S. O., Verbelen, B., Girardi, T., Vereecke, S., Rinaldi, G., … De Keersmaecker, K. (2019). The ribosomal RPL10 R98S mutation drives IRES-dependent BCL-2 translation in T-ALL. *Leukemia*, *33*(2), 319–332. https://doi.org/10.1038/s41375-018-0176-z

Kang, D.-H., Kim, S. H., Jun, J. W., Lee, Y.-W., Shin, H. B., Ahn, J. Y., … Jeon, B. R. (2012). Simultaneous Translocation of Both TCR Loci (14q11) with Rare Partner Loci (Xq22 and 12p13) in a Case of T-lymphoblastic Leukemia. *Annals of Laboratory Medicine*, *32*(3), 220–224. https://doi.org/10.3343/alm.2012.32.3.220

Karamitros, D., Stoilova, B., Aboukhalil, Z., Hamey, F., Reinisch, A., Samitsch, M., … Vyas, P. (2018). Heterogeneity of human lympho-myeloid progenitors at the single cell level. *Nature Immunology*, *19*(1), 85. https://doi.org/10.1038/S41590-017-0001-2

Karkhanis, V., Wang, L., Tae, S., Hu, Y.-J., Imbalzano, A. N., & Sif, S. (2012). Protein Arginine Methyltransferase 7 Regulates Cellular Response to DNA Damage by Methylating Promoter Histones H2A and H4 of the Polymerase δ Catalytic Subunit Gene, POLD1. *The Journal of Biological Chemistry*, *287*(35), 29801–29814. https://doi.org/10.1074/jbc.M112.378281

Katayama, Y., Battista, M., Kao, W. M., Hidalgo, A., Peired, A. J., Thomas, S. A., & Frenette, P. S. (2006). Signals from the Sympathetic Nervous System Regulate Hematopoietic Stem Cell Egress from Bone Marrow. *Cell*, *124*(2), 407–421. https://doi.org/10.1016/J.CELL.2005.10.041

Katz, J. E., Dlakić, M., & Clarke, S. (2003). Automated identification of putative methyltransferases from genomic open reading frames. *Molecular & Cellular Proteomics: MCP*, *2*(8), 525–540. https://doi.org/10.1074/mcp.M300037-MCP200

Kawagoe, H., Humphries, R. K., Blair, A., Sutherland, H. J., & Hogge, D. D. (1999). Expression of HOX genes, HOX cofactors, and MLL in phenotypically and functionally defined subpopulations of leukemic and normal human hematopoietic cells. *Leukemia*, *13*(5), 687–698. https://doi.org/10.1038/SJ.LEU.2401410

Kawakami, K., Sato, S., Ozaki, H., & Ikeda, K. (2000). Six family genes—Structure and function as transcription factors and their roles in development. *BioEssays*, *22*(7), 616–

626. https://doi.org/10.1002/1521-1878(200007)22:7<616::AID-BIES4>3.0.CO;2-R

Kendall, G., Little, M., & Wakeford, R. (2011). Numbers and proportions of leukemias in young people and adults induced by radiation of natural origin. *Leukemia Research*, *35*(8), 1039–1043. https://doi.org/10.1016/j.leukres.2011.01.023

Khanna, A., Bhushan, B., Chauhan, P. S., Saxena, S., Gupta, D. K., & Siraj, F. (2018). High mTOR expression independently prognosticates poor clinical outcome to induction chemotherapy in acute lymphoblastic leukemia. *Clinical and Experimental Medicine*, *18*(2), 221–227. https://doi.org/10.1007/s10238-017-0478-x

Kim, E., Hurtz, C., Koehrer, S., Wang, Z., Balasubramanian, S., Chang, B. Y., … Burger, J. A. (2017). Ibrutinib inhibits pre-BCR+ B-cell acute lymphoblastic leukemia progression by targeting BTK and BLK. *Blood*, *129*(9), 1155–1165. https://doi.org/10.1182/blood-2016-06-722900

Kindler, T., Cornejo, M. G., Scholl, C., Liu, J., Leeman, D. S., Haydu, J. E., … Gilliland, D. G. (2008). K-RasG12D–induced T-cell lymphoblastic lymphoma/leukemias harbor Notch1 mutations and are sensitive to γ-secretase inhibitors. *Blood*, *112*(8), 3373–3382. https://doi.org/10.1182/blood-2008-03-147587

Kinlen, L. (1988). Evidence for an infective cause of childhood leukaemia: Comparison of a Scottish new town with nuclear reprocessing sites in Britain. *Lancet (London, England)*, *2*(8624), 1323–1327. https://doi.org/10.1016/s0140-6736(88)90867-7

Kinlen, L. (2011). Childhood leukaemia, nuclear sites, and population mixing. *British Journal of Cancer*, *104*(1), 12–18. https://doi.org/10.1038/sj.bjc.6605982

Knight, T., & Irving, J. A. E. (2014). Ras/Raf/MEK/ERK Pathway Activation in Childhood Acute Lymphoblastic Leukemia and Its Therapeutic Targeting. *Frontiers in Oncology*, *4*, 160. https://doi.org/10.3389/fonc.2014.00160

Kobayashi, H., & Rowley, J. D. (1995). Identification of cytogenetically undetected 12p13 translocations and associated deletions with fluorescence in situ hybridization. *Genes, Chromosomes & Cancer*, *12*(1), 66–69. https://doi.org/10.1002/gcc.2870120112

Kobayashi, H., Satake, N., Maseki, N., Sakashita, A., & Kaneko, Y. (1996). The der(21)t(12;21) chromosome is always formed in a 12;21 translocation associated with childhood acute lymphoblastic leukaemia. *British Journal of Haematology*, *94*(1), 105–111. https://doi.org/10.1046/j.1365-2141.1996.d01-1762.x

Kobayashi, M., Toyama, R., Takeda, H., Dawid, I. B., & Kawakami, K. (1998). Overexpression of the forebrain-specific homeobox gene six3 induces rostral forebrain enlargement in zebrafish. *Development*, *125*(15), 2973–2982. https://doi.org/10.1242/dev.125.15.2973

Koch, U., & Radtke, F. (2011). Mechanisms of T Cell Development and Transformation. *Https://Doi.Org/10.1146/Annurev-Cellbio-092910-154008*, *27*, 539–562. https://doi.org/10.1146/ANNUREV-CELLBIO-092910-154008

Koh, G., Degasperi, A., Zou, X., Momen, S., & Nik-Zainal, S. (2021). Mutational signatures: Emerging concepts, caveats and clinical applications. *Nature Reviews Cancer*, *21*(10), 619–637. https://doi.org/10.1038/s41568-021-00377-7

Kollet, O., Dar, A., Shivtiel, S., Kalinkovich, A., Lapid, K., Sztainberg, Y., … Lapidot, T. (2006). Osteoclasts degrade endosteal components and promote mobilization of hematopoietic progenitor cells. *Nature Medicine 2006 12:6*, *12*(6), 657–664. https://doi.org/10.1038/nm1417

Konrad, M., Metzler, M., Panzer, S., Östreicher, I., Peham, M., Repp, R., … Panzer-Grümayer, E. R. (2003). Late relapses evolve from slow-responding subclones in

t(12;21)-positive acute lymphoblastic leukemia: Evidence for the persistence of a preleukemic clone. *Blood*, *101*(9), 3635–3640. https://doi.org/10.1182/blood-2002-10-3252

Kontro, M., Kuusanmäki, H., Eldfors, S., Burmeister, T., Andersson, E. I., Bruserud, Ø., … Heckman, C. A. (2014). Novel activating STAT5B mutations as putative drivers of T-cell acute lymphoblastic leukemia. *Leukemia*, *28*(8), 1738–1742. https://doi.org/10.1038/leu.2014.89

Kuhn, V., Diederich, L., Keller, T. C. S., Kramer, C. M., Lückstädt, W., Panknin, C., … Cortese-Krott, M. M. (2017). Red Blood Cell Function and Dysfunction: Redox Regulation, Nitric Oxide Metabolism, Anemia. *Antioxidants & Redox Signaling*, *26*(13), 718. https://doi.org/10.1089/ARS.2016.6954

Kumar, B. V., Connors, T., & Farber, D. L. (2018). Human T cell development, localization, and function throughout life. *Immunity*, *48*(2), 202–213. https://doi.org/10.1016/j.immuni.2018.01.007

Kumar, J. P. (2009). The sine oculis homeobox (SIX) family of transcription factors as regulators of development and disease. *Cellular and Molecular Life Sciences : CMLS*, *66*(4), 565–583. https://doi.org/10.1007/s00018-008-8335-4

Kuster, L., Grausenburger, R., Fuka, G., Kaindl, U., Krapf, G., Inthal, A., … Panzer-Grümayer, R. (2011). ETV6/RUNX1-positive relapses evolve from an ancestral clone and frequently acquire deletions of genes implicated in glucocorticoid signaling. *Blood*, *117*(9), 2658–2667. https://doi.org/10.1182/blood-2010-03-275347

Kwon, K., Hutter, C., Sun, Q., Bilic, I., Cobaleda, C., Malin, S., & Busslinger, M. (2008). Instructive Role of the Transcription Factor E2A in Early B Lymphopoiesis and Germinal Center B Cell Development. *Immunity*, *28*(6), 751–762. https://doi.org/10.1016/j.immuni.2008.04.014

Lai, Z., Markovets, A., Ahdesmaki, M., Chapman, B., Hofmann, O., McEwen, R., … Dry, J. R. (2016). VarDict: A novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Research*, *44*(11), e108. https://doi.org/10.1093/nar/gkw227

Larder, R., Clark, D. D., Miller, N. L. G., & Mellon, P. L. (2011). Hypothalamic dysregulation and infertility in mice lacking the homeodomain protein Six6. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(2), 426–438. https://doi.org/10.1523/JNEUROSCI.1688-10.2011

Larmonie, N. S. D., Dik, W. A., Meijerink, J. P. P., Homminga, I., van Dongen, J. J. M., & Langerak, A. W. (2013). Breakpoint sites disclose the role of the V(D)J recombination machinery in the formation of T-cell receptor (TCR) and non-TCR associated aberrations in T-cell acute lymphoblastic leukemia. *Haematologica*, *98*(8), 1173–1184. https://doi.org/10.3324/haematol.2012.082156

LaRosa, D. F., & Orange, J. S. (2008). 1. Lymphocytes. *Journal of Allergy and Clinical Immunology*, *121*(2), S364–S369. https://doi.org/10.1016/J.JACI.2007.06.016

Latchman, D. S. (1993). Transcription factors: An overview. *International Journal of Experimental Pathology*, *74*(5), 417.

Latonen, L., Afyounian, E., Jylhä, A., Nättinen, J., Aapola, U., Annala, M., … Visakorpi, T. (2018). Integrative proteomics in prostate cancer uncovers robustness against genomic and transcriptomic aberrations during disease progression. *Nature Communications*, *9*(1), 1176. https://doi.org/10.1038/s41467-018-03573-6

Lawson, S. E., Harrison, G., Richards, S., Oakhill, A., Stevens, R., Eden, O. B., & Darbyshire, P. J. (2000). The UK experience in treating relapsed childhood acute lymphoblastic leukaemia: A report on the medical research council UKALLR1 study. *British Journal of Haematology*, *108*(3), 531–543. https://doi.org/10.1046/j.1365-2141.2000.01891.x

LeBien, T. W., & Tedder, T. F. (2008). B lymphocytes: How they develop and function. *Blood*, *112*(5), 1570–1580. https://doi.org/10.1182/blood-2008-02-078071

Lee, B., Rizzoti, K., Kwon, D. S., Kim, S.-Y., Oh, S., Epstein, D. J., … Jeong, Y. (2012). Direct transcriptional regulation of Six6 is controlled by SoxB1 binding to a remote forebrain enhancer. *Developmental Biology*, *366*(2), 393–403. https://doi.org/10.1016/j.ydbio.2012.04.023

Leone, G., DeGregori, J., Zhen, Y., Laszlo, J., Seiichi, I., Sander Williams, R., & Nevins, J. R. (1998). E2F3 activity is regulated during the cell cycle and is required for the induction of S phase—PubMed. Retrieved July 31, 2023, from https://pubmed.ncbi.nlm.nih.gov/9679057/

Li, B., Brady, S. W., Ma, X., Shen, S., Zhang, Y., Li, Y., … Zhang, J. (2020). Therapy-induced mutations drive the genomic landscape of relapsed acute lymphoblastic leukemia. *Blood*, *135*(1), 41–55. https://doi.org/10.1182/BLOOD.2019002220

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, W. juan, He, Y. hui, Yang, J. jing, Hu, G. sheng, Lin, Y. an, Ran, T., … Liu, W. (2021). Profiling PRMT methylome reveals roles of hnRNPA1 arginine methylation in RNA splicing and cell growth. *Nature Communications*, *12*(1). https://doi.org/10.1038/S41467-021-21963-1

Li, Xiaoyu, & von Boehmer, H. (2011). Notch Signaling in T-Cell Development and T-ALL. *ISRN Hematology*, *2011*, 921706. https://doi.org/10.5402/2011/921706

Li, Xue, Perissi, V., Liu, F., Rose, D. W., & Rosenfeld, M. G. (2002). Tissue-Specific Regulation of Retinal and Pituitary Precursor Cell Proliferation. *Science*, *297*(5584), 1180–1183. https://doi.org/10.1126/science.1073263

Lilljebjörn, H., & Fioretos, T. (2017). New oncogenic subtypes in pediatric B-cell precursor acute lymphoblastic leukemia. *Blood*, *130*(12), 1395–1401. https://doi.org/10.1182/blood-2017-05-742643

Lilljebjörn, H., Henningsson, R., Hyrenius-Wittsten, A., Olsson, L., Orsmark-Pietras, C., von Palffy, S., … Fioretos, T. (2016). Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. *Nature Communications*, *7*, 11790. https://doi.org/10.1038/ncomms11790

Lilljebjörn, H., Soneson, C., Andersson, A., Heldrup, J., Behrendtz, M., Kawamata, N., … Fioretos, T. (2010). The correlation pattern of acquired copy number changes in 164 ETV6/RUNX1-positive childhood acute lymphoblastic leukemias. *Human Molecular Genetics*, *19*(16), 3150–3158. https://doi.org/10.1093/hmg/ddq224

Liu, F., Cheng, G., Hamard, P.-J., Greenblatt, S., Wang, L., Man, N., … Nimer, S. D. (2015). Arginine methyltransferase PRMT5 is essential for sustaining normal adult hematopoiesis. *The Journal of Clinical Investigation*, *125*(9), 3532–3544. https://doi.org/10.1172/JCI81749

Liu F, L, W., H, Z., Z, P., W, Z., & X, L. (2020). PRMT7 promotes the growth of renal cell carcinoma through modulating the β-catenin/C-MYC axis. *The International Journal of Biochemistry & Cell Biology*, *120*. https://doi.org/10.1016/J.BIOCEL.2020.105686

Liu H, Wang JY, Huang Y, Li Z, Gong W, Lehmann R, & Xu RM. (2010). Structural basis for methylarginine-dependent recognition of Aubergine by Tudor. *Genes & Development*, *24*(17), 1876–1881. https://doi.org/10.1101/GAD.1956010

Liu, Q., Li, A., Tian, Y., Liu, Y., Li, T., Zhang, C., … Wu, K. (2016). The expression profile and clinic significance of the SIX family in non-small cell lung cancer. *Journal of Hematology & Oncology*, *9*(1), 119. https://doi.org/10.1186/s13045-016-0339-1

Liu R, Gao J, Yang Y, Qiu R, Zheng Y, W, H., … Y, W. (2018). PHD finger protein 1 (PHF1) is a novel reader for histone H4R3 symmetric dimethylation and coordinates with PRMT5-WDR77/CRL4B complex to promote tumorigenesis. *Nucleic Acids Research*, *46*(13), 6608–6626. https://doi.org/10.1093/NAR/GKY461

Liu Y, Easton J, Shao Y, Maciaszek J, Wang Z, MR, W., … CG, M. (2017). The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nature Genetics*, *49*(8), 1211–1218. https://doi.org/10.1038/NG.3909

Liu, Y.-F., Wang, B.-Y., Zhang, W.-N., Huang, J.-Y., Li, B.-S., Zhang, M., … Chen, S.-J. (2016). Genomic Profiling of Adult and Pediatric B-cell Acute Lymphoblastic Leukemia. *EBioMedicine*, *8*, 173–183. https://doi.org/10.1016/j.ebiom.2016.04.038

Liu, Z., Wu, Y., Zhang, Y., Yuan, M., Li, X., Gao, J., … Zhao, Z. (2019). TIGAR Promotes Tumorigenesis and Protects Tumor Cells From Oxidative and Metabolic Stresses in Gastric Cancer. *Frontiers in Oncology*, *9*, 1258. https://doi.org/10.3389/fonc.2019.01258

Locatelli, F., Zugmaier, G., Rizzari, C., Morris, J. D., Gruhn, B., Klingebiel, T., … von Stackelberg, A. (2021). Effect of Blinatumomab vs Chemotherapy on Event-Free Survival Among Children With High-risk First-Relapse B-Cell Acute Lymphoblastic Leukemia. *JAMA*, *325*(9), 843–854. https://doi.org/10.1001/jama.2021.0987

Loeb, D. M., & Sukumar, S. (2002). The role of WT1 in oncogenesis: Tumor suppressor or oncogene? *International Journal of Hematology*, *76*(2), 117–126. https://doi.org/10.1007/BF02982573

Loh, M. L., Goldwasser, M. A., Silverman, L. B., Poon, W.-M., Vattikuti, S., Cardoso, A., … Gilliland, D. G. (2006). Prospective analysis of TEL/AML1-positive patients treated on Dana-Farber Cancer Institute Consortium Protocol 95-01. *Blood*, *107*(11), 4508–4513. https://doi.org/10.1182/blood-2005-08-3451

Lohi, O., Kanerva, J., Taskinen, M., Harila-Saari, A., Rounioja, S., Jahnukainen, K., … Vettenranta, K. (2013). [Childhood leukemia]. *Duodecim; Laaketieteellinen Aikakauskirja*, *129*(9), 939–946.

Loncarevic, I. F., Roitzheim, B., Ritterbach, J., Viehmann, S., Borkhardt, A., Lampert, F., & Harbott, J. (1999). Trisomy 21 is a recurrent secondary aberration in childhood acute lymphoblastic leukemia with TEL/AML1 gene fusion. *Genes, Chromosomes & Cancer*, *24*(3), 272–277.

Look, A. T. (1997). Oncogenic transcription factors in the human acute leukemias. *Science (New York, N.Y.)*, *278*(5340), 1059–1064. https://doi.org/10.1126/science.278.5340.1059

Lopez, R. G., Carron, C., Oury, C., Gardellin, P., Bernard, O., & Ghysdael, J. (1999). TEL Is a Sequence-specific Transcriptional Repressor *. *Journal of Biological Chemistry*, *274*(42), 30132–30138. https://doi.org/10.1074/jbc.274.42.30132

López-Ríos, J., Gallardo, M. E., Rodriguez de Córdoba, S., & Bovolenta, P. (1999). Six9 (Optx2), a new member of the six gene family of transcription factors, is expressed at early stages of vertebrate ocular and pituitary development. *Mechanisms of Development*, *83*(1–2), 155–159. https://doi.org/10.1016/s0925-4773(99)00017-9

Lucic, B., Chen, H.-C., Kuzman, M., Zorita, E., Wegner, J., Minneker, V., … Lusic, M. (2019). Spatially clustered loci with multiple enhancers are frequent targets of HIV-1 integration. *Nature Communications*, *10*(1), 4059. https://doi.org/10.1038/s41467-019-12046-3

Ma, H., Zhou, H., Song, X., Shi, S., Zhang, J., & Jia, L. (2015). Modification of sialylation is associated with multidrug resistance in human acute myeloid leukemia. *Oncogene*, *34*(6), 726–740. https://doi.org/10.1038/onc.2014.7

Ma, X., Qi, W., Pan, H., Yang, F., & Deng, J. (2018). Overexpression of USP5 contributes to tumorigenesis in non-small cell lung cancer via the stabilization of β-catenin protein. *American Journal of Cancer Research*, *8*(11), 2284–2295.

Madzo, J., Zuna, J., Muzíková, K., Kalinová, M., Krejcí, O., Hrusák, O., … Trka, J. (2003). Slower molecular response to treatment predicts poor outcome in patients with TEL/AML1 positive acute lymphoblastic leukemia: Prospective real-time quantitative reverse transcriptase-polymerase chain reaction study. *Cancer*, *97*(1), 105–113. https://doi.org/10.1002/cncr.11043

Malard, F., & Mohty, M. (2020). Acute lymphoblastic leukaemia. *The Lancet*, *395*(10230), 1146–1162. https://doi.org/10.1016/S0140-6736(19)33018-1

Mansour, M. R., Abraham, B. J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A. D., … Look, A. T. (2014). Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science (New York, N.Y.)*, *346*(6215), 1373–1377. https://doi.org/10.1126/science.1259037

Mariuzza, R. A., Agnihotri, P., & Orban, J. (2020). The structural basis of T-cell receptor (TCR) activation: An enduring enigma. *The Journal of Biological Chemistry*, *295*(4), 914–925. https://doi.org/10.1074/jbc.REV119.009411

Matsumura T, Nakamura-Ishizu A, Muddineni SSNA, Tan DQ, Wang CQ, Tokunaga K, … Suda T. (2020). Hematopoietic stem cells acquire survival advantage by loss of RUNX1 methylation identified in familial leukemia. *Blood*, *136*(17), 1919–1932. https://doi.org/10.1182/BLOOD.2019004292

McLean, T. W., Ringold, S., Neuberg, D., Stegmaier, K., Tantravahi, R., Ritz, J., … Golub, T. R. (1996). TEL/AML-1 dimerizes and is associated with a favorable outcome in childhood acute lymphoblastic leukemia. *Blood*, *88*(11), 4252–4258.

Mehtonen J, Pölönen P, Häyrynen S, Dufva O, Lin J, Liuksiala T, … Heinäniemi M. (2019). Data-driven characterization of molecular phenotypes across heterogeneous sample collections. *Nucleic Acids Research*, *47*(13). https://doi.org/10.1093/NAR/GKZ281

Mehtonen, J., Teppo, S., Lahnalampi, M., Kokko, A., Kaukonen, R., Oksa, L., … Heinäniemi, M. (2020). Single cell characterization of B-lymphoid differentiation and leukemic cell states during chemotherapy in ETV6-RUNX1-positive pediatric leukemia identifies drug-targetable transcription factor activities. *Genome Medicine*, *12*(1), 99. https://doi.org/10.1186/s13073-020-00799-2

Meldrum, C., Doyle, M. A., & Tothill, R. W. (2011). Next-Generation Sequencing for Cancer Diagnostics: A Practical Perspective. *The Clinical Biochemist Reviews*, *32*(4), 177–195.

Messina, M., Chiaretti, S., Wang, J., Fedullo, A. L., Peragine, N., Gianfelici, V., … Foà, R. (2016). Prognostic and therapeutic role of targetable lesions in B-lineage acute lymphoblastic leukemia without recurrent fusion genes. *Oncotarget*, *7*(12), 13886–13901. https://doi.org/10.18632/oncotarget.7356

Meurer, L., Ferdman, L., Belcher, B., & Camarata, T. (2021). The SIX Family of Transcription Factors: Common Themes Integrating Developmental and Cancer

Biology. *Frontiers in Cell and Developmental Biology*, *9*. Retrieved from https://www.frontiersin.org/articles/10.3389/fcell.2021.707854

Mikkola, H. K. A., & Orkin, S. H. (2006). The journey of developing hematopoietic stem cells. *Development*, *133*(19), 3733–3744. https://doi.org/10.1242/DEV.02568

Miranda, T., Miranda M, Frankel A, & Clarke S. (2004). PRMT7 is a member of the protein arginine methyltransferase family with a distinct substrate specificity. *The Journal of Biological Chemistry*, *279*(22), 22902–22907. https://doi.org/10.1074/JBC.M312904200

Miyoshi, H., Shimizu, K., Kozu, T., Maseki, N., Kaneko, Y., & Ohki, M. (1991). T(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(23), 10431–10434. https://doi.org/10.1073/PNAS.88.23.10431

Mizutani, S., Yoshida, T., Zhao, X., Nimer, S. D., Taniwaki, M., & Okuda, T. (2015). *Loss of RUNX1/AML1 arginine-methylation impairs peripheral T cell homeostasis*. https://doi.org/10.1111/bjh.13499

Mori, H., Colman, S. M., Xiao, Z., Ford, A. M., Healy, L. E., Donaldson, C., … Greaves, M. (2002). Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(12), 8242–8247. https://doi.org/10.1073/pnas.112218799

Morrow, M., Horton, S., Kioussis, D., Brady, H. J. M., & Williams, O. (2004). TEL-AML1 promotes development of specific hematopoietic lineages consistent with preleukemic activity. *Blood*, *103*(10), 3890–3896. https://doi.org/10.1182/blood-2003-10-3695

Müller-Sieburg, C. E., Cho, R. H., Karlsson, L., Huang, J. F., & Sieburg, H. B. (2004). Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood*, *103*(11), 4111–4118. https://doi.org/10.1182/BLOOD-2003-10-3448

Müller-Sieburg, C. E., Cho, R. H., Thoman, M., Adkins, B., & Sieburg, H. B. (2002). Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. *Blood*, *100*(4), 1302–1309.

Mullighan, C. G. (2012). Molecular genetics of B-precursor acute lymphoblastic leukemia. *The Journal of Clinical Investigation*, *122*(10), 3407–3415. https://doi.org/10.1172/JCI61203

Mullighan, C. G., Goorha, S., Radtke, I., Miller, C. B., Coustan-Smith, E., Dalton, J. D., … Downing, J. R. (2007). Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*, *446*(7137), 758–764. https://doi.org/10.1038/nature05690

Mullighan, C. G., Jeha, S., Pei, D., Payne-Turner, D., Coustan-Smith, E., Roberts, K. G., … Pui, C.-H. (2015). Outcome of children with hypodiploid ALL treated with risk-directed therapy based on MRD levels. *Blood*, *126*(26), 2896–2899. https://doi.org/10.1182/blood-2015-09-671131

Mullighan, C. G., Su, X., Zhang, J., Radtke, I., Phillips, L. A. A., Miller, C. B., … Downing, J. R. (2009). Deletion of IKZF1 and Prognosis in Acute Lymphoblastic Leukemia. *New England Journal of Medicine*, *360*(5), 470–480. https://doi.org/10.1056/NEJMoa0808253

Müschen, M. (2018). Autoimmunity checkpoints as therapeutic targets in B cell malignancies. *Nature Reviews. Cancer*, *18*(2), 103–116. https://doi.org/10.1038/nrc.2017.111

Nachman, J. B., Heerema, N. A., Sather, H., Camitta, B., Forestier, E., Harrison, C. J., … Janka-Schaub, G. E. (2007). Outcome of treatment in children with hypodiploid acute lymphoblastic leukemia. *Blood*, *110*(4), 1112–1115. https://doi.org/10.1182/blood-2006-07-038299

Nadeu, F., Mas-de-les-Valls, R., Navarro, A., Royo, R., Martín, S., Villamor, N., … Campo, E. (2020). IgCaller for reconstructing immunoglobulin gene rearrangements and oncogenic translocations from whole-genome sequencing in lymphoid neoplasms. *Nature Communications*, *11*, 3390. https://doi.org/10.1038/s41467-020-17095-7

Nagel, S., MacLeod, R. A. F., Pommerenke, C., Meyer, C., Kaufmann, M., & Drexler, H. G. (2018). NKL homeobox gene NKX2-2 is aberrantly expressed in Hodgkin lymphoma. *Oncotarget*, *9*(101), 37480–37496. https://doi.org/10.18632/oncotarget.26459

Nagel, S., Pommerenke, C., Scherr, M., Meyer, C., Kaufmann, M., Battmer, K., … Drexler, H. G. (2017). NKL homeobox gene activities in hematopoietic stem cells, T-cell development and T-cell leukemia. *PLOS ONE*, *12*(2), e0171164. https://doi.org/10.1371/journal.pone.0171164

Navarro, J.-M., Touzart, A., Pradel, L. C., Loosveld, M., Koubi, M., Fenouil, R., … Nadel, B. (2015). Site- and allele-specific polycomb dysregulation in T-cell leukaemia. *Nature Communications*, *6*, 6094. https://doi.org/10.1038/ncomms7094

Neveu, B., Caron, M., Lagacé, K., Richer, C., & Sinnett, D. (2018). Genome wide mapping of ETV6 binding sites in pre-B leukemic cells. *Scientific Reports*, *8*(1), 15526. https://doi.org/10.1038/s41598-018-33947-1

Ng, J. C. F., Quist, J., Grigoriadis, A., Malim, M. H., & Fraternali, F. (2019). Pan-cancer transcriptomic analysis dissects immune and proliferative functions of APOBEC3 cytidine deaminases. *Nucleic Acids Research*, *47*(3), 1178–1194. https://doi.org/10.1093/nar/gky1316

Ng, S. Y. M., Yoshida, T., Zhang, J., & Georgopoulos, K. (2009). Genome-wide lineage-specific transcriptional networks underscore Ikaros-dependent lymphoid priming in hematopoietic stem cells. *Immunity*, *30*(4), 493–507. https://doi.org/10.1016/J.IMMUNI.2009.01.014

Nikkilä, A., Erme, S., Arvela, H., Holmgren, O., Raitanen, J., Lohi, O., & Auvinen, A. (2016). Background radiation and childhood leukemia: A nationwide register-based case-control study. *International Journal of Cancer*, *139*(9), 1975–1982. https://doi.org/10.1002/ijc.30264

Nishii, R., Baskin-Doerfler, R., Yang, W., Oak, N., Zhao, X., Yang, W., … Yang, J. J. (2021). Molecular basis of ETV6-mediated predisposition to childhood acute lymphoblastic leukemia. *Blood*, *137*(3), 364–373. https://doi.org/10.1182/blood.2020006164

Nordlund, J., Bäcklin, C. L., Zachariadis, V., Cavelier, L., Dahlberg, J., Öfverholm, I., … Syvänen, A.-C. (2015). DNA methylation-based subtype prediction for pediatric acute lymphoblastic leukemia. *Clinical Epigenetics*, *7*(1), 11. https://doi.org/10.1186/s13148-014-0039-z

Norén-Nyström, U., Andersen, M. K., Barbany, G., Dirse, V., Eilert-Olsen, M., Engvall, M., … Johansson, B. (2023). Genetic Subtypes and Outcome of Patients Aged 1 to 45 Years Old With Acute Lymphoblastic Leukemia in the NOPHO ALL2008 Trial. *HemaSphere*, *7*(5), e883. https://doi.org/10.1097/HS9.0000000000000883

O'Connor, D., Enshaei, A., Bartram, J., Hancock, J., Harrison, C. J., Hough, R., … Goulden, N. (2018). Genotype-Specific Minimal Residual Disease Interpretation Improves

Stratification in Pediatric Acute Lymphoblastic Leukemia. *Journal of Clinical Oncology*, *36*(1), 34–43. https://doi.org/10.1200/JCO.2017.74.0449

Okuda, T., Van Deursen, J., Hiebert, S. W., Grosveld, G., & Downing, J. R. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell*, *84*(2), 321–330. https://doi.org/10.1016/S0092-8674(00)80986-1

Oliver, G., Mailhos, A., Wehr, R., Copeland, N. G., Jenkins, N. A., & Gruss, P. (1995). Six3, a murine homologue of the sine oculis gene, demarcates the most anterior border of the developing neural plate and is expressed during eye development. *Development*, *121*(12), 4045–4055. https://doi.org/10.1242/dev.121.12.4045

Orkin, S. H., & Zon, L. I. (2008). Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell*, *132*(4), 631–644. https://doi.org/10.1016/J.CELL.2008.01.025

Oshima, K., Khiabanian, H., da Silva-Almeida, A. C., Tzoneva, G., Abate, F., Ambesi-Impiombato, A., … Ferrando, A. A. (2016). Mutational landscape, clonal evolution patterns, and role of RAS mutations in relapsed acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(40), 11306–11311. https://doi.org/10.1073/pnas.1608420113

Oskarsson, T., Söderhäll, S., Arvidson, J., Forestier, E., Montgomery, S., Bottai, M., … Heyman, M. (2016). Relapsed childhood acute lymphoblastic leukemia in the Nordic countries: Prognostic factors, treatment and outcome. *Haematologica*, *101*(1), 68–76. https://doi.org/10.3324/haematol.2015.131680

Paik, W. K., & Kim, S. (1971). Protein Methylation. *New Series*, *174*(4005), 114–119.

Palii, C. G., Perez-Iratxeta, C., Yao, Z., Cao, Y., Dai, F., Davison, J., … Brand, M. (2011). Differential genomic targeting of the transcription factor TAL1 in alternate haematopoietic lineages. *The EMBO Journal*, *30*(3), 494–509. https://doi.org/10.1038/emboj.2010.342

Palis, J., & Yoder, M. C. (2001). Yolk-sac hematopoiesis: The first blood cells of mouse and man. *Experimental Hematology*, *29*(8), 927–936. https://doi.org/10.1016/s0301-472x(01)00669-5

Palomero, T., Lim, W. K., Odom, D. T., Sulis, M. L., Real, P. J., Margolin, A., … Ferrando, A. A. (2006). NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(48), 18261–18266. https://doi.org/10.1073/pnas.0606108103

Pang, S. H. M., de Graaf, C. A., Hilton, D. J., Huntington, N. D., Carotta, S., Wu, L., & Nutt, S. L. (2018). PU.1 is required for the developmental progression of multipotent progenitors to common lymphoid progenitors. *Frontiers in Immunology*, *9*(JUN), 11. https://doi.org/10.3389/FIMMU.2018.01264/FULL

Papaemmanuil, E., Rapado, I., Li, Y., Potter, N. E., Wedge, D. C., Tubio, J., … Campbell, P. J. (2014). RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nature Genetics*, *46*(2), 116–125. https://doi.org/10.1038/ng.2874

Park, I. K., Qian, D., Kiel, M., Becker, M. W., Pihalja, M., Weissman, I. L., … Clarke, M. F. (2003). Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature*, *423*(6937), 302–305. https://doi.org/10.1038/NATURE01587

Park, J., Botting RA, Domínguez Conde C, Popescu DM, Lavaert M, DJ, K., … SA, T. (2020). A cell atlas of human thymic development defines T cell repertoire formation. *Science (New York, N.Y.)*, *367*(6480). https://doi.org/10.1126/SCIENCE.AAY3224

Patel, S. R., Hartwig, J. H., & Italiano, J. E. (2005). The biogenesis of platelets from megakaryocyte proplatelets. *Journal of Clinical Investigation*, *115*(12), 3348. https://doi.org/10.1172/JCI26891

Patrick, K., Wade, R., Goulden, N., Mitchell, C., Moorman, A. V., Rowntree, C., … Vora, A. (2014). Outcome for children and young people with Early T-cell precursor acute lymphoblastic leukaemia treated on a contemporary protocol, UKALL 2003. *British Journal of Haematology*, *166*(3), 421–424. https://doi.org/10.1111/bjh.12882

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419. https://doi.org/10.1038/nmeth.4197

Paulsson, K., Horvat, A., Strömbeck, B., Nilsson, F., Heldrup, J., Behrendtz, M., … Johansson, B. (2008). Mutations of FLT3, NRAS, KRAS, and PTPN11 are frequent and possibly mutually exclusive in high hyperdiploid childhood acute lymphoblastic leukemia. *Genes, Chromosomes & Cancer*, *47*(1), 26–33. https://doi.org/10.1002/gcc.20502

Paulsson, K., & Johansson, B. (2009). High hyperdiploid childhood acute lymphoblastic leukemia. *Genes, Chromosomes & Cancer*, *48*(8), 637–660. https://doi.org/10.1002/gcc.20671

Peter, A., Heiden, T., Taube, T., Körner, G., & Seeger, K. (2009). Interphase FISH on TEL/AML1 positive acute lymphoblastic leukemia relapses—Analysis of clinical relevance of additional TEL and AML1 copy number changes. *European Journal of Haematology*, *83*(5), 420–432. https://doi.org/10.1111/j.1600-0609.2009.01315.x

Petrossian, T. C., & Clarke, S. G. (2011). Uncovering the Human Methyltransferasome. *Molecular & Cellular Proteomics : MCP*, *10*(1). https://doi.org/10.1074/MCP.M110.000976

Pieters, R., De Lorenzo, P., Ancliffe, P., Aversa, L. A., Brethon, B., Biondi, A., … Valsecchi, M. G. (2019). 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. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *37*(25), 2246–2256. https://doi.org/10.1200/JCO.19.00261

Pietras, E. M., Reynaud, D., Kang, Y. A., Carlin, D., Calero-Nieto, F. J., Leavitt, A. D., … Passegué, E. (2015). Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell*, *17*(1), 35–46. https://doi.org/10.1016/J.STEM.2015.05.003

Piette, C., Suciu, S., Clappier, E., Bertrand, Y., Drunat, S., Girard, S., … Cavé, H. (2018). Differential impact of drugs on the outcome of ETV6-RUNX1 positive childhood B-cell precursor acute lymphoblastic leukaemia: Results of the EORTC CLG 58881 and 58951 trials. *Leukemia*, *32*(1), 244–248. https://doi.org/10.1038/leu.2017.289

Piller, G. J. (2001). Leukaemia – a brief historical review from ancient times to 1950. *British Journal of Haematology*, *112*(2), 282–292. https://doi.org/10.1046/j.1365-2141.2001.02411.x

Pine, S. R., Wiemels, J. L., Jayabose, S., & Sandoval, C. (2003). TEL-AML1 fusion precedes differentiation to pre-B cells in childhood acute lymphoblastic leukemia. *Leukemia Research*, *27*(2), 155–164. https://doi.org/10.1016/s0145-2126(02)00183-2

Pinho, S., & Frenette, P. S. (2019). Haematopoietic stem cell activity and interactions with the niche. *Nature Reviews Molecular Cell Biology 2019 20:5*, *20*(5), 303–320. https://doi.org/10.1038/s41580-019-0103-9

Pitkäniemi, J. (2020). *Cancer in Finland*.

Pleasance, E. D., Cheetham, R. K., Stephens, P. J., McBride, D. J., Humphray, S. J., Greenman, C. D., … Stratton, M. R. (2010). A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature*, *463*(7278), 191–196. https://doi.org/10.1038/nature08658

Polak, P., Kim, J., Braunstein, L. Z., Tiao, G., Karlic, R., Rosebrock, D., … Getz, G. (2017). A mutational signature reveals alterations underlying deficient homologous recombination repair in breast cancer. *Nature Genetics*, *49*(10), 1476–1486. https://doi.org/10.1038/ng.3934

Pölönen P, Mehtonen J, Lin J, Liuksiala T, Häyrynen S, Teppo S, … Heinäniemi M. (2019). Hemap: An Interactive Online Resource for Characterizing Molecular Phenotypes across Hematologic Malignancies. *Cancer Research*, *79*(10), 2466–2479. https://doi.org/10.1158/0008-5472.CAN-18-2970

Prior, I. A., Lewis, P. D., & Mattos, C. (2012). A Comprehensive Survey of Ras Mutations in Cancer. *Cancer Research*, *72*(10), 2457–2467. https://doi.org/10.1158/0008-5472.CAN-11-2612

Pui, C.-H., Campana, D., Pei, D., Bowman, W. P., Sandlund, J. T., Kaste, S. C., … Relling, M. V. (2009). Treating Childhood Acute Lymphoblastic Leukemia without Cranial Irradiation. *New England Journal of Medicine*, *360*(26), 2730–2741. https://doi.org/10.1056/NEJMoa0900386

Pui, C.-H., & Evans, W. E. (2006). Treatment of acute lymphoblastic leukemia. *The New England Journal of Medicine*, *354*(2), 166–178. https://doi.org/10.1056/NEJMra052603

Pui, C.-H., Relling, M. V., & Downing, J. R. (2004). Acute lymphoblastic leukemia. *The New England Journal of Medicine*, *350*(15), 1535–1548. https://doi.org/10.1056/NEJMra023001

Pui, C.-H., Robison, L. L., & Look, A. T. (2008). Acute lymphoblastic leukaemia. *The Lancet*, *371*(9617), 1030–1043. https://doi.org/10.1016/S0140-6736(08)60457-2

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Quist-Paulsen, P., Toft N, Heyman M, Abrahamsson J, Griškevičius L, Hallböök H, … Schmiegelow K. (2020). T-cell acute lymphoblastic leukemia in patients 1-45 years treated with the pediatric NOPHO ALL2008 protocol. *Leukemia*, *34*(2), 347–357. https://doi.org/10.1038/S41375-019-0598-2

Raine, K. M., Van Loo, P., Wedge, D. C., Jones, D., Menzies, A., Butler, A. P., … Campbell, P. J. (2016). ascatNgs: Identifying Somatically Acquired Copy-Number Alterations from Whole-Genome Sequencing Data. *Current Protocols in Bioinformatics*, *56*, 15.9.1-15.9.17. https://doi.org/10.1002/cpbi.17

Rao, S., Cai, K. Q., Stadanlick, J. E., Greenberg-Kushnir, N., Solanki-Patel, N., Lee, S.-Y., … Wiest, D. L. (2016). Ribosomal Protein Rpl22 Controls the Dissemination of T-cell Lymphoma. *Cancer Research*, *76*(11), 3387–3396. https://doi.org/10.1158/0008-5472.CAN-15-2698

Rausch, T., Zichner, T., Schlattl, A., Stütz, A. M., Benes, V., & Korbel, J. O. (2012). DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*, *28*(18), i333–i339. https://doi.org/10.1093/bioinformatics/bts378

Reya, T., Duncan, A. W., Ailles, L., Domen, J., Scherer, D. C., Willert, K., … Weissman, I. L. (2003). A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature 2003 423:6938*, *423*(6938), 409–414. https://doi.org/10.1038/nature01593

Reya, T., O'Riordan, M., Okamura, R., Devaney, E., Willert, K., Nusse, R., & Grosschedl, R. (2000). Wnt Signaling Regulates B Lymphocyte Proliferation through a LEF-1 Dependent Mechanism. *Immunity*, *13*(1), 15–24. https://doi.org/10.1016/S1074-7613(00)00004-2

Ribeiro, D., Melão, A., & Barata, J. T. (2013). IL-7R-mediated signaling in T-cell acute lymphoblastic leukemia. *Advances in Biological Regulation*, *53*(2), 211–222. https://doi.org/10.1016/j.jbior.2012.10.005

Ribeiro, R. C., Abromowitch, M., Raimondi, S. C., Murphy, S. B., Behm, F., & Williams, D. L. (1987). Clinical and biologic hallmarks of the Philadelphia chromosome in childhood acute lymphoblastic leukemia. *Blood*, *70*(4), 948–953.

Rieger, M. A., & Schroeder, T. (2012). Hematopoiesis. *Cold Spring Harbor Perspectives in Biology*, *4*(12). https://doi.org/10.1101/CSHPERSPECT.A008250

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47. https://doi.org/10.1093/nar/gkv007

Rivera, G. K., Raimondi, S. C., Hancock, M. L., Behm, F. G., Pui, C. H., Abromowitch, M., … Williams, D. L. (1991). Improved outcome in childhood acute lymphoblastic leukaemia with reinforced early treatment and rotational combination chemotherapy. *Lancet (London, England)*, *337*(8733), 61–66. https://doi.org/10.1016/0140-6736(91)90733-6

Robb, L., Lyons, I., Li, R., Hartley, L., Köntgen, F., Harvey, R. P., … Begley, C. G. (1995). Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(15), 7075–7079. https://doi.org/10.1073/pnas.92.15.7075

Roberts, K. G., Li Y, Payne-Turner D, Harvey RC, Yang YL, Pei D, … CG, M. (2014). Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. *The New England Journal of Medicine*, *371*(11), 1005–1015. https://doi.org/10.1056/NEJMOA1403088

Robertson, S. M., Kennedy, M., Shannon, J. M., & Keller, G. (2000). A transitional stage in the commitment of mesoderm to hematopoiesis requiring the transcription factor SCL/tal-1. *Development (Cambridge, England)*, *127*(11), 2447–2459. https://doi.org/10.1242/dev.127.11.2447

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative Genomics Viewer. *Nature Biotechnology*, *29*(1), 24–26. https://doi.org/10.1038/nbt.1754

Romana, S. P., Coniat, M. L., & Berger, R. (1994). t(12;21): A new recurrent translocation in acute lymphoblastic leukemia. *Genes, Chromosomes and Cancer*, *9*(3), 186–191. https://doi.org/10.1002/gcc.2870090307

Rompaey, L. V., Potter, M., Adams, C., & Grosveld, G. (2000). Tel induces a G1 arrest and suppresses Ras-induced transformation. *Oncogene*, *19*(46), 5244–5250. https://doi.org/10.1038/sj.onc.1203899

Rothenberg, E. V., Moore, J. E., & Yui, M. A. (2008). Launching the T-cell-lineage developmental programme. *Nature Reviews Immunology 2008 8:1*, *8*(1), 9–21. https://doi.org/10.1038/nri2232

Roti G & Stegmaier K. (2014). New Approaches to Target T-ALL. *Frontiers in Oncology*, *4*. https://doi.org/10.3389/FONC.2014.00170

Roy, U., & Raghavan, S. C. (2021). Deleterious point mutations in T-cell acute lymphoblastic leukemia: Mechanistic insights into leukemogenesis. *International Journal of Cancer*, *149*(6), 1210–1220. https://doi.org/10.1002/ijc.33527

Ryan, S. L., Peden, J. F., Kingsbury, Z., Schwab, C. J., James, T., Polonen, P., … Ross, M. T. (2023). Whole genome sequencing provides comprehensive genetic testing in childhood B-cell acute lymphoblastic leukaemia. *Leukemia*, *37*(3), 518–528. https://doi.org/10.1038/s41375-022-01806-8

Sadras, T., Martin, M., Kume, K., Robinson, M. E., Saravanakumar, S., Lenz, G., … Müschen, M. (2021). Developmental partitioning of SYK and ZAP70 prevents autoimmunity and cancer. *Molecular Cell*, *81*(10), 2094-2111.e9. https://doi.org/10.1016/j.molcel.2021.03.043

Safavi, S., & Paulsson, K. (2017). Near-haploid and low-hypodiploid acute lymphoblastic leukemia: Two distinct subtypes with consistently poor prognosis. *Blood*, *129*(4), 420–423. https://doi.org/10.1182/blood-2016-10-743765

Saint-André, V., Federation, A. J., Lin, C. Y., Abraham, B. J., Reddy, J., Lee, T. I., … Young, R. A. (2016). Models of human core transcriptional regulatory circuitries. *Genome Research*, *26*(3), 385–396. https://doi.org/10.1101/gr.197590.115

Sanchez-Martin, M., & Ferrando, A. (2017). The NOTCH1-MYC highway toward T-cell acute lymphoblastic leukemia. *Blood*, *129*(9), 1124–1133. https://doi.org/10.1182/blood-2016-09-692582

Sanda, T. (2017). RUNX1 in T-ALL: Tumor suppressive or oncogenic? *Blood*, *130*(15), 1686–1688. https://doi.org/10.1182/blood-2017-08-802181

Sanda, T., Lawton, L. N., Barrasa, M. I., Fan, Z. P., Kohlhammer, H., Gutierrez, A., … Look, A. T. (2012). Core Transcriptional Regulatory Circuit Controlled by the TAL1 Complex in Human T Cell Acute Lymphoblastic Leukemia. *Cancer Cell*, *22*(2), 209–221. https://doi.org/10.1016/j.ccr.2012.06.007

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*, *33*(5), 495–502. https://doi.org/10.1038/nbt.3192

Schatz, D. G., & Swanson, P. C. (2011). V(D)J recombination: Mechanisms of initiation. *Annual Review of Genetics*, *45*, 167–202. https://doi.org/10.1146/annurev-genet-110410-132552

Schmidt, J.-A., Hornhardt, S., Erdmann, F., Sánchez-García, I., Fischer, U., Schüz, J., & Ziegelberger, G. (2021). Risk Factors for Childhood Leukemia: Radiation and Beyond. *Frontiers in Public Health*, *9*, 805757. https://doi.org/10.3389/fpubh.2021.805757

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*. https://doi.org/10.1038/nmeth.2089

Schofield, R. (1978). The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells*, *4*(1–2), 7–25.

Schrappe, M., Valsecchi, M. G., Bartram, C. R., Schrauder, A., Panzer-Grümayer, R., Möricke, A., … Conter, V. (2011). Late MRD response determines relapse risk overall and in subsets of childhood T-cell ALL: Results of the AIEOP-BFM-ALL 2000 study. *Blood*, *118*(8), 2077–2084. https://doi.org/10.1182/blood-2011-03-338707

Schultz, K. R., Carroll, A., Heerema, N. A., Bowman, W. P., Aledo, A., Slayton, W. B., … Camitta, B. (2014). Long-term follow-up of imatinib in pediatric Philadelphia chromosome-positive acute lymphoblastic leukemia: Children's Oncology Group Study AALL0031. *Leukemia*, *28*(7), 1467–1471. https://doi.org/10.1038/leu.2014.30

Schwickert, T. A., Tagoh, H., Gültekin, S., Dakic, A., Axelsson, E., Minnich, M., … Busslinger, M. (2014). Stage-specific control of early B cell development by the transcription factor Ikaros. *Nature Immunology*, *15*(3), 283–293. https://doi.org/10.1038/NI.2828

Scott, E. W., Simon, M. C., Anastasi, J., & Singh, H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science (New York, N.Y.)*, *265*(5178), 1573–1577. https://doi.org/10.1126/SCIENCE.8079170

Seeger, K., Stackelberg, A. V., Taube, T., Buchwald, D., Körner, G., Suttorp, M., … Henze, G. (2001). Relapse of TEL-AML1–Positive Acute Lymphoblastic Leukemia in Childhood: A Matched-Pair Analysis. *Journal of Clinical Oncology*, *19*(13), 3188–3193. https://doi.org/10.1200/JCO.2001.19.13.3188

Seki, M., Kimura, S., Isobe, T., Yoshida, K., Ueno, H., Nakajima-Takagi, Y., … Takita, J. (2017). Recurrent SPI1 (PU.1) fusions in high-risk pediatric T cell acute lymphoblastic leukemia. *Nature Genetics*, *49*(8), 1274–1281. https://doi.org/10.1038/ng.3900

Selenko, P., Sprangers, R., Stier, G., Bühler, D., Fischer, U., & Sattler, M. (2001). SMN Tudor domain structure and its interaction with the Sm proteins. *Nature Structural Biology*, *8*(1), 27–31. https://doi.org/10.1038/83014

Self, M., Lagutin, O. V., Bowling, B., Hendrix, J., Cai, Y., Dressler, G. R., & Oliver, G. (2006). Six2 is required for suppression of nephrogenesis and progenitor renewal in the developing kidney. *The EMBO Journal*, *25*(21), 5214–5228. https://doi.org/10.1038/sj.emboj.7601381

Shah, K., Al-Haidari, A., Sun, J., & Kazi, J. U. (2021). T cell receptor (TCR) signaling in health and disease. *Signal Transduction and Targeted Therapy*, *6*(1), 1–26. https://doi.org/10.1038/s41392-021-00823-w

Shurtleff, S. A., Buijs, A., Behm, F. G., Rubnitz, J. E., Raimondi, S. C., Hancock, M. L., … Downing, J. R. (1995). TEL/AML1 fusion resulting from a cryptic t(12;21) is the most common genetic lesion in pediatric ALL and defines a subgroup of patients with an excellent prognosis. *Leukemia*, *9*(12), 1985–1989.

Singh, V. K., Rastogi, A., Hu, X., Wang, Y., & De, S. (2020). Mutational signature SBS8 predominantly arises due to late replication errors in cancer. *Communications Biology*, *3*(1), 421. https://doi.org/10.1038/s42003-020-01119-5

Sood, R., Kamikubo, Y., & Liu, P. (2017). Role of RUNX1 in hematological malignancies. *Blood*, *129*(15), 2070–2082. https://doi.org/10.1182/BLOOD-2016-10-687830

Soulier, J., Clappier, E., Cayuela, J.-M., Regnault, A., García-Peydró, M., Dombret, H., … Sigaux, F. (2005). HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). *Blood*, *106*(1), 274–286. https://doi.org/10.1182/blood-2004-10-3900

Spits, H. (2002). Development of αβ T cells in the human thymus. *Nature Reviews Immunology*, *2*(10), 760–772. https://doi.org/10.1038/nri913

Stams, W. a. G., Beverloo, H. B., den Boer, M. L., de Menezes, R. X., Stigter, R. L., van Drunen, E., … Pieters, R. (2006). Incidence of additional genetic changes in the TEL and AML1 genes in DCOG and COALL-treated t(12;21)-positive pediatric ALL, and their relation with drug sensitivity and clinical outcome. *Leukemia*, *20*(3), 410–416. https://doi.org/10.1038/sj.leu.2404083

Starr, T. K., Jameson, S. C., & Hogquist, K. A. (2003). Positive and Negative Selection of T Cells. *Https://Doi.Org/10.1146/Annurev.Immunol.21.120601.141107*, *21*, 139–176. https://doi.org/10.1146/ANNUREV.IMMUNOL.21.120601.141107

Steelman, L. S., Franklin, R. A., Abrams, S. L., Chappell, W., Kempf, C. R., Bäsecke, J., … McCubrey, J. A. (2011). Roles of the Ras/Raf/MEK/ERK pathway in leukemia therapy. *Leukemia*, *25*(7), 1080–1094. https://doi.org/10.1038/leu.2011.66

Steliarova-Foucher, E., Colombet, M., Ries, L. A. G., Moreno, F., Dolya, A., Bray, F., … Stiller, C. A. (2017). International incidence of childhood cancer, 2001–10: A population-based registry study. *The Lancet. Oncology*, *18*(6), 719–731. https://doi.org/10.1016/S1470-2045(17)30186-9

Streubel, G., Bouchard, C., Berberich, H., Zeller, M. S., Teichmann, S., Adamkiewicz, J., … Bauer, U.-M. (2013). PRMT4 is a novel coactivator of c-Myb-dependent transcription in haematopoietic cell lines. *PLoS Genetics*, *9*(3), e1003343. https://doi.org/10.1371/journal.pgen.1003343

Sun, C., Chang, L., & Zhu, X. (2017). Pathogenesis of ETV6/RUNX1-positive childhood acute lymphoblastic leukemia and mechanisms underlying its relapse. *Oncotarget*, *8*(21), 35445–35459. https://doi.org/10.18632/oncotarget.16367

Suwabe, N., Takahashi, S., Nakano, T., & Yamamoto, M. (1998). GATA-1 Regulates Growth and Differentiation of Definitive Erythroid Lineage Cells During In Vitro ES Cell Differentiation. *Blood*, *92*(11), 4108–4118. https://doi.org/10.1182/BLOOD.V92.11.4108

Swerdlow, S. H., Campo, E., Pileri, S. A., Harris, N. L., Stein, H., Siebert, R., … Jaffe, E. S. (2016). The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*, *127*(20), 2375–2390. https://doi.org/10.1182/blood-2016-01-643569

Szewczyk, M. M., Ishikawa, Y., Organ, S., Sakai, N., Li, F., Halabelian, L., … Barsyte-Lovejoy, D. (2020). Pharmacological inhibition of PRMT7 links arginine monomethylation to the cellular stress response. *Nature Communications*, *11*(1), 2396. https://doi.org/10.1038/s41467-020-16271-z

Tan, T. K., Zhang, C., & Sanda, T. (2019). Oncogenic transcriptional program driven by TAL1 in T-cell acute lymphoblastic leukemia. *International Journal of Hematology*, *109*(1), 5–17. https://doi.org/10.1007/s12185-018-2518-z

Tavian, M., & Péault, B. (2005). Embryonic development of the human hematopoietic system. *The International Journal of Developmental Biology*, *49*(2–3), 243–250. https://doi.org/10.1387/ijdb.041957mt

Tenenbaum, D., & Maintainer, B. (2021). *Package "KEGGREST" Title Client-side REST access to the Kyoto Encyclopedia of Genes and Genomes (KEGG). R package version 1.35.0.* Bioconductor version: Release (3.14). https://doi.org/10.18129/B9.bioc.KEGGREST

Teppo, S., Laukkanen, S., Liuksiala, T., Nordlund, J., Oittinen, M., Teittinen, K., … Lohi, O. (2016). Genome-wide repression of eRNA and target gene loci by the ETV6-RUNX1 fusion in acute leukemia. *Genome Research*, *26*(11), 1468–1477. https://doi.org/10.1101/gr.193649.115

Tewary, S. K., Zheng, Y. G., & Ho, M.-C. (2019). Protein arginine methyltransferases: Insights into the enzyme structure and mechanism at the atomic level. *Cellular and Molecular Life Sciences*, *76*(15), 2917–2932. https://doi.org/10.1007/s00018-019-03145-x

Thatikonda, V., Islam, S. M. A., Autry, R. J., Jones, B. C., Gröbner, S. N., Warsow, G., … Jäger, N. (2023). Comprehensive analysis of mutational signatures reveals distinct patterns and molecular processes across 27 pediatric cancers. *Nature Cancer*, *4*(2), 276–289. https://doi.org/10.1038/s43018-022-00509-4

Till, J. E., & McCulloch, E. A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiation Research*, *14*, 213–222.

Toft, N., Birgens, H., Abrahamsson, J., Griškevičius, L., Hallböök, H., Heyman, M., … Schmiegelow, K. (2018). Results of NOPHO ALL2008 treatment for patients aged 1–45 years with acute lymphoblastic leukemia. *Leukemia*, *32*(3), 606–615. https://doi.org/10.1038/leu.2017.265

Torrano, V., Procter, J., Cardus, P., Greaves, M., & Ford, A. M. (2011). ETV6-RUNX1 promotes survival of early B lineage progenitor cells via a dysregulated erythropoietin receptor. *Blood*, *118*(18), 4910–4918. https://doi.org/10.1182/blood-2011-05-354266

Tosello, V., Mansour, M. R., Barnes, K., Paganin, M., Sulis, M. L., Jenkinson, S., … Ferrando, A. A. (2009). WT1 mutations in T-ALL. *Blood*, *114*(5), 1038–1045. https://doi.org/10.1182/blood-2008-12-192039

Toy, J., & Sundin, O. H. (1999). Expression of the optx2 homeobox gene during mouse development. *Mechanisms of Development*, *83*(1–2), 183–186. https://doi.org/10.1016/s0925-4773(99)00049-0

Treviño, L. R., Yang, W., French, D., Hunger, S. P., Carroll, W. L., Devidas, M., … Relling, M. V. (2009). Germline genomic variants associated with childhood acute lymphoblastic leukemia. *Nature Genetics*, *41*(9), 1001–1005. https://doi.org/10.1038/ng.432

Tripsianes, K., Madl, T., Machyna, M., Fessas, D., Englbrecht, C., Fischer, U., … Sattler, M. (2011). Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nature Structural & Molecular Biology*, *18*(12), 1414–1420. https://doi.org/10.1038/nsmb.2185

Tsuzuki, S., Karnan, S., Horibe, K., Matsumoto, K., Kato, K., Inukai, T., … Seto, M. (2007). Genetic abnormalities involved in t(12;21) TEL-AML1 acute lymphoblastic leukemia: Analysis by means of array-based comparative genomic hybridization. *Cancer Science*, *98*(5), 698–706. https://doi.org/10.1111/j.1349-7006.2007.00443.x

van Delft, F. W., Horsley, S., Colman, S., Anderson, K., Bateman, C., Kempski, H., … Greaves, M. (2011). Clonal origins of relapse in ETV6-RUNX1 acute lymphoblastic leukemia. *Blood*, *117*(23), 6247–6254. https://doi.org/10.1182/blood-2010-10-314674

Van der Meulen, J., Van Roy, N., Van Vlierberghe, P., & Speleman, F. (2014). The epigenetic landscape of T-cell acute lymphoblastic leukemia. *The International Journal of Biochemistry & Cell Biology*, *53*, 547–557. https://doi.org/10.1016/j.biocel.2014.04.015

van Dongen, J. J. M., van der Velden, V. H. J., Brüggemann, M., & Orfao, A. (2015). Minimal residual disease diagnostics in acute lymphoblastic leukemia: Need for sensitive, fast, and standardized technologies. *Blood*, *125*(26), 3996–4009. https://doi.org/10.1182/blood-2015-03-580027

Van Vlierberghe, P., Ambesi-Impiombato, A., Perez-Garcia, A., Haydu, J. E., Rigo, I., Hadler, M., … Ferrando, A. A. (2011). ETV6 mutations in early immature human T cell leukemias. *The Journal of Experimental Medicine*, *208*(13), 2571–2579. https://doi.org/10.1084/jem.20112239

Van Vlierberghe, P., Palomero, T., Khiabanian, H., Van der Meulen, J., Castillo, M., Van Roy, N., … Ferrando, A. (2010). PHF6 mutations in T-cell acute lymphoblastic leukemia. *Nature Genetics*, *42*(4), 338–342. https://doi.org/10.1038/ng.542

van Zelm, M. C., van der Burg, M., de Ridder, D., Barendregt, B. H., de Haas, E. F. E., Reinders, M. J. T., … van Dongen, J. J. M. (2005). Ig gene rearrangement steps are

initiated in early human precursor B cell subsets and correlate with specific transcription factor expression. *Journal of Immunology (Baltimore, Md.: 1950)*, *175*(9), 5912–5922. https://doi.org/10.4049/jimmunol.175.9.5912

Vicente, C., Conchillo, A., García-Sánchez, M. A., & Odero, M. D. (2012). The role of the GATA2 transcription factor in normal and malignant hematopoiesis. *Critical Reviews in Oncology/Hematology*, *82*(1), 1–17. https://doi.org/10.1016/J.CRITREVONC.2011.04.007

von Lintig, F. C., Huvar, I., Law, P., Diccianni, M. B., Yu, A. L., & Boss, G. R. (2000). Ras activation in normal white blood cells and childhood acute lymphoblastic leukemia. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *6*(5), 1804–1810.

Waddington, C. H. (1957). The strategy of the genes. A discussion of some aspects of theoretical biology. With an appendix by H. Kacser. *The Strategy of the Genes. A Discussion of Some Aspects of Theoretical Biology. With an Appendix by H. Kacser.*

Wang, J. H., Nichogiannopoulou, A., Wu, L., Sun, L., Sharpe, A. H., Bigby, M., & Georgopoulos, K. (1996). Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. *Immunity*, *5*(6), 537–549. https://doi.org/10.1016/S1074-7613(00)80269-1

Wang, L. C., Swat, W., Fujiwara, Y., Davidson, L., Visvader, J., Kuo, F., … Orkin, S. H. (1998). The TEL/ETV6 gene is required specifically for hematopoiesis in the bone marrow. *Genes & Development*, *12*(15), 2392–2402. https://doi.org/10.1101/gad.12.15.2392

Wang, M., & Zhang, C. (2020). Low LEF1 expression is a biomarker of early T-cell precursor, an aggressive subtype of T-cell lymphoblastic leukemia. *PLoS ONE*, *15*(5), e0232520. https://doi.org/10.1371/journal.pone.0232520

Wang, Y., Zeng, H., & Zhang, L. (2018). ETV6/RUNX1-positive childhood acute lymphoblastic leukemia in China: Excellent prognosis with improved BFM protocol. *Italian Journal of Pediatrics*, *44*, 94. https://doi.org/10.1186/s13052-018-0541-6

Ward, A. F., Braun, B. S., & Shannon, K. M. (2012). Targeting oncogenic Ras signaling in hematologic malignancies. *Blood*, *120*(17), 3397–3406. https://doi.org/10.1182/blood-2012-05-378596

Ward, G. (1917). The infective theory of acute leukaemia. *British Journal of Childhood Diseases*, *14*, 10–20.

Weber, B. N., Chi, A. W.-S., Chavez, A., Yashiro-Ohtani, Y., Yang, Q., Shestova, O., & Bhandoola, A. (2011). A critical role for TCF-1 in T-lineage specification and differentiation. *Nature*, *476*(7358), 63–68. https://doi.org/10.1038/nature10279

Wiemels, J. L., Ford, A. M., Van Wering, E. R., Postma, A., & Greaves, M. (1999). Protracted and variable latency of acute lymphoblastic leukemia after TEL-AML1 gene fusion in utero. *Blood*, *94*(3), 1057–1062.

Wilson, A., Laurenti, E., Oser, G., van der Wath, R. C., Blanco-Bose, W., Jaworski, M., … Trumpp, A. (2008). Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair. *Cell*, *135*(6), 1118–1129. https://doi.org/10.1016/j.cell.2008.10.048

Wilson, A., & Trumpp, A. (2006). Bone-marrow haematopoietic-stem-cell niches. *Nature Reviews Immunology 2006 6:2*, *6*(2), 93–106. https://doi.org/10.1038/nri1779

Wilson, N. K., Foster, S. D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., … Göttgens, B. (2010). Combinatorial transcriptional control in blood stem/progenitor cells:

Genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell*, *7*(4), 532–544. https://doi.org/10.1016/j.stem.2010.07.016

Winter, S. S., Dunsmore, K. P., Devidas, M., Wood, B. L., Esiashvili, N., Chen, Z., … Hunger, S. P. (2018). Improved Survival for Children and Young Adults With T-Lineage Acute Lymphoblastic Leukemia: Results From the Children's Oncology Group AALL0434 Methotrexate Randomization. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *36*(29), 2926–2934. https://doi.org/10.1200/JCO.2018.77.7250

Wu, K., Yang, Y., Wang, C., Davoli, M. A., D'Amico, M., Li, A., … Pestell, R. G. (2003). DACH1 Inhibits Transforming Growth Factor-β Signaling through Binding Smad4*. *Journal of Biological Chemistry*, *278*(51), 51673–51684. https://doi.org/10.1074/jbc.M310021200

Xia, X. (2020). Beyond Trees: Regulons and Regulatory Motif Characterization. *Genes*, *11*(9), 995. https://doi.org/10.3390/genes11090995

Xie, H., Hoffmann, H. M., Meadows, J. D., Mayo, S. L., Trang, C., Leming, S. S., … Mellon, P. L. (2015). Homeodomain Proteins SIX3 and SIX6 Regulate Gonadotrope-specific Genes During Pituitary Development. *Molecular Endocrinology*, *29*(6), 842–855. https://doi.org/10.1210/me.2014-1279

Xu, H.-X., Wu, K.-J., Tian, Y.-J., Liu, Q., Han, N., He, X.-L., … Wu, K.-M. (2016). Expression profile of SIX family members correlates with clinic-pathological features and prognosis of breast cancer. *Medicine*, *95*(27), e4085. https://doi.org/10.1097/MD.0000000000004085

Xu, J., & Richard, S. (2021). Cellular pathways influenced by protein arginine methylation: Implications for cancer. *Molecular Cell*, *81*(21), 4357–4368. https://doi.org/10.1016/j.molcel.2021.09.011

Xu, P.-X. (2013). The EYA-SO/SIX complex in development and disease. *Pediatric Nephrology*, *28*(6), 843–854. https://doi.org/10.1007/s00467-012-2246-1

Xu, P.-X., Zheng, W., Huang, L., Maire, P., Laclef, C., & Silvius, D. (2003). Six1 is required for the early organogenesis of mammalian kidney. *Development (Cambridge, England)*, *130*(14), 3085–3094. https://doi.org/10.1242/dev.00536

Yang, Y., & Bedford, M. T. (2013). Protein arginine methyltransferases and cancer. *Nature Reviews Cancer*, *13*(1), 37–50. https://doi.org/10.1038/nrc3409

Yao R, H, J., Y, M., L, W., L, W., J, D., … J, L. (2014). PRMT7 induces epithelial-to-mesenchymal transition and promotes metastasis in breast cancer. *Cancer Research*, *74*(19), 5656–5667. https://doi.org/10.1158/0008-5472.CAN-14-0800

Yasuda, T., Tsuzuki, S., Kawazu, M., Hayakawa, F., Kojima, S., Ueno, T., … Mano, H. (2016). Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults. *Nature Genetics*, *48*(5), 569–574. https://doi.org/10.1038/ng.3535

Ying, Z., Mei, M., Zhang, P., Liu, C., He, H., Gao, F., & Bao, S. (2015). Histone Arginine Methylation by PRMT7 Controls Germinal Center Formation via Regulating Bcl6 Transcription. *The Journal of Immunology*, *195*(4), 1538–1547. https://doi.org/10.4049/jimmunol.1500224

Yoshida, T., & Georgopoulos, K. (2014). Ikaros fingers on lymphocyte differentiation. *International Journal of Hematology*, *100*(3), 220. https://doi.org/10.1007/S12185-014-1644-5

Yu, Y., Wang, J., Khaled, W., Burke, S., Li, P., Chen, X., … Liu, P. (2012). Bcl11a is essential for lymphoid development and negatively regulates p53. *The Journal of Experimental Medicine*, *209*(13), 2467–2483. https://doi.org/10.1084/jem.20121846

Zakaria, Z., Ahid, M. F. M., Ismail, A., Keoh, T. S., Nor, N. M., Kamaluddin, N. R., … Osman, R. (2012). Chromosomal Aberrations in ETV6/RUNX1-positive Childhood Acute Lymphoblastic Leukemia using 244K Oligonucleotide Array Comparative Genomic Hybridization. *Molecular Cytogenetics*, *5*(1), 41. https://doi.org/10.1186/1755-8166-5-41

Zaliova, M., Kotrova, M., Bresolin, S., Stuchly, J., Stary, J., Hrusak, O., … Vaskova, M. (2017). ETV6/RUNX1-like acute lymphoblastic leukemia: A novel B-cell precursor leukemia subtype associated with the CD27/CD44 immunophenotype. *Genes, Chromosomes & Cancer*, *56*(8), 608–616. https://doi.org/10.1002/gcc.22464

Zanardi, E., Verzoni, E., Grassi, P., Necchi, A., Giannatempo, P., Raggi, D., … Procopio, G. (2015). Clinical experience with temsirolimus in the treatment of advanced renal cell carcinoma. *Therapeutic Advances in Urology*, *7*(3), 152–161. https://doi.org/10.1177/1756287215574457

Zenatti, P. P., Ribeiro, D., Li, W., Zuurbier, L., Silva, M. C., Paganin, M., … Barata, J. T. (2011). Oncogenic IL7R gain-of-function mutations in childhood T-cell acute lymphoblastic leukemia. *Nature Genetics*, *43*(10), 932–939. https://doi.org/10.1038/ng.924

Zhang, J., Ding, L., Holmfeldt, L., Wu, G., Heatley, S. L., Payne-Turner, D., … Mullighan, C. G. (2012). The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature*, *481*(7380), 157–163. https://doi.org/10.1038/nature10725

Zhang, J., McCastlain, K., Yoshihara, H., Xu, B., Chang, Y., Churchman, M. L., … St. Jude Children's Research Hospital–Washington University Pediatric Cancer Genome Project. (2016). Deregulation of DUX4 and ERG in acute lymphoblastic leukemia. *Nature Genetics*, *48*(12), 1481–1489. https://doi.org/10.1038/ng.3691

Zhang, X., Dong, W., Zhou, H., Li, H., Wang, N., Miao, X., & Jia, L. (2015). α-2,8-sialyltransferase is involved in the development of multidrug resistance via PI3K/Akt pathway in human chronic myeloid leukemia. *IUBMB Life*, *67*(2), 77–87. https://doi.org/10.1002/iub.1351

Zhang, Y., Huang, Y., Hu, L., & Cheng, T. (2022). New insights into Human Hematopoietic Stem and Progenitor Cells via Single-Cell Omics. *Stem Cell Reviews and Reports*, *18*(4), 1322–1336. https://doi.org/10.1007/S12015-022-10330-2/TABLES/1

Zhao, X., Jankovic, V., Gural, A., Huang, G., Pardanani, A., Menendez, S., … Nimer, S. D. (2008). *Methylation of RUNX1 by PRMT1 abrogates SIN3A binding and potentiates its transcriptional activity*. https://doi.org/10.1101/gad.1632608

Zhou, X., Li, Z., & Zhou, J. (2017). Tumor necrosis factor α in the onset and progression of leukemia. *Experimental Hematology*, *45*, 17–26. https://doi.org/10.1016/j.exphem.2016.10.005

Zhu, Jiang, & Emerson, S. G. (2002). Hematopoietic cytokines, transcription factors and lineage commitment. *Oncogene*, *21*(21), 3295–3313. https://doi.org/10.1038/sj.onc.1205318

Zhu, Junji, Li, X., Cai, X., Zha, H., Zhou, Z., Sun, X., … Xiao, W. (2021). Arginine monomethylation by PRMT7 controls MAVS-mediated antiviral innate immunity. *Molecular Cell*, *81*(15), 3171-3186.e8. https://doi.org/10.1016/j.molcel.2021.06.004

Ziegler, G. C., Almos, P., McNeill, R. V., Jansch, C., & Lesch, K.-P. (2020). Cellular effects and clinical implications of SLC2A3 copy number variation. *Journal of Cellular Physiology*, *235*(12), 9021–9036. https://doi.org/10.1002/jcp.29753

Zlotoff, D. A., & Bhandoola, A. (2011). Hematopoietic progenitor migration to the adult thymus. *Annals of the New York Academy of Sciences*, *1217*(1), 122–138. https://doi.org/10.1111/J.1749-6632.2010.05881.X

Zuber, M. E., Perron, M., Philpott, A., Bang, A., & Harris, W. A. (1999). Giant eyes in Xenopus laevis by overexpression of XOptx2. *Cell*, *98*(3), 341–352. https://doi.org/10.1016/s0092-8674(00)81963-7

Zuna, J., Ford, A. M., Peham, M., Patel, N., Saha, V., Eckert, C., … Greaves, M. (2004). TEL Deletion Analysis Supports a Novel View of Relapse in Childhood Acute Lymphoblastic Leukemia. *Clinical Cancer Research*, *10*(16), 5355–5360. https://doi.org/10.1158/1078-0432.CCR-04-0584

Zuna, J., Madzo, J., Krejci, O., Zemanova, Z., Kalinova, M., Muzikova, K., … Trka, J. (2011). ETV6/RUNX1 (TEL/AML1) is a frequent prenatal first hit in childhood leukemia. *Blood*, *117*(1), 368–369. https://doi.org/10.1182/blood-2010-09-309070

PUBLICATIONS

Publication I

Laukkanen, S., **Oksa, L.,** Nikkilä, A., Lahnalampi, M., Parikka, M., Seki, M., Takita, J., Degerman, S., de Bock, C. E., Heinäniemi, M., & Lohi, O. (2020). SIX6 is a TAL1-regulated transcription factor in T-ALL and associated with inferior outcome. *Leukemia & lymphoma*, *61*(13), 3089–3100. https://doi.org/10.1080/10428194.2020.1804560

Publication II

Oksa, L., Mäkinen, A., Nikkilä, A., Hyvärinen, N., Laukkanen, S., Rokka, A., Haapaniemi, P., Seki, M., Takita, J., Kauko, O., Heinäniemi, M., & Lohi, O. (2022). Arginine Methyltransferase PRMT7 Deregulates Expression of RUNX1 Target Genes in T cell Acute Lymphoblastic Leukemia. *Cancers*, *14*(9), 2169. https://doi.org/10.3390/cancers14092169.

Publication III

Oksa, L., Moisio S., Maqbool, K., Kramer, R. Nikkilä, A., Foroughi-Asl H., Suhonen, Krali, O., Voutilainen, M., Lahnalampi, M., J., Duque-Afonso, J., Hauer J., Vepsäläinen, K., Nordlund, J., Wirta, V., Lohi O.* & Heinäniemi M.* Genomic determinants of therapy response in *ETV6::RUNX1* leukemia. Submitted.

PUBLICATION I

SIX6 is a TAL1-regulated transcription factor in T-ALL and associated with inferior outcome

Laukkanen, S., Oksa, L., Nikkilä, A., Lahnalampi, M., Parikka, M., Seki, M., Takita, J., Degerman, S., de Bock, C. E., Heinäniemi, M., & Lohi, O.

Leukemia & lymphoma 2020, 61(13), 3089–3100.
https://doi.org/10.1080/10428194.2020.1804560

SIX6 is a TAL1-regulated transcription factor in T-ALL and associated with inferior outcome

Saara Laukkanen[1*], Laura Oksa[1], Atte Nikkilä[1], Mari Lahnalampi[2], Mataleena Parikka[3,4], Masafumi Seki[5], Junko Takita[6], Sofie Degerman[7], Charles E. de Bock[8,9], Merja Heinäniemi[2] and Olli Lohi[1,10]

1. Tampere Center for Child Health Research, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland
2. The Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland
3. BioMediTech, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland
4. Tampere University Hospital, Oral and Maxillofacial Unit, Tampere, Finland
5. Department of Cell and Molecular Biology, Karolinska Institutet, Solna, Sweden
6. Graduate School of Medicine, Kyoto University, Kyoto, Japan
7. Department of Medical Biosciences and Department of Clinical Microbiology, Umeå University, Umeå, Sweden
8. Children's Cancer Institute, Lowy Cancer Research Centre, Sydney, NSW, Australia.
9. School of Women's and Children's Health, University of New South Wales Sydney, Sydney, NSW, Australia
10. Tampere University Hospital, Tays Cancer Centre, Tampere, Finland

*Corresponding author, saara.laukkanen@tuni.fi

Leukemia & Lymphoma, 2020, https://doi.org/10.1080/10428194.2020.1804560

T-cell acute lymphoblastic leukemia (T-ALL) is a hematological malignancy driven by abnormal activity of transcription factors. Here we report an aberrant expression of the developmental transcription factor SIX6 in the TAL1-subtype of T-ALL. Our results demonstrate that the binding of TAL1 and GATA3 transcription factors into an upstream enhancer element directly regulates SIX6 expression. High expression of SIX6 was associated with inferior event-free survival within three independent patient cohorts. At a functional level, CRISPR-Cas9-mediated knockout of the SIX6 gene in TAL1 positive Jurkat cells induced changes in genes associated with the mTOR-, K-RAS-, and TNFα-related molecular signatures but did not impair cell proliferation or viability. There was also no acceleration of T-ALL development within a Myc driven zebrafish tumor model in vivo. Taken together, our results show that SIX6 belongs to the TAL1 regulatory gene network in T-ALL but is alone insufficient to influence the development or maintenance of T-ALL.

Keywords: T-cell acute lymphoblastic leukemia, SIX6, TAL1, Event-free survival, aberrant gene expression

---

1. Introduction

T-cell acute lymphoblastic leukemia (T-ALL) is a hematopoietic cancer of lymphoid lineage, which accounts for 10-15% of pediatric and 25% of adult acute lymphoblastic leukemia cases [1]. Generally, T-ALL patients have less favorable outcomes compared to B-cell acute lymphoblastic leukemia (B-ALL), but the survival rates have improved over time, especially in pediatric patients, as the most recent NOPHO ALL-2008 protocol had the five year overall survival rate of 82% in pediatric patients under 10 year old [2]. That said, in this same study, 14% of the patients in all age groups experienced relapse and only three out of 39 relapse patients were alive at the end of follow-up as the leukemic cells become resistant to chemotherapeutic drugs currently available [2]. Therefore, there remains an urgent need to better understand the molecular events underlying the development and maintenance of T-ALL to aid the development of targeted and more effective treatments.

Development and maturation of T-cells is regulated by a network of transcription factors (TF) [3-4]. Fittingly, abnormal activity and

expression of TFs is a common feature in leukemia and T-ALL cases can be grouped based on the expression of TAL1, LMO1/2, LYL1, HOXA family, NKX2-1, TLX1, and TLX3 transcription factors [3]. Aberrant expression patterns are typically a result of indel mutations or chromosomal rearrangements at coding or regulatory regions [5]. In addition, the gene expression subgroups in T-ALL were recently associated with CpG island methylation phenotype (CIMP) status, indicating a direct or indirect role of epigenetics in the regulation [6]. The abnormal activation of TFs by genetic or epigenetic mechanisms then alters the expression of multiple other genes.

Here we set out to investigate the expression of transcriptional regulators across leukemias and hematological malignancies. We uncovered a distinct expression of Sine oculis homeobox 6 (SIX6) in T-ALL, and explored its transcriptional regulation, function and the prognostic value in T-ALL.

2. Materials and methods

2.1 CRISPR-Cas9 mediated knockout of SIX6 in Jurkat-Cas9 cells

CRISPR guide RNAs were designed with Desktop genetics online tool (London, UK) and px321-GFP-gRNA plasmids were transfected to Jurkat-Cas9 cells using Nucleofector electroporation system (Lonza, Basel, Switzerland). Resulting cell pool was single cell sorted to obtain monoclonal cell cultures and resulting clones were analyzed for mutations with T7 endonuclease (New England Biolabs, Ipswich, MA, USA), Sanger sequencing and Western blotting.

See Supplementary Materials and Methods for further details on cell culture, creation and in vitro characterization of SIX6 knockout cells.

2.2 Next generation sequencing methods and data origins

For Global Nuclear Run-On sequencing (GRO-seq), the nuclei isolation was performed on untreated Jurkat, MOLT-4, CCRF-CEM and P12-Ichikawa cells (yielding ~5 × 106 nuclei per condition). The nuclear run-on reaction, library preparation, and data processing were performed as previously described [7-10].

For analyzing histone modifications at gene promoter and enhancer region, histone data from T-ALL cell lines (Jurkat, MOLT-4, CCRF-CEM and P12-Ichikawa) were acquired from NCBI SRA repository (sra-toolkit/2.9.2): GSE68976, GSE76783, GSE97514, GSE60104[11-14]. ChiP-seq data for TAL1 and GATA3 transcription factors in Jurkat cells (GSE29180) [15] as well as Hi-C data, that was retrieved for exploring interactions between genomic regions in untreated Jurkat cells (GSE122958) [16], were similarly obtained.

The RNA-seq data of SIX6 KO and wildtype Jurkat cells was generated and analyzed in this study and data is available in the Gene Expression Omnibus (GEO) repository: GSE148658.

Single-cell RNA-seq data has been previously published by Park et al. [17] and the data was utilized via Human Cell Atlas Developmental thymus data portal (https://developmentcellatlas.ncl.ac.uk/). Pre-processed and normalized expression matrix for RNA-seq data from GFP, LUC, TAL1, GATA3, MYB, E2A, HEB, or RUNX1 shRNA treated Jurkat cells were retrieved from GEO repository (GSE29179) [15].

See Supplementary Materials and methods for detailed description of the analyses.

2.3 Hemap resource

Compilation of gene expression data (Hemap) from the microarray studies of various hematological and lymphoid malignancies and healthy tissues that were performed on a uniform technical platform (Affymetrix HG U133 Plus 2.0) was utilized for comparative gene expression analyses [18, 19]. Hemap resource is available at http://hemap.uta.fi/. To study SIX6 expression specifically among T-ALL subtypes, the data was clustered as described previously [20].

2.4 Survival analysis of patient cohorts

Transcriptome and survival data of pediatric T-ALL patients in Nordic, Japanese and TARGET cohorts [3, 6, 21, 22] were analyzed by calculating Kaplan-Meier survival estimates, and the difference between the groups was examined using the Log-rank test. In all three patient cohorts, induction failure, relapse, secondary malignant disease and death by any cause were counted as events. In the Nordic cohort, resistant disease was also counted as an event. Cox proportional hazards models were fitted to

estimate the effects of individual risk factors on survival. The R-packages survminer and/or survival were applied in order to analyze and visualize the data.

2.5 Zebrafish experiments

In microinjection experiments, one cell stage embryos were injected with a mix, which contained overexpression constructs and rhodamine dextran. The study group was co-injected with linearized Rag2:mMYC and Rag2:hSIX6-TdTomato plasmids and the control group was injected with linearized Rag2:mMYC-TdTomato plasmid only. All resulting zebrafish were screened for fluorescence positivity at 21 dpf and 28 dpf. After 28 dpf, only fluorescent positive animals were followed weekly or every other week. When over 50% of a zebrafish's body turned fluorescent positive, the fish was considered to have leukemia [23].

3. Results

3.1 Expression of SIX6 is highest within the TAL1-subtype of T-ALL

We have previously compiled a data set consisting of gene expression profiles from 9 544 samples, including 385 T-ALL, 1 304 pre-B-ALL, 1 713 AML, 801 CLL, and 247 normal T lymphocytes (Hemap data) [18, 19]. We have utilized this data set to find essential transcriptional regulators and prognostic markers in hematological malignancies based on the aberrant gene expression pattern in a disease entity or subtype [18, 24]. In this in silico screen, SIX6 came out as a candidate gene with an abnormal and disease-specific expression pattern as SIX6 was expressed at low levels in precursor B-ALL, AML, CML, CLL and in healthy lymphocytes, but showed an 1.3-fold increased expression in T-ALL over healthy T lymphocytes (p-value < 0.001, Figure 1A). To examine the SIX6 expression in more detail, we clustered T-ALL samples into transcription factor defined subtypes and observed high expression of SIX6 in the TAL1 related subtypes (Figure 1B). SIX6 and TAL1 expressions also clearly correlated in this data set (Figure 1C). In a panel of T- and B-cell leukemia cell lines, and CD3+ and CD19+ lymphocytes from healthy blood donors, SIX6 expression was strong among the TAL1 subgroup of T-ALL cell lines (Jurkat, MOLT-4 and CCRF-CEM), low in CD19+ B-cells and absent in B-ALL cell lines or CD3+ T-cells (Figure 1D).

SIX6 is a developmental transcription factor that has an essential function during the eye development [25-28]. Therefore, we examined if SIX6 has a developmental role in T-cells as well by studying its expression at a certain stages of T-cell development. We utilized single-cell RNA-seq data that includes cells from the early double negative to CD4 and CD8 single positive stages, as well as T-reg, CD8αα+, and γδ T cells [17]. As shown in Figure 1E, distinct expression of several genes with a well-known function in T-cell development was observed [29-33], but SIX6 and TAL1 were not expressed in any of the cell types along the T cell differentiation pathway.

Taken together, our data shows aberrant and disease specific expression of SIX6 in T-ALL.

3.2 SIX6 expression is regulated by TAL1 through a novel upstream enhancer

Correlation of SIX6 expression to the TAL1 subgroup suggests a regulatory relationship between the TAL1 and SIX6 proteins. To address this question, we utilized gene silencing data generated by Sanda et al. [15]. Silencing of either TAL1 or GATA3, partners of a known transcription factor complex [15, 34], triggered downregulation of SIX6 expression in Jurkat cells (n = 4, Figure 1F). Downregulation of MYB had a similar effect on SIX6 expression, likely contributed by indirect regulation via TAL1 due to a mutation at the regulatory region of TAL1 in Jurkat cells that introduces a novel MYB binding site [35].

ChIP-seq data for TAL1 or GATA3[15] did not uncover any strong binding to the proximal promoter region of the SIX6 gene, suggesting that some long-range enhancer elements might be responsible for the regulation of gene expression (Figure 2). We employed GRO-seq assay, which detects nascent RNA expression including short-lived RNA molecules such as enhancer RNAs (eRNA) [36], to explore the co-expression of regulatory elements around the SIX6 gene, and found an active enhancer element 165 kb upstream of the SIX6 gene. Three of the examined cell lines showed prominent co-expression of SIX6 and the upstream eRNA (Jurkat, MOLT-4, and CCRF-CEM), while a SIX6-negative cell line P12-Ichikawa lacked signal at the enhancer locus. Furthermore, binding of the TAL1 and GATA3 proteins was prominent at this enhancer locus.

Figure 1

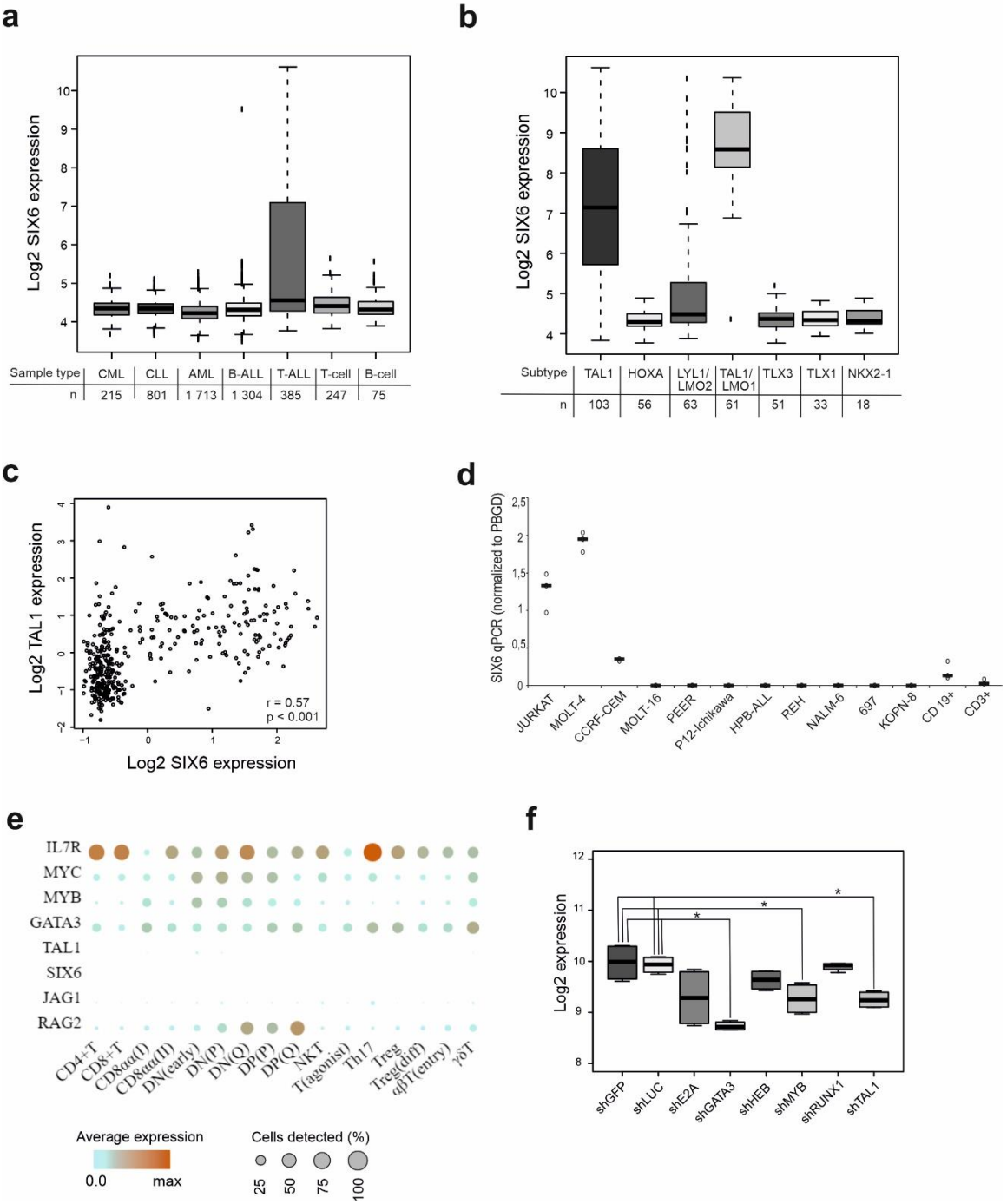


FIGURE 1. Expression of *SIX6* across hematological malignancies and healthy T-cells. A) Expression of *SIX6* in the Hemap data set. B) Expression of *SIX6* in T-ALL subtypes in the Hemap data set. C) Correlation of *TAL1* and *SIX6* gene expression in T-ALL samples from the Hemap data set. D) Expression of *SIX6* as measured by qRT-PCR in a panel of T- and B-ALL cell lines and normal lymphocytes (n=3, all experimental data points and medians are shown). E) Expression of *SIX6* and the indicated genes that are relevant to T-cell and T-ALL development in a single cell RNA-seq data of human thymic cells; the color indicates the maximum-normalized mean expression, and the size refers to the proportion of cells expressing the marker genes relative to the absolute number of cells detected in the data set. F) Expression of *SIX6* after silencing of *E2A*, *GATA3*, *HEB*, *MYB*, *RUNX1* or *TAL1* genes (n=4, * = p-value < 0.05).

Figure 2

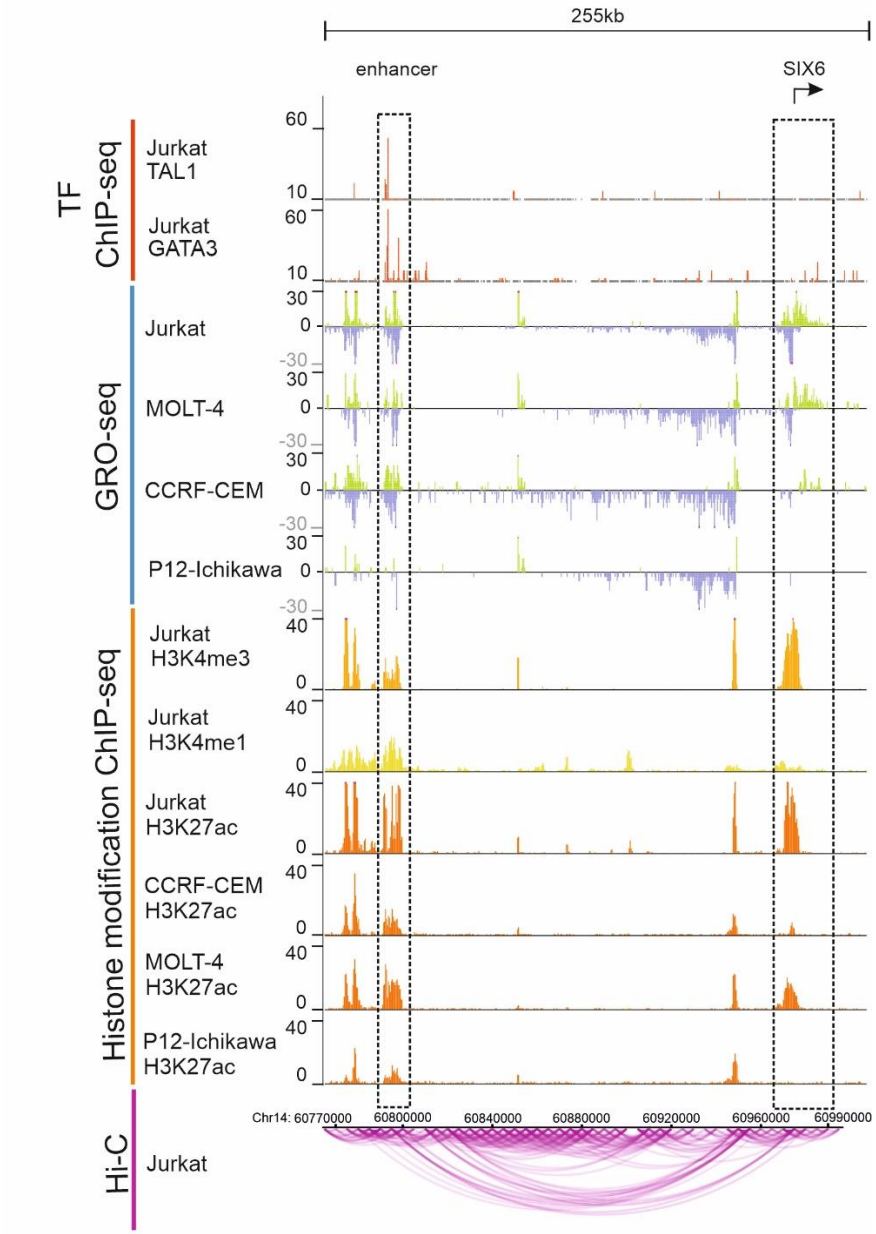


FIGURE 2. Regulation of SIX6 expression: enhancer activity and interaction with the SIX6 promoter as shown by the ChIP-seq, GRO-seq and Hi-C. Transcription factor ChIP-seq data from Jurkat cells presents the TAL1 and GATA3 binding at the proximity of the SIX6 gene. GRO-seq results show nascent transcripts, including eRNA, at this chromatin region from four T-ALL cell lines (Jurkat, MOLT-4, CCRF-CEM and P12-Ichikawa). Data above the track indicates plus strand and data below the track minus strand. ChIP-seq results of H3K4me3, H3K4me1 and H3K27ac modifications show enhancer and promoter activity at histone level in Jurkat cells as well as H3K27ac modifications in CCRF-CEM, MOLT-4 and P12-Ichikawa cells. The y-axis of sequencing data tracks shows the normalized read density. Hi-C data with long range intersection from Jurkat cells presents the interactions within the depicted chromatin region, displayed by arcs with the interaction score threshold value of 30. Color scale of the arcs runs from 0 to 400, with darker colour indicating higher interaction score. Figure covers the genomic locus Chr14:60,764,025–60,996,944 and dashed boxes highlight the SIX6 gene locus and the distal enhancer region.

H3K4me1, H3K4me3 and H3K27ac histone modifications are well established markers for enhancer and promoter activity with H3K27ac and H3K4me1 being more prominent at enhancer and H3K4me3 at promoter loci [37]. We aligned the ChIP-seq data [11-14] with the GRO-seq reads in Jurkat cells and observed a strong enrichment of H3K4me1 and H3K27ac signals at the distal enhancer element along with an enrichment of H3K4me3 signal at the SIX6 promoter. H3K27ac signal was similarly detected at the enhancer site in two other SIX6 expressing cells, MOLT-4 and CCRF-CEM, whereas only a weak signal was present at SIX6 negative P12-Ichikawa cells. To further explore whether the enhancer region associates with the SIX6 gene locus, we utilized the High-throughput Chromosome Conformation Capture (Hi-C) data[16], which confirmed a physical looping and interaction between the distal enhancer element and the promoter of SIX6 (Figure 2).

In summary, these observations suggest a direct regulation of SIX6 through an upstream enhancer element by a transcription factor complex containing TAL1 and GATA3.

3.3 High SIX6 expression defines a subgroup of T-ALL with inferior outcome

To assess the potential clinical relevance of SIX6 expression, we correlated SIX6 to clinical features in three independent patient cohorts. The first cohort consisted of 30 pediatric T-ALL patients treated according to the Nordic NOPHO ALL-2008 protocol [2, 6]. The median expression of SIX6 was used as a cut-off, and the Kaplan-Meier method was used to estimate the association of SIX6 to prognosis (Supplementary Figure S1A). Patients with higher than median SIX6 had inferior event-free (EFS) and overall survival (OS), although the difference did not reach statistical significance (Figure 3A and Supplementary Figure S2A), possibly due to the low number of cases analyzed.

The finding was replicated in a Japanese patient cohort of 119 pediatric T-ALL patients [21], where the higher than median expression of SIX6 was similarly associated with worse EFS (Log rank test, p-value = 0.056, HR 1.82, 95% CI, 0.975 to 3.39; Figure 3B).

In order to further validate the results and attain greater statistical power, we retrieved expression data and clinical information from the TARGET trial that included 264 pediatric T-ALL cases [3, 22]. We plotted SIX6 expression distribution, and categorized the data into three groups, SIX6low, SIX6medium, and SIX6high, based on the intensity of expression (see Supplementary Figure S1B). As shown in Figure 3C, SIX6 expression level was associated with EFS as SIX6low cases had EFS of 94% (95% CI, 89% to 99%), SIX6medium 90% (95% CI, 84% to 96%), and SIX6high 80% (95% CI, 70% to 92%). The difference between the SIX6low and SIX6high groups was statistically significant (Log rank test, p-value = 0.011).

The Cox proportional hazards model was fitted to the TARGET data to estimate the effect of known risk factors (age, WBC, MRD) and SIX6 and TAL1 expression. In the univariate model, high levels of SIX6 (>q80) showed significantly increased hazard (HR = 3.15, 95% CI, 1.22 to 8.14, p = 0.02). Similarly, in the multivariate model with covariates (age, WBC, MRD, SIX6, TAL1), high levels of SIX6 manifested a statistically significant predictive value independent of other clinical markers (Table 1). Notably, SIX6high was a stronger predictive marker than the end-of-induction MRD (positive but below 5%).

As the high expression of SIX6 associates with the T-ALL subgroup defined by the TAL1 gene expression, we explored whether the TAL1 had a similar (and collinear) prognostic impact. TAL1high cases showed a trend of inferior EFS (p-value = 0.31) but no difference in OS in TARGET cohort (Supplementary Figure S2D). Despite the rather robust correlation of SIX6 and TAL1 expression ($r = 0.64$, $p < 0.001$, Figure 1C), SIX6 presented as an independent risk factor in the Cox proportional hazards model, even in the presence of an SIX6/TAL1 interaction term.

3.4 Knockout of SIX6 affects critical T-ALL genes involved in mTOR, KRAS, and TNFα molecular signatures

Having established that SIX6 is associated with clinical response in T-ALL patients, we next sought to determine the potential molecular mechanism underlying this observation and employed CRISPR-Cas9 to silence the expression of SIX6 in TAL1 positive Jurkat T-ALL cell line (Figure 4A, Supplementary Figure S3).

RNA-sequencing of the isogenic wild-type and knockout Jurkat cells (SIX6-KO6) revealed 113 significantly up- or down-regulated genes with

Figure 3

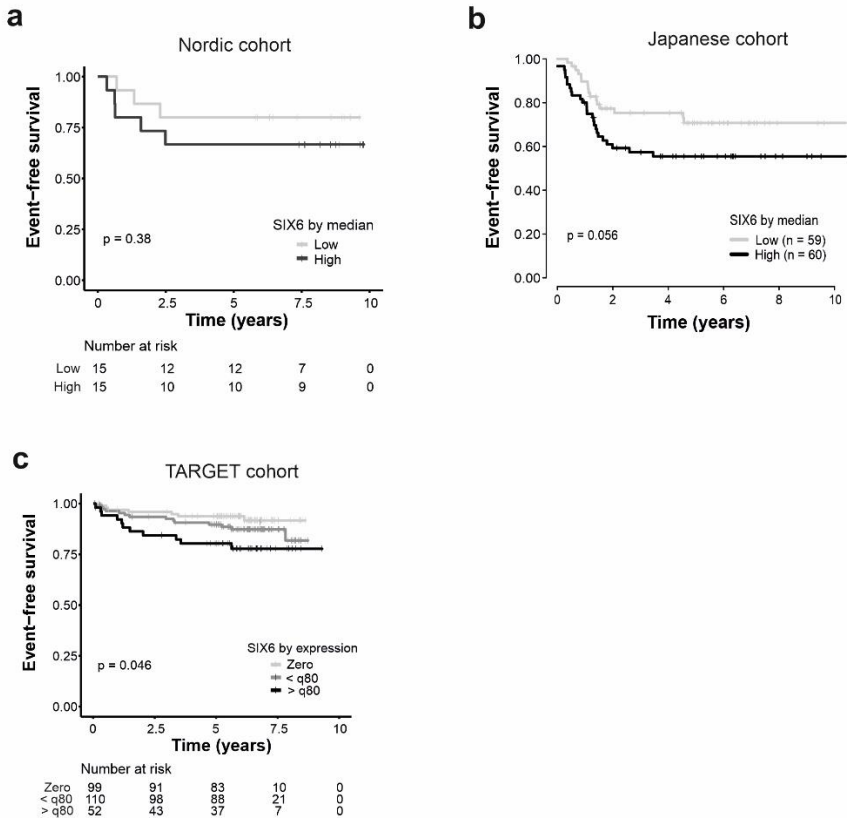


FIGURE 3. Association of *SIX6* expression with patient survival.
EFS among the SIX6^{high} and SIX6^{low} patients, based on median *SIX6* expression, in pediatric T-ALL cases from A) the Nordic cohort and B) the Japanese cohort. C) EFS among the SIX6^{low}, SIX6^{medium}, and SIX6^{high}, based on the intensity of expression, in the TARGET cohort.

LogFC of over 2 or under -2 (n=3, adjusted p-value <0.001, Figure 4B). Among the targets were interleukin 7 receptor (IL7R), which is a critical cell survival factor in T-ALL[38], Jagged Canonical Notch Ligand 1 and 2 (JAG1 and JAG2), which are part of the Notch signaling pathway[39], and cyclin D2 (CCND2). Since JAG1 is an activating ligand of Notch1, we explored whether Notch1 cleavage was affected by SIX6 knockout but found no difference between the wild-type and knockout Jurkat cells (Supplementary Figure S4).

We next sought to determine the role of SIX6 expression in the maintenance of tumor growth. Firstly, given its previous association with cell cycle regulation [25], we tested whether loss of SIX6 gene expression alone could affect cell cycle, proliferation or apoptosis. However, knockout of SIX6 did not affect any of these phenotypic features (Figure 4C-D, Supplementary Figure S5A).

Therefore, we next explored the cellular pathways affected by the SIX6 knockout by applying gene set enrichment analysis (GSEA), which manifested enrichment of mTOR-, KRAS-, and TNFα-related molecular signatures with normalized enrichment scores (NES) of 3.7, 2.0, and 2.1, respectively (FDR q-value < 0.05) (Figure 4E, Supplementary Figure S6A-B). A closer examination of the mTOR pathway uncovered statistically significant alterations in several genes (LogFC ≥ 0.5 or ≤ -0.5, n = 3, adjusted p-value < 0.001; Figure 4F). To further investigate the effect of the knockout to mTOR pathway, we treated wild-type and knockout SIX6 cells with temsirolimus. However, no significant differences in drug sensitivity were observed between the wild-type and knockout cells (Supplementary Figure S5B). We also noted statistically significant changes in the expression of KRAS- and TNFα-pathway related genes such as VEGFA, CD69, TERT, and CD1A (Supplementary Figure S6C-D). Similarly, many

Table 1. Multivariate and univariate analyses of the event-free survival based on expression level of SIX6 in pediatric T-ALL from TARGET data.

		MULTIVARIATE			UNIVARIATE		
		HR	**95% CI**	***p-value***	**HR**	**95% CI**	***p-value***
	N						
Age (years)							
<10	147	1.00 *			1.00 *		
≥10	117	0.94	0.46, 1.93	0.88	0.86	0.42, 1.74	0.67
WBC (10^9/l)							
<50	91	1.00 *			1.00 *		
≥50	173	1.32	0.58, 3.01	0.51	1.53	0.69, 3.41	0.30
MRD at EOI							
<0.1%	172	1.00 *			1.00 *		
≥0.1%	92	1.66	0.80, 3.47	0.18	1.30	0.64, 2.64	0.46
TAL1 expression							
<median [a]	132	1.00 *			1.00 *		
≥median [a]	132	1.00	0.43, 2.33	1.00	1.44	0.71, 2.92	0.31
SIX6 expression							
0	100	1.00 *			1.00 *		
>0 and ≤q80 [b]	111	1.85	0.70, 4.92	0.22	1.80	0.72, 4.45	0.21
>q80 [b]	53	3.46	1.10, 10.86	0.03 †	3.15	1.22, 8.14	0.02 †

Cox proportional-hazards regression calculated for known risk factors. SIX6 expression was treated as a factor. When the interaction term of SIX6 and TAL1 expression was introduced to the model, the effect of SIX6 remained stable.

HR: hazard ratio; CI: confidence interval; WBC: white blood cell count;
MRD: minimal residual disease; EOI: end of induction.

* marks reference groups of each categorical variable.
† marks statistically significant results ($p<0.05$).
[a] median of the TAL1 expression (Log2(FPKM + 1)) is 2.45.
[b] 80th quantile of the SIX6 expression (Log2(FPKM + 1)) is 4.85.

of the known SIX6 target genes, including CDKN2C, CDKN2D, and CDKN1B, were repressed (Supplementary Figure S6E).

As treatment resistance is a common challenge in T-ALL, we explored if the knockdown altered response to ALL chemotherapies, such as dexamethasone, prednisolone, vincristine, asparaginase and doxorubicin, as well as a tyrosine kinase inhibitor dasatinib, but observed no differences in drug responses (Supplementary Figure S5C).

SIX3 is a related member of the SIX family and shown to be able to compensate for the loss of SIX6[40]. This prompted us to determine if the lack of differential drug response was due to compensation by other SIX family members and we observed a modest increase in SIX3 expression (LogFC = 0.33, p-val = 0.019; Figure 4G), but not in others.

3.5 Overexpression of SIX6 does not accelerate progression of T-ALL in zebrafish

The SIX6 gene is an important developmental transcription factor and we therefore asked

whether ectopic expression might affect leukemia maintenance or progression. To this end, we utilized a Rag2:mMYC zebrafish model of T-ALL [41]. Expression of the endogenous zebrafish SIX6 gene was low in this model, providing a suitable tool to assess the impact of SIX6 overexpression to the progression of T-ALL (Supplementary Figure S7A). Zebrafish, co-injected with the Rag2:mMyc and the Rag2:hSIX6-TdTomato constructs at one cell stage, were compared to control zebrafish injected with the Rag2:mMYC-TdTomato construct only. The majority of the injected zebrafish in both study groups reached leukemia at 49 dpf, but no significant difference was observed between the injection groups (Figure 4H). All animals in the study developed T cell lineage leukemia (Supplementary Figure S7B).

4. Discussion

We show here that the SIX6 gene is absent in healthy differentiating or mature T-cells but aberrantly expressed in a subset of T-cell acute leukemias, the TAL1 subgroup. Its expression is regulated by the TF complex containing TAL1 and GATA3, which are a part of the "core transcriptional regulatory circuit" in T-ALL [15]. Knockdown of SIX6 caused expression changes in genes belonging to mTOR, KRAS and TNFα molecular signatures, which are closely related to T-ALL proliferation, survival and apoptosis [42-45]. However, we did not notice any major effect on cell cycle, proliferation or apoptosis in SIX6 knockout cells, or T-ALL progression in a SIX6 overexpression model in zebrafish. High levels of SIX6 were associated with an inferior EFS in three independent patient cohorts.

SIX6 is a member of SIX homeobox family of TFs [46] and these proteins can function both as transcriptional activators and repressors [25, 26, 47-52]. For example, the SIX6-DACH1 interaction is important for repressing CDKN1B, which is an inhibitor of the G1/S transition in the cell cycle [25, 52]. Under normal circumstances, SIX family TFs mainly function during eye development, but have been detected in several other developmental contexts and tissues [46, 53, 54].

Abnormal expression of SIX6 has been discovered in several cancers including T-ALL [55-61], and previous studies have associated SIX6 expression with the TAL1 subgroup [56, 59]. However, its role in T-ALL has not been thoroughly studied before. The expression of SIX6 in T-ALL exceeds that of the normal developing retina, suggesting a biologically relevant expression level with the potential to impact the phenotype of T-ALL [57]. Nagel et al. (2018) reported that overexpression of SIX6 reduced expression of the cyclin-related kinase inhibitors CDKN1A and CDKN2D in Jurkat cells. However, this did not translate into a significant change in the cell-cycle progression [58]. In our knockout model, several cyclin-related kinase inhibitors (CDKN2C, CDKN2D and CDKN1B) were repressed, but this affected neither the cell cycle nor proliferation. The deletion of SIX6 did not reduce cell viability or contribute to increased drug resistance. The redundancy of SIX6 and its subfamily member SIX3 may explain this lack of effect. SIX6 and SIX3 are together indicated as the major SIX proteins in the retina and are expressed at multiple points during the development of the eye and several other tissues [26-28, 46]. Functional and structural similarities may allow SIX6 or SIX3 to compensate for the loss of other [40].

TAL1 is a part of the core TF network in T-ALL and regulates a wide array of genes including the homeobox gene NKX3-1[57]. Previous studies have suggested NKX3-1 and its close relative, NKX3-2, as drivers of abnormal expression of SIX6 in T-ALL since the SIX6 gene has a functional binding site for these TFs in its exonic region [57, 58]. Moreover, direct regulation of NKX3-1 by the TAL1/GATA3/LMO complex has linked SIX6 mechanistically to the TAL1 subgroup. Yet, to our knowledge, the correlation between the expression of NKX3-1 and SIX6 has not been experimentally proven, and Nagel et al. (2012) reported that two out of five SIX6-positive T-ALL cell lines did not express NKX3-1. On the other hand, the silencing of NKX3-2 reduced, and upregulation increased, expression of SIX6 in T-ALL cell lines, suggesting that NKX3-2 could regulate SIX6 [58]. We present here an alternative regulatory pathway by showing that SIX6 is controlled by the TAL1 transcription factor through a direct binding to an upstream enhancer element. The same binding site is also occupied by GATA3, suggesting that the regulation takes place through a classical TF complex in T-ALL. The activity of this enhancer region was correlated with the expression of SIX6 and Hi-C data shows looping and interaction between the upstream enhancer element and the promoter of SIX6.

Figure 4

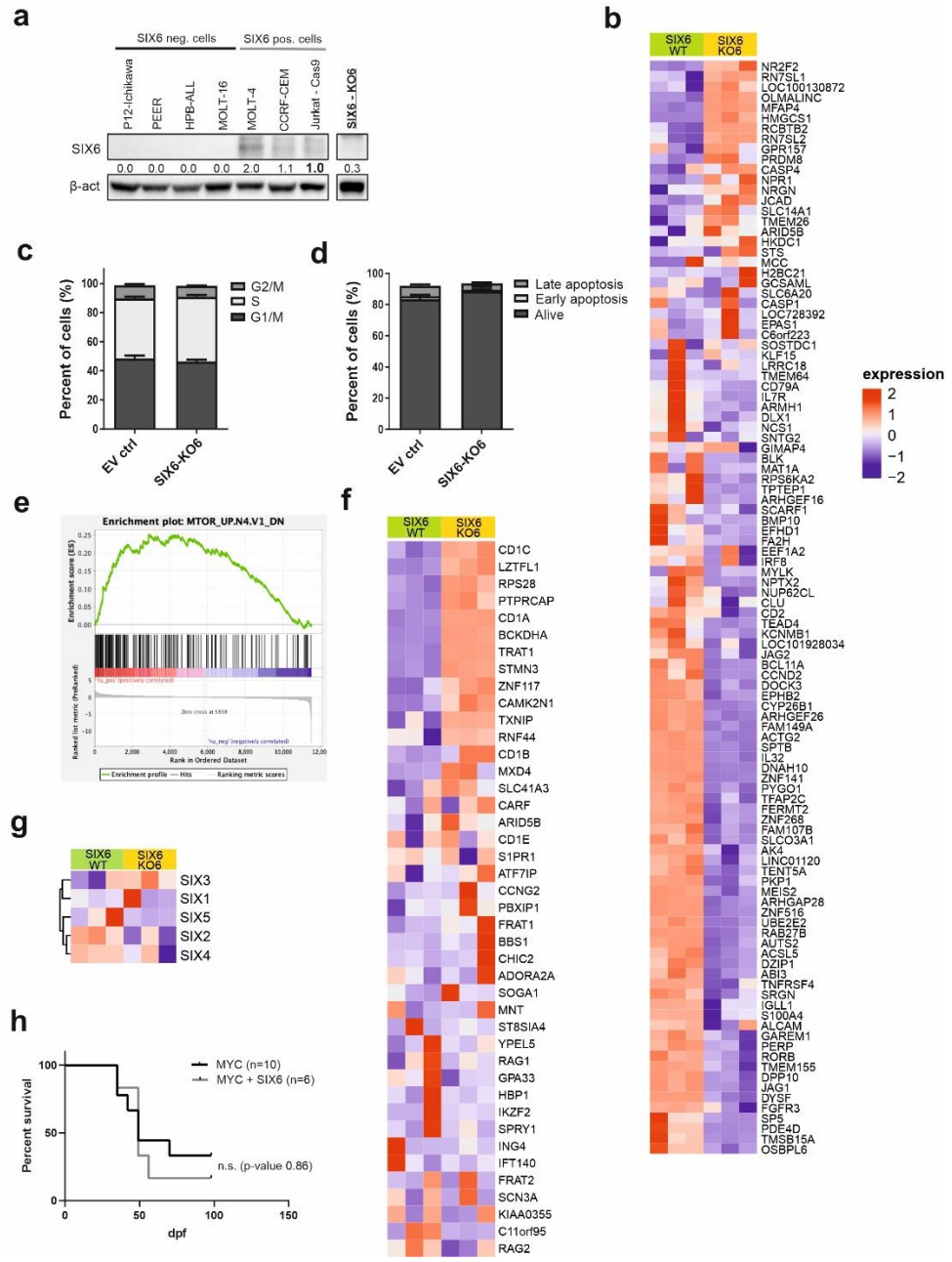


FIGURE 4. Effect of manipulating *SIX6* expression on transcriptome, proliferation and apoptosis in Jurkat cell line, and disease progression in zebrafish T-ALL model. A) Western Blot experiment showing the protein levels of SIX6 in negative control, positive control and the SIX6-KO6 knockdown cells. Quantitation is relative to the Jurkat-Cas9 cell line. B) Heatmap of genes with LogFC of ≥ 2 or ≤ -2 (n = 3, adjusted p-value < 0.001) from edgeR analyses of Jurkat wild-type and *SIX6* knockout cells. Effect of *SIX6* knockout on C) cell cycle and the number of proliferating cells (EdU assay) and D) apoptosis (AnnexinV-PI staining). Bar plots indicate the mean value and SEM (n = 3). E) Enrichment of the mTOR related molecular signature in LogFC ranked genes of *SIX6* knockout cells. F) Expression of genes from the mTOR-pathway related gene signature in the RNA-seq data from Jurkat wild-type and *SIX6* knockout cells (LogFC of ≥ 0.5 or ≤ -0.5, n = 3, adjusted p-value < 0.001. G) Expression of *SIX* family members in the RNA-seq data of Jurkat wild-type and *SIX6* knockout cells. H) Kaplan-Meier analysis of Rag2:hSIX6-TdTomato + Rag2:mMYC injected zebrafish, and Rag2:mMYC-TdTomato injected control zebrafish. Leukemic zebrafish were defined as having over 50% of their body overtaken by T-ALL cells. The number of animals analyzed per genotype is shown in parenthesis.

Furthermore, the silencing of GATA3 or TAL1 caused downregulation of SIX6 expression.

RNA-seq analysis of SIX6 knockout cells revealed changes in many known leukemia- and T-ALL-related genes, including CCND2, TFAP2C, and IL7R. In addition, several genes from the mTOR-related molecular signature showed differential expression in the SIX6 knockout cells, but they did not contribute to sensitivity to temsirolimus, an mTOR inhibitor. As SIX6 is likely part of a larger TF network, the absence of a single member of the complex may be tolerated or compensated by the cells.

Interestingly, high SIX6 was associated with inferior treatment response and prognosis in three independent T-ALL cohorts. Although the prognosis of T-ALL has improved during the past decades with fewer events such as relapse, resistant disease or death, it is still lagging behind that of B-ALL [2]. In contrast to B-ALL, where many subtypes with well-known prognostic features have been recognized, T-ALL is still lacking clinically used predictive markers [62]. In a multivariable Cox model, the three most discriminating variables were the oncogenetic classifier, MRD, and white blood cell count. Here we show that high expression of SIX6 is statistically significantly and independently associated with the EFS. Our data suggest that SIX6 could serve as a marker to flag the cases with increased risk to outcome-associated events in T-ALL.

Acknowledgements

This work was supported by grants from the Academy of Finland (O.L. and M.H. 321553, O.L., 310106, M.P. 316324 and 322010), Cancer Foundation Finland (O.L., M.H), Jane and Aatos Erkko Foundation (O.L., M.H., M.P.), Sigrid Juselius Foundation (O.L., M.H., M.P.), Finnish Hematology Association (S.L.), Tampere Tuberculosis Foundation (M.P.), The Swedish Childhood Cancer Fund (S.D.) and the Competitive State Research Financing of the Expert Responsibility Area of Tampere University Hospital (O.L. 9V033 and 9X027). Tampere University Zebrafish Core Facility was supported by funding from Biocenter Finland. We wish to thank Jorma Kulmala, Eini Eskola, Noora Hyvärinen and Jalmari Kettunen for assistance in laboratory work and bioinformatic analyses.

Author contributions

O.L., S.L, and M.H. conceived the study. S.L. and L.O. conducted experiments and analyzed the data. A.N. analyzed data. M.L processed ChIP-seq and HiC data. M.P. advised on zebrafish studies. C.D.B. generated Jurkat-Cas9 cells and provided px321-GFP plasmids. M.H. advised on bioinformatic analyses. S.D., M.S and J.T provided data and performed analyses. O.L. supervised the study. All authors reviewed and accepted the manuscript.

Disclosure of interest

The authors report no conflict of interest.

5. References

[1] C. Pui, M.V. Relling and J.R. Downing, Acute Lymphoblastic Leukemia, N.Engl.J.Med. 350 (2004), pp. 1535-1548.

[2] P. Quist-Paulsen, N. Toft, M. Heyman, et al. T-Cell Acute Lymphoblastic Leukemia in Patients 1–45 Years Treated with the Pediatric NOPHO ALL2008 Protocol, Leukemia. 34 (2020), pp. 347-357.

[3] Y. Liu, J. Easton, Y. Shao, et al. The Genomic Landscape of Pediatric and Young Adult T-Lineage Acute Lymphoblastic Leukemia, Nat.Genet. 49 (2017), pp. 1211-1218.

[4] E.V. Rothenberg, The Chromatin Landscape and Transcription Factors in T Cell Programming, Trends Immunol. 35 (2014), pp. 195-204.

[5] T. Girardi, C. Vicente, J. Cools and K. De Keersmaecker, The Genetics and Molecular Biology of T-ALL, Blood. 129 (2017), pp. 1113-1123.

[6] Z. Haider, P. Larsson, M. Landfors, et al. An Integrated Transcriptome Analysis in T-Cell Acute Lymphoblastic Leukemia Links DNA Methylation Subgroups to Dysregulated TAL1 and ANTP Homeobox Gene Expression, Cancer Med. 8 (2019), pp. 311-324.

[7] M. Heinäniemi, T. Vuorenmaa, S. Teppo, et al. Transcription-Coupled Genetic Instability Marks Acute Lymphoblastic Leukemia Structural Variation Hotspots, eLife. 5 (2016), pp. e13087.

[8] S. Teppo, S. Laukkanen, T. Liuksiala, et al. Genome-Wide Repression of eRNA and Target Gene Loci by the ETV6-RUNX1 Fusion in Acute Leukemia, Genome Res. (2016).

[9] D. Wang, I. Garcia-Bassets, C. Benner, et al. Reprogramming Transcription by Distinct Classes of Enhancers Functionally Defined by eRNA, Nature. 474 (2011), pp. 390-394.

[10] M. Kaikkonen, N. Spann, S. Heinz, et al. Remodeling of the Enhancer Landscape during Macrophage Activation is Coupled to Enhancer Transcription, Mol.Cell. 51 (2013), pp. 310-325.

[11] D. Hnisz, A.S. Weintraub, D.S. Day, et al. Activation of Proto-Oncogenes by Disruption of Chromosome Neighborhoods, Science (New York, N.Y.). 351 (2016), pp. 1454-1458.

[12] B.J. Abraham, D. Hnisz, A.S. Weintraub, et al. Small Genomic Insertions Form Enhancers that Misregulate Oncogenes, Nature communications. 8 (2017), pp. 14385.

[13] W.Z. Leong, S.H. Tan, P.C.T. Ngoc, et al. ARID5B as a Critical Downstream Target of the TAL1 Complex that Activates the Oncogenic Transcriptional Program and Promotes T-Cell Leukemogenesis, Genes Dev. 31 (2017), pp. 2343-2360.

[14] D. Orlando, M. Chen, V. Brown, et al. Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome, Cell Reports. 9 (2014), pp. 1163-1170.

[15] T. Sanda, L.N. Lawton, M.I. Barrasa, et al. Core Transcriptional Regulatory Circuit Controlled by the TAL1 Complex in Human T Cell Acute Lymphoblastic Leukemia, Cancer cell. 22 (2012), pp. 209-221.

[16] B. Lucic, H. Chen, M. Kuzman, et al. Spatially Clustered Loci with Multiple Enhancers are Frequent Targets of HIV-1 Integration, Nature communications. 10 (2019), pp. 4059.

[17] J. Park, R.A. Botting, C. Domínguez Conde, et al. A Cell Atlas of Human Thymic Development Defines T Cell Repertoire Formation, Science. 367 (2020), pp. eaay3224.

[18] J. Mehtonen, P. Pölönen, S. Häyrynen, et al. Data-Driven Characterization of Molecular Phenotypes Across Heterogeneous Sample Collections, Nucleic Acids Res. 47 (2019), pp. e76.

[19] P. Pölönen, J. Mehtonen, J. Lin, et al. Hemap: An Interactive Online Resource for Characterizing Molecular Phenotypes Across Hematologic Malignancies, Cancer Res. (2019), pp. canres.2970.2018.

[20] S. Laukkanen, T. Grönroos, P. Pölönen, et al. In Silico and Preclinical Drug Screening Identifies Dasatinib as a Targeted Therapy for T-ALL, Blood cancer journal. 7 (2017), pp. e604.

[21] M. Seki, S. Kimura, T. Isobe, et al. Recurrent SPI1 (PU.1) Fusions in High-Risk Pediatric T Cell Acute Lymphoblastic Leukemia, Nat.Genet. 49 (2017), pp. 1274-1281.

[22] K.G. Roberts, Y. Li, D. Payne-Turner, et al. Targetable Kinase-Activating Lesions in Ph-Like Acute Lymphoblastic Leukemia, N.Engl.J.Med. 371 (2014), pp. 1005-1015.

[23] J.S. Blackburn, S. Liu, D.M. Raiser, et al. Notch Signaling Expands a Pre-Malignant Pool of T-Cell Acute Lymphoblastic Leukemia Clones without Affecting Leukemia-Propagating Cell Frequency, Leukemia. 26 (2012), pp. 2069-2078.

[24] T. Grönroos, A. Mäkinen, S. Laukkanen, et al. Clinicopathological Features and Prognostic Value of SOX11 in Childhood Acute Lymphoblastic Leukemia, Scientific Reports. 10 (2020), pp. 2043.

[25] X. Li, K.A. Ohgi, J. Zhang, et al. Eya Protein Phosphatase Activity Regulates Six1–Dach–Eya Transcriptional Effects in Mammalian Organogenesis, Nature. 426 (2003), pp. 247-254.

[26] J. López-Ríos, M.E. Gallardo, S. Rodríguez de Córdoba and P. Bovolenta, Six9 (Optx2), a New Member of the Six Gene Family of Transcription Factors, is Expressed at Early Stages of Vertebrate Ocular and Pituitary Development, Mech.Dev. 83 (1999), pp. 155-159.

[27] D. Jean, G. Bernier and P. Gruss, Six6 (Optx2) is a Novel Murine Six3-Related Homeobox Gene that Demarcates the Presumptive Pituitary/Hypothalamic Axis and the Ventral Optic Stalk, Mech.Dev. 84 (1999), pp. 31-40.

[28] J. Toy and O.H. Sundin, Expression of the Optx2 Homeobox Gene during Mouse Development, Mech.Dev. 83 (1999), pp. 183-186.

[29] C.T. Kuo and J.M. Leiden, Transcriptional Regulation of T Lymphocyte Development and Function, Annu.Rev.Immunol. 17 (1999), pp. 149-187.

[30] W. Seo and I. Taniuchi, Transcriptional Regulation of Early T-Cell Development in the Thymus, Eur.J.Immunol. 46 (2016), pp. 531-538.

[31] T. Naito, H. Tanaka, Y. Naoe and I. Taniuchi, Transcriptional Control of T-Cell Development, Int.Immunol. 23 (2011), pp. 661-668.

[32] N.C. Douglas, H. Jacobs, A.L.M. Bothwell and A.C. Hayday, Defining the Specific Physiological Requirements for C-Myc in T Cell Development, Nat.Immunol. 2 (2001), pp. 307-315.

[33] M. Mingueneau, T. Kreslavsky, D. Gray, et al. The Transcriptional Landscape of Aβ T Cell Differentiation, Nat.Immunol. 14 (2013), pp. 619-632.

[34] E. Lécuyer, S. Herblot, M. Saint-Denis, et al. The SCL Complex Regulates C-Kit Expression in Hematopoietic Cells through Functional Interaction with Sp1, Blood. 100 (2002), pp. 2430-2440.

[35] M.R. Mansour, B.J. Abraham, L. Anders, et al. Oncogene Regulation. an Oncogenic Super-Enhancer Formed through Somatic Mutation of a Noncoding Intergenic Element, Science (New York, N.Y.). 346 (2014), pp. 1373-1377.

[36] A. Gardini, Global Run-on Sequencing (GRO-Seq), in Enhancer RNAs: Methods and Protocols, U.A. Ørom, ed., Springer New York, New York, NY, 2017, pp. 111-120.

[37] E. Calo and J. Wysocka, Modification of Enhancer Chromatin: What, how, and Why?, Mol.Cell. 49 (2013), pp. 825-837.

[38] M.L. Oliveira, P. Akkapeddi, D. Ribeiro, et al. IL-7R-Mediated Signaling in T-Cell Acute Lymphoblastic Leukemia: An Update, Advances in biological regulation. 71 (2019), pp. 88-96.

[39] A.A. Ferrando, The Role of NOTCH1 Signaling in T-ALL, Hematology Am Soc Hematol Educ Program. 2009 (2009), pp. 353-361.

[40] H. Xie, H.M. Hoffmann, J.D. Meadows, et al. Homeodomain Proteins SIX3 and SIX6 Regulate Gonadotrope-Specific Genes during Pituitary Development, Mol.Endocrinol. 29 (2015), pp. 842-855.

[41] D.M. Langenau, D. Traver, A.A. Ferrando, et al. Myc-Induced T Cell Leukemia in Transgenic Zebrafish, Science. 299 (2003), pp. 887.

[42] C. Evangelisti, C. Evangelisti, F. Chiarini et al. Therapeutic Potential of Targeting mTOR in T-Cell Acute Lymphoblastic Leukemia (Review), International Journal of Oncology. 45 (2014), pp. 909-918.

[43] S. Wullschleger, R. Loewith and M.N. Hall, TOR Signaling in Growth and Metabolism, Cell. 124 (2006), pp. 471-484.

[44] I.A. Prior, P.D. Lewis and C. Mattos, A Comprehensive Survey of Ras Mutations in Cancer, Cancer Res. 72 (2012), pp. 2457-2467.

[45] X. Zhou, Z. Li and J. Zhou, Tumor Necrosis Factor A in the Onset and Progression of Leukemia, Exp.Hematol. 45 (2017), pp. 17-26.

[46] J.P. Kumar, The Sine Oculis Homeobox (SIX) Family of Transcription Factors as Regulators of Development and Disease, Cellular and molecular life sciences : CMLS. 66 (2009), pp. 565-583.

[47] M. Kobayashi, K. Nishikawa, T. Suzuki and M. Yamamoto, The Homeobox Protein Six3 Interacts with the Groucho Corepressor and Acts as a Transcriptional Repressor in Eye and Forebrain Formation, Dev.Biol. 232 (2001), pp. 315-326.

[48] K.L. Kenyon, D.J. Li, C. Clouser, S. Tran and F. Pignoni, Fly Six-Type Homeodomain Proteins Sine Oculis and Optix Partner with Different Cofactors during Eye Development, Dev.Dyn. 234 (2005), pp. 497-504.

[49] C.C. Zhu, M.A. Dyer, M. Uchikawa, et al. Six3-Mediated Auto Repression and Eye Development Requires its Interaction with Members of the Groucho-Related Family of Co-Repressors, Development. 129 (2002), pp. 2835.

[50] S.J. Silver, E.L. Davies, L. Doyon and I. Rebay, Functional Dissection of Eyes Absent Reveals New Modes of Regulation within the Retinal Determination Gene Network, Mol.Cell.Biol. 23 (2003), pp. 5989-5999.

[51] J. López-Ríos, K. Tessmar, F. Loosli, et al. Six3 and Six6 Activity is Modulated by Members of the Groucho Family, Development. 130 (2003), pp. 185.

[52] X. Li, V. Perissi, F. Liu, D.W. Rose and M.G. Rosenfeld, Tissue-Specific Regulation of Retinal and Pituitary Precursor Cell Proliferation, Science. 297 (2002), pp. 1180.

[53] K. Kawakami, S. Sato, H. Ozaki and K. Ikeda, Six Family Genes—structure and Function as Transcription Factors and their Roles in Development, Bioessays. 22 (2000), pp. 616-626.

[54] R. Larder, D.D. Clark, N.L.G. Miller and P.L. Mellon, Hypothalamic Dysregulation and Infertility in Mice Lacking the Homeodomain Protein Six6, The Journal of neuroscience : the official journal of the Society for Neuroscience. 31 (2011), pp. 426-438.

[55] Q. Liu, A. Li, Y. Tian, et al. The Expression Profile and Clinic Significance of the SIX Family in

Non-Small Cell Lung Cancer, Journal of Hematology & Oncology. 9 (2016), pp. 119.

[56] B. Chen, L. Jiang, M. Zhong, et al. Identification of Fusion Genes and Characterization of Transcriptome Features in T-Cell Acute Lymphoblastic Leukemia, Proc.Natl.Acad.Sci.U.S.A. 115 (2018), pp. 373-378.

[57] S. Nagel, S. Ehrentraut, J. Tomasch, et al. Transcriptional Activation of Prostate Specific Homeobox Gene NKX3-1 in Subsets of T-Cell Lymphoblastic Leukemia (T-ALL), PloS one. 7 (2012), pp. e40747.

[58] S. Nagel, C. Meyer, M. Kaufmann, et al. Aberrant Activity of NKL Homeobox Gene NKX3-2 in a T-ALL Subset, PLOS ONE. 13 (2018), pp. e0197194.

[59] J. Soulier, E. Clappier, J. Cayuela, et al. HOXA Genes are Included in Genetic and Biologic Networks Defining Human Acute T-Cell Leukemia (T-ALL), Blood. 106 (2005), pp. 274-286.

[60] H. Xu, K. Wu, Y. Tian, et al., Expression Profile of SIX Family Members Correlates with Clinic-Pathological Features and Prognosis of Breast Cancer: A Systematic Review and Meta-Analysis, Medicine. 95 (2016), pp. e4085.

[61] M. Borssén, L. Palmqvist, K. Karrman, et al. DNA Methylation Pattern Identifies Prognostic Subgroups in Childhood T-Cell Acute Lymphoblastic Leukemia, PLOS ONE. 8 (2013), pp. e65373.

[62] R. De Smedt, J. Morscio, S. Goossens and P. Van Vlierberghe, Targeting Steroid Resistance in T-Cell Acute Lymphoblastic Leukemia, Blood Rev. 38 (2019), pp. 100591.

PUBLICATION II

Arginine Methyltransferase PRMT7 Deregulates Expression of RUNX1 Target Genes in T cell Acute Lymphoblastic Leukemia

Oksa, L., Mäkinen, A., Nikkilä, A., Hyvärinen, N., Laukkanen, S., Rokka, A., Haapaniemi, P., Seki, M., Takita, J., Kauko, O., Heinäniemi, M., & Lohi, O.

Cancers 2022, *14*(9), 2169.
https://doi.org/10.3390/cancers14092169.

cancers

MDPI

Article

Arginine Methyltransferase PRMT7 Deregulates Expression of RUNX1 Target Genes in T-Cell Acute Lymphoblastic Leukemia

Laura Oksa [1,*], Artturi Mäkinen [1,2], Atte Nikkilä [1], Noora Hyvärinen [1], Saara Laukkanen [1], Anne Rokka [3], Pekka Haapaniemi [3], Masafumi Seki [4], Junko Takita [5], Otto Kauko [3], Merja Heinäniemi [6] and Olli Lohi [1,7]

1 Tampere Center for Child, Adolescent, and Maternal Health Research, Faculty of Medicine and Health Technology, Tampere University, FI-33520 Tampere, Finland; artturi.makinen@tuni.fi (A.M.); atte.nikkila@tuni.fi (A.N.); noora.hyvarinen@tuni.fi (N.H.); saara.laukkanen@tuni.fi (S.L.); olli.lohi@tuni.fi (O.L.)
2 Fimlab Laboratories, Department of Pathology, Tampere University Hospital, FI-33520 Tampere, Finland
3 Turku Bioscience Center, University of Turku and Åbo Akademi University, FI-20014 Turku, Finland; anne.rokka@bioscience.fi (A.R.); pekka.haapaniemi@bioscience.fi (P.H.); otkauko@utu.fi (O.K.)
4 Department of Cell and Molecular Biology, Karolinska Institutet, SE-17165 Solna, Sweden; masafumi.seki@ki.se
5 Graduate School of Medicine, Kyoto University, Kyoto JP-606-8501, Japan; jtakita@kuhp.kyoto-u.ac.jp
6 The Institute of Biomedicine, University of Eastern Finland, FI-70211 Kuopio, Finland; merja.heinaniemi@uef.fi
7 Tays Cancer Center, Tampere University Hospital, FI-33520 Tampere, Finland
* Correspondence: laura.oksa@tuni.fi

Simple Summary: Approximately 15–25% of acute lymphoblastic leukemias (ALL) originate from T-lineage cells. The prognosis of T-ALL is less favorable than in B-ALL and patients experience more often relapse, which is associated with dismal survival. No established prognostic biomarkers exist for T-ALL. Here, we identified the high expression of *PRMT7* in T-ALL cells. Genetic deletion of *PRMT7* decreased the colony formation of T-ALL cells and altered arginine monomethylation patterns in protein complexes associated with RNA and DNA processing. Moreover, several proteins with an established role in the pathogenesis of T-ALL had disrupted arginine monomethylation patterns. Among them was RUNX1, whose target gene expression was consequently deregulated.

Abstract: T-cell acute lymphoblastic leukemia (T-ALL) is an aggressive hematological malignancy with no well-established prognostic biomarkers. We examined the expression of protein arginine methyltransferases across hematological malignancies and discovered high levels of *PRMT7* mRNA in T-ALL, particularly in the mature subtypes of T-ALL. The genetic deletion of *PRMT7* by CRISPR-Cas9 reduced the colony formation of T-ALL cells and changed arginine monomethylation patterns in protein complexes associated with the RNA and DNA processing and the T-ALL pathogenesis. Among them was RUNX1, whose target gene expression was consequently deregulated. These results suggest that PRMT7 plays an active role in the pathogenesis of T-ALL.

Keywords: Leukemia; T-ALL; arginine methylation; PRMT7; RUNX1

Citation: Oksa, L.; Mäkinen, A.; Nikkilä, A.; Hyvärinen, N.; Laukkanen, S.; Rokka, A.; Haapaniemi, P.; Seki, M.; Takita, J.; Kauko, O.; et al. Arginine Methyltransferase PRMT7 Deregulates Expression of RUNX1 Target Genes in T-Cell Acute Lymphoblastic Leukemia. *Cancers* **2022**, *14*, 2169. https://doi.org/10.3390/cancers14092169

Academic Editor: Ilaria Iacobucci

Received: 18 March 2022
Accepted: 23 April 2022
Published: 26 April 2022

1. Introduction

T-cell acute lymphoblastic leukemia (T-ALL) is an aggressive hematological malignancy that accounts for approximately 15–25% of pediatric and adult ALL [1]. In contrast to B-cell acute lymphoblastic leukemia (B-ALL), the prognosis of T-ALL is less favorable despite significant improvements during the last decade [2]. Five-year event-free survival (EFS) and overall survival (OS) were 74% and 75% in the NOPHO ALL2008 study [3], to 76.3% and 81.2% in the AIEOP-BFM-ALL2000 study [4], 83.8% and 89.5% in the AALL0434 study [5], and 84.6% and 90.9% in the UKALL2003 study [6]. The prognosis differs among age groups, as the five-year EFS was 80%, 74%, and 64% for age groups 1–9, 10–17, and

18–45 years in the NOPHO ALL2008 study, respectively, and showed 15–20% improvement compared to previous protocols in the study group [3]. T-ALL patients more often experience failure of induction therapy or suffer from an early relapse, which are associated with a dismal prognosis [2,3]. In the NOPHO ALL2008 study, 14% of T-ALL patients relapsed. Overall, only three out of the 39 relapsed patients were alive at the time of the last follow-up [3]. Hence, there is an evident clinical need for more efficient therapies, especially for relapsed and refractory T-ALL.

Protein Arginine Methyltransferase 7 (PRMT7) belongs to the PRMT family, which consists of nine arginine methyltransferases [7–9]. In eukaryotic cells, they catalyze reactions where a methyl group from the S-adenosyl-L-methionine is transferred to an arginine residue on protein substrates. Methylation of arginine residue increases the bulkiness of the side chain of the protein, decreasing the hydrogen bonding potential and altering the binding of methylarginine to protein modules that read the arginine methylation marks, such as the PHD finger, Tudor domain, and SH3 domain [10–14].

PRMT7 is classified as a type III arginine methyltransferase because it generates only ω-monomethyl arginine (MMA) residues to its substrates, typically at basic RXR sequences in peptides and histones. In contrast to other PRMT enzymes, PRMT7 has two AdoMet-binding sites [7] and participates in many cellular processes, including transcriptional regulation, DNA damage repair, RNA splicing, cell differentiation and proliferation, and ability to metastasize, all of which are altered in cancer [8,9,15,16]. In breast cancer, renal cell carcinoma, and non-small-cell lung cancer, the high expression of PRMT7 has been associated with metastasis or decreased survival [15–18].

PRMT7 is an attractive target protein, as there are potential inhibitors available. A cell-active chemical probe inhibited the activity of PRMT7, leading to the silencing of HSP70 family proteins [19]. Another compound, a dual PRMT5-PRMT7 inhibitor, DS-437, decreased the cell proliferation and migration of human breast cell cancer cells (MDA-MB-231) by inhibiting PRMT5 and PRMT7 activity [20].

The role of PRMT enzymes in hematological malignancies is poorly characterized. Therefore, we examined the expression of the PRMT enzymes across hematological malignancies, particularly leukemias, and focused on the role of PRMT7 in T-ALL pathogenesis and prognosis.

2. Materials and Methods

2.1. Gene Expression Analysis

The HEMAP interactive online resource (http://hemap.uta.fi/, accessed on 1 March 2019) includes gene expression data from microarray studies of various hematological and lymphoid malignancies and healthy tissues performed on a uniform technical platform (Affymetrix HG U133 Plus 2.0) [21,22]. The clustering of different T-ALL subgroups was performed as described previously by Laukkanen et al. [23,24]. The Human Cell Atlas Developmental thymus data portal (https://developmentcellatlas.ncl.ac.uk/datasets/HCA_thymus/, accessed on 9 June 2020) was used to characterize *PRMT7*, *ILR7*, *MYC*, *MYB*, *GATA3*, *TAL1*, *NOTCH1*, and *RAG2* gene expression during human T-cell development [25].

2.2. Protein Expression and Immunohistochemistry

For protein extraction, cells were lysed with CellLytic M reagent (Sigma-Aldrich, Saint Louis, MO, USA) according to the manufacturer's instructions. Protein samples were subjected to Mini-PROTEAN® TGX Stain-Free™ Precast 10% or 12% gels (Bio-Rad) and transferred to 0.2 μM PVDF membranes using the Trans-Blot Turbo Transfer Pack and Trans-Blot Turbo Transfer System (Bio-Rad). Membranes were processed using standard procedures, with PRMT7 (1:1000 in 3% BSA, #14762, Cell Signaling Technology, Danvers, MA, USA), Histone H3 (1:10,000 in 3% BSA, #9715S, Cell Signaling Technology, Danvers, MA, USA), and Horseradish peroxidase conjugated anti-rabbit (1:5000 in 3% BSA, #7074S, Cell Signaling Technology, Danvers, MA, USA) antibodies. Cell Amersham ECL Reagent (GE Healthcare, Chicago, IL, USA) was used for the chemiluminescence reaction, and

chemiluminescence was detected with ChemiDoc™ XRS+ using Image Lab™ Software (Bio-Rad, version 6.0.1, Hercules, CA, USA). PageRuler Plus prestained protein ladder (Thermo Fisher Scientific, Waltham, MA, USA) was used as a reference for protein size. Quantitation was performed using ImageJ software (version 1.59i) [26].

For the immunohistochemistry (IHC) of cell lines, cell pellets were collected from ~50 million cells via formalin fixation and paraffin embedding (FFPE). Tissue microarrays (TMA) of decalcified FFPE bone marrow trephine samples of pediatric ALL patients were used for IHC as previously described [27]. Briefly, the PRMT7 antibody (1:1000, #14762, Cell Signaling Technology, Danvers, MA, USA) and Ultraview Universal DAB kit were used for immunostaining. Then, the stained glass slides were scanned with the Hamamatsu Nanozoomer XR using a 40× magnification, and QuPath software (version 0.2.3) [28]) was applied for image analysis and the quantification of PRMT7 expression in leukemic blasts.

2.3. Cell Culture and Knockout Lines

CRISPR guide RNAs were designed with the Desktop genetics online tool (London, UK) and cloned into GFP fluorochrome marker containing vector px321-GFP (PRMT7-2: 5′-TGAACACTATGATTACCACCAGG-3′, PRMT7-3: 5′-ACCACCAGGAGATTGCAAGG-3′, PRMT7-5: 5′-AAGGCCTTGGTTCTCGACAT-3′). Jurkat cells, which stably express Cas9 enzyme (Jurkat-Cas9), were transfected with 2 μg of the plasmids (EV, PRMT7-2, PRMT7-3, and PRMT7-5) using 4D Nucleofector electroporation equipment (Lonza) and electroporation program CL-120 with the SE solution. After 24 h, GFP positive cells were single-cell sorted into 96-well plates in conditioned media that contained 20–40% of media harvested from cultured Jurkat and Jurkat-Cas9 cells and 60–80% RPMI 1640 media. *PRMT7* knockout cells for the Molt-4 line were obtained from Synthego (Synthego Corporation, Menlo Park, CA, USA). The bulk cells were single-cell sorted into 96-well plates in conditioned media containing 40% of media harvested from unmodified and modified Molt-4 cells and 60% of RPMI-1640 media. Single-cell clones were screened for mutations with T7 endonuclease (New England Biolabs, Ipswich, MA, USA) and confirmed by Sanger sequencing (DNA sequencing and genomics service, University of Helsinki, Helsinki, Finland). MutationTaster2 was used to study the functional effects of the mutations and to predict the disease-causing potential of DNA variants [29].

The Jurkat and Molt-4 cell lines and their derivatives were cultured in RPMI 1640 Medium (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) with 2 mM L-glut, 100 U penicillin, and 100 μg/mL streptomycin with 10–20% FBS (Gibco). Cas9-expressing Jurkat cells were obtained from Jan Cool's lab (VIB, Leuven, Belgium), whereas the unmodified and knockout cells from the Molt-4 line were purchased from Synthego (Synthego Corporation, Menlo Park, CA, USA). Mycoplasma testing was conducted on a regular basis for all cell lines, and STR genotyping was performed in Eurofins Genomics (Ebersberg, Germany) to authenticate the Jurkat Cas9 cell line.

2.4. Functional Studies

A Click-iT™ EdU Alexa Fluor™ 647 Flow Cytometry Assay (Invitrogen, Carlsbad, CA, USA) was used to study the cell cycle. Staining was performed according to the manufacturer's instructions, adding 1 μL of EdU to 1×10^6 cells in 1 mL and incubating at 37 °C for 1 h. Apoptosis was measured using the Annexin V Apoptosis Detection Kit-APC (eBioscience, Thermo Fisher Scientific) according to the manufacturer's instructions. EdU and apoptosis assay results were recorded using BD FACS Canto II (BD Biosciences, San Jose, CA, USA).

For cell viability assay, 10,000 cells per cell line were seeded into a 96-well plate in 100 ul of growth media and incubated for the indicated times at 37 °C in 5% CO_2. Cell viability was measured using a CellTiter-Glo Luminescent Cell Viability Assay (Promega, Madison, WI, USA) at the time points 0 h, 24 h, 48 h, and 72 h.

For colony formation analysis, MethoCult media was prepared by adding 600 μL of RPMI-1640 medium into 2.4 mL of methylcellulose-based MethoCult medium (H4230,

Stemcell Technology). Cells were harvested and added into the MethoCult medium at a concentration of 10,000 cells/mL. Thereafter, 1.1 mL of MethoCult mixture containing cells were plated into 6-well plates with 1000 cells per well. Cells were incubated at 37 °C in a 5% CO_2 incubator for 14 days. Colonies were imaged at day 10 or 14 with a Nikon AZ100 Fluorescence Macroscope (Nikon, Minato, Tokyo, Japan).

2.5. Mass Spectrometry-Based Arginine Monomethylation Analysis

For protein extraction, 1×10^8 cells were collected from J-EV, J-KO3, and J-KO5 cells (n = 6). Proteins were extracted using a Urea lysis buffer (200 mM HEPES pH 8.0 (Sigma-Aldrich), 9 M urea (Thermo Fisher Scientific), 2.5 mM sodium orthovanadate (Sigma-Aldrich), 1 mM sodium pyrophosphate (Sigma-Aldrich), 1 mM β-glycerophosphate (Cayman Chemical, Ann Arbor, MI, USA)), and lysates were sonicated (1 min, pulse 1 sec on/off with 40% amplitude, three times) using a Vibracell 500-Watt ultrasonic processor (VC 505) sonicator (3 mm microtip) (Sonics & Materials Inc., Newtown, CT, USA) to break apart the DNA. Samples were processed according to PTMScan® Mono-Methyl Arginine Motif [mme-RG] kit (Cell Signaling Technology) protocol. Shortly, proteins were digested by proteases, and the resulting peptides were purified by reversed-phase extraction. Peptides were then subjected to immunoaffinity purification. Unbound peptides were removed through washing, and the captured MMA-containing peptides were eluted with dilute acid. The enriched peptides were analyzed with a Q Exactive HF mass spectrometer (Thermo Fisher Scientific). Label free quantification was carried out in the peptide isoform level. Abundance values were calculated based on intensities of peptide precursor ions.

2.6. LC-MS/MS Analysis

The LC-ESI-MS/MS analyses were performed on a nanoflow HPLC system (Easy-nLC1200, Thermo Fisher Scientific) coupled to the Q Exactive HF mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nano-electrospray ionization source. Peptides were first loaded on a trapping column and subsequently separated inline on a 15 cm C18 column (75 µm × 15 cm, ReproSil-Pur 3 µm 120 Å C18-AQ, Dr. Maisch HPLC GmbH, Ammerbuch-Entringen, Germany). The mobile phase consisted of water with 0.1% formic acid (solvent A) and acetonitrile/water (80:20 (v/v)) with 0.1% formic acid (solvent B). Peptides were eluted with a 120 min gradient: from 2% to 39% of solvent B in 105 min, from 39% to 100% solvent B in 10 min, followed by wash for 5 min at 100% of solvent B. MS data were acquired automatically by using Thermo Xcalibur 4.1. software (Thermo Fisher Scientific). An information dependent acquisition method consisted of an Orbitrap MS survey scan of mass range m/z 300–1750 followed by HCD fragmentation for the most intense peptide ions.

Database searches were performed using Proteome Discoverer 2.4 software (Thermo Fisher Scientific) connected to in-house Mascot 2.7.0 search engine software (version 2.7.9, Matrix Science, Boston, MA, USA) against the UniProt Swiss-Prot (version 2021_02) Homo sapiens protein database [30]. Search criteria were the following: trypsin as an enzyme; cysteine carbamidomethylation as fixed modification; arginine methylation, lysine methylation, methione oxidation, and N-terminal acetylation as variable modifications; peptide mass tolerance 10 ppm; and MS/MS ion tolerance 0.02 Da. Label free quantification, validation, and filtering of the results were performed using Proteome Discoverer 2.4 software (version 2.4, Thermo Fisher Scientific, Waltham, MA, USA).

2.7. Survival Analysis

The gene expression of *PRMT7* was explored in the context of patient survival from pediatric T-ALL patients in two different cohorts (TARGET, n = 264; Japanese leukemia group, n = 119) [31–33]. The Kaplan–Meier method was used to estimate survival and the group differences were tested by using the log-rank method. The events included induction failure, disease progression, relapse, secondary malignant disease, and death by any cause. The Cox proportional hazards model was used to examine the association

of individual risk factors with patient survival. The proportionality assumption was tested with Schoenfeld residuals. The R-packages *survminer* and *survival* were used for data visualization and ordered expression graphs were plotted. The R (v. 3.6.2) software environment (R Foundation for Statistical Computing, Vienna, Austria) was used for statistical analyses.

2.8. RNA Sequencing

Library preparation and RNA sequencing of cell lines from unmodified Jurkat and Molt-4 cells and *PRMT7* knockout lines were performed by Novogene (Cambridge, UK). The quality of the raw sequencing reads was ensured with FastQC (v0.11.8). The reads were mapped to the human reference genome version hg19 using STAR aligner software (2.5.3a modified). Reads aligning to more than two locations were discarded. The alignment file was turned into tag directories, and read counts were calculated using the HOMER toolkit (v.4.11). Differential gene expression was analyzed using the quasi-likelihood F-test from edgeR R-package. Volcano plots were drawn using the R-package *ggplot2* Volcano plot function, combining the three biological replicates and using log2 fold changes values and FDR values [34].

2.9. Enrichment Analysis

Gene lists with up- and downregulated genes were analyzed using the online web server Enrichr (release March 2021) [35–37]. The analysis was performed based on gene sets from transcription factor (TF) perturbations followed by expression. In order to test whether the RUNX1 target gene expression was affected by the *PRMT7* KO, we specifically selected the significant RUNX1 related terms (GSE40155, GSE29639, GSE47375, GSE46970, GSE34292) (p-value < 0.05).

2.10. Statistical Analyses

The R (v. 4.0.2) software environment (R Foundation for Statistical Computing, Vienna, Austria) and Microsoft Excel (Redmond, WA, USA) were used in statistical analyses. The Wilcoxon test was used to compare the statistical significance of differences in gene expression of different leukemia types. All in vitro experiments were performed in three or more independent biological replicates, and the students' t-test was used to analyze the statistical significance of differences. All statistical tests were two-tailed, and a p-value < 0.05 was considered statistically significant. In mass spectrometry-based arginine monomethylation analysis, statistical testing was performed with a two-tailed Student's t-test, and p-values were adjusted following the Benjamini–Hochberg procedure [38] to account for multiple testing. Human KEGG pathway data were downloaded on 20 January 2022 from the Kyoto Encyclopedia of Genes and Genomes (KEGG) REST server using R package KEGGREST [39]. Human protein complexes were defined based on the CORUM database 3.0 release [40]. Unsupervised clustering was performed using Pearson's correlation as distance metric and Ward's linkage method.

3. Results

3.1. PRMT7 Is Strongly Expressed in Mature T-ALL

We recently published global gene expression data of various hematologic malignancies and healthy tissues comprised of 9544 expression profiles, including 4430 leukemias (385 T-ALL, 1304 pre-B-ALL, 1713 acute myeloid leukemia (AML), and 801 chronic lymphocytic leukemia (CLL), 1306 lymphomas (208 T-cell lymphoma (TCL) and 743 diffuse large B-cell lymphoma (DLBCL), and 428 healthy samples (247 T-lymphocytes and 75 B-lymphocytes) (Hemap resource (http://hemap.uta.fi/hemap/index.html, accessed on 1 March 2019), [21,22]). Using this data set, we evaluated the expression of the *PRMT* family of genes in various leukemias, lymphomas, and healthy B- and T-lymphocytes. Overall, genes of the *PRMT* family exhibited moderate and relatively even expression across hematological malignancies, with few exceptions. *PRMT3* and *PRMT5* were markedly decreased

in CLL, while *PRMT5* and *PRMT7* were increased in T-ALL (Figures 1a,b and S1a). The expression of *PRMT7* was approximately 2.6-fold higher in T-ALL compared to healthy T-lymphocytes (p-value < 0.01), whereas *PRMT5* was 1.5-fold higher in T-ALL compared to T-lymphocytes (Figure 1a and Figure S1 and Table S1).

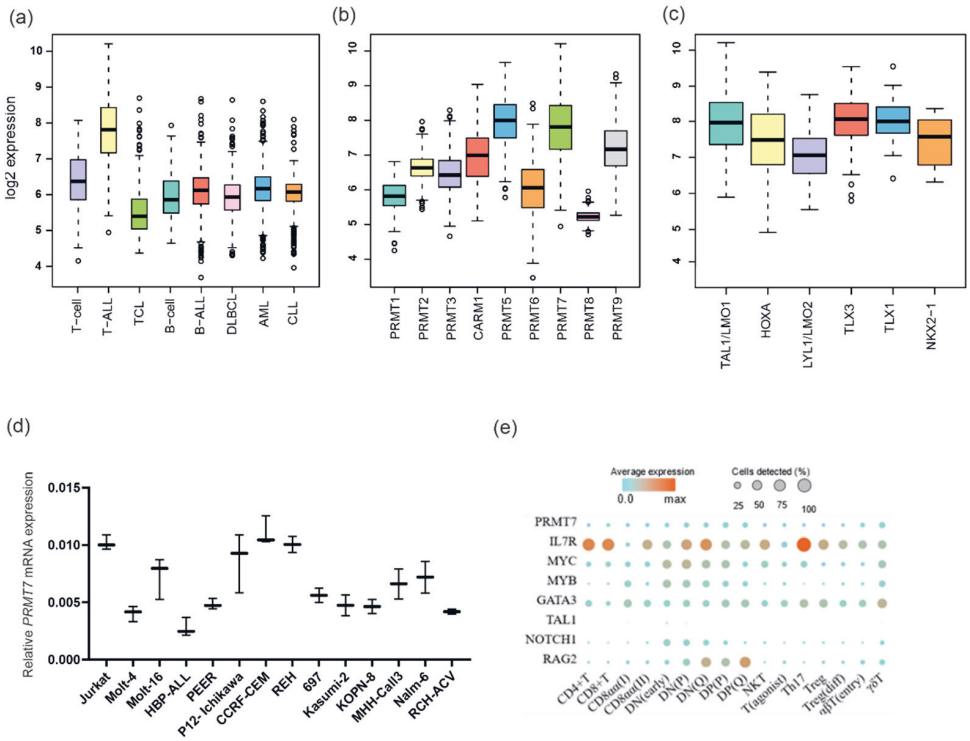


Figure 1. *PRMT7* expression in hematological malignancies and healthy T-cells. (**a**) Expression of *PRMT7* in leukemias (n = 4430), lymphomas (n = 1306), and healthy cells (n = 428) in Hemap dataset (p-values presented in Table S1); (**b**) summary of expression of all *PRMT* family genes in T-ALL; (**c**) expression of *PRMT7* in subgroups of T-ALL (*HOXA* vs. *TAL1/LMO1* p = 0.003, *HOXA* vs. *LYL1/LMO2* p = 0.026, *HOXA* vs. *TLX3* p = 0.010, *HOXA* vs. *TLX1* p = 0.006, *HOXA* vs. *NKX2-1* p = 0.845, *LYL1/LMO2* vs. *HOXA* p = 0.025, *LYL1/LMO2* vs. *TAL1/LMO1* $p = 1.098 \times 10^{-9}$, *LYL1/LMO2* vs. *TLX3* $p = 2.368 \times 10^{-7}$, *LYL1/LMO2* vs. *TLX1* $p = 9.859 \times 10^{-8}$, and *LYL1/LMO2* vs. *NKX2-1* p = 0.198, *TAL1/LMO2* n = 171, *HOXA* n = 56, *LYL1/LMO2* n = 54, *TLX3* n = 35 and *NKX2-1* n = 11); (**d**) expression of *PRMT7* mRNA in T- and B-ALL cell lines as measured by qRT-PCR (n = 3).; and (**e**) expression of *PRMT7* and the selected essential T-cell developmental genes during the T-cell differentiation and maturation process at single-cell resolution. Color indicates the mean expression and size the proportion of cells expressing the gene relative to the absolute number of cells detected in the dataset.

From here onwards, we focused our attention on *PRMT7* in T-ALL due to its strong expression in T-ALL in relation to healthy T cells. Among the genetic subgroups of T-ALL, *PRMT7* expression levels were significantly higher in the mature subtypes of T-ALL than in the more immature *LYL1/LMO2* and *HOXA*-associated T-ALL (Figure 1c). Out of the tested leukemia cell lines, T-ALL-derived cells had the highest level of *PRMT7* mRNA by RT-qPCR (Figure 1d).

At the protein level, PRMT7 expression was examined in 84 bone marrow biopsies that included T-ALL (n = 6) and B-ALL (n = 78). Two out of the six T-ALL cases showed strong staining, with 49% and 57% of the blasts positive, while two cases showed weak staining and two remained negative (Figure S1b). Weak to moderate expression of the PRMT7 protein was noted across subtypes of B-ALL (Figure S1b).

To determine whether the expression of *PRMT7* was present during the development and maturation of T-cells, we examined healthy T-lineage cells at the single-cell (sc) level. Here, we utilized a publicly available scRNA-seq data set that includes cells from the early double negative phase to the late CD4 or CD8 single positive phases, and retains mature T-regulatory, CD8αα+, and γδ T-cells [25]. Only weak expression of *PRMT7* was observed during the later stages of differentiation or maturation of T-cells (Figure 1e), suggesting that the strong expression of *PRMT7* in T-ALL is aberrant and associated with the leukemogenic process.

3.2. Association of PRMT7 Expression with Survival

We next examined whether the level of *PRMT7* mRNA was associated with patient survival using two independent patient cohorts. In the TARGET data set, which consisted of 264 T-ALL patients, *PRMT7* expression was categorized into two groups by using the median value as a cut-off. Cases with above-median expression level (*PRMT7*^{high}) had a 5-y EFS of 86% (95% CI, 81–93%), while cases with below-median level (*PRMT7*^{low}) had an EFS of 92% (95% CI, 88–97%) (p = 0.098) (Figure 2a, Table S2).

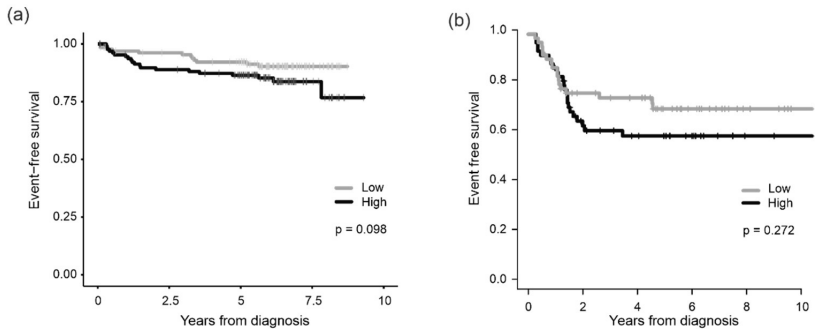


Figure 2. Association of *PRMT7* expression with patient outcome. EFS of patients in (**a**) the TARGET and (**b**) the Japanese patient cohorts. Patients were classified into two groups by using the median expression value of *PRMT7* as a cut-off.

In another patient cohort that included 119 pediatric T-ALL patients, a similar prognostic trend was observed: patients with *PRMT7*^{high} had worse EFS (Figure 2b) and increased hazard for an event (HR = 1.41, 95% CI 0.76–2.6) (p = 0.272). Neither of the survival differences reached statistical significance.

The Cox proportional hazard model was used to assess the *PRMT7* mRNA levels in conjunction with the known T-ALL risk factors (age, white blood cell count, and minimal residual disease). In the univariate model, *PRMT7*^{high} showed a trend to increased hazard for an event (HR = 1.81, 95% CI, 0.89–3.71, p = 0.1), but it did not reach statistical significance. A similar trend was observed in the multivariate model (HR = 1.77, 95% CI, 0.85–3.68, p = 0.13) (Table S2).

3.3. Reduced Colony Formation and Cell Viability after Genetic Deletion of PRMT7

Aberrant expression of *PRMT7* in T-ALL might indicate a role in leukemogenic process. Hence, we generated several knockout (KO) cells using CRISPR-Cas9 in two T-ALL cell lines, Jurkat and Molt-4. Guide RNAs targeting the first exon of the *PRMT7* gene led to indel

mutations that introduced early stop codons in the coding sequence of *PRMT7* (Figure S2). Abrogation of the PRMT7 protein expression was confirmed by immunohistochemistry and Western blotting (Figures 3a,b and S3a).

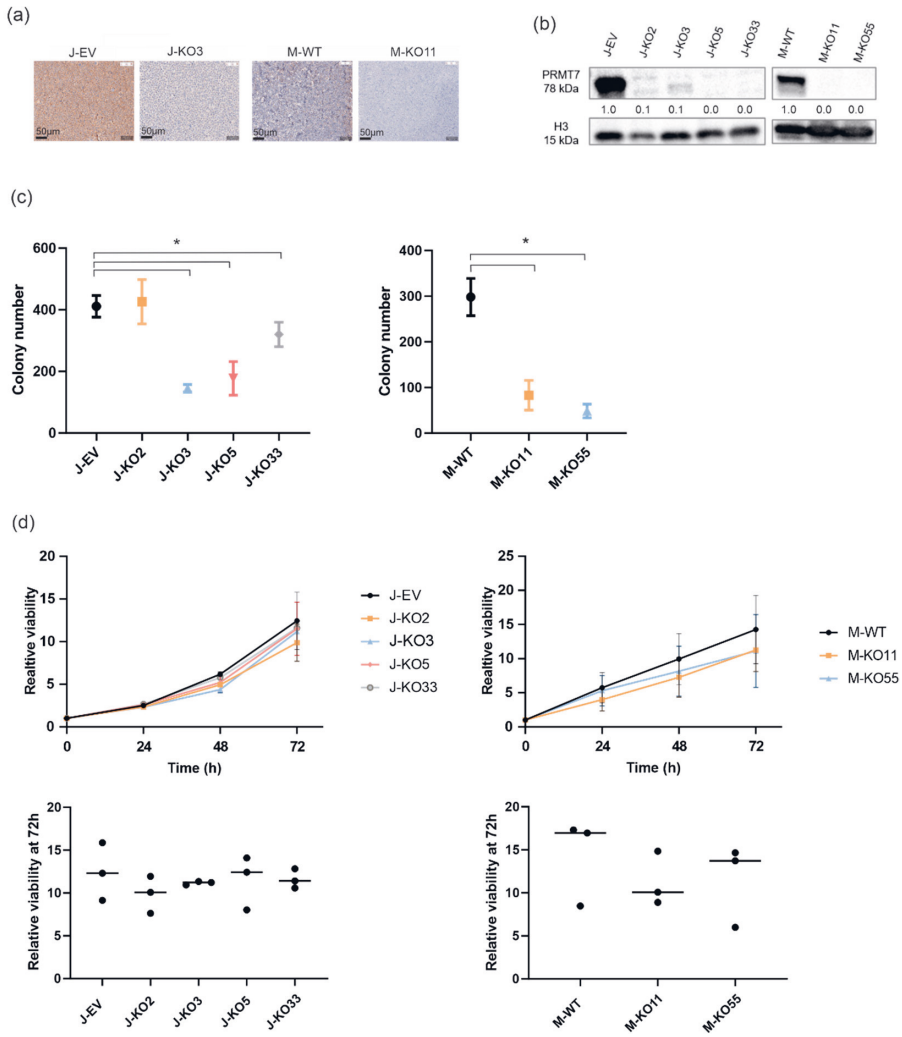


Figure 3. Functional consequences of genetic deletion of *PRMT7* in T-ALL cells. (**a**) Immunohistochemical staining (scale bar 50 µm) showing PRMT7 expression in J-EV, J-KO3, M-WT and M-KO11 cell lines. Brown color indicates positivity to the PRMT7 protein and blue marks PRMT7 negative cells. (**b**) Western blotting in unmodified (J-EV and M-WT) and knockout cell lines (marked with KO) by using the PRMT7 antibody. Quantitation in Western blot is relative to J-EV or M-EV control cell lines. J stands for Jurkat and M for Molt4 cell line, EV for empty vector and WT for wild type. The uncropped Western blots have been shown in Figure S6. Effect of the *PRMT7* knockout on (**c**) colony formation ($n = 4$) and (**d**) cell viability ($n = 3$). Cell viability results are presented as line plots for all the examined time points and as scatter plots for the 72 h time point. Asterisks denote samples with significant *p*-values (<0.05) when compared to the unmodified (J-EV and M-WT) cells.

A colony formation assay was performed to study the clonogenic capacity of Jurkat empty vector (J-EV) transduced, Molt-4 wild type (M-WT), and *PRMT7* KO cell lines. Except for one line (J-KO2), all *PRMT7* KO cells showed a significantly decreased number of colonies (p-values < 0.05) when compared to the control lines (Figures 3c and S3b). In the cell viability, apoptosis, and cell cycle assays, variation in results was evident across the Jurkat and Molt-4 derived KO lines, with no consistent changes in either direction (Figures 3d and S3b–d). The variability of the results may have been caused by residual protein expression, as shown in a report by Smith et al. (2019) [41].

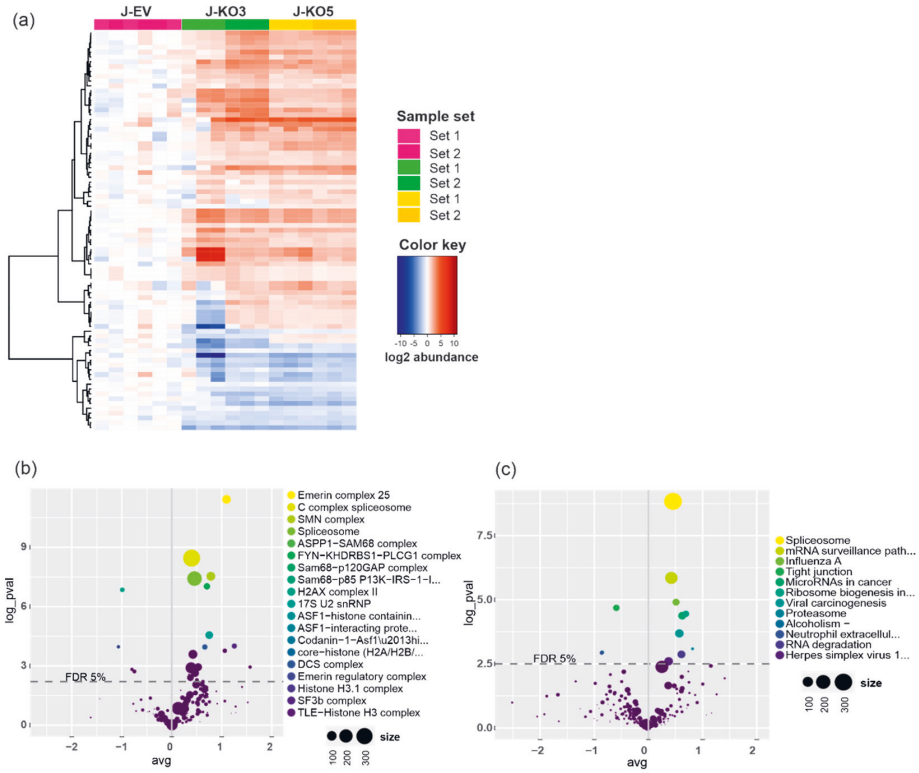
3.4. Arginine Monomethylation Is Disrupted after PRMT7 Knockout in T-ALL

PRMT7 methylates arginine residues in histone and non-histone proteins in a monomethyl manner [7]. Hence, we performed arginine monomethylation mapping in control Jurkat cells (J-EV) and Jurkat-derived *PRMT7* KO cells (J-KO3 and J-KO5). Monomethylation analysis was performed in two sets and included three biological replicates for each cell line. After the normalization of peptides to the median abundance of control, sample replicates clustered well in both sample sets and displayed high correlation, indicating the similarity of the knockout cell lines (Figures 4a and S4, Table S3). Hence, for the downstream analysis, the J-KO3 and J-KO5 data sets were combined.

We first examined global alterations in arginine monomethylation patterns in protein complexes by utilizing the Comprehensive Resource of Mammalian protein complexes (CORUM) database [40]. Average peptide methylation was statistically significantly changed in 37 protein complexes out of 243 (Benjamini–Hochberg adjusted p-value < 0.05) (Figure 4b, Table S4). Most of the differentially regulated peptides belonged to protein complexes affecting RNA and DNA processing. In total, 19 protein complexes had functional effects on RNA binding, splicing, processing, biosynthesis, or transcriptional regulation (e.g., Emerin complex 25 and C complex spliceosome), with average methylation of 18 complexes increased compared to the control cells. Complexes having a role in DNA processing, such as DNA binding, replication, repair, and synthesis, or chromatin remodeling (e.g., AFF1-histone containing and ASF1-interacting protein complex) had mostly decreased average arginine monomethylation levels. We also found methylation changes in two complexes, which play an essential role in tumor development, namely H2AX complex II and PIDDsome [42,43], both having decreased average arginine monomethylation levels in *PRMT7* KO cells.

To complement the methylation analysis, we applied the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [39]. Average peptide methylation level was significantly altered in 12 pathways (Benjamini–Hochberg adjusted p-value < 0.05), with an increase in eight and decrease in four pathways (Figure 4c, Table S5). Four of the pathways with increased methylation levels were associated with RNA processing, including spliceosome, mRNA surveillance, ribosome biogenesis, and RNA degradation pathways.

We next evaluated changes in arginine monomethylation patterns in proteins that have a well-established role in the pathogenesis of T-ALL [44]. Of particular interest were recent reports showing that changes in RUNX1 arginine methylation alter its transcriptional activity [45,46]. Indeed, *PRMT7* KO cells displayed changes in several peptides belonging to RUNX1, including a decrease in monomethylation at residue R319, a decrease or unaltered status at residue R224, and an increase at residue R223. Moreover, changes were observed in arginine monomethylation levels in peptides belonging to LEF1, LYL1, BCL11B, MYB, and NRAS proteins, all critical players in T-ALL (Figure 4d).



(d)

Methylation site	J-KO3	J-KO5
RUNX1_R319	-1.38	-1.36
LEF1_R393	-1.87	-0.38
RUNX1_R244	-1.07	0.03
LYL1_R148	0.24	0.03
BCL11B_R322	0.35	0.25
MYB_R457_R465	1.08	0.38
RUNX1_R223	0.27	0.55
NRAS_R88	0.32	0.58

Figure 4. Effects of *PRMT7* knockout on arginine monomethylation in T-ALL cells. (**a**) A heatmap of top differentially regulated peptides with altered arginine monomethylation level in *PRMT7* knockout cells (J-KO3 and J-KO5) compared to unmodified cells (J-EV) (Benjamini-Hochberg adjusted p-value ≤ 0.01). Color key indicates the average arginine monomethylation in log2 abundance. Duplicate image with row names is presented in Figure S4c. Impact of *PRMT7* knockout to arginine monomethylation by (**b**) CORUM protein complexes (Benjamini-Hochberg adjusted log2 p-value < 0.05) and (**c**) KEGG pathways (Benjamini-Hochberg adjusted log2 p-value < 0.05), with both cell lines (J-KO3 and J-KO5) combined for the analysis. (**d**) Tabulation of arginine monomethylation changes in peptides belonging to proteins associated with T-ALL pathogenesis in *PRMT7* knockout cell lines (J-KO3 and J-KO5). Red color indicates increased and blue decreased level of arginine monomethylation compared to unmodified cells.

3.5. PRMT7 Deletion Deregulates Expression of RUNX1 Target Genes

As monomethylation of arginine residues can lead to altered gene expression, we performed RNA sequencing of the knockout and control cells. In Jurkat-derived *PRMT7* KO lines, downregulation of 23 genes was noticed in all four knockout lines, and in 65 genes in at least three cell lines (log2FC $\geq \pm 1.5$, FDR < 0.05) (Figures 5a and S5a,b). Among the downregulated genes were *BCL11A* and *DNTT*, both well-known regulators of T-cell development. BCL11A acts as an oncogene in hematological malignancies, while DNTT plays role in early T-cell differentiation and V(D)J recombination [47–51]. Genes involved in propagating tumorigenesis and cell growth, namely *MDK* and *EPHA3* [52,53], or drug resistance and relapse in T-ALL, *ST8SIA6* and *NR3C2* [54–56], were silenced in the *PRMT7* KO lines. In Molt-4-derived *PRMT7* KO lines, downregulation of 24 genes was seen in both cell lines (log2FC $\geq \pm 1.5$, FDR < 0.05) (Figures 5b and S5c,d). When data on the Jurkat- and Molt-4-derived knockout lines were combined, four genes, namely *ARHGEF26* (also known as *SGEF*), *AUTS2*, *EPHA3*, and *MLC1*, were consistently silenced across all lines. Only genes in two Jurkat-derived PRMT7 KO cell lines were upregulated, i.e., MCC and NEDD9. Upregulation of MCC is associated malignant cell transformation for example in B-cells, and NEDD9 is associated to oncogenic signaling and tumor aggressiveness [57–59].

When focusing on the expression of important T-ALL associated genes, we noticed the silencing of *IL7R* gene, which harbors gain-of-function mutations in around 10% of T-ALL leading to the constitutive activation of downstream signaling [60], in five out of six *PRMT7* KO lines (log2FC $\geq \pm 1.5$, FDR < 0.05) (Figure 5c). Moreover, *PRMT7* KO led to the downregulation of cyclin D2 (*CCDN2*), which complexes with CDK4 and CDK6 to progress the cell cycle trough G1/S transition, and *NKX2-2*, a gene encoding a transcriptional activator protein (Figure 5c) [61–63].

Due to changes in the monomethylation of three arginine residues in the RUNX1 protein, we specifically asked if the expression of RUNX1 target genes was affected by the *PRMT7* deletion. Indeed, Enrichr analysis [34,35,64] revealed that in the *PRMT7* KO lines five RUNX1 target gene sets were statistically significantly changed (log2FC $\geq \pm 1.5$, *p*-value < 0.05) (Table S6).

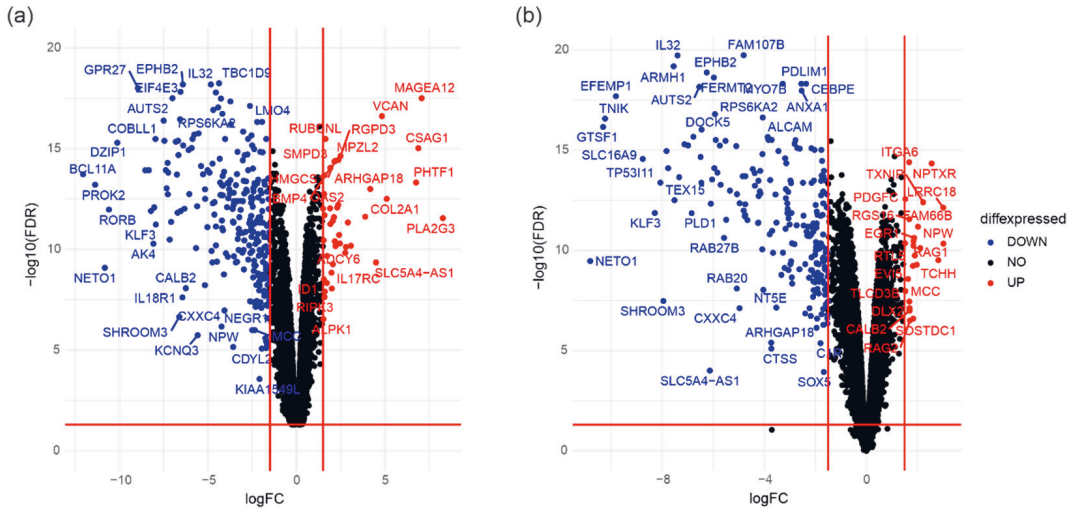


Figure 5. *Cont.*

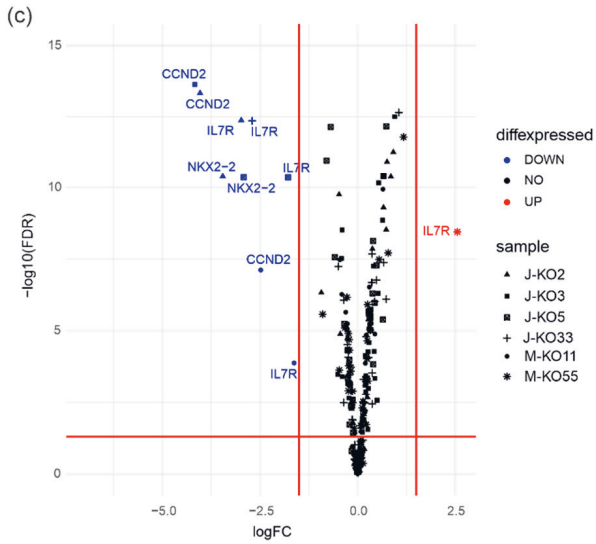


Figure 5. Effects of *PRMT7* deletion on gene expression in T-ALL cells. Volcano plots displaying the differentially expressed genes (log2FC $\geq \pm 1.5$, FDR < 0.05) in (**a**) J-KO3 and (**b**) J-KO5 cell lines. (**c**) A volcano plot displaying gene expression changes (log2FC $\geq \pm 1.5$, FDR < 0.05) in genes known to be associated with T-ALL pathogenesis and altered in all six *PRMT7* knockout cell lines labeled with different symbols. Red color indicates genes with increased and blue with decreased expression compared to unmodified cells. Three biological replicates were combined for the analysis for each cell line. J stands for Jurkat and M for Molt4 cell line, EV for empty vector and WT for wild type.

4. Discussion

T-ALL is an aggressive leukemia that accounts for approximately 10–15% of pediatric and 25% of adult ALL cases and lacks well-established prognostic biomarkers. We discovered the high expression of *PRMT7* in T-ALL compared to healthy T-cells. Jurkat and Molt-4 T-ALL cells lacking the *PRMT7* gene had diminished colony formation capability. Moreover, alterations in the monomethylation of arginine residues were observed in proteins with a well-known role in T-ALL, including RUNX1, whose target gene expression was consequently deregulated.

The proteins of the PRMT family methylate arginine residues and thereby affect many cellular processes, such as gene transcription, RNA splicing, and DNA damage response [11,14,65–69]. The methylation of arginine in histone proteins plays an important role in fine-tuning gene expression [70] and it is no surprise that it plays an important role in cancer as well. The high expression of PRMT1 has been shown to block cell differentiation and propagate leukemia, whereas PRMT4 (CRAM1) and PRMT5 affected RNA splicing in hematological malignancies [14,71]. Aberrant expression of PRMT7 has been linked with altered RNA splicing, proliferation, and colony formation in breast, colorectal, and prostate cancers [14,16]. To date, no mutations have been identified in the PRMT genes that could explain their aberrant expression in cancers [71,72].

Previously, no comprehensive studies have been conducted on *PRMT7* in leukemia and lymphoma. We found that *PRMT7* was highly expressed in T-ALL, particularly in the more differentiated subtypes when compared to healthy T-cells. Weak expression was observed during the later stages of development of healthy T-cells, indicating that the strong expression in immature T-lymphoblasts is brought about by disease associated events. Two independent patient cohorts demonstrated a decreased survival trend in patients with high levels of *PRMT7* mRNA expression, although the difference did not reach statistical significance, possibly due to the low number of events in the cohorts. It is noteworthy

that the TARGET cohort excluded patients with over 5% residual disease at the end of the induction period, thus leaving out the patients with the worst prognosis. Survival results align well with our cell modelling data, where deletion of the *PRMT7* gene decreased the oncogenic potential of the T-ALL cells as measured by colony formation assay.

As PRMT7 catalyzes the transfer of methyl groups into arginine residues, we performed global monomethyl mapping in *PRMT7* knockout cells. The most significantly affected protein complexes were involved in the RNA and DNA processing. Surprisingly, many of the complexes affecting RNA binding, splicing, biosynthesis, or transcription had increased average arginine methylation in *PRMT7* KO cells, suggesting that other enzymes with similar activity may have compensated for the loss, and perhaps in a slightly different manner. PRMT family enzymes do not exhibit strong sequence selectivity since for example PRMT5 can methylate the same proteins as PRMT7 [14]. Arginine methylation can prevent hydrogen bonding by sterically hindering the interactions between RNA and protein, thereby negatively affecting the affinity of a particular RNA-binding protein to its target. Alternatively, arginine methylation can positively regulate RNA–protein interactions, making the arginine more "hydrophobic", and thereby facilitate stacking of the bases of the RNA [73,74]. On the other hand, a majority of the complexes associated with DNA or chromatin remodeling had decreased levels of arginine monomethylation in *PRMT7* KO cells. The involvement of arginine methylation in the DNA damage response and DNA repair may cause conformational changes in DNA and trigger genomic instability [73].

We also examined the arginine monomethylation in proteins that have a well-established role in the T-ALL pathogenesis [44] and noticed minor changes in peptides belonging to LEF1, LYL1, BCL11B, MYB, and NRAS proteins. However, the most noticeable finding was associated with the RUNX1 protein, in which three arginine residues showed altered monomethylation levels. Simultaneously, the expression of the RUNX1 target genes was significantly deregulated, suggesting that modifications by PRMT7 may play an important role in fine-tuning transcriptional activity of the RUNX1 protein. This may have significant relevance in T-ALL as RUNX1 targets are involved in many critical cellular processes, especially in hematopoietic cells, such as cell differentiation, ribosome biogenesis, cell cycle regulation, and p53 and transforming growth factor beta (TGF-β) signaling pathways [29,75,76]. RUNX1 is also critical for the differentiation of lymphocytes, and its aberrant expression has been linked with various hematological malignancies [31,77]. Previously, arginine monomethylation of RUNX1 by PRMT1 in hematopoietic stem cells was associated with resistance to apoptosis and increased cell survival [46,62,78]. In myeloid cells, methylation of RUNX1 protein at residue R233 by the PRMT4 protein led to the decreased translation of the myeloid target genes, manifested as the persistent maintenance of an undifferentiated phenotype of hematopoietic progenitor cells [45].

Our results suggest that a high level of *PRMT7* may augment the aggressiveness of T-ALL. To this end, we sought an answer by looking at RNA-sequencing data in the *PRMT7* knockout cell lines. Many genes involved in drug resistance, tumor progression, or T-ALL pathogenesis were silenced in the *PRMT7* knockout cells, suggesting that the high expression of *PRMT7* is relevant for their expression. Among the silenced genes was *BCL11A*, whose overexpression is observed in many hematological cancers and mediates its oncogenic role in leukemia via target genes *Bcl2*, *Bcl2-xL*, *Mdm2*, and *p53* [47,48]. We also discovered the downregulation of three genes (*NR3C2*, *ST8SIA6*, and *ARHGEF26*) in the *PRMT7* knockout cells involved in drug resistance and relapse. NR3C2 is a glucocorticosteroid receptor and associated with NR3C1 in glucocorticoid response [56]. ST8SIA6 is a member of sialytransferases (STs), and aberrant sialylation of the cell surface glycoprotein is linked with multidrug resistance in AML and CML [54,55]. Likewise, the reduction of *ARHGEF26* sensitizes cells to chemotherapeutic agents and suppresses the colony formation of glioma cells [79]. We also observed changes in the expression of *IL7R*, *CCDN2*, and *NKX2-2*, which all have an important role in the T-ALL pathogenesis by participating in the cell cycle, growth, and transformation [60–63].

5. Conclusions

The high expression of *PRMT7* was associated with increased oncogenicity in vitro and a trend towards an inferior outcome in T-ALL in two independent patient cohorts. Alterations of arginine monomethylation after the genetic deletion of *PRMT7* were accompanied by changes in the target gene expression of the RUNX1 protein. These results suggest that PRMT7 plays an active role in the pathogenesis of T-ALL and further mechanistic studies are warranted in future.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers14092169/s1, Figure S1: Boxplots presenting expression of the *PRMT* family of genes in leukemias, lymphomas, and healthy cells, and in T- and B-lineage cell lines, Figure S2: Schematic representation of the consequences of CRISPR-Cas9-mediated genetic mutations of *PRMT7* in the Jurkat and Molt-4 cell lines, Figure S3: Images of effects of *PRMT7* knockout on colony formation, and bar plots presenting effect on apoptosis, and cell cycle, Figure S4: Plots illustrating the correlation of arginine monomethylation findings in two PRMT7 knockout cell lines (J-KO3 and J-KO5) when using the CORUM and KEGG databases for the analysis, Figure S5: Volcano plots showing effects of *PRMT7* knockout on gene expression of T-ALL cells, Figure S6: The uncropped Western blots, Table S1: Results of statistical test from HEMAP analysis where PRMT family gene expressions in T-ALL compared to other cell types, p-values are calculated using Wilcoxon rank sum test, Table S2: Multivariate and univariate analyses of the event-free survival based on the expression level of PRMT7 in pediatric T-ALL from the TARGET data set, Table S3: Top differentially regulated peptides (FDR ≤ 0.01) normalized to the median abundances of control using Pearson's correlation, Table S4: Average arginine monomethylation (abundance ratio) of CORUM protein complexes in PRMT7 KO cells (J-KO3 and J-KO5 cell lines combined for the analysis), Table S5: Average arginine methylation (abundance ratio) of KEGG pathways in PRMT7 KO cells (J-KO3 and J-KO5 cell lines combined for the analysis), Table S6: Up- and downregulated genes in PRMT7 KO cells analyzed using Enrichr. Significant terms (p-value < 0.05) summarized from TF perturbations followed by expression. Yellow color highlights the RUNX1 related gene sets.

Author Contributions: L.O., O.L. and M.H. conceived the study. L.O., S.L. and N.H. conducted the experiments and analyzed the data. A.N. participated in data analysis. A.M. collected the patient samples, performed IHC stainings, and analyzed the data. A.R. and O.K. coordinated MMA analysis. P.H. conducted the MMA experiments and O.K. analyzed the data. J.T. and M.S. provided and analyzed data. O.L. supervised the study. All authors reviewed and accepted the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grants from the Academy of Finland (O.L. and M.H. 321553, O.L., 310106 and 341540), Cancer Foundation Finland (O.L., M.H.), Jane and Aatos Erkko Foundation (O.L., M.H.), Sigrid Juselius Foundation (O.L., M.H.), Finnish Hematology Association (L.O.), and the Competitive State Research Financing of the Expert Responsibility Area of Tampere University Hospital (O.L. 9V033 and 9X027).

Institutional Review Board Statement: The study was approved by the Regional Ethics Committee of the Expert Responsibility area of Tampere University Hospital (R13109 and R19060B) and the National Supervisory Authority for Welfare and Health (Valvira, Dnro: 4243/06.01.03.01/2016 and V/3994112019).

Informed Consent Statement: Not applicable.

Data Availability Statement: The RNA-seq data generated and analyzed in the current study are available in the Gene Expression Omnibus (GEO) repository: GSE148658 (Jurkat EV control cells) GSE186238 (Jurkat Cas9 *PRMT7* KO2, Jurkat Cas9 *PRMT7* KO3, Jurkat Cas9 *PRMT7* KO5, Jurkat Cas9 *PRMT7* KO33, Molt-4 *PRMT7* KO11 and Molt-4 *PRMT7* KO55 cells). Raw MS data (Jurkat Cas9 EV control, Jurkat Cas9 *PRMT7* KO3 and Jurkat Cas9 *PRMT7* KO5 cells) is available at the jPOSTrepo (Japan ProteOme STandard Repository) with jPOSTrepo identification number JPST001540 and ProteomeXchange identification number PXD032819 [80]). Data sets reanalyzed in this study are available at: http://hemap.uta.fi/ (Hemap resource, accessed on 1 March 2019), https://portal.gdc.cancer.gov/projects (TARGET, accessed on 8 November 2019), https://developmentcellatlas.ncl.ac.uk/(Human Cell Atlas Developmental; thymus data portal accessed on 2 September 2019).

Acknowledgments: The authors acknowledge the Tampere facility of Flow Cytometry, Tampere facility of iPS Cells, Tampere Imaging Facility and the Biocenter Finland (BF) for their services. Jan Cools and Charles de Bock are acknowledged for providing the Jurkat Cas9 cells and px321-GFP -plasmids for CRISPR studies. Mass spectrometry analysis was performed at the Turku Proteomics Facility, University of Turku and Åbo Akademi University. The facility is supported by Biocenter Finland. Part of the work was carried out with the support of UEF Bioinformatics Center, University of Eastern Finland, Finland. We wish to thank Jorma Kulmala, Eini Eskola, and Gitte Gaethofs for assistance in laboratory work.

Conflicts of Interest: Declare conflicts of interest or state. The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Girardi, T.; Vicente, C.; Cools, J.; De Keersmaecker, K. The Genetics and Molecular Biology of T-ALL. *Blood* **2017**, *129*, 1113–1123. [CrossRef] [PubMed]
2. Roti, G.; Stegmaier, K. New Approaches to Target T-ALL. *Front. Oncol.* **2014**, *4*, 170. [CrossRef] [PubMed]
3. Quist-Paulsen, P.; Toft, N.; Heyman, M.; Abrahamsson, J.; Griškevičius, L.; Hallböök, H.; Jónsson, Ó.G.; Palk, K.; Vaitkeviciene, G.; Vettenranta, K.; et al. T-Cell Acute Lymphoblastic Leukemia in Patients 1–45 Years Treated with the Pediatric NOPHO ALL2008 Protocol. *Leukemia* **2020**, *34*, 347–357. [CrossRef] [PubMed]
4. Schrappe, M.; Valsecchi, M.G.; Bartram, C.R.; Schrauder, A.; Panzer-Grümayer, R.; Möricke, A.; Parasole, R.; Zimmermann, M.; Dworzak, M.; Buldini, B.; et al. Late MRD Response Determines Relapse Risk Overall and in Subsets of Childhood T-Cell ALL: Results of the AIEOP-BFM-ALL 2000 Study. *Blood* **2011**, *118*, 2077–2084. [CrossRef]
5. Winter, S.S.; Dunsmore, K.P.; Devidas, M.; Wood, B.L.; Esiashvili, N.; Chen, Z.; Eisenberg, N.; Briegel, N.; Hayashi, R.J.; Gastier-Foster, J.M.; et al. Improved Survival for Children and Young Adults with T-Lineage Acute Lymphoblastic Leukemia: Results from the Children's Oncology Group AALL0434 Methotrexate Randomization. *J. Clin. Oncol.* **2018**, *36*, 2926–2934. [CrossRef]
6. Patrick, K.; Wade, R.; Goulden, N.; Mitchell, C.; Moorman, A.V.; Rowntree, C.; Jenkinson, S.; Hough, R.; Vora, A. Outcome for Children and Young People with Early T-Cell Precursor Acute Lymphoblastic Leukaemia Treated on a Contemporary Protocol, UKALL 2003. *Br. J. Haematol.* **2014**, *166*, 421–424. [CrossRef]
7. Miranda, T.B.; Miranda, M.; Frankel, A.; Clarke, S. PRMT7 Is a Member of the Protein Arginine Methyltransferase Family with a Distinct Substrate Specificity. *J. Biol. Chem.* **2004**, *279*, 22902–22907. [CrossRef]
8. Feng, Y.; Maity, R.; Whitelegge, J.P.; Hadjikyriacou, A.; Li, Z.; Zurita-Lopez, C.; Al-Hadid, Q.; Clark, A.T.; Bedford, M.T.; Masson, J.Y.; et al. Mammalian Protein Arginine Methyltransferase 7 (PRMT7) Specifically Targets RXR Sites in Lysine- and Arginine-Rich Regions. *J. Biol. Chem.* **2013**, *288*, 37010–37025. [CrossRef] [PubMed]
9. Wang, B.; Zhang, M.; Liu, Z.; Mu, Y.; Li, K. PRMT7: A Pivotal Arginine Methyltransferase in Stem Cells and Development. *Stem Cells Int.* **2021**, *2021*, 6241600. [CrossRef]
10. Bedford, M.T.; Frankel, A.; Yaffe, M.B.; Clarke, S.; Leder, P.; Richard, S. Arginine Methylation Inhibits the Binding of Proline-Rich Ligands to Src Homology 3, but Not WW, Domains. *J. Biol. Chem.* **2000**, *275*, 16030–16036. [CrossRef]
11. Côté, J.; Richard, S. Tudor Domains Bind Symmetrical Dimethylated Arginines. *J. Biol. Chem.* **2005**, *280*, 28476–28483. [CrossRef]
12. Liu, H.; Wang, J.Y.; Huang, Y.; Li, Z.; Gong, W.; Lehmann, R.; Xu, R.M. Structural Basis for Methylarginine-Dependent Recognition of Aubergine by Tudor. *Genes Dev.* **2010**, *24*, 1876–1881. [CrossRef]
13. Liu, R.; Gao, J.; Yang, Y.; Qiu, R.; Zheng, Y.; Huang, W.; Zeng, Y.; Hou, Y.; Wang, S.; Leng, S.; et al. PHD Finger Protein 1 (PHF1) Is a Novel Reader for Histone H4R3 Symmetric Dimethylation and Coordinates with PRMT5-WDR77/CRL4B Complex to Promote Tumorigenesis. *Nucleic Acids Res.* **2018**, *46*, 6608–6626. [CrossRef] [PubMed]
14. Li, W.J.; He, Y.H.; Yang, J.J.; Hu, G.S.; Lin, Y.A.; Ran, T.; Peng, B.L.; Xie, B.L.; Huang, M.F.; Gao, X.; et al. Profiling PRMT Methylome Reveals Roles of HnRNPA1 Arginine Methylation in RNA Splicing and Cell Growth. *Nat. Commun.* **2021**, *12*, 1946. [CrossRef] [PubMed]
15. Baldwin, R.M.; Morettin, A.; Côté, J. Role of PRMTs in Cancer: Could Minor Isoforms Be Leaving a Mark? *World J. Biol. Chem.* **2014**, *5*, 115–129. [CrossRef] [PubMed]
16. Yao, R.; Jiang, H.; Ma, Y.; Wang, L.; Wang, L.; Du, J.; Hou, P.; Gao, Y.; Zhao, L.; Wang, G.; et al. PRMT7 Induces Epithelial-to-Mesenchymal Transition and Promotes Metastasis in Breast Cancer. *Cancer Res.* **2014**, *74*, 5656–5667. [CrossRef] [PubMed]
17. Liu, F.; Wan, L.; Zou, H.; Pan, Z.; Zhou, W.; Lu, X. PRMT7 Promotes the Growth of Renal Cell Carcinoma through Modulating the β-Catenin/C-MYC Axis. *Int. J. Biochem. Cell Biol.* **2020**, *120*, 105686. [CrossRef] [PubMed]
18. Cheng, D.; He, Z.; Zheng, L.; Xie, D.; Dong, S.; Zhang, P. PRMT7 Contributes to the Metastasis Phenotype in Human Non-Small-Cell Lung Cancer Cells Possibly through the Interaction with HSPA5 and EEF2. *Onco. Targets. Ther.* **2018**, *11*, 4869–4876. [CrossRef]
19. Szewczyk, M.M.; Ishikawa, Y.; Organ, S.; Sakai, N.; Li, F.; Halabelian, L.; Ackloo, S.; Couzens, A.L.; Eram, M.; Dilworth, D.; et al. Pharmacological Inhibition of PRMT7 Links Arginine Monomethylation to the Cellular Stress Response. *Nat. Commun.* **2020**, *11*, 2396. [CrossRef]

20. Smil, D.; Eram, M.S.; Li, F.; Kennedy, S.; Szewczyk, M.M.; Brown, P.J.; Barsyte-Lovejoy, D.; Arrowsmith, C.H.; Vedadi, M.; Schapira, M. Discovery of a Dual PRMT5-PRMT7 Inhibitor. *ACS Med. Chem. Lett.* **2015**, *6*, 408–412. [CrossRef]
21. Mehtonen, J.; Pölönen, P.; Häyrynen, S.; Dufva, O.; Lin, J.; Liuksiala, T.; Granberg, K.; Lohi, O.; Hautamäki, V.; Nykter, M.; et al. Data-Driven Characterization of Molecular Phenotypes across Heterogeneous Sample Collections. *Nucleic Acids Res.* **2019**, *47*, e76. [CrossRef] [PubMed]
22. Pölönen, P.; Mehtonen, J.; Lin, J.; Liuksiala, T.; Häyrynen, S.; Teppo, S.; Mäkinen, A.; Kumar, A.; Malani, D.; Pohjolainen, V.; et al. Hemap: An Interactive Online Resource for Characterizing Molecular Phenotypes across Hematologic Malignancies. *Cancer Res.* **2019**, *79*, 2466–2479. [CrossRef] [PubMed]
23. Laukkanen, S.; Grönroos, T.; Pölönen, P.; Kuusanmäki, H.; Mehtonen, J.; Cloos, J.; Ossenkoppele, G.; Gjertsen, B.; Øystein, B.; Heckman, C.; et al. In Silico and Preclinical Drug Screening Identifies Dasatinib as a Targeted Therapy for T-ALL. *Blood Cancer J.* **2017**, *7*, e604. [CrossRef] [PubMed]
24. Laukkanen, S.; Oksa, L.; Nikkilä, A.; Lahnalampi, M.; Parikka, M.; Seki, M.; Takita, J.; Degerman, S.; de Bock, C.E.; Heinäniemi, M.; et al. SIX6 Is a TAL1-Regulated Transcription Factor in T-ALL and Associated with Inferior Outcome. *Leuk. Lymphoma* **2020**, *61*, 3089–3100. [CrossRef] [PubMed]
25. Park, J.E.; Botting, R.A.; Domínguez Conde, C.; Popescu, D.M.; Lavaert, M.; Kunz, D.J.; Goh, I.; Stephenson, E.; Ragazzini, R.; Tuck, E.; et al. A Cell Atlas of Human Thymic Development Defines T Cell Repertoire Formation. *Science* **2020**, *367*, eaay3224. [CrossRef]
26. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* **2012**, *9*, 671–675. [CrossRef] [PubMed]
27. Mäkinen, A.; Nikkilä, A.; Haapaniemi, T.; Oksa, L.; Mehtonen, J.; Vänskä, M.; Heinäniemi, M.; Paavonen, T.; Lohi, O. IGF2BP3 Associates with Proliferative Phenotype and Prognostic Features in B-Cell Acute Lymphoblastic Leukemia. *Cancers* **2021**, *13*, 1505. [CrossRef] [PubMed]
28. Bankhead, P.; Loughrey, M.B.; Fernández, J.A.; Dombrowski, Y.; McArt, D.G.; Dunne, P.D.; McQuaid, S.; Gray, R.T.; Murray, L.J.; Coleman, H.; et al. QuPath: Open Source Software for Digital Pathology Image Analysis. *Sci. Rep.* **2017**, *7*, 16878. [CrossRef]
29. Schwarz, J.M.; Cooper, D.N.; Schuelke, M.; Seelow, D. MutationTaster2: Mutation Prediction for the Deep-Sequencing Age. *Nat. Methods* **2014**, *11*, 361–362. [CrossRef] [PubMed]
30. Bateman, A.; Martin, M.J.; Orchard, S.; Magrane, M.; Agivetova, R.; Ahmad, S.; Alpi, E.; Bowler-Barnett, E.H.; Britto, R.; Bursteinas, B.; et al. UniProt: The Universal Protein Knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489. [CrossRef]
31. Sood, R.; Kamikubo, Y.; Liu, P. Role of RUNX1 in Hematological Malignancies. *Blood* **2017**, *129*, 2070–2082. [CrossRef] [PubMed]
32. Seki, M.; Kimura, S.; Isobe, T.; Yoshida, K.; Ueno, H.; Nakajima-Takagi, Y.; Wang, C.; Lin, L.; Kon, A.; Suzuki, H.; et al. Recurrent SPI1 (PU.1) Fusions in High-Risk Pediatric T Cell Acute Lymphoblastic Leukemia. *Nat. Genet.* **2017**, *49*, 1274–1281. [CrossRef] [PubMed]
33. Roberts, K.G.; Li, Y.; Payne-Turner, D.; Harvey, R.C.; Yang, Y.L.; Pei, D.; McCastlain, K.; Ding, L.; Lu, C.; Song, G.; et al. Targetable Kinase-Activating Lesions in Ph-like Acute Lymphoblastic Leukemia. *N. Engl. J. Med.* **2014**, *371*, 1005–1015. [CrossRef] [PubMed]
34. Wickham, H. *Ggplot2*; Springer: New York, NY, USA, 2016. [CrossRef]
35. Chen, E.Y.; Tan, C.M.; Kou, Y.; Duan, Q.; Wang, Z.; Meirelles, G.V.; Clark, N.R.; Ma'ayan, A. Enrichr: Interactive and Collaborative HTML5 Gene List Enrichment Analysis Tool. *BMC Bioinform.* **2013**, *14*, 128. [CrossRef]
36. Kuleshov, M.V.; Jones, M.R.; Rouillard, A.D.; Fernandez, N.F.; Duan, Q.; Wang, Z.; Koplev, S.; Jenkins, S.L.; Jagodnik, K.M.; Lachmann, A.; et al. Enrichr: A Comprehensive Gene Set Enrichment Analysis Web Server 2016 Update. *Nucleic Acids Res.* **2016**, *44*, W90–W97. [CrossRef]
37. Xie, Z.; Bailey, A.; Kuleshov, M.V.; Clarke, D.J.B.; Evangelista, J.E.; Jenkins, S.L.; Lachmann, A.; Wojciechowicz, M.L.; Kropiwnicki, E.; Jagodnik, K.M.; et al. Gene Set Knowledge Discovery with Enrichr. *Curr. Protoc.* **2021**, *1*, e90. [CrossRef]
38. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. [CrossRef]
39. Tenenbaum, D.; Maintainer, B. *KEGGREST: Client-Side REST Access to the Kyoto Encyclopedia of Genes and Genomes (KEGG)*; R Package Version 1.35.0; version: Release (3.14); Bioconductor: Dortmund, Germany, 2021; Available online: https://bioconductor.org/packages/release/bioc/html/KEGGREST.html (accessed on 1 March 2019). [CrossRef]
40. Giurgiu, M.; Reinhard, J.; Brauner, B.; Dunger-Kaltenbach, I.; Fobo, G.; Frishman, G.; Montrone, C.; Ruepp, A. CORUM: The Comprehensive Resource of Mammalian Protein Complexes—2019. *Nucleic Acids Res.* **2019**, *47*, D559. [CrossRef]
41. Smits, A.H.; Ziebell, F.; Joberty, G.; Zinn, N.; Mueller, W.F.; Clauder-Münster, S.; Eberhard, D.; Fälth Savitski, M.; Grandi, P.; Jakob, P.; et al. Biological Plasticity Rescues Target Activity in CRISPR Knock Outs. *Nat. Methods* **2019**, *16*, 1087–1093. [CrossRef]
42. Meador, J.A.; Zhao, M.; Su, Y.; Narayan, G.; Geard, C.R.; Balajee, A.S. Histone H2AX Is a Critical Factor for Cellular Protection against DNA Alkylating Agents. *Oncogene* **2008**, *27*, 5662–5671. [CrossRef]
43. Manzl, C.; Peintner, L.; Krumschnabel, G.; Bock, F.; Labi, V.; Drach, M.; Newbold, A.; Johnstone, R.; Villunger, A. PIDDosome-Independent Tumor Suppression by Caspase-2. *Cell Death Differ.* **2012**, *19*, 1722. [CrossRef] [PubMed]
44. Hnisz, D.; Weintrau, A.S.; Day, D.S.; Valton, A.L.; Bak, R.O.; Li, C.H.; Goldmann, J.; Lajoie, B.R.; Fan, Z.P.; Sigova, A.A.; et al. Activation of Proto-Oncogenes by Disruption of Chromosome Neighborhoods. *Science* **2016**, *351*, 1454–1458. [CrossRef] [PubMed]
45. Mizutani, S.; Yoshida, T.; Zhao, X.; Nimer, S.D.; Taniwaki, M.; Okuda, T. Loss of RUNX1/AML1 Arginine-Methylation Impairs Peripheral T Cell Homeostasis. *Br. J. Haematol.* **2015**, *170*, 859–873. [CrossRef] [PubMed]

46. Zhao, X.; Jankovic, V.; Gural, A.; Huang, G.; Pardanani, A.; Menendez, S.; Zhang, J.; Dunne, R.; Xiao, A.; Erdjument-Bromage, H.; et al. Methylation of RUNX1 by PRMT1 Abrogates SIN3A Binding and Potentiates Its Transcriptional Activity. *Genes Dev.* **2008**, *22*, 640–653. [CrossRef] [PubMed]
47. Yin, J.; Xie, X.; Ye, Y.; Wang, L.; Che, F. BCL11A: A Potential Diagnostic Biomarker and Therapeutic Target in Human Diseases. *Biosci. Rep.* **2019**, *39*, 1–13. [CrossRef] [PubMed]
48. Yu, Y.; Wang, J.; Khaled, W.; Burke, S.; Li, P.; Chen, X.; Yang, W.; Jenkins, N.A.; Copeland, N.G.; Zhang, S.; et al. Bcl11a Is Essential for Lymphoid Development and Negatively Regulates P53. *J. Exp. Med.* **2012**, *209*, 2467. [CrossRef] [PubMed]
49. Yin, B.; Delwel, R.; Valk, P.J.; Wallace, M.R.; Loh, M.L.; Shannon, K.M.; Largaespada, D.A. A Retroviral Mutagenesis Screen Reveals Strong Cooperation between Bcl11a Overexpression and Loss of the Nf1 Tumor Suppressor Gene. *Blood* **2009**, *113*, 1075–1085. [CrossRef]
50. Saadi, W.; Kermezli, Y.; Dao, L.T.M.; Mathieu, E.; Santiago-Algarra, D.; Manosalva, I.; Torres, M.; Belhocine, M.; Pradel, L.; Loriod, B.; et al. A Critical Regulator of Bcl2 Revealed by Systematic Transcript Discovery of LncRNAs Associated with T-Cell Differentiation. *Sci. Rep.* **2019**, *9*, 4707. [CrossRef]
51. Rothenberg, E.V. Transcriptional Control of Early T and B Cell Developmental Choices. *Annu. Rev. Immunol.* **2014**, *32*, 283. [CrossRef]
52. Filippou, P.S.; Karagiannis, G.S.; Constantinidou, A. Midkine (MDK) Growth Factor: A Key Player in Cancer Progression and a Promising Therapeutic Target. *Oncogene* **2019**, *39*, 2040–2054. [CrossRef]
53. Charmsaz, S.; Al-Ejeh, F.; Yeadon, T.M.; Miller, K.J.; Smith, F.M.; Stringer, B.W.; Moore, A.S.; Lee, F.T.; Cooper, L.T.; Stylianou, C.; et al. EphA3 as a Target for Antibody Immunotherapy in Acute Lymphoblastic Leukemia. *Leukemia* **2016**, *31*, 1779–1787. [CrossRef] [PubMed]
54. Ma, H.; Zhou, H.; Song, X.; Shi, S.; Zhang, J.; Jia, L. Modification of Sialylation Is Associated with Multidrug Resistance in Human Acute Myeloid Leukemia. *Oncogene* **2015**, *34*, 726–740. [CrossRef] [PubMed]
55. Zhang, X.; Dong, W.; Zhou, H.; Li, H.; Wang, N.; Miao, X.; Jia, L. α-2,8-Sialyltransferase Is Involved in the Development of Multidrug Resistance via PI3K/Akt Pathway in Human Chronic Myeloid Leukemia. *IUBMB Life* **2015**, *67*, 77–87. [CrossRef] [PubMed]
56. Li, B.; Brady, S.W.; Ma, X.; Shen, S.; Zhang, Y.; Li, Y.; Szlachta, K.; Dong, L.; Liu, Y.; Yang, F.; et al. Therapy-Induced Mutations Drive the Genomic Landscape of Relapsed Acute Lymphoblastic Leukemia. *Blood* **2020**, *135*, 41–55. [CrossRef] [PubMed]
57. Edwards, S.K.E.; Baron, J.; Moore, C.R.; Liu, Y.; Perlman, D.H.; Hart, R.P.; Xie, P. Mutated in Colorectal Cancer (MCC) Is a Novel Oncogene in B Lymphocytes. *J. Hematol. Oncol.* **2014**, *7*, 56. [CrossRef]
58. Gabbasov, R.; Xiao, F.; Howe, C.G.; Bickel, L.E.; O'Brien, S.W.; Benrubi, D.; Do, T.V.; Zhou, Y.; Nicolas, E.; Cai, K.Q.; et al. NEDD9 Promotes Oncogenic Signaling, a Stem/Mesenchymal Gene Signature, and Aggressive Ovarian Cancer Growth in Mice. *Oncogene* **2018**, *37*, 4854–4870. [CrossRef]
59. Shagisultanova, E.; Gaponova, A.V.; Gabbasov, R.; Nicolas, E.; Golemis, E.A. Preclinical and Clinical Studies of the NEDD9 Scaffold Protein in Cancer and Other Diseases. *Gene* **2015**, *567*, 1. [CrossRef]
60. Ribeiro, D.; Melão, A.; Barata, J.T. IL-7R-Mediated Signaling in T-Cell Acute Lymphoblastic Leukemia. *Adv. Biol. Regul.* **2013**, *53*, 211–222. [CrossRef]
61. Clappier, E.; Cuccuini, W.; Cayuela, J.M.; Vecchione, D.; Baruchel, A.; Dombret, H.; Sigaux, F.; Soulier, J. Cyclin D2 Dysregulation by Chromosomal Translocations to TCR Loci in T-Cell Acute Lymphoblastic Leukemias. *Leukemia* **2006**, *20*, 82–86. [CrossRef]
62. Nagel, S.; MacLeod, R.A.F.; Pommerenke, C.; Meyer, C.; Kaufmann, M.; Drexler, H.G. NKL Homeobox Gene NKX2-2 Is Aberrantly Expressed in Hodgkin Lymphoma. *Oncotarget* **2018**, *9*, 37480. [CrossRef]
63. Nagel, S.; Pommerenke, C.; Scherr, M.; Meyer, C.; Kaufmann, M.; Battmer, K.; MacLeod, R.A.F.; Drexler, H.G. NKL Homeobox Gene Activities in Hematopoietic Stem Cells, T-Cell Development and T-Cell Leukemia. *PLoS ONE* **2017**, *12*, e0171164. [CrossRef] [PubMed]
64. Vu, L.P.; Perna, F.; Wang, L.; Voza, F.; Figueroa, M.E.; Tempst, P.; Erdjument-Bromage, H.; Gao, R.; Chen, S.; Paietta, E.; et al. PRMT4 Blocks Myeloid Differentiation by Assembling a Methyl-RUNX1-Dependent Repressor Complex. *Cell Rep.* **2013**, *5*, 1625–1638. [CrossRef] [PubMed]
65. Yang, Y.; Lu, Y.; Espejo, A.; Wu, J.; Xu, W.; Liang, S.; Bedford, M.T. TDRD3 Is an Effector Molecule for Arginine-Methylated Histone Marks. *Mol. Cell* **2010**, *40*, 1016–1023. [CrossRef] [PubMed]
66. Sims, R.J., 3rd; Rojas, L.A.; Beck, D.B.; Bonasio, R.; Schüller, R.; Drury, W.J., 3rd; Eick, D.; Reinberg, D. The C-Terminal Domain of RNA Polymerase II Is Modified by Site-Specific Methylation. *Science* **2011**, *332*, 99–103. [CrossRef] [PubMed]
67. Wu, J.; Xu, W. Histone H3R17me2a Mark Recruits Human RNA Polymerase-Associated Factor 1 Complex to Activate Transcription. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5675–5680. [CrossRef] [PubMed]
68. Lee, Y.H.; Bedford, M.T.; Stallcup, M.R. Regulated Recruitment of Tumor Suppressor BRCA1 to the P21 Gene by Coactivator Methylation. *Genes Dev.* **2011**, *25*, 176–188. [CrossRef]
69. Durant, S.T.; Cho, E.C.; La Thangue, N.B. P53 Methylation–the Arg-Ument Is Clear. *Cell Cycle* **2009**, *8*, 801–802. [CrossRef]
70. Cheung, N.; Fung, T.K.; Zeisig, B.B.; Holmes, K.; Rane, J.K.; Mowen, K.A.; Finn, M.G.; Lenhard, B.; Chan, L.C.; So, C.W. Targeting Aberrant Epigenetic Networks Mediated by PRMT1 and KDM4C in Acute Myeloid Leukemia. *Cancer Cell* **2016**, *29*, 32–48. [CrossRef]

71. Shah, M.A.; Denton, E.L.; Arrowsmith, C.H.; Lupien, M.; Schapira, M. A Global Assessment of Cancer Genomic Alterations in Epigenetic Mechanisms. *Epigenetics Chromatin* **2014**, *7*, 29. [CrossRef]
72. Wu, Q.; Schapira, M.; Arrowsmith, C.H.; Barsyte-Lovejoy, D. Protein Arginine Methylation: From Enigmatic Functions to Therapeutic Targeting. *Nat. Rev. Drug Discov.* **2021**, *20*, 509–530. [CrossRef]
73. Bedford, M.T.; Richard, S. Arginine Methylation: An Emerging Regulatorof Protein Function. *Mol. Cell* **2005**, *18*, 263–272. [CrossRef]
74. Blackwell, E.; Ceman, S. Arginine Methylation of RNA-Binding Proteins Regulates Cell Function and Differentiation. *Mol. Reprod. Dev.* **2012**, *79*, 163. [CrossRef] [PubMed]
75. Cai, X.; Gao, L.; Teng, L.; Ge, J.; Oo, Z.M.; Kumar, A.R.; Gilliland, D.G.; Mason, P.J.; Tan, K.; Speck, N.A. Runx1 Deficiency Decreases Ribosome Biogenesis and Confers Stress Resistance to Hematopoietic Stem and Progenitor Cells. *Cell Stem Cell* **2015**, *17*, 165–177. [CrossRef]
76. Chuang, L.S.H.; Ito, K.; Ito, Y. RUNX Family: Regulation and Diversification of Roles through Interacting Proteins. *Int. J. Cancer* **2013**, *132*, 1260–1271. [CrossRef] [PubMed]
77. Look, A.T. Oncogenic Transcription Factors in the Human Acute Leukemias. *Science* **1997**, *278*, 1059–1064. [CrossRef] [PubMed]
78. Matsumura, T.; Nakamura-Ishizu, A.; Muddineni, S.; Tan, D.Q.; Wang, C.Q.; Tokunaga, K.; Tirado-Magallanes, R.; Sian, S.; Benoukraf, T.; Okuda, T.; et al. Hematopoietic Stem Cells Acquire Survival Advantage by Loss of RUNX1 Methylation Identified in Familial Leukemia. *Blood* **2020**, *136*, 1919–1932. [CrossRef]
79. Fortin Ensign, S.P.; Roos, A.; Mathews, I.T.; Dhruv, H.D.; Tuncali, S.; Sarkaria, J.N.; Symons, M.H.; Loftus, J.C.; Berens, M.E.; Tran, N.L. SGEF Is Regulated via TWEAK/Fn14/NF-KB Signaling and Promotes Survival by Modulation of the DNA Repair Response to Temozolomide. *Mol. Cancer Res.* **2016**, *14*, 302–312. [CrossRef]
80. Okuda, S.; Watanabe, Y.; Moriya, Y.; Kawano, S.; Yamamoto, T.; Matsumoto, M.; Takami, T.; Kobayashi, D.; Araki, N.; Yoshizawa, A.C.; et al. JPOSTrepo: An International Standard Data Repository for Proteomes. *Nucleic Acids Res.* **2017**, *45*, D1107–D1111. [CrossRef] [PubMed]

PUBLICATION III

Genomic determinants of therapy response in *ETV6::RUNX1* leukemia

Oksa, L., Moisio S., Maqbool, K., Kramer, R. Nikkilä, A., Foroughi-Asl H., Suhonen, Krali, O., Voutilainen, M., Lahnalampi, M., J., Duque-Afonso, J., Hauer J., Vepsäläinen, K., Nordlund, J., Wirta, V., Lohi O.* & Heinäniemi M.*

Submitted manuscript