Tampere University

TONI GRÖNROOS

Transcriptional Regulation and Cell Signaling in Acute Lymphoblastic Leukemia and Hematopoiesis

Tampere University Dissertations 74

Tampere University Dissertations 74

TONI GRÖNROOS

Transcriptional Regulation and Cell Signaling in Acute Lymphoblastic Leukemia and Hematopoiesis

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty Council of the Faculty of Medicine and Health Technology
of Tampere University,
for public discussion in the Jarmo Visakorpi auditorium
of the Arvo building, Arvo Ylpön katu 34, Tampere,
on 14 June 2019, 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 Mika Rämet Tampere University Finland	
Pre-examiners	Docent Outi Kilpivaara University of Helsinki Finland	Docent Einari Niskanen University of Jyväskylä Finland
Opponent	Docent Matti Korhonen University of Helsinki Finland	

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

Cover design: Roihu Inc.

ISBN 978-952-03-1117-9 (print)
ISBN 978-952-03-1118-6 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-1118-6

PunaMusta Oy – Yliopistopaino
Tampere 2019

ACKNOWLEDGEMENTS

The research for this thesis was carried out at the Tampere Center for Child Health Research, Faculty of Medicine and Health Technology, Tampere University and Tampere University Hospital, Finland, during 2010-2019. I wish to give my deepest gratitude to all the people who have helped and advised me with my work during these years. I am very thankful for the financial support from the Väre Foundation for Pediatric Cancer Research, Finnish Cultural Foundation: Pirkanmaa Regional Fund, Tampere University, Finnish Hematology Association, Blood Disease Research Foundation and the Scientific Foundation of the City of Tampere.

I am truly grateful to my supervisors Docent Olli Lohi, MD, PhD, and Professor Mika Rämet, MD, PhD, for the opportunity to work with interesting research projects. Olli, you gave me the opportunity to join your research group and I am grateful for your support over the years. Without your help, encouragement and patience during these years this would not have been possible. I admire your enthusiasm and positivity towards scientific work. Mika, I wish to thank you for your support and guidance during my thesis years. You had always time to stop for a minute and ask me how it is going. That, combined with all the questions and comments during the years, were priceless and motivated me to go forward. Thank you both for everything.

I wish to thank the official reviewers of this thesis, Docent Outi Kilpivaara, PhD, and Docent Einari Niskanen, PhD, for their great advice and valuable comments, which improved this thesis. I also like to thank the members of my thesis supervision committee, Professor Seppo Parkkila, MD, PhD, and Professor Kim Vettenranta MD, PhD, for their guidance and advice regarding my projects. It was always a pleasure to sit down and discuss the current states of my thesis projects with you.

I wish to thank all my co-authors and collaborators regarding their contribution to the original communications of this thesis. Without your contribution, this work would not have been possible to accomplish. I would especially like to thank Juha Mehtonen, MSc, and Associate Professor Merja Heinäniemi, PhD, for their invaluable contribution and for patiently answering my numerous questions. Joel Johnson, MSc, is acknowledged for improving the language of this thesis.

From all my heart, I want to thank the present and past members of HemoRes group: Susanna Teppo, MSc, Saara Laukkanen, MSc, Laura Oksa, MSc, Artturi Mäkinen, MD, Atte Pentikäinen, MD, Miikka Voutilainen, BSc and Hanna Korkeamäki, MD. It has been a privilege to work with you all and learn from you over the years. From my early years as a HemoRessian I want to warmly thank Kaisa Teittinen, PhD, for the guidance and patience when I was merely starting my life as a researcher. You were a mentor to me.

I want to thank also all my other co-workers and colleagues that I have had the pleasure to work. ISE group members, especially Mikko Oittinen, MSc, Joel Johnson, MSc, docent Keijo Viiri, PhD, and past and present CeliRes group members, Laura Airaksinen, MSc, Minna Hietikko, MSc, Suvi Kalliokoski, PhD, Tiina Rauhavirta, PhD and Heidi Kontro, PhD, are thanked for the peer support and the fun times on and off duty. Zebrafish core facility personnel and all the other people from the fishlab are warmly acknowledged for all the help and guidance over the years. I owe my gratitude also to Associate Professor Merja Heinäniemis group members from all the collaborations on and off work. I wish to thank Juha Määttä, PhD, for inspiring and supporting me to pursuing PhD. I especially wish to express my deepest gratitude to Mr. Jorma Kulmala for all the help and support.

I have been fortunate to have so many good friends in my life who have supported me through everything. Mypa, Emppu, Teme, Tuure, Sussu, Heidi, Risto and all other fellow members of the Northern Tamperes toughest guy association (Pohjois-Tampereen Kovin Jätkä ry). Lifetime friends from my time in Jyväskylä, especially Kaitsu, Mikko, Juha, Jarski, Iko, Antti and Anu.

I am grateful of my family and loved ones. Mummi, Äit, Titta, Jussi, Pena, Jouko, Virpi, Virpin Jussi, Jukka and Marita. It is impossible to put in words how much your support and trust have ment to me through all these years. From all my heart, thank you. I especially want to thank you, Juha, for keeping us connected to each others' lives by calling and meeting me whenever you had the time and for being my brother.

Finally, I would like to thank Sanni for your never-ending support. You trusted me and carried me through the last steps of this thesis process when I had lost all the strength and hope to finalize it. I love you.

Tampere, May 2019

Toni Grönroos

ABSTRACT

Leukemia originates from blood-forming tissue and it occurs in people of all ages. Various genetic and epigenetic alterations drive the malignant transformation of progenitor cells resulting in aberrant cell proliferation, lack of differentiation and extended survival. Acute leukemias can be divided into either myeloid or lymphoid diseases, based on the blood cell lineage from where the leukemic cells arise. Based on the malignant precursor from where acute lymphoblastic leukemia (ALL) originates, ALLs can be further subdivided into B cell precursor ALL (BCP-ALL) or T cell ALL (T-ALL) which often carry either chromosomal aberrations or distinct gene expression patterns. Unlike in BCP-ALL, distinct T-ALL subtypes do not serve as prognostic markers. ALL is more common in pediatric patients, with the peak incidence at 2-5 years of age. ALL survival in pediatric patients has improved significantly during the last decades but there is still prognostic variation across subtypes so that some subtypes carry higher relapse risk. Relapse remains the most common treatment failure in pediatric ALL and the poor response to early therapy is the strongest predictor of relapse risk.

Transcriptional regulation of gene expression and cell signaling are critical players in cell differentiation, proliferation and survival. We sought to identify aberrantly expressed genes that show subtype-specific expression in ALL. Among these were *PTP4A3* and *SOX11*. PTP4A3 is a dual-specificity phosphatase that can inactivate targets by dephosphorylating either serine, threonine or tyrosine residues. From the three PTP4A family members, *PTP4A3* demonstrated a subtype-specific expression in ETV6-RUNX1 and BCR-ABL1 acute leukemias. We showed that ETV6-RUNX1 fusion regulated *PTP4A3* expression but no marked changes were seen in cell viability after knockdown of *PTP4A3* gene or inhibition of its enzymatic activity.

Developmental and neuronal transcription factor SOX11 was aberrantly expressed in ETV6-RUNX1 and TCF3-PBX1 subtypes of BCP-ALL. Knockdown of SOX11 led to alteration in the expression of genes typically associated with cell adhesion, migration and differentiation. High *SOX11* expression was associated with DNA hypomethylation at *SOX11* locus and favourable clinical outcome. These results indicate that SOX11 expression marks a group of patients with good outcome and encourage further studies of its use as a prognostic biomarker.

The prognosis of T-ALL has also improved during the last decade, but when relapsed, T-ALL patients rarely survive. We performed *in silico* drug screening across 4430 leukemia samples in order to find novel targeted therapies for T-ALL by matching leukemia gene expression data with drug target gene expressions. As a result, strong expression of *LCK*, a Src family tyrosine kinase, was detected in T-ALL samples and *in silico* screening suggested dasatinib as a targeted therapy. This was confirmed by *in vitro* and *ex vivo* experimental studies in cell lines and patient samples where 30 % of the studied cases showed a response to dasatinib, suggesting the potential of dasatinib drug in the treatment of T-ALL patients.

Finally, we studied the Sin3A-associated protein SAP30L, which is part of the Sin3A corepressor complex and regulates gene expression by stabilizing protein-protein and protein-DNA interactions. Zebrafish (*Danio rerio*) was chosen as a model organism for studying the role of SAP30L as it only expresses SAP30L and not SAP30. SAP30L showed ubiquitous expression in both embryonic and adult zebrafish tissues, and morpholino-mediated *SAP30L* knockdown resulted in cardiac abnormalities and reduced hemoglobin levels of red blood cells. A microarray analysis of *SAP30L* morphant gene expression revealed down-regulation of the heart-specific *nkx2.5* gene and genes important for hemoglobin synthesis and erythropoiesis. Results demonstrated that SAP30L regulates cardiac development and erythropoiesis in zebrafish.

In conclusion, this thesis identifies novel regulators of transcription (SOX11, SAP30L) and cell signaling (PTP4A3) in acute leukemia and hematopoiesis, and discover a novel potential targeted therapy for a subgroup of T-ALL patients.

TIIVISTELMÄ

Leukemia kehittyy verisoluja muodostavassa kudoksessa ja sitä tavataan kaiken ikäisillä potilailla. Useat geneettiset ja epigeneettiset muutokset voivat aiheuttaa häiriöitä normaalien luuytimen varhaisvaiheen solujen erilaistumisessa, jakaantumisessa ja elinkyvyssä. Akuutit leukemiat voidaan jakaa myelooiseen ja lymfaattiseen tyyppiin riippuen siitä, mistä verisolulinjasta leukeemiset solut ovat saaneet alkunsa. Akuutit lymfaattiset leukemiat (ALL) voidaan puolestaan jakaa varhaisten B-solujen (B-ALL) tai T-solujen (T-ALL) leukemioihin. B-ALL-potilaille on tunnusomaisia erilaiset kromosomaaliset poikkeavuudet, joita voidaan käyttää myös ennusteellisina merkkeinä. T-ALL:ssa puolestaan tavataan aliryhmäkohtaisia geenien ilmentymismuutoksia. ALL on yleisempi lapsipotilailla ja tyypillinen esiintymispiikki ilmenee 2-5 vuoden iässä. Vaikka ALL-potilaiden selviytymisaste on parantunut huomattavasti viime vuosikymmeninä, on eri alityyppien välisissä ennusteissa edelleen merkittäviä eroja. Lasten ALL:ssa taudin uusiutuminen onkin yleisin syy hoidon epäonnistumiselle ja hidas alkuhoitovaste ennustaa tarkinten taudin uusiutumisriskiä.

Geenien ilmentymisen säätely ja solujen signalointi ovat merkittävässä roolissa solujen erilaistumisessa, sekä niiden elin- ja kasvukyvyssä. Tarkastelimme leukemioissa ja niiden alityypeissä poikkeavasti ilmentyneitä geenejä ja tunnistimme tällaisiksi *PTP4A3* ja *SOX11* geenit. PTP4A3 on fosfataasi, joka muokkaa kohdeproteiiniensa toimintaa fosforitähteen poistolla seriini-, treoniini- tai tyrosiiniaminohaposta. Kolmesta PTP4A proteiiniperheen jäsenistä *PTP4A3* ilmentyi leukemiassa lähinnä ETV6-RUNX1- ja BCR-ABL1-alityypeissä. ETV6-RUNX1-proteiinifuusion todettiin säätelevän *PTP4A3*:n ilmentymistä, mutta solujen kasvukyvyssä ei havaittu merkittävää muutosta *PTP4A3* geenin hiljentämisen tai proteiinin entsymaattisen aktiivisuuden estämisen seurauksena.

SOX11 on geenien ilmentymisen säätelytekijä (transkriptiotekijä), jonka toiminta on erityisen tärkeää aikaisessa sikiön kehittymisessä. *SOX11* geenin ilmentymisen havaittiin olevan poikkeavaa ALL:n ETV6-RUNX1- ja TCF3-PBX1-alityypeissä. *SOX11* geenin hiljentämisen seurauksena havaittiin ilmentymisen muutoksia geeneissä, jotka liittyvät solujen kiinnittymiseen, liikkumiseen ja erilaistumiseen. *SOX11* geenin korkea ilmenemistaso oli yhteydessä *SOX11* geenialueen vähäiseen

metylaatiotasoon ja parempaan kliiniseen ennusteeseen. Tulokset viittaavat siihe, että SOX11 proteiinia voitaisi hyödyntää hyväennusteisen potilasryhmän tunnistamisessa.

T-ALL:n ennuste on parantunut viimeisen vuosikymmenen aikana, mutta taudin uusiutuessa potilaiden selviytymisennuste on heikko. Laajalle, 4430 leukemianäytettä sisältävälle aineistolle tehdyllä *in silico*-lääkeaineseulonnalla pyrittiin löytämään uusia kohdennettuja T-ALL:n lääkehoitoja. Lääkeaineseulonnassa yhdistettiin leukemiassa tavattujen geenien ilmentyminen lääkeaineen kohdegeenien ilmentymiseen. Tuloksena havaittiin Src proteiiniperheeseen kuuluvan tyrosiinikinaasin, *LCK*:n, voimakas ilmentyminen, jolle lääkeaineseulonnan perusteella mahdolliseksi kohdennetuksi lääkkeeksi soveltuisi dasatinib. *In vitro-* ja *ex vivo*-kokeet solulinjoilla ja potilasnäytteillä vahvistivat dasatinibin tehon noin 30 %:lla tutkituissa tapauksissa. Tulokset viittaisivat siihen, että dasatinibin käyttöä T-ALL:ssa kannattaa selvittää jatkotutkimuksilla.

SAP30L (Sin3A-associated protein 30-like) proteiini kuuluu Sin3A-proteiinikompleksiin, joka toimii pääasiassa geenien ilmentymisen hiljentäjänä stabiloimalla joko proteiini-proteiini tai proteiini-DNA interaktioita. Seeprakala (*Danio rerio*) ilmentää vain SAP30L proteiinia (eikä sukulaisproteiini SAP30:ta), mikä tekee siitä erinomaisen valinnan tutkittaessa kyseisten proteiinien toimintaa. Seeprakalassa SAP30L ilmentyy kaikkialla sekä poikasvaiheen että aikuisen yksilön kudoksissa. Morfoliino-välitteisen *SAP30L*-geenin hiljentämisen seurauksena havaittiin muutoksia sydämen morfologiassa sekä hemoglobiinitasoissa. DNA-mikrosiruanalyysi osoitti *SAP30L*-morfanteissa ilmentymisen heikentymistä sydänspesifisessä *nkx2.5*. geenissä, sekä hemoglobiinisynteesin ja punasolujen kehittymisen kannalta tärkeissä geeneissä. Tulokset osoittavat SAP30L säätelevän sydämen kehittymistä sekä punasolujen muodostumista seeprakalassa.

Tässä väitöskirjassa löydettiin uusia leukemian ja verenmuodostuksen geenien ilmentymisen säätelytekijöitä (SOX11, SAP30L) sekä poikkeava ilmentyminen solusignaloinnin (PTP4A3) säätelyproteiinissa. Lisäksi tutkimuksissa tunnistettiin uusi mahdollinen täsmälääkehoito T-soluiseen akuuttiin leukemiaan.

CONTENTS

List of Figures

List of Tables

List of Appendices

ABBREVIATIONS

AD	activation domain
AGM	aorta-gonad-mesonephros
ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia
B-ALL	B cell acute lymphoblastic leukemia
BCP-ALL	B cell precursor acute lymphoblastic leukemia
BCR	B cell receptor
cDNA	complementary deoxyribonucleic acid
ChIP	chromatin immunoprecipitation
CHT	caudal hematopoietic tissue
CLP	common lymphoid progenitor
CML	chronic myeloid leukemia
CMP	common myeloid progenitor
CNA	copy number alterations
DMSO	dimethyl sulfoxide
DN	double negative
DNA	deoxyribonucleic acid
DP	double positive
dpf	days post fertilization
DsigDB	drug signature database
DSS	drug signature scores
EFS	event-free survival
FFPE	formalin fixed paraffin embedded
FISH	fluorescence *in situ* hybridization
GO	gene ontology
GRO-seq	global nuclear run-on sequencing
HD	homeodomain
HDAC	histone deacetylase
HLH	helix-loop-helix
HMG	high-mobility group

hpf	hours post fertilization
HSC	hematopoietic stem cells
IC50	half-maximal growth inhibition concentration
ICM	intermediate cell mass
ID	inhibitory domain
Igh	immunoglobulin heavy chain
Igl	immunoglobulin light chain
IHC	immunohistochemical
IKZF1	IKAROS family zinc finger 1
IL7	interleukin 7
LMO	LIM-domain-only
MCL	mantle cell lymphoma
MHC	major histocompatibility complex
MO	morpholino oligonucleotides
MPP	multipotent progenitor
MRD	minimal residual disease
mRNA	messenger RNA
OS	overall survival
PNT	pointed domain
PRL	phosphatases of regenerating liver
PTP4A	protein tyrosine phosphatase 4A
RD	runt domain
RFS	relapse-free survival
RNA	ribonucleic acid
RNA-seq	ribonucleic acid sequencing
RT-qPCR	quantitative reverse transcription polymerase chain reaction
SAP	sin3A associated protein
SAP30L	sin3A associated protein 30 -like
SFK	Src family kinase
shRNA	small hairpin ribonucleic acid
siRNA	small interfering ribonucleic acid
STR	short tandem repeats
TAD	transactivation domain
TAL1	T cell acute lymphoblastic leukemia 1
T-ALL	T cell acute lymphoblastic leukemia
TCR	T cell receptor

TEC	thymic epithelial cell
WB	western blot
WGS	whole genome sequencing
WHO	World Health Organization
WISH	whole-mount *in situ* hybridization

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original communications, referred to in the text by Roman numerals (I-IV).

Publication I **Grönroos T**, Teppo S, Mehtonen J, Laukkanen S, Liuksiala T, Nykter M, Heinäniemi M, Lohi O. Overexpression of PTP4A3 in ETV6-RUNX1 acute lymphoblastic leukemia. Leuk Res. 2017. 54:1-6.

Publication II **Grönroos T***, Mäkinen A*, Laukkanen S, Mehtonen J, Nikkilä A, Oksa L, Rounioja S, Marincevic-Zuniga Y, Nordlund J, Pohjolainen V, Paavonen T, Heinäniemi M, Lohi O. High expression of SOX11 is associated with favourable prognosis in childhood acute lymphoblastic leukemia. *Submitted manuscript.*

Publication III Laukkanen S, **Grönroos T**, Pölönen P, Kuusanmäki H, Mehtonen J, Cloos J, Ossenkoppele G, Gjertsen B, Øystein B, Heckman C, Heinäniemi M, Kontro M, Lohi O. *In silico* and preclinical drug screening identifies dasatinib as a targeted therapy for T-ALL. Blood Cancer J. 2017. 7(9):e604.

Publication IV Teittinen KJ, **Grönroos T**, Parikka M, Junttila S, Uusimäki A, Laiho A, Korkeamäki H, Kurppa K, Turpeinen H, Pesu M, Gyenesei A, Rämet M, Lohi O. SAP30L (Sin3A-associated protein 30-like) is involved in regulation of cardiac development and hematopoiesis in zebrafish embryos. J Cell Biochem. 2012. 113(12):3843-3852.

* Authors contributed equally.

1 INTRODUCTION

Leukemia is liquid cancer that originates from the blood-forming cells. The clinical behavior of the disease can be either slow ("chronic") or fast ("acute"). While the chronic forms of leukemia are predominant in adults, the acute forms comprise the vast majority of cases in childhood. A cell of origin determines the type of leukemia, lymphoblastic or myeloid, and, again, there is a distinction between adults and children with the lymphoblastic form being the predominant in childhood by comprising approximately 85 % of acute leukemias (Inaba et al. 2013). Approximately 80-85 % of acute lymphoblastic leukemias (ALL) are of B cell origin while T cell ALLs are relatively rare in early childhood but becomes more common later in adolescents and adults (Tasian and Hunger, 2017). Malignant transformation of hematopoietic progenitor cells results from genetic and epigenetic alterations during lymphocyte development leading to changes in the normal cell differentiation, proliferation and survival (Van Vlierberghe and Ferrando, 2012). In childhood ALL, the first genetic alteration occurs often already *in utero*, while the disease onset requires additional genetic events after birth (Greaves and Wiemels, 2013). Current treatment of ALL consists of a long chemotherapy and survival rates in pediatric leukemias have risen to approximately 90 % level. However, some genetic subtypes still have a high risk of relapse (Oskarsson et al., 2016; Pierro et al., 2017). In addition, slow or poor early response to chemotherapy is a reliable marker for less favourable prognosis and increased risk of relapse (O`Connor et al., 2018). A number of studies have probed the genomic diversity of ALL at diagnosis, clonal evolution during therapy and their association to relapse (Mullighan et al., 2008; Hogan et al., 2011; Ma et al., 2015). While the prognosis of childhood ALL has improved markedly, there are plenty of unanswered questions related to disease pathogenesis. Clarity in pathogenesis could lead tho the design of more targeted therapies to reduce chemotherapy burden and treatment of relapsed patients. This thesis seeks to provide more detailed understanding of leukemia pathobiology to advance disease understanding and thereby lead to novel treatment opportunities.

2 REVIEW OF THE LITERATURE

2.1 Hematopoiesis

The process where blood cells are formed is called hematopoiesis, which in vertebrates is divided in two phases, primitive and definitive waves (Galloway and Zon, 2003). Primitive hematopoiesis is transitory and its primary function is erythrocyte production from the erythroid progenitors, which appear from the blood islands in the extraembryonic yolk sac in mammals. The erythroid progenitors in the primitive hematopoiesis are not pluripotent and do not have renewal capability (Orkin and Zon, 2008; Palis and Yoder, 2001). Definitive hematopoiesis follows later in the development giving rise to a multipotent hematopoietic stem cells (HSCs) which have potential to develop and differentiate into all blood cell lineages of adult organism (Cumano and Godin, 2007). In most organisms, before the actual HSCs are formed there is a transient definitive wave where erythroid-myeloid progenitors are produced (McGrath et al., 2011; Bertrand et al., 2007). The definitive wave starts in the aorta-gonad-mesonephros (AGM) region where the HSCs are born. HSCs then migrate temporarily to the fetal liver and spleen and subsequently to the bone marrow, which is the main site of HSC-derived hematopoiesis (Cumano and Godin, 2007).

2.1.1 Hematopoiesis, lymphocyte differentiation and microenvironment

HSCs are the source for all blood cell types, and the major site for the adult hematopoiesis is bone marrow. The blood system contains over ten different blood lineages with various functions, and a life span ranging from few hours to several years (Rieger and Schroeder, 2012). HSC differentiate to all blood and immune cell types through hierarchically organized system with progenitor cells and lineage restricted precursor cells (Figure 1). These various intermediates demonstrate the complexity of the blood cell development and maturation. Hematopoietic differentiation starts when long-term HSCs, which have life-long self-renewal ability, gives rise to short-term HSCs with limited self-renewal ability to sustain

hematopoietic system roughly six weeks. Short-term HSCs further gives rise to multi-potent progenitors (MPP) with ability to generate all mature blood cell types without self-renewal ability. MPPs seems to be the first step for the lineage commitment in blood cell development as it gives rise to lineage-restricted oligo-potent progenitors making the binary choice for lymphoid or myeloid progenitor (Brown and Ceredig, 2009). Previously it was believed that HSCs or MPPs make the binary choice to become either myeloid or lymphoid progenitor. If committed to a lymhpoid lineage, cell becomes a common lymphoid progenitor (CLP) with capability to differentiate into any lymphoid cell types. This type of model is referred as classical model in Figure 1 (Akashi et al., 2000; Brown and Ceredig, 2009). However, recent studies have suggested revisions to this model by showing that myeloid progenitors are not a uniform population of cells with multi-lineage differentiation potential but instead a more heterogeneous mixture of progenitors committed to specific lineages (revised model in Fig. 1) (Paul et al., 2015; Perié et al., 2015).

During the lineage commitment, hematopoietic cells lose their promiscuous expression of different lineage specific genes, which is needed in HSCs and progenitors to keep up with differentiation plasticity. After the progenitors are committed to a specific lineage, the cells need to silence differentiation genes for other alternate lineage and cell types. Silencing other differentiation genes is associated with activation of specific maturation genes to promote the lineage specific maturation events (Brown and Ceredig, 2009). As the "classical" hematological tree has been under revision, it seems likely that in the future growing knowledge and methodological development will give more information about the lineage differentiations, relationships between lineages and even recognize new intermediates.

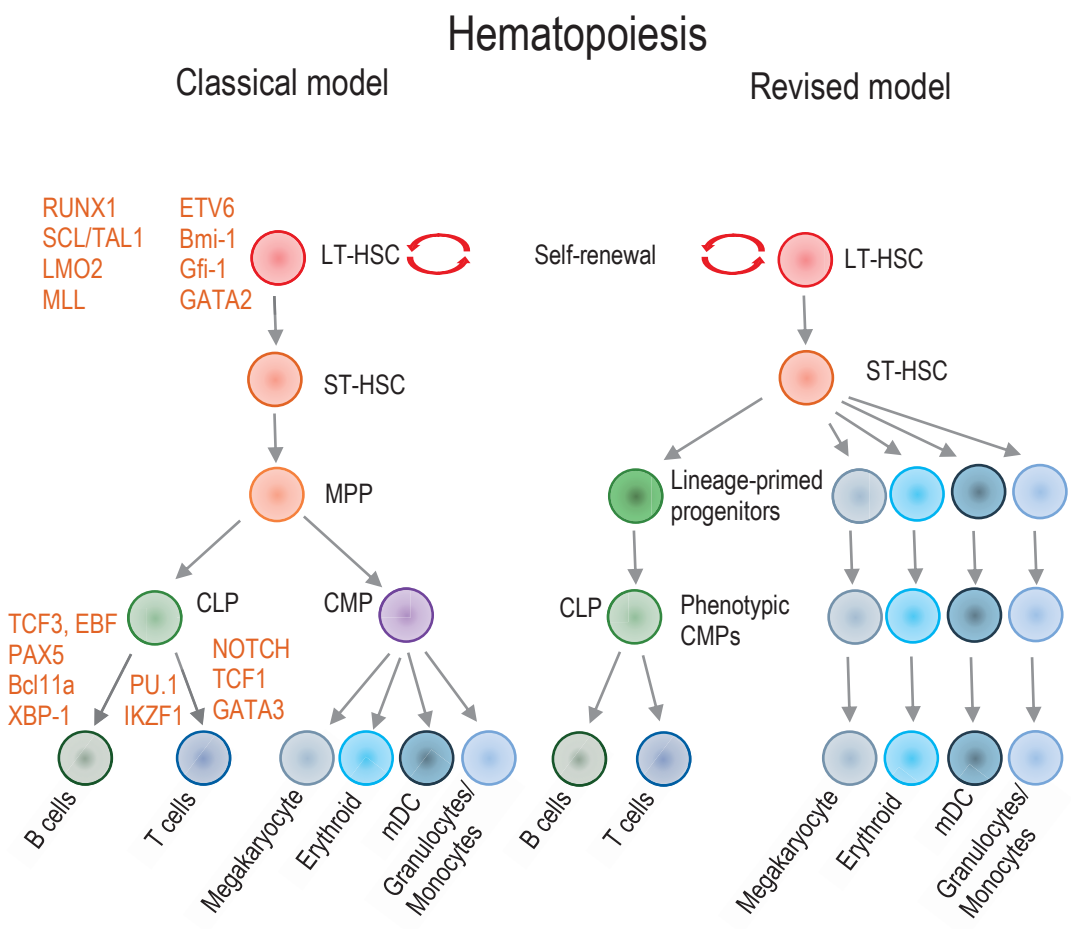


Figure 1. Models for human hematopoiesis shown together with critical transcription factors essential for HSC self-renewal and lymphoblast maturation. Classical model based on the assumption that HSCs commit either to lymphoid or myeloid fate and then progressively narrow their differentiation potential. Revised model is based on the theory of early individual cell fate commitment. Abbreviations: LT-HSC=long-term hematopoietic stem cell, ST-HSC=short-term hematopoietic stem cell, MPP=multipotent progenitor, CLP=common lymphoid progenitor, CMP=common myeloid progenitor, mDC=myeloid dendritic cell. Adapted from Orkin and Zon, 2008; Mercier and Scadden, 2015; Paul et al., 2015; Perié et al., 2015.

2.1.2 Hematological microenvironments for blood cell differentiation

Bone marrow is the major maturation site of the hematopoietic cell lines with the exception of T cells, which mature in the thymus. Bone marrow is a complex and highly organized tissue composed of heterogeneous cell populations including hematopoietic, mesenchymal and endothelial cell lineages (Ho et al., 2015). Bone marrow resides in the bone cavity surrounded by hard cortical bone and it has its specific spatialized structure with specialized microenvironments or niches, endosteal and perivascular niche, where HSCs and progenitor cells localize and receive proper signals for their differentiation and function (Brown and Ceredig, 2009; Ho et al., 2015). In the endosteal niche, stromal cells constitutes mainly from fully committed osteogenic cells called osteoblasts, which are derived from mesenchymal stem cells. Osteoblasts reside in the junction of cortical bone and bone cavity expressing adhesion molecules to retain the HSCs in the bone marrow and signal molecules to keep HSCs undifferentiated and quiescent (Brown and Ceredig, 2009; Ho et al., 2015). Differentiating blood cells migrate from the endosteal niche towards the bone marrow core to the perivascular niche where stromal cells comprise from undifferentiated mesenchymal cells, sinusoidal endothelial cells and perivascular cells. Signals and chemokines in perivascular niche supports proliferation and differentiation of HSCs and eventual egression to the peripheral circulation (Brown and Ceredig, 2009; Ho et al., 2015).

Thymus constitutes of the subcapsular zone, the cortex, the medulla, and the corticomedullary junction with specific developmentally essential microenvironments (Koch and Radtke, 2011). Earliest thymocytes differentiation events take place in the cortex, which constitutes from thymic epithelial cells (TEC), fibroblasts and macrophages. Corticomedullary junction constitutes mainly of endothelial cells and is the entry point of the thymus seeding progenitors. Medulla constitutes of dendritic cells and medullary TECs and is involved for the later events of the T cell development. TECs are mainly responsible for controlling homing, expansion, maturation and selection of developing thymocytes. Mature T-lymphocytes that are produced in thymus have a critical role in the adaptive immune system against viruses, bacteria and fungi and also in the maintenance of self-tolerance (Brown and Ceredig, 2009; Koch and Radtke, 2011).

2.1.3 Differentiation of B cells

B cells are generated in the bone marrow from HSCs through multiple developmental stages. Maturation starts with lymphoid specification and lineage-primed progenitor cells, followed by differentiation stages where B cells gain B-lymphocyte identity, and finally enters a state of maintenance ultimately leading to terminal differentiation (Fig. 2). B cell development relies on sequential expression of selected transcription factors as well as expression of cell surface receptors and adhesion molecules required for the movement in the bone marrow during the differentiation process (Brown and Ceredig 2009; Kee et al., 2000). TCF3 proteins, together with PU.1 and IKAROS family zinc finger 1 (IKZF1), promote lymphoid-primed progenitor development and it is also essential for the B cell differentiation together with EBF1 and PAX5 (Dias et al., 2008; Kee et al., 2000). V(D)J recombination is genetic recombination of antigen receptor genes occurring in developing lymphocytes during the early stages of T and B cell maturation (Brown and Ceredig, 2009). Pre-pro-B cells are the first to be fated for B-lymphocytes with DJ rearrangements of the immunoglobulin heavy chain (*Igh*). Immunoglobulin loci encode for the B cell receptor (BCR) subunits and V(D)J recombination of the immunoglobulin loci ensures diverse B cell antibody repertoire whereas the signaling through the pre-BCR/BCR is required for maturation and fate decision (Brown and Ceredig 2009). Up-regulation of PAX5 and successful rearrangement of the *Igh* locus, V(D)J rearrangement, and pre-BCR checkpoint at the pro-B stage allows the cells to proceed to the pre-B stage. Pre-B cells have a rapid proliferation and the immunoglobulin light chain (*Igl*) rearrangement also starts at this stage. After successful *Igl* rearrangement, immature B cells express a functional BCR and are ready to leave the bone marrow to secondary lymphoid organs (Brown and Ceredig 2009). Mature B-lymphocytes then wait quiescently for the appropriate antigen for activation before they proliferate and mature into plasma cells (Brown and Ceredig 2009). During the differentiation, pro-B and pre-B cells interact with the stromal cell populations at various points for specific cytokine signals such as interleukin-7 (IL7) through IL7 receptor which is important for the proliferation and survival during the B cell development (Clark et al., 2014).

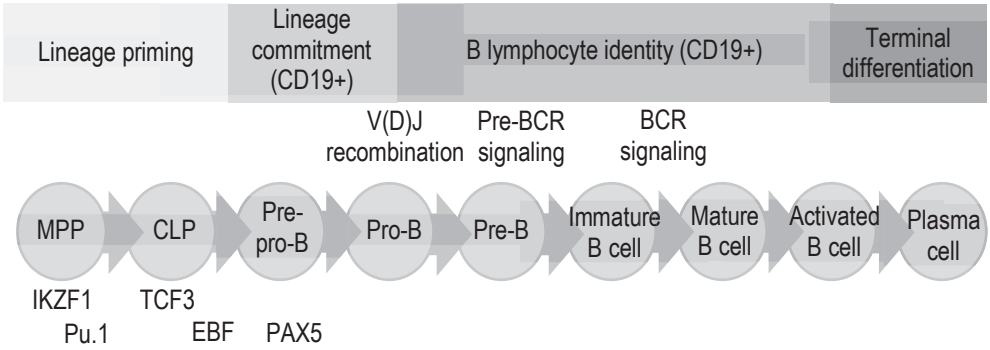


Figure 2. Core transcriptional regulators together with key checkpoints and differentiation stages of B cell differentiation. Abbreviations, MPP=multipotent progenitor, Pro-B=progenitor B cell, Pre-B=precursor B cell, BCR=B cell receptor. Adapted from Kee et al. (2000) and Brown and Ceredig (2009).

2.1.4 Differentiation of T cells

T cells differ from other blood cells as their main maturation site is the thymus and not the bone marrow. Thymic microenvironment gives T cells a unique niche to develop and differentiate from progenitor cells to different mature T cells (Brown and Ceredig, 2009; Krueger et al., 2017). T cell differentiation begins with migration of some early lymphoid progenitors to the thymus via blood circulation (Brown and Ceredig, 2009; Krueger et al., 2017). These progenitor cells enter the thymus at the corticomedullary junction and are called thymic seeding progenitors (Brown and Ceredig, 2009; Krueger et al., 2017). During the development, these thymocytes travel across the thymic microenvironment differentiating eventually to T cells (Fig. 3). Thymocytes start their differentiation as early thymic progenitors from the double negative (DN) stage ($CD4^-CD8^-$). DN thymocytes migrate from the corticomedullary junction to the outer cortex and achieve the T lineage commitment at the DN2 stage. During this stage, a somatic rearrangement of the T cell receptor (TCR) gene starts from the TCRβ, TCRδ and TCRγ loci (Brown and Ceredig, 2009; Krueger et al., 2017). TCR is responsible for the antigen recognition presented by major histocompatibility complex (MHC). During this stage T cells are also divided for TCRαβ and TCRγδ lineages based on the productive rearrangement of the TCR genes. Rearrangement of TCRδ and TCRγ loci leads to TCRγδ lineage and TCRβ rearranged cells move forward to DN3 stage and go through β-selection. Correctly rearranged TCRβ cells signal through pre-TCR, leading to a proliferation and

developmental progression towards double positive (DP) stage ($CD4^+CD8^+$), the most abundant thymocytes in the thymus. DP cells then have rearrangement of their TCRα loci, eventually producing the mature TCR and then followed by the positive-negative selection. Based on the TCR affinity the selection ensures that the produced T cells recognizes peptide:MHC ligands but do not recognize self-antigens leading to the elimination of roughly 90 % of all thymocytes. Whether the selection is finished in the cortex or medulla variates, but thereafter the T cells will locate in the medulla and are committed to either $CD4^+$ (T helper cells) or $CD8^+$ (cytotoxic T cells) positive lines based on the glycoprotein expression on their cell surface (Brown and Ceredig, 2009; Krueger et al., 2017). Mature single positive T cells then leave the thymus via post-capillary venules and enter the peripheral blood circulation (Brown and Ceredig, 2009; Krueger et al., 2017).

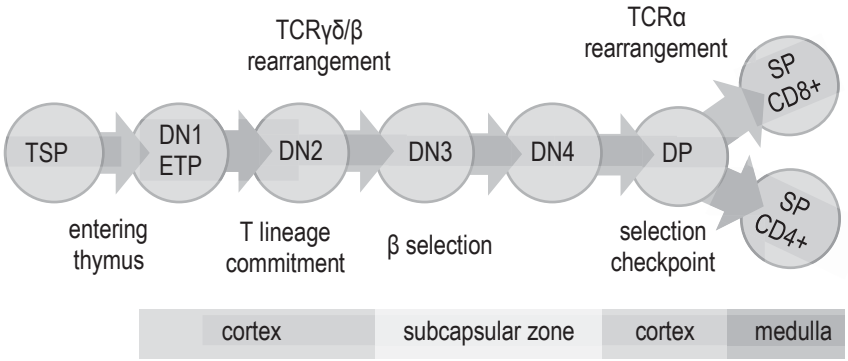


Figure 3. Developmental stages and critical checkpoints of T cell maturation. Abbreviations: TSP=thymys seeding progenitor, DN=double negative, ETP=early thymic progenitor, DP=double positive, SP=single positive, TCR=T cell receptor. Adapted from Koch and Radtke (2011).

2.2 Transcriptional regulators of hematopoiesis

Transcription factors are proteins that recognize specific deoxyribonucleic acid (DNA) sequences (DNA binding motifs) with their DNA binding domains and thereby regulate activity of gene expression. Regulation can be positive (activation) or negative (repression), and some transcription factors can exert both functions in a context-dependent way (Latchman, 2001). Some transcription factors bind to a promoter sequence located near the transcription start site and help the transcription initiation complex to form. Other transcription factors function as transcriptional

enhancers by binding to regulatory sequences, which can in mammals locate even a million or more base pairs upstream or downstream from the transcription starting site, even within the introns of neighboring genes (Levine, 2010). Transcription factors are responsible for the unique gene expression patterns in different cell types and drive cell fate decisions at different developmental stages. Mutations in transcription factors and their regulators are a foundation for numerous human diseases (Lambert et al., 2018).

Primitive and definitive hematopoieses are regulated by distinct transcription factors. In primitive hematopoiesis, the two transcription factors, GATA1 and PU.1, regulate primitive erythroid and myeloid fates with cross-inhibitory relationship (Cantor and Orkin, 2002). It has been shown that these two transcription factors interact physically and the switch towards myeloid or erythroid fate involves target gene competition between PU.1 and GATA1 (Cantor and Orkin, 2002; Cantor et al., 2002; Jagannathan-Bogdan and Zon; 2013). GATA1 acts as a master regulator of an erythrocyte development and it suppresses myeloid fate. Accordingly, in zebrafish (*Danio rerio*) GATA1 knockdown leads to myeloid specific gene expression (including *PU.1*) and switches blood cells to myeloid fate (Jagannathan-Bogdan and Zon; 2013). PU.1 is a master regulator in myeloid cell fate and *GATA1* expression increases after knockdown of PU.1 (Scott et al., 1994; Rhodes et al., 2005).

In definitive hematopoiesis, numerous transcription factors regulate self-renewal of HSCs and their differentiation to various blood cell lineages. Some of the transcription factors are essential for HSC formation, for example RUNX1, TAL1 and LMO2, while others are more relevant in a lineage-specific differentiation like TCF3, PAX5 and IKZF1 (Orkin and Zon, 2008). Many of these transcription factors are partners to translocations, deleted or mutated and these genetic alterations lead to deregulation of gene expression and failure of normal differentiation process in hematopoietic malignancies (Orkin and Zon, 2008). The most relevant transcription factors that are required for the production, survival and self-renewal of the HSCs and maturation of lymphocytes are shown in Figure 1 and the most important transcription factors in this thesis are introduced in the following chapter.

2.2.1 Hematopoietic transcription factors in definitive hematopoiesis

RUNX1 (AML1) is a member of runt family of transcription factors and is required for the normal hematopoietic development in vertebrates (Wang et al., 1996). RUNX1 can regulate gene expression either via activation or repression. RUNX1

contains a DNA-binding domain, called runt domain (RD), a transactivation domain (TAD) and an inhibitory domain (ID) from which RD and TAD are essential for *in vitro* hematopoiesis and for embryonal development in mice (Wang et al., 1996; Goyama et al., 2004; Okuda et al., 2000). In adult mouse, RUNX1 is expressed in all blood cells except erythrocytes and RUNX1 expression is regulated in both a cell line and maturation specific manner. RUNX1 expression is highest in adult bone marrow HSCs. During B cell development, RUNX1 expression is initiated in early stage of development and is expressed uniformly throughout the subsequent developmental stages (Lorsbach et al., 2004). In thymus, the RUNX1 expression is highest in the DN thymocytes while in peripheral T-lymphocytes the $CD4^+$ cells have two to three -fold higher expression level compared to $CD8^+$ T cells (Lorsbach et al., 2004; North et al., 2004). The critical role of RUNX1 in definitive hematopoiesis in early development was first shown in mice where knockout of RUNX1 led to a loss of definitive erythroid, myeloid and lymphoid cells in fetal liver hematopoiesis (Wang et al., 1996; Okuda et al., 1996). Because of the embryonic lethality of *RUNX1* knockout mice, the role of *RUNX1* in adult hematopoiesis has been studied in bone marrow chimeras, which show that the loss of *RUNX1* expression does not have a marked impact on the long term HSC maintenance (Ichikawa et al., 2004; Cai et al., 2011). In addition, *RUNX1* deletion led to an increase in myeloid and megakaryocytic progenitors and defects in lymphocyte development (Ichikawa et al., 2004; Growney et al., 2005). In zebrafish, *RUNX1* knockdown led to a loss of blood progenitors and a reduction of thymocytes at early phase of definitive hematopoiesis (Kalev-Zylinska et al., 2002; Burns et al., 2002; Gering and Patient, 2005; Burns et al., 2005).

Mutations in *RUNX1* are associated with multiple hematological malignancies such as familial platelet disorder, myelodysplastic syndrome and leukemias of myeloid and lymphoid lineages (Osato, 2004; Song et al., 1999). Chromosomal rearrangements of *RUNX1* are frequently observed in ALL, acute myeloid leukemia (AML), and chronic myelomonocytic leukemia, which indicates its importance in the formation of blood cells (Sood et al., 2017).

TAL1 (T cell acute lymphoblastic leukemia 1 / SCL, stem cell leukemia) is a transcription factor with basic helix-loop-helix (HLH) domain and it is one of the main transcription factors involved in specification of hematopoiesis in early development and regulation of hematopoietic differentiation in adults (Porcher et al., 2017). *TAL1* is expressed in various blood cell lineages in different developmental stages. Expression is seen in HSCs, myeloid progenitors, mature

myeloid cells and with lower levels in lymphoid progenitors even though the expression is absent in differentiating and mature T- and B-lymphocytes (Mouthon et al., 1993; Hall et al., 2003 Akashi et al., 2000; Lacombe et al., 2010). TAL1 has multiple functions in adult hematopoietic system that vary depending on the hematopoietic lineage and the developmental stage. In murine models, TAL1 is essential for the long-term HSCs quiescence stage as it negatively regulates their G_0/G_1 transition and maintains their long-term stem cell competence (Lacombe et al., 2010). TAL1 also plays a role in determining the blood cell lineage fate of HSCs by regulating erythrocyte differentiation, mekagaryocyte maturation and proliferation of monocyte progenitors in adult organisms (Goardon et al., 2006; Chagraoui et al., 2011; Dey et al., 2010). The regulatory role of TAL1 in adult hematopoiesis is transmitted through associations and complexes that TAL1 forms with multiple other hematopoietic transcription factors like megakaryocyte lineage priming heptad complex (GATA2, LYL1, TAL1, FLI1, ERG, RUNX1, LMO2) (Wilson et al., 2010; Pimkin et al., 2014). TAL1 is aberrantly expressed in 40-60 % of T-ALL cases and TAL1 overexpression is associated with a differentiation block in double positive ($CD4^+CD8^+$) state of thymocytes (Sanda et al., 2012).

LMO2 (LIM-only protein 2) is a member of the large zinc finger protein family, LIM-domain-only (LMO) proteins, which characteristically have two LIM domains (El Omari et al., 2011). As no interactions with DNA has been shown *in vivo*, LMO2 is hypothesized to mediate gene expression through protein-protein interactions (El Omari et al., 2011). LMO2 interacts with excessive amount of proteins and it is thought to act as a scaffold protein in multiprotein complexes forming a DNA binding complex (El Omari et al., 2011). Based on studies in mouse, LMO2 is known to be associated with TAL1 transcription factor in erythroid cells and it is essential for erythroid development (Valge-Archer et al., 1994). In the absence of LMO2, no definitive hematopoiesis is observed in mouse (Yamada et al., 1998). Enforced expression of LMO2 in T cells can result from chromosomal rearrangements including translocations t(11;14)(p13;q11) and t(7;11)(q35;p13) and deletion del(11)(p12p13) (Royer-Pokora et al., 1991; Van Vlierberghe et al., 2006). Genetic *LMO2* abnormalities are detected in 9 % of all pediatric T-ALL cases (Royer-Pokora et al., 1991; Van Vlierberghe et al., 2006; Wu et al., 2015).

ETV6 (TEL) is a nuclear protein and a member of ETS (E-Twenty-Six) family of transcription factors (Mavrothalassitis and Ghysdael, 2000). ETV6 protein constitutes of an N-terminal pointed domain (PNT) also referred to as HLH domain,

C-terminal DNA-binding domain (ETS domain) and an internal domain (De Braekeleer et al., 2012). PNT domain is responsible for the protein-protein interactions and functions as a homo-oligomerization domain. ETS domain is responsible for the protein-protein interactions and a sequence-specific DNA binding. Functional studies have shown that ETV6 acts as a strong transcriptional repressor and the repression is mediated through PNT and internal domains (De Braekeleer et al., 2012). ETV6 is required for the hematopoiesis of all blood cell lineages in the bone marrow, but it is also essential for the embryonic development and blood vessel formation in the yolk sac of developing embryos (Fenrick et al., 2000; Wang et al., 1997). ETV6 is known to be a fusion partner with multiple proteins via translocations. The most common translocation in pediatric ALL is the t(12;21) (p13;q22) creating the ETV6-RUNX1 (TEL-AML1) fusion protein (Romana et al., 1995a; Bhojwani et al., 2012; Liang et al., 2010).

TCF3 encodes for two E-proteins, HLH transcription factors E12 and E47 that arise from an alternative splicing (Slattery et al., 2008). These proteins are thought to play a crucial role in regulation of commitment, cell growth and lymphocyte differentiation (Massari and Murre, 2000; Slattery et al., 2008). TCF3 proteins form homodimers and heterodimers with other HLH transcription factors, to alter their target gene expressions. Massari and Murre (2000) and Kee et al. (2000) summarized a model where the activation of B-lineage-specific gene expression is orchestrated by TCF3 proteins. Deletion studies in mice demonstrated that TCF3 proteins are essential for the initiation of B cell development at the pro-B cell stage and that they act upstream of EBF and PAX5 (Bain et al., 1997; Kwon et al., 2008). In addition, TCF3 protein-deficient mice showed impairment in the rearrangement of the *Igh* and *Igl* in B cells (Bain et al., 1997; Kwon et al., 2008). Later studies have also suggested that TCF3 proteins are already transcriptionally active in the CLPs and support the lymphoid lineage specification from the HSCs and (Dias et al., 2008). Loss of E47 leads to a reduced CLP population which lacks V(D)J recombinase activity and D-J_H rearrangements in mice (Borghesi et al., 2005; Dias et al., 2008). *TCF3* rearrangements are present in approximately 5-7 % of pediatric BCP-ALL giving usually rise to t(1;19) (q23;p13) translocation encoding for the fusion protein TCF3-PBX1 (Barber et al., 2007; Pui et al., 2011).

PAX5 is member of paired box family of transcription factors. It plays a dual role in the hematopoiesis as it is essential in B cell diffenrentiation and maturation but also in inhibition of the differentiation of other lineages (Shahjahani et al., 2015). PAX5

is associated with multiple aspects of B cell differentiation, as it is involved in the commitment, immunoglobulin gene rearrangement and B cell survival (Shahjahani et al., 2015). Furthermore, deletions or inactivating mutations of PAX5 causes arrest in pro-B cell stage (Shahjahani et al., 2015). Somatic alterations of *PAX5* are frequent in BCP-ALL and Shah et al. (2013) have associated *PAX5* germline mutations with susceptibility to BCP-ALL (Hyde and Liu, 2013).

IKZF1 (IKAROS family zinc finger 1) is a transcription factor, which acts as a zinc finger DNA binding protein. In hematopoietic system, IKZF1 primes the potential of lymphoid lineage and the lack of it impairs the lymphocyte production in mice (Georgopoulos, 2002). IKZF1 acts in chromatin remodeling and its association with nucleosome-remodeling and histone deacetylation complexes controls lymphocyte development (Zhang et al., 2011). In addition to its importance in lymphoid differentiation, IKZF1 acts as a tumor suppressor and the loss of function mutations in IKZF1 is associated with BCP-ALL (Payne and Dovat, 2011; Kastner et al., 2013).

2.3 Zebrafish development and hematopoiesis

Developmental stages of zebrafish can be divided in four stages: embryonic, larval, juvenile and adult. In laboratory conditions (28.5 °C), embryonic development, embryogenesis, takes place in 0-72 hours post fertilization (hpf). Larval stage starts after 3 days post fertilization (dpf) and transition to juvenile stage happens at 30 dpf when their total body length is approximately 10 mm. Zebrafish reach sexual maturity in 3 months and zebrafish are considered as adults after 90 dpf (ZF-HEALTH; ZFIN). As adults, the total body length of zebrafish can be 4-5 cm and maximum lifespan in laboratory is typically 1.5-2 years (ZF-HEALTH; ZFIN).

Compared to other vertebrate model organisms zebrafish has several advantages that make it powerful system for genetic analysis, imaging and for studying hematological diseases (Paik and Zon, 2010; Teittinen et al., 2012; Lohi et al., 2013). Zebrafish embryos are transparent and eggs are fertilized externally allowing the monitoring of the developing embryos immediately after fertilization. Zebrafish also produces large amounts of embryos whose genetic manipulation is rather straightforward (Lessman, 2011; Peal et al., 2010; Amsterdam and Hopkins, 2006). Zebrafish have a diploid genome which constitutes from 25 chromosome pairs and, based on the Zv9 genome assembly, approximately 70 % of the human genes have at least one orthologue in zebrafish genome (Fish Development and Genetics, 2004;

ZFIN; Howe et al., 2013). Zebrafish genome has undergone a partial duplication during teleost development and 26 % of zebrafish genes are retained as ancestral duplicates having two orthologues for a subset of human genes (Fish Development and Genetics, 2004; Howe et al., 2013). These gene duplicates, called ohnologues, might lead to a potential problem when genes are targeted, as it is possible that both ohnologues needs to be targeted to reveal the effects on phenotype (Vacaru et al., 2014; Howe et al., 2013).

In the beginning of the 21st century, zebrafish emerged as a new system to study hematopoietic development as it closely resembles human hematopoiesis. Like in humans, the zebrafish hematopoiesis consists also from the primitive and definitive waves (Fig. 4). Primitive HSCs are born in the intermediate cell mass (ICM) and during the primitive wave, the embryo generates myeloid and erythroid cells which start circulating already after 24 hpf (Detrich et al., 1995; Paik and Zon, 2010). Definitive HSCs originate at 30 hpf from AGM progenitors in the ventral wall of the dorsal aorta and at 48 hpf these progenitors seed the caudal hematopoietic tissue (CHT) and kidney marrow. Kidney marrow, analogous to the bone marrow niche of humans and other mammals, gives rise to all blood cell lineages and, together with the thymus, produces all mature hematopoietic cells throughout zebrafish life (Davidson and Zon, 2004). CHT locates in the posterior part of the embryo and serves as an intermediate site of definitive hematopoiesis (erythropoiesis and myelopoiesis). Site of definitive hematopoiesis eventually switches to kidney marrow myelopoiesis which starts at 4 dpf followed by maturation of other blood cell lineages. Lymphopoiesis starts in thymus from 3 dpf and it serves as the site for lymphoid T cell maturation. (Paik and Zon, 2010; Jing and Zon, 2011; Chen and Zon, 2009) (Fig. 4).

Despite some unique characteristics, zebrafish and other vertebrates regulate the hematopoiesis with common genetic programs (Paik and Zon, 2010). Mutual signaling pathways and transcription factors are involved in definitive hematopoiesis of zebrafish and mammalians. Notable examples are Hedgedog, VEGF and NOTCH signaling pathways in the initiation of definitive hematopoiesis and RUNX1 transcription factor which plays a role in mediating the NOTCH pathway (Gering and Patient, 2005; Burns et al., 2005). Zebrafish mutants have also revealed that IKZF1 transcription factor is required for larvae lymphopoiesis (between 3-14 dpf) but not in adult lymphopoiesis (Schorpp et al., 2006) and *TAL1* mutants do not initiate hematopoietic development (Bussmann et al., 2007). *ETV6* knockdown studies demonstrated a regulative role in embryonic and larvae hematopoiesis impacting multiple lineages, including reducing erythrocytes in primitive and

definitive hematopoiesis and decreasing myeloid cells in definitive hematopoiesis (Rasighaemi et al., 2015).

Besides many similarities in blood cell development and regulation, there are also morphological and physiological differences between mature blood cells between zebrafish and human. While mammalian erythrocytes have discoid biconcave shape, the zebrafish erythrocytes are elliptical and have retained their nuclei whereas mammalian erythrocytes lose nuclei during the maturation (Chen and Zon, 2009). Zebrafish megakaryocytes and their nuclei-containing thrombocytes originate from prothrombocytes, whereas mammalian thrombocytes have no nucleus and are derived from megakaryocytes (Carradice and Lieschke, 2008). Lack of antibodies for specific cell surface markers is a problem in the zebrafish hematopoiesis research and sorting of hematopoietic lineages is mostly done with transgenic lines expressing a lineage-specific fluorescence marker, which limits the isolation of subpopulations. Other challenge is the limitations in the *in vitro* and *ex vivo* cell culturing and differentiation of the HSCs and progenitors (Gore et al., 2017).

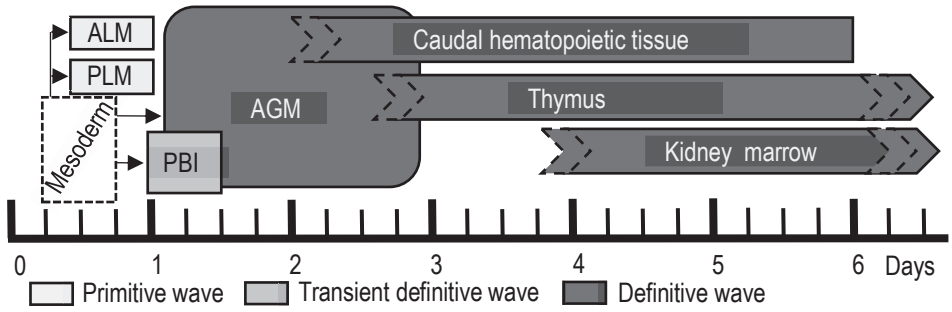


Figure 4. Sites and waves of zebrafish embryonic hematopoiesis during the early development. Adapted from Chen and Zon (2009), Paik and Zon (2010) and Jing and Zon (2011). Abbreviations: ALM=anterior lateral mesoderm. PLM=posterior lateral mesoderm, PBI=posterior blood island, AGM=aorta-gonad-mesonephros region.

2.4 Acute leukemia

Acute leukemias can be classified into two groups, myeloid and lymphoid, based on the lineage from where the leukemic cells arise. ALL occurs in people of all age but is more common in pediatric patients with peak incidence being between two and five years of age (Inaba et al., 2013). Survival in pediatric ALL patients with current treatments in developed countries is approximately 85-90 % but there is variation

between subtypes (Pui et al., 2015; Toft et al., 2018). Even though the cure rate in pediatric patients is good and has improved in adult patients, some subtypes, infants and many adults still have a marked relapse risk (Gatta et al., 2013; Hunger et al., 2012; Pui and Evans, 2006; Nachman et al., 2007). AML is the most common acute leukemia amongst the adults making up roughly 60-80 % of the cases but is relatively rare in children accounting only 15-20 % of pediatric leukemias (Yamamoto and Goodman, 2008; De Kouchkosvky and Abdul-Hay, 2016; Linabery and Ross, 2008; Puumala et al., 2013; Dores et al., 2012). AML treatments for younger patients have also led to significant improvements with approximately 60-70 % survival rates (Rubnitz, 2017), but patients over 65 years have a 70 % chance of dying within 1 year of diagnosis (Meyers et al., 2013). Chronic leukemias are very rare in children representing only 1-2 % of all childhood leukemias (Belgaumi et al., 2010).

Environmental, biological and genetic factors play a role in leukemia susceptibility and development. Exposure to infections has been widely studied as causal mechanisms for the leukemia development: lack of infections during early life cause dysregulation in immune system, which leads to an abnormal response to common infections later in childhood (Greaves, 2018). Abnormal immune response creates strong pressure on lymphoid cells and trigger the secondary mutations needed for the leukemia development (Greaves, 2018). A number of possible environmental exposures has also been linked to ALL, but only ionizing radiation is universally accepted to have a causal effect on ALL development (Greaves, 2018). Exposures during pregnancy as maternal consumption (e.g. alcohol, coffee, tea) has also been studied but without a clear association to development of ALL (Greaves, 2018). Even though leukemia is not considered as an inheritable disease, there are some genetic factors that contribute to disease onset, with Down syndrome being the most well-known with 20-fold increased risk for ALL (Lee et al., 2016). Low hypodiploid ALL has been linked to Li-Fraumeni syndrome (mutation of *TP53* gene) and is considered as a manifestation for the syndrome (Holmfeldt et al., 2013). There are also a number of genes whose mutation predispose carriers to ALL development (e.g. *ARID5B*, *CEBPE*, *GATA3*, *ETV6*, *IKZF1* and *PAX5*) (Tasian et al., 2015).

2.4.1 B cell precursor acute lymphoblastic leukemia

BCP-ALL is the most common subtype of ALL in childhood. BCP-ALL can be further divided into separate subtypes according to non-overlapping chromosol aberrations, which also serve as prognostic factors. World health organization

(WHO) classified these subtypes in 2008 by the presence of certain fusion genes, chromosomal alterations or genetic rearrangements (Vardiman et al., 2009). In 2016, two provisional entities were added to the list (Arber et al., 2016). Together these subtypes cover roughly 80 % of all BCP-ALL cases. Recently, at least four additional subtypes were identified. The main subtypes and their prevalence in pediatric patients are shown in Figure 5 (data from Lilljebjörn and Fioretos, 2017).

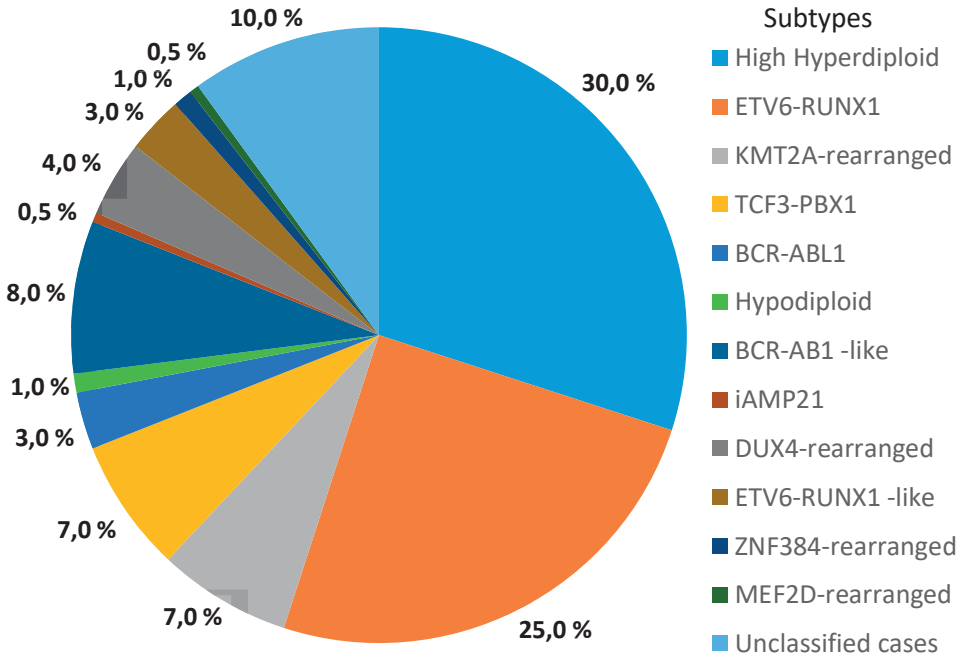


Figure 5. BCP-ALL subtypes and their estimated frequencies from Lilljebjörn and colleagues (2016) in a pediatric cohort. Data consists of 195 pediatric (< 18 years of age) BCP-ALL cases from a population-based series.

Two main types of genomic alterations present in childhood ALL are aneuploidy and chromosomal rearrangements. In aneuploidy, there is either gain or loss of whole chromosomes, whereas chromosomal rearrangements usually lead to chimeric fusion genes or aberrant expression of the rearranged gene. High hyperdiploid subtype, which encompasses approximately 25-30 % of childhood ALL, is characterized by the presence of more than 50 chromosomes and often involves chromosomes 4, 10, 14, 17 and 21 (Roberts and Mullighan, 2015). High hyperdiploid ALL is usually associated with favorable prognosis, whereas the converse event, namely the lack of chromosomes, hypodiploid ALL (less than 44 chromosomes),

carries poor prognosis. Hypodipdloid ALL can be further divided into subgroups based on the chromosome number, from which most common are near-haploid ALL (24-31 chromosomes), low-hypodiploid ALL (32-39 chromosomes) and high-hypodiploid ALL (40-43 chromosomes) (Roberts and Mullighan, 2015).

ETV6-RUNX1 fusion gene results from translocation t(12;21)(p13;q22) involving two hematopoietic transcription factors, ETV6 and RUNX1, and is the most common translocation in pediatric ALL comprising approximately 20-25 % of the cases (Romana et al., 1995a; Liang et al., 2010). The second most frequent chromosomal translocation in the pediatric ALL is t(1;19)(q23;p13) which produces *TCF3-PBX1* fusion involving two transcription factors, TCF3 and PBX1 (Hunger et al., 1991). Prevalence of the TCF3-PBX1 subtype amongst pediatric BCP-ALL patients is approximately 5-7 % and it is associated with favorable prognosis with contemporary therapies (Roberts and Mullighan, 2015; Bhojwani etl al., 2015; Lilljebjörn et al., 2016). The t(9;22)(q34;q11) translocation generates fusion of two kinases, BCR and ABL1, and is better known as a Philadelphia chromosome. *BCR-ABL1* occurs in approximately 3 % of pediatric BCP-ALL cases and has less favourable prognosis with a risk of treatment failure and relapse despite use of targeted therapy (e.g. imatinib) (Biondi et al., 2018). *KMT2A* gene (known as *mixed lineage leukemia*, *MLL*) loci at chromosome 11q23 is a frequent target of chromosomal translocations and rearrangements with multiple partners. These *KMT2A*-rearrangements manifest approximately 2-7 % of pediatric cases (Lilljebjörn et al., 2016; Roberts and Mullighan, 2015; Bhojwani etl al., 2015). In infant leukemias (patients < 1 year), the *KMT2A*-rearrangement frequency is remarkably higher than in pediatric leukemias in general as the prevalence is approximately 60-80 % (Pieters et al., 2007; Mann et al., 2010; Ross and Robison, 1997).

These genetic rearrangements alone are usually insufficient to generate an overt leukemia (*KMT2A*-rearrangement may be an exception). They serve as a "first hit" mutation and additional mutations are needed before BCP-ALL development. Such mutations can be deletions, amplifications and structural rearrangements in genes important for B-lymphocyte differentiation (e.g. *PAX5* and *IKZF1*) (Mullighan et al., 2007; Mullighan et al., 2009). This hypothesis is supported by twin studies with ETV6-RUNX1 and high hyperdiploid subtypes where identical twins have the same primary chromosomal aberration, a pre-leukemic clone, but develop the leukemia independently after birth (Bateman et al., 2015; Bateman et al., 2010; Mori et al., 2002; Greaves et al., 2013). Bateman et al. (2010) studied monozygotic twins with concordant ETV6-RUNX1 subtype and compared their copy number alterations (CNA) classified as "driver" and "passenger" mutations. All CNA drivers were

absent from their healthy co-twin, with same pre-leukemic clone, placing the driver CNAs as secondary mutations.

2.4.2 T cell acute lymphoblastic leukemia

Leukemia that arises from the malignant transformation of T cell progenitors is called T-ALL and accounts for approximately 10-15 % of pediatric and 25 % of adult ALL cases (Ferrando et al., 2002; Toft et al., 2018). Transformation of T cell precursors into malignant T-ALL lymphoblasts results from genetic and epigenetic alterations during thymocyte development that lead to changes in the normal growth, differentiation, survival and proliferation (Van Vlierberghe and Ferrando, 2012). T-ALL has been divided into distinct molecular subgroups according to the gene expression signatures (Table 1), but unlike in BCP-ALL, T-ALL subtypes have no confirmed prognostic significance (Girardi et al., 2017; Liu et al., 2017).

Transmembrane receptor protein NOTCH1 is normally involved in T cell development and activating mutations in NOTCH1 signaling are present in over 50 % T-ALLs making it one of the most significant oncogenic pathways in T-ALL development (Weng et al., 2004). Deletion in the *CDKN2A* locus at 9p21, which contains the tumor suppressor genes *p16/INK4A* and *p14/ARF*, is present in over 70 % of all T-ALL cases (Ferrando et al., 2002; Hebert et al., 1994). T-ALL is also associated with several chromosomal translocations and aberrant expressions of transcription factors like basic HLH family members TAL1, LYL1, LMO2, TLX1, TLX3 and NKX2.1 in T cell progenitors (Van Vlierberghe and Ferrando, 2012). These transcription factors are often placed under the control of strong T cell specific enhancers, which locate in the *TCRβ* (7q34) and *TCR*α/*TCR*δ (14q11) loci (Van Vlierberghe and Ferrando, 2012). Cauwelier et al. (2006) demonstrated that aberrations in these loci were involved in 33 % of their T-ALL samples.

In addition to translocated and deregulated transcription factors fusion genes are also present in T-ALL. One of these fusion genes is *SIL-TAL1* fusion where *SIL* promoter is placed in the vicinity of *TAL1* gene and induces an abnormal expression of *TAL1;* this fusion variates between 15-25 % of T-ALL cases (Brown et al., 1990; Aplan et al., 1992; Wang et al., 2013). Graux et al. (2004) observed gene fusion between *ABL1* and *NUP214* creating the second most common *ABL1* fusion, *NUP214-ABL1*, in hematopoietic malignancies with a frequency of about 5-6 % in T-ALL (Graux et al., 2009; De Braekeleer et al., 2011b). Other *NUP214* gene fusion, *SET-NUP214*, in T-ALL was discovered by Van Vlierberghe et al. (2008) where they

suggested that SET-NUP214 binds promoter regions of *HOXA* genes, which may activate the *HOXA* genes and contribute to T cell differentiation arrest.

Table 1. Major subgroups of T-ALL.

Subgroup	Associated genes	Rearrangement partners resulting in deregulated expression of target gene.
LMO2/LYL1	*RUNX1, FLT3, TCF7, NRAS*	*TRA, TRB, CSTF3-LMO2, FOXJ3-LMO2 / TRB*
HOXA	*ETV6, MLLT10, STAT5B, JAK3, EZH2, CNOT3, KMT2A*	*TRB, HOXA* insertion, *LINC01260-HOXA, POLR2E-HOXA*
TLX3	*PHF6, DNM2, WT1, SMARCA4, BCOR, KDM6A, RPL5*	*BCL11B-TLX3, CDK6-TLX3, TRB, TLX3-CACS15, TLX3-TARBP1*
TLX1	*CDKN1B, BCL11B, PTPN2, DLEU1, RB1*	*TRB, TRA, TLX1*-upstream, *LINC00502-TLX1*
NKX2.1	*RPL10, LEF1*	*TRA, NKX2-1-BCL11B, CDK6-NKX2-1, NKX2-1-DIO2*
TAL1	*PTEN, USP7, 6q del, PIK3R1*	*STIL, TRA, DHX9-TAL1, GNPAT-TAL1, TAL1-TARP*
Others *	*CDKN2A/B, NOTCH1, FBXW7, ABL1, IL7R*	

*Genetic alterations not specifically enriched in any T-ALL subtype. Data adapted from Liu et al., 2017.

2.4.3 Treatment and outcomes of acute lymphoblastic leukemia

Treatment of ALL consists of chemotherapy, which lasts approximately 2-2.5 years, and can be divided into four phases: induction of remission, deepening of the remission (consolidation), intensification and maintenance therapies (Pui et al., 2008). Treatment aims at complete eradication of disease and relapse prevention. Treatment stratification is based on biological features of leukemic cells, age, leukocyte count and treatment response (Inaba et al., 2013; Toft et al., 2018). In high risk cases, allogeneic hematopoietic stem cell transplantation is applied. With modern treatments, the cure rate for pediatric patients overall in ALL is above 85 % and, for adults, long-term survival is estimated around 75 % of the patients with the pediatric-type protocols (Toft et al., 2018; Terwilliger et al., 2017). Eventhough the overall survival (OS) in pediatric patients is reaching 85 %, relapses with poor outcome remain approximately 15-20 % of the patients (Einsiedel et al., 2005; Nguyen et al., 2008; Durinck et al., 2015). Very few targeted therapies have thus far been applied to ALL. However, one classical example is the use of tyrosine kinase inhibitors imatinib or dasatinib in BCR-ABL1-positive ALL resulting in significant

improvement in the event-free survival (EFS) and OS (Thomas et al., 2004; Schultz et al., 2009; Ravandi et al., 2010, Biondi et al.,2018).

2.4.4 Genetic diversity and relapse

Relapse remains the most common treatment failure in childhood ALL occurring in approximately 15-20 % of ALL cases in developed countries with relatively poor outcome (Einsiedel et al., 2005; Nguyen et al., 2008; Durinck et al., 2015). OS after the relapse variates between 40-70 % depending on the follow-up period and involved risk group (Oskarsson et al., 2016; Pierro et al., 2017; Saarinen-Pihkala et al., 2006). At relapse, the patients are risk stratified based on the time from initial diagnosis, anatomical site of relapse, immunophenotype and therapy response as measured by minimal residual disease (MRD) (Oskarsson et al., 2016). Additional independent unfavorable prognostic factors in relapse include age over ten years, Down syndrome, T cell immunophenotype with hyperleukocytosis and unfavorable cytogenetic aberration (Oskarsson et al., 2016). Cytogenetics of the relapsed ALL is not used in the risk stratification to individualize treatment as it is used in primary diagnosis (Oskarsson et al., 2016).

During the last decade, numerous studies have investigated the genomic diversity of ALL at diagnosis and clonal evolution during therapy in relapsed ALL. One of the first genomic studies was conducted by Mullighan and coworkers (2008). They matched diagnostic and relapsed samples from 61 patients and analyzed differences in CNAs. Interestingly, they noted that diagnosis and relapse samples were genetically related but most relapse samples had striking degree of change in the CNAs that were present already in the original diagnosis sample. In BCP-ALL cases, the mean number of CNAs was increased at relapse, but no significant changes were observed in T-ALL cases. This difference between relapse and diagnosis samples originated from a common ancestral primary leukemia clone and abnormalities acquired in the relapse cases affected preferentially genes involved in B cell development and cell cycle regulation. Mullighan et al. (2008) suggested that genomic alterations contributing to relapse might be selected during the treatment because the relapse clone was often already present in the diagnosis as a minor sub-population.

Later studies have revealed more details about the genomic aberrations in relapsed ALL. Recurrent lesions in relapsed BCP-ALL might be candidate drivers for drug resistance. For example, genes involved in glucocorticoid signaling which

acquired deletions that drive glucocorticoid resistance contributing to relapse (Hogan et al., 2011; Mullighan et al., 2008). In addition to mutations of single genes, activating mutations in certain pathways like WNT, MAPK and Ras have been associated with relapse and drug insensitivity and inhibition to these pathways have increased chemosensitivity (Hogan et al., 2011; Dandekar et al., 2014; Jones et al., 2015; Irving et al., 2014).

Ma et al. (2015) looked deeper into genetic heterogeneity and clonal evolution of pediatric BCP-ALL progression and relapse. They used high-coverage whole-exome sequencing to study 20 BCP-ALL cases from diagnosis to relapse and showed that clonal survival from diagnosis to relapse was not associated with mutation burden and that clonal diversity was comparable at diagnosis and relapse. They found frequently mutated pathways (e.g. CREBP, TP53, NRAS and IKZF1) with multiple sub-clonal mutations at diagnosis but remission-induction therapy disposed all but one. The surviving clone, usually a minor clone, acquired additional mutations and deepened the resistance to therapy and thus became a founder clone for the relapse. One example of relapse-acquired mutations due to treatment pressure are cases with *NT5C2* mutations at relapse, which cause enhanced enzymatic activity of this enzyme and confers resistance to thiopurines used in maintenance therapy of ALL (Meyer et al., 2013; Tzoneva et al., 2013). Enrichment of mutations in epigenetic regulators have also been noted (Ma et al., 2015; Mullighan et al., 2011; Mar et al., 2014). It seems that diversity of mutations most commonly towards key regulators, transcriptional or epigenetic, are contributing to relapse. Possible relapse routes from subclones or via evolutional mutations are presented in Figure 6.

Two major points seem to underline the ALL relapse treatment: (i) Leukemia heterogeneity causes drug resistance via escape of minor resistant subclones, and (ii) slow progression of drug development in overall, as very few new drugs have been developed during past decades. Studies over prognostic factors, treatment outcomes and drug resistant phenotypes have been gaining insights for pathogenic properties of leukemia and relapsed leukemia. Despite the progress, relapse rate and outcomes remained unsatisfactory and more discoveries are needed from genomic studies and drug-screenings. One very promising opening is the targeted immunotherapies to cell surface antigens such as CD19 or CD22, which are evidently providing help to many patients with relapsed leukemias (Wang et al., 2017; Kaplan et al., 2015).

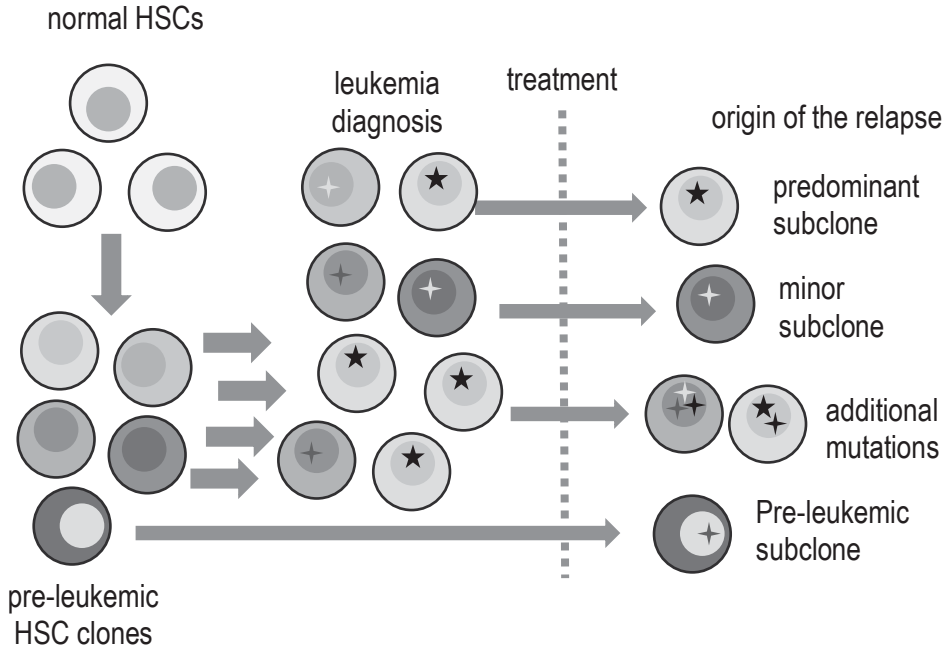


Figure 6. Model for the clonal relationships of diagnosis and relapse in acute leukemia. A schematic illustration showing relationship between pre-leukemic clones, initial leukemia diagnosis and relapse. Relapse can arise from the treatment when either an original predominant subclone or a minor subclone survives. Third possible mechanism is that the treatment induces acquisition of additional mutation(s) into a diagnosis phase subclone. Fourth possible route for the relapse in minority of the cases is that novel leukemic clone originates from the pre-leukemic HSC clones. Adapted from Mullighan et al. (2008) and Jan and Majeti (2013).

2.5 Regulators of transcription and signaling in leukemia

Chromosomal translocations can create fusion genes (hybrid genes) which are recognized as important drivers in cancer in general. Gene fusions often lead to a dysfunction of either of the partner genes, creation of an oncogenic fusion protein or truncating the gene, creating a loss of function mutant (Latysheva and Babu, 2016). In-frame fusions can generate fusion proteins such as transcription factor fusions ETV6-RUNX1 or TCF3-PBX1 in BCP-ALL. These fusions cause an aberrant transcriptional activity which leads to differentiation arrest and changes in normal proliferation and survival (O'Neil and Look, 2007; Shima and Kitabayashi, 2011; Teppo et al., 2016).

2.5.1 Transcriptional fusion proteins in BCP-ALL

ETV6-RUNX1 fusion gene is the second most common genetic rearrangement in pediatric BCP-ALL comprising about 20-25 % of the cases (Romana et al., 1995a; Liang et al., 2010). It is generated by translocation which fuses the 5' region of the *ETV6* (exons 1 to 5) (12p13.2) together with almost the entire coding region of *RUNX1* (12q22.12) creating a chimeric fusion t(12;21)(p13;q22) (Golub et al., 1995; Romana et al.,1995b; De Braekeleer et al., 2012) (Fig. 7).

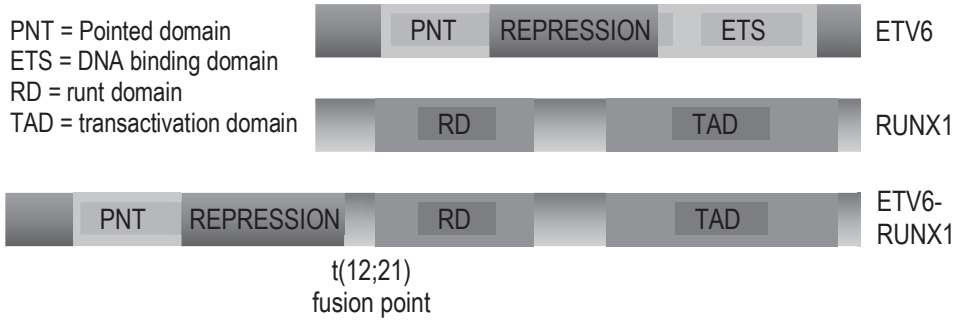


Figure 7. Schematic illustration of ETV6-RUNX1 fusion protein. Adapted from: Zelent et al. (2004) and Sood et al. (2017).

Fusion gene retains the N-terminal PNT domain and a central repression domains from the ETV6 that are fused with RD and TAD domains from RUNX1 (Zelent et al., 2004). As a result, RUNX1 loses its transactivation properties converting the fusion protein to a repressor through functional domains of ETV6. ETV6-RUNX1 fusion protein binds enhancer areas with RUNX1 motif, which leads to repression of the RUNX1 target genes (Zelent et al., 2004; Morrow et al., 2007; Teppo et al. 2016). ETV6-RUNX1 fusion can also disrupt the function of normal ETV6 by dimerization via PNT domains (Rubnitz et al., 1999; De Braekeleer et al., 2011a; Gunji et al., 2004).

Debate of the prevalence of *ETV6-RUNX1* fusion at birth has been ongoing over a decade. Mori et al. (2002) suggested that *ETV6-RUNX1* fusion was present in 1 % of cord blood in newborns (Mori et al., 2002). This was later questioned when Lausten-Thomsen et al. (2011) stated that the prevalence was actually much lower and present only in 0.01 % of newborns. Latest publication on this matter was presented by Schäfer et al. (2018) suggesting that as many as 5 % of newborns harbor the *ETV6-RUNX1* fusion. It seems likely that the prevalence supposed by Schäfer

et al. (2018) might be the most accurate as they based their results on genomic DNA opposed to ribonucleic acid (RNA)-based methods used in other studies. Higher *ETV6-RUNX1* prevalence amongst newborns compared to pediatric ETV6-RUNX1 leukemia cases is also supported by animal studies. Studies with ETV6-RUNX1 expressing zebrafish showed that roughly 3 % of the fish developed a leukemia (Sabaawy et al., 2006). In mice, ETV6-RUNX1 model did not develop leukemia under pathogen free conditions, but when exposed to infectious environment approximately 11 % of the mice developed leukemia (Rodríguez-Hernández et al., 2017).

TCF3-PBX1 fusion results from the chromosomal translocation in which the N-terminal transcriptional activation domains from the TCF3 (1;q23.3) is fused with the C-terminal DNA-binding homeodomain of PBX1 (19;p13.3) (Hunger et al., 1991). This translocation t(1;19)(q23;p13) (Fig. 8) creates a chimeric transcription factor which is found approximately 5-7 % of childhood ALL cases (Carroll et al., 1984; Roberts and Mullighan, 2015; Bhojwani etl al., 2015; Lilljebjörn et al., 2016).

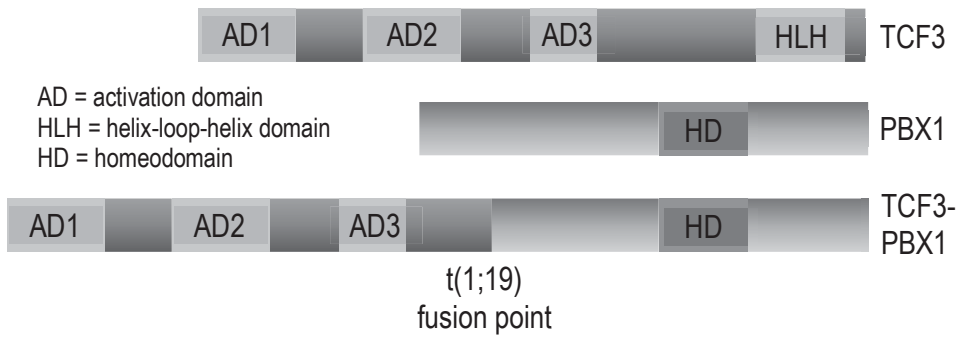


Figure 8. Schematic illustration of TCF3-PBX1 fusion protein. Adapted from: Bayly et al. (2004).

Studies have demonstrated that the fusion protein needs both of its partners for the oncogenesis and the activation domains (AD1 and AD2) are vital but the homeodomain (HD) of PBX1 is dispensable (Monica et al., 1994; Chang et al., 1997). It has been demonstrated that there is a motif in the C-terminal part after the homeodomain and this Hox cooperativity motif is required for the oncogenic potential of TCF3-PBX1 (Monica et al., 1994; Chang et al., 1997). It has been proposed that TCF3-PBX1 causes aberrant activation in genes which are normally controlled by PBX/HOX complexes. Moreover, the fusion protein is known to form heterocomplexes with Hox proteins (Chang et al., 1997; Thorsteinsdottir et al.,

1999). In addition to the interactions with other proteins, the activation domains of TCF3 portion in TCF3-PBX1 fusion protein interacts directly with CBP/p300 and induces proliferation in primary hematopoietic cells (Bayly et al., 2004; Bayly et al., 2006). One downstream target of the TCF3-PBX1 fusion is Wnt-signaling pathway and the presence of TCF3-PBX1 fusion results in aberrant regulation of cell cycle control and apoptosis, which could compose to a driving force towards propagation and maintenance of TCF3-PBX1 leukemic phenotype (Casagrande et al., 2006; Mazieres et al., 2005; McWhirter et al., 1999; Diakos et al., 2014). Several studies have begun to reveal the functional mechanisms of TCF3-PBX1, from its regulation of target genes and its effects on signaling pathways leading from pre-leukemic clone to full-blown leukemia. Even though exact mechanisms of the oncogenic function of TCF3-PBX1 fusion protein is still unclear, the fusion needs additional genetic events for the leukemia initiation, like PAX5 insufficiency or activation of signaling pathways (e.g. JAK/STAT, RAS/MAPK) (Duque-Afonso et al., 2015).

2.5.2 Sox family of transcription factors

Sox (Sry -related high-mobility group box) family of transcription factors consists of 20 members in most vertebrates and this protein family can be divided into eight subgroups (A-H) (Schepers et al., 2002; Lefebvre et al., 2007). The main feature for these SOX proteins is the DNA-binding domain called high-mobility group (HMG) box, which is similar to that of originally identified in HMG box of chromatin-associated non-histone proteins (Laudet et al., 1993; Bianchi et al., 2005). Subgroup members are highly similar to each other and most of the SOXs have central roles in cell fate determination and differentiation into specific lineages (Dy et al., 2008). In humans, the SoxC family contains three proteins, SOX4, SOX11 and SOX12, which share high degrees of similarity in the HMG box and the C-terminal region where the TAD is located (Jay et al., 1995; Jay et al., 1997; van de Wetering et al., 1993; Kuhlbrodt et al., 1998; Dy et al., 2008). Similar expression patterns of SoxC proteins suggest that these three protein members have similar functional properties and at least SOX4 and SOX11 are vital for the early development as these knockout mice die early (Dy et al., 2008).

SOX11 is normally expressed with high level in the human fetus, central and peripheral nervous system and internal organs (Jay et al., 1995; Hargrave et al., 1997; Sock et al., 2004). Upregulations of SOX4 and SOX11 expression have been observed in many cancer types (Weigle et al., 2005; Liu et al., 2006; Pramoonjago et

al., 2006). SOX11 has been previously studied in mantle cell lymphoma (MCL) and its nuclear protein expression is a specific marker for MCL and used as a marker for monitoring minimal residual disease (Ek at al. 2008; Mozos et al. 2009; Zeng et al. 2012; Szostakowska et al. 2018). Nuclear expression of SOX11 protein was also reported in two ALL studies, with mostly strong expression, and high *SOX11* messenger RNA (mRNA) expression seen in another study (Dictor et al. 2009; Mozos et al. 2009; Vegliante et al. 2011). Knockdown of SOX11 in MCL cell lines increased proliferation *in vitro* and led to more aggressive tumors *in vivo* (Conrotto et al. 2011). Overexpression of SOX11 promotes BCR signaling in B cells and produces a disease phenotype similar to human MCL (Kuo et al. 2018). SOX11 can be a useful tool in MCL diagnostics and as a prognostic marker but functional role in the malignancies remains poorly understood.

2.5.3 Epigenetic regulation of gene transcription

Epigenetic modifications can alter gene expression and therefore regulate multiple biological processes, like cell growth, differentiation, development and apoptosis. The main epigenetic mechanisms are DNA methylation and histone modification (Wu et al., 2018). DNA methylation is a process where a methyl group is added to DNA base (cytosine or adenine). Methylation at promoter region leads typically to a repression of gene transcription, whereas demethylation typically leads to a re-activation of the repressed gene (Bernstein et al., 2007; Klose and Bird, 2006; Kohli and Chang, 2013). Histone modifications include acetylation, methylation, ubiquitination, phosphorylation and sumoylation (Wu et al., 2018). These posttranslational modifications regulate gene expression by changing the chromatin structure. Acetylation and deacetylation are mediated through histone acetyltransferases and histone deacetylases (HDACs), which add or remove an acetyl group. Acetylation usually activates gene expression by opening the chromatin structure (Wu et al., 2018; Peserico and Simone, 2011). Methylation usually has an opposing effect as it tightens chromatin structure, but this depends on the modified residue (arginine or lysine) and the number of modified methyl groups so histone methylation can also lead to gene activation (Greer and Shi, 2012; Black et al., 2012). Non-coding RNAs can also regulate the gene expression by binding 3'-untranslated regions of target mRNAs, which result in cleavage, degradation or translational block of the mRNA (Winter et al., 2009; Wu et al., 2018).

2.5.4 Sin3A corepressor and associated proteins SAP30 and SAP30L

Histone acetylation and deacetylation play a role in gene expression by influencing the chromatin structure. Sin3A is an essential scaffold protein of a histone deacetylation multiprotein complex called Sin3A/HDAC corepressor complex. This multiprotein complex comprises from proteins including HDAC1, HDAC2, the histone binding proteins RbAp46 and RbAp48, the Sin3A associated proteins (SAP) 18 and 30 and SDS3 (Silverstein and Ekwall, 2005). Sin3A acts as a platform bringing together chromatin remodelers and transcriptional regulators, which alter the chromatin structure and gene transcription causing changes in key biological processes (Silverstein and Ekwall, 2005). Based on knockdown studies, Sin3 plays a role in cell proliferation, differentiation, apoptosis and cell cycle regulation (Kadamb et al., 2013). SAP30 is a known component of the Sin3A co-repressor complex and it functions as a bridging and stabilizing molecule between the complex and its corepressors or DNA-binding transcription factors. In cell culture studies, SAP30 was not necessary for Sin3 complex repression but it mediated repression by specific transcription factors (Zhang et al., 1998; Viiri et al., 2009a). SAP30 has a close homolog SAP30-like (SAP30L) with 70 % identity at protein sequence level (Lindfors et al., 2003). SAP30L interacts with Sin3A corepressor complex members and induces transcriptional repression via Sin3A recruitment (Viiri et al., 2006; Korkeamäki et al., 2008).

2.5.5 Protein tyrosine kinases and phosphatases

Phosphorylation and dephosphorylation are one of the most prevalent post-translational modifications of proteins in multicellular organisms. Group of enzymes called protein kinases and phosphatases carries out these reactions in a concerted action to control signaling that underlie a broad spectrum of biological processes ranging from cell fate control to metabolism. Deregulation of such coordination can lead to malignancies. (Paul and Mukhopadhyay, 2004; Motiwala and Jacob, 2006).

Members of the protein tyrosine phosphatase 4A (PTP4A) family, also known as phosphatases of regenerating liver (PRL), are dual specificity phosphatases with oncogenic properties (Campbell and Zhang, 2014; Park et al., 2013; Rios et al., 2013). This protein family consists of three members, PTP4A1-3, and all members are small sized proteins with molecular weight of approximately 20 kDa. PTP4A proteins share a high degree of sequence identity within the phosphatase domain (Campbell and Zhang, 2014; Stephens et al., 2005; Bessette et al., 2008; Rios et al., 2012). They

are involved in multiple cellular activities such as cell proliferation, migration and invasion (Campbell and Zhang, 2014), and they have been associated with several hematological malignancies (Zhou et al., 2011; Park et al., 2013; Beekman et al., 2011). PTP4A3 overexpression correlates with poor prognosis and progression to metastasis in breast cancer, gastric cancer, and colorectal cancers (Campbell and Zhang, 2014).

PTP4A3 is upregulated in patients with chronic myeloid leukemia (CML) and in CML cell lines (Juric et al., 2007; Zhou et al., 2012). PTP4A3 is a downstream target of BCR-ABL1 fusion, and it is upregulated in Ph+ ALL patients, but the exact mechanism is unknown (Zhou et al., 2012). In AML with normal karyotype, the high expression of *PTP4A3* has been associated with poor prognosis (Juric et al., 2007; Zhou et al., 2012; Zhou et al., 2011; Park et al., 2013; Beekman et al., 2011). Overexpression of PTP4A3 is also present in multiple myeloma where knockdown of PTP4A3 inhibits cell migration (Fagerli et al., 2008). In addition, PTP4A3 influences cell cycle regulation and depending on the expression levels the effect may either be positive or negative (Beekman et al., 2011). Knockdown studies by Hjort et al. (2017) demonstrated that PTP4A3 promoted cellular adhesion and migration in BCP-ALL cell lines. They also reported that PTP4A3 expression supported drug resistance towards cytarabine treatment.

As stated earlier, kinases work opposite to phosphatases as they catalyze the phosphate group transfer to their substrates. Perhaps the best-known example of kinase involvement in leukemias is BCR and ABL1 kinases that form a BCR-ABL1 fusion and cause acute and chronic leukemias (Quintás-Cardama and Cortes, 2009). Members of Src family kinases (SFKs) are key regulators of several fundamental cellular processes and operate via cytoplasmic signaling machinery. SFKs have a pleiotropic function as they regulate cell growth, differentiation, cell shape, migration and survival. Many of the SFK family members are identified as oncogenes. SFKs constitutes from nine members from which B cells are primarily expressing LYN, FYN and BLK and T cells LCK and FYN (Parsons and Parsons, 2004). Family members expressed in B cells are known to initiate B cell activation and to play a key role in B cell development possibly via pre-BCR signaling (Gauld and Cambier, 2004). LCK and FYN have a role in T cell activation but also in the development via TCR signal transduction, as TCR based signaling is vital at several developmental stages e.g. pre-TCR signaling (Palacios and Weiss, 2004). In addition, SFKs have been associated with development of leukemias and lymphomas (ALL, AML, CML and Burkitt's lymphoma) (Siveen et al., 2018). However, oncogenic mutations in SFKs have rarely been observed in these malignancies and the association comes

from the SFK activation, which dysregulates pathway signaling (Siveen et al., 2018). LCK has demonstrated overexpression and aberrant activation in BCP-ALL, T-ALL and chronic lymphocytic leukemia and has therefore been studied as a possible target for drug therapies (Cazzaniga et al., 2015; De Keersmaecker et al., 2014; Talab et al., 2013).

3 AIMS OF THE STUDY

The aim of the study was to identify regulators of transcription and signaling in pediatric ALL, and to search for novel targeted therapies. We also investigated the function of SAP30L in zebrafish hematopoiesis.

The specific aims were:

1) to study the expression and function of protein tyrosine phosphatase PTP4A3 in ALL (study **I**).

2) to study expression and function of SOX11 transcription factor in ALL and its use as a prognostic marker in childhood ALL (study **II**).

3) to identify novel targeted therapies for T-ALL (study **III**).

4) to study the function of SAP30L in zebrafish hematopoiesis (study **IV**).

4 MATERIALS AND METHODS

4.1 Cell lines and cell culture (I-III)

The cell lines used in the studies are listed in Table 2. Cell lines were obtained either from Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany) or American Type Culture Collection (J.Cam 1.6, Manassas, VA, USA). Cells were cultured according to cell line-specific instructions. NALM-6, REH, 697, RCH-ACV, KOPN-8, KASUMI-2, Jurkat, J.Cam1.6, MOLT-16, P12-Ichikawa, HPB-ALL and CCRF-CEM cell lines were cultured in in RPMI 1640 (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 10 % FBS (Gibco), 2 mM L-glutamine, 100 μg/ml streptomycin and 100 U penicillin at 37°C in 5 % CO_2. MOLT-4, PEER, and MHH-CALL3 cell lines were cultured in same medium with 20 % FBS. All cell lines (except J.Cam1.6), were authenticated by short tandem repeat (STR)-genotyping (Eurofins Genomics, Ebersberg, Germany). Cell lines were regularly tested for the presence of mycoplasma.

Table 2. Characteristics of cell lines used in studies I, II and III.

Cell line	Disease	Genetics	Donor age	Study
697	BCP-ALL	TCF3-PBX1	12 years	I, II, III
CCRF-CEM	T-ALL	NKX2-5 expression	3 years	I, III
HPB-ALL	T-ALL	TLX3 subgroup*	14 years	I, III
Jurkat	T-ALL	TAL1 subgroup*	14 years	I, III
KOPN-8	BCP-ALL	KMT2A-ENL	3 months	I, III
MOLT-4	T-ALL	TAL1 (with LMO2) subgroup*	19 years	III
MOLT-16	T-ALL	TAL1 subgroup*	5 years	I, III
NALM-6	BCP-ALL	ETV6/PDGFRB, DUX4/ERG**	19 years	I, II, III
P12-Ichikawa	T-ALL	LYL1/LMO2 ***	7 years	I, III
PEER	T-ALL	LYL1/LMO2 *** NUP214-ABL1****	4 years	III
RCH-ACV	BCP-ALL	TCF3-PBX1	8 years	II
REH	BCP-ALL	ETV6-RUNX1	15 years	I, II, III
J.cam6.1 (Jurkat)	T-ALL	Jurkat with deletion of exon 7	14 years	III
KASUMI-2	BCP-ALL	TCF3-PBX1	15 years	II
MHH-CALL-3	BCP-ALL	TCF3-PBX1	11 years	II

**subgroup specification was carried out in our laboratory **Zhang et al., 2016 ***mixed subgroup specificated in our laboratory ****NUP214-ABL1 fusion confirmed in our laboratory.*

4.2 Cell viability assay (I-III)

Cell viability was measured by using AlamarBlue® reagent (Life Technologies, Thermo Fisher Scientific, Waltham, MA, USA) in which the active ingredient is resazurin. In the assay, viable cells reduces resazurin to highly fluorescent resofurin and the fluorescence is measured after 2 h by using the Tecan fluorometer Infinite 200 (Tecan, Männedorf, Switzerland) with excitation of 560 nm and emission of 590 nm. Reagents are added to 10 % of the sample volume. Cell viability measurements were performed with three to four technical replicates per sample in each experiment. Depending on the cell line of interest, 10 000 to 50 000 cells per well were plated in the beginning of the experiment with culture media, supplements and drugs in 96-well plates.

4.3 Modification of gene expression

4.3.1 Transient modifications (I-III)

Transient modification of gene expression in suspension cells were performed with 4D-Nucleofector™ (Lonza, Basel, Switzerland) mediated transfection. Transient gene knockdowns were carried out with small interfering RNAs (siRNA). Target gene specific siRNAs were used either as single siRNA oligos or as mixtures of many oligos. Commercial siRNA mixes and single oligos are listed in a Table 3. Nucleofections were carried out according to manufacturer's instructions (see Table 4 for details). Depending on the chosen nucleofection method, the number of cells varied from 0.5 to 1 million per reaction. After transfections, the cells were assayed for cell viability or collected for RNA extraction. The success of knockdown was validated by quantitative reverse-transcription polymerase chain reaction (RT-qPCR) and/or western blot (WB).

Table 3. Characteristics of siRNAs used in studies I, II and III.

Target gene	Primers 5' -> 3' / Commercial name	Origin/Source	Catalog No.	Study
si*SCR* sense	GUUGCGUAGCGUACGUCGCAA	Own/Sigma		II
si*SCR* antisense	UUGCGACGUACGCUACGCAAC	Own/Sigma		II
si*SOX11* sense	GAUAAGAUGUCGUGACGCA	Kuo et al. 2015/ Sigma		II
si*SOX11* antisense	UGCGUCACGACAUCUUAUC	Kuo et al. 2015/ Sigma		II
PTP4A3	PTP4A3 (ID 11156) Trilencer-27 Human siRNA	OriGene	SR307645	I
ABL1	ABL1 (ID 25) Trilencer-27 Human siRNA	OriGene	SR300017	III
FYN	FYN (ID 2534) Trilencer-27 Human siRNA	OriGene	SR301682	III
LCK	LCK (ID 3932) Trilencer-27 Human siRNA	OriGene	SR302659	III
LYN	LYN (ID 4067) Trilencer-27 Human siRNA	OriGene	SR302752	III
MAP2K5	MAP2K5 (ID 5607) Trilencer-27 Human siRNA	OriGene	SR303764	III
MAP4K5	MAP4K5 (ID 11183) Trilencer-27 Human siRNA	OriGene	SR307671	III
Negative Control	Trilencer-27 Universal Scrambled Negative Control siRNA Duplex	OriGene	SR30004	I, III

Table 4. Nucleofection solutions and programs for the indicated cell lines.

Cell line	Nucleofection solution	Nucleofection program
REH	SF	CA-137
697	SG	CA-137
Jurkat	SE	CL-120
MOLT-4	SF	CA-137
P12-Ichikawa	SF / SG	CA-137
RCH-ACV	SF / SG	CM-137

4.3.2 Stable modifications (I)

Stable cell lines with inducible expression of ETV6-RUNX1 in NALM-6 cell line and knockdown of ETV6-RUNX1 in REH cell line were created by Teppo et al. (2016). *ETV6-RUNX1* expression was induced with 500 ng/ml doxycycline (Clontech, Fremont, CA, USA).

4.4 Drug treatments and inhibitions

4.4.1 *In vitro* drug treatments (III)

Dasatinib (9052, Cell Signaling Technology, Danvers, MA, USA) was stored in dimethyl sulfoxide (DMSO) stock solution at -20°C. Dasatinib was used as 10-fold dilutions (1-1000 nM) in 96 well plate with 5000 – 10 000 cells/well. After 72 h incubation with the drug or DMSO-control, cell viability was assayed as previously described (section 4.2.). To obtain relative cell viabilities, drug-treated samples were normalized to DMSO control sample. Experiments were performed with three or four technical and two or three biological replicates in each experiment.

4.4.2 Mononuclear cell isolation and *ex vivo* drug treatments (III)

Bone marrow aspirates or peripheral blood samples were used as a starting material for mononuclear cells in the *ex vivo* drug treatments. Cells were incubated 72 h in drug media or DMSO control media before the cell viability measurements (see more details in Laukkanen et al., 2017, supplementary information). Collected data were

normalized to positive controls (containing 100 μmol/L benzethonium chloride, effectively killing all cells) and negative control (DMSO only). Data were used to calculate the drug sensitivity scores (DSS) from the dose-response curves as previously described (Pemovska et al., 2013). Dasatinib concentration range was 0.1-1000 nM as a 10-fold dilutions and glucocorticoid (dexamethasone, prednisolone and methylprednisolone) concentrations were 1-10 000 nM.

4.4.3 Inhibition of PTP4A3 phosphatase activity (I)

PTP4A3 inhibition was performed with PRL-3 inhibitor I (Lot # 047K4614V, Sigma-Aldrich, St. Louis, MO, USA) which was stored in DMSO stock solutions at -20°C. Cells were plated in 96-well plates with the desired concentrations (0 – 50 μM) of PRL-3 inhibitor I. Cell viability measurements were performed as previously described in section 4.2. with three technical and four biological replicates.

4.5 Gene expression analysis

4.5.1 RNA extraction (I-IV)

Total RNA extractions accompanied with DNA removal from the cultured cells were performed using either GeneJET RNA Purification Kit (Thermo Fisher Scientific), DNA-free™ DNase Treatment & Removal kit (Bio-Rad, Hercules, CA, USA), or PureLink™ RNA Mini Kit with On-Column PureLink® DNase Treatment Protocol (Ambion® by Life Technologies and Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Total RNA extractions from adult zebrafish tissue and zebrafish embryos was performed with RNeasy Mini Kit (Qiagen, Venlo, Holland). The quality and concentration of the RNA were analyzed with nanodrop spectrophotometer (Thermo Fischer Scientific).

4.5.2 Quantitative PCR (I-IV)

Extracted total RNA (0.25-1 μg) was used for reverse transcription synthesis of complementary DNA (cDNA) with iScript (Bio-Rad). Synthesized cDNA was used for the RT-qPCR which was performed with SsoFast EvaGreen® Supermix (Bio-

Rad) according to manufacturer`s instructions. BioRad CFX96TM Real Time System (Bio-Rad) was run with the following program: initial denaturation at 96°C for 30 s, 39 cycles of denaturation at 96°C for 2 s, annealing/extension at 60°C for 5 s, and plate read (melt curve 60°C → 95°C). Each RT-qPCR measurement was performed in triplicate, and the relative $2^{-\Delta\Delta C_T}$ method was used for quantification (Livak et al., 2001). Chromatin immunoprecipitation (ChIP) qPCR runs from ChIP samples were performed from two independent biological replicates with duplicate runs and the samples were analyzed with percent input method to normalize the data. Used primers and their sequences are listed in Table 5.

Table 5. qPCR primer sequences.

Target	Primers 5' -> 3'	Study	Target	Primers 5'->3'	Study
SOX11 forw.	CGGTCAAGTGCGTGTTTCTG	II	*BLK* rev.	CCCCTTCATCCAGGCAGC	III
SOX11 rev.	CACTTTGGCGACGTTGTAGC	II	*LYN* forw.	CCTGGAGGAGCATGGAGAAT	III
GAPDH forw.	TCCATGACAACTTTGGTATCGTGG	I, III, IV	*LYN* rev.	CATCACCATGCACAGGGTC	III
GAPDH rev.	GACGCCTGCTTCACCACCTTCT	I, III, IV	*FRK* forw.	CCTGTGTGTCAAGCTGGGG	III
PBGD forw.	CGCATCTGGAGTTCAGGAGTA	I, II	*FRK* rev.	GTGCCTCCCTCAGGAAGTC	III
PBGD rev.	CCAGGATGATGGCACTGA	I, II	KIT forw.	GCATTGTTCTGTGGACCAGG	III
ETV6-RUNX1 forw.	TGCACCCTCTGATCCTGAAC	I	KIT rev.	GAAGTCACCGTGATGCCAG	III
ETV6-RUNX1 rev.	AACGCCTCGCTCATCTTGC	I	*PDGFRa* forw.	CTATGTGCCAGACCCAGATG	III
PTP4A3 forw.	CGCCACGCTCAGCACCTTCA	I	*PDGFRa* rev.	CGGTGGCCTCACAGATATAG	III
PTP4A3 rev.	GAACTTGGCCTTCACCAGGC	I	*PDGFRb* forw.	CTTTGTGCCAGATCCCACC	III
g*GAPDH* forw.	CACAGTCCAGTCCTGGGAAC	I	*PDGFRb* rev.	GTCCCCAATGGTGGTTTTGC	III
g*GAPDH* rev.	TAGTAGCCGGGCCCTACTTT	I	*BTK* forw.	GAGGTGAAGAGTCCAGTGAA	III
e*ITGA4* forw.1	CCAACCTCTCTGTTGGGGTC	I	*BTK* rev.	GTCTGAGAGCAGCAGAGATAC	III
e*ITGA4* rev.1	GTGAGCAGGCTAGCAGCAG	I	*TEC* forw.	GGTCATTTCAGTAGAGACGTAC	III
e*PTP4A3* e1 forw.1	ACAGGAGCCAAGGTCAGTC	I	*TEC* rev.	GGACACCACTTCACAGGAAAC	III
e*PTP4A3* e1 rev.1	AGGCCGCCCCTGGGCTGGGA	I	*BRK* forw.	GGACCCCGTGTACATCATCA	III
e*PTP4A3* e1 forw.3	GGTCTGACCCCAAGTACCCG	I	*BRK* rev.	CATGGGAGAGGTAGACGTC	III
e*PTP4A3* e1 rev.3	AAGTCGACTGGGGCTAGGTC	I	*ACK* forw.	GGATGAGTAAGGTGTTCAGTG	III
e*PTP4A3* e1 forw.4	CTCAGGAGGCCGCTGTTTC	I	*ACK* rev.	CACTTCACAGCCACACTCAC	III
e*PTP4A3* e1 rev.4	CTTCCAGCCAGGGGTGTTC	I	*EGFR* forw.	GTGACTGCTGCCACAACCA	III
e*PTP4A3* e2 forw.1	CCTGAGAGTGTCTTGGGGTT	I	*EGFR* rev.	GCCGTGATCTGTCACCACA	III
e*PTP4A3* e2 rev.1	TCACGCTCACTTCCTCTTCA	I	*DDR1* forw.	CCACAGCAGGTTGGAGAGCA	III
e*PTP4A3* e2 forw.3	GGAGGCCTTGTTTACCCCTC	I	*DDR1* rev.	CCTGAGATCACCTCCTGA CC	III
e*PTP4A3* e2 rev.3	CACAGACCTCTGAGCCGATG	I	*EPHA2* forw.	CCAGGCAGGCTACGAGAAG	III
e*PTP4A3* e2 forw.4	AGATGCTCTGCTGACGGATG	I	*EPHA2* rev.	GCAAGGCATCGACGCTGG	III
e*PTP4A3* e2 rev.4	CTCCAGATGGCTTCCTGTGG	I	*MAP2K5* forw.	CCTGAAGGTGAATACTCGGG	III
e*PTP4A3* e2 forw.5	CTTGTTTGTTGGGTGCCTGG	I	*MAP2K5* rev.	GCTTTGTAGACTGTGCCTCC	III
e*PTP4A3* e2 rev.5	AGGGACCCTCCTGAGACATC	I	*MAP4K5* forw.	GGAGCCTGGAGATGATTTTTC	III
e*PTP4A3* e2 forw.6	AATGGCCCTGAAAGGGAAGC	I	*MAP4K5* rev.	GCAAATAGGCAAGACCCTGTA	III
e*PTP4A3* e2 rev.6	TCCGTCAAGTACAGCAGCTC	I	*EPHB2* forw.	CACCGTCTGCCGAGGTTGT	III
e*PTP4A3* e3 forw.1	CTGGCTGGGCTTCCTCTTC	I	*EPHB2* rev.	GATGGTTGTGCAGGGCATG	III
e*PTP4A3* e3 rev.1	CCTGCCGCCTAAATATAGCC	I	*EPHB6* forw.	GACAAGGCCTGCCAAGCCT	III
e*PTP4A3* e3_forw.3	GGCCGCATTTCCATTTCCAA	I	*EPHB6* rev.	CCGATGGAGGACCAGTGCA	III
e*PTP4A3* e3_rev.3	TGGGGAGGCCTGGACAGA	I	*NUP214* exon 2, forw. *	ACCCGGAGATGATCCCAACAAAAT	III
e*PTP4A3* e4 forw.1	TCTTTCCCCACCCCAACTAC	I	*NUP214* exon 20, forw. *	CACCAAATCCTTGCCCAAAGTACC	III
e*PTP4A3* e4 rev.1	CCCTAGGGCCCTGTAACAG	I	*NUP214* exon 23, forw.	CAATGCTTGCCACGAAAACC	III
LCK forw.	GCATCCTGGAGCAGAGCGG	III	*NUP214* exon 28, forw.	TCACACCAACACCGTCTTCT	III
LCK rev.	GCTCTCGCTCTCCCGGATG	III	*NUP214* exon 29, forw.	TGGGTTCAGCTTTTGCCAAGCT	III
ABL1 forw.	GTGGCCAGTGGAGATAACACT	III	*NUP214* exon 31, forw.	GGTTGGTTTCTTCAGTGGCCTTGG	III
ABL1 rev.	CAGATACTCAGCGGCATTGCG	III	*NUP214* exon 34, forw.	CTGGTTTTGGATCAGGCACAGGA	III
SRC forw.	GATCCGCAAGCTGGACAGCG	III	*ABL1* exon 3, rev *	CCATTTTTGGTTTGGGCTTCACACC	III
SRC rev.	CATCCACACCTCGCCAAAGC	III	*ABL1* exon 1, forw.	GTAGCCAAAGACCATCAGCG	III
YES forw.	GCTGGTTTAACAGGTGGTGT	III	*EF1A* (dr), forw. **	CTGGAGGCCAGCTCAAACAT	IV
YES rev.	CCAAAATACCATTCTTCTGCC	III	*EF1A* (dr), rev **	ATCAAGAAGAGTAGTACCGCTAGCATTAC	IV
FYN forw.	GCTTCGGTGTGACCTCCAT	III	*nkx2.5* (dr), forw. ***	GCGAAGACCTTCAGGAGGACAAAGGCAAC	IV
FYN rev.	GAGCGGGCTTCCCACCAAT	III	*nkx2.5* (dr), rev ***	CGAGGCTTCCTCCTCTTCCTCTGCTTGGG	IV
FGR forw.	GGTGACTGGTGGGAGGC	III	*ppox* (dr), forw.	GTGATCAGTAAAGGGACAGAA	IV
FGR rev.	CTTCACATGATCGCCTCTGG	III	*ppox* (dr), rev.	CTGCATGCCCCTCTTTAG	IV
HCK forw.	GGACCAGATGGTGGTCCTA	III	*SAP30L* (dr), forw.	AACGCG TCC TTC AGC AAG AGG AT	IV
HCK rev.	TCGCACGGACAAAGAGTAGC	III	*SAP30L* (dr), rev.	GTTGACCTGAAGCTGGAACAGATC	IV
BLK forw.	CCAGGTCACTCGTCACAGG	III			

** Burmeister et al., 2006, ** Tang et al., 2007, *** Targoff et al., 2008, dr* = Danio rerio

4.5.3 Global Nuclear Run-On sequencing (I)

Primary bone marrow or blood samples from ALL, BCP-ALL (hyperdiploid, ETV6-RUNX1) and T-ALL, patients and healthy donor were used on global run-on sequencing (GRO-seq) assay together with modified REH cell lines. These samples were used to capture different transcriptional activity patterns between ALL samples, healthy donors and cell lines. Nuclei extractions and GRO-seq reactions were carried out as previously described in Heinäniemi et al. (2016).

4.5.4 RNA-sequencing sample preparation (II)

Specific siRNAs were used to knockdown expression of *SOX11* and scrambled siRNAs for control samples. After 48 h of the nucleofection, 3 million cells per sample were collected and total RNA was extracted. The efficacy of knockdown was verified by RT-qPCR and WB. Three independent biological replicants were collected per cell line with 25 ng of RNA in 40 µl volume. Library preparation and stranded mRNA sequencing was done at the Finnish Functional Genomics Centre (FFGS, Turku, Finland). See more details in original study II.

4.5.5 Chromatin immunoprecipitation (I)

DNA-protein interactions were analysed by ChIP assay where DNA fragments were immunoprecipitated from crosslinked cells by using two pooled antibodies against ETV6 (sc-166835, Santa Cruz Biotechnology Inc., Dallas, TX, USA; RRID:AB_2101020 and HPA000264, Atlas Antibodies, Bromma, Sweden; RRID:AB_611466) as described in Teppo et al. (2016). Collected ChIP samples were used in ChIP-qPCR.

4.5.6 Gene fusion detection: *NUP214-ABL1* and *SIL-TAL1* (III)

RNA from patient samples was used as template for reverse transcription to cDNA to serve as a template to *NUP214-ABL1* fusion detection. The known range of *NUP214-ABL1* fusion breakpoints was covered using seven different primer pairs (Hagermeijer et al., 2010). RNA quality was controlled using additional primer pair targeting *ABL1*. This was done regardless of the sample's fusion status. All primers

are listed in table 5. In one case the *NUP214-ABL1* and *SIL-TAL1* fusions were detected from RNA sequencing (RNA-seq) data (See details in Laukkanen et al., 2017, supplementary information).

4.6 Protein expression analysis

4.6.1 Protein extraction and western blotting (II-IV)

Cultured cells were lysed and proteins were extracted from the cell lines with M-PER reagent (Thermo Fisher Scientific) according to manufacturer's instructions. The total protein concentrations were determined using DC Protein Assay (Bio-rad). 15-20 μg of a protein per sample was loaded for the precasted 10 % Mini-PROTEAN® TGX Stain-Free™ Gels (Bio-Rad) or casted on 10-12 % SDS–polyacrylamide gel. After the gel run, Trans-Blot® Turbo™ Pack (Bio-Rad) was used to transfer proteins from the gel to the nitrocellulose membrane with Trans-Blot® Turbo™ Transfer system according to manufacturer's instructions (Bio-Rad). Proteins from casted gels were transferred to Amersham Protran Supported nitrocellulose membrane (GE Healthcare, Helsinki, Finland) under standard procedures. Prestained protein ladder PageRuler PLus (#26619, Thermo Fischer Scientific) was used as a protein size marker in WB. Chemiluminescence reaction by Amersham ECL reagent was detected with ChemiDoc™ XRS+ using Image Lab™ Software (Bio-Rad).

Zebrafish samples for WB were prepared from larvaes at 2 dpf. 10 larvae / 50μl was collected to 2 x Laemmli buffer and samples were lysed by boiling and passing through a small needle. Lysate was centrifuged and the soluble fraction was collected for the WB analysis. Antibodies and their working dilutions are listed in Table 6.

Table 6. Antibodies used for western blotting and immunofluorescence staining (Studies II, III and IV).

Antibody	Dilution (WB)	Dilution (IHC/IF)	Product #	Manufacturer	Study
H3	1:75000		ab4729	Abcam	II
SOX11	1:1000		HPA000536	Sigma	II
SOX11		1:50	MRQ-58	Sigma-Aldrich	II
LCK	1:1000		2714	Cell Signaling Technology	III
FYN	1:2000		4023	Cell Signaling Technology	III
LYN	1:500		4576	Cell Signaling Technology	III
ABL1	1:250		2862	Cell Signaling Technology	III
β-actin	1:100000		sc-1615	Santa Cruz Biotechnology	III
Nucleophosmin (NPM)	1:8000	1:2000	35-5200	Invitrogen	IV
Prohibitin (PHB)	1:1000	1:50	sc-28259	Santa Cruz Biotechnology	IV
SAP30L	1:250		-	Own; Korkeamäki et al. 2008	IV
TDP-43	1:1000	1:150	T1705	Sigma-Aldrich	IV
TDP-43-N	1:1000	1:150	SAB4200006	Sigma-Aldrich	IV
Horseradish peroxidase conjugated anti-rabbit	1:2000/1:5000		P0217	Agilent Technologies	II, III
Horseradish peroxidase conjugated anti-mouse	1:5000		P0260	Agilent Technologies	III

4.6.2 Immunohistochemistry and fluorescence *in situ* hybridization (II)

In order to analyze the expression of SOX11 protein in bone marrow biopsies, two pathologists analyzed the nuclear expression of SOX11 stained slides independently with a light microscope without knowledge of the clinical data. Antibodies and their working dilutions are listed in Table 6. In the case of different scoring, a third pathologists analyzed the data to decide the score. Formalin fixed paraffin embedded (FFPE) MCL sample was used as positive control and remission bone marrow sample as negative control. Cases with less than 20 % of immunohistochemically positive SOX11 cells were classified as negative, cases with 20-50 % of positive cells were graded positive and cases with over 50 % of nuclear positive cells were

considered strongly positive. Fluorescence *in situ* hybridization (FISH) was performed for cases that were not tested at diagnosis with current methods. Bone marrow culture or FFPE samples were used for the FISH with following probes: Metasystems E2A Break Apart Probe 19p13 (lot# 18216), Metasystems XL MLL plus Break Apart Probe 11q23 (lot 18451), Metasystems XL BCR/ABL1 plus (lot# 19082) and Metasystems XL t(12;21) (lot 19133). See more details in original study II.

4.6.3 Flow cytometry analysis (II)

Cells permeabilization was carried out with Fix&Perm reagent according to manufacturer's instructions (GAS003, Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) for flow cytometry analysis and the data collection with Beckman Coulter Navios cytometer (Beckman Coulter, Brea, CA, USA) using Red laser (638nm) and 660/20 bandpass filter. After permeabilization cells were stained with Alexa Fluor 647® conjugated rabbit anti-human SOX11 antibody [EPR8191 (2)] (ab198540, Abcam, Cambridge, UK), while CD3-APC antibody (345767) (Becton Dickinson, Franklin Lakes, NJ, USA) served as a negative control. Data were analyzed with Kaluza software (Beckman Coulter, Brea, CA, USA).

4.7 Clinical data and patient samples (II-III)

Patient data in study II included 126 pediatric acute lymphoblastic leukemia cases treated in Tampere University Hospital from the year 1990 onwards (POTTI registry). Patients under 18 years old were included in this study with additional clinical information (age at diagnosis, leukocyte count, blast percentage, gender, relapses, deaths, stem cell transplantation, duration of treatment and genetics). Case classification was based on the WHO 2017 Classification of tumors of hematopoietic and lymphoid tissues (Swerdlow et al., 2017).

For *ex vivo* drug treatments, 22 T-ALL patient samples (bone marrow or peripheral blood) were used in study III.

4.8 Zebrafish methods (IV)

4.8.1 Zebrafish maintenance

Wild type AB zebrafish were maintained at the Tampere zebrafish core facility under standard conditions at 28.5°C as described in Westerfield (1995). Animal care was carried out in accordance with the Finnish Laboratory Animal Welfare Act 62/2006 and the Laboratory Animal Welfare Ordinance 36/2006.

4.8.2 Microinjections: Morpholino knockdown and mRNA rescue

Zebrafish embryos at their 1-4 cell stage were used for the microinjections, where 1 nl of the injection mix was injected into the yolk sac. Injection mix contained morpholino, mRNA (where indicated) and rhodamine dextran as a marker color. Injections were carried out with The PV830 Pneumatic PicoPump injection system (WPI, Inc.). Two independent antisense morpholino oligonucleotides (SAP30L-MO1: CCTCCTCCGTGCTGAAGCCGTTCAT and SAP30L-MO2: TCATCCCGGAGACCGGACACCGAGC) were used to target the start site of the zebrafish *SAP30L* mRNA (Ensembl ID: ENSDARG00000030213). Morpholino targeting *TP53* (Robu et al., 2007) and the random control morpholino (RC-MO: CCTCTTACCTCAGTTACAATTTATA) were used as controls. All morpholinos were purchased from Gene-Tools, LLC (Philomath, OR, USA). The effective morpholino concentrations were determined experimentally by titration. In the mRNA rescue experiments SAP30L-MO1 or SAP30L-MO2 were co-injected with a capped synthetic mRNA encoding the full-length zebrafish *SAP30L* or full-length zebrafish *nkx2.5*. Corresponding plasmid templates and a T7 mMESSAGE mMACHINE Kit (Ambion) were used to synthesize the required mRNAs according to the manufacturer's instructions.

4.8.3 Whole-mount *in situ* hybridization

Linearized plasmids containing the full-length zebrafish *SAP30L* cDNA (in the sense or antisense orientation) under the *T7* promoter control were used as templates for the DIG-labeled probes. DIG RNA Labeling Mix (Roche Diagnostics, Espoo, Finland) was used to prepare the sense and antisense probes for the zebrafish

SAP30L (Thisse and Thisse, 2008). Whole-mount *in situ* hybridization (WISH) was performed for the collected zebrafish larvae. Larvae were collected at 12, 24, 36, 48, and 72 hpf as previously described by Thisse and Thisse (2008). WISH was performed according to Thisse and Thisse (2008) with alkaline phosphatase detection and the BM Purple substrate (Roche Diagnostics). The following modifications were used in the protocol: 2 ml eppendorf tubes were used during the entire experiment, hybridization was carried out at 65°C, sense and antisense probe concentrations were 130 ng per reaction and dilution ratio for the anti-DIG antibody (sheep anti-digoxigenin-AP Fab fragments, Roche Diagnostics) was 1: 2000. After the desired level of the staining was obtained, the reaction was stopped and dissecting microscope was used to photograph and analyze the fish. Anatomical images by Haffter et al. (1996) were used as reference to assess the localizations of the staining.

4.8.4 O-dianisidine staining

Hemoglobin in red blood cells of 3 dpf zebrafish larvae was detected with o-dianisidine staining as previously described in Ransom et al. (1996). Staining was visually scored for each specimen as normal, reduced or severely reduced based on the level of the staining. In the normal staining category, staining was comparable to the wild type larvae. Larvae with in the reduced category had o-dianisidine staining visibly decreased compared to the wild-type specimens and in the severely reduced group there was very little to no visible staining.

4.8.5 Analysis of heart morphology and function

Heart morphology of zebrafish larvae was assessed visually at 5 dpf and categorized as normal or deformed. Unanaesthetized 5 dpf larvae were observed visually to determine the heart rates in morphant and control larvaes. Heart function, ventricular and atrial performance, were analyzed with 10-20 s time-lapse videos. The function was calculated from the longitudinal fractional shortening according to the formula (LDd-LDs)/LDd, where LDd is the longitudinal diameter at diastole and LDs the diameter at systole.

4.8.6 Histological staining

The zebrafish larvae were fixed with 4 % PFA-PBS at + 4°C followed by embedding in 2 % agarose. Larvae samples were dehydrated by incubating them in rising ethanol solution series (70 %, 96 %, and abs EtOH) for 1-2 hours per solution before the final incubation xylene for at least 1 hour. Dehydration was followed with paraffin embedding which enabled cutting of 5 mm-thick sections in longitudinal orientation and fixing on glass slides. Fixed sections were then deparafinized by incubation in xylene (2 x 4 min), abs EtOH (2 x 3 min), 96 % EtOH (2 x 3 min), and 70 % EtOH followed by rinsing once in water. Samples were then hematoxylin-eosin stained by incubation with the following protocol and solutions: Mayer's hematoxylin (2 min), running water (2 min), water (1.5 min), 70 % EtOH (15 s), eosin Y (15 s), 96 % EtOH (2 x 30 s), abs EtOH (2 x 1 min), and xylene (1–4 min). Thereafter, slides were mounted and BX60 microscope with Cell^D program (Olympus) was used to take pictures.

4.9 Gene expression data and data analysis (I-IV)

Three different data sets were utilized to study differential gene expressions of *PTP4A3* and *SOX11* (study I-II). A combined microarray-based gene expression data set (GeneChip™ Human Genome U133 Plus 2.0 Array, Thermo Fisher Scientific) processed by Heinäniemi et al. (2016) and Pölönen et al. (2019), and two independent data sets with gene expression profiles from patients with pediatric ALL, GSE47051 (Nordlund et al., 2013) and from St. Jude Children`s Research Hospital's Pediatric Cancer data portal (referred from now as St. Jude data) (PeCan).

For study II, differential gene expression of SOX11 knockdown was investigated using RNA-seq (GSE123943) by using edgeR R-package and differentially expressed genes were filtered with adjusted p-value ($p<0.05$). Differentially expressed genes were illustrated with Venn diagram (Interactive Venn online tool, http://www.interactivenn.net/) and a heatmap (ComplexHeatmap R-package). Gene set enrichment analysis was performed with GSEA 3.0 software and enrichment of GO terms was studied with GSEA software and GOrilla online tool (study II). For more detailed information see original study II. *SOX11* mRNA expression levels were further studied from previously published transcriptomic data from 116 BCP-ALL patients and correlated with clinical parameters (Marincevic-

Zuniga et al., 2017). 112 samples out of the 116 also had methylation data available (GSE47051) (Marincevic-Zuniga et al., 2017).

For study III, normalized microarray data was first transformed with Principal Component Analysis to reduce uninteresting variance effect of the data. Optimized parameter *k* for k-means clustering was used to find suitable clustering complementing the subgroups of T-ALL according to chosen transcription factors (TAL1, TLX1, TLX3, HOXA10, NKX2-1 and LYL1). For more detailed information, see Laukkanen et al. (2017).

For study IV, zebrafish total RNA samples were prepared according to manufactures recommendation (Santa Clara, CA, USA) to Agilent's 4 × 44K Zebrafish V3 Gene Expression Microarray. Zebrafish microarray data were analyzed using the R programming language and environment (R Development Core Team, 2008) and its Bioconductor module (Gentleman et al., 2004). Statistical testing was performed with the Limma Package (Smyth, 2005) with filtering thresholds (fold change $FC \geq |3|$ and p-value $p \leq 0.001$). For more detailed information see Teittinen et al. (2012).

4.10 Statistical analysis (I-IV)

Expression levels of mRNA measured by RT-qPCR were evaluated with Mann-Whitney U-test (study I). Differential expression of target genes between distinct leukemia subgroups were evaluated with Kruskal-Wallis H-test and/or Mann-Whitney U-test (study I, II, III). Cell viability differences between treated and control cells were evaluated with Student's t-test (study I) or Mann-Whitney U-test and correlation coefficients with Spearman method (study III). Clinical status of the patient samples were analyzed with Kaplan-Meier survival analysis and Chi-squared test was performed on *SOX11* expression and clinicopathological prognostic variables (study II). In study IV, all quantitative data is given as mean ± standard deviation. Differences between groups, when appropriate, were evaluated with two-tailed Student t-test or the $\chi 2$ test for crosstabulation.

4.11 Ethical considerations (I-IV)

Commercial cell lines were used and gene expression data from patients were retrieved from publicly available databases (studies I-III). In study II, National

supervisory authority of welfare and health (VALVIRA, Helsinki, Finland) approved the usage of primary patient samples under the ethics committee approval code R16054. Furthermore, patient samples from Tampere University Hospital were collected under the Pirkanmaa Hospital District Ethics committee permit number R13109. For study III, collection of bone marrow and peripheral blood samples were carried out following written informed consent by protocols approved by local institutional review boards and in accordance with the Declaration of Helsinki. Zebrafish experiments in study IV were carried out at the Tampere Zebrafish core facility under the permission of the State Provincial Office of Western Finland (permission LSLH-2007-7254-/Ym-23). Zebrafish embryos and larvae that were used in these studies were under the age of seven days, so no separate permission was required according to the effective practice in Finland.

5 RESULTS

5.1 Expression and role of *PTP4A3* in BCP-ALL

5.1.1 *PTP4A3* is expressed in a subtype-specific manner in BCP-ALL

A previously described gene expression data set (Heinäniemi et al., 2016) was utilized to investigate expression of different PTP4A family member across acute leukemias and healthy cells of hematopoietic origin. To visualize the expression patterns, a Hemap cancer-map was used which projects the data in two dimensions by using a t-SNE algorithm (Pölönen et al., 2019; van der Maaten and Hinton, 2008). *PTP4A1* showed no difference in expression between disease and healthy controls, whereas *PTP4A2* showed elevated expression for some of the T-ALL samples. *PTP4A3*, on the other hand, showed markedly increased expression in BCP-ALL samples and also in a few T-ALL cases (see Fig. 1A in study I). When we examined the *PTP4A3* expression across all hematological malignancies, the strongest expression was observed in ETV6-RUNX1- and BCR-ABL1-positive ALL cases. Multiple myeloma showed overexpression but in other malignancies the expression level was low or was not detectable (Fig. 9).

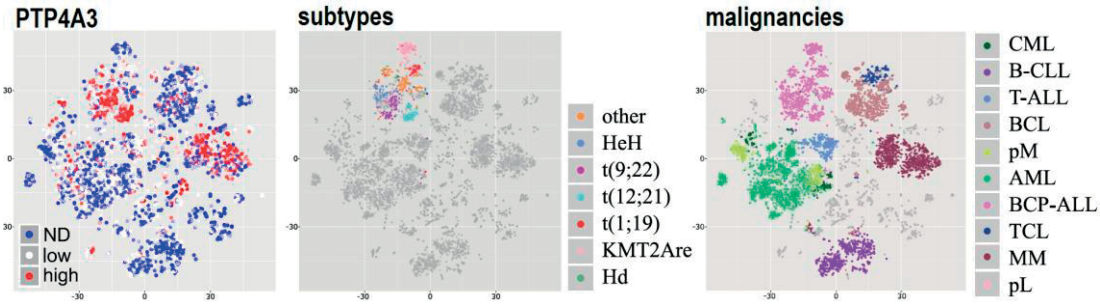


Figure 9. Expression levels *of PTP4A3* in hematological malignancies and BCP-ALL subtypes shown in a t-SNE-based two-dimensional map. Expression level is indicated with colors blue (ND = not detected), white (low expression), red (high expression). Disease subtypes are indicated in the right side of the figure with colors. Abbreviations: AML = acute myeloid leukemia, B-CLL = B cell chronic lymphocytic leukemia, BCL = B cell lymphoma, BCP-ALL = precursor B cell acute lymphoblastic leukemia, CML = chronic myeloid leukemia, Hd = hypodiploid, HeH = hyperdiploid, KMT2Are = KMT2A-rearrengement, MM = multiple myeloma, pM = proliferative myeloid, pL = proliferative lymphoid, t(1;19) = TCF3-PBX1, t(9;22) = BCR-ABL1, t(12;21) = ETV6-RUNX1, T-ALL = T cell acute lymphoblastic leukemia, TCL = T cell lymphoma.

When compared to other BCP-ALL leukemias, the ETV6-RUNX1 subtype showed three to four-fold higher expression of *PTP4A3* (see Fig. 1B in study I; Mann-Whitney U-test, p = 1.894997e-52). To further validate results, two more leukemia data sets were retrieved and analyzed. *PTP4A3* showed similarly high expression in ETV6-RUNX1 and BCR-ABL1 positive leukemia subtypes in both data sets: GSE47051 (Nordlund et al., 2013) and St. Jude data (PeCan) (see Fig. 1B in study I). Interestingly, KMT2A-rearranged leukemias showed evidence of low level of *PTP4A3* expression in all three data sets.

5.1.2 Expression of *PTP4A3* is regulated by ETV6-RUNX1 fusion protein

Since *PTP4A3* showed higher expression in ETV6-RUNX1 leukemia samples, we next sought to investigate if ETV6-RUNX1 fusion regulates expression of *PTP4A3*. As a first step, we screened expression of *PTP4A3* in a panel of leukemia cell lines with RT-qPCR. Strongest expression was demonstrated in REH cell line, which carries endogenously the ETV6-RUNX1 fusion, while moderate expression was seen in two BCP-ALL cell lines (KOPN-8 and 697) and in healthy B cells and T cells. (see Fig. 2A in study I). NALM-6 cell line showed a low level of *PTP4A3* expression. We chose the REH cell line for knockdown experiments and NALM-6

cell line for overexpression studies to see if ETV6-RUNX1 fusion regulates the *PTP4A3* expression directly (Teppo et al., 2016). In the inducible NALM-6-ETV6-RUNX1 cell line model, expression of *ETV6-RUNX1* fusion can be controlled by addition of doxycycline into the cell culture media (Teppo et al. 2016). Induction of *ETV6-RUNX1* increased expression of *PTP4A3* (see Fig. 2B in study I) while no marked downregulation was seen with the knockdown of *ETV6-RUNX1* in the REH cell line (see Fig. 2C in study I). We also investigated transcription of *PTP4A3* at primary transcript level by GRO-seq, and the signal decrease in at *PTP4A3* locus was more evident in REH-shETV6-RUNX1 as compared to control cells (see Fig. 3A in study I). GRO-seq also revealed robust expression of *PTP4A3* in patient samples so that ETV6-RUNX1 cases had notably higher expression as compared to other subtypes (see Fig. 3A in study I). These data indicated that ETV6-RUNX1 fusion protein might regulate *PTP4A3* either directly or indirectly. We looked for the possible direct interactions of ETV6-RUNX1 fusion protein in the putative regulatory regions upstream of *PTP4A3* with targeted ChIP-qPCR and ETV6 antibodies from NALM-6 cell lines with induced *ETV6-RUNX1*. All of these *PTP4A3* regulatory regions failed to detect any binding of ETV6-RUNX1 fusion into the possible binding sites (see Fig. 3A-B in study I).

5.1.3 Functional role of the PTP4A3 in ETV6-RUNX1 leukemia

It has been previously shown that depending on the expression level, *PTP4A3* may affect cell cycle regulation either positively or negatively (Beekman et al., 2011). To investigate the role of *PTP4A3* further in ALL, we knocked down expression of *PTP4A3* by transient siRNA transfection. Two different cell lines, REH and 697, were used for knockdown studies and the level of the knockdown was confirmed by RT-qPCR. After 48 h of transfection, RT-qPCR showed 42 % decrease in the expression of *PTP4A3* in REH cells and 63 % in 697 cells. Confirmation of the knockdown at protein level was not possible due to the absence of antibodies that detect PTP4A3 in WB experiments. After the knockdown, the cells were plated and cell viability was measured with Alamar Blue -assay for every 24 h up to 96 h with two biological replicates. These Alamar Blue -assays did not show any changes in cell viability. Similarly cell cycle or apoptosis did not reveal any changes after the knockdown. As our *PTP4A3* knockdown did not have any significant effect on cell viability, we sought for other possibilities to disrupt or block the PTP4A3 function in leukemia cells. A rhodamine derivative can be used to inhibit the enzymatic

activity of PTP4A3 with an IC50 value of 0.9 μM (Min et al., 2013). The potency of this PRL-3 inhibitor I to suppress proliferation of ETV6-RUNX1 positive leukemia cells was tested in an assay with increasing drug concentration and the results were compared between the ETV6-RUNX1 positive and negative leukemia cell lines. We noticed reduction in cell viability after the concentrations exceeded 10 μM (see Fig. 4 in study I). A mild resistance towards this inhibitor was observed in ETV6-RUNX1 positive cells, possibly due to higher expression of PTP4A3. In addition to this, we also tested a combination of PRL3-inhibitor I with a mixture of different leukemia drugs (vincristine, doxorubicin and prednisone). These combination experiments were done with variations of different drug and inhibitor concentrations in NALM-6 and REH cell lines, but no additive effect was seen in cell viabilities in the studied time points.

5.2 SOX11 transcription factor in BCP-ALL

5.2.1 *SOX11* overexpression in acute lymphoblastic leukemia

We sought to identify novel transcriptional regulators of BCP-ALL, and to this end we utilized a large microarray gene expression data set by Heinäniemi et al. (2016). We identified *SOX11* as an aberrantly expressed transcription factor gene with subtype-specificity. When we compared all leukemias and lymphomas, *SOX11* expression was highest in MCLs but increased expression was also observed in BCP-ALL. We saw a subtype-specific overexpression in ETV6-RUNX1 and TCF3-PBX1 subtypes when compared to other acute leukemia subtypes (Fig. 10), approximately four to five-fold higher expression was noted in ETV6-RUNX1 and TCF3-PBX1 subtypes (see Fig. 1A in study II). These results were further confirmed by two additional gene expression data sets, namely GSE47051 (Nordlund et al., 2013) and St. Jude data (PeCan) (see Fig. 1A in study II). Similar subtype specific expression pattern of *SOX11* was also seen in cell lines and patient samples (see Fig. 1B in study II).

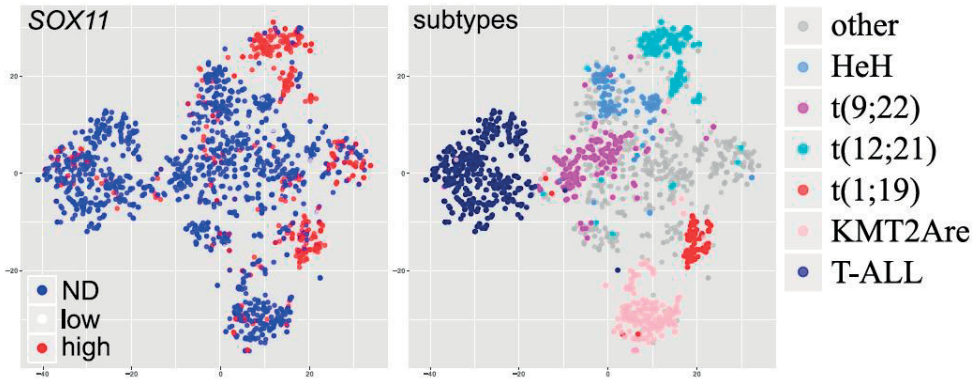


Figure 10. Expression levels of *SOX11* in acute leukemias shown in a t-SNE-based two-dimensional map. Expression level is indicated with colors blue (ND = not detected), white (low expression), red (high expression). Leukemia subtypes are indicated in the right figure with colors. Abbreviations: HeH = hyperdiploid, KMT2Are = KMT2A-rearrangement, t(1;19) = TCF3-PBX1, t(9;22) = BCR-ABL1, t(12;21) = ETV6-RUNX1, T-ALL = T cell acute lymphoblastic leukemia.

5.2.2 SOX11 expression affects survival in acute lymphoblastic leukemia

Having previously shown that *SOX11* is strongly expressed in subgroups of pediatric ALLs at the RNA level, we next sought to analyze its expression at protein level. To this end, a cohort of 119 primary FFPE bone marrow biopsy samples of childhood B cell acute lypbhoblastic leukemia (B-ALL) cases were analyzed with immunohistochemical (IHC) staining by SOX11 antibody (see Table 1 and 2 in study II). 79 out of 119 cases were deemed as negative with less than 20 % of positive cells. 35 cases were SOX11 positive (nuclear positivity 20 – 50 %) while 5 cases were deemed strongly positive (positivity over 50 %) (Fig. 11 and Table 2 in study II). From the 40 positive cases, 29 demonstrated either ETV6-RUNX1 or TCF3-PBX1 subtype with statistically significant association with ETV6-RUNX1 subgroup (p-value <0.001). *KMT2A*-rearrangement and *BCR-ABL1* translocation cases did not express SOX11 protein and majority of hyperdiploid cases were negative. In addition to IHC staining, we studied the suitability of flow cytometry to discriminate SOX11 positive cases based on the cytoplasmic expression. SOX11 antibody staining was able to readily separate high expressing (REH, RCH-ACV) from low expressing (NALM-6, KOPN-8) cell lines (see Fig. 5B in study II).

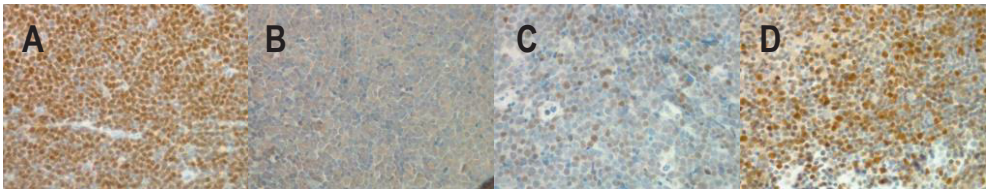


Figure 11. Immunohistological stainings of SOX11 in FFPE samples. A) Positive control, B) Negative control, C Positive BCP-ALL, D) Strong positive BCP-ALL.

Next step was to evaluate the clinical significance of SOX11 expression in B-ALL. In our patient cohort, OS was better in the SOX11 positive compared to SOX11 negative group (p=0.039) (see Fig. 2B in study II) and no deaths were reported amongst SOX11 positive group. Similar trends were also seen in EFS and relapse-free survival (RFS) but without statistical significance (see Fig. 2B in study II). Favourable trend was seen in univariate analysis of SOX11 positivity in EFS among the ETV6-RUNX1 subtype (HR = 0.27, 95 % CI 0.04, 1.60, p = 0.149) and other subtypes (HR = 0.28, 95 % CI 0.04, 2.07, p = 0.212). Good early therapy response was also associated with SOX11 positivity at the end of induction therapy (OR = 0.49, 95 % CI 0.19, 1.19, p = 0.123). Positive IHC for SOX11 remained statistically significant in multivariate analysis of EFS (see Table 3 in study II) indicating its independent value as a prognostic marker.

We next sought to validate our findings in an external data set and, therefore analyzed *SOX11* expression from transcriptomic data of a patient cohort comprising 116 BCP-ALL cases (Marincevic-Zuniga et al., 2017). This data set supported our protein expression data, as high expression of *SOX11* mRNA was associated with favourable EFS, even though no statistical significance was reached (see Fig. 3A in study II).

5.2.3 Hypomethylation of *SOX11* locus in BCP-ALL

Next, we investigated the biology behind the increased expression of *SOX11* in leukemia. When patient samples with available genomic data were analysed, no aberrant enhancer activity in primary transcription (GRO-seq) or somatic mutations in whole-genome sequencing (WGS) were observed in the nearby regions of *SOX11,* suggesting that perhaps epigenetic mechanisms could be involved. As hypomethylation partially drives increased expression of *SOX11* in MCL (Vegliante et al. 2011), we utilized available genome-wide CpG methylation data from the BCP-

ALL patient cohort (Marincevic-Zuniga et al., 2017). A total of 23 CpG sites were located within the *SOX11* locus and patients with high *SOX11* mRNA expression demonstrated a strong DNA hypomethylation at *SOX11* gene locus (see Fig. 3B and C). *SOX11* loci with hypomethylated CpG sites were associated with better EFS although no statistical significance was reached (see Fig. 3A).

5.2.4 Knockdown of SOX11 in BCP-ALL cell lines

In order to study the functional role of SOX11, we decided to knock down the expression of *SOX11* in cell lines with high expression of *SOX11* and for this, we chose siRNA oligos targeting *SOX11*. After transient transfection of siRNA oligos, we observed marked down-regulation of *SOX11* expression so that 20-40 % of expression were left, depending on the cell line, as compared to controls (see Fig. 4A-B in study II). Down-regulation was further confirmed by WB. Then we assessed the impact of knockdown of SOX11 into cell viability by using Alamar Blue -assay, but no marked effect on cell viability was observed in any of the studied cell lines (see Fig. 4B in study II). In addition to this, we investigated if knockdown had any effect on steroid (dexamethasone and prednisolone) sensitivity but no effect was seen.

SOX11 knockdown effect in gene expression was further studied with RNA-seq in REH, 697, RCH-ACV cell lines. Gene ontology (GO) annotations demonstrated altered GO-terms related to cell migration, adhesion and differentiation. Gene set enrichment analysis on the other hand implied altered expression MYC and EF2 target genes (see Fig. 4C in study II and Appendix 1). As summarized in the heatmap, after knockdown of *SOX11*, all three cell lines differentially expressed 18 genes, whereas 15 were concordantly regulated (same direction) (see Fig. 4E in study II and Appendix 2). Gene alterations were seen in genes related to leukemia cell motility, adhesion, differentiation and drug response. *SOX11* silencing led to a down-regulation of *WW Domain Binding Protein 1 Like* (*WBP1L*), a direct target of ETV6 in ALL (Neveu et al., 2016). Increased *WBP1L* expression (2.8-fold) in ETV6-RUNX1 subtype (Yeoh et al., 2002) coincides with *SOX11* overexpression and may suggest co-regulation in same subtype. *SOX11* knockdown silenced a secreted growth factor *Midkine* (*MDK*) which promotes cell growth and migration and is associated with adverse prognosis in ALL (Jia et al., 2016). Another gene regulated by SOX11 and involved with cell regulation and migration was *WD Repeat Domain 1* (*WDR1*), which is also involved in megakaryocyte maturation (Kuhns et al., 2016; Standing et al.,

2017). *MLLT11*, a gene associated with hematopoiesis (Tse et al., 1995) found to be downregulated by SOX11. *MLLT11* is involved in lymphoid regulation and is a known partner in rare leukemia translocations (Tse et al., 1995; Ney Garcia et al., 2017). Downregulated *Coiled-coil domain containing 28A* (*CCDC28A*) is similarly a known partner of NUP98 in AML. SOX11 also regulated genes associated with drug response. Association of *MKD* with adverse prognosis in ALL is possibly via increased drug resistance (Hu et al., 2010) and *Lipopolysaccharide Induced TNF Factor* (*LITAF*) was recently suggested to sensitize leukemia cells to chemotherapeutic drugs (Liu et al., 2016). *CANT1* was down regulated by SOX11 and is involved in pyrimidine metabolism and could be related with cytarabine metabolism and therapy response in ALL (Fridley et al., 2011). Only few of the genes were overlapping with previous studies with pro-B (Lord et al., 2018) and MCL cells (Vegliante et al., 2013; Conrotto et al., 2011; Kuo et al., 2015; Wang et al., 2010; Kuci et al., 2016). No marked effect was seen in *Id1* and *Tal1* as reported in pro-B cells (Lord et al., 2018) or *PAX5* in MCL (Vegliante et al., 2013). *MDK*, which is involved in cell migration and growth, was downregulated by *SOX11* knockdown in our study with similar manner of *SOX11* knockdown in MCL cell line Z138 (Conrotto et al. 2011).

5.3 Targeted therapies for T-ALL

5.3.1 *In silico* screening for potential targets and matching drugs

T-ALL has a less favorable prognosis than BCP-ALL, especially after relapse. We therefore searched for novel targeted therapies for T-ALL by utilizing the gene expression profiles of 4430 leukemia samples that were retrieved from the Gene Expression Omnibus (GEO) in an *in silico* drug target screening (Heinäniemi et al. 2016). In the target screening, expression profiles were compared to known drug targets in the Drug signature database (DsigDB) containing Food and Drug Administration approved and novel investigational drugs. This approach matched a target, LCK kinase, with a drug, dasatinib tyrosine kinase inhibitor. LCK showed strong expression in a portion of T-ALL cases as compared to other leukemias and normal T cells (see Fig. 1A in study III). In a kinase activity assay, dasatinib treatment decreased the activity of LCK kinase to 1 % at 100 nM concentration when compared to controls (Yoo et al., 2015). Elevated expression of few other potential targets of dasatinib were similarly noticed in T-ALL patients (see Fig. 1B and Suppl.

Fig. S1 in study III). These data suggested that dasatinib could serve as a potential candidate targeted therapy for T-ALL patients.

5.3.2 Dasatinib decreases cell viability of Jurkat cells and targets LCK kinase

We next evaluated the potency of dasatinib experimentally in a panel of seven T-ALL cell lines. In an assay with 10-fold concentration changes of dasatinib (1-1000 nM), cell viability was measured after 72 h treatment. From the studied cell lines, Jurkat cells showed the most significant response with 31 % decrease in cell viability at 10 nM concentration (n=3, P = 0.0039; see Fig. 1C and Suppl. Fig. S2 in study III). Three other cell lines (Molt-16, Molt-4 and PEER) responded for dasatinib with higher concentrations or milder effect on cell viability (see Suppl. Fig. S2 in study III). In a follow up study *LCK* knockdown in a dasatinib sensitive Jurkat cell line decreased cell proliferation in a significant manner (14 % decrease, P = 0.0289, n = 7, see Fig. 1E in study III). We next utilized a genetically engineered variant of Jurkat cells (J.CaM1.6), where a deletion of exon 7 of LCK kinase renders it inactive. Accordingly, these cells were insensitive to dasatinib treatment (see Fig. 1F in study III). *LCK* knockdown in a relatively dasatinib-insensitive P12-Ichikawa cell line did not cause any significant change in cell proliferation (see Suppl. Fig. S4E in study III).

As dasatinib has several targets and inhibits multiple key regulators of cellular proliferation, we next sought to study other possible target genes and analyze their expression in T-ALL cell lines. Totally, 21 possible targets (9 targets based on the *in silico* screening and 12 other well-known targets from the literature) were identified and selected for the screening with RT-qPCR and WB. Expression levels of the other possible targets in T-ALL cell lines were lower or nonexistent when compared to *LCK* expression (see Fig. 1D and Suppl. Fig. S3 in study III). Gene knockdown of other possible targets, FYN, ABL1, MAP2K5 and MAP4K5, with lower expression levels did not cause any significant effect on Jurkat cells proliferation (see Suppl. Fig. S4A-D in study III).

5.3.3 One third of T-ALL cases are sensitive to dasatinib

We next utilized an *ex vivo* drug testing to see if T-ALL patient samples respond to dasatinib treatment. 6 cases out of 22 (27 %) had a significant response to dasatinib

according to DSS (using a cutoff value of 10, see Fig. 2A in study III) (Yadav et al., 2014). Half-maximal growth inhibition concentrations (IC50) ranged between 1.3 and 16 nM, while the IC50 values for the control samples were >1000 nM (see Suppl. Fig. S5 in study III). In addition to this, we also noted a negative correlation in DSS between dasatinib and glucocorticoid (see Suppl. Fig. S6 in study III). None of our patients had the *NUP214-ABL1* fusion (Quintás-Cardama et al., 2008; Deenik et al., 2009; Crombet et al., 2012) (see Suppl. Fig. S7 in study III). From the dasatinib targets under examination, *LCK* expression was prominent in four out of five (RNA unavailable for one of the six samples) dasatinib-responsive samples whereas other targets were expressed either at low (*LYN*, *ABL1*) or medium level (*FYN*, *MAP2K5* and *MAP4K5*) (see Suppl. Fig. S8A in study III). We also noted a statistically significant correlation between *LCK* expression and the TAL1 subgroup of T-ALL (see Fig. 2C-D and Suppl. Fig. S10 in study III).

5.4 The role of SAP30L in zebrafish hematopoiesis

5.4.1 Expression of *SAP30L* in zebrafish

Sin3A corepressor complex is an important regulator of gene expression (Silverstein and Ekwall, 2005). SAP30L is part of the Sin3A co-repressor complex but its function *in vivo* has remained unknown (Viiri et al., 2006; Korkeamäki et al., 2008). For that matter, we studied *SAP30L* mRNA expression in zebrafish with RT-qPCR from whole embryos and used WISH to determine the localization of *SAP30L* expression. During the early developmental stages of zebrafish, expression of *SAP30L* was observed in tissues of embryos and larvae at several time points (see Fig. 1A-B in study IV). During the embryogenesis, the *SAP30L* mRNA was widely expressed and after 24 hpf the expression was strongest in the brain (see Fig. 1A in study IV). At 36 hpf, expression was visible in pectoral fin buds and, at 72 hpf, a strong expression was detectable in pectoral fins (black arrows in Fig. 1A in study IV). The heart area stained positively in pericardial cavity at 24 hpf and later also the heart tissue itself. After 36 hpf, the staining of *SAP30L* mRNA could be seen in the common cardinal vein (marked with white arrows in Fig. 1A in study IV). In addition, moderate mRNA expression is seen in the trunk, including ICM, and weak staining in blood islands of some embryos. Adult zebrafish shows strong

expression of *SAP30L* in the brain and moderate levels in muscle, eye, gonad and heart (see Fig. 1C in study IV).

5.4.2 SAP30L knockdown affects heart morphology and hematopoiesis

After studying the *SAP30L* expression and localization during the embryogenesis, we investigated the role of SAP30L during the early development of zebrafish. Effects of the SAP30L depletion was studied using morpholino oligonucleotides (MO) -based silencing. MOs block the translation of mRNA and can be used to silence the expression of gene of interest. Schematic illustration of the microinjection is shown in Figure 12. *SAP30L* knockdown experiments were carried out with two independent MOs (SAP30L-MO1 and SAP30L-MO2) targeting the translation initiation site and their efficacy was demonstrated *in vivo* with reducing SAP30L protein levels in MO-treated zebrafish embryos (see Fig. 2 in study IV). Embryos were injected with the morpholinos at the 1- to 4-cell stage and the phenotypic development was observed up to 5 to 6 dpf. During the observation period, the morphants displayed prominent and progressive pericardial edema and almost 90 % of them had deformed cardiac morphology (see Fig. 3A in study IV). Heart malformation was visible in a single plain as a string-like heart instead of a healthy three-dimensional complex heart (see Fig. 3B in study IV). Compared to the control larvae, the morphants already demonstrated at 3 dpf an impaired cardiac function with marked reduction in heart rates and decreased performance of the atrium and the ventricle (see Fig. 3C-D in study IV). Both morpholinos gave similar results for the phenotypes of injected zebrafish (morphants). Gene expression analysis by microarray of the SAP30L morphants revealed downregulation of *nkx2.5* gene at 12 hpf (see Fig. 4 in study IV). *Nkx2.5* is known to regulate cardiac development (Chen and Fishman, 1996; Targoff et al., 2008). *Nkx2.5* down-regulation was confirmed with RT-qPCR (see Fig. 5A in study IV). In a rescue experiment, SAP30L-MO1 was co-injected with *nkx2.5* mRNA which led to partial rescue of the deformed cardiac morphology (see Fig. 5B in study IV).

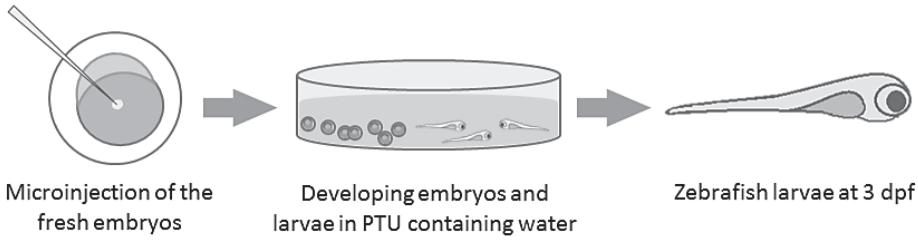


Figure 12. Schematic illustration of zebrafish embryos microinjection. Freshly fertilized embryos were injected with morpholino solution directly in to the embryos yolk sack during the 1-4 cell stage. After the injection the embryos were transferred to a petri dish for the development of the embryo to larvae. PTU=1-phenyl 2-thiourea, dpf=days post fertilization.

When observing the SAP30L morphant fish, we noticed that red blood cells appeared paler compared to control larvaes. We determined the hemoglobin levels of the morphant and control fish by o-dianisidine staining. Staining was performed in 3 dpf larvae and the intensity of staining was analyzed by visual observation. SAP30L morphants demonstrated markedly reduced staining indicating that hemoglobin levels were decreased in the red blood cells (see Fig. 3E in study IV). In microarray analysis of gene expression, downregulation of genes involved in erythropoiesis and hemoglobin synthesis were observed. For example, downregulation of *ppox* and *alas2* genes was observed at 24 hpf (Dailey and Dailey, 1996; Brownlie et al., 1998) (see Fig. 4 in study IV and Supplementary Figure S5 from Teittinen et al. 2012). These results were further confirmed with RT-qPCR where morphants showed diminished expression of *ppox* (see Fig. 5C in study IV). These results suggest a role for SAP30L in TGF-β signaling and point to a notable role in multiple transcription factor regulation (see Fig. 4 in study IV and Supplementary Figure S5 from Teittinen et al. 2012).

6 DISCUSSION

6.1 PTP4A family of phosphatases in acute lymphoblastic leukemia

The first aim of this thesis was to study the expression and functional role of PTP4A3 in ALL (study I). Normally, expression of PTP4A3 is restricted to non-hematological tissues such as the brain, skeletal muscle, heart and vasculature (Matter et al., 2001). This expression pattern suggests that PTP4A3 overexpression in hematological malignancies is driven by cancer-associated factors. Earlier findings by Zhou and colleagues (2012) supported this hypothesis when they demonstrated PTP4A3 as a novel downstream target of BCR-ABL1 signaling pathway. From other protein family members, high PTP4A1 and PTP4A2 expressions have been linked to AML, and PTP4A2 expression to AML cell proliferation (Arora et al., 2012; Kobayashi et al., 2017). Strongest expression of *PTP4A3* mRNA in our microarray data set was seen in BCR-ABL1 and ETV6-RUNX1 subtypes of BCP-ALL whereas *PTP4A1* and *PTP4A2* expression levels were not detectable. In T-ALL, we saw some patients with high expression of *PTP4A2* and *PTP4A3*. Kobayashi et al. (2014) reported similar results of *PTP4A2* and *PTP4A3* upregulation in T-ALL.

We focused on the ETV6-RUNX1 fusion and its role in the regulation of *PTP4A3* expression since the ETV6-RUNX1 subtype is the second most common subtype of childhood acute leukemia whereas BCR-ABL1 represents approximately 3 % of childhood ALL (Lilljebjörn et al., 2016). We showed that ETV6-RUNX1 fusion regulates expression of *PTP4A3*. By looking at the binding sites for ETV6-RUNX1 at the vicinity of *PTP4A3* gene, a number of potential chromatin binding sites were identified from the RUNX1 ChIP-seq data from SEM leukemia cells. This suggested for the possibility of direct DNA binding of ETV6-RUNX1 fusion protein. However, we failed to confirm the binding of ETV6-RUNX1 fusion into any of the tested sites (see Fig. 3A-B in study I) thus hinting that there is no direct regulation by ETV6-RUNX1. However, this might also be a consequence of several other factors. First, we used the NALM-6 cells, while the ChIP-seq data was generated in the SEM cells. Secondly, we only tested a few putative binding sites based on the known RUNX1 binding motifs, and this could lead to missing potential

binding sites. Thirdly, we only focused on the binding sites upstream of the transcription start site of *PTP4A3* even though the regulatory elements can reside even far away in either direction from the transcription starting site. An unbiased ChIP-seq experiment in the NALM-6 cells would solve this issue and will be needed if the function of PTP4A3 in ALL is pursued further in the future.

PTP4A family members are dual-specificity phosphatases meaning that they can inactivate targets by dephosphorylating both tyrosine and serine/threonine residues. PTP4A family members regulate several important signaling pathways such as Src and PTEN (Campbell and Zhang, 2014). We studied the role of PTP4A3 in leukemia pathogenesis by using two inhibitory approaches: knockdown of *PTP4A3* through siRNA oligos and with an *in vitro* treatment of cell by a specific inhibitor of PTP4A3 enzyme activity (PRL-3 inhibitor I). Experiments were performed in BCP-ALL cell lines since primary samples from children with ALL were scanty causing inability to perform multiple biological replicates with and without other leukemia drugs and in a time-series manner. Transient knockdown of *PTP4A3* did not have any effect on the cell viability and, similarly, the inhibitor had only modest effect on cell viability with high concentrations (μM). Similar trend with decreasing cell viability was noted in different cell lines. We also studied the effect of PRL3 inhibitor I together with conventional ALL induction drugs (prednisone, doxorubicin and vincristine), but no additive effect was evident in these experiments. These results imply that PTP4A3 does not play a major role in the pathogenesis of ETV6-RUNX1 or BCR-ABL1 subtypes of ALL. One could also argue that the inhibitor used in this study was less than optimal for targeting of PTP4A3, or that high expression of PTP4A3 and its enzymatic activity are not correlated. In the case of knockdown and cell viability experiments, it is possible that the *PTP4A3* knockdown was not efficient enough at protein level or the half-life of the siRNA was too short to make an impact on the cell viability. It is also noteworthy that cell viability assay is quite robust technique and could masquerade most subtle effects.

In summary, we demonstrated that *PTP4A3* is highly expressed in selected subtypes of pediatric ALL and regulated by the ETV6-RUNX1 fusion protein. However, we failed to elucidate the exact functional role of PTP4A3 in BCP-ALL. One caveat in our experiments is that the expression data was solely based on the level of mRNA expression. In future, the role of PTP4A3 should be studied at protein level, which takes into account also the enzymatic activity and not just the expression state. It has also been shown that the mRNA expression does not always correlate with the protein levels (Latonen et al., 2018). What effect does the change in enzymatic activity have on the leukemia cell proliferation and viability?

Intracellular cell signaling is a tightly regulated and dynamic system, which is often derailed in malignancies (Tesio et al., 2016). PTP4A3 is known to have an important role in multiple cellular functions that are related to cancer, like cell migration, adhesion and invasion. Interestingly, Hjort et al. (2017) demonstrated very recently that PTP4A3 inhibition/knockdown reduced cell migration towards an SDF-1α gradient in two BCP-ALL cell lines, and also cell adhesion towards fibronectin was reduced in REH cells. Hjort et al. (2017) inferred that PTP4A3 could represent a novel target in the treatment of BCP-ALL due to its promotion of cell resistance to cytarabine, an antimetabolite prodrug.

6.2 High expression of SOX11 in acute lymphoblastic leukemia

The second aim of this thesis was to study expression and function of SOX11 transcription factor in ALL and its use as a prognostic marker in childhood ALL (study II). First, we looked at the microarray expression data for the *SOX11* expression across all hematological malignancies. The highest expression of *SOX11* was detected in MCL patient which corroborated previous studies (Ek et al., 2008; Zeng et al., 2012; Mozos et al., 2009). Patients with BCP-ALL also showed tendency towards the elevated expression of *SOX11*, and when we looked more closely at the subtypes, we noticed higher *SOX11* expression in ETV6-RUNX1 and TCF3-PBX1 subtypes. In our patient cohort, SOX11 expression had statistically significant ($p<0.001$) association with ETV6-RUNX1 subtype. There was no statistical association with SOX11 expression and TCF3-PBX1 subtype, but this was most likely due the low number of TCF3-PBX1 cases, with three SOX11 positive out of four cases. These findings fall in line with previous studies. Dictor et al. (2009) found nuclear expression of SOX11 in majority (8 out of 9) of B cell ALLs/lymphomas but they did not specify the protein expression levels (no subtype-specific data were reported either). Mozos and colleagues (2009) identified strong nuclear SOX11 expression in five T cell lymphoblastic leukemias/lymphomas and one B cell lymphoblastic leukemia /lymphoma cases, and Vegliante et al. (2011) associated high *SOX11* mRNA expression to ETV6-RUNX1 and *TCF3*-rearranged cases in a small cohort study. Our results of SOX11 expression are thus not fully original but the breadth of the data is by far the most extensive.

We next investigated the impact of high SOX11 expression as prognostic markers in pediatric BCP-ALL. Bone marrow biopsies from primary diagnostic samples of ALL were evaluated for SOX11 protein expression by IHC and a correlation to

patient survival was evaluated. Interestingly, SOX11 positive patients had statistically significantly better OS as compared to SOX11 negative patients in Kaplan-Meier analysis (see Fig. 2A in study II). EFS and RFS also had similar trends on survival time but without statistical significance. As our patient cohort with 119 patients from which 40 were SOX11 positive was still relatively small, we sought to validate the findings in external data set. Importantly, the other data set where *SOX11* expression was based on transcriptomic expression profiles showed similar survival trends (Marincevic-Zuniga et al., 2017). Studies in MCL have shown that mRNA or nuclear protein expression of SOX11 can be used as a diagnostic marker in classical MCL morphology (Ek et al., 2008; Zeng et al., 2012; Mozos et al., 2009). Wang et al. (2008) associated nuclear positivity of SOX11 in MCL with better OS but this finding was contradicted by Fernandez et al. (2010). In our study, SOX11 positivity retained its prognostic significance in multivariate analysis, including subtype information, which suggests that (high) SOX11 expression might be a useful biomarker even in low-risk cases of BCP-ALL. It should be mentioned that our BCP-ALL case series spans over two decades and the patients were treated with three successive NOPHO ALL chemotherapy protocols (Toft et al., 2018; Schmiegelow et al., 2010). As the most recent NOPHO protocol has reported the best survival results (Toft et al., 2018), it is necessary to investigate the prognostic value of SOX11 also in the most recent protocols in future.

As bone marrow biopsies are no longer taken routinely in all treatment centres and IHC staining is time consuming, we tested flow cytometry based screening for SOX11 positivity with cell lines. As we were able to distinguish cell lines with flow cytometry based on SOX11 expression, it would be worthy use primary ALL samples to look further if SOX11 could be used as clinical biomarker and possibly include it in flow cytometry panels in future.

SOX11 is a known developmental and neural transcription factor but its functional mechanisms have remained partly unknown. For leukemia studies, it would be beneficial to collect ChIP-seq data from leukemia cell lines and primary patient samples to identify SOX11 targets in genome. This could give an insight for the matter why SOX11 positive ALLs seem to have a better prognosis. From the other point of view, it would be interesting to solve what causes the altered SOX11 expression in some leukemias and other hematological malignancies whereas adult stem cells and normal hematopoietic cells do not express SOX11. We showed that *SOX11* expression might be regulated by DNA hypomethylation of *SOX11* locus in a similar manner as Vegliante et al. (2011) reported in MCL and five ETV6-RUNX1 positive BCP-ALL cases. Vegliante et al. (2011) characterized the epigenetic

mechanism leading to the *SOX11* deregulation but they did not demonstrate which upstream transcription factors and histone modifying enzymes are behind this alteration.

In knockdown experiments, SOX11 silencing had no marked effect on cell viability in any of the studied cell lines implicating that SOX11 does not play a central role in the BCP-ALL cell viability. In addition to cell viability, we also looked for the transcriptional profiling of *SOX11* knockdown cells, to look for the target genes and alterations in pathways. Two noteworthy points for interpreting our *SOX11* knockdown studies are as follows: (i) changes in gene expression profiling were mostly from mild to moderate, (ii) other SoxC family members, SOX4 and SOX12, might mitigate the effects of *SOX11* knockdown.

Previous studies with MCL cell models have given contradictory results regarding to cell proliferation. Constitutive *SOX11* knockdown by Conrotto et al. (2011) and Kuci et al. (2016) led to an increased cell growth rate in MCL cell model. These studies are in concordance with transient siRNA-mediated knockdown studies of *SOX11* by Gustavsson et al. (2010) and Kuci et al. (2016). In contrary, Wang et al. (2010) reported that transient siRNA-mediated knockdown of *SOX11* did not have any significant effect on MCL cell line proliferation. *SOX11* overexpression repressed cell proliferation in MCL (Kuo et al., 2015) and pro-B cells (Lord et al., 2018). Balsas et al. (2017) found in their MCL xenograft transplantation studies that SOX11 positive cells had higher cell migration, invasion and proliferation over SOX11 knockdown cells. In future, animal models with cell or tissue specific knockdown and overexpression of SOX11 in BCP-ALL should be created to solve the issue.

6.3 Dasatinib as a targeted therapy for T-ALL

Approximately 15 % of childhood acute leukemias are of T cell origin. Although the prognosis of T-ALL has improved with the modern treatment protocols, it is still lagging behind the BCP-ALL, and is dismal at relapse. We therefore sought to identify novel targeted therapies for T-ALL (study III). Previously, it was reported that dasatinib had clinical utility on NUP214-ABL1 cases, which only represent 4 to 10 % of T-ALL patients (Quintás-Cardama et al., 2008; Burmeister et al., 2006). We and others (Frismantas et al. 2017) showed that approximately 30-40 % of T-ALL cases are sensitive to dasatinib, suggesting that other targets beyond ABL1 exist. Recently, Serafin et al. (2017) published an article where they had functionally

validated LCK as a therapeutic target to reverse glucocorticoid resistance in T-ALL. Co-treatments with dasatinib and dexamethasone and *LCK* knockdown with dexamethasone treatment reduced cell viability and increased cell death (Serafin et al., 2017). Our results are thus corroborated by several other recent publications.

We identified nine possible dasatinib targets based on the *in silico* screening of DSigDB, and expanded this list with 12 more targets based on the literature. The expression of these target genes was measured in T-ALL cell lines and thereby we were able to narrow down to five genes that were further tested one-by-one in a siRNA-mediated knockdown. Selected target genes were *LCK*, *FYN*, *ABL1*, *MAP2K5* and *MAP4K5*, and from these candidates, *LCK* knockdown caused the most significant decrease in cell viability (see Fig. 1E and Suppl. Fig. S4A-E in study III). Second highest expression was seen in *FYN*, and the knockdown of *FYN* had slight reduction in cell viability. Importantly, Jurkat cell line with genetic mutation of LCK kinase did not respond to dasatinib, indicating the central role played by this kinase.

We confirmed the findings of Frismantas et al. (2017) and were the first to imply LCK as the major target of dasatinib in T-ALL. Frismantas et al. (2017), in turn, associated dasatinib sensitivity to SRC kinase expression and the phosphorylation status by using a SRC inhibitor KX2-391. However, based on the DsigDB, both SRC and LCK are targets of KX2-391. We did not find any marked difference in *SRC* expression levels between normal T cell and T-ALL samples, and the overall expression level of *SRC* mRNA was low compared to *LCK*. Based on these results, we are more confident that LCK, and not SRC, plays a central role in response to dasatinib treatment. We also noted that dasatinib sensitivity was higher on T-ALL cell lines with higher *LCK* expression. However, our studies do not fully exclude other possible targets for dasatinib, which are part of the receptor signaling pathways related to cell proliferation and survival. For examples LCK and FYN are both part of the TCR signaling cascade and also have partially overlapping functions (Salmond et al., 2009).

Our results also revealed that patients belonging to the TAL1 subgroup of T-ALL are more likely to respond to dasatinib even though their response was not exclusive. We saw a statistically significant correlation between *LCK* expression and TAL1 subgroup even though TAL1 itself did not seem to directly regulate expression of *LCK* (data were reproduced from Sanda et al., 2012). Dasatinib is known to suppress healthy T cells, (Blake et al., 2008; Lee et al., 2010; Schade et al., 2008) but in our study, the patient samples had significantly higher *LCK* expression as compared to normal T cells and higher sensitivity to dasatinib was seen in T-ALL

cells when compared to bone marrow cells. It is noteworthy that our patient cohort was small and further studies are needed. Dasatinib is currently approved for Ph+ CML treatment in imatinib resistant cases and in the second-line treatment of Ph+ ALL. This finding by us and others suggest that dasatinib could be applied to a subset of T-ALL, although the exact predictive biomarker is still not known. There is a urgent need to explore the biomarker findings and deepen the knowledge of the mechanistic functions of dasatinib.

6.4 SAP30L participates in development of zebrafish cardiac system and erythropoiesis

SAP30 and SAP30L are part of the Sin3A-corepressor complex and they participate in regulation of gene expression through this complex (Laherty et al., 1998; Zhang et al., 1998; Viiri et al., 2006; 2009b). In zebrafish, only one member of the SAP30 protein family, SAP30L, has been identified (Viiri et al., 2009a). Therefore, the function of SAP30 protein family in hematopoiesis can be studied by modulating the expression of a single gene in zebrafish (study IV). The fact that zebrafish genome has only one copy of *SAP30/SAP30L* was an advantage as it enabled us to study the functional role of the SAP30L without potentially overlapping functions. We showed that *SAP30L* mRNA is widely expressed during early development of zebrafish with strongest expression in the brain, but was also visible in the cardinal vein, the heart area and in the blood islands of some of the embryos.

We also carried out a series of knock down experiments with two different MOs (MO1 and MO2). These MOs targeted the translation initiation site and blocked the translation of the *SAP30L* mRNA leading to a significantly reduced SAP30L protein expression. One advantage of using zebrafish as a disease model is its transparency so that the embryonal development can be easily followed by microscope and naked eye. The most marked effect in the morphants (MO1 or MO2 injected embryos and larvae) was the prominent pericardial edema and cardiac deformation and reduced cardiac function. Both morpholinos caused similar defects but MO2-morphants had a slightly delayed onset accompanied with SAP30L protein levels that were less decreased. This could be explained by the differences between the morpholinos. MO1 targets the downstream sequence of the start codon and MO2 the upstream region of the *SAP30L* start codon and therefore it seems that the upstream targeting is not as effective. Interestingly, Starkovich et al. (2016) recently suggested that SAP30L may contribute to the congenital heart defect in humans based on the

chromosome 5q33 deletions where *SAP30L* is located, even though they could not exclude other genes deleted from the same chromosome. In another study, a long non-coding RNA, *ENSMUST00000117266* was involved in regulation of cardiomyocyte proliferative activity and the authors suggested that this long non-coding RNA was likely to function via SAP30L, which is its neighboring coding gene (Sun et al. 2017).

The specificity of the observed phenotype can be further confirmed by performing mRNA rescue experiments where mRNA of gene of interest is co-injected with the morpholino (Eisen and Smith, 2008). The microarray of the SAP30L morphants revealed down-regulation of *nkx2.5* gene, which is known to regulate a cardiac development and therefore we used it in an mRNA rescue experiment (Chen and Fishman, 1996; Targoff et al., 2008). Co-injection of *nkx2.5* mRNA with SAP30L-MO1 rescued partially the cardiac phenotype, which suggests that SAP30L plays a role in the regulation of *nkx2.5* expression. Previous studies in other model organisms have also stated that nxk2.5 might have a role in multiple processes of cardiac development (Targoff et al., 2008). Because SAP30L is part of a repressor complex and only binds DNA nonspecifically it is likely that the effect on *nkx2.5* expression is indirect. Failure of the complete rescue may be due to multiple different reasons like short half-life of the injected mRNA or its inability to reach the right tissue. The morpholino injection only gives a transient knockdown during early development and these results would be interesting to verify and extend in a stable knockdown/knockout model. Today, this is often feasible using CRISPR-Cas9-based mutagenesis especially for those genes that are highly expressed during early developments (Uusi-Mäkelä et al., 2018).

Other interesting effect of the MO mediated knockdown was seen in the red blood cells as some of the morphants demonstrated paler red blood cells compared to the control embryos. Based on this observation we performed *o*-dianisidine staining for the embryo at 3 dpf to assess the hemoglobin levels of the morphants. Both morphant types demonstrated weaker staining intensities, which indicates reduced hemoglobin levels compared to the controls. MO1-morphants had more severely affected stainings over MO2-morphants. The reduced hemoglobin staining is in concordance with the WISH results showing *SAP30L* expression in the cardinal vein and in ICM, which is the primary site for erythroid lineage development (Davidson and Zon, 2004). Our mRNA rescue experiment with SAP30L-MO2 and *SAP30L* mRNA co-injection had no restoring effect on the reduced hemoglobin staining which might be because of the short mRNA half-life or other reasons discussed in the previous paragraph. Based on our microarray data, SAP30L

morphants demonstrated a downregulation of *ALAS2* and *PPOX* genes which both take part in heme biosynthesis. *ALAS2* gene is required for the first step in heme biosynthesis and deficiency leads to an iron accumulation and maturation arrest in primitive erythropoiesis but does not interfere definitive erythropoiesis (Harigae et al., 2003). Mutations in *ALAS2* can cause X-linked sideroblastic anemia and X-linked dominant protoporphyria in humans (Brownlie et al., 1998; Whatley et al., 2008). PPOX is a protoporphyrinogen oxidase and it catalyzes the second-to-last step of the heme biosynthesis and mutations in *PPOX* can cause variegate porphyria (Sassa and Kappas, 2000; Dooley et al., 2008).

Our results showed that *SAP30L* knockdown reduced the level of hemoglobin in the red blood cells in zebrafish and altered the expression of genes involved in the heme biosynthesis. One plausible explanation might be that SAP30L silences inhibitors in heme biosynthesis, but we cannot exclude the possibility that the hemoglobin deficiency is caused by another mechanism causing impaired erythropoiesis. An interesting experiment would be an mRNA rescue with either *ALAS2* or *PPOX* mRNA to see if either of these could restore the hemoglobin staining levels from SAP30L morphants. Dooley et al. (2008) were able to partially restore the anemia with human *PPOX* mRNA in *montalcino* mutants, which are *PPOX* deficient as well as hemoglobin deficient and anemic. Notable point from our microarray data from erythropoietic view is that it was collected during the primitive erythropoiesis at the point when erythrocytes start circulating (24 hpf) whereas definitive hematopoiesis starts around 26 hpf (Gore et al., 2018). None of the embryonic specific hemoglobin genes (*hbae1.1*, *hbae3*, *hbae5*, *hbbe1.1*, *hbbe1.2*, *hbbe1.3*, *hbbe2*, *hbbe3*) showed changes in our microarray data (Gore et al., 2018). It would also be interesting to look a similar data from later developmental stage (3-6 dpf) to see possible differences between primitive and definitive erythrocytes.

7 CONCLUSIONS

This study was carried out in order to shed light on the function of a variety of differentially expressed genes, ranging from transcription factors to kinases and phosphatases, in acute lymphoblastic leukemia and hematopoiesis. The major conclusions are as follows.

1) *PTP4A3* has a subtype specific overexpression in BCP-ALL and the *PTP4A3* expression is elevated alongside with the *ETV6-RUNX1* overexpression. The effect of transient knockdown of *PTP4A3* on cell proliferation was not different between the ETV6-RUNX1 positive and negative cell lines. However, inhibition of PTP4A3 enzymatic activity showed modest reduction on proliferation in an ETV6-RUNX1 specific manner.

2) We confirmed high expression of SOX11 in the ETV6-RUNX1 and TCF3-PBX1 subtypes of BCP-ALL. High expression of SOX11 was associated with DNA hypomethlyation in *SOX11* loci and favourable prognosis in pediatric BCP-ALL.

3) We identified that dasatinib is a potential novel targeted therapy for T-ALL, and that LCK kinase is the prime target.

4) SAP30L plays a significant role in cardiogenesis and hematopoiesis (erythropoiesis) of zebrafish embryos.

8 REFERENCES

Aplan PD, Lombardi DP, Reaman GH, Sather HN, Hammond GD and Kirsch IR. (1992). Involvement of the putative hematopoietic transcription factor SCL in T-cell acute lymphoblastic leukemia. Blood. 79(5):1327-1333.

Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, Bloomfield CD, Cazzola M and Vardiman JW. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 127(20):2391-2405.

Arora D, Köthe S, van den Eijnden M, Hooft van Huijsduijnen R, Heidel F, Fischer T, Scholl S, Tölle B, Böhmer SA, Lennartsson J, Isken F, Müller-Tidow C and Böhmer FD. (2012). Expression of protein-tyrosine phosphatases in Acute Myeloid Leukemia cells: FLT3 ITD sustains high levels of DUSP6 expression. Cell Commun Signal. 10(1):19.

Bain G, Engel I, Robanus Maandag EC, te Riele HP, Voland JR, Sharp LL, Chun J, Huey B, Pinkel D and Murre C. (1997). E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas. Mol Cell Biol. 17(8):4782-4791.

Balsas P, Palomero J, Eguileor Á, Rodríguez ML, Vegliante MC, Planas-Rigol E, Sureda-Gómez M, Cid MC, Campo E and Amador V. (2017). SOX11 promotes tumor protective microenvironment interactions through CXCR4 and FAK regulation in mantle cell lymphoma. Blood. 130(4):501-513.

Barber KE, Harrison CJ, Broadfield ZJ, Stewart AR, Wright SL, Martineau M, Strefford JC and Moorman AV. (2007). Molecular cytogenetic characterization of TCF3 (E2A)/19p13.3 rearrangements in B-cell precursor acute lymphoblastic leukemia. Genes Chromosomes Cancer. 46(5):478-486.

Bateman CM, Alpar D, Ford AM, Colman SM, Wren D, Morgan M, Kearney L and Greaves M. (2015). Evolutionary trajectories of hyperdiploid ALL in monozygotic twins. Leukemia. 29(1):58-65.

Bateman CM, Colman SM, Chaplin T, Young BD, Eden TO, Bhakta M, Gratias EJ, van Wering ER, Cazzaniga G, Harrison CJ, Hain R, Ancliff P, Ford AM, Kearney L and Greaves M. (2010). Acquisition of genome-wide copy number alterations in monozygotic twins with acute lymphoblastic leukemia. Blood. 115(17):3553-3558.

Bayly R, Chuen L, Currie RA, Hyndman BD, Casselman R, Blobel GA and LeBrun DP. (2004). E2A-PBX1 interacts directly with the KIX domain of CBP/p300 in the induction of proliferation in primary hematopoietic cells. J Biol Chem. 279(53):55362-55371.

Bayly R, Murase T, Hyndman BD, Savage R, Nurmohamed S, Munro K, Casselman R, Smith SP and LeBrun DP. (2006). Critical role for a single leucine residue in leukemia induction by E2A-PBX1. Mol Cell Biol. 26(17):6442-6452.

Beekman R, Valkhof M, Erkeland SJ, Taskesen E, Rockova V, Peeters JK, Valk PJ, Löwenberg B and Touw IP. (2011). Retroviral integration mutagenesis in mice and

comparative analysis in human AML identify reduced PTP4A3 expression as a prognostic indicator. PLoS One. 6(10):e26537.
Belgaumi AF, Al-Shehri A, Ayas M, Al-Mahr M, Al-Seraihy A, Al-Ahmari A and El-Solh H. (2010). Clinical characteristics and treatment outcome of pediatric patients with chronic myeloid leukemia. Haematologica. 95(7):1211-1215.
Bernstein BE, Meissner A and Lander ES. (2007). The mammalian epigenome. Cell. 128(4):669-681.
Bertrand JY, Kim AD, Violette EP, Stachura DL, Cisson JL and Traver D. (2007). Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo. Development. 134(23):4147-4156.
Bessette DC, Qiu D and Pallen CJ. (2008). PRL PTPs: mediators and markers of cancer progression. Cancer Metastasis Rev. 27(2):231-252.
Bhojwani D, Pei D, Sandlund JT, Jeha S, Ribeiro RC, Rubnitz JE, Raimondi SC, Shurtleff S, Onciu M, Cheng C, Coustan-Smith E, Bowman WP, Howard SC, Metzger ML, Inaba H, Leung W, Evans WE, Campana D, Relling MV and Pui CH. (2012). ETV6-RUNX1-positive childhood acute lymphoblastic leukemia: improved outcome with contemporary therapy. Leukemia. 26(2):265-270.
Bhojwani D, Yang JJ and Pui CH. (2015). Biology of childhood acute lymphoblastic leukemia. Pediatr Clin North Am. 62(1):47-60.
Bianchi ME and Agresti A. (2005). HMG proteins: dynamic players in gene regulation and differentiation. Curr Opin Genet Dev. 15(5):496-506.
Biondi A, Gandemer V, De Lorenzo P, Cario G, Campbell M, Castor A, Pieters R, Baruchel A, Vora A, Leoni V, Stary J, Escherich G, Li CK, Cazzaniga G, Cavé H, Bradtke J, Conter V, Saha V, Schrappe M, Grazia Valsecchi M. (2018). Imatinib treatment of paediatric Philadelphia chromosome-positive acute lymphoblastic leukaemia (EsPhALL2010): a prospective, intergroup, open-label, single-arm clinical trial. Lancet Haematol. 5(12):e641-e652.
Black JC, Van Rechem C, Whetstine JR. (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. Mol Cell. 48(4):491-507.
Blake S, Hughes TP, Mayrhofer G and Lyons AB. (2008). The Src/ABL kinase inhibitor dasatinib (BMS-354825) inhibits function of normal human T-lymphocytes in vitro. Clin Immunol. 127(3):330-339.
Borghesi L, Aites J, Nelson S, Lefterov P, James P and Gerstein R. (2005). E47 is required for V(D)J recombinase activity in common lymphoid progenitors. J Exp Med. 202(12):1669-1677.
Brown Geoffrey and Ceredig Rhodri (editors). Cell determination during hematopoiesis. Nova Biomedical Books, New York. 2009.
Brown L, Cheng JT, Chen Q, Siciliano MJ, Crist W, Buchanan G and Baer R. (1990). Site-specific recombination of the tal-1 gene is a common occurrence in human T cell leukemia. EMBO J. 9(10):3343-3351.
Brownlie A, Donovan A, Pratt SJ, Paw BH, Oates AC, Brugnara C, Witkowska HE, Sassa S and Zon LI. (1998). Positional cloning of the zebrafish sauternes gene: A model for congenital sideroblastic anaemia. Nat Genet. 20(3):244-250.
Brownlie A, Donovan A, Pratt SJ, Paw BH, Oates AC, Brugnara C, Witkowska HE, Sassa S and Zon LI. (1998). Positional cloning of the zebrafish sauternes gene: a model for congenital sideroblastic anaemia. Nat Genet. 20(3):244-250.

Burmeister T, Gökbuget N, Reinhardt R, Rieder H, Hoelzer D and Schwartz S. (2006). NUP214-ABL1 in adult T-ALL: the GMALL study group experience. Blood. 108(10):3556-3559.

Burns CE, Deblasio T, Zhou Y, Zhang J, Zon L and Nimer SD. (2002). Isolation and characterization of runxa and runxb, zebrafish members of the runt family of transcriptional regulators. Exp Hematol. 30(12):1381-1389.

Burns CE, Traver D, Mayhall E, Shepard JL and Zon LI. (2005) Hematopoietic stem cell fate is established by the Notch-Runx pathway. Genes Dev. 19(19):2331-2342.

Bussmann J, Bakkers J and Schulte-Merker S. (2007). Early endocardial morphogenesis requires Scl/Tal1. PLoS Genet. 3(8):e140.

Cai X, Gaudet JJ, Mangan JK, Chen MJ, De Obaldia ME, Oo Z, Ernst P and Speck NA. (2011). Runx1 loss minimally impacts long-term hematopoietic stem cells. PLoS One. 6(12):e28430.

Campbell AM and Zhang ZY. (2014). Phosphatase of regenerating liver: a novel target for cancer therapy. Expert Opin Ther Targets. 18(5):555-569.

Cantor AB and Orkin SH. (2002). Transcriptional regulation of erythropoiesis: an affair involving multiple partners. Oncogene 21(21):3368-3376.

Cantor AB, Katz SG and Orkin SH. (2002). Distinct domains of the GATA-1 cofactor FOG-1 differentially influence erythroid versus megakaryocytic maturation. Mol. Cell. Biol. 22(12):4268-4279.

Carradice D and Lieschke GJ. (2008). Zebrafish in hematology: sushi or science? Blood. 111(7):3331-3342.

Carroll AJ, Crist WM, Parmley RT, Roper M, Cooper MD and Finley WH. (1984). Pre-B cell leukemia associated with chromosome translocation 1;19. Blood. 63(3):721-472.

Casagrande G, te Kronnie G and Basso G. (2006). The effects of siRNA-mediated inhibition of E2A-PBX1 on EB-1 and Wnt16b expression in the 697 pre-B leukemia cell line. Haematologica. 91(6):765-771.

Cauwelier B, Dastugue N, Cools J, Poppe B, Herens C, De Paepe A, Hagemeijer A and Speleman F. (2006). Molecular cytogenetic study of 126 unselected T-ALL cases reveals high incidence of TCRbeta locus rearrangements and putative new T-cell oncogenes. Leukemia. 20(7): 1238-1244.

Cazzaniga V, Bugarin C, Bardini M, Giordan M, te Kronnie G, Basso G, Biondi A, Fazio G and Cazzaniga G. (2015). LCK over-expression drives STAT5 oncogenic signaling in PAX5 translocated BCP-ALL patients. Oncotarget. 6(3):1569-1581.

Chagraoui H, Kassouf M, Banerjee S, Goardon N, Clark K, Atzberger A, Pearce AC, Skoda RC, Ferguson DJ, Watson SP, Vyas P and Porcher C. (2011). SCL-mediated regulation of the cell-cycle regulator p21 is critical for murine megakaryopoiesis. Blood. 118(3):723-735.

Chang CP, de Vivo I and Cleary ML. (1997). The Hox cooperativity motif of the chimeric oncoprotein E2a-Pbx1 is necessary and sufficient for oncogenesis. Mol Cell Biol. 17(1):81-88.

Chen AT and Zon LI. (2009) Zebrafish blood stem cells. J Cell Biochem. 108(1):35-42.

Chen JN and Fishman MC. (1996). Zebrafish tinman homolog demarcates the heart field and initiates myocardial differentiation. Development. 122(12):3809-3816.

Clark MR, Mandal M, Ochiai K and Singh H. (2014). Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. Nat Rev Immunol. 14(2):69-80.

Conrotto P, Andréasson U, Kuci V, Borrebaeck CA and Ek S. (2011). Knock-down of SOX11 induces autotaxin-dependent increase in proliferation in vitro and more aggressive tumors in vivo. Mol Oncol. 5(6):527-537.

Crombet O, Lastrapes K, Zieske A and Morales-Arias J. (2012). Complete morphologic and molecular remission after introduction of dasatinib in the treatment of a pediatric patient with t-cell acute lymphoblastic leukemia and ABL1 amplification. Pediatr Blood Cancer. 59(2):333-334.

Cumano A and Godin I. (2007). Ontogeny of the hematopoietic system. Annu Rev Immunol. 25:745-785.

Dailey TA and Dailey HA. (1996). Human protoporphyrinogen oxidase: expression, purification, and characterization of the cloned enzyme. Protein Sci. 5(1):98-105.

Dandekar S, Romanos-Sirakis E, Pais F, Bhatla T, Jones C, Bourgeois W, Hunger SP, Raetz EA, Hermiston ML, Dasgupta R, Morrison DJ amd Carroll WL. (2014). Wnt inhibition leads to improved chemosensitivity in paediatric acute lymphoblastic leukaemia. Br J Haematol. 167(1):87-99.

Davidson AJ and Zon LI. (2004). The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. Oncogene. 23(43):7233-7246.

De Braekeleer E, Douet-Guilbert N, Morel F, Le Bris MJ, Basinko A and De Braekeleer M. (2012). ETV6 fusion genes in hematological malignancies: a review. Leuk Res. 36(8):945-961.

De Braekeleer E, Douet-Guilbert N, Morel F, Le Bris MJ, Férec C and De Braekeleer M. (2011a). RUNX1 translocations and fusion genes in malignant hemopathies. Future Oncol. 7(1):77-91.

De Braekeleer E, Douet-Guilbert N, Rowe D, Bown N, Morel F, Berthou C, Férec C and De Braekeleer M. (2011b). ABL1 fusion genes in hematological malignancies: a review. Eur J Haematol. 86(5):361-371.

De Keersmaecker K, Porcu M, Cox L, Girardi T, Vandepoel R, de Beeck JO, Gielen O, Mentens N, Bennett KL and Hantschel O. (2014). NUP214-ABL1-mediated cell proliferation in T-cell acute lymphoblastic leukemia is dependent on the LCK kinase and various interacting proteins. Haematologica. 99(1):85-93.

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

Deenik W, Beverloo HB, van der Poel-van de Luytgaarde SC, Wattel MM, van Esser JW, Valk PJ and Cornelissen JJ. (2009). Rapid complete cytogenetic remission after upfront dasatinib monotherapy in a patient with a NUP214-ABL1-positive T-cell acute lymphoblastic leukemia. Leukemia. (3):627-629.

Detrich HW 3rd, Kieran MW, Chan FY, Barone LM, Yee K, Rundstadler JA, Pratt S, Ransom D and Zon LI. (1995). Intraembryonic hematopoietic cell migration during vertebrate development. Proc Natl Acad Sci USA. 92(23):10713-10717.

Dey S, Curtis DJ, Jane SM and Brandt SJ. (2010). The TAL1/SCL transcription factor regulates cell cycle progression and proliferation in differentiating murine bone marrow monocyte precursors. Mol Cell Biol. 30(9):2181-2192.

Diakos C, Xiao Y, Zheng S, Kager L, Dworzak M and Wiemels JL. (2014). Direct and indirect targets of the E2A-PBX1 leukemia-specific fusion protein. PLoS One. 9(2):e87602.

Dias S, Månsson R, Gurbuxani S, Sigvardsson M and Kee BL. (2008). E2A proteins promote development of lymphoid-primed multipotent progenitors. Immunity. 29(2):217-227.

Dictor M, Ek S, Sundberg M, Warenholt J, György C, Sernbo S, Gustavsson E, Abu-Alsoud W, Wadström T and Borrebaeck C. (2009). Strong lymphoid nuclear expression of SOX11 transcription factor defines lymphoblastic neoplasms, mantle cell lymphoma and Burkitt's lymphoma. Haematologica. 94(11):1563-1568.

Dooley KA, Fraenkel PG, Langer NB, Schmid B, Davidson AJ, Weber G, Chiang K, Foott H, Dwyer C, Wingert RA, Zhou Y, Paw BH and Zon LI; Tübingen 2000 Screen Consortium. (2008). montalcino, A zebrafish model for variegate porphyria. Exp Hematol. 36(9):1132-1142.

Dores GM, Devesa SS, Curtis RE, Linet MS and Morton LM. (2012). Acute leukemia incidence and patient survival among children and adults in the United States, 2001-2007. Blood. 119(1):34-43.

Duque-Afonso J, Feng J, Scherer F, Lin CH, Wong SH, Wang Z, Iwasaki M and Cleary ML. (2015). Comparative genomics reveals multistep pathogenesis of E2A-PBX1 acute lymphoblastic leukemia. J Clin Invest. 125(9):3667-3680.

Durinck K, Goossens S, Peirs S, Wallaert A, Van Loocke W, Matthijssens F, Pieters T, Milani G, Lammens T, Rondou P, Van Roy N, De Moerloose B, Benoit Y, Haigh J, Speleman F, Poppe B and Van Vlierberghe P. (2015). Novel biological insights in T-cell acute lymphoblastic leukemia. Exp Hematol. 43(8):625-639.

Dy P Penzo-Méndez A, Wang H, Pedraza CE, Macklin WB and Lefebvre V. (2008). The three SoxC proteins--Sox4, Sox11 and Sox12--exhibit overlapping expression patterns and molecular properties. Nucleic Acids Res. 36(9):3101-3117.

Einsiedel HG, von Stackelberg A, Hartmann R, Fengler R, Schrappe M, Janka-Schaub G, Mann G, Hählen K, Göbel U, Klingebiel T, Ludwig WD and Henze G. (2005). Long-term outcome in children with relapsed ALL by risk-stratified salvage therapy: results of trial acute lymphoblastic leukemia-relapse study of the Berlin-Frankfurt-Münster Group 87. J Clin Oncol. 23(31): 7942-7950.

Eisen JS and Smith JC. (2008). Controlling morpholino experiments: Don't stop making antisense. Development. 135(10):1735-1743.

Ek S, Dictor M, Jerkeman M, Jirström K and Borrebaeck CA. (2008). Nuclear expression of the non B-cell lineage Sox11 transcription factor identifies mantle cell lymphoma. Blood. 111(2):800-805.

El Omari K, Hoosdally SJ, Tuladhar K, Karia D, Vyas P, Patient R, Porcher C and Mancini EJ. (2011). Structure of the leukemia oncogene LMO2: implications for the assembly of a hematopoietic transcription factor complex. Blood. 117(7):2146-2156.

Fagerli UM, Holt RU, Holien T, Vaatsveen TK, Zhan F, Egeberg KW, Barlogie B, Waage A, Aarset H, Dai HY, Shaughnessy JD Jr, Sundan A and Børset M. (2008). Overexpression and involvement in migration by the metastasis-associated phosphatase PRL-3 in human myeloma cells. Blood. 111(2):806-815.

Fenrick R, Wang L, Nip J, Amann JM, Rooney RJ, Walker-Daniels J, Crawford HC, Hulboy DL, Kinch MS, Matrisian LM and Hiebert SW. (2000). TEL, a putative tumor suppressor, modulates cell growth and cell morphology of ras-transformed cells while repressing the transcription of stromelysin-1. Mol Cell Biol. 20(16):5828-5839.

Fernàndez V, Salamero O, Espinet B, Solé F, Royo C, Navarro A, Camacho F, Beà S, Hartmann E, Amador V, Hernández L, Agostinelli C, Sargent RL, Rozman M, Aymerich M, Colomer D, Villamor N, Swerdlow SH, Pileri SA, Bosch F, Piris MA, Montserrat E, Ott G, Rosenwald A, López-Guillermo A, Jares P, Serrano S and Campo E. (2010). Genomic and gene expression profiling defines indolent forms of mantle cell lymphoma. Cancer Res. 70(4):1408-1418.

Ferrando AA, Neuberg DS, Staunton J, Loh ML, Huard C, Raimondi SC, Behm FG, Pui CH, Downing JR, Gilliland DG, Lander ES, Golub TR and Look AT. (2002). Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell. 1(1):75-87.

Fish Development and Genetics, The Zebraqfish and Medaka Models, Zhiyuan Gong & Vladimir Korzh, Molecular Aspects of Fish and Marine Biology - Vol. 2, 2004.

Fridley BL, Batzler A, Li L, Li F, Matimba A, Jenkins GD, Ji Y, Wang L and Weinshilboum RM. (2011). Gene set analysis of purine and pyrimidine antimetabolites cancer therapies. *Pharmacogenet Genomics*. 21(11):701-712.

Frismantas V, Dobay MP, Rinaldi A, Tchinda J, Dunn SH, Kunz J, Richter-Pechanska P, Marovca B, Pail O, Jenni S, Diaz-Flores E, Chang BH, Brown TJ, Collins RH, Uhrig S, Balasubramanian GP, Bandapalli OR, Higi S, Eugster S, Voegeli P, Delorenzi M, Cario G, Loh ML, Schrappe M, Stanulla M, Kulozik AE, Muckenthaler MU, Saha V, Irving JA, Meisel R, Radimerski T, Von Stackelberg A, Eckert C, Tyner JW, Horvath P, Bornhauser BC and Bourquin JP. (2017). Ex vivo drug response profiling detects recurrent sensitivity patterns in drug-resistant acute lymphoblastic leukemia. Blood. 129(11):e26-e37.

Galloway JL and Zon LI. (2003). Ontogeny of hematopoiesis: examining the emergence of hematopoietic cells in the vertebrate embryo. Curr. Top. Dev. Biol. 53:139-158.

Gatta G, Rossi S, Foschi R, Trama A, Marcos-Gragera R, Pastore G, Peris-Bonet R, Stiller C and Capocaccia R; EUROCARE Working Group. (2013) Survival and cure trends for European children, adolescents and young adults diagnosed with acute lymphoblastic leukemia from 1982 to 2002. Haematologica. 98(5):744-752.

Gauld SB and Cambier JC. (2004). Src-family kinases in B-cell development and signaling. Oncogene. 23(48):8001-8006.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch R, Li C, MaechlerM, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH and Zhang J. (2004). Bioconductor: Open software development for computational biology and bioinformatics. Genome Biol. 5(10):R80.

Georgopoulos K. (2002). Haematopoietic cell-fate decisions, chromatin regulation and ikaros. Nat Rev Immunol. 2(3):162-174.

Gering M and Patient R. (2005). Hedgehog signaling is required for adult blood stem cell formation in zebrafish embryos. Dev Cell. 8(3):389-400.

Girardi T, Vicente C, Cools J and De Keersmaecker K. (2017). The genetics and molecular biology of T-ALL. Blood. 129(9):1113-1123.

Goardon N, Lambert JA, Rodriguez P, Nissaire P, Herblot S, Thibault P, Dumenil D, Strouboulis J, Romeo PH and Hoang T. (2006). ETO2 coordinates cellular proliferation and differentiation during erythropoiesis. EMBO J. 25(2):357-366.

Golub TR, Barker GF, Bohlander SK, Hiebert SW, Ward DC, Bray-Ward P, Morgan E, Raimondi SC, Rowley JD and Gilliland DG. (1995). Fusion of the TEL gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia. Proc Natl Acad Sci U S A. 92(11):4917-4921.

Gore AV, Pillay LM, Venero Galanternik M and Weinstein BM. (2018). The zebrafish: A fintastic model for hematopoietic development and disease. Wiley Interdiscip Rev Dev Biol. 7(3):e312.

Goyama S, Yamaguchi Y, Imai Y, Kawazu M, Nakagawa M, Asai T, Kumano K, Mitani K, Ogawa S, Chiba S, Kurokawa M, Hirai H. (2004). The transcriptionally active form

of AML1 is required for hematopoietic rescue of the AML1-deficient embryonic para-aortic splanchnopleural (P-Sp) region. Blood. 104(12):3558-3564.

Graux C, Cools J, Melotte C, Quentmeier H, Ferrando A, Levine R, Vermeesch JR, Stul M, Dutta B, Boeckx N, Bosly A, Heimann P, Uyttebroeck A, Mentens N, Somers R, MacLeod RA, Drexler HG, Look AT, Gilliland DG, Michaux L, Vandenberghe P, Wlodarska I, Marynen P and Hagemeijer A. (2004). Fusion of NUP214 to ABL1 on amplified episomes in T-cell acute lymphoblastic leukemia. Nat Genet. 36(10):1084-1089.

Graux C, Stevens-Kroef M, Lafage M, Dastugue N, Harrison CJ, Mugneret F, Bahloula K, Struski S, Grégoire MJ, Nadal N, Lippert E, Taviaux S, Simons A, Kuiper RP, Moorman AV, Barber K, Bosly A, Michaux L, Vandenberghe P, Lahortiga I, De Keersmaecker K, Wlodarska I, Cools J, Hagemeijer A and Poirel HA; Groupe Francophone de Cytogénétique Hématologique; 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.

Greaves M. (2018). A causal mechanism for childhood acute lymphoblastic leukaemia. Nat Rev Cancer. 18(8):471-484.

Greaves MF and Wiemels J. (2013) Origins of chromosome translocations in childhood leukaemia. Nat Rev Cancer. 3(9):639-49.

Greaves MF, Maia AT, Wiemels JL, Ford AM. (2013). Leukemia in twins: lessons in natural history. Blood. 102(7):2321-2333.

Greer EL and Shi Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. Nat Rev Genet. 13(5):343-357.

Growney JD, Shigematsu H, Li Z, Lee BH, Adelsperger J, Rowan R, Curley DP, Kutok JL, Akashi K, Williams IR, Speck NA and Gilliland DG. (2005). Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. Blood. 106(2):494-504.

Gunji H, Waga K, Nakamura F, Maki K, Sasaki K, Nakamura Y and Mitani K. (2004). TEL/AML1 shows dominant-negative effects over TEL as well as AML1. Biochem Biophys Res Commun. 322(2):623-630.

Gustavsson E, Sernbo S, Andersson E, Brennan DJ, Dictor M, Jerkeman M, Borrebaeck CA and Ek S. (2010). SOX11 expression correlates to promoter methylation and regulates tumor growth in hematopoietic malignancies. Mol Cancer. 9:187.

Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA, Odenthal J, van Eeden FJ, Jiang YJ, Heisenberg CP, Kelsh RN, Furutani-Seiki M, Vogelsang E, Beuchle D, Schach U, Fabian C and Nüsslein-Volhard C. (1996). The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. Development. 123:1-36.

Hagemeijer A and Graux C. (2010). ABL1 rearrangements in T-cell acute lymphoblastic leukemia. Genes Chromosomes Cancer. 49(4):299-308.

Hall MA, Curtis DJ, Metcalf D, Elefanty AG, Sourris K, Robb L, Gothert JR, Jane SM and Begley CG. (2003). The critical regulator of embryonic hematopoiesis, SCL, is vital in the adult for megakaryopoiesis, erythropoiesis, and lineage choice in CFU-S12. Proc Natl Acad Sci U S A. 100(3):992-997.

Hargrave M, Wright E, Kun J, Emery J, Cooper L and Koopman P. (1997). Expression of the Sox11 gene in mouse embryos suggests roles in neuronal maturation and epithelio-mesenchymal induction. Dev Dyn. 210(2):79-86.

Harigae H, Nakajima O, Suwabe N, Yokoyama H, Furuyama K, Sasaki T, Kaku M, Yamamoto M and Sassa S. (2003). Aberrant iron accumulation and oxidized status of erythroid-specific delta-aminolevulinate synthase (ALAS2)-deficient definitive erythroblasts. Blood. 101(3):1188-1193.

Hebert J, Cayuela JM, Berkeley J and Sigaux F. (1994). Candidate tumor-suppressor genes MTS1 (p16INK4A) and MTS2 (p15INK4B) display frequent homozygous deletions in primary cells from T- but not from B-cell lineage acute lymphoblastic leukemias. Blood. 84(12):4038-4044.

Heinäniemi M, Vuorenmaa T, Teppo S, Kaikkonen MU, Bouvy-Liivrand M, Mehtonen J, Niskanen H, Zachariadis V, Laukkanen S, Liuksiala T, Teittinen K and Lohi O. (2016). Transcription-coupled genetic instability marks acute lymphoblastic leukemia structural variation hotspots. Elife. 5:e13087.

Hjort MA, Abdollahi P, Vandsemb EN, Fenstad MH, Lund B, Slørdahl TS, Børset M and Rø TB. (2017). Phosphatase of regenerating liver-3 is expressed in acute lymphoblastic leukemia and mediates leukemic cell adhesion, migration and drug resistance. Oncotarget. 9(3):3549-3561.

Ho MS, Medcalf RL, Livesey SA and Traianedes K. (2015). The dynamics of adult haematopoiesis in the bone and bone marrow environment. Br J Haematol. (4):472-486.

Hogan LE, Meyer JA, Yang J, Wang J, Wong N, Yang W, Condos G, Hunger SP, Raetz E, Saffery R, Relling MV, Bhojwani D, Morrison DJ and Carroll WL. (2011) Integrated genomic analysis of relapsed childhood acute lymphoblastic leukemia reveals therapeutic strategies. Blood. 118(19):5218-5226.

Holmfeldt L, Wei L, Diaz-Flores E, Walsh M, Zhang J, Ding L, Payne-Turner D, Churchman M, Andersson A, Chen SC, McCastlain K, Becksfort J, Ma J, Wu G, Patel SN, Heatley SL, Phillips LA, Song G, Easton J, Parker M, Chen X, Rusch M, Boggs K, Vadodaria B, Hedlund E, Drenberg C, Baker S, Pei D, Cheng C, Huether R, Lu C, Fulton RS, Fulton LL, Tabib Y, Dooling DJ, Ochoa K, Minden M, Lewis ID, To LB, Marlton P, Roberts AW, Raca G, Stock W, Neale G, Drexler HG, Dickins RA, Ellison DW, Shurtleff SA, Pui CH, Ribeiro RC, Devidas M, Carroll AJ, Heerema NA, Wood B, Borowitz MJ, Gastier-Foster JM, Raimondi SC, Mardis ER, Wilson RK, Downing JR, Hunger SP, Loh ML and Mullighan CG. (2013). The genomic landscape of hypodiploid acute lymphoblastic leukemia. Nat Genet. 45(3):242-252.

Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, McLaren S, Sealy I, Caccamo M, Churcher C, Scott C, Barrett JC, Koch R, Rauch GJ, White S, Chow W, Kilian B, Quintais LT, Guerra-Assunção JA, Zhou Y, Gu Y, Yen J, Vogel JH, Eyre T, Redmond S, Banerjee R, Chi J, Fu B, Langley E, Maguire SF, Laird GK, Lloyd D, Kenyon E, Donaldson S, Sehra H, Almeida-King J, Loveland J, Trevanion S, Jones M, Quail M, Willey D, Hunt A, Burton J, Sims S, McLay K, Plumb B, Davis J, Clee C, Oliver K, Clark R, Riddle C, Elliot D, Threadgold G, Harden G, Ware D, Begum S, Mortimore B, Kerry G, Heath P, Phillimore B, Tracey A, Corby N, Dunn M, Johnson C, Wood J, Clark S, Pelan S, Griffiths G, Smith M, Glithero R, Howden P, Barker N, Lloyd C, Stevens C, Harley J, Holt K, Panagiotidis G, Lovell J, Beasley H, Henderson C, Gordon D, Auger K, Wright D, Collins J, Raisen C, Dyer L, Leung K, Robertson L, Ambridge K, Leongamornlert D, McGuire S, Gilderthorp R, Griffiths C, Manthravadi D, Nichol S, Barker G, Whitehead S, Kay M, Brown J, Murnane C, Gray E, Humphries M, Sycamore N, Barker D, Saunders D, Wallis J, Babbage A, Hammond S, Mashreghi-

Mohammadi M, Barr L, Martin S, Wray P, Ellington A, Matthews N, Ellwood M, Woodmansey R, Clark G, Cooper J, Tromans A, Grafham D, Skuce C, Pandian R, Andrews R, Harrison E, Kimberley A, Garnett J, Fosker N, Hall R, Garner P, Kelly D, Bird C, Palmer S, Gehring I, Berger A, Dooley CM, Ersan-Ürün Z, Eser C, Geiger H, Geisler M, Karotki L, Kirn A, Konantz J, Konantz M, Oberländer M, Rudolph-Geiger S, Teucke M, Lanz C, Raddatz G, Osoegawa K, Zhu B, Rapp A, Widaa S, Langford C, Yang F, Schuster SC, Carter NP, Harrow J, Ning Z, Herrero J, Searle SM, Enright A, Geisler R, Plasterk RH, Lee C, Westerfield M, de Jong PJ, Zon LI, Postlethwait JH, Nüsslein-Volhard C, Hubbard TJ, Roest Crollius H, Rogers J and Stemple DL. (2013). The zebrafish reference genome sequence and its relationship to the human genome. Nature. 496(7446):498-503.

Hu R, Yan Y, Li Q, Lin Y, Jin W, Li H, Lu Y and Pang T. (2010). Increased drug efflux along with midkine gene high expression in childhood B-lineage acute lymphoblastic leukemia cells. *Int J Hematol*. 92(1):105-110.

Hunger SP, Galili N, Carroll AJ, Crist WM, Link MP and Cleary ML. (1991). The t(1;19)(q23;p13) results in consistent fusion of E2A and PBX1 coding sequences in acute lymphoblastic leukemias. Blood. 77(4):687-693.

Hunger SP, Lu X, Devidas M, Camitta BM, Gaynon PS, Winick NJ, Reaman GH and Carroll WL. (2012). Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children's oncology group. J Clin Oncol. 30(14):1663-1669.

Hyde RK and Liu PP. (2013). Germline PAX5 mutations and B cell leukemia. Nat Genet. 45(10):1104-1105.

Ichikawa M, Asai T, Saito T, Seo S, Yamazaki I, Yamagata T, Mitani K, Chiba S, Ogawa S, Kurokawa M and Hirai H. (2004). AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. Nat Med. 10(3):299-304.

Inaba H, Greaves M and Mullighan CG. (2013). Acute lymphoblastic leukaemia. Lancet. 381(9881):1943-1955.

Irving J, Matheson E, Minto L, Blair H, Case M, Halsey C, Swidenbank I, Ponthan F, Kirschner-Schwabe R, Groeneveld-Krentz S, Hof J, Allan J, Harrison C, Vormoor J, von Stackelberg A amd 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.

Jagannathan-Bogdan M and Zon LI. (2013). Hematopoiesis. Development. 140(12):2463-2467.

Jan M and Majeti R. (2013). Clonal evolution of acute leukemia genomes. Oncogene. 32(2):135-140.

Jay P, Gozé C, Marsollier C, Taviaux S, Hardelin JP, Koopman P and Berta P. (1995). The human SOX11 gene: cloning, chromosomal assignment and tissue expression. Genomics. 29(2):541-545.

Jay P, Sahly I, Gozé C, Taviaux S, Poulat F, Couly G, Abitbol M and Berta P. (1997). SOX22 is a new member of the SOX gene family, mainly expressed in human nervous tissue. Hum Mol Genet. 6(7):1069-77.

Jia M, Zhao HZ, Cheng YP, Luo ZB, Zhang JY, Li SS, Xu XJ and Tang YM. (2016). High expression of Midkine (MK) indicates poor prognosis in childhood acute lymphoblastic leukemia. *Hematology*. 21(2):69-77.

Jing L and Zon LI. (2011) Zebrafish as a model for normal and malignant hematopoiesis. Dis Model Mech. 4(4):433-438.

Jones CL, Gearheart CM, Fosmire S, Delgado-Martin C, Evensen NA, Bride K, Waanders AJ, Pais F, Wang J, Bhatla T, Bitterman DS, de Rijk SR, Bourgeois W, Dandekar S, Park E, Burleson TM, Madhusoodhan PP, Teachey DT, Raetz EA, Hermiston ML, Müschen M, Loh ML, Hunger SP, Zhang J, Garabedian MJ, Porter CC and Carroll WL. (2015). MAPK signaling cascades mediate distinct glucocorticoid resistance mechanisms in pediatric leukemia. Blood. 126(19):2202-2212.

Juric D, Lacayo NJ, Ramsey MC, Racevskis J, Wiernik PH, Rowe JM, Goldstone AH, O'Dwyer PJ, Paietta E and Sikic BI. (2007). Differential gene expression patterns and interaction networks in BCR-ABL-positive and -negative adult acute lymphoblastic leukemias. J Clin Oncol. 25(11):1341-1349.

Kadamb R, Mittal S, Bansal N, Batra H and Saluja D. (2013). Sin3: insight into its transcription regulatory functions. Eur J Cell Biol. 92(8-9):237-246.

Kalev-Zylinska ML, Horshield JA, Flores MV, Postlethwait JH, Vitas MR, Baas AM, Crosier PS and Crosier KE. (2002). Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. Development. 129(8):2015-2030.

Kaplan JB, Grischenko M and Giles FJ. (2015). Blinatumomab for the treatment of acute lymphoblastic leukemia. Invest New Drugs. 33(6):1271-1279.

Kastner P, Dupuis A, Gaub MP, Herbrecht R, Lutz P and Chan S. (2013). Function of Ikaros as a tumor suppressor in B cell acute lymphoblastic leukemia. Am J Blood Res. 3(1):1-13.

Kee BL, Quong MW and Murre C. (2000). E2A proteins: essential regulators at multiple stages of B-cell development. Immunol Rev. 175:138-149.

Klose RJ and Bird AP. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci. 31(2):89-97.

Kobayashi M, Chen S, Bai Y, Yao C, Gao R, Sun XJ, Mu C, Twiggs TA, Yu ZH, Boswell HS, Yoder MC, Kapur R, Mulloy JC, Zhang ZY and Liu Y. (2017). Phosphatase PRL2 promotes AML1-ETO-induced acute myeloid leukemia. Leukemia. 31(6):1453-1457.

Kobayashi M, Chen S, Gao R, Bai Y, Zhang ZY and Liu Y. (2014). Phosphatase of regenerating liver in hematopoietic stem cells and hematological malignancies. Cell Cycle. 13(18):2827-2835.

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

Kohli RM and Zhang Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. Nature. 502(7472):472-479.

Korkeamäki H, Viiri K, Kukkonen MK, Mäki M and Lohi O. (2008). Alternative mRNA splicing of SAP30L regulates its transcriptional repression activity. FEBS Lett. 582(2):379-384.

Krueger A, Ziętara N and Łyszkiewicz M. (2017). T Cell Development by the Numbers. Trends Immunol. 38(2):128-139.

Kuci V, Nordström L, Conrotto P and Ek S. (2016). SOX11 and HIG-2 are cross-regulated and affect growth in mantle cell lymphoma. *Leuk Lymphoma*. 57(8):1883-1892.

Kuhlbrodt K, Herbarth B, Sock E, Enderich J, Hermans-Borgmeyer I and Wegner M. (1998). Cooperative function of POU proteins and SOX proteins in glial cells. J Biol Chem. 273(26):16050-16057.

Kuhns DB, Fink DL, Choi U, Sweeney C, Lau K, Priel DL, Riva D, Mendez L, Uzel G, Freeman AF, Olivier KN, Anderson VL, Currens R, Mackley V, Kang A, Al-Adeli M, Mace E, Orange JS, Kang E, Lockett SJ, Chen, Steinbach PJ, Hsu AP, Zarember KA, Malech HL, Gallin JI and Holland SM. (2016). Cytoskeletal abnormalities and neutrophil dysfunction in WDR1 deficiency. *Blood*. 128(17):2135-2143.

Kuo PY, Jatiani SS, Rahman AH, Edwards D, Jiang Z, Ahr K, Perumal D, Leshchenko VV, Brody J, Shaknovich R, Ye BH and Parekh S. (2018). SOX11 augments BCR signaling to drive MCL-like tumor development. Blood. 131(20):2247-2255.

Kuo PY, Leshchenko VV, Fazzari MJ, Perumal D, Gellen T, He T, Iqbal J, Baumgartner-Wennerholm S, Nygren L, Zhang F, Zhang W, Suh KS, Goy A, Yang DT, Chan WC, Kahl BS, Verma AK, Gascoyne RD, Kimby E, Sander B, Ye BH, Melnick AM and Parekh S. (2015). High-resolution chromatin immunoprecipitation (ChIP) sequencing reveals novel binding targets and prognostic role for SOX11 in mantle cell lymphoma. Oncogene. 34(10):1231-1240.

Kwon K, Hutter C, Sun Q, Bilic I, Cobaleda C, Malin S and 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.

Lacombe J, Herblot S, Rojas-Sutterlin S, Haman A, Barakat S, Iscove NN, Sauvageau G and Hoang T. (2010). Scl regulates the quiescence and the long-term competence of hematopoietic stem cells. Blood. 115(4):792-803.

Laherty CD, Billin AN, Lavinsky RM, Yochum GS, Bush AC, Sun JM, Mullen JM, Davie JR, Rose DW, Glass CK, Rosenfeld MG, Ayer DE and Eisenman RN. (1998). SAP30, a component of the mSin3 corepressor complex involved in N-CoR-mediated repression by specific transcription factors. Mol Cell. 2(1):33-42.

Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR and Weirauch MT. (2018). The Human Transcription Factors. Cell. 172(4):650-665.

Latchman DS. (2001). Transcription factors: bound to activate or repress. Trends Biochem Sci. 26(4):211-213.

Latonen L, Afyounian E, Jylhä A, Nättinen J, Aapola U, Annala M, Kivinummi KK, Tammela TTL, Beuerman RW, Uusitalo H, Nykter M and Visakorpi T. (2018). Integrative proteomics in prostate cancer uncovers robustness against genomic and transcriptomic aberrations during disease progression. Nat Commun. 9(1):1176.

Latysheva NS and Babu MM. (2016). Discovering and understanding oncogenic gene fusions through data intensive computational approaches. Nucleic Acids Res. 44(10):4487-4503.

Laudet V, Stehelin D and Clevers H. (1993). Ancestry and diversity of the HMG box superfamily. Nucleic Acids Res. 21(10):2493-2501.

Laukkanen S, Grönroos T, Pölönen P, Kuusanmäki H, Mehtonen J, Cloos J, Ossenkoppele G, Gjertsen B, Øystein B, Heckman C, Heinäniemi M, Kontro M and Lohi O. (2017). In silico and preclinical drug screening identifies dasatinib as a targeted therapy for T-ALL. Blood Cancer J. 7(9):e604.

Lausten-Thomsen U, Madsen HO, Vestergaard TR, Hjalgrim H, Nersting J and Schmiegelow K. (2011). Prevalence of t(12;21)[ETV6-RUNX1]-positive cells in healthy neonates. Blood. 117(1):186-189.

Lee KC, Ouwehand I, Giannini AL, Thomas NS, Dibb NJ and Bijlmakers MJ. (2010). Lck is a key target of imatinib and dasatinib in T-cell activation. Leukemia. 24(4):896-900.

Lee P, Bhansali R, Izraeli S, Hijiya N and Crispino JD. (2016). The biology, pathogenesis and clinical aspects of acute lymphoblastic leukemia in children with Down syndrome. Leukemia. 30(9):1816-1823.

Lefebvre V, Dumitriu B, Penzo-Méndez A, Han Y and Pallavi B. (2007). Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. Int J Biochem Cell Biol. 39(12):2195-2214.

Lessman CA. (2011) The developing zebrafish (Danio rerio): a vertebrate model for high-throughput screening of chemical libraries. Birth Defects Res C Embryo Today 93(3):268-380.

Levine M. (2010). Transcriptional enhancers in animal development and evolution. Curr Biol. 20(17):R754-763.

Liang DC, Shih LY, Yang CP, Hung IJ, Liu HC, Jaing TH, Yeh TC, Liang ST, Chang CL, Lee EH, Lai CL and Chang WH. (2010). Frequencies of ETV6-RUNX1 fusion and hyperdiploidy in pediatric acute lymphoblastic leukemia are lower in far east than west. Pediatr Blood Cancer. 55(3):430-433.

Lilljebjörn H and Fioretos T. (2017). New oncogenic subtypes in pediatric B-cell precursor acute lymphoblastic leukemia. Blood. 130(12):1395-1401.

Lilljebjörn H, Henningsson R, Hyrenius-Wittsten A, Olsson L, Orsmark-Pietras C, von Palffy S, Askmyr M, Rissler M, Schrappe M, Cario G, Castor A, Pronk CJ, Behrendtz M, Mitelman F, Johansson B, Paulsson K, Andersson AK, Fontes M and Fioretos T. (2016). Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. Nat Commun. 7:11790.

Linabery AM and Ross JA. (2008). Trends in childhood cancer incidence in the U.S. (1992-2004) Cancer. 112(2):416-432.

Lindfors K, Viiri KM, Niittynen M, Heinonen TY, Maki M and Kainulainen H. (2003). TGF-beta induces the expression of SAP30L, a novel nuclear protein. BMC Genomics. 4:53.

Liu J, Peng L, Niu T, Wu Y, Li J, Wang F, Zheng Y and Liu T. (2016). PIG7 promotes leukemia cell chemosensitivity via lysosomal membrane permeabilization. *Oncotarget*. 7(4):4841-4859.

Liu P, Ramachandran S, Ali Seyed M, Scharer CD, Laycock N, Dalton WB, Williams H, Karanam S, Datta MW, Jaye DL and Moreno CS. (2006). Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells. Cancer Res. 66(8):4011-4019.

Liu Y, Easton J, Shao Y, Maciaszek J, Wang Z, Wilkinson MR, McCastlain K, Edmonson M, Pounds SB, Shi L, Zhou X, Ma X, Sioson E, Li Y, Rusch M, Gupta P, Pei D, Cheng C, Smith MA, Auvil JG, Gerhard DS, Relling MV, Winick NJ, Carroll AJ, Heerema NA, Raetz E, Devidas M, Willman CL, Harvey RC, Carroll WL, Dunsmore KP, Winter SS, Wood BL, Sorrentino BP, Downing JR, Loh ML, Hunger SP, Zhang J and Mullighan CG. (2017). The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. Nat Genet. 49(8):1211-1218.

Livak KJ and Schmittgen TD. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 25(4):402-408.

Lohi O, Parikka M and Rämet M. (2013). The zebrafish as a model for paediatric diseases. Acta Paediatr. 102(2):104-110.Lord M, Arvidsson G, Wasik AM, Christensson B, Wright AP, Grandien A and Sander B. (2018). Impact of Sox11 overexpression in Ba/F3 cells. *Haematologica*. pii: haematol.2018.197467 [Epub ahead of print].

Lorsbach RB, Moore J, Ang SO, Sun W, Lenny N and Downing JR. (2004). Role of RUNX1 in adult hematopoiesis: analysis of RUNX1-IRES-GFP knock-in mice reveals differential lineage expression. Blood. 103(7):2522-2529.

Ma X, Edmonson M, Yergeau D, Muzny DM, Hampton OA, Rusch M, Song G, Easton J, Harvey RC, Wheeler DA, Ma J, Doddapaneni H, Vadodaria B, Wu G, Nagahawatte P, Carroll WL, Chen IM, Gastier-Foster JM, Relling MV, Smith MA, Devidas M, Guidry Auvil JM, Downing JR, Loh ML, Willman CL, Gerhard DS, Mullighan CG, Hunger SP and Zhang J. (2015). Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. Nat Commun. 6:6604.

Mann G, Attarbaschi A, Schrappe M, De Lorenzo P, Peters C, Hann I, De Rossi G, Felice M, Lausen B, Leblanc T, Szczepanski T, Ferster A, Janka-Schaub G, Rubnitz J, Silverman LB, Stary J, Campbell M, Li CK, Suppiah R, Biondi A, Vora A, Valsecchi MG and Pieters R; Interfant-99 Study Group. (2010). Improved outcome with hematopoietic stem cell transplantation in a poor prognostic subgroup of infants with mixed-lineage-leukemia (MLL)-rearranged acute lymphoblastic leukemia: results from the Interfant-99 Study. Blood. 116(15):2644-2650.

Mar BG, Bullinger LB, McLean KM, Grauman PV, Harris MH, Stevenson K, Neuberg DS, Sinha AU, Sallan SE, Silverman LB, Kung AL, Lo Nigro L, Ebert BL and Armstrong SA. (2014). Mutations in epigenetic regulators including SETD2 are gained during relapse in paediatric acute lymphoblastic leukaemia. Nat Commun. 5:3469.

Marincevic-Zuniga Y, Dahlberg J, Nilsson S, Raine A, Nystedt S, Lindqvist CM, Berglund EC, Abrahamsson J, Cavelier L, Forestier E, Heyman M, Lönnerholm G, Nordlund J and Syvänen AC. (2017). Transcriptome sequencing in pediatric acute lymphoblastic leukemia identifies fusion genes associated with distinct DNA methylation profiles. J. Hematol. Oncol. 10(1):148.

Massari ME and Murre C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. Mol Cell Biol. 20(2):429-440.

Matter WF, Estridge T, Zhang C, Belagaje R, Stancato L, Dixon J, Johnson B, Bloem L, Pickard T, Donaghue M, Acton S, Jeyaseelan R, Kadambi V and Vlahos CJ. (2001). Role of PRL-3, a human muscle-specific tyrosine phosphatase, in angiotensin-II signaling. Biochem Biophys Res Commun. 283(5):1061-1068.

Mavrothalassitis G and Ghysdael J. (2000). Proteins of the ETS family with transcriptional repressor activity. Oncogene. 19(55):6524-6532.

Mazieres J, You L, He B, Xu Z, Lee AY, Mikami I, McCormick F and Jablons DM. (2005). Inhibition of Wnt16 in human acute lymphoblastoid leukemia cells containing the t(1;19) translocation induces apoptosis. Oncogene. 24(34):5396-5400.

McGrath KE, Frame JM, Fromm GJ, Koniski AD, Kingsley PD, Little J, Bulger M and Palis J. (2011). A transient definitive erythroid lineage with unique regulation of the β-globin locus in the mammalian embryo. Blood. 117(17):4600-4608.

McWhirter JR, Neuteboom ST, Wancewicz EV, Monia BP, Downing JR and Murre C. (1999). Oncogenic homeodomain transcription factor E2A-Pbx1 activates a novel WNT gene in pre-B acute lymphoblastoid leukemia. Proc Natl Acad Sci U S A. 96(20):11464-11469.

Mercier F and Scadden DT. (2015). Not All Created Equal: Lineage Hard-Wiring in the Production of Blood. Cell. 163(7):1568-1570.

Meyer JA, Wang J, Hogan LE, Yang JJ, Dandekar S, Patel JP, Tang Z, Zumbo P, Li S, Zavadil J, Levine RL, Cardozo T, Hunger SP, Raetz EA, Evans WE, Morrison DJ, Mason

CE and Carroll WL. (2013). Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia. Nat Genet. 45(3):290-294.

Meyers J, Yu Y, Kaye JA and Davis KL. (2013). Medicare fee-for-service enrollees with primary acute myeloid leukemia: an analysis of treatment patterns, survival, and healthcare resource utilization and costs. Appl Health Econ Health Policy. 11(3):275-286.

Min G, Lee SK, Kim HN, Han YM, Lee RH, Jeong DG, Han DC and Kwon BM. (2013). Rhodanine-based PRL-3 inhibitors blocked the migration and invasion of metastatic cancer cells. Bioorg Med Chem Lett. 23(13):3769-3774.

Monica K, LeBrun DP, Dedera DA, Brown R and Cleary ML. (1994). Transformation properties of the E2a-Pbx1 chimeric oncoprotein: fusion with E2a is essential, but the Pbx1 homeodomain is dispensable. Mol Cell Biol. 14(12):8304-8314.

Mori H, Colman SM, Xiao Z, Ford AM, Healy LE, Donaldson C, Hows JM, Navarrete C and Greaves M. (2002). Chromosome translocations and covert leukemic clones are generated during normal fetal development. Proc Natl Acad Sci U S A. 99(12):8242-8247.

Morrow M, Samanta A, Kioussis D, Brady HJ and Williams O. (2007). TEL-AML1 preleukemic activity requires the DNA binding domain of AML1 and the dimerization and corepressor binding domains of TEL. Oncogene. 26(30):4404-4414.

Motiwala T and Jacob ST. (2006). Role of protein tyrosine phosphatases in cancer. Prog Nucleic Acid Res Mol Biol. 81:297-329.

Mouthon MA, Bernard O, Mitjavila MT, Romeo PH, Vainchenker W and Mathieu-Mahul D. (1993). Expression of tal-1 and GATA-binding proteins during human hematopoiesis. Blood. 81(3):647-655.

Mozos A, Royo C, Hartmann E, De Jong D, Baró C, Valera A, Fu K, Weisenburger DD, Delabie J, Chuang SS, Jaffe ES, Ruiz-Marcellan C, Dave S, Rimsza L, Braziel R, Gascoyne RD, Solé F, López-Guillermo A, Colomer D, Staudt LM, Rosenwald A, Ott G, Jares P and Campo E. (2009). SOX11 expression is highly specific for mantle cell lymphoma and identifies the cyclin D1-negative subtype. Haematologica. 94(11):1555-1562.

Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, Girtman K, Mathew S, Ma J, Pounds SB, Su X, Pui CH, Relling MV, Evans WE, Shurtleff SA and Downing JR. (2007). Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature. 446(7137):758-764.

Mullighan CG, Phillips LA, Su X, Ma J, Miller CB, Shurtleff SA and Downing JR. (2008). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. Science. 322(5906):1377-1380.

Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB, Ma J, Liu W, Cheng C, Schulman BA, Harvey RC, Chen IM, Clifford RJ, Carroll WL, Reaman G, Bowman WP, Devidas M, Gerhard DS, Yang W, Relling MV, Shurtleff SA, Campana D, Borowitz MJ, Pui CH, Smith M, Hunger SP, Willman CL and Downing JR; Children's Oncology Group. (2009). Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. N Engl J Med. 360(5):470-480.

Mullighan CG, Zhang J, Kasper LH, Lerach S, Payne-Turner D, Phillips LA, Heatley SL, Holmfeldt L, Collins-Underwood JR, Ma J, Buetow KH, Pui CH, Baker SD, Brindle PK and Downing JR. (2011). CREBBP mutations in relapsed acute lymphoblastic leukaemia. Nature. 471(7337):235-239.

Nachman JB, Heerema NA, Sather H, Camitta B, Forestier E, Harrison CJ, Dastugue N, Schrappe M, Pui CH, Basso G, Silverman LB and Janka-Schaub GE. (2007). Outcome of treatment in children with hypodiploid acute lymphoblastic leukemia. Blood. 110(4):1112-1115.

Neveu B, Spinella JF, Richer C, Lagacé K, Cassart P, Lajoie M, Jananji S, Drouin S, Healy J, Hickson GR and Sinnett D. (2016). CLIC5: a novel ETV6 target gene in childhood acute lymphoblastic leukemia. *Haematologica*. 101(12):1534-1543.

Ney Garcia DR, de Souza MT, de Figueiredo AF, Othman MAK, Rittscher K, Abdelhay E, Capela de Matos RR, Meyer C, Marschalek R, Land MGP, Liehr T, Ribeiro RC and Silva MLM. (2017). Molecular characterization of KMT2A fusion partner genes in 13 cases of pediatric leukemia with complex or cryptic karyotypes. *Hematol Oncol.* 35(4):760-768.

Nguyen K, Devidas M, Cheng SC, La M, Raetz EA, Carroll WL, Winick NJ, Hunger SP, Gaynon PS and Loh ML; Children's Oncology Group. (2008). Factors influencing survival after relapse from acute lymphoblastic leukemia: a Children's Oncology Group study. Leukemia. 22(12):2142-2150.

Nordlund J, Bäcklin CL, Wahlberg P, Busche S, Berglund EC, Eloranta ML, Flaegstad T, Forestier E, Frost BM, Harila-Saari A, Heyman M, Jónsson OG, Larsson R, Palle J, Rönnblom L, Schmiegelow K, Sinnett D, Söderhäll S, Pastinen T, Gustafsson MG, Lönnerholm G and Syvänen AC. (2013). Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. Genome Biol. 14(9):r105.

North TE, Stacy T, Matheny CJ, Speck NA and de Bruijn MF. (2004). Runx1 is expressed in adult mouse hematopoietic stem cells and differentiating myeloid and lymphoid cells, but not in maturing erythroid cells. Stem Cells. 22(2):158-168.

O'Connor D, Enshaei A, Bartram J, Hancock J, Harrison CJ, Hough R, Samarasinghe S, Schwab C, Vora A, Wade R, Moppett J, Moorman AV and Goulden N. (2018). Genotype-Specific Minimal Residual Disease Interpretation Improves Stratification in Pediatric Acute Lymphoblastic Leukemia. J Clin Oncol. 36(1):34-43.

O'Neil J and Look AT. (2007). Mechanisms of transcription factor deregulation in lymphoid cell transformation. Oncogene. 26(47):6838-6849.

Okuda T, Takeda K, Fujita Y, Nishimura M, Yagyu S, Yoshida M, Akira S, Downing JR and Abe T. (2000). Biological characteristics of the leukemia-associated transcriptional factor AML1 disclosed by hematopoietic rescue of AML1-deficient embryonic stem cells by using a knock-in strategy. Mol Cell Biol. 20(1):319-328.

Okuda T, Van Deursen J, Hiebert SW, Grosweld G and Downing JR. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell. 84(2):321-330.

Orkin SH and Zon LI. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell. 132(4):631-644.

Osato M. (2004). Point mutations in the RUNX1/AML1 gene: another actor in RUNX leukemia. Oncogene. 23(24):4284-4296.

Oskarsson T, Söderhäll S, Arvidson J, Forestier E, Montgomery S, Bottai M, Lausen B, Carlsen N, Hellebostad M, Lähteenmäki P, Saarinen-Pihkala UM, Jónsson ÓG and Heyman M; Nordic Society of Paediatric Haematology and Oncology (NOPHO) ALL relapse working group. (2016). Relapsed childhood acute lymphoblastic leukemia in the Nordic countries: prognostic factors, treatment and outcome. Haematologica. 101(1):68-76.

Paik EJ and Zon LI. (2010). Hematopoietic development in the zebrafish. Int J Dev Biol. 54(6-7):1127-1137.

Palacios EH and Weiss A. (2004) Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation. Oncogene. 23(48):7990-8000.

Palis J and Yoder MC. (2001). Yolk-sac hematopoiesis: the first blood cells of mouse and man. Exp. Hematol. 29(8):927-936.

Park JE, Yuen HF, Zhou JB, Al-Aidaroos AQ, Guo K, Valk PJ, Zhang SD, Chng WJ, Hong CW, Mills K and Zeng Q. (2013). Oncogenic roles of PRL-3 in FLT3-ITD induced acute myeloid leukaemia. EMBO Mol Med. 5(9):1351-1366.

Parsons SJ and Parsons JT. (2004). Src family kinases, key regulators of signal transduction. Oncogene. 23(48):7906-7909.

Paul MK and Mukhopadhyay AK. (2004). Tyrosine kinase - Role and significance in Cancer. Int J Med Sci. 1(2):101-115.

Paul F, Arkin Y, Giladi A, Jaitin DA, Kenigsberg E, Keren-Shaul H, Winter D, Lara-Astiaso D, Gury M, Weiner A, David E, Cohen N, Lauridsen FK, Haas S, Schlitzer A, Mildner A, Ginhoux F, Jung S, Trumpp A, Porse BT, Tanay A and Amit I. (2015). Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. Cell. 163:1663-1677, correction 164(1-2):325.

Payne KJ and Dovat S. (2011). Ikaros and tumor suppression in acute lymphoblastic leukemia. Crit Rev Oncog. 16(1-2):3-12.

Peal DS, Peterson RT and Milan D. Small molecule screening in zebrafish. J Cardiovasc Transl Res. 3(5):454-460.

PeCan: St. Jude Children`s Research Hospital's Pediatric Cancer data portal; https://pecan.stjude.org/home/ (accessed 27.10.2015)

Pemovska T, Kontro M, Yadav B, Edgren H, Eldfors S, Szwajda A, Almusa H, Bespalov MM, Ellonen P, Elonen E, Gjertsen BT, Karjalainen R, Kulesskiy E, Lagström S, Lehto A, Lepistö M, Lundán T, Majumder MM, Marti JM, Mattila P, Murumägi A, Mustjoki S, Palva A, Parsons A, Pirttinen T, Rämet ME, Suvela M, Turunen L, Västrik I, Wolf M, Knowles J, Aittokallio T, Heckman CA, Porkka K, Kallioniemi O and Wennerberg K. (2013). Individualized systems medicine strategy to tailor treatments for patients with chemorefractory acute myeloid leukemia. Cancer Discov. 3(12):1416-1429.

Perié L, Duffy KR, Kok L, de Boer RJ and Schumacher TN. (2015). The Branching Point in Erythro-Myeloid Differentiation. Cell. 163(7) :1655-1662.

Peserico A and Simone C. (2011). Physical and functional HAT/HDAC interplay regulates protein acetylation balance. J Biomed Biotechnol. 2011:371832.

Pierro J, Hogan LE, Bhatla T and Carroll WL. (2017). New targeted therapies for relapsed pediatric acute lymphoblastic leukemia. Expert Rev Anticancer Ther. 17(8):725-736.

Pieters R, Schrappe M, De Lorenzo P, Hann I, De Rossi G, Felice M, Hovi L, LeBlanc T, Szczepanski T, Ferster A, Janka G, Rubnitz J, Silverman L, Stary J, Campbell M, Li CK, Mann G, Suppiah R, Biondi A, Vora A and Valsecchi MG. (2007). A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. Lancet. 370(9583):240-250.

Pimkin M, Kossenkov AV, Mishra T, Morrissey CS, Wu W, Keller CA, Blobel GA, Lee D, Beer MA, Hardison RC and Weiss MJ. (2014). Divergent functions of hematopoietic transcription factors in lineage priming and differentiation during erythro-megakaryopoiesis. Genome Res. 24(12):1932-1944.

Porcher C, Chagraoui H and Kristiansen MS. (2017). SCL/TAL1: a multifaceted regulator from blood development to disease. Blood. 129(15):2051-2060.

Pramoonjago P, Baras AS and Moskaluk CA. (2006). Knockdown of Sox4 expression by RNAi induces apoptosis in ACC3 cells. Oncogene. 25(41):5626-5639.

Pui CH and Evans WE. (2006). Treatment of acute lymphoblastic leukemia. N Engl J Med. 354(2): 166-178.

Pui CH, Carroll WL, Meshinchi S and Arceci RJ. (2011). Biology, risk stratification, and therapy of pediatric acute leukemias: an update. J Clin Oncol. 29(5):551-565.

Pui CH, Robison LL and Look AT. (2008). Acute lymphoblastic leukaemia. Lancet. 371(9617):1030-1043.

Pui CH, Yang JJ, Hunger SP, Pieters R, Schrappe M, Biondi A, Vora A, Baruchel A, Silverman LB, Schmiegelow K, Escherich G, Horibe K, Benoit YC, Izraeli S, Yeoh AE, Liang DC, Downing JR, Evans WE, Relling MV and Mullighan CG. (2015). Childhood Acute Lymphoblastic Leukemia: Progress Through Collaboration. J Clin Oncol. 33(27):2938-2948.

Puumala SE, Ross JA, Aplenc R and Spector LG. (2013). Epidemiology of childhood acute myeloid leukemia. Pediatr Blood Cancer. (5):728-733.

Pölönen P, Mehtonen J, Lin J, Liuksiala T, Häyrynen S, Teppo S, Mäkinen A, Kumar A, Malani D, Pohjolainen V, Porkka K, Heckman CA, May P, Hautamäki V, Granberg KJ, Lohi O, Nykter M and Heinäniemi M. (2019). Hemap: An interactive online resource for characterizing molecular phenotypes across hematologic malignancies. Cancer Research. Published Online First April 2, 2019. doi: 10.1158/0008-5472.CAN-18-2970

Quintás-Cardama A and Cortes J. (2009). Molecular biology of bcr-abl1-positive chronic myeloid leukemia. Blood. 113(8):1619-1630.

Quintás-Cardama A, Tong W, Manshouri T, Vega F, Lennon PA, Cools J, Gilliland DG, Lee F, Cortes J, Kantarjian H and Garcia-Manero G. (2008). Activity of tyrosine kinase inhibitors against human NUP214-ABL1-positive T cell malignancies. Leukemia. 22(6):1117-1124.

R Development Core Team. (2008). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ransom DG, Haffter P, Odenthal J, Brownlie A, Vogelsang E, Kelsh RN, Brand M, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, Mullins MC and Nüsslein-Volhard C. (1996). Characterization of zebrafish mutants with defects in embryonic hematopoiesis. Development. 123:311-319.

Rasighaemi P, Onnebo SM, Liongue C and Ward AC. (2015). ETV6 (TEL1) regulates embryonic hematopoiesis in zebrafish. Haematologica. 100(1):23-31.

Ravandi F1, O'Brien S, Thomas D, Faderl S, Jones D, Garris R, Dara S, Jorgensen J, Kebriaei P, Champlin R, Borthakur G, Burger J, Ferrajoli A, Garcia-Manero G, Wierda W, Cortes J and Kantarjian H. (2010). First report of phase 2 study of dasatinib with hyper-CVAD for the frontline treatment of patients with Philadelphia chromosome-positive (Ph+) acute lymphoblastic leukemia. Blood. 116(12):2070-2077.

Rhodes J, Hagen A, Hsu K, Deng M, Liu TX, Look AT and Kanki JP. (2005). Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. Dev Cell. 8(1):97-108.

Rieger MA and Schroeder T. (2012). Hematopoiesis. Cold Spring Harb Perspect Biol. 4(12).

Rios P, Li X and Köhn M. (2013). Molecular mechanisms of the PRL phosphatases. FEBS J. 280(2):505-524.

Roberts KG and Mullighan CG. (2015). Genomics in acute lymphoblastic leukaemia: insights and treatment implications. Nat Rev Clin Oncol. 12(6):344-357.

Robu ME, Larson JD, Nasevicius A, Beiraghi S, Brenner C, Farber SA and Ekker SC. (2007). p53 activation by knockdown technologies. PLoS Genet. 3(5):e78.

Rodríguez-Hernández G, Hauer J, Martín-Lorenzo A, Schäfer D, Bartenhagen C, García-Ramírez I, Auer F, González-Herrero I, Ruiz-Roca L, Gombert M, Okpanyi V, Fischer U, Chen C, Dugas M, Bhatia S, Linka RM, Garcia-Suquia M, Rascón-Trincado MV, Garcia-Sanchez A, Blanco O, García-Cenador MB, García-Criado FJ, Cobaleda C, Alonso-López D, De Las Rivas J, Müschen M, Vicente-Dueñas C, Sánchez-García I and Borkhardt A. (2017). Infection Exposure Promotes ETV6-RUNX1 Precursor B-cell Leukemia via Impaired H3K4 Demethylases. Cancer Res. 77(16):4365-4377.

Romana SP, Mauchauffé M, Le Coniat M, Chumakov I, Le Paslier D, Berger R and Bernard OA. (1995b). The t(12;21) of acute lymphoblastic leukemia results in a tel-AML1 gene fusion. Blood. 85(12):3662-3670.

Romana SP, Poirel H, Leconiat M, Flexor MA, Mauchauffé M, Jonveaux P, Macintyre EA, Berger R and Bernard OA. (1995a). High frequency of t(12;21) in childhood B-lineage acute lymphoblastic leukemia. Blood. 86(11):4263-4269.

Ross JA and Robison LL. (1997). MLL rearrangements in infant leukemia: is there a higher frequency in females? Leuk Res. 21(8):793-795.

Royer-Pokora B, Loos U and Ludwig WD. (1991). TTG-2, a new gene encoding a cysteine-rich protein with the LIM motif, is overexpressed in acute T-cell leukaemia with the t(11;14)(p13;q11). Oncogene. 6(10):1887-1893.

Rubnitz JE. (2017) Current Management of Childhood Acute Myeloid Leukemia. Paediatr Drugs. 19(1):1-10.

Rubnitz JE, Pui CH and Downing JR. (1999). The role of TEL fusion genes in pediatric leukemias. Leukemia. 13(1):6-13.

Saarinen-Pihkala UM, Heilmann C, Winiarski J, Glomstein A, Abrahamsson J, Arvidson J, Békássy AN, Forestier E, Jonmundsson G, Schroeder H, Vettenranta K, Wesenberg F, Gustafsson G. (2006). Pathways through relapses and deaths of children with acute lymphoblastic leukemia: role of allogeneic stem-cell transplantation in Nordic data. J Clin Oncol. 24(36):5750-5762.

Sabaawy HE, Azuma M, Embree LJ, Tsai HJ, Starost MF and Hickstein DD. (2006). TEL-AML1 transgenic zebrafish model of precursor B cell acute lymphoblastic leukemia. Proc Natl Acad Sci U S A. 103(41):15166-15171.

Salmond RJ, Filby A, Qureshi I, Caserta S and Zamoyska R. (2009). T-cell receptor proximal signaling via the Src-family kinases, Lck and Fyn, influences T-cell activation, differentiation, and tolerance. Immunol Rev. 228(1):9-22.

Sanda T, Lawton LN, Barrasa MI, Fan ZP, Kohlhammer H, Gutierrez A, Ma W, Tatarek J, Ahn Y, Kelliher MA, Jamieson CH, Staudt LM, Young RA and Look AT. (2012). Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. Cancer Cell. 22(2):209-221.

Sassa S and Kappas A. (2000). Molecular aspects of the inherited porphyrias. J Intern Med. 247(2):169-178.

Schade AE, Schieven GL, Townsend R, Jankowska AM, Susulic V, Zhang R, Szpurka H and Maciejewski JP. (2008). Dasatinib, a small-molecule protein tyrosine kinase inhibitor, inhibits T-cell activation and proliferation. Blood. 111(3):1366-1377.

Schepers GE, Teasdale RD and Koopman P. (2002). Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. Dev Cell. 3(2):167-170.

Schmiegelow K, Forestier E, Hellebostad M, Heyman M, Kristinsson J, Söderhäll S and Taskinen M; Nordic Society of Paediatric Haematology and Oncology. (2010). Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. *Leukemia*. 24(2):345-354.

Schorpp M, Bialecki M, Diekhoff D, Walderich B, Odenthal J, Maischein HM, Zapata AG and Boehm T. (2006). Conserved functions of Ikaros in vertebrate lymphocyte development: genetic evidence for distinct larval and adult phases of T cell development and two lineages of B cells in zebrafish. J Immunol. 177(4):2463-2476.

Schultz KR, Bowman WP, Aledo A, Slayton WB, Sather H, Devidas M, Wang C, Davies SM, Gaynon PS, Trigg M, Rutledge R, Burden L, Jorstad D, Carroll A, Heerema NA, Winick N, Borowitz MJ, Hunger SP, Carroll WL and Camitta B. (2009). Improved early event-free survival with imatinib in Philadelphia chromosome-positive acute lymphoblastic leukemia: a children's oncology group study. J Clin Oncol. 27(31):5175-5181.

Schäfer D, Olsen M, Lähnemann D, Stanulla M, Slany R, Schmiegelow K, Borkhardt A and Fischer U. (2018). Five percent of healthy newborns have an ETV6-RUNX1 fusion as revealed by DNA-based GIPFEL screening. Blood. 131(7):821-826.

Scott E W, Simon MC, Anastasi J and Singh H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science 265(5178):1573-1577.

Serafin V, Capuzzo G, Milani G, Minuzzo SA, Pinazza M, Bortolozzi R, Bresolin S, Porcù E, Frasson C, Indraccolo S, Basso G and Accordi B. (2017). Glucocorticoid resistance is reverted by LCK inhibition in pediatric T-cell acute lymphoblastic leukemia. Blood. 130(25):2750-2761.

Shah S, Schrader KA, Waanders E, Timms AE, Vijai J, Miething C, Wechsler J, Yang J, Hayes J, Klein RJ, Zhang J, Wei L, Wu G, Rusch M, Nagahawatte P, Ma J, Chen SC, Song G, Cheng J, Meyers P, Bhojwani D, Jhanwar S, Maslak P, Fleisher M, Littman J, Offit L, Rau-Murthy R, Fleischut MH, Corines M, Murali R, Gao X, Manschreck C, Kitzing T, Murty VV, Raimondi S, Kuiper RP, Simons A, Schiffman JD, Onel K, Plon SE, Wheeler D, Ritter D, Ziegler DS, Tucker K, Sutton R, Chenevix-Trench G, Li J, Huntsman DG, Hansford S, Senz J, Walsh T, Lee M, Hahn CN, Roberts K, King MC, Lo SM, Levine RL, Viale A, Socci ND, Nathanson KL, Scott HS, Daly M, Lipkin SM, Lowe SW, Downing JR, Altshuler D30, Sandlund JT, Horwitz MS, Mullighan CG and Offit K. (2013). A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia. Nat Genet. 45(10):1226-1231.

Shahjahani M, Norozi F, Ahmadzadeh A, Shahrabi S, Tavakoli F, Asnafi AA and Saki N. (2015). The role of Pax5 in leukemia: diagnosis and prognosis significance. Med Oncol. 32(1):360.

Shima Y and Kitabayashi I. (2011). Deregulated transcription factors in leukemia. Int J Hematol. 94(2):134-141.

Silverstein RA and Ekwall K. (2005) Sin3: a flexible regulator of global gene expression and genome stability. Curr Genet. 47(1):1-17.

Siveen KS, Prabhu KS, Achkar IW, Kuttikrishnan S, Shyam S, Khan AQ, Merhi M, Dermime S, Uddin S. (2018). Role of Non Receptor Tyrosine Kinases in

Hematological Malignances and its Targeting by Natural Products. Mol Cancer. 17(1): 31.
Slattery C, Ryan MP and McMorrow T. (2008). E2A proteins: regulators of cell phenotype in normal physiology and disease. Int J Biochem Cell Biol. 40(8):1431-1436.
Smyth GK. (2005). Limma: Linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Huber W, Irizarry R, editors. Bioinformatics and computational biology solutions using R and bioconductor. NY: Springer. 397–420.
Sock E, Rettig SD, Enderich J, Bösl MR, Tamm ER and Wegner M. (2004). Gene targeting reveals a widespread role for the high-mobility-group transcription factor Sox11 in tissue remodeling. Mol Cell Biol. 24(15):6635-6644.
Song WJ, Sullivan MG, Legare RD, Hutchings S, Tan X, Kufrin D, Ratajczak J, Resende IC, Haworth C, Hock R, Loh M, Felix C, Roy DC, Busque L, Kurnit D, Willman C, Gewirtz AM, Speck NA, Bushweller JH, Li FP, Gardiner K, Poncz M, Maris JM and Gilliland DG. (1999). Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. Nat Genet. 23(2):166-175.
Sood R, Kamikubo Y and Liu P. (2017). Role of RUNX1 in hematological malignancies. Blood. 129(15):2070-2082.
Standing ASI, Malinova D, Hong Y, Record J, Moulding D, Blundell MP, Nowak K, Jones H, Omoyinmi E, Gilmour KC, Medlar A, Stanescu H, Kleta R, Anderson G, Nanthapisal S, Melo Gomes S, Klein N, Eleftheriou D, Thrasher AJ and Brogan PA. (2017). Autoinflammatory periodic fever, immunodeficiency, and thrombocytopenia (PFIT) caused by mutation in actin-regulatory gene WDR1. *J Exp Med.* 214: 59-71.
Starkovich M, Lalani SR, Mercer CL and Scott DA. (2016). Chromosome 5q33 deletions associated with congenital heart defects. Am J Med Genet A. 170(12):3338-3342.
Stephens BJ, Han H, Gokhale V and Von Hoff DD. (2005). PRL phosphatases as potential molecular targets in cancer. Mol Cancer Ther. 4(11):1653-1661.
Sun X, Han Q, Luo H, Pan X, Ji Y, Yang Y, Chen H, Wang F, Lai W, Guan X, Zhang Q, Tang Y, Chu J, Yu J, Shou W, Deng Y and Li X. (2017). Profiling analysis of long non-coding RNAs in early postnatal mouse hearts. Sci Rep. 7:43485.
Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J (Eds): WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (Revised 4th Edition) IARC: Lyon 2017.
Szostakowska M, Szymczyk M, Badowska K, Tudek B and Fabisiewicz A. (2018). SOX11 expression as a MRD molecular marker for MCL in comparison with t(11;14) and IGH rearrangement. Med Oncol. 35(4):49.
Talab F, Allen JC, Thompson V, Lin K and Slupsky JR. (2013). LCK is an important mediator of B-cell receptor signaling in chronic lymphocytic leukemia cells. Mol Cancer Res. 11(5):541-554.
Tang R, Dodd A, Lai D, McNabb WC and Love DR. (2007). Validation of zebrafish (Danio rerio) reference genes for quantitative real-time RT-PCR normalization. Acta Biochim Biophys Sin (Shanghai). 39(5):384-390.
Targoff KL, Schell T and Yelon D. (2008). Nkx genes regulate heart tube extension and exert differential effects on ventricular and atrial cell number. Dev Biol. 322(2):314-321.
Targoff KL, Schell T and Yelon D. (2008). Nkx genes regulate heart tube extension and exert differential effects on ventricular and atrial cell number. Dev Biol. 322(2):314-321.

Tasian SK and Hunger SP. (2017) Genomic characterization of paediatric acute lymphoblastic leukaemia: an opportunity for precision medicine therapeutics. Br J Haematol. 176(6):867-882.

Tasian SK, Loh ML and Hunger SP1. (2015). Childhood acute lymphoblastic leukemia: Integrating genomics into therapy. Cancer. 121(20):3577-3590.

Teittinen KJ, Grönroos T, Parikka M, Junttila S, Uusimäki A, Laiho A, Korkeamäki H, Kurppa K, Turpeinen H, Pesu M, Gyenesei A, Rämet M and Lohi O. (2012). SAP30L (Sin3A-associated protein 30-like) is involved in regulation of cardiac development and hematopoiesis in zebrafish embryos. J Cell Biochem. 113(12):3843-3852.

Teppo S, Laukkanen S, Liuksiala T, Nordlund J, Oittinen M, Teittinen K, Grönroos T, St-Onge P, Sinnett D, Syvänen AC, Nykter M, Viiri K, Heinäniemi M and Lohi O. (2016). Genome-wide repression of eRNA and target gene loci by the ETV6-RUNX1 fusion in acute leukemia. Genome Res. 26(11):1468-1477.

Terwilliger T and Abdul-Hay M. (2017). Acute lymphoblastic leukemia: a comprehensive review and 2017 update. Blood Cancer J. 7(6):e577.

Tesio M, Trinquand A, Macintyre E and Asnafi V. (2016). Oncogenic PTEN functions and models in T-cell malignancies. Oncogene. 35(30):3887-3896.

Thisse C and Thisse B. (2008). High-resolution in situ hybridization to wholemount zebrafish embryos. Nat Protoc. 3(1):59-69.

Thomas DA, Faderl S, Cortes J, O'Brien S, Giles FJ, Kornblau SM, Garcia-Manero G, Keating MJ, Andreeff M, Jeha S, Beran M, Verstovsek S, Pierce S, Letvak L, Salvado A, Champlin R, Talpaz M and Kantarjian H. (2004). Treatment of Philadelphia chromosome-positive acute lymphocytic leukemia with hyper-CVAD and imatinib mesylate. Blood. 103(12):4396-4407.

Thorsteinsdottir U, Krosl J, Kroon E, Haman A, Hoang T and Sauvageau G. (1999). The oncoprotein E2A-Pbx1a collaborates with Hoxa9 to acutely transform primary bone marrow cells. Mol Cell Biol. 19(9):6355-6366.

Toft N, Birgens H, Abrahamsson J, Grišķevičius L, Hallböök H, Heyman M, Klausen TW, Jónsson ÓG, Palk K, Pruunsild K, Quist-Paulsen P, Vaitkeviciene G, Vettenranta K, Åsberg A, Frandsen TL, Marquart HV, Madsen HO, Norén-Nyström U and Schmiegelow K. (2018). Results of NOPHO ALL2008 treatment for patients aged 1-45 years with acute lymphoblastic leukemia. Leukemia. 32(3):606-615.

Tse W, Zhu W, Chen HS and Cohen A. (1995). A novel gene, AF1q, fused to MLL in t(1;11)(q21;q23), is specifically expressed in leukemic and immature hematopoietic cells. *Blood.* 85(3):650-656.

Tzoneva G, Perez-Garcia A, Carpenter Z, Khiabanian H, Tosello V, Allegretta M, Paietta E, Racevskis J, Rowe JM, Tallman MS, Paganin M, Basso G, Hof J, Kirschner-Schwabe R, Palomero T, Rabadan R and Ferrando A. (2013). Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. Nat Med. 19(3):368-371.

Uusi-Mäkelä MIE, Barker HR, Bäuerlein CA, Häkkinen T, Nykter M and Rämet M. (2018). Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (Danio rerio). PLoS One. 13(4):e0196238.

Vacaru AM, Unlu G, Spitzner M, Mione M, Knapik EW and Sadler KC. (2014). In vivo cell biology in zebrafish - providing insights into vertebrate development and disease. J Cell Sci. 127(Pt 3):485-495.

Valge-Archer VE, Osada H, Warren AJ, Forster A, Li J, Baer R and Rabbitts TH. (1994). The LIM protein RBTN2 and the basic helix-loop-helix protein TAL1 are present in a complex in erythroid cells. Proc Natl Acad Sci U S A. 91(18):8617-8621.

van de Wetering M, Oosterwegel M, van Norren K and Clevers H. (1993). Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. EMBO J. 12(10):3847-3854.

Van der Maaten L and Hinton G. (2008). Visualizing Data Using t-SNE. J Mach Learn Res. 9:2579-2605.

Van Vlierberghe P and Ferrando A. (2012). The molecular basis of T cell acute lymphoblastic leukemia. J Clin Invest. 122(10):3398-3406.

Van Vlierberghe P, van Grotel M, Beverloo HB, Lee C, Helgason T, Buijs-Gladdines J, Passier M, van Wering ER, Veerman AJ, Kamps WA, Meijerink JP and Pieters R. (2006). The cryptic chromosomal deletion del(11)(p12p13) as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia. Blood. 108(10):3520-3529.

Van Vlierberghe P, van Grotel M, Tchinda J, Lee C, Beverloo HB, van der Spek PJ, Stubbs A, Cools J, Nagata K, Fornerod M, Buijs-Gladdines J, Horstmann M, van Wering ER, Soulier J, Pieters R and Meijerink JP. (2008). The recurrent SET-NUP214 fusion as a new HOXA activation mechanism in pediatric T-cell acute lymphoblastic leukemia. Blood. 111(9):4668-4680.

Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, Harris NL, Le Beau MM, Hellström-Lindberg E, Tefferi A and Bloomfield CD. (2009). The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 114(5):937-951.

Vegliante MC, Palomero J, Pérez-Galán P, Roué G, Castellano G, Navarro A, Clot G, Moros A, Suárez-Cisneros H, Beà S, Hernández L, Enjuanes A, Jares P, Villamor N, Colomer D, Martín-Subero JI, Campo E and Amador V. (2013). SOX11 regulates PAX5 expression and blocks terminal B-cell differentiation in aggressive mantle cell lymphoma. *Blood.* 121(12):2175-2185.

Vegliante MC, Royo C, Palomero J, Salaverria I, Balint B, Martín-Guerrero I, Agirre X, Lujambio A, Richter J, Xargay-Torrent S, Bea S, Hernandez L, Enjuanes A, Calasanz MJ, Rosenwald A, Ott G, Roman-Gomez J, Prosper F, Esteller M, Jares P, Siebert R, Campo E, Martín-Subero JI and Amador V. (2011). Epigenetic activation of SOX11 in lymphoid neoplasms by histone modifications. PLoS One. 6(6):e21382.

Viiri KM, Heinonen TY, Maki M and Lohi O. (2009a). Phylogenetic analysis of the SAP30 family of transcriptional regulators reveals functional divergence in the domain that binds the nuclear matrix. BMC Evol Biol. 9:149.

Viiri KM, Janis J, Siggers T, Heinonen TY, Valjakka J, Bulyk ML, Maki M and Lohi O. (2009b). DNA-binding and -bending activities of SAP30L and SAP30 are mediated by a zinc-dependent module and monophosphoinositides. Mol Cell Biol. 29(2):342-356.

Viiri KM, Korkeamaki H, Kukkonen MK, Nieminen LK, Lindfors K, Peterson P, Maki M, Kainulainen H and Lohi O. (2006). SAP30L interacts with members of the Sin3A corepressor complex and targets Sin3A to the nucleolus. Nucleic Acids Res. 34(11):3288-3298.

Wang D, Zhu G, Wang N, Zhou X, Yang Y, Zhou S, Xiong J, He J, Jiang L, Li C, Xu D, Huang L and Zhou J. (2013). SIL-TAL1 rearrangement is related with poor outcome: a study from a Chinese institution. PLoS One. 8(9):e73865.

Wang LC, Kuo F, Fujiwara Y, Gilliland DG, Golub TR and Orkin SH. (1997). Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. EMBO J. 16(14):4374-4383.

Wang Q, Stacy T, Binder M, Marin-Padilla M, Sharpe AH and Speck NA. (1996). Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. Proc Natl Acad Sci USA. 93(8):3444-3449.

Wang X, Asplund AC, Porwit A, Flygare J, Smith CI, Christensson B and Sander B. (2008). The subcellular Sox11 distribution pattern identifies subsets of mantle cell lymphoma: correlation to overall survival. Br J Haematol. 143(2):248-252.

Wang X, Björklund S, Wasik AM, Grandien A, Andersson P, Kimby E, Dahlman-Wright K, Zhao C, Christensson B and Sander B. (2010). Gene expression profiling and chromatin immunoprecipitation identify DBN1, SETMAR and HIG2 as direct targets of SOX11 in mantle cell lymphoma. PLoS One. 5(11):e14085.

Wang Z, Wu Z, Liu Y and Han W. (2017). New development in CAR-T cell therapy. J Hematol Oncol. 10(1):53.

Weigle B, Ebner R, Temme A, Schwind S, Schmitz M, Kiessling A, Rieger MA, Schackert G, Schackert HK and Rieber EP. (2005). Highly specific overexpression of the transcription factor SOX11 in human malignant gliomas. Oncol Rep. 13(1):139-144.

Weng AP, Ferrando AA, Lee W, Morris JP 4th, Silverman LB, Sanchez-Irizarry C, Blacklow SC, Look AT and Aster JC. (2004). Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science. 306(5694):269-271.

Westerfield M. (1995). The zebrafish book: A guide to the laboratory use of zebrafish. Eugene OR: University of Oregon Press.

Whatley SD, Ducamp S, Gouya L, Grandchamp B, Beaumont C, Badminton MN, Elder GH, Holme SA, Anstey AV, Parker M, Corrigall AV, Meissner PN, Hift RJ, Marsden JT, Ma Y, Mieli-Vergani G, Deybach JC and Puy H. (2008). C-terminal deletions in the ALAS2 gene lead to gain of function and cause X-linked dominant protoporphyria without anemia or iron overload. Am J Hum Genet. 83(3):408-414.

Wilson NK, Foster SD, Wang X, Knezevic K, Schütte J, Kaimakis P, Chilarska PM, Kinston S, Ouwehand WH, Dzierzak E, Pimanda JE, de Bruijn MF and 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.

Winter J, Jung S, Keller S, Gregory RI and Diederichs S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol. 11(3):228-234.

Wu H, Deng Y, Feng Y, Long D, Ma K, Wang X, Zhao M, Lu L and Lu Q1. (2018). Epigenetic regulation in B-cell maturation and its dysregulation in autoimmunity. Cell Mol Immunol. 15(7):676-684.

Wu L, Xu Y, Wang Q, Ruan C, Drexler HG, Wu D, MacLeod RA, Chen S. (2015). High frequency of cryptic chromosomal rearrangements involving the LMO2 gene in T-cell acute lymphoblastic leukemia. Haematologica. 100(6):e233-236.

Yadav B, Pemovska T, Szwajda A, Kulesskiy E, Kontro M, Karjalainen R, Majumder MM, Malani D, Murumägi A, Knowles J, Porkka K, Heckman C, Kallioniemi O, Wennerberg K and Aittokallio T. (2014). Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. Sci Rep. 4:5193.

Yamada Y, Warren AJ, Dobson C, Forster A, Pannell R and Rabbitts TH. (1998). The T cell leukemia LIM protein Lmo2 is necessary for adult mouse hematopoiesis. Proc Natl Acad Sci U S A. 95:3890-3895.

Yamamoto JF and Goodman MT. (2008). Patterns of leukemia incidence in the United States by subtype and demographic characteristics, 1997-2002. Cancer Causes Control. 19(4):379-390.

Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, Raimondi SC, Relling MV, Patel A, Cheng C, Campana D, Wilkins D, Zhou X, Li J, Liu H, Pui CH, Evans WE, Naeve C and Wong L, Downing JR. (2002). Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell.* 1(2):133-143.

Yoo M, Shin J, Kim J, Ryall KA, Lee K, Lee S, Jeon M, Kang J and Tan AC. (2015). DSigDB: drug signatures database for gene set analysis. Bioinformatics. 31(18):3069-3071.

Zaliova M, Meyer C, Cario G, Vaskova M, Marschalek R, Stary J, Zuna J and Trka J. (2011). TEL/AML1-positive patients lacking TEL exon 5 resemble canonical TEL/AML1 cases. Pediatr Blood Cancer. 56(2):217-225.

Zelent A, Greaves M and Enver T. (2004). Role of the TEL-AML1 fusion gene in the molecular pathogenesis of childhood acute lymphoblastic leukaemia. Oncogene. 23(24):4275-4283.

Zeng W, Fu K, Quintanilla-Fend L, Lim M, Ondrejka S and Hsi ED. (2012). Cyclin D1-negative blastoid mantle cell lymphoma identified by SOX11 expression. Am J Surg Pathol. 36(2):214-219.

ZF-HEALTH: http://www.zf-health.org/information/factsheet.html (visited 24.7.2018)

ZFIN: https://zfin.org/zf_info/zfbook/stages/index.html (visited 24.7.2018)

Zhang J, Jackson AF, Naito T, Dose M, Seavitt J, Liu F, Heller EJ, Kashiwagi M, Yoshida T, Gounari F, Petrie HT and Georgopoulos K. (2011). Harnessing of the nucleosome-remodeling-deacetylase complex controls lymphocyte development and prevents leukemogenesis. Nat Immunol. 13(1):86-94.

Zhang J, McCastlain K, Yoshihara H, Xu B, Chang Y, Churchman ML, Wu G, Li Y, Wei L, Iacobucci I, Liu Y, Qu C, Wen J, Edmonson M, Payne-Turner D, Kaufmann KB, Takayanagi SI, Wienholds E, Waanders E, Ntziachristos P, Bakogianni S, Wang J, Aifantis I, Roberts KG, Ma J, Song G, Easton J, Mulder HL, Chen X, Newman S, Ma X, Rusch M, Gupta P, Boggs K, Vadodaria B, Dalton J, Liu Y, Valentine ML, Ding L, Lu C, Fulton RS, Fulton L, Tabib Y, Ochoa K, Devidas M, Pei D, Cheng C, Yang J, Evans WE, Relling MV, Pui CH, Jeha S, Harvey RC, Chen IL, Willman CL, Marcucci G, Bloomfield CD, Kohlschmidt J, Mrózek K, Paietta E, Tallman MS, Stock W, Foster MC, Racevskis J, Rowe JM, Luger S, Kornblau SM, Shurtleff SA, Raimondi SC, Mardis ER, Wilson RK, Dick JE, Hunger SP, Loh ML, Downing JR and Mullighan CG; St. Jude Children's Research Hospital–Washington University Pediatric Cancer Genome Project. (2016). Deregulation of DUX4 and ERG in acute lymphoblastic leukemia. Nat Genet. 48(12):1481-1489.

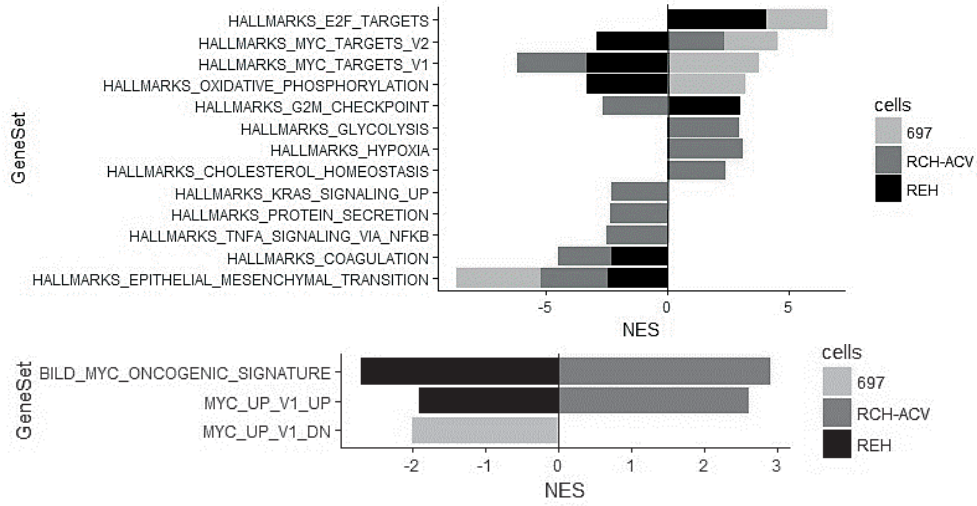
Zhang Y, Sun ZW, Iratni R, Erdjument-Bromage H, Tempst P, Hampsey M and Reinberg D. (1998). SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. Mol Cell. 1(7):1021-1031.

Zhou J, Bi C, Chng WJ, Cheong LL, Liu SC, Mahara S, Tay KG, Zeng Q, Li J, Guo K, Tan CP, Yu H, Albert DH and Chen CS. (2011). PRL-3, a metastasis associated tyrosine phosphatase, is involved in FLT3-ITD signaling and implicated in anti-AML therapy. PLoS One. 6(5):e19798.

Zhou J, Cheong LL, Liu SC, Chong PS, Mahara S, Bi C, Ong KO, Zeng Q and Chng WJ. (2012). The pro-metastasis tyrosine phosphatase, PRL-3 (PTP4A3), is a novel mediator of oncogenic function of BCR-ABL in human chronic myeloid leukemia. Mol Cancer. 11:72.

APPENDICES

Appendix 1. Supplementary information 4 from study II.



Supplementary information 4. A. Enrichment of hallmark gene sets retrieved from molecular signature database in REH, RCH-ACV and 697 cell lines. Adjusted p-value of enrichment for all gene sets < 0.01. B. Enrichment of founder gene sets of "MYC targets V2" gene set in REH, RCH-ACV and 697 cell lines. Adjusted p-value of enrichment for all gene sets < 0.02.

Appendix 2: 15 concordantly regulated genes in all three ALL cell lines (summary from Supplementary information 3 from study II).

	REH					697					RCH-ACV				
gene	logFC	logCPM	F	PValue	FDR	logFC	logCPM	F	PValue	FDR	logFC	logCPM	F	PValue	FDR
WDR1	-0,3407	8,4099	83,4624	4,98E-09	2,95E-06	-1,0331	6,6984	314,497	5,02E-09	5,48E-05	-0,6439	8,21727	177,463	9,58E-10	3,66E-06
CGGBP1	-0,3998	7,98295	72,3239	1,76E-08	8,16E-06	-0,7027	7,95881	150,267	1,88E-07	0,00041	-0,5597	7,81641	84,1519	1,47E-07	8,93E-05
WBP1L	-0,3955	7,81478	108,765	4,37E-10	6,32E-07	-0,7452	4,95497	95,007	1,66E-06	0,0014	-0,6636	5,09424	82,9899	1,61E-07	8,93E-05
MDK	-0,4053	8,10154	71,4734	1,95E-08	8,23E-06	-0,4793	7,69239	68,0679	7,68E-06	0,00419	-0,4254	7,41657	78,7754	2,25E-07	0,00011
WDTC1	-0,5848	5,89587	108,474	4,49E-10	6,32E-07	-1,141	5,88663	199,552	4,74E-08	0,00026	-0,6806	5,50482	73,4476	3,51E-07	0,00013
HNRNPA1P10	-0,2398	7,59992	35,8014	4,62E-06	0,00059	-0,5533	7,88301	126,665	4,28E-07	0,00058	-0,397	7,5856	71,0026	4,35E-07	0,00015
MLLT11	-0,3063	7,39516	51,5505	2,96E-07	7,27E-05	-0,3478	7,08044	35,2671	0,00013	0,03452	-0,376	7,4938	69,4908	4,98E-07	0,00016
ADAM10	-0,5884	7,54735	203,183	9,48E-13	1,07E-08	-0,6484	8,16369	128,758	3,95E-07	0,00058	-0,3947	8,0346	67,4845	5,98E-07	0,00017
HNRNPA1	-0,1907	8,22254	28,7393	2,05E-05	0,00178	-0,521	8,48788	129,34	3,87E-07	0,00058	-0,3358	8,161	51,3883	3,14E-06	0,00054
CANT1	-0,5202	5,5641	69,0745	2,62E-08	1,02E-05	-0,8958	5,22313	116,348	6,41E-07	0,00078	-0,573	5,59439	51,331	3,16E-06	0,00054
CSK	0,20797	7,35127	16,9347	0,00044	0,01534	0,40476	6,75399	32,0603	0,00019	0,04132	0,54493	5,65576	47,9497	4,72E-06	0,00067
LZIC	-0,4079	6,54633	60,8332	7,69E-08	2,62E-05	-0,5902	5,35301	51,946	2,54E-05	0,00992	-0,4922	5,56815	43,501	8,28E-06	0,001
SNX1	-0,1565	7,13668	12,5922	0,00176	0,04058	-0,404	7,1981	54,3197	2,09E-05	0,00879	-0,2676	7,3088	35,5648	2,54E-05	0,00206
CCDC28A	-0,7873	4,48833	59,4705	9,28E-08	2,68E-05	-1,1541	3,6776	46,2301	4,20E-05	0,01391	-0,7316	3,66147	34,8929	2,82E-05	0,00223
LITAF	0,31011	6,87898	42,001	1,44E-06	0,00024	0,9207	4,50898	53,228	2,29E-05	0,00925	0,45824	3,69564	16,7647	0,00095	0,02062

PUBLICATIONS

PUBLICATION I

Overexpression of PTP4A3 in ETV6-RUNX1 acute lymphoblastic leukemia

Grönroos T, Teppo S, Mehtonen J, Laukkanen S, Liuksiala T, Nykter M, Heinäniemi M, Lohi O.

Leukemia Research. 2017. 54:1-6. doi: 10.1016/j.leukres.2016.12.005. Epub 2016 Dec 26.

Leukemia Research 54 (2017) 1–6

Contents lists available at ScienceDirect

Leukemia Research

journal homepage: www.elsevier.com/locate/leukres

Overexpression of PTP4A3 in ETV6-RUNX1 acute lymphoblastic leukemia

Toni Grönroos[a,*], Susanna Teppo[a], Juha Mehtonen[b], Saara Laukkanen[a], Thomas Liuksiala[a,c], Matti Nykter[c,d], Merja Heinäniemi[b], Olli Lohi[a]

[a] Tampere Center for Child Health Research, University of Tampere and Tampere University Hospital, Tampere, Finland
[b] Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio, Finland
[c] Institute of Biosciences and Medical Technology, University of Tampere, Finland
[d] Department of Signal Processing, Tampere University of Technology, Tampere, Finland

ARTICLE INFO

Article history:
Received 3 August 2016
Received in revised form 25 October 2016
Accepted 18 December 2016
Available online 26 December 2016

Keywords:
ETV6-RUNX1
Leukemia
PTP4A3

ABSTRACT

Cell signalling, which is often derailed in cancer, is a network of multiple interconnected pathways with numerous feedback mechanisms. Dynamics of cell signalling is intimately regulated by addition and removal of phosphate groups by kinases and phosphatases. We examined expression of members of the PTP4A family of phosphatases across acute leukemias. While expression of *PTP4A1* and *PTP4A2* remained relatively unchanged across diseases, *PTP4A3* showed marked overexpression in ETV6-RUNX1 and BCR-ABL1 subtypes of precursor B cell acute lymphoblastic leukemia. We show that *PTP4A3* is regulated by the ETV6-RUNX1 fusion, but noticed no marked impact on cell viability either after *PTP4A3* silencing or treatment with a PTP4A3 inhibitor. Regulation of *PTP4A3* expression is altered in specific subgroups of acute leukemias and this is likely brought about by expression of the aberrant fusion genes.

1. Introduction

Acute leukemia is the most common cancer in childhood. *ETV6-RUNX1* (E/R) fusion gene is present in around 25% of cases of childhood precursor B cell acute lymphoblastic leukemia (pre-B-ALL) [1]. *E/R* fusion is generated by translocation between chromosomes 12 and 21 (t[12,21][p13;q22]) and creates an aberrant fusion transcription factor which induces repression of regulatory sequences and gene transcripts genome-wide [2]. Presence of *E/R* fusion is associated with good prognosis so that the overall survival exceeds 90% [3].

Adept regulation of cellular signalling is essential in controlling cellular proliferation. Kinases and phosphatases maintain a precise balance of intracellular signalling activity, and this balance is often deregulated in malignancies. The protein tyrosine phosphatase 4A (PTP4A) (also known as phosphatases of regenerating liver, PRL) are small, approximately 20 kDa proteins with dual specificity phosphatase activity [4]. The family consists of three members (PTP4A1-3) that share a high degree of sequence identity and harbour a phosphatase domain. Further, they contain a C-terminal consensus prenylation motif CaaX, which is important for localization to the plasma membrane and early endosomal compartments [4–7]. Phosphatases of PTP4A family participate in a wide range of cellular activities, including cell proliferation, migration and invasion [4].

In colorectal, breast, and gastric cancers, *PTP4A3* overexpression is correlated with poor prognosis and progression to metastasis [4]. In a murine model of colon carcinogenesis, *PTP4A3* knockout mice exhibited significantly fewer tumors than wild-type mice [8]. High expression of *PTP4A3* is associated with poor prognosis in acute myeloid leukemia (AML) with normal karyotype [9–11]. It is upregulated in human chronic myeloid leukemia (CML) cell lines, primary CML patient samples and Ph + ALL patients [12,13]. PTP4A3 is a downstream target of BCR-ABL1 fusion although the mechanism of action is unknown [13]. In multiple myeloma, *PTP4A3* is overexpressed and its knockdown inhibits cell migration [14]. Depending on expression level, *PTP4A3* may exhibit either positive or negative effects on cell cycle regulation [11].

Members of the PTP4A family of phosphatases have been associated with several hematological malignancies. This motivated us

Abbreviations: E/R, ETV6-RUNX1; pre-B-ALL, precursor B cell acute lymphoblastic leukemia; PTP4A, the protein tyrosine phosphatase 4A; PRL, the phosphatases of regenerating liver; AML, acute myeloid leukemia; CML, chronic myeloid leukemia; shRNA, short hairpin RNA; GRO-seq, global nuclear run-on sequencing.

* Corresponding author at: University of Tampere, Tampere Center for Child Health Research, P.O. Box 100, 33014 University of Tampere, Tampere, Finland.
E-mail address: toni.gronroos@staff.uta.fi (T. Grönroos).

http://dx.doi.org/10.1016/j.leukres.2016.12.005
0145-2126/

to investigate their expression across acute leukemias, and specifically the role of PTP4A3 in ALL.

2. Materials and methods

2.1. Cell lines and cell culture

Previously created stable cell line models with an inducible *E/R* fusion in NALM-6 cells and *E/R* short hairpin RNA (shRNA) knockdown in REH cells were used [2]. Cells were cultured in RPMI (Gibco, Thermo Fisher Scientific, Waltham, MA, USA), 2 mM L-glut, 100 U penicillin, 100 μg/ml streptomycin and 10% Tet System Approved FBS (Clontech) (NALM-6-Tet-cells) or 10% FBS (Gibco) (REH, NALM-6) at 37 °C in 5% CO_2. The expression of *E/R* was induced with 500 ng/ml doxycycline (Clontech). The shRNA knockdown and the induced expression of *E/R* were confirmed with RT-qPCR (Ssofast EvaGreen, BioRad, Hercules, CA, USA) with fusion gene specific primers (Supplemental Table 1).

2.2. Quantitative PCR

Total RNA extractions were performed using either GeneJET RNA Purification Kit (Thermo Fisher Scientific) and DNA-free™ DNase Treatment & Removal kit (BioRad), or using PureLink™ RNA Mini Kit with On-Column PureLink® DNase Treatment Protocol (Ambion® by Life Technologies). 1 μg of RNA was used as a starting material to perform cDNA synthesis with iScpript (BioRad). RT-qPCR reactions were performed with SsoFast EvaGreen® Supermix (BioRad) according to manufacturer's instructions. BioRad CFX96™ Real Time System (BioRad) was run with the following program: initial denaturation at 96 °C for 30 s, 39 cycles of denaturation at 96 °C for 2 s, annealing at 60 °C for 5 s, and plate read. RT-qPCR measurements were repeated from three independent experiments performed in triplicate, and the relative $2^{-\Delta\Delta C_T}$ method was used for quantification [15]. ChIP-qPCR was performed in duplicate runs from two independent biological replicates and the percent input method was used to normalize the data (2% of starting material as input). All qPCR primer sequences are listed in Supplemental Table 1. The level of *PTP4A3* mRNA expression between cell lines was evaluated with Mann-Whitney *U* test.

2.3. Chromatin immunoprecipitation, GRO-seq assay, and microarray data sets

Cells were crosslinked and DNA-protein interactions were immunoprecipitated using two pooled antibodies against ETV6 (sc-166835, Santa Cruz Biotechnology Inc., Dallas, TX, USA; RRID:AB_2101020 and HPA000264, Atlas Antibodies, RRID:AB_611466) as described in Teppo et al. [2]. Nuclei were extracted and the global nuclear run-on (GRO-seq) reaction was performed as described in Heinäniemi et al. [16]. To study differential gene expression, a previously combined and processed database containing microarray samples was utilized [16]. Two other data sets containing gene expression profiles in patients with pediatric ALL, GSE47051 [17], and data from Pediatric Cancer data portal (St. Jude Children's Research Hospital) [18], were retrieved. Differential expression of *PTP4A3* in distinct subgroups was evaluated with Kruskal-Wallis *H* test and Mann-Whitney *U* test.

2.4. Cell viability assays and inhibition of PTP4A3

Fresh media was changed on the REH and NALM-6 cells, and cells were plated at 1 million/ml on 6 well plates. To induce *E/R* or luciferase expression in NALM-6 cells, doxycycline (Clontech) was added at 500 ng/ml. After incubating plated cells for 24 h, 10 000 cells per well were plated in 96 wells with the desired concentration of PRL-3 inhibitor I (Sigma, Lot # 047K4614V) concentrations. The doxycycline treatment was continued in E/R induction samples during the entire experiment. Cells were allowed to grow up to 96 h. 10 μl Alamar Blue reagent (Life technologies) was added to each well at each time points. After 2 h, fluorescence was measured with excitation of 560 nm and emission of 590 nm using the Tecan fluorometer Infinite 200 (Tecan, Switzerland). Three technical replicates per sample were included in each proliferation experiment. Difference in cell viability between inhibitor cultured cells and control cells was evaluated with Student's *t*-test.

3. Results

3.1. High expression of PTP4A3 in the t(12;21) and t(9;22) subtypes of ALL

We investigated expression of members of PTP4A family of phosphatases across acute leukemias and healthy cells of hematopoietic origin. To this end, a large collection of previously combined and processed microarray samples was utilized [16]. Subgroup consisting of leukemic and normal healthy samples was selected and a "leukemia-map" was generated by using t-SNE algorithm to visualize the data in two dimensions [19], and expression of the *PTP4A* family members was examined. As shown in Fig. 1A, *PTP4A1* was expressed indifferently across diseases and healthy cells, whereas *PTP4A2* showed increased expression among some T-ALL samples. In contrast, *PTP4A3* showed markedly increased expression among pre-B-ALL and a few T-ALL cases (Fig. 1A). Strong expression of *PTP4A3* co-occurred mostly with the fusion positive subtypes of ALL, namely the E/R and BCR-ABL1/t(9;22) cases.

The expression of *PTP4A3* in the E/R subtype was approximately three to four-fold higher as compared to other pre-B-ALL leukemias (Fig. 1B), and this difference was statistically significant (Mann-Whitney *U* test p = 1.894997e-52 when E/R is compared to other subtypes, see Supplemental Table 2 for details). As further validation, we retrieved two more leukemia data sets (GSE47051 [17] and PeCan [18]), and *PTP4A3* showed similarly high expression among E/R and BCR-ABL1 positive leukemias in both data sets (Fig. 1B). At the same time, MLL-rearranged leukemias exhibited low level of *PTP4A3* across all three data sets.

3.2. PTP4A3 expression is regulated by E/R fusion

We next sought to determine if *E/R* fusion directly regulates *PTP4A3*. To this end, a panel of leukemia cell lines was screened by RT-qPCR for the expression of *PTP4A3* (Fig. 2A). REH cells, which endogenously carry the *E/R* fusion, exhibited strongest expression while KOPN-8, 697 and healthy T-cells and B-cells showed moderate expression. Another pre-B cell line NALM-6 evidenced low expression along with majority of the tested T-ALL lines. We generated two additional cell line models: a NALM-6 cell line with an inducible *E/R* fusion (NALM-6-E/R) and a REH cell line with knockdown of endogenous *E/R* fusion (REH-shE/R) [2]. As shown in Fig. 2B, induction of *E/R* in NALM-6 cells increased *PTP4A3* mRNA level. Knockdown of endogenous *E/R* fusion in REH cells did not significantly decrease expression of *PTP4A3* at the level of mature mRNA transcript (Fig. 2C). However, when the transcription was examined at primary transcript level, as assayed by GRO-seq, the decrease in signal intensity was more evident (Fig. 3A). When leukemia patient samples were profiled using the GRO-seq assay, *PTP4A3* showed robust expression in E/R cases and also to varying degrees among other subtypes (Fig. 3A).

Our data suggested that *PTP4A3* expression may be regulated by the E/R fusion itself, either directly or indirectly, and this moti-

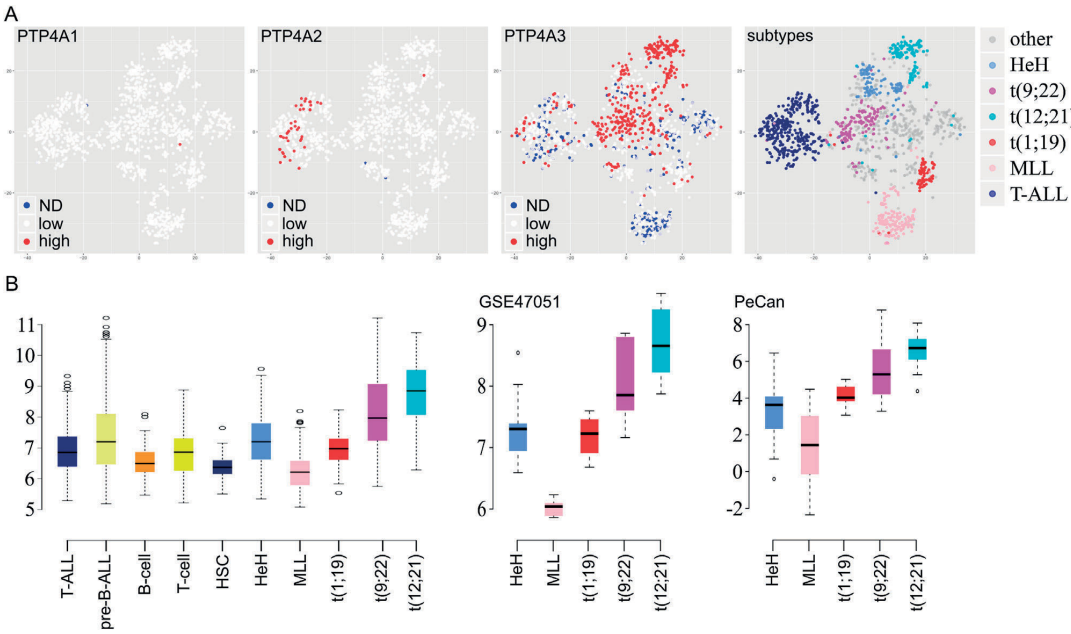


Fig. 1. *PTP4A3* is overexpressed in specific subtypes of precursor B-ALL.
A) Expression levels of the members of *PTP4A* family in leukemias is shown in a t-SNE-based two dimensional map. Expression level is indicated with colours: blue (ND = not detected), white (low expression), and red (high expression). Disease subtypes are indicated in the right-most figure with colours. HeH = high hyperdiploid. B) Boxplots of *PTP4A3* expression in various leukemias, disease subtypes, and normal cells from three independent data sources: Combined microarray data set [16], GEO series GSE47051 [17], and publicly available BCP-ALL data from Pediatric Cancer data portal [18]. For sample sizes see Supplemental Table 3.

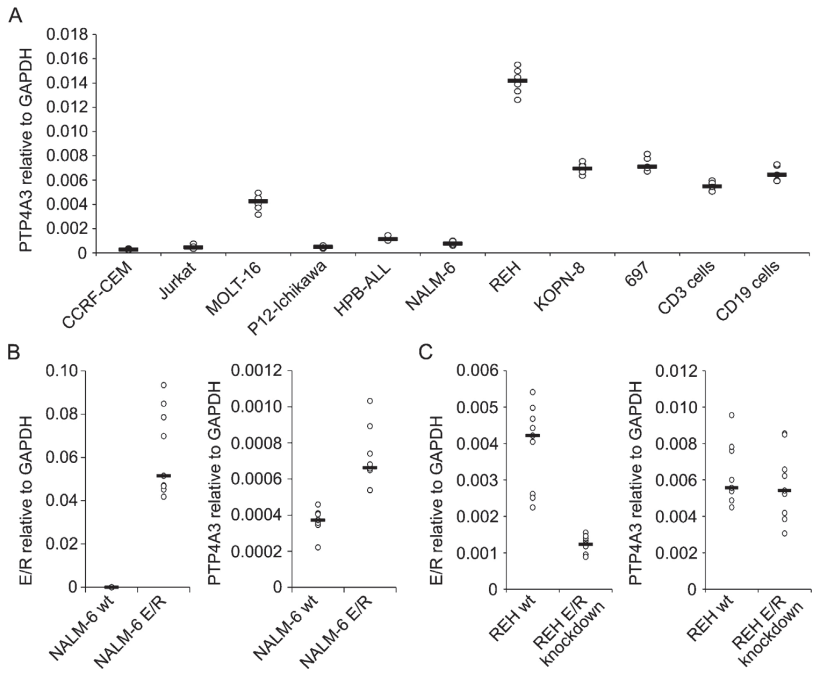


Fig. 2. *PTP4A3* expression is associated with expression of the E/R fusion.
A) Expression of *PTP4A3* in ALL cell lines and healthy B- and T-cells as measured by RT-qPCR. B) and C) *E/R* and *PTP4A3* expression in NALM-6 and REH cells with either induced overexpression (NALM-6-E/R) or knockdown (REH-shE/R) of E/R, respectively, as measured by RT-qPCR. Expression of inducible *E/R* fusion (NALM-6-E/R) increases expression of *PTP4A3* as compared to control cells (NALM-6-LUC). *) Mann-Whitney test, p-val. 0.00042. Expression level is shown as relative to a housekeeping gene (*GAPDH*).

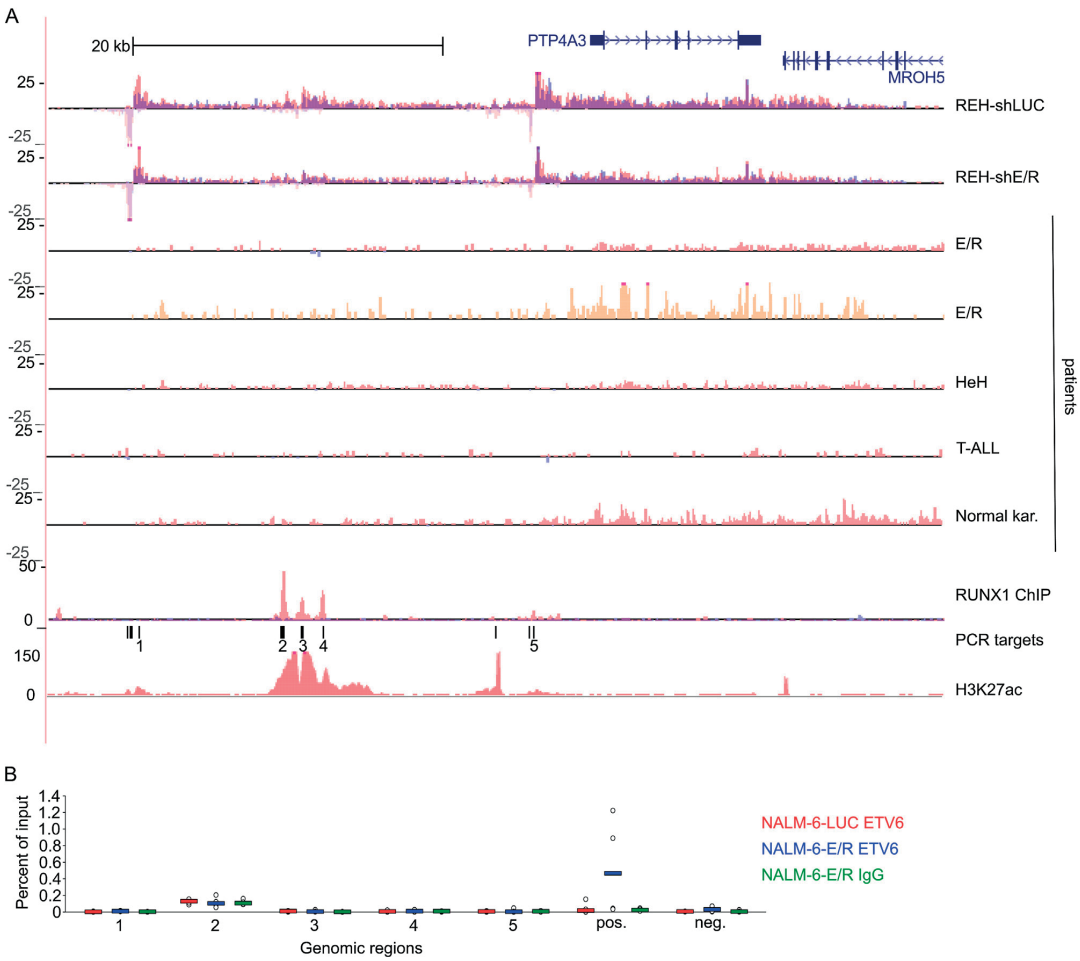


Fig. 3. GRO-seq indicates expression of *PTP4A3* transcript in patient samples.
A) GRO-seq signal tracks at *PTP4A3* locus. Two biological replicates of REH-shLUC and REH-shE/R samples are shown with different shades of color. Two E/R-positive patient samples, and one high hyperdiploid (HeH), T-ALL, and normal karyotype samples are shown. Signals above and below the axis indicate plus and minus strands, respectively. RUNX1 ChIP peaks in SEM cells (GSE42075 [20]) are shown in light red overlaid with the input control in purple. ChIP-seq track for H3K27ac in GM12878 B-cells indicates active sites and is retrieved from the ENCODE [25]. PCR regions tested in ChIP-qPCR are shown as strokes. B) No direct binding of E/R was detected by ChIP-qPCR at putative regulatory regions (1–5 as marked in A) upstream of *PTP4A3* after induction of *E/R* expression in NALM-6 cells. Negative antibody control (IgG-ChIP) was performed in NALM-6-E/R cells. The promoter region of *ITGA4* was used as a positive control and *GAPDH* as a negative control (see Supplemental Table 1 for primer sequences).

vated further analyses on regulatory regions. As E/R fusion binds its target sequences predominantly through RUNX1 sites [2], we retrieved RUNX1 ChIP-seq peaks in SEM (GSE42075 [20]) and HSC cells (GSE45144 [21]) and aligned them jointly with the GRO-seq data on cell lines and patient samples (Fig. 3A). Guided with this information, we performed targeted ChIP-qPCR assays using ETV6 antibodies, but failed to detect direct binding of E/R fusion into these regions (Fig. 3B).

3.3. Knockdown and inhibition of PTP4A3 in leukemia cells

To investigate the role of *PTP4A3* in leukemia, we knocked down its expression by transient transfection using three distinct siRNA oligos (see Supplemental methods). Since none of the tested antibodies against PTP4A3 performed reliably, we confirmed the knockdown by RT-qPCR, showing a decrease in expression by approximately 42% (SD 7.8%) in REH cells and 63% (SD 4.0%) in 697 cells after 48 h of transfection. In two biological replicates, knockdown of *PTP4A3* showed no effect on cell viability (data not shown).

PTP4A3 is a dual phosphatase whose enzymatic activity can be inhibited by a rhodamine derivative with an IC_{50} value of 0.9 μM [22]. We therefore examined the potency of this inhibitor to inhibit proliferation in leukemia cells. In an assay with 10-fold increases in drug concentrations, cell viability was reduced after concentrations exceeding 10 uM (Fig. 4). At these concentrations, we noticed a statistically significant difference between cells that have high or low expression of *E/R* fusion, possibly implying increased resistance among cells with high expression. We also tested potential synergy with conventional ALL induction chemotherapy (prednisone, doxorubicin and vincristine) but did not find additive effect (data not shown).

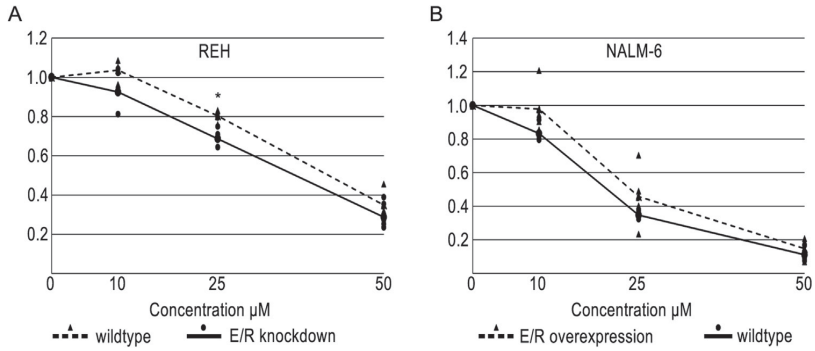


Fig. 4. Reduced cell viability at high concentrations of PTP4A3 inhibitor.
E/R positive cells are less sensitive to PTP4A3 inhibition as compared to control cells. Cell viability was assessed by Alamar Blue and is shown relative to cells treated with the vehicle control. *) Student's *t*-test, p-val. 0.00936.

4. Discussion

Intracellular cell signalling is tightly regulated network with multiple interconnected pathways and elaborate feedback systems. Addition and removal of phosphate groups by kinases and phosphatases, respectively, dynamically alters signalling intensity and direction, and are often derailed in malignancies with PTEN being the prime example [23]. Here we focused on members of the PTP4A family of phosphatases and examined their expression across acute leukemias. While expression of *PTP4A1* and *PTP4A2* remained relatively unchanged across diseases, *PTP4A3* showed marked overexpression in E/R and BCR-ABL1 subtypes of precursor B-ALL. We provide evidence of regulation of *PTP4A3* by the E/R fusion. The functional consequences of *PTP4A3* overexpression in ALL remain a subject of further studies since no marked impact on cell viability was evident either after *PTP4A3* silencing or treatment with a PTP4A3 inhibitor.

The expression results presented here are based on publicly available microarray data deposited in GEO. In order to diminish variation between various datasets, we utilized data that were generated using a single array platform [16]. It has been previously reported that *PTP4A3* is either downregulated [11] or upregulated in AML [10], and upregulated in CML and Ph+ ALL [9,11–13]. We found similarly high expression of *PTP4A3* in high expression of PTP4A3 in BCR-ABL1 fusion positive cases.

In healthy tissues, *PTP4A3* expression is primarily restricted to the heart, skeletal muscle, vasculature, and brain [24], suggesting that overexpression in hematological malignancies is driven by cancer-associated factors. This was already suggested by Zhou et al. [13] who identified *PTP4A3* as a downstream target of BCR-ABL1. We provide evidence that E/R fusion, which is a fusion of two hematopoietic transcription factors, regulates expression of *PTP4A3* either directly or indirectly. A RUNX1 ChIP data from SEM leukemia cells suggested direct binding. Our inability to prove direct binding of E/R fusion to these sites in NALM-6 cells may be due to low sensitivity of ChIP-qPCR assay. We tested potential binding sites only upstream of the transcription start site and since the location of regulatory elements may reside far away from the host gene in both directions, we may have lost the real site(s).

The expression data presented here are at the level of mRNA. As proteomic approaches develop, PTP4A3 protein levels and even enzymatic activity assays would be worth studying in hematological malignancies. Dual-specificity phosphatases can dephosphorylate both tyrosine and serine/threonine residues, and thereby inactivate targets. Several important signalling pathways are regulated by PTP4A enzymes, including Src/ERK1/2 and PTEN/PI3K/AKT [4]. We reasoned that if PTP4A3 played an essential role in leukemia pathogenesis, it would be evident in the cell viability assay. However, transient knockdown had no effect on cell viability and a specific inhibitor of PTP4A3 displayed only modest effect at high concentrations. This may imply that PTP4A3 does not play a central role in precursor-B ALL. Alternatively, the efficacy of inhibitor could be less than optimal as this class of enzymes is hard to effectively target [4]. Thirdly, it is possible that high expression is not accompanied by similarly increased enzymatic activity, or PTP4A3 could have function(s) independent of its enzymatic activity. Finally, since PTP4A3 has been shown to be important in cell adhesion, invasion, migration, and metastasis, functional studies should be directed towards these processes in future.

Acknowledgements

The work was supported by grants from the Academy of Finland (project no. 277816, OL; 276634, MH), The Foundation for Pediatric Research (OL), Jane and Aatos Erkko Foundation (OL), Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital (Grant numbers 9S032, OL), The Finnish Cultural Foundation (Interdisciplinary Science Workshops, MH), and The Väre Foundation for Pediatric Cancer Research (TG).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.leukres.2016.12.005.

References

[1] H. Inaba, M. Greaves, C.G. Mullighan, Acute lymphoblastic leukaemia, Lancet 381 (2013) 1943–1955, http://dx.doi.org/10.1016/S0140-6736(12)62187-4.
[2] S. Teppo, S. Laukkanen, T. Liuksiala, J. Nordlund, M. Oittinen, K. Teittinen, T. Grönroos, A.-C. Syvänen, M. Nykter, K. Viiri, M. Heinäniemi, O. Lohi, Genome-wide repression of eRNA and target gene loci by the ETV6-RUNX1 fusion in acute leukemia, Genome Res. 26 (11) (2016) 1468–1477.
[3] D. Bhojwani, D. Pei, J.T. Sandlund, S. Jeha, R.C. Ribeiro, J.E. Rubnitz, S.C. Raimondi, S. Shurtleff, M. Onciu, C. Cheng, E. Coustan-Smith, W.P. Bowman, S.C. Howard, M.L. Metzger, H. Inaba, W. Leung, W.E. Evans, D. Campana, M.V. Relling, C.H. Pui, ETV6-RUNX1-positive childhood acute lymphoblastic leukemia: improved outcome with contemporary therapy, Leukemia 26 (2012) 265–270, http://dx.doi.org/10.1038/leu.2011.227.
[4] A.M. Campbell, Z.Y. ZhangY, Phosphatase of regenerating liver: a novel target for cancer therapy, Expert Opin. Ther. Targets 18 (2014) 555–569, http://dx.doi.org/10.1517/14728222.2014.892926.
[5] B.J. Stephens, H. Han, V. Gokhale, D.D. Von Hoff, PRL: phosphatases as potential molecular targets in cancer, Mol. Cancer Ther. 4 (2005) 1653–1661.

[6] D.C. Bessette, D. Qiu, C.J. Pallen, PRL PTPs: mediators and markers of cancer progression, Cancer Metastasis Rev. 27 (2008) 231–252, http://dx.doi.org/10.1007/s10555-008-9121-3.
[7] P. Rios, X. Li, M. Köhn, Molecular mechanisms of the PRL phosphatases, FEBS J. 280 (2013) 505–524, http://dx.doi.org/10.1111/j.1742-4658.2012.08565.x.
[8] M.W. Zimmerman, G.E. Homanics, J.S. Lazo, Targeted deletion of the metastasis-associated phosphatase Ptp4a3 (PRL-3) suppresses murine colon cancer, PLoS One 8 (2013) e58300, http://dx.doi.org/10.1371/journal.pone.0058300.
[9] J. Zhou, C. Bi, W.J. Chng, L.L. Cheong, S.C. Liu, S. Mahara, K.G. Tay, Q. Zeng, J. Li, K. Guo, C.P. Tan, H. Yu, D.H. Albert, C.S. Chen, PRL-3, a metastasis associated tyrosine phosphatase, is involved in FLT3-ITD signaling and implicated in anti-AML therapy, PLoS One 6 (2011) e19798, http://dx.doi.org/10.1371/journal.pone.0019798.
[10] J.E. Park, H.F. Yuen, J.B. Zhou, A.Q. Al-Aidaroos, K. Guo, P.J. Valk, S.D. Zhang, W.J. Chng, C.W. Hong, K. Mills, Q. Zeng, Oncogenic roles of PRL-3 in FLT3-ITD induced acute myeloid leukaemia, EMBO Mol. Med. 5 (2013) 1351–1366, http://dx.doi.org/10.1002/emmm.201202183.
[11] R. Beekman, M. Valkhof, S.J. Erkeland, E. Taskesen, V. Rockova, J.K. Peeters, P.J. Valk, B. Löwenberg, I.Å. Touw, Retroviral integration mutagenesis in mice and comparative analysis in human AML identify reduced PTP4A3 expression as a prognostic indicator, PLoS One 6 (2011) e26537, http://dx.doi.org/10.1371/journal.pone.0026537.
[12] D. Juric, N.J. Lacayo, M.C. Ramsey, J. Racevskis, P.H. Wiernik, J.M. Rowe, A.H. Goldstone, P.J. O'Dwyer, E. Paietta, B.I. Sikic, Differential gene expression patterns and interaction networks in BCR-ABL-positive and -negative adult acute lymphoblastic leukemias, J. Clin. Oncol. 25 (2007) 1341–1349.
[13] J. Zhou, L.L. Cheong, S.C. Liu, P.S. Chong, S. Mahara, C. Bi, K.O. Ong, Q. Zeng, W.J. Chng, The pro-metastasis tyrosine phosphatase, PRL-3 (PTP4A3), is a novel mediator of oncogenic function of BCR-ABL in human chronic myeloid leukemia, Mol. Cancer 11 (2012) 72, http://dx.doi.org/10.1186/1476-4598-11-72.
[14] U.M. Fagerli, R.U. Holt, T. Holien, T.K. Vaatsveen, F. Zhan, K.W. Egeberg, B. Barlogie, A. Waage, H. Aarset, H.Y. Dai, J.D. Shaughnessy Jr., A. Sundan, M. Børset, Overexpression and involvement in migration by the metastasis-associated phosphatase PRL-3 in human myeloma cells, Blood 111 (2008) 806–815.
[15] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, Methods 4 (2001) 402–408.
[16] M. Heinäniemi, T. Vuorenmaa, S. Teppo, M.U. Kaikkonen, M. Bouvy-Liivrand, J. Mehtonen, H. Niskanen, V. Zachariadis, S. Laukkanen, T. Liuksiala, K. Teittinen, O. Lohi, Transcription-coupled genetic instability marks acute lymphoblastic leukemia structural variation hotspots, Elife 5 (2016), http://dx.doi.org/10.7554/elife.13087.
[17] J. Nordlund, C.L. Bäcklin, P. Wahlberg, S. Busche, E.C. Berglund, M.L. Eloranta, T. Flaegstad, E. Forestier, B.M. Frost, A. Harila-Saari, M. Heyman, O.G. Jónsson, R. Larsson, J. Palle, L. Rönnblom, K. Schmiegelow, D. Sinnett, S. Söderhäll, T. Pastinen, M.G. Gustafsson, G. Lönnerholm, A.-C. Syvänen, Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia, Genome Biol. 14 (2013) r105, http://dx.doi.org/10.1186/gb-2013-14-9-r105.
[18] https://pecan.stjude.org/home/ (Accessed 27 October 2015).
[19] L. van der Maaten, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (2008) 2579–2605.
[20] A.C. Wilkinson, E. Ballabio, H. Geng, P. North, M. Tapia, J. Kerry, D. Biswas, R.G. Roeder, C.D. Allis, A. Melnick, M.F. de Bruijn, T.A. Milne, RUNX1 is a key target in t(4;11) leukemias that contributes to gene activation through an AF4-MLL complex interaction, Cell Rep. 3 (2013) 116–127, http://dx.doi.org/10.1016/j.celrep.2012.12.016.
[21] D. Beck, J.A. Thoms, D. Perera, J. Schütte, A. Unnikrishnan, K. Knezevic, S.J. Kinston, N.K. Wilson, T.A. O'Brien, B. Göttgens, J.W. Wong, J.E. Pimanda, Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes, Blood 122 (2013) e12–e22, http://dx.doi.org/10.1182/blood-2013-03-490425.
[22] G. Min, S.K. Lee, H.N. Kim, Y.M. Han, R.H. Lee, D.G. Jeong, D.C. Han, B.M. Kwon, Rhodanine-based PRL-3 inhibitors blocked the migration and invasion of metastatic cancer cells, Bioorg. Med. Chem. Lett. 23 (2013) 3769–3774, http://dx.doi.org/10.1016/j.bmcl.2013.04.092.
[23] M. Tesio, A. Trinquand, E. Macintyre, V. Asnafi, Oncogenic PTEN functions and models in T-cell malignancies, Oncogene 35 (30) (2016) 3887–3896, http://dx.doi.org/10.1038/onc.2015.462.
[24] W.F. Matter, T. Estridge, C. Zhang, R. Belagaje, L. Stancato, J. Dixon, B. Johnson, L. Bloem, T. Pickard, M. Donaghue, S. Acton, R. Jeyaseelan, V. Kadambi, C.J. Vlahos, Role of PRL-3 a human muscle-specific tyrosine phosphatase, in angiotensin-II signaling, Biochem. Biophys. Res. Commun. 283 (2001) 1061–1068.
[25] ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome, Nature 489 (2012) 57–74, http://dx.doi.org/10.1038/nature11247.

PUBLICATION II

High expression of SOX11 is associated with favourable prognosis in childhood acute lymphoblastic leukemia

Grönroos T, Mäkinen A, Laukkanen S, Mehtonen J, Nikkilä A, Oksa L, Rounioja S, Marincevic-Zuniga Y, Nordlund J, Pohjolainen V, Paavonen T, Heinäniemi M, Lohi O.

Submitted manuscript

PUBLICATION III

In silico and preclinical drug screening identifies dasatinib as a targeted therapy for T-ALL

Laukkanen S, Grönroos T, Pölönen P, Kuusanmäki H, Mehtonen J, Cloos J, Ossenkoppele G, Gjertsen B, Øystein B, Heckman C, Heinäniemi M, Kontro M, Lohi O.

Blood Cancer Journal. 2017. 7(9):e604. doi: 10.1038/bcj.2017.87.

OPEN

Citation: Blood Cancer Journal (2017) **7**, e604; doi:10.1038/bcj.2017.87

www.nature.com/bcj

LETTER TO THE EDITOR

In silico and preclinical drug screening identifies dasatinib as a targeted therapy for T-ALL

Blood Cancer Journal (2017) **7,** e604; doi:10.1038/bcj.2017.87; published online 8 September 2017

Systematic development of combination chemotherapies has significantly improved the prognosis of acute lymphoblastic leukemia (ALL). Patients with T-cell ALL (T-ALL) still have less favorable outcomes, and the prognosis of relapsed T-ALL is dismal.[1] In order to find novel targeted therapies for T-ALL, we retrieved the gene expression profiles of 4430 leukemia samples from the Gene Expression Omnibus (GEO) and performed an *in silico* drug target screening, where the expression profiles were compared with known drug targets in the Drug signature database (DsigDB).[2,3] This compound library contains both FDA (Food and Drug Administration)–approved and novel investigational drugs. Strong expression of Src family tyrosine kinase *LCK* was detected in T-ALL samples, exceeding that of other leukemias and normal T cells (Figure 1a), whereas a LCK-targeting drug, dasatinib, decreased the kinase activity of LCK to 1% in comparison with the control at a 100 nM concentration.[4] Dasatinib is known to have multiple intracellular targets, and we noticed that some of them had a similarly elevated expression in T-ALL patients (Figure 1b and Supplementary Figure S1). Therefore, our combinatorial drug/target screening suggests dasatinib as a candidate targeted therapy for T-ALL patients.

To experimentally evaluate the potency of dasatinib, we treated T-ALL cell lines with increasing concentrations of dasatinib (1–1000 nM). After 72 h, the most significant response was seen in

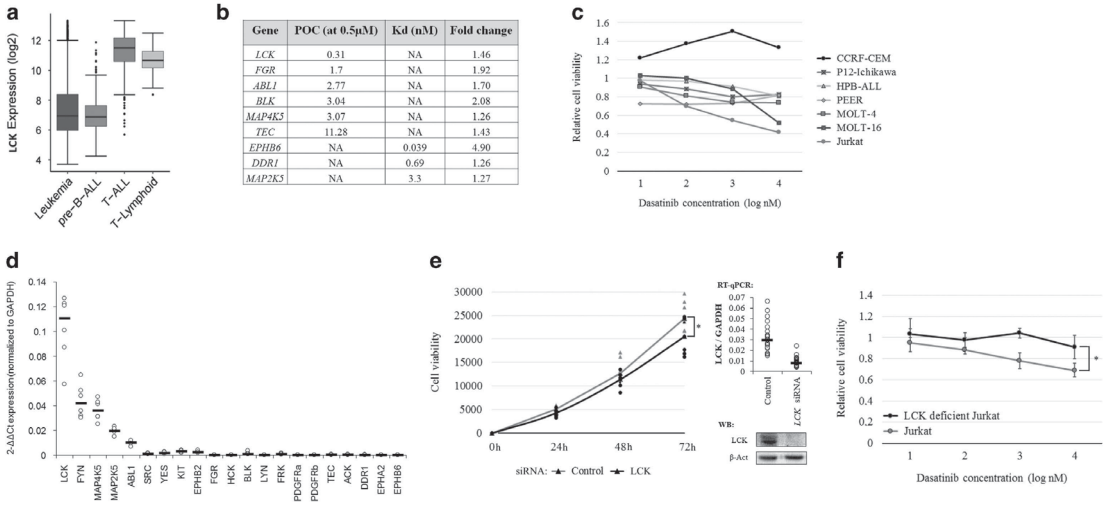


Figure 1. The *in silico* and *in vitro* results indicate dasatinib as a potential drug for T-ALL with *LCK* as its prime target. (**a**) *LCK* expression in different hematological sample groups: acute leukemia ($n = 4430$), pre-B-ALL ($n = 1304$), T-ALL ($n = 385$) and T lymphoids ($n = 247$). (**b**) The filtered list of targets of dasatinib from *in silico* screening. The list contains targets with a lower expression in normal cells (myeloid, B lymphoid and T lymphoid) in comparison with their leukemic counterparts; a significant expression difference between T-ALL and T-lymphoid samples (adjusted $P < 0.001$ and a >1.25-fold change) and dasatinib is capable of inhibiting them with high efficiency. Percentage of control (POC) or Kd values were used to estimate the efficacy of dasatinib against its targets. The POC value indicates the percentage of remaining activity after inhibitor treatment in comparison with the untreated control sample. In addition to 0.5 μM, POC values at 100 nM concentrations were also available for LCK and ABL1, and they were 1 and 0, respectively. (**c**) Effect of dasatinib for cell viability in several T-ALL cell lines measured by alamarBlue assay after 72 h of incubation in 10-fold dasatinib dilution series (1–1000 nM). Values are relative cell viabilites in comparison with dimethyl sulfoxide (DMSO) control and results are median values from three independent experiments performed in triplicate, except for CCRF-CEM and HPB-ALL that are from two independent experiments. (**d**) The expression of 21 potential dasatinib targets in Jurkat cells. Bars indicate median values. (**e**) The effect of *LCK* knockdown for Jurkat cell proliferation measured in time series (0h, 24h, 48 and 72 h) with alamarBlue assay. Proliferation trend lines are drawn through median values. At time point 72 h, the proliferation had decreased by 14% in comparison with the mock-treated control ($P = 0.0289$, Mann–Whitney *U*-test). The data consist of seven individual experiments performed in triplicate, and each time point is normalized to the 0 h time point. RT-qPCR and western blot results show the efficiency of *LCK* knockdown. (**f**) The effect of dasatinib on cell viability in the LCK-deficient Jurkat cell line in comparison with the normal Jurkat cell line measured by alamarBlue assay after 72 h of incubation in a 10-fold dasatinib dilution series (1–1000 nM). The difference between the two cell lines was statistically significant already at a 10 nM concentration ($P = 0.014$, Mann–Whitney *U*-test). The values are relative cell viabilities in comparison with the DMSO control, and the results are the median values from three independent experiments performed in triplicate. The error bars indicate 95% confidence intervals.

Jurkat cells, with a 31% viability decrease at a 10 nM concentration ($n=3$, $P=0.0039$; Figure 1c and Supplementary Figure S2). As dasatinib inhibits several kinases that are key regulators of cellular proliferation and viability, we chose a panel of likely candidates (9 targets based on the *in silico* screening and 12 other well-known targets from the literature), and analyzed their expression in T-ALL cell lines by quantitative reverse transcriptase-PCR (RT-qPCR) and western blotting. *LCK* was the most prominently expressed gene in T-ALL cell lines, whereas *FYN*, *ABL1*, *MAP2K5*, *MAP4K5* and *LYN* were expressed at lower levels (Figure 1d and Supplementary Figure S3). Knockdown of *LCK* in a dasatinib-sensitive cell line (Jurkat) significantly decreased cell proliferation (14% decrease, $P=0.0289$, $n=7$, Figure 1e), whereas knockdown of *FYN*, *ABL1*, *MAP2K5* and *MAP4K5* had no significant effect (Supplementary Figures S4a–d). Importantly, Jurkat cells with reduced LCK activity due to a deletion of exon 7 (cell line J.CaM1.6) lost dasatinib sensitivity (Figure 1f). Moreover, *LCK* knockdown did not cause statistically significant decrease of proliferation in relatively dasatinib-insensitive P12-Ichikawa cell line (Supplementary Figure S4e). These results suggest that LCK is the prime target of dasatinib in T-ALL.

We next performed *ex vivo* drug testing of 22 primary T-ALL samples. In 6 cases (27%), the response to dasatinib was significant based on drug sensitivity scores (DSS, using a cutoff value of 10, Figure 2a).[5] Half-maximal growth inhibition concentrations (IC_{50}) ranged between 1.3 and 16 nM, whereas the control samples had an IC_{50} of >1000 nM (Supplementary Figure S5). We also noted a negative correlation between dasatinib and glucocorticoid DSS scores (Supplementary Figure S6). Previously, dasatinib sensitivity has been reported in T-ALL cases with *NUP214-ABL1* fusion.[6–8] In contrast, none of the dasatinib responders in our sample set carried the fusion gene based on either genomic PCR or RNA-sequencing analysis (Supplementary Figure S7). *LCK* was strongly expressed in four out of the five dasatinib-responsive patient samples, whereas the expression of other potential targets varied from a low (*LYN*, *ABL1)* to medium level (*FYN*, *MAP2K5* and *MAP4K5*, Supplementary Figure S8a). As *LCK* was also relatively strongly expressed in dasatinib-insensitive patient samples, no correlation between dasatinib response and *LCK* expression was observed (Supplementary Figure S8b).

As T-ALL subgroups can be separated by expression of specific transcription factors, we explored whether dasatinib sensitivity was associated with any specific subgroups.[9] In addition to Jurkat cells that belong to the TAL1 subgroup, five out of six dasatinib-sensitive patient samples showed either prominent *TAL1* expression (Figure 2b) or carried *SIL-TAL1* fusion (data not shown). There was also increased expression of *LMO2* and *HOXA9/10* genes in some samples (Supplementary Figure S9). 6q deletions have also

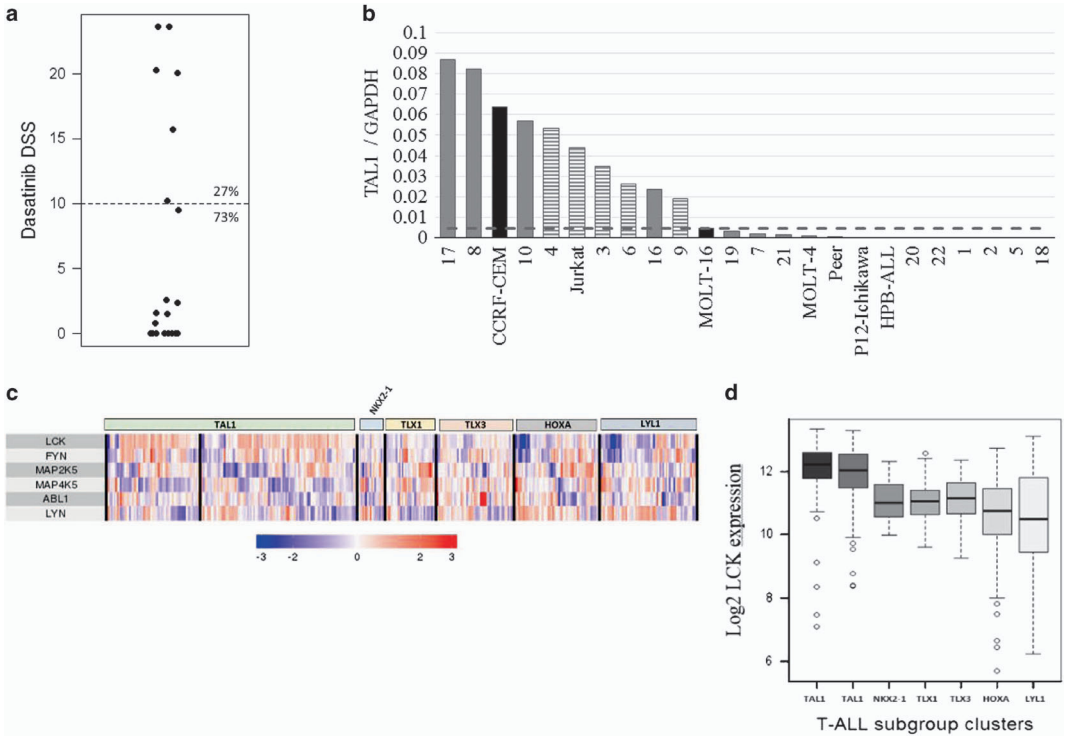


Figure 2. Dasatinib-sensitive subgroup of T-ALL samples. (**a**) Drug sensitivity scores (DSS) of dasatinib in a cohort of 22 patient samples. These DSS values are calculated from growth inhibition measurements after 72 h of treatment in a 10-fold dasatinib dilution series (0.1–1000 nM), and a DSS value of 10 was used as the threshold for dasatinib sensitivity. (**b**) The expression of the T-ALL subtype defining transcription factor *TAL1* in T-ALL patient samples and cell lines measured by RT-qPCR. The threshold for ectopic *TAL1* expression (dashed line) is defined by the expression of *TAL1* in SIL-TAL1 fusion-positive cell lines, indicated with black columns, and the striped columns represent the dasatinib-sensitive samples.[17] Patient sample 4 was processed in a separate RT-qPCR batch. (**c**) The heat map of *LCK*, *FYN*, *MAP2K5*, *MAP4K5*, *ABL1* and *LYN* expression in a separate GEO-based T-ALL sample data set ($n=385$). Samples are clustered based on the expression of T-ALL subtype-defining transcription factors: *TAL1* ($n=61+103$), *NKX2-1* ($n=18$), *TLX1* ($n=33$), *TLX3* ($n=51$), *HOXA* ($n=56$) and *LYL* ($n=63$). (**d**) *LCK* expression in T-ALL subtype clusters. The difference between TAL1 clusters and any other cluster was statistically significant ($P<0.001$, Mann–Whitney *U*-test).

been associated with the TAL1 subgroup, and three out of six samples carried the 6q deletion in cytogenetic analyses (Supplementary Table S1).[9] When *LCK* expression was correlated with the subgroup information in our GEO-based gene expression data set, a statistically significant correlation was found between *LCK* expression and the TAL1 subgroup (Figures 2c and d and Supplementary Figure S10). However, no direct transcriptional regulation of *LCK* by TAL1 was seen when the expression of *TAL1* was knocked down (Supplementary Figure S11, data reproduced from Sanda *et al.*[10]).

Taken together, the combination of *in silico*, *in vitro* and *ex vivo* data indicate that dasatinib exerts clinical utility beyond the originally suggested *NUP214-ABL1* cases that only represent 4–10% of T-ALL patients.[6,11] Our results are in line with the data by Frismantas *et al.*[12] who identified a slightly higher percentage of dasatinib-sensitive patients (30–40%) in their T-ALL cohorts. In addition to confirming their main findings, we expand on them in two different ways. First, we identified *LCK* tyrosine kinase as the potential prime target of dasatinib, whereas Frismantas *et al.*[12] did not recognize any recurrent gene fusions, mutations or transcriptome profiles associated with dasatinib sensitivity; rather, they associated *SRC* expression and phosphorylation status with dasatinib sensitivity. They supported their hypothesis by reporting a positive correlation between dasatinib and SRC inhibitor KX2-391 responses in T-ALL patient samples.[12] However, DsigDB identifies both SRC and LCK as targets of KX2-391. In our GEO-based gene expression data set and T-ALL cell lines, the expression of *SRC* was very low when compared with *LCK*, and there was no difference in the expression levels of T-ALL and normal T-cell samples (Supplementary Figure S1). A low level of *SRC* expression was also observed in the data set of Frismantas *et al.*[12] Importantly, the *LCK*-mutant derivative of the Jurkat cell line failed to respond to dasatinib, thus strongly suggesting the central role played by LCK itself. We failed to detect an association between the phosphorylation status of LCK kinase and dasatinib sensitivity in T-ALL cell lines (data not shown).

Second, we show here that patients belonging to the TAL1 subgroup are the most likely to respond to dasatinib, although not exclusively. A high expression of *LCK* associates with the TAL1 subgroup, but TAL1 does not directly regulate *LCK* expression. As up to 60% of T-ALL patients belong to the TAL1 subgroup and only 30–40% respond to dasatinib, there is clearly a need for more accurate biomarkers to be identified.

Dasatinib is known to suppress proliferation of healthy T cells by LCK inhibition.[13–15] We noticed a significantly higher expression of *LCK* in T-ALL samples compared with healthy T cells (a 1.46-fold change, $P < 0.001$), and a higher sensitivity of T-ALL cells toward dasatinib compared with healthy bone marrow cells. Although we identified LCK as the main target of dasatinib, our results do not exclude the contribution by other kinases, including other Src family kinases. Many of the target candidates function in receptor signaling pathways related to cell proliferation and survival. LCK and FYN are components of the T-cell receptor (TCR) signaling cascade.[16] During T-cell development, LCK is required for the normal development of thymocytes, whereas in mature T cells, FYN is capable of activating several TCR signaling pathways in the absence of LCK, including the Ras/extracellular signal-regulated kinase and phosphatidylinositol 3-kinase pathways.[16] In knockdown experiments, we observed reciprocal feedback mechanisms between LCK and FYN (data not shown), in agreement with their known partially overlapping functions.

Dasatinib is a tyrosine kinase inhibitor that is currently approved for imatinib-resistant Philadelphia chromosome-positive (Ph+) chronic myeloid leukemia and the second-line treatment of Ph+ ALL. We report here the potential utility of dasatinib in the treatment of a subset of T-ALL. As our patient cohort is relatively small, further studies are needed to explore the biomarker findings and to deepen the mechanistic basis before embarking on animal and human studies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

S Laukkanen[1], T Grönroos[1], P Pölönen[2], H Kuusanmäki[3], J Mehtonen[2], J Cloos[4], G Ossenkoppele[5], B Gjertsen[6,7], B Øystein[6,7], C Heckman[3], M Heinäniemi[2], M Kontro[8] and O Lohi[1]

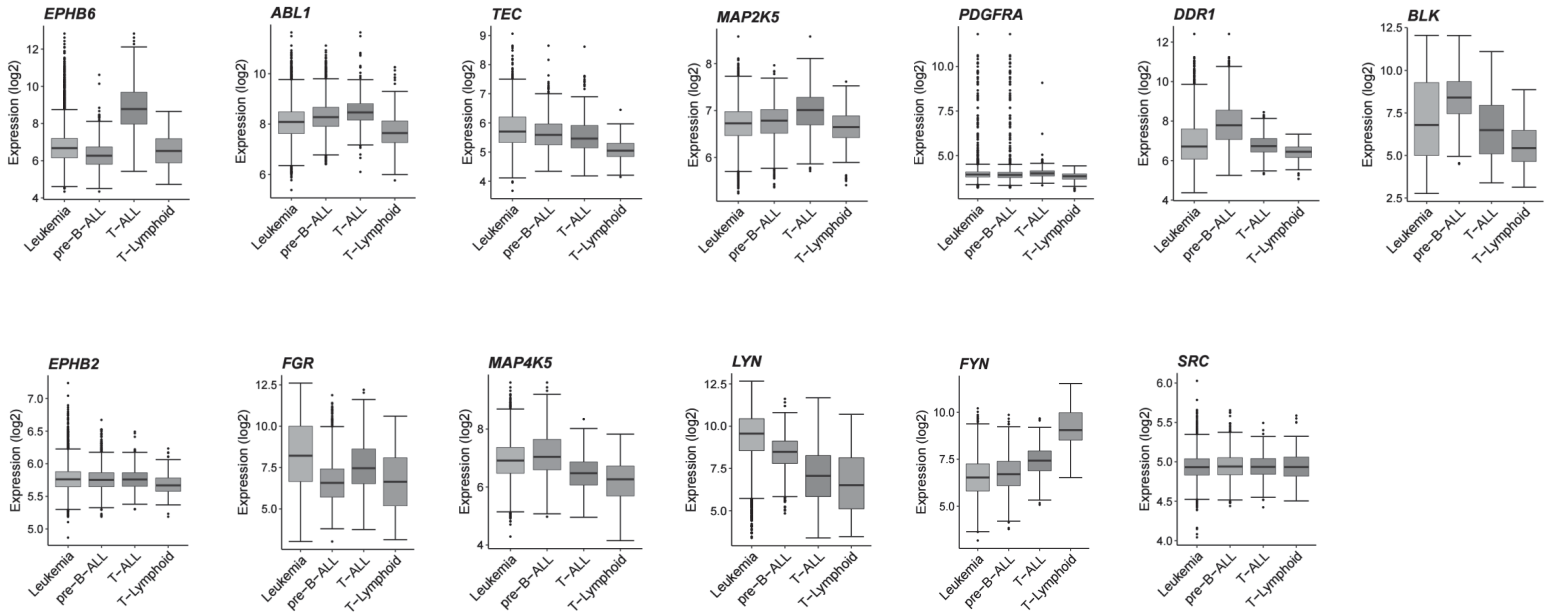
[1]*Tampere Center for Child Health Research, University of Tampere, and Tampere University Hospital, Tampere, Finland;*
[2]*Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio, Finland;*
[3]*Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland;*
[4]*Department of Pediatric Oncology/Hematology, VU University Medical Center, Amsterdam, The Netherlands;*
[5]*Department of Hematology, VU University Medical Center, Amsterdam, The Netherlands;*
[6]*Department of Clinical Science, Hematology Section, University of Bergen, Bergen, Norway;*
[7]*Department of Internal Medicine, Hematology Section, Haukeland University Hospital, Bergen, Norway and*
[8]*Department of Hematology, Hematology Research Unit Helsinki, University of Helsinki, and Helsinki University Central Hospital Cancer Center, Helsinki, Finland*
E-mail: saara.laukkanen@staff.uta.fi

REFERENCES

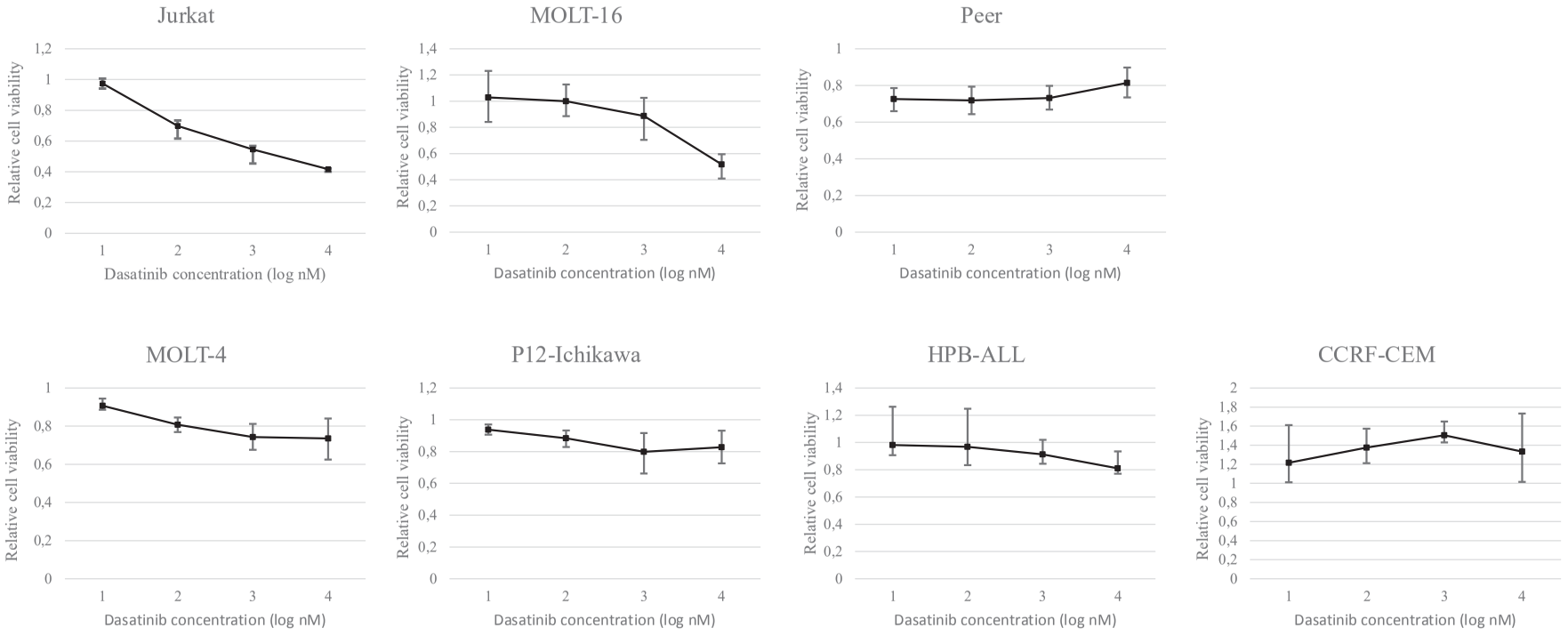
1 Goldberg JM, Silverman LB, Levy DE, Dalton VK, Gelber RD, Lehmann L *et al.* Childhood T-cell acute lymphoblastic leukemia: the Dana-Farber Cancer Institute acute lymphoblastic leukemia consortium experience. *J Clin Oncol* 2003; **21**: 3616–3622.
2 Liuksiala T, Teittinen KJ, Granberg K, Heinäniemi M, Annala M, Mäki M *et al.* Overexpression of SNORD114-3 marks acute promyelocytic leukemia. *Leukemia* 2014; **28**: 233–236.
3 Heinäniemi M, Vuorenmaa T, Teppo S, Kaikkonen MU, Bouvy-Liivrand M, Mehtonen J *et al.* Transcription-coupled genetic instability marks acute lymphoblastic leukemia structural variation hotspots. *eLife* 2016; **5**: e13087.
4 Yoo M, Shin J, Kim J, Ryall KA, Lee K, Lee S *et al.* DSigDB: drug signatures database for gene set analysis. *Bioinformatics* 2015; **31**: 3069–3071.
5 Yadav B, Pemovska T, Szwajda A, Kulesskiy E, Kontro M, Karjalainen R *et al.* Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. *Sci Rep* 2014; **4**: 5193.
6 Quintás-Cardama A, Tong W, Manshouri T, Vega F, Lennon PA, Cools J *et al.* Activity of tyrosine kinase inhibitors against human NUP214-ABL1-positive T cell malignancies. *Leukemia* 2008; **22**: 1117–1124.
7 Deenik W, Beverloo HB, van der Poel-van de Luytgaarde SC, Wattel MM, van Esser JW, Valk PJ *et al.* Rapid complete cytogenetic remission after upfront dasatinib monotherapy in a patient with a NUP214-ABL1-positive T-cell acute lymphoblastic leukemia. *Leukemia* 2009; **23**: 627–629.
8 Crombet O, Lastrapes K, Zieske A, Morales-Arias J. Complete morphologic and molecular remission after introduction of dasatinib in the treatment of a pediatric patient with t-cell acute lymphoblastic leukemia and ABL1 amplification. *Pediatr Blood Cancer* 2012; **59**: 333–334.
9 Girardi T, Vicente C, Cools J, De Keersmaecker K. The genetics and molecular biology of T-ALL. *Blood* 2017; **129**: 1113–1123.
10 Sanda T, Lawton LN, Barrasa MI, Fan ZP, Kohlhammer H, Gutierrez A *et al.* Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. *Cancer Cell* 2012; **22**: 209–221.
11 Burmeister T, Gökbuget N, Reinhardt R, Rieder H, Hoelzer D, Schwartz S. NUP214-ABL1 in adult T-ALL: the GMALL study group experience. *Blood* 2006; **108**: 3556–3559.
12 Frismantas V, Dobay MP, Rinaldi A, Tchinda J, Dunn SH, Kunz J *et al.* Ex vivo drug response profiling detects recurrent sensitivity patterns in drug-resistant acute lymphoblastic leukemia. *Blood* 2017; **129**: e26–e37.
13 Blake S, Hughes TP, Mayrhofer G, Lyons AB. The Src/ABL kinase inhibitor dasatinib (BMS-354825) inhibits function of normal human T-lymphocytes in vitro. *Clin Immunol* 2008; **127**: 330–339.

14 Lee KC, Ouwehand I, Giannini AL, Thomas NS, Dibb NJ, Bijlmakers MJ. Lck is a key target of imatinib and dasatinib in T-cell activation. *Leukemia* 2010; **24**: 896–900.
15 Schade AE, Schieven GL, Townsend R, Jankowska AM, Susulic V, Zhang R *et al.* Dasatinib, a small-molecule protein tyrosine kinase inhibitor, inhibits T-cell activation and proliferation. *Blood* 2008; **111**: 1366–1377.
16 Salmond RJ, Filby A, Qureshi I, Caserta S, Zamoyska R. T-cell receptor proximal signaling via the Src-family kinases, Lck and Fyn, influences T-cell activation, differentiation, and tolerance. *Immunol Rev* 2009; **228**: 9–22.
17 Nagel S, Venturini L, Meyer C, Kaufmann M, Scherr M, Drexler HG *et al.* Multiple mechanisms induce ectopic expression of LYL1 in subsets of T-ALL cell lines. *Leuk Res* 2010; **34**: 521–528.

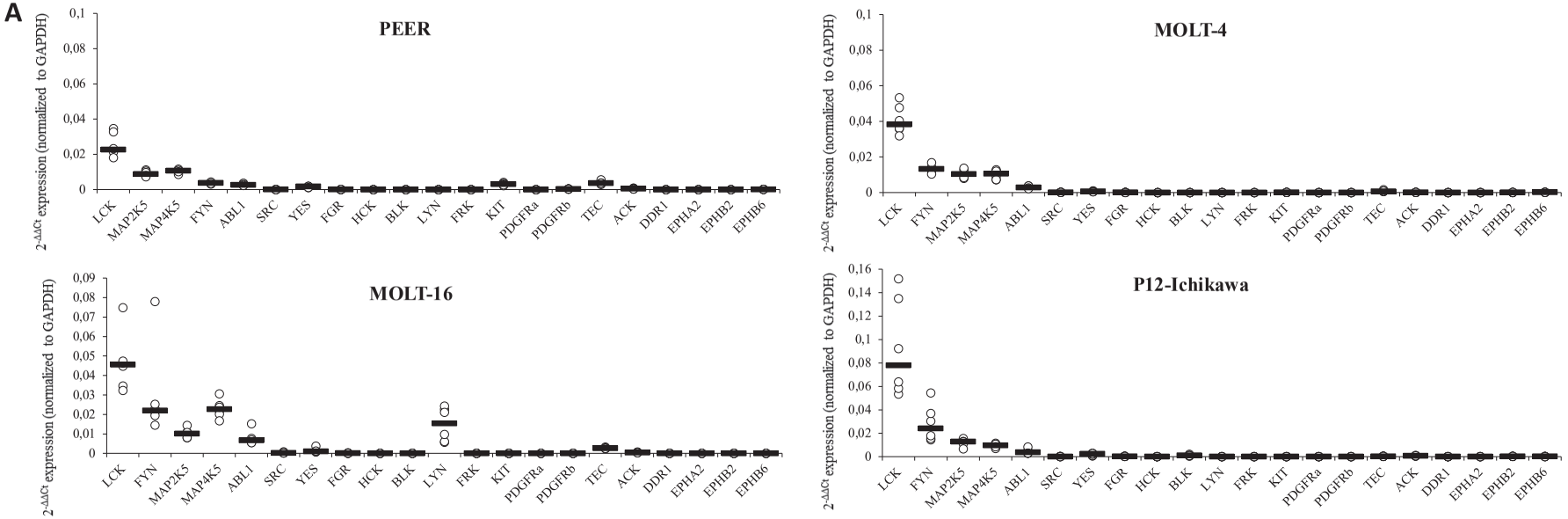
Supplementary Information accompanies this paper on Blood Cancer Journal website (http://www.nature.com/bcj).

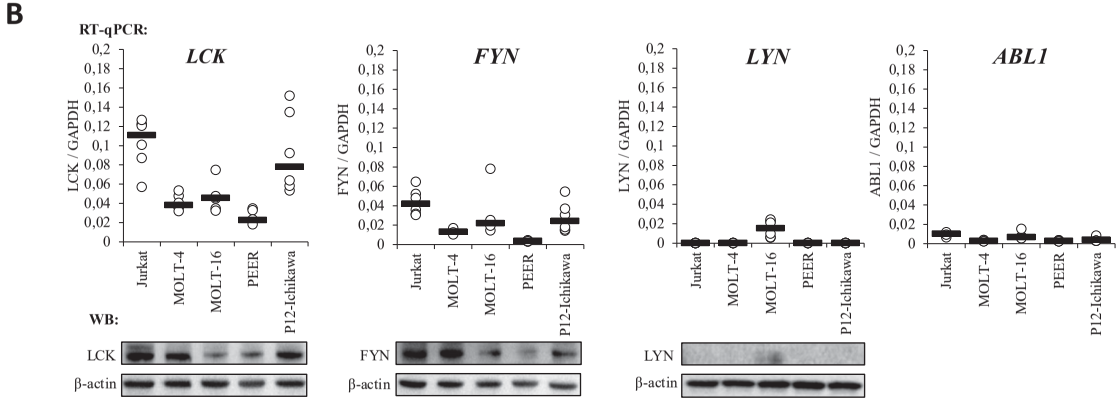


Supplementary Figure S1. The gene expression boxplots of several dasatinib targets in different hematological sample groups: acute leukemia (n=4,430), B-ALL (n=1,304), T-ALL (n=385), and T lymphoids (n=247).

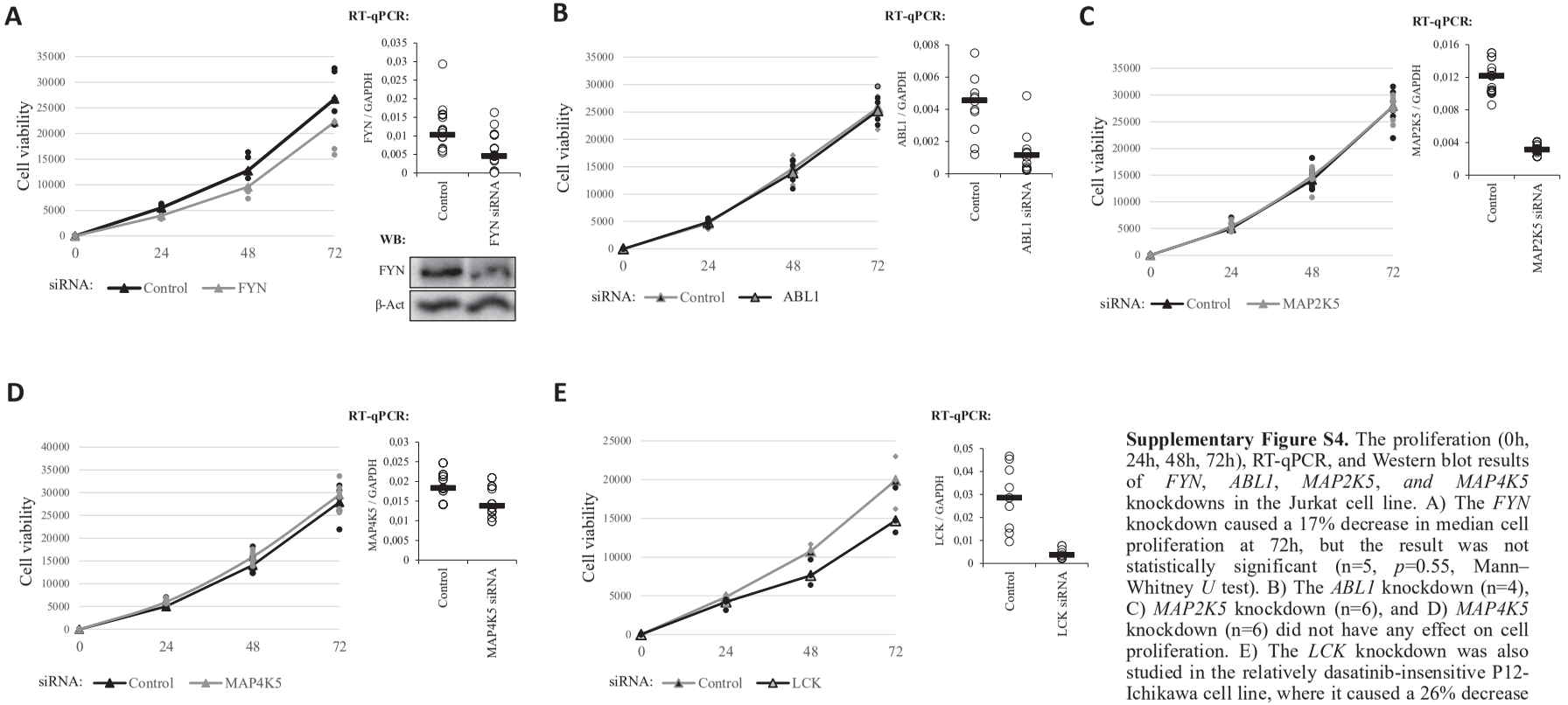


Supplementary Figure S2. The relative cell viabilities of dasatinib treated T-ALL cell lines measured by alamarBlue assay after 72h incubation in a 10-fold dasatinib dilution series (1–1000 nM). The values are relative cell viabilities in comparison to the DMSO control, and the results are the median values from three independent experiments (except for CCRF-CEM and HPB-ALL, which are from two independent experiments). The error bars indicate 95% confidence intervals.

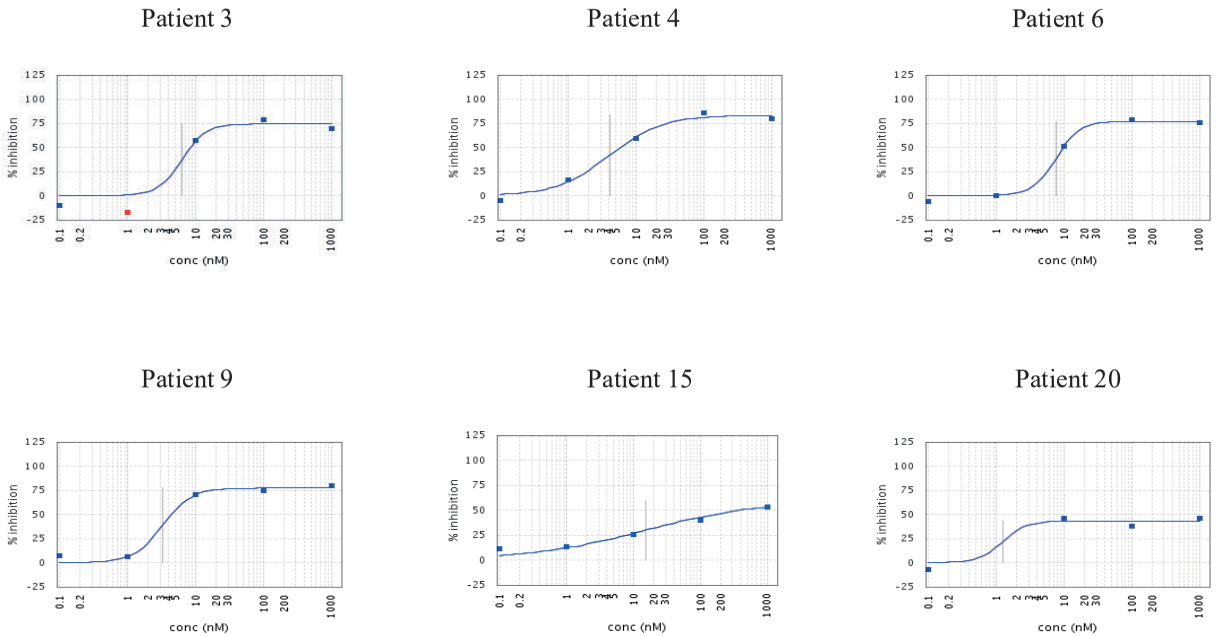




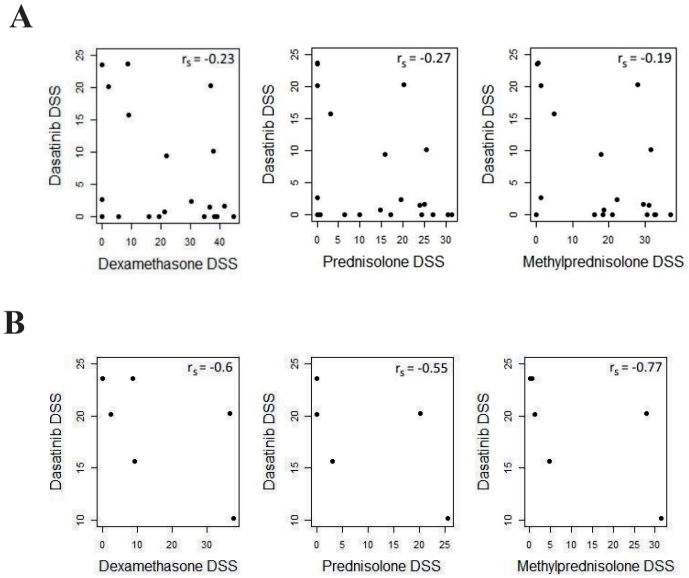
Supplementary Figure S3. A) The RT-qPCR of 21 known dasatinib targets in several T-ALL cell lines. The values are normalized to GAPDH housekeeping control, and the bars indicate the median values. B) The expression of interesting dasatinib-target candidates (*LCK*, *FYN*, *LYN*, and *ABL1*) in five T-ALL cell lines measured by RT-qPCR and Western blotting. Wt ABL1 expression was not detectable by the antibody that recognized the BCR-ABL1 fusion protein in the control sample (2862, Cell Signaling Technology), suggesting a low expression of the protein. The bars in the RT-qPCR plots indicate the median values.



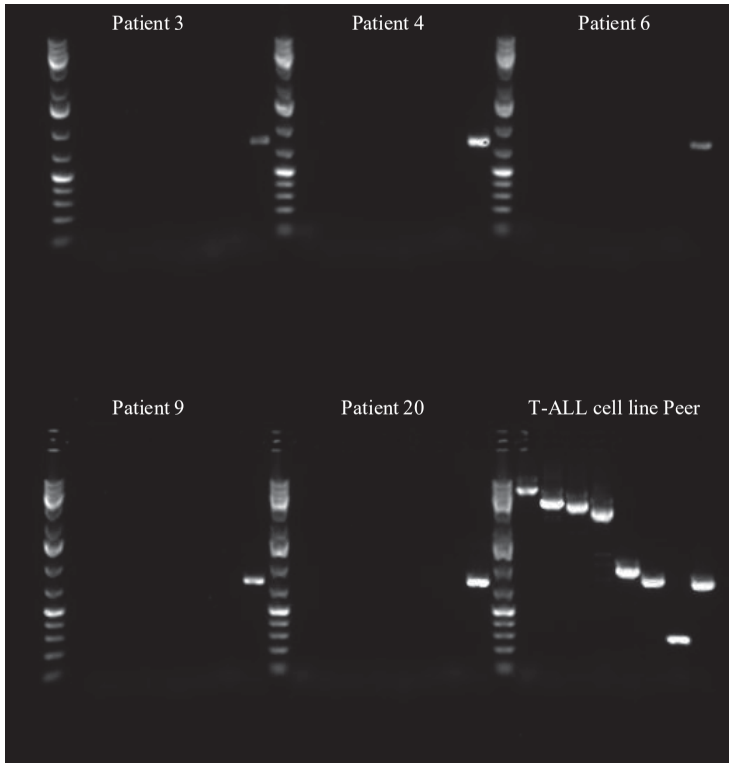
Supplementary Figure S4. The proliferation (0h, 24h, 48h, 72h), RT-qPCR, and Western blot results of *FYN*, *ABL1*, *MAP2K5*, *and MAP4K5* knockdowns in the Jurkat cell line. A) The *FYN* knockdown caused a 17% decrease in median cell proliferation at 72h, but the result was not statistically significant (n=5, p=0.55, Mann–Whitney U test). B) The *ABL1* knockdown (n=4), C) *MAP2K5* knockdown (n=6), and D) *MAP4K5* knockdown (n=6) did not have any effect on cell proliferation. E) The *LCK* knockdown was also studied in the relatively dasatinib-insensitive P12-Ichikawa cell line, where it caused a 26% decrease in median cell proliferation at 72h, but the result was not statistically significant (n=3, p=0.2, Mann–Whitney U test).



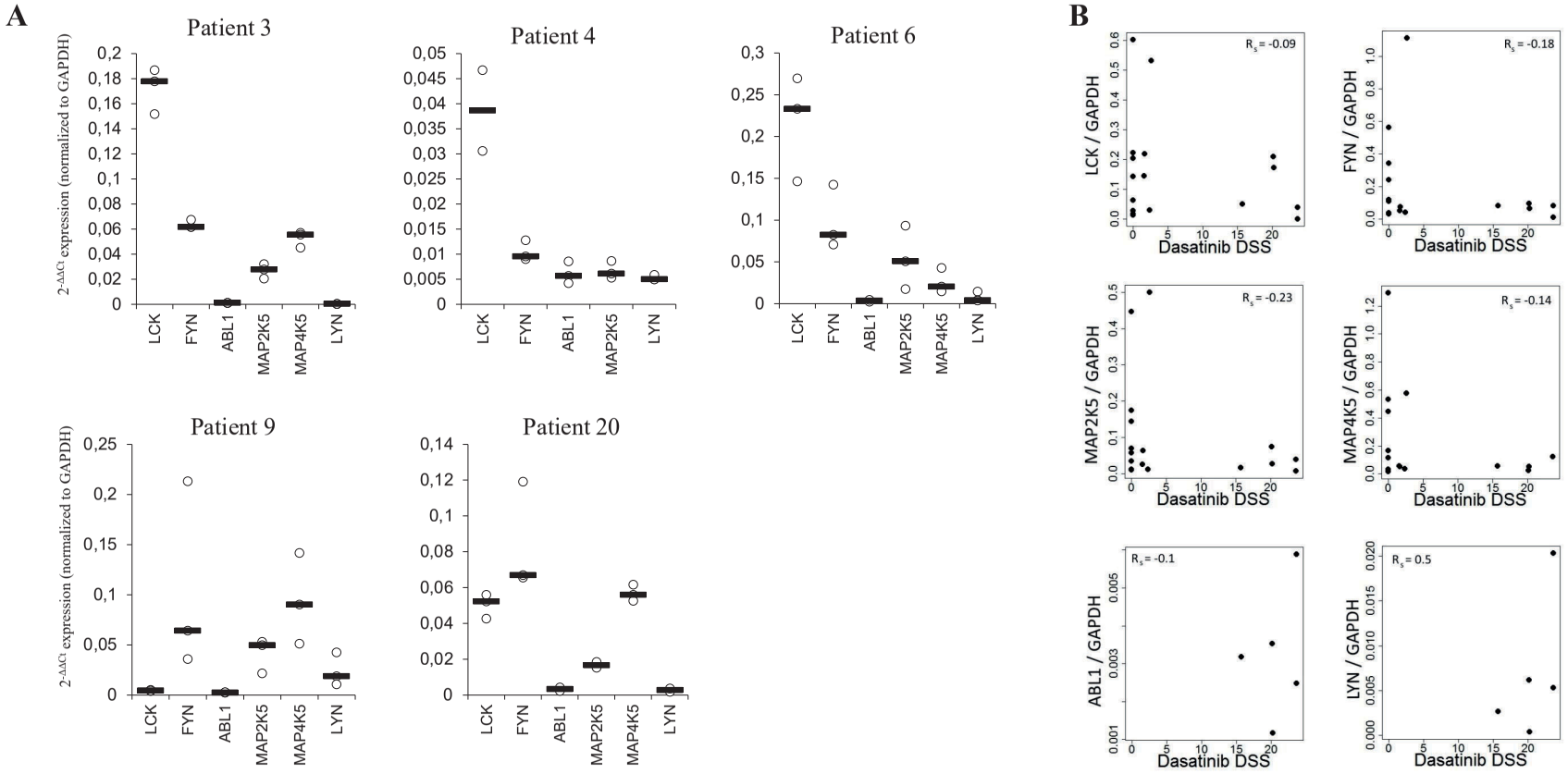
Supplementary Figure S5. Growth inhibition in dasatinib-sensitive patient samples after 72h of treatment. The data are normalized to negative (DMSO) and positive (benzethonium chloride) controls. The gray vertical lines indicate half maximal growth inhibition concentrations.



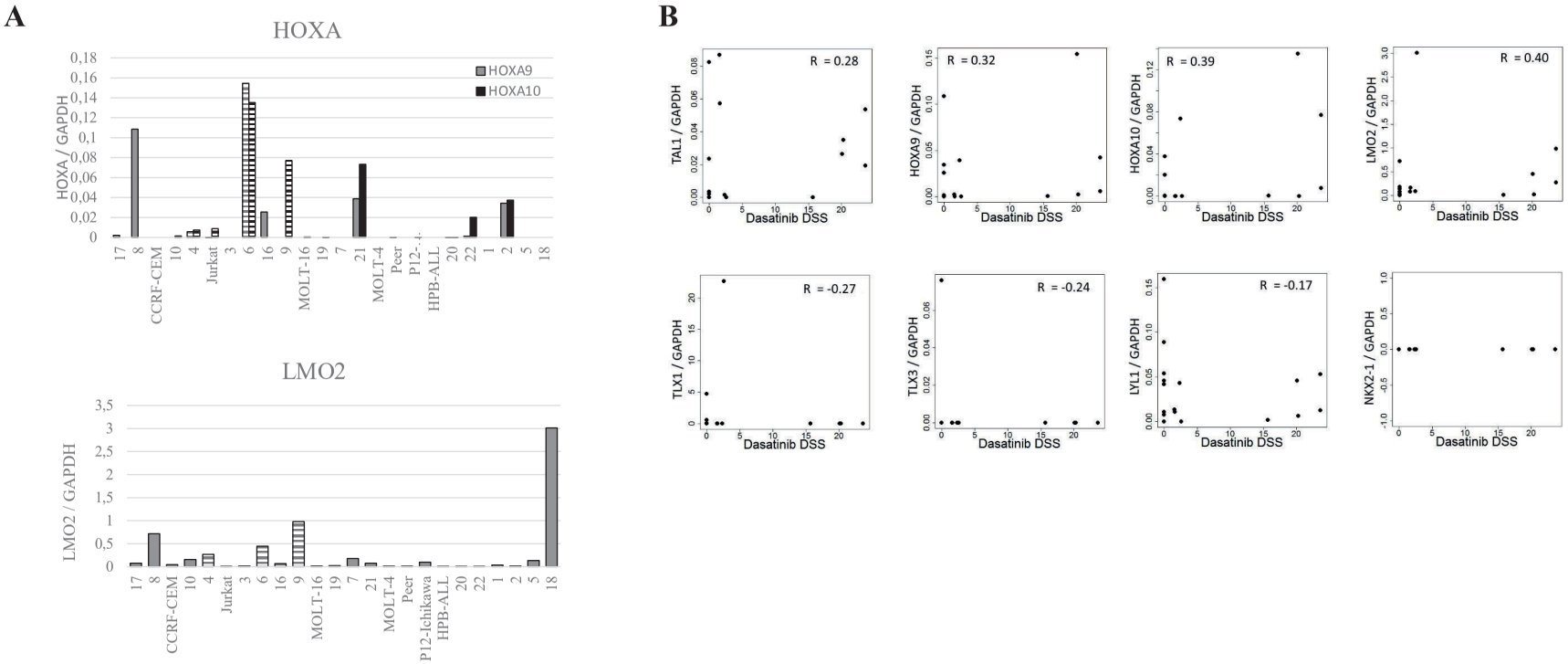
Supplementary Figure S6. A) The negative correlation in the effects of dasatinib and glucocorticoids (prednisolone and methylprednisolone) in a cohort of 22 T-ALL patient samples. B) The negative correlation in the effects of dasatinib and glucocorticoids if subsetting only dasatinib sensitive patient samples (n=6). Spearman's correlation coefficients are shown.



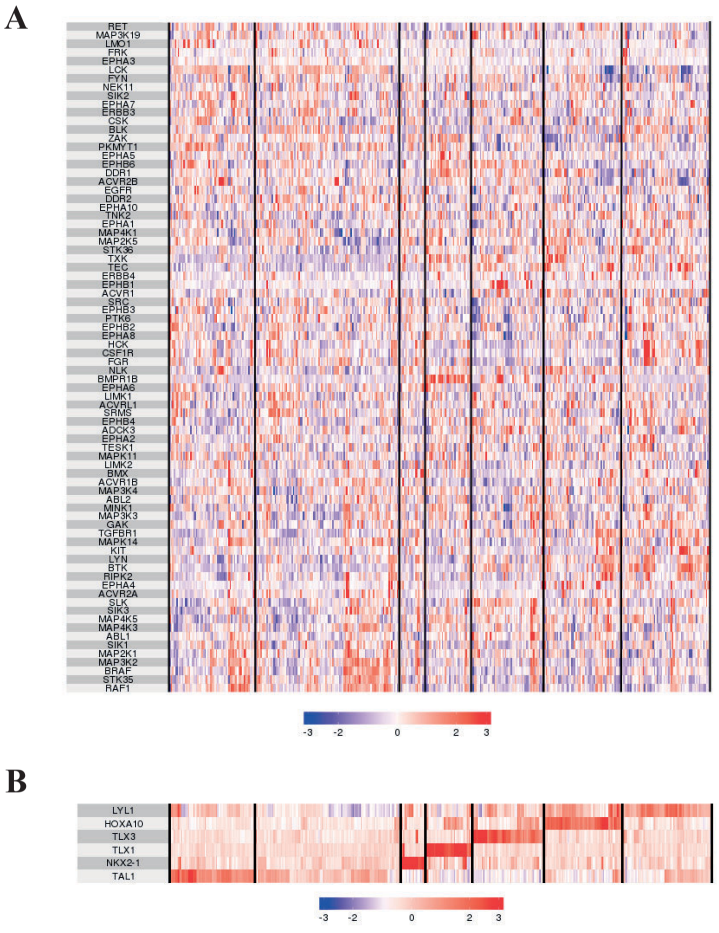
Supplementary Figure S7. The RT-PCR analysis of patient samples for *NUP214-ABL1* fusion. Lanes 1–7 present PCR primers for fusion with different breakpoints, and lane 8 presents the control reaction for RNA quality with PCR primers for *ABL1*. T-ALL cell line Peer was used as a positive control.



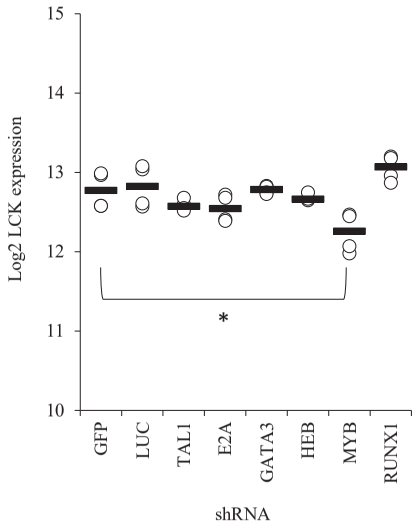
Supplementary Figure S8. A) The expression of *LCK*, *FYN*, *ABL1*, *MAP2K5*, *MAP4K5*, and *LYN* in five dasatinib-responsive patient samples. B) The correlation between Dasatinib DSS and *LCK*, *FYN*, *MAP2K5*, *MAP4K5*, *ABL1*, and *LYN* expression. *LCK*, *FYN*, *MAP2K5*, and *MAP4K5* expression is studied in both dasatinib-sensitive and insensitive samples (n=17). *ABL1* and *LYN* are studied only in dasatinib-sensitive patient samples (n=5). The *ABL1* and *LYN* of all samples and *FYN* of patient 4 were processed in a separate RT-qPCR batch.



Supplementary Figure S9. A) The expression of T-ALL subtype-defining transcription factors *HOXA9/10* and *LMO2* in T-ALL patient samples, and the cell lines measured by RT-qPCR. The striped columns indicate the dasatinib-sensitive samples. B) The correlation between dasatinib response and expression of T-ALL subtype-defining transcription factors: *TAL1*, *HOXA9/HOXA10*, *LMO2*, *TLX1*, *TLX3*, *LYL1*, and *NKX2-1* in patient samples (n=17). Patient sample 4 was processed in a separate RT-qPCR batch, and none of the samples expressed *NKX2-1*.



Supplementary Figure S10. A) Expression of 83 possible dasatinib targets in T-ALL subgroup clusters that are based on subtype-defining transcription factors: *TAL1* (n=61+103), *NKX2-1* (n=18), *TLX1* (n=33), *TLX3* (n=51), *HOXA* (n=56), and *LYL* (n=63). B) The heat map of T-ALL subtype clustering.



Supplementary Figure S11. *LCK* expression in the Jurkat cell line after shRNA knockdowns of *TAL1* and several other transcription factors related to T-ALL (data reproduced from Sanda et al. 2012, *Cancer Cell*). The only significant expression difference was observed between the *MYB* and *GFP/LUC* control knockdowns (*GFP* vs. *MYB*, p=0.021, Mann–Whitney U test).

Supplementary Table S1. Clinical information of patient samples. Highlighted samples were dasatinib responsive.

Sample ID	Age at diagnosis	Sample (diagnosis/relapse)	Cytogenetics	Immunophenotype	Additional immunophenotype information
1	28	Diagnosis	NA	T-ALL	Mature
2	40	Relapse	NA	T-ALL	Cortical
3	**17**	**Diagnosis**	**46<XY,del(6)(q21q25)[4]/46,XY[16]**	**T-ALL**	**Cortical**
4	**50**	**Diagnosis**	**46,XY,t(11;14)(p13;q11)[4]/46,XY[16]**	**T-ALL**	**Cortical**
5	72	Diagnosis	40~44,X,-X,del)3_(q21),-5,add(6),q2?3),add(7)(pl?),-8,-	T-ALL	Immature
6	**75**	**Diagnosis**	**46,XY**	**T-ALL**	**NA**
7	16	Diagnosis	NA	T-ALL	Immature
8	13	Diagnosis	NA	T-ALL	Cortical
9	**26**	**Diagnosis**	**46, XY, del(6q)**	**T-ALL**	**Cortical**
10	63	Diagnosis	46,YY,t(1;14)(p32;q11), deletio 6q matching with T-ALL, add(11)	T-ALL	Mature
11	18	Relapse	Hyperdiploid 48, +8,+18	T-ALL	NA
12	55	Diagnosis	hyperd.48; t(1,16), trisomies 10 and 19	T-ALL	NA
13	38	Diagnosis	46,XY,t(3;19)(p?21;p13),t(10;11)(p13;q14ü21) .ish t(3;19)(E2A+;E2A),t(10;11)(WCP10+, WCP11+,MLL+; WCP11+,MLL-,WCP10+)	T-ALL	NA
14	61	Diagnosis	NA	T-ALL	NA
15	**27**	**Diagnosis**	**46,XY, del 6q21**	**T-ALL**	**NA**
16	29	Diagnosis	NA	T-ALL	NA
17	23	Diagnosis	46,XY	T-ALL	NA
18	NA	NA	NA	T-ALL	NA
19	NA	NA	NA	T-ALL	NA
20	**NA**	**NA**	**NA**	**T-ALL**	**NA**
21	NA	NA	NA	T-ALL	NA
22	NA	NA	NA	T-ALL	NA

PUBLICATION IV

SAP30L (Sin3A-associated protein 30-like) is involved in regulation of cardiac development and hematopoiesis in zebrafish embryos

Teittinen KJ, Grönroos T, Parikka M, Junttila S, Uusimäki A, Laiho A, Korkeamäki H, Kurppa K, Turpeinen H, Pesu M, Gyenesei A, Rämet M, Lohi O.

Journal of Cellular Biochemistry. 2012. 113(12):3843-3852. doi: 10.1002/jcb.24298.

ARTICLE
Journal of Cellular Biochemistry 113:3843–3852 (2012)

Journal of Cellular Biochemistry

SAP30L (Sin3A-Associated Protein 30-Like) is Involved in Regulation of Cardiac Development and Hematopoiesis in Zebrafish Embryos

Kaisa J. Teittinen,[1*] Toni Grönroos,[1] Mataleena Parikka,[2] Sini Junttila,[3] Annemari Uusimäki,[1,2] Asta Laiho,[3] Hanna Korkeamäki,[1] Kalle Kurppa,[1,4] Hannu Turpeinen,[2] Marko Pesu,[2,5] Attila Gyenesei,[3] Mika Rämet,[2,4] and Olli Lohi[1,4]

[1] *Paediatric Research Centre, University of Tampere School of Medicine and Tampere University Hospital, 33520 Tampere, Finland*

[2] *Institute of Biomedical Technology, BioMediTech, University of Tampere, 33520 Tampere, Finland*

[3] *Turku Centre for Biotechnology and the Finnish Microarray and Sequencing Centre, University of Turku and Åbo Akademi University, 20520 Turku, Finland*

[4] *Department of Paediatrics, Tampere University Hospital, 33521 Tampere, Finland*

[5] *Fimlab Laboratories, Tampere University Hospital, 33521 Tampere, Finland*

ABSTRACT

The Sin3A-associated proteins SAP30 and SAP30L share 70% sequence identity and are part of the multiprotein Sin3A corepressor complex. They participate in gene repression events by linking members of the complex and stabilizing interactions among the protein members as well as between proteins and DNA. While most organisms have both SAP30 and SAP30L, the zebrafish is exceptional because it only has SAP30L. Here we demonstrate that SAP30L is expressed ubiquitously in embryonic and adult zebrafish tissues. Knockdown of SAP30L using morpholino-mediated technology resulted in a morphant phenotype manifesting as cardiac insufficiency and defective hemoglobinization of red blood cells. A microarray analysis of gene expression in SAP30L morphant embryos revealed changes in the expression of genes involved in regulation of transcription, TGF-beta signaling, Wnt-family transcription factors, and nuclear genes encoding mitochondrial proteins. The expression of the heart-specific nkx2.5 gene was markedly down-regulated in SAP30L morphants, and the cardiac phenotype could be partially rescued by nkx2.5 mRNA. In addition, changes were detected in the expression of genes known to be important in hemoglobin synthesis and erythropoiesis. Our results demonstrate that SAP30L regulates several transcriptional pathways in zebrafish embryos and is involved in the development of cardiac and hematopoietic systems. J. Cell. Biochem. 113: 3843–3852, 2012.

KEY WORDS: ZEBRAFISH; SAP30L; TRANSCRIPTION; MICROARRAY; HEMATOPOIESIS; CARDIOGENESIS

Acetylation and other covalent modifications of histones play fundamental roles in chromatin dynamics and regulation of gene expression. Deacetylation of histones is carried out by a multiprotein corepressor complex in which Sin3A is an essential scaffold protein. Sin3A is composed of domains that mediate protein–protein interactions and thereby forms a platform to which several enzymes, DNA-binding transcription factors, and other bridging proteins bind [Silverstein and Ekwall, 2005]. The core

The authors have no conflict of interest to declare.

Additional supporting information may be found in the online version of this article.

Grant sponsor: Academy of Finland Research Council for Health; Grant numbers: 115260, 121003, 1286223, 135980, 135736, 139225; Grant sponsor: Foundation for Paediatric Research in Finland; Grant sponsor: Finnish Medical Foundation; Grant sponsor: Competitive Research Funding of Tampere University Hospital; Grant numbers: 9J062, 9K073, 9M052, 9L070, 9K093, 9L075, 9M080, 9M093; Grant sponsor: Marie Curie International Reintegration Grant within the 7th European Community Framework Programme; Grant sponsor: Emil Aaltonen Foundation; Grant sponsor: Sigrid Juselius Foundation; Grant sponsor: Nona and Kullervo Väre Foundation; Grant sponsor: Päivikki and Sakari Sohlberg Foundation.

*Correspondence to: Kaisa J. Teittinen, Paediatric Research Centre, University of Tampere School of Medicine, FIN-33520 Tampere, Finland E-mail: kaisa.teittinen@uta.fi

Manuscript Received: 28 May 2012; Manuscript Accepted: 16 July 2012

Accepted manuscript online in Wiley Online Library (wileyonlinelibrary.com): 20 July 2012

DOI 10.1002/jcb.24298 • © 2012 Wiley Periodicals, Inc.

Sin3A-HDAC corepressor complex contains the histone deacetylases HDAC1 and HDAC2, the histone-binding proteins RbAp46 and RbAp48, the Sin3A-associated protein 18 (SAP18), SAP30, and SDS3 [Silverstein and Ekwall, 2005].

The Sin3A-associated proteins SAP30 and SAP30L (SAP30-like) have been shown to be present in various corepressor complexes [Laherty et al., 1998; Zhang et al., 1998; Viiri et al., 2006]. We have previously reported that these proteins bind DNA directly and this association is regulated by nuclear monophospholipids [Viiri et al., 2009b]. SAP30 and SAP30L are small and basic, and contain nuclear (NLS) and nucleolar (NoLS) localization signals [Viiri et al., 2006]. The main difference between the two proteins lies in their N-termini, where SAP30 contains a 38 amino acid insertion relative to SAP30L. Both proteins are able to localize to the nucleus or the nucleolus, and they can direct Sin3A to the nucleolus [Viiri et al., 2006].

While a range of biochemical and cell biological results implicate SAP30 and SAP30L in repression of gene expression, the function(s) of these proteins in the setting of a living organism has not been established. In order to study this, we chose the zebrafish as a model organism. The zebrafish has only one of these proteins, SAP30L [Viiri et al., 2009a], enabling us to study the role of SAP30L without interference of potentially overlapping functions of SAP30. Our results suggest an important role for SAP30L in diverse transcriptional pathways in zebrafish embryos.

MATERIALS AND METHODS

ZEBRAFISH MAINTENANCE AND MICROINJECTIONS

Wild type AB zebrafish were maintained under standard conditions at 28.5°C as described [Westerfield, 1995]. In microinjection experiments, embryos at the 1-4-cell stage were injected into the yolk sac with 1 nl of injection mix, which contained the morpholino, mRNA (where indicated) and rhodamine dextran. The PV830 Pneumatic PicoPump injection system (WPI, Inc.) was used, and the injection volume was calibrated to 1 nl using 1 mm Micrometer scale (S48, Ted Pella Inc.). The care of the animals was in accordance with the Finnish Laboratory Animal Welfare Act 62/2006 and the Laboratory Animal Welfare Ordinance 36/2006.

WHOLE-MOUNT IN SITU HYBRIDIZATION (WISH)

DIG-labeled sense and antisense probes for zebrafish SAP30L were prepared using the DIG RNA Labeling Mix (Roche Diagnostics) according to the manufacturer's instructions, as described [Thisse and Thisse, 2008]. Linearized plasmids containing the full-length SAP30L cDNA (in the sense or antisense orientation) under the control of the T7 promoter were used as templates. Zebrafish larvae were collected and processed as described [Thisse and Thisse, 2008] at 12, 24, 36, 48, and 72 hpf. Whole-mount in situ hybridization (WISH) using alkaline phosphatase detection with the BM Purple substrate (Roche Diagnostics) was carried out according to Thisse and Thisse [2008] with the following modifications: 2 ml tubes were used during the entire procedure, hybridization was carried out at +65°C, 130 ng of sense or antisense probe was used per reaction and a dilution ratio of 1:2,000 was used for the anti-DIG antibody (sheep anti-digoxigenin-AP Fab fragments, Roche Diagnostics). The stained fish were analyzed and photographed under a dissecting microscope. The localization of the staining was assessed by comparing to the anatomical reference images by Haffter et al. [1996].

RNA EXTRACTION AND QUANTITATIVE REAL-TIME PCR

Expression of the SAP30L, nkx2.5 and ppox genes was assayed by quantitative real-time PCR (qRT-PCR). Total RNA was extracted from zebrafish embryos and adult fish tissues with an RNeasy Mini Kit (Qiagen), and converted to cDNA using an iScript Select cDNA Synthesis Kit (Bio-Rad) and random primers according to the manufacturers' instructions. The housekeeping gene EF1A (Ensembl ID: ENSDARG00000020850) [Tang et al., 2007] was used as a normalization control. The primer sequences are described in the Supplementary data. qRT-PCR was performed by using an Evagreen Ssofast supermix kit (Bio-Rad) as instructed by the manufacturer. An average threshold cycle (Ct) value was calculated from two or three replicate samples. The mRNA levels (mean ± SD) are expressed relative to those of the housekeeping gene, EF1A. The PCR runs were repeated twice, and no significant differences between replicate runs were observed (data not shown). For the microarray experiments, the total RNA was extracted from whole embryos collected at 12 or 24 hpf.

ANTISENSE MORPHOLINO EXPERIMENTS

The translational start site of the zebrafish SAP30L mRNA (Ensembl ID: ENSDARG00000030213) was targeted with two independent antisense morpholinos: SAP30L-MO1 and SAP30L-MO2 (Fig. 2a, Supplementary Table I). Their sequences exhibited no significant similarity to other loci in a search of a zebrafish database. The random control morpholino (RC, see Supplementary Table I) has no gene targets or observable biological activity in the zebrafish. The p53-targeting morpholino was also used (Supplementary Table I). All morpholinos were purchased from Gene-Tools, LLC. The effective concentrations of morpholinos were experimentally titrated and are indicated in the figures or figure legends.

WESTERN BLOTTING

The efficacy of MO1 and MO2 in blocking the translation of the SAP30L mRNA was analyzed using a human anti-SAP30L antibody [Korkeamaki et al., 2008]. MO1-, MO2-, or RC-injected larvae were collected at 2 dpf (days post-fertilization) directly into 2× Laemmli buffer (10 larvae/50 μl), and lysed by boiling and passing through a pipette vigorously. The lysates were centrifuged for 5 min to remove the debris, and the soluble fraction was used for further analysis. The proteins were resolved on 12% SDS–PAGE gels and transferred to nitrocellulose membranes (Amersham Biosciences). The immunoblots were probed with an anti-SAP30L antibody, and the proteins were detected by enhanced chemiluminescence. In all zebrafish lysates, the anti-SAP30L-antibody recognized only a single band, the molecular weight of which is a few kilodaltons lower than in human control (lysed cells transfected with Myc-His-tagged SAP30L), as expected due to the presence of the tag and the larger size of human SAP30L.

ANALYSIS OF HEART MORPHOLOGY AND FUNCTION

Heart morphology at 5 dpf was assessed visually and categorized as normal or deformed, examples of which are shown in Supplementary Figure S3. Heart rates were determined visually in unanaesthetized morphant and control larvae at 5 dpf. Ventricular and atrial performance was analyzed on 10–20 s time-lapse videos. The longitudinal fractional shortening was calculated according to the formula $(LDd - LDs)/LDd$, where *LDd* is the longitudinal diameter at diastole and *LDs* the diameter at systole.

HISTOLOGICAL STAINING

The morphant and control larvae were fixed in a 4% formaldehyde–PBS solution at +4°C and embedded in 2% agarose. The samples were dehydrated by incubating them in a series of alcohol solutions (70%, 96%, and absolute ethanol) for 1–2 h in each solution, and finally in xylene for at least 1 h. The samples were then embedded in paraffin and 5 μm-thick sections were cut in longitudinal orientation. The sections were fixed on glass slides and deparaffinised by incubating them twice in xylene (á 4 min), twice in absolute ethanol (á 3 min), twice in 96% ethanol (á 3 min), once in 70% ethanol, and rinsing once with water. Hematoxylin–eosin staining was performed by incubating the samples in Mayer's hematoxylin (2 min), running water (2 min), water (1.5 min), 70% ethanol (15 s), eosin Y (15 s), 96% ethanol (30 s, twice), absolute ethanol (1 min, twice), and xylene (1–4 min). After this the slides were mounted and pictures were taken under a BX60 microscope using the Cell^D program (Olympus).

O-DIANISIDINE STAINING

Detection of hemoglobin by *o*-dianisidine staining was performed as described previously [Ransom et al., 1996]. Larvae stained at 3 dpf were scored as normal, reduced, or severely reduced based on their qualitative level of staining which was assessed visually. In the normal category, staining was judged to be equivalent to that of wild type larvae, reduced staining was visibly decreased as compared to wild type, and in the severely reduced category, the larvae showed little or no staining.

mRNA RESCUE

The mRNA rescue experiments were performed by injecting simultaneously SAP30L-MO1 or SAP30L-MO2 and a capped synthetic mRNA encoding the gene of interest. We used various doses of mRNA in the injections, and the data were pooled for analysis. The mRNA was synthesized using a T7 mMESSAGE mMACHINE Kit (Ambion) according to the manufacturer's instructions. Briefly, a plasmid template linearized with *Hind*III was mixed thoroughly with components of the kit and incubated at +37°C for 2 h, after which the template DNA was removed by a DNase treatment. The synthesized mRNA was purified using a MEGAclear kit (Ambion) according to the manufacturer's protocol and further concentrated by ammonium acetate precipitation. The templates used in the synthesis reactions had been created by cloning the desired cDNA into the pcDNA3.1/*myc*His(−)A-vector (Invitrogen) using *Eco*RI and *Hind*III restriction sites. The constructs contained full-length zebrafish SAP30L (drSAP30L) or full-length zebrafish nkx2.5 (Nkx2.5), each without a stop codon.

STATISTICAL ANALYSES

Quantitative data are given as mean ± standard deviation. When appropriate, differences between groups were compared by using the two-tailed Student *t*-test or the χ^2 test for cross tabulations. A *P*-value of less than 0.05 was considered statistically significant. Statistical calculations were performed using PASW Statistics version 18.

MICROARRAYS

200 ng of total RNA was amplified and labeled with Cy3 using a Low Input Quick Amp Labeling kit (Agilent Technologies). The samples were processed using an RNA Spike In kit (Agilent Technologies). 1.65 μg of each Cy3-labeled sample was hybridized to Agilent's 4 × 44K Zebrafish V3 Gene Expression Microarray overnight at 65°C in the buffers of Agilent's Gene Expression Hybridization kit. The arrays were washed according to the manufacturer's instructions and scanned with an Agilent Technologies' scanner (model G2565CA), using scan profile AgilentHD_GX_1Color. Numerical data were extracted with Agilent's Feature Extraction software, version 10.7.1., using grid 026437_D_F_20100719 and protocol GE1_107_Sep09.

MICROARRAY DATA ANALYSIS

The R programming language and environment [R Development Core Team, 2008] and its Bioconductor module [Gentleman et al., 2004] were used for analysis of data from the microarrays. The raw data were normalized with quantile normalization, and the quality of the data was checked with several quality control methods. Statistical testing was performed with the Limma package [Smyth, 2005], and stringent filtering thresholds were used for both time point comparisons (fold change $FC \geq |3|$ and *P*-value $P \leq 0.001$).

RESULTS AND DISCUSSION

SAP30L mRNA IS EXPRESSED DURING EMBRYOGENESIS AND IN ADULT ZEBRAFISH TISSUES

SAP30 and SAP30L have been implicated in the regulation of gene repression through the Sin3A-corepressor complex [Laherty et al., 1998; Zhang et al., 1998; Viiri et al., 2006, 2009b]. Their function in vivo remains unknown and we therefore set out to examine the function of SAP30L using the zebrafish as a model. Contrary to the human and mouse genomes, only one member of the SAP30 family has thus far been identified in the zebrafish genome, and its derived protein sequence more closely resembles human SAP30L than SAP30 (Supplementary Fig. S1; Viiri et al., 2009a).

We studied the expression of the SAP30L gene during zebrafish embryogenesis in whole embryos by qRT-PCR, and determined the localization of the expression by WISH at 12, 24, 36, 48, and 72 h post-fertilization (hpf) (Fig. 1a and b). The results show that SAP30L mRNA is widely expressed during embryogenesis (Fig. 1a and b). The expression is strongest in the brain, including forebrain, midbrain, and hindbrain, at the 24, 36, 48, and 72 hpf time points (Fig. 1a). The pectoral fin buds (at 36 and 48 hpf) and pectoral fins (72 hpf) also show strong expression of the mRNA. Next to the pectoral fin buds, staining is seen also in the region of common cardinal vein (see dorsal view in Fig. 1a) starting at 36 hpf. Furthermore, the heart area

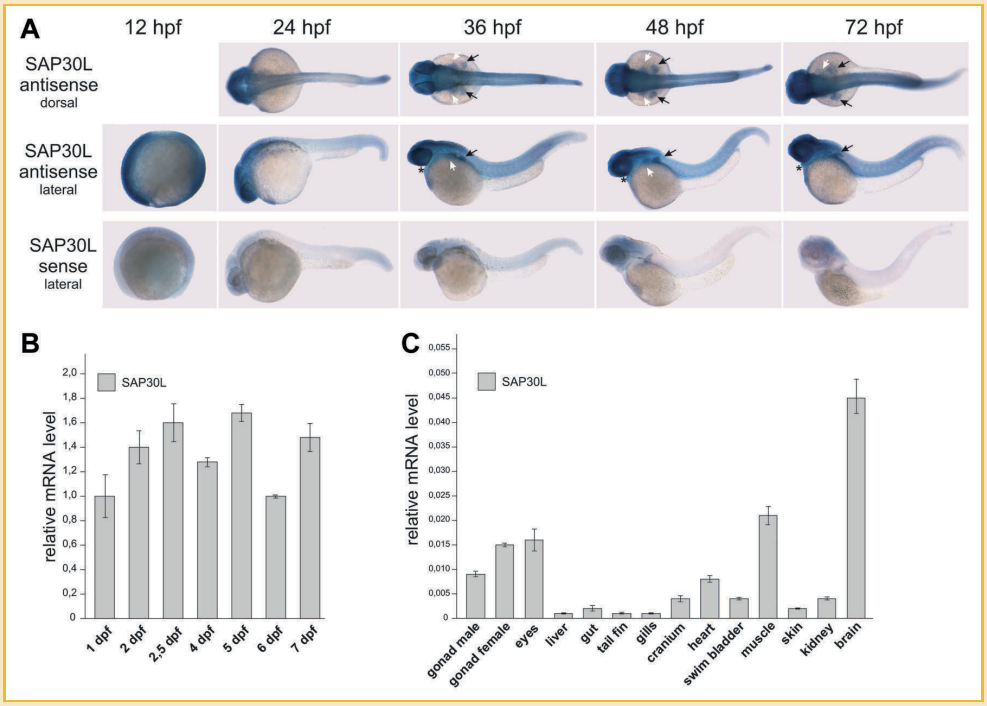


Fig. 1. Expression of SAP30L mRNA in zebrafish embryos, larvae, and adult tissues. A: WISH analysis of SAP30L mRNA expression during embryogenesis at indicated time points. Staining with an antisense probe is shown in dorsal (top row) and lateral (middle row) views, and a control sense-strand probe is shown in lateral view (bottom row). The black asterisk points out the heart next to it, the black arrow indicates the pectoral fin bud/pectoral fin, and the white arrow indicates the region of the common cardinal vein. B: qRT-PCR analysis of SAP30L mRNA in whole embryos during embryogenesis. SAP30L mRNA levels are normalized to the level at 1 day post-fertilization (dpf). C: qRT-PCR analysis of SAP30L mRNA in adult tissues. SAP30L mRNA levels (mean ± SD) are expressed relative to those of the housekeeping gene, EF1A. [Color figure can be seen in the online version of this article, available at http://wileyonlinelibrary.com/journal/jcb]

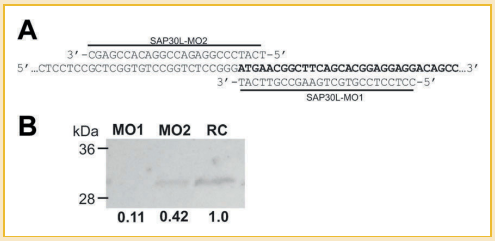


Fig. 2. The morpholinos used to target the SAP30L mRNA. A: Sequence of the coding strand of the zebrafish SAP30L gene, showing the 5′-untranslated region (normal font) and part of the first exon (boldface starting from the ATG-codon). Complementarity of the sequences of the morpholinos SAP30L-MO1 and SAP30L-MO2 to the coding strand of the gene is illustrated. B: Western blotting analysis of SAP30L protein levels in 2 dpf larvae injected with SAP30L-MO1 (MO1), SAP30L-MO2 (MO2), or random control morpholino (RC). Morpholino concentrations used: 400 μM for MO1, 600 μM for MO2, 600 μM for RC. The numbers under the lanes indicate band intensity relative to RC. The lysates contain identical numbers of larvae per volume of lysis buffer, and equal volume of sample was loaded on the gel in each case.

shows positive staining, first seen in the pericardial cavity (24 hpf) and later in the heart itself. In addition, moderate expression of SAP30L mRNA is seen in the trunk (including the area of the intermediate cell mass, ICM), and some embryos show weak staining in the blood island at 24 hpf. In adult fish, strong expression is seen in the brain and moderate expression in muscle, eye, gonad, and heart tissues (Fig. 1c).

DESIGN AND VALIDATION OF THE SAP30L-TARGETING MORPHOLINOS

In zebrafish embryos, morpholino oligonucleotides (MO) can be used to knock down expression of genes of interest. We designed two translation-blocking morpholinos (MO1 and MO2) which target the translation initiation site in the SAP30L mRNA (Fig. 2a). In an in vitro translation assay, a specific and dose-dependent inhibition of translation of SAP30L mRNA was detected with the MO1 but not with RC control (Supplementary Fig. S2). The ability of morpholinos to block translation of SAP30L was further investigated by using the SAP30L antibody and Western blotting. A significant reduction in SAP30L protein levels was detected in MO1- and MO2-treated zebrafish embryos at 2 dpf, demonstrating the efficacy of these

morpholinos in down-regulating the expression of the SAP30L protein in vivo (Fig. 2b).

KNOCKDOWN OF SAP30L EVOKES CARDIAC DEFECTS AND REDUCED HEMOGLOBIN LEVELS

In order to investigate the role of SAP30L in zebrafish development, we injected morpholinos (MO1 or MO2) into 1- to 4-cell stage embryos, and observed their development for 5–6 dpf. The SAP30L morphants showed prominent and progressive pericardial edema. Importantly, most of the SAP30L morphants exhibited deformed cardiac morphology (85–89%; Fig. 3a, Supplementary Fig. S3), with the hearts remaining string-like and visible on a single plane, in contrast to the three-dimensional complexity of the healthy hearts in the control larvae (Fig. 3b). Cardiac function was assessed by determining the heart rates and the contractility of the chambers. The results demonstrate significantly reduced heart rates in MO1- and MO2-injected morphants (Fig. 3c), with the differences detectable already at 3 dpf. Similarly, a substantial decrease in the longitudinal fractional shortening values was seen in both the atrium and the ventricle, as compared to control hearts (Fig. 3d). Videos illustrating the diminished cardiac function can be found in the Supplementary data (SV1-3). The results are compatible with the WISH and qRT-PCR analyses which showed expression of SAP30L in the developing and adult heart (Fig. 1a and c), and suggest that the phenotype results specifically from the knockdown of SAP30L.

In some SAP30L morphants, the red blood cells seemed paler as compared to those in control larvae. In a Sin3B-knockout mouse, impaired erythroid differentiation has been reported, with a 43% reduction in the number of circulating red blood cells and a 37% reduction in hemoglobin levels [David et al., 2008]. In order to determine hemoglobin levels in the SAP30L morphants, *o*-dianisidine staining was performed on 3 dpf embryos (see examples in Supplementary Fig. S4). As summarized in Figure 3e, the morphants exhibited reduced staining, indicating reduced hemoglobin levels in the red blood cells. The observed effect on erythroid differentiation is consistent with the fact that SAP30L expression is seen at the area of the common cardinal vein and ICM,

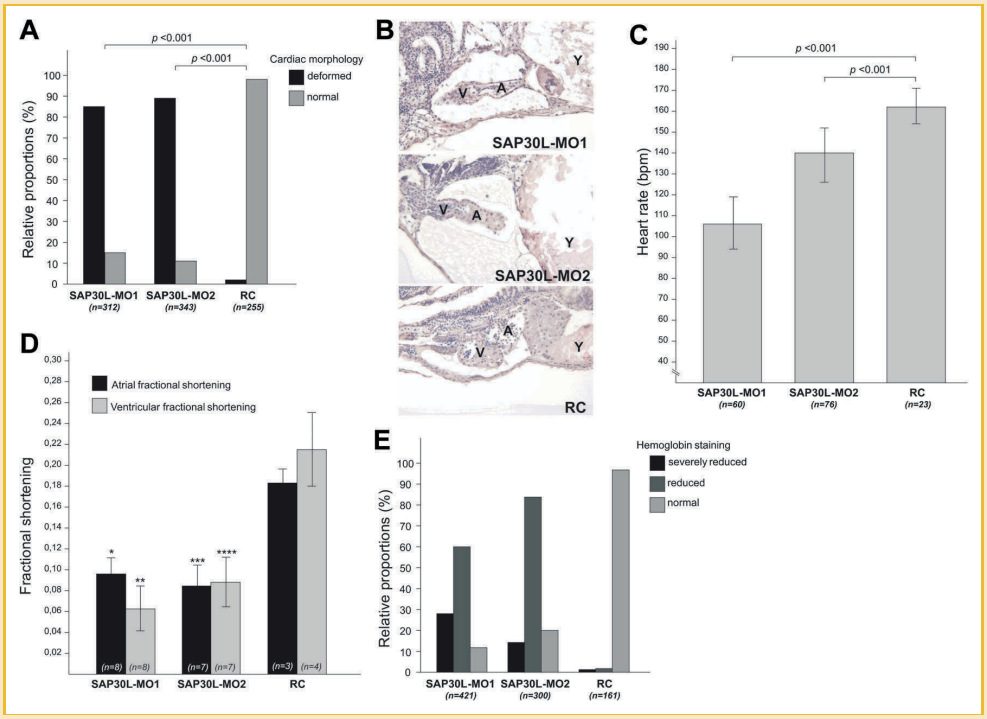


Fig. 3. The phenotype induced by knockdown of SAP30L involves defects in the heart and in hemoglobinization of red blood cells. A: The majority of the SAP30L morphant larvae have deformed hearts (see close-up images of normal and deformed hearts in Supplementary Fig. S2). B: Histological analysis of cardiac morphology. Hematoxylin–eosin staining of sections reveals string-like appearance of the heart in the MO1- and MO2-injected larvae, in contrast to the RC-injected controls. V, ventricle; A, atrium; Y, yolk sac. C: SAP30L morphant larvae exhibit reduced heart rates. Heart rates (beats per minute, bpm) are represented as mean bpm ± SD. D: Both atrial and ventricular performance in SAP30L morphant hearts are impaired, as shown by the decreased longitudinal fractional shortening values (represented as mean ± SD). $^{*}P = 0.002$; $^{**}P < 0.001$; $^{***}P = 0.005$; $^{****}P = 0.006$, as compared to RC control. The morpholino concentrations used in A–D: 300–750 μM for MO1, 500–600 μM for MO2, 400–750 μM for RC. E: SAP30L morphants exhibit defects in hemoglobinization of red blood cells (see examples in Supplementary Fig. S4). MO1-, MO2-, or RC-injected larvae are scored as normal, reduced, or severely reduced based on hemoglobin staining by *o*-dianisidine. Morpholino-concentrations used: 400 μM for MO1, 600 μM for MO2, 600 μM for RC. [Color figure can be seen in the online version of this article, available at http://wileyonlinelibrary.com/journal/jcb]

where developing erythrocytes are present [Davidson and Zon, 2004].

In addition, some morphant larvae were slightly smaller than control larvae, and showed edema in the head and brain (data not shown).

The phenotypes obtained with both SAP30L-targeting morpholinos were essentially similar, with MO2-morphants showing slightly delayed onset, which concurs with the less decreased level of SAP30L protein in these morphants as compared to MO1-morphants (Fig. 2b). Since morpholinos are known to induce off-target p53-mediated effects [Robu et al., 2007], we co-injected both morpholinos with the p53-morpholino. Larvae injected with SAP30L-MO2 alone or together with the p53-morpholino, did not differ in phenotype. SAP30L-MO1 morphants exhibited a slightly alleviated but essentially similar phenotype when co-injected with the p53-morpholino (data not shown).

One way to demonstrate the specificity of the observed phenotype is to rescue it by co-injecting the morpholino with the target mRNA [Eisen and Smith, 2008]. In our case, this strategy could be exploited only with MO2, since it targets the SAP30L mRNA upstream of the start codon. No rescue of the reduced hemoglobin staining but a tendency towards a less severe cardiac phenotype could be observed in the rescued morphant larvae (data not shown). The incomplete rescue may be due to various reasons, such as the short half-life of the injected mRNA, strict temporal or spatial constraints on the expression at certain developmental stages, inability of the mRNA to reach the right tissue, or the fine balance between under- and over-expression of the protein in the embryos and larvae.

SAP30L IS INVOLVED IN MULTIPLE TRANSCRIPTIONAL PATHWAYS DURING ZEBRAFISH EMBRYOGENESIS

SAP30L is a part of the Sin3A-HDAC multiprotein corepressor complex [Viiri et al., 2006]. In order to elucidate the transcriptional pathways in which SAP30L participates during zebrafish embryogenesis, we performed a microarray analysis of embryos in which expression of SAP30L was knocked down. To avoid any secondary effects on gene expression, we chose early time points at 12 and 24 hpf. It has been reported previously that most developmentally regulated genes reach their peak expression by 24 hpf, with a sharp increase in transcript accumulation at 12 hpf [Mathavan et al., 2005]. In our analysis, 205 genes were up-regulated and 174 genes down-regulated at the 12 hpf time point, whereas at 24 hpf, the respective numbers were 143 and 64 (Fig. 4, Supplementary Fig. S5). Of the affected genes, 46 were identified at both time points and showed similar alterations in expression. Since SAP30L is part of a transcriptional repressor complex, its knockdown may be expected to lead to more up-regulation than down-regulation of genes. Our results bear out this expectation and are in line with the published microarray data for Sin3, another member of the same repressor complex. RNAi-mediated knockdown of *Drosophila* SIN3 resulted in induction of 364 genes but only 35 genes were repressed [Pile et al., 2003]. Similarly, in mouse cells in which Sin3A had been knocked out, the majority of the differentially expressed genes were up-regulated [Dannenberg et al., 2005]. The results of knockdown of *Drosophila* SIN and the mouse mSin3A were similar in that mainly genes involved in cytosolic and mitochondrial energy-generating pathways were affected. Our data on knockdown of SAP30L in *Danio rerio* show no marked changes in expression of the genes reported in the *Drosophila* SIN and mouse Sin3A data sets. Instead, expression of several nuclear genes encoding mitochondrial proteins was affected by knockdown of SAP30L. These include timm8a, cyb5b, ppox, alas2 and fars2 (Fig. 4, Supplementary Fig. S5).

SAP30L was originally identified in our laboratory as a TGF-beta-inducible transcript in a cell culture system that models differentiation of the intestinal epithelium [Lindfors et al., 2003]. Interestingly, down-regulation of SAP30L in the zebrafish resulted in strong induction of Smad2, a mediator in the TGF-beta signaling pathway. The 150 kDa TGF-beta-1-binding protein was similarly up-regulated at 12 hpf, whereas the bone morphogenetic factor 5 (BMP5), a member of the TGF-beta superfamily, was slightly down-regulated. These results taken together with our previous findings suggest a role for SAP30L in TGF-beta signaling events.

In gene ontology analysis, the pathway with the most hits is associated with regulation of transcription. The same result was obtained when the analysis was performed on the genes that showed the largest differences in their expression. Indeed, several transcription factors show altered expression in our microarray analysis (Fig. 4, Supplementary Fig. S5). Among them are the Wnt-family transcription factor Wnt7b, which is involved in the development of vasculature, lung, and kidney [Shu et al., 2002; Lobov et al., 2005; Rajagopal et al., 2008; Yu et al., 2009; Lin et al., 2010], and Wnt2, which has been shown to be associated with cardiogenesis and several other developmental processes [Alexandrovich et al., 2006; Wang et al., 2007; Goss et al., 2009, 2011; Sousa et al., 2010]. In addition, Wisp3, a protein involved in modulation of Wnt and BMP signaling [Nakamura et al., 2007], was affected.

NKX2.5 EXPRESSION IS REDUCED IN SAP30L KNOCKDOWN EMBRYOS

One of the genes down-regulated at 12 hpf in the microarray analysis of SAP30L morphants, Nkx2.5 (Fig. 4), is a well-known regulator of heart development [Chen and Fishman, 1996; Targoff et al., 2008; Reamon-Buettner and Borlak, 2010]. The down-regulation of the Nkx2.5 gene was confirmed by qRT-PCR, which showed a threefold decrease in its expression in SAP30L-MO1-treated embryos, and a sixfold decrease in SAP30L-MO2-treated embryos (Fig. 5a). Interestingly, co-injection of Nkx2.5 mRNA with SAP30L-MO1 could partially rescue the deformed cardiac morphology in the SAP30L morphants (Fig. 5b). Previous studies in other model organisms have implicated Nkx2.5 in various cardiac developmental processes, including progenitor specification and proliferation, heart tube extension and looping, and chamber morphogenesis [Targoff et al., 2008]. Our results suggest that transcription of Nkx2.5 is regulated by a SAP30L-containing repressor complex in zebrafish embryos. As SAP30L is a member of a large corepressor complex and binds DNA only nonspecifically, its observed effect on expression of Nkx2.5 is likely to be indirect. Interestingly, some members of the Sin3A corepressor complex are known to regulate expression of genes critical for cardiac morphogenesis and function in both the zebrafish and mice. These include HDAC1 [Pillai et al., 2004; Montgomery et al., 2007], HDAC2

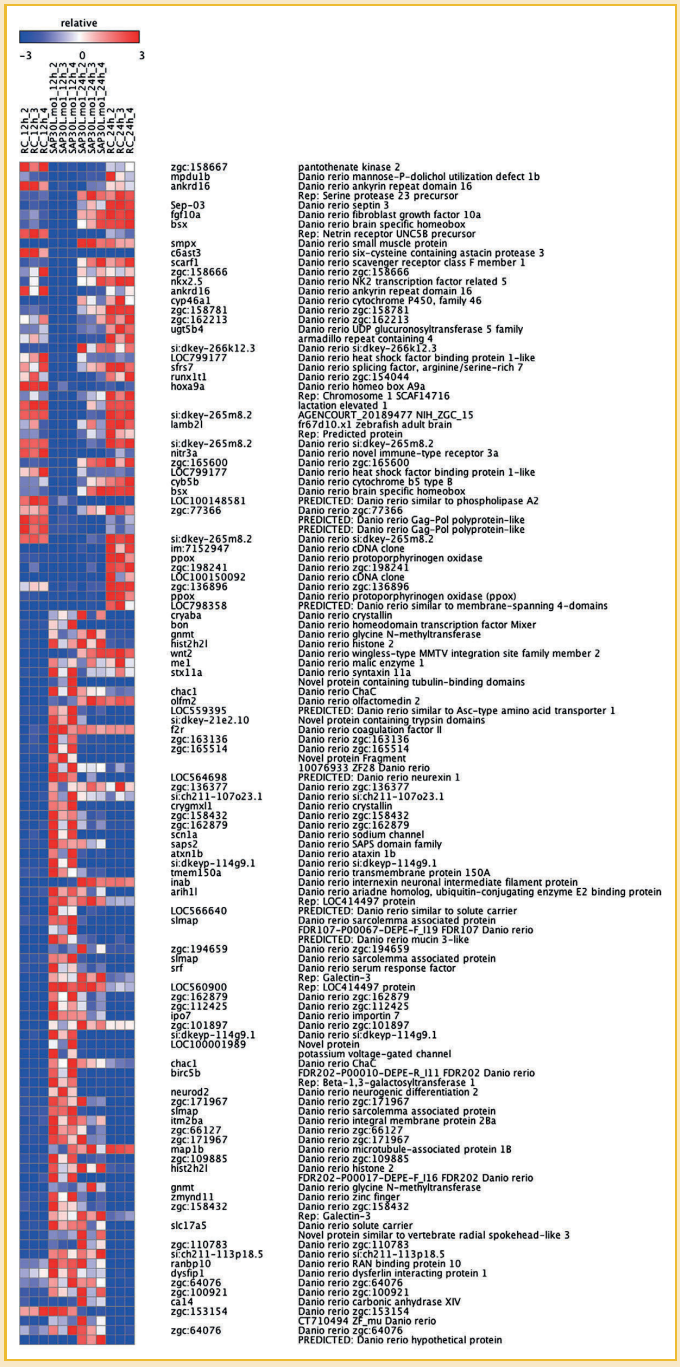


Fig. 4. Microarray analyses of zebrafish embryos (at 12 and 24 hpf) treated with SAP30L-MO1 or random control morpholino (RC). The 104 genes showing the largest differences in expression, ranked according to both the fold-change value and the *P*-value are shown for both time points and in all three parallel samples (2, 3, 4). Blue color denotes down-regulation, red up-regulation, and white no change. [Color figure can be seen in the online version of this article, available at http://wileyonlinelibrary.com/journal/jcb]

[Montgomery et al., 2007], HDAC3 [Farooq et al., 2008], and mSDS [Amsterdam et al., 2004].

GENES INVOLVED HEMOGLOBIN SYNTHESIS ARE DOWN-REGULATED BY SAP30L KNOCKDOWN

Consistent with the finding of diminished hemoglobin staining in the morphant larvae, we observed a marked reduction in expression of genes involved in hemoglobin synthesis and erythropoiesis (ppox, alas2, and ba1l globin-like gene) at 24 hpf (Fig. 4, Supplementary Fig. S5). The zebrafish ppox, a protoporphyrinogen oxidase, is expressed in the ICM as early as 22 hpf, similar to other genes involved in hematopoiesis, including alas2, gata1, and β-spectrin [Detrich et al., 1995; Brownlie et al., 1998; Liao et al., 2000]. Down-regulation of ppox was confirmed by qRT-PCR, which showed a sixfold decrease in its expression in SAP30L-MO1-treated embryos and a 4.8-fold decrease in SAP30L-MO2-treated embryos (Fig. 5c). In the future, it remains to be elucidated whether the lower hemoglobin levels observed after knockdown of SAP30L are due to impaired erythropoiesis, a block in heme biosynthesis, or some other mechanism. Our results suggest that one function of SAP30L may be to silence the expression of a factor that inhibits heme synthesis or erythropoiesis.

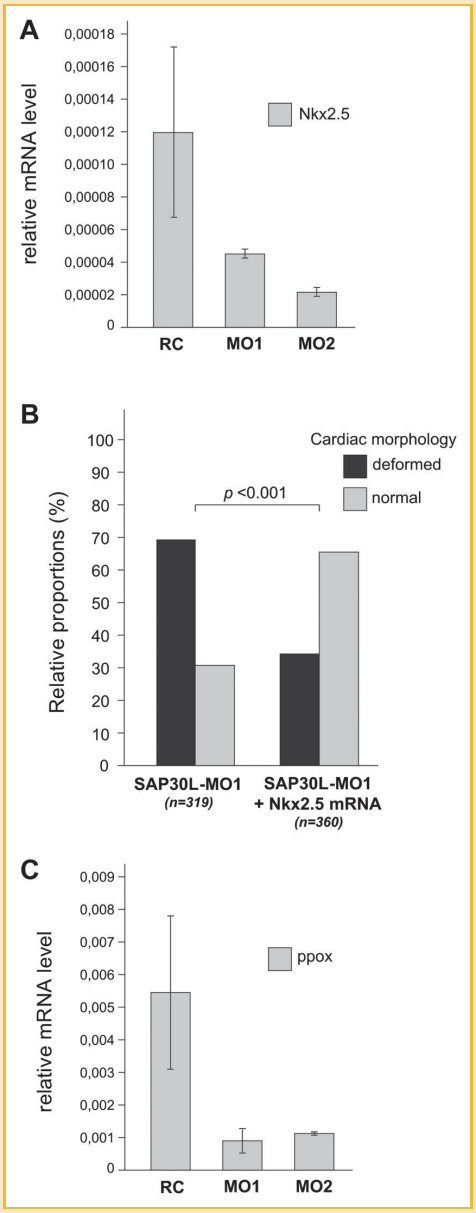


Fig. 5. Expression of Nkx2.5 and ppox mRNAs is down-regulated during SAP30L knockdown. A: Nkx2.5 mRNA expression is reduced in both SAP30L-MO1- and SAP30L-MO2-treated embryos at 12 hpf relative to RC-treated embryos. Nkx2.5 mRNA levels (mean ± SD) are expressed relative to those of the housekeeping gene, EF1A. B: The number of larvae exhibiting cardiac deformation is significantly reduced when treated with both SAP30L-MO1 (400 μM) and Nkx2.5 mRNA (20–50 pg per embryo), as compared to larvae treated with SAP30L-MO1 only. C: ppox mRNA expression is reduced in SAP30L-MO1- and SAP30L-MO2-treated embryos at 24 hpf compared to RC-treated embryos. ppox mRNA levels (mean ± SD) are expressed relative to those of the housekeeping gene, EF1A. The morpholino concentrations used in A and C: 400 μM for MO1, 600 μM for MO2, and 600 μM for RC.

CONCLUSION

SAP30 and SAP30L, members of the well-known Sin3A corepressor complex, have been previously implicated in repression of gene expression. However, the function(s) of these proteins in the setting of a living organism has not been established. Zebrafish was chosen as a model organism because it has only one of these proteins, SAP30L, enabling us to study its role without interference of potentially overlapping functions. Here we have demonstrated that SAP30L is ubiquitously expressed in the zebrafish and participates in the regulation of transcription of multiple genes. Morpholino-mediated knockdown of SAP30L protein results in disruption of cardiac development and function, and a reduction in hemoglobinization of red blood cells. Our results give important in vivo evidence of the role and function of SAP30L during the zebrafish embryogenesis.

ACKNOWLEDGMENTS

We are grateful to Mrs. Leena Mäkinen, Ms. Matilda Martikainen, and Ms. Sanna-Kaisa Harjula (IBT, University of Tampere, Finland) for technical assistance and advice. Dr. Annemarie H. Meijer (Leiden University, the Netherlands) is thanked for technical advice concerning WISH. The microarray analysis was carried out in the Finnish Microarray and Sequencing Centre at the Turku Centre for Biotechnology.

The zebrafish work was carried out in the Tampere zebrafish core facility supported by Biocenter Finland, Tampere Tuberculosis

Foundation and Emil Aaltonen Foundation. The funding sources had no involvement in the study.

REFERENCES

Alexandrovich A, Arno M, Patient RK, Shah AM, Pizzey JA, Brewer AC. 2006. Wnt2 is a direct downstream target of GATA6 during early cardiogenesis. Mech Dev 123:297–311.

Amsterdam A, Nissen RM, Sun Z, Swindell EC, Farrington S, Hopkins N. 2004. Identification of 315 genes essential for early zebrafish development. Proc Natl Acad Sci USA 101:12792–12797.

Brownlie A, Donovan A, Pratt SJ, Paw BH, Oates AC, Brugnara C, Witkowska HE, Sassa S, Zon LI. 1998. Positional cloning of the zebrafish sauternes gene: A model for congenital sideroblastic anaemia. Nat Genet 20:244–250.

Chen JN, Fishman MC. 1996. Zebrafish tinman homolog demarcates the heart field and initiates myocardial differentiation. Development 122:3809–3816.

Dannenberg JH, David G, Zhong S, van der Torre J, Wong WH, Depinho RA. 2005. mSin3A corepressor regulates diverse transcriptional networks governing normal and neoplastic growth and survival. Genes Dev 19:1581–1595.

David G, Grandinetti KB, Finnerty PM, Simpson N, Chu GC, Depinho RA. 2008. Specific requirement of the chromatin modifier mSin3B in cell cycle exit and cellular differentiation. Proc Natl Acad Sci USA 105:4168–4172.

Davidson AJ, Zon LI. 2004. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. Oncogene 23:7233–7246.

Detrich HW III, Kieran MW, Chan FY, Barone LM, Yee K, Rundstadler JA, Pratt S, Ransom D, Zon LI. 1995. Intraembryonic hematopoietic cell migration during vertebrate development. Proc Natl Acad Sci USA 92:10713–10717.

Eisen JS, Smith JC. 2008. Controlling morpholino experiments: Don't stop making antisense. Development 135:1735–1743.

Farooq M, Sulochana KN, Pan X, To J, Sheng D, Gong Z, Ge R. 2008. Histone deacetylase 3 (hdac3) is specifically required for liver development in zebrafish. Dev Biol 317:336–353.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch R, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH, Zhang J. 2004. Bioconductor: Open software development for computational biology and bioinformatics.

Goss AM, Tian Y, Tsukiyama T, Cohen ED, Zhou D, Lu MM, Yamaguchi TP, Morrisey EE. 2009. Wnt2/2b and beta-catenin signaling are necessary and sufficient to specify lung progenitors in the foregut. Dev Cell 17:290–298.

Goss AM, Tian Y, Cheng L, Yang J, Zhou D, Cohen ED, Morrisey EE. 2011. Wnt2 signaling is necessary and sufficient to activate the airway smooth muscle program in the lung by regulating myocardin/Mrtf-B and Fgf10 expression. Dev Biol 356:541–552.

Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA, Odenthal J, van Eeden FJ, Jiang YJ, Heisenberg CP, Kelsh RN, Furutani-Seiki M, Vogelsang E, Beuchle D, Schach U, Fabian C, Nusslein-Volhard C. 1996. The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. Development 123:1–36.

Korkeamaki H, Viiri K, Kukkonen MK, Maki M, Lohi O. 2008. Alternative mRNA splicing of SAP30L regulates its transcriptional repression activity. FEBS Lett 582:379–384.

Laherty CD, Billin AN, Lavinsky RM, Yochum GS, Bush AC, Sun JM, Mullen JM, Davie JR, Rose DW, Glass CK, Rosenfeld MG, Ayer DE, Eisenman RN. 1998. SAP30, a component of the mSin3 corepressor complex involved in N-CoR-mediated repression by specific transcription factors. Mol Cell 2:33–42.

Liao EC, Paw BH, Peters LL, Zapata A, Pratt SJ, Do CP, Lieschke G, Zon LI. 2000. Hereditary spherocytosis in zebrafish riesling illustrates evolution of erythroid beta-spectrin structure, and function in red cell morphogenesis and membrane stability. Development 127:5123–5132.

Lin SL, Li B, Rao S, Yeo EJ, Hudson TE, Nowlin BT, Pei H, Chen L, Zheng JJ, Carroll TJ, Pollard JW, McMahon AP, Lang RA, Duffield JS. 2010. Macrophage Wnt7b is critical for kidney repair and regeneration. Proc Natl Acad Sci USA 107:4194–4199.

Lindfors K, Viiri KM, Niittynen M, Heinonen TY, Maki M, Kainulainen H. 2003. TGF-beta induces the expression of SAP30L, a novel nuclear protein. BMC Genomics 4:53.

Lobov IB, Rao S, Carroll TJ, Vallance JE, Ito M, Ondr JK, Kurup S, Glass DA, Patel MS, Shu W, Morrisey EE, McMahon AP, Karsenty G, Lang RA. 2005. WNT7b mediates macrophage-induced programmed cell death in patterning of the vasculature. Nature 437:417–421.

Mathavan S, Lee SG, Mak A, Miller LD, Murthy KR, Govindarajan KR, Tong Y, Wu YL, Lam SH, Yang H, Ruan Y, Korzh V, Gong Z, Liu ET, Lufkin Tl. 2005. Transcriptome analysis of zebrafish embryogenesis using microarrays. PLoS Genet 1:260–276.

Montgomery RL, Davis CA, Potthoff MJ, Haberland M, Fielitz J, Qi X, Hill JA, Richardson JA, Olson EN. 2007. Histone deacetylases 1 and 2 redundantly regulate cardiac morphogenesis, growth, and contractility. Genes Dev 21:1790–1802.

Nakamura Y, Weidinger G, Liang JO, Aquilina-Beck A, Tamai K, Moon RT, Warman ML. 2007. The CCN family member Wisp3, mutant in progressive pseudorheumatoid dysplasia, modulates BMP and wnt signaling. J Clin Invest 117:3075–3086.

Pile LA, Spellman PT, Katzenberger RJ, Wassarman DA. 2003. The SIN3 deacetylase complex represses genes encoding mitochondrial proteins: Implications for the regulation of energy metabolism. J Biol Chem 278:37840–37848.

Pillai R, Coverdale LE, Dubey G, Martin CC. 2004. Histone deacetylase 1 (HDAC-1) required for the normal formation of craniofacial cartilage and pectoral fins of the zebrafish. Dev Dyn 231:647–654.

R Development Core Team. 2008. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rajagopal J, Carroll TJ, Guseh JS, Bores SA, Blank LJ, Anderson WJ, Yu J, Zhou Q, McMahon AP, Melton DA. 2008. Wnt7b stimulates embryonic lung growth by coordinately increasing the replication of epithelium and mesenchyme. Development 135:1625–1634.

Ransom DG, Haffter P, Odenthal J, Brownlie A, Vogelsang E, Kelsh RN, Brand M, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, Mullins MC, Nusslein-Volhard C. 1996. Characterization of zebrafish mutants with defects in embryonic hematopoiesis. Development 123:311–319.

Reamon-Buettner SM, Borlak J. 2010. NKX2-5: An update on this hypermutable homeodomain protein and its role in human congenital heart disease (CHD). Hum Mutat 31:1185–1194.

Robu ME, Larson JD, Nasevicius A, Beiraghi S, Brenner C, Farber SA, Ekker SC. 2007. P53 activation by knockdown technologies. PLoS Genet 3:e78.

Shu W, Jiang YQ, Lu MM, Morrisey EE. 2002. Wnt7b regulates mesenchymal proliferation and vascular development in the lung. Development 129:4831–4842.

Silverstein RA, Ekwall K. 2005. Sin3: A flexible regulator of global gene expression and genome stability. Curr Genet 47:1–17.

Smyth GK. 2005. Limma: Linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Huber W, Irizarry R, editors. Bioinformatics and computational biology solutions using R and bioconductor. NY: Springer. pp 397–420.

Sousa KM, Villaescusa JC, Cajanek L, Ondr JK, Castelo-Branco G, Hofstra W, Bryja V, Palmberg C, Bergman T, Wainwright B, Lang RA, Arenas E. 2010. Wnt2 regulates progenitor proliferation in the developing ventral midbrain. J Biol Chem 285:7246–7253.

Tang R, Dodd A, Lai D, McNabb WC, Love DR. 2007. Validation of zebrafish (*Danio rerio*) reference genes for quantitative real-time RT-PCR normalization. Acta Biochim Biophys Sin (Shanghai) 39:384–390.

Targoff KL, Schell T, Yelon D. 2008. Nkx genes regulate heart tube extension and exert differential effects on ventricular and atrial cell number. Dev Biol 322:314–321.

Thisse C, Thisse B. 2008. High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat Protoc 3:59–69.

Viiri KM, Korkeamaki H, Kukkonen MK, Nieminen LK, Lindfors K, Peterson P, Maki M, Kainulainen H, Lohi O. 2006. SAP30L interacts with members of the Sin3A corepressor complex and targets Sin3A to the nucleolus. Nucleic Acids Res 34:3288–3298.

Viiri KM, Heinonen TY, Maki M, Lohi O. 2009a. Phylogenetic analysis of the SAP30 family of transcriptional regulators reveals functional divergence in the domain that binds the nuclear matrix. BMC Evol Biol 9:149.

Viiri KM, Janis J, Siggers T, Heinonen TY, Valjakka J, Bulyk ML, Maki M, Lohi O. 2009b. DNA-binding and -bending activities of SAP30L and SAP30 are mediated by a zinc-dependent module and monophosphoinositides. Mol Cell Biol 29:342–356.

Wang H, Gilner JB, Bautch VL, Wang DZ, Wainwright BJ, Kirby SL, Patterson C. 2007. Wnt2 coordinates the commitment of mesoderm to hematopoietic, endothelial, and cardiac lineages in embryoid bodies. J Biol Chem 282:782–791.

Westerfield M. 1995. The zebrafish book: A guide to the laboratory use of zebrafish. Eugene OR: University of Oregon Press.

Yu J, Carroll TJ, Rajagopal J, Kobayashi A, Ren Q, McMahon AP. 2009. A Wnt7b-dependent pathway regulates the orientation of epithelial cell division and establishes the cortico-medullary axis of the mammalian kidney. Development 136:161–171.

Zhang Y, Sun ZW, Iratni R, Erdjument-Bromage H, Tempst P, Hampsey M, Reinberg D. 1998. SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. Mol Cell 1:1021–1031.