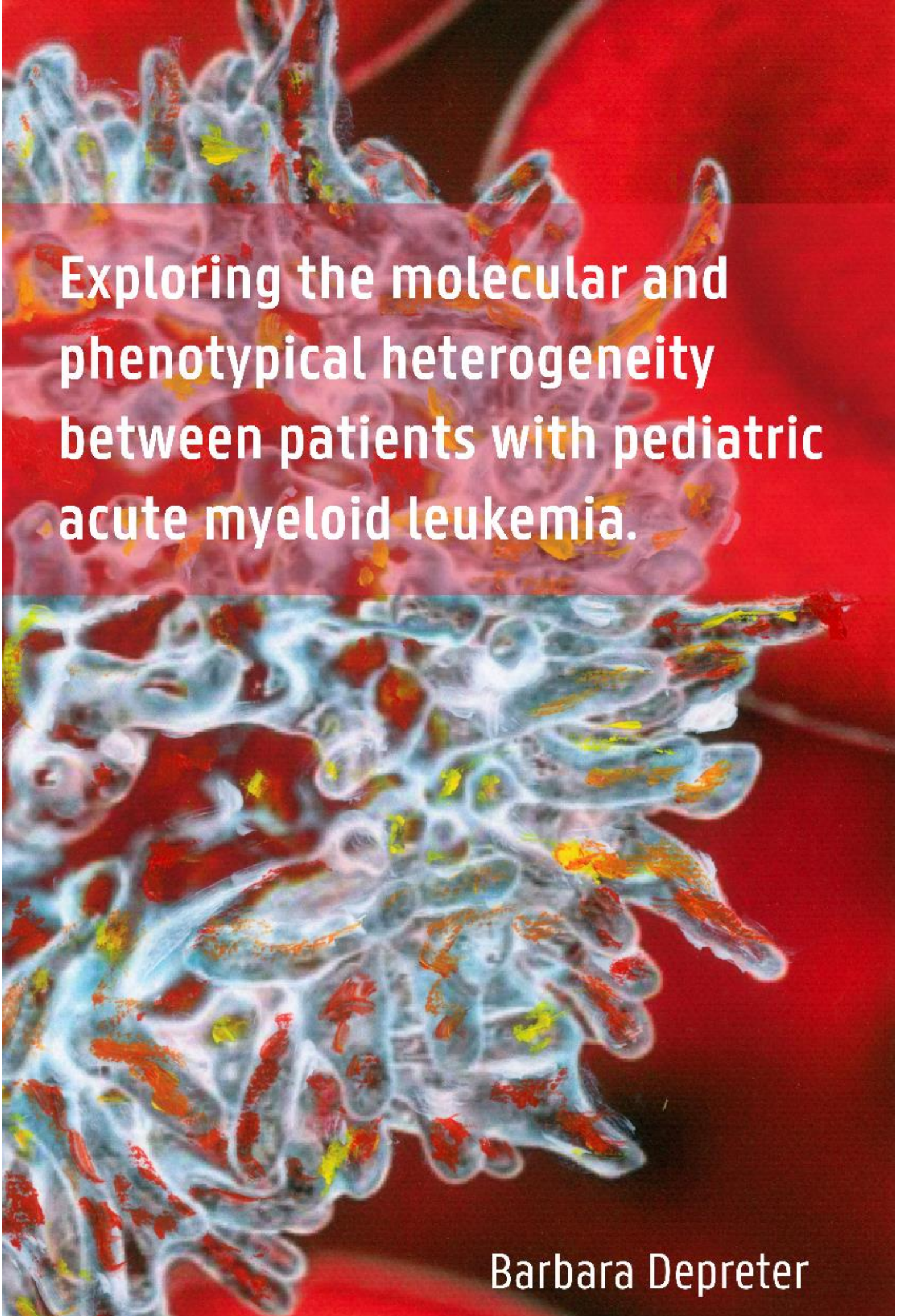

Exploring the molecular and phenotypical heterogeneity between patients with pediatric acute myeloid leukemia.
Barbara Depreter

FACULTY OF MEDICINE AND HEALTH SCIENCES

Exploring the molecular and phenotypical heterogeneity between patients with pediatric acute myeloid leukemia.

Word count: 107 477

Barbara Depreter
Student number: 00602798
Ghent University
Academic year: 2019 – 2020

Supervisors: Prof. dr. Tim Lammens, Prof. dr. Jan Philippé

A dissertation submitted to Ghent University in fulfilment of the requirements for the degree of Doctor in Medical sciences

Promotors:

Prof. dr. Tim Lammens

Prof. dr. Jan Philippé

Additional members of the Doctoral Advisory Committee:

Prof. dr. Barbara De Moerloose

Prof. dr. Bart Vandekerckhove

Prof. dr. Pieter Van Vlierberghe

Barbara Depreter was supported by a grant from the Research-Foundation-Flanders (Fonds voor Wetenschappelijk Onderzoek Vlaanderen, FWO).

This research was further supported by the Belgian Foundation against Cancer (grant 2014–265), FOD-KankerPlan (Actie29, grant to prof. dr. Jan Philippé) and vzw Kinderkankerfonds (grant to prof. dr. Tim Lammens).

The research in this thesis was conducted at Ghent University, Ghent, Belgium.

Members of the examination committee:

Prof. dr. Tom Taghon (Ghent University, Ghent, Belgium)

Dr. Julie Morscio (Ghent University, Ghent, Belgium)

Prof. dr. Steven Goossens (Ghent University, Ghent, Belgium)

Prof. dr. Katleen De Preter (Ghent University, Ghent, Belgium)

Prof. dr. Jacqueline Cloos (VU University Medical Center, Amsterdam, the Netherlands)

Prof. dr. Gert-Jan Kaspers (VU University Medical Center, Amsterdam, the Netherlands)

Prof. dr. Johan Vande Walle, chairman (Ghent University, Belgium)

Table of contents

Abbreviation list

ADxE	Ara-C, liposomal daunorubicin, etoposide
ADC	Antibody drug conjugate
AIET	AraC, idarubicin, etoposide and 6-thioguanine
ALDH	Aldehydedehydrogenase
Allo-SCT	Allogeneic donor stem cell transplantation
AM	Ara-C, mitoxantrone
AMKL	Acute megakaryoblastic leukemia
AML	Acute myeloid leukemia
AR	Allelic ratio
Ara-C	Cytarabine
B-ALL	B-cell precursor acute lymphoblastic leukemia
BiTE	bispecific T-cell engager
BM	Bone marrow
CAR	Chimeric antigen receptor
CB	Cord blood
CBF	Core binding factor
C-blast	Control myeloblast
cCAR	Compound CAR
CD	Cluster of differentiation
CEBPa	CCAAT/enhancer-binding protein alpha
CFC	Colony-forming cells
CIR	Cumulative incidence of relapse
CloEC	Clofarabine, etoposide, cyclophosphamide
CLP	Common lymphoid progenitor
CMP	Common myeloid progenitor
CN	Cytogenetic normal
CNA	Copy-number alterations
CNRQ	Calibrated normalised relative quantity
CNS	Central nerve system
COG	Children's Oncology Group
CR	Complete remission
CTL	Cytotoxic T-cell
CTNNB1	β-catenin
CvN	Cancer versus normal
DART	Dual-affinity re-targeting molecules
DB-AML01	Dutch-Belgian pediatric AML
DEG	Differentially expressed gene
DfN	Different from Normal
DFS	Disease-free survival
DxEC	Liposomal daunorubicin, etoposide, Ara-C
ECM	Extracellular matrix proteins
EFS	Event-free survival
ELN	European LeukemiaNet

EMP	Erythro-myeloid progenitor
EoBP	Eosinophil/basophil progenitors
EZH	Enhancer of zeste homolog
F	Female
FAB	French–American–British
FACS	Fluorescence-activated cell sorting
FDR	False discovery rate
FISH	Fluorescence in situ hybridization
FLA	Fludarabine, high dose HD Ara-C
FLA-Dx	Fludarabine, high-dose AraC and liposomal daunorubicin
FLT3	Fms-like tyrosine kinase receptor-3
FMO	Fluorescence-minus-one
FRβ	Folate receptor beta
FSC	Forward scatter
GEP	Gene expression profile
GMP	Granulocyte-macrophage progenitors
GM-SCF	Granulocyte-macrophage colony-stimulating factor
GPR56	G protein-coupled receptor 56
GSEA	Gene set enrichment analysis
HA	High dose cytarabine
HAE	High dose cytarabine and etoposide
HAM	High dose cytarabine and mitoxantrone
Hg	Hedgehog
HLA	Histocompatibility leukocyte antigen
HR	High risk
HSC	Hematopoietic stem cell
HSPC	Hematopoietic stem and progenitor cell
HTR	5-hydroxytryptamine (serotonin) receptors
IDT	Internal tandem duplication
Int	Intermediate
iPSC	Induced pluripotent stem cell
ITD	Internal tandem duplication
JAM-C	Junctional Adhesion Molecule-C
JMML	Juvenile myelomonocytic leukemia
LAIP	Leukemia-associated immunophenotype
L-blast	Leukemic blast
LFS	Leukemia-free survival
LIC	Leukemia-initiating cell
LMPP	Lymphoid-primed multipotent progenitor
lncRNA	Long non-coding RNA
LOH	Loss of heterozygosity
LSC	Leukemic stem cell
LT	Long-term
LT-IC	Long-term culture initiating cell
M	Male

mAb	Monoclonal antibody
MACE	Amsacrine, cytarabine, etoposide ± Gemtuzumab
MDS	Myelodysplasia
MEC	Mitoxantrone, etoposide and intermediate-dose Ara-C
MEP	Megakaryocyte-erythrocyte progenitors
MFI	Median fluorescence intensity
MLP	Multilymphoid progenitors
MNC	Mononuclear cells
mPBSC	Mobilized peripheral blood stem cells
MPP	Multipotential progenitor
MRD	Minimal/measurable residual disease
mRNA	Messenger RNA
MSC	Mesenchymal stem cell
MUT	Mutated
NBM	Normal bone marrow
NCI	National Cancer Institute
ncRNA	Non-coding RNA
NK	Natural killer
NKG2D-L	Natural killer group 2D ligand
NOD	Non-obese diabetic
NOPHO-DBH	Nordic society for paediatric haematology and oncology (NOPHO), BSPHO, DCOG, Estonia and Hong-Kong
NPM1	Nucleophosmin
NRQ	Normalised relative quantity
NS	NOD/SCID
NSG	NOD/SCID gamma
NSGS	NOD/SCID gamma-SGM3
NSS	NOD/SCID modified by expression of human cytokines
OS	Overall survival
P	P-value
PB	Peripheral blood
pCD	Plasmacytoid dendritic cells
pedAML	Pediatric AML
pLSC	Putative leukemic stem cell
PTD	Partial tandem duplications
qPCR	Real-time quantitative polymerase chain reaction
RBM	Reactive bone marrow
RFS	Relapse-free survival
ROS	Reactive oxygen species
RQ	Relative quantity
SCF	Stem cell factor
SCID	Severe combined immunodeficiency
scRNA-seq	Single-cell RNA sequencing
SP	Side population
SR	Standard risk

SSC	Sideward scatter
ST	Short-term
TARP	TCR γ chain alternate reading frame protein
TCR	T-cell receptor
TKD	Tyrosine kinase domain
TSG	Tumor suppressor gene
VAF	Variant allele frequency
VEGF	Vascular endothelial growth factor
VH	Variable heavy-chain
VL	Variable light-chain
WBC	White blood cell
WGS	Whole-genome sequencing
WHO	World Health Organization
WT	Wild type
WT1	Wilms' tumor 1
yr.	Year
yrs.	Years

1

CHAPTER I: Introduction

A part of the introduction was used for publication in the Belgian Journal of Haematology.

B. Depreter, B. De Moerloose, T. Lammens and J. Philippé*. Leukemic stem cells in AML: where are we now? An update on recent findings and detection. Accepted for publication as review in the Belgian Journal of Haematology.

(*shared senior authorship).

1. Haematopoiesis

1.1. Hematopoietic stem cell and hematopoietic tree

In the beginning of the 20th century, haematopoietic cells were postulated to consist out of various classes of progenitor cells that differentiate into mature white blood cells (WBCs) in a rigorously organised hierarchical structure, with one common ancestor, the hematopoietic stem cell (HSC) [1]. This hypothesis could explain the morphological diversity observed in peripheral blood by French and English morphologists as early as in the mid-1800s [2, 3]. The hypothetical existence of a HSC was supported half a century later by the observation that injection of non-irradiated donor cells is able to reconstitute normal haematopoiesis in lethally irradiated mice [4]. A decade later, direct functional evidence was provided by the development of clonal repopulation assays [5]. These foundations are still today the pillars of our current perception on haematopoiesis, with multipotent HSCs residing at the apex and terminally differentiated WBCs on the bottom **(Fig. 1)**.

During normal myeloid development, stem cells undergo a process of commitment to multipotential progenitors, which in turn give rise to mature blood cells. Studies on the regulation of normal myeloid development has showed that transcription factors play a major role in both myeloid differentiation and leukemogenesis. Some of these myeloid factors were already known as transcription factors, i.e. homeobox genes, and other have been identified through their abnormal expression and involvement in leukemia and leukemic translocations, i.e. PU.1, AML1 and PLZF [6]. The CEBP transcription factor plays a crucial role in haematopoiesis and is indispensable for granulocytic differentiation. Regulation of lineage-specific gene expression is known to occur via direct interaction with the basal transcriptional apparatus (TBF/TFIIB), but also via interaction with the SWI/SNF chromatin remodelling complex that epigenetically modulates gene expression. Several lines of evidence indicate a role for the retinoic acid receptor (RAR) in myeloid differentiation. Egr-1, on the other hand, was identified as an inducer of monocytic differentiation and exerts a strong interplay with WT1 in promoting monocytic differentiation.

Progenitor and mature WBCs can be distinguished from each other based on the differential expression of cluster of differentiation (CD) antigens measured by flow cytometry. Following terminal differentiation, lymphocytes, erythrocytes, granulocytes, monocytes and blood platelets move from the bone marrow (BM) to the lymphoid tissues and peripheral blood (PB).

HSCs include both long-term (LT) and short-term (ST) HSCs that are roughly defined by a CD34+/CD38-/CD90+ phenotype and harbour positive levels for CD117 (c-kit) and Sca1 (Ly-6A/E) [7-9]. Upon loss of CD90, HSCs differentiate into multipotential progenitors (MPPs). In the initial model of the hematopoietic tree **(Fig. 1)** [10], it was reasoned that maturation is strictly separated between myeloid-erythroid and lymphoid lineages, with a common lymphoid progenitor (CLP) and common myeloid progenitor (CMP) as respective ancestors [10, 11]. CMPs were further described to have granulocyte-macrophage progenitors (GMP) and megakaryocyte-erythrocyte progenitors (MEPs) as downstream progeny [10]. CD34+/CD38+ progenitors CLP, CMP, GMP and MEP can further be separated based on the differential expression of CD123, CD110 and CD45RA. It took until 2011 to discover the existence of an intermediate stage progenitor with lymphoid-, myeloid- and macrophage-differentiation capacity, defined as lymphoid-primed multipotent progenitor (LMPP) cells [12]. This seminal observation triggered a revision of the hematopoietic tree in 2013 [13], in which LMPPs give rise to multilymphoid progenitors (MLPs), capable of generating lymphocytes, monocytes and dendritic cells, and granulocyte-macrophage progenitors (GMPs), capable of generating neutrophils and monocytes/macrophages. Progenitors with megakaryocyte-erythroid potential are defined as erythro-myeloid progenitor (EMPs), and give on their turn rise to eosinophil/basophil progenitors (EoBPs) and MEPs, distinguishable based on CD133 expression [13].

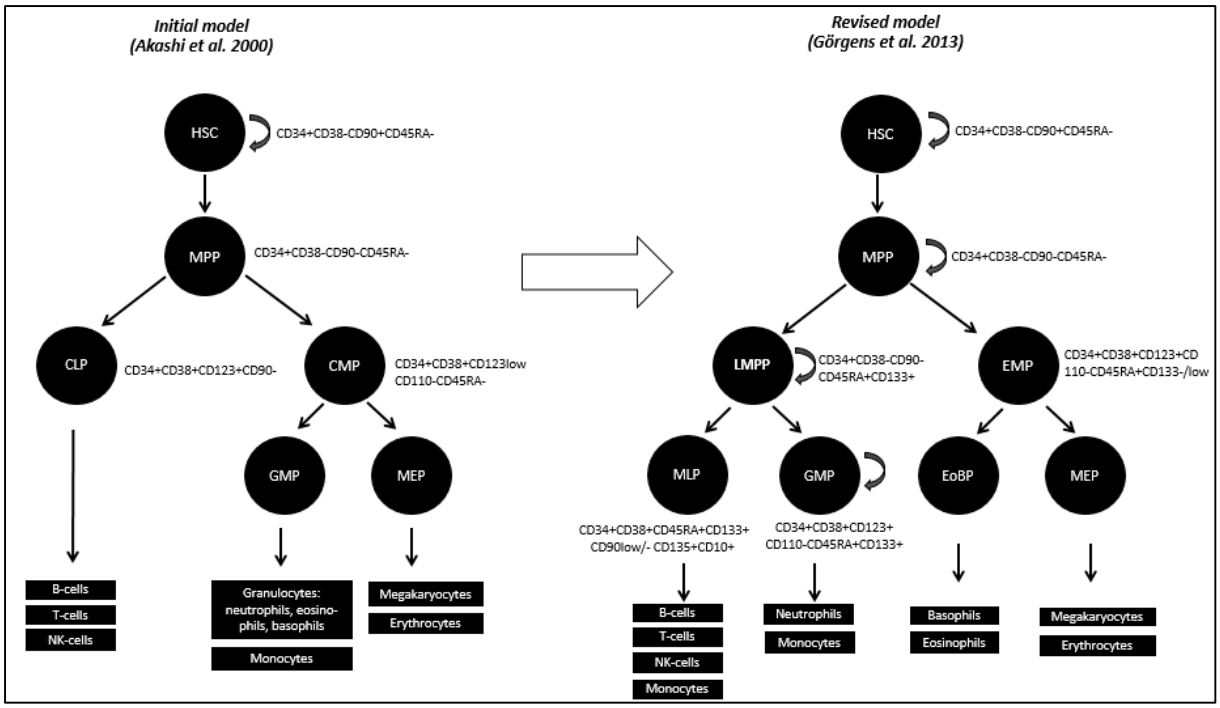


Fig. 1. Evolution of the hematopoietic tree. The seminal discovery of the existence of a LMPP imposed a revision of the hematopoietic tree, as conducted by Görgens et al. in 2013. Progenitors with oligo-/multipotential clonogenic capacity are illustrated by circles, terminally differentiated WBCs by squares. Whilst it was initially presumed that only HSCs have self-renewal capacities, indicated by recycling arrow symbols, leukemia-initiating capacities were later on attributed to HSC, LMPP, GMP and MPP progenitor populations (chapter 3.1.).

1.2. Hematopoietic stem cell niche

Haematopoiesis occurs in the marrow of axial, skeletal and long bones. These bones are characterized by the presence of supportive tissue, also referred to as the marrow “stroma”, “niche” or “micro-environment” **(Fig. 2)**. This micro-environment consists of different pawns in a large game of interacting networks capable of maintaining and regulating stem cell activity. The central piece is the perivascular mesenchymal stem cell (MSC), a multipotent bone marrow stromal cell that gives rise to a plethora of niche-regulating components e.g. fibroblasts, stem cell factor (SCF)-expressing cells, α-smooth muscle actin-expressing macrophages, non-myelinated Schwann-like cells, adipocytes (formation of adipose tissue), endothelial cells (formation of sinusoids), chondrocytes, myocytes and osteoblasts-osteoclasts that line the endosteal surface [14] [15]. One of the most important MSC progenitors is the CXCL12-abundant reticular cell that produces the essential cytokines SCF and CXCL12, and is capable of adipogenesis and osteogenesis [16, 17]. Cytokines, chemokines, adhesion factors and growth factors secreted from the micro-environment are crucial to maintain the equilibrium between proliferative, apoptotic and differention stimuli. CCL3, a proinflammatory cytokine, mediates osteogenesis [18] and the CXCR4/CXCL12 axis were shown to be crucial for homing based on the interplay between HSCs that are chemo-attracted to (CXCL)12-abundant reticular cells and MSCs. Adhesion molecules LFA-1, VLA-4, and VLA-5 support stem cell adhesion to fibronectin and are upregulated by the chemokine SDF-1 [19].

HSCs are non-randomly organised along the central bone axis, where they attach to osteoblasts aided by adherens junction molecules N-cadherin and β-catenin (defined as endosteal niches), or to

sinusoidal endothelium (defined as vascular niches) [14, 20]. Osteoblasts were shown to regulate HSCs in the stem cell niche by, among others, the expression of Notch ligand JAG1. Stem cells that home towards endosteal and vascular niches are most likely physiologically involved in different functions, e.g. transplantation stress (endosteal) and homeostasis (perivascular) [21]. Several lines of evidence showed that there exists only a limited number of niches, which need to be exhausted by myeloablative therapies or by competing exogenous stem cell populations, in order for new engraftment to take place [22, 23]. This view was recently challenged, as a large number of vacant HSC niches, located distantly from filled niches, were discovered to be available for supporting proliferation and engraftment of endogenous/exogenous HSCs in normal mice not exposed to myeloablation [24]. Research on the purpose of these 'back-up' niches may strengthen our understanding of physiological engrafting and increase future transplantation success rates.

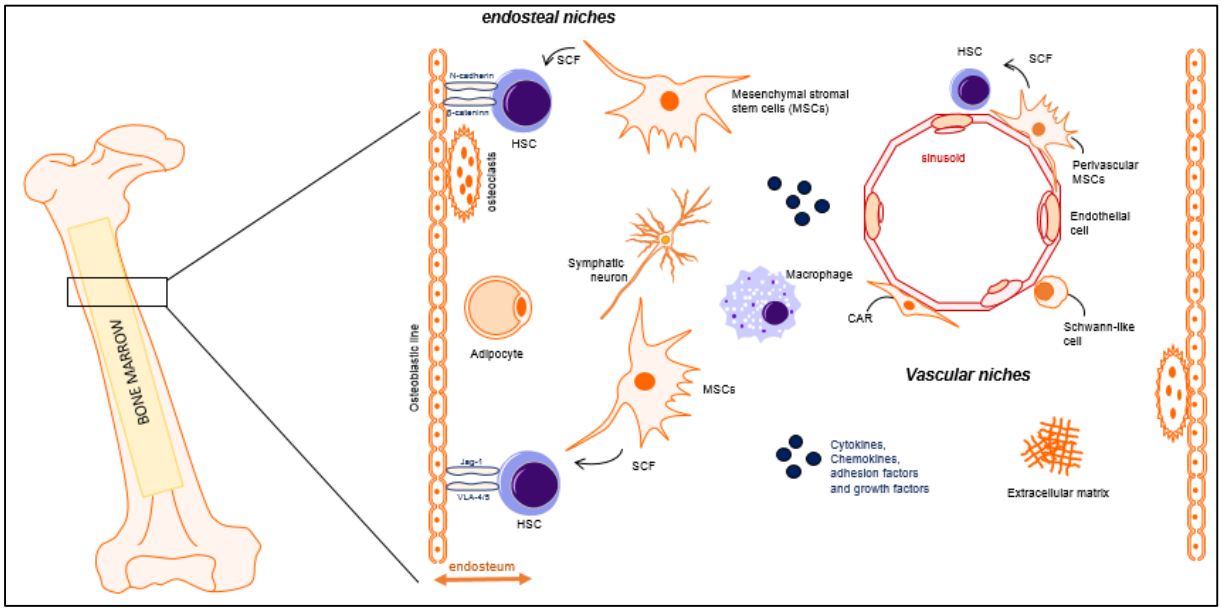


Fig. 2. Line-up of the bone marrow micro-environment. Supporting cells are coloured in orange, white blood cells in purple, sinusoids and endothelial cells in red and secreted supporting factors and linking molecules in dark blue. The mesenchymal stem cell (MSC) is crucial for maintaining and regulating stem cell activity in the micro-environment, and is defined as a multipotent bone marrow stromal cell that give rise to fibroblasts, SCF-expressing cells, α-smooth muscle actin-expressing macrophages, megakaryocytes, non-myelinated Schwann cells, adipocytes, arteriolar and sinusoidal endothelial cells, chondrocytes, myocytes, osteoblasts and osteoclasts.

2. Pediatric acute myeloid leukemia

2.1. Epidemiology

Acute myeloid leukemia (AML) is a heterogeneous haematological clonal disorder defined by the accumulation of immature myeloid cells in the BM, denominated as leukemic blasts (L-blasts). Pediatric AML (pedAML) accounts for 20% of all leukemia in children [25]. The incidence of childhood AML in the US is estimated at 7-8 cases per million (0-14 years) [26]. In the Netherlands and Belgium, approximately 30-35 children are annually diagnosed with AML, yielding an average of just over one per million [27].

The vast majority of the diagnoses occur *de novo* and the underlying reason remains unknown. However, in some cases, AML can develop secondary to another hematologic malignancy. Antecedent

myelodysplasia (MDS), i.e. refractory cytopenia or monosomy 7, and juvenile myelomonocytic leukemia (JMML) are prone to evolve into overt AML [28]. In addition, some pedAML patients are burdened by predisposition and environmental factors. Prenatal leukemia is characterised by high mortality rates [29], and efforts have been made to elucidate underlying predisposition patterns [30]. Germline predisposing gene mutations were found in 4.4 – 11.8% of the pedAML patients [30, 31]. Germline defects in DNA repair mechanisms (e.g. Fanconi anemia), telomere biology disorder, cell cycle and differentiation defects (e.g. Noonan syndrome), predisposition genes related to myeloid neoplasm, presence of hereditary transcription factor syndromes (e.g. *CEBPA* deficiency) and aneuploidy (e.g. Down syndrome) increase the risk for developing AML [31-33]. A recent study investigating the prevalence of predisposition gene mutations in a pediatric setting revealed a mean number of 0.6 mutated germline predisposition genes [31]. Patients also often co-carried somatic mutations (75.0%, 1.5 per patients) and in the vast majority of the cases showed abnormal karyotypes (37.5% complex karyotypes). Early-onset AML, also referred to as congenital AML, is defined by presentation within one month after birth with the exclusion of other differential diagnoses i.e. transient myeloproliferative disorders and leukemoid reactions [29]. Congenital AML is typified by hepatosplenomegaly, leukaemia cutis, high WBC counts, *KMT2A*-rearrangements and French–American–British (FAB) classifications M4/M5, but the underlying pathological mechanisms still needs to be unravelled [34]. Within this group, rare recurring abnormalities i.e. t(6;17)(q23;q11.2) have been reported [35]. The diagnosis of familial pedAML is appropriate when (i) one or more first-degree relatives have an AML anamnesis, (ii) patients have normal karyotypes, (iii) there is no evidence of chromosomal fragility and (iv) a single gene mutation was inherited in an autosomal dominant manner [36, 37]. Leukemic onset normally occurs between the age of 3-19 yr., and mutated *RUNX1* and *CEBPA* genes have been identified as causative in families inheriting AML. In addition, prenatal exposure to carcinogenic (i.e. pesticide propoxur) and therapeutic agents (i.e. topoisomerase II inhibitors) were shown to be culprit for *in utero* occurrence of type II mutations [33, 38].

2.2. Pathogenesis

Inspired by the double hit model of Knudson in familial cancer syndromes almost half a century ago [39], Gilliland and colleagues suggested that the key oncogenic events that trigger AML onset can be modelled in type I and type II mutations [40, 41]. Type II mutations concern non-random somatically acquired chromosomal aberrancies involving transcription factors, such as core binding factor (CBF) subunits *RUNX1* and *CBFβ*, the retinoic acid Receptor α (*RARα*), *HOX* family members, transcriptional coactivators e.g. *KMT2A*, etc. These events impair myeloid differentiation, block apoptosis, and are considered to be the initiating pre-leukemic lesion in AML. However, a secondary event is required to progress to overt AML, e.g. type I mutation. Type I mutations often reflect mutation hot spots in receptor tyrosine kinases (FMS-like tyrosine kinase-3 (*FLT3*) and *KIT*), receptor tyrosine kinases signal transduction target genes (*CBL*, *JAK2*), or in RAS pathway (*NRAS*, *KRAS*, *PTPN11* and *NF1*), and increase the proliferation and/or survival signaling of the cells [42]. Hence, class I mutations confer a proliferation or survival advantage, while class II mutations block differentiation and promote self-renewability.

It took another decade to build a comprehensive overview of the various type I and type II aberrations in pediatric acute myeloid leukemia (pedAML). Balgobind et al. evaluated the mutational status of type I and type II aberrations in children with AML, and evaluated their relation to clinical outcome [43]. *KMT2A*-rearrangements comprised the highest share of type II mutations, followed by t(8;21)(q22;q22), inv(16)(p13q22) and t(15;17)(q22;q21). *FLT3*-ITD mutations were identified as the most prevalent type I mutations, followed by *N/KRAS*, *KIT* and *CEBPA* double mutations. Remarkably,

no mutations were found below the age of two years (yr.), and type I mutations *NPM1*, *CEBPA* and *MLL*-PTD were mutually exclusive from the other cytogenetic subgroups. Type II fusion gene aberrancies were lacking in 30% of the total cohort, whilst 40% of the children older than two yr. did not harbour *FLT3*-internal tandem duplications (ITD), *FLT3*- tyrosine kinase domain (TKD) mutations nor type I aberrations in *NRAS*, *KRAS*, *PTPN11*, *KIT* or *WT1*. The authors revealed non-random associations between type I and II mutations, e.g. *KIT* mutations and CBF-AML, or *FLT3*-ITD and t(15;17)(q22;q21), as well as exclusivity, e.g. lack of type I aberrations in *KMT2A*-rearranged AML.

The following section summarizes the efforts made in unravelling the molecular landscape of pedAML. It must be noted that although numerous karyotypic, molecular and somatic alterations have been identified, that often cluster within distinct clinical entities, their prognostic value is not always clear. Also, patterns of coding messenger RNA (mRNA) and non-coding (microRNA (miRNA) and long non-coding RNA (lncRNA)) expression profiles exert significant impact on clinical outcome, and, although still in their infancy, will most likely gain more importance in the future.

2.3. The molecular landscape

2.3.1. Chromosomal aberrations

The chromosomal changes observed by karyotyping can roughly subdivide patients into good, intermediate and poor prognostic risk groups. Cytogenetic normal (CN) profiles are found in 15-30% of *de novo* pedAML and typified by an intermediate prognosis. In the majority of the cases (70-85%), chromosome abnormalities may be numerical, structural or both [37]. Imbalanced large chromosomal gains or losses mostly confer to a poor prognosis, with deletion of chromosome 5⁄del(5q), deletion chromosome 7⁄del(7q), del(9q), del(11q) and trisomy of chromosomes 8, 11, and 13 as most common [44]. Complex karyotypes, defined by three or more chromosomal abnormalities, contribute to an unfavourable risk stratification. Cell clones with complex karyotype were found in approximately 14% of pedAML patients, with the highest rate below the age of 3 yr. [45].

Recurrent genomic rearrangements **(Table 1)** are considered to be the most valuable prognostic determinant in pedAML, and can readily be identified through conventional karyotyping in about 50% of the patients [46]. Rearrangements of the *MLL* gene located on chromosome 11q23, defined as *KMT2A* by the updated HUGO gene nomenclature, are reported as the most frequent chromosomal aberrancy [46]. *KMT2A*-rearrangements characterize a high-risk (HR) pedAML subclass with clinical outcomes depending on their fusion partner. The five most prevalent *KMT2A* partner genes, i.e. *MLLT3*, *MLLT10*, *MLLT4*, *MLLT1*, and *ELL*, show slightly different incidences below or above the age of one yr. [47, 48]. Multivariate analysis of a large multicentre *KMT2A*–rearranged pedAML cohort (n=756) showed that only the fusion genes *KMT2A-MLLT11* (t(1;11)(q21;q23)), *KMT2A-AFDN* (t(6;11)(q27;q23)), *KMT2A-MLLT10* (t(10;11)(p12;q23)) and *KMT2A-ABI1* (t(10;11)(p11.2;q23) have a significant independent negative impact on prognosis [49]. The favourable prognostic impact of *KMT2A-MLLT3* (t(9;11)(p22;q23)) was found to be controversial [49, 50]. In addition, *MLL*-partial tandem duplications (PTDs) are rare in childhood, only identified above >1 yr., and confer to a worse prognosis with shortened overall survival (OS) and event-free survival (EFS) [47, 51]. Next, CBF leukemias account for 30% of all pedAML cases, and are characterized by gene fusions with CBF transcription factor subunits, e.g. *RUNX1-RUNX1T1* in t(8;21) and *CBFB-MYH11* in inv(16)/t(16;16) abnormalities. Together with *PML-RARα* fusion genes, these cytogenetic abnormalities are defined as good risk.

In addition to these four large groups, also more sporadically occurring recurrent chromosomal abnormalities have been described. Inv(3)/t(3;3)] abnormalities, with *EVI1* (*MECOM*) overexpression as molecular hallmark, occur in approximately 1% of the pedAML cases [46]. *EVI1* overexpression is present in 30% of the patients and identified as a compelling poor prognostic marker [52]. Implementation of *EVI1* overexpression can be helpful to further risk stratify CN pedAML patients. It was suggested, based on their stem cell promoting capacities in normal HSCs, that *EVI1* is involved in the quiescence maintenance and chemoresistance of leukemic stem cells (LSCs) [53]. T(1;22)(p13;q13) and t(6;9)(p23;q34) share a frequency between 0-3%, and are in the vast majority restricted to acute megakaryoblastic leukemia (AMKL). T(6;9)(p23;q34) has an unclear prognostic significance due to its low prevalence and significant association with other molecular abnormalities, i.e. *FLT3*-ITD. Seldom detected chromosomal abnormalities are t(8;16)(p11;p13), associated with FAB classifications M4/M5 and t(16;21)(p11;q22), also identified in AMKL.

Next to the group of recurrent abnormalities readily identified by karyotyping, cryptic rearrangements rarely occur in pedAML and require fluorescence in situ hybridization (FISH) for accurate diagnosis. Within this group, nucleoporin 98 (*NUP98*) gene rearrangements occur in 2-4% of the childhood AML cases, with tens of different fusion partner genes [54]. Most *NUP98*-rearranged pedAML cases have normal karyotypes, poor prognosis, and harbour in 70% of the cases a *FLT3*-ITD activating mutation [55]. The cryptic translocation t(5;11)(q35;p15.5) is the most prevalent *NUP98*-rearrangement (*NUP98-NSD1*), but often missed by cytogenetics due to its terminal 5q location [46]. Because of similar reasons, the cryptic t(7;12)(q36;p13) with fusion protein *ETV6-HLXB9* can be easily misdiagnosed as del(12p). T(7;12) almost exclusively occurs in infants and is associated with an extreme detrimental outcome (3-yr. EFS of 0%) [56].

In addition, also chromosomal instability may impact outcome. Genomic profiling showed that DNA copy-number alterations (CNA) and loss of heterozygosity (LOH) are uncommon in children. Novel pediatric-specific chromosomal CNA, including focal deletions in *MBNL1*, *ZEB2*, *ELF1* and *IL9R*, were identified through the Children's Oncology Group NCI Target AML Initiative [57]. The presence of CNA is significantly associated with decreased 3-yr. OS, EFS and a higher relapse risk [44].

Chromosomal aberrancy	Fusion protein	Incidence (%)	Age (yr.)	Relevant characteristics	Prognosis	References
t(8;21)(q22;q22)	*RUNX1-RUNX1T1*	7–16	0-18	CBF leukemia, correlated with FAB M2.	Good	[25, 58, 59]
inv(16)(p13q22)／t(16;16)(p13;q22)	*CBFB-MYH11*	3-12	0-18	CBF leukemia, pathognomonic of acute myelomonocytic leukemia with abnormal eosinophils (M4-eo).	Good	[25, 58, 59]
t(15;17)(q22;q21)	*PML-RARA*	2–10	0-18	Cure rates of 70–90%, higher incidence in Central and South America.	Good	[25, 58-60]
t(11q23) rearrangements	*KMT2A-x (>90 fusion partners)*	14–22	<1 (50%), 1-18 (50%)	Frequent CNS involvement and high WBC counts. A total of 94 fusion partners have been identified at the molecular level.	Depends on fusion partner	[25, 33, 47]
inv(3)(q21q26) / t(3;3)(q21;q26)	*RPN1-MECOM*	1–5		Also other 3q21–3q26 abnormalities have been described, e.g. t(3;21), 3q gain and 3q loss. Predominately associates with megakaryoblastic leukemia and trilineage dysplasia.	Poor	[61]
t(1;22)(p13;q13)	*RBM15-MKL1*	0-3	<2 (median 0.3)	Most specific and recurring chromosomal abnormality in pediatric AMKL (70 to 100% in infants).	Poor	[62]
t(6;9)(p23;q34)	*DEK-CAN*	1	2-18 (median 11)	Associated with mild to moderate bilinear dysplasia.	Intermediate/ poor	[60, 63, 64]
t(7;12)(q36;p13)	*ETV6-HLXB9*	3-5	<1	Cryptic in the majority of cases by conventional cytogenetic analysis, therefore often misdiagnosed as del(12p). Exclusively present in infant AML and associated with poor outcome (3- yr EFS of 0%).	Poor	[56]
t(8;16)(p11;p13)	*KAT6A-REBBP*	<1	0-17 (median 1.2)	Unique cytomorphological characteristics (parallel positive MPO and NSE staining and erythrophagocytosis) and immunophenotype.	Intermediate/ poor	[58, 59, 65, 66]
t(9;22)(q34;q11.2)	*BCR-ABL1*	1	0-18	Mostly M1 or M2 AML, chromosomal anomaly disappears during remission.	(very) poor	[67]
t(16;21)(p11;q22)	*TLS-FUS-ERG*	1-4	0-18	Leukemic blasts exhibit hemophagocytosis.	Poor/ unknown	[68, 69]
NUP98 (11p15) rearrangements	*NUP98-x (>20 fusion partners)*	2-4	1-17 (median 11)	Homeobox and non-homeobox Fusion partner genes. Characterised by high WBC counts and FAB M4/M5 classifications.	Poor/ unknown	[70, 71]

Table 1. Overview of the main cytogenetic aberrancies in cytogenetic abnormal pedAML patients. Other aberrancies, also occurring in cytogenetic normal patients, are discussed in 2.3.2.

2.3.2. Somatic mutations

Risk categorisation solely based on cytogenetics is insufficient, as 15-30% of the pediatric patients lack chromosomal aberrations [37, 72]. Efforts were made to better stratify CN pedAML using recurrent duplications and somatic mutations **(Fig. 3)**.

PedAML is characterised by a low somatic mutational burden, with a median rate from 0.17 mutations per million base pairs [73]. *NRAS*, *KRAS*, *FLT3*, *RUNX1*, *KMT2A*, *WT1*, *CBFB* and *KIT* mutations are the most prevalent ones, whilst *PTPN11*, *NPM1* and *CEBPA* mutations occur seldom. Co-existing *NPM1* and *CEBPA* double mutations were shown to be independent factors for favourable EFS [43, 74]. *FLT3*-ITD occur in 10-15% of the *de novo* pedAML and represents a highly significant, independent prognostic factor for detrimental outcome [75, 76]. Interestingly, patients with high allelic ratios (AR), defined as ITD-AR^{high}, were shown to have an even worse EFS compared to ITD-AR^{low} pedAML, based on a variant allele frequency (VAF) cut-off of 0.4 [77]. This correlation was confirmed in a large multicentre retrospective trial by the AIEOP AML-2002 study group, who showed that ITD-AR^{high} pedAML have a lower EFS compared to ITD-AR^{low} (19.2% vs 63.5%, respectively, cut-off set at 0.51) [75]. The impact of other ITD physical characteristics, e.g. clone size and number of clones, remains disputable. Very recently, targeted deep sequencing revealed that *FLT3*-ITD mutations in *de novo* CN pedAML are in the vast majority of the cases (33/34, 97%) accompanied by co-existing chimeric transcripts (19/34, 56%), e.g. *NUP98-NSD1*, and/or additional mutations, e.g. *MLL*-PTDs and somatic mutations in *NPM1*, *WT1*, *RAD21*, *CEBPα*, *IDH2*, *DNMT3A*, and *PLCG2* (27/34, 79%), with VAFs between 0.14-0.55 [78]. Mutations in the methylation cytosine pathway components, e.g. *DNMT3A* and *IDH1/2*, generally do not occur in pedAML or at very low frequencies, except for *TET2* [79]. *TET2* mutations by average occur in 1.4-6% of children with AML [79-81].

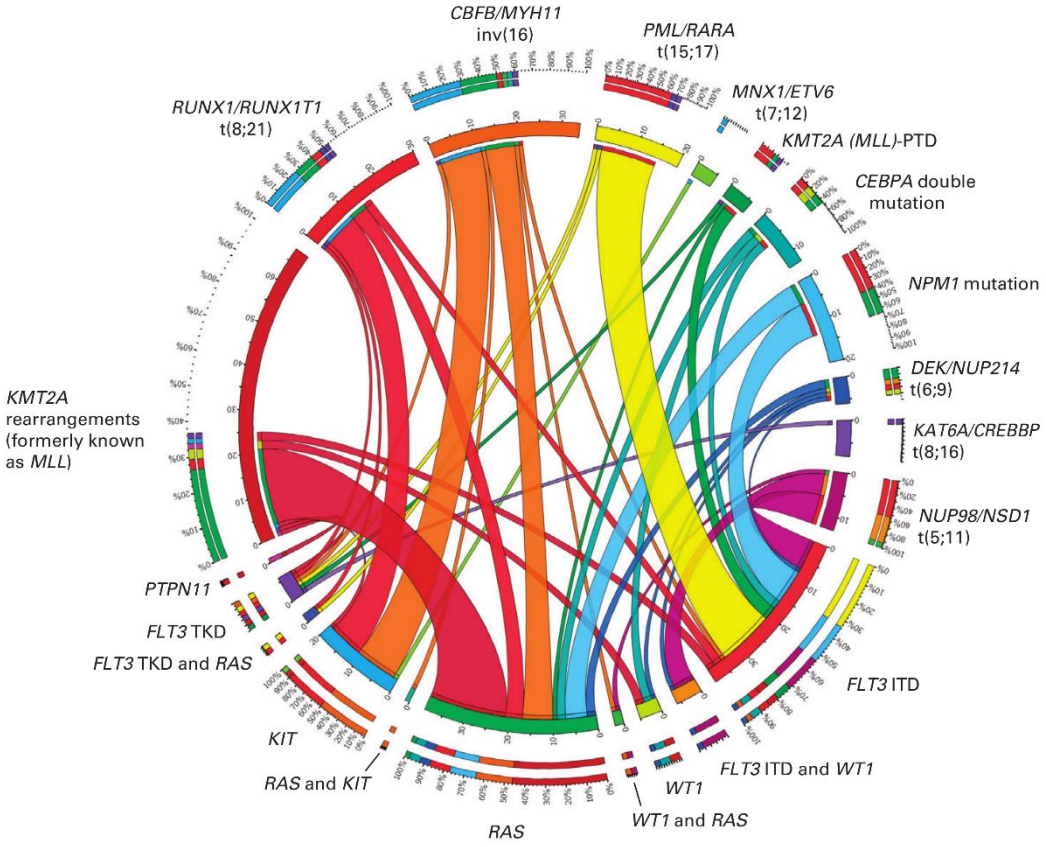


Fig. 3. Circos plot illustrating type I and II mutations in *de novo* pedAML patients, adapted from Creutzig et al. The length of the arch corresponds with the frequency of the type II mutation, and the width of the ribbon corresponds with the percentage of patients who have a specific type I mutation or combination of type I mutations.

Altogether, besides *FLT3*-ITD and double-mutated *CEBPA*, the prognostic impact of most other mutations remains unclear in a pediatric setting and hampers implementation in clinical practice.

2.3.3. Non-coding RNA landscape

Since merely 1–2% of the human transcriptome encodes for proteins, whilst 75% is transcribed into RNA, the majority consists of non-coding RNAs (ncRNA) with mainly regulatory functions amongst other in cell fate decisions [82]. NcRNA can roughly be categorised based on their length, into long ncRNAs (>200 nucleotides, abbreviated as lncRNA) and a heterogeneous group of small ncRNA (sncRNAs), comprising less than 200 nucleotides.

2.3.3.1. Small non-coding RNAs

The role in cell functionality for some of these sncRNAs has been well-known for decades as they are involved in the protein machinery, i.e. transfer RNAs (tRNAs); ribosomal RNAs (rRNAs), required for mRNA translation; small nuclear RNAs (snRNAs), essential for splicing; and small nucleolar RNAs (snoRNAs), involved in RNA modification. Others have only been discovered these last years to play essential roles in gene expression and transposon silencing, i.e. microRNAs (miRNAs), small interfering RNAs (siRNAs) and piwi interacting RNAs (piRNAs).

The miRNome has earned increased attention in pedAML this past decade. MiRNA molecules range between 21-25 nucleotides and are known to reside in introns of their pre-mRNA host genes and share their regulatory elements, primary transcript, and have a similar expression profile. They were shown to be involved in the pathogenesis of several molecular subgroups, play a role in chemoresistance and their expression has been correlated to survival [83-88]. Physiologically, miRNAs interact with mRNAs by complementary base pairing to 3'-UTR domains and can inhibit their translation with or without degradation, depending on the level of complementarity [89].

The aberrant expression of miRNAs can exert oncogenic or tumor suppressive effects on the leukemia, depending on the downstream targets. For instance, *EVI1* overexpression is associated with miR-9 promoter methylation and subsequent miR-9 downregulation [90]. MiRNA expression can hold prognostic information. Low expression of microRNA-34b was correlated to poor prognosis [91]. On the other hand, increased miR-155 expression was repeatedly found in *FLT3*-ITD mutated pedAML, although statistical significance lacked in one study [92], and correlated to OS and EFS [87], regardless of the *FLT3*-ITD mutational status [93]. Also high miR-196b expression was associated with dismal outcome [94]. The identification of pedAML-associated miRNA profiles, and their integration in a clinical setting as a prognostic parameter, was examined using a miRNA microarray platform [95]. Seventeen upregulated, e.g. miR-100, miR-125b, miR-335, miR-146a, and miR-99a, and 18 downregulated miRNAs were identified. Unsupervised miRNA clustering correlates with FAB classicisation, which suggest that miRNAs are differentially expressed depending on the degree of maturation of the leukemia, and with particular molecular subgroups, exemplified by miR-126 and *AML1-ETO*.

2.3.3.2. Long non-coding RNAs

While the role of miRNAs in leukemic haematopoiesis has been thoroughly investigated these last years, the relevance of lncRNAs is still poorly documented. Their low expression levels (often less than one copy per cell) and poorly conserved sequences have mislead many scientists [96]. LncRNAs may

act as competing endogenous RNA molecules at the level of translation and regulate protein expression [97]. They can also control gene expression in the nucleus by influencing the level of transcription, epigenetic modifications and chromatin state.

LncRNAs can be categorized in three different ways; (i) based on their location relative to the mRNA genomic locus (i.e. sense, antisense, intronic, bidirectional and intergenic); (ii) based on their special characteristics (i.e. enhancer RNAs, lncRNA-activating genes, transcribed ultraconserved regions, pseudogenes, telomere-associated ncRNAs and circular RNAs (circRNAs)); and (iii) based on their mode of action in a *cis*- or *trans*-regulatory mode [97]. In the first situation, lncRNAs recruit transcription factors and epigenetic modifiers to their locus, and enhance or repress expression of proximal coding genes, or, assist in the formation of chromatin conformations. In the second situation, lncRNA transcripts diffuse away from their loci and regulate distant genes, either by interfacing with nuclear regulatory complexes or by triggering the assembly of transcriptional machinery. However, clear division based on this latter feature has been blurred by recent evidence that lncRNAs may act in both *cis*- and *trans*-acting modes [98].

Although data are still limited, recent publications have led to the hypothesis that many lncRNAs are key regulators of development and play relevant roles in cell homeostasis and proliferation. It was shown that lncRNA can exert both tumor suppressive and oncogenic effects [96, 97]. A summary of 20 lncRNAs proven to play regulatory roles in (mainly adult) AML was recently compiled [97]. Among these, the lncRNA H19 was reported to exert a crucial role in *BCR-ABL* mediated leukemogenesis [99] and showed a significant association with intermediate karyotypes, *FLT3*-ITD, *DNMT3A* mutations and lower CR and OS rates [100]. Noteworthy, H19 was previously found to be overexpressed in about 50% of JMML patients, with expression levels regulated by *LIN28B* [101]. LncRNA expression data in pedAML are scarce, with only six publications since 2016 [97, 102-106]. Cao et al. detected and validated, through micro-array profiling and qPCR respectively, 12 up- and 12 downregulated lncRNAs between pedAML leukemic cells and healthy controls, that apparently were involved in cell cycle progression and immune response regulation [106]. Yin et al. [103] detected for a small cohort of 27 CN-pedAML that the expression of six lncRNAs negatively (4/6) or positively (2/6) associated with OS. Within this 6-lncRNA set, *CRNDE* expression also correlated with WBC counts and percentages of blasts in BM and PB. Lnc-SOX6-1 was shown to correlate with clinicopathological features, cell proliferation and apoptosis [104]. CASC15, previously annotated as LINC00340, appeared to be enriched in t(8;21)-translocated pedAML [105]. UCA1 expression increased upon treatment, as evaluated in 27 pedAML cases, and attributes to the chemoresistance of AML leukemic cells [102].

Altogether, although the relevance of lncRNAs in AML is booming, their prognostic relevance in a clinically well-defined and molecularly heterogeneous large pediatric cohort remains to be evaluated. One of the most promising newcomers are the circRNAs, lncRNA molecules that originate from reverse splicing in which the acceptor splice site located downstream binds to a donor splice site upstream [107, 108]. CircRNAs lack 5' or 3' ends and are hence resistant to exonucleases, which tremendously increases their stability and their ability for long-lasting interactions with miRNAs and proteins. Based on their high stability against degradation and high abundance in body fluids and exosomes, circRNA-based diagnostic and therapeutic strategies hold great promise in AML [108]. In order to move the field forward, streamlining of circRNA databases will be necessary [109].

2.4. Differences between pediatric and adult AML

These last years, collaborative studies have focussed on linking the degree of genetic heterogeneity of AML to clinical outcome [25, 57, 73]. The herein observed genetic and epigenetic differences within

between adults and children has highlighted their distinct pathogenesis (**Table 2**). One of the most important studies was published in 2012 by the Children's Oncology Group (COG)–National Cancer Institute (NCI) TARGET AML initiative [57]. The genomic landscape of pedAML was characterized using whole-genome sequencing (WGS) in 197 patients, and targeted capture sequencing at read depths averaging 500× for validation of mutations identified by WGS in 800 patients.

This study has shed light on some of the important similarities between pediatric and adult AML. First, the mutational rate of AML cells is lower than for most other cancers. Second, almost all samples had at least one mutation in one of the nine following genes: transcription factor fusions, NPM1, tumor suppressors, DNA-methylation-related genes, signaling mediators, chromatin modifiers, myeloid transcription factors and cohesin genes and spliceosome complex. Thirdly, there was a high degree in overlap between recurrently impacted genes, with functional cooperation between class II and class I mutations. Also critical differences were observed between pediatric and adult AML. As it is well-known that the mutational burden increases with age, overall somatic mutation frequency in pediatric AML was lower than that in adult patients. Fusions and focal copy number aberrations were more common in younger patients, while smaller sequence variants are more frequent in older individuals. The authors also identified novel pediatric-specific chromosomal copy number changes, including focal deletions in *MBNL1*, *ZEB2*, *ELF1* or *IL9R*. Furthermore, distinct combinations of co-occurring alterations, such as the *NUP98–NSD1* fusion were observed, significantly affecting disease. In spite of their recurrence, the frequency of certain mutations often differed between pediatric and adult patients. Alterations in signaling mediators (*NRAS* or *KRAS)* and the receptor tyrosine kinases (*KIT* or *FLT3*) are more prevalent in children, while *DNMT3A*, *IDH1/2*, *NPM1*, *TP53* and *CEBPA* mutations occur more in adults. More specific, very recent data based report on the occurrence of NPM1 mutations in 7.6% pedAML patients (n=869), compared to 25.4-41% in adult AML [110, 111]. Somatic *PTPN11* mutations are exclusively reserved for childhood leukemia [25, 112].

Also the prognostic impact can either differ depending on the age of onset. *KIT* mutations lack prognostic significance in a large series of pediatric patients with CBF AML [113], while contributing to a higher rate of relapse in CBF adult AML [114], though not at a significant basis in all studies [111]. *WT1* mutations had no independent prognostic significance in pedAML [115], but conferred to a poor prognosis in adult AML [116]. Biallelic *CEBPA* mutations, on the other hand, represent a beneficial outcome in both adult and pediatric patients [74, 111]. The prognostic value of *NPM1* mutations, remains controversial. In adult AML, two groups reported a good prognosis for *NPM1*-mutated patients in the absence of *FLT3*-ITD and *DNMT3A* mutations [111, 117]. This argument was advocated by Straube et al., who reported that *NPM1*-mutated AML patients > 60 yrs. may not be considered as favourable risk [118]. However, all studies agreed that co-occurrence of *NPM1* and *FLT3*-ITD mutations is translated into an unfavourable outcome. This latter finding was proven not to be true in a pediatric setting; Xu et al. [110] analyzed the impact of *NPM1* mutations in 869 pediatric AML patients from the TAGET dataset. They found that *NPM1* mutations confer an independent favourable prognostic impact in pediatric AML despite of *FLT3*-ITD mutations.

Hence, age-separated studies are absolutely required for risk stratification in contemporary pedAML treatment protocols. Subsequently, the World Health Organization (WHO) classification established in adult AML may not regardless be applied in a pediatric setting [119].

	Adult AML		Pediatric AML	
	Frequency/special features	Prognosis 5-y survival (age 18-60 y)	Frequency/special features/occurrences	Prognosis 5-y survival (age < 18 y)
Epidemiology	12 (20-39 y) to	—	2 (< 1 y)	—
Incidence rate (age-adjusted/100 person-years)[206]	30 (40-59 y)		0.7 (5-19 y)	
Biology				
De novo AML	83%	30%-40%	> 95%	60%-75%
Secondary (MDS-related) AML[207,208]	17%	Adverse	1%	Adverse
Abnormal karyotypes	55%	Depends on subgroups	70%-80%	Depends on subgroups
WHO classification				
AML with recurrent genetic abnormalities[14,15,209-211]				
t(8;21)(q22;q22)/*RUNX1-RUNX1T1*	8%	Favorable	12%-14%	Favorable
inv(16)(p13.1q22)/*CBFB-MYH11*	5%	Favorable	8%	Favorable
t(15;17)(q22;q21)/*PML-RARA*	5%-10%	Favorable	6%-10%	Favorable
t(9;11)(p22;q23)/*MLLT3-MLL*	2%	50%	7%	Intermediate or favorable (63%-77%)
t(10;11)(p12;q23)/*MLLT10-MLL*	1%	Unclear	3%, mainly infants	Adverse
AML with t(6;9)(p23;q34)/*DEK-NUP214*	1%	Adverse	< 2%	Adverse
AML with inv(3)(q21q26.2) or t(3;3)(q21;q26.2)/*RPN1-EVI1*	1%	Adverse	< 1%	Adverse
AML (megakaryoblastic) with t(1;22)(p13;q13)/*RBM15-MKL1*	< 1%	Unclear	AMKL only; infants	Intermediate
Provisional entity: AML with mutated *NPM1*[33,212,213]	35% (53% in CN)	Favorable in case of *FLT3*-ITD negative	5%-10% (14%-22% in CN), type A mutation dominant, increasing by age	Favorable
Provisional entity: AML with mutated *CEBPA*[212]	10% in CN	Favorable	5% (14% in CN)	Favorable
FLT3-ITD[27,30,213-215]	20%-40% (~ 50% in CN)	Adverse	10% (18% in CN)	Context dependent*
Not included in WHO 2008				
Adverse karyotypes/mutations	32%		5%-6%	
t(7;12)(q36;p13)/t(7;12)(q32;p13)	Not in adults	—	Infants	Adverse
t(5;11)(q35;p15.5)/*NUP98/NSD1*[22]	Not in adults	—	Mostly in CN	Adverse
New mutations				
N-RAS[35,40,216]	10%	No prognostic significance	20% (3% in CN)	No prognostic significance
MLL-PTD[45]	3%-5%	Adverse	3%	Not yet defined
c-KIT[40,43,87,217]	17% in CBF leukemia	Adverse in t(8;21)	25% in t(8;21)	Unclear
WT1[31,36,214]	10% in CN	Adverse combined with *FLT3*-ITD	13% in CN	Adverse combined with *FLT3*-ITD
PTPN11[42]	Not in adults	—	5%-21%, infants only	Unclear
IDH1/2[51]	16%	Context-dependent mutation	Rare, 2%-3%	No prognostic significance
TET2[218]	8%-17%	Unclear	Very rare	Unclear
DNMT3A[219]	20%	Unclear	Rare	Unclear
WHO classification continued				
AML with myelodysplasia-related changes[220,221]	48%	Unfavorable	Low	No prognostic significance
Therapy-related myeloid neoplasms[213]	6%	Adverse	3.5%	Adverse
Acute myeloid leukemia, not otherwise specified (NOS)[213,222,27]	~ 17%	Intermediate	~ 15%	Intermediate
Myeloid sarcoma (syn.: extramedullary myeloid tumor; granulocytic sarcoma; chloroma)	1%	—	2%-4%, leukemia cutis (infants), chloroma	Favorable in case of favorable karyotypes
Myeloid proliferations related to DS				
Transient abnormal myelopoiesis (syn.: transient myeloproliferative disorder)[223]	Not in adults	—	In 5% of the newborns with DS	Favorable
Myeloid leukemia associated with Down syndrome (ML-DS)	Not in adults	—	400-fold increased risk for AMKL	90%
Blastic plasmacytic dendritic cell neoplasia	Rare	Adverse	Not seen	—
Acute leukemias of ambiguous lineage				—
MPAL[134]	Rare	Adverse	4.5%	65%
Therapy				
Percentage in clinical trials	< 40%	—	50%-96%	—
HSCT in first CR	~ 60%	—	13%-30%	—
CNS	Not analyzed	—	Intrathecal prophylaxis	—

Table 2. Adapted from Creutzig et al. 2012. Comparison of biologic properties and genetic abnormalities in pediatric (children and adolescents < 18 yr.) and adult (age < 60 yr.) AML. Favourable indicates 5-year survival > 60% in adults and > 70% in children; intermediate, 23%-60% in adults and 50%-70% in children; and adverse, < 23% and < 50%, respectively, in children and adults. Favourable, intermediate, and adverse were defined according to the definitions referred to by Creutzig et al.—indicates not applicable; and AMKL, acute megakaryoblastic leukemia. *Different prognosis in combination with different other mutations.

2.5. Standard therapeutic regimens

In the early 1970s, the vast majority of pedAML patients failed to achieve long-term survival due to early and late treatment-related complications as well as refractory or relapsed disease [120]. Survival rates improved remarkably over the past decades, partly as a consequence of improved risk stratification [25, 121, 122]. The current mainstay for risk stratification and appropriate therapy allocation relies on a three-tiered approach, e.g. the cytogenetic and molecular subtype of the disease, the presence of acquired (coding) gene mutations, and the response to induction therapy by determining the measurable residual disease (MRD) status i.e. the minute 'rest' population of L-blasts [121, 122].

Determining the MRD status to evaluate chemotherapy response and to define remission was initially based on morphology [121, 123, 124]. Since 2004, a patient-tailored detection of a leukemia-associated immunophenotype (LAIP) of the L-blasts by flow cytometry (FCM) has become the new standard [122, 125]. Such LAIP consists of backbone markers to define the immature myeloid compartment, e.g. CD34 and CD117, in combination with aberrantly expressed (combinations of) markers that can be subdivided into 4 categories: (i) cross-lineage expression, (ii) asynchronous expression, (iii) lack and (iv) overexpression of markers [126, 127]. Another strategy, advocated by the Children's Oncology Group (COG), is the use of a different from normal (DfN) approach to avoid false negative results in case of phenotypic shifts (addressed in chapter 2.5.) [128, 129]. MRD detection by flow cytometry endows a higher sensitivity compared to morphology (range 10^{-3} down to 10^{-6}, depending of the total number of events measured) and is applicable in approximately 80% of the children with AML [125]. In about 20% of the cases, the immunophenotype of the leukemic blasts cannot be distinguished from normal myeloid blasts, mainly in monocytic AML, and is therefore non-informative [125]. Another high-sensitive technique is the molecular detection of MRD using real-time quantitative polymerase chain reaction (qPCR), or, next generation sequencing techniques [130].

Patient-tailored (de-)intensification of the chemotherapeutic regimen (number of days and/or doses) was able to decrease the number of fatal therapy complications e.g. infections, bleeding, graft-versus-host disease, secondary malignancies and cardiac failure [131]. Unfortunately, the achievement of complete remission (CR) is not consistently translated into cure. The CONCORD-2 study reported between 2004-2009 5-yr. survival rates in Belgium between 44–70%, depending on the risk group, with a median survival of 57% [132]. Low, standard en intermediate risk definitions differ between protocols and are therefore collectively categorised as standard risk (SR), comprising 30-40% of the total cohort [133]. Nowadays, therapeutic regimens for pedAML consist of 4-5 courses based on cytarabine and anthracyclines. In relapsed or refractory patients, two blocks of reinduction therapy and allogeneic donor stem cell transplantation (allo-SCT) directly or after consolidation has shown to be most effective [27]. The detailed protocol of the two most recent trials in Belgium, DB-AML01 [122] and NOPHO-DBH AML2012 protocol, is discussed further in detail below and illustrated in **Fig. 4**.

The DB-AML01 study (EudraCT 2009-014462-26) was in part based on the NOPHO 2004 study [125] with the following modifications: omission of one course of chemotherapy (high-dose AraC combined with mitoxantrone (HAM)) and hence reduction of the cumulative anthracycline dose and no allo-SCT in first CR. All patients received an initial six-day induction course with AIET (AraC, idarubicin, etoposide and 6-thioguanine). At day 15, patients with < 5% L-blasts were considered good responders and stratified for continuing therapy after full blood cell recovery. Children with BM L-blasts ≥5% on day 15 continued with the second course immediately unless life-threatening complications necessitated a delay. Second induction course consisted out of either AM (AraC combined with mitoxantrone, until April 2011), or FLA-Dx (starting from April 2011, fludarabine, high-dose AraC (HA) and liposomal daunorubicin) for patients with a blast count ≥5% at day 15 or a t(8;21) translocation. If remission was

obtained after the second induction course, three consolidation courses were given based on HA without anthracyclines. Etoposide was additionally administered in the first and the third consolidation course (HAE). Three-year EFS and OS were 52.6% (95% CI [42.9–61.3%]) and 74.0% (95% CI [64.8–81.2%]), respectively, with an overall remission rate equal to 93.5% [122]. The relapse rate after achieving CR was around 40%, and occurred averagely after nine months. About half of the children with relapsed AML could be rescued with second-line treatment including allo-SCT.

The ongoing NOPHO-DBH AML2012 trial (EudraCT 2012-002934-35) [134] initially consisted of two randomisation groups, DxEC (liposomal daunorubicin, etoposide, Ara-C) or MEC (Mitoxantrone, etoposide and intermediate-dose Ara-C). The first randomisation was closed earlier due to significant superior results of MEC and all patients now receive MEC. Flow cytometric MRD evaluation in BM is performed on day 22. Patients with >5% L-blasts proceed immediately to the second induction course, with randomisation between ADxE and FLADx. In patients with leukemic blasts <5% haematological recovery is allowed with weekly BM evaluation. After the second course, flow cytometric quantification of the L-blasts is again performed. Patients are defined as HR if they achieve CR after two induction courses (otherwise resistant disease) and have (i) *FLT3*-ITD/*NPM1* WT profiles, (ii) poor response after induction 1 (i.e. ≥15% leukemic cells at day 22 or at any subsequent evaluation prior to course 2) or (iii) intermediate response after induction 2 (i.e. 0.1%-4.9% leukemic cells before consolidation) [135]. Patients in CR with inv16(p13q22) and SR profiles receive consolidation by HA3E and FLA, whereas the other patients in CR receive HAM, followed by HA3E and FLA for SR patients or allo-SCT for HR patients. If CR is not achieved, salvage therapy is given. PedAML with resistant disease are eligible for the MACE study (amsacrine, cytarabine, etoposide ± Gemtuzumab). If patients respond, treatment should aim at allo-SCT, even for patients who respond poorly and continue to have significant residual disease.

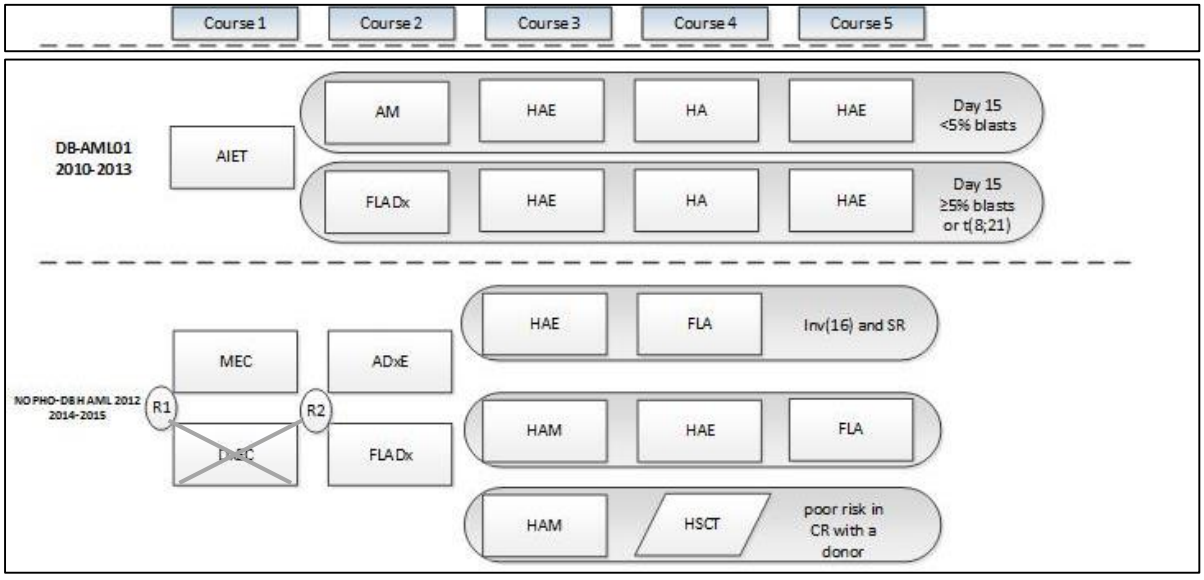


Fig. 4. Figure partly adapted from A. M. J. Reedijk et al., 2019, Leukemia [27]. Schematic visualisation of the DB-AML01 (top) and NOPHO-DBH AML2012 (bottom) treatment protocol. **AIET**, Ara-C, idarubicin, etoposide, 6-thioguanine; **AM**, Ara-C, mitoxantrone; **FLADx**, fludarabine, HD Ara-C, liposomal daunorubicin; **HAE**, HD Ara-C, etoposide; **HA**, HD Ara-C; **MEC**, Mitoxantrone, etoposide, intermediate-dose Ara-C; **DxEC**, liposomal daunorubicin, etoposide, Ara-C; **ADxE**, Ara-C, liposomal daunorubicin, etoposide; **HAM**, HD Ara-C and mitoxantrone; **ADxE**, Ara-C, liposomal daunorubicin, etoposide; **FLA**, fludarabine, high dose HD Ara-C; **FLADx**, fludarabine, HD Ara-C, liposomal daunorubicin.

Despite well-recognized improvements in outcome with current 5-yr. OS rates of 70%, still 30-40% of the good responders relapse [122]. Primary refractory and relapsed disease are held responsible for a significant proportion of the morbidity and mortality rates [136]. During the past decade, ample evidence was provided showing that relapse is promoted by the persistence of LSCs. Chapter 3 focusses on their role in AML, and gives an overview of the current-state-of-the-art methods for their detection, characterization and targeting by novel (immuno-)therapeutic strategies.

2.6. Heterogeneity, clonal evolution and single-cell analysis

PedAML patients who present identical cytogenetic and molecular profiles may still have variable outcomes. This phenomenon is due to the high heterogeneity of the disease. AML heterogeneity may be evident at diagnosis due to the co-existence of multiple (minor) subclones, or, can develop during treatment due to instability of the tumor clone **(Fig. 5)** [137]. Shifts between subclones and clonal progression (e.g. transformation from 'diagnostic clone' to 'relapse clone') can occur spontaneously or under therapeutic pressure, in which the most chemoresistant subpopulations will conquer. It is important to underline that type I mutations almost exclusively occur in independent clones, and the existence of two or more clones affecting the same signaling pathways is referred to as clonal interference. The presence of two or more independent type I mutations (e.g. *KIT*, *NRAS*, *KRAS*, *FLT3*, *JAK2* and *CBL*) in pediatric CBF-AML (n=73) independently adversely impacts EFS [42]. Larger cohorts of paired diagnostic-relapsed pedAML samples subjected to deep sequencing, able to discriminate clonal progression from subclonal shifts based on mutational VAFs, are needed to further elucidate the prognostic impact of high intratumoural heterogeneity at diagnosis.

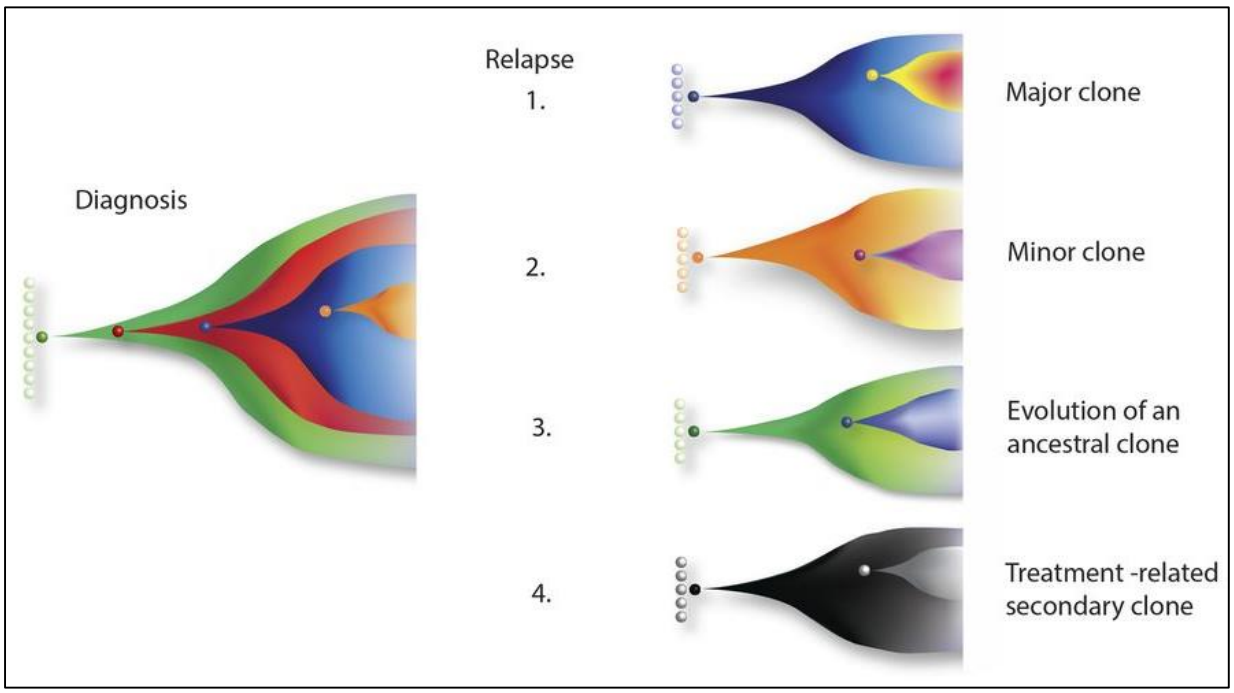


Fig. 5. Figure adapted from D. Grimwade et al., 2016, Blood [138]. Heterogeneous clonal patterns of relapse in AML. Potential patterns of relapse from the hypothetical tumor in panel A are shown in the Fish plots: 1, relapse of the dominant clone at diagnosis; 2, relapse of a subclone present at diagnosis; 3, relapse from an ancestrally related clone; 4, "apparent relapse," where the new tumor is not clonally related to the initial leukemia, such as might happen in therapy-related AML.

Cytogenetic, molecular and immunophenotypic shifts may be observed at relapse and render false negative results during MRD monitoring [137]. Only few studies have researched these shifts in a pediatric setting, and conflicting data have been reported.

Phenotypical shifts between diagnosis and relapse was evaluated in 16 pedAML patients [139]. Here, markers aberrantly expressed at diagnosis prevalently remained expressed at relapse, suggesting that aberrant expression extended to virtually all subclones within the leukemic cell population. Among others, the expression of CD96 and CLL-1 appeared to be highly stable. On the other hand, the emergence of novel markers (e.g. CD44) occurred in 7.7% of the cases, whilst lack of the initial aberrant marker (e.g. CD47) was observed in 5.4% of the cases.

Mutational shifts were also investigated in a limited set of paired diagnosis-relapse pedAML samples, and generally associate with a shorter time to relapse [140, 141]. Cloos et al. [140] revealed that *FLT3*-ITDs undergo notable changes at relapse in approximately one fifth of the pedAML patients (n=14/80). Within this group, *FLT3*-ITD mutated clones were gained, lost or experienced physical changes e.g. ITD length or number of clones. One decade later, Farrar et al. [142] supported these findings using whole exome capture sequencing (n=20). In their observations, clonal stability was predominantly determined by the VAF, as the majority of dominant variants (>0.4 for *FLT3*-ITD) persisted at relapse, whilst subclonal variants were lost. Besides *FLT3*-ITD, Bachas et al. showed that also other type I and II mutational gene expression profiles may alter upon relapse, as type I/II mutations appeared to be instable in 35% of the patients using mutation-specific techniques (n=23) [143]. In contrast to these data, Buelow et al. [78] recently reported a high clonal stability of mutational and fusion gene patterns in *FLT3*-ITD mutated CN pedAML using targeted deep sequencing. None of the ITDs disappeared at relapse, and *FLT3*-ITD VAFs tended to be lower at relapse in 5/7 patients. Comparably, the co-existing *WT1*, *RAD21*, *SMARCA2* and *TYK2* mutations present at diagnosis were without exception retained at relapse. However, care must be taken as the investigated patient cohort was very limited (n=7) and no details regarding ITD length or number of clones were shared.

So far, patterns that link immunophenotypic and molecular shifts in pedAML have not yet been identified. If so, flow cytometry could be of value to guide identification of molecular and cytogenetic aberrancies or vice versa. Meta-analysis is required to perform an integrative analysis of possible links between phenotypical, mutational and cytogenetic alterations.

Another approach to unravel the heterogeneous complexity of AML, and focus on subclonal evolutions on a per patient basis, is the integration of single-cell analysis. Gene expression data originating from bulk leukemic samples may (mis-)lead to “expression averaging” in case of heterogeneous cell populations, e.g. subclones, and heterogeneous sample mixtures, e.g. different cell types or cell developmental stages [144, 145]. Consequently, (high-throughput) single-cell RNA sequencing (scRNA-seq) platforms were developed and represent nowadays a powerful tool to investigate transcriptomic cell-to-cell variations, developmental processes and transcriptional stochasticity [146, 147]. The current available platforms differ in sensitivity, but all show an excellent accuracy [147]. A single-cell qPCR micro-fluidic platform was recently used to map clonal architecture in AML, and could differentiate secondary acquired mutations from shifted subclonal ratios [148]. However, the advances of scRNA-seq platforms has introduced new challenges, as these techniques are time-consuming, require high-expertise for data analysis i.e. building of high-dimensional data mining techniques and bioinformatic pipeline frameworks, and are expensive. The use of single-cell isolation methods prior to expression profiling, e.g. by microscopic manipulation or fluorescence-activated cell sorting (FACS), has become a widely used technique. However, antibody labelling and physical shearing during the sorting process causes stress to the cells, causing a significant cell and subsequent RNA loss [149].

3. Leukemic stem cells

3.1. Conceptualisation of the (non-)hierarchical leukemic model

The morphological and proliferation differences observed within the leukemic cells of AML patients had triggered hematopathologists to reason that leukemic cells, akin to the normal hematopoietic tree model, are structured in a hierarchical fashion [150, 151] **(Fig. 6)**. The first *in vitro* support for this hypothesis was provided in the late 1960's - early 1970's, based on the outgrowth of leukemic blast colonies in semi-solid media, defined as AML colony-forming cells (CFCs) [152, 153]. More than a decade later, the implementation of feeder layers in culture systems led to the identification of a CFC-subpopulation with lower frequencies but higher survival capacity, defined as long-term culture initiating cells (LTC-ICs) [154]. *In vivo* confirmation followed from serial xenotransplantation experiments, and led to the definition of a leukemia-initiating cell (LIC) [155]. Three years later, 2 to 8 weeks cultured CFCs and engrafted LICs were shown to mainly harbour positive CD34 expression and negative CD90 expression [156]. That same year, Bonnet and Dick [157] postulated the concept of AML being hierarchically organised with LSCs, defined as a rare CD34+/CD38- cell population, residing at the apex. Some months later, the addition of human cytokines to the feeder layers in the culture models enabled *in vitro* demonstration of LTC-ICs in AML [158].

The model of Bonnet and Dick, in which they define LSCs as scarce leukemic cells able to serial engraft based on their self-renewal capacities, and able to manifest leukemia by their proliferation/differentiation capacities, remains the cornerstone of our hierarchical AML model today. In addition, LSCs were endowed with multiple characteristics that are biologically crucial for disease propagation and chemotherapy resistance, e.g. relative quiescence by increased G0 phase, increased drug efflux and apoptosis-resistance mechanisms [159]. *In vivo* xenograft experiments performed in the 21th century confirmed that LSCs mostly reside in the CD34+/CD38- compartment [12, 159-161]. The CD34+/CD38- compartment contains both leukemic and normal stem cells, which can be segregated based on differential CD90 expression [156].

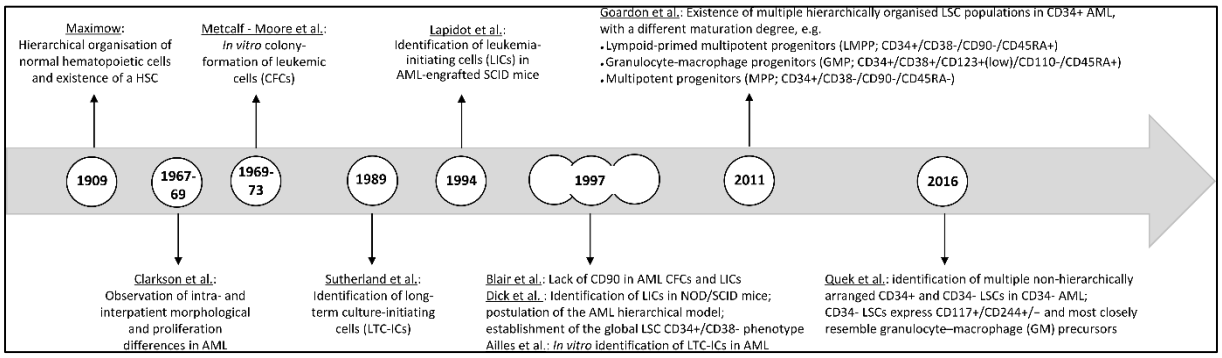


Fig. 6. Timeline describing the conceptualisation and evolution of the leukemic model in CD34+ and CD34- AML.

Initially, it was suggested that HSCs are the target population for leukemic transformation, as they share fundamental characteristics with LSCs, e.g. phenotype, scarcity and self-renewal ability [157]. In the following decade, conflicting data were published in view of this hypothesis. First, more mature GMPs were shown to produce LSCs in mice [162-164]. Second, LSCs identified in a blast-phase chronic myeloid leukemia (CML) patient were shown to resemble GMP cells [165]. Third, LSCs were generated from committed GMPs through the introduction of the *MLL-AF9* fusion protein and were shown to be capable of serial engrafting by the re-activation of self-renewal mechanisms, which originated the

infant idea of a “leukemic progenitor cells with stem cell potential” [162]. Fourth, also CD34+/CD38- and CD34+/CD38+ populations isolated from primary AML patients were able to engraft [166].

The major breakthrough was provided by Goardon et al., who evaluated the relation between normal and leukemic stem and progenitor cells in 74 primary diagnostic CD34+ AML patients and eight age-matched controls [12]. In 87% of the cases, LSCs demonstrated a ‘LMPP-like’ (CD34+/CD38-/CD90-/CD45RA+) phenotype, and the CD34+/CD38+ populations in this group was defined as ‘GMP-like’ (CD34+/CD38+/CD123+(low)/CD110-/CD45RA+) **(Fig. 7)**. In the minority of the cases, there was a ‘MPP-like’ LSC population (CD34+/CD38-/CD90-/CD45RA-), and the CD34+/CD38+ cells herein present shared the immunophenotype of normal CMPs. Segregation of AML patients based on LMPP-like and MPP-like LSC profiles did not correlate with morphology or genetic markers. They further focussed on the dominant group, and revealed that LMPP- and GMP-like populations are hierarchically organised, possess LIC activity and are able to serial engraftment in xenograft models, with LMPP-like LSCs being more potent than GMP-like LSCs. Although they could not be morphologically distinguished, both populations were genetically distinct and presented distinct gene expression profiles (GEPs), with LMPP-like LSCs resembling a more immature phenotype (clustering with AML FAB M0/1) than GMP-like LSCs (clustering with AML FAB M4/5). Moreover, the more immature and potent LMPP cells were shown to be capable of converting into more mature GMP-like LSCs, whilst the opposite versatility could not be demonstrated. This observation provided proof that LSC populations are hierarchically organised in CD34+ AML. The given that LSCs incline more towards progenitor cells than towards HSCs, both phenotypically and molecularly, implies that they most likely originate from progenitors that re-gain self-renewal capacities upon leukemic transformation. In concert with these findings, the LSC population was also shown to reside in the MPP compartment in a *TET2*/*FLT3*-ITD mutated mouse AML model [167]. However, conflicting data were published concluding that LSCs bear a higher molecular resemblance to HSCs, and the expression of genes found in progenitors is negatively correlated [161, 168]. In conclusion, further research is warranted to conclude whether leukemic transformation occurs in various stem and/or progenitor cells at different maturation stages, or, whether distinct LSCs with different phenotypes identified within one patient have evolved from one single LSC clone through the acquisition of additional hits.

Although the CD34+/CD38- compartment is most commonly ascertained to evaluate stem cell activity, *in vitro* and *in vivo* xenograft experiments showed that LSC activity also resides in CD34+/CD38+ and CD34- compartments [12, 160, 169]. Especially in CD34- AML, LSCs were shown to exhibit a highly heterogeneous phenotypic profile which can vary even within a single patient. Both CD34+ and CD34- cell populations, identified in AML cases with <10% or <1% CD34 expression, illustrated a significant leukemia-initiating potential in immunocompromised mice [166, 170]. Interestingly, CD34+ cells derived from a primary CD34- AML patient often lacked CD34 expression upon primary and secondary engraftment [166, 170]. Some even postulated that the CD34+/CD38- compartment in CD34- AML is of non-leukemic nature and that all LICs are concentrated in the CD34- compartment [171, 172].

Of utmost importance was the discovery by Quek et al. [173] who showed that, in contrast to CD34+ AML where one LSC population (~LMPP-like) may give rise to the other (~GMP-like), in CD34- AML, CD34+ and CD34- LSCs are non-hierarchically organised, co-exist next to each other, and can be regarded as nearly equivalent in terms of transcriptional and engraftment features. Hence, they propose two models in CD34- AML: (i) CD34- leukemic blasts arise from CD34+ stem or progenitor cells (~CD34+ LSCs) that lose CD34 expression through oncogenic events or induced by the micro-environment, and (ii) CD34- leukemic blasts derive from CD34- myeloid cells that have gained self-renewal potential (~CD34- LSC). Furthermore, they illustrated that CD34- LSCs most closely resemble

GM precursors and are situated in the CD34-/CD117+ compartment showing plasticity in CD244 expression **(Fig. 7)**. In contrast to the LSC populations described in CD34+ AML in terms of CD38, no versatility is observed between CD244+ and CD244- populations.

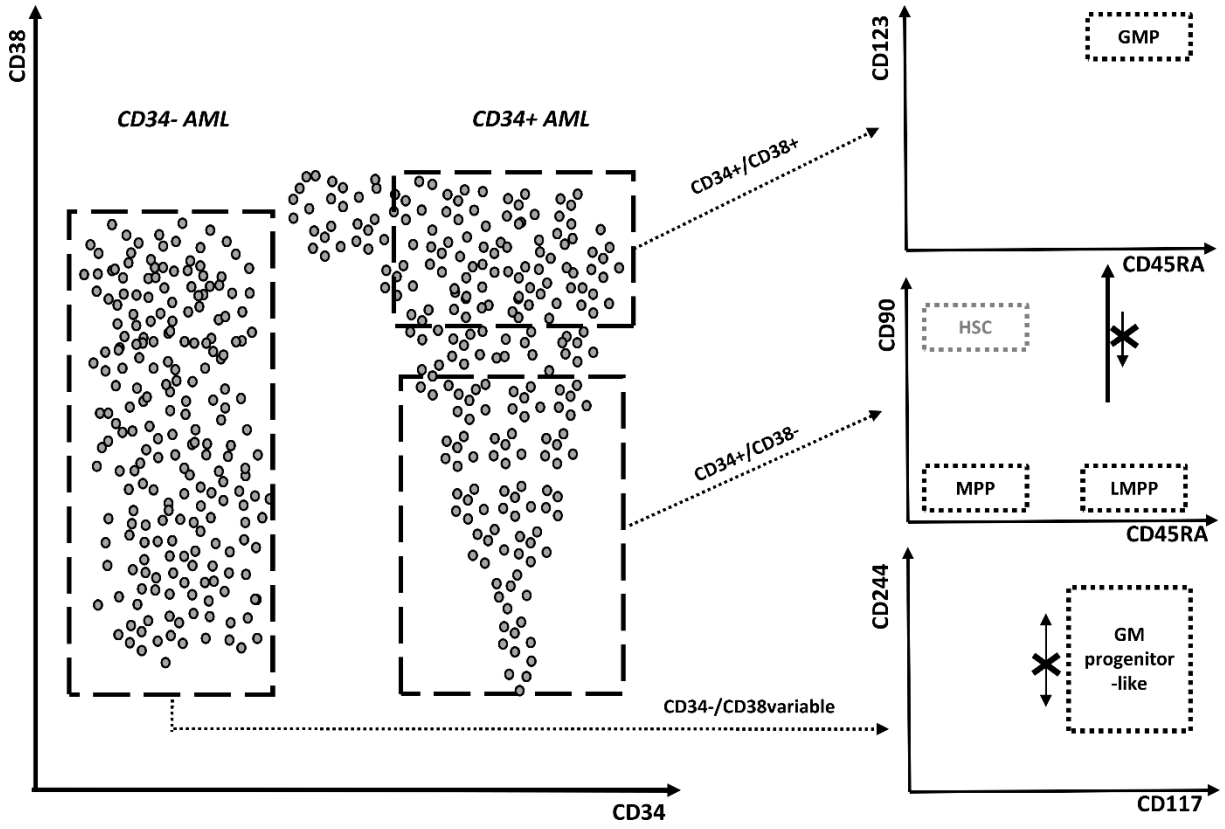


Fig. 7. Cartoon illustrating the distinct LSC phenotypes in CD34+ and CD34- AML. Bulk leukemic population for CD34- and CD34+ AML are depicted on the left side, LSC populations for each group separately on the right side. CD34+ AML is characterised by distinct hierarchically organised LSC phenotypes, among which the most dominant CD34+/CD38- LMPP-like LSC population can give rise to GMP-like LSCs, but not vice versa. CD34- AML is characterised by multiple CD34+ and CD34- LSC populations, and CD34- LSCs are characterised by CD117 expression and versatile CD244 expression. The phenotype of normal HSCs is depicted in grey. AML; acute myeloid leukemia; GMP, granulocyte-macrophage progenitors; LMPP, lymphoid-primed multipotent progenitor; MPP, myeloid progenitor cell; HSC, hematopoietic stem cell.

3.2. Detection and characterization of leukemic stem cells

3.2.1. Xenografting

The gold standard for detecting cells with genuine leukemia-initiating capacity in CD34+ and CD34- AML remains *in vivo* serial engraftment in immunocompromised mice. However, current xenotransplantation assays experience difficulties with engrafting sorted primary AML populations [174]. For instance, it was suggested that immune reactivity of the CD38 antibody used in earlier studies likely caused lack of engraftment of CD38+ populations [175]. Variable success rates between 0-70% were reported based on four different mouse models i.e. non-obese diabetic severe combined immunodeficient (NOD/SCID, NS) mice, NOD/SCID mice modified by expression of human cytokines (NSS), NOD/SCID gamma (NSG) mice and NOD/SCID gamma-SGM3 (NSGS) mice [12, 159, 176-178]. Collectively, it can be stated that 0.1% human leukemic chimerism, considered the clinically relevant cut-off for engraftment, is not reached in about half of the primary AML cases. Only one study reported 80% engraftment in mice (45/56 cases), most likely a biased result since relapsed samples (34%) or

samples from patients not in clinical remission after 1st cycle of induction (16%) were used in half of the cases [179]. Conflicting data were published concerning the impact of the innate immunity of the mouse model *in se* on engraftment success rates. One study showed that using NSG instead of NS mice did not increase the niche permissiveness for exogenous stem cells [176]. By contrast, others showed that NSG and NSGS mice initiate more successful grafts than NS(S) mice [177]. Increasing the number of cells to engraft, or applying a more intensive conditioning, appeared not to affect the engraftment potential [180]. Measuring PB or BM as matrix to evaluate engraftment generated the same conclusion [177]. Apart from the disappointing engraftment successes, a significant heterogeneity was noted in the phenotype and frequency of LSCs amongst patients and mice [179]. In addition, uncertainty exists on whether engrafted populations truly represent founder clones and whether they are responsible for relapse [181]. Hence, the relation between engraftment potential and clinical relevance of LSCs is difficult to elucidate. In spite of these question marks, one study did find a significant correlation between prognosis and the ability of AML to engraft in NS mice [180].

Cumbersome patient-derived xenograft mouse modeling, together with its accompanying long timeframe needed (8-20 weeks), has been held responsible for the slow progress made in understanding the self-renewal and chemoresistance mechanisms of LSCs. Hopes are now being set on *in vitro* culture systems that mimic the micro-environment and are able to maintain LSCs sufficiently long enough to allow self-renewal. The research group of Bonnet et al. developed a niche-like environment based on different stromal feeder layers and hypoxic culture conditions able to maintain LSCs for three weeks *ex vivo* [182]. An alternative niche-mimicking model was developed by researchers of the Bambino Gesù Children's Hospital. Their three-dimensional culture system consists of engineered hydroxyapatite and collagen I to imitate the micro-environment together with patient-derived MSCs and leukemic cells, and was able to maintain clonogenic potential for three weeks [183]. In this experimental set-up, the use of patient-derived MSCs was shown to have benefit over MSCs from unrelated donors [183]. Noteworthy, maintenance of bulk leukemic cells, including LSCs, up to six weeks in an *in vitro* stromal/LSC co-culture system using unrelated human MSCs was previously reported to be successful [184].

Although these novel niche-like *ex vivo* models do seem promising, they are still at an early stage. Therefore, more pragmatic surrogated LSC definitions, based on the immunophenotype and functional characteristics or molecular profile, are currently more used.

3.2.2. Flow cytometric detection

3.2.2.1. Membrane markers

3.2.2.1.1. LSC-specific and –associated markers in the $CD34^{+}/CD38^{-/dim}$ compartment

In spite of novel emerging data **(Fig. 7)**, the CD34+/CD38- compartment is still accepted to contain the highest fraction of LSCs, especially in CD34+ AML, but is also acknowledged to contain normal HSCs [156]. To allow for prospective discrimination between leukemic and normal stem cells in the CD34+/CD38- compartment, a plethora of differential membrane, cytoplasmic, and nuclear cell surface proteins that can easily be detected by flow cytometry, has been proposed **(Table 3)**. These markers are referred to as LAIP markers. Many of them acts as drug transporters, cytokine receptors and signaling pathway members and have been identified based on expression on cells with engraftment potential in immunodeficient mouse models or based on transcriptome profiling after cell-sorting.

Table 3. Current knowledge on 14 LSC-associated and –specific LAIP markers.

Marker	Synonym	Normal function	HSC expression	Expression in healthy tissue	Expression in adult LSC	Expression in pediatric LSC
CD45RA	Tyrosine phosphatase receptor type C (PTPRC).	220 kDa isoform of the leukocyte common antigen and member of the protein tyrosine phosphatase family, involved in regulating cell growth, differentiation, mitosis, and oncogenic transformation. Essential regulator of T- and B-cell antigen receptor signaling.	Negative, only on progenitor cells i.e. lymphoid-primed multipotential progenitors (LMPPs) [12].	Hematopoietic system (myeloid progenitors [12] subset of T-cells [185] (naive CD4+ and CD8+ T-cells, subsets of B- and NK-cells)), heart, nervous system, muscle, skin, bone, eye, kidney, pancreas and stomach.	By average 50-80% LSC, used to address the maturation stadium of LSC and to correlate with scatter properties [12, 186, 187].	Not reported.
CD56	Neural Cell Adhesion Molecule 1 (N-CAM1), MSK39.	Homophilic binding glycoprotein: role in immune surveillance (expansion T cells and dendritic cells) and development nervous system.	Absent [187, 188].	Hematopoietic system (NK-cells, activated T-cells), nervous system, muscle, liver, kidney, eye, heart, adrenal gland, thyroid gland and intestine.	31% AML [187].	Strong expression in leukemic cells, including LSCs, of pedAML with RAM phenotype (2-3%, n=821) [189], aberrant expression, frequencies not specified (undistinguishable from CD2, n=73) [190].
CD96	Tactile.	Member of the immunoglobulin superfamily, involved in antigen presentation and immune response.	Weakly expressed in adults (4.9 ± 1.6% [191], 4% ± 0.99% [192]) and children (20-28% [193]).	Activated T-cells, resting and activated NK-cells [194].	Various percentages reported: 65% of patients >40% CD96+ and 39% patients ≥90% CD96+ in CD34+/CD38- cells[191], 74 ± 25.3% [191], 34% [187], 7.32 ± 0.40% [192].	Significantly higher compared to NBM (n=12) [193].
CD7	GP40, TP41, LEU-9.	Transmembrane protein, member of the immunoglobulin superfamily. Role in T- /B-cell interaction during early lymphoid development.	Absent [187, 188].	Hematopoietic system (thymocytes and mature T cells).	43% [187], contributes to drug resistance and poor prognosis [192].	Aberrant expression, frequencies not specified (n=73) [190].
CD22	Sialic Acid Binding Ig-Like Lectin 2, SIGLEC-2.	Member of immunoglobulin superfamily, regulatory molecule that prevents the over activation of the immune system and the development of autoimmune diseases, inhibitory receptor for B-cell receptor signaling.	Absent [187, 188].	B-cells.	51% [187].	Not reported.
CD11b	Integrin Alpha-M, Mac-1 α subunit, CR3A, ITGAM, MAC-1, SLEB6, MO1A.	Responsible for cell-cell interactions, transmigration, phagocytosis of infectious agents [195].	Absent [187, 188].	Myeloid progenitors, monocytes, granulocytes, activated T-cells [195].	55%[187], bulk high expression significantly correlates with poor cytogenetic risk and therapy resistance[195].	Not reported.

TIM3	CD366, T-cell Ig mucin-3, Kidney Injury Molecule-3 (KIM-3), Hepatitis A Virus Cellular Receptor 2 (HAVcr-2).	Negative regulator Th-1 cell immunity, Toll-like receptor signaling (phagocytosis of apoptotic cells).	Absent [187].	Myeloid progenitors, reactive T-cells, dendritic cells, monocytes and macrophages, mast cells [194, 196].	Present in AML, except for AML M3[197-199], 62% (mainly CEBPA-mutated and CBF-AML[187, 199]) - 64%[187].	Not reported.
CLL-1	C-type lectin-like molecule-1, CLEC12A, CD371, Myeloid Inhibitory C-Type Lectin-Like Receptor (MICL), Dendritic Cell-Associated Lectin 2 (DCAL-2).	Involved in cell adhesion, cell-cell signaling, glycoprotein turnover, and roles in inflammation and immune response[194].	Virtually absent [188]; low expression in NBM, RBM, and mPBSC [200].	Restricted to the myeloid cell lineage, including progenitors [201].	70-87%[186, 200] [187, 188, 200, 202, 203], stable at relapse (n=6) [188].	71.4% (n=7), 85.7% stability at relapse [204].
CD123	IL-3 receptor α chain (IL-3Rα).	Part of the heterodimeric IL3R cytokine receptor, required for IL-3 activity, controlling proliferation and differentiation of hematopoietic cells[205]	Depends on the control population: Absent or weak (12%, n=6) in adult HSC [187, 188, 206, 207], 50% in regenerating HSC [188], high expression in CB [208], 15-20% in pediatric HSC [193].	Hematopoietic system (restricted to the myeloid cell lineage, including progenitors [201]), and nervous system.	Substantial higher (66-100%) [187, 206, 207, 209, 210], though not always statistical significant. Associated with *FLT3*-ITD and/or *NPM1* mutations, poor prognosis, adverse outcome[188, 211, 212] , with lower CR rate and shorter survival[211], and with constitutionally active NFκB[213]/STAT5 [212] activity. Also in combination with hMICL positivity significzant impact on DFS.[214, 215] By contrast, reported to have no significant impact on OS [210]. Stable at relapse (n=6) [188].	Aberrant expression in 67% (n=73) [190].
CD49d	α chain of α4β1 integrin (VLA-4), α4 integrin, ITGA4.	Cell-surface receptor involved in adhesion, i.e. niche homing of HSCs by binding stromal adhesion molecules fibronectin and VCAM-1 [19, 216, 217], and by migration through CXCR-4 cooperation [218].	Abundant expression.	Ubiquitously expressed.	Not reported, only in bulk (high expression negatively impacts OS and CR [219-221])	Not reported, only in bulk (low expression negatively impacts DFS [222]).
CD2	T-cell surface antigen T11/Leu-5, LFA-2, LFA-3 receptor.	Cell adhesion molecule, aids in immune recognition.	Absent [187, 188].	Hematopoietic system (restricted to the T-cells and NK-cells).	Present [187, 188].	Aberrant expression, frequencies not specified (undistinguishable from CD56) [190].
CD15	Fucosyltransferase 4, ELAM-1 Ligand Fucosyltransferase (ELFT), (FCT3A).	Involved in neutrophil functions, i.e. cell-cell interactions, phagocytosis, stimulation of degranulation, and respiratory burst, although the function of CD15 is not clear [223].	Absent [187].	Central nervous system (neurons, glial cells)[223, 224], stomach, epithelial cells of intestinal tissues and hematopoietic system (subset NK cells and (activated) T-cells, all myeloid	30%[187], expression in bulk correlated with monoblastic phenotype [227].	Not reported.

				cells from promyelocytes onwards) [225, 226]		
NG2	Neural/glial antigen 2, Chondroitin Sulfate Proteoglycan (CSPG4).	G2 proteoglycan promotes endothelial cell motility and angiogenesis via engagement of galectin-3 and alpha3beta1 integrin.	Minority of CD34+CD38+ HPCs from CB (2.1±2.4%), NBM (0.83±2.2%) and mPB (1.3%), and CD34+ pDC [228].	Niche mesenchymal stromal cells and pericytes [229], oligodendrocyte progenitor cells [230], esophagus, placenta, uterus and malignant melanoma cells.	Not reported, only in bulk: ~90% IN 11q23/MLLr leukemias [231].	Not reported, only in bulk: ~90% IN 11q23/MLLr leukemias [231].
HLA-DR	Major Histocompatibility Complex, Class II DR	Central role in the immune system by extracellular peptide presentation.	High expression [188].	Ubiquitously expressed.	Aberrant if absent [187, 188, 201, 208].	Not reported, only in bulk (absent in pedAML with RAM phenotype (2-3%, n=821)[189].

Legend to Table 3. pedAML, pediatric AML; NBM, normal bone marrow; RBM, reactive bone marrow; mPBSCs, mobilised peripheral blood cells; mPB, mobilised peripheral blood; CB, cord blood; pCD, plasmacytoid dendritic cells.

However, it should be noted that the quest for the Holy Grail of the LSC markers is still ongoing, as no marker has yet been identified that ticks all the boxes: (i) provide a clear-cut discrimination between leukemic and normal stem cells within the CD34+/CD38- compartment, allowing a correct correlation between LSC load and prognosis (ii) universal expression in all LSCs, (iii) exclusive expression in LSCs whilst absent in leukemic blasts, (iv) able to identify LSCs regardless of the inter- and intra-patient intrinsic heterogeneity that characterises AML and (v) show a stable expression over time and do not fluctuate upon relapse. Taking all these criteria into account, broad application of a single target antigen for diagnostic purposes, as well as for future LSC-directed antibody therapies, seems highly unlikely.

3.2.2.1.2. LSC markers outside the $CD34^{+}/CD38^{-/dim}$ compartment

Quek et al. showed that CD34 is not a fixed marker for leukemia-initiating capacity, as in CD34- AML, both CD34+ and CD34- LSC populations exhibited a similar frequency of engraftment [173]. Goardon et al. [12] illustrated that both CD38+ and CD38- LSC populations exist and can initiate leukemia. Hence, insufficient characterisation of the LSC compartment by the CD34+/CD38- phenotype prompted to pursue a search for new surrogate membrane markers that are also readily identifiable by flow cytometry. Below, we summarize recent findings concerning three novel candidate LSC membrane markers that, regardless of the CD34/CD38 expression profile, are suggested to (i) be able to capture all leukemia-initiating cells, (ii) discriminate stemness states between normal and leukemic stem cells, and (iii) (partly) relate to the chemoresistant capacities of LSCs.

First, the G protein-coupled receptor 56 (GPR56) encodes a member of the G protein-coupled receptor family and regulates diverse neurobiological processes. The first association between GPR56 and stem cell biology was made in the neurological system, showing the highest expression in neuronal stem and progenitor cells but a marked decreased expression in mature differentiated cells i.e. astrocytes [232]. Within the central nerve system (CNS), GPR56 was demonstrated to bind the ligand collagen III, which plays a role in cortical development and lamination [233, 234]. As collagen III is also a crucial ECM protein involved in niche homing of leukemic and normal stem cells, Saito et al. [53] attributed GPR56 a role in homing of HSCs in mice. Moreover, they demonstrated that GPR56 is physiologically involved in adhesion, migration, homing and mobilization of AML LSCs through the RhoA signaling pathway, especially in *EVI1*-overexpressed AML. Here, *EVI1* binds the *GPR56* promoter region and GPR56 knockdown renders leukemic cells more susceptible to chemotherapy.

Pabst et al. [235] were the first to postulate GPR56 as a LSC-specific marker in CD34+ and CD34- AML. Three major profiles could be distinguished based on CD34/GPR56 combinatorial flow cytometry, and GPR56 positive subpopulations showed the highest engrafting potential. Although CD34+/GPR56+ and CD34+/CD38- compartments identified the same populations in some samples, GPR6 generally subdivided the CD34+/CD38- compartment. Two third of the CD34+/CD38- cells harbored GPR56 positive expression (as evaluated for one patient). In other words, GPR56 positivity provides a novel LSC compartment with high repopulating potential *in vivo* irrespective of the CD34/CD38 status. Moreover, they observed an enrichment of GPR56+ populations in patients with bad outcome, and proposed GPR56 as a marker to differentiate between good and bad prognosis. This feature was later on confirmed by profiling data [161, 236], showing that GPR56 expression correlates with a LSC signature and is associated with a detrimental outcome in adult AML. Also within the CD34+/CD38- compartment, high GPR56 expression correlates with LSC signatures and decreased OS in patients intensively treated with chemotherapy [236]. Very recently, a six-gene LSC score was published specific

for pedAML which also included GPR56 [237]. This pLSC6 score holds great promise in redefining initial risk-stratification and identifying poor risk AML.

However, concomitant presence of GPR56 in normal HSCs represents a major disadvantage in targeted therapy. This crucial feature was evaluated by the group of Daria et al. [238], who illustrated that GPR56 expression is the highest in HSPCs, and expression in AML leukemic cells is substantial though lower. Conflicting data with this report were recently published, as Daga et al. [236] found slightly, but non-significant, lower GPR56 median fluorescence intensity (MFI) values in HSPC counterparts. The authors provide proof-of-concept that in a homeobox-driven AML model GPR56 robustly contributes to the development of AML. Of note, the authors only evaluated expression within three healthy controls, but confirmed these findings in a larger micro-array dataset.

GPR56 was shown to be vital for the emergence of HSCs in mice embryos, though would be dispensable in adult HSC [53, 239]. However, conflicting data were published by a second independent research group of the St. Jude Children's Research Hospital [240], showing that GPR56 knockdown resulted into loss of HSC repopulation in mice. Although Saito et al. illustrated that GPR56 knockout results into a decreased adhesion to stroma and ECM proteins *in vitro* [53], this finding could not be confirmed by Daria et al. using retroviral GPR56-transduced progenitor cells [238].

Second, the Junctional Adhesion Molecule-C (JAM-C) is an adhesion molecule expressed by normal hematopoietic stem and progenitor cells (HSPCs) that interacts with JAM-B expressed by stromal cells in the micro-environment. JAM-C was initially identified as a functional marker for engraftment of HSCs in NSG mice [241]. The JAM-C/JAM-B axis was shown to contribute to the homeostasis of HSC [242, 243]. The same research group revealed that JAM-C is also expressed by leukemic cells in about half of the CD34+ AML patients, irrespective of the CD34/CD38 expression profile, although higher percentages were observed within CD34+/CD38-/CD123+ cells [244]. The JAM-C fluorescence intensity was higher in leukemic cells than in HSPCs isolated from healthy donors, providing a clear-cut distinction between LSC and HSC. *De novo* AML patients (n=71) could be dichotomised in JAM-C^{high} (n=41) and JAM-C^{low} (n=30) patients, based on the frequency of JAM-C-positive cells determined by flow cytometry in respect to a 0.42% cut-off. High frequency of CD34+/CD38-/CD123+/JAM-C+ cells was identified as an independent prognostic marker for disease outcome in terms of OS and leukemia-free survival (LFS). Also a significantly reduced cumulative incidence of relapse (CIR) could be demonstrated, but was not confirmed by multivariate analysis. In addition, JAM-C+ leukemic cells harbored leukemia-initiating capacity upon serial engraftment. In our opinion, the prognostic value of JAM-C as LSC marker is of particular interest as (i) expression can readily be identified by flow cytometry, (ii) expression allows discrimination between HSC and LSC and (ii) since expression between PB and BM was not statistically different, PB could be used as sample matrix and avoid invasive BM aspirations.

Third, natural killer group 2D ligands (NKG2D-Ls) are rarely expressed by healthy cells in normal conditions, but, expression can be induced at the cell surface when the cell is stressed as the result of a viral infection or malignant transformation; they are therefore called "induced-self" ligands [245, 246]. For example, the expression of NKG2D-Ls is induced by DNA damage, a characteristic of tumor transformation, which leads to the activation of the DNA repair pathways. Expression of NKG2D-Ls is regulated by several mechanisms, which may be transcriptional, translational or post-translational.

NKG2D-L-positive cells are detected and eliminated by the immune system. Hence, suppressed expression of NKG2D-L could aid in immune evasion.

Very recently, Paczulla et al. [247] demonstrated that stemness and immune evasion are closely intertwined in AML. They showed that NKG2D-Ls are generally expressed on bulk AML cells whilst absent on the surface of LSCs. NKG2D-L negative LSCs were able to serially engraft, initiate leukaemia and survive chemotherapy in patient-derived xenotransplant models. So forth, absent NKG2D-L expression may function as a surrogate marker to identify leukemic cells with LSC features in both CD34+ and CD34- AML patients.

3.2.2.2. Light scatter properties

Terwijn et al. [201] suggested that evaluating aberrant physical characteristics of LSCs based on light scatter aberrancies, e.g. size (forward scatter (FSC)) and complexity (sideward scatter (SSC)), has a two-tiered added value in defining LSCs. On the one hand, LAIP-positive LSCs can be further delineated by a tightly clustered cell population with higher FSC and SSC. On the other hand, LAIP-negative but molecularly aberrant CD34+/CD38- cells could be additionally picked up based on aberrant scatter properties. Kersten et al. illustrated that high percentages (>90%) of CD45RA positive LSC reflected higher scatter properties (factor 1.68 higher compared to HSC, P<.001) [186]. However, scatter properties did not significantly differ between LSC with intermediate (10-90%) or low (<10%) CD45RA expression, implying that scatter properties may not replace the conventional use of CD-directed monoclonal antibodies (mAb) to discriminate LSC from HSC/MPP populations. Combining 13 LAIP markers with scatter abnormalities in adult AML increased the detection of aberrant $CD34^{+}/CD38^{-/dim}$ LSC from 70-75% [188] to 87% [187]. Combining aberrant scatter properties with LSC-associated and –specific markers markedly elevated the prognostic impact of the LSC load, most likely due to elimination of ‘‘HSC contamination’’ of the leukemic CD34+/CD38- compartment. Furthermore, both CD45RA expression and light scatter properties were shown to be indicators for maturation, and a more mature CD34+/CD38- phenotype reflected a more favourable cytogenetic/molecular profile, as also previously shown by Goardon et al. [12].

3.2.2.3. Functional markers

3.2.2.3.1. Side population

The side population (SP) is characterized by high efflux of the fluorescent dye Hoechst 33342 via the ABCG2 multidrug resistance-mediating transporter and can be determined by flow cytometry [248]. Based on their chemoresistant capacities, and the fact that nearly all stem cells in normal tissues and solid tumours reside in this compartment, the SP was proposed as a surrogate compartment for LSCs in AML.

The relationship between the SP and CD34+/CD38- compartment has only been addressed in a handful reports. In patients with both CD34+/CD38- and SP cells present, the CD34+/CD38- compartment appeared to be most represented [249]. Indeed, it was previously shown that most SP cells reside in the CD34+/CD38+ compartment due to the presence of progenitors herein [248]. In this last report, lack of a significant clinical impact of the SP turned the authors in favour of the immunophenotypic CD34+/CD38- LSC definition. Moshaver et al. performed an in-depth analysis in CD34+ and CD34- AML patients regarding the distribution of all four CD34/CD38 compartments, containing both leukemic and normal cells, in SP and non-SP compartments [250]. Interestingly, in both CD34+ and CD34- AML, SP cells exerted the highest stem cell activity, e.g. > 500 times higher, than non-SP cells based on a

clonogenic CFU assay. This result implies that the SP assay is of added value to define genuine LIC. Furthermore, integration of both immunophenotypic (CD34/CD38) and functional characteristics (SP) into a more restricted, single LSC definition tremendously (>factor 500) decreased the putative LIC frequency in both CD34+ and CD34- AML. Interestingly, only a minute subfraction of the SP compartment harboured expression of the aberrant LSC marker CLL-1 (median 0.0016% [249]).

3.2.2.3.2. Aldehyde dehydrogenase

In normal haematopoiesis, aldehyde dehydrogenase (ALDH) enzymes shield HSCs from the destructive properties of oxidative aldehydes deriving from lipid peroxidation, reduce oxidative stress by metabolising reactive oxygen species (ROS), and are involved in drug resistance and detoxification mechanisms [251]. The aldehyde dehydrogenase family consists of a group of highly homologous enzymes located in the cytosol that exhibit chaperone activities and are involved in various metabolic processes [252]. The first reports describing elevated ALDH activity in stem cells date from two decades ago, using the fluorescent substrate BODIPY aminoacetaldehyde [253, 254]. A decade later, it became clear that the levels of the ALDH1A1 isoform are most upregulated in immature cells, and that expression diminishes during differentiation [255]. Hitherto, 19 ALDH isoforms have been described in human hematopoietic and solid cancers [251, 256, 257]. The distinct ALDH1 isoforms, e.g. ALDH1A1, ALDH1A2, ALDH1A3, ALDH1B1, ALDH1L1 and ALDH1L2, were shown to exhibit different prognostic values in solid cancers. The ALDH3 family comprises four members, among which ALDH3A1 is an isoform repeatedly reported to be involved in chemoresistance [254, 257].

The interest in measuring ALDH levels in AML LSCs has significantly increased this last decade.

First, ALDH expression can be used as a surrogate marker to distinguish HSCs from LSCs, both present in the CD34+/CD38- compartment of AML patients, and can be applied in CD34- AML. Gerber et al. evaluated ALDH expression in the CD34+/CD38- fraction isolated from normal bone marrow and AML patients, i.e. HSC and LSCs, respectively [258]. HSCs had in more than 75% of the cases $ALDH^{high}$ expression, and readily generated serial engraftment when transplanted into NSG mice. In AML however, only 8.3% of the LSCs expressed $ALDH^{high}$ levels, and those who did, lacked previously defined cytogenetic abnormalities. Nearly all LSCs expressed $ALDH^{low}$ or $ALDH^{intermediate(int)}$ levels. Interestingly, CD34+/CD38-/$ALDH^{low}$ cells failed to serially engraft, even at high cell doses, whereas engraftment of CD34+/CD38-/$ALDH^{int}$ was already successful using only 1000 cells. The hypothesis that $ALDH^{int}$ cells are more clinically relevant in AML was further supported by a significant correlation between the number of CD34+/CD38-/$ALDH^{int}$ cells after treatment and the occurrence of relapse [258]. Later on, these data were confirmed by a second research group and applicability of the ALDH assay was also proven for CD34- AML [259]. The findings that CD34+/$ALDH^{int}$ cells are not consistently located in the CD34+/CD38- compartment and that ALDH expression levels can discriminate between HSCs and LSCs [254], suggest that ALDH expression analysis may increase the sensitivity and specificity of LSC detection. However, applicability of the assay showed a high inter-patient variability. Whilst ALDH levels are suppressed in most CD34+ AML patients of the two aforementioned studies (20/24 cases; 83% and 17/19; 89%), a more recent study showed only presence of a CD34+/CD38-/$ALDH^{int}$ population in half of the patients (47/80, 58%) [260].

Second, ALDH levels may provide relevant information regarding the chemoresistance of the disease, although conflicting data have been reported. In solid tumours, high ALDH expression is generally accepted to associate with therapy refractoriness and detrimental outcome [256]. Indeed, AML cases displaying $ALDH^{int}$ levels showed increased survival rates versus $ALDH^{high}$ patients, who showed significantly lower CR, EFS and OS [260]. However, these data are in conflict with the previous report

by Pearce et al., who showed that ALDH^{high} stem cells are non-leukemic in roughly one third of the cases [254].

Controversy also exists on the role of ALDH isoforms in the development of leukemia. Knockout or deletion of ALDH2, ALDH1A1 and ALDH3A1 isoforms promoted the development of acute leukemia both *in vivo* and *in vitro* [257, 261]. On the other hand, also increased ALDH expression levels were shown to enhance engraftment of primary AML leukemic cells [262]. A total of nine isoforms were recently demonstrated to react with the commercial Aldefluor activity test, including the stem cell expressed ALDH1A1 and ALDH3A1 iso-enzymes [263]. This finding emphasizes that the measured ALDH activity can be caused by different ALDH isoforms, depending on the cancer type and leukemic or normal setting. Pearce et al. showed that LSC from 2/3 of AML samples do not express detectable ALDH1, whereas normal HSC always do [254]. Hence, increased ALDH activity in these patients is most likely caused by increased ALDH-3 expression, and requires a therapeutic reagent specifically targeting ALDH3. Future research is needed to elucidate the isoform distribution within HSCs and LSCs in a pediatric leukemic setting, and to investigate which specific ALDH isoforms impact prognosis.

3.2.3. Molecular detection of targets and signatures

Xenografting to assess leukemia-initiating capacity is not feasible in a clinical setting, and flow cytometric analysis is hampered by the multiple CD34/CD38 compartments in which LSCs reside and the lack of an 'ultimate' marker that captures all LSCs. These limitations triggered researchers to explore molecular LSC identification strategies.

In adult AML, LSC-specific gene expression signatures were recently generated [161, 168, 264-269], and some were shown to hold significant prognostic value in independent cohorts. The expression for the set of genes included in the LSC-related (n=42) and HSC-related (n=121) signature by Eppert et al. was correlated to clinical outcome using three large bulk leukemic gene expression data sets [161]. Also the LSC17 gene signature developed by Ng et al. correlated with high-risk profiles (adverse cytogenetics and *FLT3*-ITD mutations) and detrimental outcome [264]. The authors also developed a LSC3-signature which could further stratify intermediate-risk patients (normal karyotype, *NPM1* mutated and *FLT3* WT). Hence, a high LSC-score (high expression measured for the genes included in the LSC-specific signature) can theoretically function as a surrogate marker for high LSC activity.

In pedAML, most GEPs were established based on bulk leukemic samples, and failed to identify critical genes and pathways characterizing the LSC compartment [143, 270-272]. Until recently, no pedAML LSC-signatures were available, although applicability of the LSC17 score was demonstrated in childhood AML [273, 274]. However, Lamba and collegues [237] just now published a pLSC6-score with a highly significant prognostic value in poor risk pedAML patients.

In addition to the use of LSC signatures, Yassin et al. identified only a couple of months ago the ERG gene enhancer+85 (ERG+85) region as a molecular biomarker representative for the stemness-state in LSCs and HSCs [275]. By integrating genomic and functional analysis, and tagging endogenous stemness-regulatory regions, they discovered a HSC-specific superenhancer residing in the ERG+85 region and showed a positive correlation between its activity and the degree of stemness *in vitro* and *in vivo*. Moreover, they developed a fluorescent lentiviral reporter that can accurately recapitulate the endogenous stemness activity. This novel reporter allows detection of the cellular stemness state in normal and leukemic hematopoietic cells, and furthermore, is able to predict disease outcome and drug sensitivity.

3.3. Clinical impact of the leukemic stem cell frequency at diagnosis

Defining the LSC load at diagnosis is considered to be a tremendous leap forward in the prognostication of adult AML [276-278]. The frequency of CD34+/CD38- cells was also extensively shown to impact prognosis (overview shown in **Table 4**). Initially, the LSC load was defined solely based on the CD34+/CD38- expression profile. One of the first reports was by Van Rhenen et al. [206], who retrospectively showed that high CD34+/CD38- frequency at diagnosis significantly correlates with a high MRD status, especially after the third course of chemotherapy, and directly correlates with poor survival. Two years later, the same group hypothesised that LSCs can bear abnormal immunophenotypes, in parallel with the abnormal phenotype of the leukemic blasts used to define MRD [188, 200]. Very recently, the results of a large adult AML prospective study showed that the LSC frequency is an independent prognostic factor for OS in multivariate analysis [279].

To the best of our knowledge, only two studies have been described in pedAML so far. Although they both illustrate a significant impact on EFS and/or relapse-free survival (RFS), some notable differences are present. First, the definition of a "LSC" seemingly differs. In the first study by Witte et al. [280], LSCs were defined as the total CD34+/CD38- population, whilst in the study by Hanekamp et al. [190], the aberrant expression of LAIP markers CD123, CD56, CD2 or CD7 was required to name a stem cell as "leukemic". Second, a different cut-off was established to distinguish LSC^{low} from LSC^{high} patients. Witte et al. quantified the LSC population in regard to the total WBC population, and showed that patients with LSC load ≤0.68% have a significant better EFS than those with LSCs ≥0.83%. Hanekamp et al. determined that a LSC population of 17.2% within the CD34+ compartment is the most specific cut-off to dichotomise pedAML patients based on their LSC load. Of note, the cut-offs proposed by Witte et al. was confirmed in the study of Hanekamp et al. to significantly impact OS (P=.002), but could not confirm impact on RFS at a significant level. Altogether, regardless of the cut-off value and markers used, the CD34+/CD38- burden at diagnosis also seems to be a prognostic factor in pedAML. Moreover, complete absence of this fraction also holds a strong prognostic value: CD34- status, characterized by the <1% presence of leukemic CD34+ cells, turned out to be an independent beneficial prognostic factor compared to patients with high or low CD34+/CD38− LSC frequencies [190]. Future retrospective and prospective studies are needed to evaluate whether measurement of the LSC burden at diagnosis could result into an improved risk stratification in pedAML.

Table 4. Overview of different cut-offs used to define high versus low LSC load at diagnosis in adult (n=8) and pediatric (n=3) AML and their impact on clinical outcome.

Age group	Patients (n)	Definition LSC	Denominator	Median ± SD (range)	Cut-off	Prognostic value	Ref
adult AML	594	CD34+/CD38-	White blood cells	0.0079% (0-19.88%)	Low: < 0.03% High: > 0.03% (based on [201])	LSC frequency was independent prognostic for OS in multivariate analysis (P<.0001), significant impact on achievement of CR (P<.001), though not in a multivariate model. LSC^{high} patients showed a higher portion of *FLT3*-ITD/*NPM1* WT and *EVI1* overexpression (detrimental outcome), and lower portion of *CEBPA* double mutations (beneficial outcome).	[279]
adult AML	92	CD34+/CD38-	(1) CD34+ cells (2) Leukemic compartment (CD45low/SSClow)	Fresh samples (n=56): 3.2% (0.1-82.5%). Thawed samples (n = 36): 3.7% (0.1-95.1%).	(1) 3.5% (2) 7.5%	High LSC frequency significantly correlates with low OS (n=92, P=.02), low DFS (n=60, P=.06), and low RFS (n=60, P=.005)	[206]
adult AML	63	CD34+/CD38-	Leukemic compartment (CD45low/SSClow)	1.3% (0.0 - 33.1%)	ND	Significantly higher LSC frequencies in therapy refractory patients compared to those achieving CR (0.7% vs 6.9%, P=.006)	[210]
adult AML	88	CD34+/CD38-	White blood cells	CR: 0.036%. No CR: 0.225%	1) 0.03% 2) Low: <0.005%, Intermediate: 0.005-0.1%, High: >0.1%	Independent prognostic factor on RFS and OS: significant higher CD34+CD38- compartment in therapy-refractory patients (n=18, 0.225%) compared to those who achieved CR (n=70, 0.036%) (sixfold, P=.041).	[201]
adult AML	30	Lin−/CD34+/CD38−	Not mentioned	11% (1-99%)*	ND	ND	[281]
adult AML	101	ALDH bright, irrespective of CD34 expression	White blood cells	Not mentioned	Low: <0.36%, High: ≥0.36%	Significant correlation between frequency of ALDHbright cells and the failure to achieve CR (P=.0025, n=84/101); significant impact on OS and RFS in univariate model (P=.0436 and P=.0125, respectively, n=100/101), strongest prognostic marker affecting OS in a multivariate model (P=.0095)	[282]
adult AML	22	CD34+/CD38-, CD34+/CD38dim and SP	White blood cells	CD38-: total: 0.011% (0 - 0.92%), responders: 0.001%, no response: 0.096%. CD38dim: 1.3% (0.02 - 17.4%), responders 0.260%, no response: 3.996%. SP: 0.04% (0 - 13.2%).	ND	SP: no significant impact on response. CD34+/CD38- and CD34+/CD38dim: significant anticorrelated with CR (P≤0.01) and longer CR duration (P≤0.01).	[248]
adult AML	111	CD34+/CD38- and dim/CD123+	Leukemic compartment (CD45low/SSClow)	2.8% (0.01 - 67%)	CR: 15%. Adverse event: 1%.	Significant correlation between LSChigh >15% and lack of CR (P<.05); Significant correlation between LSChigh >1% and DFS (P<.0001) and OS (P<.0001)	[211]
pedAML	17	CD34+/CD38-	Mononuclear white blood cells	Total: 1.02% ± 1.28%. Remission: 0.45% ± 0.61%. Relapse/death: 1.52% ± 1.52%	Low: ≤0.68%, High: ≥0.83%.	LSC frequency significantly higher in patients with relapse and/or death versus those with 5-year DFS rates (P<.05).	[280]
pedAML	12	CD34+/CD38-	Mononuclear white blood cells	0.23% (0.02-0.7)	ND	ND	[193]
pedAML	68	CD34+/CD38-, combined with CD123, CD56, CD2 and/or CD7	CD34+ blasts	0.99% (0 - 85.7%)	Low: ≤ 17.2%, High: >17.2%	LSChigh patients show significantly more risk of developing relapse compared to LSClow patients (P<.05).	[190]

PedAML, pediatric AML; DFS, disease-free survival; RFS, relapse-free survival; OS, overall survival; CR, complete remission; ALDH, aldehydedehydrogenase. .*Asterisk indicates biased calculation, since only CD34 positive AML patients were taken into account in whom the CD34+/CD38- populations of putative leukemic stem cells was >0.5% of the total AML leukemic compartment.

3.4. Metabolic, signaling and micro-environment alterations in leukemic versus normal hematopoietic stem cells

3.4.1. Metabolic dysregulations

LSC quiescence inherently follows from a decreased oxidative phosphorylation rate, with subsequent lower oxygen consumption and lower reactive oxygen species (ROS) levels. Furthermore, aerobic glycolysis seems to be impaired in these ROS-low LSCs. The low glycolytic capacity obligates LSCs to heavily rely on alternate energy sources to regulate the balance between quiescence and proliferation, i.e. anaerobic glycolysis, glutamine metabolism, fatty acid metabolism, bioenergetic signalling and the nutrient-sensitive AKT-mTOR pathway [283, 284]. LSCs residing in the adipose tissue of mouse models were shown to be capable of upregulating CD36, a fatty acid transporter, supporting lipolysis if needed [285]. Next to an intrinsic lower susceptibility due to their quiescent state and increased drug efflux, LSCs can also escape oxidative stress induced by chemotherapy through shifts within and upregulation of the glutathione metabolism [286].

3.4.2. Intrinsic signaling pathways

LSCs rely on a number of evolutionarily conserved signaling pathways that are redundant for normal adult HSCs, e.g. NF-κB, Notch and Wnt/β-catenin, Hedgehog (Hg) and IGF2/IGF1R/Nanog Signaling pathways [287]. NF-κB is a transcription factor that promotes cell growth and inhibits apoptosis, and is constitutively activated in LSCs whilst not in HSCs [288]. Although Notch signaling is essential during embryonic blood formation, it seems to be otiose in adult haematopoiesis. The canonical Wnt pathway, with β-catenin (Ctnnb1) as central player, plays a critical role in embryonic and adult generation of normal HSCs and exerts variable effects depending on the magnitude of activation. Both pathways promote leukemogenesis upon overexpression in the BM micro-environment [287]. Similarly, Nanog is a crucial factor for maintaining self-renewal and pluripotency of mouse embryonic stem cells, but dispensable in adult haematopoiesis whilst upregulated in LSCs [289]. Hh signaling preserves and increases the ST-HSC population in normal haematopoiesis. Normal haematopoiesis is not impaired by Hh signaling inhibition, but deletion of the Hh gene does suppress self-renewal capacities. Interestingly, suppressed Hh signaling reduced leukemic cell proliferation and colony formation *in vitro*, suggesting that the pathway is indispensable for LSCs [287, 290]. However, the leukemic effect of aberrant Hh signaling and the molecular mechanisms responsible for Hh activation in pedAML still need to be elucidated [291].

In addition, aberrant signaling of a plethora of pathways essential for both LSCs and normal HSCs has also been described, e.g. PI3/AKT/mTOR, JAK/STAT and MAP/ERK pathways [292]. PI3K/AKT/mTOR pathway plays a central role in key survival processes of HSCs and LSCs, e.g. NF-kB expression and self-renewal through the Wnt/b-catenin pathway. The frequent and constitutively increased activation of PI3K/AKT and mTOR pathways, respectively, observed in AML renders LSCs a competitive survival advantage over HSCs [293].

3.4.3. Extrinsic micro-environment

The altered interactions between LSCs and the micro-environment in the advantage of the leukemogenic process are bidirectional [21]. The micro-environment is able to modulate the proliferation, self-renewal and differentiation capacities of LSCs. The accumulative alterations in LSCs shift the balance between proliferative and apoptotive signals, yielding a hyper proliferative state and

promoting leukemia survival. Reversely, LSCs may provoke changes that mold the BM niche and turn the micro-environment into in “hostile hostel”, in which LSCs are able to outcompete HSCs. The interactions between LSCs and the regulatory niche appears to be critical in promoting therapeutic resistance through facilitation of immune evasion and progressive acquisition of (epi-)genetic changes within leukemic cells [294].

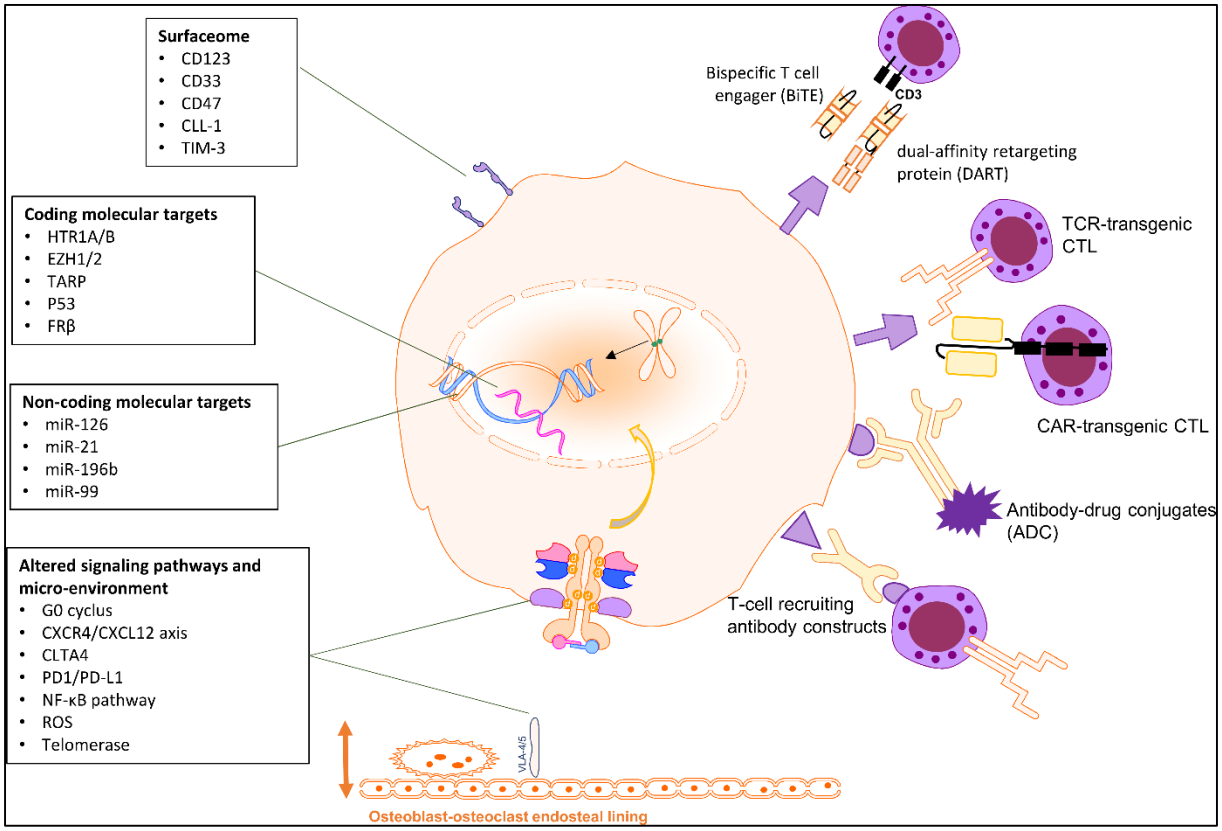
The hypoxic environment stimulates leukemic cells to release vascular endothelial growth factor (VEGF), osteopontin, and several CCL-and CXCL-chemokines with proangiogenic effects [295]. High stromal levels of CXCL12, also referred to as SDF-1 and produced by CXCL12-abundant reticular cells and MSCs, enhances homing of LSCs by interacting with its ligand CXCR4 [290, 296]. Comparably, stromal upregulation of the CD44 receptor facilitates stroma-dependent surveillance of the leukemic cells [296]. The ECM proteins hyaluronan, osteopontin and E-selectin also attract CD44 expressed on the LSC surface [21]. In addition, high soluble VCAM-1 levels increase homing though binding of the surface VLA-4 molecule (integrin α4β1), an integrin dimer composed of CD49d (α4) and CD29 (β1) [220]. In addition, an *in vivo* AML mouse model demonstrated that leukemic blasts express higher levels of the chemokine CCL3. CCL3 is responsible for a rigorous regulation of the osteoblast-osteoclast balance. CCL3 overexpression results in a leukemia-dependent bone loss, creating a less hospitable environment for normal HSCs and, at the same time, an opportunity for LSCs to hijack the niche [297]. Also osteoblasts play a crucial role, as it was shown that LSCs preferentially home to and expand within the osteoblast-rich endosteal niches **(Fig. 2)** [159]. Osteoblast ablation results into loss of LSC quiescence, reduces long-term engraftment, reduces self-renewal and accelerates leukemia development [298]. Osteoblast mutations in the Wnt signaling molecule CTNNB1 caused a constitutive activation of the Notch pathway, and induced leukemogenesis in HSPCs [299]. *In vitro* experiments proved that osteoblasts support leukemic development in a direct and indirect manner, e.g. by enhancing AML blast proliferation and through angiogenesis, respectively [300]. Angiogenesis is held responsible for the circulation of angioregulatory mediators and cytokines, e.g. SCF, that exhibit growth-enhancing effects on leukemic blasts and therefore also significantly contributes to the pathogenesis of AML [301].

3.5. LSC-targeted therapy

Based on multidrug chemotherapy, optimal stratification for allo-SCT and advances in supportive care, the current therapeutic regimens have reached a plateau around 70%. Excessive toxicity rates are a hurdle for further treatment intensifications [122]. Therefore, future perspectives need to focus on the development of novel strategies, such as targeted therapy. However, the high clinical heterogeneity within the pedAML landscape counteracts the establishment of patient subgroup-specific therapeutic strategies. Aside from the successful *FLT3* inhibitor-based therapies [302, 303] and the CD33-targeting agents [120, 304], targeted strategies have not yet found their way into the treatment of pedAML patients.

Relapse is the most frequent event in pedAML [120], and high frequency of chemotherapy-resistant LSCs is thought to be (partly) responsible. Therefore, targeting LSCs might represent a very interesting strategy to improve long-term outcome. However, the answer to the million dollar question ‘will it provide the path to cure AML?’ remains a check to be cashed. Translating the success of targeted immunotherapy from ALL, exemplified by Blinatumomab, to AML is challenged by the cumbersome identification of suitable target antigens. Most likely, combinatorial treatment approaches will be required to improve the effectiveness and safety of immunotherapy in AML. Below, we summarize recent advances made regarding LSC-targeted immunotherapeutic strategies **(Fig. 8)**.

Fig. 8. Cartoon of LSC-specific targets (left) and immunotherapeutic strategies (right).



3.5.1. Immunotherapeutic targeted strategies: what's in a name?

3.5.1.1. Antibodies

Monoclonal antibodies (mAbs) conjugated to toxins (antibody-drug conjugates, ADCs) or other molecules can exert site- or cell-specific toxic effects. Antibody-dependent cellular cytotoxicity of natural killer (NK) cells largely contributes to the success of mAb treatment in cancer. In the haematological malignancies field, therapeutic mAbs are especially relevant owing to accessibility to tumor cells, facilitating *in vitro* studies of targets [305]. Of particular interest are bispecific antibodies that combine antigen-binding specificities for target cells (e.g. LSCs) and effector cells (e.g. T- and NK-cells) and redirect effector cells to engage chosen targets on tumor cells. When the effector cells are NK-cells, both antibody-dependent and antibody-independent mechanisms are exploited. When the effectors are T-cells, direct tumor cytotoxicity can be engaged followed by a potential vaccination effect [306]. To overcome production challenges with respect to quantity, quality, and stability of bispecific antibodies, bispecific T-cell engagers (BiTEs) were developed [307]. The BiTE format recombinantly links the four variable domains of heavy and light chains required for the two antigen-binding specificities, e.g. LSCs and effector cells, and combine tumor targeting with selective T-cell activation at low picomolar concentration. Their success stimulated further research for similar downgraded antibody formats, and led to the development of dual-affinity of re-targeting molecules (DART) [308]. DARTs have a diabody format that separates cognate variable domains of heavy (VH) and light chains (VL) of the two antigen binding specificities on two separate polypeptide chains. The scFv dimers, consisting out of VH and VL regions connected by a small peptide linker, are additionally stabilized by a C-terminal disulfide bridge.

3.5.1.2. Adoptive cellular immunotherapy

Genetic engineering of cytotoxic effector cells, e.g. T-cells and NK-cells, is a fast‑developing field in cancer therapeutics. Stable transduction of T-cell receptors (TCRs) and chimeric antigen-receptors (CARs) was shown to be effective in re-directing effector cells towards tumor-associated antigens [309]. A variety of effector cell types and sources have been described. In general, autologous or allogeneic cytotoxic T-cells (CTLs) are *ex-vivo* selected, activated and genetically modified, followed by expansion and then transferred back to patients, who hereby adopt a specific immunity. In order to maximise the tumor targeting versus toxicity ratio, alternative T-cell sources have been resourced that reduce the need for autologous cell manufacturing and crosses histocompatibility barriers, such as induced pluripotent stem cells (iPSCs) and embryonic stem cells [310]. In spite of these encouraging results, we may not overlook the worrisome finding that iPSCs acquire mutations during infinite expansion *in vitro*, e.g. in TP53 [311].

TCRs are physiologically engaged to bind histocompatibility leukocyte antigen (HLA)-peptide complexes. This requisite remains one of the major concerns, together with the on-target/off-tumor recognition of antigens in healthy tissues [312]. A resolution came with the development of CARs. CARs are synthetic receptors that have an antibody-based external structure and cytosolic domains encoding for the signal transduction axis of a TCR. Unlike physiological TCRs, they bind to native cell surface molecules and do not require antigen processing nor HLA expression for recognition. Subsequently, they are less prone to fail in case of tumor immune escape mechanisms, such as downregulated HLA expression or suppressed proteasomal antigen processing. The first generation CARs consisted out of antigen-recognition domains that were fused to the CD3ζ activation chain of the TCR complex, and showed only limited antitumor efficacy *in vivo* [313]. Subsequent second and third generation CAR iterations included one and two, respectively, additional secondary costimulatory signals, most commonly CD28 and 4-1BB, which improved *in vivo* efficacy tremendously. Also, their design allows a double-pronged targeting approach. Implementation of two discrete scFv domains directed against two different antigens increases tumor selectivity. Recently, a novel CAR technology termed UniCAR was developed that reduces the risk for on-target/off-tumor side effects by a rapid and reversible control of the reactivity of the CAR T-cells. UniCARs are able to target more than one antigen, hereby reducing the risk for failure in case of antigen-loss under treatment. Their efficacy in targeting CD33- and/or CD123-positive AML blasts was recently proven *in vitro* and *in vivo*. [314].

Nevertheless, there still remain some challenges. The extreme potency of CAR T-cells may result into life-threatening cytokine-release syndromes, and require temporal remote control mechanisms that restrains cytolytic activity [315]. Compared to TCRs, CARs require higher absolute antigen expression to activate effector functions. Like TCR-transgenic T-cells, the clinical use of CARs is counteracted by the induction of “on-target/off-tumor” toxicities. So far, the main successes were obtained towards cell surface expressed antigens, and the experience with CARs targeting intracellular antigens that are presented on the cell surface is limited. However, recent successful targeting of the *WT1* oncoprotein by a TCR-mimicking CAR will possibly pave the way for other attractive intracellular targets [316].

3.5.2. The LSC target-ome

In ideal circumstances, the target should be universally and highly expressed in LSCs whilst absent in healthy vital tissues and HSCs. In case of immunotherapy, targets should also be absent in T-/NK-cells to avoid fratricide elimination. The challenge lies in finding a target with optimal high on-tumor/low off-target capacities. An increased success rates by integrating transcriptomic and proteomic data was convincingly demonstrated [317]. We can roughly subdivide the LSC targetable landscape, aka “LSC

target-ome", into four categories **(Fig. 8)**: (1) CD antigen membrane markers, (2) Coding molecular factors, (3) Non-coding molecular targets e.g. lncRNA and miRNA and (4) altered signaling and micro-environment.

3.5.2.1. Membrane antigen markers (LSC surface-ome)

The aberrant expression of CD markers on the membrane of LSCs provides important hints for the development of LSC-targeted therapy. Several attractive cell-surface targets amenable for immunotherapy have been identified. Of note, the occurrence of phenotypical shifts questions therapeutic strategies that rely on targeting single antigens, especially in relapse setting were most novel agents are employed, and ask for combinatorial approaches [318].

CSL360, a CD123-directed recombinant chimeric immunoglobulin, was evaluated in a phase I clinical trials for treatment of relapsed, refractory and high-risk AML (#NCT00401739) [319]. Although well-tolerated, responsiveness appeared unfortunately to be low (5%, 2/40 cases), suggesting that the anti-CD123 leukemic effect is insufficient to eradicate LSCs. Data on the clinical effectiveness of a recently developed CD123-specific CAR, proven to have an anti-leukemic activity *in vitro* and *in vivo*, is still awaiting [320]. In addition, anti-CD3/anti-CD123 DART (MGD006/S80880) induced a dose-dependent killing of AML cell lines and primary AML blasts *in vitro* and *in vivo*, but needs to be evaluated in a clinical trial.

CD33 has long been of interest as target for anti-AML therapy because of its highly prevalent expression in leukemic blasts. However, concomitant expression in normal HSPCs can cause considerable haematological toxicity and calls for caution. So far, four CD33-targeting molecules have entered clinical trials. Gemtuzumab ozogamicin (GO, Mylotarg®), a CD33-targeted antibody-calicheamicin conjugate on the market since 2000, has regained interest since its FDA withdrawal in 2010, and was recently approved for treatment of CD33-expressing adult AML in mono- and combination therapy [321]. Retrospective studies has convincingly shown the benefit of GO (combinational) bridging therapy in order to create a window for successful transplantation results in pedAML [322]. A handful of clinical studies, investigating GO alone, in combination or as front-line therapy, are ongoing to prove improved outcomes prospectively [322, 323]. Convincing data were provided by Gamis et al., who showed in a large cohort of over 1000 AML patients between 0-29 yrs. that combining chemotherapy with GO significantly reduces the number of relapses, and despite higher post remission toxic mortality, leads to a significant higher EFS [324].

Recently, another anti-CD33 Ab linked to the toxic DNA binding agent pyrrlobenzodiazpine was developed, referred to as vadastuximab talirine (SGN-33A), and found to be safe in AML patients when combined with standard chemotherapy [325]. In addition, the CD33/CD3 BiTE construct AMG 330 was proven to be very effective in recruiting and activating T-cells towards primary AML cells *ex vivo* [326, 327]. Moreover, AMG 330-mediated anti-leukemic cytotoxicity could be further enhanced by blockade of the PD-1/PD-L1 axis. Lastly, Lintuzumab (HuM195), a humanized anti-CD33 antibody, was shown to achieve CR in about one quarter of adult AML patients [328]. Conjugation to the alpha-particle-emitting radionuclide bismuth-213 (213Bi-lintuzumab) increased potency and was tolerable in combination with cytarabine [329]. Recently, a two-tiered compound CAR (cCAR) directed towards CD33 and CD123 was designed, including a safety-switch for rapid termination, able to eliminate AML bulk disease and LSCs [330]

Anti-CD47 monoclonal antibodies that block CD47/SIRPα signaling, and soluble SIRPα-Fc fusion proteins that neutralize CD47, disrupt the "do not eat" signal sent out by between LSCs and phagocytic cells, respectively, and activate the innate immune response to eliminate LSCs [331, 332].

Due to its high specificity, CLL-1 is considered as a very promising anti-LSC target. A series of monoclonal antibodies against the extracellular domain of CLL-1 showed successful *in vitro* and *in vivo* results [333]. In addition, a CLL-1/CD3 BiTE construct was recently proven to effectively target AML in monkeys [334], and its safety, tolerability and preliminary efficacy in humans is currently being evaluated in a phase I clinical trials (MCLA-117, NCT03038230). Alternatively, nanomicelles decorated with CLL1-targeting peptides are able to specifically bind CLL-1 expressing LSCs and directly deliver daunorubicin [335]. Last year, anti-CLL-1 targeting CAR T-cells were developed and shown to specifically lyse CLL-1+ cell lines and primary AML patient samples *in vitro* [336].

T-cell Ig mucin-3 (TIM3), also known as CD366 or Kidney Injury Molecule-3 (KIM3), is normally expressed on various cells of the immune system and enables phagocytosis of apoptotic cells. TIM3 is mostly expressed in CBF-leukemia and *CEBPA*-mutated leukemia, including adult AML leukemic cells, while absent in normal HSC [199]. Therefore, anti-TIM3 monoclonal antibodies such as ATIK2a do not impair normal haematopoiesis while blocking engraftment of human AML in mice. Administration of TIM3 targeting antibodies deprived LSC from their leukemia-initiating capacity in secondary mice. Hence, ATIK2a holds promise to eradicate LSCs [337].

Anti-CD38 therapy such as Daratumumab has been suggested to be of value in adult AML patients, especially in combination with all-trans retinoic acid (ATRA) [338]. However, one could argue whether these response rates would be durable, as most of the LSC fraction, especially in CD34+ AML, will remain untargeted.

3.5.2.2. Coding molecular targets

Several targets have been proposed to be eligible for immunotherapy based on their highly differential expression in LSC compared to HSC, e.g. serotonin receptors type 1 (*HTR1*) A/B [339], T-cell receptor γ chain alternate reading frame protein (*TARP*) [340], enhancer of zeste (*EZH*) 1 and *EZH2* [341], p53 [342] and folate receptor beta (*FRβ*) [343].

LSC-associated signatures provide an excellent source to select and screen putative LSC specific targets [268, 344]. In order to develop novel LSC targeted therapies, Saito et al. [268] selected LSC-specific molecules based on global mRNA expression patterns between LSCs and HSCs, followed by functional characterization in a mouse xenotransplantation model. LSC signatures also allow to evaluate the impact of existing compounds on the chemoresistant nature of LSCs and to identify alternate LSC targeting indications. *In silico* analysis of prognostic LSC gene expression signatures and drug–gene interaction datasets, followed by *in vitro* evaluation of the selected LSC candidate drugs for anti-leukemic activity, suggested that antihistamines, cardiac glycosides and glucocorticoids hold great promise as LSC-targeting candidates [344]. Lovastatin, an anti-cholesterol drug, was shown to inhibit LSCs whilst sparing HSPCs *in vitro* and *in vivo* [345]. Fenretinide, a vitamin A derivate conventionally used for treatment of acne and psoriasis, induced LSC apoptosis in an experimental setting by affecting several key signaling pathways, e.g. NF-κB and ERK [346]. Moreover, the ERK pathways was identified together with MSK/Sp1/c-Myc axis as a critical regulator of survivin expression in LSCs, offering a new potential target for LSCs therapy [347]. Confirmatively, others showed that suppression of survivin leads to decreased cell proliferation and self-renewal in LSCs [348].

3.5.2.3. Non-coding molecular targets

Modulation of miRNAs and subsequently their downstream targets can successfully eliminate LSCs by inducing apoptosis or by increasing chemosensitivity [349]. Since HSCs and LSCs behave oppositely upon miR-126 knockdown, attenuation of microRNA-126 expression leads to tumor eradication, and at the same time, promotes HSC recovery [269]. Another strategy is to inhibit miR-21 and miR-196b simultaneously using antagomiRs. MiRNA-21/-196b suppression inhibited *in vitro* leukemic colony forming activity and depleted *in vivo* leukemia-initiating cell activity of 11q23-rearranged and other HOXA9-related AML, which led to leukemia-free survival and delayed disease onset in xenograft models [350, 351]. *CDKN1B* was identified as a direct miR-196b target whose repression enhanced a stem cell-like signature and was associated with increased LSCs *in vivo*. MiR-99 was shown to be significantly upregulated in CD34+ cells with versus without leukemia-initiating capacity in xenografts. Interestingly, miR-99a levels were significantly higher in LSCs at relapse compared to diagnosis. Further research on the role of miR-99 suggested its involvement in chemotherapy resistance and leukemic proliferation [352].

The distinct lncRNA profiles observed in LSCs might also be relevant in order to identify new therapeutic targets. Within this perspective, *DANCR* was proposed as valuable target in adult AML, shown to be highly overexpressed in functionally validated LSCs. *DANCR* knockdown resulted in decreased LSC renewal and quiescence *in vitro*, and prolonged the survival of mice after serial transplantation *in vivo* [353]

3.5.2.4. Altered signaling pathways and micro-environment

The altered conditions in which LSC reside and survive provide multiple therapeutic targets for LSC eradication.

One of the main tricks of LSCs to circumvent damage by chemotherapeutic agents is the low responsiveness as a consequence from their quiescent state. Subsequently, pushing LSCs into an active cell cycle, e.g. by the addition of growth factors such as granulocyte (macrophage) colony-stimulating factor (G(M)-CSF) [354], would theoretically make LSCs more chemosusceptible. A recent meta-analysis based on 11 high-quality studies including a total of 5076 patients (age 10-60 yr.), of which 2446 patients received G-CSF in conjunction to chemotherapy and 2610 did not, showed that adding G-CSF only significantly improved DFS and decreased CIR in *de novo* AML patients. No effect was observed in terms of OS and CR, and none of these clinical endpoints were significantly different within relapsed/refractory AML patients [355].

Another approach is to target signalling pathways that are expressed in LSC whilst otiose in HSCs, or, aberrantly activated in LSCs. Parthenolide, MG-132 and BAY 11-7082 are three molecules that function as NF-κB inhibitors, a pathway which is constitutionally activated in LSCs and induce apoptosis by activation of p53 and increased ROS production [288, 356]. Alternatively, targeting the altered micro-environment by disturbing the CXCR4/CXCL12 signaling axis e.g. through administration of anti-CXCR4 neutralising antibodies, anti-CXCR4 antagonists (AMD3100), or anti-CXCR4 antibodies (Ulocuplumab, BMS-936564), could result into LSC eradication whilst sparing normal HSCs. Indeed, blocking of CXCR4 dramatically decreased engraftment of leukemic cells but did not significantly affect HSPC engraftment in NS mice [357]. In addition, mAbs that block immune checkpoint receptors CTLA4 (Ipilimumab) or PD1/PD-L1 (Nivolumab, Pembrolizumab and Avelumab) are well tolerated, beside the risk on immune-related adverse events, and have entered phase I clinical trials.

Recently, intracellular ROS and telomerase depletion were proposed as novel methods for LSC eradication. Since LSC have an intrinsic ROS^{low} status, agents that alter their redox status by inducing/modulating ROS production could result into a selective LSC ablation [358]. Also, telomerase depletion tackles the unlimited self-renewal capacities of LSCs and can cause irreversible damage [359]. Noteworthy, up to date, these therapeutic paths have not yet been explored in a pediatric setting.

References

1. A. M: Der Lymphozyt als gemeinsame Stammzelle der verschiedenen Blutelemente in der embryonalen Entwicklung und im postfetalen Leben der Säugetiere. Folia Haematol 1909, 8:125-134.
2. Andral G: Essai d'Hématologie Pathologique. Fortin, Masson & Cie, Paris, 1843.
3. Addison W: Experimental and Practical Researches on Inflammation and on the Origin and Nature of Tubercles of the Lung. J Churchill, London, 1843.
4. Lorenz E, Uphoff D, Reid TR, Shelton E: Modification of irradiation injury in mice and guinea pigs by bone marrow injections. Journal of the National Cancer Institute 1951, 12(1):197-201.
5. Till JE, McCulloch EA, Siminovitch L: A Stochastic Model of Stem Cell Proliferation, Based on the Growth of Spleen Colony-Forming Cells. Proceedings of the National Academy of Sciences of the United States of America 1964, 51:29-36.
6. Tenen DG, Hromas R, Licht JD, Zhang DE: Transcription factors, normal myeloid development, and leukemia. Blood 1997, 90(2):489-519.
7. Morrison SJ, Weissman IL: The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. Immunity 1994, 1(8):661-673.
8. Gunji Y, Nakamura M, Osawa H, Nagayoshi K, Nakauchi H, Miura Y, Yanagisawa M, Suda T: Human primitive hematopoietic progenitor cells are more enriched in KITlow cells than in KIThigh cells. Blood 1993, 82(11):3283-3289.
9. Yin AH, Miraglia S, Zanjani ED, Almeida-Porada G, Ogawa M, Leary AG, Olweus J, Kearney J, Buck DW: AC133, a novel marker for human hematopoietic stem and progenitor cells. Blood 1997, 90(12):5002-5012.
10. Akashi K, Traver D, Miyamoto T, Weissman IL: A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature 2000, 404(6774):193-197.
11. Reya T, Morrison SJ, Clarke MF, Weissman IL: Stem cells, cancer, and cancer stem cells. Nature 2001, 414(6859):105-111.
12. Goardon N, Marchi E, Atzberger A, Quek L, Schuh A, Soneji S, Woll P, Mead A, Alford KA, Rout R et al: Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer cell 2011, 19(1):138-152.
13. Gorgens A, Radtke S, Mollmann M, Cross M, Durig J, Horn PA, Giebel B: Revision of the human hematopoietic tree: granulocyte subtypes derive from distinct hematopoietic lineages. Cell reports 2013, 3(5):1539-1552.
14. Zhang J, Niu C, Ye L, Huang H, He X, Tong WG, Ross J, Haug J, Johnson T, Feng JQ et al: Identification of the haematopoietic stem cell niche and control of the niche size. Nature 2003, 425(6960):836-841.
15. Asada N, Takeishi S, Frenette PS: Complexity of bone marrow hematopoietic stem cell niche. International journal of hematology 2017, 106(1):45-54.
16. Omatsu Y, Seike M, Sugiyama T, Kume T, Nagasawa T: Foxc1 is a critical regulator of haematopoietic stem/progenitor cell niche formation. Nature 2014, 508(7497):536-540.
17. Omatsu Y, Sugiyama T, Kohara H, Kondoh G, Fujii N, Kohno K, Nagasawa T: The essential functions of adipo-osteogenic progenitors as the hematopoietic stem and progenitor cell niche. Immunity 2010, 33(3):387-399.
18. Calvi LM: Osteolineage cells and regulation of the hematopoietic stem cell. Best practice & research Clinical haematology 2013, 26(3):249-252.
19. Peled A, Kollet O, Ponomaryov T, Petit I, Franitza S, Grabovsky V, Slav MM, Nagler A, Lider O, Alon R et al: The chemokine SDF-1 activates the integrins LFA-1, VLA-4, and VLA-5 on immature human CD34(+) cells: role in transendothelial/stromal migration and engraftment of NOD/SCID mice. Blood 2000, 95(11):3289-3296.

20. Lord BI, Testa NG, Hendry JH: The relative spatial distributions of CFUs and CFUc in the normal mouse femur. Blood 1975, 46(1):65-72.
21. Krause DS, Scadden DT: A hostel for the hostile: the bone marrow niche in hematologic neoplasms. Haematologica 2015, 100(11):1376-1387.
22. Bhattacharya D, Rossi DJ, Bryder D, Weissman IL: Purified hematopoietic stem cell engraftment of rare niches corrects severe lymphoid deficiencies without host conditioning. The Journal of experimental medicine 2006, 203(1):73-85.
23. Czechowicz A, Kraft D, Weissman IL, Bhattacharya D: Efficient transplantation via antibody-based clearance of hematopoietic stem cell niches. Science 2007, 318(5854):1296-1299.
24. Shimoto M, Sugiyama T, Nagasawa T: Numerous niches for hematopoietic stem cells remain empty during homeostasis. Blood 2017, Apr 13(129(15)):2124-2131.
25. Creutzig U, van den Heuvel-Eibrink MM, Gibson B, Dworzak MN, Adachi S, de Bont E, Harbott J, Hasle H, Johnston D, Kinoshita A et al: Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. Blood 2012, 120(16):3187-3205.
26. Howlader N NA, Krapcho M, et al. : SEER Cancer Statistics Review, 1975-2009 (Vintage 2009 Populations) National Cancer Institute; Bethesda, MD. 2012.
27. Reedijk AMJ, Klein K, Coebergh JWW, Kremer LC, Dinmohamed AG, de Haas V, Versluijs AB, Ossenkoppele GJ, Beverloo HB, Pieters R et al: Improved survival for children and young adolescents with acute myeloid leukemia: a Dutch study on incidence, survival and mortality. Leukemia 2019, 33(6):1349-1359.
28. Niemeyer CM, Baumann I: Myelodysplastic syndrome in children and adolescents. Seminars in hematology 2008, 45(1):60-70.
29. Sande JE, Arceci RJ, Lampkin BC: Congenital and neonatal leukemia. Seminars in perinatology 1999, 23(4):274-285.
30. Zhang J, Walsh MF, Wu G, Edmonson MN, Gruber TA, Easton J, Hedges D, Ma X, Zhou X, Yergeau DA et al: Germline Mutations in Predisposition Genes in Pediatric Cancer. The New England journal of medicine 2015, 373(24):2336-2346.
31. Jeong D, Lee DS, Kim N, Choi S, Kim K, Kim SM, Im K, Park HS, Yun J, Lim KM et al: Prevalence of germline predisposition gene mutations in pediatric acute myeloid leukemia: Genetic background of pediatric AML. Leukemia research 2019, 85:106210.
32. Seif AE: Pediatric leukemia predisposition syndromes: clues to understanding leukemogenesis. Cancer genetics 2011, 204(5):227-244.
33. Hall GW: Childhood myeloid leukaemias. Best practice & research Clinical haematology 2001, 14(3):573-591.
34. Bresters D, Reus AC, Veerman AJ, van Wering ER, van der Does-van den Berg A, Kaspers GJ: Congenital leukaemia: the Dutch experience and review of the literature. British journal of haematology 2002, 117(3):513-524.
35. Ferguson EC, Talley P, Vora A: Translocation (6;17)(q23;q11.2): a novel cytogenetic abnormality in congenital acute myeloid leukemia? Cancer genetics and cytogenetics 2005, 163(1):71-73.
36. Owen C, Barnett M, Fitzgibbon J: Familial myelodysplasia and acute myeloid leukaemia--a review. British journal of haematology 2008, 140(2):123-132.
37. Manola KN: Cytogenetics of pediatric acute myeloid leukemia. European journal of haematology 2009, 83(5):391-405.
38. Lafiura KM, Bielawski DM, Posecion NC, Jr., Ostrea EM, Jr., Matherly LH, Taub JW, Ge Y: Association between prenatal pesticide exposures and the generation of leukemia-associated T(8;21). Pediatric blood & cancer 2007, 49(5):624-628.
39. Knudson AG, Jr.: Mutation and cancer: statistical study of retinoblastoma. Proceedings of the National Academy of Sciences of the United States of America 1971, 68(4):820-823.
40. Gilliland DG: Molecular genetics of human leukemias: new insights into therapy. Seminars in hematology 2002, 39(4 Suppl 3):6-11.

41. Gilliland DG, Griffin JD: The roles of FLT3 in hematopoiesis and leukemia. Blood 2002, 100(5):1532-1542.
42. Itzykson R, Duployez N, Fasan A, Decool G, Marceau-Renaut A, Meggendorfer M, Jourdan E, Petit A, Lapillonne H, Micol JB et al: Clonal interference of signaling mutations worsens prognosis in core-binding factor acute myeloid leukemia. Blood 2018, 132(2):187-196.
43. Balgobind BV, Hollink IH, Arentsen-Peters ST, Zimmermann M, Harbott J, Beverloo HB, von Bergh AR, Cloos J, Kaspers GJ, de Haas V et al: Integrative analysis of type-I and type-II aberrations underscores the genetic heterogeneity of pediatric acute myeloid leukemia. Haematologica 2011, 96(10):1478-1487.
44. Vujkovic M, Attiyeh EF, Ries RE, Goodman EK, Ding Y, Kavcic M, Alonzo TA, Wang YC, Gerbing RB, Sung L et al: Genomic architecture and treatment outcome in pediatric acute myeloid leukemia: a Children's Oncology Group report. Blood 2017, 129(23):3051-3058.
45. Fleishman EV, Sokova OI, Kirichenko OP, Konstantinova LN, Metel'kova NF, Popa AV, Shneider MM: [Complex karyotype abnormalities in pediatric acute myeloid leukemia]. Vestnik Rossiiskoi akademii meditsinskikh nauk 2008(5):3-7.
46. Struski S, Lagarde S, Bories P, Puiseux C, Prade N, Cuccuini W, Pages MP, Bidet A, Gervais C, Lafage-Pochitaloff M et al: NUP98 is rearranged in 3.8% of pediatric AML forming a clinical and molecular homogenous group with a poor prognosis. Leukemia 2017, 31(3):565-572.
47. Meyer C, Burmeister T, Groger D, Tsaur G, Fechina L, Renneville A, Sutton R, Venn NC, Emerenciano M, Pombo-de-Oliveira MS et al: The MLL recombinome of acute leukemias in 2017. Leukemia 2018, 32(2):273-284.
48. Pigazzi M, Masetti R, Bresolin S, Beghin A, Di Meglio A, Gelain S, Trentin L, Baron E, Giordan M, Zangrando A et al: MLL partner genes drive distinct gene expression profiles and genomic alterations in pediatric acute myeloid leukemia: an AIEOP study. Leukemia 2011, 25(3):560-563.
49. Balgobind BV, Raimondi SC, Harbott J, Zimmermann M, Alonzo TA, Auvrignon A, Beverloo HB, Chang M, Creutzig U, Dworzak MN et al: Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. Blood 2009, 114(12):2489-2496.
50. Rubnitz JE, Raimondi SC, Tong X, Srivastava DK, Razzouk BI, Shurtleff SA, Downing JR, Pui CH, Ribeiro RC, Behm FG: Favorable impact of the t(9;11) in childhood acute myeloid leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2002, 20(9):2302-2309.
51. Shimada A, Taki T, Tabuchi K, Taketani T, Hanada R, Tawa A, Tsuchida M, Horibe K, Tsukimoto I, Hayashi Y: Tandem duplications of MLL and FLT3 are correlated with poor prognoses in pediatric acute myeloid leukemia: a study of the Japanese childhood AML Cooperative Study Group. Pediatric blood & cancer 2008, 50(2):264-269.
52. Jo A, Mitani S, Shiba N, Hayashi Y, Hara Y, Takahashi H, Tsukimoto I, Tawa A, Horibe K, Tomizawa D et al: High expression of EVI1 and MEL1 is a compelling poor prognostic marker of pediatric AML. Leukemia 2015, 29(5):1076-1083.
53. Saito Y, Kaneda K, Suekane A, Ichihara E, Nakahata S, Yamakawa N, Nagai K, Mizuno N, Kogawa K, Miura I et al: Maintenance of the hematopoietic stem cell pool in bone marrow niches by EVI1-regulated GPR56. Leukemia 2013, 27(8):1637-1649.
54. Mercher T, Schwaller J: Pediatric Acute Myeloid Leukemia (AML): From Genes to Models Toward Targeted Therapeutic Intervention. Frontiers in pediatrics 2019, 7:401.
55. Thanasopoulou A, Tzankov A, Schwaller J: Potent co-operation between the NUP98-NSD1 fusion and the FLT3-ITD mutation in acute myeloid leukemia induction. Haematologica 2014, 99(9):1465-1471.
56. von Bergh AR, van Drunen E, van Wering ER, van Zutven LJ, Hainmann I, Lonnerholm G, Meijerink JP, Pieters R, Beverloo HB: High incidence of t(7;12)(q36;p13) in infant AML but not in infant ALL, with a dismal outcome and ectopic expression of HLXB9. Genes, chromosomes & cancer 2006, 45(8):731-739.

57. Bolouri H, Farrar JE, Triche T, Jr., Ries RE, Lim EL, Alonzo TA, Ma Y, Moore R, Mungall AJ, Marra MA et al: The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nature medicine 2018, Jan(24(1)):103-112.
58. Forestier E, Heim S, Blennow E, Borgstrom G, Holmgren G, Heinonen K, Johannsson J, Kerndrup G, Andersen MK, Lundin C et al: Cytogenetic abnormalities in childhood acute myeloid leukaemia: a Nordic series comprising all children enrolled in the NOPHO-93-AML trial between 1993 and 2001. British journal of haematology 2003, 121(4):566-577.
59. Grimwade D, Ivey A, Huntly BJ: Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance. Blood 2016, Jan 7(127(1)):29-41.
60. Meshinchi S, Arceci RJ: Prognostic factors and risk-based therapy in pediatric acute myeloid leukemia. Oncologist 2007, 12(3):341-355.
61. Haltrich I, Kost-Alimova M, Kovacs G, Klein G, Fekete G, Imreh S: Multipoint interphase FISH analysis of chromosome 3 abnormalities in 28 childhood AML patients. European journal of haematology 2006, 76(2):124-133.
62. Huret J: t(1;22)(p13;q13). Atlas Genet Cytogenet Oncol Haematol 2001, 5(3):196-197.
63. Sandahl JD, Kjeldsen E, Abrahamsson J, Ha SY, Heldrup J, Jahnukainen K, Jonsson OG, Lausen B, Palle J, Zeller B et al: Ploidy and clinical characteristics of childhood acute myeloid leukemia: A NOPHO-AML study. Genes, chromosomes & cancer 2014, 53(8):667-675.
64. Tarlock K, Meshinchi S: Pediatric acute myeloid leukemia: biology and therapeutic implications of genomic variants. Pediatric clinics of North America 2015, 62(1):75-93.
65. Coenen EA, Zwaan CM, Reinhardt D, Harrison CJ, Haas OA, de Haas V, Mihal V, De Moerloose B, Jeison M, Rubnitz JE et al: Pediatric acute myeloid leukemia with t(8;16)(p11;p13), a distinct clinical and biological entity: a collaborative study by the International-Berlin-Frankfurt-Munster AML-study group. Blood 2013, 122(15):2704-2713.
66. Smol T C-RM: t(8;16)(p11;p13) KAT6A/CREBBP. Atlas Genet Cytogenet Oncol Haematol in press.
67. Huret J: t(9;22)(q34;q11) in AML. Atlas Genet Cytogenet Oncol Haematol 1997, 1(1):29-31.
68. Imashuku S, Hibi S, Sako M, Lin YW, Ikuta K, Nakata Y, Mori T, Iizuka S, Horibe K, Tsunematsu Y: Hemophagocytosis by leukemic blasts in 7 acute myeloid leukemia cases with t(16;21)(p11;q22): common morphologic characteristics for this type of leukemia. Cancer 2000, 88(8):1970-1975.
69. Kong XT, Ida K, Ichikawa H, Shimizu K, Ohki M, Maseki N, Kaneko Y, Sako M, Kobayashi Y, Tojou A et al: Consistent detection of TLS/FUS-ERG chimeric transcripts in acute myeloid leukemia with t(16;21)(p11;q22) and identification of a novel transcript. Blood 1997, 90(3):1192-1199.
70. Romana SP, Radford-Weiss I, Ben Abdelali R, Schluth C, Petit A, Dastugue N, Talmant P, Bilhou-Nabera C, Mugneret F, Lafage-Pochitaloff M et al: NUP98 rearrangements in hematopoietic malignancies: a study of the Groupe Francophone de Cytogenetique Hematologique. Leukemia 2006, 20(4):696-706.
71. Hollink IH, van den Heuvel-Eibrink MM, Arentsen-Peters ST, Pratcorona M, Abbas S, Kuipers JE, van Galen JF, Beverloo HB, Sonneveld E, Kaspers GJ et al: NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid leukemia with a distinct HOX gene expression pattern. Blood 2011, 118(13):3645-3656.
72. von Neuhoff C, Reinhardt D, Sander A, Zimmermann M, Bradtke J, Betts DR, Zemanova Z, Stary J, Bourquin JP, Haas OA et al: Prognostic impact of specific chromosomal aberrations in a large group of pediatric patients with acute myeloid leukemia treated uniformly according to trial AML-BFM 98. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2010, 28(16):2682-2689.
73. Ma X, Liu Y, Alexandrov LB, Edmonson MN, Gawad C, Zhou X, Li Y, Rusch MC, Easton J, Huether R et al: Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. Nature 2018, 555(7696):371-376.

74. Hollink IH, van den Heuvel-Eibrink MM, Arentsen-Peters ST, Zimmermann M, Peeters JK, Valk PJ, Balgobind BV, Sonneveld E, Kaspers GJ, de Bont ES et al: Characterization of CEBPA mutations and promoter hypermethylation in pediatric acute myeloid leukemia. Haematologica 2011, 96(3):384-392.
75. Manara E, Basso G, Zampini M, Buldini B, Tregnago C, Rondelli R, Masetti R, Bisio V, Frison M, Polato K et al: Characterization of children with FLT3-ITD acute myeloid leukemia: a report from the AIEOP AML-2002 study group. Leukemia 2017, 31(1):18-25.
76. Meshinchi S, Woods WG, Stirewalt DL, Sweetser DA, Buckley JD, Tjoa TK, Bernstein ID, Radich JP: Prevalence and prognostic significance of Flt3 internal tandem duplication in pediatric acute myeloid leukemia. Blood 2001, 97(1):89-94.
77. Meshinchi S, Alonzo TA, Stirewalt DL, Zwaan M, Zimmerman M, Reinhardt D, Kaspers GJ, Heerema NA, Gerbing R, Lange BJ et al: Clinical implications of FLT3 mutations in pediatric AML. Blood 2006, 108(12):3654-3661.
78. Buelow DR, Pounds SB, Wang YD, Shi L, Li Y, Finkelstein D, Shurtleff S, Neale G, Inaba H, Ribeiro RC et al: Uncovering the genomic landscape in Newly Diagnosed and Relapsed Pediatric Cytogenetically normal FLT3-ITD AML. Clinical and translational science 2019, Nov (12(6)):641–647.
79. Ho PA, Kutny MA, Alonzo TA, Gerbing RB, Joaquin J, Raimondi SC, Gamis AS, Meshinchi S: Leukemic mutations in the methylation-associated genes DNMT3A and IDH2 are rare events in pediatric AML: a report from the Children's Oncology Group. Pediatric blood & cancer 2011, 57(2):204-209.
80. Langemeijer SMC, Jansen JH, Hooijer J, van Hoogen P, Stevens-Linders E, Massop M, Waanders E, van Reijmersdal SV, Stevens-Kroef MJPL, Zwaan CM et al: TET2 mutations in childhood leukemia. Leukemia 2011, 25(1):189-192.
81. Li MJ, Yang YL, Lee NC, Jou ST, Lu MY, Chang HH, Lin KH, Peng CT, Lin DT: Tet oncogene family member 2 gene alterations in childhood acute myeloid leukemia. Journal of the Formosan Medical Association = Taiwan yi zhi 2015.
82. Schwarzer A, Emmrich S, Schmidt F, Beck D, Ng M, Reimer C, Adams FF, Grasedieck S, Witte D, Kabler S et al: The non-coding RNA landscape of human hematopoiesis and leukemia. Nature communications 2017, 8(1):218.
83. Kumar S, Bakhshi S: Diagnostic & prognostic role of microRNAs in paediatric acute myeloid leukaemia. The Indian journal of medical research 2016, 144(6):807-814.
84. Danen-van Oorschot AA, Kuipers JE, Arentsen-Peters S, Schotte D, de Haas V, Trka J, Baruchel A, Reinhardt D, Pieters R, Zwaan CM et al: Differentially expressed miRNAs in cytogenetic and molecular subtypes of pediatric acute myeloid leukemia. Pediatric blood & cancer 2012, 58(5):715-721.
85. Daschkey S, Rottgers S, Giri A, Bradtke J, Teigler-Schlegel A, Meister G, Borkhardt A, Landgraf P: MicroRNAs distinguish cytogenetic subgroups in pediatric AML and contribute to complex regulatory networks in AML-relevant pathways. PloS one 2013, 8(2):e56334.
86. Obulkasim A, Katsman-Kuipers JE, Verboon L, Sanders M, Touw I, Jongen-Lavrencic M, Pieters R, Klusmann JH, Michel Zwaan C, van den Heuvel-Eibrink MM et al: Classification of pediatric acute myeloid leukemia based on miRNA expression profiles. Oncotarget 2017, 8(20):33078-33085.
87. Ramamurthy R, Hughes M, Morris V, Bolouri H, Gerbing RB, Wang YC, Loken MR, Raimondi SC, Hirsch BA, Gamis AS et al: miR-155 expression and correlation with clinical outcome in pediatric AML: A report from Children's Oncology Group. Pediatric blood & cancer 2016, 63(12):2096-2103.
88. Gabra MM, Salmena L: microRNAs and Acute Myeloid Leukemia Chemoresistance: A Mechanistic Overview. Frontiers in oncology 2017, 7:255.
89. Ciccone M, Calin GA: MicroRNAs in Myeloid Hematological Malignancies. Current genomics 2015, 16(5):336-348.

90. Mittal N, Li L, Sheng Y, Hu C, Li F, Zhu T, Qiao X, Qian Z: A critical role of epigenetic inactivation of miR-9 in EVI1(high) pediatric AML. Molecular cancer 2019, 18(1):30.
91. Qi HX, Cao Q, Zhou GP, Sun XZ, Zhou WD, Hong Z, Hu J, Juan CX, Li S, Kuai WX: MicroRNA 34b inhibits cell proliferation in pediatric acute myeloid leukemia via regulating LDHA. European review for medical and pharmacological sciences 2019, 23(12):5351-5359.
92. Shivarov V, Stoimenov A, Spassov B, Angelova S, Niagolov M, Ivanova M: Patient-specific microRNA expression profiles as a marker for minimal residual disease in acute myeloid leukemia. Hematology 2014, 19(1):18-21.
93. Xu LH, Guo Y, Cen JN, Yan WY, He HL, Niu YN, Lin YX, Chen CS, Hu SY: Overexpressed miR-155 is associated with initial presentation and poor outcome in Chinese pediatric acute myeloid leukemia. European review for medical and pharmacological sciences 2015, 19(24):4841-4850.
94. Xu L, Guo Y, Yan W, Cen J, Niu Y, Yan Q, He H, Chen CS, Hu S: High level of miR-196b at newly diagnosed pediatric acute myeloid leukemia predicts a poor outcome. EXCLI journal 2017, 16:197-209.
95. Zhang H, Luo XQ, Zhang P, Huang LB, Zheng YS, Wu J, Zhou H, Qu LH, Xu L, Chen YQ: MicroRNA patterns associated with clinical prognostic parameters and CNS relapse prediction in pediatric acute leukemia. PloS one 2009, 4(11):e7826.
96. Garitano-Trojaola A, Agirre X, Prosper F, Fortes P: Long non-coding RNAs in haematological malignancies. Int J Mol Sci 2013, 14(8):15386-15422.
97. Ng M, Heckl D, Klusmann JH: The Regulatory Roles of Long Noncoding RNAs in Acute Myeloid Leukemia. Frontiers in oncology 2019, 9:570.
98. Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, Fraser P: The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science 2008, 322(5908):1717-1720.
99. Guo G, Kang Q, Chen Q, Chen Z, Wang J, Tan L, Chen JL: High expression of long non-coding RNA H19 is required for efficient tumorigenesis induced by Bcr-Abl oncogene. FEBS letters 2014, 588(9):1780-1786.
100. Zhang TJ, Zhou JD, Zhang W, Lin J, Ma JC, Wen XM, Yuan Q, Li XX, Xu ZJ, Qian J: H19 overexpression promotes leukemogenesis and predicts unfavorable prognosis in acute myeloid leukemia. Clinical epigenetics 2018, 10:47.
101. Helsmoortel HH, De Moerloose B, Pieters T, Ghazavi F, Bresolin S, Cave H, de Vries A, de Haas V, Flotho C, Labarque V et al: LIN28B is overexpressed in specific subtypes of paediatric leukaemia and regulates long non-coding RNA H19. Haematologica 2016, 101(6):240-244.
102. Zhang Y, Liu Y, Xu X: Knockdown of LncRNA-UCA1 suppresses chemoresistance of pediatric AML by inhibiting glycolysis through the microRNA-125a/hexokinase 2 pathway. Journal of cellular biochemistry 2018, 119(7):6296-6308.
103. Yin X, Huang S, Zhu R, Fan F, Sun C, Hu Y: Identification of long non-coding RNA competing interactions and biological pathways associated with prognosis in pediatric and adolescent cytogenetically normal acute myeloid leukemia. Cancer cell international 2018, 18:122.
104. Guan X, Wen X, Xiao J, An X, Yu J, Guo Y: Lnc-SOX6-1 upregulation correlates with poor risk stratification and worse treatment outcomes, and promotes cell proliferation while inhibits apoptosis in pediatric acute myeloid leukemia. International journal of laboratory hematology 2019, 41(2):234-241.
105. Fernando TR, Contreras JR, Zampini M, Rodriguez-Malave NI, Alberti MO, Anguiano J, Tran TM, Palanichamy JK, Gajeton J, Ung NM et al: The lncRNA CASC15 regulates SOX4 expression in RUNX1-rearranged acute leukemia. Molecular cancer 2017, 16(1):126.
106. Cao L, Xiao PF, Tao YF, Hu SY, Lu J, Zhao WL, Li ZH, Wang NN, Wang J, Feng X et al: Microarray profiling of bone marrow long non-coding RNA expression in Chinese pediatric acute myeloid leukemia patients. Oncology reports 2016, 35(2):757-770.
107. Ledford H: Circular RNAs throw genetics for a loop. Nature 2013, 494(7438):415.

108. Jamal M, Song T, Chen B, Faisal M, Hong Z, Xie T, Wu Y, Pan S, Yin Q, Shao L et al: Recent Progress on Circular RNA Research in Acute Myeloid Leukemia. Frontiers in oncology 2019, 9:1108.
109. Vromman M, Vandesompele J, Volders PJ: Closing the circle: current state and perspectives of circular RNA databases. 00 2020, (00):1-10.
110. Xu LH, Fang JP, Liu YC, Jones AI, Chai L: Nucleophosmin mutations confer an independent favorable prognostic impact in 869 pediatric patients with acute myeloid leukemia. Blood cancer journal 2020, 10(1):1.
111. Metzeler KH, Herold T, Rothenberg-Thurley M, Amler S, Sauerland MC, Gorlich D, Schneider S, Konstandin NP, Dufour A, Braundl K et al: Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. Blood 2016, 128(5):686-698.
112. Dahl NA, Michaels ST, McMasters RL, Chandra S, O'Brien MM: Azacitidine and Sorafenib Therapy in a Pediatric Patient With Refractory Acute Myeloid Leukemia With Monosomy 7 and Somatic PTPN11 Mutation. Pediatric blood & cancer 2015, Mar;63(3):551-553.
113. Pollard JA, Alonzo TA, Gerbing RB, Ho PA, Zeng R, Ravindranath Y, Dahl G, Lacayo NJ, Becton D, Chang M et al: Prevalence and prognostic significance of KIT mutations in pediatric patients with core binding factor AML enrolled on serial pediatric cooperative trials for de novo AML. Blood 2010, 115(12):2372-2379.
114. Paschka P, Marcucci G, Ruppert AS, Mrozek K, Chen H, Kittles RA, Vukosavljevic T, Perrotti D, Vardiman JW, Carroll AJ et al: Adverse prognostic significance of KIT mutations in adult acute myeloid leukemia with inv(16) and t(8;21): a Cancer and Leukemia Group B Study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2006, 24(24):3904-3911.
115. Ho PA, Zeng R, Alonzo TA, Gerbing RB, Miller KL, Pollard JA, Stirewalt DL, Heerema NA, Raimondi SC, Hirsch B et al: Prevalence and prognostic implications of WT1 mutations in pediatric acute myeloid leukemia (AML): a report from the Children's Oncology Group. Blood 2010, 116(5):702-710.
116. Hou HA, Huang TC, Lin LI, Liu CY, Chen CY, Chou WC, Tang JL, Tseng MH, Huang CF, Chiang YC et al: WT1 mutation in 470 adult patients with acute myeloid leukemia: stability during disease evolution and implication of its incorporation into a survival scoring system. Blood 2010, 115(25):5222-5231.
117. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, Potter NE, Heuser M, Thol F, Bolli N et al: Genomic Classification and Prognosis in Acute Myeloid Leukemia. The New England journal of medicine 2016, 374(23):2209-2221.
118. Straube J, Ling VY, Hill GR, Lane SW: The impact of age, NPM1(mut), and FLT3(ITD) allelic ratio in patients with acute myeloid leukemia. Blood 2018, 131(10):1148-1153.
119. Sandahl JD, Kjeldsen E, Abrahamsson J, Ha SY, Heldrup J, Jahnukainen K, Jonsson OG, Lausen B, Palle J, Zeller B et al: The applicability of the WHO classification in paediatric AML. A NOPHO-AML study. British journal of haematology 2015, 169(6):859-867.
120. Kaspers GJ: Pediatric acute myeloid leukemia. Expert review of anticancer therapy 2012, 12(3):405-413.
121. Abrahamsson J, Forestier E, Heldrup J, Jahnukainen K, Jonsson OG, Lausen B, Palle J, Zeller B, Hasle H: Response-guided induction therapy in pediatric acute myeloid leukemia with excellent remission rate. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2011, 29(3):310-315.
122. De Moerloose B, Reedijk A, de Bock GH, Lammens T, de Haas V, Denys B, Dedeken L, van den Heuvel-Eibrink MM, Te Loo M, Uyttebroeck A et al: Response-guided chemotherapy for pediatric acute myeloid leukemia without hematopoietic stem cell transplantation in first complete remission: Results from protocol DB AML-01. Pediatric blood & cancer 2019:e27605.
123. Creutzig U, Zimmermann M, Lehrnbecher T, Graf N, Hermann J, Niemeyer CM, Reiter A, Ritter J, Dworzak M, Stary J et al: Less toxicity by optimizing chemotherapy, but not by addition of granulocyte colony-stimulating factor in children and adolescents with acute myeloid

leukemia: results of AML-BFM 98. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2006, 24(27):4499-4506.
124. Rubnitz JE, Inaba H, Dahl G, Ribeiro RC, Bowman WP, Taub J, Pounds S, Razzouk BI, Lacayo NJ, Cao X et al: Minimal residual disease-directed therapy for childhood acute myeloid leukaemia: results of the AML02 multicentre trial. The Lancet Oncology 2010, 11(6):543-552.
125. Tierens A, Bjorklund E, Siitonen S, Marquart HV, Wulff-Juergensen G, Pelliniemi TT, Forestier E, Hasle H, Jahnukainen K, Lausen B et al: Residual disease detected by flow cytometry is an independent predictor of survival in childhood acute myeloid leukaemia; results of the NOPHO-AML 2004 study. British journal of haematology 2016.
126. Ossenkoppele GJ, van de Loosdrecht AA, Schuurhuis GJ: Review of the relevance of aberrant antigen expression by flow cytometry in myeloid neoplasms. British journal of haematology 2011, 153(4):421-436.
127. Reading CL, Estey EH, Huh YO, Claxton DF, Sanchez G, Terstappen LW, O'Brien MC, Baron S, Deisseroth AB: Expression of unusual immunophenotype combinations in acute myelogenous leukemia. Blood 1993, 81(11):3083-3090.
128. Sievers EL, Lange BJ, Alonzo TA, Gerbing RB, Bernstein ID, Smith FO, Arceci RJ, Woods WG, Loken MR: Immunophenotypic evidence of leukemia after induction therapy predicts relapse: results from a prospective Children's Cancer Group study of 252 patients with acute myeloid leukemia. Blood 2003, 101(9):3398-3406.
129. Loken MR, Alonzo TA, Pardo L, Gerbing RB, Raimondi SC, Hirsch BA, Ho PA, Franklin J, Cooper TM, Gamis AS et al: Residual disease detected by multidimensional flow cytometry signifies high relapse risk in patients with de novo acute myeloid leukemia: a report from Children's Oncology Group. Blood 2012, 120(8):1581-1588.
130. Inaba H, Coustan-Smith E, Cao X, Pounds SB, Shurtleff SA, Wang KY, Raimondi SC, Onciu M, Jacobsen J, Ribeiro RC et al: Comparative analysis of different approaches to measure treatment response in acute myeloid leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012, 30(29):3625-3632.
131. Kaspers GJ, Zwaan CM: Pediatric acute myeloid leukemia: towards high-quality cure of all patients. Haematologica 2007, 92(11):1519-1532.
132. Bonaventure A, Harewood R, Stiller CA, Gatta G, Clavel J, Stefan DC, Carreira H, Spika D, Marcos-Gragera R, Peris-Bonet R et al: Worldwide comparison of survival from childhood leukaemia for 1995-2009, by subtype, age, and sex (CONCORD-2): a population-based study of individual data for 89 828 children from 198 registries in 53 countries. Lancet Haematol 2017, 4(5):e202-e217.
133. Zwaan CM, Kolb EA, Reinhardt D, Abrahamsson J, Adachi S, Aplenc R, De Bont ES, De Moerloose B, Dworzak M, Gibson BE et al: Collaborative Efforts Driving Progress in Pediatric Acute Myeloid Leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2015, 33(27):2949-2962.
134. Protocol N-DA: Research study for treatment of children and adolescents with acute myeloid leukaemia 0-18 years. . EudraCT nr 2012-002934-35.
135. Chair) JAS: Research study for treatment of children and adolescents with acute myeloid leukaemia 0-18 years. . NOPHO-DBH AML 2012 protocol v21 2013, EUdract number 2012-002934-35.
136. Davila J, Slotkin E, Renaud T: Relapsed and refractory pediatric acute myeloid leukemia: current and emerging treatments. Paediatric drugs 2014, 16(2):151-168.
137. Zeijlemaker W, Gratama JW, Schuurhuis GJ: Tumor heterogeneity makes AML a "moving target" for detection of residual disease. Cytometry Part B, Clinical cytometry 2014, 86(1):3-14.
138. Grimwade D, Ivey A, Huntly BJ: Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance. Blood 2016, 127(1):29-41.

139. Coustan-Smith E, Song G, Shurtleff S, Yeoh AE, Chng WJ, Chen SP, Rubnitz JE, Pui CH, Downing JR, Campana D: Universal monitoring of minimal residual disease in acute myeloid leukemia. JCI insight 2018, 3(9).
140. Cloos J, Goemans BF, Hess CJ, van Oostveen JW, Waisfisz Q, Corthals S, de Lange D, Boeckx N, Hahlen K, Reinhardt D et al: Stability and prognostic influence of FLT3 mutations in paired initial and relapsed AML samples. Leukemia 2006, 20(7):1217-1220.
141. Bachas C, Schuurhuis GJ, Hollink IH, Kwidama ZJ, Goemans BF, Zwaan CM, van den Heuvel-Eibrink MM, de Bont ES, Reinhardt D, Creutzig U et al: High-frequency type I/II mutational shifts between diagnosis and relapse are associated with outcome in pediatric AML: implications for personalized medicine. Blood 2010, 116(15):2752-2758.
142. Farrar JE, Schuback HL, Ries RE, Wai D, Hampton OA, Trevino LR, Alonzo TA, Guidry Auvil JM, Davidsen TM, Gesuwan P et al: Genomic profiling of pediatric acute myeloid leukemia reveals a changing mutational landscape from disease diagnosis to relapse. Cancer research 2016, Apr 15(76(8)):2197-2205.
143. Bachas C, Schuurhuis GJ, Zwaan CM, van den Heuvel-Eibrink MM, den Boer ML, de Bont ES, Kwidama ZJ, Reinhardt D, Creutzig U, de Haas V et al: Gene expression profiles associated with pediatric relapsed AML. PloS one 2015, 10(4):e0121730.
144. Rapin N, Bagger FO, Jendholm J, Mora-Jensen H, Krogh A, Kohlmann A, Thiede C, Borregaard N, Bullinger L, Winther O et al: Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. Blood 2014, 123(6):894-904.
145. Zhong JF, Chen Y, Marcus JS, Scherer A, Quake SR, Taylor CR, Weiner LP: A microfluidic processor for gene expression profiling of single human embryonic stem cells. Lab on a chip 2008, 8(1):68-74.
146. Svensson V: Moore ' s Law in Single Cell Transcriptomics. 2017.
147. Svensson V, Natarajan KN, Ly LH, Miragaia RJ, Labalette C, Macaulay IC, Cvejic A, Teichmann SA: Power analysis of single-cell RNA-sequencing experiments. Nature methods 2017, 14(4):381-387.
148. Potter N, Miraki-Moud F, Ermini L, Titley I, Vijayaraghavan G, Papaemmanuil E, Campbell P, Gribben J, Taussig D, Greaves M: Single cell analysis of clonal architecture in acute myeloid leukaemia. Leukemia 2019, 33(5):1113-1123.
149. Kalisky T, Quake SR: Single-cell genomics. Nature methods 2011, 8(4):311-314.
150. Clarkson BD: Review of recent studies of cellular proliferation in acute leukemia. National Cancer Institute monograph 1969, 30:81-120.
151. Clarkson B, Ohkita T, Ota K, Fried J: Studies of cellular proliferation in human leukemia. I. Estimation of growth rates of leukemic and normal hematopoietic cells in two adults with acute leukemia given single injections of tritiated thymidine. The Journal of clinical investigation 1967, 46(4):506-529.
152. Metcalf D, Moore MA, Warner NL: Colony formation in vitro by myelomonocytic leukemic cells. Journal of the National Cancer Institute 1969, 43(4):983-1001.
153. Moore MA, Williams N, Metcalf D: In vitro colony formation by normal and leukemic human hematopoietic cells: interaction between colony-forming and colony-stimulating cells. Journal of the National Cancer Institute 1973, 50(3):591-602.
154. Sutherland HJ, Eaves CJ, Eaves AC, Dragowska W, Lansdorp PM: Characterization and partial purification of human marrow cells capable of initiating long-term hematopoiesis in vitro. Blood 1989, 74(5):1563-1570.
155. Lapidot T, Sirard C, Vormoor J, Murdoch B, Hoang T, Caceres-Cortes J, Minden M, Paterson B, Caligiuri MA, Dick JE: A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature 1994, 367(6464):645-648.
156. Blair A, Hogge DE, Ailles LE, Lansdorp PM, Sutherland HJ: Lack of expression of Thy-1 (CD90) on acute myeloid leukemia cells with long-term proliferative ability in vitro and in vivo. Blood 1997, 89(9):3104-3112.

157. Bonnet D, Dick JE: Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nature medicine 1997, 3(7):730-737.
158. Ailles LE, Gerhard B, Hogge DE: Detection and characterization of primitive malignant and normal progenitors in patients with acute myelogenous leukemia using long-term coculture with supportive feeder layers and cytokines. Blood 1997, 90(7):2555-2564.
159. Ishikawa F, Yoshida S, Saito Y, Hijikata A, Kitamura H, Tanaka S, Nakamura R, Tanaka T, Tomiyama H, Saito N et al: Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. Nat Biotechnol 2007, 25(11):1315-1321.
160. Sarry JE, Murphy K, Perry R, Sanchez PV, Secreto A, Keefer C, Swider CR, Strzelecki AC, Cavelier C, Recher C et al: Human acute myelogenous leukemia stem cells are rare and heterogeneous when assayed in NOD/SCID/IL2R gamma c-deficient mice. Journal of Clinical Investigation 2011, 121(1):384-395.
161. Eppert K, Takenaka K, Lechman ER, Waldron L, Nilsson B, van Galen P, Metzeler KH, Poeppl A, Ling V, Beyene J et al: Stem cell gene expression programs influence clinical outcome in human leukemia. Nature medicine 2011, 17(9):1086-1093.
162. Krivtsov AV, Twomey D, Feng Z, Stubbs MC, Wang Y, Faber J, Levine JE, Wang J, Hahn WC, Gilliland DG et al: Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature 2006, 442(7104):818-822.
163. Cozzio A, Passegue E, Ayton PM, Karsunky H, Cleary ML, Weissman IL: Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors. Genes & development 2003, 17(24):3029-3035.
164. So CW, Karsunky H, Passegue E, Cozzio A, Weissman IL, Cleary ML: MLL-GAS7 transforms multipotent hematopoietic progenitors and induces mixed lineage leukemias in mice. Cancer cell 2003, 3(2):161-171.
165. Jamieson CH, Ailles LE, Dylla SJ, Muijtjens M, Jones C, Zehnder JL, Gotlib J, Li K, Manz MG, Keating A et al: Granulocyte-macrophage progenitors as candidate leukemic stem cells in blast-crisis CML. The New England journal of medicine 2004, 351(7):657-667.
166. Taussig DC, Vargaftig J, Miraki-Moud F, Griessinger E, Sharrock K, Luke T, Lillington D, Oakervee H, Cavenagh J, Agrawal SG et al: Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34(-) fraction. Blood 2010, 115(10):1976-1984.
167. Shih AH, Jiang Y, Meydan C, Shank K, Pandey S, Barreyro L, Antony-Debre I, Viale A, Socci N, Sun Y et al: Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. Cancer cell 2015, 27(4):502-515.
168. Gal H, Amariglio N, Trakhtenbrot L, Jacob-Hirsh J, Margalit O, Avigdor A, Nagler A, Tavor S, Ein-Dor L, Lapidot T et al: Gene expression profiles of AML derived stem cells; similarity to hematopoietic stem cells. Leukemia 2006, 20(12):2147-2154.
169. Kreso A, Dick JE: Evolution of the cancer stem cell model. Cell stem cell 2014, 14(3):275-291.
170. Martelli MP, Pettirossi V, Thiede C, Bonifacio E, Mezzasoma F, Cecchini D, Pacini R, Tabarrini A, Ciurnelli R, Gionfriddo I et al: CD34+ cells from AML with mutated NPM1 harbor cytoplasmic mutated nucleophosmin and generate leukemia in immunocompromised mice. Blood 2010, 116(19):3907-3922.
171. Zeijlemaker W, Kelder A, Wouters R, Valk PJ, Witte BI, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: Absence of leukaemic CD34 cells in acute myeloid leukaemia is of high prognostic value: a longstanding controversy deciphered. British journal of haematology 2015, Oct(171(2)):227-238.
172. van der Pol MA, Feller N, Roseboom M, Moshaver B, Westra G, Broxterman HJ, Ossenkoppele GJ, Schuurhuis GJ: Assessment of the normal or leukemic nature of CD34+ cells in acute myeloid leukemia with low percentages of CD34 cells. Haematologica 2003, 88(9):983-993.
173. Quek L, Otto GW, Garnett C, Lhermitte L, Karamitros D, Stoilova B, Lau IJ, Doondeea J, Usukhbayar B, Kennedy A et al: Genetically distinct leukemic stem cells in human CD34- acute

myeloid leukemia are arrested at a hemopoietic precursor-like stage. The Journal of experimental medicine 2016, 213(8):1513-1535.
174. Goyama S, Wunderlich M, Mulloy JC: Xenograft models for normal and malignant stem cells. Blood 2015, 125(17):2630-2640.
175. Taussig DC, Miraki-Moud F, Anjos-Afonso F, Pearce DJ, Allen K, Ridler C, Lillington D, Oakervee H, Cavenagh J, Agrawal SG et al: Anti-CD38 antibody-mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. Blood 2008, 112(3):568-575.
176. Vargaftig J, Taussig DC, Griessinger E, Anjos-Afonso F, Lister TA, Cavenagh J, Oakervee H, Gribben J, Bonnet D: Frequency of leukemic initiating cells does not depend on the xenotransplantation model used. Leukemia 2012, 26(4):858-860.
177. Wunderlich M, Chou FS, Link KA, Mizukawa B, Perry RL, Carroll M, Mulloy JC: AML xenograft efficiency is significantly improved in NOD/SCID-IL2RG mice constitutively expressing human SCF, GM-CSF and IL-3. Leukemia 2010, 24(10):1785-1788.
178. Ailles LE, Gerhard B, Kawagoe M, Hogge DE: Growth characteristics of acute myelogenous leukemia progenitors that initiate malignant hematopoiesis in nonobese diabetic/severe combined immunodeficient mice. Blood 1999, 94(5):1761-1772.
179. Wang K, Sanchez-Martin M, Wang X, Knapp KM, Koche R, Vu L, Nahas MK, He J, Hadler M, Stein EM et al: Patient-derived xenotransplants can recapitulate the genetic driver landscape of acute leukemias. Leukemia 2017, 31(1):151-158.
180. Pearce DJ, Taussig D, Zibara K, Smith LL, Ridler CM, Preudhomme C, Young BD, Rohatiner AZ, Lister TA, Bonnet D: AML engraftment in the NOD/SCID assay reflects the outcome of AML: implications for our understanding of the heterogeneity of AML. Blood 2006, 107(3):1166-1173.
181. Klco JM, Spencer DH, Miller CA, Griffith M, Lamprecht TL, O'Laughlin M, Fronick C, Magrini V, Demeter RT, Fulton RS et al: Functional heterogeneity of genetically defined subclones in acute myeloid leukemia. Cancer cell 2014, 25(3):379-392.
182. Griessinger E, Anjos-Afonso F, Pizzitola I, Rouault-Pierre K, Vargaftig J, Taussig D, Gribben J, Lassailly F, Bonnet D: A niche-like culture system allowing the maintenance of primary human acute myeloid leukemia-initiating cells: a new tool to decipher their chemoresistance and self-renewal mechanisms. Stem cells translational medicine 2014, 3(4):520-529.
183. Giulia Borella VB, Ambra Da Ros, Claudia Tregnago, Maddalena Benetton, Elisabetta Campodoni, Silvia Panseri, Monica Sandri, Monica Montesi, Stefano Cairo, Franco Locatelli, Martina PigazziEHA Library. Borella G. : Development of innovative preclinical in vitro and in vivo tools for an effective therapeutic strategy in pediatric acute myeloid leukemia. 24th EHA congress, Amsterdam, the Netherlands 2019, Jun 15(266843; PS1226).
184. Ito S, Barrett AJ, Dutra A, Pak E, Miner S, Keyvanfar K, Hensel NF, Rezvani K, Muranski P, Liu P et al: Long term maintenance of myeloid leukemic stem cells cultured with unrelated human mesenchymal stromal cells. Stem cell research 2015, 14(1):95-104.
185. Hoffmann P, Eder R, Boeld TJ, Doser K, Piseshka B, Andreesen R, Edinger M: Only the CD45RA+ subpopulation of CD4+CD25high T cells gives rise to homogeneous regulatory T-cell lines upon in vitro expansion. Blood 2006, 108(13):4260-4267.
186. Kersten B, Valkering M, Wouters R, van Amerongen R, Hanekamp D, Kwidama Z, Valk P, Ossenkoppele G, Zeijlemaker W, Kaspers G et al: CD45RA, a specific marker for leukaemia stem cell sub-populations in acute myeloid leukaemia. British journal of haematology 2016, 173(2):219-235.
187. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, Scholten WJ, Snel AN, Veldhuizen D, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. Leukemia 2016, 30(2):439-446.
188. van Rhenen A, Moshaver B, Kelder A, Feller N, Nieuwint AW, Zweegman S, Ossenkoppele GJ, Schuurhuis GJ: Aberrant marker expression patterns on the CD34+CD38- stem cell

compartment in acute myeloid leukemia allows to distinguish the malignant from the normal stem cell compartment both at diagnosis and in remission. Leukemia 2007, 21(8):1700-1707.
189. Eidenschink Brodersen L, Alonzo TA, Menssen AJ, Gerbing RB, Pardo L, Voigt AP, Kahwash SB, Hirsch B, Raimondi S, Gamis AS et al: A recurrent immunophenotype at diagnosis independently identifies high-risk pediatric acute myeloid leukemia: a report from Children's Oncology Group. Leukemia 2016, 30(10):2077-2080.
190. Hanekamp D, Denys B, Kaspers GJL, Te Marvelde JG, Schuurhuis GJ, De Haas V, De Moerloose B, de Bont ES, Zwaan CM, de Jong A et al: Leukaemic stem cell load at diagnosis predicts the development of relapse in young acute myeloid leukaemia patients. British journal of haematology 2018, 183(3):512-516.
191. Hosen N, Park CY, Tatsumi N, Oji Y, Sugiyama H, Gramatzki M, Krensky AM, Weissman IL: CD96 is a leukemic stem cell-specific marker in human acute myeloid leukemia. Proceedings of the National Academy of Sciences of the United States of America 2007, 104(26):11008-11013.
192. Du W, Hu Y, Lu C, Li J, Liu W, He Y, Wang P, Cheng C, Hu YU, Huang S et al: Cluster of differentiation 96 as a leukemia stem cell-specific marker and a factor for prognosis evaluation in leukemia. Molecular and clinical oncology 2015, 3(4):833-838.
193. Chavez-Gonzalez A, Dorantes-Acosta E, Moreno-Lorenzana D, Alvarado-Moreno A, Arriaga-Pizano L, Mayani H: Expression of CD90, CD96, CD117, and CD123 on different hematopoietic cell populations from pediatric patients with acute myeloid leukemia. Archives of medical research 2014, 45(4):343-350.
194. Pelosi E, Castelli G, Testa U: Targeting LSCs through membrane antigens selectively or preferentially expressed on these cells. Blood cells, molecules & diseases 2015, 55(4):336-346.
195. Graf M, Reif S, Kroll T, Hecht K, Nuessler V, Schmetzer H: Expression of MAC-1 (CD11b) in acute myeloid leukemia (AML) is associated with an unfavorable prognosis. American journal of hematology 2006, 81(4):227-235.
196. Han G, Chen G, Shen B, Li Y: Tim-3: an activation marker and activation limiter of innate immune cells. Frontiers in immunology 2013, 4:449.
197. Kikushige Y, Miyamoto T: Identification of TIM-3 as a Leukemic Stem Cell Surface Molecule in Primary Acute Myeloid Leukemia. Oncology 2015, 89 Suppl 1:28-32.
198. Kikushige Y, Shima T, Takayanagi S, Urata S, Miyamoto T, Iwasaki H, Takenaka K, Teshima T, Tanaka T, Inagaki Y et al: TIM-3 is a promising target to selectively kill acute myeloid leukemia stem cells. Cell stem cell 2010, 7(6):708-717.
199. Jan M, Chao MP, Cha AC, Alizadeh AA, Gentles AJ, Weissman IL, Majeti R: Prospective separation of normal and leukemic stem cells based on differential expression of TIM3, a human acute myeloid leukemia stem cell marker. Proceedings of the National Academy of Sciences of the United States of America 2011, 108(12):5009-5014.
200. van Rhenen A, van Dongen GAMS, Kelder A, Rombouts EJ, Feller N, Moshaver B, Stigter-van Walsum M, Zweegman S, Ossenkoppele GJ, Schuurhuis GJ: The novel AML stem cell-associated antigen CLL-1 aids in discrimination between normal and leukemic stem cells. Blood 2007, 110(7):2659-2666.
201. Terwijn M, Zeijlemaker W, Kelder A, Rutten AP, Snel AN, Scholten WJ, Pabst T, Verhoef G, Lowenberg B, Zweegman S et al: Leukemic Stem Cell Frequency: A Strong Biomarker for Clinical Outcome in Acute Myeloid Leukemia. PloS one 2014, 9(9).
202. Bill M, Aggerholm A, Kjeldsen E, Roug AS, Hokland P, Nederby L: Revisiting CLEC12A as leukaemic stem cell marker in AML: highlighting the necessity of precision diagnostics in patients eligible for targeted therapy. British journal of haematology 2018, 184(5):769-781.
203. Bakker AB, van den Oudenrijn S, Bakker AQ, Feller N, van Meijer M, Bia JA, Jongeneelen MA, Visser TJ, Bijl N, Geuijen CA et al: C-type lectin-like molecule-1: a novel myeloid cell surface marker associated with acute myeloid leukemia. Cancer research 2004, 64(22):8443-8450.
204. Depreter B, Hanekamp, D., Lammens, T. et al. : Role of CLEC12A in the LSC compartment of pediatric acute myeloid leukemia at diagnosis versus relapse. 23rd Congress of EHA June 14-

17, 2018, Stockholm, Sweden 2018, Topic: Acute myeloid leukemia - Biology & Translational Research:Abstract Code: PB1693.

205. Thomas D, Vadas M, Lopez A: Regulation of haematopoiesis by growth factors - emerging insights and therapies. Expert Opin Biol Ther 2004, 4(6):869-879.
206. van Rhenen A, Feller N, Kelder A, Westra AH, Rombouts E, Zweegman S, van der Pol MA, Waisfisz Q, Ossenkoppele GJ, Schuurhuis GJ: High stem cell frequency in acute myeloid leukemia at diagnosis predicts high minimal residual disease and poor survival. Clinical Cancer Research 2005, 11(18):6520-6527.
207. Jordan CT, Upchurch D, Szilvassy SJ, Guzman ML, Howard DS, Pettigrew AL, Meyerrose T, Rossi R, Grimes B, Rizzieri DA et al: The interleukin-3 receptor alpha chain is a unique marker for human acute myelogenous leukemia stem cells. Leukemia 2000, 14(10):1777-1784.
208. Taussig DC, Pearce DJ, Simpson C, Rohatiner AZ, Lister TA, Kelly G, Luongo JL, Danet-Desnoyers GA, Bonnet D: Hematopoietic stem cells express multiple myeloid markers: implications for the origin and targeted therapy of acute myeloid leukemia. Blood 2005, 106(13):4086-4092.
209. Majeti R: Monoclonal antibody therapy directed against human acute myeloid leukemia stem cells. Oncogene 2011, 30(9):1009-1019.
210. Hwang K, Park CJ, Jang S, Chi HS, Kim DY, Lee JH, Lee KH, Im HJ, Seo JJ: Flow cytometric quantification and immunophenotyping of leukemic stem cells in acute myeloid leukemia. Annals of hematology 2012, 91(10):1541-1546.
211. Vergez F, Green AS, Tamburini J, Sarry JE, Gaillard B, Cornillet-Lefebvre P, Pannetier M, Neyret A, Chapuis N, Ifrah N et al: High levels of CD34+CD38low/-CD123+ blasts are predictive of an adverse outcome in acute myeloid leukemia: a Groupe Ouest-Est des Leucemies Aigues et Maladies du Sang (GOELAMS) study. Haematologica 2011, 96(12):1792-1798.
212. Testa U, Riccioni R, Militi S, Coccia E, Stellacci E, Samoggia P, Latagliata R, Mariani G, Rossini A, Battistini A et al: Elevated expression of IL-3Ralpha in acute myelogenous leukemia is associated with enhanced blast proliferation, increased cellularity, and poor prognosis. Blood 2002, 100(8):2980-2988.
213. Guzman ML, Neering SJ, Upchurch D, Grimes B, Howard DS, Rizzieri DA, Luger SM, Jordan CT: Nuclear factor-kappaB is constitutively activated in primitive human acute myelogenous leukemia cells. Blood 2001, 98(8):2301-2307.
214. Larsen HO, Roug AS, Just T, Brown GD, Hokland P: Expression of the hMICL in acute myeloid leukemia-a highly reliable disease marker at diagnosis and during follow-up. Cytometry Part B, Clinical cytometry 2012, 82(1):3-8.
215. Roug AS, Larsen HO, Nederby L, Just T, Brown G, Nyvold CG, Ommen HB, Hokland P: hMICL and CD123 in combination with a CD45/CD34/CD117 backbone - a universal marker combination for the detection of minimal residual disease in acute myeloid leukaemia. British journal of haematology 2014, 164(2):212-222.
216. Aiuti A, Webb IJ, Bleul C, Springer T, Gutierrez-Ramos JC: The chemokine SDF-1 is a chemoattractant for human CD34+ hematopoietic progenitor cells and provides a new mechanism to explain the mobilization of CD34+ progenitors to peripheral blood. The Journal of experimental medicine 1997, 185(1):111-120.
217. Liu T, Liu X, Xiang J, Zou P, Zhou J, Chen Y, Yu D, Li C: [Study on the relationship between the expression of adhesion molecules and the invasiveness of acute myeloid leukemia cells]. Zhonghua xue ye xue za zhi = Zhonghua xueyexue zazhi 1997, 18(1):29-31.
218. Burger JA, Spoo A, Dwenger A, Burger M, Behringer D: CXCR4 chemokine receptors (CD184) and alpha4beta1 integrins mediate spontaneous migration of human CD34+ progenitors and acute myeloid leukaemia cells beneath marrow stromal cells (pseudoemperipolesis). British journal of haematology 2003, 122(4):579-589.
219. Tavernier-Tardy E, Cornillon J, Campos L, Flandrin P, Duval A, Nadal N, Guyotat D: Prognostic value of CXCR4 and FAK expression in acute myelogenous leukemia. Leukemia research 2009, 33(6):764-768.

220. Becker PS, Kopecky KJ, Wilks AN, Chien S, Harlan JM, Willman CL, Petersdorf SH, Stirewalt DL, Papayannopoulou T, Appelbaum FR: Very late antigen-4 function of myeloblasts correlates with improved overall survival for patients with acute myeloid leukemia. Blood 2009, 113(4):866-874.
221. Matsunaga T, Takemoto N, Sato T, Takimoto R, Tanaka I, Fujimi A, Akiyama T, Kuroda H, Kawano Y, Kobune M et al: Interaction between leukemic-cell VLA-4 and stromal fibronectin is a decisive factor for minimal residual disease of acute myelogenous leukemia (vol 9, pg 1158, 2003). Nature medicine 2005, 11(5):578-578.
222. Walter RB, Alonzo TA, Gerbing RB, Ho PA, Smith FO, Raimondi SC, Hirsch BA, Gamis AS, Franklin JL, Hurwitz CA et al: High expression of the very late antigen-4 integrin independently predicts reduced risk of relapse and improved outcome in pediatric acute myeloid leukemia: a report from the children's oncology group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2010, 28(17):2831-2838.
223. Nakayama F, Nishihara S, Iwasaki H, Kudo T, Okubo R, Kaneko M, Nakamura M, Karube M, Sasaki K, Narimatsu H: CD15 expression in mature granulocytes is determined by alpha 1,3-fucosyltransferase IX, but in promyelocytes and monocytes by alpha 1,3-fucosyltransferase IV. The Journal of biological chemistry 2001, 276(19):16100-16106.
224. Buttler D, Mai JK, Ashwell KW, Andressen C: Clonogenic CD15 immunoreactive radial glial cells from the developing human lateral ganglionic eminence. Current pharmaceutical biotechnology 2013, 14(1):29-35.
225. Ball ED, Schwarz LM, Bloomfield CD: Expression of the CD15 antigen on normal and leukemic myeloid cells: effects of neuraminidase and variable detection with a panel of monoclonal antibodies. Molecular immunology 1991, 28(9):951-958.
226. Chadburn A, Inghirami G, Knowles DM: The kinetics and temporal expression of T-cell activation-associated antigens CD15 (LeuM1), CD30 (Ki-1), EMA, and CD11c (LeuM5) by benign activated T cells. Hematologic pathology 1992, 6(4):193-202.
227. Griffin JD, Davis R, Nelson DA, Davey FR, Mayer RJ, Schiffer C, McIntyre OR, Bloomfield CD: Use of surface marker analysis to predict outcome of adult acute myeloblastic leukemia. Blood 1986, 68(6):1232-1241.
228. Bueno C, Montes R, Martin L, Prat I, Hernandez MC, Orfao A, Menendez P: NG2 antigen is expressed in CD34+ HPCs and plasmacytoid dendritic cell precursors: is NG2 expression in leukemia dependent on the target cell where leukemogenesis is triggered? Leukemia 2008, 22(8):1475-1478.
229. Agarwala S, Tamplin OJ: Neural Crossroads in the Hematopoietic Stem Cell Niche. Trends in cell biology 2018, 28(12):987-998.
230. Stallcup WB, Cohn M: Correlation of surface antigens and cell type in cloned cell lines from the rat central nervous system. Experimental cell research 1976, 98(2):285-297.
231. Lopez-Millan B, Sanchez-Martinez D, Roca-Ho H, Gutierrez-Aguera F, Molina O, Diaz de la Guardia R, Torres-Ruiz R, Fuster JL, Ballerini P, Suessbier U et al: NG2 antigen is a therapeutic target for MLL-rearranged B-cell acute lymphoblastic leukemia. Leukemia 2019.
232. Salzman GS, Ackerman SD, Ding C, Koide A, Leon K, Luo R, Stoveken HM, Fernandez CG, Tall GG, Piao X et al: Structural Basis for Regulation of GPR56/ADGRG1 by Its Alternatively Spliced Extracellular Domains. Neuron 2016, 91(6):1292-1304.
233. Luo R, Jeong SJ, Yang A, Wen M, Saslowsky DE, Lencer WI, Arac D, Piao X: Mechanism for adhesion G protein-coupled receptor GPR56-mediated RhoA activation induced by collagen III stimulation. PloS one 2014, 9(6):e100043.
234. Luo R, Jeong SJ, Jin Z, Strokes N, Li S, Piao X: G protein-coupled receptor 56 and collagen III, a receptor-ligand pair, regulates cortical development and lamination. Proceedings of the National Academy of Sciences of the United States of America 2011, 108(31):12925-12930.
235. Pabst C, Bergeron A, Lavallee VP, Yeh J, Gendron P, Norddahl GL, Krosl J, Boivin I, Deneault E, Simard J et al: GPR56 identifies primary human acute myeloid leukemia cells with high repopulating potential in vivo. Blood 2016, 127(16)(16):2018-2027.

236. Daga S, Rosenberger A, Quehenberger F, Krisper N, Prietl B, Reinisch A, Zebisch A, Sill H, Wolfler A: High GPR56 surface expression correlates with a leukemic stem cell gene signature in CD34-positive AML. Cancer medicine 2019, 8(4):1771-1778.
237. Elsayed AH, Rafiee R, Cao X, Raimondi S, Downing JR, Ribeiro R, Fan Y, Gruber TA, Baker S, Klco J et al: A six-gene leukemic stem cell score identifies high risk pediatric acute myeloid leukemia. Leukemia 2019, doi: 10.1038/s41375-019-0604-8. [Epub ahead of print].
238. Daria D, Kirsten N, Muranyi A, Mulaw M, Ihme S, Kechter A, Hollnagel M, Bullinger L, Dohner K, Dohner H et al: GPR56 contributes to the development of acute myeloid leukemia in mice. Leukemia 2016, 30(8):1734-1741.
239. Rao TN, Marks-Bluth J, Sullivan J, Gupta MK, Chandrakanthan V, Fitch SR, Ottersbach K, Jang YC, Piao X, Kulkarni RN et al: High-level Gpr56 expression is dispensable for the maintenance and function of hematopoietic stem and progenitor cells in mice. Stem cell research 2015, 14(3):307-322.
240. Holmfeldt P, Ganuza M, Marathe H, He B, Hall T, Kang G, Moen J, Pardieck J, Saulsberry AC, Cico A et al: Functional screen identifies regulators of murine hematopoietic stem cell repopulation. The Journal of experimental medicine 2016, 213(3):433-449.
241. Arcangeli ML, Bardin F, Frontera V, Bidaut G, Obrados E, Adams RH, Chabannon C, Aurrand-Lions M: Function of Jam-B/Jam-C interaction in homing and mobilization of human and mouse hematopoietic stem and progenitor cells. Stem Cells 2014, 32(4):1043-1054.
242. Praetor A, McBride JM, Chiu H, Rangell L, Cabote L, Lee WP, Cupp J, Danilenko DM, Fong S: Genetic deletion of JAM-C reveals a role in myeloid progenitor generation. Blood 2009, 113(9):1919-1928.
243. Arcangeli ML, Frontera V, Bardin F, Obrados E, Adams S, Chabannon C, Schiff C, Mancini SJ, Adams RH, Aurrand-Lions M: JAM-B regulates maintenance of hematopoietic stem cells in the bone marrow. Blood 2011, 118(17):4609-4619.
244. De Grandis M, Bardin F, Fauriat C, Zemmour C, El-Kaoutari A, Serge A, Granjeaud S, Pouyet L, Montersino C, Chretien AS et al: JAM-C Identifies Src Family Kinase-Activated Leukemia-Initiating Cells and Predicts Poor Prognosis in Acute Myeloid Leukemia. Cancer research 2017, 77(23):6627-6640.
245. Gasser S, Orsulic S, Brown EJ, Raulet DH: The DNA damage pathway regulates innate immune system ligands of the NKG2D receptor. Nature 2005, 436(7054):1186-1190.
246. Frazao A, Rethacker L, Messaoudene M, Avril MF, Toubert A, Dulphy N, Caignard A: NKG2D/NKG2-Ligand Pathway Offers New Opportunities in Cancer Treatment. Frontiers in immunology 2019, 10:661.
247. Paczulla AM, Rothfelder K, Raffel S, Konantz M, Steinbacher J, Wang H, Tandler C, Mbarga M, Schaefer T, Falcone M et al: Absence of NKG2D ligands defines leukaemia stem cells and mediates their immune evasion. Nature 2019, 572(7768):254-259.
248. Roshal M, Chien S, Othus M, Wood BL, Fang M, Appelbaum FR, Estey EH, Papayannopoulou T, Becker PS: The proportion of CD34(+)CD38(low or neg) myeloblasts, but not side population frequency, predicts initial response to induction therapy in patients with newly diagnosed acute myeloid leukemia. Leukemia 2013, 27(3):728-731.
249. Moshaver B, van Rhenen A, Kelder A, van der Pol M, Terwijn M, Bachas C, Westra AH, Ossenkoppele GJ, Zweegman S, Schuurhuis GJ: Identification of a small subpopulation of candidate leukemia-initiating cells in the side population of patients with acute myeloid leukemia. Stem Cells 2008, 26(12):3059-3067.
250. Moshaver B, Wouters RF, Kelder A, Ossenkoppele GJ, Westra GAH, Kwidama Z, Rutten AR, Kaspers GJL, Zweegman S, Cloos J et al: Relationship between CD34/CD38 and side population (SP) defined leukemia stem cell compartments in acute myeloid leukemia. Leukemia research 2019, 81:27-34.
251. Yang X, Yao R, Wang H: Update of ALDH as a Potential Biomarker and Therapeutic Target for AML. BioMed research international 2018, 2018:9192104.

252. Yang CK, Wang XK, Liao XW, Han CY, Yu TD, Qin W, Zhu GZ, Su H, Yu L, Liu XG et al: Aldehyde dehydrogenase 1 (ALDH1) isoform expression and potential clinical implications in hepatocellular carcinoma. PloS one 2017, 12(8):e0182208.
253. Storms RW, Trujillo AP, Springer JB, Shah L, Colvin OM, Ludeman SM, Smith C: Isolation of primitive human hematopoietic progenitors on the basis of aldehyde dehydrogenase activity. Proceedings of the National Academy of Sciences of the United States of America 1999, 96(16):9118-9123.
254. Pearce DJ, Taussig D, Simpson C, Allen K, Rohatiner AZ, Lister TA, Bonnet D: Characterization of cells with a high aldehyde dehydrogenase activity from cord blood and acute myeloid leukemia samples. Stem Cells 2005, 23(6):752-760.
255. Armstrong L, Stojkovic M, Dimmick I, Ahmad S, Stojkovic P, Hole N, Lako M: Phenotypic characterization of murine primitive hematopoietic progenitor cells isolated on basis of aldehyde dehydrogenase activity. Stem Cells 2004, 22(7):1142-1151.
256. Chang PM, Chen CH, Yeh CC, Lu HJ, Liu TT, Chen MH, Liu CY, Wu ATH, Yang MH, Tai SK et al: Transcriptome analysis and prognosis of ALDH isoforms in human cancer. Scientific reports 2018, 8(1):2713.
257. Gasparetto M, Smith CA: ALDHs in normal and malignant hematopoietic cells: Potential new avenues for treatment of AML and other blood cancers. Chemico-biological interactions 2017, 276:46-51.
258. Gerber JM, Smith BD, Ngwang B, Zhang H, Vala MS, Morsberger L, Galkin S, Collector MI, Perkins B, Levis MJ et al: A clinically relevant population of leukemic CD34(+)CD38(-) cells in acute myeloid leukemia. Blood 2012, 119(15):3571-3577.
259. Schuurhuis GJ, Meel MH, Wouters F, Min LA, Terwijn M, de Jonge NA, Kelder A, Snel AN, Zweegman S, Ossenkoppele GJ et al: Normal hematopoietic stem cells within the AML bone marrow have a distinct and higher ALDH activity level than co-existing leukemic stem cells. PloS one 2013, 8(11):e78897.
260. Gerber JM, Zeidner JF, Morse S, Blackford AL, Perkins B, Yanagisawa B, Zhang H, Morsberger L, Karp J, Ning Y et al: Association of acute myeloid leukemia's most immature phenotype with risk groups and outcomes. Haematologica 2016, 101(5):607-616.
261. Langevin F, Crossan GP, Rosado IV, Arends MJ, Patel KJ: Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. Nature 2011, 475(7354):53-58.
262. Cheung AM, Wan TS, Leung JC, Chan LY, Huang H, Kwong YL, Liang R, Leung AY: Aldehyde dehydrogenase activity in leukemic blasts defines a subgroup of acute myeloid leukemia with adverse prognosis and superior NOD/SCID engrafting potential. Leukemia 2007, 21(7):1423-1430.
263. Zhou L, Sheng D, Wang D, Ma W, Deng Q, Deng L, Liu S: Identification of cancer-type specific expression patterns for active aldehyde dehydrogenase (ALDH) isoforms in ALDEFLUOR assay. Cell biology and toxicology 2019, 35(2):161-177.
264. Ng SW, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, Arruda A, Popescu A, Gupta V, Schimmer AD et al: A 17-gene stemness score for rapid determination of risk in acute leukaemia. Nature 2016, Dec 15(540(7633)):433-437.
265. Majetl R, Becker MW, Tian Q, Lee TLM, Yan XW, Liu R, Chiang JH, Hood L, Clarke MF, Weissman IL: Dysregulated gene expression networks in human acute myelogenous leukemia stem cells. Proceedings of the National Academy of Sciences of the United States of America 2009, 106(9):3396-3401.
266. Forsberg EC, Passegue E, Prohaska SS, Wagers AJ, Koeva M, Stuart JM, Weissman IL: Molecular signatures of quiescent, mobilized and leukemia-initiating hematopoietic stem cells. PloS one 2010, 5(1):e8785.
267. Gentles AJ, Plevritis SK, Majeti R, Alizadeh AA: Association of a leukemic stem cell gene expression signature with clinical outcomes in acute myeloid leukemia. Jama 2010, 304(24):2706-2715.

268. Saito Y, Kitamura H, Hijikata A, Tomizawa-Murasawa M, Tanaka S, Takagi S, Uchida N, Suzuki N, Sone A, Najima Y et al: Identification of therapeutic targets for quiescent, chemotherapy-resistant human leukemia stem cells. Science translational medicine 2010, 2(17):17ra19.
269. de Leeuw DC, Denkers F, Olthof MC, Rutten AP, Pouwels W, Schuurhuis GJ, Ossenkoppele GJ, Smit L: Attenuation of microRNA-126 expression that drives CD34+38- stem/progenitor cells in acute myeloid leukemia leads to tumor eradication. Cancer research 2014, 74(7):2094-2105.
270. Yagi T, Morimoto A, Eguchi M, Hibi S, Sako M, Ishii E, Mizutani S, Imashuku S, Ohki M, Ichikawa H: Identification of a gene expression signature associated with pediatric AML prognosis. Blood 2003, 102(5):1849-1856.
271. Ross ME, Mahfouz R, Onciu M, Liu HC, Zhou X, Song G, Shurtleff SA, Pounds S, Cheng C, Ma J et al: Gene expression profiling of pediatric acute myelogenous leukemia. Blood 2004, 104(12):3679-3687.
272. Andersson A, Ritz C, Lindgren D, Eden P, Lassen C, Heldrup J, Olofsson T, Rade J, Fontes M, Porwit-Macdonald A et al: Microarray-based classification of a consecutive series of 121 childhood acute leukemias: prediction of leukemic and genetic subtype as well as of minimal residual disease status. Leukemia 2007, 21(6):1198-1203.
273. Duployez N, Preudhomme C, Cheok M: A 17-gene-expression profile to improve prognosis prediction in childhood acute myeloid leukemia. Oncotarget 2018, 9(74):33869-33870.
274. Duployez N, Marceau-Renaut A, Villenet C, Petit A, Rousseau A, Ng SWK, Paquet A, Gonzales F, Barthelemy A, Lepretre F et al: The stem cell-associated gene expression signature allows risk stratification in pediatric acute myeloid leukemia. Leukemia 2018, 33(2):348-357.
275. Yassin M, Aqaqe N, Yassin AA, van Galen P, Kugler E, Bernstein BE, Koren-Michowitz M, Canaani J, Nagler A, Lechman ER et al: A novel method for detecting the cellular stemness state in normal and leukemic human hematopoietic cells can predict disease outcome and drug sensitivity. Leukemia 2019, 33(8):2061-2077.
276. De Kouchkovsky I, Abdul-Hay M: 'Acute myeloid leukemia: a comprehensive review and 2016 update'. Blood cancer journal 2016, 6(7):e441.
277. Hope KJ, Jin L, Dick JE: Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity. Nat Immunol 2004, 5(7):738-743.
278. Shlush LI, Mitchell A, Heisler L, Abelson S, Ng SWK, Trotman-Grant A, Medeiros JJF, Rao-Bhatia A, Jaciw-Zurakowsky I, Marke R et al: Tracing the origins of relapse in acute myeloid leukaemia to stem cells. Nature 2017, 547(7661):104-108.
279. Zeijlemaker W, Grob T, Meijer R, Hanekamp D, Kelder A, Carbaat-Ham JC, Oussoren-Brockhoff YJM, Snel AN, Veldhuizen D, Scholten WJ et al: CD34(+)CD38(-) leukemic stem cell frequency to predict outcome in acute myeloid leukemia. Leukemia 2018, 33:1102–1112.
280. Witte KE, Ahlers J, Schafer I, Andre M, Kerst G, Scheel-Walter HG, Schwarze CP, Pfeiffer M, Lang P, Handgretinger R et al: High Proportion of Leukemic Stem Cells at Diagnosis Is Correlated with Unfavorable Prognosis in Childhood Acute Myeloid Leukemia. Pediatr Hemat Oncol 2011, 28(2):91-99.
281. Garg S, Ghosh K, Madkaikar M: Antigen expression on a putative leukemic stem cell population and AML blast. International journal of hematology 2016, 103(5):567-571.
282. Ran D, Schubert M, Taubert I, Eckstein V, Bellos F, Jauch A, Chen H, Bruckner T, Saffrich R, Wuchter P et al: Heterogeneity of leukemia stem cell candidates at diagnosis of acute myeloid leukemia and their clinical significance. Experimental hematology 2012, 40(2):155-165 e151.
283. Lagadinou ED, Sach A, Callahan K, Rossi RM, Neering SJ, Minhajuddin M, Ashton JM, Pei S, Grose V, O'Dwyer KM et al: BCL-2 inhibition targets oxidative phosphorylation and selectively eradicates quiescent human leukemia stem cells. Cell stem cell 2013, 12(3):329-341.
284. Ito K, Suda T: Metabolic requirements for the maintenance of self-renewing stem cells. Nature reviews Molecular cell biology 2014, 15(4):243-256.
285. Ye H, Adane B, Khan N, Sullivan T, Minhajuddin M, Gasparetto M, Stevens B, Pei S, Balys M, Ashton JM et al: Leukemic Stem Cells Evade Chemotherapy by Metabolic Adaptation to an Adipose Tissue Niche. Cell stem cell 2016, 7(19(1)):23-37.

286. Pei S, Minhajuddin M, Callahan KP, Balys M, Ashton JM, Neering SJ, Lagadinou ED, Corbett C, Ye H, Liesveld JL et al: Targeting aberrant glutathione metabolism to eradicate human acute myelogenous leukemia cells. The Journal of biological chemistry 2013, 288(47):33542-33558.
287. Heidel FH, Arreba-Tutusaus P, Armstrong SA, Fischer T: Evolutionarily conserved signaling pathways: acting in the shadows of acute myelogenous leukemia's genetic diversity. Clinical cancer research : an official journal of the American Association for Cancer Research 2015, 21(2):240-248.
288. Guzman ML, Neering SJ, Upchurch D, Grimes B, Howard DS, Rizzieri DA, Luger SM, Jordan CT: Nuclear factor-kappa B is constitutively activated in primitive human acute myelogenous leukemia cells. Blood 2001, 98(8):2301-2307.
289. Xu DD, Wang Y, Zhou PJ, Qin SR, Zhang R, Zhang Y, Xue X, Wang J, Wang X, Chen HC et al: The IGF2/IGF1R/Nanog Signaling Pathway Regulates the Proliferation of Acute Myeloid Leukemia Stem Cells. Frontiers in pharmacology 2018, 9:687.
290. Hu Y, Li S: Survival regulation of leukemia stem cells. Cellular and molecular life sciences : CMLS 2015.
291. Pession A, Lonetti A, Bertuccio S, Locatelli F, Masetti R: Targeting Hedgehog pathway in pediatric acute myeloid leukemia: challenges and opportunities. Expert opinion on therapeutic targets 2019, 23(2):87-91.
292. Garg S, Shanmukhaiah C, Marathe S, Mishra P, Babu Rao V, Ghosh K, Madkaikar M: Differential antigen expression and aberrant signaling via PI3/AKT, MAP/ERK, JAK/STAT, and Wnt/beta catenin pathways in Lin-/CD38-/CD34+ cells in acute myeloid leukemia. European journal of haematology 2016, 96(3):309-317.
293. Park S, Chapuis N, Tamburini J, Bardet V, Cornillet-Lefebvre P, Willems L, Green A, Mayeux P, Lacombe C, Bouscary D: Role of the PI3K/AKT and mTOR signaling pathways in acute myeloid leukemia. Haematologica 2010, 95(5):819-828.
294. Boyd AL, Campbell CJV, Hopkins CI, Fiebig-Comyn A, Russell J, Ulemek J, Foley R, Leber B, Xenocostas A, Collins TJ et al: Niche displacement of human leukemic stem cells uniquely allows their competitive replacement with healthy HSPCs. Journal of Experimental Medicine 2014, 211(10):1925-1935.
295. Hatfield KJ, Bedringsaas SL, Ryningen A, Gjertsen BT, Bruserud O: Hypoxia increases HIF-1alpha expression and constitutive cytokine release by primary human acute myeloid leukaemia cells. European cytokine network 2010, 21(3):154-164.
296. Basak P, Chatterjee S, Das M, Das P, Pereira JA, Dutta RK, Chaklader M, Chaudhuri S, Law S: Phenotypic Alteration of Bone Marrow HSC and Microenvironmental Association in Experimentally Induced Leukemia. Curr Stem Cell Res T 2010, 5(4):379-386.
297. Frisch BJ, Ashton JM, Xing L, Becker MW, Jordan CT, Calvi LM: Functional inhibition of osteoblastic cells in an in vivo mouse model of myeloid leukemia. Blood 2012, 119(2):540-550.
298. Bowers M, Zhang B, Ho Y, Agarwal P, Chen CC, Bhatia R: Osteoblast ablation reduces normal long-term hematopoietic stem cell self-renewal but accelerates leukemia development. Blood 2015, 125(17):2678-2688.
299. Kode A, Manavalan JS, Mosialou I, Bhagat G, Rathinam CV, Luo N, Khiabanian H, Lee A, Murty VV, Friedman R et al: Leukaemogenesis induced by an activating beta-catenin mutation in osteoblasts. Nature 2014, 506(7487):240-244.
300. Bruserud O, Ryningen A, Wergeland L, Glenjen NI, Gjertsen BT: Osteoblasts increase proliferation and release of pro-angiogenic interleukin 8 by native human acute myelogenous leukemia blasts. Haematologica 2004, 89(4):391-402.
301. Bruserud O, Glenjen N, Ryningen A: Effects of angiogenic regulators on in vitro proliferation and cytokine secretion by native human acute myelogenous leukemia blasts. European journal of haematology 2003, 71(1):9-17.
302. Annesley CE, Brown P: The Biology and Targeting of FLT3 in Pediatric Leukemia. Frontiers in oncology 2014, 4:263.

303. Levis M: FLT3 mutations in acute myeloid leukemia: what is the best approach in 2013? Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program 2013, 2013:220-226.
304. Yu MG, Zheng HY: Acute Myeloid Leukemia: Advancements in Diagnosis and Treatment. Chinese medical journal 2017, 130(2):211-218.
305. Cuesta-Mateos C, Alcaraz-Serna A, Somovilla-Crespo B, Munoz-Calleja C: Monoclonal Antibody Therapies for Hematological Malignancies: Not Just Lineage-Specific Targets. Frontiers in immunology 2017, 8:1936.
306. Tiller KE, Tessier PM: Advances in Antibody Design. Annual review of biomedical engineering 2015, 17:191-216.
307. Baeuerle PA, Reinhardt C: Bispecific T-cell engaging antibodies for cancer therapy. Cancer research 2009, 69(12):4941-4944.
308. Johnson S, Burke S, Huang L, Gorlatov S, Li H, Wang W, Zhang W, Tuaillon N, Rainey J, Barat B et al: Effector cell recruitment with novel Fv-based dual-affinity re-targeting protein leads to potent tumor cytolysis and in vivo B-cell depletion. Journal of molecular biology 2010, 399(3):436-449.
309. Sadelain M, Riviere I, Brentjens R: Targeting tumours with genetically enhanced T lymphocytes. Nature reviews Cancer 2003, 3(1):35-45.
310. Themeli M, Riviere I, Sadelain M: New cell sources for T cell engineering and adoptive immunotherapy. Cell stem cell 2015, 16(4):357-366.
311. Trounson A: Potential Pitfall of Pluripotent Stem Cells. The New England journal of medicine 2017, 377(5):490-491.
312. Maher J, Davies ET: Targeting cytotoxic T lymphocytes for cancer immunotherapy. British journal of cancer 2004, 91(5):817-821.
313. Carpenito C, Milone MC, Hassan R, Simonet JC, Lakhal M, Suhoski MM, Varela-Rohena A, Haines KM, Heitjan DF, Albelda SM et al: Control of large, established tumor xenografts with genetically retargeted human T cells containing CD28 and CD137 domains. Proceedings of the National Academy of Sciences of the United States of America 2009, 106(9):3360-3365.
314. Cartellieri M, Feldmann A, Koristka S, Arndt C, Loff S, Ehninger A, von Bonin M, Bejestani EP, Ehninger G, Bachmann MP: Switching CAR T cells on and off: a novel modular platform for retargeting of T cells to AML blasts. Blood cancer journal 2016, 6(8):e458.
315. Magee MS, Snook AE: Challenges to Chimeric Antigen Receptor (CAR)-T Cell Therapy for Cancer. Discovery medicine 2014, 100:265-271.
316. Rafiq S, Purdon TJ, Daniyan AF, Koneru M, Dao T, Liu C, Scheinberg DA, Brentjens RJ: Optimized T-cell receptor-mimic chimeric antigen receptor T cells directed toward the intracellular Wilms Tumor 1 antigen. Leukemia 2017, 31(8):1788-1797.
317. Perna F, Berman SH, Soni RK, Mansilla-Soto J, Eyquem J, Hamieh M, Hendrickson RC, Brennan CW, Sadelain M: Integrating Proteomics and Transcriptomics for Systematic Combinatorial Chimeric Antigen Receptor Therapy of AML. Cancer cell 2017, 32(4):506-519 e505.
318. Haubner S, Perna F, Kohnke T, Schmidt C, Berman S, Augsberger C, Schnorfeil FM, Krupka C, Lichtenegger FS, Liu X et al: Coexpression profile of leukemic stem cell markers for combinatorial targeted therapy in AML. Leukemia 2019, 33(1):64-74.
319. He SZ, Busfield S, Ritchie DS, Hertzberg MS, Durrant S, Lewis ID, Marlton P, McLachlan AJ, Kerridge I, Bradstock KF et al: A Phase 1 study of the safety, pharmacokinetics and anti-leukemic activity of the anti-CD123 monoclonal antibody CSL360 in relapsed, refractory or high-risk acute myeloid leukemia. Leukemia & lymphoma 2015, 56(5):1406-1415.
320. Mardiros A, Dos Santos C, McDonald T, Brown CE, Wang X, Budde LE, Hoffman L, Aguilar B, Chang WC, Bretzlaff W et al: T cells expressing CD123-specific chimeric antigen receptors exhibit specific cytolytic effector functions and antitumor effects against human acute myeloid leukemia. Blood 2013, 122(18):3138-3148.
321. Jen EY, Ko CW, Lee JE, Del Valle PL, Aydanian A, Jewell C, Norsworthy KJ, Przepiorka D, Nie L, Liu J et al: FDA Approval: Gemtuzumab Ozogamicin for the Treatment of Adults with Newly

Diagnosed CD33-Positive Acute Myeloid Leukemia. Clinical cancer research : an official journal of the American Association for Cancer Research 2018, 24(14):3242-3246.

322. Niktoreh N, Lerius B, Zimmermann M, Gruhn B, Escherich G, Bourquin JP, Dworzak M, Sramkova L, Rossig C, Creutzig U et al: Gemtuzumab ozogamicin in children with relapsed or refractory acute myeloid leukemia: a report by Berlin-Frankfurt-Munster study group. Haematologica 2019, 104(1):120-127.
323. Parigger J, Zwaan CM, Reinhardt D, Kaspers GJ: Dose-related efficacy and toxicity of gemtuzumab ozogamicin in pediatric acute myeloid leukemia. Expert review of anticancer therapy 2016, 16(2):137-146.
324. Gamis AS, Alonzo TA, Meshinchi S, Sung L, Gerbing RB, Raimondi SC, Hirsch BA, Kahwash SB, Heerema-McKenney A, Winter L et al: Gemtuzumab ozogamicin in children and adolescents with de novo acute myeloid leukemia improves event-free survival by reducing relapse risk: results from the randomized phase III Children's Oncology Group trial AAML0531. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2014, 32(27):3021-3032.
325. Stein EM, Walter RB, Erba HP, Fathi AT, Advani AS, Lancet JE, Ravandi F, Kovacsovics T, DeAngelo DJ, Bixby D et al: A phase 1 trial of vadastuximab talirine as monotherapy in patients with CD33-positive acute myeloid leukemia. Blood 2018, 131(4):387-396.
326. Krupka C, Kufer P, Kischel R, Zugmaier G, Bogeholz J, Kohnke T, Lichtenegger FS, Schneider S, Metzeler KH, Fiegl M et al: CD33 target validation and sustained depletion of AML blasts in long-term cultures by the bispecific T-cell-engaging antibody AMG 330. Blood 2014, 123(3):356-365.
327. Krupka C, Kufer P, Kischel R, Zugmaier G, Lichtenegger FS, Kohnke T, Vick B, Jeremias I, Metzeler KH, Altmann T et al: Blockade of the PD-1/PD-L1 axis augments lysis of AML cells by the CD33/CD3 BiTE antibody construct AMG 330: reversing a T-cell-induced immune escape mechanism. Leukemia 2016, 30(2):484-491.
328. Raza A, Jurcic JG, Roboz GJ, Maris M, Stephenson JJ, Wood BL, Feldman EJ, Galili N, Grove LE, Drachman JG et al: Complete remissions observed in acute myeloid leukemia following prolonged exposure to lintuzumab: a phase 1 trial. Leukemia & lymphoma 2009, 50(8):1336-1344.
329. Rosenblat TL, McDevitt MR, Mulford DA, Pandit-Taskar N, Divgi CR, Panageas KS, Heaney ML, Chanel S, Morgenstern A, Sgouros G et al: Sequential cytarabine and alpha-particle immunotherapy with bismuth-213-lintuzumab (HuM195) for acute myeloid leukemia. Clinical cancer research : an official journal of the American Association for Cancer Research 2010, 16(21):5303-5311.
330. Petrov JC, Wada M, Pinz KG, Yan LE, Chen KH, Shuai X, Liu H, Chen X, Leung LH, Salman H et al: Compound CAR T-cells as a double-pronged approach for treating acute myeloid leukemia. Leukemia 2018, 32:1317–1326.
331. Majeti R, Chao MP, Alizadeh AA, Pang WW, Jaiswal S, Gibbs KD, van Rooijen N, Weissman IL: CD47 Is an Adverse Prognostic Factor and Therapeutic Antibody Target on Human Acute Myeloid Leukemia Stem Cells. Cell 2009, 138(2):286-299.
332. Theocharides AP, Jin L, Cheng PY, Prasolava TK, Malko AV, Ho JM, Poeppl AG, van Rooijen N, Minden MD, Danska JS et al: Disruption of SIRPalpha signaling in macrophages eliminates human acute myeloid leukemia stem cells in xenografts. The Journal of experimental medicine 2012, 209(10):1883-1899.
333. Zhao X, Singh S, Pardoux C, Zhao J, Hsi ED, Abo A, Korver W: Targeting C-type lectin-like molecule-1 for antibody-mediated immunotherapy in acute myeloid leukemia. Haematologica 2010, 95(1):71-78.
334. Leong SR, Sukumaran S, Hristopoulos M, Totpal K, Stainton S, Lu E, Wong A, Tam L, Newman R, Vuillemenot BR et al: An anti-CD3/anti-CLL-1 bispecific antibody for the treatment of acute myeloid leukemia. Blood 2017, 129(5):609-618.

335. Zhang H, Luo J, Li Y, Henderson PT, Wang Y, Wachsmann-Hogiu S, Zhao W, Lam KS, Pan CX: Characterization of high-affinity peptides and their feasibility for use in nanotherapeutics targeting leukemia stem cells. Nanomedicine : nanotechnology, biology, and medicine 2012, 8(7):1116-1124.
336. Wang J, Chen S, Xiao W, Li W, Wang L, Yang S, Wang W, Xu L, Liao S, Liu W et al: CAR-T cells targeting CLL-1 as an approach to treat acute myeloid leukemia. Journal of hematology & oncology 2018, 11(1):7.
337. Kikushige Y, Miyamoto T: TIM-3 as a novel therapeutic target for eradicating acute myelogenous leukemia stem cells. International journal of hematology 2013, 98(6):627-633.
338. Naik J, Themeli M, de Jong-Korlaar R, Ruiter RWJ, Poddighe PJ, Yuan H, de Bruijn JD, Ossenkoppele GJ, Zweegman S, Smit L et al: CD38 as a therapeutic target for adult acute myeloid leukemia and T-cell acute lymphoblastic leukemia. Haematologica 2019, 104(3):e100-e103.
339. Etxabe A, Lara-Castillo MC, Cornet-Masana JM, Banus-Mulet A, Nomdedeu M, Torrente MA, Pratcorona M, Diaz-Beya M, Esteve J, Risueno RM: Inhibition of serotonin receptor type 1 in acute myeloid leukemia impairs leukemia stem-cell functionality: A promising novel therapeutic target. Leukemia 2017, 31:2288–2302.
340. Depreter B, Weening KE, Vandepoele K, Essand M, De Moerloose B, Themeli M, Cloos J, Hanekamp D, Moors I, I DH et al: TARP is an immunotherapeutic target in acute myeloid leukemia expressed in the leukemic stem cell compartment. Haematologica 2019, 104:xxx(doi: 10.3324/haematol.2019.222612. [Epub ahead of print]).
341. Fujita S, Honma D, Adachi N, Araki K, Takamatsu E, Katsumoto T, Yamagata K, Akashi K, Aoyama K, Iwama A et al: Dual inhibition of EZH1/2 breaks the quiescence of leukemia stem cells in acute myeloid leukemia. Leukemia 2018, 32(4):855-864.
342. Abraham SA, Hopcroft LE, Carrick E, Drotar ME, Dunn K, Williamson AJ, Korfi K, Baquero P, Park LE, Scott MT et al: Dual targeting of p53 and c-MYC selectively eliminates leukaemic stem cells. Nature 2016, 534(7607):341-346.
343. Lynn RC, Feng Y, Schutsky K, Poussin M, Kalota A, Dimitrov DS, Powell DJ, Jr.: High-affinity FRbeta-specific CAR T cells eradicate AML and normal myeloid lineage without HSC toxicity. Leukemia 2016, 30(6):1355-1364.
344. Laverdiere I, Boileau M, Neumann AL, Frison H, Mitchell A, Ng SWK, Wang JCY, Minden MD, Eppert K: Leukemic stem cell signatures identify novel therapeutics targeting acute myeloid leukemia. Blood cancer journal 2018, 8(6):52.
345. Hartwell KA, Miller PG, Mukherjee S, Kahn AR, Stewart AL, Logan DJ, Negri JM, Duvet M, Jaras M, Puram R et al: Niche-based screening identifies small-molecule inhibitors of leukemia stem cells. Nat Chem Biol 2013, 9(12):840-848.
346. Zhang H, Mi JQ, Fang H, Wang Z, Wang C, Wu L, Zhang B, Minden M, Yang WT, Wang HW et al: Preferential eradication of acute myelogenous leukemia stem cells by fenretinide. Proceedings of the National Academy of Sciences of the United States of America 2013, 110(14):5606-5611.
347. Zhang Y, Chen HX, Zhou SY, Wang SX, Zheng K, Xu DD, Liu YT, Wang XY, Wang X, Yan HZ et al: Sp1 and c-Myc modulate drug resistance of leukemia stem cells by regulating survivin expression through the ERK-MSK MAPK signaling pathway. Molecular cancer 2015, 14:56.
348. Zhang Y, Zhou SY, Yan HZ, Xu DD, Chen HX, Wang XY, Wang X, Liu YT, Zhang L, Wang S et al: miR-203 inhibits proliferation and self-renewal of leukemia stem cells by targeting survivin and Bmi-1. Scientific reports 2016, 6:19995.
349. Martianez Canales T, de Leeuw DC, Vermue E, Ossenkoppele GJ, Smit L: Specific Depletion of Leukemic Stem Cells: Can MicroRNAs Make the Difference? Cancers 2017, 9(7).
350. Velu CS, Chaubey A, Phelan JD, Horman SR, Wunderlich M, Guzman ML, Jegga AG, Zeleznik-Le NJ, Chen J, Mulloy JC et al: Therapeutic antagonists of microRNAs deplete leukemia-initiating cell activity. The Journal of clinical investigation 2014, 124(1):222-236.

351. Meyer SE, Muench DE, Rogers AM, Newkold TJ, Orr E, O'Brien E, Perentesis JP, Doench JG, Lal A, Morris PJ et al: miR-196b target screen reveals mechanisms maintaining leukemia stemness with therapeutic potential. The Journal of experimental medicine 2018, 215(8):2115-2136.
352. Si X, Zhang X, Hao X, Li Y, Chen Z, Ding Y, Shi H, Bai J, Gao Y, Cheng T et al: Upregulation of miR-99a is associated with poor prognosis of acute myeloid leukemia and promotes myeloid leukemia cell expansion. Oncotarget 2016, 22(7(47)):78095–78109.
353. Bill M, Papaioannou D, Karunasiri M, Kohlschmidt J, Pepe F, Walker CJ, Walker AE, Brannan Z, Pathmanathan A, Zhang X et al: Expression and functional relevance of long non-coding RNAs in acute myeloid leukemia stem cells. Leukemia 2019, 33(9):2169-2182.
354. Butturini A, Santucci MA, Gale RP, Perocco P, Tura S: GM-CSF incubation prior to treatment with cytarabine or doxorubicin enhances drug activity against AML cells in vitro: a model for leukemia chemotherapy. Leukemia research 1990, 14(9):743-749.
355. Feng X, Lan H, Ruan Y, Li C: Impact on acute myeloid leukemia relapse in granulocyte colony-stimulating factor application: a meta-analysis. Hematology 2018, 23(9):581-589.
356. Guzman ML, Rossi RM, Karnischky L, Li X, Peterson DR, Howard DS, Jordan CT: The sesquiterpene lactone parthenolide induces apoptosis of human acute myelogenous leukemia stem and progenitor cells. Blood 2005, 105(11):4163-4169.
357. Tavor S, Petit I, Porozov S, Avigdor A, Dar A, Leider-Trejo L, Shemtov N, Deutsch V, Naparstek E, Nagler A et al: CXCR4 regulates migration and development of human acute myelogenous leukemia stem cells in transplanted NOD/SCID mice. Cancer research 2004, 64(8):2817-2824.
358. Zhang H, Fang H, Wang K: Reactive oxygen species in eradicating acute myeloid leukemic stem cells. Stem cell investigation 2014, 1:13.
359. Kuo YH, Bhatia R: Pushing the limits: defeating leukemia stem cells by depleting telomerase. Cell stem cell 2014, 15(6):673-675.

2

CHAPTER II: Research objectives.

The topic of this research is the oncopathogenesis of AML in children. Most pediatric AML (pedAML) patients will achieve excellent clinical remission rates after the first induction course, nowadays over >90%, but this achievement is unfortunately not consistently translated into high cure rates. Still 30-40% of the good responders relapse. During the past decade, ample evidence in adult AML showed that relapse is promoted by the persistence of leukemic stem cells (LSC). This cell fraction has an unlimited self-renewal capacity and is capable to clonally expand and propagate leukemia at any unexpected moment. Within this project, we focussed on the role and molecular and flow cytometric characteristics of LSC in pedAML. This way, we aimed to better understand the development and the molecular biology of AML in children and the emergence of relapses. Overall, we aimed at increasing the knowledge on the concepts of LSC biology and to deliver specific markers to predict relapse, guide therapy and potentially deliver novel drug targets. This dissertation is built on four major goals.

First, we performed an in-depth evaluation of the immunophenotype of the LSC fraction in pedAML. During the study period, diagnostic and relapse samples from children with AML in Belgium were selected for immunophenotyping of the LSC compartment based on sample availability and CD34 expression. To this extent, backbone markers CD38, CD34, CD45 were combined with a set of markers formerly investigated in pedAML (CD123, CD7, CD56, CD2 and CD96) in addition to markers prior to this research only explored in an adult AML setting (CD22, CD11b, TIM-3, CLL-1, CD15, NG2, CD49d and GPR56). Next to these descriptive analyses, we aimed to evaluate the prognostic impact of the LSC load at diagnosis. The result of this research is documented in **chapter III**.

Second, we explored the molecular heterogeneity of pedAML leukemic subpopulations. The currently available LSC signatures present some drawbacks, as the (background) expression in HSC and leukemic blasts (L-blast) is not or only partially explored, and they mostly lack the inclusion of downregulated targets. The information gathered in our first research question allowed us to set an immunophenotypic strategy for sorting LSC and L-blast populations from the same cohort of consecutively collected patients. A molecular characterization of these sorted subpopulations was conducted to gather more insights in the pedAML stem cell biology, following a multi-tiered approach. Micro-array analysis was performed on LSC and leukemic blasts sorted from a limited pedAML patient cohort and profiled together with normal counterparts sorted from healthy controls. In addition, we re-analyzed a publicly available micro-array dataset containing expression profiles of LSCs and HSCs from adult AML and control patients, respectively. By using a cancer vs normal approach, we focussed on the coding and non-coding transcript differences with the patients' leukemic cell populations versus their normal counterparts. Differential expression analysis and subsequent qPCR validation of the coding transcripts was performed in order to identify a novel set of overexpressed and downregulated genes in leukemic stem cells and in leukemic blasts. These data revealed targets for further functional analyses and targeted therapy. By performing differential analysis in regard to the non-coding targets, we aimed to discover novel lncRNAs in the enigmatic LSC fraction from pedAML patient samples. The results of this research are addressed in **chapters IV and V**.

Third, we explored novel assays, which should be applicable in a routine clinical diagnostic setting, for biomarker detection in heterogeneous cell populations. Most gene expression data in pediatric malignancies available through literature represent bulk population averages. Tumours are complex biological systems constituting of a heterogeneous mix of cancer and normal cells. Hence, cell population averaging might obscure prognostic and therapeutically relevant information hidden in low-frequent subpopulations. Flow cytometry has become a powerful approach to dissect heterogeneous cell populations based on the phenotypic characteristics of single-cells. Unfortunately,

directly conjugated antibodies are not available for RNA expression evaluation. This limitation has driven the development of molecular techniques enabling single-cell expression analysis, and single-cell isolation methods such as cell-sorting. However, the significant cell and subsequent RNA loss in cell sorted-populations represents a major challenge. As sample availability often poses an issue in pediatric malignancies, techniques other than cell-sorting that allow single-cell RNA expression analysis are of high interest. We evaluated whether the PrimeFlow^{TM} RNA assay could serve as an innovative toolbox for target detection and disease monitoring in pediatric malignancies. This assay allows multiplex gene expression analysis of high and low abundant (non-)coding mRNAs at a single cell level, i.e. small numbers of LSCs. We aimed to perform a technical validation, followed by the evaluation of its applicability in detecting key target mRNAs in AML subpopulations. The combined identification of cellular subpopulations through flow cytometry with mRNA expression investigation would theoretically decrease the analytical turn-around-time compared to a classic two-step approach of FACS followed by genomic technologies. The results of our findings are discussed in **chapter VI**.

Fourth, we searched for immunotherapeutic targets in pedAML. The success of immunotherapy mainly relies on the choice of the targetable leukemia-associated antigen. Clinical applicability largely depends on the on-target/off-tumor effects. LSC-associated MHC-presented antigens that are absent in hematopoietic stem and progenitor populations are of high interest. Preliminary data were available, at the start of this doctoral dissertation, on the specific LSC expression of *TARP*, while negative in HSC. *TARP* was previously unexplored within the field of AML, but proven to harbour prognostic relevance in androgen-dependent prostate and breast carcinoma. We first assessed *TARP* expression in the leukemic blasts and LSCs of various leukemic cell lines and adult and pediatric patient samples. *TARP* expression in HSC and normal myeloblasts from healthy controls were evaluated as a control population. The research group of Uppsala (Sweden) had developed a TCR directed against the HLA-A*0201-restricted synthetic peptide TARP(P5L)$_{4–13}$, which specifically killed HLA-A*0201+ prostate and breast cancer cells. Through collaboration, we used this plasmid to generate lenti- and retroviral transgenic TARP-TCR CTLs. These engineered CTLs were subsequently *in vitro* evaluated for cytotoxicity toward AML cell lines and patients leukemic cells. Based on these promising results, we subsequently performed a more in-depth evaluation of *TARP* in a larger pedAML patient cohort. The result of this research is discussed in **chapter VII**.

In **chapter VIII**, we discuss future perspectives regarding our research results.

3

CHAPTER III: Results:

Exploring the immunophenotype of LSC in pediatric AML.

Manuscript in preparation.

Abstract

Introduction.
Despite excellent clinical remission rates, pediatric acute myeloid leukemia (pedAML) patients experience high relapse rates. Relapse is thought to arise from therapy-resistant leukemic stem cells (LSC). Hitherto, flow cytometric (FCM) data on LSC in pedAML is limited to five markers.

Materials and Methods.
Immunophenotyping of $CD34^{+}/CD38^{-/dim}$ cells was performed for 35 CD34+ pedAML patients, including 32 diagnostic and nine relapsed patients, with six diagnosis-relapse couples. For each patient and sample matrix (bone marrow (BM) and/or peripheral blood (PB)), backbone markers CD34, CD38 and CD45 were combined with at least one of the 11 LSC-specific leukemia-associated immunophenotype (LAIP) markers CD45RA, CD56, CD96, CD7, CD22, CD11b, TIM3, CD123, CLL-1, CD2, CD15 and NG2, and/or two of the LSC-associated markers CD49d and GPR56. We calculated the *total* LSC load, i.e. $CD34^{+}/CD38^{-/dim}$, and the *LAIP+* LSC load, taking into account $CD34^{+}/CD38^{-/dim}$ cells that harbor LAIP markers. Samples from 24 healthy subjects were used to evaluate expression in normal hematopoietic stem cells (HSC). FCM analysis was performed on a FACSCanto II flow cytometer (BD Biosciences) with instrument set-up according to EuroFlow. Gating was performed in Infinicyt (Cytognos).

Results.
Combining multiple LAIP markers is a necessity to detect genuine LSC, as roughly one third of the *total* $CD34^{+}/CD38^{-/dim}$ compartment lacks aberrant LAIP and/or secondary markers. The *total* $CD34^{+}CD38^{-/dim}$ cells was significantly lower at relapse compared to diagnosis, while *LAIP+* $CD34^{+}CD38^{-/dim}$ cells did not significantly differ. *Total* and *LAIP+* $CD34^{+}/CD38^{-/dim}$ LSC percentages, LAIP positivity or MFI values did not significantly differ between BM and PB. LSC^{high} patients, defined as having LAIP+ LSC loads ≥4.78% of the white blood cells (WBCs) or ≥17.39% of the CD34+ cells, tended to present higher WBC counts and more *WT1* overexpression, *FLT3*-ITD mutations, abnormal karyotypes and non-CBF translocations. Within patients for whom LSC load was determined and clinical outcome data were available (NOPHO-DBH AML2012 study (n=18), DB-AML01 study (n=4) or treated otherwise (n=1), a trend towards a lower event-free survival (EFS) was observed for LSC^{high} patients. CLL-1 and CD45RA were the most frequent and strongest expressed LAIP markers, whereas CD7, TIM3 and CD15 were positive in half of the patients, and CD56 and CD22 in one third of the patients.

Discussion and conclusion.
Approximately two third of the $CD34^{+}/CD38^{-/dim}$ compartment harbours aberrant expression. LSC phenotypic profiles appear to be quite similar in a pediatric and adult setting, and show a high intra- and inter-patient heterogeneity. PB may be an advantageous sample matrix to determine the LSC load at diagnosis. More patients are needed to evaluate the prognostic impact. These data pave the way for LSC-targeted therapies in pedAML, and show that future strategies will benefit from combinatorial approaches.

Introduction

Pediatric acute myeloid leukemia (pedAML) accounts for 20% of all leukemias in children [1-3]. Patients exhibit a high risk of relapse despite good clinical remission rates [4]. This relapse rate is thought to arise from a therapy-resistant cell fraction with unlimited self-renewal capacities, denominated as leukemic stem cells (LSC) [5-10]. The $CD34^{+}/CD38^{-/dim}$ compartment was shown to be most LSC-enriched in CD34+ AML [11]. In adult AML, a high LSC load at diagnosis is a significant adverse prognostic factor in regard to overall survival (OS), event-free survival (EFS) and therapy responsiveness [12-16]. Only two studies have addressed the prognostic impact of the LSC load in pedAML [17, 18]. Both studies proposed cut-offs to define high LSC frequencies and showed a significant impact of the diagnostic LSC load on the occurrence of relapse. However, still a considerable number of LSC^{low} patients experience relapse (30-39%) [17, 18]. Hitherto, FCM characterization of LSC in pedAML has only been performed by five markers [18, 19].

A more detailed phenotypic characterization using flow cytometry (FCM) could aid in identifying LSC key characteristics and their heterogeneity and allow better discrimination from hematopoietic stem cells (HSC). In adult AML, a plethora of surface proteins have been described, referred to as 'leukemia-associated immunophenotype' (LAIP) markers. Also secondary parameters, i.e. cell size (forward scatter, FSC) and cell granularity (sideward scatter, SSC), and differential backbone marker expression (CD34, CD45), are of value to distinguish LSC from HSC [12, 20, 21]. Combining 13 LAIP markers with scatter abnormalities in adult AML increased the detection of aberrant $CD34^{+}/CD38^{-/dim}$ LSC from 70-75% [22] to 87% [23].

However, still 13% of the LSC population remains undetectable with this approach, fuelling the need for identifying novel markers. In adult AML, GPR56+ subpopulations were shown to have high engrafting capacities [24], regardless of the CD34/CD38 phenotype, and correlated with a LSC signature and detrimental outcome [25, 26]. The cell-surface receptor VLA-4 is an integrin molecule, composed of α4 (CD49d) and β1 (CD29) chain, and involved in niche homing [27-30].

We here present a detailed FCM characterization of $CD34^{+}/CD38^{-}$ and $CD34^{+}/CD38^{dim}$ compartments in 35 CD34+ pedAML patients at diagnosis (n=32) and relapse (n=9) using a different-to-normal (DfN) approach. The expression of 11 LSC-specific LAIP markers was investigated, next to two LSC-associated markers CD49d and GPR56, with the vast majority previously unreported in a pediatric setting. Expression was additionally evaluated in healthy controls (n=24). We evaluated whether molecular characteristics and clinical outcome differed between CD34+ pedAML patients with high versus low LSC loads. In addition, we investigated whether the sample matrix is an important factor, as sampling peripheral blood (PB) is less invasive than obtaining bone marrow (BM). Finally, we explored associations between each of the LAIP markers with scatter properties and molecular aberrancies.

Materials en Methods

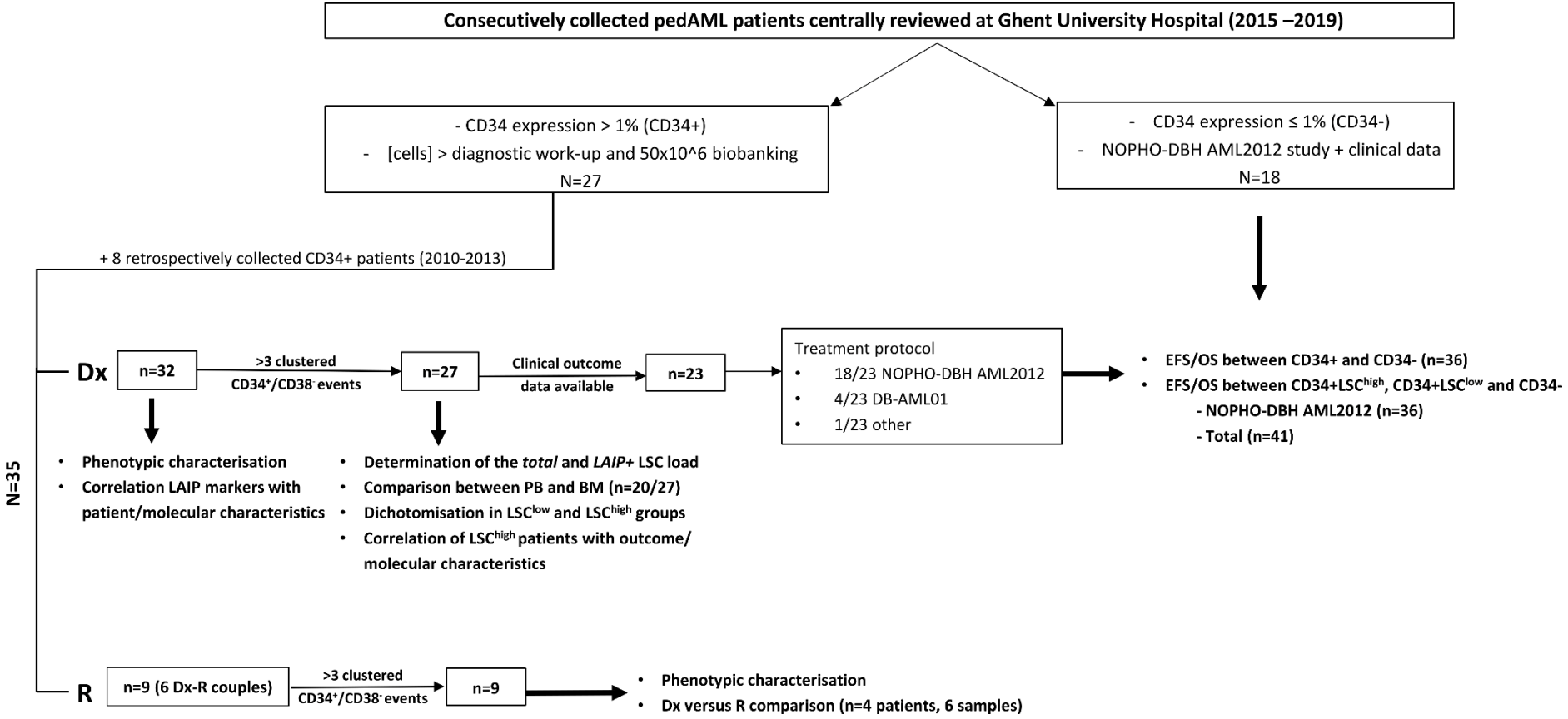
Patients and controls

Diagnostic (Dx) and relapse (R) samples from children with AML in Belgium were centrally reviewed at the University Hospital of Ghent during the study period (2015-2019) (Fig. 1). The total number of cells in PB and BM after diagnostic work-up needed to exceed $50x10^{6}$ for patients to be included in the study. Samples were processed within 48 h upon arrival and stored at room temperature. This consecutively collected patient cohort was complemented with cryopreserved samples from eight patients diagnosed between 2010-2013. Overall, the total cohort consisted out of 35 CD34+ pedAML

patients. Diagnostic CD34+ pedAML patients (n=32, Table 1) were included in the NOPHO-DBH AML2012 study (n=18/32), DB-AML01 study (n=8/32) or treated otherwise (n=6/32: 4/6 secondary AML, one APL patient and one patient treated in an adult protocol (17 years (yrs.)). Paired relapse samples were available for 6/32 patients, next to three patients only evaluated at relapse (Table 2). Clinical outcome data were additionally available for 18/19 *de novo* CD34- pedAML patients treated in the NOPHO-DBH AML2012 study (Table S1). Details on treatment protocols, definitions for EFS and OS, statistical assays and data processing can be found in a data supplement.

Samples from 24 healthy subjects were prospectively collected. Normal bone marrow (NBM, n=11) was sampled from posterior iliac crest of pediatric patients (4-18 yrs.) undergoing scoliosis surgery. Cord blood (CB, n=13) was obtained after vaginal deliveries at full term deliveries. Approval was issued by the ethical committee in accordance with the declaration of Helsinki (EC2015-1443 and EC2019-0294). All healthy subjects and patients and/or their guardians gave informed consent.

Fig. 1. Flowchart of the number of patients used for the different analyses.



Cell concentrations in PB and BM were measured by an automatic cell counter (Sysmex XP300). A higher number of diagnostics patients could be included for phenotypical characterisation (n=32) compared to LSC load determinations (n=27), as this latter analysis required a $CD34^+/CD38^-$ compartment >3 clustered events. Results of the different analyses are shown in the manuscript. The number of patients evaluated per marker are shown in Table S10. pedAML, pediatric AML; LSC, leukemic stem cell; LAIP, leukemia-associated immunophenotype; Dx, diagnosis, R, relapse; n, number; PB, peripheral blood; BM, bone marrow.

Table 1. Demographics of CD34+ pedAML patients at diagnosis (n=32).

	Mean (Range)	
Age, years	9.63 (0 - 17)	
WBC count, x 10^9/L	59.67 (3 - 336)*	
Morphological blast count		
BM, %	62.77 (5 - 96)†	
PB, %	45.34 (1 - 95)	
	N	%
Time point		
Only diagnosis (Dx)	26	81.3%
Paired Dx-R couples	6	18.8%
Sample		
Only BM	7	21.9%
Only PB	4	12.5%
Paired BM-PB couples	21	65.6%
Gender		
F	16	50.0%
M	16	50.0%
Primary/secondary AML		
Primary	28	87.5%
Secondary	4	12.5%
Treatment protocol		
DB AML-01	8	25.0%
NOPHO-DBH AML2012	18	56.3%
Other	6	18.8%
Occurence of event		
No	13	40.6%
Yes	14	43.8%
Death in CR1	1	
Relapse	10	
Resistant disease	3	
Unknown	5	15.6%
Status		
Alive	22	68.8%
Dead	4	12.5%
Death	2	
Death after relapse	2	
Unknown	6	18.8%
WT1 overexpression		
Yes	21	65.6%
No	9	28.1%
Unknown	2	6.3%

	N	%
Translocation		
Yes	15	46.9%
No	16	50.0%
Unknown	1	3.1%
Core-binding factor leukemia		
Yes	11	34.4%
AML1-ETO + *C-KIT*^{WT}	1	
AML1-ETO + *C-KIT*^{MUT}	2	
AML1-ETO + *C-KIT*ND	2	
CBFB-MYH11	6	
No	20	62.5%
Unknown	1	3.1%
NPM1		
Mutated	0	0.0%
Wild type	30	93.8%
Unknown	2	6.3%
FLT3		
ITD	9	28.1%
ITD+TKD	**1**	3.1%
Wild type	21	65.6%
Unknown	1	3.1%
CEBPA		
Double mutated	1	3.1%
Wild type	29	90.6%
Unknown	2	6.3%
Karyotype		
Abnormal	18	56.3%
Normal	10	31.3%
Unknown	4	12.5%
CNS involvement		
Yes	3	9.4%
No	25	78.1%
Unknown	4	12.5%
Risk classification		
HR	5	15.6%
SR	21	65.6%
Unknown	6	18.8%
FAB classification		
Immature (M0 - M1)	3	9.4%
Mature (M2-M7)	27	84.4%
Unknown	2	6.3%

Characteristics of 32 diagnostic CD34+ pedAML patients used flow cytometric evaluation of the LSC compartment. *WT1* overexpression was interpreted in regard to in-house or published (Cilloni et al. 2009) cut-offs. Superscripts indicate one (*) or two (†) missing data.

PedAML, pediatric acute myeloid leukemia; LSC, leukemic stem cell; Dx, diagnosis; R, relapse; F, female; M, male; WBC, white blood cell; BM, bone marrow; PB, peripheral blood; *FLT3*, fms-like tyrosine kinase receptor-3; *NPM1*, nucleophosmin; CEBPA, CCAAT/enhancer-binding protein alpha; ITD, internal tandem duplication; WT, wild type; MUT, mutated; FAB, French-British-American; CBF, core-binding factor; *WT1*, Wilms' tumor 1; CNS, central nerve system; HR, high risk; SR, standard risk; CR1, first complete remission.

Table 2. Demographics of CD34+ pedAML patients at relapse (n=9).

	Mean (Range)	
Age, years	8.67 (2 - 17)	
WBC count, x 10^9/L	42.44 (1 - 188)§	
Morphological blast count		
BM, %	39.83 (12-80)‡	
PB, %	37.8 (1-94)§	
	N	%
Time point		
Only relapse (R)	3	33.3%
Paired Dx-R couples	6	66.7%
Sample		
Only BM	0	0.0%
Only PB	1	11.1%
Paired BM-PB couples	8	88.9%
Gender		
F	3	33.3%
M	6	66.7%
Former treatment protocol		
DB AML-01	1	11.1%
NOPHO-DBH AML2012	7	77.8%
Other	1	11.1%
Status		
Alive	2	22.2%
Dead after relapse	3	33.3%
Unknown	4	44.4%
WT1 overexpression		
Yes	4	44.4%
No	2	22.2%
Unknown	3	33.3%

Translocation		N	%
	Yes	3	33.3%
	No	6	66.7%
Core-binding factor leukemia			
	Yes	1	11.1%
	AML1-ETO	0	
	CBFB-MYH11	1	
	No	8	88.9%
NPM1			
	Mutated	0	0.0%
	Wild type	8	88.9%
	Unknown	1	11.1%
FLT3			
	ITD	2	22.2%
	Wild type	6	66.7%
CEBPA			
	Double mutated	0	0.0%
	Wild type	8	88.9%
	Unknown	1	11.1%
Karyotype			
	Abnormal	1	11.1%
	Normal	1	11.1%
	Unknown	7	77.8%
CNS involvement			
	Yes	1	11.1%
	No	1	11.1%
	Unknown	7	77.8%
FAB classification			
	Immature (M0 - M1)	0	0.0%
	Mature (M2-M7)	6	66.7%
	Unknown	3	33.3%

Characteristics of nine relapsed CD34+ pedAML patients used flow cytometric evaluation of the LSC compartment. *WT1* overexpression was interpreted in regard to in-house or published (Cilloni et al. 2009) cut-offs. Superscripts indicate three (‡) or four (§) missing data.

PedAML, pediatric acute myeloid leukemia; LSC, leukemic stem cell; Dx, diagnosis; R, relapse; F, female; M, male; WBC, white blood cell; BM, bone marrow; PB, peripheral blood; *FLT3*, fms-like tyrosine kinase receptor-3; *NPM1*, nucleophosmin; CEBPA, CCAAT/enhancer-binding protein alpha; ITD, internal tandem duplication; WT, wild type; MUT, mutated; FAB, French-British-American; CBF, core-binding factor; *WT1*, Wilms' tumor 1; CNS, central nerve system.

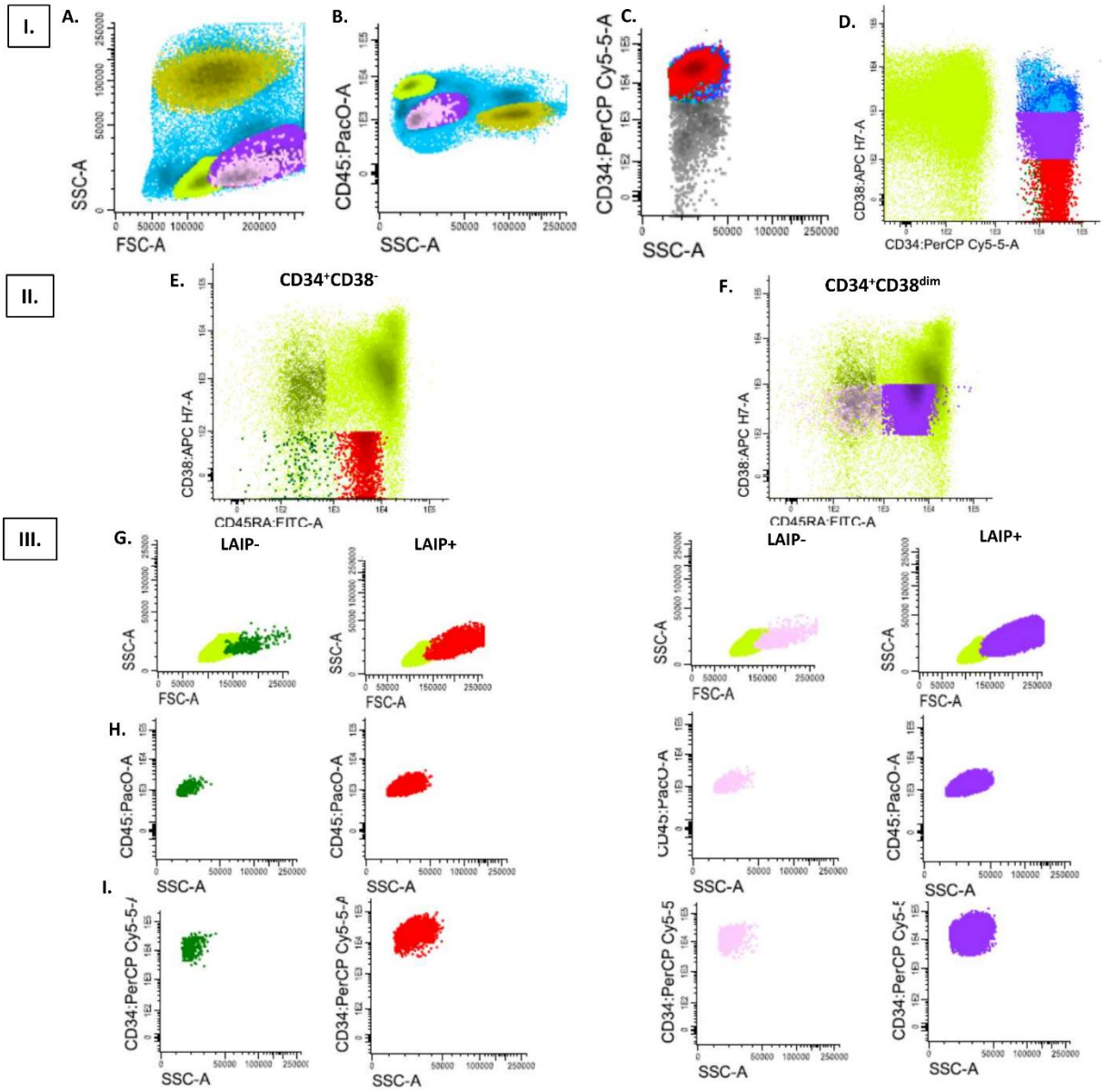
Flow cytometry and gating

LSC immunophenotyping was performed on fresh samples, except for retrospectively selected patients from whom cryopreserved mononuclear cells (MNCs) were thawed. Immunophenotyping of HSC from healthy controls was performed on freshly isolated MNCs. A monoclonal antibody (mAb) cocktail consisting out of backbone markers CD34, CD38 and CD45 was added to each tube, followed by the addition of at least one of the LSC-specific (n=11) or LSC-associated (n=2) LAIP markers (Table S2). Multiple tubes were analyzed per patient and sample matrix containing different markers of interest. Samples and mAb were incubated for 20 min in the dark, followed by 10 min red blood cell (RBC) lysis (2 mL FACSLysing solution, Becton Dickinson, BD, San Jose, CA, U.S.A.). Labelled cells were spinoculated (400 g, 5 min) and washed with 2 mL wash buffer (PBS/0.05% azide-0.1%HSA). FCM analysis was performed on a FACSCanto II flow cytometer (BD Biosciences, San Jose, CA, USA), and instrument set-up was performed according to EuroFlow [31, 32].

Gating was performed in Infinicyt (v.1.8, Cytognos, Salamanca, Spain) according to a strategy shown in Fig 2. CD34+ cases were defined as those with >1% CD34 expression [20, 33]. CD38 expression was gated based on fixed thresholds, i.e. $CD38^-$ ($<10^2$) and $CD38^{dim}$ (10^2-10^3). These thresholds were proven to be valid in multicenter studies when following EuroFlow recommendations [18, 20, 23]. Expression of each marker was scored individually in $CD34^+/CD38^-$ and/or $CD34^+/CD38^{dim}$ compartments. At least three clustered events were required to evaluate marker expressions and scatters [18]. Cells that scored positive for the LAIP marker of interest (i.e. positive fluorescence shift compared to the relevant negative reference population) were defined as 'LAIP positive' (LAIP+), and LAIP negative (LAIP-) otherwise. Lymphocytes and RBCs were used as negative reference population, except for CD49d, where neutrophils were used. LAIP+ and LAP- LSC populations were backgated on FSC/SSC, CD34/SSC and CD45/SSC plots to exclude non-specific events.

The LSC load was calculated in two different ways. First, LSC load was computed irrespective of the expression of aberrant markers, referred to as the *total* LSC load. Second, LSC load was computed by taking into account only cells that harbor at least one of the 11 evaluated LSC-specific LAIP markers, referred to as the *LAIP+* LSC load. Both the total white blood cell compartment (WBC) and CD34+ compartment were evaluated as denominator. An overview of the number of events per compartment is shown in Table S3.

Fig. 2. Gating strategy.



I: Gating of CD34^{+}/CD38^{-} and CD34^{+}/CD38^{dim} compartments. Doublets and cell debris were excluded based on FSC-A/FSC-H (not shown) and (A) SSC-A/FSC-A. (B) WBC were identified based on positive CD45 expression. Lymphocytes (CD45^{high}/SSC^{low}), granulocytes (CD45^{+}/SSC^{high}) and the immature leukemic compartment (CD45^{low}/SSC^{low}) was defined within the WBC compartment. (C) CD34 expression was evaluated in reference to the red cells-fraction and lymphocytes. (D) The CD38^{low} fraction within the total CD34+ population was gated based on fixed thresholds, i.e. CD38^{dim} (10^{2}-10^{3}) and CD38^{-} (<10^{2}) population. If the CD38^{-} population was difficult to assess due to very low frequencies, the cut-off was uniformly set at 10^{3}, taking into account CD38^{-} and CD38^{dim} subpopulations simultaneously. Gating steps A-D were optimised for each patient, and linked to the analysis profile as a patient-specific gating strategy using Infinicyt software v.1.8 (Cytognos, Salamanca, Spain). This approach prevented subjective bias to occur and guaranteed comparability of the analyses over time.

II: Identification of LAIP+ and LAIP- putative LSC populations. CD38^{-} (E) and CD38^{dim} (F) compartments were evaluated for each marker of interest (shown here for CD45RA). Cells that scored positive (i.e. positive shift of fluorescence compared to the negative reference population) were defined as LAIP positive (LAIP+), and LAIP negative (LAIP-) otherwise.

III: Backgating of LAIP+ and LAIP- LSC populations. LAIP+ and LAIP- CD38^{-} and CD38^{dim} LSC populations were backgated to assure that homogeneous scattered cells with expected physical characteristics were investigated, and not debris or clumps. Backgating was performed for each LAIP+ and LAIP- population based on three plots: (G) SSC/FSC, (H) CD45/SSC and (I) CD34/SSC.

Results

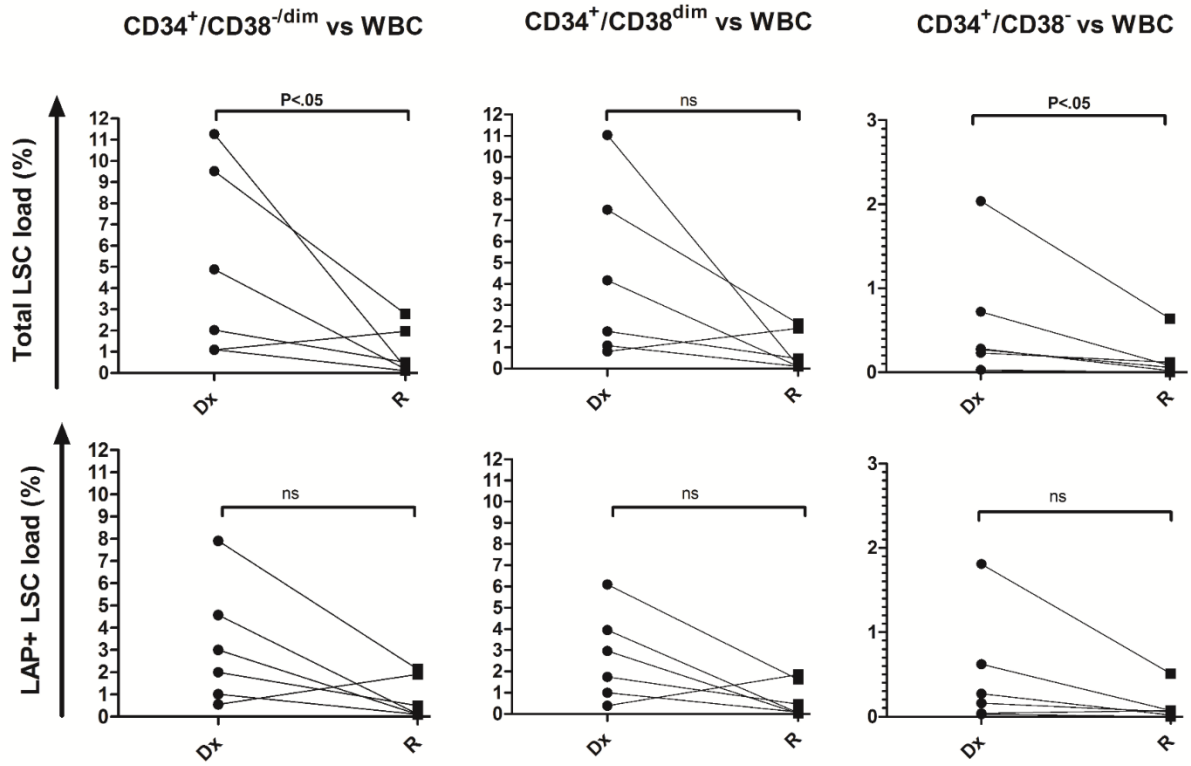
Evaluation of the LSC load at diagnosis and relapse in BM and PB

Size of the $CD34^{+}/CD38^{-/dim}$ fraction within WBCs was determined for CD34+ pedAML patients at diagnosis and relapse. Twenty-seven of the 32 diagnostic patients, and all nine relapsed patients, could be evaluated based on the minimum required $CD34^{+}/CD38^{-}$ events.

At diagnosis, median CD34 positivity was comparable between BM (36.6%, 95% confidence interval 20.3-46.4%) and PB (23.7%, 14.5-29.5%) (P>.05, Table S4). The median *total* $CD34^{+}/CD38^{-/dim}$ LSC load was 4.5% in BM (range 0.05 – 55.2%) and 4.3% in PB (0.04 – 20.0%) (Table S5). The median *LAIP+* LSC load, i.e. frequency of only those $CD34^{+}/CD38^{-/dim}$ cells harbouring LAIP expression, was 2.1% in BM (range 0.02 – 53.6%) and 3.8 % in PB (range 0.04 – 18.1%). So thus, frequency of the *LAIP+* LSC load was averagely 1.65-fold lower compared to the *total* LSC load (P<.0001). *Total* and *LAIP+* LSC loads were 9.04- and 11.3-fold higher in the $CD38^{dim}$ fractions (CD38 expression between 10^{2}-10^{3}) compared to the $CD38^{-}$ fractions (<10^{2}), meaning that the $CD38^{dim}$ fraction comprises the highest share of the LSC compartment (Table S5).

Within patients who relapsed, CD34 positivity was also comparable between BM and PB (P>.05). The median *total* $CD34^{+}CD38^{-/dim}$ LSC load measured in BM was 0.37% (range 0.002 – 7.4%), and 0.96% (0.18 – 2.0%) in PB. The median *LAIP+* LSC load was again significantly lower, i.e. 0.24% (range 0.001 – 6.2%) in BM and 0.33% (0.02 – 1.9%) in PB (P<.05). These data suggest that the LSC frequency decreases at relapse compared to diagnosis. We further investigated this hypothesis by comparing the *total* and *LAIP+* LSC load between diagnostic and relapsed samples on a per patient basis (Table S6). The *total* LSC load was significantly lower at relapse (median 0.38%) compared to diagnosis (median 3.46%) (P<.05, Fig. 3). The same observation was true for the *LAIP+* LSC load, but, not at a significant level (0.31% versus 2.5%, respectively, P=.075). Significance of the difference was mainly located in the $CD34^{+}/CD38^{-}$ fraction, and was borderline for the $CD34^{+}/CD38^{dim}$ fraction (P=.075). Hence, these data further suggest that although the size of LSC load decreases, LSC harbor an equally of even more aberrant phenotype at relapse compared to diagnosis.

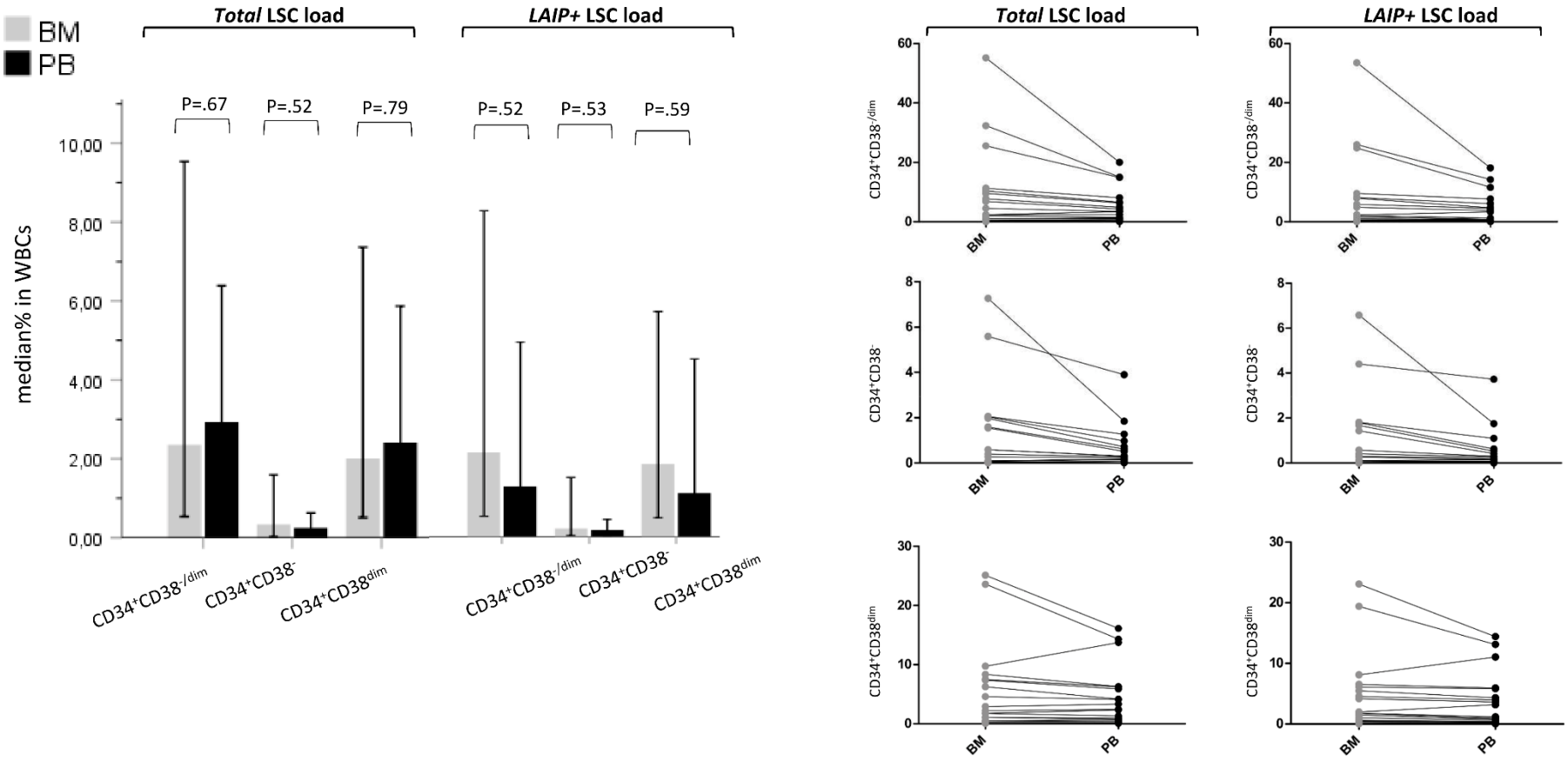
Fig. 3. Comparison of the *total* and *LAIP+* LSC load between paired Dx-R samples.



Comparison between *total* and *LAIP+* LSC load in a limited number (n=4) of patients with Dx-R couples available (n=6 samples). Four of the six CD34+ patients with Dx-R couples available fulfilled the minimal required number of events at both time points. Paired BM and PB samples were available for 2/4 patients. BM and PB measurements were evaluated as independent data points to increase statistical power. The total LSC load significantly decreased upon relapse, while the *LAIP+* compartment decreased, but not at a significant level. Horizontal square brackets represent statistical comparisons (Wilcoxon signed-rank test). Significant P-values are indicated in bold.
BM, bone marrow; PB, peripheral blood; WBC, white blood cells; Dx, diagnosis; R, relapse; LSC, leukemic stem cells; LAIP, leukemia-associated immunophenotype.

We next evaluated whether BM and PB showed significant differences regarding the LSC load. Paired BM-PB samples could be evaluated for 20 pedAML patients (Table S7). Although LSC loads tended to be somewhat higher in PB, not the *total* and nor the *LAIP+* $CD34^{+}/CD38^{-/dim}$, $CD34^{+}/CD38^{-}$ or $CD34^{+}/CD38^{dim}$ LSC frequencies within the WBC compartment did not significantly differ between BM and PB (Fig. 4). Taking into account both BM and PB, on average 64.9% of the *total* LSC load harbored the expression of at least one LAIP marker at diagnosis (range 41.5 – 87.7%). So forth, based on a total set of 11 LSC-specific markers, roughly two third of the *total* $CD34^{+}/CD38^{-/dim}$ compartment can be defined as *LAIP+.* These data also suggest that PB is at least as suited as BM to determine the phenotype and frequency of the LSC compartment. Based on this knowledge, and the fact that LAIP combinations hold a higher specificity in PB, we further used the *LAIP+* LSC load in PB to evaluate associations with molecular characteristics and clinical outcome.

Fig. 4. Comparison between *total* and *LAIP+* LSC loads in BM compared to PB in pedAML patients at Dx.



Paired BM-PB samples were available for 21 Dx CD34+ pedAML patients. One patient was excluded due to a too low number of events in one of both sample matrixes. Hence, the *total* and the *LAIP+* LSC load calculated in within the WBC compartment could be compared between BM and PB in 20 pedAML patients. Measurements in PB are indicated in grey, and BM measurements are shown in black. Height of the bars are representative for the median values and error bars indicate ± SEM. Horizontal square brackets represent statistical comparisons (Wilcoxon signed-rank test). Two-sided P-values were calculated by the Wilcoxon signed-rank test. No significant differences were observed. Dx, diagnosis; BM, bone marrow; PB, peripheral blood; pedAML, pediatric AML; LSC, leukemic stem cells; LAIP, leukemia-associated immunophenotype.

Patients with high LSC loads show a trend towards detrimental outcome

CD34+ patients for whom the LSC load could be evaluated (n=27/32) were stratified into LSC^{high} and LSC^{low} groups based on two cut-offs calculated by ROC curve analysis. The *LAIP+* LSC load determined in PB, or the BM *LAIP+* LSC loads if PB was not available (n=3/27), was used as numerator. Two denominators were evaluated i.e. WBCs and CD34+ cells, in order to allow comparison with current literature. The cut-off to discriminate LSC^{high} from LSC^{low} patients within CD34+ cells (17.39%) appeared to be almost identical to the one previously established by Hanekamp et al. (17.2%) [18]. By contrast, the ROC-based cut-off using the WBCs as denominator (4.78%) was sixfold higher compared to the cut-off published by Witte et al. (≤0.68% or ≥0.83%) [17].

Irrespective of the cut-off used, LSC^{high} patients showed a higher share of *WT1* overexpression, *FLT3*-ITD mutations, abnormal karyotypes and (non-CBF) translocations compared to LSC^{low} patients, although not at a significant level (P>.05) (Table S8). LSC^{high} patients classified according to the 4.78% cut-off also showed a trend towards higher WBC counts. These data suggest that pedAML patients with high LSC loads tend to present molecular profiles associated with detrimental outcome.

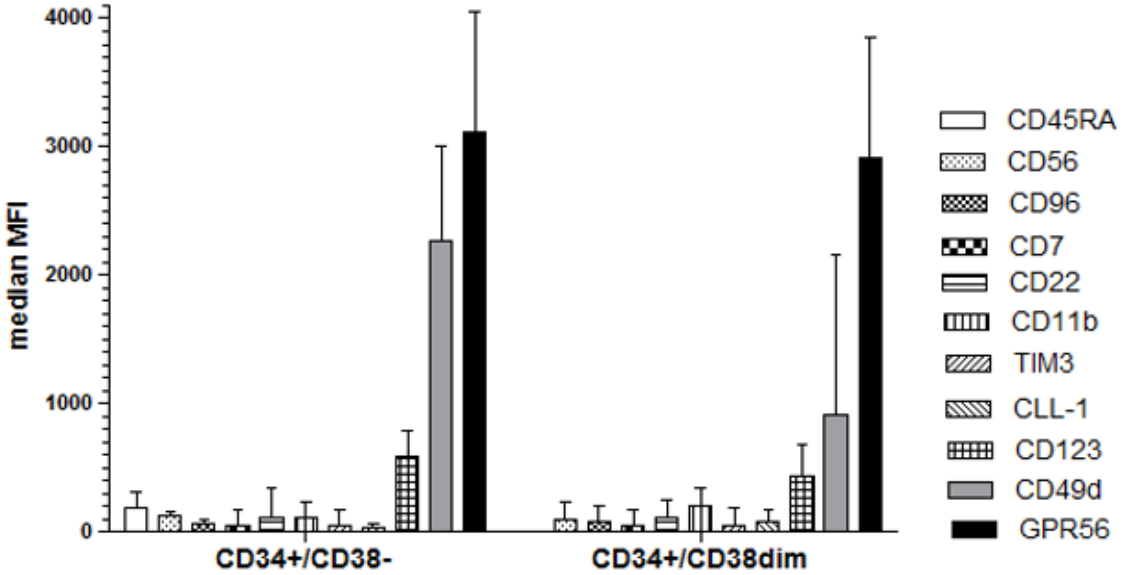
We questioned whether the diagnostic *LAIP+* LSC load was able to further risk stratify CD34+ pedAML patients. We selected those patients for whom clinical outcome data were available (n=23/27), treated according to the NOPHO-DBH AML2012 study (n=18/23), DB-AML01 study (n=4/27) or otherwise (n=1/27). Clinical outcome data were also available for 18/19 CD34- patients treated in the NOPHO-DBH AML2012 study.

Within NOPHO-DBH AML2012-treated patients (n=36), CD34+ patients showed a trend towards a lower EFS compared to CD34- patients (55.6% vs 77.8%; P>.05, Fig. S1A), while OS was highly comparable (88.9% vs 83.3%, respectively, Fig. S1B, P>.05). Difference in EFS was then evaluated following dichotomization of CD34+ patients (18/36) into LSC^{high} and LSC^{low} groups. Dichotomisation based on the cut-off of 4.78% within WBC resulted into 13 LSC^{low} and 5 LSC^{high} patients. The median LSC frequency within the LSC^{low} group was 0.71% (95% CI 0.23-3.00%) and 8.20% within the LSC^{high} group (95% CI 7.68-18.15%). Dichotomization based on a cut-off of 17.39% within the CD34+ compartment generated similar results , i.e. 15 LSC^{low} (median 5.67%, 95% CI 2.37-10.28%) and 3 LSC^{high} (median 25.81%, 22.34-51.48%) patients. However, no significant difference in EFS was observed between LSC^{low} and LSC^{high} patients, irrespective of the cut-off used (Fig. S2 A-B, Table S9).

Most likely, the total number of patients with an event (n=8) was too low to yield a statistical significant result. This assumption was strengthened by including also four DB-AML01- and one otherwise-treated patients, of which 4/5 showed an event, yielding a total of 12 patients with an event. Patients then classified as LSC^{high} (n=8) based on the *LAIP+* LSC load within WBC showed a clear trend towards lower EFS compared to LSC^{low} patients (n=15) (53.3% versus 37.5%, P=.083) (Fig. S2C). The same observation could be made by classifying patients based on a *LAIP+* LSC load within CD34+ cells (50.0% (n=7) vs 42.9% (n=16), P=.22, Fig. S2D). Hence, these data will most likely be more convincing upon completion of the NOPHO-DBH AML2012 study and await further investigation.

Fig. 5. Expression of LSC-specific and –associated markers in HSC.

Marker	Healthy controls (medians)			
	$CD34^+/CD38^-$	$CD34^+/CD38^{dim}$	negative reference	positive reference
CD56	121	92	91	1074
CD96	64	73	109	1467
CD7	50	41	39	4288
CD22	112	109	36	3669
CD11b	114	203	142	8163
TIM3	46	45	50	2895
CLL-1	32	72	33	2656
CD123	580	430	41	1272
CD49d	2261	911	213	1370
GPR56	3120	2909	111	2794



Healthy controls consist out of cord blood (CB) and normal pediatric bonne marrow samples (NBM). The number of controls evaluated differs per marker, as shown in Table S9. CD45RA was excluded from this comparison since normal lymphoid-primed multipotent progenitor (LMPP) cells with CD38low expression also harbour positive CD45RA expression [34]. HSC, hematopoietic stem cell; LSC, leukemic stem cells.

Phenotypic characterization of the LSC compartment

We next determined which LAIP markers are present in the LSC fraction of CD34+ pedAML patients. Expression of eleven LSC-specific LAIP markers CD45RA, CD56, CD96, CD7, CD22, CD11b, TIM3, CD123, CLL-1, CD15 and NG2 was evaluated in the $CD34^{+}/CD38^{-/dim}$ compartment. CD15 and NG2 were only measured in BM when patients were included in the NOPHO-DBH AML2012 study. If material availability was too low to evaluate all markers, a set four markers conjugated to different fluorochromes were combined in one tube (CD123, CD56+2 and CD7). In addition, we explored differential expression of two LSC-associated markers CD49d and GPR56. Background expression of 9/11 LSC-specific LAIP markers, and CD49d and GPR56, was measured in a total cohort of 24 healthy controls. An overview of the number of patients and samples analyzed per marker is shown in Table S10.

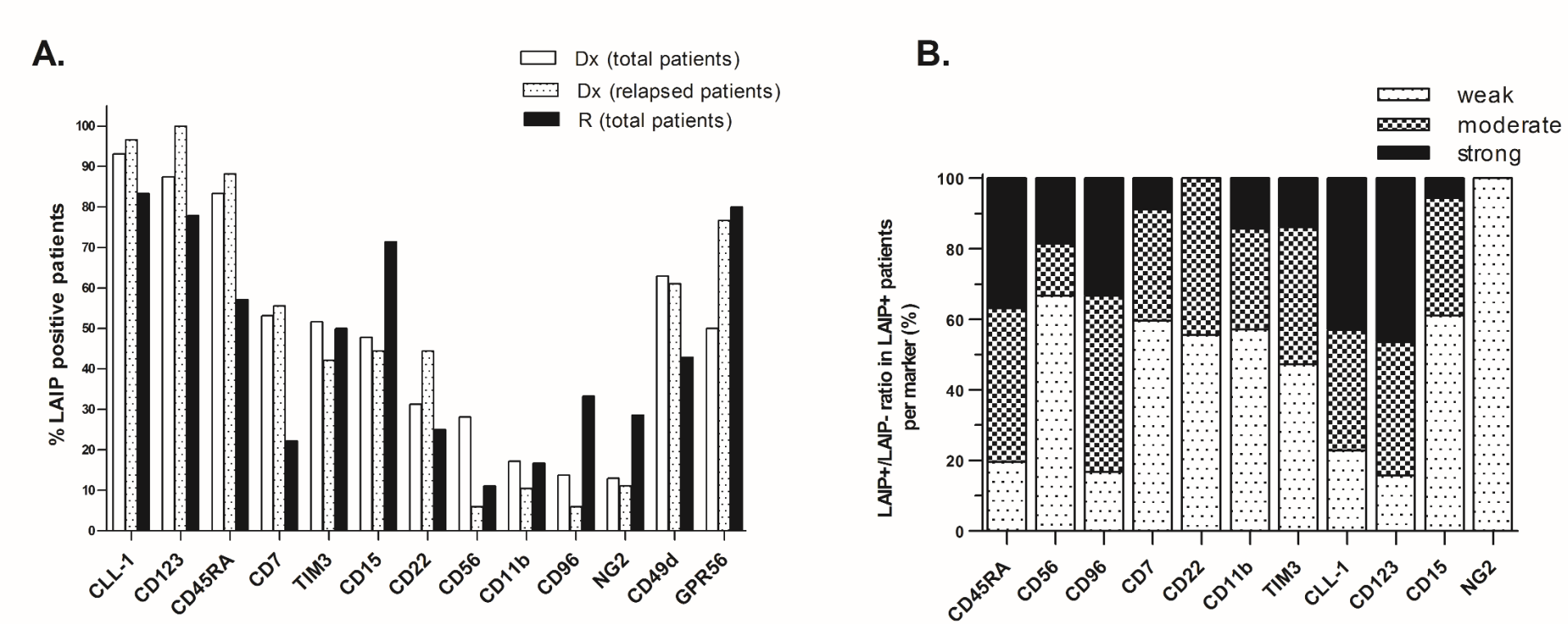
Specificity, prevalence and sensitivity of LSC-specific and –associated LAIP markers

For a marker to be LSC-specific, hematopoietic stem and progenitor cell (HSPC) populations are required to be negative. MFI values measured in HSC from healthy controls were lower or comparable to those observed for the relevant negative reference population for all evaluated LSC-specific LAIP markers, except for CD123 (Fig. 5). Median CD123 expression appeared to be more than tenfold higher in HSC compared to the negative lymphocyte subpopulation. CD49d and GPR56 were ubiquitously expressed in HSC.

We evaluated the prevalence of LAIP markers in diagnostic and relapsed CD34+ pedAML patients. CLL-1 (93.1%), CD123 (87.5%) and CD45RA (83.3%) showed the highest prevalence at diagnosis (Fig. 6A). CD7, TIM3 and CD15 were positive in about half of the patients, and CD56 and CD22 in about one third. CD49d and GPR56 overexpression was observed in approximately half of the cases. Evaluation of only the diagnostic patients who eventually relapsed showed a much lower frequency of CD56 (5.90%), suggesting that CD56 is not a good marker to predict relapse. By contrast, a higher frequency of GPR56 expression was observed at relapse (76.6%). At relapse, CLL-1 (83.3%) and CD123 (77.8%) remained most represented, followed by CD15 (71.4%) instead of CD45RA (57.1%). Expression of TIM3 and CD11b remained stable, while CD96, NG2 and GPR56 showed a minor increase, and CD7, CD56 and CD49d slightly decreased.

We next determined the coverage of each LAIP LSC marker within the LSC compartment (i.e. sensitivity), by determining the ratio of LAIP+ events versus LAIP- events. LAIP+/LAIP- ratios expressions were categorized into three groups, i.e. weak (ratio 0.05 – 1), moderate (ratio 1 – 10) and strong (ratio ≥10). LAIP markers that occurred most frequently, also seemed to occur at the highest ratios (Fig. 6B). LAIP+/LAIP- ratios determined for CLL-1, CD123 and CD45RA were ≥10 in 43.0%, 46.5% and 37.0% of the cases, respectively. CD96+ LSCs were at least equally represented as CD96- LSCs in half of the cases. The other LAIP+ LSC populations were in the vast majority not more than tenfold higher represented than their respective LAIP- compartment.

Fig. 6. Prevalence of LSC-specific and –associated LAIP markers in pedAML at Dx and R.



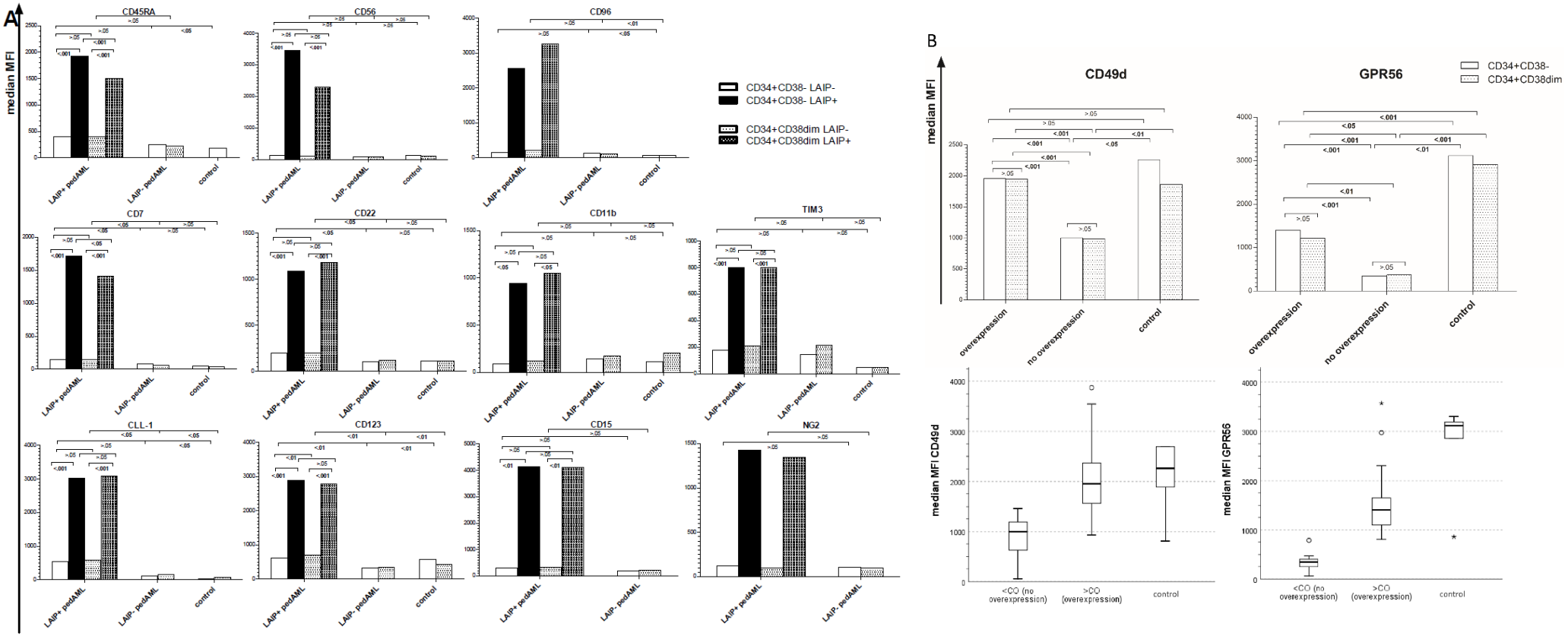
(A) Percentage of patients harbouring positive expression for each evaluated LSC-specific and –associated marker. Prevalence was determined in all diagnostic patients, a subcohort of the diagnostic patients i.e. only those who ultimately experienced an event (Dx relapsed), and relapsed patients. An overview of the number of patients evaluated per marker is shown in Table S10. **(B)** Determination of the relative size of the LSC fraction harbouring a LAIP marker within LAIP+ patients (percentage per marker shown in Table S8). The ratio of the number of LAIP+ versus LAIP- LSC was determined for each marker individually. LAIP+/LAIP- ratios expressions were categorized into three groups, i.e. weak (ratio 0.05 – 1), moderate (ratio 1 – 10) and strong (ratio ≥10). Sample matrixes were pooled, since the number of LAIP+ LSC in PB and BM appeared not to be significantly different. Dx, diagnosis; R, relapse; LSC, leukemic stem cells; LAIP, leukemia-associated immunophenotype.

Expression strength (MFI) of the LAIP+ LSC population was significantly higher than observed for the concomitantly present LAIP- LSC population within the same patient for each marker (Fig. 7A). No significant MFI differences were observed between LAIP+ $CD38^{-}$ and $CD38^{dim}$ populations, except for CD45RA, CD7 and CD123. Higher CD45RA expression (P<.01) in the $CD34^{+}/CD38^{dim}$ compartment can be explained by the admixture of lymphoid-primed multipotent progenitor (LMPP) cells [34]. Higher CD123 expression observed in CD123- $CD34^{+}/CD38^{dim}$ cells (P<.01) is likely due to contamination with myeloid precursors. No explanation was readily available for the higher CD7 expression in $CD34^{+}/CD38^{-}$ compared to $CD34^{+}/CD38^{dim}$ cells (P=.026). Altogether, we conclude that LSCs with an aberrant phenotype are equally distributed along the CD38 axis.

MFI values measured in LAIP- LSCs, from patients harbouring a LAIP+ LSC population, did not significantly differ from those from LAIP- patients. Only CD7, CD22 and CD123 showed a significant difference (P=.014, .049 and .002, respectively), i.e. higher in LAIP- LSC from LAIP+ patients than LAIP- patients. It is likely that the LAIP+ compartment was underestimated, causing 'spill over' in the subsequently slightly overestimated LAIP- compartment.

MFI values were also significantly higher in LSC populations of patients defined as having CD49d and GPR56 overexpression (P<.001 or <.01, Fig. 7B). PedAML patients had lower CD49d and GPR56 baseline expression compared to healthy controls. Expression levels in patients defined as having CD49d overexpression did not significantly differ from healthy controls, in contrast to GPR56, where baseline HSC expression was twofold higher than observed in GPR56-overexpressed patients.

Fig. 7. Comparison of MFI values between LAIP+ and LAIP- LSC subpopulations within LAIP+ pedAML, next to LAIP- LSC subpopulations between LAIP- patients and healthy controls.



(A) Differential expression of the median fluorescence intensity (MFI) per LSC-specific marker between LAIP+ and LAIP- LSC subpopulations ($CD34^+/CD38^-$ and $CD34^+/CD38^{dim}$) within LAIP+ patients, next to the median expression levels measured in $CD34^+/CD38^-$ and $CD34^+/CD38^{dim}$ subpopulations from LAIP- patients and healthy controls. Significance was calculated within LAIP+ patients, except for CD96 and NG2, due to a too low number of positive samples (n=4 and 3, respectively).
(B) Top: Differential expression (MFI) in $CD34^+/CD38^-$ and $CD34^+/CD38^{dim}$ subpopulations per LSC-associated marker within patients with overexpression versus those without, and compared to healthy controls. Bottom: spread of the MFI values for LSC-associated markers measured in the total $CD34^+CD38^{-/dim}$ compartment between patients with overexpression, without overexpression, and in healthy controls (cord blood and normal pediatric bone marrow).

LSC phenotypic profiles are highly heterogeneous and can be linked to molecular subgroups

All but one of the 27 CD34+ pedAML patients expressed at least one LAIP marker at diagnosis. Approximately half of the patients showed presence of five (24.2%) or six markers (24.2%), and 9.1% presented more than six markers. The number of positive LAIP markers in relapsed patients was also highly variable. Most patients showed presence of four (22.2%) or five (33.3%) markers, and one patient showed no LAIP positivity. No recurrent immunophenotypes could be identified at diagnosis or relapse. Interestingly, patients with >4 LAIP markers expressed in the LSC compartment harbored more frequently CBF-leukemia than patients with ≤4 LAIP markers (56.3% versus 13.3%, P=.013). The occurrence of an event did not significantly differ between both groups.

Patients with CD56+ expression in the LSC compartment were significantly older (70.0% >10 yrs. versus 30.0% <10 yrs., P<.05) and showed significantly more abnormal than normal karyotypes (90% versus 10%, respectively, P<.05). Translocations were mainly observed in patients lacking CD7 expression (11.7% CD7+ versus 88.2% CD7-, P<.001), including CBF-leukemia (8.3% CD7+ versus 91.7% CD7-, P<.01). *FLT3*-ITD mutations were more frequent in patients harbouring CD7 LSC expression (88.9% CD7+ versus 11.1% CD7-, P=.01). TIM3+ patients borderline significantly showed more *WT1* overexpression than TIM3- patients (72.2% versus 27.8%, P=.05). Sixty-nine percent of the patients with LSC CD49d overexpression showed WBC counts <30x10^9/L, compared to only 31% of the pedAML with normal CD49d expression (P=.07). *FLT3*-ITD mutations were significantly more found in patients with GPR56 overexpression (85.7% GPR56+ versus 14.3% GPR56-, P<.05). CBF-leukemias were exclusively observed in patients without GPR56 overexpression (P<.01).

LAIP-marker associated scatter properties and backbone marker expressions

FSC/SSC positions and CD34/CD45 expression were determined for LAIP+ and LAIP- LSC populations as a ratio versus the median values measured in lymphocytes.

FSC/SSC/CD34/CD45 ratios showed only minor variation in HSC from healthy controls (Table S11B). In pedAML patients, CD45RA expression was characterized by significantly higher FSC, SSC, CD34 and CD45 ratios. CD45RA+ LSCs showed 1.13-fold and 1.21–fold higher FSC and SSC, respectively, than those lacking CD45RA expression within the same patients (P<.001, Table S11A). Properties of CD45RA- LSCs within patients harbouring CD45RA expression were similar to those observed in patients without CD45RA expression. Also other LAIP+ LSC subpopulations harbored significantly higher FSC and/or SSC ratios, i.e. CLL-1, CD123, CD15, CD56, CD7 and TIM3. As these markers were often co-expressed with CD45RA, it remains to be elucidated whether these aberrant scatter properties follow from concomitant CD45RA expression, or from the respective LAIP marker itself. Increased CD34 and/or CD45 ratios were observed for CD56+, CD7+, CD11b+, TIM3+ and CD15+ LSC populations.

Discussion

The $CD34^{+}/CD38^{-}/CD45^{low}/FSC^{low}/SSC^{low}$ fraction is a helpful surrogate to estimate the LSC burden, but also contains normal HSC. Gating of the $CD38^{low}$ compartment can be challenging as the leukemic population often represents a continuum of $CD34^{+}/CD38^{+/-}$ cells. Recent guidelines detail how to define $CD38^{-/dim}$ populations, and can theoretically be adapted by other centers when following EuroFlow guidelines [31, 35]. We here confirm that the fixed 10^2 and 10^3 cut-offs described in these guidelines, which have been multicenter validated in large studies, can be easily implemented and without a hitch when respecting all pre-analytical and analytical EuroFlow requirements. Although $CD34^{+}/CD38^{dim}$ populations (10^2 - 10^3) are by average tenfold larger than $CD34^{+}/CD38^{-}$ fractions ($<10^2$), the expression of aberrant LAIP markers was shown to be equally distributed along the CD38 axis. Therefore, the entire $CD38^{low}$ population $<10^3$ may be considered in pedAML, which will increase robustness and sensitivity of the analysis.

The median *total* $CD34^{+}/CD38^{-/dim}$ LSC load in BM (4.50%) and PB (4.60%) was higher than those previously reported in a pediatric setting (mean 1.02 ± 1.28% [17] and median 0.23% [19], both evaluated in BM). Also the *LAIP*+ LSC fraction was larger than previously reported (2.05% (BM) and 4.21% (PB) compared to 0.99% (BM) [18]). This latter observation can be explained by the inclusion of more LAIP markers in the present study. We observed in a small cohort of paired diagnosis-relapse couples that the size of the total $CD34^{+}CD38^{-/dim}$ compartment decreased at relapse compared to diagnosis, while the number of cells harbouring aberrant LAIP expression did not. More diagnosis-relapse couples are needed to confirm this finding.

Previous studies investigating the diagnostic LSC load often mixed BM and PB depending on availability [12, 14], or were limited to BM [17, 18]. We demonstrated that nor the size nor the LAIP positivity of the *total* and *LAIP+* $CD34^{+}CD38^{-/dim}$ compartment significantly differs between BM and PB. Hence, PB may be considered as a valid alternative for LSC immunophenotyping when BM aspiration fails ('dry-tab') or material is limited. The lower background observed in PB due to virtual absence of myeloid progenitor cells increases specificity of the LSC analysis. The lack of sensitivity, as LSC mostly reside in the BM niche, does not seem to be problematic at diagnosis. Validation of our findings in a larger cohort, and in a follow-up setting, are needed, as (traumatic) BM aspirates could hence be replaced by less invasive PB samples.

The cut-off determined in our study to discriminate LSC^{high} from LSC^{low} patients based on the *LAIP+* LSC load within the CD34+ compartment (17.39%) was almost identical to the one previously established by Hanekamp et al. (17.2%) [18]. We did not detect a significant association with EFS for pedAML patients classified as LSC^{high} compared to LSC^{low} patients using two different thresholds. As the difference in EFS became more significant by including more patients, though treated differently, it is likely that our patient cohort was too small to detect genuine significant differences. Hence, evaluation of these cut-offs will be of high interest at the end of the NOPHO-DBH AML2012 study.

Molecular subgroups appeared not to be significantly enriched in LSC^{high} patients, although a higher share of WBC counts, *WT1* overexpression, *FLT3*-ITD mutations, abnormal karyotypes and non-CBF translocations was observed.

In the present study, LSC in CD34+ pedAML were characterized by a DfN approach using 11 LSC-specific markers. Approximately two third of the $CD34^{+}/CD38^{-/dim}$ compartment harbored an aberrant phenotype, which is lower than described in adult AML (87%) [23]. Since not all LSC-specific markers available through literature were assayed, i.e CD44 [36], we cannot entirely exclude a leukemic origin

for LAIP- cells. Therefore, we reserved the term “HSC” for $CD34^{+}/CD38^{-/dim}$ cells from healthy controls, and “LAIP- LSC” for $CD34^{+}/CD38^{-/dim}$ cells in pedAML patients without (dateable) aberrant marker expression.

We showed that LAIP markers CD45RA, CD56, CD96, CD7, CD22, CD11, TIM3 and CLL-1 are genuinely LSC-specific and may serve to distinguish putative LSC populations from HSC. CD123, on the other hand, should be considered as a LSC-associated marker. These results are in agreement with previous findings in adult NBM [37]. We found CD96 to be negative in HSC, in contrast to a previously reported weak expression in small adult [38] (n=3) and pediatric [19] (n=12) NBM cohorts.

CLL-1 and CD45RA were the most frequent and strongest expressed LAIP markers. CD7, TIM3 and CD15 were positive in about half of the patients, and CD56 and CD22 in about one third. These data further underline the added value of combining LSC antigen-directed monoclonal antibodies (mAb) labelled to the same fluorochrome into a DfN cocktail approach, and provide proof-of-concept that the markers included in the one-tube assay of the Dutch LSC research group are also of diagnostic value in pedAML [23]. These data also illustrate that, although pediatric and adult AML are two distinct entities [39], the LSC compartment seems to bear comparable phenotypical aberrancies.

The various immunophenotypic profiles confirm high inter- and intra-patient heterogeneity regarding the LSC compartment, and therefore the need for a combinatorial approach. We showed that aberrant CD45RA expression is also accompanied by increased SSC/FSC properties in a pediatric setting (1.13-fold and 1.21–fold, respectively, P<.001). Increased CD34 and/or CD45 ratios were also observed for CD56+, CD7+, CD11b+, TIM3+ and CD15+ LSC populations. These findings are in line with those reported for adult AML [12, 20] and confirm that FSC/SSC positions and backbone marker expressions are helpful in fine-tuning neoplastic LSC populations. Patients with CD56+ LSC expression appeared to be older (P<.05) and harbored more abnormal karyotypes (P<.01). This finding is in line with van Solinge et al. [40], who showed a significant correlation between CD56 blast expression and monosomal karyotypes. We also found *FLT3* more frequently mutated in patients harbouring CD7 expression (P=.01). CD7 expression in blasts has been associated with *FLT3*-ITD [41, 42]. A borderline significant association between TIM3 and *WT1* overexpression was found, which awaits validation (P=.05).

CD49d overexpression was observed in the LSC compartment of approximately half of the diagnostic patients. This is the first report addressing CD49d expression in LSCs of pedAML. Controversy exists between adults and children with AML concerning the prognostic impact of CD49d (VLA-4 α-chain). The Children’s Oncology Group found high VLA-4 expression to be associated with younger age, lower *FLT3*-ITD prevalence and significantly beneficial impacted relapse in a multivariate model [43]. We found a borderline significant correlation between LSC CD49d overexpression and lower WBC counts (<30x10^9/L, P=.07). In adult AML, VLA-4 expression in bulk leukemic cells negatively impacted OS [44] and CR [45], with significant higher relapse rates in the latter study.

GPR56 LSC overexpression showed mutual exclusivity with CBF-leukemia (P<.01) and a significant higher proportion of patients harboured *FLT3*-ITD mutations (P<.05). GPR56 overexpression was observed in 50.0% of the patients at diagnosis, and increased at relapse (80.0%). The percentage of patients with GPR56+ LSC overexpression at diagnosis who experienced an event was also higher (76.6%) than those who ultimately showed no event. Altogether, these observations suggest that GPR56 overexpression could be associated with a detrimental outcome. Conflicting data were reported on GPR56 expression in HSPC compared to leukemic cells [25, 46]. We here illustrate that baseline GPR56 expression in HSC from CB and NBM of pediatric patients is significantly higher than in

LSC of pedAML patients. It will be interesting to evaluate if GPR56 overexpression in the LSC compartment of pedAML is also linked to *EVI1* overexpression, as documented for adult AML [47].

It is important to acknowledge that the 27/32 CD34+ pedAML patients, in whom LSC load could be determined, also contained three secondary AML cases transformed after MDS/JMML. Knowledge on whether LSC phenotypic aberrancies differ between primary and secondary pedAML will require larger cohorts.

In conclusion, we here describe a detailed immunophenotypical characterization of the LSC compartment in CD34+ pedAML patients at diagnosis and relapse. Narrowing down the $CD34^{+}/CD38^{-/dim}$ compartment to only those cells that harbour aberrant LAIP expression is crucial for detecting genuine LSCs. Approximately two third of the $CD34^{+}/CD38^{-/dim}$ compartment in pedAML harbors an aberrant phenotype. LSC aberrancies appear to be quite similar in a pediatric and adult setting, and show a high inter- and intra-patient heterogeneity. We demonstrate that PB is an advantageous alternative sample matrix at diagnosis to determine the LSC load. We consolidate that the previously reported cut-off by Hanekamp et al. allows optimal separation of LSC^{high} from LSC^{low} patients based on the *LAIP+* LSC load at diagnosis within the CD34+ compartment. More patients treated according to the NOPHO-DBH AML2012 protocol are needed to confirm a significant prognostic impact of the diagnostic LSC load. These data pave the way for LSC-targeted therapies in pedAML, and show that future strategies will benefit from combinatorial approaches.

Author contributions

JP, TL and BD designed the research study. BD performed the research and analyzed the data. DH and JC assisted in data analysis. BD, JP, TL and BDM wrote the paper. BDN and KV contributed molecular analyses. AU, AVD, LD, MFD and JVDW contributed patient samples. All authors reviewed the manuscript.

Acknowledgements

We thank Prof. F. Plasschaert for providing NBM, the medical staff of the Department of Pediatric Oncology and Haematology of the Ghent University Hospital (Ghent, Belgium) for providing patient material, C. Matthys of the CB Bank of Ghent University Hospital and all technicians of the Childhood Cancer Foundation and Laboratory of Haematology of the Ghent University Hospital (Ghent, Belgium) for their assistance. The authors thank all patients and their parents for their participation in the study, as well as the data managers involved in the clinical trials. This research was supported by the Belgian Foundation against Cancer (grant 2014–265), FOD-KankerPlan (Actie29, grant to JP) and vzw Kinderkankerfonds (grant to TL). The Research Foundation - Flanders (Fonds voor Wetenschappelijk Onderzoek Vlaanderen, FWO) supported and BD (grants 1113117 and V433317N). This work is submitted in partial fulfilment of the requirement for the PhD of candidate BD at Ghent University.

References

1. Rasche M, Zimmermann M, Borschel L, Bourquin JP, Dworzak M, Klingebiel T, Lehrnbecher T, Creutzig U, Klusmann JH, Reinhardt D: Successes and challenges in the treatment of pediatric acute myeloid leukemia: a retrospective analysis of the AML-BFM trials from 1987 to 2012. Leukemia 2018, 32(10):2167-2177.
2. Creutzig U, van den Heuvel-Eibrink MM, Gibson B, Dworzak MN, Adachi S, de Bont E, Harbott J, Hasle H, Johnston D, Kinoshita A et al: Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. Blood 2012, 120(16):3187-3205.
3. Zwaan CM, Kolb EA, Reinhardt D, Abrahamsson J, Adachi S, Aplenc R, De Bont ES, De Moerloose B, Dworzak M, Gibson BE et al: Collaborative Efforts Driving Progress in Pediatric Acute Myeloid Leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2015, 33(27):2949-2962.
4. De Moerloose B, Reedijk A, de Bock GH, Lammens T, de Haas V, Denys B, Dedeken L, van den Heuvel-Eibrink MM, Te Loo M, Uyttebroeck A et al: Response-guided chemotherapy for pediatric acute myeloid leukemia without hematopoietic stem cell transplantation in first complete remission: Results from protocol DB AML-01. Pediatric blood & cancer 2019:e27605.
5. Bonnet D, Dick JE: Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nature medicine 1997, 3(7):730-737.
6. Hope KJ, Jin L, Dick JE: Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity. Nat Immunol 2004, 5(7):738-743.
7. Shlush LI, Mitchell A, Heisler L, Abelson S, Ng SWK, Trotman-Grant A, Medeiros JJF, Rao-Bhatia A, Jaciw-Zurakowsky I, Marke R et al: Tracing the origins of relapse in acute myeloid leukaemia to stem cells. Nature 2017, 547(7661):104-108.
8. Ishikawa F, Yoshida S, Saito Y, Hijikata A, Kitamura H, Tanaka S, Nakamura R, Tanaka T, Tomiyama H, Saito N et al: Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. Nat Biotechnol 2007, 25(11):1315-1321.
9. Thomas D, Majeti R: Biology and relevance of human acute myeloid leukemia stem cells. Blood 2017, 129(12):1577-1585.
10. Griessinger E, Anjos-Afonso F, Pizzitola I, Rouault-Pierre K, Vargaftig J, Taussig D, Gribben J, Lassailly F, Bonnet D: A niche-like culture system allowing the maintenance of primary human acute myeloid leukemia-initiating cells: a new tool to decipher their chemoresistance and self-renewal mechanisms. Stem cells translational medicine 2014, 3(4):520-529.
11. Ng SW, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, Arruda A, Popescu A, Gupta V, Schimmer AD et al: A 17-gene stemness score for rapid determination of risk in acute leukaemia. Nature 2016, Dec 15(540(7633)):433-437.
12. Terwijn M, Zeijlemaker W, Kelder A, Rutten AP, Snel AN, Scholten WJ, Pabst T, Verhoef G, Lowenberg B, Zweegman S et al: Leukemic Stem Cell Frequency: A Strong Biomarker for Clinical Outcome in Acute Myeloid Leukemia. PloS one 2014, 9(9).
13. van Rhenen A, Feller N, Kelder A, Westra AH, Rombouts E, Zweegman S, van der Pol MA, Waisfisz Q, Ossenkoppele GJ, Schuurhuis GJ: High stem cell frequency in acute myeloid leukemia at diagnosis predicts high minimal residual disease and poor survival. Clinical Cancer Research 2005, 11(18):6520-6527.
14. Zeijlemaker W, Grob T, Meijer R, Hanekamp D, Kelder A, Carbaat-Ham JC, Oussoren-Brockhoff YJM, Snel AN, Veldhuizen D, Scholten WJ et al: CD34(+)CD38(-) leukemic stem cell frequency to predict outcome in acute myeloid leukemia. Leukemia 2018, 33:1102–1112.
15. Hwang K, Park CJ, Jang S, Chi HS, Kim DY, Lee JH, Lee KH, Im HJ, Seo JJ: Flow cytometric quantification and immunophenotyping of leukemic stem cells in acute myeloid leukemia. Annals of hematology 2012, 91(10):1541-1546.

16. Roshal M, Chien S, Othus M, Wood BL, Fang M, Appelbaum FR, Estey EH, Papayannopoulou T, Becker PS: The proportion of CD34(+)CD38(low or neg) myeloblasts, but not side population frequency, predicts initial response to induction therapy in patients with newly diagnosed acute myeloid leukemia. Leukemia 2013, 27(3):728-731.
17. Witte KE, Ahlers J, Schafer I, Andre M, Kerst G, Scheel-Walter HG, Schwarze CP, Pfeiffer M, Lang P, Handgretinger R et al: High Proportion of Leukemic Stem Cells at Diagnosis Is Correlated with Unfavorable Prognosis in Childhood Acute Myeloid Leukemia. Pediatr Hemat Oncol 2011, 28(2):91-99.
18. Hanekamp D, Denys B, Kaspers GJL, Te Marvelde JG, Schuurhuis GJ, De Haas V, De Moerloose B, de Bont ES, Zwaan CM, de Jong A et al: Leukaemic stem cell load at diagnosis predicts the development of relapse in young acute myeloid leukaemia patients. British journal of haematology 2018, 183(3):512-516.
19. Chavez-Gonzalez A, Dorantes-Acosta E, Moreno-Lorenzana D, Alvarado-Moreno A, Arriaga-Pizano L, Mayani H: Expression of CD90, CD96, CD117, and CD123 on different hematopoietic cell populations from pediatric patients with acute myeloid leukemia. Archives of medical research 2014, 45(4):343-350.
20. Kersten B, Valkering M, Wouters R, van Amerongen R, Hanekamp D, Kwidama Z, Valk P, Ossenkoppele G, Zeijlemaker W, Kaspers G et al: CD45RA, a specific marker for leukaemia stem cell sub-populations in acute myeloid leukaemia. British journal of haematology 2016, 173(2):219-235.
21. Harada N, Okamura S, Kubota A, Shimoda K, Ikematsu W, Kondo S, Harada M, Niho Y: Analysis of acute myeloid leukemia cells by flow cytometry, introducing a new light-scattering classification. Journal of cancer research and clinical oncology 1994, 120(9):553-557.
22. van Rhenen A, Moshaver B, Kelder A, Feller N, Nieuwint AW, Zweegman S, Ossenkoppele GJ, Schuurhuis GJ: Aberrant marker expression patterns on the CD34+CD38- stem cell compartment in acute myeloid leukemia allows to distinguish the malignant from the normal stem cell compartment both at diagnosis and in remission. Leukemia 2007, 21(8):1700-1707.
23. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, Scholten WJ, Snel AN, Veldhuizen D, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. Leukemia 2016, 30(2):439-446.
24. Pabst C, Bergeron A, Lavallee VP, Yeh J, Gendron P, Norddahl GL, Krosl J, Boivin I, Deneault E, Simard J et al: GPR56 identifies primary human acute myeloid leukemia cells with high repopulating potential in vivo. Blood 2016, 127(16)(16):2018-2027.
25. Daga S, Rosenberger A, Quehenberger F, Krisper N, Prietl B, Reinisch A, Zebisch A, Sill H, Wolfler A: High GPR56 surface expression correlates with a leukemic stem cell gene signature in CD34-positive AML. Cancer medicine 2019, 8(4):1771-1778.
26. Eppert K, Takenaka K, Lechman ER, Waldron L, Nilsson B, van Galen P, Metzeler KH, Poeppl A, Ling V, Beyene J et al: Stem cell gene expression programs influence clinical outcome in human leukemia. Nature medicine 2011, 17(9):1086-1093.
27. Aiuti A, Webb IJ, Bleul C, Springer T, Gutierrez-Ramos JC: The chemokine SDF-1 is a chemoattractant for human CD34+ hematopoietic progenitor cells and provides a new mechanism to explain the mobilization of CD34+ progenitors to peripheral blood. The Journal of experimental medicine 1997, 185(1):111-120.
28. Peled A, Kollet O, Ponomaryov T, Petit I, Franitza S, Grabovsky V, Slav MM, Nagler A, Lider O, Alon R et al: The chemokine SDF-1 activates the integrins LFA-1, VLA-4, and VLA-5 on immature human CD34(+) cells: role in transendothelial/stromal migration and engraftment of NOD/SCID mice. Blood 2000, 95(11):3289-3296.
29. Liu T, Liu X, Xiang J, Zou P, Zhou J, Chen Y, Yu D, Li C: [Study on the relationship between the expression of adhesion molecules and the invasiveness of acute myeloid leukemia cells]. Zhonghua xue ye xue za zhi = Zhonghua xueyexue zazhi 1997, 18(1):29-31.

30. Burger JA, Spoo A, Dwenger A, Burger M, Behringer D: CXCR4 chemokine receptors (CD184) and alpha4beta1 integrins mediate spontaneous migration of human CD34+ progenitors and acute myeloid leukaemia cells beneath marrow stromal cells (pseudoemperipolesis). British journal of haematology 2003, 122(4):579-589.
31. Cloos J, Harris JR, Janssen J, Kelder A, Huang F, Sijm G, Vonk M, Snel AN, Scheick JR, Scholten WJ et al: Comprehensive Protocol to Sample and Process Bone Marrow for Measuring Measurable Residual Disease and Leukemic Stem Cells in Acute Myeloid Leukemia. Journal of visualized experiments : JoVE 2018(133).
32. Kalina T, Flores-Montero J, van der Velden VH, Martin-Ayuso M, Bottcher S, Ritgen M, Almeida J, Lhermitte L, Asnafi V, Mendonca A et al: EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. Leukemia 2012, 26(9):1986-2010.
33. Zeijlemaker W, Kelder A, Wouters R, Valk PJ, Witte BI, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: Absence of leukaemic CD34 cells in acute myeloid leukaemia is of high prognostic value: a longstanding controversy deciphered. British journal of haematology 2015, Oct(171(2)):227-238.
34. Goardon N, Marchi E, Atzberger A, Quek L, Schuh A, Soneji S, Woll P, Mead A, Alford KA, Rout R et al: Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer cell 2011, 19(1):138-152.
35. Zeijlemaker W, Kelder A, Cloos J, Schuurhuis GJ: Immunophenotypic Detection of Measurable Residual (Stem Cell) Disease Using LAIP Approach in Acute Myeloid Leukemia. Current protocols in cytometry / editorial board, J Paul Robinson, managing editor [et al] 2019, 91(1):e66.
36. Jin L, Hope KJ, Zhai Q, Smadja-Joffe F, Dick JE: Targeting of CD44 eradicates human acute myeloid leukemic stem cells. Nature medicine 2006, 12(10):1167-1174.
37. van Rhenen A, van Dongen GAMS, Kelder A, Rombouts EJ, Feller N, Moshaver B, Stigter-van Walsum M, Zweegman S, Ossenkoppele GJ, Schuurhuis GJ: The novel AML stem cell-associated antigen CLL-1 aids in discrimination between normal and leukemic stem cells. Blood 2007, 110(7):2659-2666.
38. Hosen N, Park CY, Tatsumi N, Oji Y, Sugiyama H, Gramatzki M, Krensky AM, Weissman IL: CD96 is a leukemic stem cell-specific marker in human acute myeloid leukemia. Proceedings of the National Academy of Sciences of the United States of America 2007, 104(26):11008-11013.
39. Bolouri H, Farrar JE, Triche T, Jr., Ries RE, Lim EL, Alonzo TA, Ma Y, Moore R, Mungall AJ, Marra MA et al: The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nature medicine 2018, Jan(24(1)):103-112.
40. van Solinge TS, Zeijlemaker W, Ossenkoppele GJ, Cloos J, Schuurhuis GJ: The interference of genetic associations in establishing the prognostic value of the immunophenotype in acute myeloid leukemia. Cytometry Part B, Clinical cytometry 2018, 94(1):151-158.
41. Baqai J, Crisan D: Correlation of FLT3 mutations with expression of CD7 in acute myeloid leukemia. Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry 2015, 23(2):104-108.
42. Rausei-Mills V, Chang KL, Gaal KK, Weiss LM, Huang Q: Aberrant expression of CD7 in myeloblasts is highly associated with de novo acute myeloid leukemias with FLT3/ITD mutation. American journal of clinical pathology 2008, 129(4):624-629.
43. Walter RB, Alonzo TA, Gerbing RB, Ho PA, Smith FO, Raimondi SC, Hirsch BA, Gamis AS, Franklin JL, Hurwitz CA et al: High expression of the very late antigen-4 integrin independently predicts reduced risk of relapse and improved outcome in pediatric acute myeloid leukemia: a report from the children's oncology group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2010, 28(17):2831-2838.
44. Tavernier-Tardy E, Cornillon J, Campos L, Flandrin P, Duval A, Nadal N, Guyotat D: Prognostic value of CXCR4 and FAK expression in acute myelogenous leukemia. Leukemia research 2009, 33(6):764-768.

45. Matsunaga T, Takemoto N, Sato T, Takimoto R, Tanaka I, Fujimi A, Akiyama T, Kuroda H, Kawano Y, Kobune M et al: Interaction between leukemic-cell VLA-4 and stromal fibronectin is a decisive factor for minimal residual disease of acute myelogenous leukemia (vol 9, pg 1158, 2003). Nature medicine 2005, 11(5):578-578.
46. Daria D, Kirsten N, Muranyi A, Mulaw M, Ihme S, Kechter A, Hollnagel M, Bullinger L, Dohner K, Dohner H et al: GPR56 contributes to the development of acute myeloid leukemia in mice. Leukemia 2016, 30(8):1734-1741.
47. Saito Y, Kaneda K, Suekane A, Ichihara E, Nakahata S, Yamakawa N, Nagai K, Mizuno N, Kogawa K, Miura I et al: Maintenance of the hematopoietic stem cell pool in bone marrow niches by EVI1-regulated GPR56. Leukemia 2013, 27(8):1637-1649.

Detailed patient characteristics and clinical trial information

De novo pediatric (< 18 years (yrs.)) AML patients diagnosed between October 2015 and March 2019 were consecutively collected. During this period, patients were included in the NOPHO-DBH AML2012 trial (EudraCT 2012-002934-35) unless they showed exclusion criteria i.e. isolated central nervous system, extramedullary leukemia, previous chemo- or radiotherapy, AML secondary to a previous bone marrow failure syndrome, myeloid leukemia in Down syndrome with age <5 or ≥5 yrs. with *GATA1* mutation, acute promyelocytic leukemia (APL), juvenile myelomonocytic leukemia (JMML), myelodysplastic syndrome (MDS), Fanconi anemia or a positive pregnancy test. NOPHO-DBH AML 2012-treated patients received two intensive induction courses with randomisation, followed by risk-adapted consolidation with three courses of conventional chemotherapy for standard risk (SR) patients and allogeneic stem cell transplantation for high-risk (HR) patients. If excluded, patients are referred to as 'otherwise-treated' patients.

De novo pediatric (<18 yrs.) AML patients diagnosed between 2010 and 2013 were retrospectively selected. These patients were treated in the DB-AML01 study according to a protocol as described elsewhere [1]. Selection was based on the number of cells frozen (>10x10^6) and CD34 positivity. At time of diagnosis, mononuclear cells (MNCs) were obtained by Ficoll density gradient (Axis-shield) within 48 h and cryopreserved in 10% dimethylsulfoxide. Samples were thawed by short incubation in a pre-heated water batch (42°C), followed by 30 min incubation at room temperature (RT) in 20 mL RPMI with 20% FCS, 200 µL DNase I (1 mg/mL, grade II bovine pancreas) and 200 µL MgCl2 (1 M) (both by Sigma-Aldrich). After incubation, cells were spinoculated (10 min, 400 rpm) and washed once more with RPMI/20% fetal calf serum (FCS).

Our test cohort is unintentionally biased since not all pedAML patients consecutively diagnosed during the study period could be included due to limited sample availability, and eight pedAML patients were retrospectively selected.

The definition of HR depends on the treatment protocol. In the DB-AML01 study [1], patients were considered as HR if ≥15% blasts persisted after the first induction course and ≤5% blasts after the second course (≥5% blasts after the second course was defined as refractory disease). In the NOPHO-DBH AML2012 study, patients were defined as HR if they achieved CR after two induction courses and had (i) *FLT3*-ITD/*NPM1* WT profiles, (ii) poor response after induction 1 (i.e. ≥15% leukemic cells at day 22 or at any subsequent evaluation prior to the second course) or (iii) intermediate response after the second induction (i.e. 0.1%-4.9% leukemic cells before consolidation) [2].

For estimates of event-free survival (EFS), an event was defined as failure to achieve complete remission (CR) (defined as resistant disease), the occurrence of relapse or the occurrence of death in CR1. EFS was calculated from date of diagnosis to the date of the first event, with failure to achieve CR calculated as an event at t = 0. Overall survival (OS) was calculated from date of diagnosis to the date of last follow-up or time of death due to any cause. Follow-up time was censored at the last follow-up visit if no failure was observed.

Statistical assays, cut-off selection and data analysis

Statistical analyses were performed using IBM SPSS Statistics software (version 25.0.0.2, Inc., Chicago, IL) and GraphPad Prism version 5.04 for Windows (GraphPad Software, La Jolla California USA).
Comparison between two groups of continuous variables was done using the Mann-Whitney test in unpaired samples and the Wilcoxon signed-rank test in paired samples (i.e. bone marrow (BM) versus peripheral blood (PB), *total* versus *LAIP+* LSC load and diagnosis (Dx) versus relapse (R). To compare

categorical variables, the Pearson's Chi-Square test was used if the expected count was higher or equal to five. Fisher's exact test was used if the expected count was below five. The Kaplan-Meier log-rank test was used to evaluate associations with the time before the occurrence of an event or death. Survival curves were generated through the Kaplan-Meier method. Two-sided p-values of less than .05 were considered as significant.

ROC-curve analysis was performed in order to select a cut-off that maximally discriminated patients based on the LSC load within the total CD45, $CD45^{low}/SSC^{low}$ and CD34+ compartment. The LSC frequencies determined in all CD34+ pedAML patients was used as input (n=28). A cut-off equal to 4.78% within the WBC compartment showed the highest sensitivity and specificity, equal to 58.3% and 47.4%, respectively (AUC 0.553). cut-off. A cut-off equal to 17.39% within the CD34+ compartment showed a sensitivity equal to 75.0% and specificity equal to 52.6% (AUC 0.592).

Cut-offs used to dichotomize patient characteristics (immature FAB, age >10 yrs., WBC count $>30x10^9/L$, PB blasts >50% or BM blasts >70%) were based on previous studies [3, 4]. Since CD49d and GPR56 are normally expressed in HSC, COs were also calculated to define overexpression in LSC based on the median values measured in CD34+ pedAML. Median MFI values were highly comparable between $CD34^+/CD38^-$ and $CD34^+/CD38^{dim}$ subpopulations (CD49d: 1512 and 1649; GPR56: 1011 and 757, respectively). Therefore, a threshold was set in between both values to define overexpression in pedAML LSC populations (MFI 1581 for CD49d and MFI 884 for GPR56).

A variable number of events was recorded amongst patients and tubes (Table S3). To minimise this shortcoming, we only proceeded evaluation/statistical calculations if the total number of WBCs exceeded 100 000 events, which we considered sufficient in a diagnostic setting. However, we cannot exclude that LAIP- LSC populations actually harbour minute aberrant expression.

Table S1. Demographics on diagnostic CD34- patients included in the NOPHO-DBH AML2012 study.

Demographics CD34-pedAML at Dx (n=21)	Mean (Range)	
Age, years	8.12 (0 - 16)	
WBC count, x 10^9/L	47.14 (0 - 356)*	
Morphological blast count		
BM, %	60.55 (24 - 96)*	
PB, %	37.25 (0 - 96)*	
	N	%
Time point		
Only diagnosis (Dx)	20	95,2%
Paired Dx-R couples	1	4,8%
Sample		
Only BM	3	14,3%
Only PB	1	4,8%
Paired BM-PB couples	17	81,0%
Gender		
F	9	42,9%
M	12	57,1%
Primary/secondary AML		
Primary	21	100,0%
Secondary	0	0,0%
Treatment protocol		
DB AML-01	0	0,0%
NOPHO-DBH AML2012	19	90,5%
Other	2	9,5%
Occurence of event		
No	14	66,7%
Yes	4	19,0%
Death in CCR	1	
Relapse	3	
Unknown	3	14,3%
Status		
Alive	15	71,4%
Dead	3	14,3%
Death	1	
Death in CCR	1	
Death after relapse	1	
Unknown	3	14,3%

	Mean (Range)	
(continuation)	N	%
WT1 overexpression		
Yes	3	14.3%
No	17	81.0%
Unknown	1	4.8%
Translocation		
Yes	11	52,4%
No	10	47,6%
Core-binding factor leukemia		
Yes	0	0.0%
No	21	100,0%
NPM1		
Mutated	2	9,5%
Wild type	18	85.7%
Unknown	1	4,8%
FLT3		
ITD	2	9,5%
ITD+TKD	**1**	4.8%
Wild type	17	81,0%
Unknown	1	4,8%
CEBPA		
Double mutated	0	0.0%
Wild type	20	95.2%
Unknown	1	4,8%
Karyotype		
Abnormal	10	47,6%
Normal	5	23,8%
Unknown	6	28.6%
CNS involvement		
Yes	5	23,8%
No	13	61,9%
Unknown	3	14,3%
Risk classification		
HR	4	19.0%
SR	11	52.4%
Unknown	6	28.6%
FAB classification		
Immature (M0 - M1)	1	4.8%
Mature (M2-M7)	20	95.2%

Characteristics of *de novo* CD34- pedAML patients at diagnosis. *WT1* overexpression was interpreted in regard to in-house or published (Cilloni et al. 2009) cut-offs. Superscripts indicate one (*) or two (†) missing data. PedAML, pediatric acute myeloid leukemia; LSC, leukemic stem cell; Dx, diagnosis; R, relapse; F, female; M, male; WBC, white blood cell; BM, bone marrow; PB, peripheral blood; *FLT3*, fms-like tyrosine kinase receptor-3; *NPM1*, nucleophosmin; CEBPA, CCAAT/enhancer-binding protein alpha; ITD, internal tandem duplication; WT, wild type; MUT, mutated; FAB, French-British-American; CBF, core-binding factor; *WT1*, Wilms' tumor 1; CNS, central nerve system; HR, high risk; SR, standard risk; CR1, first complete remission.

Table S2. Overview of the used antibodies.

Type of marker		Antigen	Fluorochrome	Clone	Source	Catalogue no.	Reference
Backbone markers		CD34	PerCp-Cy5.5	8G12	BD Bioscience	347222	(Hanekamp et al. 2018)
		CD38	APC-H7	HB7	BD Bioscience	6467896	(Hanekamp et al. 2018)
		CD45	PacO	HI30	Invitrogen	MHCD4530	(Hanekamp et al. 2018)
		CD45	V500	HI30	BD Bioscience	560777	Routine diagnostics
LSC-specific and -associated LAIP markers	All patients	CD45RA	FITC	HI100	BD Bioscience	555488	(Kersten et al. 2016)
		CD56	PE	MY31	BD Bioscience	345810	(Zeijlemaker et al. 2016)
		CD96	PE	NK92.39	eBioscience	12-0969-42	(Chavez-Gonzalez et al. 2014)
		CD7	PE	M-T701	BD Bioscience	555362	(Zeijlemaker et al. 2016)
		CD22	PE	S-HCL1	BD Bioscience	337899	(Zeijlemaker et al. 2016)
		TIM-3	PE	344823	R&D Systems	FAB2365P	(Zeijlemaker et al. 2016)
		CD11b	PE	D12	BD Bioscience	333142	(Zeijlemaker et al. 2016)
		CLL-1	PE	50C1	BD Bioscience	562566	(Zeijlemaker et al. 2016)
		CD123	PE	6H6	eBioscience	12-1239-42	(Chavez-Gonzalez et al. 2014)
		CD49d	PE	9F10	BD Bioscience	556635	(Walter et al. 2010)
		GPR56	PE	CG4	Biolegend	358203	(Pabst et al. 2016)
	NOPHO-DBH AML2012 study and/or limited sample material	CD123	FITC	AC145	MiltenyiBiotec	130-090-897	(Hanekamp et al. 2018)
		CD56*	PE	C5.9	Cytognos	CYT-56PE	(Hanekamp et al. 2018)
		CD2*	PE	39C1.5	Beckman Coulter	A07744	(Hanekamp et al. 2018)
		CD7	APC	124-1D1	eBioscience	17-0079-42	(Hanekamp et al. 2018)
		CD15	FITC	MM7	BD Bioscience	332778	Routine diagnostics
		NG2	PE	7.1	Beckman Coulter	IM3454U/B92429	Routine diagnostics
		CD22	APC	S-HCL1	BD Bioscience	333145	Routine diagnostics

Table S3. Number of events measured per compartment in CD34+ pedAML patients at Dx and R.

		Total events					WBC events					$CD45^{low}/SSC^{low}$ events				
		Median	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**
Dx (n=27)	BM (n=23)	520350	303071	1000000	50000	3961243	466405	221376	801845	33388	3848685	127687	95930	515869	4332	3283453
	PB (n=24)	209152	132350	419386	63850	3661968	170814	114656	332294	36576	3444150	53397	21858	112827	5025	1891245
R (n=9)	BM (n=8)	279097	76250	574517	49962	1052775	209306	49640	532263	29528	884021	28825	6611	415579	2466	446641
	PB (n=6)	87892	50000	155975	33442	859475	58092	40764	88149	28342	725802	5720	2248	36291	1638	138934

		Total $CD34^{+}/CD38^{-}$ events					*Total* $CD34^{+}/CD38^{dim}$ events					*Total* $CD34^{+}/CD38^{-/dim}$ events				
		Median	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**
Dx (n=27)	BM (n=23)	835	197	4363	9	279734	7937	2092	48302	99	966817	16288	2286	56599	127	1246551
	PB (n=24)	662	268	1517	9	42389	4210	2156	17684	118	113647	4986	2455	19201	133	124564
R (n=9)	BM (n=8)	115	4	335	4	1031	571	33	2634	5	4506	671	37	3665	9	4841
	PB (n=6)	51	16	269	15	376	543	104	679	42	3940	625	173	921	75	4209

		LAIP+ CD34+/CD38- events					*LAIP+* CD34+/CD38dim events					*LAIP+* CD34+/CD38-/dim events				
		Median	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**
Dx (n=27)	BM (n=23)	565	135	3887	6	253066	6414	1821	21074	43	747128	7890	1956	24083	55	1000194
	PB (n=24)	554	196	1495	5	30514	3874	2039	17482	43	103216	4789	2235	18977	49	111515
R (n=9)	BM (n=8*)	141	10	818	4	818	500	54	3103	28	3103	629	195	3307	32	3307
	PB (n=6*)	27	16	228	10	228	39	26	3630	7	3630	92	53	3858	17	3858

The tube concordant to the marker with the highest LAIP positivity within LSC was selected for calculation of the number of events. If the LSC compartment showed no aberrant expression for one of the markers (LAIP-), then the tube with the highest WBC events was used. The diagnostic (Dx) patient cohort contained 23 BM and 24 PB samples, with 20 paired BM-PB couples. The relapse (R) cohort contained 8 BM and 6 PB samples, with 6 paired BM-PB couples. *In regard to *LAIP+* $CD34^{+}/CD38^{-}$ events measured at relapse, the total number of patients was reduced from

nine to seven (7 BM samples, 5 PB samples, 5 PB-BM couples), as two patients showed <3 clustered $CD34^+/CD38^-$ events. BM, bone marrow; PB, peripheral blood; WBC, white blood cells; Dx, diagnosis; R, relapse; LCL, lower confidence limit; UCL, upper confidence limit; min, minimum; max, maximum; LSC, leukemic stem cells; LAIP, leukemia-associated immunophenotype.

Table S4. Number of CD34+ events, and the calculated CD34% within WBC (based on Table S3), in CD34+ pedAML patients at Dx and R.

		CD34+ events					**CD34% within WBC compartment**					
		Median	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**P**
Dx (n=27)	BM (n=23)	95848	70973	404150	4223	2690835	**36.6**	**20.3**	**46.4**	1.7	77.3	P>.05 (.197)
	PB (n=24)	25393	16799	73571	3783	1075536	**23.7**	**14.5**	**29.5**	3.2	79.1	
R (n=9)	BM (n=8)	5125	1959	99037	1407	404940	8.9	2.1	22.3	0.4	76.1	P>.05 (0.519)
	PB (n=6)	1501	1027	35069	203	107022	4.1	1.3	14.8	0.5	74.7	

The diagnostic (Dx) patient cohort contained 23 BM and 24 PB samples, with 20 paired BM-PB couples. The relapse (R) cohort contained 8 BM and 6 PB samples, with 6 paired BM-PB couples. BM, bone marrow; PB, peripheral blood; WBC, white blood cells; Dx, diagnosis; R, relapse; LCL, lower confidence limit; UCL, upper confidence limit; min, minimum; max, maximum; LSC, leukemic stem cells; LAIP, leukemia-associated immunophenotype.

Table S5. Overview of *total* and *LAIP+* LSC load determined CD34+ pedAML at Dx and R.

			CD34^{+}/CD38$^{-/dim}$ versus WBC (%)							CD34^{+}/CD38^{-} versus WBC (%)					CD34^{+}/CD38^{dim} versus WBC (%)				
			Median	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**FC**	**P**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**
Dx (n=27)	BM (n=23)	*Total*	4.50	1.15	11.27	0.05	55.2	**2.2**	**<.0001**	0.32	0.09	1.59	0.00	50.6	2.91	1.08	8.33	0.04	25.1
		LAIP+	2.05	1.02	7.90	0.02	53.6			0.11	0.04	1.43	0.00	49.4	1.77	1.00	5.45	0.02	23.1
	PB (n=24)	*Total*	4.30	1.46	8.08	0.04	20.0	**1.1**	**<.0001**	0.31	0.17	0.94	0.01	4.2	4.13	1.31	6.28	0.04	16.1
		LAIP+	3.84	0.89	7.59	0.04	18.1			0.29	0.08	0.62	0.00	3.7	3.79	0.81	5.80	0.04	14.4
R (n=9)	BM (n=8)	*Total*	0.37	0.03	2.78	0.002	7.4	**1.6**	**P<.05**	0.03	0.00	0.64	0.00	2.1	0.31	0.03	2.13	0.00	5.3
		LAIP+	0.24	0.02	2.13	0.001	6.2			0.02	0.00	1.65	0.00	1.6	0.22	0.02	1.62	0.00	4.5
	PB (n=6)	*Total*	0.96	0.20	1.48	0.18	2.0	**2.9**	**P<.05**	0.07	0.04	0.08	0.03	0.5	0.67	0.12	1.45	0.10	1.9
		LAIP+	0.33	0.10	0.73	0.02	1.9			0.04	0.03	0.07	0.01	0.1	0.28	0.04	0.73	0.01	1.9

The *total* and *LAIP+* LSC load was determined based on (A) CD34^{+}CD38$^{-/dim}$, (B) CD34^{+}CD38^{-} and (C) CD34^{+}CD38^{dim} cells as a numerator versus the WBC compartment as denominator. The tube concordant to the marker with the highest LAIP positivity was selected for *total* and *LAIP+* LSC load measurements. If the LSC compartment showed no aberrant expression for one of the markers (LAIP-), then the tube with the highest WBC events was used. The diagnostic (Dx) patient cohort contained 23 BM and 24 PB samples, with 20 paired BM-PB couples. The relapse (R) cohort contained 8 BM and 6 PB samples, with 6 paired BM-PB couples. *In case of *LAIP+* LSC loads determined based on the CD34+/CD38- fraction, the total number of patients was reduced from nine to seven (7 BM samples, 5 PB samples, 5 PB-BM couples), as two patients showed <3 clustered events within the CD34^{+}/CD38^{-} compartment. Significant P-values are indicated in bold. BM, bone marrow; PB, peripheral blood; WBC, white blood cells; Dx, diagnosis; R, relapse; LCL, lower confidence limit; UCL, upper confidence limit; min, minimum; max, maximum; LSC, leukemic stem cells.

Table S6. Comparison of *total* and *LAIP+* LSC load between paired Dx-R samples.

	LSC load	Time point	CD34⁺/CD38$^{-/dim}$ versus WBC (%)				CD34⁺/CD38^{-} versus WBC (%)				CD34⁺/CD38^{dim} versus WBC (%)			
			Median	**Min**	**Max**	**P**	**Median**	**Min**	**Max**	**P**	**Median**	**Min**	**Max**	**P**
Paired Dx-R comparison (BM=4, PB=2)	***Total***	**Dx**	3.46	1.09	11.27	**P<.05**	0.28	0.03	2.04	**P<.05**	2.96	0.82	11.04	P>.05
		R	0.38	0.10	2.78		0.07	0.00	0.64		0.31	0.10	2.13	
	LAIP+	**Dx**	2.51	0.55	7.90	P>.05	0.22	0.03	1.81	P>.05	2.35	0.39	6.09	P>.05
		R	0.31	0.10	2.13		0.07	0.00	0.51		0.28	0.03	1.85	

Four of the six CD34+ patients with Dx-R couples available fulfilled the minimal required number of events at both time points. Paired BM and PB samples were available for 2/4 patients. BM and PB measurements were evaluated as independent data points to increase statistical power which led to a pairwise comparison of six Dx-R samples. The LSC load was calculated based on the CD34⁺CD38$^{-/dim}$, CD34⁺CD38^{-} and CD34⁺CD38^{dim} population versus the WBC compartment as denominator. Two-sided P-values were calculated by the Wilcoxon signed-rank test, and indicated in bold if significant. BM, bone marrow; PB, peripheral blood; WBC, white blood cells; Dx, diagnosis; R, relapse; min, minimum; max, maximum; LSC, leukemic stem cells.

Table S7. Comparison between the *total* and *LAIP+* LSC load in BM versus PB.

		CD34⁺/CD38$^{-/dim}$ versus WBC (%)						CD34⁺/CD38^{-} versus WBC (%)						CD34⁺/CD38^{dim} versus WBC (%)					
		Median	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**P**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**P**	**Median**	**95% LCL median**	**95% UCL median**	**Min**	**Max**	**P**
Total	BM (n=20)	2.34	0.53	9.53	0.05	55.17	>.05	0.33	0.03	1.59	0.00	50.59	>.05	2.00	0.51	7.36	0.04	25.12	>.05
	PB (n=20)	2.93	0.74	6.38	0.03	20.01		0.25	0.08	0.62	0.01	3.90		2.41	0.66	5.87	0.03	16.12	
LAIP+	BM (n=20)	2.03	0.49	7.90	0.02	53.57	>.05	0.19	0.03	1.43	0.00	49.40	>.05	1.76	0.45	5.45	0.02	23.10	>.05
	PB (n=20)	1.22	0.34	4.72	0.01	18.15		0.16	0.02	0.41	0.00	3.72		1.05	0.32	4.30	0.00	14.42	
***LAIP+/ Total* (%)**	**BM**	**86.7%**						**57.6%**						**87.7%**					
	PB	**41.5%**						**63.8%**						**43.6%**					

Paired BM-PB samples were available for 21 Dx CD34+ pedAML patients, but one patient was excluded due to a too low number of events in one of both sample matrixes. Hence, the *total* and the *LAIP+* LSC load was compared between BM and PB in 20 pedAML patients. Two-sided P-values were calculated by the Wilcoxon signed-rank test. No significant differences could be observed. Dx, diagnosis; BM, bone marrow; PB, peripheral blood; WBC, white blood cells; pedAML, pediatric AML; LCL, lower confidence limit; UCL, upper confidence limit; min, minimum; max, maximum; LSC, leukemic stem cells.

Table S8. Correlation of LSC^{high} and LSC^{low} classified patients with molecular and patient characteristics.

Cut-off	Classifi-cation	n	FAB immature		Age >10 yr.		WBC >30x10^9/L		BM blasts>70%		PB blasts>50%		Translocation	
			N	%	N	%	N	%	N	%	N	%	N	%
< or ≥ 17.39% within CD34+	LSC^{low}	**17**	1	5.9	6	35.3	7	41.2	7	43.8	10	58.8	9	**52.9**
	LSC^{high}	**10**	1	10.0	3	30.0	3	33.3	2	22.2	2	20.0	5	**55.6**
< or ≥ 4.78% within WBC	LSC^{low}	**17**	1	5.9	5	29.4	5	**31.3**	6	35.3	7	41.2	8	**50.0**
	LSC^{high}	**10**	1	10.0	4	40.0	5	**50.0**	3	37.5	5	50.0	6	**60.0**

Cut-off	Classifi-cation	n	CBF		*WT1* overexpression		*FLT3*-ITD		Abnormal Karyotype		CNS invasion		HR	
			N	%	N	%	N	%	N	%	N	%	N	%
< or ≥ 17.39% within CD34+	LSC^{low}	**17**	7	**41.2**	10	**62.5**	4	**23.5**	9	**56.3**	1	6.3	3	20.0
	LSC^{high}	**10**	3	**33.3**	8	**88.9**	2	**33.3**	6	**75.0**	1	12.5	1	14.3
< or ≥ 4.78% within WBC	LSC^{low}	**17**	6	37.5	10	**66.7**	3	**18.8**	8	**57.1**	1	7.1	2	14.3
	LSC^{high}	**10**	4	40.0	8	**80.0**	4	**40.0**	7	**70.0**	1	10.0	2	25.0

Twenty-seven out of the 32 CD34+ patients for whom LSC load could be evaluated were dichotomised as LSC^{high} or LSC^{high} according to two different cut-offs determined by ROC curve analysis (Supplemental data). Classification is based on the *LAIP+* LSC load measured in PB (n=24/27), except for 3/27 patients lacking PB for whom BM was used. Unknown data were considered as missing values and excluded for percentage calculation. The Pearson's Chi-Square test (n>5) or Fisher's exact test (n<5) was used was used for statistical comparison (all P>.05). BM, bone marrow; PB, peripheral blood; WBC, white blood cells; pedAML, pediatric AML; LSC, leukemic stem cell; CNS, central nerve system; yrs., year; FLT3, fms-like tyrosine kinase receptor-3; ITD, internal tandem duplication; FAB, French-British-American; CBF, core-binding factor; *WT1*, Wilms' tumor 1; CNS, central nerve system; HR, high risk.

Table S9. Correlation of LSC^{high} and LSC^{low} classified patients with clinical outcome.

				***LAIP+* LSC load ($CD34^+/CD38^{-/dim}$)**		
Denominator	**Cut-off value**	**No. Patients**	**Variables**	**LSC^{low}**	**LSC^{high}**	P
WBC compartment	Low: < 4.78% High: ≥ 4.78%	NOPHO-DBH AML2012 treated patients (n=18)	n (patients)	13	5	
			median (95% CI)	0.71 (0.23-3.00)	8.20 (7.68-18.15)	
			EFS (%)	53.8	60.0	0.54
		NOPHO-DBH AML2012 (n=18), DB-AML01 (n=4) and other (n=1)	n (patients)	15	8	
			median (95% CI)	0.89 (0.34 - 3.00)	8.63 (6.38-18.15)	
			EFS (%)	**53.3**	**37.5**	0.083
CD34+ compartment	Low: < 17.39% High: ≥ 17.39%	NOPHO-DBH AML2012 treated patients (n=18)	n (patients)	15	3	
			median (95% CI)	5.67 (2.37-10.28)	25.81 (22.34-51.48)	
			EFS (%)	46.7	100.0	0.17
		NOPHO-DBH AML2012 (n=18), DB-AML01 (n=4) and other (n=1)	n (patients)	16	7	
			median (95% CI)	5.53 (2.31-9.41)	30.80 (24.23-94.40)	
			EFS (%)	**50.0**	**42.9**	0.22

Table S10. Overview of the number of patients and samples analyzed per LAIP marker.

Marker*		CD34+ pedAML at Dx				CD34+ pedAML at R				Healthy controls			
		patients		sample matrix (BM and PB)		patients		sample matrix (BM and PB)		NBM		CB	
		n	% (total)	n	% (total)	n	% (total)	n	% (total)	n	% (total)	n	% (total)
LSC-specific	**CD45RA**	30	94%	49	92%	7	78%	11	65%	11	100%	13	100%
	CD56	32	100%	53	100%	6	67%	10	59%	8	73%	13	100%
	CD96	29	91%	47	89%	6	67%	10	59%	8	73%	13	100%
	CD7	32	100%	53	100%	6	67%	10	59%	8	73%	13	100%
	CD22	32	100%	53	100%	6	67%	10	59%	8	73%	13	100%
	CD11b	29	91%	47	89%	6	67%	10	59%	8	73%	13	100%
	TIM3	29	91%	47	89%	6	67%	10	59%	8	73%	13	100%
	CLL-1	29	91%	47	89%	5	56%	9	53%	9	82%	13	100%
	CD123	32	100%	53	100%	5	56%	9	53%	9	82%	13	100%
	CD2	23	72%	43	81%	8	89%	15	88%	0	0%	0	0%
	CD15	23	72%	43	81%	7	78%	7	41%	0	0%	0	0%
	NG2	23	72%	43	81%	7	78%	7	41%	0	0%	0	0%
LSC-associated	**CD49d**	27	84%	44	83%	6	67%	9	53%	8	73%	13	100%
	GPR56	20	63%	30	57%	5	56%	8	47%	8	73%	7	54%
Total		**32**		**53**		**9**		**17**		**11**		**13**	

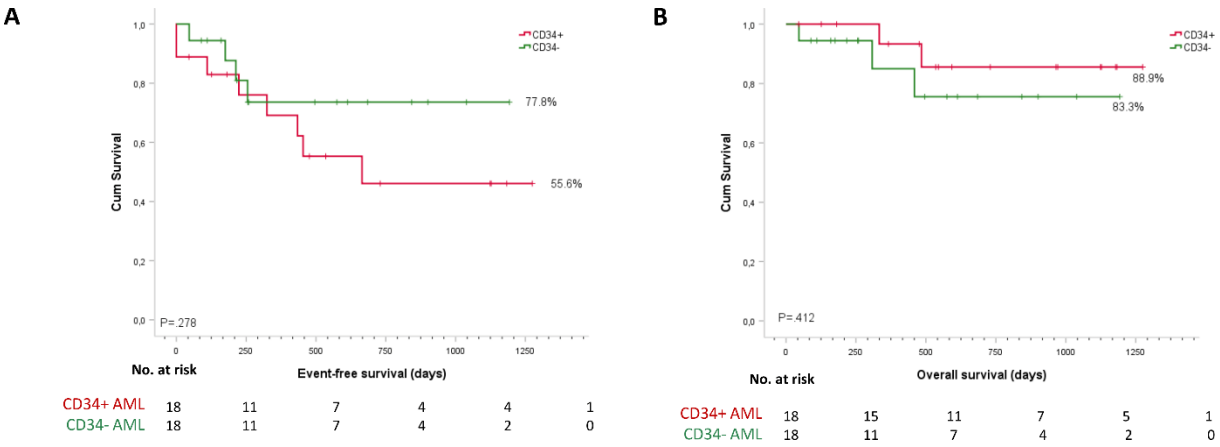
*The number of evaluated patients differs per marker due to consecutive inclusion of the patients, sample availability, and the requirement to have at least three clustered $CD34^{+}/CD38^{-/dim}$ events for each marker individually. BM, bone marrow; PB, peripheral blood; WBC, white blood cells; pedAML, pediatric AML; LSC, leukemic stem cell, Dx, diagnosis; R, relapse; NBM, normal bone marrow; CB, cord blood.

Table S11. FSC, SSC, CD34 and CD45 ratios in (A) LAIP+ pedAML, containing both LAIP+ and LAIP- subpopulations, and (B) healthy controls.

Marker		A. LAIP+ pedAML												B. Healthy controls (LAIP-)			
		FSC ratio			SSC ratio			CD34 ratio			CD45 ratio			FSC ratio	SSC ratio	CD34 ratio	CD45 ratio
		LAIP-	LAIP+	P	LAIP-	LAIP+	P	LAIP-	LAIP+	P	LAIP-	LAIP+	P				
CD45RA	median	1,51	1,70	**<.001**	1,68	2,04	**<.001**	178,68	303,88	**<.05**	0,18	0,25	**<.001**	1,44	1,50	225,22	0,20
	range	0.50-2.87	1.18-3.78		0.90-3.13	1.10-3.14		45-7870	51-7940		0.06-0.43	0.08-0.49		1.21-1.67	1.20-2.04	33-641	0.11-0.34
CD56	median	1,41	1,52	>.05	1,99	2,03	**<.01**	277,07	416,00	**<.05**	0,21	0,19	>.05	1,41	1,24	308,92	0,20
	range	0.94-2.12	0.96-2.38		1.53-2.30	1.69-2.65		0-8298	0-13561		0.06-0.46	0.06-0.53		1.36-1.52	1.19-1.66	151-1553	0.11-0.33
CD96	median	1,24	1,55	NC	1,99	2,54	NC	331,29	472,81	NC	0,19	0,20	NC	1,41	1,24	317,11	0,20
	range	1.03-1.38	1.05-1.59		1.25-2.08	1.26-2.57		43-1838	25-2350		0.15-0.24	0.15-0.44		1.36-1.58	1.19-1.66	151-1553	0.11-0.33
CD7	median	1,58	1,65	>.05	1,81	1,97	**<.01**	113,49	132,72	>.05	0,24	0,30	**<.001**	1,42	1,34	287,60	0,20
	range	0.35-2.91	0.33-2.89		0.97-3.61	0.75-3.42		0-11859	0-9942		0.04-1.36	0.07-1.35		1.31-1.74	1.18-1.93	0-6348	0.14-0.34
CD22	median	1,60	1,54	>.05	1,87	1,91	>.05	169,52	189,77	NC	0,25	0,25	NC	1,36	1,38	287,60	0,20
	range	0.35-2.26	0.33-2.59		0.80-3.34	0.87-3.71		0-1445	0-1668		0.04-0.40	0.11-0.60		1.31-1.74	1.18-1.93	0-6347.57	0.14-0.32
CD11b	median	1,20	1,36	NC	1,24	1,45	NC	91,74	100,89	>.05	0,17	0,22	**<.05**	1,35	1,38	397,69	0,20
	range	0.42-1.37	0.49-1.75		1.01-1.76	1.06-2.45		4-494	6-692		0.06-0.20	0.11-0.26		1.31-1.55	1.21-1.59	161-653	0.14-0.33
TIM3	median	1,58	1,77	>.05	1,76	2,11	**<.05**	298,95	325,12	**<.05**	0,22	0,29	**<.001**	1,34	1,33	378,15	0,20
	range	0.42-2.81	0.49-2.21		1.01-3.62	1.06-3.10		4-3888	6-2655		0.06-0.47	0.11-0.54		1.31-1.55	1.21-1.59	161-653	0.14-0.33
CLL-1	median	1,34	1,61	**<.001**	1,60	1,92	**<.001**	175,40	233,23	>.05	0,20	0,25	**<.001**	1,37	1,29	276,82	0,21
	range	0.43-2.82	0.45-2.58		0.92-2.78	0.62-3.28		3.24-1989	6-1137		0.06-0.44	0.07-0.54		0.59-1.69	0.14-2.10	150-442	0.12-0.55
CD123	median	1,35	1,62	**<.001**	1,55	1,97	**<.001**	190,09	252,58	**<.01**	0,19	0,25	**<.001**	1,35	1,36	307,46	0,21
	range	0.43-2.82	0.45-2.94		0.92-3.55	0.82-3.87		0-12096	0-15234		0.06-0.44	0.06-0.67		0.95-1.69	1.11-2.29	133-509	0.12-0.31
CD15	median	1,59	1,75	**<.05**	2,01	2,27	>.05	353,59	392,12	>.05	0,24	0,30	**<.01**	/	/	/	/
	range	0.86-2.54	0.90-3.06		1.29-3.95	1.62-3.41		49-1452	76-3015		0.13-0.29	0.17-0.76		/	/	/	/
NG2	median	1,80	1,74	NC	2,44	2,34	NC	1,75	1,81	NC	2,35	2,59	NC	/	/	/	/
	range	1.65-2.32	1.71-2.99		1.79-3.89	2.13-4.17		1.62-2.06	1.71-2.60		1.84-3.45	2.14-4.56		/	/	/	/

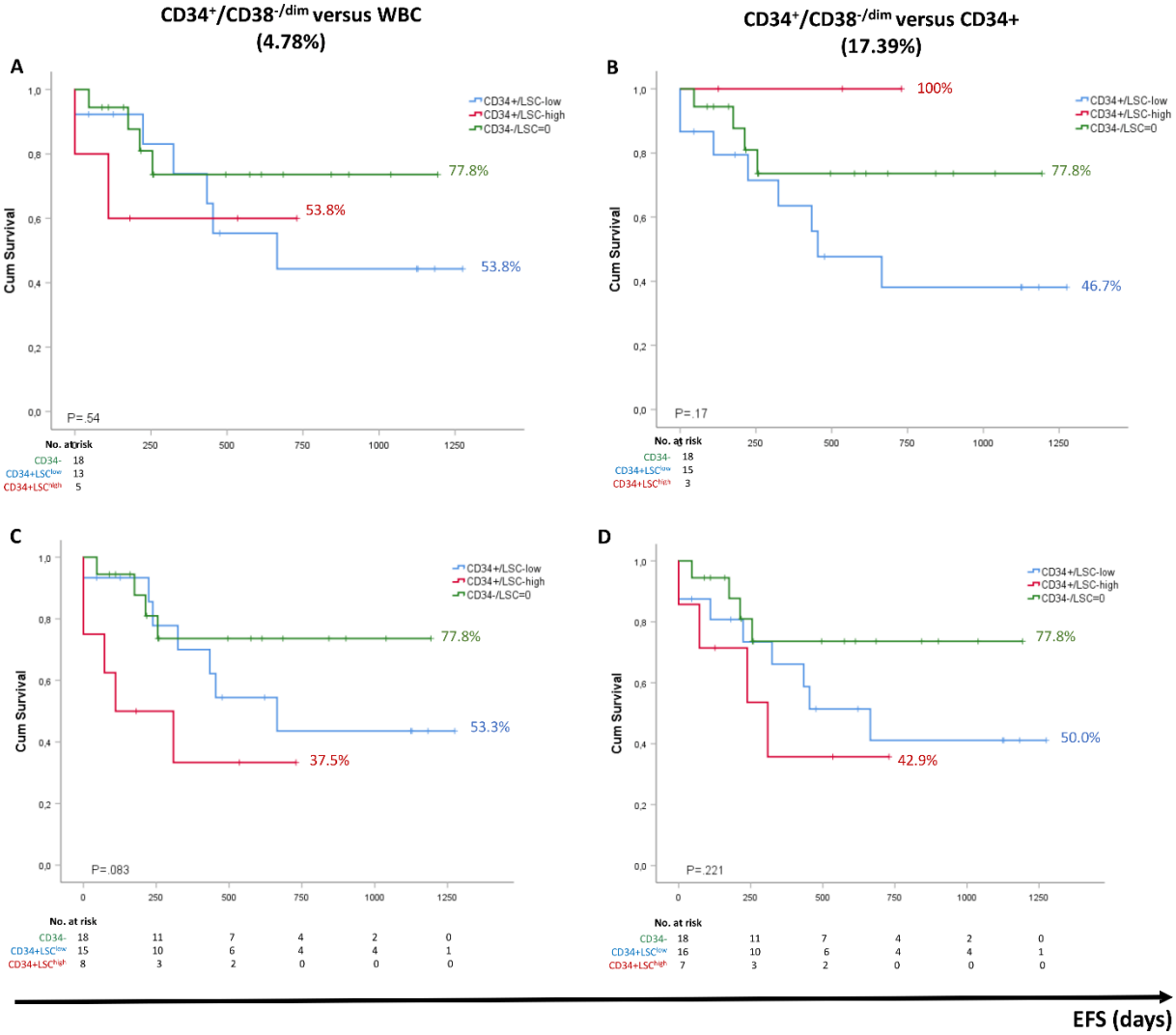
FSC/SSC/CD34/CD45 ratios were calculated for LAIP+ and/or LAIP- LSC subpopulations as a ratio versus the median values measured in lymphocytes. For *LAIP+* pedAML patients (A), ratios were statistically compared between LAIP+ and LAIP- subpopulations (Wilcoxon signed-rank test). Significant P-values are indicated in bold. The number of LAIP+ LSC subpopulations was too low for CD96 (n=4), CD11b (n=5) and NG2 (n=3) in order to perform statistical analysis. Healthy controls (B) only harbored LAIP- subpopulations, and were not evaluated for CD15 and NG2. Negative FI values were set equal to zero. NC, not calculated; pedAML, pediatric AML; LSC, leukemic stem cell; LAIP, leukemia-associated immunophenotype; FSC, forward scatter; sideward scatter, SSC

Fig. S1. EFS and OS in CD34+ and CD34- pedAML patients treated in the NOPHO-DBH AML2012 study.



Event-free survival (EFS) and overall survival (OS) were calculated for CD34+ (n=18) and CD34- (n=18) patients treated in the NOPHO-DBH AML2012 trial. Definitions of an event and survival are explained in supplemental data. P-values were calculated by the Kaplan-Meier logrank test. CD34+ patients showed a trend towards lower EFS compared to CD34- patients (Fig. S1A, P>.05), while OS was highly comparable (Fig. S1B, P>.05).

Fig. S2. EFS in CD34+LSC^{high}, CD34+LSC^{low} and CD34- pedAML patients.



Eighteen CD34- AML were treated in the NOPHO-DBH AML2012 study and categorized as LSC=0, in agreement with previous work [5]. CD34+ AML were dichotomized into LSC^{high} and LSC^{low} groups based on the *LAIP+* LSC load determined in PB (BM if not available) according to two different ROC-based cut-offs. The *LAIP+* LSC load within the WBC compartment was evaluated versus a cut-off of 4.78% (A and C), and the *LAIP+* LSC load within CD34+ cells was evaluated versus a cut-off of 17.39% (B and D). P-values were calculated by the Kaplan-Meier logrank test. EFS, event-free survival; BM, bone marrow; PB, peripheral blood; WBC, white blood cells; pedAML, pediatric AML; LSC, leukemic stem cell; CO, cut-off.

(A-B) Evaluation of CD34+ patients treated according to the NOPHO-DBH AML2012 protocol (n=18).
(C-D) Evaluation of CD34+ patients treated according to the NOPHO-DBH AML2012 protocol (n=18/23), next to 4/23 patients treated in the DB-AML01 study and 1/23 otherwise treated patient. Two of these five non NPHO-DBH AML2012-treated patients experiences relapse, one showed resistant disease and one patient died during first complete remission.

Supplemental references

1. De Moerloose B, Reedijk A, de Bock GH, Lammens T, de Haas V, Denys B, Dedeken L, van den Heuvel-Eibrink MM, Te Loo M, Uyttebroeck A et al: Response-guided chemotherapy for pediatric acute myeloid leukemia without hematopoietic stem cell transplantation in first complete remission: Results from protocol DB AML-01. Pediatric blood & cancer 2019:e27605.
2. Chair) JAS: Research study for treatment of children and adolescents with acute myeloid leukaemia 0-18 years. . NOPHO-DBH AML 2012 protocol v21 2013, EUdract number 2012-002934-35.
3. Goardon N, Marchi E, Atzberger A, Quek L, Schuh A, Soneji S, Woll P, Mead A, Alford KA, Rout R et al: Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer cell 2011, 19(1):138-152.
4. Elsayed AH, Rafiee R, Cao X, Raimondi S, Downing JR, Ribeiro R, Fan Y, Gruber TA, Baker S, Klco J et al: A six-gene leukemic stem cell score identifies high risk pediatric acute myeloid leukemia. Leukemia 2019, doi: 10.1038/s41375-019-0604-8. [Epub ahead of print].
5. Zeijlemaker W, Grob T, Meijer R, Hanekamp D, Kelder A, Carbaat-Ham JC, Oussoren-Brockhoff YJM, Snel AN, Veldhuizen D, Scholten WJ et al: CD34(+)CD38(-) leukemic stem cell frequency to predict outcome in acute myeloid leukemia. Leukemia 2018, 33:1102–1112.

4

CHAPTER IV: Results:

Deciphering molecular heterogeneity in pediatric AML using a cancer vs normal transcriptomic approach.

Manuscript ready for submission.

Abstract awarded with an ASH Abstract Achievement Award on the 61st ASH Annual meeting.

Abstract

Background.
Still 30-40% of pediatric acute myeloid leukemia (pedAML) patients with initial good remission rates experience relapse. Delineation of the transcriptomic profile of leukemic subpopulations could aid in a better understanding of pedAML biology and provide novel biomarkers.

Materials and Methods.
Micro-array profiling and qPCR validation was performed on sorted leukemic stem cells (LSC) and leukemic blasts (L-blast) from 24 and 25 pedAML patients, respectively. Transcript expression was compared with expression measured in normal hematopoietic stem cells (HSC) and control blasts (C-blast) sorted from 20 and 19 healthy controls, respectively. Gene set enrichment analysis of the differentially expressed genes (DEGs) with unsupervised clustering was performed to identify relevant gene set enrichment signatures, and functional protein associations were identified by STRING analysis.

Results.
Highly significantly overexpressed genes in LSC and L-blast were identified, with the vast majority not studied in AML. *CDKN1A*, *CFP* and *CFD* (LSC targets) and *HOMER3*, *CTSA* and *GADD45B* (L-blast targets) represent potentially interesting biomarkers or therapeutic targets based on their role in tumorigenesis, immune regulation, apoptosis, adhesion, or signaling. Eleven LSC downregulated targets were identified that potentially qualify as novel tumor suppressor genes, with *MYCT1*, *PBX1* and *PTPRD* of highest interest. Inflammatory and immune dysregulation were critical biological networks perturbed in LSC of pedAML patients, whereas L-blast showed metabolic dysregulated profiles, compared to their normal counterparts.

Conclusion.
Our study illustrates the power of taking into account cell population heterogeneity in pedAML expression profiling and reveals a set of relevant targets for functional studies and targeted therapy.

Introduction

Pediatric acute myeloid leukemia (pedAML) is a rare haematological disease that accounts for 20% of all pediatric leukemias [1]. Cytogenetic risk stratification combined with response-guided therapeutic decisions considerably improved prognostication [2, 3]. However, still 30-40% of the good responders experience relapse [2]. Since initial good clinical remission rates are not consistently translated into cure, children nowadays show 50% 5-year (yr.) event-free survival (EFS) and 70% overall survival (OS) rates [2]. During the past decade, ample evidence showed that relapse is associated with a high leukemic stem cell (LSC) load at diagnosis and LSC persistence during apparent remission [4-8]. However, persistent relapse in pedAML patients with low diagnostic LSC loads emphasizes the need for a more profound molecular and phenotypic characterization of LSC [7].

Hitherto, most pedAML gene expression profiles (GEPs) were established in bulk leukemic samples [9-15], not taking into account cellular heterogeneity, and thus fail to identify critical LSC-specific genes and pathways. These last years, LSC gene signatures were developed for adult AML patients [16-23]. Interestingly, the LSC17 signature by Ng. et al. [16] also held a prognostic value in pedAML [24, 25]. Moreover, it was recently used to develop a pediatric-specific LSC6 score, able to identify high-risk (HR) pedAML patients [26]. However, current LSC signatures contain genes that are also expressed in hematopoietic stem cells (HSC) and lack the inclusion of downregulated targets [27]. Adding *PCDH17*, a LSC-specific downregulated tumor suppressor gene (TSG), to the LSC17 score enabled an improved risk stratification [28]. Hence, the identification of novel differentially expressed genes (DEGs) in pedAML leukemic subpopulations could aid in a better understanding of the molecular biology and provide novel biomarkers for risk stratification, follow-up and targeted therapy.

Here, we describe novel LSC and leukemic blast (L-blast) targets in pedAML discovered by micro-array profiling followed by quantitative PCR (qPCR) validation according to a cancer versus normal (CvN) approach. We reveal pathways that are deregulated in LSC and L-blast, which have not yet been addressed in children, and aid in a further understanding of the pedAML molecular biology.

Materials and Methods

Patients and controls

Bone marrow (BM) and/or peripheral blood (PB) from a total of 28 pedAML patients was selected based on cell availability (>50x10^6 after routine work-up) and CD34 positivity (≥1%). For 21/28 patients, both LSC and L-blast fractions were available, whereas for 3/28 and 4/28 patients, only LSCs or L-blasts, respectively, could be evaluated. Demographics of patients used for LSC (n=24) and L-blast (n=25) expression evaluation are shown in Table 1, details on treatment protocols and outcome definitions are described in Supplementary. In addition, samples were collected from 20 healthy controls, with HSC and control blasts (C-blast) fractions available in 20/20 and 19/20 controls, respectively. Pediatric normal bone marrow (NBM, n=9, 12-18 yr.) was collected from posterior iliac crest during scoliosis surgery. Cord blood (CB, n=11) was obtained after full-term delivery. All subjects gave their informed

consent for inclusion before participation. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the University Hospital of Ghent (EC2015-1443 and EC2019-0294).

Cell sorting

Mononuclear cells were isolated by Ficoll density gradient (Axis-shield), complemented by CD34 isolation if CD34 expression was <50% (CD34 MicroBead Kit, Milteny). Cell-sorting was performed to isolate CD34+/CD38- and CD34+/CD38+ cells from patients and controls, defined as LSC and HSC, and L-blast and C-blast, respectively. Availability of both PB and BM in 11/24 and 10/25 patients evaluated for LSC and L-blast expression, respectively, yielded a total of 35 fractions in each cohort. In addition, the LSC compartment of seven patients could be phenotypically subdivided into two fractions based on different expression of CD45RA (n=3), CLL-1 (n=2), CD123 (n=1) or GPR56 (n=1), yielding a total of 42 LSC fractions. Staining and sorting strategy are described in Supplementary, monoclonal antibodies (mAb) are described in Table S1. Sorted cells were collected in RPMI supplemented with 50% FCS and a post-sort purity of >90% was reached. Sorted cells were spun down (10 min, 3000 rpm, 4° C) and resuspended in 700 µL TRIzol for RNA extraction.

Table 1. Demographics of the total pedAML patient cohort evaluated by qPCR.

		LSC		**L-blast**	
No. patients (fractions)		**24 (42)**		**25 (35)**	
		Median (Range)		Median (Range)	
Age, yrs.		8 (1-17)		9 (1-17)	
WBC count, x 10^9/L		24.3 (3.1-336)*		25.5 (3.4-336)*	
Morphological blast count					
	BM, %	65 (27-96)†		70 (31-96)†	
	PB, %	48.5 (1-95)		51 (1-95)	
		N	%	N	%
Time point					
	Diagnosis (Dx)	22	92%	24	96%
	Relapse (R)	2	8%	1	4%
Treatment protocol					
	DB-AML01	5	21%	9	36%
	NOPHO-DBH AML2012	14	58%	13	52%
	Other	5	21%	3	12%
Gender					
	F	11	46%	12	48%

	M	13	54%	13	52%
Sample					
	BM and PB	11	46%	10	40%
	Only BM	7	29%	9	36%
	Only PB	6	25%	6	24%
LSC phenotype couples		7	29%	NA	NA
WT1 overexpression		15	63%	17	68%
Fusion transcript		14	58%	14	56%
CBF leukemia		9	38%	10	40%
	AML1-ETO + *C-KIT*^{WT}	1	4%	1	4%
	AML1-ETO + *C-KIT*^{MUT}	2	8%	2	8%
	AML1-ETO + *C-KIT*ND	2	8%	3	12%
	CBFB-MYH11	4	17%	4	16%
NPM1					
	MUT	0	0%	0	0%
	WT	23	96%	24	96%
	Unknown	1	4%	1	4%
FLT3					
	ITD	8	33%	10	40%
	WT	16	67%	15	60%
CEBPA					
	Single MUT	0	0%	0	0%
	Double MUT	1	4%	1	4%
	WT	22	92%	23	92%
	Unknown	1	4%	1	4%
Karyotype					
	Normal	6	25%	6	24%
	Abnormal	12	50%	15	60%
	Unknown	6	25%	4	16%
CNS involvement					
	Yes	3	13%	4	16%
	No	16	67%	18	72%
	Unknown	5	21%	3	12%
Risk classification					
	SR	14	58%	16	64%
	HR	4	17%	5	20%

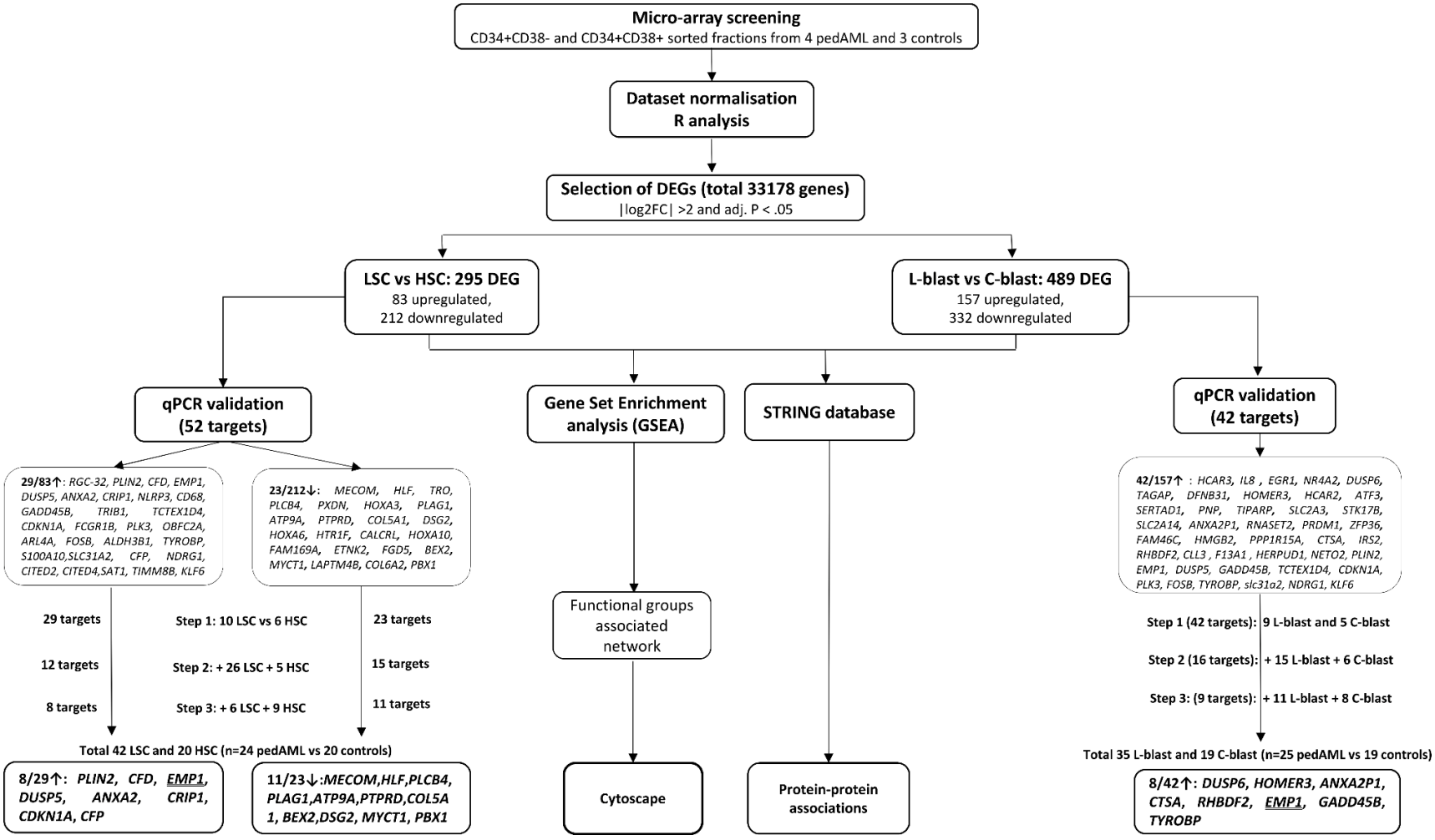
Unknown	6	25%	4	16%
FAB classification				
M0	0	0%	1	4%
M1	3	13%	3	12%
M2	8	33%	8	32%
M3	1	4%	1	4%
M4	5	21%	6	24%
M5	6	25%	5	20%
M7	1	4%	1	4%

Characteristics of pedAML patients used for sorting and qPCR analysis of the most significant DEGs identified between LSC and HSC, and between L-blasts and C-blasts. Superscripts indicate one (*) or two (†) missing data. PedAML, pediatric acute myeloid leukemia; yr., years LSC, leukemic stem cell; L-blast, leukemic blast; F, female; M, male; WBC, white blood cell; BM, bone marrow; PB, peripheral blood; *FLT3*, fms-like tyrosine kinase receptor-3; *NPM1*, nucleophosmin; *CEBPA*, CCAAT/enhancer-binding protein alpha; ITD, internal tandem duplication; WT, wild type; MUT, mutated; FAB, French-British-American; CBF, core-binding factor; *WT1*, Wilms' tumor 1; CNS, central nerve system.

Micro-array downstream analyses

Micro-array analysis was performed on LSC and L-blast sorted from 3/24 and 4/25 pedAML patients (Table S2), respectively, next to 2 HSC and 3 C-blast fractions sorted from CB. Technical details are described in Supplementary. DEGs were identified based on |log2FC|>2 and adjusted P-values (adj. P)<.05. Functional networks between protein–protein associations encoded by DEGs were identified by STRING at a high evidence level [29]. Only KEGG annotated pathways were derived from significant pathway analysis. Gene set enrichment analysis (GSEA) was performed by combining independent omics datasets trough pathway enrichment meta-analysis in order to obtain gene set enrichment signatures [30]. Unsupervised organization and visualization of enriched gene sets was performed in Cytoscape [31]. At least two clustered gene sets based on P < .05, false discovery rate (FDR) < .25 and Jaccard overlap combined index 0.375 were required for node visualization. Selection of DEGs for qPCR validation was based on the magnitude and significance of differential expression and feasibility of primer development. Fig. 1. schematically summarizes all data processing steps. In addition, we re-analyzed the publicly available GSE 17054 micro-array dataset [17], containing GEPs of nine LSC of adult AML patients and four HSC from healthy adults. The top significant DEGs and (anti-)correlated gene sets for each comparison separately are discussed below, with full lists shown in Tables S4-S13.

Fig. 1. Experimental setup and data processing steps of micro-array analysis.



Step-by-step workflow illustrating the experimental and data processing steps pursued to filter out a selection of highly significantly DEGs in leukemic subpopulations starting from a micro-array profiling dataset. PedAML, pediatric acute myeloid leukemia; CB, cord blood; FC, fold change; P, P-value; DEGs, differentially expressed genes; LSC, leukemic stem cell; HSC, hematopoietic stem cell; L-blast, leukemic blast; C-blast, control blast.

Real-time quantitative PCR

qPCR was performed for 94 DEGs using a stepwise approach; available cell fractions were subdivided into three cohorts based on sample availability, while respecting a balanced distribution of the genetic variations, and measured by qPCR in three steps. Sorted cell fractions with only limited material were reserved for the most significantly DEGs, evaluated in the third step, hereby avoiding paucity of sample material to present as an issue. For the 52 DEGs identified between LSC and HSC (29 upregulated and 23 downregulated), 10 LSC and 6 HSC fractions were evaluated in the first step, followed by 26 and 5 additional LSC and HSC fractions in the second step, and another 6 LSC and 9 HSC fractions in the third step (total of 42 LSCs and 20 HSCs). For the 42 upregulated DEGs identified in L-blast versus C-blast, 9 L-blast and 5 C-blast were evaluated in the first step, followed by 15 L-blast and 6 C-blast fractions in the second step, and another 11 L-blast and 8 C-blast fractions in the third step (total of 35 L-blasts and 19 C-blasts).

Technical and analytical details of qPCR experiments are described in Supplementary, primers are shown in Table S3. Data analysis was performed according to state-of-the-art methods [30]. Ct values were corrected for primer efficiency and expressed as relative quantities. Normalized relative quantities (NRQ) were calculated against the expression of housekeeping genes *GAPD*, *HPRT1* and *TBP*. To allow inter-run comparison, calibrated NRQ values (CNRQ) were generated by taking into account the expression of an inter-run calibrator. Target-specific cut-offs for overexpression were calculated based on the average expression plus two standard deviations measured in the respective normal counterparts.

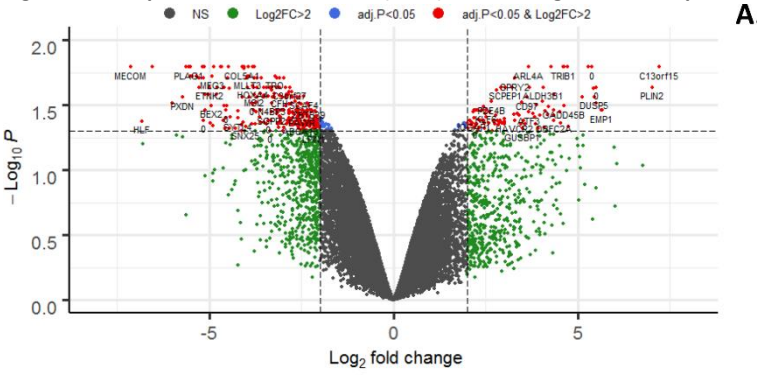
Results

Immune dysregulation separates LSC in pedAML from HSC

The expression of 295 targets significantly differed between LSC and HSC, with 83 targets up- and 212 downregulated in LSC. The top 50 ranked DEGs is shown in Fig. 2A. Well-known oncogenes were present amongst the highest LSC overexpressed targets (e.g. *CFD*, *ANXA2*, *NLRP3*) next to genes with yet undefined roles in AML (e.g. *PLIN2*, *CRIP1*). The top 10 most downregulated genes also contained targets for which no role was yet reported in pedAML (e.g. *ATP9A*, *PLCB4*, *COL5A1*).

Analysis of functional protein associations (STRING) showed upregulated pathways in LSC related to (breast) cancer, osteoclast differentiation and apoptosis, whereas transcriptional misregulation, Th17 cell differentiation, Rap1/MAPK signaling were downregulated (Table 2). Myeloid cell activation networks involved in immune response were enriched in LSC (FDR 4.6e-6), whereas networks related to stimuli responses, signaling and cell communication were suppressed (FDR 5.9e-3). From a total of 3650 signatures available through GSEA, 240 and 18 gene sets were significantly enriched or suppressed in LSC, respectively. The top 10 LSC-enriched gene sets involved LSC signatures, inflammatory response, apoptosis, immune suppression and adipogenesis, whereas HSC-signatures were anti-correlated (Fig. 2B). Unsupervised visualization (Cytoscape) identified LSC-upregulated pathways related to abnormal cell division, quiescence, autoimmune regulation and environmental stress, whilst gene sets involved in normal quiescence and cell death signaling were downregulated (Fig. S1A). Altogether, these data suggest that dysregulation of the immune system contributes to the leukemic transformation of stem cells in pedAML.

Fig. 2. Transcriptional differences and (anti-)correlated gene set and pathways between LSC and HSC.



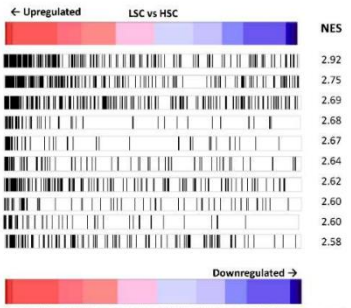
A.

Top 50 LSC upregulated targets		
Gene Symbol	log2FC	adj.P.Val
RGC-32	7.20	0.016
PLIN2	7.02	0.023
CFD	5.65	0.035
EMP1	5.64	0.035
ANXA2	5.50	0.031
DUSP5	5.44	0.027
CRIP1	5.28	0.016
NLRP3	5.13	0.028
CD68	4.75	0.046
42795	4.64	0.016
GADD45B	4.63	0.032
TRIB1	4.62	0.016
TCTEX1D4	4.54	0.043
CDKN1A	4.51	0.036
FCGR1B	4.36	0.038
ANXA2P1	4.35	0.026
PLK3	4.34	0.035
OBFC2A	4.34	0.042
ARL4A	4.27	0.016
SLED1	4.20	0.033
FOSB	4.18	0.037
ALDH3B1	4.08	0.023
ANXA1	4.05	0.030
TYROBP	3.96	0.038
S100A10	3.93	0.048
SLC31A2	3.87	0.032
CFP	3.85	0.044
ANXA2P3	3.83	0.039
NDRG1	3.69	0.023
ATF3	3.66	0.036
CD97	3.63	0.028
CRNDE	3.57	0.044
NAMPT	3.52	0.042
GUSBP1	3.50	0.048
CITED4	3.47	0.030
TOB1	3.45	0.044
SAT1	3.41	0.038
DOK2	3.37	0.037
HAVCR2	3.30	0.042
SPRY2	3.29	0.020
DUSP6	3.28	0.038
STK17B	3.22	0.023
SCPEP1	3.03	0.023
AEN	3.01	0.042
LOC439911	3.01	0.043
EMILIN2	3.00	0.041
SLC2A14	3.00	0.044
OGFRL1	2.94	0.043
TUBB4B	2.90	0.042
TUBB6	2.86	0.043

Top 50 LSC downregulated targets		
Gene Symbol	log2FC	adj.P.Val
MECOM	-7.14	0.016
HLF	-6.82	0.042
TRO	-6.53	0.016
PLCB4	-6.01	0.031
PXDN	-5.72	0.028
HOXA3	-5.62	0.016
PLAG1	-5.56	0.016
ATP9A	-5.51	0.019
PTPRD	-5.50	0.016
COL5A1	-5.36	0.019
DSG2	-5.29	0.023
HOXA6	-5.28	0.020
HTR1F	-5.16	0.020
CALCRL	-5.16	0.020
NKAIN2	-5.13	0.016
HOXA10	-5.12	0.026
FAM169A	-5.02	0.026
ETNK2	-5.00	0.023
FGD5	-4.98	0.043
BEX2	-4.94	0.032
MEG3	-4.92	0.019
MYCT1	-4.90	0.046
LAPTM4B	-4.87	0.016
LOC100132526	-4.82	0.026
ZNF204P	-4.68	0.016
COL6A2	-4.68	0.016
WHAMMP3	-4.62	0.023
PBX1	-4.61	0.042
GPR56	-4.60	0.032
TMEM163	-4.58	0.023
LRRC16A	-4.58	0.048
CALN1	-4.55	0.040
BACH2	-4.52	0.032
SLC22A15	-4.50	0.020
RAPGEF2	-4.47	0.016
OPTN	-4.45	0.023
TSPYL5	-4.45	0.048
HBG1	-4.43	0.039
AIG1	-4.28	0.048
SYTL4	-4.22	0.040
BTBD11	-4.22	0.035
ANKS6	-4.13	0.045
HOPX	-4.12	0.050
SCML4	-4.11	0.027
LIN28B	-4.09	0.043
FAM3C	-4.05	0.017
SNX25	-4.02	0.047
GATA3	-3.98	0.020
MLLT3	-3.97	0.019
HOXA9	-3.92	0.019

B.

N	Gene set	CE genes	NES
1	Enriched genes in dental carious lesions (McLachlan et al.)	98/219	2.92
2	HSC downregulated (Jaatinen et al.)	97/201	2.75
3	LSC enriched (Gal et al.)	83/219	2.69
4	Upregulated genes in CD4+ T helper cells in MF (Hahtola et al.)	35/56	2.68
5	Enriched CBFB-MYH11 genes in AML (Ross et al.)	26/47	2.67
6	Enriched inflammatory response and apoptosis (Galindo et al.)	32/75	2.64
7	Enriched myeloid development (Brown et al.)	64/143	2.62
8	Enriched immunosuppressive targets (Rhein et al.)	31/69	2.60
9	Enriched adipogenesis (Burton et al.)	9/46	2.60
10	Downregulated dendritic cell maturation (Lenaour et al.)	40/118	2.58
1	HSC enriched (Jaatinen et al.)	146/293	-3.05
2	HSC enriched (Eppert et al.)	65/112	-2.81
3	HPC enriched (Haddad et al.)	43/76	-2.74
4	HSC enriched (Eppert et al.)	28/41	-2.72
5	HSC enriched (Georgantas et al.)	28/67	-2.53
6	LSC enriched (Gentles et al.)	19/27	-2.15
7	miR-21 deregulated targets in glioma (Gabriely et al.)	129/278	-2.15
8	Genes enriched in cytoplasmic vs nuclear NPM1 (Alcalay et al.)	50/126	-2.12
9	AML upregulated genes (Valk et al.)	13/32	-2.10
10	FLT3-ITD predicting genes (Valk et al.)	18/39	-2.09



(A) Visualization of DEGs identified between LSC (n=3) and HSC (n=2). Genes are plotted in a volcano plot as log2FC values against -Log10 adj. P-values. Thresholds |log2FC| > 2 and -log10 adj. P < .05 are shown as dashed lines. Genes selected as significantly different are highlighted in red. The top 50-ranked downregulated genes (left) and upregulated genes (right), annotated with gene symbols, are sorted by log2FC values. **(B)** Top 10 most correlated (top) and anti-correlated (bottom) gene sets identified through GSEA. The number of concordantly expressed (CE) genes/total genes and normalized enrichment scores (NES) is shown for each gene set individually. FC, fold change; DEGs, differentially expressed genes; LSC, leukemic stem cell; HSC, hematopoietic stem cell.

As pediatric and adult AML represent two genetically distinct diseases [32], we wondered if this heterogeneity is also reflected in the stem cell transcriptome. To this end, we re-analyzed the GSE 17054 micro-array dataset from Majeti et al. [17] and identified 486 significant DEGs between adult LSC and HSC (Fig. S2A). Comparing the set of LSC-HSC DEGs identified in pediatric versus adult AML revealed 71 common downregulated targets (Fig. S2B), which was translated into mutual repressed pathways, e.g. tight junction and MAPK signaling [17] (Table 2). In sharp contrast, only three common LSC-upregulated transcripts were identified (*TYROBP*, *CFP* and *PTH2R*).

Table 2. Enriched and suppressed pathways identified by STRING functional protein associations.

DEGs between LSC and HSC		DEGs between L-blast and C-blast
Enriched pathways (DEG=83)	**Suppressed pathways (DEG=212)**	**Enriched pathways (DEG=157)**
Pathways in cancer	*Tight junctions*	FoxO signaling pathway
Cancer transcriptional misregulation	Cancer transcriptional misregulation	Cytokine-cytokine receptor interaction
MAPK signaling pathway	Rap1 signaling pathway	Influenza A
Colorectal cancer	*MAPK signaling pathway*	HTLV-I infection
Osteoclast differentiation	Th17 cell differentiation	Osteoclast differentiation
NOD-like receptor signaling	Calcium signaling pathway	Epstein-Barr virus infection
Apoptosis		Cancer transcriptional misregulation
Breast cancer		
FoxO signaling pathway		

KEGG annotated pathways were derived by STRING analysis based on significant DEGs (number shown between brackets) identified for each comparison. LSC pathways that overlap with the top 10 ranked dysregulated pathways in adult AML (GSE 17054) are indicated in italic. DEGs, differentially expressed genes; LSC, leukemic stem cell; HSC, hematopoietic stem cell; L-blast, leukemic blast; C-blast, control blast.

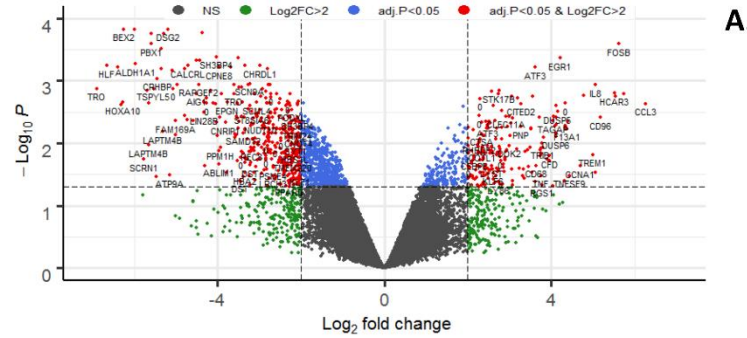
Metabolic changes in pedAML L-blasts enhance proliferation compared to normal counterparts

We identified 157 and 332 significantly up- and downregulated targets in L-blast versus C-blast. The top 50 upregulated transcripts is shown in Fig. 3A. Pathways enriched in L-blasts involved cancer transcriptional misregulation, FoxO signaling, and cytokine-cytokine receptor interaction (Table 2).

Functional protein network analysis (STRING) illustrated a significant enrichment in L-blasts of stimuli responses (6.8e-11) and metabolic processes (FDR 1.4e-05). GSEA identified 163 enriched and 23 suppressed gene sets in L-blast versus C-blast. The top-ranked adipogenesis gene set correlates with metabolic dysregulation, whereas increased EGF signaling and decreased stemness signatures relate to high proliferation (Fig. 3B). Among others, upregulated cancer and EGFR signaling, and downregulated death signaling, were confirmed by unsupervised clustering (Fig. S1B).

Interestingly, the top-ranked (anti-)correlated gene sets identified in LSC (Fig. 2B) and L-blast (Fig. 3B) partially overlapped, and also enriched and suppressed pathways were recurrent (Table 2). Therefore, we sought similarities in the DEGs identified in LSC and L-blast. From the 83 and 157 significantly upregulated genes in LSC and L-blasts, respectively, 49 genes appeared to be common (Fig. S3A). On the other hand, 134 targets were mutually downregulated from a total set of 212 and 332 transcripts, respectively (Fig. S3B). Analysis of functional protein associations (STRING) of the mutual up- and downregulated DEGs illustrated their involvement in general cancer-related pathways. Taken together, we concluded that LSC and L-blast share pan-leukemic molecular aberrancies compared to their normal counterparts.

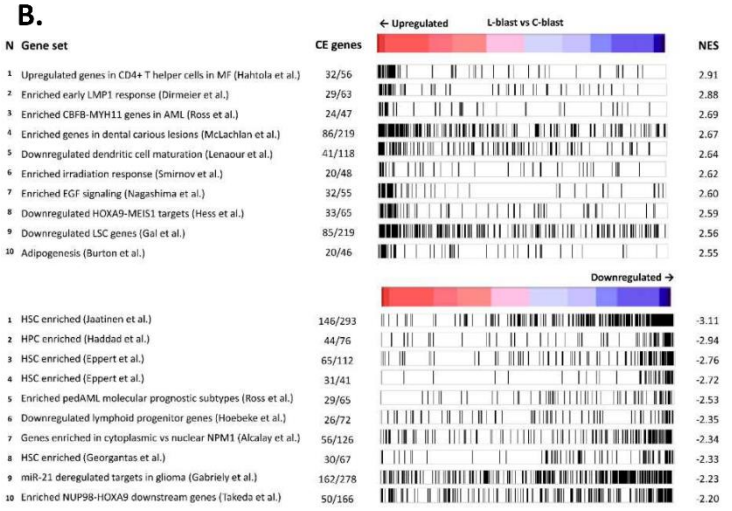
Fig. 3. Transcriptional differences and (anti-)correlated gene set and pathways between L-blast and C-blast.



A.

Top 50 L-blast upregulated targets		
Gene Symbol	log2FC	adj.P.Val
CCL3	6.26	0.0023
C13orf15	5.74	0.0016
FOSB	5.62	0.0003
PLIN2	5.54	0.0018
HCAR3	5.53	0.0016
CD96	5.18	0.0039
SGK1	5.07	0.0294
IL8	5.06	0.0012
TREM1	5.00	0.0152
GADD45B	4.77	0.0017
CCNA1	4.71	0.0232
TNFSF9	4.43	0.0328
TCTEX1D4	4.40	0.0044
F13A1	4.37	0.0058
CDKN1A	4.34	0.0052
HSPA1B	4.32	0.0404
EGR1	4.22	0.0004
DUSP5	4.15	0.0031
C11orf96	4.12	0.0051
NR4A2	4.12	0.0038
EMP1	4.12	0.0025
DUSP6	4.11	0.0079
HOMER3	4.11	0.0069
PLK3	4.09	0.0052
HILPDA	4.04	0.0488
TAGAP	4.04	0.0044
SNAI1	3.96	0.0193
CFD	3,.96	0.0160
P2RY2	3.90	0.0179
DFNB31	3.81	0.0080
TRIB1	3.79	0.0115
RGS1	3.77	0.0466
IL13RA1	3.76	0.0149
SLED1	3.75	0.0146
TNF	3.75	0.0332
S100P	3.75	0.0121
HCAR2	3.72	0.0039
CD68	3.63	0.0228
HSPA8	3.61	0.0410
ATF3	3.61	0.0006
SERTAD1	3.55	0.0018
DNAJB1	3.53	0.0291
TNFAIP3	3.49	0.0116
NLRP3	3.45	0.0238
CD83	3.44	0.0128
CDKN2C	3.38	0.0135
CACNA2D4	3.36	0.0367
CLEC12A	3.33	0.0262
PLAU	3.31	0.0332
PNP	3.28	0.0054

B.



N	Gene set	CE genes	NES
1	Upregulated genes in CD4+ T helper cells in MF (Hahtola et al.)	32/56	2.91
2	Enriched early LMP1 response (Dirmeier et al.)	29/63	2.88
3	Enriched CBFB-MYH11 genes in AML (Ross et al.)	24/47	2.69
4	Enriched genes in dental carious lesions (McLachlan et al.)	86/219	2.67
5	Downregulated dendritic cell maturation (Lenaour et al.)	41/118	2.64
6	Enriched irradiation response (Smirnov et al.)	20/48	2.62
7	Enriched EGF signaling (Nagashima et al.)	32/55	2.60
8	Downregulated HOXA9-MEIS1 targets (Hess et al.)	33/65	2.59
9	Downregulated LSC genes (Gal et al.)	85/219	2.56
10	Adipogenesis (Burton et al.)	20/46	2.55
1	HSC enriched (Jaatinen et al.)	146/293	-3.11
2	HPC enriched (Haddad et al.)	44/76	-2.94
3	HSC enriched (Eppert et al.)	65/112	-2.76
4	HSC enriched (Eppert et al.)	31/41	-2.72
5	Enriched pedAML molecular prognostic subtypes (Ross et al.)	29/65	-2.53
6	Downregulated lymphoid progenitor genes (Hoebeke et al.)	26/72	-2.35
7	Genes enriched in cytoplasmic vs nuclear NPM1 (Alcalay et al.)	56/126	-2.34
8	HSC enriched (Georgantas et al.)	30/67	-2.33
9	miR-21 deregulated targets in glioma (Gabriely et al.)	162/278	-2.23
10	Enriched NUP98-HOXA9 downstream genes (Takeda et al.)	50/166	-2.20

(A) Visualization of genes identified to be upregulated in L-blast (n=4) compared to C-blast (n=3). Genes are plotted in a volcano plot as log2FC values against -Log10 adj. P-values. Thresholds |log2FC| > 2 and -log10 adj. P < .05 are shown as dashed lines, DEGs are highlighted in red. The top 50-ranked upregulated genes, sorted by log2FC values, are annotated with gene symbols on the right. **(B)** Top 10 most correlated (top) and anti-correlated (bottom) gene sets identified through GSEA, based on the DEGs between L-blast and C-blast. The number of concordantly expressed (CE) genes/total genes and normalized enrichment scores (NES) is shown for each gene set individually. FC, fold change; DEGs, differentially expressed genes; L-blast, leukemic blast; C-blast, control blast.

Novel candidate targets in pedAML leukemic subpopulations validated by qPCR

The top-ranked up- and downregulated targets in LSC versus HSC, and upregulated targets in L-blast versus C-blast, identified by micro-array profiling were validated by qPCR (29/83, 23/212 and 42/157, respectively). All 94 targets were evaluated according to a three-step exclusion strategy, allowing the most significantly DEGs to be evaluated in the highest number of cell fractions. Patients were dichotomized as high and low for the targets with the highest significant differential expression. Per target, overexpression was correlated to cytogenetic and molecular markers, and if treated in the NOPHO-DBH AML2012 study, with clinical outcome.

First, significant differential expression was confirmed for 24/29 LSC upregulated targets by qPCR analysis of 10 LSC versus 6 HSC fractions (Fig. S4A, P<.05). Moreover, for 12/24 targets, differential expression was significant at P<.01 with concomitant low expression in HSC. Expression of these 12 targets was further evaluated using additional LSC and HSC fractions (Fig. S4B). Too low LSC expression,

or too high HSC expression, led to the exclusion of 4/12 targets. Finally, expression of *PLIN2*, *CFD*, *EMP1*, *DUSP5*, *ANXA2*, *CRIP1*, *CDKN1A* and *CFP* was evaluated in all fractions and shown to be highly significantly overexpressed in LSC (n=42) compared to HSC (n=20) (Table 3, Fig. S4C). Overexpression of these eight targets was observed in 38-67% of the patients, with *CDKN1A*, *CRIP1*, *CFP* and *CFD* overexpressed in more than half of the patients. LSC expression was averagely 4- too 12-fold higher compared to HSC, with *CFD* and *CRIP1* as most upregulated targets. *CFP* overexpression significantly correlated to *FLT3*-ITD mutations, e.g. 46% in *CFP*-high (n=14) versus 10% in *CFP*-low (n=10) patients (P=.043). High *ANXA2* levels beneficially impacted EFS at a borderline significant level (P=.061, 4 *ANXA2*-high versus 10 *ANXA2*-low patients, Fig. S5A). These targets showed no significant expression differences between BM and PB (n=11), except for *CFP* (5-6 fold higher in PB, P=.018), nor between phenotypically different LSC sorted from the same patient (n=7).

Second, qPCR confirmed significant upregulation of 16/42 targets in 9 L-blast versus 5 C-blast fractions (Fig. S6A, P<.05). Further analysis using more samples showed that expression was too low in L-blasts, or too high in C-blasts, for 7/16 targets (Fig. S6B). Evaluation of the remaining nine targets in all samples illustrated that *DUSP6*, *HOMER3*, *ANXA2P1*, *CTSA*, *RHBDF2*, *EMP1*, *GADD45B*, *TYROBP* and *PNP* transcripts were highly significantly overexpressed in L-blasts (n=36) versus C-blasts (n=19) (Fig. S6C, Table 3). Again, expression levels did not significantly differ between PB and BM (n=10). Five out of nine targets (*RHBDF2*, *HOMER3*, *ANXA2P1*, *GADD45B* and *TYROBP*) were overexpressed in more than two third of the patients. *HOMER3* showed the highest differential expression (14-FC higher in L-blast compared to C-blast). Interestingly, *HOMER3*-high cases (n=21) showed significantly less inv(16)(p13q22) (P=.014) and FAB M4 (P=.031), associated with a beneficial outcome, compared to *HOMER3*-low pedAML (n=4). *DUSP6*, overexpressed in one third of the patients, was previously shown to be significantly associated with *FLT3*-ITD in adult AML [33], which we could not confirm in a pediatric setting (P=.49). Amongst NOPHO-DBH AML2012-treated pedAML (n=13/25), *PNP*-high patients showed a significant lower EFS (Fig. S5B), which was confirmed by Cox log-rank univariate analysis (hazard ratio 9.24, P=.04), but did not remain significant in multivariate analysis taking into account other demographics shown in Table 1 (P>.05).

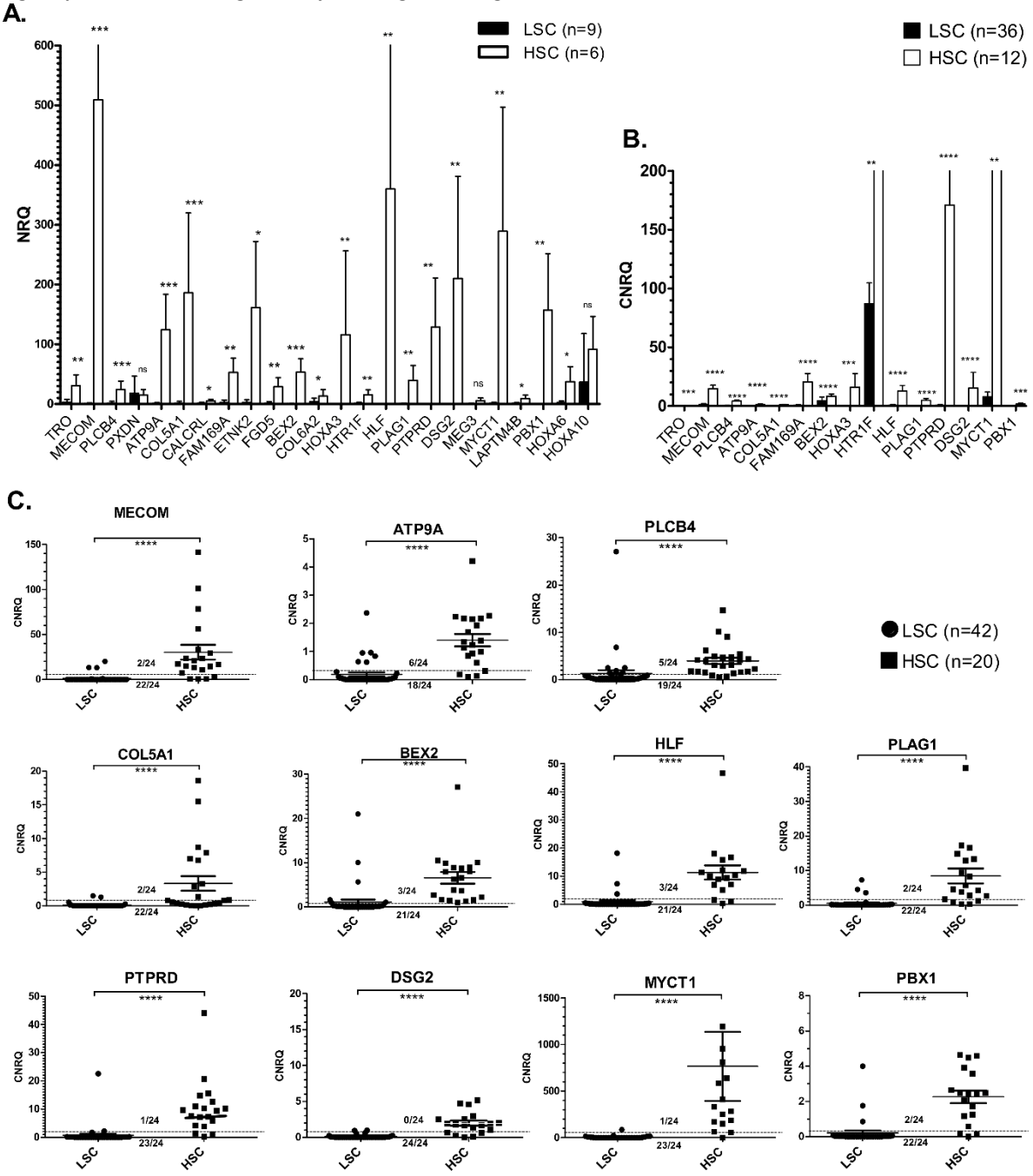
Third, 23 LSC downregulated targets were evaluated by qPCR (Fig. 4A). Twenty-one targets showed significant differential expression at P<.05. Among these, 15/21 targets were even significantly downregulated at P<.01, with virtual absent expression in LSC. Including more sample fractions led to the exclusion of 4/15 targets due to too low differential expression levels (Fig. 4B). Evaluation of all samples illustrated that *MECOM, HLF, PLCB4, PLAG1, ATP9A, PTPRD, COL5A1, BEX2, DSG2, MYCT1* and *PBX1* transcripts are highly significantly repressed in LSC compared to HSC (P<.0001, Fig. 4C). These targets appeared to be uniformly suppressed, i.e. between 75-100% of the patients. *BEX2* downregulation was significantly anticorrelated to *KMT2A*-rearrangements in pedAML LSC fractions (P<.01), as previously demonstrated in cell lines [34]. On the other hand, previously reported associations between *PLAG1* and inv(16)(p13q22), or between *MYCT1* and FAB M1/M5/M6, were not confirmed (P>.05) [35, 36]. Interestingly, 9/11 pedAML LSC-downregulated targets were also significantly suppressed in adult LSC compared to HSC (Fig. S1B). Furthermore, *PBX1*, *MYCT1*, *HLF*, *ATP9A* and *PLCB4* appeared to be also suppressed in other pediatric hematological malignancies, whereas *PTPRD*, *COL5A1* and *MECOM* downregulation was rather AML-specific (Fig. S7).

Table 3. Frequency and magnitude of overexpression in LSC and L-blast upregulated targets.

	LSC upregulated targets								
Expression per fraction	***PLIN-2***	***CFD***	***EMP1***	***DUSP5***	***ANXA2***	***CRIP1***	***CDKN1A***	***CFP***	
HSC (n=20)									
mean	0.88	0.59	0.53	0.16	0.11	0.18	0.08	1.35	
SD	0.85	0.80	0.83	0.33	0.23	0.25	0.05	1.80	
Cut-off	2.58	2.20	2.18	0.82	0.57	0.69	0.17	4.95	
LSC (n=24)									
mean	5.95	7.06	2.80	0.96	0.55	2.03	0.32	7.93	
SD	13.90	14.00	4.76	1.10	0.48	2.02	0.35	10.30	
PedAML with overexpression	**46%**	**54%**	**38%**	**38%**	**38%**	**63%**	**67%**	**58%**	
Ratio LSC/ HSC	**6.79**	**11.89**	**5.28**	**5.99**	**4.83**	**11.34**	**3.78**	**5.87**	
	L-blast upregulated targets								
	DUSP6	***HOMER3***	***PNP***	***ANXA2P1***	***CTSA***	***RHBDF2***	***EMP1***	***GADD45B***	***TYROBP***
C-blast (n=19)									
mean	0.66	0.03	0.64	0.19	0.43	0.72	0.45	0.19	9.33
SD	0.65	0.04	0.34	0.13	0.27	0.32	0.47	0.14	12.80
Cut-off	1.97	0.11	1.32	0.44	0.97	1.37	1.39	0.47	34.93
L-blast (n=25)									
mean	3.22	0.45	1.31	0.67	1.42	3.30	3.57	0.78	63.80
SD	2.87	0.52	0.81	0.54	0.81	4.04	6.71	0.65	39.70
PedAML with overexpression	**36%**	**84%**	**47%**	**72%**	**44%**	**88%**	**68%**	**72%**	**72%**
Ratio L-blast/ C-blast	**4.85**	**13.62**	**2.04**	**3.52**	**3.29**	**4.56**	**8.02**	**4.16**	**6.84**

Mean expression and standard deviations (SD) calculated for the eight most significantly upregulated targets in LSC versus HSC and the nine most significantly upregulated targets in L-blast versus C-blast. Cut-offs were calculated based on the mean expression measured in the healthy counterparts plus 2xSD. Percentage of patients classified as having overexpression are shown, together with the magnitude of overexpression, expressed as the fold change (FC) of the mean expression in the leukemic versus normal compartment.

Fig. 4. qPCR validation of significantly downregulated targets in LSC versus HSC.



Targets with significant downregulated expression in LSC compared to HSC, shown by micro-array analysis, were evaluated by qPCR analysis in three steps and expressed as (calibrated) normalized relative quantities (CNRQ). **(A)** 23/212 LSC downregulated targets were selected for qPCR validation. **(B)** 15 out of these 23 targets were further evaluated in a second step. **(C)** 11/15 targets were evaluated in all available LSC (n=42) and HSC (n=20) fractions. Mean values are shown by horizontal lines, error bars indicate the 95% confidence interval of the mean. The dotted line indicates the number of patients showing underexpression. P-values (one-tailed) were calculated by the Mann-Whitney U test, and one, two, three or four asteriks indicate the level of significance (P<.05, P<.01, P<.001 and P<.0001, respectively). LSC, leukemic stem cell; HSC, hematopoietic stem cell; DEG, differentially expressed gene.

Discussion

We here describe a novel set of differentially expressed targets in LSC and L-blast of pedAML patients identified using a CvN approach. Moreover, we reveal previously unexplored deregulated pathways in these leukemic subpopulations sorted from children with AML.

Eight targets were found highly significantly overexpressed in pedAML LSC compared to HSC. *CDKN1A*, *ANXA2, EMP1, CFD* were previously linked to leukemogenesis, whereas the role of *PLIN2*, *DUSP5*, *CRIP1*, and *CFP* in AML remains elusive. *CDKN1A*, *CRIP1*, *CFP* and *CFD* were found to be most frequently overexpressed (>50% of the patients), with *CRIP1* and *CFD* showing the highest differential expression compared to HSC. *CDKN1A* might represent an interesting target for LSC eradication, since elevated expression was reported to maintain LSC activity [37, 38], and *CDKN1A* knockdown indirectly reversed stem cell quiescence [39]. Overexpression of *CFP* and *CFD* in LSC of pedAML patients suggest a disturbed complement pathway regulation. *CFP*, involved in regulation of the alternative complement pathway, might be an interesting target due to the significant correlation with *FLT3*-ITD (P=.043) and concomitant overexpression in adult AML LSC. *CFD* expression was previously linked to poor outcome in adult AML [40], and its prognostic value in children with AML awaits validation. Lastly, we found that patients with high *ANXA2* LSC expression (38%, 4.8-fold higher expression than HSC) show a trend towards prolonged EFS (P=.061), in agreement with the previously reported favorable prognostic effect of *ANXA2* in bulk pedAML cells [41]. Therefore, *ANXA2* holds promise as a biomarker for risk stratification. Although only the top-ranked DEGs were validated by qPCR in our study, micro-array analysis additionally revealed interesting targets for flow cytometric validation. Our data confirm a previously reported high CD96 expression in pedAML LSC [42]. A potential role for targeting CD180 and CD68 in LSC, or their qualification as follow-up marker, deserves further attention. High CLEC12A expression in LSC suggests a role for anti-CLL-1 therapy in pedAML [43, 44].

Suppression of LSC-specific downregulated targets was highly consistent across the different genetic subgroups, as only very few patients showed comparable expression to their healthy counterparts. Since several of these targets endow tumor suppressor roles in other cancer entities, further investigation whether hypomethylating therapy could result into LSC eradication in pedAML is warranted. *MYCT1* was already identified as a TSG in AML [36], and deserves the highest attention since expression levels are 100-fold lower in LSC compared to HSC. *PBX1* was previously reported to act as an oncogene and tumor suppressor [45]. Suppressed *PTPRD* levels in LSC correlates with the previous report that *PTPRD* is suppressed in bulk leukemic cells of 2/3 pedAML patients [46].

Transcriptional misregulation in cancer, osteoclastogenesis and tight junction pathways were dysregulated in LSC. The ‘cancer transcriptional misregulation’ pathway is associated with myeloid leukemogenesis, AML cell functions and held responsible for tumorigenic epigenetic abnormalities [47, 48]. Distortion of osteoclastogenesis and tight junction pathways might hypothetically provide LSC an advantage over HSC during homing towards the endosteal-vascular niche. The observed immune dysregulation, separating LSC from HSC, strokes with a previous statement that multiple inflammatory signaling pathways are involved in the generation of pre-LSCs [49].

L-blast upregulated targets were often overexpressed in a larger portion of the patient cohort compared to LSC-upregulated targets (38-67%, median 50% versus 36-88%, median 72%, respectively). Among these, only *DUSP6* and *HOMER3* were previously addressed in AML. *DUSP6* is an important cellular signaling regulator overexpressed in AML [33]. *HOMER3* relates to the occurrence and

development of AML [50], and increased levels significantly associated with favorable cytogenetics in adult AML [51]. Since *HOMER3* also showed the highest differential expression compared to C-blasts, targeting could be of therapeutic value. *CTSA* also represent an interesting target, since several other cathepsins (*CTSB*, *CTSG*) were shown to have a diagnostic, prognostic or therapeutic significance in AML [52-55]. *GADD45B* is known to be involved in negative growth control during myeloid differentiation [56], and suggested to play a role in the tumorigenesis of colorectal carcinoma [57]. Finally, high *PNP* expression was significantly associated with a worse EFS. Studies with PNP inhibitors in relapsed and refractory leukemias are ongoing [58], and investigation of their applicability in pedAML might be worthwhile.

Deregulated pathways identified in L-blast compared to their normal counterparts suggest that disturbed regulation of cell cycling, apoptosis, glycolysis/gluconeogenesis and oxidative stress resistance promote the maintenance and proliferation of blasts in a leukemic setting. Indeed, adapting to hypoxic conditions and switching from oxidative phosphorylation towards glycolysis was shown to correlate with an aggressive disease course in solid cancers [59]. Therefore, blocking gluconeogenesis was proposed as an anti-tumor therapeutic strategy [60].

Only *EMP1* appeared to be highly significantly upregulated in LSC and L-blast (38% and 68% of patients, respectively). Although Ng et al. did not elaborate on its role, *EMP1* was included in the recently established LSC17 score [16, 25], but not retained in the pedAML-specific LSC6-score [26]. However, based on the here observed CD38-independent overexpression, and the previously reported *in vitro* targetability of *EMP1* in B-ALL [61], its role as a therapeutic target should be further explored.

We also detected a high molecular heterogeneity between pediatric and adult AML at the stem cell level. LSC populations from both entities shared 71 suppressed transcripts, but only three mutual upregulated targets (*TYROBP*, *CFP*, *PTH2R*) were identified. *TYROBP* and *CFP* have not been functionally associated with AML, and their role in LSC transformation should be further explored. *PTH2R*, on the other hand, is known to be upregulated in AML and MDS [62], including adult AML LSC [63]. Further research is warranted to evaluate whether these three targets could serve as pan-LSC targets, irrespective of the age of onset. Although these findings further underline the distinct biology between pediatric and adult AML [32], it should be taken into account that, due to the small number of patients evaluated (four pedAML and nine adult AML patients), the genetic subtypes might also play a role besides the age of the patients.

Although promising, these data need confirmation in larger, preferentially multicenter trials, as survival analyses were performed in only a limited number patients. It is important to acknowledge that the pedAML cohort included one secondary AML evolved from juvenile myelomonocytic leukemia (JMML), one acute promyelocytic leukemia (APL) and two relapsed patients, although they were all excluded from outcome analysis. Knowledge on whether expression differs between primary and secondary pedAML, and between APL and non-APL cases, will be required.

In conclusion, we here report an unique set of LSC and L-blast specific overexpressed genes in pedAML. Most targets have not been studied in AML, and are involved in immune regulation, apoptosis, adhesion, or intracellular signaling, making them attractive candidates for functional studies, refining signatures and targeted therapy. Inflammatory pathways and immune regulation are critical biological networks perturbed in pedAML LSC, and L-blast present a high proliferative cell cycle activity combined with metabolic dysregulation. In addition, we identified novel LSC-specific downregulated targets,

often described as TSGs in solid tumours, of which some are also relevant in adult AML LSC or other pediatric haematological diseases.

Acknowledgments

Our gratitude goes to dr. F. Plasschaert, C. Matthys of the Cord Blood Bank and the Department of Pediatric Haematology and Oncology of the Ghent University Hospital (Ghent, Belgium) for providing samples. The authors thank all patients and their parents for their participation in the study, as well as the data managers involved in the clinical trials. This research was funded by the Belgian Foundation against Cancer (grant 2014–265), FOD-KankerPlan (Actie29, grant to JP), vzw Kinderkankerfonds (grant to TL) and the Research Foundation - Flanders (Fonds voor Wetenschappelijk Onderzoek Vlaanderen, FWO, grant 1113117 to BD). This work is submitted in partial fulfilment of the requirement for the PhD of candidate BD at Ghent University.

References

1. Rasche M, Zimmermann M, Borschel L, Bourquin JP, Dworzak M, Klingebiel T, Lehrnbecher T, Creutzig U, Klusmann JH, Reinhardt D: Successes and challenges in the treatment of pediatric acute myeloid leukemia: a retrospective analysis of the AML-BFM trials from 1987 to 2012. Leukemia 2018, 32(10):2167-2177.
2. De Moerloose B, Reedijk A, de Bock GH, Lammens T, de Haas V, Denys B, Dedeken L, van den Heuvel-Eibrink MM, Te Loo M, Uyttebroeck A et al: Response-guided chemotherapy for pediatric acute myeloid leukemia without hematopoietic stem cell transplantation in first complete remission: Results from protocol DB AML-01. Pediatric blood & cancer 2019:e27605.
3. Abrahamsson J, Forestier E, Heldrup J, Jahnukainen K, Jonsson OG, Lausen B, Palle J, Zeller B, Hasle H: Response-guided induction therapy in pediatric acute myeloid leukemia with excellent remission rate. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2011, 29(3):310-315.
4. De Kouchkovsky I, Abdul-Hay M: Acute myeloid leukemia: a comprehensive review and 2016 update. Blood cancer journal 2016, 6(7):e441.
5. Hope KJ, Jin L, Dick JE: Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity. Nat Immunol 2004, 5(7):738-743.
6. Shlush LI, Mitchell A, Heisler L, Abelson S, Ng SWK, Trotman-Grant A, Medeiros JJF, Rao-Bhatia A, Jaciw-Zurakowsky I, Marke R et al: Tracing the origins of relapse in acute myeloid leukaemia to stem cells. Nature 2017, 547(7661):104-108.
7. Hanekamp D, Denys B, Kaspers GJL, Te Marvelde JG, Schuurhuis GJ, De Haas V, De Moerloose B, de Bont ES, Zwaan CM, de Jong A et al: Leukaemic stem cell load at diagnosis predicts the development of relapse in young acute myeloid leukaemia patients. British journal of haematology 2018, 183(3):512-516.
8. Witte KE, Ahlers J, Schafer I, Andre M, Kerst G, Scheel-Walter HG, Schwarze CP, Pfeiffer M, Lang P, Handgretinger R et al: High Proportion of Leukemic Stem Cells at Diagnosis Is Correlated with Unfavorable Prognosis in Childhood Acute Myeloid Leukemia. Pediatr Hemat Oncol 2011, 28(2):91-99.
9. Bachas C, Schuurhuis GJ, Zwaan CM, van den Heuvel-Eibrink MM, den Boer ML, de Bont ES, Kwidama ZJ, Reinhardt D, Creutzig U, de Haas V et al: Gene expression profiles associated with pediatric relapsed AML. PloS one 2015, 10(4):e0121730.
10. Pigazzi M, Masetti R, Bresolin S, Beghin A, Di Meglio A, Gelain S, Trentin L, Baron E, Giordan M, Zangrando A et al: MLL partner genes drive distinct gene expression profiles and genomic

alterations in pediatric acute myeloid leukemia: an AIEOP study. Leukemia 2011, 25(3):560-563.

11. Jo A, Mitani S, Shiba N, Hayashi Y, Hara Y, Takahashi H, Tsukimoto I, Tawa A, Horibe K, Tomizawa D et al: High expression of EVI1 and MEL1 is a compelling poor prognostic marker of pediatric AML. Leukemia 2015, 29(5):1076-1083.
12. Andersson A, Ritz C, Lindgren D, Eden P, Lassen C, Heldrup J, Olofsson T, Rade J, Fontes M, Porwit-Macdonald A et al: Microarray-based classification of a consecutive series of 121 childhood acute leukemias: prediction of leukemic and genetic subtype as well as of minimal residual disease status. Leukemia 2007, 21(6):1198-1203.
13. Balgobind BV, Van den Heuvel-Eibrink MM, De Menezes RX, Reinhardt D, Hollink IH, Arentsen-Peters ST, van Wering ER, Kaspers GJ, Cloos J, de Bont ES et al: Evaluation of gene expression signatures predictive of cytogenetic and molecular subtypes of pediatric acute myeloid leukemia. Haematologica 2011, 96(2):221-230.
14. Zangrando A, Dell'orto MC, Te Kronnie G, Basso G: MLL rearrangements in pediatric acute lymphoblastic and myeloblastic leukemias: MLL specific and lineage specific signatures. BMC medical genomics 2009, 2:36.
15. Ross ME, Mahfouz R, Onciu M, Liu HC, Zhou X, Song G, Shurtleff SA, Pounds S, Cheng C, Ma J et al: Gene expression profiling of pediatric acute myelogenous leukemia. Blood 2004, 104(12):3679-3687.
16. Ng SW, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, Arruda A, Popescu A, Gupta V, Schimmer AD et al: A 17-gene stemness score for rapid determination of risk in acute leukaemia. Nature 2016, Dec 15(540(7633)):433-437.
17. Majetl R, Becker MW, Tian Q, Lee TLM, Yan XW, Liu R, Chiang JH, Hood L, Clarke MF, Weissman IL: Dysregulated gene expression networks in human acute myelogenous leukemia stem cells. Proceedings of the National Academy of Sciences of the United States of America 2009, 106(9):3396-3401.
18. Eppert K, Takenaka K, Lechman ER, Waldron L, Nilsson B, van Galen P, Metzeler KH, Poeppl A, Ling V, Beyene J et al: Stem cell gene expression programs influence clinical outcome in human leukemia. Nature medicine 2011, 17(9):1086-1093.
19. Forsberg EC, Passegue E, Prohaska SS, Wagers AJ, Koeva M, Stuart JM, Weissman IL: Molecular signatures of quiescent, mobilized and leukemia-initiating hematopoietic stem cells. PloS one 2010, 5(1):e8785.
20. Gal H, Amariglio N, Trakhtenbrot L, Jacob-Hirsh J, Margalit O, Avigdor A, Nagler A, Tavor S, Ein-Dor L, Lapidot T et al: Gene expression profiles of AML derived stem cells; similarity to hematopoietic stem cells. Leukemia 2006, 20(12):2147-2154.
21. Gentles AJ, Plevritis SK, Majeti R, Alizadeh AA: Association of a leukemic stem cell gene expression signature with clinical outcomes in acute myeloid leukemia. Jama 2010, 304(24):2706-2715.
22. Saito Y, Kitamura H, Hijikata A, Tomizawa-Murasawa M, Tanaka S, Takagi S, Uchida N, Suzuki N, Sone A, Najima Y et al: Identification of therapeutic targets for quiescent, chemotherapy-resistant human leukemia stem cells. Science translational medicine 2010, 2(17):17ra19.
23. de Leeuw DC, Denkers F, Olthof MC, Rutten AP, Pouwels W, Schuurhuis GJ, Ossenkoppele GJ, Smit L: Attenuation of microRNA-126 expression that drives CD34+38- stem/progenitor cells in acute myeloid leukemia leads to tumor eradication. Cancer research 2014, 74(7):2094-2105.
24. Duployez N, Preudhomme C, Cheok M: A 17-gene-expression profile to improve prognosis prediction in childhood acute myeloid leukemia. Oncotarget 2018, 9(74):33869-33870.
25. Duployez N, Marceau-Renaut A, Villenet C, Petit A, Rousseau A, Ng SWK, Paquet A, Gonzales F, Barthelemy A, Lepretre F et al: The stem cell-associated gene expression signature allows risk stratification in pediatric acute myeloid leukemia. Leukemia 2018, 33(2):348-357.
26. Elsayed AH, Rafiee R, Cao X, Raimondi S, Downing JR, Ribeiro R, Fan Y, Gruber TA, Baker S, Klco J et al: A six-gene leukemic stem cell score identifies high risk pediatric acute myeloid leukemia. Leukemia 2019, doi: 10.1038/s41375-019-0604-8. [Epub ahead of print].

27. Richard-Carpentier G, Sauvageau G: Bringing a Leukemic Stem Cell Gene Signature into Clinics: Are We There Yet? Cell stem cell 2017, 20(3):300-301.
28. Xu ZJ, Ma JC, Zhou JD, Wen XM, Yao DM, Zhang W, Ji RB, Wu DH, Tang LJ, Deng ZQ et al: Reduced protocadherin17 expression in leukemia stem cells: the clinical and biological effect in acute myeloid leukemia. Journal of translational medicine 2019, 17(1):102.
29. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P et al: STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic acids research 2019, 47(D1):D607-D613.
30. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 2005, 102(43):15545-15550.
31. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, Wadi L, Meyer M, Wong J, Xu C et al: Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nature protocols 2019, 14(2):482-517.
32. Bolouri H, Farrar JE, Triche T, Jr., Ries RE, Lim EL, Alonzo TA, Ma Y, Moore R, Mungall AJ, Marra MA et al: The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nature medicine 2018, Jan(24(1)):103-112.
33. Arora D, Kothe S, van den Eijnden M, Hooft van Huijsduijnen R, Heidel F, Fischer T, Scholl S, Tolle B, Bohmer SA, Lennartsson J et al: Expression of protein-tyrosine phosphatases in Acute Myeloid Leukemia cells: FLT3 ITD sustains high levels of DUSP6 expression. Cell communication and signaling : CCS 2012, 10(1):19.
34. Rohrs S, Dirks WG, Meyer C, Marschalek R, Scherr M, Slany R, Wallace A, Drexler HG, Quentmeier H: Hypomethylation and expression of BEX2, IGSF4 and TIMP3 indicative of MLL translocations in acute myeloid leukemia. Molecular cancer 2009, 8:86.
35. Landrette SF, Kuo YH, Hensen K, Barjesteh van Waalwijk van Doorn-Khosrovani S, Perrat PN, Van de Ven WJ, Delwel R, Castilla LH: Plag1 and Plagl2 are oncogenes that induce acute myeloid leukemia in cooperation with Cbfb-MYH11. Blood 2005, 105(7):2900-2907.
36. Fu S, Fu Y, Chen F, Hu Y, Quan B, Zhang J: Overexpression of MYCT1 Inhibits Proliferation and Induces Apoptosis in Human Acute Myeloid Leukemia HL-60 and KG-1a Cells in vitro and in vivo. Frontiers in pharmacology 2018, 9:1045.
37. Peng L, Tang Y, Zhang Y, Guo S, Ye L, Wang Y, Jiang Y: Structural maintenance of chromosomes 4 is required for leukemia stem cell maintenance in MLL-AF9 induced acute myeloid leukemia. Leukemia & lymphoma 2018, 59(10):2423-2430.
38. Andries V, Vandepoele K, Staes K, Berx G, Bogaert P, Van Isterdael G, Ginneberge D, Parthoens E, Vandenbussche J, Gevaert K et al: NBPF1, a tumor suppressor candidate in neuroblastoma, exerts growth inhibitory effects by inducing a G1 cell cycle arrest. BMC cancer 2015, 15:391.
39. Korthuis PM, Berger G, Bakker B, Rozenveld-Geugien M, Jaques J, de Haan G, Schuringa JJ, Vellenga E, Schepers H: CITED2-mediated human hematopoietic stem cell maintenance is critical for acute myeloid leukemia. Leukemia 2015, 29(3):625-635.
40. Laverdiere I, Boileau M, Herold T, Rak J, Berdel WE, Wormann B, Hiddemann W, Spiekermann K, Bohlander SK, Eppert K: Complement cascade gene expression defines novel prognostic subgroups of acute myeloid leukemia. Experimental hematology 2016, 44(11):1039-1043 e1010.
41. Niu Y, Yang X, Chen Y, Jin X, Xie Y, Tang Y, Li L, Liu S, Guo Y, Li X et al: Distinct prognostic values of Annexin family members expression in acute myeloid leukemia. Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico 2019.
42. Chavez-Gonzalez A, Dorantes-Acosta E, Moreno-Lorenzana D, Alvarado-Moreno A, Arriaga-Pizano L, Mayani H: Expression of CD90, CD96, CD117, and CD123 on different hematopoietic

cell populations from pediatric patients with acute myeloid leukemia. Archives of medical research 2014, 45(4):343-350.
43. Darwish NH, Sudha T, Godugu K, Elbaz O, Abdelghaffar HA, Hassan EE, Mousa SA: Acute myeloid leukemia stem cell markers in prognosis and targeted therapy: potential impact of BMI-1, TIM-3 and CLL-1. Oncotarget 2016.
44. Hanekamp D, Cloos J, Schuurhuis GJ: Leukemic stem cells: identification and clinical application. International journal of hematology 2017, 105(5):549-557.
45. Dardaei L, Longobardi E, Blasi F: Prep1 and Meis1 competition for Pbx1 binding regulates protein stability and tumorigenesis. Proceedings of the National Academy of Sciences of the United States of America 2014, 111(10):E896-905.
46. Song L, Jiang W, Liu W, Ji JH, Shi TF, Zhang J, Xia CQ: Protein tyrosine phosphatases receptor type D is a potential tumour suppressor gene inactivated by deoxyribonucleic acid methylation in paediatric acute myeloid leukaemia. Acta paediatrica 2016, 105(3):e132-141.
47. Jasielec J, Saloura V, Godley LA: The mechanistic role of DNA methylation in myeloid leukemogenesis. Leukemia 2014, 28(9):1765-1773.
48. Yin X, Huang S, Zhu R, Fan F, Sun C, Hu Y: Identification of long non-coding RNA competing interactions and biological pathways associated with prognosis in pediatric and adolescent cytogenetically normal acute myeloid leukemia. Cancer cell international 2018, 18:122.
49. Hemmati S, Haque T, Gritsman K: Inflammatory Signaling Pathways in Preleukemic and Leukemic Stem Cells. Frontiers in oncology 2017, 7:265.
50. Li Z, Qiu HY, Jiao Y, Cen JN, Fu CM, Hu SY, Zhu MQ, Wu DP, Qi XF: Growth and differentiation effects of Homer3 on a leukemia cell line. Asian Pacific journal of cancer prevention : APJCP 2013, 14(4):2525-2528.
51. Stirewalt DL, Meshinchi S, Kopecky KJ, Fan W, Pogosova-Agadjanyan EL, Engel JH, Cronk MR, Dorcy KS, McQuary AR, Hockenbery D et al: Identification of genes with abnormal expression changes in acute myeloid leukemia. Genes, chromosomes & cancer 2008, 47(1):8-20.
52. Jain M, Bakhshi S, Shukla AA, Chauhan SS: Cathepsins B and L in peripheral blood mononuclear cells of pediatric acute myeloid leukemia: potential poor prognostic markers. Annals of hematology 2010, 89(12):1223-1232.
53. Pandey G, Bakhshi S, Thakur B, Jain P, Chauhan SS: Prognostic significance of cathepsin L expression in pediatric acute myeloid leukemia. Leukemia & lymphoma 2018, 59(9):2175-2187.
54. Alatrash G: Targeting cathepsin G in myeloid leukemia. Oncoimmunology 2013, 2(4):e23442.
55. Pandey G, Bakhshi S, Kumar M, Thakur B, Jain P, Kaur P, Chauhan SS: Prognostic and therapeutic relevance of cathepsin B in pediatric acute myeloid leukemia. American journal of cancer research 2019, 9(12):2634-2649.
56. Liebermann DA, Hoffman B: Myeloid differentiation (MyD) primary response genes in hematopoiesis. Oncogene 2002, 21(21):3391-3402.
57. Wang L, Xiao X, Li D, Chi Y, Wei P, Wang Y, Ni S, Tan C, Zhou X, Du X: Abnormal expression of GADD45B in human colorectal carcinoma. Journal of translational medicine 2012, 10:215.
58. Korycka A, Lech-Maranda E, Robak T: Novel purine nucleoside analogues for hematological malignancies. Recent patents on anti-cancer drug discovery 2008, 3(2):123-136.
59. Calle EE, Kaaks R: Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nature reviews Cancer 2004, 4(8):579-591.
60. Zi F, Zi H, Li Y, He J, Shi Q, Cai Z: Metformin and cancer: An existing drug for cancer prevention and therapy. Oncology letters 2018, 15(1):683-690.
61. Aries IM, Jerchel IS, van den Dungen RE, van den Berk LC, Boer JM, Horstmann MA, Escherich G, Pieters R, den Boer ML: EMP1, a novel poor prognostic factor in pediatric leukemia regulates prednisolone resistance, cell proliferation, migration and adhesion. Leukemia 2014, 28(9):1828-1837.

62. Zhang Z, Zhao L, Wei X, Guo Q, Zhu X, Wei R, Yin X, Zhang Y, Wang B, Li X: Integrated bioinformatic analysis of microarray data reveals shared gene signature between MDS and AML. Oncology letters 2018, 16(4):5147-5159.
63. Bonardi F, Fusetti F, Deelen P, van Gosliga D, Vellenga E, Schuringa JJ: A proteomics and transcriptomics approach to identify leukemic stem cell (LSC) markers. Molecular & cellular proteomics : MCP 2013, 12(3):626-637.

Supplemental Materials and Methods

Treatment protocols and outcome definitions

Patients used for LSC evaluation were included in the DB-AML01 trial (n=5/24), NOPHO-DBH AML2012 trial (n=14/24) or received another treatment (n=5/24, i.e. two relapse cases, one acute promyelocytic leukemia (APL), one patient treated according to an adult protocol (17 year (yr.)) and one secondary AML from juvenile myelomonocytic leukemia (JMML)). Patients used for L-blast evaluation were included in the DB-AML01 trial (n=9/25), NOPHO-DBH AML2012 trial (n=13/25) or received another treatment (n=3/25, i.e. one relapsed patient, one APL and one adult protocol-treated patient (17 yr.)).

Patients included in the DB-AML01 trial (EudraCT 2009-014462-26) concerned children (< 18 yrs.) diagnosed with *de novo* AML between 2010 and 2013. The treatment protocol is described elsewhere [1]. Patients were considered as high-risk (HR) if ≥ 15% blasts persisted after the first induction course and ≤ 5% blasts after the second course (≥5% blasts after the second course was defined as refractory disease).

Patients diagnosed between 2015 and 2019 were included in the NOPHO-DBH AML2012 trial (EudraCT 2012-002934-35), unless they did not fulfill the inclusion criteria. Patients with isolated central nervous system (CNS), extramedullary leukemia, previous chemo- or radiotherapy, AML secondary to a previous bone marrow failure syndrome, myeloid leukemia in Down syndrome with age <5 or ≥5 yrs. with *GATA1* mutation, APL, JMML, myelodysplastic syndrome, Fanconi anemia and/or positive pregnancy test were excluded. NOPHO-DBH AML2012-treated patients received two intensive induction courses, followed by risk-adapted consolidation with three courses of conventional chemotherapy for standard risk (SR) patients and allogeneic stem cell transplantation for HR patients. Patients were defined as HR if they achieved complete remission (CR) after two induction courses and had (i) *FLT3*-ITD/*NPM1* WT profiles, (ii) poor response after induction 1 (i.e. ≥15% leukemic cells at day 22 or at any subsequent evaluation prior to the second course) or (iii) intermediate response after induction 2 (i.e. 0.1%-4.9% leukemic cells before consolidation) [2].

For estimates of EFS, an event was defined as failure to achieve CR, induction death, relapse, development of a second malignancy, or death due to any cause, whichever occurred first. EFS was calculated from date of diagnosis to the date of first event, with failure to achieve CR calculated as an event at t = 0. OS was calculated from date of diagnosis to the date of last follow-up or time of death due to any cause. Follow-up time was censored at the last follow-up visit if no failure was observed.

Staining and sorting strategy

Leukemic cells from patients diagnosed between 2015-2019 (n=19) were immediately processed upon arrival. Leukemic cells from pedAML patients diagnosed between 2010-2013 (n=9) were cryopreserved at time of diagnosis in 90% fetal calf serum (FCS) and 10% dimethylsulfoxide. Cryopreserved samples were thawed by short incubation in a 42°C pre-heated water bath, followed by 30 min incubation at room temperature (RT) in 20 mL RPMI with 20% FCS, 200 µL DNase I (1 mg/mL, grade II bovine pancreas) and 200 µL $MgCl_2$ (1 M) (Sigma-Aldrich). After incubation, cells were spinoculated (10 min, 400 rpm) and washed with 15 mL RPMI/20% FCS.

Freshly collected and thawed mononuclear cells (MNCs) were spinoculated (5 min, 1500 rpm). Monoclonal antibodies were added to the cell pellet (mAb, Table S1). After 20 min incubation in the dark at RT, cell pellets were washed with PBS+2% BSA. Next, labeled cells were resuspended in 50%

RPMI/50%FCS medium and sorted on a FACSAria III with red, blue, and ultraviolet lasers (BD Biosciences).
All scatters were devoid of cell debris and doublets based on propidium iodide exclusion and FSC-H vs FSC-A plots, respectively. The immature myeloid compartment was defined by CD34, CD45 and scatter properties. CD34-positive cases were identified as those with >1% of CD34+ blasts in the leukemic $CD45^{low}/SSC^{low}$ compartment [3, 4]. CD34+/CD38+ blasts and CD34+/CD38- stem cells were gated as previously described [5]. Lymphocytes and fluorescence-minus-one (FMO) controls were used to determine expression cut-offs for CD38 and LSC aberrant markers. Delineated cell populations were backgated on FSC-A/SSC-A and CD45/SSC-A scatter plots to exclude non-specific events and assure homogeneous scatters.

Micro-array experimental settings and data analysis

Total RNA was extracted from sorted cells, resuspended in TRIzol, using the miRNeasy Mini Kit (Qiagen) in combination with on-column DNase I digestion (RNase-Free DNase set, Qiagen) according to manufacturer's instructions. RNA quality and concentrations were measured by Agilent 2100 Bioanalyzer (Agilent) and Qubit (ThermoFisher Scientific), respectively. Mean RNA integrity number of all sorted fractions was 9.3 (range 8.6 – 9.9). Three LSC and four L-blast fractions from pedAML patients were profiled. The LSC fraction for one pedAML patient is lacking, as the RNA fraction experienced a technical issue during profiling. Two 2 HSC and three C-blast fractions from healthy controls were profiled. The RNA yield from one HSC fraction appeared to be too low during the experiment, and was therefore excluded.
Profiling was performed by Biogazelle using a custom 8x60K human Gene expression micro-array (Agilent), containing probes for 33178 human protein-coding genes. To this end, 20 ng RNA was pre-amplified using the Complete Whole Transcriptome Amplification Kit (Sigma-Aldrich). Amplified RNA was subsequently labelled using the Genomic DNA ULS Labeling Kit (Agilent) and hybridized to the array in combination with CGH blocking to reduce background signaling. Micro-arrays were analyzed using an Agilent micro-array scanner and Feature Extraction software (v12.0). Probe intensities were background subtracted, quantile-normalized and log2-based probe intensities were calculated. A target was present if the log2 expression value exceeded the cut-off set at 6.75, based on the dark corner control probe value plus 1. Raw micro-array data are available under the accession number GSE 128103 (released on March 9, 2022).
Data processing of the GSE 128103 dataset, and the publicly available GSE 17054 dataset by Majeti et al. [6], was performed in R using packages EnhancedVolcano and EdgeR (R Bioconductor). If a gene was represented by multiple probes, the highest expression was selected. Raw P-values (P) were adjusted using a Benjamini–Hochberg multiple testing correction. Differential expression analyses were performed between LSC and HSC, and L-blast and C-blast, using unpaired t-test analyses with EdgeR (R Bioconductor).

qPCR technical analysis

RNA was extracted from sorted cells using the miRNeasy Mini or Micro Kit (Qiagen) in combination with on-column DNase I digestion (RNase-Free DNase set, Qiagen) according to manufacturer's instructions. RNA concentrations were measured by Nanodrop (ThermoFisher Scientific) or Qubit RNA HS Assay (Invitrogen). cDNA synthesis was performed after an additional in-solution gDNase elimination step (Heat&Run gDNA removal kit, ArcticZymes), using the 5x PrimeScriptTM RT Master Mix (Takara Bio Europe S.A.S.) in a final volume of 12.5 µL. cDNA was diluted until a final concentration

of 2.38 ng cDNA/µL. qPCR reactions were carried out in 96-well plates using 0.3 µM primers, 2x Takyon Low ROX SYBR 2X MasterMix (Eurogentec), 2.38 ng cDNA and H_2O (Sigma-Aldrich) in a 10 µL reaction. Samples were run in duplicate after a heat-activation step (3 min 95 °C) by a 2-step real-time protocol of 45 cycles (95 °C 15 sec, 60 °C 60 sec) on a Viia7 analyzer (ThermoFisher), combined with melting curve analysis (65 °C to 95 °C, gradually increasing with 0.5 °C/5 sec). Ct thresholds were automatically determined by the QuantStudio™ Real-Time PCR Software.

Primers sequences were extracted from PrimerBank (https://pga.mgh.harvard.edu/primerbank/) or in-house developed using the NCBI primer pick tool (overview in Table S3). All primers for transcript evaluation were purchased at IDT Technologies. Primers were *a priori* evaluated on in-house generated reference material, also used as interrun calibrator (IRC), composed by cDNA mixtures of 20 different leukemic cell lines. Primer efficiency was calculated by the LinRegPCR software (AMC, University of Amsterdam, the Netherlands) using a fivefold dilution serie, and approved if efficiency ranged between 90% and 110%. Housekeeping genes for normalisation were *a priori* selected in a pilot study [7].

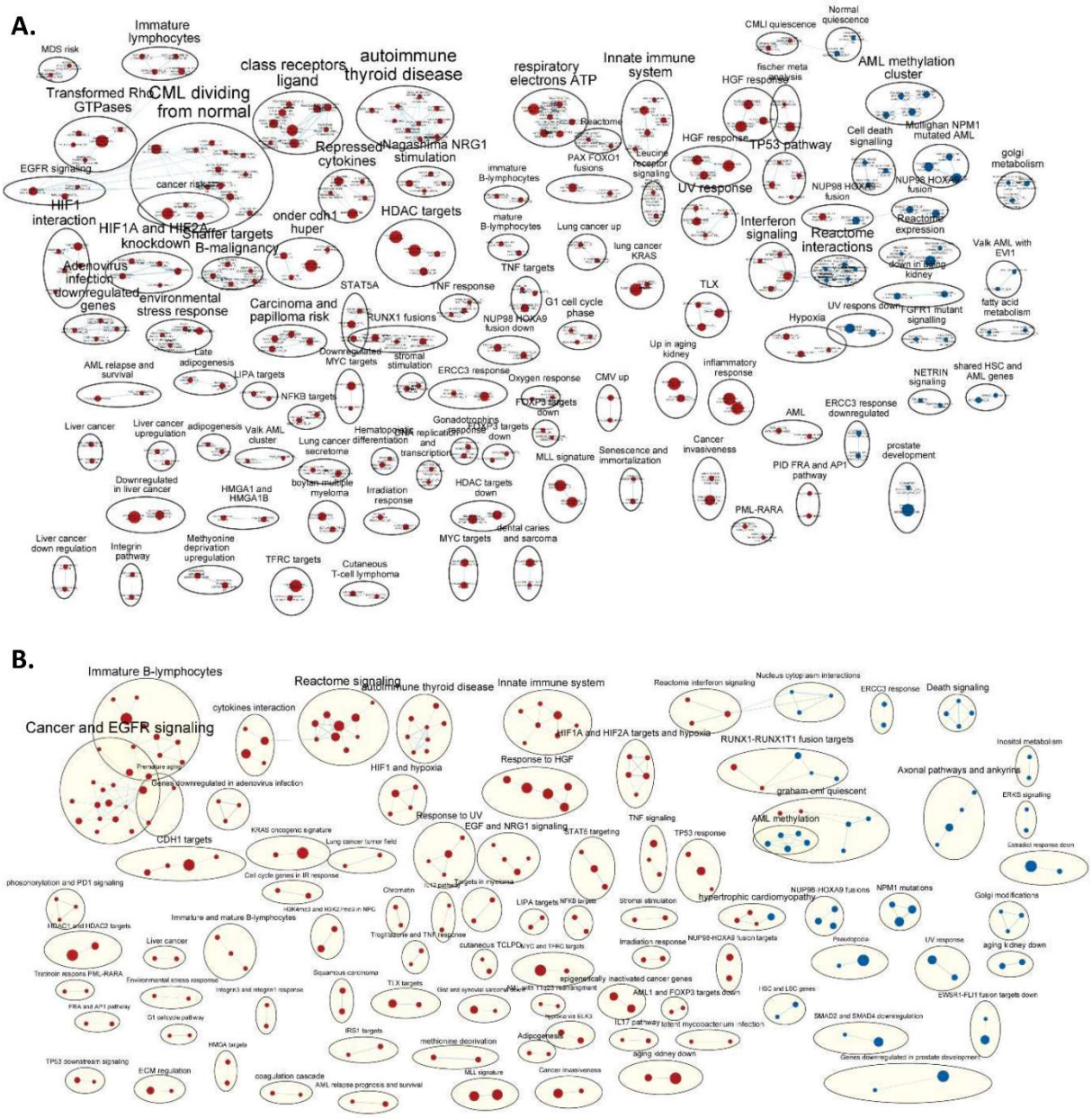
Statistical analysis

Graphics and statistical calculations were made in GraphPad Prism (version 5.04, La Jolla California USA) or SPSS (version 25.0.0.2, Inc., Chicago, IL). The Mann-Whitney U test was applied as a non-parametric test for independent samples from two groups. Wilcoxon matched-pairs signed rank tests was used to compare expression levels between PB and BM and between LSC phenotypes. The Pearson chi squared test ($n \geq 5$) or the Fisher's exact test ($n < 5$) was used to evaluate associations between the expression of DEGs and continuous and dichotomous variables. P-values calculated were one- or two-tailed, depending on the comparison.

To assess the prognostic impact of DEGs expression, e.g. impact on event-free survival (EFS) and overall survival (OS), only pedAML patients treated according to the NOPHO-DBH AML2012 protocol were included (n=14 for LSC targets and n=13 for L-blast targets). The Kaplan Meier method was used to estimate the survival probabilities for EFS and OS. Follow-up time was censored at the last follow-up visit if no failure was observed. Univariate regression analysis was performed by the Kaplan-Meier log-rank test, and confirmed by the univariate COX regression log-rank test if significant ($P < .05$), also used to calculated hazard ratio's. Confirmation of significance in univariate models was performed by a multivariate Cox proportional hazard model, taking into account all aforementioned continuous and dichotomous variables.

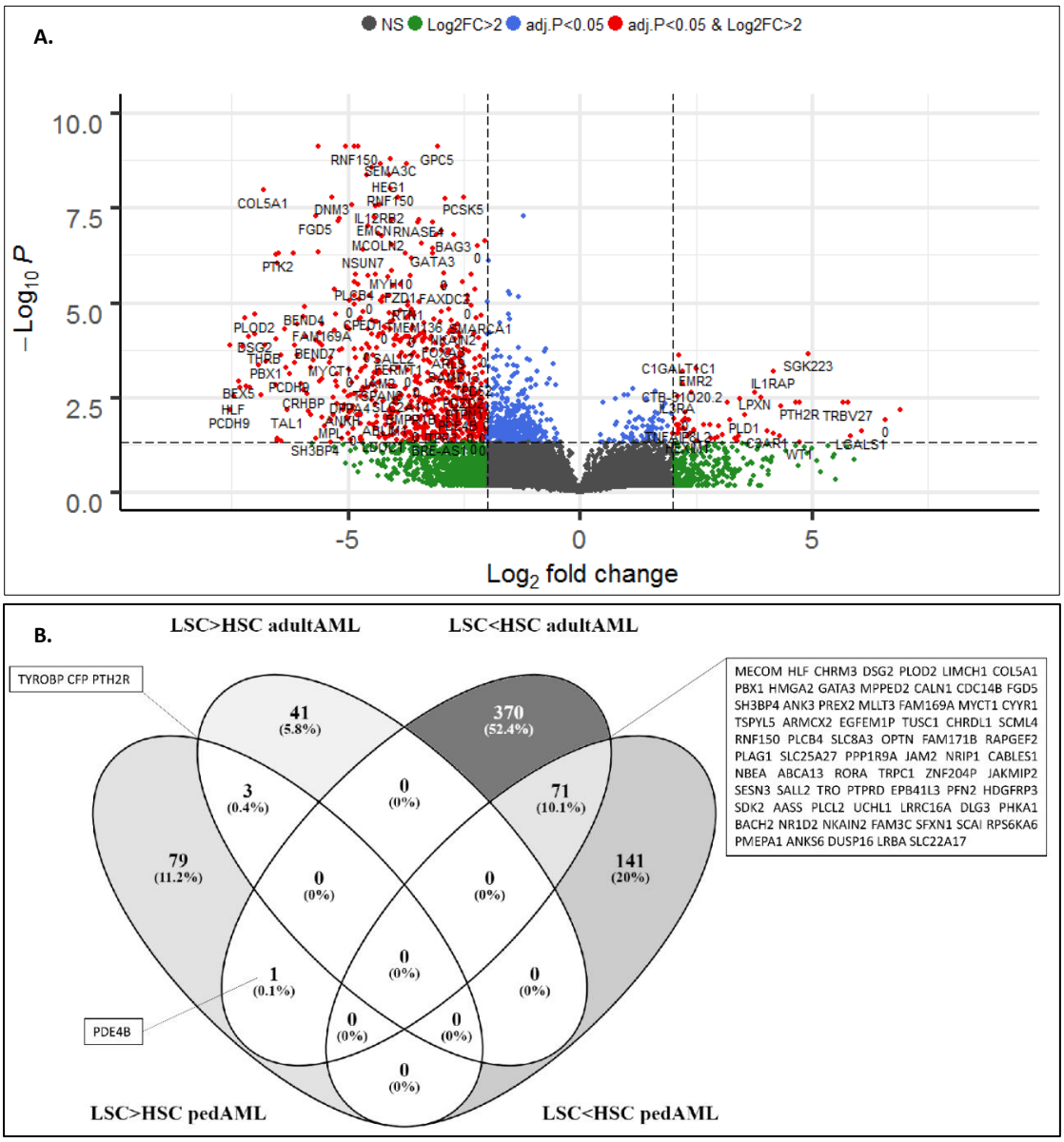
Supplemental Figures

Fig. S1. Unsupervised organization and visualization (Cytoscape) of significant gene sets identified through GSEA.



Enriched gene sets are indicated in red, suppressed gene sets are indicated in blue. (A) Leukemic stem cells (LSC) versus hematopoietic stem cells (HSC). (B) Leukemic blasts (L-blast) versus control blasts (C-blast). GSEA, gene set enrichment analysis.

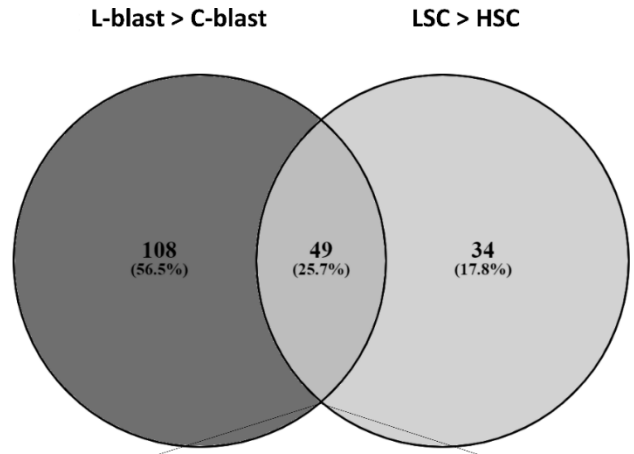
Fig. S2. Comparison of significantly up- and downregulated genes between LSC and HSC in pedAML versus adult AML.



(A) DEGs identified in adult leukemic stem cells (LSC, n=9) and hematopoietic stem cells (HSC, n=4) based on the dataset from Majeti et al. (GSE 17054). Thresholds |log2FC| > 2 and adj. P-value < .05 are shown as dashed lines, and genes selected as significantly different are highlighted in red. An overview of the total number of significant up- and downregulated transcripts is shown in Table S7-8. **(B)** Intersection of significant DEGs identified in pedAML (83 upregulated, 212 downregulated) and adult AML (44 upregulated, 442 downregulated). PedAML indicates pediatric AML; DEGs, differentially expressed genes; FC, fold change; LSC, leukemic stem cell; HSC, hematopoietic stem cell; L-blast, leukemic blast; C-blast, control blast; >, upregulated; <, downregulated.

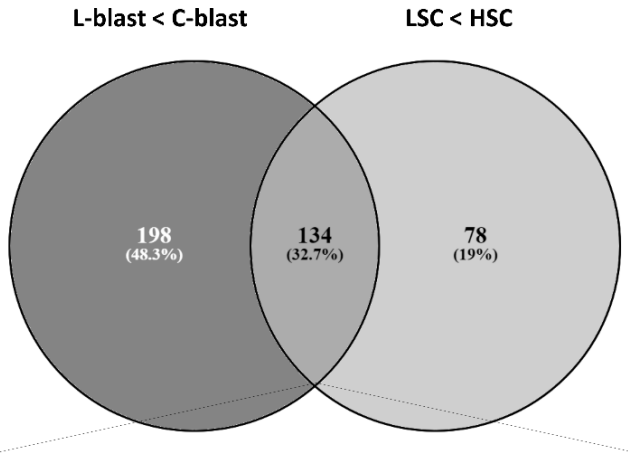
Fig. S3. Intersection of (A) mutual upregulated and (B) mutual downregulated DEGs identified, based on the CvN approach, in both LSC and L-blast compartments of pedAML (discovery cohort).

A.



C13orf15 FOSB PLIN2 GADD45B TCTEX1D4 CDKN1A DUSP5 EMP1 DUSP6 PLK3 CFD TRIB1 SLED1 CD68 ATF3 SERTAD1 NLRP3 CITED2 SCPEP1 CITED4 TYROBP DOK2 NDRG1 ARL4A OBFC2A ALDH3B1 MILR1 SLC31A2 TOB1 CFP STK17B TUBB4B ANXA2 SLC2A14 ANXA2P1 CRNDE KLF6 JUN CTSA LOC439911 FAM166A PTH2R SPRY2 LRPAP1 PTGER2 ANXA2P3 CRIP1 CD97 GPX1

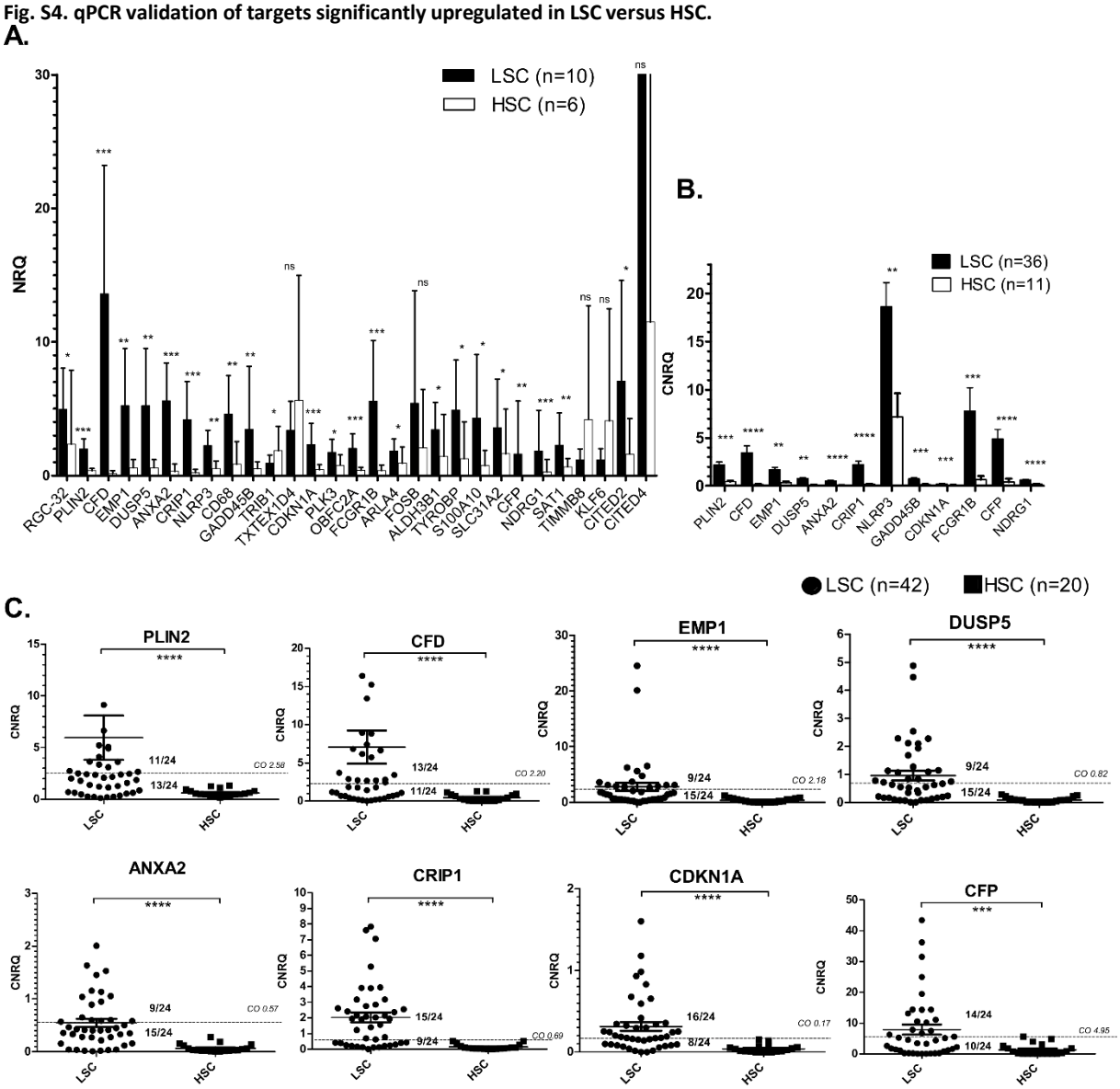
B.



MECOM HOXA9 TRO HLF HTR1F FGD5 HOXA10 BEX2 PBX1 PLCB4 HOXA3 TSPYL5 PTPRD LRRC16A LAPTM4B MYCT1 DSG2 FAM169A LOC100132526 HOXA6 PLAG1 CALCRL BACH2 AIG1 RAPGEF2 NKAIN2 UCHL1 LIN28B SLC22A15 COL5A1 SH3BP4 MEG3 CPNE8 ST8SIA6 ETNK2 SYTL4 HOXA4 AASS SGPP2 SH3BGRL2 DLG3 SAMD12 HDGFRP3 HMGA2 SCN9A ZC3H12C PREX2 TMEM163 CYYR1 ZNF204P SCML4 DUSP16 MLLT3 CHRDL1 NUDT10 SOX12 PLCL2 PPP1R9A PPP4R1L PAWR CNN3 SCAI PAFAH1B1 C3orf27 JAM2 EGFEM1P LOC254057 SNX25 NUDT9P1 TUSC1 BCL11A UBAP2 CFH FAM3C SESN3 TPTE RAPGEF6 PLEKHG6 RNF150 CCDC136 HIVEP2 SFXN1 FAM171B PLOD2 NBEA LINC00304 TRIM50 KIAA1257 GOSR1 NFATC2 RC3H2 EXT1 OPA1 TAOK3 LRBA TRIM24 SLC8A3 CNOT4 DENND5A EPB41L3 NRSN2 DMRTC1 JAM3 SPG20 CABLES1 MPPED2 ARMCX2 PMEPA1 PLA2G10 OPTN WHAMMP3 SBNO1 LOC642776 TTN CDC14B PFN2 LPIN1 MSI2 C17orf69 RXFP4 FARP1 PCF11 COL6A2 CRYGD USP7 TRPC1 N4BP3 CHRM3 BTBD11 FLJ44342 ATP9A HBG1 CALN1 CYLD

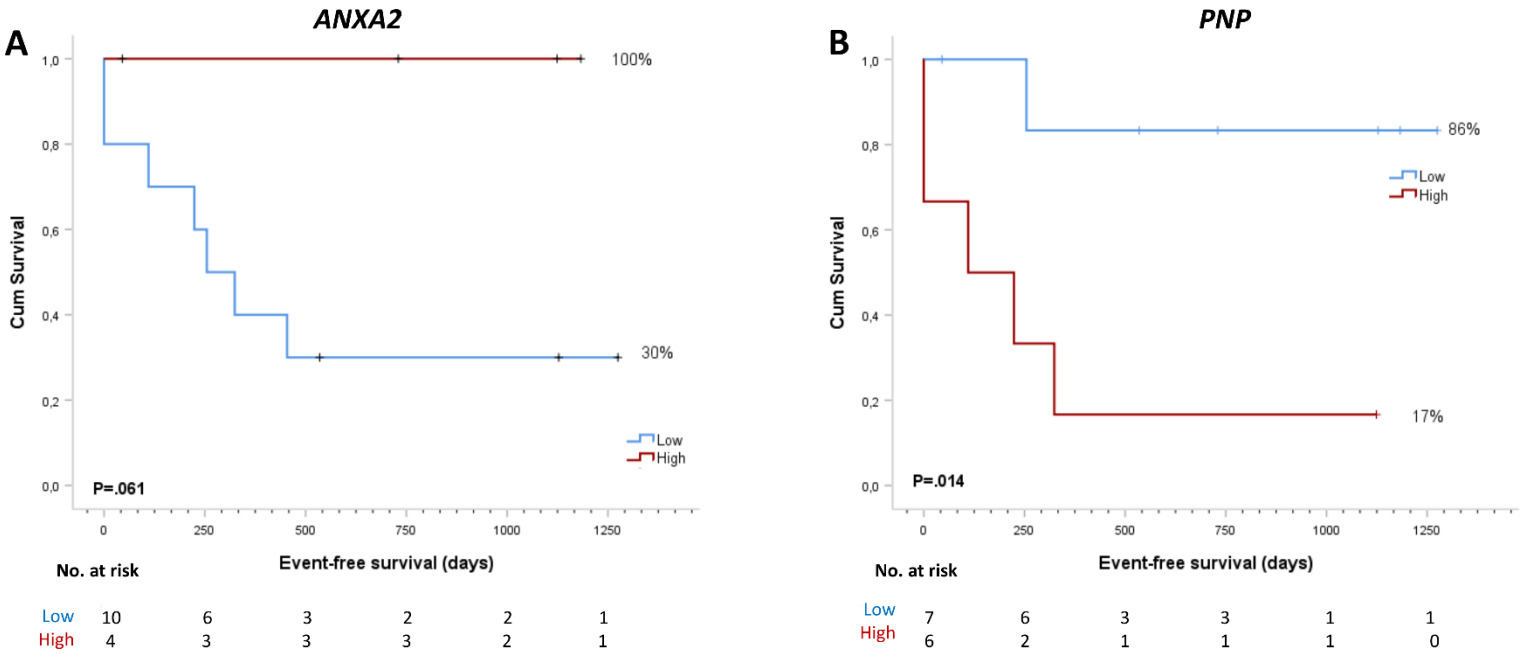
PedAML indicates pediatric AML; DEGs, differentially expressed genes; CvN, Cancer versus Normal approach; LSC, leukemic stem cell; L-blast, leukemic blast.

Fig. S4. qPCR validation of targets significantly upregulated in LSC versus HSC.



Significantly upregulated targets in LSC versus HSC, identified by micro-array analysis, were validated by qPCR in three steps and expressed as (calibrated) normalized relative quantities (CNRQ). (A) 29/83 LSC upregulated targets were selected for qPCR validation. (B) 12 out of these 29 targets were further evaluated in a second step. (C) 8/12 targets were selected for evaluation in the third step, using all available LSC (n=42) and HSC (n=20) fractions. Mean values are shown by horizontal lines, error bars indicate the 95% confidence interval of the mean. The dotted line indicates the cut-off (CO) used to define overexpression, per target, with the respective numbers of patients classified as high or low indicated above or below the line, respectively. P-values (one-tailed) were calculated by the Mann-Whitney U test, and one, two, three or four asteriks indicate the level of significance (P<.05, P<.01, P<.001 and P<.0001, respectively). LSC, leukemic stem cell; HSC, hematopoietic stem cell; DEG, differentially expressed gene.

Fig. S5. Association between EFS and *ANXA2* and *PNP* expression.

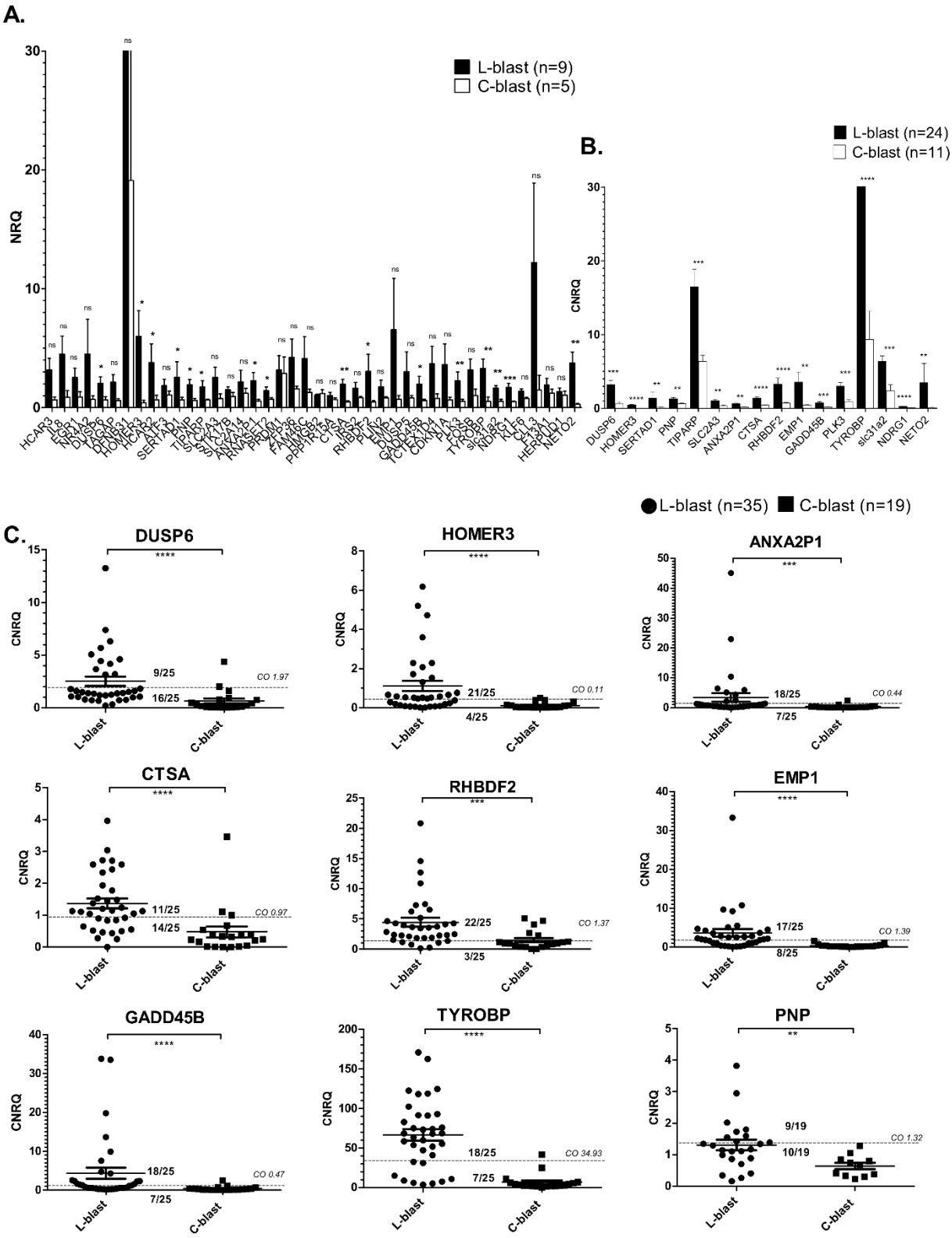


In both (A) and (B), the number of patients and P-values are shown at the bottom of each curve, and EFS percentages are shown next to the curves. Patients were dichotomized as 'high' (red) and 'low' (blue). EFS indicates event-free survival; LSC, leukemic stem cell; HSC, hematopoietic stem cell; L-blast, leukemic blast; C-blast, control blast; P, P-value; CNRQ, calibrated normalized relative quantity.

(A) EFS was calculated for NOPHO-DBH AML2012 treated pedAML (n=14/24, Table 1) and correlated to *ANXA2* expression measured in the LSC compartment (Kaplan-Meier log-rank test).

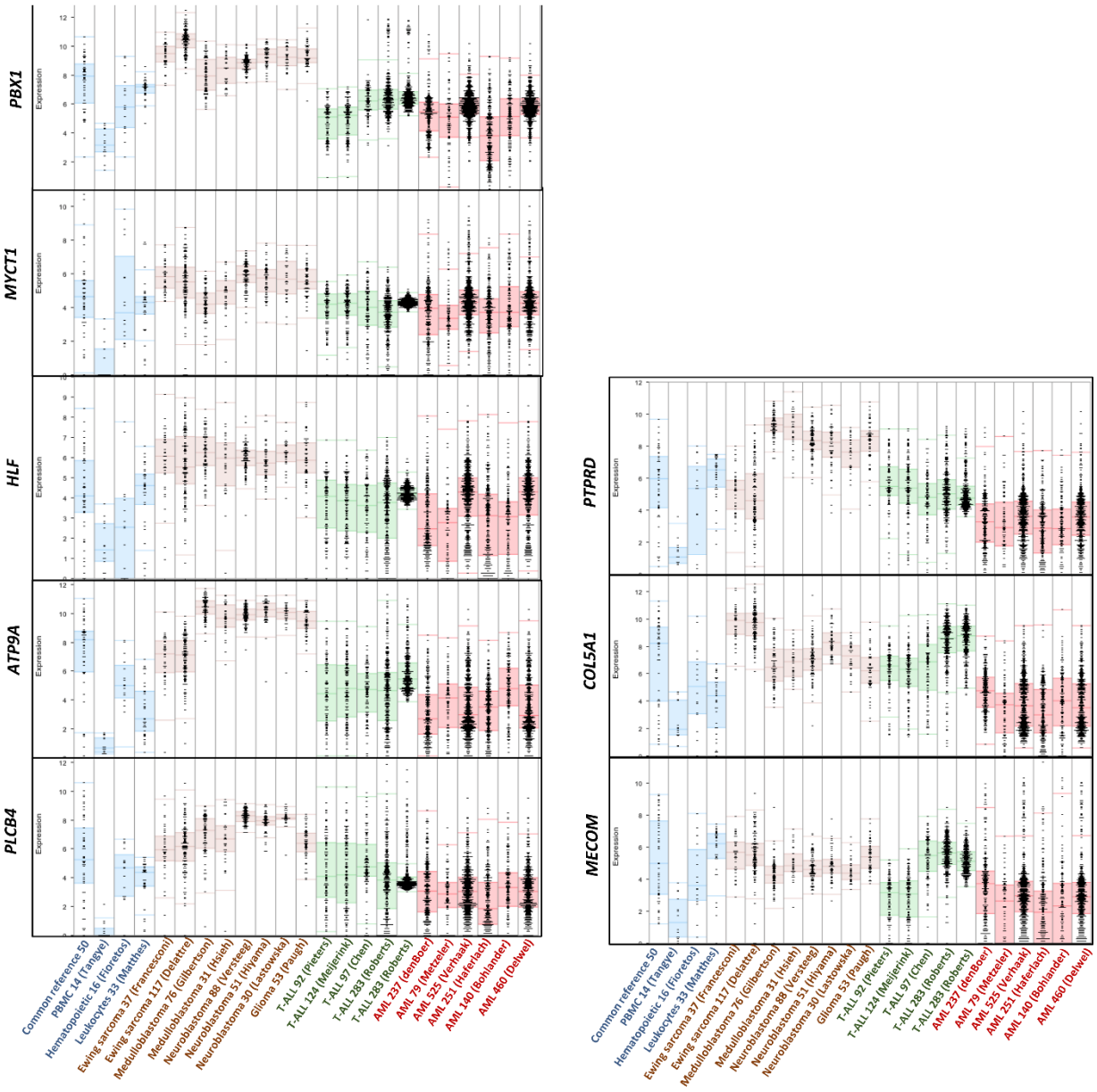
(B) EFS was calculated for NOPHO-DBH AML2012 treated pedAML (n=13/25, Table 1) and correlated to *PNP* expression measured in the L-blast compartment (Kaplan-Meier log-rank test).

Fig. S6. qPCR validation of targets significantly upregulated in L-blast versus C-blast.



Targets with significant upregulated expression in L-blast compared to C-blast, demonstrated by micro-array analysis, were evaluated by qPCR analysis in three steps and expressed as (calibrated) normalized relative quantities (CNRQ). (A) 42/157 L-blast upregulated DEGs were selected for qPCR validation. (B) 16 out of these 42 targets were further evaluated in a second step. (C) 9/16 targets were selected for evaluation in the third step, using all available L-blast (n=35) and C-blast (n=19) fractions. P-values (one-tailed) were calculated by the Mann-Whitney U test. Mean values are shown by horizontal lines, error bars indicate the 95% confidence interval of the mean. One, two, three or four asteriks indicate the level of significance (P<.05, P<.01, P<.001 and P<.0001, respectively). L-blast, leukemic blast; C-blast, control blast.

Fig. S7. Expression of qPCR-validated LSC-downregulated targets across 23 publicly available micro-array datasets.



Affymetrix datasets (u133p2, MAS 5.0) derived from R2 Genomics Analysis and Visualization Platform Expression were evaluated in normal tissues (n=4, blue) next to pediatric T-cell acute lymphoblastic leukemia (T-ALL) (n=5, green), pediatric solid cancers (n=8, brown; 2 ewing sarcoma, 2 medulloblastoma, 3 neuroblastoma, 1 glioma) and adult acute myeloid leukemia (n=6, red). Log2 transformed normalized values are shown.

Supplemental Tables

Table S1. Overview of monoclonal antibodies used for membrane staining and sorting.

Antibody	Fluorochrome	Clone	Supplier	Category no.
CD34	PerCP-Cy5.5	8G12	BD Biosciences	333146
CD38	APC-H7	HB7	BD Biosciences	656646
CD45	PacO	HI30	Invitrogen	MHCD4530
CD45RA	FITC	HI100	BD Biosciences	4343710
CLL-1	PE	50C1	BD Biosciences	6266719
CD123	PE	6H6	ThermoFisher Scientific	4290222
GPR56	PE	CG4	BioLegend	B200744

Table S2. Demographics micro-array cohort.

		Median (Range)
Age, years		14 (10-15)
WBC count, x 10^9/L		79 (58.1-118)
Morphological blast count		
	BM, %	88 (34-95)
	PB, %	74 (38-78)
		N
Gender		
	F	3
	M	1
WT1 overexpression		2
Fusion transcript		2
CBF leukemia		2
	AML1-ETO + *C-KIT*^{WT}	1
	CBFB-MYH11	1
FLT3		
	ITD	2
	WT	2
Karyotype		
	Normal	1
	Abnormal	3
FAB classification		
	M0	1
	M2	1
	M4	2

Characteristics of four *de novo* pedAML patients used for micro-array profiling of CD34+/CD38- and CD34+/CD38+ sorted cell fractions. Patients were diagnosed in Belgium, classified as standard risk, presented no central nerve system invasion and were treated in the DB AML-01 study. No *NPM1* or *CEBPA* mutations were detected. PedAML, pediatric acute myeloid leukemia; F, female; M, male; WBC, white blood cell; BM, bone marrow; PB, peripheral blood; CBF, core-binding factor; *FLT3*, fms-like tyrosine kinase receptor-3; *NPM1*, nucleophosmin; *CEBPA*, CCAAT/enhancer-binding protein alpha; ITD, internal tandem duplication; WT, wild type; MUT, mutated; FAB, French-British-American; *WT1*, Wilms' tumor 1.

Table S3. Overview of primers used and their respective sequences.

No.	Transcript	Primer category	Forward (5'-3')	Reverse (5'-3')	Reference	Primerbank ID
1	***C13orf15***	Upregulated in LSC versus HSC	AGGAACAGCTTCAGCTTCAG	GCTAAAGTTTTGTCAAGATCAGCA	Kruszewski et al. (2015)	/
2	***PLIN2***	Upregulated in LSC versus HSC	ATGGCATCCGTTGCAGTTGAT	GGACATGAGGTCATACGTGGAG	Primerbank	327199305c1
3	***CFD***	Upregulated in LSC versus HSC	GACACCATCGACCACGACC	GCCACGTCGCAGAGAGTTC	Primerbank	42544238c1
4	***EMP1***	Upregulated in LSC versus HSC	GTGCTGGCTGTGCATTCTTG	CCGTGGTGATACTGCGTTCC	Primerbank	197313788c1
5	***DUSP5***	Upregulated in LSC versus HSC	GCCAGCTTATGACCAGGGTG	GTCCGTCGGGAGACATTCAG	Primerbank	62865889c2
6	***ANXA2***	Upregulated in LSC versus HSC	TCTACTGTTCACGAAATCCTGTG	AGTATAGGCTTTGACAGACCCAT	Primerbank	216547999c1
7	***CRIP1***	Upregulated in LSC versus HSC	CCTGCCTGAAGTGCGAGAAAT	CCTTTAGGCCCAAACATGGC	Primerbank	188595726c1
8	***NLRP3***	Upregulated in LSC versus HSC	GATCTTCGCTGCGATCAACAG	CGTGCATTATCTGAACCCCAC	Primerbank	208879435c1
9	***CD68***	Upregulated in LSC versus HSC	CTTCTCTCATTCCCCTATGGACA	GAAGGACACATTGTACTCCACC	Primerbank	91199547c2
10	***GADD45B***	Upregulated in LSC versus HSC	TACGAGTCGGCCAAGTTGATG	GGATGAGCGTGAAGTGGATTT	Primerbank	299782594c1
11	***TRIB1***	Upregulated in LSC versus HSC	CTTAGGAAGTTCGTCTTCTCCAC	GGCAGCCATGTTTGTCTGAC	Primerbank	40788016c3
12	***TCTEX1D4***	Upregulated in LSC versus HSC	CTAACAGCCTTTAACCTTCTCAGC	GAGTCTTTGGCATTCTCCTCCT	In-house	/
13	***CDKN1A***	Upregulated in LSC versus HSC	TGGAGACTCTCAGGGTCGAAA	AGAAGATGTAGAGCGGGCCT	In-house	/
14	***FCGR1B***	Upregulated in LSC versus HSC	TGGGTCAGCGTGTTCCAAG	GTCACTTCGCCCTGAGAGAC	Primerbank	292023a3
15	***PLK3***	Upregulated in LSC versus HSC	CCCTGAGGCGGATGTATGG	GTCAGCCGTCTCAAAGGGAG	Primerbank	41872373c3
16	***OBFC2A***	Upregulated in LSC versus HSC	CAGTGAACCCAACCCAGATTATC	GGTCCAAATGTACCTGTACCCAT	Primerbank	362999017c1
17	***ARL4A***	Upregulated in LSC versus HSC	TGTTGTGGACTCTGTTGATGTC	AGCTCAGTTCACCCATTGCTA	Primerbank	306482559c1
18	***FOSB***	Upregulated in LSC versus HSC	GCTGCAAGATCCCCTACGAAG	ACGAAGAAGTGTACGAAGGGTT	Primerbank	166999782c1
19	***ALDH3B1***	Upregulated in LSC versus HSC	GCCCTGGAACTATCCGCTG	CGTTCTTGCTAATCTCCGATGG	Primerbank	71773289c1
20	***TYROBP***	Upregulated in LSC versus HSC	ACTGAGACCGAGTCGCCTTAT	ATACGGCCTCTGTGTGTTGAG	Primerbank	291045270c1
21	***S100A10***	Upregulated in LSC versus HSC	GGCTACTTAACAAAGGAGGACC	GAGGCCCGCAATTAGGGAAA	Primerbank	115298655c1
22	***slc31a2***	Upregulated in LSC versus HSC	TACAGCGGTGCTTCTGTTTG	GGAGCACCAACACCGAAAGG	In-house	/
23	***CFP***	Upregulated in LSC versus HSC	CTACCAGAAACGTAGTGGTGGG	CTCAGAGCACGTCACCGAAC	Primerbank	223671862c2
24	***NDRG1***	Upregulated in LSC versus HSC	CCAACAAAGACCACTCTCCTC	CCATGCCCTGCACGAAGTA	Primerbank	207028746c3
25	***CITED4***	Upregulated in LSC versus HSC	CCTGGCATACGGCTCCTTC	AGACTGCAGGTGCGTGCTAC	Bezzerides et al. (2016)	/
26	***CITED2***	Upregulated in LSC versus HSC	CCTAATGGGCGAGCACATACA	GGGGTAGGGGTGATGGTTGA	Primerbank	51807294c1
27	***SAT1***	Upregulated in LSC versus HSC	ACCCGTGGATTGGCAAGTTAT	TGCAACCTGGCTTAGATTCTTC	Primerbank	239835753c1
28	***TIMM8B***	Upregulated in LSC versus HSC	TCTCGCACTGAAAATTGTCTCTC	GCAAACCGACTGGTGATGG	Primerbank	256773259c2
29	***KLF6***	Upregulated in LSC versus HSC	GGCAACAGACCTGCCTAGAG	CTCCCGAGCCAGAATGATTTT	Primerbank	236460554c1
30	***MECOM***	Downregulated in LSC versus HSC	TGA TCT GCT AGT TCA GCC TTA	GTT CAT GAA GAG CGA AGA CT	In-house	/
31	***HLF***	Downregulated in LSC versus HSC	CCCTCGGTCATGGACCTCA	ACTTGGTGTATTGCGGTTTGC	Primerbank	94983916c3
32	***TRO***	Downregulated in LSC versus HSC	TCATCAGCACTCAGGAATCCTCT	GAATCACCATGAGGAGACCCAG	In-house	/
33	***PLCB4***	Downregulated in LSC versus HSC	TTGACAGATACGAGGAGGAATCC	GAGGGAGCATTCTAGCACCTG	Primerbank	289547593c1
34	***PXDN***	Downregulated in LSC versus HSC	AATCAGAGAGATCCAACCTGGG	AATGCTCCACTAGGTATCCTCTT	Primerbank	109150415c1

35	***HOXA3***	Downregulated in LSC versus HSC	ATGCAAAAAGCGACCTACTACG	TACGGCTGCTGATTGGCATTA	Primerbank	84043947c1
36	***PLAG1***	Downregulated in LSC versus HSC	ATCACCTCCATACACACGACC	AGCTTGGTATTGTAGTTCTTGCC	Primerbank	167857795c1
37	***ATP9A***	Downregulated in LSC versus HSC	CAGTTTGTTCCCGAAATGAGACT	GCACGTAGCATCGGATCTCC	Primerbank	65301138c2
38	***PTPRD***	Downregulated in LSC versus HSC	TACAACCCTTACGGACTCCGA	GGAATTTGATCTTCCCGCAAAAC	Primerbank	283484023c2
39	***COL5A1***	Downregulated in LSC versus HSC	TACCCTGCGTCTGCATTTCC	GCTCGTTGTAGATGGAGACCA	Primerbank	89276750c2
40	***DSG2***	Downregulated in LSC versus HSC	TTGTTGGGTCTGTTGAAGAGTTG	TTCAGGGTATTGGGCTCATCT	Primerbank	189181754c2
41	***HOXA6***	Downregulated in LSC versus HSC	ACTACCTGCACTTTTCTCCCG	GCTCGTGTACTTCCGGTCG	Primerbank	334724446c2
42	***HTR1F***	Downregulated in LSC versus HSC	TGTGATCGCTGCAATTATTGTGA	TGTGACTGCAAGGGAACAAAT	Primerbank	148277594c2
43	***CALCRL***	Downregulated in LSC versus HSC	TCCTGAGGACTCAATTCAGTTGG	CTGTTGCAGTAAACGCCTTCT	Primerbank	144953883c1
44	***HOXA10***	Downregulated in LSC versus HSC	TGGCTCACGGCAAAGAGTG	GCTGCGGCTAATCTCTAGGC	Primerbank	83977447c2
45	***FAM169A***	Downregulated in LSC versus HSC	CTCGAAGTGGTAATCTAAAGCGG	CGTGCAGTAGACTCCAATTTGT	Primerbank	380861664c3
46	***ETNK2***	Downregulated in LSC versus HSC	GCCCCGGCTTTTCAGGTTAAT	GGCTGGGGTTGATCTCGTT	Primerbank	186659524c1
47	***FGD5***	Downregulated in LSC versus HSC	GCTTCAAGAGACCCCAGTGT	TAGCAGTTCCTGTGCGATGAC	In-house	/
48	***BEX2***	Downregulated in LSC versus HSC	CGACGCACCTCACGTC	TCATCTTTTTCATCATTTTCCTGGT	In-house	/
49	***MYCT1***	Downregulated in LSC versus HSC	CAATCGGGCTGGTACTTGGAG	CGTGGGTGTAAGAAGACCTAGA	Primerbank	156151388c1
50	***LAPTM4B***	Downregulated in LSC versus HSC	GCCCGGAGCGATGAAGATG	CAACAGTACCACAGCATTGATGA	Primerbank	74271830c1
51	***COL6A2***	Downregulated in LSC versus HSC	GACTCCACCGAGATCGACCA	CTTGTAGCACTCTCCGTAGGC	Primerbank	115527065c1
52	***PBX1***	Downregulated in LSC versus HSC	GACAACTCAGTGGAGCATTCA	CTCTCGCAGGAGATTCATCAC	Primerbank	326320046c2
53	***HCAR3***	Upregulated in L-blast versus C-blast	GCGTTCAGACTGGAAGTTTGG	TCGTGCCACCGGAAGGTAT	Primerbank	157738693c1
54	***IL8***	Upregulated in L-blast versus C-blast	ACTGAGAGTGATTGAGAGTGGAC	AACCCTCTGCACCCAGTTTTC	Primerbank	10834978a2
55	***EGR1***	Upregulated in L-blast versus C-blast	CTACGAGCACCTGACCGC	GTGGTTTGGCTGGGGTAACT	In-house	/
56	***NR4A2***	Upregulated in L-blast versus C-blast	AGTCTGATCAGTGCCCTCGT	GATAGTCAGGGTTCGCCTGG	In-house	/
57	***DUSP6***	Upregulated in L-blast versus C-blast	GCAAATCCCCATCTCGGATCA	TGCCAGCCAAGCAATGTACC	In-house	/
58	***TAGAP***	Upregulated in L-blast versus C-blast	CCAACCTCCTGCTACTCAAGC	ACCTTGTTGTTCAGGTCCTTCT	Primerbank	23199970c1
59	***DFNB31***	Upregulated in L-blast versus C-blast	CTGGATGAGTACCGTGGTGG	GTGGTCGAAGCGTTCTAGGTC	Primerbank	290746375c2
60	***HOMER3***	Upregulated in L-blast versus C-blast	AGGGAGCAGCCAATCTTCAG	CCCACTGCCCGAACTTCTG	Primerbank	224809413c1
61	***HCAR2***	Upregulated in L-blast versus C-blast	ATGTTGGCTATGAACCGCCAG	GCTGCTGTCCGATTGGAGA	Primerbank	41152145c1
62	***ATF3***	Upregulated in L-blast versus C-blast	CCTCTGCGCTGGAATCAGTC	TTCTTTCTCGTCGCCTCTTTTT	Primerbank	346223459c1
63	***SERTAD1***	Upregulated in L-blast versus C-blast	TGATTGGTTGTGGGTGGCTA	CGTTTCCGCTTCAGACCCTT	In-house	/
64	***PNP***	Upregulated in L-blast versus C-blast	ATGGAGAACGGATACACCTATGA	GAGGTCGGTGCTTAGTGTGAG	Primerbank	270288734c1
65	***TIPARP***	Upregulated in L-blast versus C-blast	AATTTGACCAACTACGAAGGCTG	CAGACTCGGGATACTCTCTCC	Primerbank	296080690c3
66	***SLC2A3***	Upregulated in L-blast versus C-blast	GCTGGGCATCGTTGTTGGA	GCACTTTGTAGGATAGCAGGAAG	Primerbank	221136810c1
67	***STK17B***	Upregulated in L-blast versus C-blast	GCCTGTGTTTACCTGAGTTGG	TGTCCCCGAGAGGGTATATGC	Primerbank	217416412c1
68	***SLC2A14***	Upregulated in L-blast versus C-blast	AGATCTCGCCTACTGCCCTG	GTAAAGCCTAATAGCACCGGC	in-house	/
69	***ANXA2P1***	Upregulated in L-blast versus C-blast	GCCCTGGCAAAGGGTAGAAG	CTCGGTCATGACGCTGATCC	In-house	/
70	***RNASET2***	Upregulated in L-blast versus C-blast	GCGAGAAAATTCAAAACGACTGT	CCTTCACTTTTATCGGGCCATAG	Primerbank	156071486c1

71	***PRDM1***	Upregulated in L-blast versus C-blast	AACTTCTTGTGTGGTATTGTCGG	CAGTGCTCGGTTGCTTTAGAC	Primerbank	353249929c2
72	***ZFP36***	Upregulated in L-blast versus C-blast	CTCATGGCCAACCGTTACACC	TCCATGGTCGGATGGCACG	In-house	/
73	***FAM46C***	Upregulated in L-blast versus C-blast	TACCAGGGATTGCATGTCCTT	CGTCCGTGGATAGGTACAACTT	Primerbank	96975056c1
74	***HMGB2***	Upregulated in L-blast versus C-blast	GCTCGCTATGACAGGGAGATG	GCGATGTTCAGAGCAAAACAGG	Primerbank	194688134c3
75	***PPP1R15A***	Upregulated in L-blast versus C-blast	CCTGAGGGTGAGATGAACGC	ACTGGGGACAGGAGGAAGAA	In-house	/
76	***CTSA***	Upregulated in L-blast versus C-blast	TGGTCTACTTTGCCTACTACCAT	CACACGGGGCATAGAGATTG	Primerbank	189163484c3
77	***IRS2***	Upregulated in L-blast versus C-blast	CGCCAGCATTGACTTCTTGTC	AAACAGTGCTGAGCGTCTTCT	In-house	/
78	***RHBDF2***	Upregulated in L-blast versus C-blast	CACCCAACCTCTCCATCACC	GAGGCTGACGCTCTTCAAGT	In-house	/
79	***CCL3***	Upregulated in L-blast versus C-blast	CGGCAGATTCCACAGAATTTCA	GCCGGCTTCGCTTGG	In-house	/
80	***F13A1***	Upregulated in L-blast versus C-blast	AGATGGGACACTAACAAGGTGG	CTGCACATAGAAAGACTGCCC	Primerbank	119395708c2
81	***HERPUD1***	Upregulated in L-blast versus C-blast	TGCTGGTTCTAATCGGGGACA	CCAGGGGAAGAAAGGTTCCG	Primerbank	58530858c2
82	***NETO2***	Upregulated in L-blast versus C-blast	AGATGGGCCATTTGGTTTCTC	TGCTCGAAATCCCAGTCCTTC	Primerbank	319996650c1
83	***PLIN2***	Upregulated in LSC versus HSC	ATGGCATCCGTTGCAGTTGAT	GGACATGAGGTCATACGTGGAG	Primerbank	327199305c1
84	***EMP1***	Upregulated in LSC versus HSC	GTGCTGGCTGTGCATTCTTG	CCGTGGTGATACTGCGTTCC	Primerbank	197313788c1
85	***DUSP5***	Upregulated in LSC versus HSC	GCCAGCTTATGACCAGGGTG	GTCCGTCGGGAGACATTCAG	Primerbank	62865889c2
86	***GADD45B***	Upregulated in L-blast versus C-blast	TACGAGTCGGCCAAGTTGATG	GGATGAGCGTGAAGTGGATTT	Primerbank	299782594c1
87	***TCTEX1D4***	Upregulated in L-blast versus C-blast	CTAACAGCCTTTAACCTTCTCAGC	GAGTCTTTGGCATTCTCCTCCT	In-house	
88	***CDKN1A***	Upregulated in L-blast versus C-blast	TGGAGACTCTCAGGGTCGAAA	AGAAGATGTAGAGCGGGCCT	In-house	
89	***PLK3***	Upregulated in L-blast versus C-blast	CCCTGAGGCGGATGTATGG	GTCAGCCGTCTCAAAGGGAG	Primerbank	41872373c3
90	***FOSB***	Upregulated in L-blast versus C-blast	GCTGCAAGATCCCCTACGAAG	ACGAAGAAGTGTACGAAGGGTT	Primerbank	166999782c1
91	***TYROBP***	Upregulated in L-blast versus C-blast	ACTGAGACCGAGTCGCCTTAT	ATACGGCCTCTGTGTGTTGAG	Primerbank	291045270c1
92	***slc31a2***	Upregulated in L-blast versus C-blast	TACAGCGGTGCTTCTGTTTG	GGAGCACCAACACCGAAAGG	In-house	/
93	***NDRG1***	Upregulated in L-blast versus C-blast	CCAACAAAGACCACTCTCCTC	CCATGCCCTGCACGAAGTA	Primerbank	207028746c3
94	***KLF6***	Upregulated in L-blast versus C-blast	GGCAACAGACCTGCCTAGAG	CTCCCGAGCCAGAATGATTTT	Primerbank	236460554c1
/	***GAPD***	Housekeeping genes	TGCACCACCAACTGCTTAGC	GGCATGGACTGTGGTCATGAG	Vandesompele et al. (2002)	/
/	***HPRT1***	Housekeeping genes	TGACACTGGCAAAACAATGCA	GGTCCTTTTCACCAGCAAGCT	Vandesompele et al. (2002)	/
/	***TBP***	Housekeeping genes	CGGCTGTTTAACTTCGCTTC	CACACGCCAAGAAACAGTGA	Bieche et al. (1999)	/

Supplemental references

1. De Moerloose B, Reedijk A, de Bock GH, Lammens T, de Haas V, Denys B, Dedeken L, van den Heuvel-Eibrink MM, Te Loo M, Uyttebroeck A et al: Response-guided chemotherapy for pediatric acute myeloid leukemia without hematopoietic stem cell transplantation in first complete remission: Results from protocol DB AML-01. Pediatric blood & cancer 2019:e27605.
2. Chair) JAS: Research study for treatment of children and adolescents with acute myeloid leukaemia 0-18 years. . NOPHO-DBH AML 2012 protocol v21 2013, EUdract number 2012-002934-35.
3. Kersten B, Valkering M, Wouters R, van Amerongen R, Hanekamp D, Kwidama Z, Valk P, Ossenkoppele G, Zeijlemaker W, Kaspers G et al: CD45RA, a specific marker for leukaemia stem cell sub-populations in acute myeloid leukaemia. British journal of haematology 2016, 173(2):219-235.
4. Zeijlemaker W, Kelder A, Wouters R, Valk PJ, Witte BI, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: Absence of leukaemic CD34 cells in acute myeloid leukaemia is of high prognostic value: a longstanding controversy deciphered. British journal of haematology 2015, Oct(171(2)):227-238.
5. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, Scholten WJ, Snel AN, Veldhuizen D, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. Leukemia 2016, 30(2):439-446.
6. Majetl R, Becker MW, Tian Q, Lee TLM, Yan XW, Liu R, Chiang JH, Hood L, Clarke MF, Weissman IL: Dysregulated gene expression networks in human acute myelogenous leukemia stem cells. Proceedings of the National Academy of Sciences of the United States of America 2009, 106(9):3396-3401.
7. Depreter B, Weening KE, Vandepoele K, Essand M, De Moerloose B, Themeli M, Cloos J, Hanekamp D, Moors I, I DH et al: TARP is an immunotherapeutic target in acute myeloid leukemia expressed in the leukemic stem cell compartment. Haematologica 2019, 104:xxx(doi: 10.3324/haematol.2019.222612. [Epub ahead of print]).

5

CHAPTER V: Results:

Long non-coding RNA expression in pediatric AML subpopulations.

Introduction

Long non-coding RNA transcripts (lncRNA) are non-coding RNAs that consist of >200 nucleotides and exert regulatory functions in gene transcription, protein translation and epigenetic regulation [1]. Their role in leukemogenesis is emerging, as it was shown that lncRNA can exert both tumor suppressive and oncogenic effects [2, 3]. However, lncRNA expression data in pediatric AML (pedAML) are scarce, with only six available publications [3-8]. Recently, a leukemic stem cell (LSC)-specific signature of 111 lncRNAs was identified using publicly available RNA sequencing datasets of cytogenetic normal (CN) adult AML [9]. In addition, it was shown that the altered lncRNA signatures present in LSC are functionally relevant and may harbour potentially interesting novel targets for therapy. However, to the best of our knowledge, such research has not yet been performed in a pediatric setting.

In the following section, we explored lncRNA expression in pedAML subpopulations (LSC and leukemic blasts (L-blast)) and their normal counterparts (hematopoietic stem cells (HSC) and control myeloblasts (C-blast), respectively). Most differentially expressed (DE)-lncRNA here identified await further validation through qPCR. Nevertheless, these preliminary data may aid in further unravelling the role of lncRNAs in the molecular pathogenesis of pedAML and pave the way for more lncRNA-orientated research.

Materials and Methods

Patient cohort and micro-array analysis

The differential expression of 25839 lncRNAs was evaluated in CD34+/CD38- and CD34+/CD38+ sorted subpopulations from four pedAML patients (LSC and L-blast, respectively) and three cord blood (CB) samples as a control population (HSC and C-blast, respectively). PedAML patient characteristics are summarized in Table 1. Cells were profiled on a custom 8x60K human gene expression micro-array, containing probes for all human protein-coding genes with lncRNA content based on LNCipedia 2.15 and performed by Biogazelle. Details on microarray hybridization are described in the Supplementary of chapter IV). Data processing was performed using EdgeR package (R Bioconductor). Principle component analysis (PCA) showed that the different subpopulations cluster correctly together based on their lncRNA content, and divergent from healthy controls (Fig.1 A-B). Moreover, LSCs and L-blasts also clustered on a per patient basis (Fig.1 C).

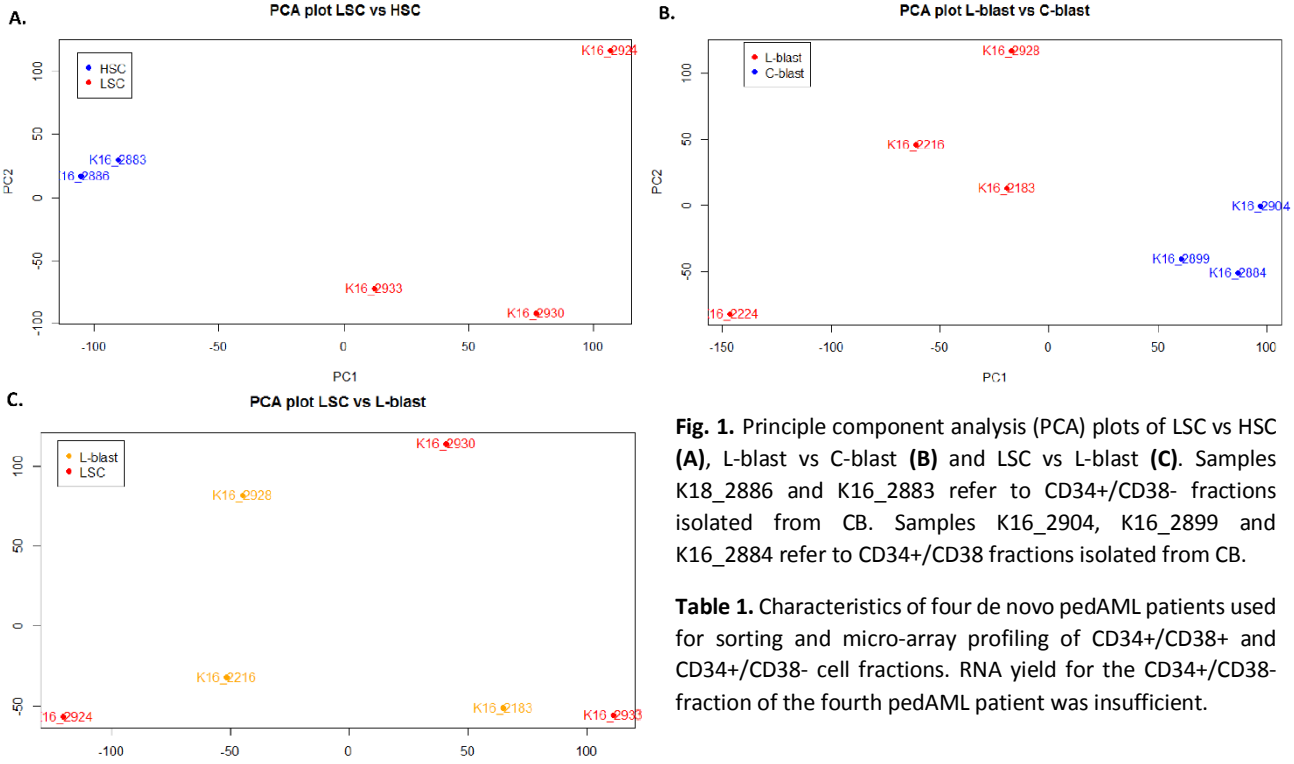


Fig. 1. Principle component analysis (PCA) plots of LSC vs HSC **(A)**, L-blast vs C-blast **(B)** and LSC vs L-blast **(C)**. Samples K18_2886 and K16_2883 refer to CD34+/CD38- fractions isolated from CB. Samples K16_2904, K16_2899 and K16_2884 refer to CD34+/CD38 fractions isolated from CB.

Table 1. Characteristics of four de novo pedAML patients used for sorting and micro-array profiling of CD34+/CD38+ and CD34+/CD38- cell fractions. RNA yield for the CD34+/CD38- fraction of the fourth pedAML patient was insufficient.

Table 1. Patient characteristics.

Patients	Sorted population	Code	Age (yrs.)	*FLT3*	*WT1*	Fusion transcript	Karyotype	FAB
pedAML1	LSC L-blast	K16_2930 K16_2928	9.7	WT	overexpression	*CBFB-MYH11*	abnormal	M4
pedAML2	LSC L-blast	K16_2924 K16_2216	14.6	ITD	overexpression	no	normal	M0
pedAML3	LSC L-blast	K16_2933 K16_2183	14.5	WT	no overexpression	*AML1-ETO* (*C-KIT*^{WT})	abnormal	M4
pedAML4	L-blast	K16_2224	12.5	ITD	no overexpression	no	abnormal	M2

Significant DE-lncRNAs between LSC and HSC, and between L-blast and C-blast, were selected based on an |log2FC|>2 and adj.P<.05. LNCipedia (version 5.2) was consulted to derive essential information regarding the lncRNAs, such as genomic location and their putative regulatory functions [10].

Pathway and gene set enrichment analysis

We used the R package LncPath to identify pathways associated with the here described DE-lncRNA sets [11]. A normalized enrichment score (NES), with concomitant false discovery rate (FDR) q-values were calculated through a permutation analysis for each pathway. A FDR q-value <.01 was considered as significant.

Pre-ranked gene set enrichment analysis (GSEA) was used to predict the functional relevance of the DE-lncRNAs. First, spearman's rho values (ρ value) were calculated between the top most significant DE-lncRNAs (Table 2) and all protein coding genes (packages "Hmisc" and "psych", R Bioconductor). Preranked gene set lists were generated for each lncRNA of interest based on this correlation matrix. These files were subsequently used for enrichment analysis using GSEA software and the C2 Molecular Signature Database with 4729 curated gene sets (http://www.broad.mit.edu/gsea/) as previously described [12].

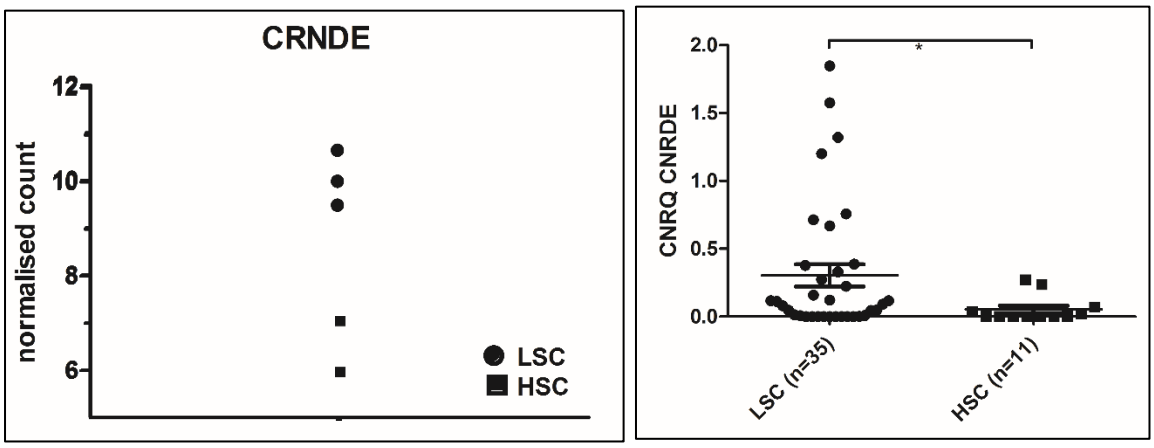
Results and discussion

Identification of differentially expressed lncRNAs in pedAML subpopulations

First, 86 significantly DE-lncRNAs were identified in LSC compared to HSC, of which 16 were upregulated and 70 downregulated. Table 2 shows the top 10 in each category. The *lnc-GSG1-1* was the highest upregulated lncRNA and was encoded by two different probes within the top 10 (log2FC 5.66 and 4.77). *Lnc-EVX1-7* (log2FC -5.85) was the most downregulated lncRNA, followed by lnc-HOXA9-1, also encoded by two different probes (log2FC -5.12 and -4.48).

In addition, we discovered that the lncRNA *CRNDE* (Colorectal Neoplasia Differentially Expressed) is overexpressed in LSC compared to HSC in pedAML patients (log2FC = 3.57, adj. P<.05) (Fig. 2A). This observation was confirmed by qPCR using a total of 35 LSC versus 11 HSC fractions (5.7-fold higher expression, P<.05) (Fig. 2B). Increased *CRNDE* expression was previously described in tumorigenic proliferating tissues [13-17]. Two recent meta-analyses confirmed the role of *CRNDE* as oncogene in multiple tumours, and showed that the level of expression significantly correlates with poor prognosis and advanced tumor progression [15, 16]. Also in haematological malignancies, i.e. AML, multiple myeloma and T-cell leukemia, *CRNDE* was reported to act as oncogene [2, 5, 18]. A recent study focussing on pediatric and adolescent CN-AML revealed that *CRNDE* expression directly correlates with the percentage of blasts (PB and BM) and WBC count, and negatively associates with overall survival (OS) [5]. Moreover, *CRNDE* was recently suggested to have therapeutic potential. In liver cancer [19], indirect silencing of *CRNDE* promoted apoptosis and inhibited proliferation, migration, invasion and drug resistance. Hence, suppression of *CRNDE* within pedAML LSC fractions as a novel targeted therapeutic strategy merits further attention.

Fig. 2. Significantly higher CRNDE expression in pedAML LSC compared to HSC.



(A) CRNDE normalized counts determined by micro-array analysis in 3 LSC (3 pedAML) versus 2 HSC (CB) fractions. (B) CRNQ values determined by qPCR in 35 LSC (17 pedAML) versus 11 HSC (7 CB, 4 NBM) fractions. Asterisk is representative for P<.05. Primer sequences for *CRNDE* were adapted from [18]: *CRNDE*-fwd: TGGATGCTGTCAGCTAAGTTCAC; *CRNDE*-rev: TTCCAGTGGCATCCTCCTTATC. Expression were normalized against housekeeping genes *GAPD*, *HPRT1* and *TBP*.

Second, we identified 146 significant DE-lncRNAs in L-blast compared to C-blast, of which 42 were upregulated and 104 downregulated, with the respective top 10 ranked lncRNAs shown in Table 2. L*nc-CD96-1* is significantly upregulated in L-blast. Taken into account the cis-regulatory effect lncRNAs might exert on neighbouring genes, this observation is in agreement with our previous finding that *CD96* mRNA expression is significantly upregulated in L-blast versus C-blast (log2FC 5.18, adj. P<.01, GSE 128103, see chapter IV).

Table 2. Top 10 significantly up- and downregulated lncRNAs in LSC and L-blast.

Probe ID	Lncipedia Transcript ID	Lncipedia Gene ID	log2FC	adj.P.Val
LSC > HSC				
PVD_LNCIPEDIA_2013_7311	lnc-GSG1-1:1	**lnc-GSG1-1**	5.66	1.0E-02
PVD_LNCIPEDIA_2013_13102	lnc-RNFT2-4:1	**lnc-RNFT2-4**	5.17	4.2E-02
PVD_LNCIPEDIA_2013_23031	lnc-RGMA-1:1	**lnc-RGMA-1**	4.88	3.1E-02
PVD_LNCIPEDIA_2013_20744	lnc-GSG1-1:1	**lnc-GSG1-1**	4.77	1.4E-02
PVD_LNCIPEDIA_2013_8971	lnc-LHFPL3-1:1	**lnc-LHFPL3-1**	4.44	1.4E-02
PVD_LNCIPEDIA_2013_6482	lnc-FOS-2:2	**lnc-FOS-2**	4.14	4.8E-02
PVD_LNCIPEDIA_2013_11894	lnc-PLD1-2:1	**lnc-PLD1-2**	3.53	4.2E-02
PVD_LNCIPEDIA_2013_8225	lnc-ITSN1-2:2	**lnc-ITSN1-2**	3.10	4.2E-02
PVD_LNCIPEDIA_2013_23873	lnc-CHST2-2:1	**lnc-CHST2-2**	3.01	4.2E-02
PVD_LNCIPEDIA_2013_13128	lnc-ROM1-3:1	**lnc-ROM1-3**	2.94	4.2E-02
LSC < HSC				
PVD_LNCIPEDIA_2013_5803	lnc-EVX1-7:1	**lnc-EVX1-7**	-5.85	3.1E-02
PVD_LNCIPEDIA_2013_23465	lnc-HOXA9-1:4	**lnc-HOXA9-1**	-5.12	3.4E-02
PVD_LNCIPEDIA_2013_3136	lnc-CALCOCO2-7:1	**lnc-CALCOCO2-7**	-5.06	3.1E-02
PVD_LNCIPEDIA_2013_2328	lnc-C15orf2-3:1	**lnc-C15orf2-3**	-4.87	3.1E-02
PVD_LNCIPEDIA_2013_5107	lnc-DLK1-8:10	**lnc-DLK1-8**	-4.86	5.0E-02
PVD_LNCIPEDIA_2013_5787	lnc-ETV3-2:1	**lnc-ETV3-2**	-4.81	1.4E-02
PVD_LNCIPEDIA_2013_22814	lnc-FAM43A-11:1	**lnc-FAM43A-11**	-4.71	3.1E-02
PVD_LNCIPEDIA_2013_21273	lnc-NRIP1-2:1	**lnc-NRIP1-2**	-4.49	2.7E-02
PVD_LNCIPEDIA_2013_7684	lnc-HOXA9-1:1	**lnc-HOXA9-1**	-4.48	2.7E-02
PVD_LNCIPEDIA_2013_4475	lnc-CSMD1-1:2	**lnc-CSMD1-1**	-4.47	4.8E-02

L-blast > C-blast				
PVD_LNCIPEDIA_2013_19559	lnc-RNFT2-4:1	**lnc-RNFT2-4**	5.71	2.9E-03
PVD_LNCIPEDIA_2013_13850	lnc-RTN2-1:1	**lnc-RTN2-1**	5.64	3.4E-04
PVD_LNCIPEDIA_2013_7311	lnc-GSG1-1:1	**lnc-GSG1-1**	5.26	6.1E-04
PVD_LNCIPEDIA_2013_3533	lnc-CD96-1:1	**lnc-CD96-1**	5.06	4.5E-03
PVD_LNCIPEDIA_2013_13102	lnc-RNFT2-4:1	**lnc-RNFT2-4**	4.93	2.6E-04
PVD_LNCIPEDIA_2013_13100	lnc-RNFT2-2:1	**lnc-RNFT2-2**	4.49	2.8E-04
PVD_LNCIPEDIA_2013_4946	lnc-DENR-2:1	**lnc-DENR-2**	4.43	7.8E-03
PVD_LNCIPEDIA_2013_23031	lnc-RGMA-1:1	**lnc-RGMA-1**	4.20	6.2E-04
PVD_LNCIPEDIA_2013_8971	lnc-LHFPL3-1:1	**lnc-LHFPL3-1**	3.91	3.0E-03
PVD_LNCIPEDIA_2013_20744	lnc-GSG1-1:1	**lnc-GSG1-1**	3.84	1.7E-03
L-blast < C-blast				
PVD_LNCIPEDIA_2013_4674	lnc-CXCL2-1:1	**lnc-CXCL2-1**	-6.97	3.5E-04
PVD_LNCIPEDIA_2013_5803	lnc-EVX1-7:1	**lnc-EVX1-7**	-6.14	2.5E-05
PVD_LNCIPEDIA_2013_10848	lnc-NRIP1-2:1	**lnc-NRIP1-2**	-5.84	2.9E-03
PVD_LNCIPEDIA_2013_15941	lnc-THADA-4:1	**lnc-THADA-4**	-5.08	3.4E-04
PVD_LNCIPEDIA_2013_23465	lnc-HOXA9-1:4	**lnc-HOXA9-1**	-4.76	8.4E-04
PVD_LNCIPEDIA_2013_5107	lnc-DLK1-8:10	**lnc-DLK1-8**	-4.74	5.2E-03
PVD_LNCIPEDIA_2013_7684	lnc-HOXA9-1:1	**lnc-HOXA9-1**	-4.57	2.5E-04
PVD_LNCIPEDIA_2013_11034	lnc-OBFC2A-5:1	**lnc-OBFC2A-5**	-4.47	1.3E-03
PVD_LNCIPEDIA_2013_5108	lnc-DLK1-8:1	**lnc-DLK1-8**	-4.28	7.0E-03
PVD_LNCIPEDIA_2013_9456	lnc-MAP7-1:1	**lnc-MAP7-1**	-4.14	2.8E-04

Legend to Table 2. Available information was researched on LNCipedia (version 5.2). "<" means "underexpressed compared to", ">" means "overexpressed compared to". Representation of the lncRNA by two different probes within the top 10 is indicated by underlining. NA, not available.

In conclusion, pedAML LSC and L-blast populations show altered lncRNA expression compared to their healthy counterparts. Importantly, several of the DE-lncRNAs identified are found in both comparisons. *Lnc-GSG1-1*, *lnc-RNFT2*-4, *lnc-RGMA-1* and *lnc-LHFPL3-1* were upregulated in both LSC and L-blast compartment, whilst *lnc-EVX1-7*, *lnc-HOXA9-1*, *lnc-DLK1-8* and *lnc-NRIP1-2* were mutually downregulated. Targetability of these lncRNAs merits further investigation.

Identification of pathways regulated by LSC and L-blast specific lncRNAs

As the functional role of most lncRNAs still needs to be elucidated [3], we aimed to identify pathways associated with the here described DE-lncRNA sets.

The LSC upregulated lncRNA-gene set identified nine significantly enriched KEGG annotated pathways (NES range 1.17 – 2.59), with the majority (n=6/9) involved in general cell metabolisms, i.e. "ribosome", "basal transcription factors", "DNA replication", "spliceosome", "nucleotide excision repair" and "neuroactive ligand receptor interaction". The top 10 upregulated lncRNAs in L-blast affected six pathways at a significant level (NES range 1.37-1.92), with the "glycolysis/gluconeogenesis" pathway showing the highest enrichment score. This result is in agreement with our previous finding based on the identified differentially expressed mRNA transcripts. In chapter IV, we show that L-blasts undergo unique metabolic alterations that enhance proliferation compared to their normal counterparts. The set of lncRNAs significantly downregulated in LSC and L-blast were both shown to regulate 14 KEGG pathways (LSC: NES range 0.94-1.85, L-blast: 1.00-2.00), with "snare interactions in vesicular transport", "cell adhesion molecules CAMs" showing the highest NES score in both categories, next to the regulation of haematopoiesis ("hematopoietic cell lineage", NES 1.64).

Also here, mutual differential expression of lncRNAs in LSC and in L-blast was noticeable. Fourteen pathways were identified in more than one comparison (Fig. 3). For instance, the KEGG annotated pathway “oxidative phosphorylation” was identified at a significant level based on the top 10 downregulated lncRNAs in LSC and L-blast (NES 1.53 and 1.67, respectively), and the top 10 upregulated lncRNAs in LSC (NES 1.40). This result met our expectations, as it is well-known that hematopoietic stem and progenitor cells show low oxidative phosphorylation rates, both related to cell-specific mechanisms as well as the hypoxic environment [20].

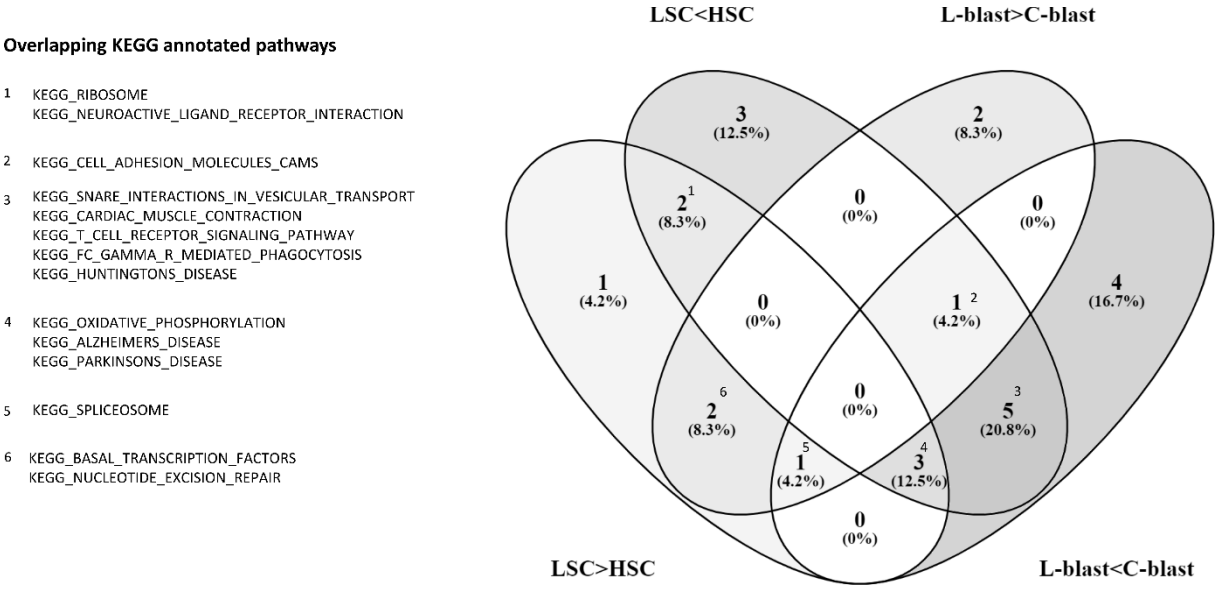


Fig. 3. A Venn plot illustrating the overlap between KEGG annotated pathways. The number of pathways shown were significantly affected by the set of DE-lncRNAs. “<” means “downregulated”, “>” means “upregulated”.

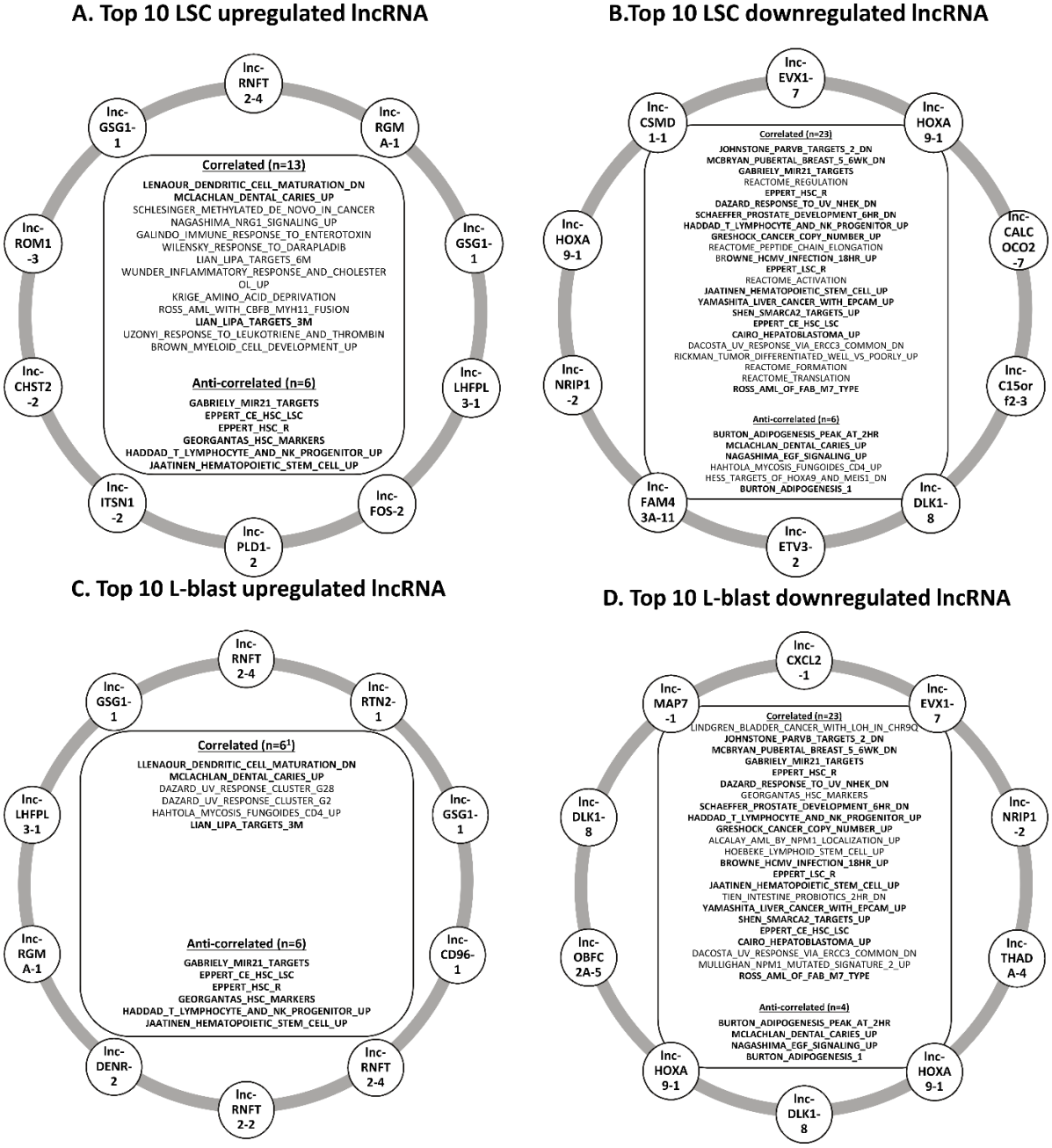
Next, we aimed to predict functional relevance of the DE-lncRNAs through pre-ranked gene set enrichment analysis (GSEA). We focused on the gene sets that were significantly (FDR q-value < .01) (anti-) correlated for each of the top 10 differentially expressed lncRNAs per comparison. An overview is shown in Fig. 4 A-D.

Gene sets that uniformly correlated with all top 10-ranked LSC-upregulated lncRNAs (n=13, Fig. 4A) were associated with inflammation and immune response (n=7/13), myeloid maturation/development (n=2/13) and cancer (n=2/13), whilst uniformly anti-correlated gene sets (n=6) mainly referred to healthy hematopoietic stem and progenitor populations (n=5/6). Gene sets significantly correlating with the top 10 LSC-downregulated lncRNAs (Fig. 4B, n=23) mainly referred to genes with downregulated expression in cancer (n=5/23), hematopoietic stem and progenitor gene expression (n=4/23) and reactome pathways (n=6/23), whilst the six anti-correlating gene sets referred to an increased inflammation status (n=4/6). Altogether, GSEA illustrates that the differentially expressed lncRNAs identified in LSC are likely to be involved in deregulated immune and inflammatory responses, and are able to discriminate between previously published LSC and HSC gene expression datasets.

Six unique gene sets were mutually significantly correlated and anti-correlated with the top 10 upregulated lncRNAs in L-blasts (Fig. 4C). The six correlating gene sets were associated with dendritic cell maturation, inflammation and inflammation-triggered cell growth and response to UV irradiation,

i.e. anti-apoptotic and cell survival factors. Regarding the top 10 downregulated lncRNAs in L-blasts (Fig. 4D), 23 and four unique gene sets were found to be mutually correlated and anti-correlated, respectively. These gene sets reflect in the vast majority genes that are downregulated in cancer or related to stem and progenitor cells.

Since the top differentially expressed lncRNAs identified in LSC showed similarity with those identified in L-blast, subsequently, also significantly (anti-)correlated gene sets were recurrent in both groups (indicated in bold in Fig. 4 A-D).



Legend to Fig. 4. Overview of (anti-)correlating gene sets with DE-lncRNAs identified in LSC (A-B) and L-blast (C-D-. Gene sets mutually (anti-) correlated in LSC and L-blast are indicated in bold (A vs C: upregulated, B and D: downregulated). Three gene sets were found to significantly correlate with upregulated lncRNAs identified in both LSC (Fig. 4A, total of 13) and L-blast (Fig. 4C, total of 6), . On the other hand, six significantly anti-correlated gene sets were identical in both groups. Regarding downregulated lncRNAs identified in LSC and L-blast (Fig. 4 B vs D), half of the significantly correlating gene sets, and 4/6 anti-correlating gene sets, were mutual between both groups. Interestingly, reactome gene sets were specifically correlated to LSC downregulated lncRNAs.

Altogether, we here provide clear proof that lncRNA profiles strongly differ between leukemic and healthy cell populations in pedAML. Expression profiles identified in LSC or L-blast compared to their respective normal counterparts show high similarity. This feature could be seen as beneficial, as both LSC and L-blasts can concomitantly be eradicated through targeted therapy. *Lnc-GSG1-1*, *lnc-RNFT2*-4, *lnc-RGMA-1* and *lnc-LHFPL3-1* deserve further attention in terms of targeted therapy.

Validation through qPCR and further research on lncRNA expression alterations in larger pedAML cohorts are needed to confirm these findings. *In vitro* and *in vivo* monitoring of the effects of targeting LSC-overexpressed lncRNAs confirmed by qPCR, i.e. *CRNDE*, could provide proof-of-concept whether lncRNA perturbation is a valid therapeutic strategy for LSC eradication whilst guaranteeing salvage of HSC.

References

1. Schwarzer A, Emmrich S, Schmidt F, Beck D, Ng M, Reimer C, Adams FF, Grasedieck S, Witte D, Kabler S et al: The non-coding RNA landscape of human hematopoiesis and leukemia. Nature communications 2017, 8(1):218.
2. Garitano-Trojaola A, Agirre X, Prosper F, Fortes P: Long non-coding RNAs in haematological malignancies. Int J Mol Sci 2013, 14(8):15386-15422.
3. Ng M, Heckl D, Klusmann JH: The Regulatory Roles of Long Noncoding RNAs in Acute Myeloid Leukemia. Frontiers in oncology 2019, 9:570.
4. Zhang Y, Liu Y, Xu X: Knockdown of LncRNA-UCA1 suppresses chemoresistance of pediatric AML by inhibiting glycolysis through the microRNA-125a/hexokinase 2 pathway. Journal of cellular biochemistry 2018, 119(7):6296-6308.
5. Yin X, Huang S, Zhu R, Fan F, Sun C, Hu Y: Identification of long non-coding RNA competing interactions and biological pathways associated with prognosis in pediatric and adolescent cytogenetically normal acute myeloid leukemia. Cancer cell international 2018, 18:122.
6. Guan X, Wen X, Xiao J, An X, Yu J, Guo Y: Lnc-SOX6-1 upregulation correlates with poor risk stratification and worse treatment outcomes, and promotes cell proliferation while inhibits apoptosis in pediatric acute myeloid leukemia. International journal of laboratory hematology 2019, 41(2):234-241.
7. Fernando TR, Contreras JR, Zampini M, Rodriguez-Malave NI, Alberti MO, Anguiano J, Tran TM, Palanichamy JK, Gajeton J, Ung NM et al: The lncRNA CASC15 regulates SOX4 expression in RUNX1-rearranged acute leukemia. Molecular cancer 2017, 16(1):126.
8. Cao L, Xiao PF, Tao YF, Hu SY, Lu J, Zhao WL, Li ZH, Wang NN, Wang J, Feng X et al: Microarray profiling of bone marrow long non-coding RNA expression in Chinese pediatric acute myeloid leukemia patients. Oncology reports 2016, 35(2):757-770.
9. Bill M, Papaioannou D, Karunasiri M, Kohlschmidt J, Pepe F, Walker CJ, Walker AE, Brannan Z, Pathmanathan A, Zhang X et al: Expression and functional relevance of long non-coding RNAs in acute myeloid leukemia stem cells. Leukemia 2019, 33(9):2169-2182.
10. Volders PJ, Anckaert J, Verheggen K, Nuytens J, Martens L, Mestdagh P, Vandesompele J: LNCipedia 5: towards a reference set of human long non-coding RNAs. Nucleic acids research 2019, 47(D1):D135-D139.
11. Han J, Liu S, Sun Z, Zhang Y, Zhang F, Zhang C, Shang D, Yang H, Su F, Xu Y et al: LncRNAs2Pathways: Identifying the pathways influenced by a set of lncRNAs of interest based on a global network propagation method. Scientific reports 2017, 7:46566.
12. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al: Gene set enrichment analysis: a knowledge-based

approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 2005, 102(43):15545-15550.
13. He TY, Li SH, Huang J, Gong M, Li G: Prognostic value of long non-coding RNA CRNDE in gastrointestinal cancers: a meta-analysis. Cancer management and research 2019, 11:5629-5642.
14. Zhou Y, Wang R, Xu T, Xie P, Zhang Y, Zhang A, Wang X, Yang C, Yang H, Zhu S: Prognostic Value of Long Noncoding RNA CRNDE as a Novel Biomarker in Solid Cancers: An Updated Systematic Review and Meta-Analysis. Journal of Cancer 2019, 10(11):2386-2396.
15. Wang W, Yuan F, Xu J: The prognostic role of long noncoding RNA CRNDE in cancer patients: a systematic review and meta-analysis. Neoplasma 2019, 66(1):73-82.
16. Hongzhen Z, Yanyu L, Xuexiang L, Meiyu D, Xiaoli C, Yun G, Jingfan C, Shengming D: The diagnostic and prognostic significance of long non-coding RNA CRNDE in pan-cancer based on TCGA, GEO and comprehensive meta-analysis. Pathology, research and practice 2019, 215(2):256-264.
17. Zhou Y, He X, Liu R, Qin Y, Wang S, Yao X, Li C, Hu Z: LncRNA CRNDE regulates the proliferation and migration of vascular smooth muscle cells. Journal of cellular physiology 2019.
18. Wang Y, Zhou Q, Ma JJ: High expression of lnc-CRNDE presents as a biomarker for acute myeloid leukemia and promotes the malignant progression in acute myeloid leukemia cell line U937. European review for medical and pharmacological sciences 2018, 22(3):763-770.
19. Han S, Han B, Li Z, Sun D: Downregulation of long noncoding RNA CRNDE suppresses drug resistance of liver cancer cells by increasing microRNA-33a expression and decreasing HMGA2 expression. Cell Cycle 2019, 18(19):2524-2537.
20. Ito K, Suda T: Metabolic requirements for the maintenance of self-renewing stem cells. Nature reviews Molecular cell biology 2014, 15(4):243-256.

6

CHAPTER VI: Results: Cancer-Related mRNA Expression Analysis Using a Novel Flow Cytometry-Based Assay.

B. Depreter, J. Philippé, M. Meul, B. Denys, K. Vandepoele, B. De Moerloose* and T. Lammens*.
(*shared senior authorship)
Cytometry B Clin Cytom. 2018 Jul;94(4):565-575.

Abstract

Background.
Cancer-related gene expression data mostly originate from unfractionated bulk samples, leading to 'expression averaging' of heterogeneous populations. Multicolour flow cytometry (FCM) may distinguish heterogeneous populations based on the phenotypic characterization of single-cells, but is not applicable for RNA targets. Here, we evaluated the PrimeFlow™ RNA assay, a novel FCM-based assay designed to measure gene expressions, in two cancer entities with high and low RNA target levels.

Methods.
Neuroblastoma (NB) cell lines were studied for *MYCN* gene expression by PrimeFlow™ and compared with the gold standard, RT-qPCR. Dilution series of NB cells (0.10-11%) were prepared to evaluate performance in small cell populations. Diagnostic material of *de novo* acute myeloid leukaemia (AML) patients was used to measure *Wilms' tumor 1* (*WT1*) expression in bulk leukemic cells and rare subsets, e.g. leukemic stem cells (LSCs). FCM analysis was performed on a FACSCanto II (BD Biosciences) using Infinicyt™ (Cytognos®) for data analysis. mRNA expression was reported by normalized mean fluorescence intensity (MFI) values and staining indices.

Results.
MYCN mRNA quantified by PrimeFlow™ significantly correlated with RT-qPCR and remained detectable in small (0.1%) populations. Using PrimeFlow™, *WT1* levels were shown to be significantly higher in AML patient samples with *WT1* overexpression, previously defined by RT-qPCR. Moreover, *WT1* overexpression was distinguishable between heterogeneous cell populations and remained measurable in rare LSCs.

Conclusion.
PrimeFlow™ is a sensitive technique to investigate mRNA expressions, with high concordance to RT-qPCR. High (*MYCN*) and subtle (*WT1*) overexpressed mRNA targets can be quantified in heterogeneous and rare subpopulations e.g. LSCs.

Introduction

Gene expression analysis of protein-coding (mRNA) and non-coding RNA has become indispensable in cancer research [1, 2]. Real-time quantitative polymerase chain reaction (RT-qPCR) has long been, and still is, the most widely used method for studying the expression of genes of interest [3]. However, the development of advanced molecular techniques, e.g. micro-arrays and RNA sequencing, has enabled whole transcriptome analysis in one single experiment [4]. Cancer-related expression data have been shown to be of paramount importance in therapeutic decision-making, providing prognostic information, minimal residual disease (MRD) monitoring and identification of novel therapeutic targets. In neuroblastoma (NB), a paediatric solid tumour with high clinical heterogeneity, amplification status of the *MYCN* gene, and its resulting overexpression at diagnosis, is considered to be the strongest genetic marker for risk stratification and helps defining those patients eligible for intensive therapy [5, 6]. Acute myeloid leukaemia (AML) is a clinically, phenotypically and molecularly heterogeneous haematological malignancy, consisting out of differing leukemic cell populations organized in a hierarchical fashion, with leukemic stem cells (LSCs) residing at the apex [7]. Both adult and paediatric AML are characterized by multiple prognostic factors, among which gene transcription levels play an important role [7, 8]. Within this context, expression of the *Wilms' tumor 1 (WT1*) gene, encoding a transcription factor indispensable for normal hematopoietic development, is frequently upregulated in AML with a bad prognosis, and was identified as a suitable target for MRD monitoring [9, 10].

Unfortunately, gene expression data mostly represent averages, as tumours are complex biological systems constituting of a heterogeneous mix of cancer and normal cells. Such 'cell population averaging' might obscure prognostic and therapeutically relevant information hidden in rare populations [11, 12]. Indeed, although overall survival of paediatric AML reaches 70% with contemporary protocols, relapse rates remain as high as 40%, most likely as a result of persisting LSCs which are not (efficiently) eradicated by current therapies [13, 14]. A better understanding of cellular processes, e.g. gene expressions, in this rare LSC fraction might guide the development of novel therapeutic agents. Within this context, *WT1* has been proposed as a promising LSC-specific therapeutic target [4, 15].

Multicolour flow cytometry (FCM) has become a powerful approach to dissect heterogeneous cell populations based on the phenotypic characterization of single-cells using fluorochrome-conjugated antibodies directed against cluster of differentiation antigens [16]. Within this context, the EuroFlow Consortium provides (pre-)analytical guidelines to achieve maximal standardization among laboratories [17]. Rigorously performed compensations prevent spectral overlap and background signalling in other fluorescence channels (18, 19). Algorithms, such as staining index calculation, normalise the brightness of the fluorescence intensity of the positive population using the data spread on the negative population, to which it is compared. Unfortunately, directly conjugated antibodies are not available for RNA expression evaluation. This limitation has driven the development of molecular techniques enabling single-cell expression analysis, and single-cell isolation methods such as fluorescence-activated cell sorting (FACS) [11, 12]. However, the significant cell and subsequent RNA loss in FACS-obtained cell populations represents a major challenge and makes RNA pre-amplification methods obligatory, with the risk of introducing bias [11].

Recently, the PrimeFlow™ RNA assay (eBioscience, San Diego, CA, USA) was introduced, allowing direct measurement of RNA expressions in diverse cell phenotypes. Herein, the single-cell resolution of FCM is combined with a hybridization-based branched DNA (bDNA) technology, allowing intracellular RNA detection followed by exponential signal amplification. Only a handful of reports have documented its application and, to the best of our knowledge, cancer-related research has not yet been addressed

[18-23]. Here, we evaluated the feasibility of PrimeFlow™ RNA assay in detecting key target mRNAs in NB and AML, using cell lines and patient samples, respectively.

Materials and Methods

Cell lines

Six human NB cell lines (SK-N-BE(2)-C, STA-NB-10, SJ-NB10, SK-N-SH, SH-SY5Y and SH-EP), with documented *MYCN* amplification status, were obtained from the American Type Culture Collection. Culture conditions are described in the Supplementary Design and Methods.

Patient samples

Diagnostic peripheral blood (PB) or bone marrow (BM) samples of four *de novo* AML patients, for whom *WT1* expression was previously determined by RT-qPCR, were selected based on availability and size of the LSC fraction, as evaluated by standard FCM. Detailed information is provided in Table S1. The study was performed in accordance with the Helsinki Declaration. The use of patient material for improved diagnostic work-up (EC/2015/1338) and the isolation of peripheral blood mononuclear cells (PBMCs) from healthy donors (EC/2015/1338) was approved by the Ethical Committee of the Ghent University Hospital. Mononuclear cell (MNC) and white blood cell (WBC) isolation, cryopreservation and thawing was performed as described in the Supplementary Design and Methods.

PrimeFlow™ RNA assay

After thawing or cell harvesting, a live-dead staining using propidium iodide was performed on a separate aliquot of each sample, to assure sufficient viability at the start of the PrimeFlow™ assay. Membrane staining was performed prior to initiation of the PrimeFlow™ protocol by labelling cells with the appropriate antibodies for 30 min in the dark at 2-8 °C (CD45 and CD56 for NB; CD34, CD38 and CD45 for AML) followed by one wash step to remove unlabelled antibodies. Detailed information concerning the used antibodies is provided in Table S2. PrimeFlow™ was performed according to the manufacturer's instructions. One million cells were used as input, except for patient samples ($3x10^6$) and NB cell lines STA-NB-10 ($0.5x10^6$) and SJ-NB10 ($0.2x10^6$) in the RT-qPCR correlation experiment. Two out of three target probes types were evaluated: type 1 (suited for Alexafluor (AF)647-tagged label probes) and type 6 (suited for AF750-tagged label probes). Target probe sets were ordered for *MYCN* (VA1-18174), *TCR gamma alternate reading frame protein* (*TARP,* VA1-19674), *WT1* (VA1-12570) and two reference genes (*ribosomal protein L13A* (*RPL13A,* VA4-13187) and *glyceraldehyde-3-phosphate dehydrogenase* (*GAPD,* VA4-10641).

Samples were analyzed on a 3-laser FACSCanto II (BD Biosciences, San Jose, CA, USA) with instrument set-up strictly performed according to EuroFlow guidelines [17]. The APC and APC-Cy7 channels were used to detect mRNA expression signal by type 1 and type 6 target probe hybridization, respectively. Samples were analyzed using Infinicyt software v.1.8 (Cytognos, Salamanca, Spain). Fluorescence-minus-one (FMO) controls (lack of one surface staining antibody) and no-probe controls (lack of target probe addition) were analyzed for each marker to set gates for positivity. To distinguish LSCs (CD34+CD38-) from blasts (CD34+CD38+), CD38 gating was performed in respect to a 1E3 cut-off as previously described [24]. Gating strategies were developed for each patient and type of experiment in order to minimize gating variation. mRNA expression was reported based on the normalized mean fluorescence intensity (MFI) value (MFI target/MFI no-probe control) or staining index ((MFI target - MFI no-probe control)/(2xSD no-probe control)) [25, 26] [25, 26].

Real-time quantitative PCR

A detailed overview on the method and reagents used can be found in the Supplementary Design and Methods. Normalized relative quantities (NRQ) were calculated by the delta-delta Cq method, using *GAPD* and *beta-2 microglobulin* (*B2M*) for *MYCN* normalisation and *Abelson* (*ABL1)* for *WT1* normalisation. For *MYCN* expression analysis, $\log_2$ fold changes were calculated against the sample with the lowest expression. For *WT1*, an in-house validated NRQ cut-off of 0.50, based on the expression in normal hematopoietic cells of healthy individuals [10], was used to define overexpression.

Statistical analysis

Statistical analysis was performed with MedCalc (Mariakerke, Belgium). Correlation between PrimeFlow™ and RT-qPCR was performed by regression analysis and the Spearman rank sum test. The Mann-Whitney test was used to compare *WT1* levels between AML patients. P-values ≤ 0.05 were considered to be statistically significant.

Results

Technical assessment

PrimeFlow™ does not change standard flowcytometric membrane staining.

NB is phenotypically characterized by the absence of the common leukocyte antigen CD45 and the surface expression of the neural cell adhesion molecule CD56 [27]. As PrimeFlow™ was designed to simultaneously detect five to seven cell surface antigens, together with one to three RNA targets, we first evaluated whether the PrimeFlow™ procedure impacts the MFI values obtained after only standard membrane staining. Expression of both markers in SK-N-BE(2)-C was shown not to be influenced by the PrimeFlow™ procedure (Fig. S1 A-B). Importantly, a high degree of aspecific fluorescence of SK-N-BE(2)-C cells in the FITC-channel was revealed (Fig. S1 C), which was absent in the PE-channel (Fig. S1 D), resulting into a 60-fold lower staining index for CD45-FITC (2.25) compared to CD56-PE (141.36), which underlines the importance of FMO controls in every new experimental set-up.

mRNA expressions may be evaluated in small cell populations.

In order to evaluate the performance of PrimeFlow™ in small cell populations, dilution series of two different NB cell lines were prepared in PBMCs. The first dilution series concerned theoretically 0.63 - 11.0% spiked SK-N-BE(2)-C cells (five dilutions), and was followed by a deeper investigation of the detection limit by spiking STA-NB-10 cells until a final ratio as low as 0.10% in a second dilution series (theoretical range 0.10 - 7.5%, six dilutions). Dilutions were prepared from biological duplicates for each cell line, and the theoretical mean percentage of spiked NB cells was plotted against the experimentally detected mean percentage of spiked NB cells. For both dilution series, a high correlation was observed between the theoretical and experimental number of spiked NB cells (Fig. 1). Our experiments confirm the manufacturer's limit of detection (1%) using two different NB cell lines (SK-N-BE(2)-C and STA-NB-10). Moreover, we were able to demonstrate a deeper limit of detection equal to 0.11%. Mean coefficient of variation (CV, %) in MFI values between the biological duplicates was comparable for both dilution series (9.4% for SK-N-BE(2)-C, 10.2% for STA-NB-10). Importantly, mRNA expression levels remained quite stable between the lowest and highest dilution (CV 9.51% for SK-N-BE(2)-C, 23.9% for STA-NB-10), although MFI values slightly decreased in accordance to the number of spiked cells (R^2=0.93 and R^2=0.82, respectively).

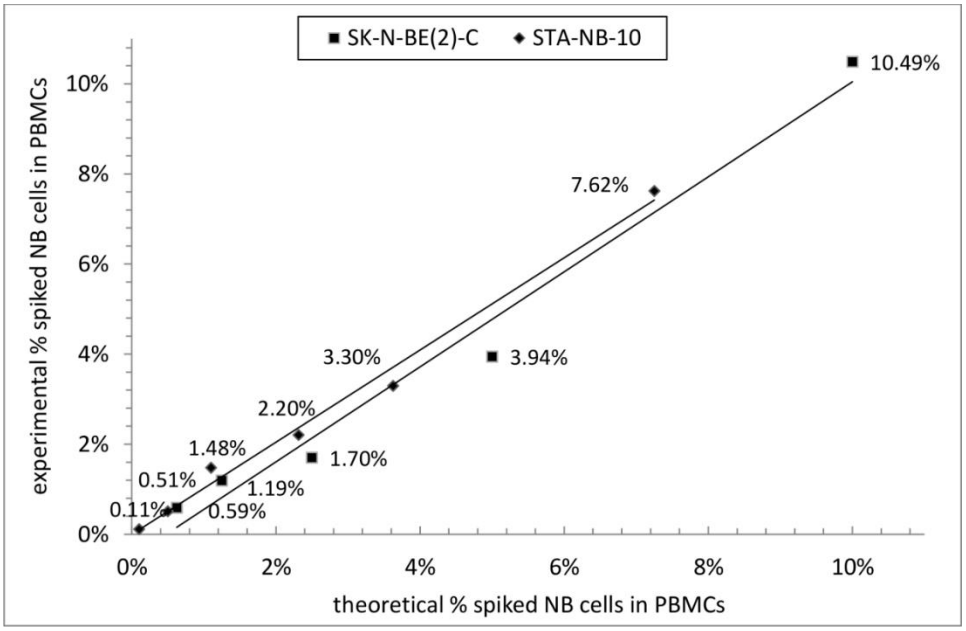


Fig. 1. Detection of small cell populations by PrimeFlow™. Cell dilution series of NB cells in PBMCs, more specific 0.63-11.0% spiked SK-N-BE(2)-C cells (five dilutions) and 0.10-7.5% spiked STA-NB-10 cells (six dilutions), were evaluated for *MYCN* expression. The theoretical mean percentage of spiked NB cells were plotted against the experimental mean percentage of spiked NB cells. For each data point, the experimentally detected mean percentage of spiked NB cells is shown. Abbreviations: MFI: mean fluorescence intensity, NB: neuroblastoma, PBMCs: peripheral blood mononuclear cells.

Comparable results using fresh, short-term and long-term cryopreserved samples.

Next, we determined the influence of short-term (11 days, two independent experiments) and long-term (three and 11 months, single experiment) cryopreservation on *MYCN* expression in SK-N-BE(2)-C cells, compared to freshly harvested cells, based on the percentage of positive events in respect to the no-probe control. Expression analysis after short-term storage reproducibly showed a comparable number of cells with high *MYCN* mRNA levels in comparison to freshly harvested cells (87% vs. 94% and 73% vs. 84%, respectively). Long-term storage, e.g. three (70%) and 11 months (73%), left the number of cells with high *MYCN* mRNA levels virtually unchanged compared to short-term storage (73%).

Influence of preservation in storage buffer depends on target probe type and storage time.

According to the manufacturer's instructions, samples may be stored in the dark at 2-8 °C after the addition of storage buffer for 72 h. We evaluated kinetics of the amplification signal after type 1 (*RPL13A* and *MYCN*) and type 6 (*RPL13A*) target probe hybridization in SK-N-BE(2)-C cells, by repeating FCM analysis up to 360 h after label probe hybridization. Regarding type 1 target probes, we detected similar MFI variations after 72 h of storage for *RPL13A* and *MYCN* expression (2.2% and 5.8%, respectively) as well as after a prolonged storage time of 360 h (2.5% and 6.3%, respectively). Using type 6 target probes for *RPL13A* expression evaluation led to a slightly higher MFI variation (7.9% after 72 h, 16.1% after 360 h).

PrimeFlow™ is not prone to off-target effects.

Expression of a NB-irrelevant target gene *TARP*, described in prostate and breast carcinoma (21), showed a highly comparable fluorescence intensity to the background fluorescence of the no-probe control (Fig. S2 A versus S2 B, respectively) in SK-N-BE(2)-C cells, that could clearly be distinguished from specific *MYCN* expression, hereby excluding off-target effects.

Correlation between PrimeFlow™ and RT-qPCR for gene expression analysis

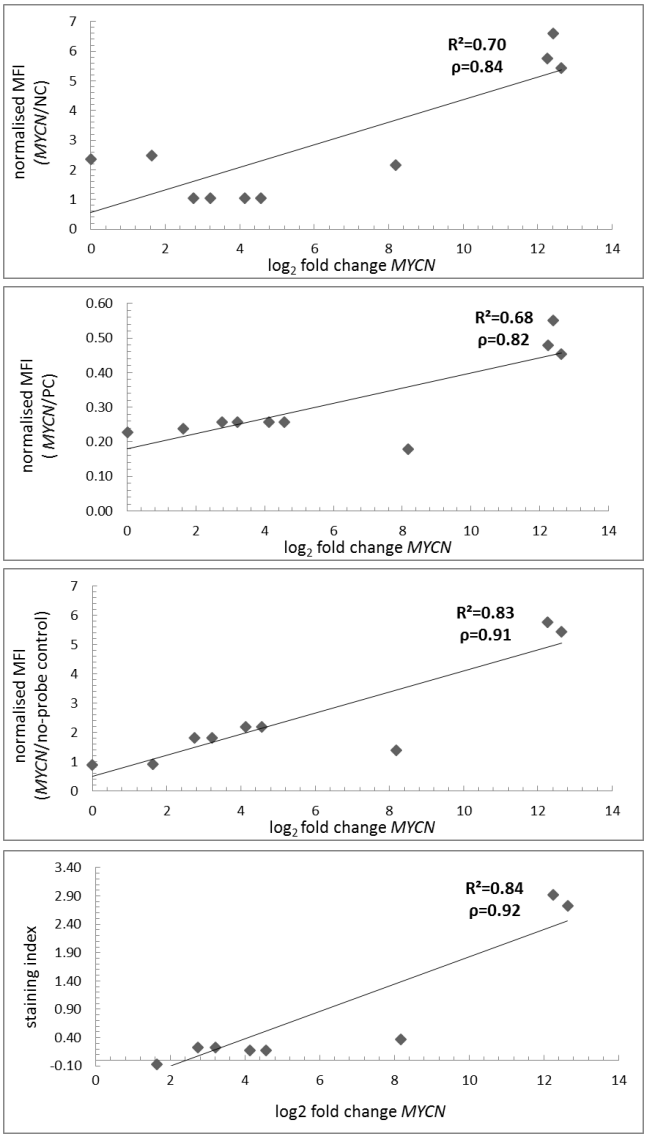


Fig. 2. A–D. Correlation between PrimeFlowTM and RT-qPCR. Correlation between normalized MFI values (A-C) or staining indices (D) obtained by PrimeFlow™ and log_2 fold changes obtained by RT-qPCR. MFI were normalized against the negative control (NC) (A), positive control (PC) (B) and no-probe control (C). Abbreviations: MFI: mean fluorescence intensity, R^2: coefficient of determination, ρ: spearman rank correlation coefficient, RT-qPCR: real-time quantitative polymerase chain reaction.

Target gene expression

We correlated the *MYCN* gene expression in six NB cell lines as evaluated by PrimeFlow™ and by RT-qPCR. Log_2 fold changes measured by RT-qPCR for all six cell lines are shown in Fig. S3. For PrimeFlow™, *RPL13A* expression in SK-N-BE(2)-C was used as positive control (PC) and SK-N-BE(2)-C without target probe as negative control (NC). Additionally, each cell line was individually evaluated for *RPL13A* expression and background fluorescence (defined as no-probe control). First, we evaluated four different MFI normalisation strategies, by expressing the observed MFI for *MYCN* as a ratio against the MFI measured in the same APC-channel for the NC, PC, no-probe control and *RPL13A* expression in the respective cell line. Regarding the first three MFI normalisation methodologies (Fig. 2 A-C), correlation between the normalized *MYCN* MFI obtained by PrimeFlow™ and the log_2 fold changes obtained by RT-qPCR, was significantly strong (R^2=[0.68 - 0.83], ρ=[0.82 - 0.91], P<0.05). Normalisation in respect to the *RPL13A* expression within each cell line, on the other hand, showed a rather low correlation (R^2=0.24, ρ=0.49, P>0.05). Following, we calculated the staining index, which also showed a significant high correlation to the log_2 fold changes obtained by RT-qPCR (Fig. 2 D; R^2=0.84, ρ=0.92, P<0.001). Subsequently, the normalized MFI value, using the no-probe control MFI as denominator, and the staining index, were further used for data reporting. Noteworthy, also the raw MFI values correlated highly significantly with the log_2 fold changes (Fig S4, R^2=0.92, ρ=0.94 (P<0.05)).

Next, we explored the applicability of PrimeFlow™ to evaluate gene expressions in heterogeneous samples and rare cell populations therein. To this end, we evaluated *WT1* expression in diagnostic samples from four AML patients, two of which were shown to overexpress *WT1* by RT-qPCR in routine

diagnostics (Table S1). MFI values were inspected for each sample (n=6) and normalized MFI values and staining indices were calculated for three out of the four patients (n=3) (Table 1).

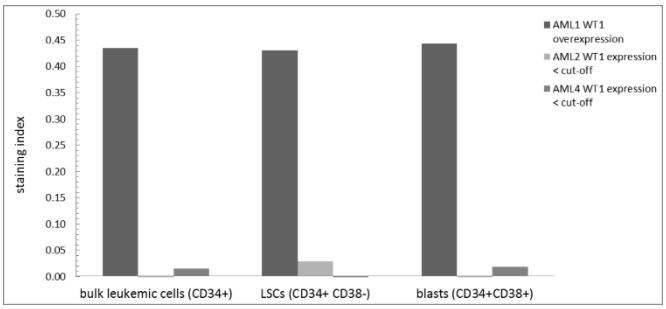
To start, MFI values representative for *WT1* mRNA in the bulk leukemic cells (CD34+) showed a significant 51.9% upregulation if *WT1* was overexpressed, according to RT-qPCR, compared to normal expression (Fig S5 A). In case of *WT1* overexpression, MFI values showed a 20.5% upregulation versus the no-probe control, compared to only 1.03% by average in case of normal *WT1* expression (Table 1). Normalized MFI values (Fig 3 A) and staining indices (Fig. 3 B) calculated for *WT1* expression in the bulk leukemic cells had a strong discriminative power in separating patients with and without *WT1* overexpression.

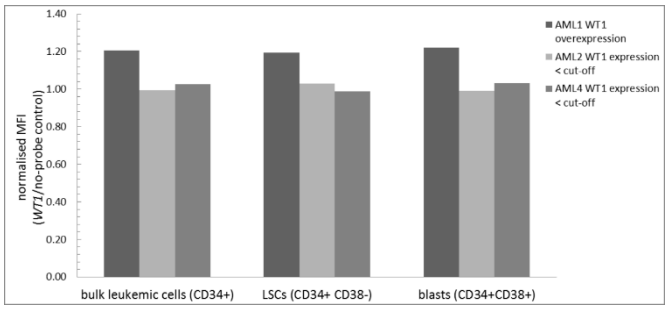
Following, we investigated whether differential *WT1* mRNA levels remained detectable by PrimeFlow™ within the fractionated populations, e.g. blasts (CD34+CD38+) and LSCs (CD34+CD38-). Again, MFI values were by average 63.4% (Fig S5 B) and 61.0% (Fig S5 C) upregulated, respectively, in samples with *WT1* overexpression compared to normal expression, although only statistically significant in the blasts. Target MFI values remained upregulated in blasts and LSCs compared to the no-probe controls (range 19.4 - 22.1%, Table 1). In agreement with the results obtained within the bulk leukemic cells, normalized MFI values (Fig 3 A) and staining indices (Fig 3 B) were able to discriminate high from normal *WT1* expression levels in blasts and LSCs. Moreover, staining indices showed thirty- to forty-fold higher values in case of *WT1* overexpression (Fig 3 B), whereas the non-normalized fluorescence intensities in blast and LSC populations (Fig. 3 C-D) showed only minor shifts. This minor shift is also illustrated by the representative dot plots (Fig. 3 E-F), showing a higher *WT1* expression, judged in respect to the 1E3 cut-off, in the blasts (50.9% versus 14.3%) and LSCs (42.2% and 2.77%) of a patient with *WT1* overexpression (Fig. 3 E) versus normal expression (Fig. 3 F) as defined by RT-qPCR.

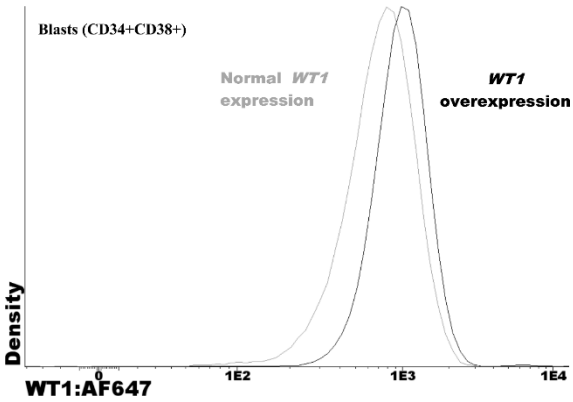
Table 1. MFI values, normalized MFI values and staining indices for *WT1* expression in AML patient samples.

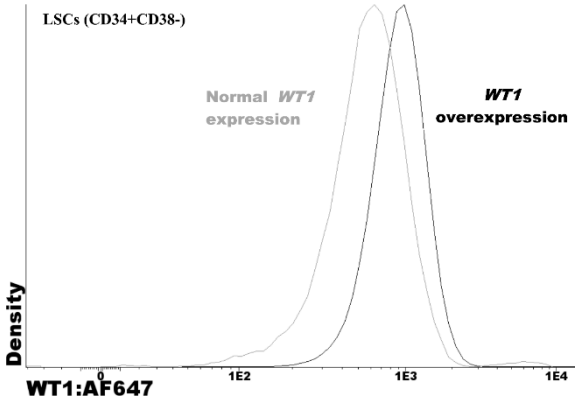
sample no.	Cell population	*WT1* PrimeFlow™ analysis					*WT1* RT-qPCR
		MFI *WT1* (AF647)	MFI no-probe control (AF647)	upregulation vs. no-probe control	normalised MFI[b]	staining index[b]	NRQ result (cut-off 0.50[c])
AML 1 exp 1[a]	bulk leukemic cells (CD34+)	957.4	NA	NC	NC	NC	over-expression (0.90)
	LSCs (CD34+ CD38-)	931.8	NA	NC	NC	NC	
	blasts (CD34+CD38+)	998.3	NA	NC	NC	NC	
AML 1 exp2	bulk leukemic cells (CD34+)	1065.7	884.7	20.5%	1.20	0.43	over-expression (0.90)
	LSCs (CD34+ CD38-)	1020.5	854.4	19.4%	1.19	0.43	
	blasts (CD34+CD38+)	1152.8	943.8	22.1%	1.22	0.44	
AML 3 exp 1[a]	bulk leukemic cells (CD34+)	855.2	NA	NC	NC	NC	over-expression (0.96)
	LSCs (CD34+ CD38-)	634.8	NA	NC	NC	NC	
	blasts (CD34+CD38+)	974.8	NA	NC	NC	NC	
AML 2 exp 1[a]	bulk leukemic cells (CD34+)	529.2	NA	NC	NC	NC	normal expression (< cut-off)
	LSCs (CD34+ CD38-)	415.0	NA	NC	NC	NC	
	blasts (CD34+CD38+)	543.2	NA	NC	NC	NC	
AML 2 exp 2	bulk leukemic cells (CD34+)	743.5	748.0	-0.60%	0.99	-0.01	normal expression (< cut-off)
	LSCs (CD34+ CD38-)	566.2	550.6	2.84%	1.03	0.03	
	blasts (CD34+CD38+)	748.9	756.4	-0.98%	0.99	-0.01	
AML 4 exp 2	bulk leukemic cells (CD34+)	621.7	605.5	2.67%	1.03	0.01	normal expression (< cut-off)
	LSCs (CD34+ CD38-)	625.4	633.8	-1.32%	0.99	-0.01	
	blasts (CD34+CD38+)	621.2	601.9	3.21%	1.03	0.02	

[a]No-probe controls were not evaluated in the first experiment due to lack of a sufficient amount of material available. [b]Based on the MFI values for target expression and no-probe controls, normalized MFI values ([MFI target/MFI no-probe control]) and staining indices ([(MFI target - MFI no-probe control)/(2xSD no-probe control)]) were calculated. [c]An in-house validated NRQ cut-off of 0.50, based on the expression in normal hematopoietic cells of healthy individuals, was used to define overexpression. Different experiments for one patient are separated by a dashed line, different patient samples by a full line, different and patient samples with *WT1* overexpression versus normal expression by a thick line. For detailed information on patient samples, see Table S1. Abbreviations: no.: number, NRQ: normalized relative quantity (versus *ABL1*), *WT1*: Wilms'

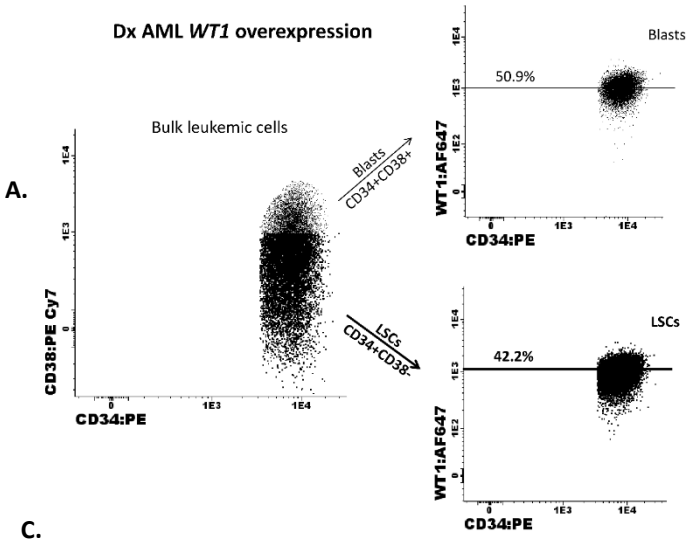


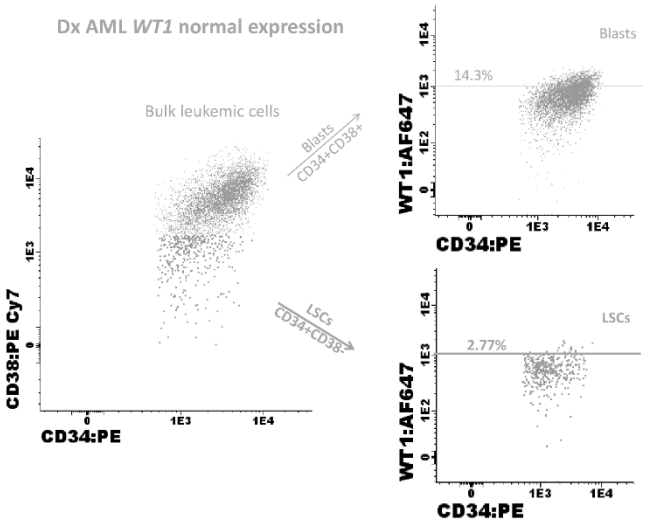






tumor 1, AML: acute myeloid leukemia, LSCs: leukemic stem cells, AF647: AlexaFluor 647, NA: not analyzed, NC: not calculable, SD: standard deviation, RT-qPCR: real-time quantitative polymerase chain reaction.





So forth, in concert with the previous data obtained in NB cell lines, using normalized MFI values and more pronounced staining indices, PrimeFlow™ is able to discriminate subtle mRNA expression differences, e.g. *WT1* overexpression, in heterogeneous samples and the fractionated populations herein, even for the rare LSC population.

Fig. 3 A-F. *WT1* expression by PrimeFlow™ in *de novo* AML patients.
WT1 expression was evaluated in AML patients, previously defined by RT-qPCR as having *WT1* overexpression or normal *WT1* expression (detailed overview in Table S1). (A) Normalized MFI values, representative for *WT1* expression, evaluated in the bulk leukemic cells (CD34+, left), LSCs (CD34+CD38-, middle) and blasts (CD34+CD38+, right) in three AML patients (1/3 *WT1* overexpression, 2/3 *WT1* normal expression). (B) Staining indices, representative for *WT1* expression, evaluated in the bulk leukemic cells (CD34+, left), LSCs (CD34+CD38-, middle) and blasts (CD34+CD38+, right) in three AML patients (1/3 *WT1* overexpression, 2/3 *WT1* normal expression). (C - D) Histograms illustrating a modest increase in AF647 fluorescence intensity, representative for *WT1* expression, in the blasts (C) and LSCs (D) of a patient with *WT1* overexpression (black curve, right) compared to the patient with normal *WT1* expression (grey curve, left). (E - F) Dot plots presented for one patient within each group, e.g. *WT1* overexpression (E) and *WT1* normal expression (F), which are indicated by a different font (black and grey, respectively). Bulk leukemic cells (CD34+, left) are separated into blasts (thin dots, top right) and LSCs (bold dots, bottom right) based on CD38 expression. Patient with RT-qPCR defined *WT1* overexpression (E) shows an upregulated number of cells with *WT1* mRNA fluorescence intensity above 1E3 (horizontal cut-off line) in both the blasts (50.9% vs. 14.3%) and LSCs (42.2% vs. 2.77%) compared to a patient with normal *WT1* expression (F). Abbreviations: *WT1*: Wilms' tumor 1, LSC: leukemic stem cell, AF647: AlexaFluor 647, MFI: mean fluorescence intensity, RT-qPCR: real-time quantitative polymerase chain reaction

Reference gene expression

As reference gene selection has been proven to be a crucial experimental design aspect in RT-qPCR [28], we evaluated the expression of *RPL13A* and *GAPD* by PrimeFlow™ in two disease entities, e.g. NB and AML, using the SK-N-BE(2)-C cell line and two patient samples, respectively. In SK-N-BE(2)-C, the expression of *GAPD* and *RPL13A* was highly comparable, as evaluated by both normalized MFI values (Fig. 4 A) and staining indices (Fig. 4 B). In contrast, *GAPD* and *RPL13A* expression in AML patient samples showed notable differences in the bulk leukemic cells (CD34+), as well as in the fractionated blasts (CD34+CD38+) and LSCs (CD34+CD38-) separately. Remarkably, reference gene expression was consistently higher in blasts compared to LSCs for each patient (Fig. 4 A-B). These data, although limited, suggest unpredictable cancer-related perturbations in heterogeneous cell populations within one disease entity and underscore the high heterogeneity in AML.

Multiplexing target and reference gene expression

We evaluated the effect of multiplexing reference gene and target gene probes in one single experiment, as this strategy reduces the number of tests and hands-on time. To this end, we compared the normalized MFI values (Fig. 5 A) and staining indices (Fig. 5 B) for the expression of *RPL13A* (type 6 target probe) and *MYCN* (type 1 target probe) measured separately and simultaneously. Expression variation due to multiplexing was higher for type 1 target probes compared to type 6 target probes, but remained within 15% limits of the target MFI [17]. No consistent pattern concerning up- or downregulation was observed.

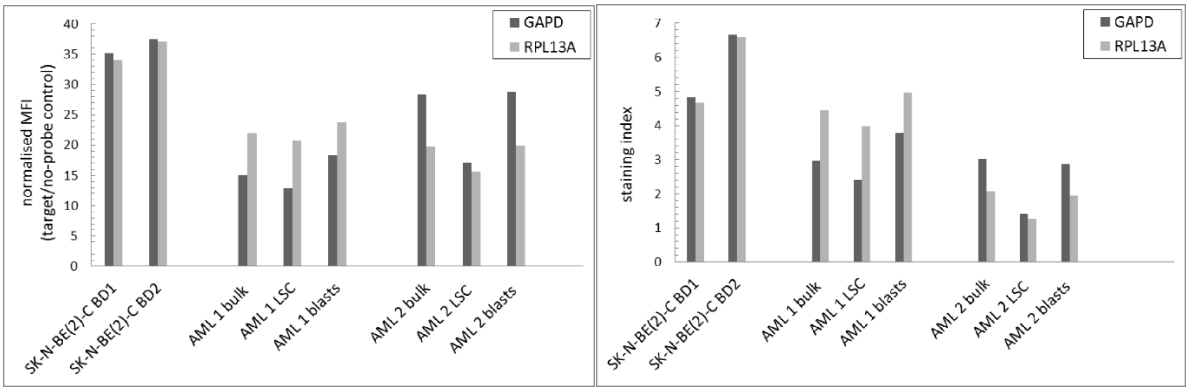


Fig. 4 A-B. Reference gene evaluation. Expression of *RPL13A* and *GAPD* was evaluated in SK-N-BE(2)-C (biological duplicates) and two AML patient samples (single analysis), the latter subdivided in bulk leukemic cells (CD34+), LSCs (CD34+CD38-) and blasts (CD34+CD38+). Normalized MFI values (A) and staining indices (B) were calculated. Abbreviations: AML: acute myeloid leukaemia, *RPL13A*: ribosomal protein L13a, *GAPD*: glyceraldehyde-3-phosphate dehydrogenase, LSC: leukemic stem cell, MFI: mean fluorescence intensity, BD: biological duplicate.

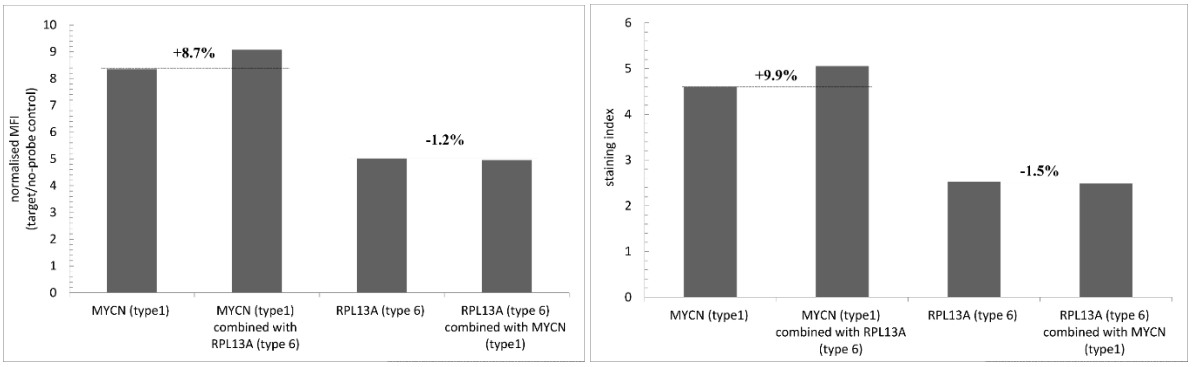


Fig. 5 A-B. Multiplexing of reference gene and target gene expressions. The expression of *RPL13A* (type 6 target probe) and *MYCN* (type 1 target probe) obtained by PrimeFlow™ by individual target probe hybridisation (dashed line) was compared to the expression after hybridisation of both probes types simultaneously. Normalized MFI values (A) and staining indices (B) were calculated. Abbreviations: *RPL13A*: ribosomal protein L13A, MFI: mean fluorescence intensity.

Discussion

Although gene expression studies using RT-qPCR, or high-throughput methods e.g. microarrays or RNA sequencing, provide crucial information concerning cancer biology, risk-stratification and prognosis, they mostly deliver 'bulk data' on heterogeneous cancer cell populations and lack the ability to detect single-cell or rare cell population-specific gene expression changes [12]. In both adult and paediatric AML, persistence of rare LSCs are thought to be causative for the high relapse rates [13, 14]. Therefore, an increasing need for advanced instrument analytical tools, able to elucidate coding and non-coding gene expressions at single-cell level, has emerged [29]. Here, we illustrate the applicability of the PrimeFlow™ RNA assay in detecting key mRNA target expressions in two cancer entities, NB and AML, using cell lines and patient samples.

Both highly overexpressed (*MYCN*) and subtle overexpressed (*WT1*) mRNA targets could be identified, even in rare subpopulations e.g. LSCs. Tamaki and Ogawa showed that the number of *WT1* transcripts in AML significantly increase at relapse compared to diagnosis, suggesting the outgrowth of persisting

WT1-overexpressing LSCs [30, 31]. Indeed, *WT1* expression in LSCs was shown to inversely correlate with survival time [32]. Here, we show that the PrimeFlow™ RNA assay is able to detect *WT1* overexpression in *de novo* AML, even in heterogeneous cell populations, e.g. blasts and LSCs, suggesting a role in MRD monitoring and LSC-specific expression analysis (Fig. 3 A-). Moreover, differential mRNA expression was detected between the cell populations within one patient (Fig. 3 A-F), making the technique of interest for intraleukemic heterogeneity evaluation [24].

To the best of our knowledge, this is the first report describing this technique in a cancer-related setting, and only a handful reports have addressed its application worldwide [18-23]. Subsequently, guidelines are lacking on how to interpret the mRNA amplification signal. Previous publications have reported raw MFI values, normalized MFI values (comparable to the field of RT-qPCR) and percentages of positivity in respect to the total number of analyzed cells [18-23]. If MFI values are used, standardised compensation settings are required to allow for inter-experiment comparisons [17]. However, we here show that MFI normalisation algorithms increase the discriminative power between mRNA expression levels and contribute to a more robust data interpretation. In this study, mRNA levels expressed as normalized MFI values or staining indices both showed a significant high correlation to RT-qPCR (Fig. 2 C-D), hereby excluding non-specific background fluorescence.

Our proof-of-principle experiments confirm that PrimeFlow™ may be successfully applied to cryopreserved material and that FCM analysis may be reliably postponed for more than 72 h, if samples are correctly preserved in storage buffer. Expression detection within small (0.1%) populations was feasible (Fig. 1), regardless of the target gene mRNA expression level, as evaluated by two different NB cell lines. Noteworthy, combined detection of multiple targets is feasible, although results may vary depending on the type of target probe used (Fig. 5 A-B).

Notwithstanding the abovementioned results pave the way for routine diagnostic applications, the high labour-intensity of this technique remains a hurdle [19]. However, postponing FCM analysis for 72 h or even more, confirmed by our experimental data, makes it more user-friendly. Several caveats were identified, such as the viability of the cells, which need to be in an active growth phase (*in vitro* culture) or isolated within 4 h of tissue collection (*ex vivo* primary cells) to preserve RNA integrity. Aspecific fluorescence needs to be critically evaluated during the initial experimental setup (Fig. S1 C-D). PrimeFlow™ escapes the high costs of rigorous antibody development and validation, but validation of the antibodies used for membrane staining is still required (Fig. S1 A-B), as epitopes might be denatured by the methanol permeabilisation step [18]. In concert with the field of RT-qPCR, we showed that reference gene expression differs between disease entities, and even between cell populations, depending on the heterogeneity of the disease (Fig 4 A-B) [28]. Therefore, reference gene selection should be performed a priori during experimental set-up and not extrapolated from RT-qPCR.

In conclusion, PrimeFlow™ RNA assay is a sensitive FCM technique for the investigation of mRNA expressions in heterogeneous samples and rare subpopulations herein, without the need for cell-sorting, showing a significant correlation to RT-qPCR. The combined identification of cellular subpopulations, through immunophenotypical staining, with mRNA expression investigation within these subpopulations, decreases the analytical turn-around-time tremendously compared to a classic two-step approach of cell-sorting methods, e.g. FACS, followed by genomic technologies. Future studies should investigate its application for routine diagnostic purposes, such as risk stratification and MRD monitoring.

Acknowledgments

This work was supported by the FWO (grant number 1113117, BDP), the Foundation against Cancer (grant number 2014-265, TL), the Cancer Plan, action 29 (KP_29_020, JP) and vzw Kinderkankerfonds – a non-profit childhood cancer foundation under Belgain law (TL). The authors would like to thank Katrien Vandemeulebroecke, Eva Terras and Inge D'hont for technical assistance. BDP is a PhD candidate at Ghent University and this work is submitted in partial fulfilment of the requirement for the PhD.

References

1. Rapin N, Bagger FO, Jendholm J, Mora-Jensen H, Krogh A, Kohlmann A, Thiede C, Borregaard N, Bullinger L, Winther O et al: Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. Blood 2014, 123(6):894-904.
2. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA et al: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999, 286(5439):531-537.
3. VanGuilder HD, Vrana KE, Freeman WM: Twenty-five years of quantitative PCR for gene expression analysis. Biotechniques 2008, 44(5):619-626.
4. Saito Y, Kitamura H, Hijikata A, Tomizawa-Murasawa M, Tanaka S, Takagi S, Uchida N, Suzuki N, Sone A, Najima Y et al: Identification of therapeutic targets for quiescent, chemotherapy-resistant human leukemia stem cells. Science translational medicine 2010, 2(17):17ra19.
5. Huang M, Weiss WA: Neuroblastoma and MYCN. Cold Spring Harbor perspectives in medicine 2013, 3(10):a014415.
6. Cohn SL, Pearson AD, London WB, Monclair T, Ambros PF, Brodeur GM, Faldum A, Hero B, Iehara T, Machin D et al: The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2009, 27(2):289-297.
7. De Kouchkovsky I, Abdul-Hay M: 'Acute myeloid leukemia: a comprehensive review and 2016 update'. Blood cancer journal 2016, 6(7):e441.
8. Sandahl JD, Kjeldsen E, Abrahamsson J, Ha SY, Heldrup J, Jahnukainen K, Jonsson OG, Lausen B, Palle J, Zeller B et al: The applicability of the WHO classification in paediatric AML. A NOPHO-AML study. British journal of haematology 2015, 169(6):859-867.
9. Inoue K, Ogawa H, Sonoda Y, Kimura T, Sakabe H, Oka Y, Miyake S, Tamaki H, Oji Y, Yamagami T et al: Aberrant overexpression of the Wilms tumor gene (WT1) in human leukemia. Blood 1997, 89(4):1405-1412.
10. Cilloni D, Renneville A, Hermitte F, Hills RK, Daly S, Jovanovic JV, Gottardi E, Fava M, Schnittger S, Weiss T et al: Real-time quantitative polymerase chain reaction detection of minimal residual disease by standardized WT1 assay to enhance risk stratification in acute myeloid leukemia: a European LeukemiaNet study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2009, 27(31):5195-5201.
11. Zhong JF, Chen Y, Marcus JS, Scherer A, Quake SR, Taylor CR, Weiner LP: A microfluidic processor for gene expression profiling of single human embryonic stem cells. Lab on a chip 2008, 8(1):68-74.
12. Kalisky T, Quake SR: Single-cell genomics. Nature methods 2011, 8(4):311-314.
13. Terwijn M, Zeijlemaker W, Kelder A, Rutten AP, Snel AN, Scholten WJ, Pabst T, Verhoef G, Lowenberg B, Zweegman S et al: Leukemic Stem Cell Frequency: A Strong Biomarker for Clinical Outcome in Acute Myeloid Leukemia. PloS one 2014, 9(9).
14. Pollyea DA, Jordan CT: Therapeutic targeting of acute myeloid leukemia stem cells. Blood 2017.
15. Rosenfeld C, Cheever MA, Gaiger A: WT1 in acute leukemia, chronic myelogenous leukemia and myelodysplastic syndrome: therapeutic potential of WT1 targeted therapies. Leukemia 2003, 17(7):1301-1312.
16. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, Scholten WJ, Snel AN, Veldhuizen D, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. Leukemia 2016, Feb(30(2)):439-446.

17. Kalina T, Flores-Montero J, van der Velden VH, Martin-Ayuso M, Bottcher S, Ritgen M, Almeida J, Lhermitte L, Asnafi V, Mendonca A et al: EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. Leukemia 2012, 26(9):1986-2010.
18. Porichis F, Hart MG, Griesbeck M, Everett HL, Hassan M, Baxter AE, Lindqvist M, Miller SM, Soghoian DZ, Kavanagh DG et al: High-throughput detection of miRNAs and gene-specific mRNA at the single-cell level by flow cytometry. Nature communications 2014, 5:5641.
19. Henning AL, Sampson JN, McFarlin BK: Measurement of Low-Abundance Intracellular mRNA Using Amplified FISH Staining and Image-Based Flow Cytometry. Current protocols in cytometry / editorial board, J Paul Robinson, managing editor [et al] 2016, 76:7 46 41-48.
20. Henning AL, McFarlin BK: Consumption of a high-fat, high-calorie meal is associated with an increase in intracellular co-localization of PPAR-gamma mRNA and protein in monocytes. Methods 2016.
21. Soh KT, Tario JD, Jr., Colligan S, Maguire O, Pan D, Minderman H, Wallace PK: Simultaneous, Single-Cell Measurement of Messenger RNA, Cell Surface Proteins, and Intracellular Proteins. Current protocols in cytometry / editorial board, J Paul Robinson, managing editor [et al] 2016, 75:7 45 41-47 45 33.
22. Martrus G, Niehrs A, Cornelis R, Rechtien A, Garcia-Beltran W, Lutgehetmann M, Hoffmann C, Altfeld M: Kinetics of HIV-1 Latency Reversal Quantified on the Single-Cell Level Using a Novel Flow-Based Technique. Journal of virology 2016, 90(20):9018-9028.
23. Henning AL, Levitt DE, Vingren JL, McFarlin BK: Measurement of T-Cell Telomere Length Using Amplified-Signal FISH Staining and Flow Cytometry. Current protocols in cytometry / editorial board, J Paul Robinson, managing editor [et al] 2017, 79:7 47 41-47 47 10.
24. Bachas C, Schuurhuis GJ, Assaraf YG, Kwidama ZJ, Kelder A, Wouters F, Snel AN, Kaspers GJL, Cloos J: The role of minor subpopulations within the leukemic blast compartment of AML patients at initial diagnosis in the development of relapse. Leukemia 2012, 26(6):1313-1320.
25. Maecker HT, Frey T, Nomura LE, Trotter J: Selecting fluorochrome conjugates for maximum sensitivity. Cytometry Part A : the journal of the International Society for Analytical Cytology 2004, 62(2):169-173.
26. Baumgarth N, Bigos M: Optimization of emission optics for multicolor flow cytometry. Methods in cell biology 2004, 75:3-22.
27. Bozzi F, Gambirasio F, Luksch R, Collini P, Brando B, Fossati-Bellani F: Detecting CD56+/NB84+/CD45- immunophenotype in the bone marrow of patients with metastatic neuroblastoma using flow cytometry. Anticancer research 2006, 26(5A):3281-3287.
28. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology 2002, 3(7):RESEARCH0034.
29. Preffer F, Dombkowski D: Advances in complex multiparameter flow cytometry technology: Applications in stem cell research. Cytometry Part B, Clinical cytometry 2009, 76(5):295-314.
30. Tamaki H, Ogawa H, Inoue K, Soma T, Yamagami T, Miyake S, Oka Y, Oji Y, Tatekawa T, Tsuboi A et al: Increased expression of the Wilms tumor gene (WT1) at relapse in acute leukemia. Blood 1996, 88(11):4396-4398.
31. Ogawa H, Tamaki H, Ikegame K, Soma T, Kawakami M, Tsuboi A, Kim EH, Hosen N, Murakami M, Fujioka T et al: The usefulness of monitoring WT1 gene transcripts for the prediction and management of relapse following allogeneic stem cell transplantation in acute type leukemia. Blood 2003, 101(5):1698-1704.
32. Xu J, Wang HW, Yang T, Tan YH, Zhang L: [Expression of WT1 gene in CD34(+)CD38(-)CD123(+) AML stem cells and its significance analysis.]. Zhonghua xue ye xue za zhi = Zhonghua xueyexue zazhi 2010, 31(3):172-175.

Supplementary Design and Methods

Cell culture

Cell lines were cultured as specified by the American Type Culture Collection in a 37 °C 5% CO_2 humidified incubator. Cells were harvested after washing (1:5000 Versene, Invitrogen, CA, USA) and short incubation with 0.05% Trypsin-EDTA (Invitrogen). Harvested cells were immediately diluted with RPMI 1640 medium (Invitrogen) to a final volume of 15 mL and centrifuged at 300 g for 10 minutes (min) at 37 °C. Finally, cells were resuspended in 1 mL RPMI 1640 medium and, after automatic cell counting (XE-5000, Sysmex UK Ltd, Milton Keynes, UK), incubated at 37 °C until further analysis.

Mononuclear cell/white blood cell isolation and cryopreservation

Ficoll-based density gradient centrifugation or white blood cell (WBC) isolation (Erythrocyte Lysis Buffer, QIAGEN, Hilden, Germany) was performed on peripheral blood (PB) (healthy donors and 1/4 AML patients) or bone marrow (BM) samples (3/4 AML patients) to isolate mononuclear cells (MNCs) or WBCs, respectively. Erythrocytes were lysed during Ficoll-based density gradient centrifugation by short incubation with ammonium chloride at room temperature [17]. MNCs/WBCs were frozen in 90% heat-inactivated fetal calf serum (FCS, Invitrogen) supplemented with 10% dimethylsulfoxide (DMSO) (Riedel-de Haen, Seelze, Germany) and cryopreserved in liquid nitrogen until thawing.

Thawing procedure

Viably frozen cells were shortly thawed at 37 °C while gently and continuous mixing with 37 °C preheated RPMI 1640 medium (Invitrogen) supplemented with 20% FCS (Invitrogen) to a final volume of 20 mL. Addition of 200 µL DNase solution (DNaseI, Roche, Sigma-Aldrich, St.Louis, Missouri, USA) and 200 µL $MgCl_2$ (Sigma-Aldrich) was followed by 30 min acclimatisation at room temperature. Next, cells were centrifuged at 400 g for 10 min at 4 °C, resuspended with the same medium until a final volume of 20 mL and again centrifuged (400 g, 10 min, 4 °C), Finally, cells were resuspended in 1 mL FCS and, after automatic cell counting (XE-5000, Sysmex), incubated at 37 °C until further analysis.

Real-time quantitative PCR (RT-qPCR)

WT1

Isolated WBCs or MNCs from *de novo* AML patients were used for RNA extraction (RNeasy Mini Kit, Qiagen) in combination with an on-column DNase digestion (RNase-Free DNase, Qiagen). cDNA synthesis was performed by random primer-mediated reverse transcription PCR using the Invitrogen SuperScript® II cDNA synthesis kit (ThermoFisher) on a thermic cycler with 5 - 20 µL RNA (concentrations 81 - 479 ng/µL) as input. RT-qPCR experiments were performed in white 96-well skirted PCR plates (Bio-Rad Technologies) on a CFX96 Real-Time PCR Detection System (Bio-Rad, Berkeley, California, USA) according to MIQE guidelines [33]. A hydrolysis probe-based RT-qPCR reaction with 5 µL cDNA as input was performed for the *WT1* and *Abelson 1* (*ABL1)* gene in a 25 µL reaction, comprising 10x PCR buffer, 50 mM $MgCl_2$, 5 mM deoxyribonucleoside triphosphates, 0.125 µL Hot Goldstar enzyme, 500 nM forward and reverse primers, 200 nM hydrolysis probe and H_2O. For *WT1* and *ABL1*, primer and probe sequences were described by Ogawa et al. [31] and Beillard et al. [34], respectively. Amplification of each sample was performed in duplicate and carried out, after 10

min at 95°C to activate Hot Goldstar enzyme, by 50 cycles at 95°C for 15 s and 60°C for 60 s. All reactions were performed in duplicate and Cq values were averaged if delta Cq ≤ 1.5. Nuclease-free H_2O (Sigma) was used as NTC and an in-house prepared cDNA dilution of the AML cell line k562 as positive control.

MYCN

Harvested cells from *in vitro* cultured neuroblastoma (NB) cell lines were centrifuged (400 g, 5 min, room temperature), resuspended in trizol (Life Technologies) and after one freeze-thaw cycle used as input for RNA extraction (miRNeasy Mini Kit, Qiagen). RNA was quantified using Nanodrop (ThermoFisher Scientific) with concentrations between 16 - 372 ng/µL and high purity ($A_{260}/A_{280} \geq 1.8$). cDNA synthesis was performed by random primer-mediated reverse transcription PCR using the Invitrogen SuperScript® III cDNA synthesis kit (ThermoFisher) on a Veriti™ Thermal Cycler (Applied Biosystems™) with 400 ng RNA as input. RT-qPCR experiments were conducted in a clear 96-well plate (FrameStar FastPlate 96 well, Bioké, Leiden, The Netherlands) on a Viia7 instrument (Life Technologies, Carlsbad, CA) according to MIQE guidelines [33]. A hydrolysis probe-based RT-qPCR reaction with 1 µL cDNA as input was performed for *MYCN (*Hs00232074_m1*)*, beta-2 microglobulin (*B2M*) (Hs00187842_m1) and glyceraldehyde-3-phosphate dehydrogenase (*GAPD*) (Hs02786624_g1) in a 5 µL reaction, comprising 0.25 µL TaqMan® primer-probe mix (Applied Biosystems™), TaqMan Fast Universal Master Mix (Applied Biosystems™) and nuclease-free H_2O (Sigma). Amplification was carried out, after 10 min at 95 °C to activate TaqMan enzyme, by 40 cycles at 95 °C for 15 seconds (s) and 60 °C for 60 s. All reactions were performed in duplicate, except for STA-NB-10 and SJ-NB10, and quantification cycle (Cq) values were averaged if delta Cq ≤ 1.5. Nuclease-free H_2O (Sigma) was used as no-template control (NTC).

Supplemental references

1. Rapin N, Bagger FO, Jendholm J, Mora-Jensen H, Krogh A, Kohlmann A, Thiede C, Borregaard N, Bullinger L, Winther O et al: Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. Blood 2014, 123(6):894-904.
2. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA et al: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999, 286(5439):531-537.
3. VanGuilder HD, Vrana KE, Freeman WM: Twenty-five years of quantitative PCR for gene expression analysis. Biotechniques 2008, 44(5):619-626.
4. Saito Y, Kitamura H, Hijikata A, Tomizawa-Murasawa M, Tanaka S, Takagi S, Uchida N, Suzuki N, Sone A, Najima Y et al: Identification of therapeutic targets for quiescent, chemotherapy-resistant human leukemia stem cells. Science translational medicine 2010, 2(17):17ra19.
5. Huang M, Weiss WA: Neuroblastoma and MYCN. Cold Spring Harbor perspectives in medicine 2013, 3(10):a014415.
6. Cohn SL, Pearson AD, London WB, Monclair T, Ambros PF, Brodeur GM, Faldum A, Hero B, Iehara T, Machin D et al: The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2009, 27(2):289-297.
7. De Kouchkovsky I, Abdul-Hay M: Acute myeloid leukemia: a comprehensive review and 2016 update. Blood cancer journal 2016, 6(7):e441.
8. Sandahl JD, Kjeldsen E, Abrahamsson J, Ha SY, Heldrup J, Jahnukainen K, Jonsson OG, Lausen B, Palle J, Zeller B et al: The applicability of the WHO classification in paediatric AML. A NOPHO-AML study. British journal of haematology 2015, 169(6):859-867.
9. Inoue K, Ogawa H, Sonoda Y, Kimura T, Sakabe H, Oka Y, Miyake S, Tamaki H, Oji Y, Yamagami T et al: Aberrant overexpression of the Wilms tumor gene (WT1) in human leukemia. Blood 1997, 89(4):1405-1412.
10. Cilloni D, Renneville A, Hermitte F, Hills RK, Daly S, Jovanovic JV, Gottardi E, Fava M, Schnittger S, Weiss T et al: Real-time quantitative polymerase chain reaction detection of minimal residual disease by standardized WT1 assay to enhance risk stratification in acute myeloid leukemia: a European LeukemiaNet study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2009, 27(31):5195-5201.
11. Zhong JF, Chen Y, Marcus JS, Scherer A, Quake SR, Taylor CR, Weiner LP: A microfluidic processor for gene expression profiling of single human embryonic stem cells. Lab on a chip 2008, 8(1):68-74.
12. Kalisky T, Quake SR: Single-cell genomics. Nature methods 2011, 8(4):311-314.
13. Terwijn M, Zeijlemaker W, Kelder A, Rutten AP, Snel AN, Scholten WJ, Pabst T, Verhoef G, Lowenberg B, Zweegman S et al: Leukemic Stem Cell Frequency: A Strong Biomarker for Clinical Outcome in Acute Myeloid Leukemia. PloS one 2014, 9(9).
14. Pollyea DA, Jordan CT: Therapeutic targeting of acute myeloid leukemia stem cells. Blood 2017.
15. Rosenfeld C, Cheever MA, Gaiger A: WT1 in acute leukemia, chronic myelogenous leukemia and myelodysplastic syndrome: therapeutic potential of WT1 targeted therapies. Leukemia 2003, 17(7):1301-1312.
16. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, Scholten WJ, Snel AN, Veldhuizen D, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. Leukemia 2016, 30(2):439-446.
17. Kalina T, Flores-Montero J, van der Velden VH, Martin-Ayuso M, Bottcher S, Ritgen M, Almeida J, Lhermitte L, Asnafi V, Mendonca A et al: EuroFlow standardization of flow

cytometer instrument settings and immunophenotyping protocols. Leukemia 2012, 26(9):1986-2010.
18. Porichis F, Hart MG, Griesbeck M, Everett HL, Hassan M, Baxter AE, Lindqvist M, Miller SM, Soghoian DZ, Kavanagh DG et al: High-throughput detection of miRNAs and gene-specific mRNA at the single-cell level by flow cytometry. Nature communications 2014, 5:5641.
19. Henning AL, Sampson JN, McFarlin BK: Measurement of Low-Abundance Intracellular mRNA Using Amplified FISH Staining and Image-Based Flow Cytometry. Current protocols in cytometry / editorial board, J Paul Robinson, managing editor [et al] 2016, 76:7 46 41-48.
20. Henning AL, McFarlin BK: Consumption of a high-fat, high-calorie meal is associated with an increase in intracellular co-localization of PPAR-gamma mRNA and protein in monocytes. Methods 2016.
21. Soh KT, Tario JD, Jr., Colligan S, Maguire O, Pan D, Minderman H, Wallace PK: Simultaneous, Single-Cell Measurement of Messenger RNA, Cell Surface Proteins, and Intracellular Proteins. Current protocols in cytometry / editorial board, J Paul Robinson, managing editor [et al] 2016, 75:7 45 41-47 45 33.
22. Martrus G, Niehrs A, Cornelis R, Rechtien A, Garcia-Beltran W, Lutgehetmann M, Hoffmann C, Altfeld M: Kinetics of HIV-1 Latency Reversal Quantified on the Single-Cell Level Using a Novel Flow-Based Technique. Journal of virology 2016, 90(20):9018-9028.
23. Henning AL, Levitt DE, Vingren JL, McFarlin BK: Measurement of T-Cell Telomere Length Using Amplified-Signal FISH Staining and Flow Cytometry. Current protocols in cytometry / editorial board, J Paul Robinson, managing editor [et al] 2017, 79:7 47 41-47 47 10.
24. Bachas C, Schuurhuis GJ, Assaraf YG, Kwidama ZJ, Kelder A, Wouters F, Snel AN, Kaspers GJL, Cloos J: The role of minor subpopulations within the leukemic blast compartment of AML patients at initial diagnosis in the development of relapse. Leukemia 2012, 26(6):1313-1320.
25. Maecker HT, Frey T, Nomura LE, Trotter J: Selecting fluorochrome conjugates for maximum sensitivity. Cytometry Part A : the journal of the International Society for Analytical Cytology 2004, 62(2):169-173.
26. Baumgarth N, Bigos M: Optimization of emission optics for multicolor flow cytometry. Methods in cell biology 2004, 75:3-22.
27. Bozzi F, Gambirasio F, Luksch R, Collini P, Brando B, Fossati-Bellani F: Detecting CD56+/NB84+/CD45- immunophenotype in the bone marrow of patients with metastatic neuroblastoma using flow cytometry. Anticancer research 2006, 26(5A):3281-3287.
28. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology 2002, 3(7):RESEARCH0034.
29. Preffer F, Dombkowski D: Advances in complex multiparameter flow cytometry technology: Applications in stem cell research. Cytometry Part B, Clinical cytometry 2009, 76(5):295-314.
30. Tamaki H, Ogawa H, Inoue K, Soma T, Yamagami T, Miyake S, Oka Y, Oji Y, Tatekawa T, Tsuboi A et al: Increased expression of the Wilms tumor gene (WT1) at relapse in acute leukemia. Blood 1996, 88(11):4396-4398.
31. Ogawa H, Tamaki H, Ikegame K, Soma T, Kawakami M, Tsuboi A, Kim EH, Hosen N, Murakami M, Fujioka T et al: The usefulness of monitoring WT1 gene transcripts for the prediction and management of relapse following allogeneic stem cell transplantation in acute type leukemia. Blood 2003, 101(5):1698-1704.
32. Xu J, Wang HW, Yang T, Tan YH, Zhang L: [Expression of WT1 gene in CD34(+)CD38(-)CD123(+) AML stem cells and its significance analysis.]. Zhonghua xue ye xue za zhi = Zhonghua xueyexue zazhi 2010, 31(3):172-175.
33. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL et al: The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clinical chemistry 2009, 55(4):611-622.
34. Beillard E, Pallisgaard N, van der Velden VH, Bi W, Dee R, van der Schoot E, Delabesse E, Macintyre E, Gottardi E, Saglio G et al: Evaluation of candidate control genes for diagnosis

and residual disease detection in leukemic patients using 'real-time' quantitative reverse-transcriptase polymerase chain reaction (RQ-PCR) - a Europe against cancer program. Leukemia 2003, 17(12):2474-2486.

Supplementary Tables and Figures

Table S1. Detailed information on AML patient samples.

no.	Sexe	Age at Dx (yr.)	Diagnostic material	Technique	Cell population	fraction (%) of WBC	*WT1* RT-qPCR NRQ result (cut-off 0.50[a])
AML 1	M	50	PB	morphology	total leukemic cells	59.0	overexpression (0.90)
				FCM	bulk (CD34+)	46.6	
					LSCs (CD34+CD38-)	31.2	
					blasts (CD34+CD38+)	15.3	
AML 2	M	78	BM	morphology	total leukemic cells	54.0	normal expression (< cut-off)
				FCM	bulk (CD34+)	49.5	
					LSCs (CD34+CD38-)	1.00	
					blasts (CD34+CD38+)	48.50	
AML 3	M	68	BM	morphology	total leukemic cells	84.5	overexpression (0.96)
				FCM	bulk (CD34+)	58.2	
					LSCs (CD34+CD38-)	15.0	
					blasts (CD34+CD38+)	43.3	
AML 4	M	4	BM	morphology	total leukemic cells	96.0	normal expression (< cut-off)
				FCM	bulk (CD34+)	99.0	
					LSCs (CD34+CD38-)	11.5	
					blasts (CD34+CD38+)	87.6	

[a]An in-house validated NRQ cut-off of 0.50, based on the expression in normal hematopoietic cells of healthy individuals, was used to define overexpression. Abbreviations: AML: acute myeloid leukemia, FCM: flow cytometry, Dx: diagnosis, PB: peripheral blood, BM: bone marrow, WBC: white blood cells, *WT1*: Wilms' tumor 1, LSCs: leukemic stem cells, NRQ: normalized relative quantity (versus *Abelson* (*ABL1*)).

Table S2. Detailed characteristics of the used antibodies.

Antigen	Fluorochrome	Clone	Dilution	Company
CD45	FITC	2D1	5 µL / 10^6 cells	BD Biosciences
CD56	PE	C5.9	1.25 µL / 10^6 cells	Cytognos
CD45	Pacific Orange	HI30	1.25 µL / 10^6 cells	Life Technologies
CD45	V500	HI30	1.25 µL / 10^6 cells	BD Biosciences
CD34	PE	8G12	1.25 µL / 10^6 cells	BD Biosciences
CD38	PE-Cy7	HB7	1.25 µL / 10^6 cells	BD Biosciences

Fig. S1. Flowcytometric evaluation of CD45 (A) and CD56 (B) expression in SK-N-BE(2)-C cells prior to PrimeFlow™.

Fig. S1 A

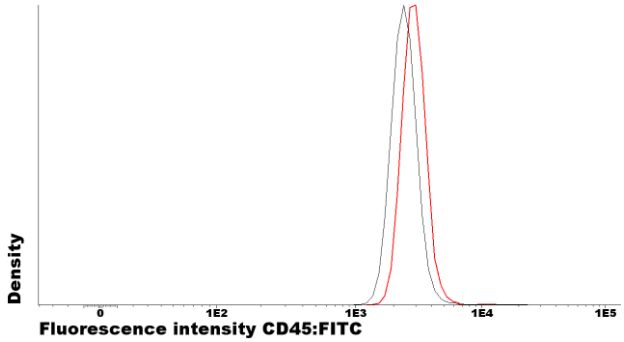


Fig. S1 B

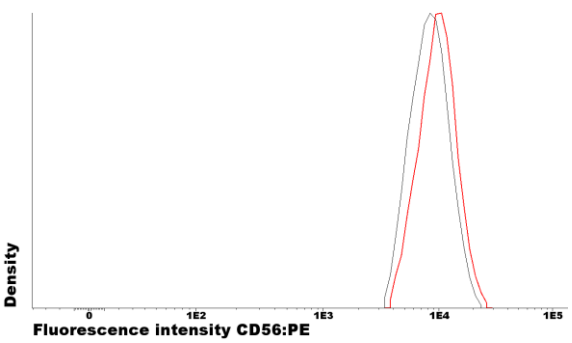


Fig. S1 C

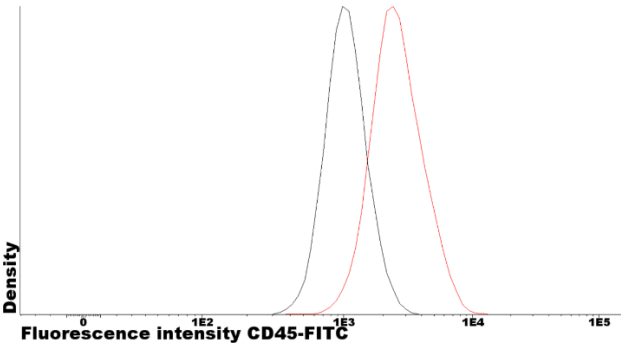
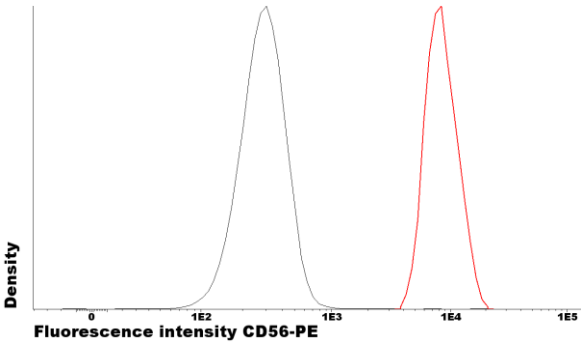


Fig. S1 D



(A) Head-to-head comparison between CD45 expression measured by standard membrane staining (grey) and by membrane staining followed by PrimeFlow™ RNA assay (red), showing a neglectable shift in fluorescence intensity.
(B) Head-to-head comparison between CD56 expression measured by standard membrane staining (grey) and by membrane staining followed by PrimeFlow™ RNA assay (red), showing a neglectable shift in fluorescence intensity.
(C) Comparison of the fluorescence intensity in the FITC-channel without (MFI 1096) and with (MFI 2820) CD45 membrane staining, illustrated by a grey and red curve respectively, shows a high degree of aspecific fluorescence regarding the unstained (grey) population.
(D) Comparison of the fluorescence intensity in the PE-channel without (MFI 363, SD 127) and with (MFI 36258) CD56 membrane staining, illustrated by a grey and red curve respectively, shows no aspecific fluorescence regarding the unstained (grey) population.

Fig. S2. Evaluation of the off-target effect.

Fig. S2 A

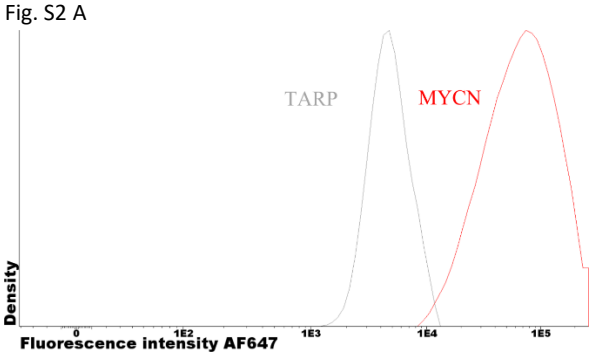
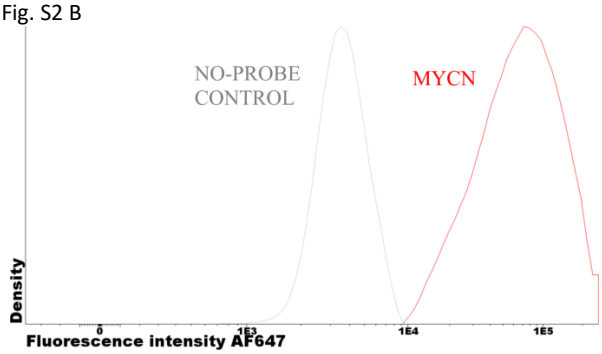
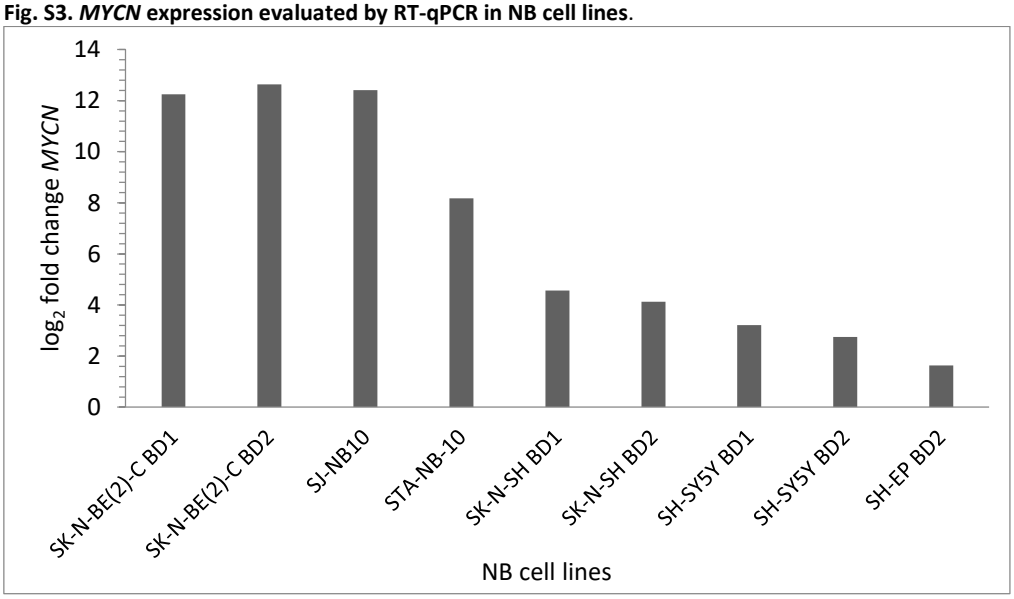


Fig. S2 B



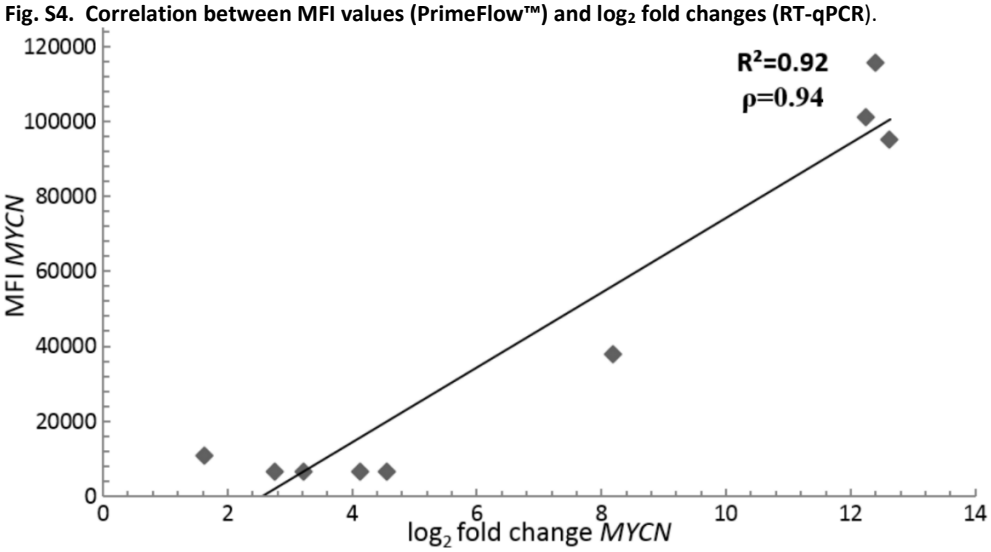
(A) Comparison of AF647 fluorescence intensity in SK-N-BE(2)-C cells hybridised with *TARP* type 1 target probe (non-specific expression; grey (MFI 4022)) and *MYCN* type 1 target probe (specific target expression; red (MFI 79927)).
(B) Comparison of AF647 fluorescence intensity in SK-N-BE(2)-C cells deriving from the no-probe control (background expression; grey (MFI 3789)) or *MYCN* type 1 target probe (specific target expression; red (MFI 79927)).
Abbreviations: AF: AlexaFluor, TARP: TCR gamma alternate reading frame protein.

Fig. S3. *MYCN* expression evaluated by RT-qPCR in NB cell lines.



Log_2 fold changes (calculated against sample with lowest expression, SH-EP BD1) for *MYCN* expression in six NB cell lines: SK-N-BE(2)-C, SJ-NB10, STA-NB-10, SK-N-SH, SH-SY5Y and SH-EP. Biological duplicates (BD) were analyzed, except for SJ-NB10 and STA-NB-10. Abbreviations: RT-qPCR: real-time quantitative polymerase chain reaction, NB: neuroblastoma.

Fig. S4. Correlation between MFI values (PrimeFlow™) and log_2 fold changes (RT-qPCR).



Abbreviations: MFI: mean fluorescence intensity, R^2: coefficient of determination, ρ: spearman rank correlation coefficient, RT-qPCR: real-time quantitative polymerase chain reaction.

Fig. S5. MFI values regarding *WT1* expression by PrimeFlow™ in de novo AML patients.

Fig. S5 A.

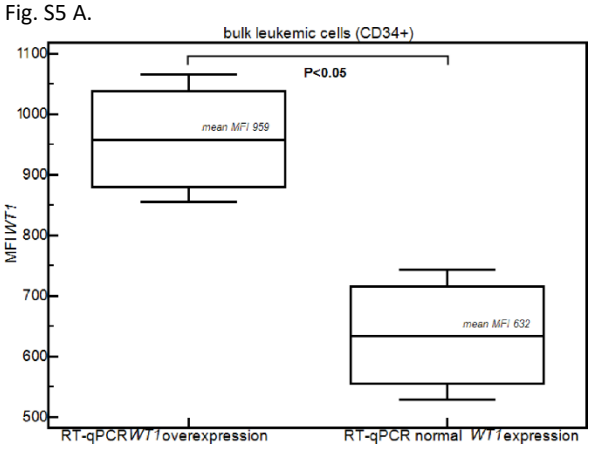


Fig. S5 B.

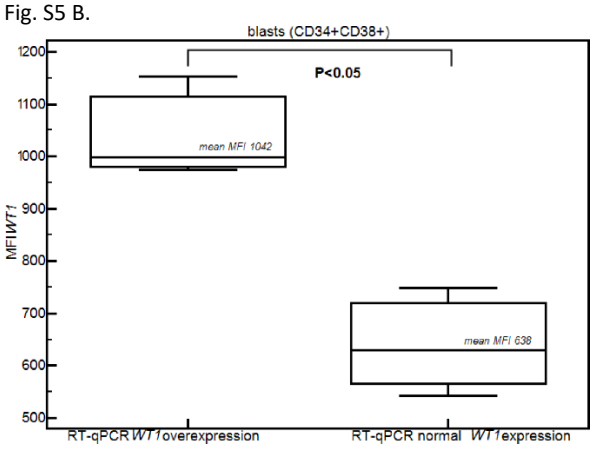
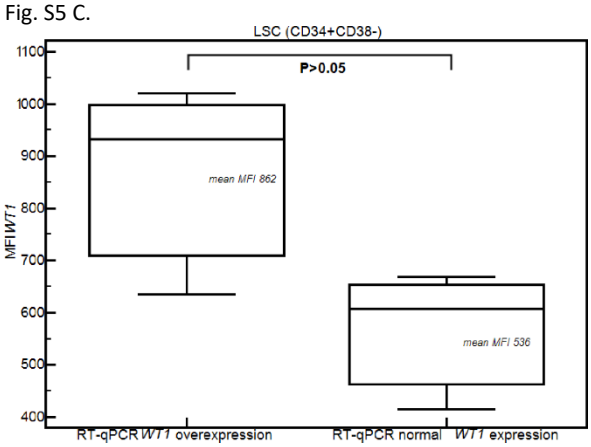


Fig. S5 C.



(A) *WT1* expression was shown to be statistically significant upregulated in the bulk leukemic cells (CD34+) from AML patients, previously defined by RT-qPCR as having *WT1* overexpression (left) versus normal *WT1* expression (right) (mean MFI±SD 959±105 vs. 632±91, P<0.05). Detailed overview patient samples is shown in Table S1. Statistical analysis was performed by the Mann-Whitney U test (MedCalc).
(B) *WT1* expression was shown to be statistically significant upregulated in the blasts (CD34+CD38+) from AML patients, previously defined by RT-qPCR as having *WT1* overexpression (left) versus normal *WT1* expression (right) (mean MFI±SD 1042±97 vs. 638±88, P<0.05). Detailed overview patient samples is shown in Table S1. Statistical analysis was performed by the Mann-Whitney U test (MedCalc).
(C) *WT1* expression was shown to be upregulated in the LSCs (CD34+CD38-) from AML patients, previously defined by RT-qPCR as having *WT1* overexpression (left) versus normal *WT1* expression (right), although not at a significant level (mean MFI±SD 862±202 vs. 536±94, P<0.05). Detailed overview patient samples is shown in Table S1. Statistical analysis was performed by the Mann-Whitney U test (MedCalc).

VII.1.

Introduction:

The discovery of *TARP* in androgen-dependent prostate and breast carcinoma.

A lot of discoveries in science involve chance circumstances in a particularly salient way. If Sir Isaac Newton had not returned home to Woolsthorpe because his university was shut down due to the plague, he would (probably) not have had the opportunity to reflect on his orchard and observe that particular apple fall from a tree. Although it took another 20 years before his theory of gravity was established, this tree helped him provoke the idea [1].

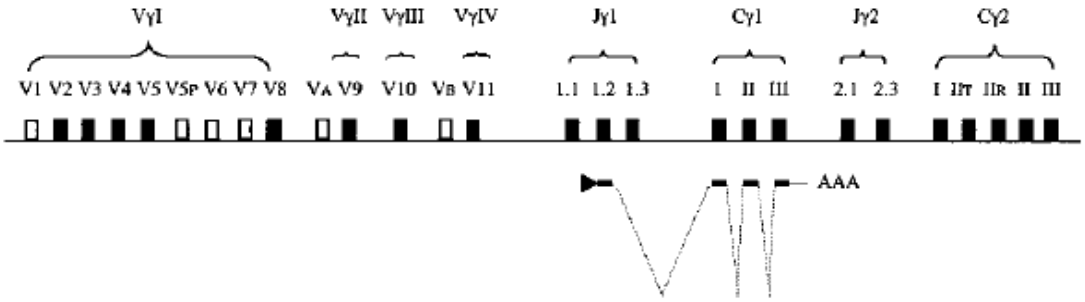
Less people are familiar with the discovery of Essand et al. in Uppsala [2]. In 1999, this research group was exploring the database of expressed sequence tags (dbEST) of Genbank, containing sequences originating from cDNA libraries prepared from different cell types. They observed that many human prostate ESTs comprised untranslated 3'end cDNA sequences of the T-cell receptor (TCR) gamma (γ) chain, but lacked ESTs from TCR delta (δ), TCR alpha (α) and TCR beta (β) chains. This was a remarkable finding, since TCRγ and TCRδ chains physiologically interact through disulphide bridging during the formation of a TCRγδ. Diversity of the TCRγ locus in T-cells is generated through somatic recombination of variable (V), joining (J) and constant (C) exons. TCRγδ T-cells are much less common than TCRαβ T-cells, i.e. 2% of the total T-cell repertoire. The highest share of TCRγδ- T-cells can be found in the gut mucosa, i.e. intra-epithelial lymphocytes.

Alignment of prostate TCRγ ESTs with TCRγ ESTs from peripheral blood T-cells in the databank confirmed their identical composite sequence. Hence, the authors concluded that the TCRγ gene is highly transcribed in tumorigenic and healthy human prostate tissue. Two explanations could account for this discovery: (i) the TCRγ transcripts derived from infiltrating TCRγδ T-cells, or, (ii) TCRγ transcripts had a non-lymphocytic prostate cell origin. RNA dot blot and northern blot hybridisation performed on normal prostate tissue confirmed the presence of TCRγ transcripts without TCRδ or CD3 transcripts. Two TCRγ transcripts were observed in normal prostate tissue, i.e. a strongly expressed 1.1 kb fragment and a less represented 2.8 kb fragment. When evaluating prostate adenocarcinoma cell lines, only the 1.1 kb TCRγ transcript was found in the androgen-dependent LNCaP cell line. No TCRγ transcript was found in the androgen-insensitive PC-3 cell line. By contrast, the transcript from TCRγδ T-cells was shown to be 1.5 kb in size. Since dbEST analysis had confirmed identical 3' end sequences, the authors hypothesized that transcript size differences must be situated upstream.

They then decided to characterize the predominant 1.1kb prostate-specific TCRγ transcript isolated from LNCaP. The different size between the prostate- and T-cell TCRγ transcript appeared to be attributable to several features: (i) The prostate-specific TCRγ locus underwent no recombination process and consequently lacked VJ rearrangements, (ii) transcription of the prostate-specific TCRγ locus was initiated in the intronic sequence of the Jγ1.2 gene segment, and (iii) the prostate-specific TCRγ locus contains a polyadenylation signal and poly(A) sequence at the 3'end.

In conclusion, these experiments revealed that the prostate-specific TCRγ transcript comprised 1023 bp, i.e. 53bp of the Jγ1.2 gene, correctly spliced to three CγI exons (519 bp) followed by 448 bp of untranslated sequence containing a polyadenylation signal and poly(A) sequence at the 3'end (Fig. 1A). Hence, the prostate-specific TCR(JC)γ transcript can be considered as a truncated TCRγ transcript that differs from the TCRγ transcripts observed in lymphoid cells. *In vitro* transcription-coupled translation of the prostate-specific TCRγ cDNA revealed that the transcript is fully functional and encodes two proteins. Different open reading frames (ORFs) were active is prostate tissue compared to T-cells, which led to a prominent 6.8 - 8 kDa-protein (third underlined ATG in Fig. 1B) and a small 13 kDa-protein (first double underlined ATG in Fig. 1B).

A



B

```
<---------------------- J gamma 1.2 ----------------------->
GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA

<-------------------- C gamma 1 (exon CI) -------------------
GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT

----------------------------------------------------------->>
GAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGAT

-------------------------------------------------------------
GTTATTAAGATACATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGG

-------------------------------------------------------------
       M  K  T  N  D  T  Y  M  K  F  S  W  L  T  V  P  E  K
AACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG

-------------------------------------------------------------
 S  L  D  K  E  H  R  C  I  V  R  H  E  N  N  K  N  G  V  D
TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGAT

------------------------------><---- C gamma 1 (exon CII) ---
 Q  E  I  I  F  P  P  I  K  T  D  V  I  T  M  D  P  K  D  N
CAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAAT

-----------------><---------- C gamma (exon CIII) ----------
 C  S  K  D  A  N  D  T  L  L  L  Q  L  T  N  T  S  A  Y  Y
TGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTAC

-------------------------------------------------------------
 M  Y  L  L  L  L  L  K  S  V  V  Y  F  A  I  I  T  C  C  L
ATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTG

------------------------------------>
 L  R  R  T  A  F  C  C  N  G  E  K  S
CTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCATAACAGACGGTGGCACAAGGA

GGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAGCGTCTTCTGAGGATCTAGTTGGGC

TTTCTTTCTGGGTTTGGGCCATTTCAGTTCTCATGTGTGTACTATTCTATCATTATTGTA

TAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGAAT

AGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAACCC

AAATAGCGCCTGCTATAGTGTAGACATCCTGCGGCTTCTAGCCTTGTCCCTCTCTTAGTG

TTCTTTAATCAGATAACTGCCTGGAAGCCTTTCATTTTACACGCCCTGAAGCAGTCTTCT

TTGCTAGTTGAATTATGTGGTGTGTTTTTCCGTAATAAGCAAAATAAATTTAAAAAAATG

AAAAGTT
```

Figure 1. The prostate-specific TCR(JC)γ transcript, adapted from Essand et al.[2]. (A) The transcript consists of a Jγ1.2 segment, the three exons of Cγ1, followed by untranslated sequence. (B) Nucleotide sequence of the TCRγ transcript as obtained from LNCaP cDNA. The starting point of transcription (underlined) is within the 10 first nucleotides of the Jγ1.2 segment. The four translational initiation codons (ATG) in the original TCRγ reading frame are double underlined.

One year later, the authors defined which start codons are responsible for the TCR(JC)γ protein [3]. The prostate-specific protein was formed by the ORF located upstream of the conventional TCRγ ORF. This ORF contained two independent start codons, and only mutagenesis of both start codons led to loss of the protein. This short ORF encoded a 7-kDa protein with no resemblance to any published protein sequence in Genbank. From that moment on, this protein was referred to as the "TCRγ alternate reading frame protein" (TARP). Furthermore, they detected that the TARP transcript is also expressed in breast cancer cell lines and breast cancer tissues, while absent in neuro- and glioblastoma, colon, gastric and kidney cell lines. Normal breast tissues did not harbour TARP transcript expression, in contrast to normal prostate tissue, suggesting that expression in breast tissue is mostly likely increased by oncogenic transformation.

Beside its alternate ORF, and the fact that either start codons in this ORF can initiate protein synthesis, the TARP protein harboured even more unusual features: (i) despite its small size, this protein was not secreted, (ii) lacks a good Kozak sequence, and (iii) the sequence contained five leucines in heptad repeats, suggesting that TARP possibly contains a leucine zipper dimerization motif. It was hypothesised that the basic region in TARP functions as a nuclear DNA-binding site and had a potential role as transcription repressor.

Establishment of a TARP overexpression constructs, followed by cDNA micro-array analysis, showed that high TARP levels promote proliferation and motility and inhibit apoptosis (increased amphiregulin expression), promotes metastatic growth (increased caveolin 1 expression) and promotes tumor progression (increased chemokine C-X-C motif ligand 1) [4]. These results suggest that TARP has a role in regulating growth and gene expression in prostate cancer cells. *TARP* expression was also shown to be transcriptionally regulated by androgen receptor (AR) binding to androgen response element (ARE) in the proximal *TARP* promoter site, and upregulated upon androgen stimulation [5, 6].

In 2004, the subcellular localization of TARP was refined using a monoclonal TARP antibody (TP1) [6]. Cell fractionation, Western blotting and immunocytochemistry revealed that TARP is located in the outer mitochondrial membrane. Immunohistochemistry using the human prostate cancer cell line LNCaP showed that TP1 reacted in a dot-like cytoplasmic pattern, consistent with presence in mitochondria (Fig. 2). In the cells labelled with anti-HSP60 antibody, the same cytoplasmic distribution was observed and the merged images showed marked yellow signals, indicating co-localization of TARP and HSP60. This result demonstrates that the endogenous TARP protein is localized in mitochondria.

Altogether, in only five years of research (1999-2004), TARP was discovered to be a novel tumor-associated antigen in androgen-dependent prostate carcinoma and breast carcinoma. TARP can be defined as a prostate-specific, truncated TCRγ transcript that is upregulated by androgen stimulation and encodes a 7 kDa protein (58 amino acids) localized in the mitochondria. These findings had aroused the interest of researchers familiar with the development of immunotherapeutic strategies.

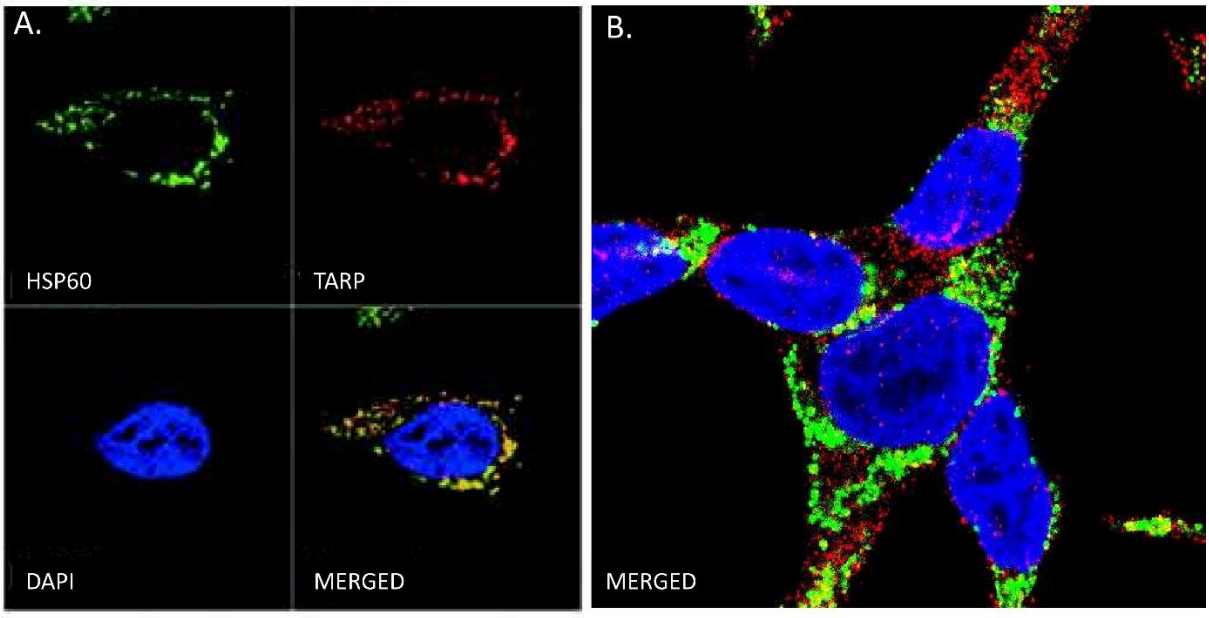


Fig. 2. Localization of endogenous TARP in LNCaP by confocal microscopy. HSP60, a mitochondrial marker, is indicated in green and TARP is indicated in red. Overlap between HSP60 and TARP is visualised by yellow fusion signals. Nuclei were stained with 4,6-diamidino-2-phenylindole (DAPI, blue). **(A)** Image adapted from Maeda et al. [6], showing a high overlap between HSP60 and TARP. **(B)** Image generated in own experiments, showing occasional overlap between HSP60 and TARP.

The group of Essand et al. [7] evaluated which human leukocyte antigen (HLA)-A*0201-restricted TARP peptides were capable of triggering cytotoxic T lymphocytes (CTL). Dendritic cells were pulsed either with wild-type (WT) or synthetic peptides having enhanced affinity for HLA-A*0201. CTLs generated after affinity-enhanced TARP(P5L)$_{4-13}$ peptide stimulation could recognize target cells displaying WT and mutant peptides. TARP-specific CTLs were then tetramer-sorted and expanded by general T cell stimulation while retaining specificity and activity. Using *in vitro* cytotoxicity experiments, TARP(P5L)$_{4-13}$-directed CTLs exerted moderate lysis of the prostate cancer cell line LNCaP and breast cancer cell line MCF-7. These data indicated that the TARP(4-13) epitope is endogenously processed and presented HLA-presented. The therapeutic value of targeting the HLA-A*0201-enhanced affinity TARP(P5L)$_{4-13}$ peptide in prostate and breast cancer immunotherapy was confirmed in a second follow-up study [7], using transgenic T-cells transduced with a cloned TCR directed against the TARP(P5L)$_{4-13}$ epitope.

The research group of Epel et al. used an antibody (Ab)-based approach [8]. They isolated recombinant Fab antibodies with peptide-specific, MHC-restricted TCR-like reactivity directed toward HLA-A2-restricted T cell epitopes from TARP. A recombinant Ab-toxin fusion molecule was generated by linking a pseudomonas endotoxin to the TCR-like antibody, and demonstrated to eradicate tumor cells in a peptide-specific, MHC-restricted manner.

In conclusion, TARP had emerged in less than a decade as a very promising immunotherapeutic target in androgen-dependent prostate and breast carcinoma with minimal ‘off-target’ effect.

References

1. Gaughan R: Accidental Genius: The World's Greatest By-Chance Discoveries. Metro Books 2010.
2. Essand M, Vasmatzis G, Brinkmann U, Duray P, Lee B, Pastan I: High expression of a specific T-cell receptor gamma transcript in epithelial cells of the prostate. Proceedings of the National Academy of Sciences of the United States of America 1999, 96(16):9287-9292.
3. Wolfgang CD, Essand M, Vincent JJ, Lee B, Pastan I: TARP: A nuclear protein expressed in prostate and breast cancer cells derived from an alternate reading frame of the T cell receptor gamma chain locus. Proceedings of the National Academy of Sciences of the United States of America 2000, 97(17):9437-9442.
4. Wolfgang CD, Essand M, Lee B, Pastan I: T-cell receptor gamma chain alternate reading frame protein (TARP) expression in prostate cancer cells leads to an increased growth rate and induction of caveolins and amphiregulin. Cancer research 2001, 61(22):8122-8126.
5. Cheng WS, Giandomenico V, Pastan I, Essand M: Characterization of the androgen-regulated prostate-specific T cell receptor gamma-chain alternate reading frame protein (TARP) promoter. Endocrinology 2003, 144(8):3433-3440.
6. Maeda H, Nagata S, Wolfgang CD, Bratthauer GL, Bera TK, Pastan I: The T cell receptor gamma chain alternate reading frame protein (TARP), a prostate-specific protein localized in mitochondria. The Journal of biological chemistry 2004, 279(23):24561-24568.
7. Hillerdal V, Nilsson B, Carlsson B, Eriksson F, Essand M: T cells engineered with a T cell receptor against the prostate antigen TARP specifically kill HLA-A2+ prostate and breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America 2012, 109(39):15877-15881.
8. Epel M, Carmi I, Soueid-Baumgarten S, Oh S, Bera T, Pastan I, Berzofsky J, Reiter Y: Targeting TARP, a novel breast and prostate tumor-associated antigen, with T cell receptor-like human recombinant antibodies. European journal of immunology 2008, 38(6):1706-1720.

VII.2.

Results:

TARP is an immunotherapeutic target in acute myeloid leukemia expressed in the leukemic stem cell compartment.

B. Depreter, K. E. Weening, K. Vandepoele, M. Essand, B. De Moerloose, M. Themeli, J. Cloos, D. Hanekamp, I. Moors, I. D' hont, B. Denys, A. Uyttebroeck, A. Van Damme, L. Dedeken, S. Snauwaert, G. Goetgeluk, S. De Munter, T. Kerre, B. Vandekerckhove, T. Lammens*, J. Philippé *

(*shared senior authorship)

Haematologica August 2019: haematol.2019.222612; Doi:10.3324/haematol.2019.222612

Abstract awarded with a Travel Grant Award on the 24th EHA Congress (2019).

Abstract

Immunotherapeutic strategies targeting the rare leukemic stem cell compartment might provide salvage to the high relapse rates currently observed in acute myeloid leukemia. We applied gene expression profiling for comparison of leukemic blasts and leukemic stem cells with their normal counterparts. Here, we show that the T-cell receptor γ chain alternate reading frame protein (TARP) is overexpressed in *de novo* pediatric (n=13) and adult (n=17) AML sorted leukemic stem cells and blasts compared to hematopoietic stem cells and normal myeloblasts (15 healthy controls). Moreover, TARP expression was significantly associated with a fms-like tyrosine kinase receptor-3 internal tandem duplications in pediatric AML. TARP overexpression was confirmed in acute myeloid leukemia cell lines (n=9), and was found to be absent in B-cell acute lymphocytic leukemia (n=5) and chronic myeloid leukemia (n=1). Sequencing revealed that both a classical TARP transcript, as described in breast and prostate adenocarcinoma, and an acute myeloid leukemia-specific alternative TARP transcript, were present. Protein expression levels mostly matched transcript levels. TARP was shown to reside in the cytoplasmic compartment and showed sporadic endoplasmic reticulum co-localization. TARP-TCR engineered cytotoxic T-cells *in vitro* killed AML cell lines and patient leukemic cells co-expressing TARP and HLA-A*0201. In conclusion, TARP qualifies as a relevant target for immunotherapeutic T-cell therapy in AML.

Introduction

Acute myeloid leukemia (AML) is a heterogeneous hematological malignancy, accounting for 80% of adult [1-4] and 20% of pediatric [5-7] leukemia. Despite initial clinical remission rates between 60-90%[2, 5, 6], patients exhibit a high relapse risk and therapy-related mortality, resulting in a 5-year overall survival of 30% in adult AML[1, 3] and 65-70% in pediatric AML (pedAML)[5, 8]. Especially the prognosis of patients with fms-like tyrosine kinase receptor-3 internal tandem duplications (*FLT3*-ITD) remains extremely poor [2, 8, 9]. The high relapse rate is thought to arise from a chemotherapy-resistant cell fraction with unlimited self-renewal capacities, denominated as leukemic stem cells (LSCs)[4, 10-14]. In CD34+ AML, stem cell characteristics were shown to be present in all four CD34/CD38 phenotypic compartments, though with the CD34+CD38- fraction being most LSC-enriched[15]. Moreover, a high LSC load at diagnosis was shown to be a significant adverse prognostic factor [16-19]. Unfortunately, current chemotherapeutic regimens were shown to be inadequate towards LSC eradication[14] and induce important toxicity[5, 6, 20]. Also hematopoietic stem cell transplantation, performed in high-risk (HR) patients or as salvage therapy, carries a high mortality and morbidity risk [2, 5], highlighting the need for alternative treatments. Thus, identifying LSC aberrations is crucial to tackle the high relapse rate and to develop therapeutic targeting strategies for LSC elimination, while ensuring salvage of normal hematopoietic stem cells (HSCs).

Targeted therapy has led to a remarkable progress in the survival rates of multiple cancers. The introduction of tyrosine kinase inhibitors in the treatment of chronic myeloid leukemia (CML) accomplished a major breakthrough, and CD19-directed chimeric antigen receptor (CAR) therapy has improved survival in relapsed/refractory pediatric ALL tremendously [21, 22]. These successes paved the way for the exploration of the clinical applicability of targeting antibodies and CAR- or T-cell receptor (TCR)-transgenic cytotoxic T-cells (CTLs) in AML [2, 23-28]. Although an increasing number of LSC-specific membrane markers have been identified the past years[18, 23, 29, 30], only few reports address the molecular abnormalities of LSC compared to HSC [15, 31-37], especially in pedAML.

Here, we identified the T-cell receptor (TCR) γ chain alternate reading frame protein (TARP) as an AML-specific target, expressed in the LSCs and blasts of pediatric and adult AML, while absent in their normal counterparts. TARP transcript expression was associated with FLT3-ITD in pedAML. In addition, we provide *in vitro* evidence that TARP may serve as a novel immunotherapeutic target in AML for TARP-TCR engineered CTLs.

Methods

Patients

We retrospectively selected diagnostic material from 13 pedAML and 17 adult AML patients based on the sample availability, LSC load, CD34 positivity, FLT3 mutational status and HLA-status (Table 1, Table S1). At diagnosis, mononuclear cells (MNC) were isolated from bone marrow (BM) or peripheral blood (PB) by Ficoll density gradient (Axis-shield) and cryopreserved in 90% fetal calf serum (FCS) and 10% dimethylsulfoxide. Samples were thawed, followed by 30 min incubation at room temperature (RT) in 20 mL RPMI with 20% FCS, 200 µL DNase I (1 mg/mL, grade II bovine pancreas) and 200 µL $MgCl_2$ (1 M) (Sigma-Aldrich). After incubation, cells were spinoculated (10 min, 400 rpm) and washed once more with RPMI/20% FCS.

In addition, we prospectively collected material from 15 healthy subjects. Normal bone marrow (NBM, n=6) was collected from posterior iliac crest of pediatric patients (4-18 years) undergoing scoliosis surgery. Umbilical cord blood (CB, n=7) was obtained after normal vaginal deliveries at full term. Mobilized peripheral blood stem cells (mPBSC, n=2) were collected by apheresis of adult donors pre-allotransplant. All patients or their guardians gave their informed consent and approval was obtained by the ethical committee, in accordance to the declaration of Helsinki. Buffy coats from donors were obtained from the Red Cross (Mechelen, Belgium) and used for CTL isolation and the preparation of feeder cell medium.

Flow cytometry (FCM) analysis and cell sorting

Cell pellets were surface stained (Table S2), followed by 20 min incubation at 4 °C and washing with PBS+2% BSA. For cell-sorting, labeled cells were resuspended in medium and sorted on a FACSAria III with red, blue, and ultraviolet lasers (BD Biosciences). For FCM analysis, cells were resuspended in PBS+2% BSA and analyzed on a LSR II or a FACSCanto II, equipped with four or three solid-state lasers, respectively (both BD Biosciences). All scatters were devoid of doublets based on FSC-H/FSC-A, and propidium iodide (PI) was used to exclude dead cells. Sorting strategies are described in Supplementary data 2.2. Regarding FCM-based cytotoxicity and cytokine assays (Supplementary data 2.9), living cells were selected using a LIVE/DEAD staining (1:10000 dilution, ThermoFisher Scientific) instead of PI. Target cells were stained with a Violet CellTrace™ (VT) Cell Proliferation Kit (5 mM, 1:10000 dilution, ThermoFisher Scientific) prior to incubation with TCR-engineered CTLs. After incubation and before surface staining, Flow-Count™ Fluorospheres (1:20 diluted, Beckman Coulter) were added to each well to enable target quantification (measurement of minimum 1000 Fluorospheres/well).

Transcript expression

Details on micro-array profiling, RNA isolation, cDNA synthesis, (quantitative) PCR conditions and primers can be found in Supplementary data (2.3, 2.4, 2.5) and Table S3. qPCR data analysis was performed according to state-of-the-art methods [38, 39]. Relative quantity (RQ) values were normalized against housekeeping genes *GAPD*, *HPRT1* and *TBP*. For TARP expression, normalized relative quantities were calibrated (CNRQ) versus a single calibrator to allow interrun comparison. For the investigation of the subcellular localization of TARP, delta (d) Ct between cytoplasmic and nuclear compartments were calculated and compared to *MALAT1* and *TBP* expression. Functional *TCRG* gene rearrangements were excluded if sufficient material remained using DNA TCRG GeneScan analysis [40] and/or TRGV(J)C qPCR (Table S4).

Protein detection

Details on Western blotting and confocal microscopy are provided in Supplementary data 2.6.

Viral transduction of AML cell lines and generation of TCR-transgenic CTLs

All transfer and helper plasmids used and procedures for transformation, plasmid isolation, transfection and transduction are described in Supplementary data 2.7 and 2.8.

Six AML cell lines (HL-60, Kg-1a, MOLM-13, HL-60-Luc, MOLM-13-Luc and MV4;11-Luc) were transduced with HLA-A*0201 MHC-I encoding retrovirus, hereafter defined by the suffix A2+. Transgenic TARP overexpression (OE) cell lines were generated for OCI-AML3 and THP-1, next to mock controls. TARP was knocked down in 4 TARP-high AML cell lines (HL-60, Kg-1a, MV4;11 and THP-1) using three different shRNA, next to mock controls.

TARP-TCR engineered CTLs were generated by transduction with lentiviral (LV) or retroviral (RV) particles encoding a TCRA8-T2A-TCRB12 sequence directed against the HLA-A*0201-restricted synthetic TARP peptide TARP(P5L)$_{4–13}$. Regarding RV transduced TARP-TCR CTLs, mock CTLs were used to correct for non-TARP mediated lysis, and CMV-TCR transduced CTLs to evaluate aspecific killing.

		pediatric AML (n=13)		adult AML (n=17)	
		Median (Range)		Median (Range)	
Age, years		10 (2-16)		48 (20 - 76)	
WBC count, x 10^9/L		66 (2.7-336)		15 (6-274)†	
Morphological blast count					
BM, %		81 (34-96)		77 (28-90)†	
PB, %		67 (1-95)		73 (7-93)‖	
		N	%	N	%
Gender					
F		7	53.8%	9	52.9%
M		6	46.2%	8	47.1%
Sample					
BM		8	61.5%	11	64.7%
PB		5	38.5%	6	35.3%
CD34 positivity		13	100.0%	15	88.2%
Fusion transcript		6	46.2%	3*	18.8%
CBF leukemia		4	30.8%	2	11.8%
WT1 overexpression		10	76.9%	10‡	71.4%
Mutation status	NPM1	0	0.0%	5‡	35.7%
	FLT3-ITD	8	61.5%	9†	60.0%
FAB classification	M0	1	7.7%	0	0.0%
	M1	1	7.7%	6	35.3%
	M2	4	30.8%	6	35.3%
	M3	1	7.7%	2	11.8%
	M4	3	23.1%	0	0.0%
	M5	2	15.4%	0	0.0%
	M6	0	0.0%	0	0.0%
	M7	1	7.7%	0	0.0%
	ND	0	0.0%	3	17.6%

Table 1. Characteristics of de novo AML patients used for sorting CD34+CD38+ and CD34+CD38- cell fractions and qPCR evaluation. PedAML patients were diagnosed in Belgium and treated according the DB AML-01 (n=9, 69%) or NOPHO-DBH AML 2012 (n=4, 31%) protocol. Pediatric patients were classified as standard risk (SR=7) or high risk (HR=5) according to a previously published risk stratification (data lacking for one patient). Adult AML samples were from patients treated at the Ghent University Hospital, Ghent, Belgium (n=12, 71%) or VUmc, Amsterdam, the Netherlands (n=5, 29%). Belgian patients were treated according to local and international guidelines, whereas Dutch patients were included in the HOVON 102 (n=3) or HOVON 132 (n=2) study. Adult AML patients were categorized into favourable (n=4), intermediate-I/II (n=7) or adverse (n=3) prognostic risk groups according to the ELN 2010 guidelines[1] (data lacking for three Dutch patients). *WT1* overexpression was interpreted in regard to in-house or published (Cilloni et al. 2009) cut-offs. CBF-positive leukemias comprised AML with t(8;21)(q22;q22) (pedAML=3) and inv(16)(p13q22) (pedAML=1, adult AML=2). Other fusion transcripts detected were DEK-NUP214 (pedAML=1) and PML-RARA (pedAML=1, adult AML=1). Superscripts indicate one (*), two (†), three (‡) or five (‖) missing data. BM indicates bone marrow; CBF, core-binding factor; F, female; FAB, French American British; M, male; NPM1, nucleophosmin; PB, peripheral blood; WBC, white blood cell; *WT1*, Wilms' tumor 1.

Results

Discovery of TARP transcript expression in AML

In order to identify LSC-specific antigens, we re-analyzed the GSE 17054 micro-array dataset from Majeti et al. [31], which included gene expression profiles of CD34+CD38- sorted fractions of four healthy adults (HSC) and nine adult AML patients (LSC). *TARP* ranked first amongst the top differentially expressed genes, with all four probes in the top 20 (range log2-FC 5.13-6.92), showing a significantly higher expression in LSC compared to HSC (P<0.01, Fig. S1). *TARP* was previously identified as a truncated TCRγ transcript expressed in androgen-sensitive prostate and breast adenocarcinoma (Fig. S2)[41, 42]. We further explored *TARP* expression in pedAML by micro-array profiling CD34+CD38+ (n=4, leukemic blast) and CD34+CD38- (n=3, LSC) sorted cell populations from four pedAML patients (2 *FLT3*-ITD and 2 *FLT3* WT, Table S1). In addition, sorted CD34+CD38+ (n=3) and CD34+CD38- (n=2) cells from CB were profiled to examine the expression in their normal counterparts (Fig. S3). *TARP* appeared to be higher expressed in leukemic blasts and LSCs from *FLT3*-ITD patients compared to *FLT3* WT patients and CB (Fig. 1A). This finding suggested that *TARP* might represent a LSC-associated target within HR pedAML patients harbouring *FLT3*-ITD.

To validate these data in a larger patient group, we sorted CD34+CD38+ and CD34+CD38- cell populations from 9 additional pedAML (resulting in 13 pedAML patients), 17 adult AML (Table 1) and 15 control samples consisting out of 7 CB, 6 NBM and 2 mPBSC. qPCR analysis using *TARP* short primers (Table S3, Fig. S2) showed that *TARP* transcripts were consistently low in HSCs and myeloblasts sorted from CB, NBM and mPBSC (Fig. 1B), although blasts from NBM showed a marginally higher expression compared to CB (mean CNRQ 0.12 vs. 0.045, P=0.049). In sharp contrast, LSCs and blasts from pediatric and adult AML showed significantly (P<0.01) higher expressions compared to their normal counterparts. Paired comparison between LSCs and blasts on a per patient basis showed no significant differences (Fig. 1C).

A cut-off for elevated *TARP* expression was determined based on the highest expression in control fractions plus two times the standard deviation. Classification of patients into *TARP*-high (8 pedAML, 13 adult AML) and *TARP*-low (5 pedAML, 4 adult AML) revealed that *FLT3*-ITD (P<0.001), CNS involvement and HR profile (P<0.05) were exclusively present in *TARP*-high pedAML patients (Fig. 1D). *TARP* expression was shown to be significantly higher in sorted LSCs (P<0.01) and blasts (P<0.0001) from *FLT3*-ITD compared to *FLT3* WT pedAML (Fig. 1E). In adult AML, high *TARP* expression was not restricted to *FLT3*-ITD. On the other hand, all pediatric (Fig. 1D) and adult (Fig. S4A) core-binding factor (CBF) leukemia were classified as *TARP*-low patients (P<0.01). *TARP*-low pedAML patients were included in the standard risk (SR) groups (P<0.05). No significant differences in age, WBC count, or blast percentages were observed between TARP-high and -low pediatric or adult AML patients (Fig. S4 B-C). We thus conclude that *TARP* is highly and specifically expressed in AML leukemic cells from both adults and children, showing a significant association with *FLT3*-ITD in pedAML.

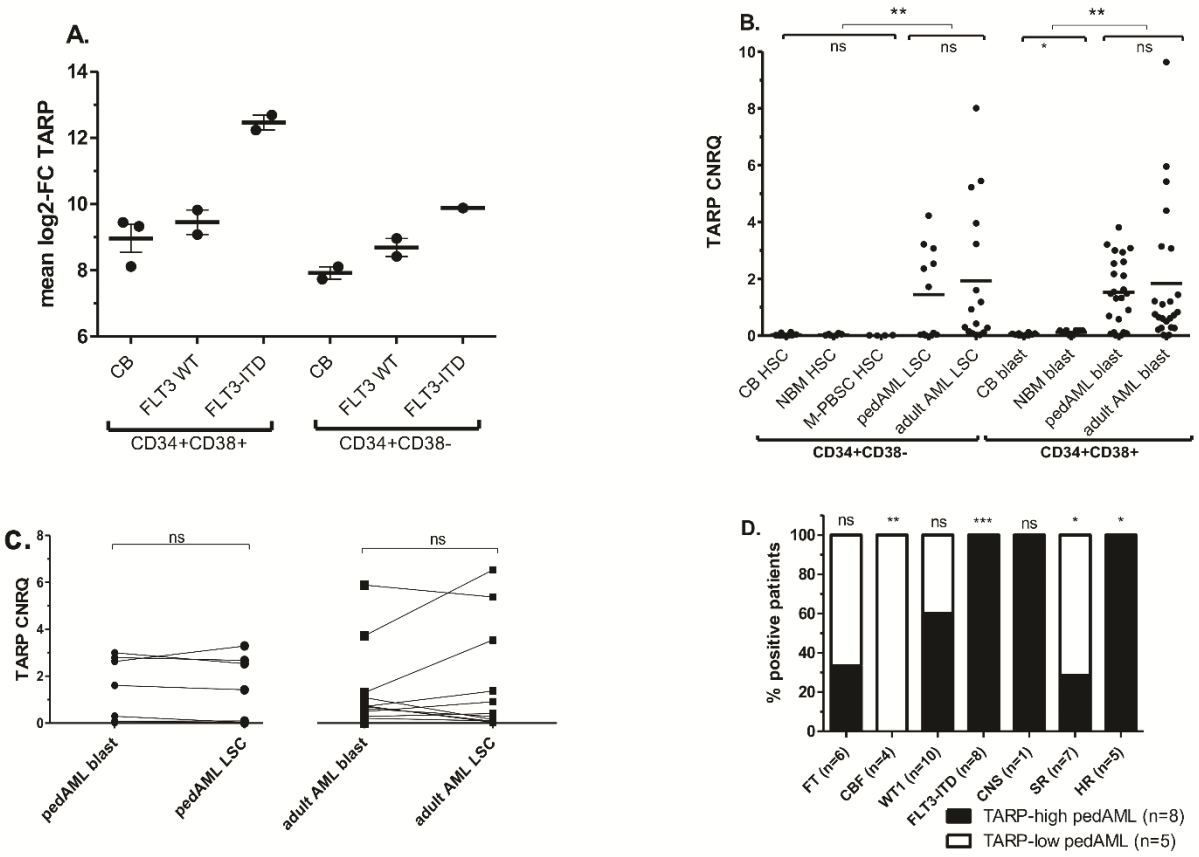
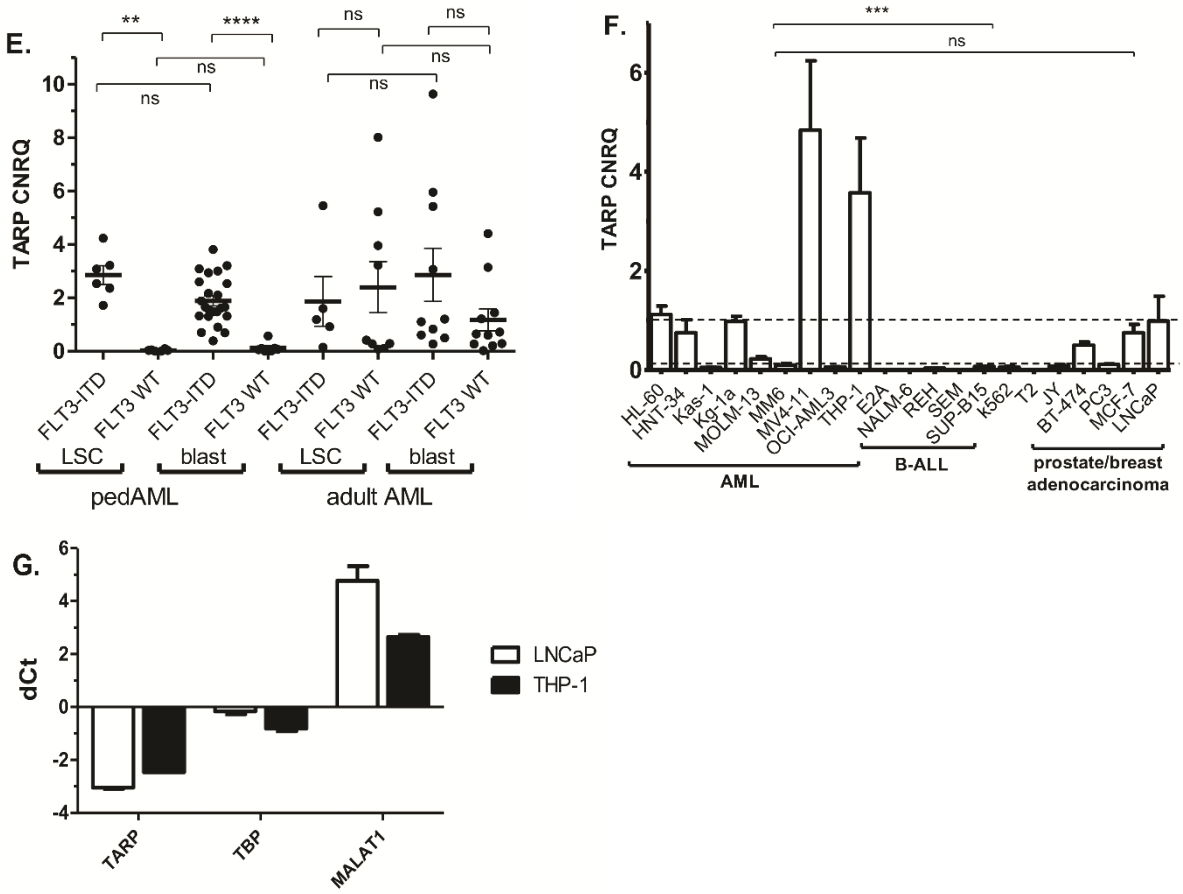


Fig. 1. *TARP* transcript expression in pedAML and adult AML leukemic cells and cell lines. For *TARP* qPCR, CNRQ values were calculated using LNCaP (prostate adenocarcinoma cell line) as interrun calibrator. Biological replicates, e.g. cells sorted from the same patient in different runs and independent cDNA syntheses, were depicted as independent data points. Horizontal bars indicate means and error bars indicate ±SEM. Horizontal square brackets represent statistical comparisons, and one, two, three or four asterisks are indicative for the level of significance (P<0.05, P<0.01, P<0.001 and P<0.0001, respectively).
(A) *TARP* expression was determined in CD34+CD38+ (n=4) and CD34+CD38- (n=3) cell fractions from four pedAML patients (2 FLT3-ITD, 2 FLT3 WT; Table S1) by micro-array profiling. Sorted CD34+CD38+ (n=3) and CD34+CD38- (n=2) cells from CB were used as control populations. Mean log2-FC values (y-axis) were calculated based on both *TARP* probes included in the array, the x-axis represents the different sample groups. (B) *TARP* expression was significantly higher in CD34+CD38- and CD34+CD38+ cell fractions from AML patients (13 pedAML and 17 adult AML) compared to healthy controls (7 CB, 6 NBM and 2 mPBSC) (P<0.01, Mann Whitney U test). Blasts from NBM showed a marginally higher expression compared to CB (P=0.049). (C) Comparison of TARP expression between LSCs and blasts within pedAML (circles, n=10) and adult AML (squares, n=12) on a per patient basis showed no significant differences (P>0.05, Paired sample T-test). (D) Bars display the percentage of patients (%), harbouring the characteristic shown in the x-axis (dichotomous variables, details shown in Table 1), for *TARP*-high (black, n=8) and *TARP*-low (white, n=5) pedAML patients. The total number of patients positive for each characteristic is shown between parentheses. Patients without central nervous system (CNS) involvement all showed negative lumbar punctures. Data on CNS involvement and risk profile is lacking for one patient. The number of patients harbouring *FLT3*-ITD (P<0.001) and HR profiles (P<0.05) were significantly higher in the *TARP*-high group, whereas *TARP*-low pedAML patients included significantly more CBF-leukemia (P<0.01)

Next, we evaluated *TARP* transcript levels in cell lines of various origin. Expression in breast and prostate adenocarcinoma (PC3, BT-474, LNCaP and MCF-7) was in agreement with previous findings[42] (Fig. 1F). No expression was detected in five B-ALL cell lines, CML cell line K562, EBV-immortalized B-cell line JY and T2 cell line. Expression in AML cell lines, on the other hand, was significantly increased (P<0.001, One-way ANOVA). The highest expression was observed in HL-60, HNT-34, Kg-1a, MV4;11 and THP-1 (median CNRQ 1.12, range 0.75-4.84), whereas low transcript levels were observed in Kas-1, MOLM-13, MONO-MAC6 and OCI-AML3 (median CNRQ 0.080, range 0.049-0.22). Furthermore, fractionation revealed a mainly cytoplasmic localization of *TARP* mRNA in THP-1 (Fig. 1G), as previously shown in LNCaP [43].

To evaluate whether the *TARP* transcript detected in AML is identical to previous reports, we sequenced the *TRGC* region of different *TARP* amplicons obtained by qPCR for AML cell lines and pedAML leukemic cells. Using *TARP* long primers, we observed a single band for Kg-1a, which was similar to the LNCaP and *TRGC1* reference sequence (Fig. S5A). Unexpectedly, three fragments were observed in the sorted blasts and LSCs from *TARP*-high pedAML patients and the MV4;11 cell line. Cloning and sequencing of each fragment (Fig. S5B) revealed that the largest fragments were artificial heteroduplexes [44], whereas the smallest fragments were identical to the fragments from Kg-1a and LNCaP. Middle sized fragments were consistently 48 bp longer, and showed the same size as the HSB-2 amplicon, a T-ALL cell line with functional *TRGC2* rearrangements [45]. As *TRGC2* contains a duplicated second exon (48 bp) compared to *TRGC1* [45] (Fig. S2), we hypothesized that an alternative *TARP* transcript might exist in AML. Indeed, most AML cell lines, but none of the prostate and breast adenocarcinoma cell lines, showed *TRGC1* as well as *TRGC2* amplicons (Fig. S5 C-E). Single bands for exon 3 and exon 1 amplicons in all cell lines provided evidence that the occurrence of the second transcript is related to the *TRGC2* duplicated second exon. Altogether, *TARP* was highly expressed in about half of the AML cell lines evaluated, and both *TRGC1*- and *TRGC2*-related transcripts co-exist.



(legend Fig.1. continued) (E) Differential TARP expression between LSCs and blasts sorted from pediatric and adult AML patients with FLT3-ITD versus FLT3 WT. Only for FLT3-ITD positive pedAML patients, a significant higher TARP expression was detected in LSCs (P<0.01) and blasts (P<0.0001) (Mann Whitney U test). (F) TARP expression in nine AML cell lines, five B-ALL cell lines, the CML cell line K562, the EBV-immortalized B-cell line JY and T2 cell line, next to two breast (BT-474, MCF-7) and two prostate (LNCaP, PC3) adenocarcinoma cell lines. Dashed lines indicate the expression observed in PC3 and LNCaP, serving as low and high reference, respectively, in agreement with previous literature[41]. (G) Delta (d) Ct values were calculated for TARP, *MALAT1* and *TBP* between cytoplasmic and nuclear compartments of THP-1 and LNCaP, in order to examine the subcellular location of TARP. THP-1 showed a cytoplasmic residence for TARP, in agreement with LNCaP. FC indicates fold change; FT, fusion transcript; Kas-1, Kasumi-1; MM-6, MONO-MAC6; SEM, standard error of the mean.

TARP protein is expressed in AML cell lines and patient leukemic cells

We generated TARP transgenic cell lines in order to optimize Western blot experiments and evaluate TARP protein expression in AML. THP-1 and OCI-AML3 OE cell lines showed a significant higher *TARP* transcript expression (P<0.01) compared to mock controls (Fig. S6A). Western blotting confirmed presence of TARP and GFP proteins in both OE cell lines, with a size around 20 kDa and 27 kDa, respectively (Fig. 2A). Concordantly, the OCI-AML3 mock cell line, negative for TARP, only showed a 27 kDa GFP protein. WT AML cell lines HL-60, MV4;11, THP-1 and MOLM-13, as well as LNCaP, also showed a 20 kDa TARP protein, with expression corresponding to the transcript levels (Fig. 2B). TARP knockdown (KD) cell lines were generated for HL-60, Kg-1a, MV4;11 and THP-1 using TARP-targeting shRNA, next to mock controls. Transcript levels were efficiently downregulated (Fig. S9), and KD cell lines for HL-60, MV4;11 and THP-1 showing the highest transcript downregulation were selected for Western blotting (Fig. 2C). Protein levels were efficiently downregulated in HL-60 transduced with shRNA 3 (19.4% compared to mock). This downregulation was less clear in MV4;11 and THP-1 (116% (shRNA 3) and 108% (shRNA 3)/63% (shRNA 2), respectively).

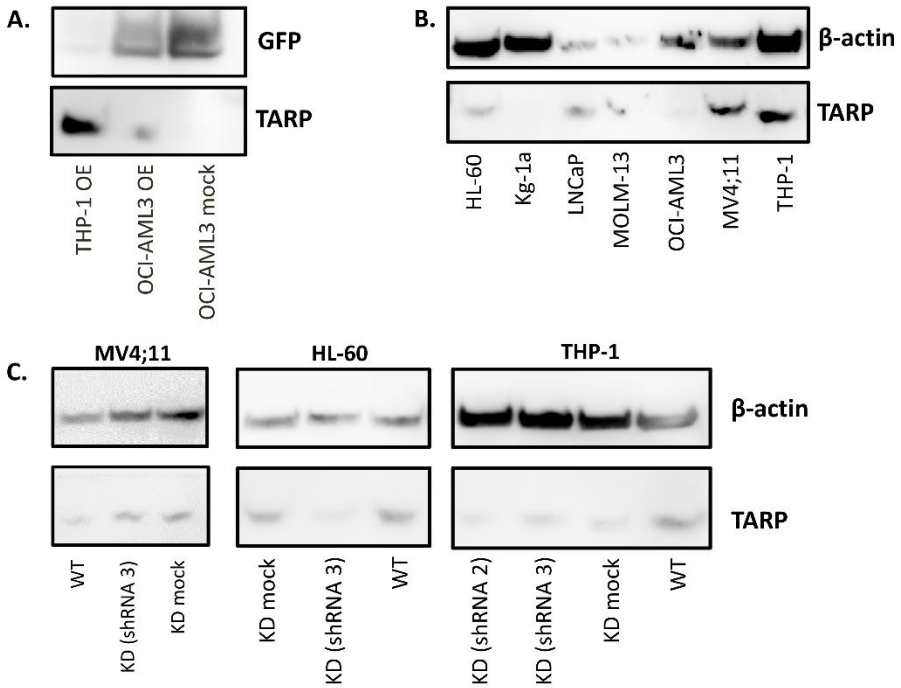


Fig. 2. TARP protein expression in cell lines evaluated by Western blotting. Whole-blot images with ladders used for size estimation are shown in Supplementary data (Fig. S7). (A) TARP transgenic (OE) cell lines generated for OCI-AML3 and THP-1 showed a 27 kDa protein for GFP and a 15-25 kDa protein for TARP. In agreement with low TARP transcript levels, the OCI-AML3 mock cell line only showed a 27 kDa GFP protein. TARP expression in THP-1 OE was higher than OCI-AML3 OE, most likely resulting from both transgenic and cognate TARP protein expression, since THP-1 was categorized by qPCR as a TARP-high AML cell line. (B) Immunoblotting of TARP and β-actin in AML cell lines (HL-60, Kg-1a, MOLM-13, OCI-AML3, MV4;11 and

THP-1) next to LNCaP. Protein expression mostly matched transcript levels, except for Kg-1a, although confocal microscopy did allow TARP protein staining in Kg-1a. β-actin expression appeared to be lower for LNCaP and MOLM-13, although equal amounts of protein were loaded. (C) Immunoblotting of TARP and β-actin in selected shRNA-mediated knockdown (KD) AML cell lines for MV4;11, HL-60 and THP-1, next to their respective mock and WT cell line. For HL-60, a stable knockdown was introduced by shRNA 3 (19.4% compared to mock). β-actin indicates beta-actin; KD, knockdown; OE, overexpression.

To confirm Western blot data and determine the subcellular location of TARP, confocal microscopy was performed using TARP antibodies combined with mitochondrial (HSP-60) and endoplasmic reticulum (ER, calnexin) staining. The overexpressing OCI-AML3 and THP-1 cell lines (Fig. S6 B-C) and TARP-high WT AML cell lines showed a perinuclear membranous-type TARP staining pattern (Kg-1a (Fig. 3), HL-60, MV4;11 and THP-1 (Fig. S8)). This finding was in contrast to the barrel-shaped TARP pattern with mitochondrial co-localization reported in LNCaP [43]. Co-localization with calnexin, presenting as a speckled pattern throughout the ER, was more abundant in some cell lines, e.g. Kg-1a, showing TARP enrichment at the cells' protrusions. TARP-low cell lines concordantly showed weak or negative TARP protein staining (Fig. S8). Importantly, the leukemic cells sorted from a TARP-high and TARP-low pedAML patient also illustrated differential TARP protein expression in agreement with the transcript levels, again showing limited mitochondrial overlap (Fig. 3).

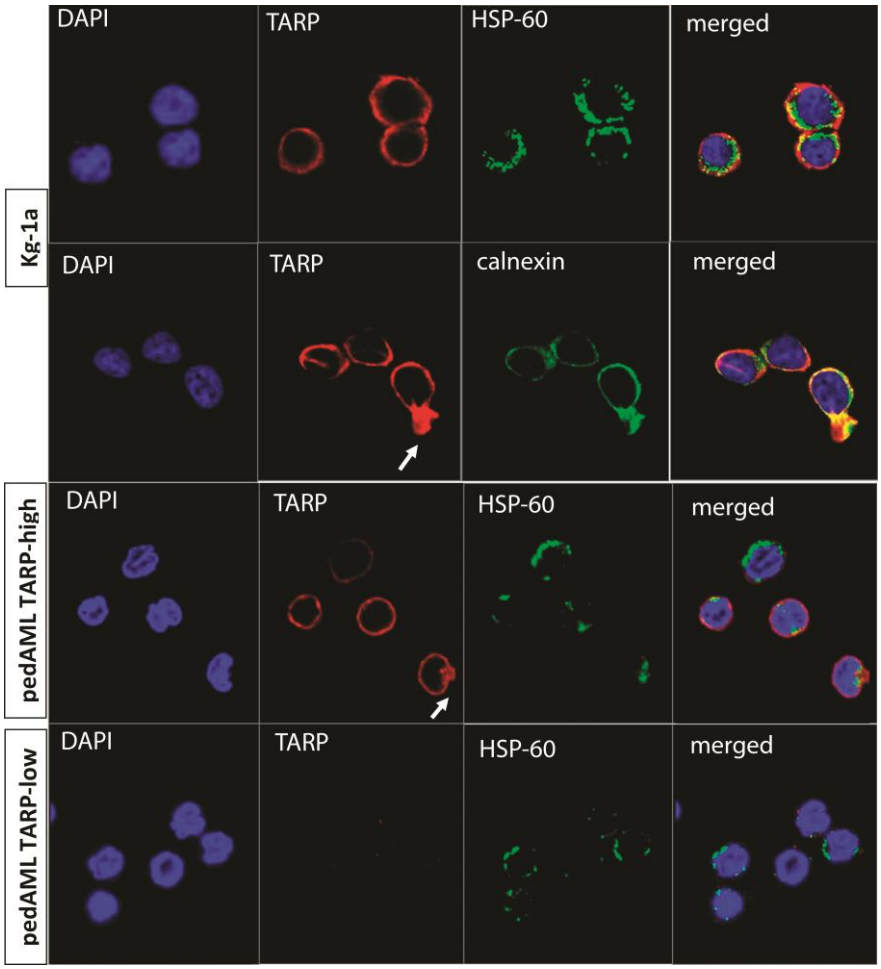


Fig. 3. TARP protein detection in Kg-1a and patient leukemic cells by confocal microscopy. Merged patterns visualize TARP (red) and HSP-60 (top lane) or calnexin (bottom lane) (both in green) co-localization (yellow fusion signals) together with DAPI nuclear counterstaining (blue). Leukemic cells were sorted from two pedAML patients, classified as TARP-high and TARP-low by qPCR. Calnexin staining was not performed on primary cells due to lack of material. Within Kg-1a and the sorted TARP-high leukemic cells, TARP expression was enriched at the cells' protrusions, indicated by arrows.

TARP-TCR transgenic CTLs display specific anti-leukemic activity

To explore if TARP might represent an immunotherapeutic target in AML, we evaluated the cytokine and cytotoxicity response of TARP-TCR transgenic CTLs, encoding a previously developed TCRA8-T2A-TCRB12 sequence targeting the HLA-A2 enhanced affinity TARP(P5L)$_{4\text{-}13}$ epitope[46, 47]. As concomitant HLA-A*0201 and TARP expression is required to trigger TCR-mediated killing, HLA-A*0201 transgenic cell lines were generated for 3 WT cell lines (HL-60, Kg-1a and MOLM-13) and 3 Luc-positive cell lines (HL-60-Luc, MOLM-13-Luc, MV4;11-Luc).

First, target specificity of the TARP-TCR was examined in a non-competitive environment using T2 cells (endogenous HLA-A*0201+) pulsed with exogenous peptides (Table S3). As expected, we found stronger cytokine responses (Fig. S10A) and higher killing rates (Fig. S10B-C) towards the TARP(P5L)$_{4\text{-}13}$ than to the cognate TARP$_{4\text{-}13}$ peptide for both RV and LV transduced CTLs, with LV TARP-TCR CTLs reacting stronger, in general. T2 cells pulsed with non-TARP related peptides (INF, CMV) were not affected, although CMV-pulsed T2 cells were efficiently recognized by CMV-TCR CTLs, indicating a high specificity of the TARP-TCR.

Second, we explored the immunogenicity of cell lines with endogenous HLA-A*0201 presentation. Exposure to LNCaP and THP-1 appeared to be insufficient to trigger cytokine release for both LV and RV transduced TARP-TCR CTLs (Fig. 4A). Using a chromium^{51} release assay, we observed a lytic response by LV transduced TARP-TCR CTLs starting from effector to target ratio (E/T) 10/1, with a maximal average response at 50/1 (LNCaP 10%, THP-1 24%), whereas RV transduced TARP-TCR CTLs performed best at 10/1 (THP-1 12%) (Fig. 4B). The TARP-low AML cell line OCI-AML3 remained unaffected at all conditions. Altogether, as the TARP-TCR targets the enhanced HLA-A2 binding peptide, we observed weaker responses against endogenous TARP expressing cell lines compared to pulsed T2 cells.

Third, lysis of TARP-high HLA-A*0201-negative cell lines was evaluated versus their HLA-A*0201 transgenic counterparts in a 48-h FCM-cytotoxicity assay. In addition, killing of TARP transgenic or TARP-pulsed HLA-A*0201-positive cell lines was compared to the respective TARP-low WT cell line (Fig. 4C). A non-TARP mediated lysis by LV TARP-TCR CTLs of maximal 20% was observed (indicated by dashed line). Stable transduction of HLA-A*0201 increased killing for Kg-1a compared to the WT cell line (29% vs. 13%), whereas killing of MOLM-13, with lower TARP expression levels, remained unaffected when HLA-A*0201 was introduced. Transgenic TARP OE and TARP(P5L)$_{4\text{-}13}$ pulsed OCI-AML3 cells were prone to a higher lysis than the WT cell line (44% and 55% vs. 24%). Killing of TARP OE/pulsed THP-1 cells was only marginally upregulated, most likely due to an already high endogenous expression. These data were confirmed using RV TARP-TCR CTLs, and corrected for non-TARP mediated lysis using mock CTLs. HLA-A*0201 expression again increased killing of Kg-1a (46% vs. -4%) and HL-60 (40% vs. 15%) compared to the WT cell line. Upregulated killing of transgenic TARP OE THP-1 cells was again limited. For OCI-AML3, lysis was upregulated after pulsing, but remained low for the TARP OE transgenic cell line. Killing by LV TARP-TCR CTLs was additionally evaluated in a bioluminescence imaging (BLI)-based assay using Luc-positive AML cell lines with high TARP expression (HL-60 and MV4;11) and low TARP expression (MOLM-13 and OCI-AML3) (Fig. 4D). A higher lysis was observed for HL-60-Luc and MV4;11-Luc when expressing HLA-A*0201 at 48 h and 56 h, indicating that also in long-term cytotoxicity experiments HLA-A*0201 restricted TARP specific killing could be detected.

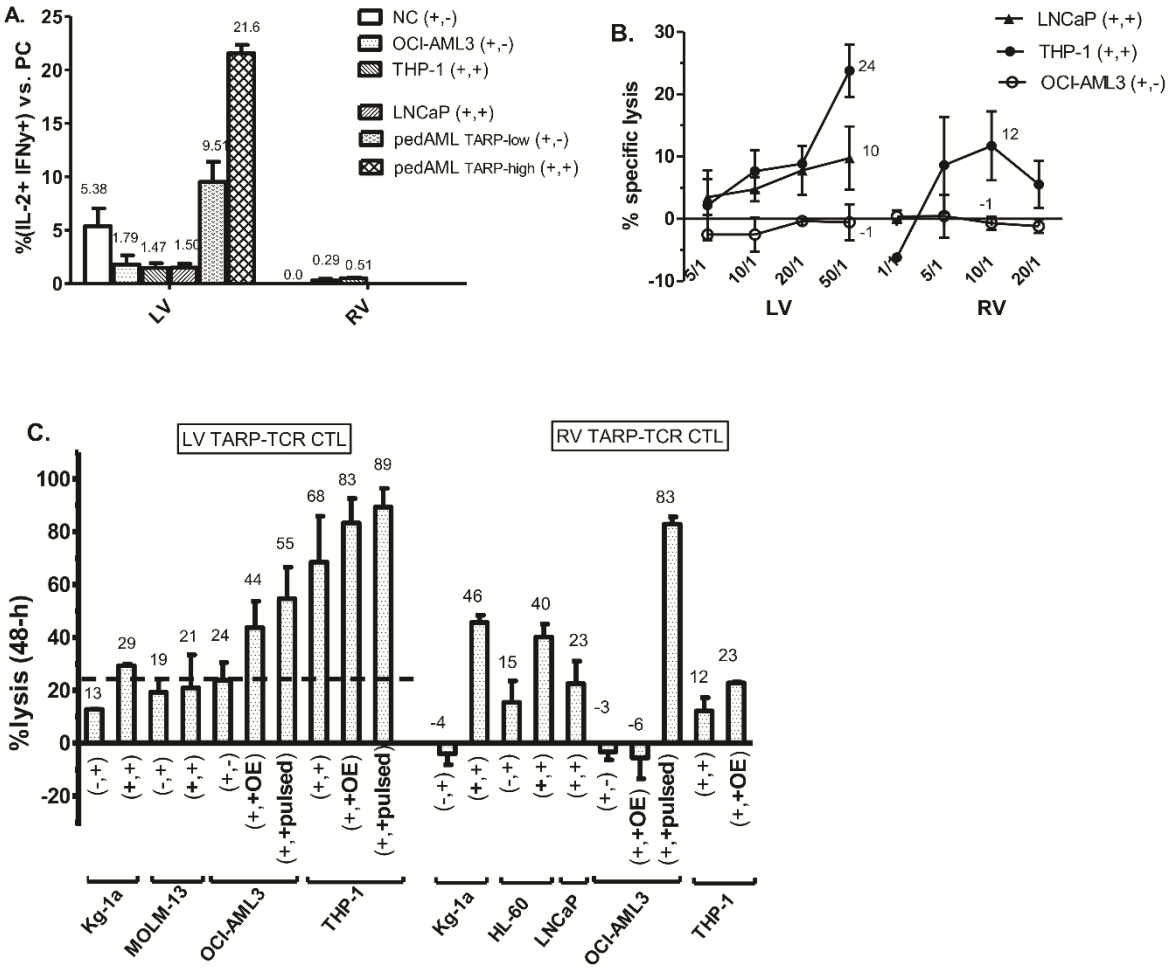
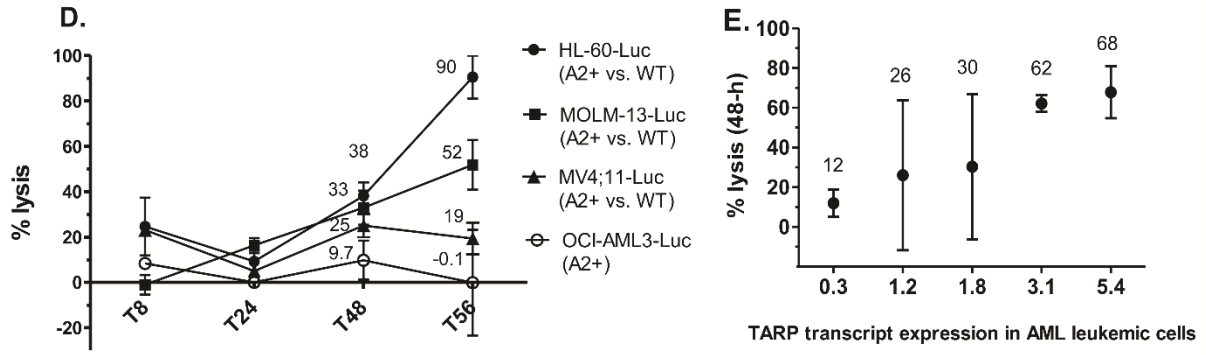


Figure 4. Functional evaluation of TCR-transgenic CTLs towards cognate and modified cell lines and patient leukemic cells. (A) Cytokine response (IFN-γ/IL-2 expression within the CD3+/CD8+ compartment) by co-incubation (1 h) with OCI-AML3 and THP-1 was evaluated by both LV and RV TARP-TCR CTLs. LNCaP and patient leukemic cells (single experiment) were only evaluated by LV transduced TARP-TCR CTLs. For each target, positive (+) or negative (-) HLA-A*0201 and TARP expression, in this respective order, is indicated between brackets. HLA-A*0201 and TARP co-expressing cell lines (LNCaP and THP-1) were unable to trigger higher cytokine release than OCI-AML3 with low TARP expression. Leukemic cells from a TARP-high pedAML patient triggered a twofold higher cytokine release compared to a TARP-low pedAML patient. (B) Lytic response of LV and RV TARP-TCR CTLs versus HLA-A*0201-positive TARP-high (black symbols) and TARP-low (white symbols) targets, measured by a chromium^{51} release assay after 4 h. Highest lysis of TARP-high cell lines was observed at E/T ratio 50/1 for LV and 10/1 for RV TARP-TCR CTLs (percentages indicated between brackets), whereas OCI-AML3 (HLA-A*0201+, TARP-) remained unaffected. (C) Lytic response of LV and RV TARP-TCR CTLs versus towards WT, transgenic and pulsed AML cell lines, measured by a 48-h FCM-based cytotoxicity assay. The dashed line indicates the highest level of non-TARP mediated background killing observed for LV TARP-TCR CTLs, as no mock CTLs could be constructed. Positive (+) or negative (-) expression for HLA-A*0201 and TARP is shown, in this respective order, between brackets. Bold symbols indicate the expression differing from wild-type, either by retroviral transduction or pulsing. HLA-A*0201 transgenic AML cell lines were more efficiently lysed compared to their HLA-A*0201-negative counterparts (Kg-1a, MOLM-13, HL-60). Higher lysis was observed for transgenic TARP OE or peptide-pulsed cell lines compared to the wild-type cell line (OCI-AML3, THP-1), except for killing of TARP OE OCI-AML3 cell line by RV TARP-TCR CTLs.

Finally, we explored the targetability of primary leukemic cells by LV TARP-TCR CTLs. Co-incubation with blasts sorted from a TARP-high pedAML patient resulted into a twofold higher IFN-γ and IL-2 production compared to a TARP-low pedAML patient (22% vs. 10%) (Fig. 4A). Moreover, TARP-TCR CTLs were also capable of killing leukemic cells from *de novo* adult AML patients (n=5) (Fig. 4E). Lysis ranged between 12-68% and borderline correlated to TARP transcript levels (Spearman's coefficient 0.82, P=0.089).



(legend Fig.4. continued) (D) Lysis by LV TARP-TCR CTLs, measured at different time points (8h, 24h, 48h and 56h, as indicated on x-axis), based on the luminescence release by transgenic HLA-A*0201-expressing TARP-high AML cell lines in respect to the HLA-A*0201 WT cell line (HL-60-Luc, MOLM-13-Luc and MV4;11-Luc: black symbols). In addition, lysis of the TARP-low, cognate HLA-A*0201-positive OCI-AML3 cell line was evaluated (white symbols). Mean lysis (%) observed after 48 h is indicated next to whiskers, representing the ±SEM. (E) 48-h FCM-based cytotoxicity assay evaluating lysis of primary leukemic cells (adult AML=5, all FLT3-ITD mutated) by LV TARP-TCR transduced CTLs (biological duplicates). TARP transcript expression (CNRQ) is shown in the x-axis for each target.
CTL indicates cytotoxic T-cells; IFN-γ, interferon gamma; IL-2, interleukin-2; INF, influenza; LV, lentiviral; RV, retroviral.

Discussion

We demonstrated increased *TARP* expression in AML LSCs (CD34+CD38-) and blasts (CD34+CD38+) from primary patients compared to their normal counterparts as well as AML cell lines. We also showed that TARP proteins are expressed in primary AML leukemic cells and are adequately presented on HLA molecules, which makes the cells targetable for immunotherapy.

TARP expression has only been investigated in prostate tissue and androgen-sensitive prostate adenocarcinoma and breast adenocarcinoma [42, 43, 48], next to a single report on salivary adenoid cystic carcinoma[49]. We found that *TARP* was significantly (P<0.001) higher expressed in *FLT3*-ITD compared to *FLT3* WT pedAML patients at diagnosis, whereas no significant difference was observed in adult AML. Importantly, the genomic landscape in adult and pediatric AML has been shown to be highly different [50, 51], potentially explaining part of the differential associations observed in our cohorts. The association between *TARP* expression and a bad prognosis is in agreement with a recent report, investigating the association between transcript expression and clinical outcome in pedAML, ranking *TARP* within the top significantly genes associated with a detrimental outcome [52]. To shed light on the link between *FLT3*-ITD and *TARP*, mRNA sequencing of the transgenic OE and KD cell lines compared to their wild-type cell line is ongoing. As it was recently shown that the *FLT3*-ITD regions encode immunogenic, HLA-presented neo-epitopes [53], the benefit of CTL therapy targeting both leukemogenic molecules in pedAML could be of high interest. On the other hand, CBF leukemias, representing a favourable cytogenetic subgroup [2, 8], were exclusively present (P<0.01) in the *TARP*-low group for both pediatric and adult patients. AML cell lines derived from pediatric cases (MV4;11, THP-1) and LSC-enriched cell lines (Kg-1a, HNT-34), showed the highest *TARP* levels, confirming a relation between TARP, the LSC compartment and pedAML, although also HL-60 (adult, CD34-) showed high expression. Whether *TARP* remains differentially expressed within LSCs outside the predominant CD34+CD38- compartment, as within CD34- AML [15, 54], needs to be explored further. In addition, we showed that transcripts differ from these in solid tumors and are derived from both the *TRGC1* and *TRGC2* coding regions. Sequencing analysis indicated the presence of a second, AML-exclusive, *TARP* transcript encoding *TRGC2* instead of *TRGC1*.

TARP protein expression was in agreement with transcript levels, showing a 15-25 kDa fragment in AML cell lines. In breast and prostate adenocarcinoma, TARP was previously defined as a 7 kDa protein [42, 48], although also a 9 kDa fragment was reported in MCF-7 [42, 48]. Fritzsche et al. detected protein sizes in prostate carcinoma of 20-25 kDa [55], comparable to our findings, whereas Yue et al. reported a 15 kDa protein [49]. Next to its size, the subcellular localization of TARP in AML needs to be refined. qPCR analysis revealed cytoplasmic localization, and confocal microscopy showed sporadic ER overlap, in contrast to previous reported mitochondrial co-localization[43]. We observed an enrichment of TARP at the cells' protrusions in Kg-1a and sorted leukemic cells. Protrusions are kinetic cytoskeletal abnormalities formed during chemokine-induced cell migration, e.g. homing of CD34+ HSCs towards the bone marrow niche [56]. The presence of molecular abnormalities in CD34+ progenitor cells was shown to increase protrusion formation [57]. Indeed, LSCs were reported to duel with HSCs for endosteal niche engraftment, where they are protected from chemotherapy-induced apoptosis [12, 58]. Whether TARP interferes in homing and chemoprotection of leukemic AML cells in the BM microenvironment needs to be elucidated. Although protein expression was readily upregulated in TARP transgenic cell lines, shRNA-mediated knockdown appeared to be less efficient. Possible explanations are the presence of escape mechanisms and alternative translation pathways during silencing or a very high stability of the TARP protein, persisting in the cell for a long period of time.

To explore TARP as an immunotherapeutic target in AML, we evaluated the cytokine release and cytotoxic killing capacities of TARP-TCR transgenic CTLs *in vitro*. TARP and HLA-A*0201 co-expressing cell lines were efficiently lysed, and although evaluated on a limited number of patients (n=5), TARP-TCR CTLs were able to kill primary leukemic cells with a borderline correlation to the TARP transcript expression. Noteworthy, weaker responses were observed for the cognate $TARP_{4\text{-}13}$ peptide, since the TCR is directed against the HLA-A*0201 enhanced affinity $TARP(P5L)_{4\text{-}13}$ peptide. Moreover, pulsed T2 cells appeared to be more susceptible than AML cells. This finding is in concert with previous data [47, 59], and several reasons may account for this phenomenon. First, peptide processing, transport and/or MHC-I presentation may be disturbed in leukemic cells [60]. Second, high and stable HLA-A*0201 expression is vital for triggering lytic responses, and transgenic expression might diminish during culture. Therefore, we cannot exclude that HLA-A*0201-mediated TARP presentation within the transgenic OCI-AML3 cell line had diminished during long-term culture. Third, competition between transgenic and endogenous MHC-I molecules might block HLA-A*0201-guided peptide presentation. Indeed, the TARP-TCR was shown to suffer from low MHC-I avidity compared to foreign epitope-directed TCRs[61]. Cloning the $TARP_{4\text{-}13}$-TCR sequence into a retroviral construct enabled higher transduction efficiencies and the generation of mock CTLs to correct non-TARP mediated lysis, which are lacking in previous reports[37, 49]. As promoters driving TCR expression differed between constructs, and functional activity is known to correlate with TCR cell-surface expression [62], different killing rates between LV and RV transduced CTL were not surprising. In addition, intrinsic reactivity, HLA status and endogenous TCR repertoire of each donor as such might have an impact [62]. In addition, comparing reactivity by effectors from an allogeneic versus autologous setting will be implemented in future experiments.

In conclusion, we showed that *TARP* is highly expressed in AML leukemic cells, including the CD34+CD38- LSC compartment, while absent in normal counterparts. Moreover, *TARP* expression was associated with FLT3-ITD in pedAML. We provide *in vitro* evidence that TARP-directed CTLs effectively kill TARP and HLA-A*0201 co-expressing cell lines and primary leukemic cells, and thus hold great promise for immunotherapeutic T-cell therapy.

Acknowledgements

Our gratitude goes to dr. F. Plasschaert, the staff of the Department of adult Haematology and Pediatric Haematology and Oncology of the Ghent University Hospital (Ghent, Belgium), and C. Matthys of the Cord Blood Bank, for providing samples. The authors thank all patients and their parents for their participation in the study, as well as the data managers involved in the clinical trials. We are indebted to S. Vermaut for cell sorting and all technicians of the Childhood Cancer Foundation and Laboratory of the Ghent University Hospital (Ghent, Belgium). We thank our collaborators from the LL Biology Working Group for their relevant contributions, in particular prof. dr. GJ Kaspers for taking interest in our research. This research was supported by the Belgian Foundation against Cancer (grant 2014–265), FOD-KankerPlan (Actie29, grant to JP) and vzw Kinderkankerfonds (grant to TL). The Research Foundation - Flanders (Fonds voor Wetenschappelijk Onderzoek Vlaanderen, FWO) supported TK (grant 1831312N) and BD (grants 1113117 and V433317N). This work is submitted in partial fulfilment of the requirement for the PhD of candidate BD at Ghent University.

References

1. Dohner H, Estey EH, Amadori S, Appelbaum FR, Buchner T, Burnett AK, Dombret H, Fenaux P, Grimwade D, Larson RA et al: Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. Blood 2010, 115(3):453-474.
2. Dohner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Buchner T, Dombret H, Ebert BL, Fenaux P, Larson RA et al: Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood 2017, 129(4):424-447.
3. De Kouchkovsky I, Abdul-Hay M: Acute myeloid leukemia: a comprehensive review and 2016 update. Blood cancer journal 2016, 6(7):e441.
4. Bonnet D, Dick JE: Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nature medicine 1997, 3(7):730-737.
5. Rasche M, Zimmermann M, Borschel L, Bourquin JP, Dworzak M, Klingebiel T, Lehrnbecher T, Creutzig U, Klusmann JH, Reinhardt D: Successes and challenges in the treatment of pediatric acute myeloid leukemia: a retrospective analysis of the AML-BFM trials from 1987 to 2012. Leukemia 2018, 32(10):2167-2177.
6. Creutzig U, van den Heuvel-Eibrink MM, Gibson B, Dworzak MN, Adachi S, de Bont E, Harbott J, Hasle H, Johnston D, Kinoshita A et al: Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. Blood 2012, 120(16):3187-3205.
7. von Neuhoff C, Reinhardt D, Sander A, Zimmermann M, Bradtke J, Betts DR, Zemanova Z, Stary J, Bourquin JP, Haas OA et al: Prognostic impact of specific chromosomal aberrations in a large group of pediatric patients with acute myeloid leukemia treated uniformly according to trial AML-BFM 98. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2010, 28(16):2682-2689.
8. De Moerloose B, Reedijk A, de Bock GH, Lammens T, de Haas V, Denys B, Dedeken L, van den Heuvel-Eibrink MM, Te Loo M, Uyttebroeck A et al: Response-guided chemotherapy for pediatric acute myeloid leukemia without hematopoietic stem cell transplantation in first complete remission: Results from protocol DB AML-01. Pediatric blood & cancer 2019:e27605.
9. Meshinchi S, Woods WG, Stirewalt DL, Sweetser DA, Buckley JD, Tjoa TK, Bernstein ID, Radich JP: Prevalence and prognostic significance of Flt3 internal tandem duplication in pediatric acute myeloid leukemia. Blood 2001, 97(1):89-94.
10. Hope KJ, Jin L, Dick JE: Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity. Nat Immunol 2004, 5(7):738-743.

11. Shlush LI, Mitchell A, Heisler L, Abelson S, Ng SWK, Trotman-Grant A, Medeiros JJF, Rao-Bhatia A, Jaciw-Zurakowsky I, Marke R et al: Tracing the origins of relapse in acute myeloid leukaemia to stem cells. Nature 2017, 547(7661):104-108.
12. Ishikawa F, Yoshida S, Saito Y, Hijikata A, Kitamura H, Tanaka S, Nakamura R, Tanaka T, Tomiyama H, Saito N et al: Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. Nat Biotechnol 2007, 25(11):1315-1321.
13. Thomas D, Majeti R: Biology and relevance of human acute myeloid leukemia stem cells. Blood 2017, 129(12):1577-1585.
14. Griessinger E, Anjos-Afonso F, Pizzitola I, Rouault-Pierre K, Vargaftig J, Taussig D, Gribben J, Lassailly F, Bonnet D: A niche-like culture system allowing the maintenance of primary human acute myeloid leukemia-initiating cells: a new tool to decipher their chemoresistance and self-renewal mechanisms. Stem cells translational medicine 2014, 3(4):520-529.
15. Ng SW, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, Arruda A, Popescu A, Gupta V, Schimmer AD et al: A 17-gene stemness score for rapid determination of risk in acute leukaemia. Nature 2016, Dec 15(540(7633)):433-437.
16. Terwijn M, Zeijlemaker W, Kelder A, Rutten AP, Snel AN, Scholten WJ, Pabst T, Verhoef G, Lowenberg B, Zweegman S et al: Leukemic Stem Cell Frequency: A Strong Biomarker for Clinical Outcome in Acute Myeloid Leukemia. PloS one 2014, 9(9).
17. van Rhenen A, Feller N, Kelder A, Westra AH, Rombouts E, Zweegman S, van der Pol MA, Waisfisz Q, Ossenkoppele GJ, Schuurhuis GJ: High stem cell frequency in acute myeloid leukemia at diagnosis predicts high minimal residual disease and poor survival. Clinical Cancer Research 2005, 11(18):6520-6527.
18. Hanekamp D, Denys B, Kaspers GJL, Te Marvelde JG, Schuurhuis GJ, De Haas V, De Moerloose B, de Bont ES, Zwaan CM, de Jong A et al: Leukaemic stem cell load at diagnosis predicts the development of relapse in young acute myeloid leukaemia patients. British journal of haematology 2018, 183(3):512-516.
19. Witte KE, Ahlers J, Schafer I, Andre M, Kerst G, Scheel-Walter HG, Schwarze CP, Pfeiffer M, Lang P, Handgretinger R et al: High Proportion of Leukemic Stem Cells at Diagnosis Is Correlated with Unfavorable Prognosis in Childhood Acute Myeloid Leukemia. Pediatr Hemat Oncol 2011, 28(2):91-99.
20. Annesley CE, Brown P: The Biology and Targeting of FLT3 in Pediatric Leukemia. Frontiers in oncology 2014, 4:263.
21. Hunger SP, Lu X, Devidas M, Camitta BM, Gaynon PS, Winick NJ, Reaman GH, Carroll WL: Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children's oncology group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012, 30(14):1663-1669.
22. Maude SL, Laetsch TW, Buechner J, Rives S, Boyer M, Bittencourt H, Bader P, Verneris MR, Stefanski HE, Myers GD et al: Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. The New England journal of medicine 2018, 378(5):439-448.
23. Hanekamp D, Cloos J, Schuurhuis GJ: Leukemic stem cells: identification and clinical application. International journal of hematology 2017, 105(5):549-557.
24. Felipe Rico J, Hassane DC, Guzman ML: Acute myelogenous leukemia stem cells: from Bench to Bedside. Cancer letters 2013, 338(1):4-9.
25. Perna F, Berman SH, Soni RK, Mansilla-Soto J, Eyquem J, Hamieh M, Hendrickson RC, Brennan CW, Sadelain M: Integrating Proteomics and Transcriptomics for Systematic Combinatorial Chimeric Antigen Receptor Therapy of AML. Cancer cell 2017, 32(4):506-519 e505.
26. Pollyea DA, Gutman JA, Gore L, Smith CA, Jordan CT: Targeting acute myeloid leukemia stem cells: a review and principles for the development of clinical trials. Haematologica 2014, 99(8):1277-1284.
27. Jen EY, Ko CW, Lee JE, Del Valle PL, Aydanian A, Jewell C, Norsworthy KJ, Przepiorka D, Nie L, Liu J et al: FDA Approval: Gemtuzumab Ozogamicin for the Treatment of Adults with Newly

Diagnosed CD33-Positive Acute Myeloid Leukemia. Clinical cancer research : an official journal of the American Association for Cancer Research 2018, 24(14):3242-3246.
28. Parigger J, Zwaan CM, Reinhardt D, Kaspers GJ: Dose-related efficacy and toxicity of gemtuzumab ozogamicin in pediatric acute myeloid leukemia. Expert review of anticancer therapy 2016, 16(2):137-146.
29. van Rhenen A, Moshaver B, Kelder A, Feller N, Nieuwint AW, Zweegman S, Ossenkoppele GJ, Schuurhuis GJ: Aberrant marker expression patterns on the CD34+CD38- stem cell compartment in acute myeloid leukemia allows to distinguish the malignant from the normal stem cell compartment both at diagnosis and in remission. Leukemia 2007, 21(8):1700-1707.
30. Bonardi F, Fusetti F, Deelen P, van Gosliga D, Vellenga E, Schuringa JJ: A proteomics and transcriptomics approach to identify leukemic stem cell (LSC) markers. Molecular & cellular proteomics : MCP 2013, 12(3):626-637.
31. Majetl R, Becker MW, Tian Q, Lee TLM, Yan XW, Liu R, Chiang JH, Hood L, Clarke MF, Weissman IL: Dysregulated gene expression networks in human acute myelogenous leukemia stem cells. Proceedings of the National Academy of Sciences of the United States of America 2009, 106(9):3396-3401.
32. Eppert K, Takenaka K, Lechman ER, Waldron L, Nilsson B, van Galen P, Metzeler KH, Poeppl A, Ling V, Beyene J et al: Stem cell gene expression programs influence clinical outcome in human leukemia. Nature medicine 2011, 17(9):1086-1093.
33. Forsberg EC, Passegue E, Prohaska SS, Wagers AJ, Koeva M, Stuart JM, Weissman IL: Molecular signatures of quiescent, mobilized and leukemia-initiating hematopoietic stem cells. PloS one 2010, 5(1):e8785.
34. Gal H, Amariglio N, Trakhtenbrot L, Jacob-Hirsh J, Margalit O, Avigdor A, Nagler A, Tavor S, Ein-Dor L, Lapidot T et al: Gene expression profiles of AML derived stem cells; similarity to hematopoietic stem cells. Leukemia 2006, 20(12):2147-2154.
35. Gentles AJ, Plevritis SK, Majeti R, Alizadeh AA: Association of a leukemic stem cell gene expression signature with clinical outcomes in acute myeloid leukemia. Jama 2010, 304(24):2706-2715.
36. Saito Y, Kitamura H, Hijikata A, Tomizawa-Murasawa M, Tanaka S, Takagi S, Uchida N, Suzuki N, Sone A, Najima Y et al: Identification of therapeutic targets for quiescent, chemotherapy-resistant human leukemia stem cells. Science translational medicine 2010, 2(17):17ra19.
37. de Leeuw DC, Denkers F, Olthof MC, Rutten AP, Pouwels W, Schuurhuis GJ, Ossenkoppele GJ, Smit L: Attenuation of microRNA-126 expression that drives CD34+38- stem/progenitor cells in acute myeloid leukemia leads to tumor eradication. Cancer research 2014, 74(7):2094-2105.
38. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J: qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome biology 2007, 8(2):R19.
39. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology 2002, 3(7):RESEARCH0034.
40. van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL, Delabesse E, Davi F, Schuuring E, Garcia-Sanz R et al: Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia 2003, 17(12):2257-2317.
41. Essand M, Vasmatzis G, Brinkmann U, Duray P, Lee B, Pastan I: High expression of a specific T-cell receptor gamma transcript in epithelial cells of the prostate. Proceedings of the National Academy of Sciences of the United States of America 1999, 96(16):9287-9292.
42. Wolfgang CD, Essand M, Vincent JJ, Lee B, Pastan I: TARP: A nuclear protein expressed in prostate and breast cancer cells derived from an alternate reading frame of the T cell receptor gamma chain locus. Proceedings of the National Academy of Sciences of the United States of America 2000, 97(17):9437-9442.

43. Maeda H, Nagata S, Wolfgang CD, Bratthauer GL, Bera TK, Pastan I: The T cell receptor gamma chain alternate reading frame protein (TARP), a prostate-specific protein localized in mitochondria. The Journal of biological chemistry 2004, 279(23):24561-24568.
44. Qi C, Jia X, Lu L, Ma P, Wei M: HEK293T Cells Are Heterozygous for CCR5 Delta 32 Mutation. PloS one 2016, 11(4):e0152975.
45. Lefranc MP, Forster A, Rabbitts TH: Genetic polymorphism and exon changes of the constant regions of the human T-cell rearranging gene gamma. Proceedings of the National Academy of Sciences of the United States of America 1986, 83(24):9596-9600.
46. Carlsson B, Totterman TH, Essand M: Generation of cytotoxic T lymphocytes specific for the prostate and breast tissue antigen TARP. The Prostate 2004, 61(2):161-170.
47. Hillerdal V, Nilsson B, Carlsson B, Eriksson F, Essand M: T cells engineered with a T cell receptor against the prostate antigen TARP specifically kill HLA-A2+ prostate and breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America 2012, 109(39):15877-15881.
48. Wolfgang CD, Essand M, Lee B, Pastan I: T-cell receptor gamma chain alternate reading frame protein (TARP) expression in prostate cancer cells leads to an increased growth rate and induction of caveolins and amphiregulin. Cancer research 2001, 61(22):8122-8126.
49. Yue H, Cai Y, Song Y, Meng L, Chen X, Wang M, Bian Z, Wang R: Elevated TARP promotes proliferation and metastasis of salivary adenoid cystic carcinoma. Oral surgery, oral medicine, oral pathology and oral radiology 2017, 123(4):468-476.
50. Farrar JE, Schuback HL, Ries RE, Wai D, Hampton OA, Trevino LR, Alonzo TA, Guidry Auvil JM, Davidsen TM, Gesuwan P et al: Genomic profiling of pediatric acute myeloid leukemia reveals a changing mutational landscape from disease diagnosis to relapse. Cancer research 2016, Apr 15(76(8)):2197-2205.
51. Bolouri H, Farrar JE, Triche T, Jr., Ries RE, Lim EL, Alonzo TA, Ma Y, Moore R, Mungall AJ, Marra MA et al: The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nature medicine 2018, Jan(24(1)):103-112.
52. Lamba JK, Cao X, Raimondi SC, Rafiee R, Downing JR, Lei S, Gruber T, Ribeiro RC, Rubnitz JE, Pounds SB: Integrated epigenetic and genetic analysis identifies markers of prognostic significance in pediatric acute myeloid leukemia. Oncotarget 2018, 9(42):26711-26723.
53. Graf C, Heidel F, Tenzer S, Radsak MP, Solem FK, Britten CM, Huber C, Fischer T, Wolfel T: A neoepitope generated by an FLT3 internal tandem duplication (FLT3-ITD) is recognized by leukemia-reactive autologous CD8+ T cells. Blood 2007, 109(7):2985-2988.
54. Quek L, Otto GW, Garnett C, Lhermitte L, Karamitros D, Stoilova B, Lau IJ, Doondeea J, Usukhbayar B, Kennedy A et al: Genetically distinct leukemic stem cells in human CD34- acute myeloid leukemia are arrested at a hemopoietic precursor-like stage. The Journal of experimental medicine 2016, 213(8):1513-1535.
55. Fritzsche FR, Stephan C, Gerhardt J, Lein M, Hofmann I, Jung K, Dietel M, Kristiansen G: Diagnostic and prognostic value of T-cell receptor gamma alternative reading frame protein (TARP) expression in prostate cancer. Histol Histopathol 2010, 25(6):733-739.
56. van Buul JD, Voermans C, van Gelderen J, Anthony EC, van der Schoot CE, Hordijk PL: Leukocyte-endothelium interaction promotes SDF-1-dependent polarization of CXCR4. The Journal of biological chemistry 2003, 278(32):30302-30310.
57. Salgia R, Li JL, Ewaniuk DS, Pear W, Pisick E, Burky SA, Ernst T, Sattler M, Chen LB, Griffin JD: BCR/ABL induces multiple abnormalities of cytoskeletal function. The Journal of clinical investigation 1997, 100(1):46-57.
58. Boyd AL, Campbell CJV, Hopkins CI, Fiebig-Comyn A, Russell J, Ulemek J, Foley R, Leber B, Xenocostas A, Collins TJ et al: Niche displacement of human leukemic stem cells uniquely allows their competitive replacement with healthy HSPCs. Journal of Experimental Medicine 2014, 211(10):1925-1935.

59. Oh S, Terabe M, Pendleton CD, Bhattacharyya A, Bera TK, Epel M, Reiter Y, Phillips J, Linehan WM, Kasten-Sportes C et al: Human CTLs to wild-type and enhanced epitopes of a novel prostate and breast tumor-associated protein, TARP, lyse human breast cancer cells. Cancer research 2004, 64(7):2610-2618.
60. Coulie PG, Van den Eynde BJ, van der Bruggen P, Boon T: Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. Nature reviews Cancer 2014, 14(2):135-146.
61. Hillerdal V, Boura VF, Bjorkelund H, Andersson K, Essand M: Avidity characterization of genetically engineered T-cells with novel and established approaches. BMC immunology 2016, 17(1):23.
62. Heemskerk MH, Hagedoorn RS, van der Hoorn MA, van der Veken LT, Hoogeboom M, Kester MG, Willemze R, Falkenburg JH: Efficiency of T-cell receptor expression in dual-specific T cells is controlled by the intrinsic qualities of the TCR chains within the TCR-CD3 complex. Blood 2007, 109(1):235-243.

Supplementary

1. Data processing and statistical assays

Flow cytometric (FCM) data were analyzed using Infinicyt software v.1.8 (Cytognos, Salamanca, Spain) or DIVA software (BD Biosciences, San Jose, CA, USA). Graphs were generated in GraphPad Prism version 5.04 for Windows (GraphPad Software, La Jolla California USA), Excel or PowerPoint (Windows). Images from gel electrophoresis, Western blotting and confocal microscopy were processed by ImageJ, Fiji and GIMP2 (free software packages available at Ghent University, Ghent, Belgium). Nucleotide sequence chromatograms were evaluated in BioEdit Sequence Alignment Editor for Windows (Ghent University). Reference mRNA sequences and annotations were derived from the University of California Santa Cruz (UCSC) Genome Browser Web-based tool using the GRCh38/hg38 Assembly. Post-sequencing alignment between samples and with UCSC reference sequences was performed in Vector NTI using the AlignX tool (Life Technologies).

Statistical calculations were performed in GraphPad Prism version 5.04 or MedCalc version 12.3.0.0 (Mariakerke, Belgium), with the exception of Chi Square test, for which MedCalc (version 18.11.3) was used. The Spearman's coefficient rank correlation coefficient was used to correlate cytotoxic killing rates with TARP transcript expression. Data were tested for normal distribution using the d'Agostino-Pearson test. One-way ANOVA with Tukey's Multiple Comparison post-test was performed to evaluate TARP transcript expressions between more than two groups. The Mann-Whitney U test was applied as a non-parametric test for independent samples from two groups. Paired sample T-test (Gaussian distribution) or Wilcoxon matched-pairs signed rank tests (non-Gaussian distribution) was used to compare expression levels before and after transduction, between different time points after transduction, and between LSCs and blasts sorted from the same patient. P-values calculated were two-tailed, and one, two, three or four asteriks are indicative the level of significance, set to 5% (0.05), 1% (0.01), 0.1% (0.001) and 0.01% (0.0001), respectively.

2. Material and methods

2.1. Culture, pulsing, human leukocyte antigen typing and fractionation of cell lines

Breast (BT-474, MCF-7) and prostate (PC3, LNCaP) adenocarcinoma cell lines were a gift from the Laboratory of Experimental Cancer Research (Ghent University Hospital, Ghent, Belgium). Four luciferase-expressing AML cell lines (HL60-luc, MOLM-13-luc, MV4;11-luc and OCI-AML-3-luc) were kindly provided by RWJ Groen and HJ Prins from the Cancer Center Amsterdam (CCA) (Vrije Universiteit Medical Center (VUmc), Amsterdam, the Netherlands). All remaining cell lines were purchased at ATCC or DMSZ. These included nine AML cell lines (HL-60, HNT-34, Kasumi-1, Kg-1a, MOLM-13, MONO-MAC6, MV4;11, OCI-AML3, THP-1), five B-ALL cell lines (E2A, REH, NALM-6, SEM, SUPB15), the CML cell line k562, the EBV-immortalized B-cell line JY and the T-ALL cell line HSB-2. Cell lines were grown in media according to supplier instructions at 37 °C in 5% or 7% CO_2 incubators. DMEM, IMDM and RPMI media (Invitrogen) were supplemented with 10% or 20% Fetal Calf Serum (FCS, Hyclone or ThermoFisher Scientific), 100 U/mL Penicillin/Streptomycin (10000 U/ml, Invitrogen) and 100 µg/mL L-Glutamine (200 mM, Invitrogen). For THP-1, medium was additionally supplied with 0.05 mM β-mercaptoethanol.

T2, a human leukocyte antigen (HLA)-A*0201-positive, TAP-deficient cell line, was used for *in vitro* pulsing with antigenic peptides (GenScript HK Limited (Hongkong), overview in Table S3). Per pulsing experiment, one million cells were incubated overnight (O/N) at 37 °C in IMDM supplemented with 1% human serum and 10 µg peptide solubilized in DMSO.

Human leukocyte antigen (HLA)-A sequencing was performed at the Red Cross (Mechelen, Belgium).

Subcellular compartmentalization of cell lines into nuclear and cytoplasmic fractions was performed according to the protocol of Gagnon et al.[1] Total nuclear and cytoplasmic RNA was resuspended in TRIzol and evaluated by qPCR, as described in 2.5.

2.2. Sorting strategy

All scatters were devoid of cell debris and doublets based on propidium iodide (PI) exclusion and FSC-H vs FSC-A, respectively. Sorting strategies were applied depending on the population of interest.

Mononuclear cells (MNC) collected from AML patients and healthy controls were used to sort CD34+CD38+ and CD34+CD38- populations. CD34-postive AML scoring was done as previously defined[2, 3], identifying CD34-positive cases as those with > 1% of CD34+ blasts in the leukemic cells. If the number of CD34-positive cells concerned less than 50% of the total white blood cell (WBC) population, CD34-isolation was performed using the CD34 MicroBead Kit (Milteny). The immature myeloid compartment was defined by CD34, CD45 and scatter properties. CD34+CD38+ blasts and CD34+CD38- stem cells were gated as previously described[4]. Lymphocytes and fluorescence-minus-one (FMO) controls were used to determine CD38 expression cut-offs. Lymphocytes were sorted based on high CD45 expression and low SSC-A. Delineated cell populations were backgated on FSC-A/SSC-A and CD45/SSC-A scatter plots to exclude non-specific events, amongst other myeloid precursor populations. Sorted cells were collected in RPMI supplemented with 50% FCS and a post-sort purity of >90% was reached. Following, cells were spun down (10 min, 3000 rpm, 4° C) and resuspended in TRIzol for RNA and/or DNA extraction (see 2.5).

Transgenic AML cell lines were sorted based on HLA-A*0201, eGFP or Zsgreen expression, depending on the transduction experiment. Sorted cells were collected in RPMI supplemented with 50% FCS, with post-sort purities described in 2.7.4. Sorted cells were further propagated in culture or resuspended in 700 µL TRIzol for RNA and/or DNA extraction (see 2.4).

TCR-transgenic cytotoxic T-cells (CTLs) were sorted directly into 96 well-plates based on CD3/CD8 expression, in combination with being positive for mTCRab or eGFP for LV or RV transduced CTLs, respectively. Sorted cells were expanded on irradiated allogeneic feeder cell medium (see 2.8.4).

2.3. Micro-array profiling

CD34+CD38+ (n=4) and CD34+CD38- (n=3) cell fractions, and lymphocytes (n=4), were sorted from four *de novo* pedAML patients and used for profiling (two FLT3-ITD, two FLT3 WT, Table S1). As control, CD34+CD38+ (n=3) and CD34+CD38- (n=2) cells were sorted from cord blood (CB).

RNA was extracted using the miRNeasy Mini Kit (Qiagen) in combination with on-column DNase I digestion (RNase-Free DNase set, Qiagen) according to manufacturer's instructions. RNA quality and concentrations were measured by Agilent 2100 Bioanalyzer (Agilent) and Qubit (ThermoFisher Scientific), respectively. Mean RIN of all sorted fractions was 9.3 (95% CI 9.1-9.5). Cells were profiled on a custom 8x60K human Gene expression micro-array, containing probes for all human protein-coding genes with lncRNA content based on LNCipedia 2.1[5] (Biogazelle), as follows: 20 ng RNA was pre-amplified using the Complete Whole Transcriptome Amplification Kit (Sigma-Aldrich). Amplified RNA was subsequently labelled using the Genomic DNA ULS Labeling Kit (Agilent) and hybridized to the array in combination with CGH blocking to reduce background signaling. Micro-arrays were analyzed using an Agilent micro-array scanner and Feature Extraction software (v12.0). Probe intensities were background subtracted, quantile-normalized and log2-based probe intensities were calculated. A target was present if the log2 expression value exceeded the cut-off set at 6.75, based on the dark corner control probe value plus 1. Data processing was performed in R using packages EnhancedVolcano, ggplot2, plotMDS, Bioconductor and limma. Principle component analysis showed

clustering of the sorted fractions according to their biological classes (Fig. S3 A, C) and according to their respective patient origin (Fig. S3 E).

2.4. RNA and DNA isolation, cDNA synthesis and (quantitative) PCR

Cell lines ($1x10^6$) and sorted cells (variable cell number) used for qPCR or FLT3-ITD mutational screening were resuspended in 700 µL TRIzol and frozen at -80 °C until further processing. DNA was extracted manually[6] or automatically using the QIAamp DNA Blood Mini Kit on a QIAcube platform (Qiagen) if cells were preserved in 75% ethanol. RNA was extracted using the miRNeasy Mini or Micro Kit (Qiagen) in combination with on-column DNase I digestion (RNase-Free DNase set, Qiagen) according to manufacturer's instructions. RNA concentrations were measured by Nanodrop (ThermoFisher Scientific) or Qubit RNA HS Assay (Invitrogen). cDNA synthesis was performed by Invitrogen SuperScript III or II Reverse Transcriptase (Invitrogen) according to the supplier`s recommendations.

All primers for TARP transcript evaluation were purchased at IDT Technologies (Table S3). The TARP transcript was previously described to initiate within the TRGJP intronic gene segment, followed by TRGC1 coding regions and an untranslated and poly(A) sequence at the end[7] (Fig. S2). Primers for TARP were chosen based on previous literature, different coding regions and the inclusion of the immunogenic $TARP_{4-13}$ coding region. For general TARP transcript evaluation, we used primers targeting the first exon of the TRGC gene segments, referred to as "TARP short". For sequencing, four different primer pairs were selected, targeting the entire (TARP long) or part of (TARP short) all TCRG exons, or exon 1 and exon 3 individually, as illustrated in Fig. S2. To elucidate the subcellular localization of TARP, we used primers directed against *MALAT1* (nuclear-retained non-coding RNA) and *TBP* (cytoplasmic TATA box binding protein). Reference genes for normalisation were *a priori* selected in a pilot study according to the state-of-the-art. The expression of eight housekeeping genes[8] (*GAPD, HMBS, HPRT1, RPL13A, SDHA, TBP, UBC, YWHAZ*) and Alu repeats[9] was investigated in 11 cell lines, chosen based on a broad genetic repertoire and pediatric origin. Adhering to a strict M-value ≤ 0.5 and V-value < 0.15[10], six out of eight housekeeping genes (ranged from highest to lowest stability: *TBP, GAPD, HPRT1, HMBS, SDHA* and *YWHAZ)* were advised for gene of interest normalisation. Due to practical considerations, the three most stable housekeeping genes *GAPD*, *HPRT1* and *TBP* (V-value 0.202, respective M-values 0.45, 0.45 and 0.57) were selected.

qPCR reactions were carried out in 96-well plates using 0.3 µM primers, Sso Fast Evagreen master mix (Bio-Rad), 2.38 ng cDNA and H_2O (Sigma-Aldrich) in a 10 µL reaction. Samples were run in duplicate by a 2-step real-time protocol (2 min 98 °C, followed by 45 cycles (98 °C 5 sec, 60 °C 20 sec)) on a Viia7 analyzer (ThermoFisher), combined with melting curve analysis (65 °C to 95 °C, gradually increasing with 0.5 °C/5 sec). Ct thresholds were automatically determined by the QuantStudio™ Real-Time PCR Software. Efficiency-corrected Ct values were calculated for each primer pair using the LinRegPCR software (AMC, University of Amsterdam, the Netherlands).

As TARP was previously defined as a truncated TCRG transcript[7], contaminating TCRγδ+ lymphocytes during sorting or the rare occurrence of cross-lineage TCRG locus recombinations in AML[11], could impede interpretation. To exclude this, we performed DNA TCRG GeneScan analysis[12] and a qPCR with primers targeting VγI, VγII, VγIII and VγIV gene segments to exclude functional TCRG gene rearrangements (cycling protocol described above).

FLT3 mutational screening assay was developed by Molecular Diagnostics.be (Belgium). Briefly, a multiplex PCR was performed for fms-like tyrosine kinase receptor-3 (*FLT3*), nucleophosmin (*NPM1*) and CCAAT/enhancer binding protein alpha (*CEBPA*) mutations, followed by fragment analysis on an ABI 3130XL Genetic Analyzer using a ROX 500 internal standard (Applied BioSystems).

2.5. Post-qPCR analysis; gel electrophoresis, cloning, purification and sequencing

Amplicons generated by qPCR were run on 3% agarose gels for 2-3 h at 110 V, followed by staining of gels with 0.5 µg/mL ethidium bromide (EtBr) for 20 min. Images were captured using the bio-imager system Gel Doc XR+ (Bio-Rad). An overview of all agarose gel images is shown in Fig. S5. Following quick imaging, separate bands were isolated for gel extraction.

Sequencing of amplicons was performed at Eurofins GATC Biotech (Constance, Germany). Single bands visualized on gel were purified directly from 96-well plates using the Wizard SV PCR Clean-up system (Promega). If multiple bands were present, DNA was extracted from each band per amplicon (indicated as A, B or C; Fig. S5 A) using the MinElute kit (Qiagen) and measured by Nanodrop. If post-purification concentrations were insufficient to allow for accurate sequencing (< 20 ng/µL), the purified amplicons were ligated into a pCR®-Blunt vector according to manufactory instructions. Between 10-100 ng ligation product was transformed into One Shot® TOP10 *E. Coli* bacteria (for method, see 2.7.2). Following O/N incubation, 2-6 colony-forming units (CFU) were picked from each plate and incubated for 16 h at 37 °C in 5 mL Luria-Bertani (LB) broth whilst shaking for plasmid isolation. Day after, miniprep cultures were pelleted for 5 min at 2000 rpm and plasmids were isolated by QIAprep® miniprep kit (Qiagen), according to manufacturer's instructions. A restriction verification digest was done with *Eco*R1 (0.5 µL, 20000 U/mL, NEB) for each ligation product, followed by agarose gel electrophoresis, to select positive transformants.

2.6. Protein detection

2.6.1. Western blotting

Whole-cell protein extracts were prepared by washing 5 million cells three times with 10 mL ice-cold PBS. Cell pellets were resuspended in 1 mL RIPA buffer supplemented with Halt™ Protease and Phosphatase Inhibitor (1/100 diluted, both ThermoFisher Scientific), sheared fifteen times through a 27G needle, snap frozen and incubated on ice for 45 min with regular vortexing. Afterwards, protein lysates were spun down (13000 rpm, 10 min, 4 °C) and supernatant was frozen at -20 °C. Protein concentrations were measured at 560 nm using a GloMax Explorer (Promega) and PierceTM BCA Protein Assay Kit (ThermoFisher Scientific).

For immunoblotting, 40 µg protein extracts were denatured with 4x LDS and 10x DTT reducing agent and run (100 V, 70 min) on a 4-12% BIS-Tris Plus gels in 1x MES running buffer (Life Technologies). Proteins were transferred to a 0.2 µm pore PVDF (Invitrogen) membrane in 1x transfer buffer for 34 min at 20 V, and blocked with 5% milk TBS/0.05% Tween (TBS-T) for 2 h at room temperature (RT). Primary antibodies (anti-huTARP (TP1 a.k.a 1F8, abcam: 1/3000), anti-β-actin monoclonal (BA3R, Invitrogen: 1/2500) or anti-GFP (B-2, sc-9996, Santa Cruz Biotechnology: 1/500)) were diluted in 5% milk TBS-T and incubated O/N at 4 °C whilst shaking. Blots were cut just above the 38 kDa ladder fragment to allow for simultaneous TARP staining and β-actin staining, avoiding reprobing. Following, blots were washed three times with TBS-T and incubated with HRP-linked sheep-anti-mouse secondary antibody (GE Healthcare Life Sciences: 1/5000) in TBS-T/5% milk for 1 h at RT. After washing, proteins were detected by SuperSignal West Femto Substrate (ThermoFisher Scientific) and an ImageQuant Las4000 with CCD camera. Percentage of TARP expression in knockdown cell lines was calculated relative to the respective mock and in respect to β-actin expression using ImageJ.

2.6.2. Confocal microscopy

Adherent cells were plated at $1x10^4$ cells/well in 96 optical glass plates (Perkin Elmer), followed by O/N incubation. Suspension cells were plated at $2x10^4$ cells/well on poly-L-Lysine hydrobromide (PLL, Sigma, 0.1 mg/mL)-coated wells at incubated for another 2 h at 37 °C. After incubation, cells were centrifuged (1200 rpm, 5 min), fixated and permeabilized with 4% paraformaldehyde for 15 min (RT)

and 0.25% Triton X-100 for 10 min (RT), respectively. In between, wells were washed three times with PBS. Blocking was done at 4 °C O/N in PBS/1% BSA/5% goat serum (GS) whilst shaking. The day after, primary antibodies (anti-huTARP: 1/100, anti-HSP60 rabbit polyclonal (H-300, Santa Cruz Biotechnology): 1/50, anti-calnexin rabbit polyclonal (H-70, SC-1139, Santa Cruz): 1/75) were diluted in PBS/1% BSA/5% GS and incubated for 1 h at RT whilst shaking. After three washes with PBS/1%BSA, cells were incubated with secondary antibodies (goat anti-mouse IgG2a Alexa 647, chicken anti-rabbit IgG Alexa 488 or goat anti-mouse IgG1 Alexa488: 1/250) diluted in PBS/1% BSA for 1 h at RT in the dark whilst shaking. Cells were stained with TARP alone or combined with HSP-60 or calnexin. Cell nuclei were counterstained with DAPI and cells were visualized with a confocal fluorescence microscope combined with bright field imaging (Leica SPE).

2.7. Retroviral transduction of AML cell lines

2.7.1. Plasmids and glycerol stocks

The HLA-A*0201 plasmid was kindly provided by T. Mutis (CCA, VUmc, Amsterdam, the Netherlands) and used to generate HLA-A*0201-positive transgenic AML cell lines (see 2.7.4).

pDNR-Lib plasmids expressing end-sequenced TARP cDNA were used to isolate the TARP open-reading frame (ORF) cDNA, as previously described in prostate and breast adenocarcinoma[7] (TARP BC 105589, provided as glycerol stock, catalog# 6339-1, transOMIC technologies (Table S3)). The retroviral pMSCV-Puro IRES GFP (PIG) vector, a kind gift from Joshua Mendel lab, was used to express TARP under transcriptional control of the PGK promotor instead of puromycin, with GFP useful for isolation of positively transduced cells. To create pMSCV-TARP IRES GFP, the entire TARP ORF cDNA sequence was cloned into the pMSCV-PIG vector using *Eco*RI/*Nsi*I restriction enzyme sites. First, two PCR reactions were performed; one to create overlapping ends in the pMSCV-PIG vector with the TARP ORF cDNA (primers P1-P2, Table S3), and one to add *Eco*RI/*Nsi*I restriction enzyme sequences to each side of the TARP ORF cDNA (primers P3-P4, Table S3). PCR reactions were carried out using Phusion High-Fidelity PCR Master Mix (25 µL, New England Biolabs (NEB)), 0.2 µM primers (P1-P2 or P3-P4), 1.5 µL DMSO and 100 ng of pMSCV-PIG or pDNR-Lib plasmid DNA, with H_2O in a final 50 µL volume. The cycling protocol (98 °C 30 sec, followed by 35 cycles (98 °C 10 sec, 60 °C 30 sec, 72 °C 20 sec) and 10 min at 72 °C) was performed on a MasterCycler pro S (ManualShelf). Amplicon sizes were checked by agarose gel electrophoresis (0.9%, 40 min, 45 V) and purified from gel using the NucleoSpin® gel and PCR clean-up kit (Macherey-Nagel™). Second, an overlap PCR was performed to align P1-P2 and P3-P4 amplicons, and create a final 723 bp insert, in which the TARP ORF is under transcriptional control of the PGK promotor and flanked by *Eco*R1 and *Nsi*I restriction enzyme sites. Fifty ng of P1-P2 and P3-P4 amplicons was added to a Phusion master mix (25 µL), containing DMSO (1.5 µL) and H_2O (final volume 50 µL) and amplified without primers (98 °C 30 sec, 10x (98 °C 10 sec, 60 °C 30 sec, 72 °C 20 sec). Then, P1-P4 primers (0.2 µM) were added and amplification continued for another 25 cycles. Amplicon size of the final insert cDNA was checked by gel electrophoresis, followed by excision and purification from gel. Third, 700 ng insert cDNA and pMSCV-PIG vector were double digested using restriction enzymes *Eco*R1 (0.5 µL, 20 000 U/mL, NEB #R3101) and *Nsi*I (1 µL, 10 000 U/mL, NEB #R0127S), 2 µL CutSmart® buffer (NEB) and H_2O in a 20 µL reaction volume. Digest mixtures were incubated for 2 h at 37 °C, followed by 20 min at 80 °C. Digested products were run by agarose gel electrophoresis, and after purification, the cutted insert cDNA and backbone vector were ligated O/N at 20 °C (3:1 molar ratio) using T4 DNA ligase, 10x buffer (NEB) and H_2O (final volume 30 µL). The final ligated vector, containing the TARP ORF cDNA, is referred to as pMSCV-TARP-IRES-GFP, and was subsequently used for transformation (see 2.9.2.).

To generate TARP knockdown (KD) AML cell lines, three shERWOOD UltramiR short hairpin (shRNA) targeting TARP, and one non-targeting shRNA, were purchased as glycerol stocks at transOMIC

technologies (TRHSU2000-445347, Table S3). The pLMN backbone allowed for selection of retroviral integration based on Zsgreen fluorescent expression and Geneticin-supplemented medium (G418 Sulfate, ThermoFisher Scientific).

2.7.2. Plasmid transformation and isolation

pDNR-Lib and pLMN-shRNA 1, 2, 3 and NT plasmids were provided as glycerol stocks. Bacteria were gently scratched from the glycerol stock into LB broth supplemented with chloramphenicol (50 µg/mL, Sigma-Aldrich) and ampicillin (100 µg/mL, Sigma-Aldrich) for pDNR-Lib and pLMN plasmids, respectively. Bacteria were cultured at 33 °C whilst shaking, and after 16 h, bacterial cultures were pelleted (5 min, 2000 rpm) and plasmids were isolated by maxiprep (ZymoPURE™ Plasmid Maxiprep Kit, ZymoResearch) or midiprep (NucleoBond® Xtra Midi, Machery-Nagel), according to manufacturer's instructions.

For HLA-A*0201, pMSCV-TARP-IRES-GFP and pMSCV-PIG vectors were transformed into Stbl3 bacteria (ThermoFisher Scientific) according to manufacture instructions. Briefly, 10-100 ng plasmid DNA, or 1-5 µL H_2O (negative control), were added to 25 µL bacteria. After gentle mixing, tubes were incubated on ice for 30 min. Bacteria were heat-shocked in a 42 °C water bath for 30 sec, and placed on ice. SOC outgrowth medium (975 µL heated at 37 °C, NEB) was added to each tube and vigorously shaken at 37 °C for 1 h at 225 rpm. Afterwards, 100 µL of each bacterial suspension was plated on ampicilline LB agar plates (100 µg/mL) and incubated O/N at 37 °C. Day after, CFU were counted in respect to the control plate, and between 1-6 CFU were further grown at 33 °C in ampicilline- supplemented LB for 16 h whilst vigorous shaking. Day after, bacterial suspensions were centrifuged (5 min, 2000 rpm) and plasmids were isolated by midiprep (NucleoBond® Xtra Midi, Machery-Nagel) or QIAprep® miniprep (Qiagen).

2.7.3. Virus production

Retrovirus encoding MHC-I HLA-A*0201, transgenic TARP for overexpression (pMSCV-TARP-IRES-GFP) and TARP-targeting shRNA for knockdown (pLMN-shRNA1, 2 and 3), together with non-targeting mock controls (pMSCV-PIG and pLMN-NT shRNA, respectively), were generated as follows: HEK 293T packaging cells (Clontech) were seeded at 70% density in 10 cm dishes pre-coated with 0.1% gelatin, and incubated O/N in DMEM supplemented with 10% FCS at 37 °C (5% CO_2). The day after, medium was refreshed 1 h before transfection, and a 25 µg DNA mix was prepared in 450 µL H_2O (3:1:1 ratio; transfer plasmid:pAmpho (Clontech):pHit60 (Roche)). Cells were calcium phosphate transfected (50 µL $CaCl_2$, 500 µL 2XHBS buffer, both by InVitrogen) and incubated O/N at 37 °C (5% CO_2). Medium was refreshed after 16 h, and viral supernatant was collected 48 h and 72 h post-transfection. Viral supernatant was centrifuged to remove cell debris (1800 rpm, 5 min) and immediately used for transduction or frozen in 500 µL aliquots at -80 °C.

2.7.4. Viral transduction

Retroviral transduction of AML cell lines in order was performed as follows: 24-well plates were coated with retronectin (6 µg/well, TaKaRa) O/N at 4 °C. Coated wells were blocked with PBS/2% BSA (30 min, RT) and washed with PBS. Target cells were plated in medium at a target-dependent multiplicity of infection ratio, ranging between 0.5-0.75 x10^6 cells/well, and virus supernatant was added in a 2:1 virus:target ratio in the presence of polybrene (6 µg/well, Sigma-Aldrich). Plates were spinoculated for 90 min at 2300 rpm (32 °C) and incubated at 37 °C (5% CO_2). The next day, transduction was repeated with the exception that after eight hours half the medium was refreshed after which the plates were incubated for 72 h at 37 °C (5% CO_2). The number of biological replicates, transduction efficiency and post-transduction selection method differed between transduction experiments, and are discussed below.

Six AML cell lines (HL-60, Kg-1a, MOLM-13, HL-60-Luc, MOLM-13-Luc and MV4;11-Luc,) were transduced to express HLA-A*0201 MHC-I molecules. Transduction was performed in triplicate per cell line and transduction efficiency ranged between 6.2-89.2 %. Positively transduced cells were selected via HLA-A*0201 APC antibody staining and EasySep™ APC Positive Selection Kit (STEMCELL Technologies) according to manufacturer's instructions. After EasySep selection, HLA-A*0201 transduced cells were on average 96% of the total viable cell population. TARP transcript levels were not influenced by HLA-A*0201 transduction (not shown). The number of HLA-A*0201 expressing cells remained >96% during short-term culture (1 month) but diminished over a period of six months for HL-60 and Kg-1a (60.0% and 55.7%, respectively).

Two HLA-A*0201-positive AML cell lines, OCI-AML3 and THP-1 with low and high endogenous TARP expression respectively, were transduced to overexpress TARP (OE), alongside a mock. Transduction was performed in triplicate and transduction efficiency was evaluated based on eGFP expression after 72 h and 144 h using FCM (range 16.0-57.9%). Efficiencies were shown to be highly comparable between replicates, without significantly difference between both time points ($P>0.05$, Wilcoxon signed-rank test). Transgenic TARP OE and mock OCI-AML3 and THP-1 replicates were pooled, sorted at >90% purity, and eGFP expression sustained >92% during culture afterwards. Post-transduction TARP transcript and protein expression levels were measured after sorting (Fig. S6).

Four TARP-high AML cell lines (HL-60, Kg-1a, MV4;11 and THP-1) were transduced with three different TARP-targeting shRNA, alongside a non-targeting shRNA as mock (single replicate). Seventy-two hours post-transduction, cells were collected and cultured in Geneticin-supplemented medium to increase the percentage of positive retroviral integrants. Geneticin sensitivity of each cell line was determined before transduction, with optimal concentrations of 400 µg/mL, 500 µg/mL, and 300 µg/mL for HL-60, MV4;11 and THP-1, respectively (Kg-1a was unsuccessful). After two weeks on selection, positively transduced cells showed an average Zsgreen expression of 39% (95% CI 27.0-49.6%), 53% (95% CI 49.6-56.4%) and 84% (95% CI 81.5-86.5%) for HL-60, MV4;11 and THP-1, respectively. As selection was not possible for Kg-1a, a much lower percentage of positively transduced cells (mean 6%, 95% CI 2.2-9.3%) was obtained. KD and mock cell lines were subsequently sorted based on Zsgreen expression and further cultured in Geneticin-supplemented medium. Zsgreen expression post-sort was >91 % (95% CI 88.8 - 93.8), except for Kg-1a (mean 24%, 95% CI 15.8-31.5%), and remained stable during culture. Post-transduction TARP transcript (Fig. S9) and protein (Fig. 3C) expression values were measured after sorting.

2.8. Retro- and lentiviral transduction of cytotoxic T-cells (CTLs)

2.8.1. Plasmids

The pBMN(TARP$_{4\text{-}13}$-TCR) target plasmid and pLP1, pLP2 and pLP/VSVG helper plasmids for LV transduction were provided by M. Essand.

The pBMN(TARP$_{4\text{-}13}$-TCR) plasmid encodes a TCRA8-T2A-TCRB12 sequence directed against the HLA-A*0201-restricted synthetic TARP peptide TARP(P5L)$_{4–13}$[13, 14]. However, in our setting, LV transduction appeared to be inconvenient. To start, selection of positively transfected packaging cells was hampered by the lack of an antibiotic resistance or a fluorescent marker in the pBMN backbone. Consequently, transduction efficiencies topped at 10 %, and mouse TCR constant domain (mTCRab) antibody staining was required for the evaluation of transduction efficiencies and sorting out positively transduced CTLs. In addition, the backbone plasmid was not suitable for the generation of non-targeting viral particles and cytotoxic killing experiments consequently lacked mock CTLs to correct for non-TARP mediated lysis. Therefore, the TCR coding region was amplified from the pBMN vector and cloned into the retroviral LZRS-IRES-GFP (LIE) vector.

To this end, primers were designed to amplify the $TARP_{4-13}$-TCR sequence from the pBMN($TARP_{4-13}$-TCR) vector, which incorporate the restriction enzyme sites *Bam*HI and *Xho*I present in the LIE plasmid (acceptor), including a Kozak sequence. These primers (P5 and P6 (0.5 µM), Table S3) were added to 20 ng pBMN plasmid DNA in a 20 µL Phusion High-Fidelity reaction mixture (5x Phusion Green buffer, Phusion hot start II Polymerase (2 U/L, 0.2 µL), dNTPS (0.4 µL, 10 mM) (all by ThermoFisher Scientific), and H_2O). PCR was run in eightfold on a Veriti 96-Well Thermal Cycler (ThermoFisher Scientific) (cycling protocol 98 °C 1 min, followed by 40 cycles (98 °C 7 sec, 69 °C 15 sec, 72 °C 10 sec) and 7 min at 72 °C). Amplicons were pooled, checked by gel electrophoresis, and the $TARP_{4-13}$-TCR sequence with LIE-overlapping ends, hereafter referred to as insert, was purified using the QIAquick Gel Extraction Kit (Qiagen). Subsequently, both insert (150 ng) and LIE backbone vector (75 ng) were double digested with *Bam*HI (0.5 µL, 20 000 U/mL, NEB #R0136S) and *Xho*I (0.5 µL, 20 000 U/mL, NEB #R0146S) restriction enzymes and ligated at 50 °C for 1 h in 10 µL Gibson Assembly® master mix (NEB). The resulting ligated vector, referred to as TARP-TCR LIE plasmid, was used for transformation and plasmid isolation (see 2.10.2).

In addition, we used the empty LIE vector as control, hereafter defined as mock. As positive control, we used an in-house created CMV-TCR encoding LIE vector, using a TCR-sequence provided by Leiden University Medical Center (Leiden, Netherlands) as previously described[15], defined as CMV-TCR LIE plasmid.

2.8.2. Plasmid transformation and isolation

Twenty-five µL DH10B bacteria (NEB) were thawed on ice, and per reaction, 10-100 ng plasmid DNA (pBMN($TARP_{4-13}$-TCR), pLP1, pLP2, pLP/VSVG and LIE plasmids encoding TARP-TCR, CMV-TCR or empty) was added, next to H_2O (15 µL) as control. Transformation was performed as described in 2.7.2. Afterwards, 100 µL of each bacterial suspension was plated on ampicilline LB agar plates (100 µg/mL) and incubated O/N at 37 °C. Day after, CFU were counted in respect to the control plate, and further grown for 16 h at 33 °C in ampicilline-supplemented LB (100 µg/mL) whilst shaking. Day after, bacterial suspensions were spinoculated (5 min, 2000 rpm) and plasmids were isolated by midiprep (NucleoBond® Xtra Midi, Machery-Nagel) or miniprep (QIAprep® Miniprep), according to manufacturer's instructions.

2.8.3. Virus production

TARP-TCR encoding lentivirus was produced by seeding HEK293FT packaging cells ($6.6x10^6$) in T175 flasks containing DMEM with 10% FCS, followed by O/N incubation at 37 °C (5% CO_2). Day after, medium was refreshed 3 h before transfection and a 20 µg DNA mix (2:1:1:1 ratio; pBMN:pLP1:pLP2:pLPVSVG) was prepared. Cells were transfected using jetPEI/NaCl (PolyPlus transfection) and 16 h later, medium was refreshed with reduced serum (1% FCS) DMEM. Viral supernatant was collected 48 h and 72 h post-transfection, immediately placed on ice, spun down (1500 rpm, 7 min, 4 °C) and filtered through a 45 µM low-binding PVDF filter. Subsequently, the filtered supernatant was 10X concentrated using Amicon® Ultra-15 Centrifugal Filters (EMD Millipore™, ThermoFisher Scientific) and aliquoted at -80 °C until use.

TARP-TCR LIE, CMV-TCR LIE or mock encoding retrovirus were produced by seeding phoenix-A packaging cells ($1x10^6$) in 6 cm dishes containing IMDM with 10% FCS, followed by O/N incubation at 37 °C (7% CO_2). Medium was refreshed day after and cells were calcium-phosphate transfected (36 µL $CaCl_2$ (2M) and 300 µL 2XHBS buffer, both by InVitrogen) with 10 µg plasmid DNA after the addition of 1 µL chloroquine (200 mM, Sigma). Medium was refreshed 16 h post-transfection, and after 72 h, transfected packaging cells were transferred to medium supplemented with puromycin (2 µg/ml). Puromycin-supplemented medium was exchanged every three days, and after two weeks, viral

supernatant (eGFP expression >99%) was collected on ice, followed by double spinoculation (1500 rpm, 7 min, 4 °C) and storage at -80 °C until further use.

2.8.4. Viral transduction

Buffy coats for CTL isolation were obtained from the Red Cross (Mechelen, Belgium) and derived from HLA-A*0201-negative donors (n=4, used for LV transduction) or HLA-A*0201 positive donors (n=2, used for RV transduction). MNC were purified from buffy coats using standard Ficoll Density gradient (Axis-shield). CD8-positive CTLs were isolated using CD8-biotine (OKT8) and streptavidin MicroBeads (Milteny Biotech) in PBS supplemented with 2 mM EDTA and 2% FCS. Before transduction, CD8+ CTLs were stimulated with CD3/CD28 T-cell activation Dynabeads (Life Technologies) in the presence of IL-12 (10 ng/mL) and IL-2 (30 IU/mL), or ImmunoCult™ Human CD3/CD28/CD2 T Cell Activator (Stem Cell Technologies) supplemented IL-2 (20 IU/mL), respectively, following manufacturer's guidelines. Seventy-two hours post-stimulation, CD8+ CTLs were collected and plated in retronectin-coated, 2% BSA-PBS blocked 24-well plates. For LV TARP-TCR transduction, $2x10^5$ stimulated CD8+ CTLs were incubated with 10X concentrated TARP-TCR encoding lentivirus per well (1:1 ratio) in the presence of polybrene (4 µg/well) and IL-2 (40 IU/mL). For RV transduction, $2.5x10^5$ stimulated CD8+ CTLs were incubated with TARP-TCR LIE, CMV-TCR LIE or LIE empty retrovirus supplemented with IL-2 (10 IU/mL). Plates were centrifuged for 90 min (2300 rpm, 32 °C) and afterwards incubated at 37 °C (5% CO2). Two rounds of transduction were performed and cells were collected 48 h after the second transduction hit. Transduction efficiencies were by average 5.0% (95% CI 1.88-11.9%) for LV transduction and 58.9% (95% CI 38.3-79.5%) for RV transduction.

Positively transduced CTLs were cell-sorted, as described in 2.3, and expanded on irradiated feeders. Feeders consisted out of 40-Gy irradiated MNC isolated from buffy coats from healthy donors and 50-Gy irradiated JY cells (ratio 10:1) in the presence of 2 µg/mL PHA and 40 IU/mL IL-2. RV transduced CTLs were expanded once whereas LV transduced CTLs required two rounds. Sort purities ranged below 90% for LV transduced CTLs, but mTCRab expression remained stable during feeder expansion (median 50.8%, 95% CI 39.1-60.1%). RV transduced CTLs showed post-sort purities >95% with stable eGFP expression during feeder expansion (median 76.2%, 95% CI 62.3-89.0%). Reactivity was assessed 10 days post-expansion at earliest using freshly collected CTLs or thawed CTLs (O/N incubation at 37 °C in IMDM with 10% FCS and 50 IU/mL IL-2).

2.9. Cytotoxicity assays

2.9.1. Flow cytometry-based assays

2.9.1.1. Cytokine release assay

Feeder-expanded CTLs were incubated with violet tracer (VT)-labelled targets in 96-well U-bottom plates for 1 h (E/T ratio 2/1), followed by the addition of GolgiStop™ (1/750, BD Biosciences)[16]. Effector CTLs incubated in medium or with JY cells (untreated wells) or aCD3(home-made)/aCD28 (BD) coated wells were used as negative (NC) and positive control (PC), respectively. Fifty thousand target cells were added to each well, and conditions were performed in duplicate (quadruplicate for NC and PC). After an additional 16 h of stimulation, cells were harvested into 96-well V-bottom plates and Flow-Count™ Fluorospheres (1:20 diluted, Beckman Coulter) were added to each well. Bead/target/effector cell mixtures were surface stained for 30 min (CD13 or CD33/CD34/CD45 for target cells and CD3/CD8/CD45 for CTLs), washed with PBS, and subsequently fixated and permeabilized (Fixation and Cell Permeabilization Kit, ThermoFisher) followed by intracellular IFN-γ and IL-2 staining. Cytokine release was calculated based on IFN-γ/IL-2 double positive cell populations within the CD3+/CD8+ compartment, evaluated compared to FMO controls, and expressed relatively to the positive control (PC).

2.9.1.2. Cytotoxicity lysis assay

VT-labelled targets were incubated in medium (untreated) or with CTLs (killing) in 96-well U-bottom plates for 48 h (LV, E/T ratio 10/1) or 24 h (RV, E/T ratio 1.25/1, 2.5/1 or 5/1, depending on the experiment). Ten thousand target cells were added to each well, and conditions were performed at least in duplicate (quadruplicate for untreated controls). After incubation, cells were harvested into 96-well V-bottom plates and Flow-Count™ Fluorospheres (1:20 diluted, Beckman Coulter) were added to each well. Bead/target/effector cell mixtures were cell-surface stained as described in 2.11.1.1. In case of HLA-A*0201-positive targets and LV TARP-TCR transduced CTLs, cell mixtures were additionally surface stained with HLA-A2 and mTCRab antibodies, respectively. To evaluate killing, Flow-Count-equalized absolute target cell numbers were calculated and lysis was determined as [1 − (viable target cells (killing)/viable target cells (untreated)] × 100%. For RV transduced CTLs, killing percentages were corrected versus the killing observed by mock CTLs.

2.9.2. ^{51}Chromium release assay

Target cells ($0.5x10^6$) were labelled with 50 µCi ^{51}Chromium (Perkin Elmer, NEZ030005MC) during 90 min at 37 °C (5 % CO_2), and washed twice with 7.5 mL IMDM with 10% FCS before incubation with CTLs. Different E/T ratios were used for LV (5/1, 10/1, 20/1 and 50/1) and RV (1/1, 5/1, 10/1 and 20/1) transduced CTLs and all killing conditions were performed in duplicate. Each target was incubated in medium with or without 2% Triton (Sigma-Aldrich) as PC or NC, respectively (quadruplicate analysis). Following 4 h incubation (37 °C, 5% CO_2), supernatant was harvested and measured in optiphase HISAFE 3 (1:3 ratio) by a 1450 LSC&Luminescence Counter (both by Perkin Elmer). Specific lysis was calculated as [(release (killing) – release (NC))/(release (PC) – release (medium))] x100%.

2.9.3. Bioluminescence imaging-based cell lysis assays

Luciferase (Luc)-positive cell lines were incubated with medium (untreated) or with LV TARP-TCR engineered CTLs (killing) in 96-well U-bottom plates for 8 h, 24 h, 48 h and 56 h (E/T ratio 10/1). Ten thousand target cells were added to each well, and conditions were performed at least in duplicate (quadruplicate for untreated controls). After incubation and adding beetle luciferin (125 µg/mL), luciferase emission (relative light units, RLU) was determined as a measure of target cell viability within 10 min on a GloMax196 Microplate Luminometer (both by Promega) at 420 nm. Baseline luminescence, measured at time point zero, was set at 100% viability for each target. Lysis was calculated as [(RLU (untreated) – RLU (killing)/RLU (untreated)] and specific lysis was calculated for TARP-high cell lines based on the difference between killing of HLA-A*201-positive versus WT targets.

3. Supplemental Tables

Table S1. Characteristics of *de novo* pedAML patients used for sorting CD34+CD38+ and CD34+CD38- cell fractions and micro-array profiling.

		Median (Range)
Age, years		14 (10-15)
WBC count, x 10^9/L		79 (58.1-118)
Morphological blast count		
BM, %		88 (34-96)
PB, %		74 (38-78)
		N (%)
Gender		
F		3 (75%)
M		1 (25%)
CBF leukemia		2 (50%)
WT1 overexpression		2 (50%)
Mutation status	NPM1	0
	FLT3-ITD	2 (50%)

All four pedAML patients were diagnosed in Belgium, classified as standard risk and included for treatment in the Dutch-Belgian (DB AML-01) protocol. Morphological evaluation according to the FAB classification categorized patients as M0 (n=1), M2 (n=1) or M4 (n=2). Bone marrow with CD34+ leukemic cells was used for sorting. *WT1* overexpression was interpreted in regard to in-house or published (Cilloni et al. 2009) cut-offs. CBF-positive leukemias comprised AML with t(8;21)(q22;q22) (n=1) and inv(16)(p13q22) (n=1).

Table S2. Overview of antibodies.

Antibody	**Clone**	**Supplier**	**Category no.**
CD3 BV421	SK7	BD Biosciences	563798
CD3 PE-Cy7	UCHT1	eBioscience	25-0038-42
CD8 APC	SK1	Biolegend	344722
CD8a APC FIRE750	SK1	Biolegend	344746
HLA-A2 FITC	BB7.2	BD Biosciences	343304
HLA-A2 PE	BB7.2	BD Biosciences	558570
HLA-A2 APC	BB7.2	BD Biosciences	561341
IFNγ PE	25723.11	BD Biosciences	340452
IL-2 APC	MQ1-17H12	Biolegend	500310
TCR α/β PE	BW242/412	Miltenyi	130-091-236
CD13 PE	L138	BD Biosciences	347406
CD33 PE	P67.6	BD Biosciences	345799
CD34 PerCP-Cy5.5	8G12	BD Biosciences	333146
CD38 APC-H7	HB7	BD Biosciences	656646
CD45 PacO	HI30	Invitrogen	MHCD4530
LIVE/DEAD® Fixable Near-IR Dead Cell Stain Kit	/	ThermoFisher Scientific	L10119
Anti-Mouse TCR β Chain (mTCRαβ)	H57-597	BD Biosciences	561081

Antibody concentrations were applied as recommended by supplier. No. indicates number.

Table S3. Overview of the nucleic acid sequences of primers and plasmids, and amino acid sequences of peptides.

Name	Type	Nucleic acid (primers/shRNA) or amino acid (peptides) sequence s	
TARP long	primers TARP	GATAAACAACTTGATGCAGATGTTTCC	TTATGATTTCTCTCCATTGCAGCAG
TARP short	primers TARP	ACGGTGCCAGAAAAGTCACTGG	GGGAAACATCTGCATCAAGTTGTTTAT
TARP exon 1	primers TARP	GATAAACAACTTGATGCAGATGTTTCC	CTCAAGAAGACAAAGGTATGTTCCAGC
TARP exon 3	primers TARP	ATACACTACTGCTGCAGCTCACAAACA	TTATGATTTCTCTCCATTGCAGCAG
TCR V_γI-C_γ	primers TRCV(J)C	AACTTGGAAGGGRGAACRAAGTCAGTC	GGGAAACATCTGCATCAAGTTGTTTAT
TCR V_γII-C_γ	primers TRCV(J)C	CGGCACTGTCAGAAAGGAATC	GGGAAACATCTGCATCAAGTTGTTTAT
TCR V_γIII-C_γ	primers TRCV(J)C	TTGGACTTGGATTATCAAAAGTGG	GGGAAACATCTGCATCAAGTTGTTTAT
TCR V_γIV-C_γ	primers TRCV(J)C	TTGGGCAGTTGGAACAACCTGAAA	GGGAAACATCTGCATCAAGTTGTTTAT
MALAT1	primers localisation	GGATTCCAGGAAGGAGCGAG	ATTGCCGACCTCACGGATTT
GAPD	primers housekeeping gene	GGCATGGACTGTGGTCATGAG	GGCATGGACTGTGGTCATGAG
HPRT1	primers housekeeping gene	TGACACTGGCAAAACAATGCA	GGTCCTTTTCACCAGCAAGCT
TBP	primers housekeeping gene/localisation	CGGCTGTTTAACTTCGCTTC	CACACGCCAAGAAACAGTGA
P1	primers for pMSCV-TARP-IRES-GFP	ATCTCTCGAGGTTAACGAAT	/
P2	primers for pMSCV-TARP-IRES-GFP	/	GAAGAAAAATAGTGGGCTTGGGGGAAACATCTG CATCGAAAGGCCCGGA
P3	primers for pMSCV-TARP-IRES-GFP	ATGCAGATGTTTCCCCCAAG	/
P4	primers for pMSCV-TARP-IRES-GFP	/	AATACGTAATGCATATGCATTCATGGTGTTCCCCT
P5	primers for TARP4-13-TCR LIE	GGGTGGACCATCCTCTAGACTGCCGGATCCGCCACCA TGCTGCTGCTG	/
P6	primers for TARP4-13-TCR LIE	/	CGTAGCGGCCGCGGCGCGCCGGCCCTCGAGTTAG CTGTTCTTCTTCTTCACCATGG
FLT3-ITD	primers FLT3	AGCAATTTAGGTATGAAAGCCAGC (FAM)	gtttcttCATCTTTGTTGCTGTCCTTCCAC
pLMN shRNA 1 (ULTRA-3326563)	Knockdown shRNA 1	TGCTGTTGACAGTGAGCGCAAAGATGCAAATGATACACTATAGTGAAGCCACAGATGTATAGTGTATCATTT GCATCTTTTTGCCTACTGCCTCGGA	
pLMN shRNA 2 (ULTRA-3326566)	Knockdown shRNA 2	TGCTGTTGACAGTGAGCGCCACAAACACCTCTGCATATTATAGTGAAGCCACAGATGTATAATATGCAGAGG TGTTTGTGAT GCCTACTGCCTCGGA	
pLMN shRNA 3 (ULTRA-3326567)	Knockdown shRNA 3	TGCTGTTGACAGTGAGCGCCACAATGGATCCCAAAGACAATAGTGAAGCCACAGATGTATTGTCTTTGGGAT CCATTGTGAT GCCTACTGCCTCGGA	
pLMN NT shRNA	Knockdown shRNA mock	TGCTGTTGACAGTGAGCGaaggcagaagtatgcaaagcatTAGTGAAGCCACAGATGTAatgctttgcatacttctgcctgT GCCTACTGCCTCGGA	
pDNR-Lib TARP	plasmid cDNA for transgenic TARP expressing cell lines (TARP ORF cDNA indicated in bold: 176 nucleotides)	GGGGGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGACAAATAAAACACC AAAAGCTTTAAGTTATTTGATTTGTGGAGCAACAGAACTTGTTATGAGCAAAATGAACCAGGACTGGAACCC TGGTCTTTTGAGAATCCCAGACCACCAGAATTTGAAGAACTCAGGGAAACTGAATTAGAGTTTTTGATATGG ACTGAATCACTGTGGAATTATTATAAGAAACTCTTTGGCAGTGGAACAACACTTGTTGTCACAGATAAACAA CTTG**ATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCCAGAAGGCTG GAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATACATTGGCAAGAAAAGAAGAGC AACACGATTCTGGGATCCCAGGAGGGGAACACCATG**AAGACTAACGACACATACATGAAATTTAGCTGGT TAACGGTGCCAGAAAAGTCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGG AGTTGATCAAGAAATTATCTTTCCTCCAATAAAGACAGATGTCATCACAATGGATCCCAAAGACAATTGTTCA AAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTACATGTACCTCCTCCTGCTCC TCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTGCTTAGAAGAACGGCTTTCTGCTGCAATGGAGA GAAATCATAACAGACGGTGGCACAAGGAGGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAGCGTCTTCTG AGGATCTAGTTGGGCTTTCTTTCTGGGTTTGGGCCATTTCAGTTCTCATGTGTGTACTATTCTATCATTATTGTAT AACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGAATAG	
$TARP_{4-13}$	peptide for pulsing	FPPSPLFFFL	
$TARP(P5L)_{4-13}$	peptide for pulsing	FLPSPLFFFL	
$TARP(V28L)_{27-35}$	peptide for pulsing	FLFLRNFSL	
INF_{58-66}	peptide for pulsing	GILGFVFTL	
CMV pp65	peptide for pulsing	NLVPMVATV	

Nucleotide sequences are shown from 5' to 3' direction and the genomic location of the first eight primer pairs is illustrated in Fig. S2. Primers were developed in-house or adapted from the MolecularDiagnostics.be assay (unpublished) or previous reports. Forward (left column) and reverse (right column) sequences from one primer pair are shown in the same row. For peptides, N-terminal anchor residues were modified into leucine (L) for TARP4-13 (proline (P) at position 5) and $TARP_{27-35}$ (valine (V) at position 28) peptides, to increase the ability of the peptide-MHC complex to activate cytotoxic T-cells. TARP indicates T-cell receptor γ chain alternate reading frame protein; INF, influenza A; CMV, cytomegalovirus.

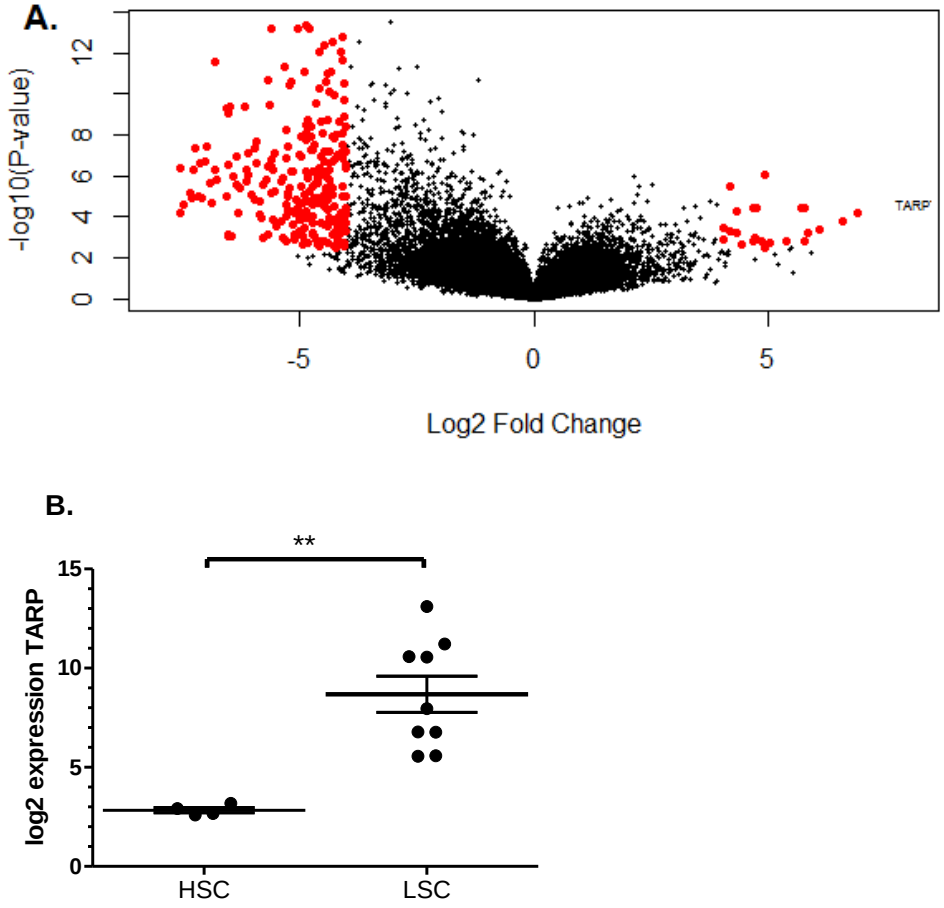
Table S4. Exclusion of functional TCRG recombinations.

Type of sample	Sample code / cell line name	Method	Result
adult AML blast	adult AML 2	DNA TCRG GeneScan analysis / qPCR	NP/neg
adult AML blast	adult AML 6	DNA TCRG GeneScan analysis / qPCR	NP/neg
adult AML blast	adult AML 8	DNA TCRG GeneScan analysis / qPCR	NP/neg
adult AML blast	adult AML 12	DNA TCRG GeneScan analysis / qPCR	NP/neg
adult AML blast	adult AML 13	qPCR	neg
adult AML blast	adult AML 14	qPCR	neg
adult AML blast	adult AML 15	qPCR	neg
adult AML blast	adult AML 16	qPCR	neg
adult AML LSC	adult AML 12	qPCR	neg
AML cell line	HL-60	DNA TCRG GeneScan analysis / qPCR	NP/neg
AML cell line	HNT-34	DNA TCRG GeneScan analysis / qPCR	NP/neg
AML cell line	Kasumi-1	qPCR	neg
AML cell line	Kg-1a	DNA TCRG GeneScan analysis / qPCR	NP/neg
AML cell line	MOLM-13	DNA TCRG GeneScan analysis / qPCR	NP/neg
AML cell line	MONO-MAC6	DNA TCRG GeneScan analysis / qPCR	NP/neg
AML cell line	MV4;11	DNA TCRG GeneScan analysis / qPCR	NP/neg
AML cell line	OCI-AML3	qPCR	neg
AML cell line	THP-1	qPCR	neg
B-ALL cell line	E2A	qPCR	neg
B-ALL cell line	NALM-6	qPCR	neg
B-ALL cell line	REH	qPCR	neg
B-ALL cell line	SEM	qPCR	neg
B-ALL cell line	SUP-B15	qPCR	neg
breast adenocarcinoma cell line	BT-474	DNA TCRG GeneScan analysis / qPCR	NP/neg
breast adenocarcinoma cell line	MCF7	qPCR	neg
CB blast	Pool CB	qPCR	neg
CB blast	pooled CB 1	qPCR	neg
CB blast	pooled CB 2	qPCR	neg
CB HSC	Pool CB	qPCR	neg
CB HSC	pooled CB 1	qPCR	neg
CB HSC	pooled CB 2	qPCR	neg
NBM blast	NBM 3	DNA TCRG GeneScan analysis / qPCR	NP/neg
NBM blast	NBM 5	qPCR	neg
NBM HSC	NBM 3	qPCR	neg
NBM HSC	NBM 5	qPCR	neg
pedAML blast	pedAML 1	qPCR	neg
pedAML blast	pedAML 11	DNA TCRG GeneScan analysis / qPCR	NP/neg
pedAML blast	pedAML 2	DNA TCRG GeneScan analysis / qPCR	NP/neg
pedAML blast	pedAML 3	qPCR	neg
pedAML blast	pedAML 4	qPCR	neg
pedAML blast	pedAML 6	DNA TCRG GeneScan analysis / qPCR	NP/neg
pedAML LSC	pedAML 11	DNA TCRG GeneScan analysis / qPCR	NP/neg
pedAML LSC	pedAML 2	DNA TCRG GeneScan analysis / qPCR	NP/neg
pedAML LSC	pedAML 4	qPCR	neg
pedAML lymfo	pedAML 4	Heteroduplex	polyclonal
pedAML lymphocytes	pedAML 2	DNA TCRG GeneScan analysis / qPCR	polyclonal/pos
pedAML lymphocytes	pedAML 4	qPCR	pos
prostate adenocarcinoma cell line	LNCaP	DNA TCRG GeneScan analysis / qPCR	NP/neg
prostate adenocarcinoma cell line	PC3	qPCR	neg

Sorted patient samples and cell lines were analyzed for functional TRGV(J)C rearrangements using DNA TCRG GeneScan analysis and/or TRGV(J)C qPCR to exclude the presence of contaminating TCRγδ+ lymphocytes or the rare occurrence of cross-lineage TCRγ locus rec36ombinations, which could impede TARP expression. NP indicates no product; neg, negative (no signal).

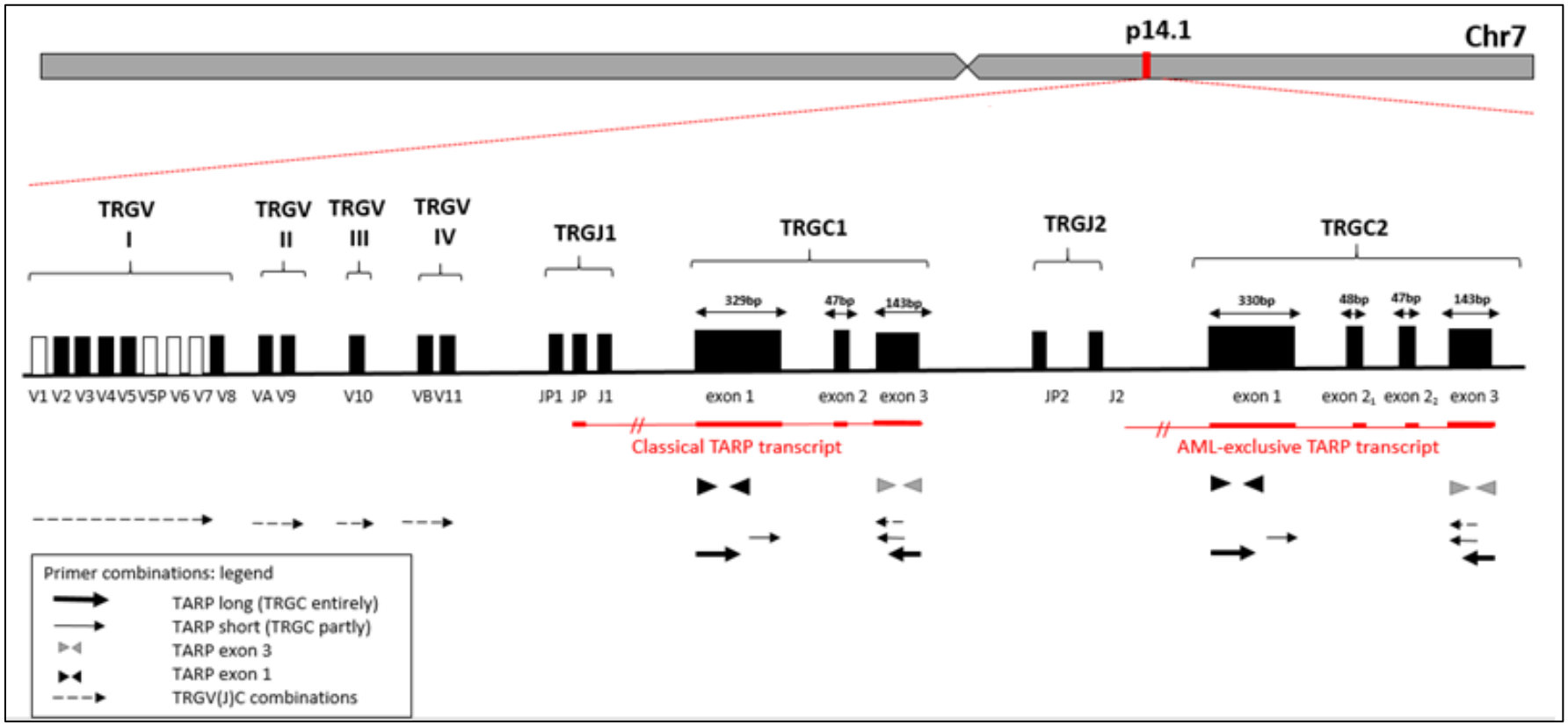
4. Supplemental figures

Figure S1. Differential gene expression in leukemic stem cells (LSC) and hematopoietic stem cells (HSC) based on the dataset from Majeti et al.



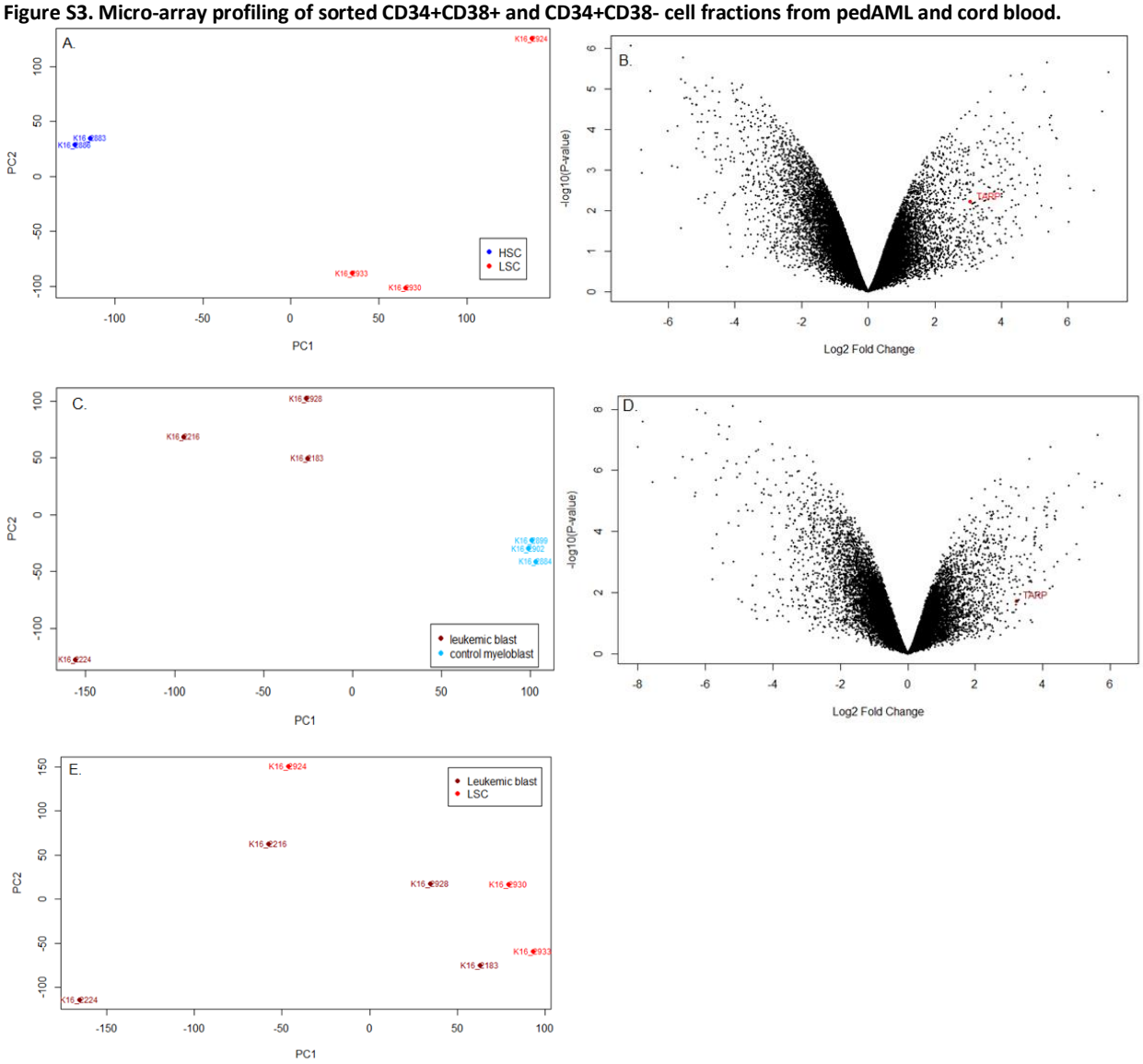
Array included four probes for the detection of TARP (216920_s_at, 209813_x_at, 215806_x_at and 211144_x_at) within LSCs (n=9) and HSCs (n=4). (A) Volcano plot showing differentially up- and down regulated genes between LSC and HSC, with 4-fold or higher expression differences (P<0.01) indicated in red. TARP was the most differentially upregulated gene (maximal log2-FC 6.92). (B) The y-axis represents the TARP mRNA log2 expression values, the x-axis the different sample groups. Horizontal bars indicate means and whiskers are representative for the ±SEM. TARP transcript expression is significantly higher in LSC versus HSC (P<0.01, Mann Whitney U test). FC indicates fold change.

Figure S2. The T-cell receptor γ chain alternate reading frame protein (TARP) at genomic level.



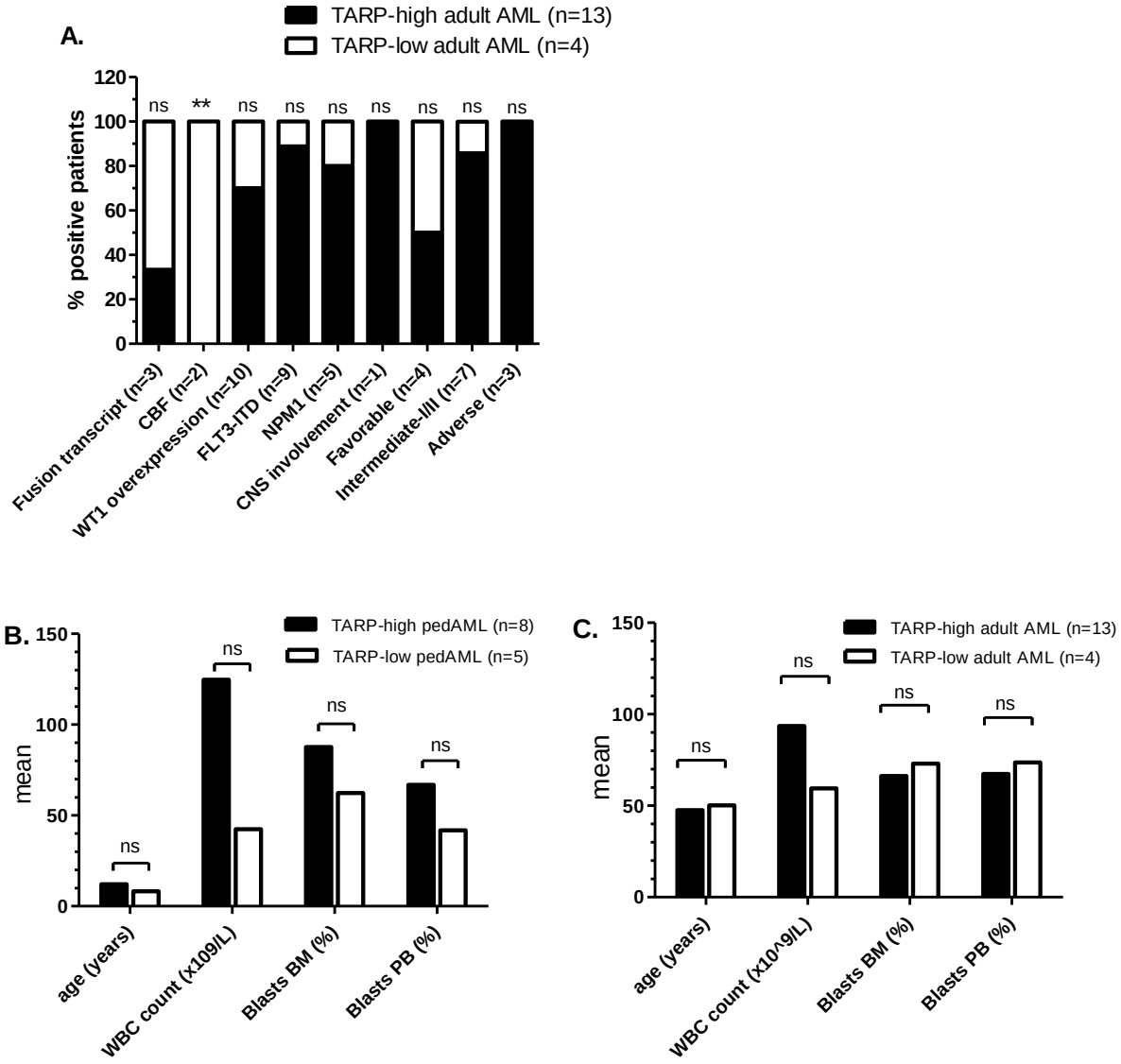
Reference mRNA sequences and annotations were derived from previous work in prostate and breast adenocarcinoma and from the University of California Santa Cruz (UCSC) Genome Browser. The T-cell receptor gamma (TRG) locus, located at 7p15-p14, spans 160 kb and consists out of 12-15 TRG variable (TRGV) genes upstream of a duplicated TRG joining (TRGJ)/TRG constant (TRGC) domain cluster. The TRGV genes are located centromeric with the TRGC2 gene located telomeric in the locus. The first TRGJ/TRGC cluster is composed out of three TRGJ genes (TRGJP1/TRGJP/TRGJ1), each consisting of one exon, and the TRGC1 gene, consisting out of three exons. The second TRGJ/TRGC cluster is separated 16 kb from the first cluster and consists out of two TRGJ genes (TRGJP2/TRGJ2, one exon) and the TRGC2 gene consisting out of four exons. TRGC2 is described to carry a duplicated second exon compared to TRGC1, and both gene segments additionally differ by the presence of single nucleotide variations. The number of base pairs are indicated above each exon of the TRGC locus. Primers used to target part of all TRGC exons (TARP short) are indicated by thin arrows, primers targeting the entire TRGC coding region (TARP long) by thick arrows, and primers targeting solely exon 1 (black) and exon 3 (grey) are indicated by arrow heads. Primers to detect functional TRGV(J)C gene rearrangements are indicated by arrows with dashed lines. mRNA transcripts for the classical and alternative AML-exclusive TARP transcript are indicated in red.

Figure S3. Micro-array profiling of sorted CD34+CD38+ and CD34+CD38- cell fractions from pedAML and cord blood.



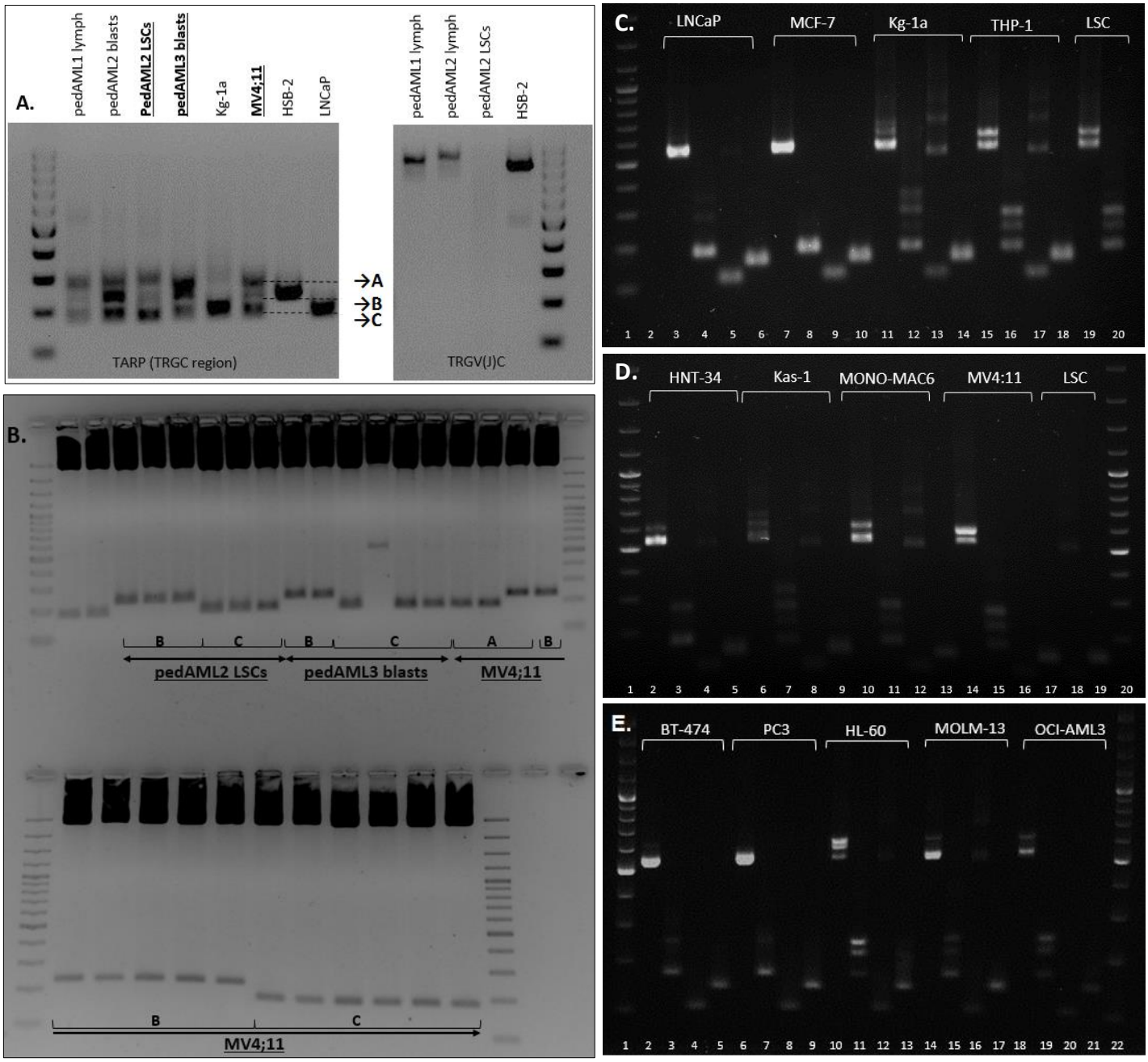
CD34+CD38+ (n=4, leukemic blast) and CD34+CD38- (n=3, LSC) cell fractions sorted from four pedAML patients (2 FLT3-ITD, 2 FLT3 WT; Table S1) were used for micro-array profiling, next to sorted CD34+CD38+ (n=3, control myeloblast) and CD34+CD38- (n=2, HSC) cells from cord blood as control populations. (A) PCA illustrates a clear separation between LSCs (red: K16_2924, K16_2933 and K16_2930) and HSC (blue: K16_2883 and K16_2886). (B) Volcano plot showing differentially up- and downregulated genes between LSCs and HSCs, with TARP indicated in red (log2-FC 3.06, P<0.01). (C) PCA illustrates a clear separation between leukemic blasts (dark red: K16_2216, K16_2183, K16_2928, K16_2224) and control myeloblasts (light blue: K16_2884, K16_2899, K16_2902). (D) Volcano plot showing differentially up- and downregulated genes between leukemic and control myeloblasts, with TARP indicated in dark red (log2-FC 3.22, P<0.05).(E) PCA illustrating of sorted LSCs (CD34+CD38- (red)) and leukemic blasts (CD34+CD38+ (dark red)) showed a correct clustering on a per patient basis (FLT3 WT pedAML: K16_2183/K16_2933 and K16_2928/K16_2930, FLT3-ITD pedAML: K16_2216/K16_2924). For one FLT3 WT pedAML patient, sorted LSCs were lacking (leukemic myeloblasts K16_2224 indicated in dark red). PCA indicates Principle Component Analysis.

Figure S4. Comparison of characteristics between TARP-high and TARP-low patients.



Patients were categorized as TARP-high or TARP-low by qPCR evaluation. Details on the patient characteristics are shown in Table 1. (A) Bars display the percentage of patients harbouring the characteristic shown in the x-axis (dichotomous variables), for TARP-high (black, n=13) and TARP-low (white, n=4) adult AML patients. The total number of patients positive for each characteristic is shown between parentheses. Patients without central nerve system (CNS) involvement either had no clinical manifestations or negative lumbar punctures. The number of CBF-leukemia was significantly ($P<0.01$) higher in TARP-low adult AML patients (Chi Square test). (B-C) Bars display the mean value for the characteristic, shown in the x-axis (continuous variables), calculated for TARP-high (black; pedAML=8 and adult AML=13) and TARP-low (white; pedAML=5 and adult AML=4) patients. No significant differences were detected between adult nor pediatric TARP-high and TARP-low patients (Mann Whitney U test).

Figure S5. TARP transcript analysis.



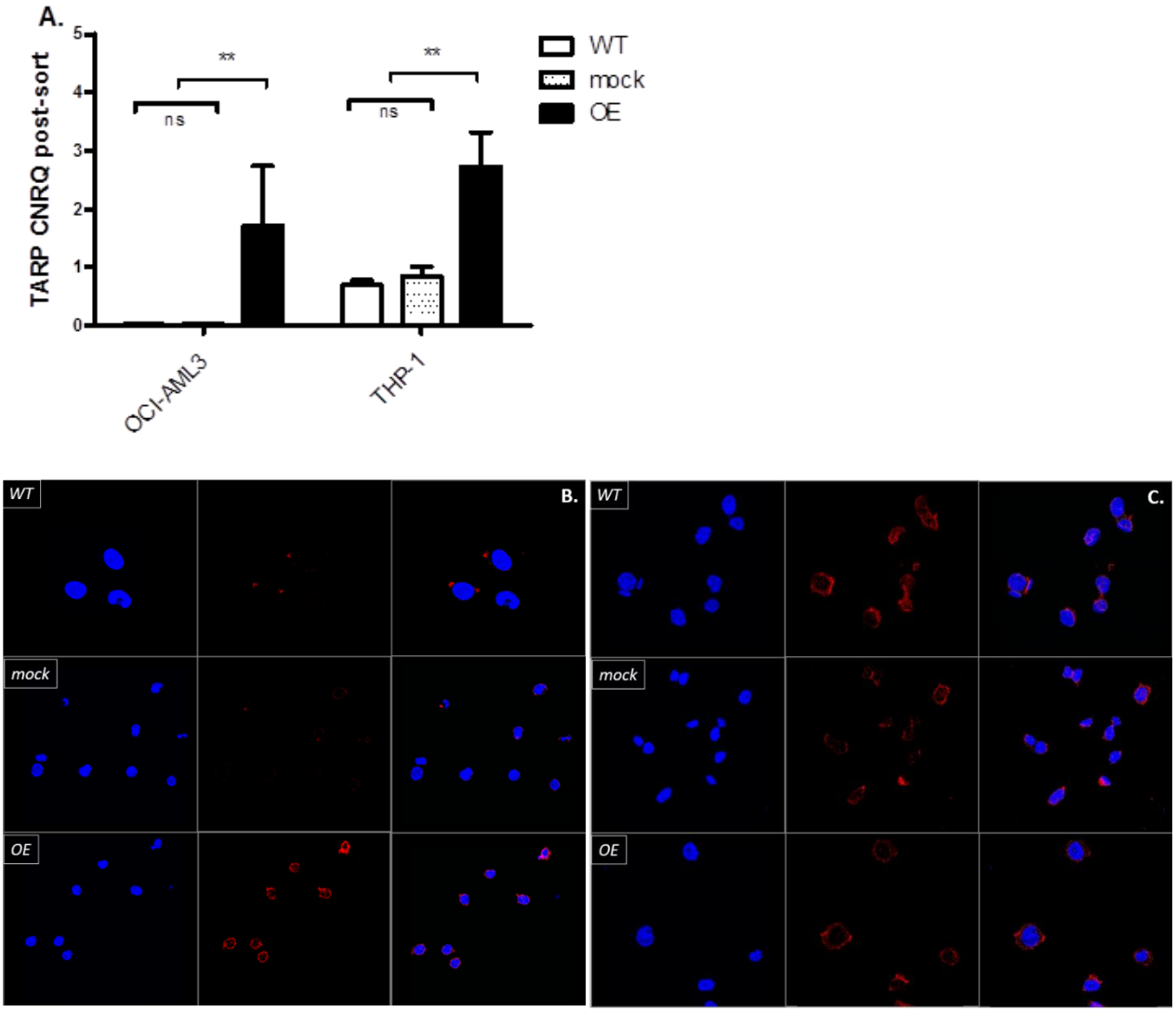
A GeneRuler 100 bp Plus DNA ladder (range 100 – 3000 bp, 500 and 1000 bp fragment highlighted, ThermoFisher Scientific) was used for size estimation. Origin of the amplicons are indicated at the top or bottom of each lane.

(A) Amplicons obtained by TARP long primers shown for (left to right): pedAML1 lymphocytes, pedAML2 blasts, pedAML2 LSCs, pedAML3 blasts, Kg-1a (AML), MV4;11 (AML), HSB-2 (T-ALL) and LNCaP (prostate adenocarcinoma). PedAML2 en pedAML3 carried FLT3-ITD mutations and were categorized as TARP-high by qPCR, whereas pedAML1 was categorized as TARP-low. Kg-1a, HSB-2 and LNCaP showed single bands, whereas triple bands were present for sorted patients fractions and MV4;11. The smallest fragments (C) were 172 bp in size, middle sized fragments (B) were 48 bp longer (220 bp) and the largest fragments (A) were approximately 250-300 bp. Bands detected for pedAML lymphocytes were proven to be part of functional TCRγ recombinations, comparable to HSB-2 cell line. By contrast, TRGV(J)C rearrangements were excluded for the AML cell line MV4;11 and sorted blasts and LSCs from pedAML patients (Table S4).

(B) DNA was extracted from A, B and C bands for pedAML2 LSCs, pedAML3 blasts and MV4;11 (underlined and bold in (A)), and all purified amplicons were ligated into a pCR®-Blunt vector, followed by transformation, colony picking, miniprep and *Eco*R1 restriction digestion. The digested plasmid of each CFU (2-6 per plasmid), containing either the A, B or C band, was run by gel electrophoresis. Digested products from A fell apart into B or C, and B-fragments were consistently 48 bp longer than C.

(C-E) Two prostate adenocarcinoma cell lines (LNCaP and PC3), two breast adenocarcinoma cell lines (BT-474 and MCF-7), nine AML cell lines (MOLM-13, MV4;11, Kg-1a, THP-1, HNT-34, Kas-1, MONO-MAC6, HL-60 and OCI-AML3) and pedAML2 LSCs were targeted for amplification by four different primer pairs e.g. TARP short, TARP long, TARP exon 1 and TARP exon 3 (genomic location primers shown in Fig. S2). Amplicons are shown per cell line in this respective order, adjacent to each other, except for pedAML LSC amplicons which were divided over two gels (first two amplicons C, last two amplicons in D). Band intensities agreed with expression levels determined by qPCR (Fig. 1F). CFU indicates colony-forming-units; Kas-1, Kasumi-1.

Figure S6. Differential TARP transcript and protein expression levels observed in transgenic OE, mock and WT OCI-AML3 and THP-1 cell lines.



(A) Transgenic TARP OE cell lines showed significantly higher expression levels compared to the mock and WT parental lines ($P<0.01$, Mann-Whitney U test). TARP expression was based on biological triplicates, and error bars indicate ±SEM.
(B-C) Comparison of TARP protein levels between WT (top), mock (middle) and transgenic OE (bottom) cell lines generated for OCI-AML3 (B) and THP-1 (C). The first column illustrates only DAPI counterstaining, the second column only TARP staining, and the third column merged images. TARP protein levels were quite comparable between the OE, TC and WT cell lines for THP-1. By contrast, the transgenic OCI-AML3 OE cell line showed an increased TARP protein expression compared to the mock and WT cell line. OE indicates overexpression; WT: wild-type.

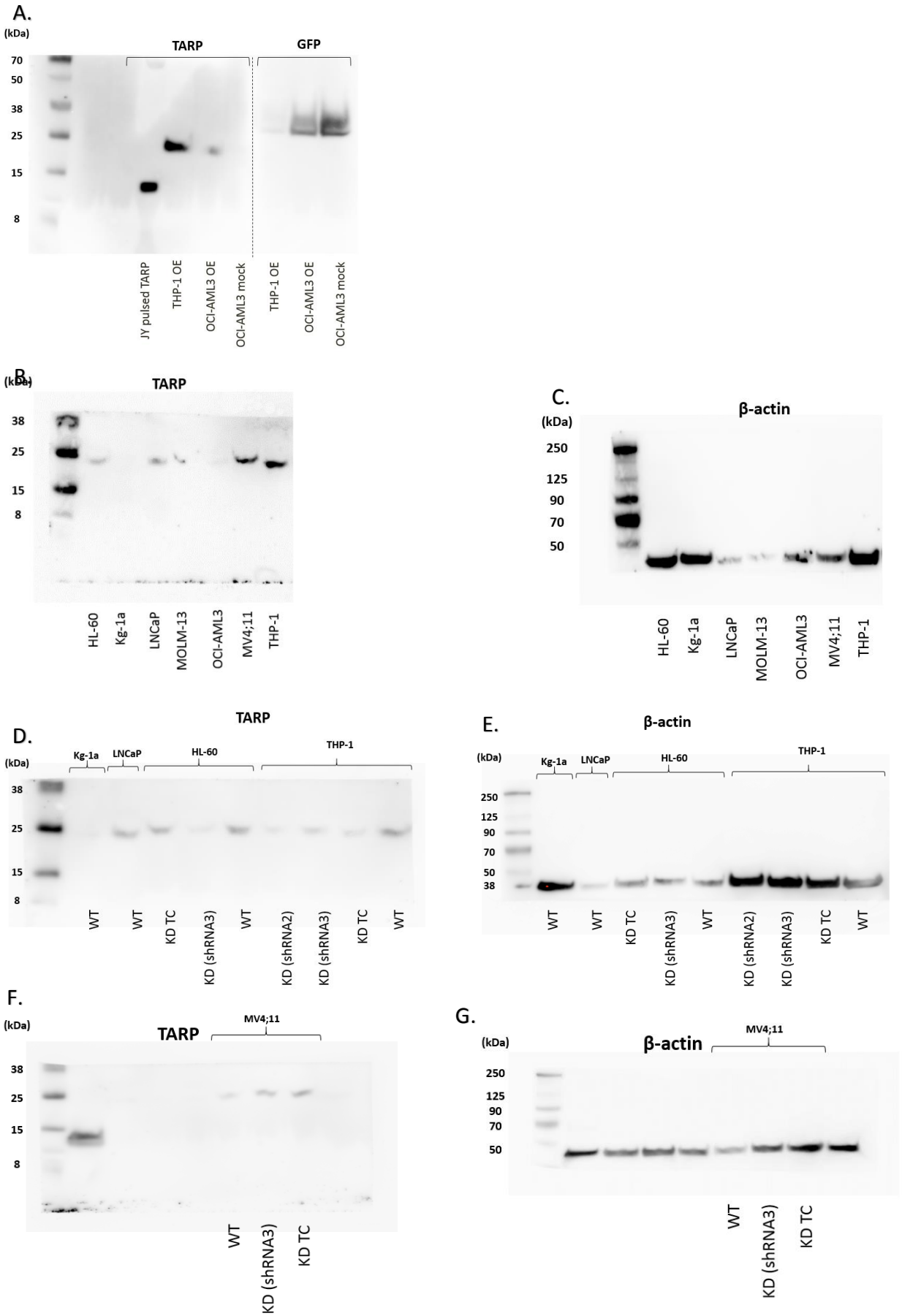


Figure S7. Whole-blot Western blot images. Overview of the whole-blot images, including a pre-stained ladder (WesternSure), which were used to generate Fig. 2 in the main document. B-C, D-E and F-G derived from the same blot and were cut just above the 38 kDa ladder fragment to allow simultaneous TARP and β-actin staining and avoid reprobing.

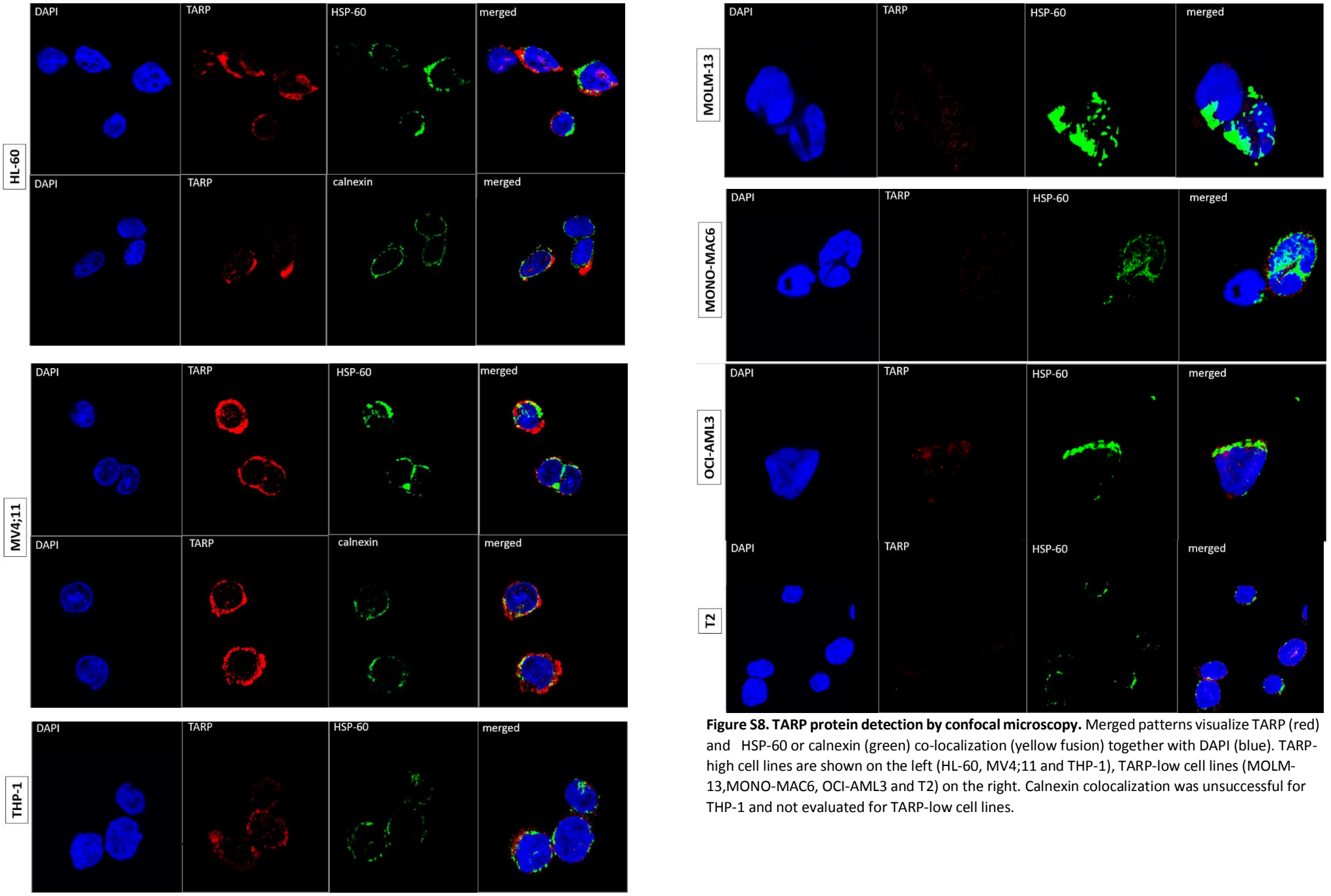
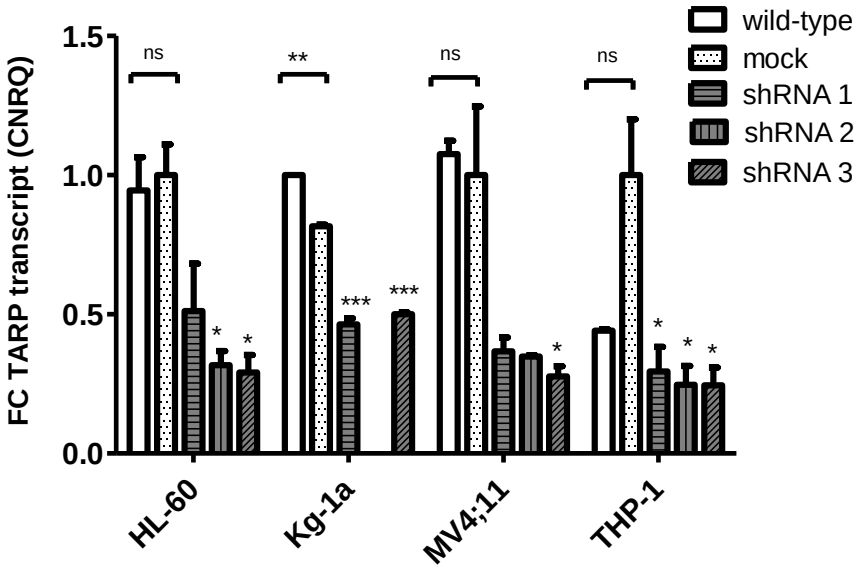


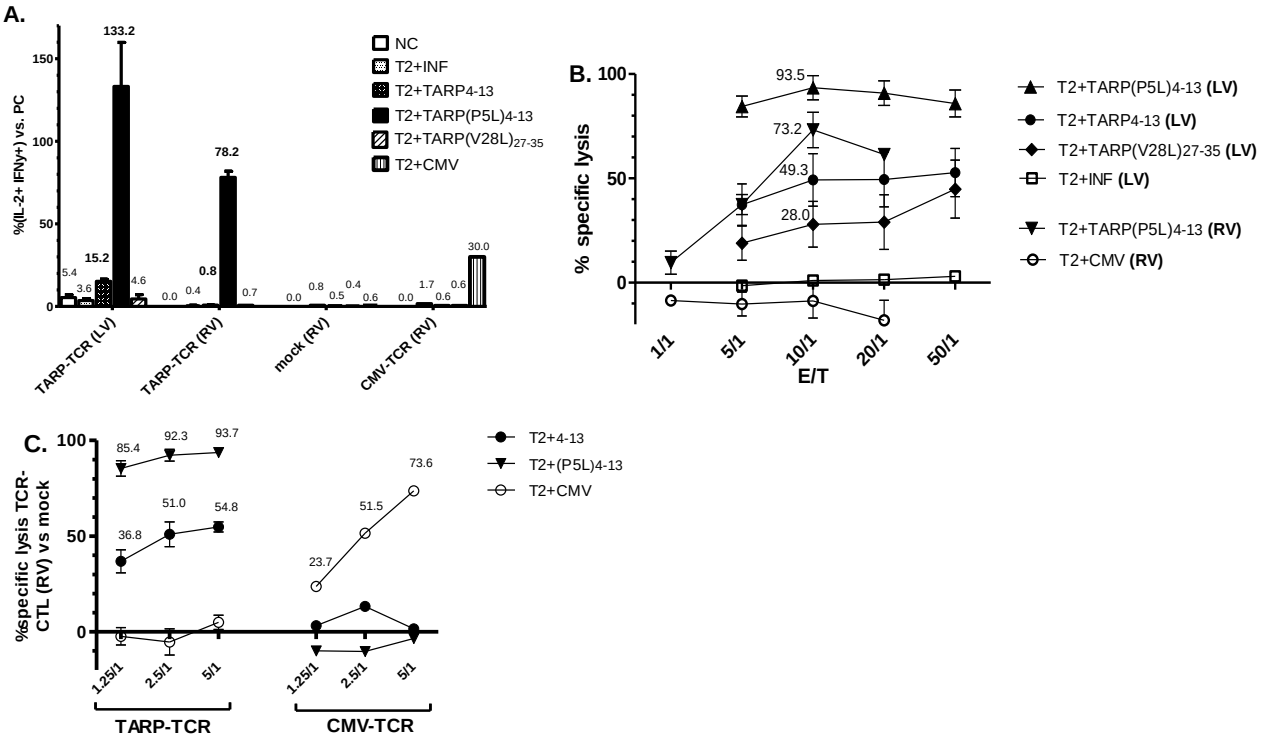
Figure S8. TARP protein detection by confocal microscopy. Merged patterns visualize TARP (red) and HSP-60 or calnexin (green) co-localization (yellow fusion) together with DAPI (blue). TARP-high cell lines are shown on the left (HL-60, MV4;11 and THP-1), TARP-low cell lines (MOLM-13,MONO-MAC6, OCI-AML3 and T2) on the right. Calnexin colocalization was unsuccessful for THP-1 and not evaluated for TARP-low cell lines.

Figure S9. Differential TARP transcript expressions between KD, mock and wild-type cell lines.



Knockdown (shRNA 1, 2, 3) and mock cell lines were generated for HL-60, Kg-1a, MV4;11 and THP-1. Transcript expression levels (CNRQ) are based on biological duplicates and expressed as fold change (FC) versus mock (HL-60, MV4;11 and THP-1) or WT (Kg-1a) cell line, with error bars indicative for ±SEM. Due to low viability during culture, Kg-1a transduced with shRNA 2 could not be evaluated. No significant differences were observed between TC and WT cell lines, although THP-1 TC was marginally upregulated, except for Kg-1a ($P<0.01$, Mann-Whitney U test). Significant downregulation was therefore calculated between each KD cell line and the respective TC (HL-60, MV4;11, THP-1), or the WT cell line in case of Kg-1a, using One-way ANOVA with Tukey's Multiple Comparison post-test. If significant, one ($P<0.05$), two ($P<0.01$) or three ($P<0.001$) asterisks are indicated above the respective chart. KD indicates knockdown.

Figure S10. Functional evaluation of TCR-transgenic CTLs towards pulsed T2 cells.



An overview of the peptides used for pulsing T2 cells is shown in Table S3. Error bars indicate ±SEM based on two, four or six biological replicates, depending on the experiment, and mean values are shown above error bars if applicable.
(A) Feeder-expanded LV and RV TARP-TCR CTLs were incubated with pulsed T2 cells. Cytokine release was calculated based on IFN-γ/IL-2 double positive cell populations within the CD3+/CD8+ compartment, and expressed relatively to the positive control. TARP-TCR CTLs exerted a peptide-specific IFN-γ and IL-2 production, reacting stronger against the TARP(P5L)$_{4\text{-}13}$ than the cognate TARP$_{4\text{-}13}$ peptide. RV transduced TARP-TCR CTLs appeared to be less responsive than LV transduced CTLs, though remained specific, as mock CTLs did not respond and CMV-TCR CTLs only reacted against the CMV peptide.
(B) Chromium^{51} release assay evaluated killing by LV TARP-TCR CTLs (E/T ratio 5/1, 10/1, 20/1, 50/1: upper part of legend) and by RV TARP-TCR CTLs (E/T ratio 1/1, 5/1, 10/1, 20/1: bottom part of legend). TARP(P5L)$_{4\text{-}13}$ pulsed T2 cells were more efficiently lysed than TARP$_{4\text{-}13}$ pulsed T2 cells. For both targets, lysis started from an E/T ratio 5/1 and peaked at E/T 10/1. Killing evaluated by RV transduced CTLs was lower compared to LV transduced CTLs (e.g. for TARP(P5L)$_{4\text{-}13}$: 93.5% vs. 73.2% at ratio 10/1).
(C) FCM-based cytotoxicity assay (24-h) illustrating cytotoxic killing of TARP-related peptides (black symbols) and TARP-unrelated peptides (white symbols). Llysis started from E/T 1.25/1, with again higher killing towards TARP(P5L)$_{4\text{-}13}$ versus TARP$_{4\text{-}13}$ (mean 90.5% vs. 47.5%), while non-TARP related peptides remained unaffected.
CTL indicates cytotoxic T-cells; FCM, flow cytometry; IFN-γ, interferon gamma; IL-2, interleukin-2; LV, lentiviral; RV, retroviral.

5. Supplemental references

1. Gagnon KT, Li L, Janowski BA, Corey DR: Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading. Nature protocols 2014, 9(9):2045-2060.
2. Kersten B, Valkering M, Wouters R, van Amerongen R, Hanekamp D, Kwidama Z, Valk P, Ossenkoppele G, Zeijlemaker W, Kaspers G et al: CD45RA, a specific marker for leukaemia stem cell sub-populations in acute myeloid leukaemia. British journal of haematology 2016, 173(2):219-235.
3. Zeijlemaker W, Kelder A, Wouters R, Valk PJ, Witte BI, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: Absence of leukaemic CD34 cells in acute myeloid leukaemia is of high prognostic value: a longstanding controversy deciphered. British journal of haematology 2015, Oct(171(2)):227-238.
4. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, Scholten WJ, Snel AN, Veldhuizen D, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. Leukemia 2016, 30(2):439-446.
5. Volders PJ, Helsens K, Wang X, Menten B, Martens L, Gevaert K, Vandesompele J, Mestdagh P: LNCipedia: a database for annotated human lncRNA transcript sequences and structures. Nucleic acids research 2013, 41(Database issue):D246-251.
6. Technical Bulletin, TRI Reagent (T9424), Sigma-Aldrich.
7. Wolfgang CD, Essand M, Vincent JJ, Lee B, Pastan I: TARP: A nuclear protein expressed in prostate and breast cancer cells derived from an alternate reading frame of the T cell receptor gamma chain locus. Proceedings of the National Academy of Sciences of the United States of America 2000, 97(17):9437-9442.
8. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology 2002, 3(7):RESEARCH0034.
9. Rihani A, Van Maerken T, Pattyn F, Van Peer G, Beckers A, De Brouwer S, Kumps C, Mets E, Van der Meulen J, Rondou P et al: Effective Alu repeat based RT-Qpcr normalization in cancer cell perturbation experiments. PloS one 2013, 8(8):e71776.
10. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J: qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome biology 2007, 8(2):R19.
11. Boeckx N, Willemse MJ, Szczepanski T, van der Velden VH, Langerak AW, Vandekerckhove P, van Dongen JJ: Fusion gene transcripts and Ig/TCR gene rearrangements are complementary but infrequent targets for PCR-based detection of minimal residual disease in acute myeloid leukemia. Leukemia 2002, 16(3):368-375.
12. van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL, Delabesse E, Davi F, Schuuring E, Garcia-Sanz R et al: Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia 2003, 17(12):2257-2317.
13. Carlsson B, Totterman TH, Essand M: Generation of cytotoxic T lymphocytes specific for the prostate and breast tissue antigen TARP. The Prostate 2004, 61(2):161-170.
14. Hillerdal V, Nilsson B, Carlsson B, Eriksson F, Essand M: T cells engineered with a T cell receptor against the prostate antigen TARP specifically kill HLA-A2+ prostate and breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America 2012, 109(39):15877-15881.
15. Heemskerk MH, Hagedoorn RS, van der Hoorn MA, van der Veken LT, Hoogeboom M, Kester MG, Willemze R, Falkenburg JH: Efficiency of T-cell receptor expression in dual-specific T cells is controlled by the intrinsic qualities of the TCR chains within the TCR-CD3 complex. Blood 2007, 109(1):235-243.

16. Bacher P, Scheffold A: Flow-cytometric analysis of rare antigen-specific T cells. Cytometry Part A : the journal of the International Society for Analytical Cytology 2013, 83(8):692-701.

VII.3.

Results:

Clinical significance of *TARP* expression in pediatric acute myeloid leukemia.

B. Depreter, B. De Moerloose, K. Vandepoele, A. Uyttebroeck, A. Van Damme, B. Denys, L. Dedeken, M-F Dresse, J. Van der Werff Ten Bosch, M. Hofmans, T. Kerre, B. Vandekerckhove, J. Philippé* and T. Lammens*

(*shared senior authorship)

Manuscript accepted for publication in HemaSphere.

Abstract

Introduction.

We previously identified *TARP* as an immunotherapeutic target in AML, also describing a small cohort of pedAML patients. Data on the clinical impact of *TARP* expression, and on *in vivo* recognition of TARP peptide-MHC complexes on leukemic cells, are lacking.

Materials and Methods.

Four wild type AML cell lines were profiled by mRNA sequencing, together with two transgenic TARP-knockdown cell lines. Twenty-four leukemic stem cell and 29 leukemic blasts from 29 pedAML patients, and 25 hematopoietic stem cell and 28 controls blasts from 32 healthy subjects, were evaluated by real-time qPCR for *TARP* expression. For 11/29 patients, lymphocytes were stained by HLA-A*0201-restricted tetramers with TARP(P5L)$_{4\text{-}13}$-epitope specificity.

Results.

mRNA sequencing confirmed that T-cell receptor gamma constant domain 1 and 2 encoded *TARP* transcripts are expressed in AML. TARP-specific cytotoxic T-cells (CTLs) were observed in HLA-A*0201+/*TARP*+ pedAML (n=3), suggesting *in vivo* TARP-specific CTL expansion. *TARP* overexpression was demonstrated in 44.8% (13/29) patients, including 11/11 *FLT3*-ITD mutated pedAML and also 2/18 *FLT3* WT pedAML with poor outcome. High *TARP* expression was significantly associated with a lower event-free survival (EFS) in NOPHO-DBH AML2012-treated patients (n=15/29, 6 *TARP*-high and 9 *TARP*-low, hazard ratio=8.41, P=.015) in univariate, though not in multivariate, analysis.

Conclusion.

TARP transcript expression has a clinical impact in pedAML, as expression significantly inversely correlates with EFS. TARP presentation on leukemic cells activates CTL *in vivo*. A possible role for *TARP* as marker for AML subtypes with poor outcome, irrespective of *FLT3*-ITD mutational status, deserves further attention.

Pediatric acute myeloid leukemia (pedAML) is a rare hematological disease accounting for 20% of all pediatric leukemias [1]. Current chemotherapeutic regimens have reached a survival plateau around 70% [2, 3]. Still 30-40% of the good responders experience relapse, and especially patients with fms-like tyrosine kinase receptor-3 internal tandem duplications (*FLT3*-ITD) show a detrimental outcome [2]. These observations have driven the development of alternate therapeutic strategies, including targeting antibodies and chimeric antigen receptor (CAR)- or T-cell receptor (TCR)-transgenic cytotoxic T-cells (CTLs). However, besides *FLT3* inhibitor-based therapies [4, 5] and CD33-directed agents [6], targeted strategies have not yet found their way into treatment protocols.

We recently identified the TCR γ chain alternate reading frame protein (TARP) as an immunotherapeutic target in leukemic blasts (L-blast) and in leukemic stem cells (LSC) of adults and children with AML [7]. TARP was previously only reported in androgen-sensitive prostate and breast adenocarcinoma [8].

Although *TARP* was upon its discovery described as a truncated TCRγ transcript encoding the first TCR γ chain constant domain (*TRGC1*), we found that an AML-exclusive, *TRGC2*-encoding *TARP* transcript co-exists in AML [7]. The high sequence homology between *TRGC1* and *TRGC2* hampers distinction through conventional techniques. Here, we used mRNA sequencing to demonstrate both *TARP* transcripts in four wild type (WT) AML cell lines with documented *TARP* expression [7]. We confirmed that *TRGC1* and *TRGC2* transcript are highly expressed in MV4;11, next to a moderate expression in HL-60 and THP-1, while negative in OCI-AML3 (Fig. S1A). To gain insight into the biological relevance of both transcripts, transgenic TARP knockdown (TARP-KD) cell lines were generated for 2/4 AML cell lines, HL-60 and MV4;11, by retroviral transduction of TARP-targeting short-hairpin (shRNA) encoding viral particles, next to a mock construct [7]. In MV4;11, both *TRGC1* and *TRGC2* transcripts were suppressed upon *TARP* knockdown. In HL-60, only the *TRGC2* transcript was significantly downregulated compared to mock and WT, while *TRGC1* showed a two-fold decrease (Fig. S1B, Table S1). Altogether, these data confirm that both *TRGC1*- and *TRGC2*-encoded *TARP* transcripts co-exist in AML cell lines, and are targetable.

Our previous data suggested that TARP peptides are adequately MHC-presented, as leukemic cells could be targeted *in vitro* by cytotoxic T-cells (CTLs) retrovirally transduced with a TCR directed against the HLA-A2 enhanced affinity TARP(P5L)$_{4\text{-}13}$ epitope [7]. However, whether recognition of tumor *TARP* peptide-MHC complexes (p-MHCs) by TCR-bearing CTLs is able to trigger antigen-specific immune responses *in vivo* still needs to be elucidated. To this end, TARP-TCR expression on CTLs isolated from pedAML patients was measured using HLA-A*0201-restricted, PE-conjugated tetramers directed against the TARP(P5L)$_{4\text{-}13}$ epitope, kindly provided by the NIH Tetramer Core Facility. Lymphocytes from *TARP*-high (n=5, 3/5 HLA-A*0201 positive) and *TARP*-low (n=6, 5/6 HLA-A*0201 positive) pedAML patients were surface- and tetramer-stained and measured on a FACSCanto II flow cytometer (BD Biosciences).

All three *TARP*-high/HLA-A*0201-positive pedAML patients showed a positive tetramer staining within the CD3+/CD8+ compartment (median 2.4%, median MFI 357) (Fig. 1). MFI values were comparable between BM and PB as evaluated for 2/3 patients. TARP-TCR expression was higher at relapse compared to diagnosis in both PB and BM, as measured in one patient. Higher intensity tetramer staining can be associated with a higher TCR and/or co-receptor expression, and/or to a higher TCR avidity. This observation may relate to an achieved immunity for the tumor epitope, i.e. the presence of resting CD8+ memory T-cells that rapidly reactivate and upregulate TCR expression upon p-MHC re-stimulation [9], or the presence of a T-stem cell pool generated during the initial immune response [10]. If this hypothesis holds true, CTL priming by *TARP* vaccination after/during chemotherapy could

be a therapeutic strategy. More diagnosis-relapse couples are needed to confirm this finding. Tetramer-positive CTLs from pedAML patients showed a fivefold lower TARP-TCR expression than the positive control (median MFI = 1793). No tetramer-positive population could be measured in the CTL compartment from *TARP*-high/HLA-A*0201 negative patients (n=2), *TARP*-low/HLA-A*0201 positive patients (n=5) or *TARP*-low/HLA-A*0201 negative (n=1) patients.

Altogether, these data suggest that a pedAML patient's native immune response is triggered by HLA-presentation of *TARP* antigenic peptides *in vivo*, but, apparently insufficiently to eradicate the leukemic cells. Both leukemic cell resistance and lymphocyte quiescence may account for this finding [11]. Also, due to T-cell ignorance, tumor-specific CTLs may be present but not primed by the antigen, or priming may be inefficient [12]. Gaining a deeper understanding of the underlying mechanisms may contribute to therapeutic targeting of TARP in pedAML.

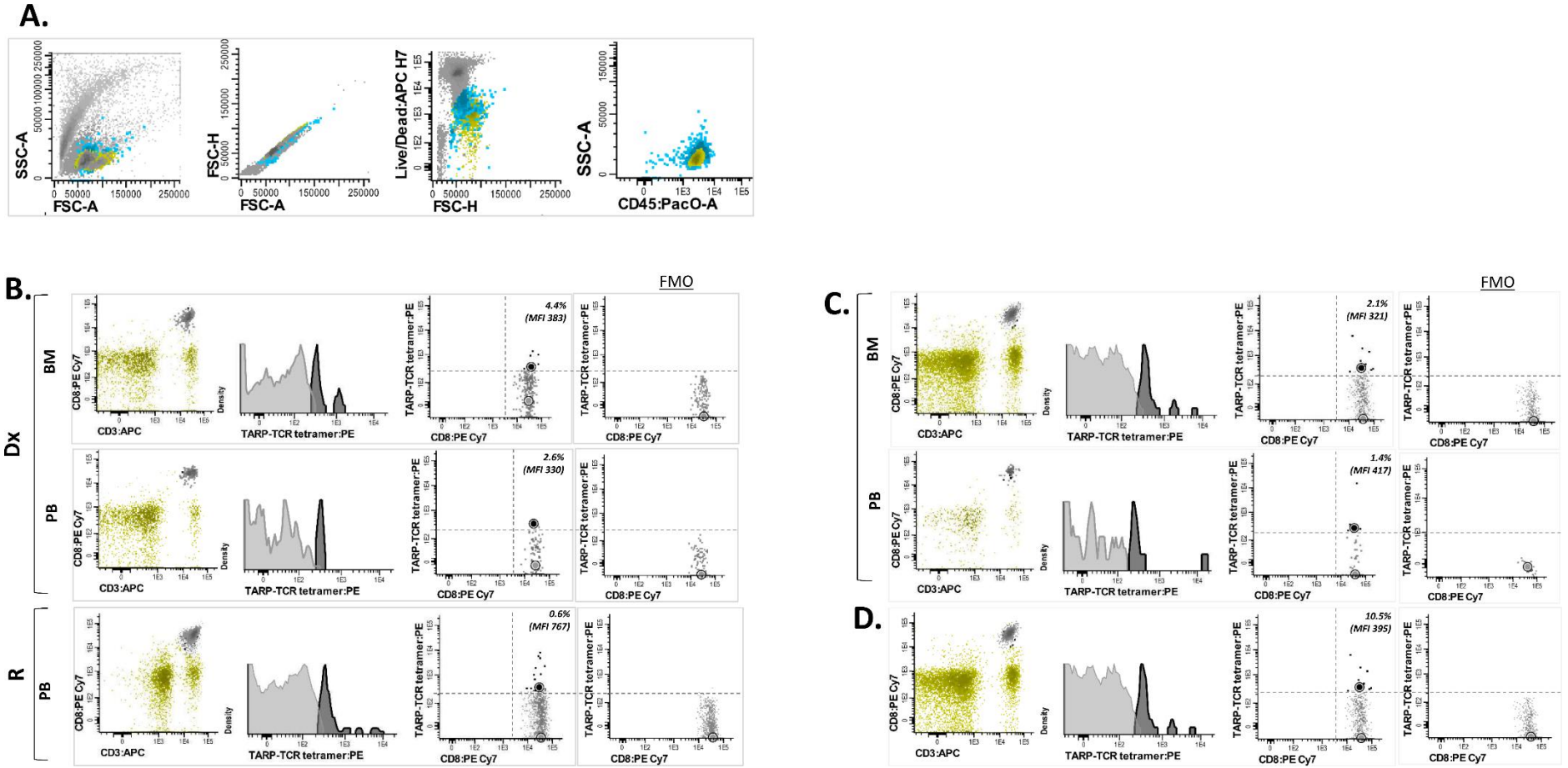


Fig. 1. Tetramer staining of TARP-TCR+ CTLs from pedAML patients.

The lymphocyte gating strategy, described in Supplemental Materials, is illustrated in (A). The total lymphocyte compartment is indicated in green, other white blood cells in blue en non-viable cells or doublets in dark grey. Subsequent patient-individual gating of the CD3+/CD8+ compartment, and tetramer staining within this compartment, is shown for three *TARP*-high/HLA-A*0201 positive patients in (B-D). Tetramer-positive events, defining TARP-TCR+ CTLs, are indicated in black, and tetramer-negative CTLs in light grey. Non-CD3+/CD8+ cells within the lymphocyte gate are indicated in green. Median MFI values are indicated by circles. Tetramer positivity was gated for each patient individually based on sample-specific FMO controls. For two out of the three patients (B-C), both BM and PB were analyzed. For one of the three patients (B), lymphocytes from both diagnosis and relapse could be evaluated.

TARP, T-cell receptor γ chain alternate reading frame protein; TCR, T-cell receptor; CTL, cytotoxic T-cell; FMO, fluorescence-minus-one; BM, bone marrow; PB, peripheral blood.

We previously showed that *TARP* expression is associated with *FLT3*-ITD in pedAML. Unfortunately, the initial pediatric sample cohort was too small to evaluate a possible clinical impact of *TARP* expression. We here evaluated *TARP* expression in a larger cohort of LSC (n=24) and L-blast (n=29) cells sorted from pedAML patients using real-time quantitative PCR (qPCR), and compared expression levels to those measured in HSC (n=25) and C-blast (n=28) sorted from healthy controls. Data analysis is described in Supplemental, and expression values were expressed as calibrated normalized relative quantities (CNRQ).

A significantly increased *TARP* expression (P<.0001) was demonstrated in LSC and L-blast compared to their healthy counterparts (Fig. S2A). PedAML patients were dichotomized as *TARP*-high (n=13/29, 44.8%) and *TARP*-low (n=16/29, 55.2%), using a cut-off based on the average expression measured in healthy controls plus two times the standard deviation (Table S2). The presence of translocations, including core-binding factor leukemia, was significantly inversely correlated to *TARP* expression (P<.01 and <.0001, respectively). *FLT3*-ITD mutations (P<.0001) and HR profiles (P<.01) were exclusively observed in *TARP*-high patients (Fig. 2A). Within the *TARP*-high group, a significantly higher proportion of patients showed WBC counts >30x10^9/L and blasts >70% in BM and >50% in PB (P<.05).

In concordance with our previous work, *TARP* transcript expression was significantly increased in LSC and L-blast (P<.0001) from *FLT3*-ITD pedAML (n=11) compared to *FLT3* WT pedAML (n=18) (Fig. 2B). Sixty-three percent of *FLT3*-ITD positive patients harbored a single ITD, 27% two ITDs and one patient presented four ITDs (Table S3). The length of the duplicated region ranged between 20 and 96 base pairs (bp) (median 33 bp), and allelic ratios (ARs) varied from 2.7 to 70.8% (median 17.7%). There was no significant association between the number of ITDs and the level of *TARP* expression. One patient presented only a single *FLT3*-ITD clone with an AR of 3.0%, but was still classified as *TARP*-high. Paired comparison of *TARP* expression measured in LSCs and L-blasts sorted from 9/11 *FLT3*-ITD mutated pedAML demonstrated a significantly higher expression in the latter compartment (P=.041, Fig. S2B). This finding is in agreement with our published micro-array data of paired LSC and L-blast couples (GSE 128103).

Two *FLT3* WT pedAML patients showed elevated *TARP* expression in LSCs and L-blasts (pedAML18 and pedAML25). The presence of *TRGC* rearrangements, able to confound *TARP* expression due to common TRGC coding regions and sporadically observed in AML [13], was excluded. *FLT3*-ITD analysis was repeated on L-blast subpopulations to exclude a possible false negative result (not performed in LSC due to too low DNA concentration). One patient (pedAML18) harboured a rare *KMT2A-SEPT9* fusion protein, presented central nerve system (CNS) invasion and was classified as HR. The other patient showed a normal karyotype and *WT1* overexpression. Both *TARP*-high/*FLT3* WT patients with and without *KMT2A-SEPT9* relapsed after 8.4 months and presented with resistant disease, respectively. The first patient died 15.1 months after diagnosis.

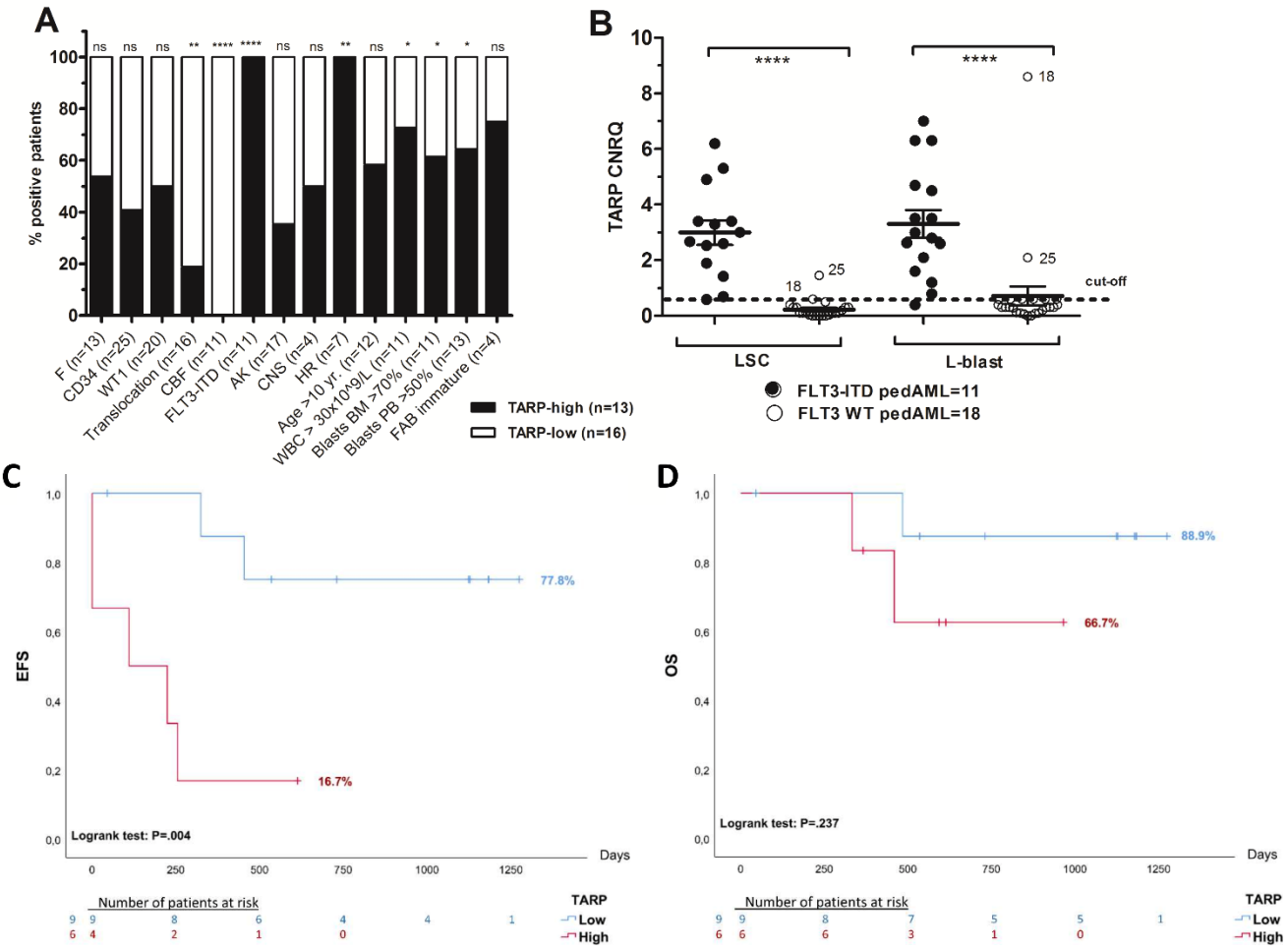


Fig. 2. *TARP* transcript expression in pedAML in relation to subgroups and outcome.
Correlation between patient characteristics and outcome between pedAML patients dichotomized as *TARP*-low (n=16) and *TARP*-high (n=13). *TARP* expression was measured by qPCR, and CNRQ values were interpreted against a cut-off calculated based on the expression in healthy controls (see Supplemental Materials). P-values <.05 were considered as significant. One, two, three or four asterisks are indicative for the level of significance (P<.05, P<.01, P<.001 and P<.0001, respectively).
(A) Bars display the percentage of patients (%), harbouring the characteristic shown in the x-axis, for *TARP*-high (black) and *TARP*-low (white) pedAML. The total number of patients positive for each characteristic is shown between parentheses.
(B) Differential *TARP* expression between *FLT3*-ITD mutated and *FLT3* WT pedAML patients measured in the LSC and L-blast compartment. *FLT3*-ITD mutated pedAML showed a significantly higher *TARP* expression in both LSC (P<.0001) and L-blast (P<.0001). Thirteen out of the 29 pedAML patients were classified as *TARP*-high, i.e. 11/11 *FLT3*-ITD pedAML and 2/18 *FLT3* WT pedAML (encoded by "18" and "25"). Horizontal bars indicate means, error bars indicate ±SEM, horizontal square brackets represent statistical comparisons and the dotted line represent the cut-off for elevated TARP expression.
(C-D) Kaplan-Meier EFS and OS survival plots based on 15 pedAML treated in the NOPHO-DBH AML2012 protocol, dichotomized as TARP-high (n=6, 4/6 *FLT3* ITD and 2/6 *FLT3* WT) or TARP-low (n=9, 9/9 *FLT3* WT). The number of days is shown on the x-axis, and the percentage as a ratio (100% equals 1.0) on the y-axis. Drop-outs of the patients are indicated at the bottom per block of 250 days. **(C)** EFS was significantly lower in *TARP*-high versus *TARP*-low patients (16.7% versus 77.8%, respectively, P<.01). **(D)** OS was lower in *TARP*-high versus *TARP*-low patients (66.7% versus 88.9%, respectively), though at a non-significant level (P>.05).
PedAML, pediatric acute myeloid leukemia; qPCR, quantitative polymerase chain reaction; LSC, leukemic stem cell; L-blast, leukemic blast; F, female; M, male; WBC, white blood cell; BM, bone marrow; PB, peripheral blood; TARP, T-cell receptor γ chain alternate reading frame protein; FAB, French-British-American; *FLT3*, fms-like tyrosine kinase receptor-3; *NPM1*, nucleophosmin; *CEBPA*, CCAAT/enhancer-binding protein alpha; ITD, internal tandem duplication; WT, wild type; MT, mutated; *WT1*, Wilms' tumor 1; CNS, central nerve system; SR: standard risk; HR: high risk; AK, abnormal karyotype; NK, normal karyotype; SEM, standard error of the mean; CNRQ, calibrated normalized relative quantity; yr., years.

Among NOPHO-DBH AML2012-treated patients (n=15), *TARP*-high pedAML (n=6) showed a significantly lower event-free survival (EFS) compared to *TARP*-low pedAML (n=9) (16.7% versus 77.8%, respectively, logrank P<.01, Fig. 2C). Relapse occurred in 3/6 *TARP*-high patients and 2/6 showed resistant disease, with an estimated time to event of 6.6 months. Two of the nine *TARP*-low patients also relapsed, with an estimated time to event of 34.6 months. Univariate Cox regression analysis confirmed this finding, showing a significant hazard ratio of 8.41 (95% confidence interval (CI) 1.52 – 46.6, P=.015) for the occurrence of an event in *TARP*-high patients. Multivariate analysis did not show an association between *TARP* expression and EFS (Table S4). However, association with EFS did remain significant when including all diagnostic pedAML patients, irrespective of the treatment protocol (n=27/29, hazard ratio 3.83 (95% CI 1.1 – 13.0), P=.032). No significant correlation was found between the level of *TARP* expression and OS, (hazard ratio 3.90 (95% CI 0.35-43.9), P=.27, Fig. 2D). Although data need to validated in a larger, preferentially multicentre cohort, these findings suggest that *TARP* expression negatively impacts outcome in pedAML.

In conclusion, we here confirm that both *TRGC1*- and *TRGC2*-encoded *TARP* transcripts co-exist in AML. We demonstrate that TARP presentation on leukemic cells may induce beneficial immune responses in pedAML patients. We consolidate our previous finding that all *FLT3*-ITD positive pedAML patients display *TARP* overexpression, though also conclude that *TARP* overexpression is not exclusive for *FLT3*-ITD mutated patients. Furthermore, investigation on the role of FLT3 inhibitors in *TARP*-high pedAML patients, and their impact on *TARP* expression levels, is warranted. *TARP* expression was significantly inversely correlated with EFS in a small cohort of NOPHO-DBH AML2012-treated patients. The hypothesis that *TARP*+/*FLT3*-WT pedAML patients may define a (till now undetectable) poor prognosis group with HR genetic lesions (*KMT2A-SEPT9*) and poor outcome will require further evaluation. Although promising, these data need confirmation in larger, preferentially multicenter cohorts.

Authorship

BDP, JP and TL participated in research design. BDP, JP, TL, BDM, MH, KV, LD, BVK and TK participated in the writing of the paper. BDP performed the research and the data analysis. KV and BDN contributed the molecular analyses. TK, BVK, AU, AV, LD, MFD, JV contributed reagents or samples. Each author has reviewed the manuscript, believes it represents valid work, and approves it for submission.

Disclosures

All authors declare to have no possible conflicts of interest in the manuscript, including financial, consultant, institutional and other relationships that might lead to bias or a conflict of interest.

Funding

This research was supported by the Belgian Foundation against Cancer (grant 2014–265), FOD-KankerPlan (Actie29, grant to JP), vzw Kinderkankerfonds (grant to TL) and the Research Foundation - Flanders (Fonds voor Wetenschappelijk Onderzoek Vlaanderen, FWO, grant 1113117 to BD and grant 1831312N to TK).This work is submitted in partial fulfilment of the requirement for the PhD of candidate BD at Ghent University.

Acknowledgement

Our gratitude goes to prof. dr. F. Plasschaert of the orthopedic surgery, C. Matthys of the Cord Blood Bank, and the staff of the Department of Pediatric Haematology and Oncology of the Ghent University Hospital (Ghent, Belgium) for providing samples. We acknowledge the NIH Tetramer Core Facility for providing tetramers. We acknowledge the Nucleomics Core for performing mRNA sequencing and data analysis (www.nucleomics.be).

References

1. Rasche M, Zimmermann M, Borschel L, Bourquin JP, Dworzak M, Klingebiel T, Lehrnbecher T, Creutzig U, Klusmann JH, Reinhardt D: Successes and challenges in the treatment of pediatric acute myeloid leukemia: a retrospective analysis of the AML-BFM trials from 1987 to 2012. Leukemia 2018, 32(10):2167-2177.
2. De Moerloose B, Reedijk A, de Bock GH, Lammens T, de Haas V, Denys B, Dedeken L, van den Heuvel-Eibrink MM, Te Loo M, Uyttebroeck A et al: Response-guided chemotherapy for pediatric acute myeloid leukemia without hematopoietic stem cell transplantation in first complete remission: Results from protocol DB AML-01. Pediatric blood & cancer 2019:e27605.
3. Zwaan CM, Kolb EA, Reinhardt D, Abrahamsson J, Adachi S, Aplenc R, De Bont ES, De Moerloose B, Dworzak M, Gibson BE et al: Collaborative Efforts Driving Progress in Pediatric Acute Myeloid Leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2015, 33(27):2949-2962.
4. Annesley CE, Brown P: The Biology and Targeting of FLT3 in Pediatric Leukemia. Frontiers in oncology 2014, 4:263.
5. Levis M: FLT3 mutations in acute myeloid leukemia: what is the best approach in 2013? Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program 2013, 2013:220-226.
6. Kaspers GJ: Pediatric acute myeloid leukemia. Expert review of anticancer therapy 2012, 12(3):405-413.
7. Depreter B, Weening KE, Vandepoele K, Essand M, De Moerloose B, Themeli M, Cloos J, Hanekamp D, Moors I, I DH et al: TARP is an immunotherapeutic target in acute myeloid leukemia expressed in the leukemic stem cell compartment. Haematologica 2019, 104:xxx(doi: 10.3324/haematol.2019.222612. [Epub ahead of print]).
8. Essand M, Vasmatzis G, Brinkmann U, Duray P, Lee B, Pastan I: High expression of a specific T-cell receptor gamma transcript in epithelial cells of the prostate. Proceedings of the National Academy of Sciences of the United States of America 1999, 96(16):9287-9292.
9. Lefrancois L, Obar JJ: Once a killer, always a killer: from cytotoxic T cell to memory cell. Immunological reviews 2010, 235(1):206-218.
10. Gattinoni L, Speiser DE, Lichterfeld M, Bonini C: T memory stem cells in health and disease. Nature medicine 2017, 23(1):18-27.
11. Coulie PG, Connerotte T: Human tumor-specific T lymphocytes: does function matter more than number? Current opinion in immunology 2005, 17(3):320-325.
12. Gajewski TF, Schreiber H, Fu YX: Innate and adaptive immune cells in the tumor microenvironment. Nat Immunol 2013, 14(10):1014-1022.
13. Manara E, Basso G, Zampini M, Buldini B, Tregnago C, Rondelli R, Masetti R, Bisio V, Frison M, Polato K et al: Characterization of children with FLT3-ITD acute myeloid leukemia: a report from the AIEOP AML-2002 study group. Leukemia 2017, 31(1):18-25.

Supplemental materials and methods

Patient characteristics

Bone marrow (BM) and/or peripheral blood (PB) samples from 29 pedAML patients were evaluated for *TARP* expression (Table S2). Patients were selected based on sample availability (>50x10^6 cells) and CD34 expression (≥1%). Patients were treated according to the DB-AML01 protocol (n=9), NOPHO-DBH AML2012 protocol (n=15) or treated otherwise (n=5).

In addition, samples were collected from healthy subjects. Pediatric normal bone marrow (NBM, n=13, 12-18 yr.) was collected from posterior iliac crest during scoliosis surgery. Cord blood (CB, n=15) was obtained after full-term vaginal deliveries. Mobilized peripheral blood stem cells (mPBSCs) were collected from apheresis of adult donors pre-allotransplant (n=4). All healthy subjects and patients and/or their guardians gave informed consent. Approval was issued by the ethical committee in accordance with the declaration of Helsinki (EC2015-1443 and EC2019-0294).

Treatment protocols

De novo pediatric (< 18 yrs.) patients diagnosed with AML between 2010 and 2013 were included in the DB-AML01 trial (EudraCT 2009-014462-26). The treatment protocol of the DB-AML01 study is described elsewhere [1].

Patients diagnosed between 2015 and 2019 were included in the NOPHO-DBH AML2012 trial (EudraCT 2012-002934-35) unless they did not fulfill the inclusion criteria. NOPHO-DBH AML 2012-treated patients received two intensive induction courses, followed by risk-adapted consolidation with three courses of conventional chemotherapy for standard risk (SR) patients and allogeneic stem cell transplantation for high-risk (HR) patients. Patients with isolated central nervous system (CNS) or extramedullary leukemia, previous chemo- or radiotherapy, AML secondary to a previous bone marrow failure syndrome, myeloid leukemia in Down syndrome with age <5 or ≥5 yrs. with GATA1 mutation, acute promyelocytic leukemia (APL), juvenile myelomonocytic leukemia (JMML), myelodysplastic syndrome, Fanconi anemia and/or positive pregnancy test were excluded. Based on these exclusion criteria, 2/27 diagnostic patients i.e. 1 APL, 1 secondary AML evolved from JMML, were treated otherwise, and 1/27 patients (17 yrs.) was treated according to a similar adult protocol.

The definition of HR depends on the treatment protocol. In the DB-AML01 study [1], patients were considered as HR if ≥ 15% blasts persisted after the first induction course and ≤ 5% blasts after the second course (≥5% blasts after the second course was defined as refractory disease). In the NOPHO-DBH AML2012 study, patients were defined as HR if they achieved CR after two induction courses and had (i) *FLT3*-ITD/*NPM1* WT profiles, (ii) poor response after induction 1 (i.e. ≥15% leukemic cells at day 22 or at any subsequent evaluation prior to course 2) or (iii) intermediate response after induction 2 (i.e. 0.1%-4.9% leukemic cells before consolidation) [2].

Outcome definitions

Otherwise-treated patients (n=3), relapsed patients (n=2) and patients treated in the DB-AML01 study (n=9), were excluded from outcome analysis. For estimates of EFS, an event was defined as failure to achieve complete remission (CR), induction death, relapse, development of a second malignancy, or death due to any cause, whichever occurred first. EFS was calculated from date of diagnosis to the date of first event, with failure to achieve CR calculated as an event at t = 0. OS was calculated from date of diagnosis to the date of last follow-up or time of death due to any cause. Follow-up time was censored at the last follow-up visit if no failure was observed.

Cell-sorting of patients samples and healthy controls

Mononuclear cells (MNCs) were isolated from pedAML patients and healthy controls by Ficoll density gradient (Axis-shield). A CD34-enrichment procedure was performed if the number of CD34-positive cells was <50% (CD34 MicroBead Kit, Milteny). MNCs were spinoculated (5 min, 1500 rpm) followed by the addition of monoclonal antibodies (mAb) and incubated in the dark at room temperature (RT) for 20 min (Table S5). After incubation, cell pellets were washed with PBS+2% BSA (5 min, 1500 rpm), resuspended in 50%RPMI/50%FCS and used for cell-sorting on a FACSAria III (BD Biosciences).

All scatters were devoid of cell debris and doublets based on propidium iodide exclusion and FSC-H versus FSC-A plots, respectively. The immature myeloid compartment was defined by CD34, CD45 and scatter properties. CD34+/CD38+ and CD34+/CD38- compartments were gated to isolate control blasts (C-blasts) and hematopoietic stem cells (HSCs) from healthy controls, or leukemic blasts (L-blasts) and leukemic stem cells (LSC) from pedAML patients, respectively. Fluorescence-minus-one (FMO) controls were used to determine the CD38 expression cut-off. Lymphocytes were sorted based on high CD45 expression and low SSC-A properties. Delineated cell populations were backgated on FSC-A/SSC-A and CD45/SSC-A scatter plots to exclude non-specific events and other myeloid precursor populations. Post-sort purities exceeded 90%. Sorted cells were collected in 50%RPMI/50%FCS, spun down (10 min, 3000 rpm, 4° C) and either resuspended in 700 µL TRIzol for RNA extraction (LSCs and L-blasts from pedAML patients, HSCs and C-blasts from healthy controls), or cryopreserved in 90% FCS/10% dimethylsulfoxide (lymphocytes from pedAML patients).
Cell-sorting of CD34+/CD38+ and CD34+/CD38- cells from 29 pedAML patients yielded a total number of 29 L-blast and 24 LSC fractions, respectively. In addition, CD34+/CD38+ sorting of NBM (n=13) and CB (n=15) yielded a sufficient number of C-blasts for RNA isolation (n=28). The number of sorted HSCs was only sufficient in 13/15 CB and 7/13 NBM samples, and therefore complemented with four CD34+/CD38- fractions sorted from mobilized peripheral blood (mPBSC=4, total of 25 HSC fractions).

Tetramer staining

An in-house developed tetramer and membrane staining protocol was used, based on previous literature and in-house optimization. The volume of cells needed to stain 50 000 cells was isolated and washed with 500 µL Hank's Balanced Salt Solution buffer (HBSS, GIBCO®) supplemented with 1%FCS (1500 rpm, 5 min). After spinoculation, supernatant was removed, pellets were resuspended at a final dilution of 50 000 cells in 30 µL and stained with 5 µL 1/75 diluted PE-label tetramers, or 5 µL HBSS/1%FCS in case of FMO controls, for 30 min at RT in the dark. Surface staining immediately proceeded tetramer labeling, using an antibody cocktail of mAb against CD3, CD8, CD45 and HLA-A2 combined with a LIVE/DEAD (L/D) staining (Table S5). Samples were further incubated for 20 min at RT in the dark, washed with 1 mL HBSS/1%FCS medium (1500 rpm, 5 min) and subsequently measured on a FACSCanto II (BD Biosciences, San Jose, CA, USA) flow cytometer, equipped with three solid-state lasers, and instrument set-up performed strictly according to EuroFlow guidelines [3]. Retroviral transduced TARP-TCR transgenic T-cells, manufactured as previously described [4], were used as positive control.

Gating was performed using Infinicyt software v.1.8 (Cytognos, Salamanca, Spain). All scatters were devoid of doublets based on FSC-H/FSC-A, and living cells were selected based exclusion of the L/D marker. Lymphocytes were selected based on SSClow/FSClow/CD45++ expression, and cytotoxic T-cells (CTLs) were subsequently gated based on CD3+/CD8+ expression. Tetramer positivity was judged based on sample-specific FMO controls. Delineated tetramer-positive events were backgated on FSC-

A/SSC-A and CD45/SSC-A scatter plots to exclude non-specific events. All samples were analyzed in duplicate (technical duplicates), and 5/12 samples (n=4/11 pedAML patients) were analyzed in two different experiments, starting from the same batch of cryopreserved cells (biological duplicates).

mRNA sequencing experimental settings and data analysis

Wild-type (WT) cell lines MV4;11, HL-60, THP-1 and OCI-AML3 were purchased at ATCC or DMSZ, and grown in RPMI medium (Invitrogen) supplemented with 10% or 20% Fetal Calf Serum (FCS, Hyclone or ThermoFisher Scientific), according to supplier instructions, together with 100 U/mL Penicillin/Streptomycin (10000 U/ml, Invitrogen) and 100 µg/mL L-Glutamine (200 mM, Invitrogen). For THP-1, medium was additionally supplied with 0.05 mM β-mercaptoethanol. Cell lines were incubated at 37 °C in 5% CO_2 incubators.

Total RNA was extracted from $1x10^6$ wild-type and transgenic cell lines using the miRNeasy Mini or Micro Kit (Qiagen) in combination with on-column DNase I digestion (RNase-Free DNase set, Qiagen). For each condition (wild-type, knockdown and mock) and cell line (OCI-AML3, HL-60, THP-1 and MV4;11), three biological replicates were used, except for the mock where two biological replicates were used for technical reasons. RNA concentrations were measured by Nanodrop and fulfilled the following criteria to pass quality control: A260/A280 ratio [1.8 - 2.1] and A260/A230 ratio [1.8 -2.0], as measured by Nanodrop, and RNA integrity number (RIN) >8.0, as measured by Agilent 2100 Bioanalyzer. RIN ranged between 8.2 – 10.0, with a median value of 9.9 (95% CI 9.8 – 9.9).
Library prep and sequencing was performed by the VIB Nucleomics Core Leuven on a HiSeq 3000/HiSeq 4000 System, using the TruSeq and Illumina NextSeq500 High Output 75 kits, respectively. Pools were analyzed in two NextSeq500 single-end 75bp runs, increasing the number of pass filtered (PF) reads (~800M PF reads + single-end 75bp reads), yielding very high quality data (>95% Avg%Q30). The data of raw counts were merged with the reference gene Homo sapiens Ensembl.GRCh38.88 annotations. Genes with less than one counts-per-million (absent genes) were removed, retaining 17585 genes. A within- and between-sample normalization was performed to avoid sample-specific effects. Within samples, GC-content was corrected by full quantile normalization on bins of GC-content (EDASeq package, Bioconductor). Between samples, a correction for library size and RNA composition was performed by full quantile normalization (EDASeq package, Bioconductor). Normalized counts were divided by the total number of counts (in millions), and the obtained scaled counts were divided by the gene length (in kbp) in order to obtain the number of Fragments Per Kilobase of gene sequence and per Million fragments of library size (FPKM).
Log2 fold change (FC) values were calculated for each pairwise comparison of interest based on the FPKM values. We adopted the criterion from the MAQC-I study and considered P-values <.001 as significant [5].

qPCR experimental settings and data analysis

RNA was extracted from sorted cells, from pedAML patients or from healthy controls, using the miRNeasy Mini or Micro Kit (Qiagen) in combination with on-column DNase I digestion (RNase-Free DNase set, Qiagen) according to manufacturer's instructions. RNA concentrations were measured by Nanodrop (ThermoFisher Scientific) or Qubit RNA HS Assay (Invitrogen). For patients' cell-sorted fractions, cDNA synthesis was performed after an additional in-solution gDNase elimination step (Heat&Run gDNA removal kit, ArcticZymes), using the 5x PrimeScriptTM RT Master Mix (Takara Bio Europe S.A.S.) in a final volume of 12.5 µL. For cell lines, cDNA synthesis was performed by Invitrogen SuperScript III Reverse Transcriptase (Invitrogen) according to the supplier`s recommendations. cDNA

from both patients and cell lines was diluted until a final concentration of 2.38 ng cDNA/µL. qPCR reactions were carried out in 96-well plates using 0.3 µM primers, 2x Takyon Low ROX SYBR 2X MasterMix (Eurogentec), 2.38 ng cDNA and H_2O (Sigma-Aldrich) in a 10 µL reaction. Samples were run in duplicate after a heat-activation step (3 min 95 °C) by a 2-step real-time protocol of 45 cycles (95 °C 15 sec, 60 °C 60 sec) on a Viia7 analyzer (ThermoFisher), combined with melting curve analysis (65 °C to 95 °C, gradually increasing with 0.5 °C/5 sec). Ct thresholds were automatically determined by the QuantStudio™ Real-Time PCR Software.

Primer sequences for *TARP* and housekeeping genes (*GAPD*, *HPRT1* and *TBP*) were previously described [4]. Ct values generated for each target were corrected for primer pair efficiency and expressed as relative quantities (RQ). Normalized relative quantities (NRQ) were calculated by normalising RQ values against the expression of housekeeping genes *GAPD*, *HPRT1* and *TBP* (NRQ). To allow inter-run comparison, calibrated NRQ values (CNRQ) were generated by taking into account the expression of a single inter-run calibrator (IRC), evaluated in each run by the respective primer pair. A cut-off for *TARP* overexpression was based on the average expression plus two standard deviations measured in the normal counterparts (HSC=25 and C-blast=28, respectively), and set to 0.59.

Statistical analysis

Graphics and statistical calculations were made in GraphPad Prism (version 5.04, La Jolla California USA) and SPSS (version 25.0.0.2, Inc., Chicago, IL). The Mann-Whitney U test was used to evaluate significant differential *TARP* expression between *FLT3*-ITD and *FLT3* WT pedAML patients. Association between *TARP* expression and dichotomous variables were evaluated by the Pearson's two-sided Chi-Square test if the expected count was ≥5, and the Fisher's exact test (two sided) otherwise.

The Kaplan Meier method was used to estimate the survival probabilities for EFS and OS. Univariate regression analysis was performed by the Kaplan-Meier log-rank test, and confirmed by the univariate COX regression log-rank test if significant, also used to calculate hazard ratios. Confirmation of significance in univariate models was performed by a multivariate Cox proportional hazard model, taking into account all aforementioned continuous and dichotomous variables.

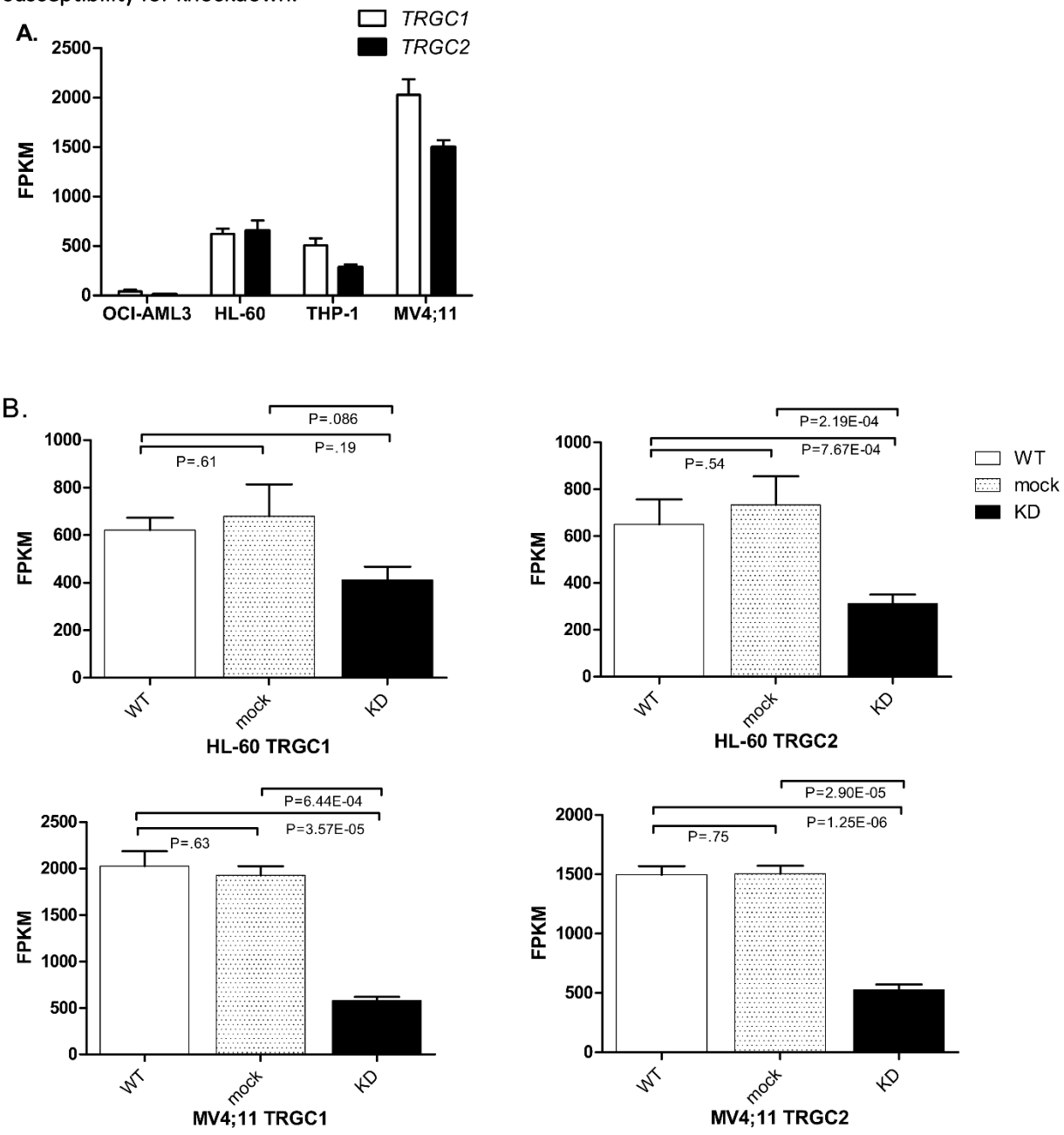
P-values <.05 were considered significant, and the number of asterisks indicate the level of significance (one, P<.05; two, P<.01; three, P<.001 and four, P<.0001).

Supplemental references

1. De Moerloose B, Reedijk A, de Bock GH, Lammens T, de Haas V, Denys B, Dedeken L, van den Heuvel-Eibrink MM, Te Loo M, Uyttebroeck A *et al*: Response-guided chemotherapy for pediatric acute myeloid leukemia without hematopoietic stem cell transplantation in first complete remission: Results from protocol DB AML-01. *Pediatric blood & cancer* 2019:e27605.
2. Chair) JAS: Research study for treatment of children and adolescents with acute myeloid leukaemia 0-18 years. . *NOPHO-DBH AML 2012 protocol v21* 2013, EUdract number 2012-002934-35.
3. Kalina T, Flores-Montero J, van der Velden VH, Martin-Ayuso M, Bottcher S, Ritgen M, Almeida J, Lhermitte L, Asnafi V, Mendonca A *et al*: EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. *Leukemia* 2012, 26(9):1986-2010.
4. Depreter B, Weening KE, Vandepoele K, Essand M, De Moerloose B, Themeli M, Cloos J, Hanekamp D, Moors I, I DH *et al*: TARP is an immunotherapeutic target in acute myeloid leukemia expressed in the leukemic stem cell compartment. *Haematologica* 2019.
5. Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, Collins PJ, de Longueville F, Kawasaki ES, Lee KY *et al*: The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 2006, 24(9):1151-1161.

Supplemental Figures

Figure S1. Confirmation of *TRGC1*- and *TRGC2*-encoding *TARP* transcripts in AML cell lines and their susceptibility for knockdown.

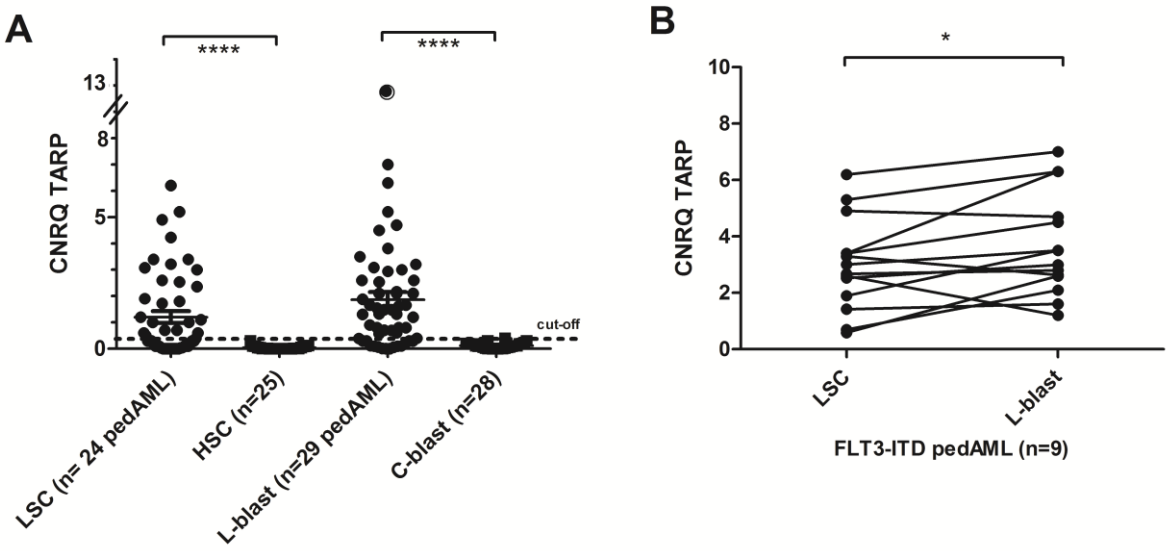


TRGC1 and *TRGC2* transcript expression was measured by mRNA sequencing and expressed as Fragments Per Kilobase per Million fragments of library size (FPKM). The log2FC values between cell lines, with accompanying P-values (significant if <.001), are shown per *TRGC* transcript in Table S1. *TRGC1*, TCR γ chain constant domain 1; *TRGC2*, TCR γ chain constant domain 2; *TARP*, T-cell receptor γ chain alternate reading frame protein; WT, wild type. KD, knockdown; FC, fold change; AML, acute myeloid leukemia.

(A) Differential *TRGC1* and *TRGC2* expression computed by mRNA sequencing in four WT AML cell lines, i.e. MV4;11, THP-1, HL-60 and OCI-AML3.

(B) Differential *TRGC1* and *TRGC2* expression upon TARP-knockdown (KD), compared to the respective WT and mock, for HL-60 (top) and MV4;11 (bottom).

Figure S2. Differential *TARP* expression between leukemic and normal stem cell and blast populations.



(A) Evaluation of *TARP* expression in 24 LSC versus 25 HSC (13 CB, 8 NBM and 4 mPBSC), and 29 L-blast versus 28 C-blast (15 CB, 13 NBM). Approximately half of the pedAML patients (n=13/29, 44.8%) showed a significant increased *TARP* expression (P<.0001) in LSC and L-blast compared to their healthy counterparts. Cut-off is indicated by a dotted line. Y-axis is interrupted between CNRQ values 8 and 13, indicated by double brackets, to allow better comparison with (B).
(B) Paired comparison of *TARP* expression measured in LSC and L-blast sorted from 9/11 *FLT3*-ITD mutated pedAML demonstrated a marginal significant higher expression in the latter compartment (P=.041).
PedAML, pediatric acute myeloid leukemia; *FLT3*, fms-like tyrosine kinase receptor-3; ITD, internal tandem duplication; T-cell receptor γ chain alternate reading frame protein; CNRQ, calibrated normalized relative quantity; LSC, leukemic stem cell; HSC, hematopoietic stem cell; L-blast, leukemic blast; C-blast, control myeloblast; CB: cord blood; NBM, normal bone marrow; mPBSC, mobilised peripheral blood stem cells.

Supplemental Tables

Table S1. Differential TRGC1 and TRGC2 transcript expression between WT and transgenic TARP-knockdown AML cell lines.

Pairwise comparison	*TRGC1*		*TRGC2*	
	log2FC	**P**	**log2FC**	**P**
MV4;11 WT vs HL-60 WT	**2.19**	**6.77E-05**	**1.34**	**4.41E-05**
MV4;11 WT vs OCI-AML3 WT	**6.27**	**5.06E-21**	**7.35**	**2.48E-58**
MV4;11 WT vs THP-1 WT	**2.43**	**1.21E-05**	**2.50**	**3.05E-13**
OCI-AML3 WT vs HL-60 WT	**-4.08**	**1.45E-11**	**-6.01**	**1.22E-42**
OCI-AML3 WT vs THP-1 WT	**-3.84**	**1.32E-10**	**-4.85**	**7.74E-30**
THP-1 WT vs HL-60 WT	-0.24	6.53E-01	**-1.16**	**4.65E-04**
MV4;11 TARP-KD vs mock	**-1.99**	**6.44E-04**	**-1.49**	**2.90E-05**
MV4;11 TARP-KD vs WT	**-2.28**	**3.57E-05**	**-1.61**	**1.25E-06**
MV4;11 mock vs WT	-0.29	6.30E-01	-0.11	7.54E-01
HL-60 TARP-KD vs mock	-1.00	8.58E-02	**-1.33**	**2.19E-04**
HL-60 TARP-KD vs WT	-0.70	1.86E-01	**-1.11**	**7.67E-04**
HL-60 mock vs WT	0.30	6.10E-01	0.22	5.44E-01

Four WT AML cell lines were subjected to mRNA sequencing, together with their respective TARP-KD and mock transgenic cell line variants for HL-60 and MV4;11. For each pairwise comparison of interest, log2FC values with accompanying P-values are shown. Significant differential log2FC values (P<.001) are indicated in bold. FC, fold change; KD, knockdown; WT, wild-type; TARP, T-cell receptor γ chain alternate reading frame protein.

Table S2. Demographics pedAML patients used for *TARP* expression evaluation by qPCR.

PedAML patient characteristics (n=29)	***TARP*-high (n=13)**		***TARP*-low (n=16)**		**P***
	Mean (Range)		Mean (Range)		
Age, years	11.7 (7-17)		8.1 (1-16)		0,22
WBC count, x 10^9/L	109.5 (2.7-336)*		25.7 (3.1-118)		**0,02**
Morphological blast count					
BM, %	77.8 (31-96)†		61.8 (27-88)		**0,03**
PB, %	58.7 (9-95)		39.2 (1-78)		**0,04**
	N	%	N	%	
Gender					0,38
F	7	53,8%	6	37,5%	
M	6	46,2%	10	62,5%	
Time point					
Dx	12	92,3%	15	93,8%	
R	1	7,7%	1	6,3%	
Sample					
BM and PB couples	4	30,8%	7	43,8%	
Only BM	4	30,8%	8	50,0%	
Only PB	5	38,5%	1	6,3%	
Primary/secondary AML					
Primary	13	100,0%	15	93,8%	
Secondary	0	0,0%	1	6,3%	
Treatment protocol					
DB AML-01	4	30,8%	5	31,3%	
NOPHO-DBH AML2012	6	46,2%	9	56,3%	
Other (Dx)	2	15,4%	1	6,3%	
Other (R)	1	7,7%	1	6,3%	
CD34 positivity	11	84,6%	16	100,0%	0,19
WT1 overexpression					0,45
Yes	10	76,9%	10	62,5%	
No	3	23,1%	6	37,5%	
Translocation					**0,00**
Yes	3	23,1%	13	81,3%	
No	10	76,9%	3	18,8%	
Core-binding factor leukemia	0	0,0%	11	68,8%	**<.0001**
AML1-ETO + *C-KIT*^{WT}	0	0,0%	1	6,3%	
AML1-ETO + *C-KIT*^{MUT}	0	0,0%	2	12,5%	
AML1-ETO + *C-KIT*ND	0	0,0%	3	18,8%	
CBFB-MYH11	0	0,0%	5	31,3%	
FLT3					**<.0001**
ITD	11	84,6%	0	0,0%	
WT	2	15,4%	16	100,0%	
CEBPA					NA
Double mutated	0	0,0%	1	6,3%	
WT	12	92,3%	15	93,8%	
Unknown	1	7,7%	0	0,0%	
Karyotype					0,26
Normal	6	46,2%	4	25,0%	
Abnormal	6	46,2%	11	68,8%	
Unknown	1	7,7%	1	6,3%	
CNS involvement					1,00
Yes	2	15,4%	2	12,5%	
No	9	69,2%	14	87,5%	
Unknown	2	15,4%	0	0,0%	

Risk classification					**0,00**
SR	4	30,8%	14	87,5%	
HR	7	53,8%	0	0,0%	
Unknown	2	15,4%	2	12,5%	
FAB classification					0,30
Immature (M0 - M1)	3	23,1%	1	6,3%	
Mature (M2-M7)	10	76,9%	15	93,8%	

Characteristics of 29 pedAML patients used for *TARP* qPCR analysis. Patients were dichotomized as TARP-high (n=13/29, 44.8%) and -low (n=16/29, 55.2%), based on a cut-off calculated as described in Supplemental Materials. Two-sided P-values are representative for the significance of the differential characteristics between *TARP*-high and -low patients and indicated in bold if significant. Continuous variables (WBC count, age and percentage blasts in BM and PB) were dichotomized as described in Supplemental Methods. *WT1* overexpression was interpreted in regard to in-house or published (Cilloni et al. 2009) cut-offs. None of the patients harbored *NPM1* or *FLT3*-TKD mutations. Secondary AML implies AML after MDS/JMML. Superscripts indicate one (*) or two (†) missing data. NA indicates not applicable (number of positive cases too low).
PedAML, pediatric acute myeloid leukemia; qPCR, quantitative PCR; Dx, diagnosis; R, relapse; F, female; M, male; WBC, white blood cell; BM, bone marrow; PB, peripheral blood; FAB, French-British-American; *FLT3*, fms-like tyrosine kinase receptor-3; *NPM1*, nucleophosmin; *CEBPA*, CCAAT/enhancer-binding protein alpha; ITD, internal tandem duplication; WT, wild type; MT, mutated; *WT1*, Wilms' tumor 1; CNS, central nerve system; SR: standard risk; HR: high risk; MDS, myelodysplasia; JMML, juvenile myelomonocytic leukemia.

Table S3. PedAML *FLT3*-ITD physical characteristics correlated to *TARP* expression.

***FLT3*-ITD+ pedAML (n=11) analysed by *TARP* qPCR**	**ITD characteristics**			**TARP expression (CNRQ)**		
	No.	**Length (bp)**	**AR (%)**	**Dichotomized**	**LSC**	**L-blast**
pedAML2	4	21/30/33/39	6.6/19.1/5.5/15.2	High	3.3	2.6
pedAML5	1	33	47.0	High	ND	1.5
pedAML6	1	33	43.0	High	1.4	1.6
pedAML9	1	96	62.0	High	2.5	3.0
pedAML11	2	29/38	14.5/9.6	High	2.7	2.8
pedAML12	2	20/23	44.8/17.1	High	3.4	4.9
pedAML13	1	23	42.0	High	4.9	4.7
pedAML14	2	31/75	70.8/2.7	High	6.2	7.0
pedAML15	1	84	23.0	High	ND	1.4
pedAML20	1	78	3.0	High	0.5	2.4
pedAML29	1	23	15.4	High	3.0	2.9
median	1	33	17.7		3.0	2.8
95% CI median	1 - 2	23 - 39	9.7 - 43.0		1.6 - 4.7	1.6 - 3.7
range	1 - 4	20 - 96	2.7 - 70.8		0.5 - 6.2	1.4 - 7.0

Representation of ITD physical characteristics, i.e. number of clones, length of ITD(s) and AR, for 11 pedAML patients (numeric coded), with concomitant *TARP* expression values (CNRQ) measured in LSC and/or L-blast by qPCR. Median values with 95% confidence intervals (CI), and minimum-maximum values, are summarized at the bottom.
FLT3, fms-like tyrosine kinase receptor-3; ITD, internal tandem duplication; pedAML, pediatric AML; *TARP*, T-cell receptor γ chain alternate reading frame protein; no., number; bp, base pairs; AR, allelic ratio; CNRQ, calibrated normalized relative quantity; ND, not determined; LSC, leukemic stem cell; L-blast, leukemic blast.

Table S4. Multivariate analysis between dichotomous and continuous variables and EFS.

Variables in the Equation	B	SE	Wald	df	Sig.	Exp(B)	95,0% CI for Exp(B)	
							Lower	Upper
TARP	,000	38,130	,000	1	1,000	1,000	,000	2.9E+35
FLT3-ITD	,000	22,947	,000	1	1,000	1,000	,000	3.4E+22
Sexe	,000	2,891	,000	1	1,000	1,000	,003	2.9E+05
CD34	,000	24,284	,000	1	1,000	1,000	,000	4.7E+23
Age	,000	,326	,000	1	1,000	1,000	,527	1.9E+03
WBC count PB	,000	,079	,000	1	1,000	1,000	,857	1.2E+03
Percentage blasts BM	,000	,060	,000	1	1,000	1,000	,889	1.1E+03
Percentage blasts PB	,000	,096	,000	1	1,000	1,000	,828	1.2E+03
Translocation	,000	35,566	,000	1	1,000	1,000	,000	1.9E+33
Core-binding factor	,000	1,400	,000	1	1,000	1,000	,064	1.6E+04
WT1 overexpression	,000	2,190	,000	1	1,000	1,000	,014	7.3E+04
CEBPA	,000	11,609	,000	1	1,000	1,000	,000	7.6E+12
Karyotype			.	0a	.			
CNS invasion			.	0a	.			
Risk classification			.	0a	.			

Covariate Means	
Variables in the Equation	Mean
*TARP*_dich_final	,308
FLT3-ITD_presence_dich	,154
Sexe	,385
CD34_dich	,846
Age	9,077
WBC_count_PB	51,790
Perc_blast_BM	66,000
Perc_blast_PB	44,000
translocation_dich	,692
CBF	1,308
*WT1*_overexpression	,615
CEBPA	,231
Karyotype_normal	1,154
CNS_dich	,231
Risk_high_dich	,231

a Degree of freedom reduced because of constant or linearly dependent covariates

TARP expression (high versus low) did not significantly impacted EFS in a multivariate analysis model (SPSS version 25.0.0.2, Inc., Chicago, IL). *FLT3*, fms-like tyrosine kinase receptor-3; ITD, internal tandem duplication; pedAML, pediatric AML; TARP, T-cell receptor γ chain alternate reading frame protein; EFS, event-free survival; PB, peripheral blood; BM, bone marrow; CNS, central nerve system; *WT1*, Wilms' tumor 1.

Table S5. Overview of the used antibodies.

Antibody	**Fluorochrome**	**Clone**	**Supplier**	**Cat. no.**	**dilution**
CD34	PerCP-Cy5.5	8G12	BD Biosciences	333146	1/20
CD38	APC-H7	HB7	BD Biosciences	656646	1/80
CD45	PacO	HI30	Invitrogen	MHCD4530	1/100
CD8	PE-Cy7	SFCI21Thy2D3	Beckman Coulter	737661	1/50
CD3	APC	SK7	BD Biosciences	345767	1/100
LIVE/DEAD® Fixable Near-IR Dead Cell Stain	APC-H7	/	ThermoFisher Scientific	L10119	1/500
HLA-A2	FITC	BB7.2	BD Biosciences	343304	1/100
Anti-Mouse TCR β Chain (mTCRαβ)	PE	H57-597	BD Biosciences	561081	1/100

8

CHAPTER VIII: Discussion and future perspectives.

Although most pediatric AML (pedAML) patients nowadays achieve excellent clinical remission rates, still 30-40% will relapse. Ample evidence in adult AML showed that relapse is promoted by the persistence of leukemic stem cells (LSC). Hitherto, literature describing the phenotype, molecular features and clinical value of LSC in pedAML is far more scarce compared to adult AML. In this doctoral dissertation, we contribute to an improved molecular and flow cytometric characterisation of LSC in pedAML, hereby identifying novel targets for therapy, and aid in further unravelling the high inter-patient heterogeneity.

Flowcytometry is considered as one of the most powerful techniques to unravel the phenotypic heterogeneity in AML, as it allows polychromatic analysis of single cells at high flow rates.

In **chapter III**, we demonstrate a high prevalence of LSC-specific markers previously described in adult AML, i.e. CD45RA, CLL-1, CD56 and TIM3, of which some are eligible for targeted therapy [1-5]. Further research is needed concerning GPR56 expression in pedAML and the potential prognostic value of this LSC-associated marker. Hitherto, flow cytometric data on GPR56 expression have only been reported in adult AML [6-9]. In our pediatric patient cohort, GPR56 overexpression showed a significant association with *FLT3*-ITD mutations and mutual exclusivity with CBF-leukemia. Approximately half of the patients showed GPR56 overexpression at diagnosis, and higher percentages were observed in those who ultimately relapsed (77%) and in relapsed patients themselves (80%). These findings suggest that GPR56 could represent an interesting immunophenotypic marker able to predict detrimental outcome. This hypothesis is in agreement with a very recent report where GPR56 was allocated as one of the six genes in a LSC signature with a high prognostic value in pedAML [10]. It would also be of interest to explore whether GPR56 overexpression in the LSC compartment of pedAML is linked to *EVI1* overexpression, as documented in adult AML [7]. Unfortunately, GPR56 is not a suited target for therapy, as expression in normal HSC is apparently higher than in pedAML LSC.

We also demonstrate in **chapter III** that narrowing down the $CD34^{+}/CD38^{-/dim}$ compartment to only those cells that harbour aberrant LAIP expression is crucial for detecting the genuine LSC population. Patients classified as LSC^{high} based on *LAIP+* $CD34^{+}CD38^{-/dim}$ cell frequencies in peripheral blood (PB) (≥4.78% of the white blood cells (WBCs) or ≥17.39% of the CD34+ cells) tended to present a worse event-free survival (EFS) compared to LSC^{low} patients. The cut-off established in our study to discriminate LSC^{high} from LSC^{low} patients within the CD34 compartment (17.39%) appeared to be almost identical to the one previously established by Hanekamp et al. (17.2%) [11]. This finding provides proof-of-concept that cut-offs and gating strategies validated by large cohort studies in laboratories working according to EuroFlow guidelines can be adapted by other (smaller) centers when strictly adhering to the pre-analytical and analytical requirements [12, 13]. Hanekamp et al. showed that pedAML patients with ≥17.2% *LAIP+* LSCs within the CD34+ compartment had significantly more risk of developing relapse compare to patients with *LAIP+* LSC loads <17.2% (77.8% vs. 42.4%, Plogrank < .05). Although we did observe a trend towards lower EFS when higher LSC loads were present at diagnosis, our patient cohort was likely too small to detect genuine significant differences. Therefore, re-evaluation upon completion of the NOPHO-DBH AML2012 study is necessary.

Despite the use of a high number of LSC-specific LAIP markers compared to previous research [11], still a considerable number of LSC^{low} patients experienced relapse. This finding suggests that a high LSC load at diagnosis is not the only feature triggering relapse, or, that current flow cytometric methodology still performs inadequate in detecting the genuine LSC fraction. It will therefore be

interesting to evaluate whether use of more advanced and sensitive flow cytometric techniques will improve risk stratification of pedAML patients based on the diagnostic LAIP+ LSC load.

Flow cytometric characterisation of LSC in (ped)AML and their implementation in routine diagnostics will most certainly benefit from the technological advances made this past decade, both on the level of instrumentation and antigens. Whereas a FACSCanto II (BD Biosciences) equipped with three lasers was used in our study, allowing simultaneous measurement of eight colours, two 10+ colour flow cytometers have been recently approved by the EU directive for In-vitro-Diagnostics use, i.e. the DxFLEX (13-colour, Beckman Coulter) and BD FACSLyric™ (12 colour, BD Biosciences). Such polychromatic flow cytometers generate multi-dimensional and complex data sets. This complexity has emerged the need for novel data analysis strategies beyond the conventional single and bi-dimensional plotting that allow objective interpretation. More advanced software such as Infinicyt (Cytognos SL, Salamanca, Spain) allows linking patient-specific gating strategies to analysis profiles, which will decrease the inter-operator variability during manual gating. Standardization of flow cytometry in routine clinical laboratories took a leap forward by the implementation of pattern-guided automated principal component analysis (PCA) [14]. Based on a mathematical algorithm incorporated as a tool within analysis software, the diagnosis can be facilitated based on reference imaging.

A next-level approach would be automated identification of cell populations by pattern classification-based approaches, hereby entirely eliminating manual gating and operator bias. The Flow self-organising map (SOM) approach by the group of Saeys and Van Gassen holds great promise, as it serves as a visualization technique and an automatic clustering algorithm [15]. The FlowSOM algorithm used a four-step approach, consisting out of (i) data read, (ii) construction of a SOM, (ii) construction of a minimal spanning tree and (iv) meta-clustering of the results. However, this approach cannot be considered as fully automated, as data file pre-processing (removal of debris and dead cells) is still a requisite. Full automation of data analysis in AML will most likely remain extremely challenging.

In addition to the progress made in conventional flow cytometry, novel high dimensional single-cell cytometric techniques are emerging. Spectral cytometry combines ultrafast optical spectroscopy with flow cytometry [16, 17]. The unique optical configuration and algorithm allow high sensitivity, automatic analysis and real-time spectral unmixing. Compensation handlings that remove the signal from off-target detectors in conventional flow cytometry is replaced by the deconvolution in spectral technology. Another advantage is the ability to filter out auto-fluorescence. Standard flow cytometers are limited in the simultaneous detection of multiple fluorescence signals, depending upon the instrument. Spectral cytometry, on the other hand, was reported to simultaneously detect beyond 30 fluorescent signals from single particles, including fluorochromes with emission spectra in close proximity to each other [18]. The Sony SP6800 was the first commercial spectral cytometer developed about a decade ago [19]. In solid tissues, hematopoietic subsets could be discriminated by combining 19 fluorochromes in a single 21-parameter analysis [20]. However, the applicability of spectral cytometry in LSC research has not yet been explored. Although the field of spectral flow cytometry is moving forward very slowly, it might replace polychromatic flow cytometry entirely in the future [19].

Also the repertoire of antigens-of-interest has expanded. Phospho flow cytometry, a.k.a. phospho-flow or phospho-signal profiling, is a novel technique capable of analysing hallmark phospho-proteins in leukemic blasts. Phospho-flow was recently applied on 166 BM samples of children with AML [21], and proven to act complementary to current genomic approaches for prognostication and prediction of treatment response. By correlating expression levels of activated phospho-proteins measured at diagnosis with the occurrence of relapse, the group of Dworzak observed that high pSTAT5 activation

and low pJNK activation increases the risk on relapse. Future analysis of phospho-proteomic profiles of LSC is warranted.

The value of monitoring (minute) LAIP+ LSC subpopulation during therapy, e.g. after first induction and before consolidation, in children with AML still needs to be explored. It was proposed by the ELN Working Party that minimal (measurable) residual disease (MRD) measurements of the leukemic blasts in adult AML is most ideally based on a patient-specific LAIP, established at diagnosis, in combination with a 'Different from Normal' (DfN) strategy, based on aberrant differentiation/maturation profiles [22]. The DfN approach combines a plethora of LAIP markers and allows multiple antibodies to be conjugated to the same fluorochromes [23, 24]. The major advantage of the DfN approach is that detection will not fail in case of phenotypic shifts due to clonal evolution or subclonal selection. The DfN approach has also been shown to be suited for LSC measurements during follow-up in adult AML [23]. Our research data in **chapter III** further underline the added value of combining LSC antigen-directed monoclonal antibodies (mAb) labelled to the same fluorochrome into a DfN cocktail approach, and provide proof-of-concept that the markers included in the one-tube assay of the Dutch LSC research group would also be applicable in pedAML [23].

In adult AML, proof-of-concept was provided that a post-treatment approach in which MRD monitoring of the leukemic blasts ('MRD') is combined with determination of the LSC frequency ('LSC') is able to better predict outcome than based on MRD alone, and reveals novel patient subgroups with distinct survival rates [25]. Patients with no residual LSC nor leukemic blast present (LSC-/MRD-) showed the highest survival rates, whilst LSC+/MRD+ patients had a significant poorer outcome. Remarkably, LSC+/MRD- patients showed lower survival rates than LSC-/MRD+ patients although not at a significant level. In a follow-up study, prognostic relevance of *LAIP+* $CD34^{+}CD38^{-}$ LSC frequencies at diagnosis and after induction therapy was validated prospectively [26]. 'LSC' and 'MRD' frequencies determined after the second induction course identified a LSC+/MRD+ adult AML subcohort with 100% treatment failure probability (0% 3-yrs. OS, although based on a limited number of 20 patients) [26]. Hence, it is convincing that the successes booked by therapy response-guided risk stratification in adult AML could be further improved by the inclusion of LSC measurements during therapy.
In a pediatric setting, MRD status of leukemic blasts after the first induction and before consolidation is the most important factor to predict relapse [27-32]. MRD measurements are currently based on the detection of leukemic blasts according to an LAIP-phenotype established at diagnosis. However, still a considerable percentage of the MRD- pedAML patients relapse, a phenomenon which is possibly attributable to the persistence of LAIP+ LSC. Subsequently, combining flow cytometric measurements of residual LSC and leukemic blasts at critical time points might provide a better risk stratification of MRD-/relapse+ children. Within this perspective, a small pilot-study is currently ongoing in which we explore the added prognostic value of LSC measurements during follow-up based on a limited set of markers (CD123, CD56+2 and CD7). Seventeen pedAML patients treated in the NOPHO-DBH AML2012 study were included, and data analysis from LSC measurements during follow-up at critical time points, i.e. after induction and before consolidation, is ongoing.

Incorporation of the MRD status determined by flow cytometry as a clinical end point has been proposed as a surrogate marker to evaluate effectiveness of AML therapy. Late-phase clinical trials currently use overall survival (OS) and EFS as clinical end points, which implies long follow-up periods [33]. Using MRD as (one of the) clinical end point(s) could possibly facilitate drug development and drug approval compared to the traditional use of the achievement of a complete remission [34]. The increasing number of LSC-targeted therapies, and the convincing data on the prognostic value of LSC

at diagnosis and follow-up, suggest that residual LSC frequencies and MRD of the leukemic blasts both have high potential to be included as a surrogate intermediate endpoint for survival.

Next to flow cytometry, gene expression profiling has also contributed to unravelling heterogeneity in AML by exposing transcriptome alterations between leukemic and normal stem cells and blasts. This has led to the development of LSC-specific signatures in adult AML [35-42], of which some were shown to hold an independent prognostic value. Despite their proven utility in clinical practice, there seems to be only minor consensus between the different LSC signatures. Furthermore, targets included in these signatures are not consistently LSC-specific as they are sometimes expressed in HSC. Within a pedAML setting, only one LSC signature has been developed [10]. Although the prognostic value of this LSC6 score was confirmed in two independent datasets, this signature is biased as only the genes included in the LSC17 signature by Ng. et al. were evaluated [43]. Since pediatric and adult AML have been convincingly shown to be two biologically distinct diseases [44], valid targets are potentially being overlooked in this LSC6 score.

In **chapter IV**, we aim to improve the validity of pedAML-specific LSC signatures and researched significantly overexpressed genes in LSC and L-blast of pedAML patients, which are not expressed in their respective normal counterparts. This way, we identified novel potentially interesting biomarkers and therapeutic targets involved in tumorigenesis, immune regulation, apoptosis, adhesion and signaling, i.e. *CDKN1A*, *CFP* and *CFD* in LSC subpopulations, and *HOMER3*, *CTSA* and *GADD45B* in L-blast subpopulations. Although promising, these data need confirmation in larger, preferentially multicenter trials, since survival analysis was performed in only a limited number patients. Further exploration of the protein expression of the selected mRNA targets identified will also be required.
Although work still needs to be done, we envision that reporting these subpopulation-derived gene expression alterations will stimulate further studies to understand the association between LSC activity and molecular characteristics with a well-established value in prognostication. These novel LSC- and L-blast-specific targets might eventually translate into a better classification of AML, and hopefully will give rise to novel targeted therapies. In the future, data integration by different research groups is warranted, so that pedAML-specific LSC signatures are not biased due to institutional setting, treatment protocols or genetic variation.
In this chapter, we demonstrate that the molecular heterogeneity between pediatric and adult AML is also present at the earliest stem cell level. Only three mutual upregulated targets (*TYROBP*, *CFP*, *PTH2R*) were identified. Further research is warranted to evaluate whether these three targets could serve as pan-LSC targets, irrespective of the age of onset. Although these findings further underline the distinct biology between pediatric and adult AML [44], it should be taken into account that, due to the small number of patients evaluated (four children and nine adults), the genetic subtypes might be a confounding factor.

The results addressed in **chapter V**, although preliminary, represent pioneer research regarding the expression of long non-coding RNAs (lncRNAs) in LSC and L-blast subpopulations of pedAML patients. To the best of our knowledge, current literature only addresses the expression of lncRNAs in bulk pedAML leukemic cells [45, 46], or, the expression of lncRNAs in LSC sorted from adult AML [47]. We show a particular interest for *Lnc-GSG1-1*, *lnc-RNFT2-*4, *lnc-RGMA-1* and *lnc-LHFPL3-1*, as their expression is increased in both LSC and L-blast, which would theoretically allow complete tumour eradication by single-hit targeting.

For these targets, we envision to design antisense LNA™ GapmeRs, potent antisense oligonucleotides used for highly efficient inhibition of lncRNA function. GapmeR-targeting of lncRNA molecules was recently proven to be successful in other pediatric haematological malignancies, i.e. juvenile myelomonocytic leukemia [48]. The effect of GapmeR treatment on lncRNA expression levels and cell viability will be evaluated by qPCR and flow cytometry viability assays, respectively. One of the first lncRNA targets that will be examined is *CRNDE*. We have already confirmed elevated *CRNDE* expression in LSC compared to HSC using qPCR at a significant level, and *CRNDE* was previously attributed oncogenic properties [46, 49, 50]. If LNA™ GapmeRs-induced lncRNA perturbation would turn out to be an effective strategy for LSC eradication in pedAML *in vitro*, whilst guaranteeing salvage of HSC, the safety and efficacy of this therapeutic approach will be validated in xenograft models.

It is important to emphasize that the (sub)clonal evolutional and mutational behaviour of leukemic subpopulations under chemotherapeutic pressure has not been evaluated in this doctoral dissertation. Important information remains hidden in this field, which can be pivotal in further understanding the biology of the disease. Taking into account the recent findings by Goardon and Quek et al. [51, 52], it would be interesting in future research to explore whether leukemic transformation in pedAML occurs in stem and/or progenitor cells at different maturation stages, or, whether distinct LSC with different phenotypes identified within one patient have evolved from one single clone through the acquisition of additional hits. It could be interesting to sort the different LSC phenotypes described up to now at diagnosis, and evaluate by xenograft experiments whether they consistently possess leukemia-initiating capacity (LIC), or, whether the LIC potential differs between phenotypes. Another point of interest could be to sort the same stem cell and progenitor populations at diagnosis and follow-up, and evaluate the expression of flow cytometric and molecular markers at different time points in order to evaluate their subjectivity to chemotherapeutic pressure. Especially differential expressions at relapse versus diagnosis could be of guidance to select targets for therapy.

Within this perspective, we started a collaboration with the group of Linda Fögelstrand (Uppsala, Sweden), which is one of the NOPHO-DBH AML2012 study coordinators. In their study, leukemia-specific mutations are identified by exome sequencing of DNA isolated from sorted leukemic blasts at diagnosis, as well as lymphocytes at diagnosis or from bulk mononuclear cells at remission state. A targeted amplicon-based sequencing assay is then designed for each of these mutations (aim for 3 mutations per case) and performed at different follow-up time points. Both the deep sequencing technique [53] and patient-tailored targeted technique [54, 55] have been extensively validated, and tested in a small cohort of six pedAML patients [55]. Although the main research question of this project is 'Is a patient-tailored next-generation sequencing-based MRD analysis able to predict relapses that are not predicted by flow cytometric MRD?', these data will also shed light on (i) the co-existence of multiple leukemic subclones at diagnosis, (ii) the mutational behaviour of leukemic blasts during chemotherapy and (iii) whether shifts between subclones and clonal progression (e.g. transformation from 'diagnostic' clone to 'relapse' clone). As for one patient also relapsed material was sent, exome sequencing data will provide information whether the clones at the initial diagnosis remain present at relapse, and are accompanied by many additional (driver) mutations.

One of the techniques that also could be of service for deeper investigation of intra-patient molecular heterogeneity is the PrimeFlow™ assay. PrimeFlow™ is able to combine flow cytometric and molecular expression measurement of mRNA and miRNA targets. In **chapter VI**, we illustrated that the PrimeFlow™ assay is a sensitive technique that allows investigation of mRNA expression level, with high concordance to qPCR, in heterogeneous samples and populations. This technique eliminates the

need for cell-sorting and subsequent RNA loss. However, the high labour-intensity of this technique, and the difficult standardization due to multiple manual handlings, counteracts its implementation in a routine diagnostic setting. Hence, as long as no LSC-specific LAIP marker is identified that covers all LSC, i.e. maximal sensitivity, and no molecular marker is found that serves as an ideal target for LSC-targeted therapy, i.e. expressed in all pedAML patients irrespective of their risk profiles and genetic subtypes, this technique is not ready for a routine clinical setting.

Current therapeutic regimens in pedAML have reached a plateau around 70% based on multidrug chemotherapy, optimal stratification for allo-SCT and advances in supportive care. Excessive toxicity rates counteract further treatment intensification. Future perspectives need to focus on the development of novel targeted therapy with low 'off-tumor' effects. However, the high heterogeneity within the pedAML landscape counteracts the establishment of patient subgroup-specific therapeutic strategies. Aside from the successful *FLT3* inhibitor-based therapies and the CD33-targeting agents, targeted strategies have not yet found their way into the treatment of pedAML.

Translating the success of targeted immunotherapy from ALL, exemplified by Blinatumomab, to AML is challenged by the cumbersome identification of suitable targets. Therefore, one of the main goals nowadays in AML research is to identify novel targetable therapies for those patients who experience relapse under conventional chemotherapy. Albeit LSC-targeting represents a very interesting strategy to improve long-term outcome, the field is moving forward slowly. In solid cancer, more progress has been made these past decades. Of particular interest was the discovery by researchers in Uppsala that the T-cell receptor γ alternate reading frame protein (TARP) is a tumor-associated antigen (TAA) expressed in androgen-dependent prostate and breast carcinoma with low "on-target/off-tumour" effects (addressed in **chapter VII.1**).

In **chapter VII.2**, we demonstrate that TARP has a high potential as an immunotherapeutic target in pediatric and adult AML. A disadvantage is that transgenic CTLs harbouring TAA-directed TCRs suffer from a low affinity in comparison to foreign viral epitope TCRs [56]. Currently lacking is the proof of *in vivo* functionality by xenograft mouse models. Whether the tumor-specific killing by TARP-TCR directed CTLs outweighs the possible graft-versus-host disease effectuated by the donors' endogenous TCR repertoire, which could severely compromise clinical practicability, is yet to be explored. If successful, one should keep in mind that implementation in clinical therapy studies will also be challenged by autologous T-cell engineering. The troublesome diverse allogeneic T-cell repertoire may also be circumvented by transducing T-cells generated from hematopoietic progenitor cells [57]. Further studies are needed to evaluate whether TARP-directed adoptive T-cell therapy may act synergistically with concomitant chemotherapy. Another disadvantage is that TCR-engineered T-cells suffer from a MHC-restricted affinity. Development of a single chain fragment variable fragment that recognizes the TARP(P5L)$_{4\text{-}13}$ epitope or other high-affinity TARP peptides, and integration into a TCR-mimicking CAR, could provide an alternative targeting strategy.

In **chapter VII.3,** we consolidate our findings from chapter VII.2 using a larger cohort of pedAML patients, and attribute a prognostic value for *TARP* expression in pedAML. We here illustrate that TARP presentation on leukemic cells may induce an *in vivo* immune response in pedAML patients. It might be interesting in future experiments to evaluate whether we can increase specific T-cell immune responses directed against leukemic cells by administration of immune-checkpoint inhibitors directed against PD-1. Using a multi-institutional approach, we envision to further investigate whether *TARP*-

high/*FLT3*-WT pedAML patients define a (till now undetectable) poor prognosis group with high-risk genetic lesions and poor outcome.

Despite these novel data, the role of TARP in leukemic development is still far from understood. We have not yet elucidated why TARP is upregulated in AML, nor what the relation is with *FLT3*-ITD in a pediatric onset. In order to get a grip on the biological role of TARP in AML, we used the mRNA sequencing generated from TARP-knockdown transgenic cell lines, as described in **chapter VII.3**, to identify TARP-downstream signaling pathways and molecules. The sequencing data of transgenic HL-60 (*FLT3* WT AML) and MV4;11 (*FLT3*-ITD AML) cell lines were compared with those established in the respective mock and WT cell lines.
This analysis showed that targets identified as significantly differentially expressed after TARP-knockdown in HL-60 and MV4;11 showed only limited overlap. Some of the mutual upregulated transcripts have already been validated by qPCR (P<.05), i.e. *MYLIP*, *STK17B*, *BTG3*, *TIPARP* and KREMEN1 and *MIR181HG*. Interestingly, five out of these six TARP-downstream targets were previously linked to prostate and/or breast cancer. Although further exploration of these data using a higher number of biological replicates is needed, it is possible that these targets are genuinely related to TARP downstream signaling.
Based on several lines of evidence, we have developed a theory in which increased *KREMEN1* and *STK17B* expression upon TARP silencing involves a central mechanism that revolves around the canonical Wnt pathway. In line with this hypothesis, it was previously demonstrated in prostate cancer that *TARP* is a biologically relevant *HOXC6* downstream target, and that *HOXC6* indirectly influences the Wnt signaling pathway *in vivo* [58, 59]. Although promising, these data are currently to preliminary and require further validation with more biological replicates.

In conclusion, this doctoral dissertation paves the way for novel LSC-targeting therapeutic strategies in pediatric AML. We address the prevalence of LSC-specific membrane markers previously unreported in a pediatric setting, and discovered novel LSC- and leukemic blast-specific targets with highly differential expression compared to their normal counterparts. Furthermore, we report on a novel immunotherapeutic target for which TARP-TCR transgenic T-cells have been developed.
However, there are also some weaknesses in our study. First, validation of the leukemia-initiating capacity of the used cell fractions by xenograft studies is lacking. Second, leukemic stem cell properties of these subpopulation were not tested by functional assays such as colony replating assays, ALDH assays and side populations, which were proven to be extremely valid in identifying the genuine LSC populations [60, 61]. The largest disadvantage of our studies, but most likely also the one with the easiest solution, is the low number of patient samples included. Although AML is the second most frequent type of leukemia in children, only 10 children are diagnosed on average per year in Belgium. The number of patients with generated qPCR and flow cytometric data of the LSC compartment was larger than initially hoped for (n=24 and n=35 CD34+ pedAML patients, respectively), but still less than 15 patients with well-characterised LSC fractions were treated according to the same protocol. Therefore, the here described analyses will be continued for all future NOPHO-DBH AML2012-treated patients at diagnosis and relapse. A big strength also lies in our collaboration with the Dutch research group of Amsterdam (G.J. Schuurhuis and J. Cloos) and Rotterdam (V. van der Velden). As standardization of the flow cytometric procedures across centers allows an easy exchange of raw data, collaboration could pave the way to more significant test results.

Despite all efforts, the LSC compartment still remains an enigmatic fraction. Future research will benefit from integrating multiple approaches in order to implement LSC measurements into the current risk stratification strategy. We envision integration of the here discussed techniques and targets for LSC identification, i.e. membrane markers (flow cytometry), and coding (mRNA) and non-coding (lncRNA) targets apparently over- or underexpressed in LSC (micro-array profiling and qPCR). Most likely, combinatorial treatment approaches will be required to improve the effectiveness and safety of immunotherapy in children with AML.

References

1. Leong SR, Sukumaran S, Hristopoulos M, Totpal K, Stainton S, Lu E, Wong A, Tam L, Newman R, Vuillemenot BR et al: An anti-CD3/anti-CLL-1 bispecific antibody for the treatment of acute myeloid leukemia. Blood 2017, 129(5):609-618.
2. Lin TY, Zhu Y, Li Y, Zhang H, Ma AH, Long Q, Keck J, Lam KS, Pan CX, Jonas BA: Daunorubicin-containing CLL1-targeting nanomicelles have anti-leukemia stem cell activity in acute myeloid leukemia. Nanomedicine : nanotechnology, biology, and medicine 2019.
3. Ma H, Padmanabhan IS, Parmar S, Gong Y: Targeting CLL-1 for acute myeloid leukemia therapy. Journal of hematology & oncology 2019, 12(1):41.
4. Kikushige Y, Miyamoto T: TIM-3 as a novel therapeutic target for eradicating acute myelogenous leukemia stem cells. International journal of hematology 2013, 98(6):627-633.
5. Kikushige Y, Shima T, Takayanagi S, Urata S, Miyamoto T, Iwasaki H, Takenaka K, Teshima T, Tanaka T, Inagaki Y et al: TIM-3 is a promising target to selectively kill acute myeloid leukemia stem cells. Cell stem cell 2010, 7(6):708-717.
6. Daga S, Rosenberger A, Quehenberger F, Krisper N, Prietl B, Reinisch A, Zebisch A, Sill H, Wolfler A: High GPR56 surface expression correlates with a leukemic stem cell gene signature in CD34-positive AML. Cancer medicine 2019, 8(4):1771-1778.
7. Saito Y, Kaneda K, Suekane A, Ichihara E, Nakahata S, Yamakawa N, Nagai K, Mizuno N, Kogawa K, Miura I et al: Maintenance of the hematopoietic stem cell pool in bone marrow niches by EVI1-regulated GPR56. Leukemia 2013, 27(8):1637-1649.
8. Pabst C, Bergeron A, Lavallee VP, Yeh J, Gendron P, Norddahl GL, Krosl J, Boivin I, Deneault E, Simard J et al: GPR56 identifies primary human acute myeloid leukemia cells with high repopulating potential in vivo. Blood 2016, 127(16)(16):2018-2027.
9. Daria D, Kirsten N, Muranyi A, Mulaw M, Ihme S, Kechter A, Hollnagel M, Bullinger L, Dohner K, Dohner H et al: GPR56 contributes to the development of acute myeloid leukemia in mice. Leukemia 2016, 30(8):1734-1741.
10. Elsayed AH, Rafiee R, Cao X, Raimondi S, Downing JR, Ribeiro R, Fan Y, Gruber TA, Baker S, Klco J et al: A six-gene leukemic stem cell score identifies high risk pediatric acute myeloid leukemia. Leukemia 2019, doi: 10.1038/s41375-019-0604-8. [Epub ahead of print].
11. Hanekamp D, Denys B, Kaspers GJL, Te Marvelde JG, Schuurhuis GJ, De Haas V, De Moerloose B, de Bont ES, Zwaan CM, de Jong A et al: Leukaemic stem cell load at diagnosis predicts the development of relapse in young acute myeloid leukaemia patients. British journal of haematology 2018, 183(3):512-516.
12. Cloos J, Harris JR, Janssen J, Kelder A, Huang F, Sijm G, Vonk M, Snel AN, Scheick JR, Scholten WJ et al: Comprehensive Protocol to Sample and Process Bone Marrow for Measuring Measurable Residual Disease and Leukemic Stem Cells in Acute Myeloid Leukemia. Journal of visualized experiments : JoVE 2018(133).
13. Zeijlemaker W, Kelder A, Cloos J, Schuurhuis GJ: Immunophenotypic Detection of Measurable Residual (Stem Cell) Disease Using LAIP Approach in Acute Myeloid Leukemia. Current protocols in cytometry / editorial board, J Paul Robinson, managing editor [et al] 2019, 91(1):e66.
14. Costa ES, Pedreira CE, Barrena S, Lecrevisse Q, Flores J, Quijano S, Almeida J, del Carmen Garcia-Macias M, Bottcher S, Van Dongen JJ et al: Automated pattern-guided principal component analysis vs expert-based immunophenotypic classification of B-cell chronic lymphoproliferative disorders: a step forward in the standardization of clinical immunophenotyping. Leukemia 2010, 24(11):1927-1933.
15. Van Gassen S, Callebaut B, Van Helden MJ, Lambrecht BN, Demeester P, Dhaene T, Saeys Y: FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. Cytometry Part A : the journal of the International Society for Analytical Cytology 2015, 87(7):636-645.

16. Nolan JP, Condello D: Spectral flow cytometry. Current protocols in cytometry / editorial board, J Paul Robinson, managing editor [et al] 2013, Chapter 1:Unit1 27.
17. Preffer F, Dombkowski D: Advances in complex multiparameter flow cytometry technology: Applications in stem cell research. Cytometry Part B, Clinical cytometry 2009, 76(5):295-314.
18. Gregori G, Patsekin V, Rajwa B, Jones J, Ragheb K, Holdman C, Robinson JP: Hyperspectral cytometry at the single-cell level using a 32-channel photodetector. Cytometry Part A : the journal of the International Society for Analytical Cytology 2012, 81(1):35-44.
19. Robinson JP: Spectral flow cytometry-Quo vadimus? Cytometry Part A : the journal of the International Society for Analytical Cytology 2019, 95(8):823-824.
20. Schmutz S, Valente M, Cumano A, Novault S: Spectral Cytometry Has Unique Properties Allowing Multicolor Analysis of Cell Suspensions Isolated from Solid Tissues. PloS one 2016, 11(8):e0159961.
21. Schumich A, Prchal-Murphy M, Maurer-Granofszky M, Hoelbl-Kovacic A, Muhlegger N, Potschger U, Fajmann S, Haas OA, Nebral K, von Neuhoff N et al: Phospho-Profiling Linking Biology and Clinics in Pediatric Acute Myeloid Leukemia. HemaSphere 2020, 4(1):e312.
22. Schuurhuis GJ, Heuser M, Freeman S, Bene MC, Buccisano F, Cloos J, Grimwade D, Haferlach T, Hills RK, Hourigan CS et al: Minimal/measurable residual disease in AML: consensus document from ELN MRD Working Party. Blood 2018.
23. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, Scholten WJ, Snel AN, Veldhuizen D, Cloos J, Ossenkoppele GJ, Schuurhuis GJ: A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. Leukemia 2016, 30(2):439-446.
24. van Rhenen A, Feller N, Kelder A, Westra AH, Rombouts E, Zweegman S, van der Pol MA, Waisfisz Q, Ossenkoppele GJ, Schuurhuis GJ: High stem cell frequency in acute myeloid leukemia at diagnosis predicts high minimal residual disease and poor survival. Clinical Cancer Research 2005, 11(18):6520-6527.
25. Terwijn M, Zeijlemaker W, Kelder A, Rutten AP, Snel AN, Scholten WJ, Pabst T, Verhoef G, Lowenberg B, Zweegman S et al: Leukemic Stem Cell Frequency: A Strong Biomarker for Clinical Outcome in Acute Myeloid Leukemia. PloS one 2014, 9(9).
26. Zeijlemaker W, Grob T, Meijer R, Hanekamp D, Kelder A, Carbaat-Ham JC, Oussoren-Brockhoff YJM, Snel AN, Veldhuizen D, Scholten WJ et al: CD34(+)CD38(-) leukemic stem cell frequency to predict outcome in acute myeloid leukemia. Leukemia 2018, 33:1102–1112.
27. Campana D: Status of minimal residual disease testing in childhood haematological malignancies. British journal of haematology 2008, 143(4):481-489.
28. Abrahamsson J, Forestier E, Heldrup J, Jahnukainen K, Jonsson OG, Lausen B, Palle J, Zeller B, Hasle H: Response-guided induction therapy in pediatric acute myeloid leukemia with excellent remission rate. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2011, 29(3):310-315.
29. van der Velden VH, van der Sluijs-Geling A, Gibson BE, te Marvelde JG, Hoogeveen PG, Hop WC, Wheatley K, Bierings MB, Schuurhuis GJ, de Graaf SS et al: Clinical significance of flowcytometric minimal residual disease detection in pediatric acute myeloid leukemia patients treated according to the DCOG ANLL97/MRC AML12 protocol. Leukemia 2010, 24(9):1599-1606.
30. Loken MR, Alonzo TA, Pardo L, Gerbing RB, Raimondi SC, Hirsch BA, Ho PA, Franklin J, Cooper TM, Gamis AS et al: Residual disease detected by multidimensional flow cytometry signifies high relapse risk in patients with de novo acute myeloid leukemia: a report from Children's Oncology Group. Blood 2012, 120(8):1581-1588.
31. Inaba H, Coustan-Smith E, Cao X, Pounds SB, Shurtleff SA, Wang KY, Raimondi SC, Onciu M, Jacobsen J, Ribeiro RC et al: Comparative analysis of different approaches to measure treatment response in acute myeloid leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012, 30(29):3625-3632.

32. Lapillonne H, Renneville A, Auvrignon A, Flamant C, Blaise A, Perot C, Lai JL, Ballerini P, Mazingue F, Fasola S et al: High WT1 expression after induction therapy predicts high risk of relapse and death in pediatric acute myeloid leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2006, 24(10):1507-1515.
33. Medeiros BC: Interpretation of clinical endpoints in trials of acute myeloid leukemia. Leukemia research 2018, 68:32-39.
34. Estey E, Othus M, Lee SJ, Appelbaum FR, Gale RP: New drug approvals in acute myeloid leukemia: what's the best end point? Leukemia 2016, 30(3):521-525.
35. Ng SW, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, Arruda A, Popescu A, Gupta V, Schimmer AD et al: A 17-gene stemness score for rapid determination of risk in acute leukaemia. Nature 2016, Dec 15(540(7633)):433-437.
36. Majetl R, Becker MW, Tian Q, Lee TLM, Yan XW, Liu R, Chiang JH, Hood L, Clarke MF, Weissman IL: Dysregulated gene expression networks in human acute myelogenous leukemia stem cells. Proceedings of the National Academy of Sciences of the United States of America 2009, 106(9):3396-3401.
37. Eppert K, Takenaka K, Lechman ER, Waldron L, Nilsson B, van Galen P, Metzeler KH, Poeppl A, Ling V, Beyene J et al: Stem cell gene expression programs influence clinical outcome in human leukemia. Nature medicine 2011, 17(9):1086-1093.
38. Forsberg EC, Passegue E, Prohaska SS, Wagers AJ, Koeva M, Stuart JM, Weissman IL: Molecular signatures of quiescent, mobilized and leukemia-initiating hematopoietic stem cells. PloS one 2010, 5(1):e8785.
39. Gal H, Amariglio N, Trakhtenbrot L, Jacob-Hirsh J, Margalit O, Avigdor A, Nagler A, Tavor S, Ein-Dor L, Lapidot T et al: Gene expression profiles of AML derived stem cells; similarity to hematopoietic stem cells. Leukemia 2006, 20(12):2147-2154.
40. Gentles AJ, Plevritis SK, Majeti R, Alizadeh AA: Association of a leukemic stem cell gene expression signature with clinical outcomes in acute myeloid leukemia. Jama 2010, 304(24):2706-2715.
41. Saito Y, Kitamura H, Hijikata A, Tomizawa-Murasawa M, Tanaka S, Takagi S, Uchida N, Suzuki N, Sone A, Najima Y et al: Identification of therapeutic targets for quiescent, chemotherapy-resistant human leukemia stem cells. Science translational medicine 2010, 2(17):17ra19.
42. de Leeuw DC, Denkers F, Olthof MC, Rutten AP, Pouwels W, Schuurhuis GJ, Ossenkoppele GJ, Smit L: Attenuation of microRNA-126 expression that drives CD34+38- stem/progenitor cells in acute myeloid leukemia leads to tumor eradication. Cancer research 2014, 74(7):2094-2105.
43. Ng M, Heckl D, Klusmann JH: The Regulatory Roles of Long Noncoding RNAs in Acute Myeloid Leukemia. Frontiers in oncology 2019, 9:570.
44. Bolouri H, Farrar JE, Triche T, Jr., Ries RE, Lim EL, Alonzo TA, Ma Y, Moore R, Mungall AJ, Marra MA et al: The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nature medicine 2018, Jan(24(1)):103-112.
45. Cao L, Xiao PF, Tao YF, Hu SY, Lu J, Zhao WL, Li ZH, Wang NN, Wang J, Feng X et al: Microarray profiling of bone marrow long non-coding RNA expression in Chinese pediatric acute myeloid leukemia patients. Oncology reports 2016, 35(2):757-770.
46. Yin X, Huang S, Zhu R, Fan F, Sun C, Hu Y: Identification of long non-coding RNA competing interactions and biological pathways associated with prognosis in pediatric and adolescent cytogenetically normal acute myeloid leukemia. Cancer cell international 2018, 18:122.
47. Bill M, Papaioannou D, Karunasiri M, Kohlschmidt J, Pepe F, Walker CJ, Walker AE, Brannan Z, Pathmanathan A, Zhang X et al: Expression and functional relevance of long non-coding RNAs in acute myeloid leukemia stem cells. Leukemia 2019, 33(9):2169-2182.
48. Hofmans M, Depreter B, Lammens T, al. e: Long Non-Coding RNAs As Novel Therapeutic Targets in Juvenile Myelomonocytic Leukemia: Proof of Concept Study. Blood 2019, 134(Supplement_1):1701-1701.

49. Wang Y, Zhou Q, Ma JJ: High expression of lnc-CRNDE presents as a biomarker for acute myeloid leukemia and promotes the malignant progression in acute myeloid leukemia cell line U937. European review for medical and pharmacological sciences 2018, 22(3):763-770.
50. Garitano-Trojaola A, Agirre X, Prosper F, Fortes P: Long non-coding RNAs in haematological malignancies. Int J Mol Sci 2013, 14(8):15386-15422.
51. Quek L, Otto GW, Garnett C, Lhermitte L, Karamitros D, Stoilova B, Lau IJ, Doondeea J, Usukhbayar B, Kennedy A et al: Genetically distinct leukemic stem cells in human CD34- acute myeloid leukemia are arrested at a hemopoietic precursor-like stage. The Journal of experimental medicine 2016, 213(8):1513-1535.
52. Goardon N, Marchi E, Atzberger A, Quek L, Schuh A, Soneji S, Woll P, Mead A, Alford KA, Rout R et al: Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer cell 2011, 19(1):138-152.
53. Delsing Malmberg E, Johansson Alm S, Nicklasson M, Lazarevic V, Stahlman S, Samuelsson T, Lenhoff S, Asp J, Ehinger M, Palmqvist L et al: Minimal residual disease assessed with deep sequencing of NPM1 mutations predicts relapse after allogeneic stem cell transplant in AML. Leukemia & lymphoma 2018:1-9.
54. Malmberg EB, Stahlman S, Rehammar A, Samuelsson T, Alm SJ, Kristiansson E, Abrahamsson J, Garelius H, Pettersson L, Ehinger M et al: Patient-tailored analysis of minimal residual disease in acute myeloid leukemia using next-generation sequencing. European journal of haematology 2017, 98(1):26-37.
55. Delsing Malmberg E, Rehammar A, Pereira MB, Abrahamsson J, Samuelsson T, Stahlman S, Asp J, Tierens A, Palmqvist L, Kristiansson E et al: Accurate and Sensitive Analysis of Minimal Residual Disease in Acute Myeloid Leukemia Using Deep Sequencing of Single Nucleotide Variations. The Journal of molecular diagnostics : JMD 2019, 21(1):149-162.
56. Hillerdal V, Boura VF, Bjorkelund H, Andersson K, Essand M: Avidity characterization of genetically engineered T-cells with novel and established approaches. BMC immunology 2016, 17(1):23.
57. Van Caeneghem Y, De Munter S, Tieppo P, Goetgeluk G, Weening K, Verstichel G, Bonte S, Taghon T, Leclercq G, Kerre T et al: Antigen receptor-redirected T cells derived from hematopoietic precursor cells lack expression of the endogenous TCR/CD3 receptor and exhibit specific antitumor capacities. Oncoimmunology 2017, 6(3):e1283460.
58. McCabe CD, Spyropoulos DD, Martin D, Moreno CS: Genome-wide analysis of the homeobox C6 transcriptional network in prostate cancer. Cancer research 2008, 68(6):1988-1996.
59. Ramachandran S, Liu P, Young AN, Yin-Goen Q, Lim SD, Laycock N, Amin MB, Carney JK, Marshall FF, Petros JA et al: Loss of HOXC6 expression induces apoptosis in prostate cancer cells. Oncogene 2005, 24(1):188-198.
60. Moshaver B, Wouters RF, Kelder A, Ossenkoppele GJ, Westra GAH, Kwidama Z, Rutten AR, Kaspers GJL, Zweegman S, Cloos J et al: Relationship between CD34/CD38 and side population (SP) defined leukemia stem cell compartments in acute myeloid leukemia. Leukemia research 2019, 81:27-34.
61. Moshaver B, van Rhenen A, Kelder A, van der Pol M, Terwijn M, Bachas C, Westra AH, Ossenkoppele GJ, Zweegman S, Schuurhuis GJ: Identification of a small subpopulation of candidate leukemia-initiating cells in the side population of patients with acute myeloid leukemia. Stem Cells 2008, 26(12):3059-3067.

Summary

Acute myeloid leukemia (AML) is a heterogeneous haematological malignancy, accounting for 20% of all pediatric leukemias. Most children with AML (pediatric AML, pedAML) will achieve excellent clinical remission rates after the first induction course, nowadays over >90%, but this achievement is unfortunately not consistently translated into high cure rates. Still 30-40% of the good responders relapse. Relapsed or refractory pedAML patients are eligible for salvage therapy by allogeneic donor stem cell transplantation (allo-SCT) in first remission, before or after consolidation. This procedure can rescue about half of the patients. The high relapse rate, together with early and late therapy-related mortality, results in 5-year event free survival (EFS) and overall survival (OS) rates of 50% and 70%, respectively.

During the past decade, ample evidence in adult AML and a handful reports in pedAML showed that relapse is promoted by the persistence of leukemic stem cells (LSCs). The cancer stem cell model postulates that LSCs endow a leukemia reservoir and harbour chemotherapy-resistant capacities based on their quiescence, unlimited self-renewal and multidrug resistance. The CD34+/CD38-/dim compartment was shown to be most LSC-enriched in CD34+ AML. In addition, the LSC frequency at diagnosis was proposed to have a prognostic value. However, high persistent relapse in LSC^{low} pedAML patients suggests that current methodologies possibly underestimate the LSC fraction. This emphasizes the need for a more profound molecular and phenotypic characterization of LSCs.

In **chapter III**, we aimed to improve the detection and quantification of LSCs by establishing a more refined immunophenotype. Aberrant expression of a total 12 LSC-specific markers with a well-established value in adult AML LSC quantification was evaluated in the $CD34^{+}/CD38^{-/dim}$ compartment of pedAML patients, next to the LSC-associated markers CD49d and GPR56, and compared to their expression in healthy counterparts i.e. HSCs in normal pediatric bone marrow and cord blood. We demonstrated that $CD34^{+}/CD38^{-}$ and $CD34^{+}/CD38^{dim}$ populations show no significant differences in terms of positivity or MFI values in the vast majority of the leukemia-associated immunophenotype (LAIP) markers. Therefore, taking the entire CD38 population $<10^{3}$ into account may be considered, which will increase robustness and sensitivity of the analysis. We showed that aberrant markers aid in a better establishment of the real frequency of the genuine LSC population. We demonstrated that PB may be considered as a valid alternative sample for LSC phenotyping. Finally, we conclude that future LSC-directed antibody therapies in a pediatric setting will benefit from combinatorial therapy, as LSC immunophenotypic profiles show high inter-patient heterogeneity.

Next to deciphering the phenotypic heterogeneity based on cell membrane markers, we evaluated the transcriptome of LSCs and leukemic blasts (L-blasts) using a cancer versus normal (CvN) approach. In **chapter IV**, we focussed on the coding transcriptome and uncovered a novel set of overexpressed transcripts in LSCs and in L-blast, as well as LSC-specific downregulated targets, of which most have not been studied in the context of AML. The vast majority of the transcripts are involved in immune regulation, apoptosis, adhesion or intracellular signaling, making those attractive candidates for functional studies, the establishment of pediatric-specific LSC-signatures and targeted therapy. In addition, gene set enrichment analysis of LSC and L-blast gene expression profiles revealed interesting pathways which might contribute to further understanding of the leukemogenesis in pedAML. In **chapter V**, we initiated pioneer research regarding the role of long non-coding RNAs (lncRNAs) in pedAML subpopulations. To the best of our knowledge, current literature only addresses the expression of lncRNAs in bulk pedAML leukemic cells, or, the expression of lncRNAs in adult AML LSCs. Here, we summarize the most differentially expressed lncRNAs between LSC and HSCs of pedAML patients. Functional analysis of differentially expressed lncRNAs was performed through a lncRNA-

mRNA interaction network. These preliminary data pave the way for introducing lncRNAs into the field of clinical diagnostics and aids in a better understanding of AML leukemogenesis.

In **chapter VI**, we validated the performance of a novel assay that integrates the power of immunophenotypic and molecular methodologies into one single assay, and is suited for the detecting of rare subpopulations. Technical validation of the PrimeFlow™ RNA assay was followed by the evaluation of its feasibility in detecting key target mRNAs in AML subpopulations, e.g. LSCs, of primary patient samples. From this research, we conclude that PrimeFlow™ RNA assay is a sensitive flow cytometric technique for the investigation of coding mRNA expressions in heterogeneous samples and rare subpopulations herein, without the need for cell-sorting, showing a significant correlation to the gold standard real-time qPCR. In spite of these successful results, the limited number of mRNA targets detectable in one run (maximum 3) and the high labour intensity of this technique represent a hurdle for implementation in routine diagnostic applications.

The expression dataset described in chapter IV did not only provided us a more detailed view on the molecular heterogeneity of pedAML, but also allowed us to discover novel therapeutic targets. The current high relapse rates with conventional chemotherapeutic regimens highlight the need for the development of new (targeted) therapeutic strategies. Targeted therapy has led to a remarkable progress in the survival rates of multiple cancers. These successes paved the way for exploring efficacy of monoclonal antibodies (mAb), T-cell receptor (TCR)- or chimeric antigen receptor (CAR)- transgenic cytotoxic T-cells (CTLs) in AML. Development of immunotherapeutic strategies targeting the rare LSC compartment might be a crucial step in further increasing the cure rates.

Within this perspective, we describe in **chapter VII.1** the relevance of the TCR γ chain alternate reading frame protein (TARP) as a novel immunotherapeutic target in androgen-dependent prostate and breast carcinoma. In **chapter VII.2**, we describe its discovery as an AML-specific target expressed in the LSCs and L-blasts of pediatric and adult AML, while absent in their normal counterparts. *TARP* transcript expression was significantly associated with FMS-like tyrosine kinase-3 internal tandem duplication (*FLT3*-ITD) in pedAML. We provide *in vitro* evidence that TARP may serve as a novel immunotherapeutic target in AML for TARP-TCR engineered CTLs. Cytokine release and cytotoxic killing capacities of TARP-TCR transgenic CTLs were *in vitro* evaluated against TARP+/HLA-A*0201+ co-expressing cell lines and primary patient samples. A more in-depth evaluation of *TARP* in a pediatric setting is provided in **chapter VII.3**. Using cell-sorting and TARP(P5L)$_{4\text{-}13}$ directed HLA-A*0201 tetramers, we demonstrated the presence of *TARP* expression in leukemic cells from *de novo* pedAML patients is able to activate CTLs *in vivo*. Furthermore, we confirmed the existence of a second, AML-exclusive *TARP* transcript. In addition, we demonstrated that *TARP* expression is significantly anti-correlated to EFS in an univariate, though not multivariate, analysis based on a small cohort of pedAML patients treated according to the NOPHO-DBH AML2012 protocol.

Finally, in **chapter VIII**, we discuss the results of this doctoral dissertation in terms of future perspectives, and discuss some preliminary data that need to be further investigated.

In conclusion, this doctoral dissertation aids in the further unravelling of AML pathogenesis in children, using in-depth immunophenotypic and molecular studies of LSC and leukemic blasts subpopulations. We discovered novel coding and non-coding targets and obtained *in vitro* data on a novel immunotherapeutic target in AML, which will be able to move the field of targeted therapy in pedAML forward.

Samenvatting

Acute myeloïde leukemie (AML) is een sterk heterogene hematologische ziekte die voorkomt in 20% van de kinderleukemieën. De overgrote meerderheid (>90%) van de kinderen met AML (pediatrische AML, pedAML) zullen een klinische remissie (ziektevrije status) bereiken na de eerste therapie kuur. Jammer genoeg wordt deze goede therapierespons niet consistent vertaald in hoge genezingscijfers. Nog steeds 30-40% van de kinderen met een goeie therapierespons zullen hervallen. PedAML patiënten die hervallen, of een refractaire ziekte vertonen (d.w.z. geen goede therapierespons) komen in aanmerking voor een allogene donor stamcel transplantatie. (alloSCT). Deze therapie kan ongeveer de helft van de patiënten nog genezing geven, maar, kent ook een niet te onderschatten vroege en late morbiditeit mortaliteit. Alles samengenomen leidt dit ertoe dat slechts 50% van de pedAML patiënten minimum 5 jaar ziektevrij blijven, en 70% van de patiënten nog in leven zijn na 5 jaar.

De afgelopen 10 jaar werd er veel evidentie verzameld in volwassen patiënten met AML, naast enkele publicaties in pedAML, dat deze hoge hervalcijfers gepromoot worden door de blijvende aanwezigheid van leukemische stamcellen (LSC). Het kanker stamcel model stelt dat leukemische stamcellen een reservoir van leukemische ziekte zijn, en chemotherapieresistente eigenschappen hebben. Voorgaand uitgebreid onderzoek heeft kunnen aantonen dat witte bloedcellen (WBC) met het fenotype $CD34^{+}/CD38^{-/dim}$ de grootste LSC activiteit hebben. Bijkomend werd het aangetoond dat het aantal leukemische stamcellen (de LSC lading, of de LSC frequentie) bij diagnose een prognostische waarde hebben. Desondanks blijven ook patiënten met een lage LSC lading hervallen. Dit suggereert dat de huidige methodologie mogelijks niet toereikend is en de LSC fractie onderschat. Deze hypothese benadrukt de nood voor een betere en meer gedetailleerde moleculaire en fenotypische karakterisatie van de leukemische stamcellen in kinderen met AML.

In **hoofdstuk III** onderzochten we of de detectie en kwantificatie van leukemische stamcellen kan verbeterd door middel van een betere fenotypische karakterisatie van celmembraan antigenen (cluster differentiatie (CD) antigenen) door middel van flowcytometrie. De aberrante expressie van 11 'leukemie-geassocieerde immuun fenotypische (LAIP)' merkers, specifiek voor LSC met een reeds gevestigde waarde bij volwassenen met AML, werden geëvalueerd in het $CD34^{+}/CD38^{-/dim}$ compartiment van pedAML patiënten, naast ook twee LSC-geassocieerde merkers CD49d en GPR56. De expressie in pedAML patiënten werd vergeleken met de expressie in gezonde controles, zijnde hematopoietische stamcellen (HSC) in normaal pediatrisch beenmerg en navelstrengbloed. Wij konden aantonen dat de $CD34^{+}/CD38^{-}$ en $CD34^{+}/CD38^{dim}$ populatie geen significante verschillen tonen op het vlak van positiviteit of mediane fluorescentie intensiteit in het merendeel van de LAIP merkers. Daarom is het aangewezen om de volledige stamcelpopulatie met zwakke tot negatieve CD38 expressie in rekening te brengen, wat de robuustheid en sensitiviteit van de analyse zal verhogen. We toonden aan dat de fenotypische analyse van aberrante merkers bijdraagt tot een betere bepaling van de echte LSC lading. We konden aantonen dat perifeer bloed een geschikte alternatieve staaltype is voor het fenotypering van leukemische stamcellen. Perifeer bloed bleek de beste staaltype te zijn voor de correlatie tussen de LSC lading bij diagnose en de ziektevrije overleving en de hervalvrij overleving. Finaal kunnen we concluderen dat toekomstige LSC-gerichte therapieën in een pediatrische setting baat zullen hebben bij combinatie therapieën, gezien vaak meerdere LAIP merkers aanwezig zijn, en de immuunfenotypische profielen sterk verschillen tussen de verschillende patiënten.

Naast het beter in kaart brengen van de fenotypische heterogeniteit, hebben we ook het transcriptoom van de leukemische stamcellen en leukemische blasten geëvalueerd door middel van het vergelijken van zieke patiënten ten opzichte van gezonde controles. In **hoofdstuk IV** hebben we ons gefocusseerd op het coderende transcriptoom. Zo hebben we genen ontdekte in leukemische

stamcellen en leukemische blasten die significant sterker tot expressie worden gebracht dan in hun gezonde tegenhangers. Daarnaast hebben we ook genen ontdekt die significant onderdrukt worden in leukemische stamcellen. De meeste van deze genen werden nog niet onderzocht in de context van (kinderen met) AML. De overgrote meerderheid van deze genen zijn betrokken in het controleren van de immuunrespons, celdood (apoptose), adhesie en intracellulaire communicatie, wat hen aantrekkelijke kandidaten maakt voor functionele studies, het opstellen van pediatrie-specifieke LSC-handtekeningen en LSC-gerichte therapieën. Hiernaast hebben we expressieprofielen bepaald van de leukemische stamcel en leukemische blast, en op deze manier interessante processen ontdekt die mogelijks zullen bijdragen tot het beter begrijpen van het ontstaan van pediatrische AML. In **hoofdstuk V** hebben we pioniersonderzoek verricht naar de rol van 'long non-coding RNA' (lncRNA) moleculen in deze leukemische stamcellen en leukemische blasten subpopulaties. Voor zover onze kennis reikt, zijn er vandaag de dag in de literatuur enkel expressiedata van lncRNA in bulk leukemische cellen te vinden, of in de leukemische stamcellen van volwassenen met AML. In dit hoofdstuk vatten we de meest differentieel tot expressie gebrachte lncRNA in leukemische stamcellen en blasten van pedAML patiënten ten opzichte van dezelfde cel fracties bij gezonde controles. Een functionele analyse van deze differentieel tot expressie gebrachte lncRNA werd uitgevoerd door middel van een niet coderend (lncRNA) – coderende (mRNA) RNA interactie netwerk. Deze preliminaire data maken de weg vrij voor het introduceren van lncRNA moleculen in het veld van klinische follow-up en therapeutische behandeling, en helpen in het beter begrijpen van het ontstaan van AML bij kinderen.

In **hoofdstuk VI** hebben we de performantie gevalideerd van een nieuwe test (PrimeFlow™ RNA assay) die de kracht van immuunfenotypische detectie bundelt met een moleculaire detectie van coderende messenger RNA (mRNA) moleculen in één enkele test. Deze test is geschikt voor het detecteren van zeldzame subpopulaties in heterogene cel mengsels, zoals LSC bij kinderen met AML. Een technische validatie werd gevolgd door de evaluatie van de toepasbaarheid van deze test in het detecteren van ziekte-specifieke transcript moleculen in primaire AML patiënten stalen, als in cellijnen representatief voor de pediatrische ziekte neuroblastoma. Uit dit onderzoek besluiten we dat de PrimeFlow™ RNA assay een gevoelige flowcytometrische techniek is voor het onderzoeken van coderende mRNA moleculen in heterogene stalen en zeldzame populaties hierin. Deze test elimineert de nood tot het sorteren van de cellen, en toonde een significante correlatie toont tot de gouden standaard 'real-time quantitative polymerase chain reaction' (qPCR). Desondanks deze beloftevolle resultaten, zijn het aantal mRNA targets die in één test geanalyseerd kunnen worden zeer beperkt (maximum 3), en bemoeilijkt de arbeidsintensiviteit van deze techniek een potentiële implementatie in een routine setting.

De expressiedata beschreven in **hoofdstuk IV** lieten ons ook toe om mogelijks interessante targets voor immuuntherapie te ontdekken. De hedendaagse hoge hervalcijfers met conventionele chemotherapeutische behandelingen benadrukken de nood voor het ontwikkelen van nieuwe (target-gerichte) therapieën. Target-gerichte therapie heeft een zeer grote vooruitgang teweeggebracht in de overlevingscijfers van verschillende kanker. Deze successen hebben de weg vrijgemaakt voor het onderzoeken van hun efficaciteit van monoklonale antilichamen, T-cel receptor (TCR)- of chimere antigen receptor (CAR)- getransduceerde cytotoxische T-cellen in de behandeling van AML. Het ontwikkelen van immuuntherapeutische strategieën die gericht zijn tegen zeldzame leukemische stamcellen zou mogelijks een cruciale stap kunnen betekenen in het verbeteren van de genezingscijfers.

In dit opzicht, beschrijven we in **hoofdstuk VII.1** de rol van de 'TCR γ chain alternate reading frame proteïne' (TARP) als immuuntherapeutisch target in androgeengevoelige prostaat- en borsttumoren. In **hoofdstuk VII.2** beschrijven we de ontdekking van TARP als een AML-specifiek target in de

leukemische stamcellen en leukemische blasten van kinderen en volwassenen met AML. *TARP* is afwezig in de normale hematopoietische stamcellen en blasten van gezonde controles. *TARP* transcript expressie was significant geassocieerd met FMS-like tyrosine kinase-3 internal tandem duplicatie (*FLT3*-ITD) in pedAML. We tonen *in vitro* bewijs dat TARP kan gedetecteerd worden door TARP-TCR getransduceerde cytotoxische T-cellen. Cytokine vrijgave en cytotoxische afdoding door deze TARP-TCR getransduceerde cytotoxische T-cellen werd *in vitro* geëvalueerd ten opzichte van cellijnen en primaire patiënten cellen die TARP presenteren op HLA-A2 MHC klasse II moleculen. Een uitgebreidere evaluatie van TARP in een pediatrische AML setting wordt beschreven in **hoofdstuk VII.3**. Door middel van cel-sortering en tetrameer-analyse tonen we hier aan dat de presentatie van TARP op leukemische cellen van pedAML patiënten bij diagnose in staat is om cytotoxische T-cellen *in vivo* te activeren. We bevestigen ook door middel van een andere techniek dan in hoofdstuk VII dat er een tweede, AML-specifiek transcript bestaat. Verder konden we een prognostische waarde aantonen voor TARP in een cohorte van 15 pedAML patiënten die allemaal behandeld waren volgens hetzelfde NOPHO-DBH AML2012 protocol. Een hoge expressie van TARP was significant gecorreleerd met lagere ziektevrije overleving in een univariate, maar niet multivariate, analyse.

Tenslotte bediscussiëren we in **hoofdstuk VIII** onze onderzoeksresultaten in het opzicht van toekomst-perspectieven, en bespreken we enkele preliminaire data die verder onderzocht zullen worden.

In conclusie, deze doctorale dissertatie draagt bij tot een beter begrijpen van het ontstaan van AML in kinderen door middel van uitgebreid flowcytometrisch en moleculair onderzoek van de leukemische stamcellen. We ontdekten nieuwe coderende en niet-coderende genen, en genereerden *in vitro* data omtrent een nieuw immuuntherapeutische target in AML. Deze resultaten zullen bijdragen tot het implementeren van target-gerichte therapie in de behandeling van kinderen met AML.

Dankwoord

Vier jaar geleden nam mijn leven een onverwachte wending door het starten van een doctoraat. Hoewel het 'no walk in the park' is geweest, ben ik vandaag zeer blij om toen die keuze gemaakt te hebben. Ik had al veel gehoord dat een doctoraat bloed, zweet en tranen kost. Ik kan nu goed begrijpen dat vele doctoraatsstudenten zich stroef doorheen het eerste jaar worstelen, of er zelfs mee stoppen. Deze periode van onzekerheid, trial-and-error, veel input en weinig output, was mentaal lastiger dan wat ik me had voorgesteld. Ondanks dit alles had ik het leerproces dat hiermee gepaard gaat, de ervaring die je opbouwt, de evolutie die je doormaakt als wetenschapper en als mens, en de finale voldoening evenmin zo groot geacht. Deze laatste vier jaar heb ik het beste van mezelf gegeven, en ik beschouw vandaag dan ook als het bereiken van een mijlpaal. Ik acht mezelf gelukkig om deze afgelopen jaren zoveel passie voor mijn werk te hebben mogen ervaren.

Talrijke mensen hebben deze weg samen met mij bewandeld. De hieruit vloeiende positieve, maar ook negatieve ervaringen hebben mij inzichten gegeven op professioneel vlak die ik zal blijven onthouden. Hoewel ik iedereen die een positieve bijdrage geleverd heeft oprecht dankbaar ben, wil ik graag enkele personen in het bijzonder bedanken.

Aan mijn beide promotoren: bedankt om mij zoveel veel vrijheid te geven, om mijn inspiratie de vrije loop te laten gaan, en om mijn enthousiasme nooit te hebben willen temperen of teniet te doen als een negatieve eigenschap.

Life is like riding a bicycle. To keep your balance, you must keep moving (Albert Einstein). Jan Philippé, de term mentor komt hier eigenlijk meer tot zijn recht dan promotor. Je slaagde er altijd in om tijd te maken voor mij. Je menselijkheid en oprechtheid waren voor mij een bron van steun en vertrouwen. Je bodemloze interesse en toewijding inspireren om altijd nieuwe uitdagingen aan te gaan, en niet te snel tevreden te zijn. Daarnaast heb je mij van aan de zijlijn mijn sterktes en mijn zwaktes leren kennen. Je kon me met weinig woorden veel inzicht geven. Ik zal onze (ochtendlijke) fietstochtjes niet snel vergeten en hoop dat er in de toekomst nog volgen.

One thing is important in science: only courageous people win (Max Planck). Tim Lammens, hoe jij je geeft voor de wetenschap, doet niemand je na. Elk uur van de dag gaf je mij een antwoord op mijn vragen. Je wist me ook tijdig 'on hold' te zetten, en niet te kort door de bocht te laten gaan. Ik hoop dat je altijd de nodige waardering en respect zal krijgen voor het vele werk dat je verricht.

There is no one giant step that does it. It's a lot of little steps (Peter Cohen).

Mattias (Clemens Rita Maria) Hofmans, samen gingen wij sneller én verder. Onze samenwerking en inspirerende discussies hebben mijn doctoraat tot een hoger niveau getild en voor een onbreekbare vriendschapsband gezorgd. Voor mij is onze verstandhouding het professionele toonbeeld van 2 uitersten die elkaar aantrekken. Met spijt in het hart neem ik afscheid van jou als co(ncu)llega.

Karl Vandepoele, hoewel je het officieel niet was, heb je je deze laatste vier jaar wel gedragen als een promotor. Je hebt me enorm veel geleerd, was altijd oprecht betrokken en hield me alert om kritisch te blijven. Ongelooflijk hoe diep je kennis reikt, en hoe je keer op keer gelijk had.

Dank ook aan alle laboratoriumtechnologen van de dienst Hematologie, CMD hematologie en PHO research. Jullie wisten mij steeds te helpen waar nodig, hielden rekening met mij op drukke dagen, en wisten me te plezieren met een babbeltje wanneer ik talloze uren aan de flowcytometer moest vertoeven. Ik hoop dat ik op mijn beurt jullie ook veel heb kunnen bijbrengen. Dit geldt ook voor het

volledige laboratoriumteam van het UZ Gent. Velen onder jullie vergaten mij niet, en maakten op regelmatige basis eens een babbeltje tijdens de vroege/late uren labowerk. Na zes jaar in het UZ te vertoeven, kan ik oprecht spreken van een vriendschapsband met velen onder jullie.

Ook alle laboratoriumkrachten en PhD studenten van het MRB II en Cancer Center Amsterdam van het VUmc ziekenhuis wil ik bedanken om mij zo op te nemen in jullie groep en mij zoveel te willen leren. Jullie vormen het voorbeeld dat samenwerking de wetenschap tot een hoger niveau tilt.

Aan alle collega's-in-opleiding, met wie ik in het AZ St-Jan en UZ Gent mijn assistentschap heb volbracht: bedankt voor jullie collegialiteit en voor de aangename groepssfeer. Jullie zorgden ervoor dat ik eens kon ontsnappen uit mijn doctoraatsbubbel. In het bijzonder denk ik hier aan Anne-Sophie De Koninck. Je was voor mij onverwachts een grote steun in mijn eerste jaar doctoraat, en dit is geëvolueerd tot een hechte vriendschap.

Dit doctoraat heeft me ook eigenschappen bijgebracht die ik mijn ganse toekomstige carrière kunnen gebruiken. Vooreerst zal ik trachten om met een open visie in het wetenschappelijk leven te staan. Ik heb geleerd dat betrokken zijn bij anderen hun werk leidt tot ideeën en het openen van deuren die anders gesloten zouden blijven. Daarnaast werd ik geïnspireerd door zij die succes niet verwerven ten koste van anderen, maar bereiken door een diepe passie en gedrevenheid en hierbij samenwerking nastreven. Barbara De Moerloose, Bart Vandekerchove, Tessa Kerre en Jacqueline Cloos, ik dank jullie hiervoor. Jullie hebben mij getoond dat bewonderenswaardige academische prestaties niet inherent gepaard hoeven te gaan met een bureaucratische academische attitude. Katrien Devreese, jouw streng maar rechtvaardig beleid heeft mij veel geleerd, en daarnaast slaagde je er ook in om als diensthoofd begrip te tonen wanneer nodig. Ook bedankt, Veronique Stove, om in de hoedanigheid van laboratoriumdirecteur mij als doctoraatstudent en als assistent het nodige vertrouwen te geven.

Naast hard werken, is ook de nodige ontspanning vereist in een doctoraat. Lieve Brugge-vriendinnen en -vrienden, met velen onder jullie deel ik al meer dan 15 jaar lief en leed. Ook zij met wie de vriendschap recenter is (een nieuwe relatie komt met een nieuwe vriendenkring…) liggen mij nauw aan het hart. Jullie hebben altijd veel respect en begrip getoond voor het groot aandeel werk in mijn leven. Jullie namen het mij nooit kwalijk als ik geen tijd had of te laat kwam opdagen. Tegelijkertijd stonden jullie op de eerste rij wanneer ik wel tijd had, en dit met een onvoorwaardelijke liefde. Ik vind jullie elk zo fantastisch anders, en jullie verrijken werkelijk mijn leven. Lieve Farma-vriendinnen, ik vergeet nooit het moment tijdens artsenij-practicum dat enkele onder jullie bijna enthousiaster dan mij waren met mijn toelating tot klinische biologie. Onze vriendschap is een voorbeeld dat "uit het oog, uit het hart" niet waar hoeft te zijn. Elke keer dat we afspreken, blijkt die vriendschap onveranderd, en ik ben trots op jullie allemaal voor wat jullie bereikt hebben. Lieve SASK-ertjes, jullie zijn werkelijk elk uniek. Bedankt voor jullie hechte vriendschap, en dit al bijna 20 jaar.

Mijn laatste dankwoord is voor mijn familie, die alle moeilijke wateren samen met mij doorzwom. Voor Gabi, die terug een stabiele thuis gecreëerd heeft. Voor mijn marraine Kathleen, zonder wie ik nu niet zover zou gestaan hebben. Voor mijn broer, die er onvoorwaardelijk is voor zowel een lach, een traan als een luisterend oor. Voor mijn mama, die me naast vele andere dingen heeft geleerd dat opgeven nooit de juiste keuze is. Ik wou dat je erbij was, en hoop dat jij, als laborante, trots zou geweest zijn. Papa, *a good hairdresser is cheaper than a good shrink*. Jij kent me soms beter dan dat ik mezelf ken. Jij zou tot het einde van de wereld gaan voor mij (maar ik hoop dat Frankrijk toch het verste is dat je ooit zal moeten rijden). Bedankt om me de betekenis van toewijding en doorzetting te leren.

Tot slot, voor mijn man Frederik, bedankt voor je eindeloos begrip, geduld en relativeringsvermogen. Maar vooral bedankt om mijn dagdagelijkse en grootste steun te zijn.

Curriculum Vitae

Personalia

Naam	Barbara Depreter
Adres	Lange Molenstraat 143, 8200 Sint-Andries Brugge
Geboortedatum	21/06/1988
Geboorteplaats	Brugge
Email	depreter_barbara@hotmail.com

Loopbaan - Diploma's

2020-heden	Resident tijdelijk kader op de dienst Klinische Biologie, afdeling Hematologie.	UZ Brussel
2019	Machtiging als apotheker klinisch bioloog.	Universiteit Gent
2019-heden	Board member Belgian Society for the Advancement of Cytometry.	
2016-2019	Doctoraat aspirant (FWO grant 1113119N) vakgroep Inwendige Ziekten en Pediatrie. Doctoraatsonderwerp: de leukemische stamcel als oorzaak van het therapeutisch falen bij pediatrische AML.	Universiteit Gent
2015-heden	Gastprofessor 3de bachelor Medische Laboratorium Technologie. Praktijkonderdeel bij de lessen hematologie II.	HOWEST, Brugge
2013	Behalen van de academische graad 'Master-na-master in de Klinische Biologie' met grootste onderscheiding (864/1000).	Universiteit Gent
2012	Behalen certificaat bloedafname voor apothekers-klinisch biologen.	Universiteit Gent
2011	Behalen van de academische graad 'Master in de Geneesmiddelenontwikkeling, Farmaceutische Wetenschappen' met grote onderscheiding (2009-2010: 843/1000 (ranking 1/172); 2010-2011: 783/1000 (ranking 11/168)).	Universiteit Gent
2009	Behalen van de academische graad 'Bachelor in de Farmaceutische Wetenschappen' met grote onderscheiding (2006-2007: 733/1000 (ranking 48/188); 2007-2008: 829/1000 (ranking 1/152); 2008-2009: 818/1000 (ranking 4/110)).	Universiteit Gent
2006	Behalen diploma secundair onderwijs, afstudeerrichting 'Wetenschappen-Wiskunde (6u)'.	Sint-Andreas Lyceum, Brugge

Stageplan opleiding klinische biologie

2014-2015	Afdeling Hematologie: flowcytometrie, morfologie en moleculaire biologie (6 maand), speciale stolling (6 maand).	UZ Gent
2013-2014	Afdeling Microbiologie: bacteriologie en virologie (9 maand).	UZ Gent
2013	Afdeling Scheikunde: 24-uurs laboratorium (3 maand).	UZ Gent
2012-2013	Afdeling Hematologie: routine hematologie, flowcytometrie en morfologie (12 maand).	AZ St-Jan ziekenhuis Brugge
2011-2012	Afdeling Scheikunde: routine en speciale scheikunde (12 maand).	AZ St-Jan ziekenhuis Brugge

Extracurriculaire opleiding

2018-2019	Laboratorium management.	UA, Antwerpen
2018	Cursus R voor SPSS gebruikers.	Universiteit Gent
2018	Advanced qPCR course.	Biogazelle, Gent
2018	Doctoral Schools specialist course on Cancer.	Universiteit Gent
2017-2018	Onderzoeksstage in het Cancer Center Amsterdam (3 maand).	VUmc, NL
2016	Digital PCR specialist course.	Universiteit Gent
2016	Introduction to R.	Universiteit Gent
2016	Behalen van attest in cursus 'Advanced Writing Skills'.	Universiteit Gent
2015	Behalen van attest in cursus 'Statistische analyse in SPSS'.	Universiteit Gent
2014	Statistical Genome analysis - Advanced Statistics methods.	IVPV, Zwijnaarde
2014	Behalen van attest in 'Conversatiegericht Frans'.	Perspectief, Gent
2013	Behalen van attest in BHS educational course in hematologie.	BHS, Brussel

Prijzen

2019	ASH Abstract Achievement Award voor het abstract "Deciphering Molecular Heterogeneity in Pediatric AML Using a Cancer Vs Normal Transcriptomic Approach".
2019	Travel Grant voor het abstract "TARP as immunotherapeutic target in AML expressed in the LSC compartment." 24th congress of the European Hematology Association, Amsterdam, Nederland.
2014	Posterprijs uitgereikt door de Belgian Society for Clinical Biology voor abstract "Treponemal tests for serodiagnosis of syphilis: comparative evaluation of five assays on Architect syphilis positive sera."
2012	National laureate van de 'Young Scientist Society', uitgereikt door de European Federation of Clinical Chemistry and Laboratory Medicine (EFLM), gebaseerd op het artikel "Accuracy of three automated 25-hydroxyvitamin D assays in haemodialysis patients."

Literatuurlijst

Publicaties

During PhD, included in this doctoral dissertation

Depreter B, De Moerloose B, Hanekamp D, Hofmans M, Uyttebroeck A, Van Damme A, Denys B, Vandepoele K, Dedeken L, Dresse MF, Van der Werff Ten Bosch J, Cloos J, Philippé J and Lammens T. Exploring the immunophenotype of LSC in pediatric AML. Manuscript in preparation.

Depreter B, De Moerloose B, Vandepoele K, Uyttebroeck A, Van Damme A, Ferster A, Van der Werff Ten Bosch J, Dresse MF, Dedeken L, Hofmans M, Philippé J and Lammens T. Deciphering Molecular Heterogeneity in Pediatric AML Using a Cancer Vs Normal Transcriptomic Approach. Manuscript ready for submission.

Depreter B, De Moerloose B, Vandepoele K, Terras E, Uyttebroeck A, Van Damme A, Ferster A, Van der Werff Ten Bosch J, Dresse MF, Hofmans M, Kerre T, Vandekerckhove B, Philippé J and Lammens T. Clinical significance of TARP expression in pediatric acute myeloid leukemia. Accepted for publication in HemaSphere.

Depreter B, Weening K, Vandepoele K, Essand M, De Moerloose B, Themeli M, Cloos J, Hanekamp D, Moors I, D'hont I, Denys B, Uyttebroeck A, Van Damme A, Dedeken L, Snauwaert S, Goetgeluk G, De Munter S, Kerre T, Vandekerckhove B, Lammens T, Philippé J. TARP is an immunotherapeutic target in AML expressed in the leukemic stem cell compartment. Accepted for Haematologica (doi: 10.3324/haematol.2019.222612)– pre-published online.

Depreter B, Philippé J, Meul M, Denys B, Vandepoele K, De Moerloose B, Lammens T. Cancer-related mRNA expression analysis using a novel flow cytometry-based assay. Cytometry B Clin Cytom. 2017 Oct 5. doi: 10.1002/cyto.b.21593.

During PhD, not included in this doctoral dissertation

Hanekamp D, Denys B, Kaspers GJL, te Marvelde JG, Schuurhuis GJ, De Haas V, De Moerloose B, de Bont ES, Zwaan CM, de Jong A, **Depreter B**, Lammens T, Philippé J, Cloos J, van der Velden VHK. Leukaemic Stem Cell Load at Diagnosis Predicts the Development of Relapse in Paediatric Acute Myeloid Leukaemia. Published in Br J Haematol. 2017 Oct 26. doi: 10.1111/bjh.14991.

Depreter B, Devreese K. M. J. Dilute Russell's viper venom time reagents in lupus anticoagulant testing: a well-considered choice. Published in Clin Chem Lab Med 2016. DOI 10.1515/cclm-2016-0245.

Depreter B, Stove V, Delanghe J. Sampling on ice does not yield correct uric acid monitoring in rasburicase-treated patients. Published in Clinical Biochemistry 49 (2016) 1390–1395.

Depreter B, Devreese K. M. J. Differences in lupus anticoagulant final conclusion through clotting time or Rosner index for mixing test interpretation. Published in Clinical Chemistry and Laboratory Medicine 2016, Aug16,vol:54, iss:9.

Van Cauwenberge MG, **Depreter B**, Dumoulin EN, Emmerechts J, Nollet F, Vanopdenbosch LJ. Bing-Neel Syndrome: two unexpected cases and review of the literature. Published in Journal of the Neurological Sciences 2015, Sep 15;356(1-2):19-26.

Depreter B, A. C. Heijboer, M. R. Langlois. Accuracy of three automated 25-hydroxyvitamin D assays in haemodialysis patients. Published in Clinica Chimica Acta 2013, volume 415, 255-260.

Depreter B, Dumoulin E, Billiet J, Cauwelier B, Maes B, Matthys C, Van Droogenbroeck J, Nollet F, Emmerechts J. Biclonal biphenotypical B-cell lymphocytosis: a CLL in disguise? Published in Belgian Journal of Hematology 2014;5(4):143-7.

Depreter B, Dumoulin E, Selleslag D, Emmerechts J, Billiet J. Sebastian Syndrome, de kunst van het kijken. Published in Nederlands Tijdschrift voor Hematologie 2014;10:33-4. Peer-reviewed (A2).

Masterproef Farmaceutische Wetenschappen

Evaluation and validation of the Tacrolimus assay on a Dimension, Cobas 6000 and Architect platform and comparison with LC-MS/MS. Begeleiding: apr. klin. Biol. Hilde Vanpoucke (AZ Delta, Roeselare).

Abstracts geselecteerd voor presentatie op (inter-)nationale meetings

Dilute Russell's viper venom time reagents in lupus anticoagulant testing: a well-considered choice. Annual meeting of the Royal Belgian Society of Laboratory Medicine and Joint meeting with the Belgian Bone club on Bone Turn-over and Bone Resorption Markers (15-16 October 2016, Liège, Belgium).

Evaluation of seven real-time PCR mixes for bacterial DNA contamination in 16S rRNA PCR (oral session: 'PCR-based diagnosis: what's new?'). 25th European Congress of Clinical Microbiology and Infectious Diseases (25-29 April 2015, Copenhagen, Denmark).

The value of SOX11 overexpression in the differential diagnosis of mantle cell lymphoma. Belgian Hematology Society Annual Meeting (30-31 January 2014, Ghent, Belgium).

Importance of accurate 25(OH)D monitoring in haemodialysis patients. 20th IFCC-EFLM European Congress of Clinical Chemistry and Laboratory Medicine (20-24 May 2013, Milano, Italy).

Wetenschappelijke abstracts/posters op (inter-)nationale meetings

Depreter B, Hofmans M, Terras E, Uyttebroeck A, Van Damme A, Ferster A, Van der Werff Ten Bosch J, Dresse MF, Vandepoele K, De Moerloose B, Philippé J and Lammens T. Deciphering Molecular Heterogeneity in Pediatric AML Using a Cancer Vs Normal Transcriptomic Approach. 61st Annual Meeting and Exposition (December 7-10, 2019) in Orlando, FL, USA.

Hofmans M, **Depreter B**, Lammens T, Philippe J, De Moerloose B. Long non-coding RNAs as novel therapeutic targets in juvenile myelomonocytic leukemia: proof of concept study. 61st Annual Meeting and Exposition (December 7-10, 2019) in Orlando, FL, USA.

Depreter B, Weening K, Vandepoele K, Essand M, De Moerloose B, Themeli M, Cloos J, Hanekamp D, Moors I, D'hont I, Denys B, Uyttebroeck A, Van Damme A, Dedeken L, Snauwaert S, Goetgeluk G, De Munter S, Kerre T, Vandekerckhove B, Lammens T, Philippé J. TARP as immunotherapeutic target in AML expressed in the LSC compartment. 24th EHA Congress.2019. Amsterdam, the Netherlands.

B. Depreter, B. Denys, M.H.C. Bakkus, B. De Moerloose, J. Philippé, T. Lammens, L. Willems, K. Vandepoele, V. Mondelaers. MRD discrepancy during follow-up of p190 BCR-ABL1+ childhood B-ALL: diagnosis of a 'CML-like' ALL subgroup. Belgian Hematology Society Annual Meeting. 2019. Brussels, Belgium.

B. Depreter, D. Hanekamp, T. Lammens, B. De Moerloose, E. Sonneveld, B. Denys, A. De Jong, GJ. Kaspers, GJ. Schuurhuis, J. Philippé, J. Cloos. Role of CLEC12A in the LSC compartment of pediatric acute myeloid leukemia at diagnosis versus relapse. Abstract only. 23nd European Hematology Association (EHA) Congress 2018, Stockholm, Sweden.

B. Depreter, B. Denys, J. Philippé. Distinction of leukemic B-lymphoblasts from normal B-cell precursors according to EuroFlow-based cut-offs for CD24, CD38, CD58 and CD66c. 6th ESLHO symposium, Leiden, The Netherlands.

B. Depreter, K. Vandepoele, B. De Moerloose, B. Denys, J. Philippé, T. Lammens. Leukemic stem cell-related mRNA expression analysis using a novel flow cytometry-based assay. 22nd European Hematology Association (EHA) Congress 2017, Madrid, Spain.

B. Depreter, M. Meul, B. Denys, B. De Moerloose, E. Terras, K. Vandepoele, J. Philippé, T. Lammens. PrimeFlow™ RNA Assay: a novel technique for single-cell expression analysis in heterogeneous cell populations. Belgian Hematology Society Annual Meeting. 2017. Brussels, Belgium.

B. Depreter, K. Devreese. Differences between dilute Russell's viper venom time reagents in lupus anticoagulant testing. European Congress on Thrombosis and Haemostasis (ECTH). 2016. The Hague, the Netherlands.

B. Depreter, L. Coorevits, P. Cools, J. Boelens, I. Leroux-Roels, E. Padalko, K. Vandepoele, G. Claeys, B. Verhasselt. Implementation of selection criteria to identify samples for which 16S rRNA gene amplicon sequencing is clinically relevant. 26th European Congress of Clinical Microbiology and Infectious Diseases. 2016. Amsterdam, the Netherlands.

B. Depreter, J. Philippé, B. Denys. Distinguishing leukemic B-lymphoblasts from normal B-cell precursors based on the expression of CD24, CD38, CD58 and CD66c. 10th Biennial Childhood Leukemia Symposium. 2016. Athens, Greece.

B. Depreter, J. Philippé, B. Denys. Distinguishing leukemic B-lymphoblasts from normal B-cell precursors using cut-off values for expression of CD24, CD38, CD58 and CD66c. Belgian Hematology Society Annual Meeting. 2016. Brussels, Belgium.

B. Depreter, K. M. J. Devreese. Clotting time or Rosner index for mixing test interpretation makes a difference in the overall lupus anticoagulant test result. Belgian Society for Thrombosis and Haemostatis Annual Meeting. 2015. Brussels, Belgium.

B. Depreter, V. Stove, J. Delanghe. Even sampling on ice does not entirely prevent temperature-dependent uricolysis in rasburicase-treated patients. Royal Belgian Society of Laboratory Medicine Annual Meeting. 2015. Brussels, Belgium.

H. Stepman, K. Maelegheer, **B. Depreter**, F. Haerynck, S. De Schepper, C. Dhooge, V. Bordon, C. Bonroy. Diagnosing Chédiak-Higashi, added value of the laboratory. International society for laboratory hematology, 19 - 21 May 2015. Chicago Illinois, USA

B. Depreter, S. Deridder, E. Padalko. Comparison of the BioPlex®2200 EBV assay with a classical two-tiered approach. Category Virology non-HIV/non-hepatitis. 25th European Congress of Clinical Microbiology and Infectious Diseases. 25– 29 April 2015. Copenhagen, Denmark.

S. Jonckheere, A. Boel, **B. Depreter**, M. Van Esbroeck, E. Padalko. Treponemal tests for serodiagnosis of syphilis: comparative evaluation of five assays on Architect syphilis positive sera.– Royal Belgian Society of Laboratory Medicine Annual Meeting. 18/10/2014. Leuven, Belgium.

E. Dumoulin*, **B. Depreter***, L. Vanopdenbosch, J. Emmerechts, F. Nollet, D. Selleslag, A. Van Hoof, J. Billiet . Bing-Neel syndrome: report of two cases. Belgian Hematology Society Annual Meeting. 30/01/2014. Ghent, Belgium. (*both authors contributed equally)

B. Depreter, A. C. Heijboer, M. R. Langlois. Accuracy of three automated 25-hydroxyvitamin D assays in haemodialysis patients. Royal Belgian Society of Laboratory Medicine Annual Meeting. 15/10/2012. Liège, Belgium.

B. Depreter, J. Emmerechts, B. Cauwelier. Evaluation of CD200 expression in the differential diagnosis of CD5+ B cell NHLs. Belgian Hematology Society Annual Meeting. January 2012. Ghent, Belgium.

B. Depreter, M. Hutsebaut, M. Langlois, S. Roggeman, M. Hidajat. EliA™ CTD screen: enzyme fluoroimmunoassay for ANA detection. 8th International Congress on Autoimmunity. 2012. Granada, Spain.

B. Depreter, A. Vermeersch, M. Hutsebaut. The antiphospholipid syndrome: ACA analysis on ImmunoCAP250 with EliA compared to a manual ELISA. 8th International Congress on Autoimmunity. 2012. Granada, Spain (accepted abstract, no poster).

B. Depreter, K. Croes, J. Van Bocxlaer, H. Vanpoucke. Evaluation of 3 immunoassays for the quantification of Tacrolimus and Cyclosporin A and correlation with LC-MS. Royal Belgian Society of Laboratory Medicine Annual Meeting. 22/04/2010. Antwerp, Belgium.

Wetenschappelijke presentaties

Immunotherapeutische targets bij leukemische stamcellen bij acute myeloide leukemie. Onderdeel van de MaNaMa Klinische Biologie Brussel, met als topic: Kankerimmunotherapie: update rond een snel evoluerend domein. Bowling Stones, 2020.

Leukemic stem cells in acute myeloid leukemia. Presentatie tijdens de Laboratory sessions of BSAC, Belgian Hematology Society Annual Meeting, 14-15 February 2020, Dole-la-Hulpe, Belgium.

22nd EHA congress: Key messages voor de praktijk. Postgraduaatvergadering UZ Gent, 2017. Compensaties en voltages in flowcytometrie. UZ Gent 2016.

Het flowcytometrisch onderscheid tussen hematogonen en leukemische B-ALL blasten o.b.v.de expressie van CD24, CD38, CD58 en CD66c. Stafvergadering UZ Gent 2015.

Cytomorfologische analyse van klieren en vochten. UZ Gent 2015.

Microbiologische casuïstiek bij de zwangere vrouw & pasgeborene. Stafvergadering UZ Gent 2015.

Nieuwe geïntroduceerde test: 16S rDNA real-time PCR. Stafvergadering UZ Gent 2014.

Exampla rara uit de hematologie. Stafvergadering UZ Gent 2014.

Master-na-master Klinische biologie voor apothekers. Postgraduaatvergadering UZ Gent, 2013.

Minimal residual disease (MRD) detection AML/ALL (flowcytometrie en moleculaire analyses). Postgraduaatvergadering UZ Gent, 2013.

Minimale residuele ziekte detectie bij AML met flowcytometrie. AZ St-Jan Brugge, 2013.

Morfologie van perifeer bloed en beenmerg. AZ St-Jan Brugge, 2013.

Nieuwe inzichten in zowel de pre-analytische, analytische als post-analytische fase bij diagnostiek van broncho-alveolaire lavages. AZ St-Jan Brugge, 2013.

Kritische evaluatie van een laboratorium onderzoek: Creatinine en zijn problematiek. AZ St-Jan Brugge, 2012.

Nieuwe richtlijnen bij xanthochromie bepaling: principe en belang. AZ St-Jan Brugge, 2012.

Plaats van immunoassays binnen toxicologische screening. UZ Gent. Toelatingsproef Klinische biologie.