Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

ONCOLOGIA E PATOLOGIA SPERIMENTALE

Ciclo XXVII

Settore Concorsuale di afferenza: 06/G1

Settore Scientifico disciplinare: MED/38

IDENTIFICATION OF NOVEL GENETIC ALTERATIONS IN PEDIATRIC CYTOGENETICALLY NORMAL ACUTE MYELOID LEUKEMIA BY NEXT-GENERATION SEQUENCING

Presentata da:
Marco Togni

Coordinatore Dottorato
Chiar.mo Prof. Pier Luigi Lollini

Relatore
Chiar.mo Prof. Andrea Pession

Esame Finale Anno 2015

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Pediatric acute myeloid leukemia (AML) is a molecularly heterogeneous disease that arises from genetic alterations in pathways that regulate self-renewal and myeloid differentiation. While the majority of patients carry recurrent chromosomal translocations, almost 20% of childhood AML do not show any recognizable cytogenetic alteration and are defined as cytogenetically normal (CN)-AML. CN-AML patients have always showed a great variability in response to therapy and overall outcome, underlining the presence of unknown genetic changes, not detectable by conventional analyses, but relevant for pathogenesis, dynamics, and outcome of AML.

During the last decade, the development of novel genome-wide techniques such as next-generation sequencing (NGS), have tremendously improved our ability to interrogate the cancer genome, raising, for the first time, the possibility to have a complete picture of all the genetic alterations that could be present in a cancer cell. With the aim of investigating the mutational landscape of CN-AML, lately, several studies have taken advantage of the novel NGS technologies, and novel important insights in the molecular pathogenesis of CN-AML have been achieved, especially in adults. Despite that, mutations find to be recurrent in adult AML appear to be rare or absent in childhood AML, straighten the significant differences between AML in elderly and young patients. Based on this background, the overall aim of this research study was to investigate the mutational landscape of pediatric CN-AML patients negative for all the currently known somatic mutations reported in AML through whole-transcriptome sequencing (RNA-seq).

RNA-seq performed on diagnostic leukemic blasts from 19 pediatric CN-AML cases revealed a considerable incidence of cryptic chromosomal rearrangements in this type of leukemia, with the identification of 21 putative fusion genes. Moreover, and more importantly, several of the fusion genes that were identified in this study are recurrent and might have a prognostic and/or therapeutic relevance. A paradigm of that is the *CBFA2T3-GLIS2* fusion, which has been demonstrated to be a novel common (8.4%) alteration in pediatric CN-AML,

predicting poor outcome. Noteworthy, due to its prognostic relevance, the presence of this fusion gene will be included in the stratification of patients in the new AIEOP LAM 2012/01 Italian clinical trial.

Beside the role of this fusion gene in defying the outcome of pediatric CN-AML patients, important findings have been obtained in the identification of novel molecular alterations that can be therapeutically targeted for more effective treatment strategies. Within the epigenetic regulators, the identification of *NUP98-JARID1A* (1.2%), and *NUP98-PHF23* (2.4%) fusions suggest the use of disulfiram (Food and Drug Administration-approved drug) or other small molecules that inhibit the binding of the plant homology domain to histone 3 trimehylated-lysine 4 and preliminary functional results support this strategy. Alterations activating tyrosine kinases, such as the fusion *TNIP1-PDGFRB* identified in one CN-AML case support the use of already approved and widely used tyrosine kinase inhibitors (e.g. imatinib, dasatinib) to specifically inhibit leukemia cells. This hypothesis has been investigated and confirmed in *in vitro* studies providing a strong rational for subsequent *in vivo* investigations. In conclusion, this study provides new insights in the knowledge of genetic alterations underlying pediatric AML, defines novel prognostic markers and putative therapeutic targets, and prospectively ensures a correct risk stratification and risk-adapted therapy also for the "all-neg" AML subgroup.

CHAPTER I – Introduction

1.1 Pediatric acute myeloid leukemia

Acute myeloid leukemia (AML) is a group of genetically heterogeneous hematopoietic disorders of the myeloid lineage characterized by the uncontrolled growth and clonal expansion without complete differentiation of a hematopoietic/stem progenitor[1].

1.1.1 Epidemiology and overall outcome

In the pediatric setting, AML accounts for approximately 20% of leukemias with seven newly diagnosed children and adolescents (< 19 years-old) out of one million of patients in Italy, annually[2]. Overall, the incidence is slightly higher in male (7.4/million cases in male compared to 6.5/million in female) with a peak in both genders under 1 year of life (*infant*-AML) [2,3] (Figure 1). In the last two decades, the prognosis of pediatric AML has been significantly improved[4,5]. In the 1970s, the survival rates of children younger than 20 years with a newly diagnosed AML was lover than 20%.[6] Nowadays, according to the AIEOP AML2002/01 trial and other international studies, the probability of 8-years event-free-survival (EFS) is approximately up to 50-70%[7–13]. Different factors contributed to this remarkable result, including: i) use of more effective anti-leukemic agents, such as cytarabine and anthracyclines, ii) better risk-group stratification based on cytogenetic and molecular analysis, iii) optimization in induction therapy and improvements in post remission/supportive treatment, and more importantly, iv) the broad use of allogeneic (ALLO) hematopoietic stem cell transplantation (HSCT) in high-risk (HR) patients[12,14].

1.1.2 Etiopathogenesis

Currently, AML is broadly classified into two main categories*:* *de-novo* AML and *secondary* AML. While a *de-novo* AML develops without any exposure to known risk factors and the cause of development remains not known, the rise of a *secondary* AML is subsequent to a prior

exposure to cytotoxic agents, or antecedent hematopoietic insufficiency (myelodysplastic syndrome [MDS], marrow failure, *etc*…)[15]. Exposure to high dose radiation, chemical compounds (benzene, herbicides, pesticides), and commonly used antitumor chemotherapy drugs such as alkylating agents, platinum (including cyclophosphamide, mechlorethamine, procarbazine, chlorambucil, melphalan, busulfan, carmustine, cisplatin, and carboplatin) and treatment with topoisomerases II inhibitors (etoposide) have been linked to an increased risk of AML (*therapy-related* AML, tAML) [16–18]. Additionally, higher risk to develop AML is associated with several congenital syndromes. Among them, the Down syndrome increase the probability to develop AML (particularly the FAB M7 subtype) of 20-fold compared to a pediatric healthy population, even if the leukemia often undergoes to a spontaneous resolution[19]. The Shwachman-Diamond syndrome[20], Kostmann syndrome (severe congenital neutropenia)[21] and the Fanconi anemia[22] are other conditions associated with an increased risk to develop AML in pediatrics.

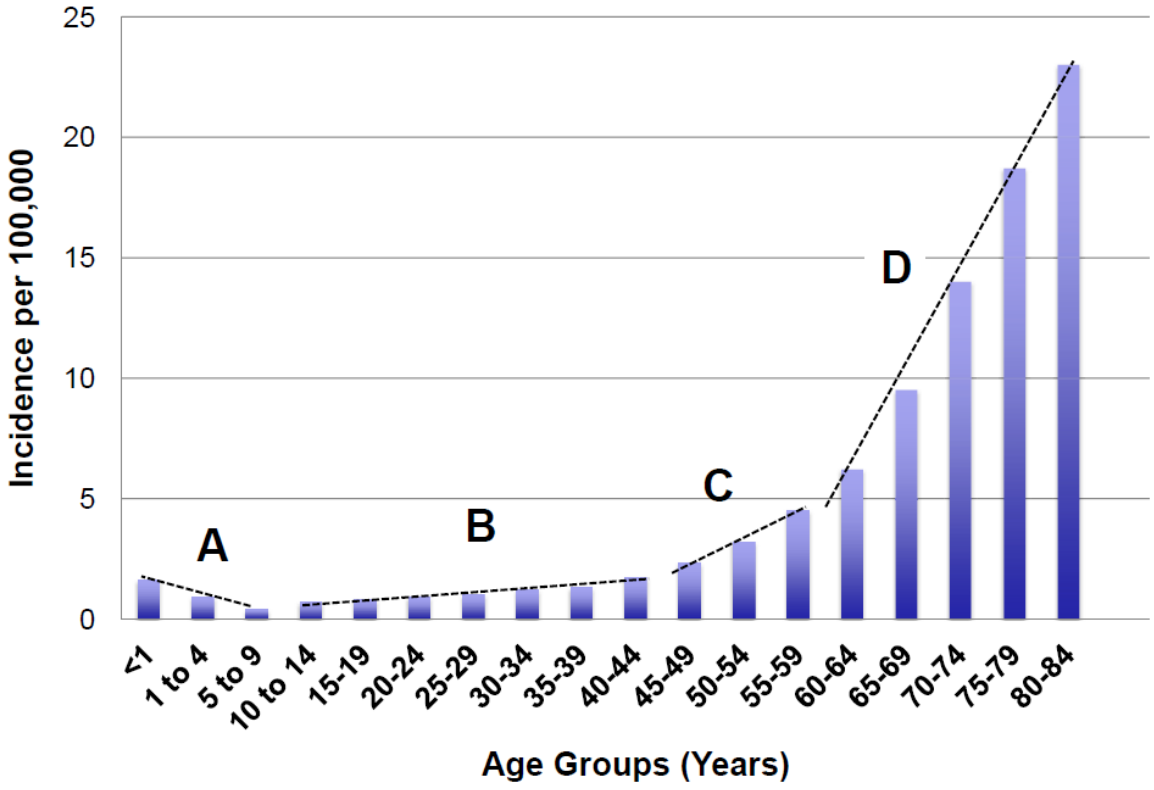


Figure 1. Incidence of acute myeloid leukemia (AML) by age. *From Tarlock K, Meshinchi S. Pediatric Acute Myeloid Leukemia: Biology and Therapeutic Implications of Genomic Variants. Pediatr Clin North Am. 2015 Feb;62(1):75-93.*

1.1.3. Classification

Historically, the first reliable and systemic attempt of a uniform system for classification and nomenclature of AML was provided by the French-American-British (FAB) classification[23,24]. This classification system was based on the morphological and histological characteristics of the leukemic cells and divide the AML by the predominant lineage and stage of differentiation of the leukemic blasts[24] (Table 1) . Although the FAB classification remains useful and commonly used nowadays for an overall description of myoblasts morphology, non-random recurrent cytogenetic and/or molecular abnormalities that are well-known to have a prognostic value in AML (Table 3) are not considered in this classification system. Thus, the need of considering the prognostic importance of these cytogenetic and/or molecular lesions has provided the major reason to the design of a novel AML classification in which these parameters are taken into account. The most recent World Health Organization (WHO) AML classification system is based on the systematic delineation of distinct clinical pathologic entities in which are take in account several parameters including morphology, cell surface, cytogenetic and/or molecular markers of the leukemic cells[25] (Table 2) .

Table 1. French-America-British (FAB) classification for AML

M0	AMLwith no Romanowsky or cytochemical evidence of differentiation
M1	Myeloblastic leukemia with litter maturation
M2	Myeloblastic leukemia with maturation
M3	Acute promyelocytic leukemia
M3h	Acute promyelocytic leukemia, hypergranular variant
M3v	Acute promyelocytic leukemia, microgranular variant
M4	Actue myelomonocytic leukemia
M4eo	Actue myelomonocytic leukemia with dysplastic marrow meosinophils
M5	Acute monoblastic leukemia
M5a	Acute monoblastic leukemia, poorly differentiated
M5b	Acute monoblastic leukemia, differentiated
M6	Erytroleukemia
M7	Acute megakaryoblastic leukemia

Modified from Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, Sultan C. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. Br J Haematol. 1976 Aug;33(4):451-8.

Table 2. The World Health Organization (WHO) classification of acute myeloid leukemia

Acute myeloid leukemia with recurrent genetic abnormalities

AML with balanced translocations/inversions

Acute myeloid leukemia with t(8;21)(q22;q22); *AML/ETO*

Acute myeloid leukemia with inv(16)(p13;q22) or t(16;16)(p13;q22); *CBFB/MYH11*

Acute myeloid leukemia with t(15;17)(q22;q21); *PML/RARα*

Acute myeloid leukemia with t(9;11)(p22;q23); *MLL/AF9*

Acute myeloid leukemia (megakaryoblastic) with t(1;22)(p13;q13); *OTT/MAL*

Acute myeloid leukemia with inv(3)(q21;q26.2) or t(3;3)(q21;q26.2); *RPN1/EVL1*

Acute myeloid leukemia with t(6;9)(p23;q34); *DEK/NUP124*

AML with gene mutations

Mutation affecting FLT3, NPM1, C/EBPα, KIT, MLL, WT1, NRAS and KRAS

Acute myeloid leukemia with myelodisplasia-related changes

Acute leukemia with 20% or more peripheral blood or bone marrow blasts with morphological features of myelodysplasia or a prior history of a myelodyplastic syndrome (MDS) or myelodyplastyic/myeloproliferative neoplasm (MDS/MPN), or MDS-related cytogenetic abnormalities, and absence of the specific genetic abnormalities of AML.

Therapy-related myeloid neoplasms

Therapy-related acute myeloid leukemia (tAML), myelodysplastic syndrome (tMDS), and myelodyplastyic/myeloproliferative neoplasm (tMDS/MPN) occurring as late complications of cytotoxic chemotherapy and/or radiation therapy administered for a prior neoplastic or non-neoplastic disorder.

Acute myeloid leukemia, not otherwise specified

FAB classification (M0, M1, M2, M3, M3h, M3v, M4, M4eo, M5, M5a, M5b, M6, M7.

Myeloid sarcoma

Tumor mass consisting of myeloid blasts with/without maturation, occurring at an anatomical site other than the bone marrow.

Myeloid proliferations related to Down syndrome

Transient abnormal myelopoiesis

Myeloid leukemia associated with Down syndrome

Blastic plasmacytoid dendritic cell neoplasm

Clinically aggressive tumor derived from the precursors of plasmacytoid dendritic cells, with a high frequency of cutaneous and bone marrow involvement and leukemic dissemination.

Modified from the Vardiman JW et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009 Jul 30;114(5):937-51.

1.1.4 Molecular Pathogenesis

AML is characterized by uncontrolled cells growth and stop of differentiation of an hematopoietic stem precursor.[1] Despite the heterogeneous pattern of genetic lesions at the base of the development of AML, several evidences reported in literature over the years sustain a shared "two-hits" model in which the leukemia arises from the cooperation between two different classes of genetic alterations: those that confer a proliferative and/or survival advantage to hematopoietic progenitors, and those that impair hematopoietic differentiation and confer properties of self-renewal to the hematopoietic cell at a particular stage of differentiation[26].

1.1.4.1 Proliferation and/or pro-survival alterations in AML (Class I)

Typically, mutations belonging to this class of genetic lesions occur in genes coding for proteins involved in signaling pathways and in the transduction of extracellular cell proliferation signals, such as *FLT3*, *c-KIT*, *RAS*, *PTPN11*, *NF1* genes[26].

FLT3 activating mutations

Located on human chromosome 13 and coding for a type III tyrosine kinase receptor (TKR), *FMS-like tyrosine kinase 3* (*FLT3*) gene is the most recurrent mutated gene in AML, with up to 30% - 35% of AML patients carrying an activation mutation in this genetic locus.[27] Physiologically, the function of FLT3 receptor is strictly regulated and the activation of his signaling pathways is crucial for the proliferation of hematopoietic stem cells. However, in AML, two different types of mutations could occur, leading to the coding of a receptor constitutively activated that sustains an aberrant proliferation[27]. The most common (20% - 25% of cases) lesion consists in internal tandem duplications (ITD) in the juxtamembrane domain of FLT3 receptor ranging in size from several to > 50 amino acids. These mutations are always in frame, occur exclusively within a tyrosine-kinase inhibitory loop (27-ammino acid stretch) and are all predicted to destroy the inhibitory function of this domain, resulting in a subsequent constitutive activation of the tyrosine kinase domain localized in the C-termini of the receptor in the cytosol[28]. In addition to ITD, single nucleotide mutations leading to missense ammino acid changes or small InDels have been reported in the activation loop of FLT3 in about 5% - 10% of AML cases[29]. Likewise the ITD, these mutations in the tyrosine kinase activation loop lead to the

constitutive activation of the receptor (Figure 2) . Interestingly, examples of both ITD and activation loop mutations in the same allele of *FLT3* have been reported, suggesting that the combination of mutations may hyperactivate the kinase and provide added proliferative advantage to cells that harbor both mutations[26]. Expression of FLT3-ITDs in pre-B murine IL3-dependent Ba/F3 cells has been demonstrated to activate several signal transduction pathways known to confer proliferative and/or survival advantage, including the RAS/ MAPK, STAT, and PI3K/AKT pathways, and are sufficient to confer cytokine independent growth[30,31]. Most studies reported that alteration of the *FLT3* gene are particularly recurrent in AML FAB M3v (acute promyelocytic leukemia, microgranular variant) and in patient with a normal karyotype.[32] More important, activating mutations in *FLT3* gene are well known to be associated with a poor prognosis and relapse in AML, although they pave the way to the broad use of novel tyrosine-kinase inhibitor (TKIs) that can specifically block the constitutive activation of this receptor[33].

c-KIT activating mutations

c-KIT is a proto-oncogene located on human chromosome 4, encoding for a type III TKR. Upon stimulation by stem cell factor, KIT undergoes to dimerization and consequential transphosphorylation and activation of a downstream signaling pathways leading to proliferation, differentiation, and survival of hematopoietic stem cells. Ligand independent activation of KIT is driven by mutations that could occur in the extracellular domain of the receptor, in transmembrane/juxtamembrane domain and in the activation loop of the tyrosine kinase domain. Although the incidence of KIT activating mutations is relevant in adult patients with CBF-AML (up to 48%)[34] and associate with a worst prognosis, the recurrence of these lesion in the pediatric setting seems to be lower (up to 19% of CBF-AML)[35] and not associated with poor prognosis, even if further studies with larger cohort are required for fully assess the prognostic power of *KIT* mutations in pediatric AML and the possible therapeutic relevance as target for TKIs[35].

RAS activating mutations

RAS oncogenes encode for a family of guanine nucleotide-binding proteins that regulate signal transduction on binding to a variety of membrane receptors, including c-KIT and FLT3, and mutations in these proteins are extremely recurrent in a large variety of human cancer, including leukemias. Typically the mutations consist in one single base change in codons 12, 13 or 61, which abrogate intrinsic RAS GTPase activity and confer constitutive activation of RAS proteins

and downstream effectors, such as RAF and MAPK/ERK kinases. Approximately 10–15 and 5% of all AML patients carry mutations in *N-RAS* and *K-RAS*, respectively[36]. No significant difference in recurrence of these mutations have been reported according to age, gender, diagnosis leukocytosis, type of AML (*de-novo* or t-AML); although an increase incidence is recorded in AML with inv(16)/t(16;16) and inv(3)/t(3;3) (up to 35%).[36] Interestingly, only 2% of FLT-ITD-positive AML patients have been found positive also for RAS mutations supporting the "two-hit" model of AML pathogenesis. However, likewise in other human neoplasms, RAS activating mutations are not considered as primary event in the leukemogenesis, and is broadly diffuse the idea that these mutations occur later in de leukemic cells probably due to genome instability[36].

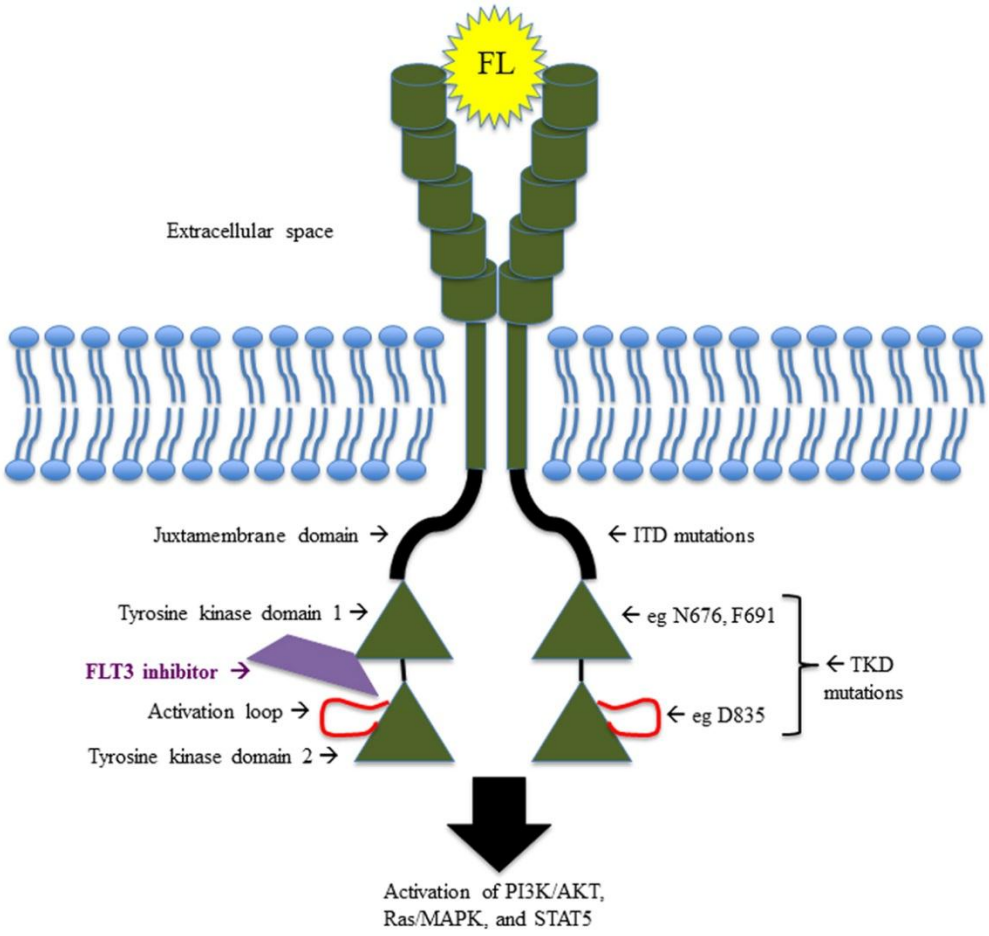


Figure 2. Schematic illustrating the structure and function of FLT3, including the sites of the most common activating mutations. *From Annesley CE, Brown P. The Biology and Targeting of FLT3 in Pediatric Leukemia. Front Oncol. 2014 Sep 23;4:263.*

PTPN11 activating mutations

Protein tyrosine standard phosphatase non-receptor 11 (*PTPN11*) gene is located on chromosome 12 and encodes a cytoplasmic protein tyrosine phosphatase (PTP) called SHP-2. SHP-2 participates in signal transduction downstream of growth factor, cytokines (KIT-ligand, IL3, IL6, granulocyte-macrophage colony-stimulating factor (GM- CSF) and EPO), hormones, and cell adhesion molecules. It have an important role in the RAS signaling pathway and it is highly expressed in hematopoietic.[37] Germ-line mutations in *PTPN11* have been reported in an autosomal dominant disorder characterized by facial dysmorphy, skeletal malformations, an broad spectrum of heart defects (Noonan's syndrome) [38]. Moreover, Somatic mutations in *PTPN11* are found in about 35% of juvenile myelomonocytic leukemia (JMML) cases[37]. Approximately 4.4% (31/702) of pediatric AML cases show *PTPN11* mutations and the 22% of these carry also mutation in *FLT3* and *RAS* genes[39]. In adults the frequency seems to be lower since only 2.6% (9/340) [37]. No significant prognostic relevance has been so far established for mutation in *PTPN11* gene in pediatric or adult AML[39].

NPM1 activating mutations

Recurrent mutations have been reported also in the *nucleolar phospho-protein B23* (*NPM1*) gene in both pediatric and adult AML, with a incidence of 5% - 10% and 35%, respectively[40]. NPM1 is a ubiquitously expressed nucleolar phosphoprotein that continuously shuttles between the nucleus and cytoplasm with predominant nucleolar localization. It is involved in the prevention of protein aggregation in the nucleolus and regulates the assembly and transport of pre-ribosomal particles through the nuclear membrane and DNA polymerase activity. Moreover, NPM1 is known to have a role in cell cycle progression, response to stress and oncogenic stimuli and regulation of the alternate reading frame protein (ARF)-p53 tumor suppressor pathway[41]. The mutations of NPM1 typically destroy the ability to shuttle from the cytosol to the nucleolus with a constitutive retention of NPM1 in the first, which can easily be detected by immunohistochemistry. The cytoplasmic localization of NPM1 mutated protein is probably critical for its putative role in leukemogenesis[36]. *NPM1* mutations are significantly associated with both pediatric and adult cytogenetically normal AML (CN-AML) (35% and up to 60%, respectively)[40] and with *FLT3*- ITD, *FLT3*-TKD mutations. Conversely, decreased prevalence of *NPM1* mutations is detected in presence of *CEBPA* mutations and MLL-partial tandem duplication (PTD). Prognostically, CN-AML patients carrying *NPM1*-mutated/*FLT3*-ITD negative have a significantly better response to induction therapy, better EFS and better overall-

survival (OS) than those without NPM1 mutation, although the presence of also FLT3 mutations predicts for a worst outcome[42] .

Janus kinase 2 (JAK2) activation mutations

Playing a key role in signal transduction initiated by multiple growth factor-receptors (type II TKR, such as erythropoietin receptor [EPOR], thrombopoietin receptor [TPOR], etc…) and cytokines, Janus Kinase 2 (JAK2) is a cytoplasmic tyrosine kinase with a crucial role in hematopoiesis[43]. The recurrent JAK2 V617F mutation causes replacement of a key valine residue leading to a constitutive activation of this kinase and the consequence hyperactivation of downstream signaling components, STAT5, PI3K/Akt, ERKs. Over 90% of polycythemia vera (PV) cases, 50% of essential thrombocythemia (ET), and in about 70% of patients with AML secondary to other myeloproliferative disorders show the JAK2 V617F mutation, although, conversely, it is a rare event (1.6%) in *de-novo* AML.[36,44–46]

1.1.4.2 Genetic lesions impairing the myeloid differentiation (Class II)

Accordingly to the "two-hits" model for leukemogenesis, mutations conferring pro-survival/proliferative advantage need to be associated with mutations that impair the ability to differentiate of the hematopoietic progenitors for a complete outset of acute myeloid leukemia. Belonging to this class of lesion, chromosomal translocations (as well as inversion, duplication and loss of whole chromosomes) affecting genes coding for transcriptional regulators, components of the transcriptional activation complex, and chromatin modifiers are an hallmark of pediatric AML with over the 75% patients showing these alterations at diagnosis[47,48].

Core-binding factor (CBF) alterations

The core-binding factor (CBF) is a heterodimeric transcription factor crucial for myeloid differentiation, and it has been demonstrated that both the component of this complex (coded by *AML1* and *CBF* genes), are essential for normal hematopoiesis[49]. Indeed, disruptions of the AML1/CBF complex could cause stop of myeloid differentiation and subsequent leukemic transformation. Recurrent and multiple chromosomal translocations as well as non-sense/frame-shift mutations and duplications affecting the *AML1* (also known as *RUNX1*) and *CBF* genes have been described in AML. The most frequent chromosomal alterations involving these two

genes are the t(8;21) (generating the fusion protein AML1/ETO) and the inv(16)/t(16;16) (generating the fusion protein CBFB/SMMHC) which are present in approximately 10% and 5% of AML patients, respectively[50]. Although through two different mechanisms, the result of these fusions proteins is the creation of an inactive CBF complex who is not able to start the transcription of important genes target required for the myeloid differentiation[50] (Figure 3). Moreover, the *AML1* gene is also involved in another translocation t(16;21) in which it results to be *in-frame* fused with the gene *CBFA2T3* localized on chromosome 16. Interestingly, *CBFA2T3* gene encode for a protein of the same family of ETO and, functionally, these two chimeric protein could be considered equal[51]. Alterations of the CBF are prognostic markers of good prognosis, and a specific subtype of AML, known as core-binding factor AML (CBF-AML), has been define base on that. Nevertheless, the CBF-AML patients are predicted to have a good outcome as long as there are no other negative prognostic markers such as *FLT3*-ITD or *KIT* mutations[48].

CCAAT/enhancer binding protein α (C/EBPα) loss of function mutations

Another important transcriptional factor required for the myeloid differentiation is encoded form the *CCAAT/enhancer binding protein α* (*C/EBPα*) gene, localize on the chromosome 19[36]. Frameshift mutations in the N-Termini portion of the gene, as well as small InDels in the C-Termini have been reported in approximately 4% - 6% of AML patients and, typically, both mutations could co-exist. Interestingly, the presence of *C/EBPα* loss-of-function lesions is typically associated with a better outcome, likewise the previously reported CBF-AML cases.[52]

PML/RARα and acute promyelocytic leukemia

Exclusively associated with the acute promyelocytic leukemia (APL) the extremely recurrent chromosomal translocation t(15;17) generates gene fusions involving the *retinoic acid receptor α* (*RAR*) gene and several different partner genes. The most common fusion is *PML/RARα* in which PML inhibits the function of RARα by a dominant negative effect mediated by the recruiting of co-repressor complex, likewise described in the AML1/ETO, CBFB/SMMHC and ETV6/AML1 fusions[53]. Remarkably, over the 90% of PML/RARα-positive cases could be successfully treated with all-trans retinoic acid (ATRA) that is able to induce the differentiation of this leukemic cells into mature granulocytes[48].

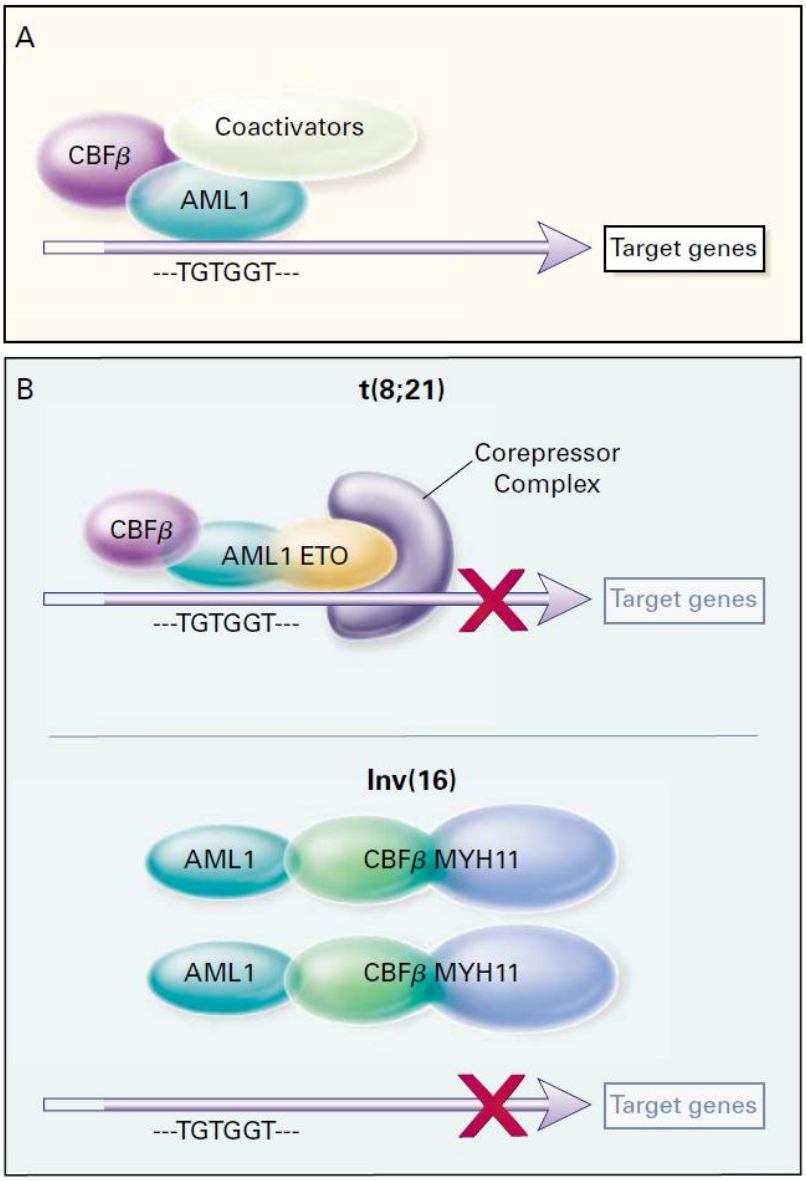


Figure 3. CBF complex genes rearrangements in acute myeloid leukemia (AML). (A) In normal cells, heterodimeric AML1-CBFβ transcription-factor complex binds to the DNA sequence TGTGGT in the transcriptional regulatory region of AML1-regulated target genes and activates transcription through the recruitment of coactivators. **(B)** In AML cells with the t(8;21) translocation, the N-terminal part of AML1 fuses with the C-terminal portion of ETO. The resultant chimeric protein continues to interact with CBFβ and to bind to the core enhancer sequence; however, ETO recruits a nuclear corepressor complex and results in the dominant repression of AML1-regulated target genes. Similarly, the CBFβ-MYH11 chimeric protein encoded by the inv(16) mutation continues to interact with AML1; however, instead of allowing AML1 to interact with DNA, this chimeric protein recruits AML1 into functionally inactive complexes in the cytoplasm. *From Löwenberg B, Downing JR, Burnett A. Acute myeloid leukemia. N Engl J Med. 1999 Sep 30;341(14):1051-62.*

Mixed lineage leukemia (MLL) gene alterations

Localized on chromosome 11q23, *the mixed lineage leukemia* (*MLL*) gene, also called *ALL-1*, *HTRX* or *HRX*, is one of the most frequently rearranged gene in AML. About 20% of pediatric AML patients show alterations of *MLL* gene, and the incidence rise up to 40% - 65% within the FAB subtype M4/M5, in the *infant*-AML, and in tAML secondary to topoisomerases II inhibitor treatment[48,54]. Encoding for a histone methyltransferase that plays an essential role in early development and hematopoiesis, the *MLL* gene could be involved in chromosomal translocations that lead to *in-frame* fusion with more than 60 different partner genes[55] (Figures 4, 5). Additionally, partial tandem duplication (PTD) within the *MLL* gene were also reported (Figure 5). The predicted effect in both cases is the deregulation of the function of the MLL protein complex that results in an aberrant expression of target genes crucial for the correct hematopoiesis, such as the *HOX* genes. Generally, chromosomal abnormalities involving *MLL* gene are indicators of poor prognosis of the disease[55].

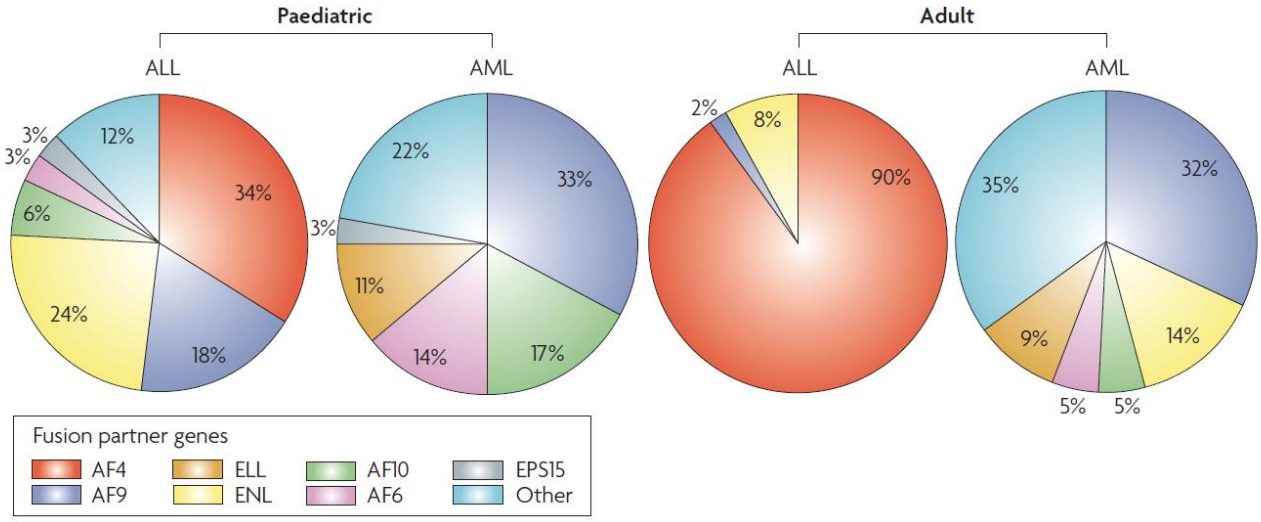


Figure 4. Distribution of major *MLL* fusion partner genes in *de-novo* childhood and adult leukemias. Mixed lineage leukaemia (MLL) rearrangements are found in approximately 5% of acute lymphoblastic leukemias (ALL), approximately 5–10% of acute myeloid leukemias (AML) and virtually all cases of mixed lineage (or biphenotypic) leukemias (MLL). Major MLL fusion partner genes are AF4, which is predominantly found in ALL; AF9, which is predominantly found in AML; and ENL, which is found in both ALL and AML. *From : AV Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development Nature Reviews Cancer. 2007; 7, 823-833.*

OTT/MAL and acute megakaryoblastic leukemia

Specifically recurrent in the pediatric FAB M7 AML, the translocation t(1;22) has been overall reported in approximately 3% of childhood AML patients. Among them, almost all the patient are *infant* (< 18 months of life). Due to this chromosomal rearrangement, the *RNA Binding Motif Protein 15* (*RBM15*, also named *OTT*) gene is *in-frame* fused with the *Mitogen-Activated Protein Kinase Kinase* (*MLK1*, also named *MAL*) gene. The result is the expression in the leukemic cells of a chimeric OTT/MLK1 fusion protein which de-regulates the NOTCH1 signaling leading to the arrest of the differentiation[56]. The presence of this genetic lesions is typically associated with a worst outcome[48].

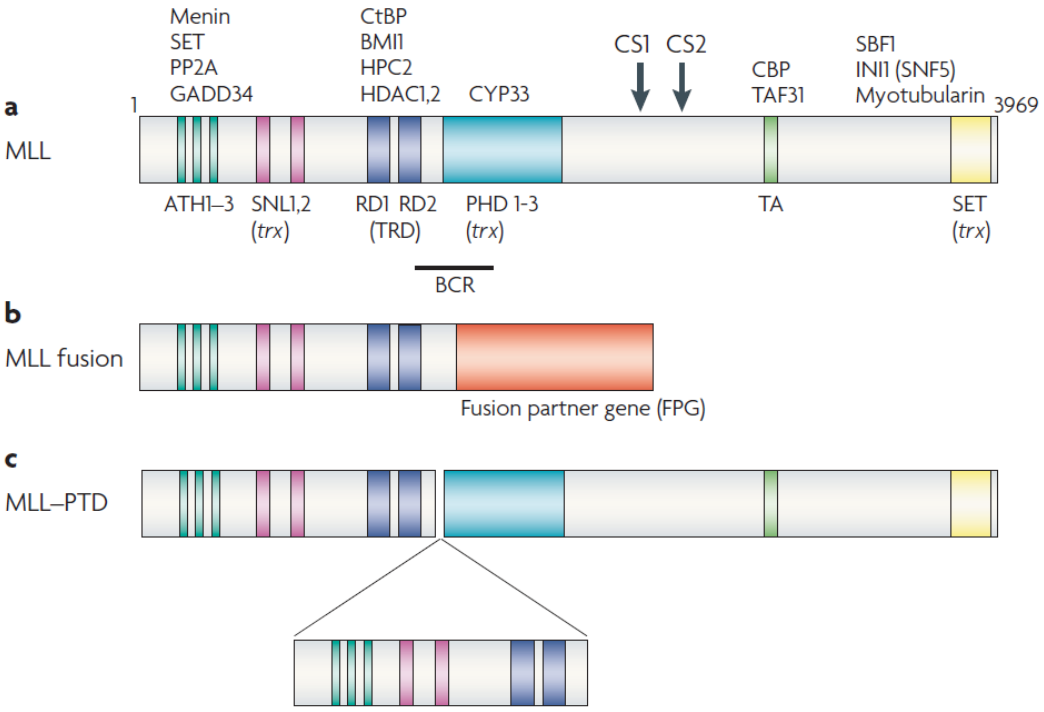


Figure 5. Schematic representation of the MLL protein and MLL fusions. **(A)** The mixed lineage leukaemia (MLL) gene is approximately 89 kb long, consists of 37 exons14,and encodes a 3,969 amino acid nuclear protein with a complex domain structure (unique domains are highlighted). The mature MLL protein consists of two non- covalently associated subunits (MLLN (300 kDa) and MLLC (180 kDa)) produced by cleavage of nascent MLL by taspase 1 after amino acid residues 2,666 (cleavage site 1 (CS1)) and 2,718 (CS2)18. Proteins that bind to specific domains are noted above each domain. The N-terminus contains three short AT-hook motifs (ATH 1–3), which are thought to mediate binding to the minor groove of AT-rich genomic DNA sequences. There are two speckled nuclear localization sites (SNL1 and SNL2) immediately C-terminal to the AT-hooks that are followed by a transcriptional repression domain (TRD) consisting of two functional subunits, RD1 and RD2. RD1 contains a DNA methyltransferase (DMT) homology domain that includes a CxxC zinc-finger motif that may recruit proteins such as HPC2 and the transcriptional co-repressor CtBP34. RD2 recruits histone deacetylases

HDAC1 and HDAC2 (REF. 34). The plant homology domain (PHD) zinc-finger motifs may mediate binding of the cyclophilin, CYP33, and potentially other proteins. The transcriptional activation (TA) domain recruits the transcriptional co-activator CBP (CREB-binding protein) and precedes a C-terminal SET (Su(var)3-9, enhancer-of-zeste, trithorax) domain that possesses histone H3 lysine 4 (H3K4) methyltransferase activity22,28 and is structurally homologous to Drosophila melanogaster trithorax. The breakpoint cluster region (BCR) spans exons 8–13. **(B)** Structure of MLL fusion proteins generated by MLL translocations. A typical MLL fusion protein contains the N terminus of MLL encoded by the first 8 to 13 exons and the C terminus of one of over 50 fusion partner genes (FPGs). **(C)** A unique MLL rearrangement results in MLL–partial tandem duplication (MLL–PTD). MLL–PTD contains a varied number of exons 5 to 12 duplicated and inserted before exon 11 or 12.. *From : AV Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development Nature Reviews Cancer. 2007; 7, 823-833.*

1.1.5 Prognostic markers and risk stratification

The heterogeneity in response to therapy and overall outcome of the different type of pediatric AML have always pushed for the identification of prognostic markers useful for stratifying the patients in the correct risk classes[4]. Nowadays, due to the novel insight, traditional prognostic markers such as age ad diagnosis, weight, white blood count, FAB subtype, have been replaced by novel, and more reliable markers based on cytogenetic and molecular features, as well as flow cytometric assessment of minimal residual disease (MRD) [4]. Table 3 summarizes the most recurrent genetic abnormalities identified in pediatric AML that have a prognostic relevance, together with those likely to have clinical relevance, and others with unknown (but potential) relevance.

Therapy response and MRD monitoring

Even if several cytogenetic and genetic prognostic markers are now available, still a large number of children with AML lack clinically informative karyotypes and are not eligible to risk-based therapy allocation based solely on cytogenetic subgroups.[15] Overall, clinically significant cytogenetic lesions and somatic mutations account for only about 35% of pediatric AML[15]. Furthermore, the recently discovered recurrent mutations affecting *Isocitrate Dehydrogenase 1* [*IDH1*], *Isocitrate Dehydrogenase 2* [IDH2] and *DNA (cytosine-5-)-methyltransferase 3 alpha* [*DNMT3A*] genes in adult AML (demonstrated to have a prognostic relevance) are no present in the pediatric AML[57–59]. So far, despite the constant increase in the knowledge of AML mutational landscape, the most important prognostic factor in AML is the response to therapy

assessed by MRD[60]. RNA-based polymerase chain reaction (PCR) analysis of leukaemia-specific gene fusions, quantitative PCR (qPCR), and flow cytometric detection of aberrant immunophenotypes[60] are the most reliable methods through with it is possible to monitor the response to the treatment and early-reveal the rise of a relapse. Although RT-PCR detection of fusion transcripts is sensitive to a level of 0.01–0.001%, it can be used in only about 50% of cases. Moreover, persistent expression of some gene fusion (such as *AML1/CBFA2T3* and *CBFβ/MYH11*) can be present also in patients who are in long-term remission, indicating the possibility to detect false-positive. Conversely, the sensitivity of flow-based MRD assays is only 0.1–0.01%, but this technique can be applied to more than 90% of cases[6]. In the ongoing Children's Oncology Group (COG) *de-novo* AML trial, the combination of diagnostic molecular risk factors with the post-induction response assessment by MRD detection has allowed risk assessment in all patients with AML, demonstrating the ability to correctly classify even those patients with clinically non-informative karyotypes[15] (Figure 6) .

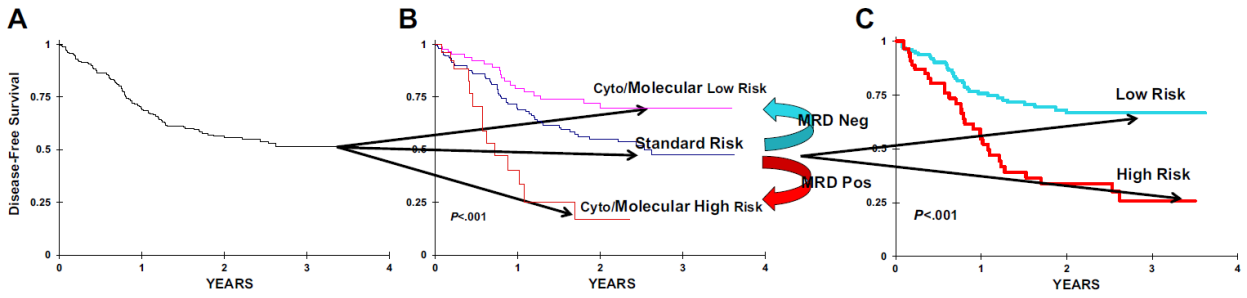


Figure 6. Newly devised risk stratification for AML. (A) Overall disease-free survival for all patients is presented. **(B)** Incorporation of known cytogenetic and molecular risk factors creates a 3-tier risk-stratification schema, allocating 35% of patients to cyto/molecular high risk (10%) or low risk (25%), with 65% of patients without informative cyto/molecular markers remaining in the standard-risk group. **(C)** Incorporation of disease assessment by multidimensional flow cytometry after the initial induction allows for identification of risk groups within the standard-risk cohort, providing a 2-tier risk stratification schema by combining the cyto/molecular and minimal residual disease (MRD) data. *From Tarlock K, Meshinchi S. Pediatric Acute Myeloid Leukemia: Biology and Therapeutic Implications of Genomic Variants. Pediatr Clin North Am. 2015 Feb;62(1):75-93.*

Table 3. Genetic lesions in pediatric acute myeloid leukemia

Karyotype	Affected genes	%[δ]	Clinical significance
t(8;21)(q22;q22)	*RUNX1-RUNX1T1*	15	Favourable prognosis; Not candidates for HSCT
inv(16)(p13.1;q22) t(16;16)(p13.1;q22)	*CBFB-MYH11*	10	Favourable prognosis Not candidates for HSCT
−7	Unknown	1	Poor prognosis
11q23	*MLL* rearrangements	20	
	MLL-MLLT11		Favourable prognosis
t(1;11)(q21;q23)	*MLL-MLLT3*	1	Intermediate prognosis
t(9;11)(p12;q23)	*MLL-MLLT4*	8	Poor prognosis
t(6;11)(q27;q23)	*MLL-MLLT10*	1	Poor prognosis
t(10;11)(p12;q23)		1	Intermediate prognosis
Others		9	
t(1;22)(p13;q13)	*RBM15-MKL1*	1	Megakaryoblastic leukaemia; Unknown prognosis
t(6;9)(p23;q34)	*DEK-NUP214*	1	Poor prognosis[γ]
t(8;16)(p11;p13)	*KAT6A-CREBBP*	1	Poor prognosis[γ]
t(16;21)(q24;q22)	*RUNX1-CBFA2T3*	1	Poor prognosis[γ]
Normal[ε]	*FLT3*-ITD	12	Poor prognosis if high ratio of mutant to wild-type allele; May benefit from HSCT or treatment with FLT3 inhibitors
Normal[ε]	*NPM1*	8	Favourable prognosis except in cases with *FLT3*-ITD
Normal[ε]	*CEBPA*	5	Favourable prognosis probably limited to cases with biallelic mutations
	WT1		
Normal[ε]	Mutation	10	Unknown
	SNP rs16754	25	May be associated with favourable outcome
	IDH1 and *IDH2*		
Normal[ε]	Mutation	4	Unknown
	IDH1 SNP rs11554137	10	Unknown
Normal[ε]	*RUNX1*	Rare	Unknown
Normal[ε]	*TET2*	Rare	Unknown
Normal[ε]	*DNMT3A*	Rare	Unknown
t(15;17)(q22;q12)	*PML-RARA*	NA	Observed only in APL; Favourable outcome

HSCT, haematopoietic stem-cell transplantation; ITD, internal tandem duplication; SNP, single nucleotide polymorphism; NA, not available; APL acute promyelocytic leukaemia. δ, Percent of non-APL cases with each abnormality. Note that some alterations co-exist, such as NPM1 mutation and FLT3-ITD. γ, The poor prognosis of these rare translocations has been firmly established only in adult AML. ε, These mutations often occur in cases with normal karyotypes, but are also seen in cases with other abnormalities. *From Rubnitz JE, Inaba H. Childhood acute myeloid leukaemia. Br J Haematol. 2012 Nov;159(3):259-76.*

1.2 Mutational landscape of cytogenetically normal acute myeloid leukemia (CN-AML) as drawn by next-generation sequencing, state of art

The first insights into AML molecular pathogenesis originated from cloning the breakpoints of recurrent chromosomal translocations detected by conventional cytogenetics[61]. The identification of recurrent chromosomal rearrangements and the association of those with clinical data have been crucial for risk stratification of patients and development of risk-adapted therapies[62]. However, approximately 20% of pediatric and 43% of adult AML patients do not show any known chromosomal aberrations and, so far, these are generically define as cytogenetically normal AML (CN-AML) [6] (Figure 7). Historically belonging to the intermediate-risk class, the CN-AML patients have always showed a great variability in therapy response and overall outcome, underlining the presence of other genetic changes not detectable by conventional cytogenetic analysis but relevant for pathogenesis, dynamics, and outcome of AML[61]. With the aim of investigating the mutational landscape of CN-AML, lately, several studies have taken advantage of the novel next-generation sequencing (NGS) technologies, and novel important insights in the molecular pathogenesis of CN-AML have been achieved, especially in adults[57,63–65].

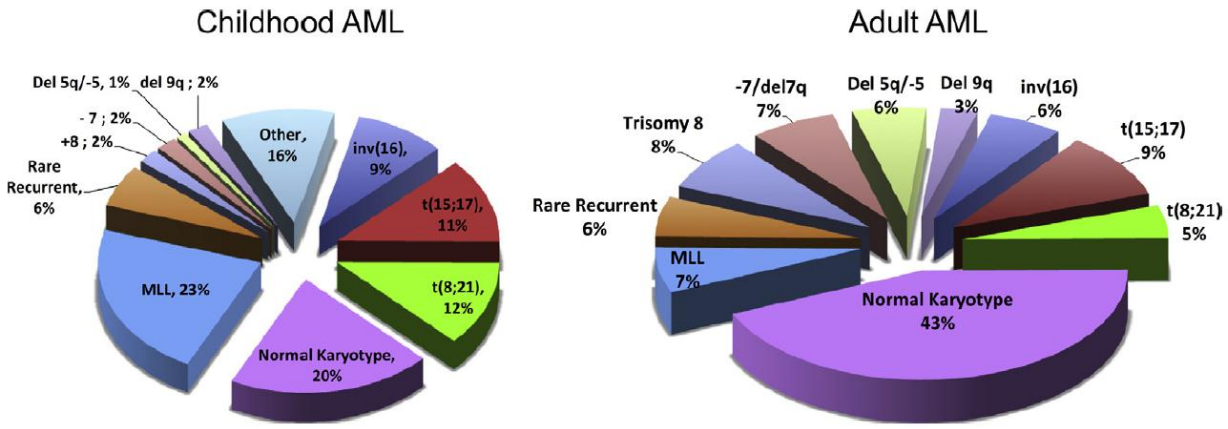


Figure 7. Karyotypic alterations in both pediatric and adult AML. Approximately 20% of pediatric and 43% of adult AML patients do not show any known chromosomal aberrations and, so far, these are generically define as cytogenetically normal AML (CN-AML). *From Tarlock K, Meshinchi S. Pediatric Acute Myeloid Leukemia: Biology and Therapeutic Implications of Genomic Variants. Pediatr Clin North Am. 2015 Feb;62(1):75-93.*

1.2.1 The modern genomic era and AML gene discovery[61]

In 2008, only 8 years later the conclusion of the Human Genome Project, the first human cancer genome was sequenced. The DNA from leukemic blasts of an adult patient with CN-AML (FAB M1) was analyzed by whole-genome sequencing in two years[63]. Subsequently, a second AML M1 genome[64] (in 2009) and an APL M3 genome[66] (in 2011) were analyzed within the following three years (2009-2011), in no more than six weeks in the latter case.

1.2.1.1 Re-definying the molecular profile of adult AML

Beyond the single somatic mutations identified by the above mentioned studies, two important aspects need to be here considered. First, these studies described for the first time that in adult AML, somatic mutations are frequent in genes encoding for epigenetic modifiers (eg. *TET2*, *DNMT3A*, *Additional Sex Combs Like 1*[*ASXL1]*), revealing a novel class of genes commonly altered in AML (Class III) other than pro-survival and impairing myeloid differentiation as described in the "two-hit" model of leukemogenesis[61] (Figure 8) . Additionally, these studies represented a technical paradigm of how the next-generation sequencing could be successfully used for mutations discovery studies[67]. Indeed, starting from here, in the last years we were witness of a constantly increase of studies in which next-generation sequencing has been used for mutation discovery in leukemia as well as in other cancers. The decrease in time and costs will soon allow a broadly use of this technology for routinely diagnostic and in clinics, paving the way to a complete personalized therapy in cancer treatment. So far, the bigger effort to investigate the mutational landscape of adult CN-AML has been coordinated by *The AML Cancer Genome Atlas project (TCGA)* study in which the most modern sequencing and genotyping techniques have been used to investigate the somatic genome of 50 patients with *de-novo* AML and the somatic exome of additional 150 cases[68]. Briefly, this study discovered that adult AML genomes have far fewer mutations compared to other types of adult cancers, with an average of 13 mutations per individual. A total of 23 genes were reported as being substantially altered (*NPM1*, *FLT3*, *DNMT3A*, *IDH1* or *IDH2*, *NRAS* or *KRAS*, *AML1*, *TET2*, *TP53*, *C/EBPα*, *WT1*, and *KIT*) [68] (Figure 9) . Additionally, the study confirmed the recurrence of previously reported chromosomal alterations, such as t(15;17), t(8;21), inv(16), abn11q23, monosomy 5, and monosomy 7. Remarkably, the study was also able to identified novel translocations below

the detection level of conventional cytogenetics, underlining the important role of fusion genes in the development of AML[61] (Figure 10) .

Analysis	Before 2008 Cytogenetic and molecular genetic analysis	2008–12 Next-generation sequencing approaches	From 2013 The Cancer Genome Atlas project	Prevelence in AML (%)
Functional groups	Class I: activated signalling—eg, *FLT3*, *KIT*, *RAS* mutations	Class I: activated signalling—eg, *FLT3*, *KIT*, *RAS* mutations	Class 1: transcription factor fusions—eg, t(8;21), t(16;16), t(15;17), *MLL* fusions	18%
			Class 2: nucleophosmin 1, *NPM1* mutations	27%
			Class 3: tumour suppressor genes—eg, *TP53*, *WT1*, *PHF6* mutations	16%
	Class II: transcription and differentiation—eg, t(8;21), t(16;16), t(15;17) *CEBPA* mutations	Class II: transcription and differentiation—eg, t(8;21), t(16;16), t(15;17), *CEBPA*, *RUNX1* mutations	Class 4: DNA-methylation-related genes: DNA hydroxymethylation—eg, *TET2*, *IDH1*, *IDH2* DNA methyltransferases eg, *DNMT3A*	44%
			Class 5: activated signalling genes—eg, *FLT3*, *KIT*, *RAS* mutations	59%
			Class 6: chromatin-modifying genes, eg, *ASXL1*, *EZH2* mutations, *MLL* fusions, *MLL* partial tandem duplications	30%
		Class III: epigenetic modifiers—eg, *TET2*, *DNMT3A*, *ASXL1* mutations	Class 7: myeloid transcription factor genes—eg, *CEBPA*, *RUNX1* mutations	22%
			Class 8: cohesin-complex genes—eg, *STAG2*, *RAD21*, *SMC1*, *SMC2* mutations	13%
			Class 9: spliceosome-complex genes—eg, *SRSF2*, *U2AF35*, *ZRSR2* mutations	14%

Figure 8. Complementation groups of genetic changes in AML. *From Meyer SC, Levine RL. Translational implications of somatic genomics in acute myeloid leukaemia. Lancet Oncol. 2014 Aug;15(9):e382-94.*

Based on the overall finding from the TCGA study, several interesting comments can be made: i) the traditional concept of Class I and Class II mutations for leukemogenesis has been radically modified (Figure 8), considering that just 59% of AML cases analyzed showed mutations in oncogenic signaling components (which were previously believed to be essential for leukaemic transformation) [61]. ii) This study, as well as the previous NGS studies on adult AML, strongly demonstrate the recurrence of mutations in epigenetic factor genes such as *TET2* and *DNMT3A*, defying a novel class of genes that could be mutated in adult AML with possible implications in prognostic and/or therapeutic. iii) Rare but not-random mutations in cohesion genes and genes encoding for component of the spliceosome-complex (so far not associated with this pathological setting) has been described, paving the way to the definition of additional mutations class in adult AML.

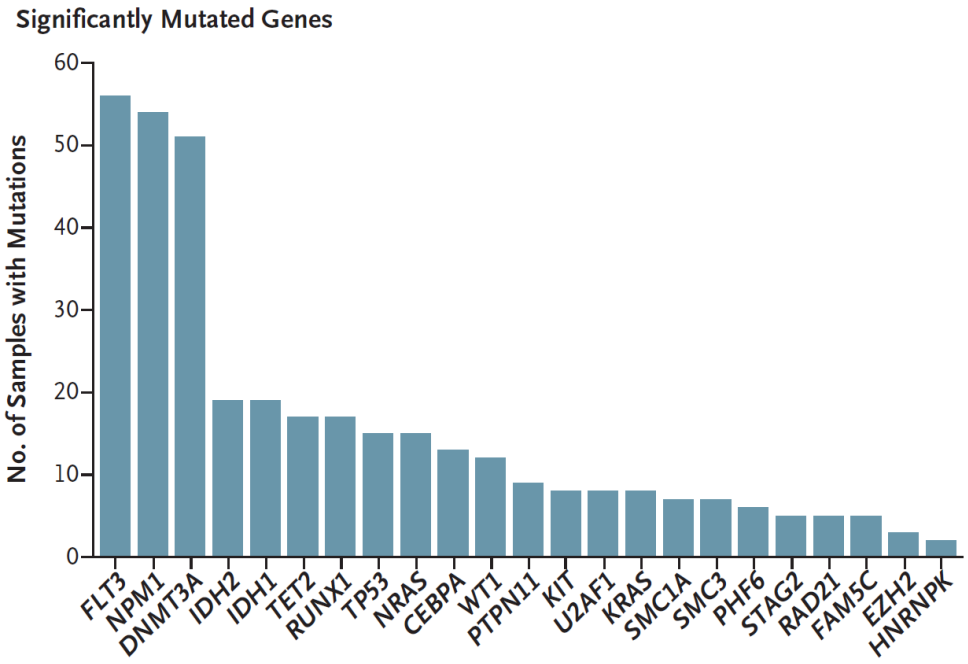


Figure 9. Characterization of mutations in adul AML. Significantly mutated genes found by TCGA study in adult AML. *From Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013 May 30;368(22):2059-74.*

1.2.1.1 The landscape of genetic alterations in childhood AML

Despite the insights recently achieved by the TCGA study on the mutational profile of adult AML, and the confirmation of the recurrence and clinical utility of several somatic mutations in adult AML (such as *DNMT3A* and *IHD1*), these mutations appear to be rare or absent in childhood AML,[58,59,69] underlying once again the significant differences between AML in elderly and young patients[15] (Figure 11 and Figure 12). Based on that evidence, two different cooperative studies are ongoing with the overall aim to investigate the genomic and epigenomic profiles for the spectrum of childhood AML:[1,70] the St. Jude/Washington University Pediatric Cancer Genome Project (PCGP) (http://www.pediatriccancergenomeproject.org/site/), and the COG/National Cancer Institute (NCI) Therapeutically Applicable Research to Generate Effective Treatments (TARGET) AML initiative (https://ocg.cancer.gov/programs/target). Early data from these two studies have already provided new biological insights in pediatric AML. Among them, analyzing by RNA-seq the whole-transcriptome from leukemic cells of a cohort of pediatric

patients with acute megakaryoblastic leukemia (AML FAB M7), recently, the PCGP discovered a novel

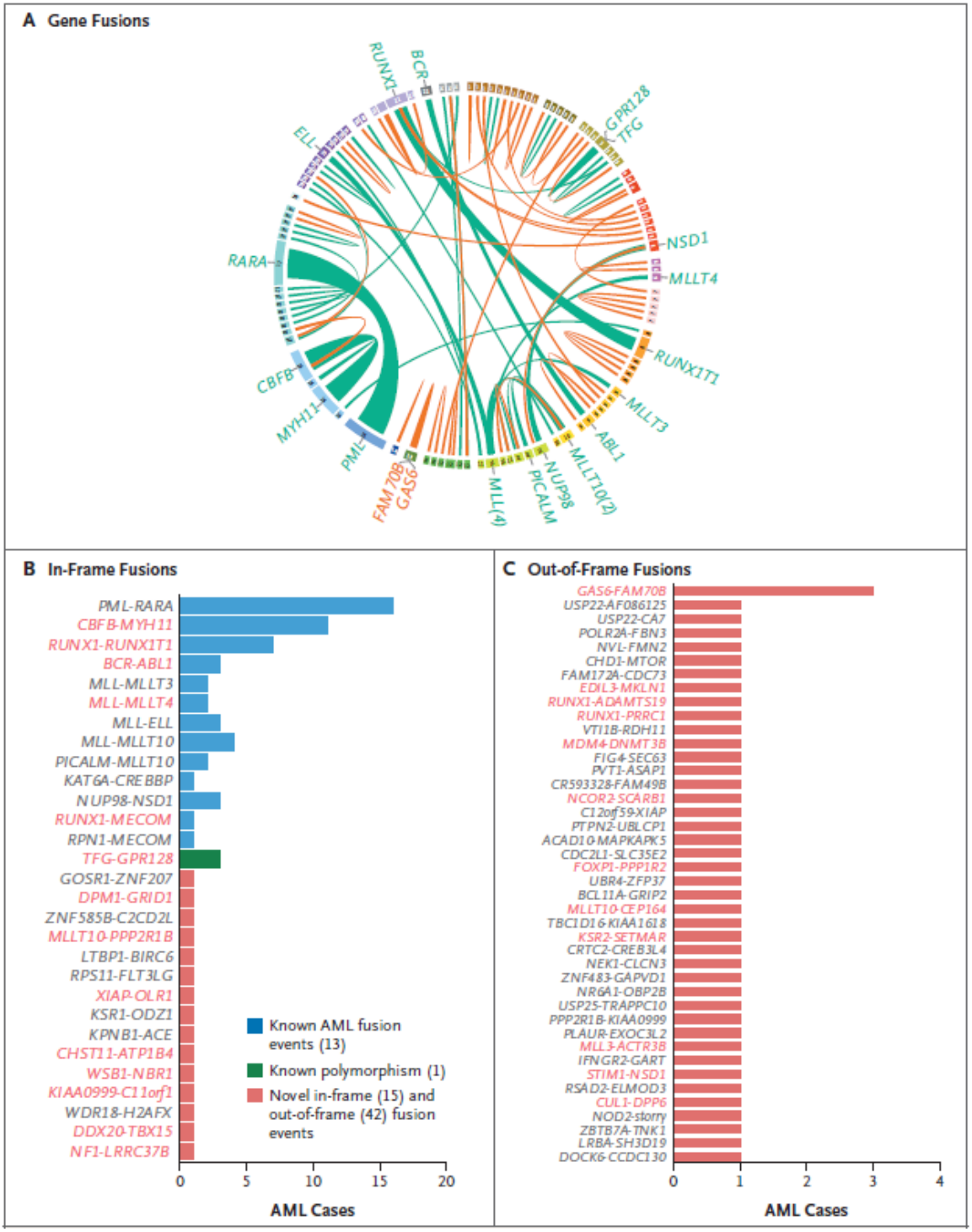


Figure 10. AML gene fusions. (A) is a plot created with the use of Circos software29 showing in-frame (green) and out-of-frame (orange) gene fusions detected in the AML cohort in the Cancer Genome Atlas

(TCGA) with the use of Trans-ABySS software. Ribbon widths are proportional to the frequency of a fusion event. Chromosomes are individually colored and are arranged clockwise from chromosome 1 to X, starting with chromosome 1 at 12 o'clock. No rearrangements involved the Y chromosome. The frequencies of in-frame and out-of-frame gene fusions are shown in **(B)** and **(C)**, respectively. For gene names shown in red, one of the partner genes in that fusion was found to be mutated in at least one other AML sample from this data set. On the basis of chromosomal aberrations and genomic variants annotated in the Mitelman database from the Cancer Genome Anatomy Project (CGAP) (http://cgap.nci.nih.gov/Chromosomes/Mitelman), all previously identified gene fusions are shown in blue, a single known polymorphic fusion is shown in green, and all novel events are shown in red. *From Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013 May 30;368(22):2059-74.*

cryptic inversion of the chromosome 16 leading to the *in-frame* fusion of the *CBFA2T3* and *GLIS2* genes. Remarkably, within the M7 subtype, this alterations seems to be extremely recurrence (about 30% of children with FAB M7 AML, not associate with Down syndrome) and predicts for a poorer outcome[71]. On the other side, the TARGT AML initiative, by the means of whole-genome sequencing and RNA-seq, analyzed an initial cohort of more than 200 children with AML (diagnosis, remission and relapse matched analysis), discovering an high frequency of *ETV6* mutations in childhood CN-AML (up to 6% of cases) which are significantly associated with a worst outcome. Moreover, this study also revealed a significant clonal evolution of the diagnostic predominant leukemic clone between diagnosis and relapse, with approximately a third of total mutations identified at diagnosis that persisted at relapse together with the acquisition of a lot of novel mutations[1]. Nevertheless, so far, the mutational landscape of pediatric CN-AML remains hazy, and additional coordinated efforts need to be made in order to fully characterize the molecular heterogeneity of this particular subtype of childhood leukemia. The rise of NGS has provided the tools for comprehensive interrogation of the cancer genome and the emerging data will be crucial to define new biomarkers and drug-targeted mutations, leading, ultimately, to make personalized medicine a reality in the near future[1,15].

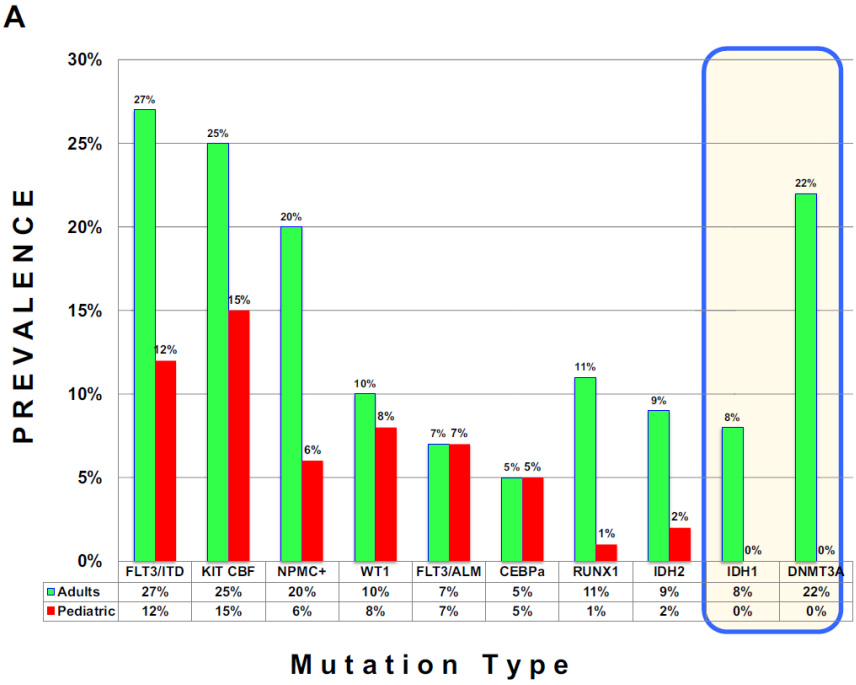


Figure 11. Somatic mutations in adults and childhood AML. (A) Prevalence of AML-associated mutations in pediatric versus adult AML, demonstrating lower incidence of mutations in pediatric AML. Bordered panel shows 2 newly discovered mutations in adults that are absent in pediatric AML.

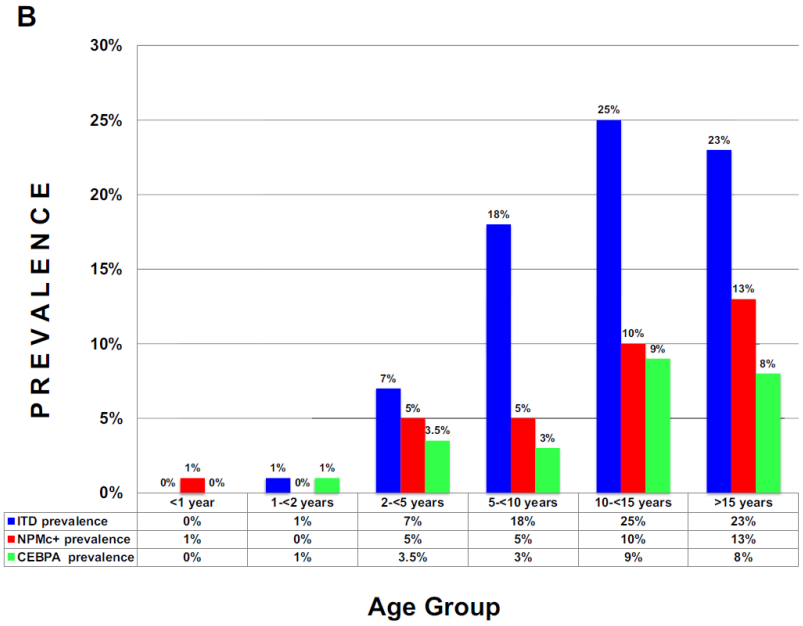


Figure 12. Age-based somatic mutations groups in adults and childhood AML. (B) Age-based prevalence of common AML-associated mutations.

AIMS

The overall aim of this research project was to investigate through RNA-seq the mutational landscape of pediatric cytogenetically normal AML patients negative for all the currently known somatic mutations reported in CN-AML.

The rationale to analyze this specific subtype of pediatric acute myeloid leukemia was driven by the fact that these patients do not show any cytogenetic or molecular aberrations so far known in AML, representing the “all-neg” AML subgroup.

Due to the lacking of any prognostic marker, “all-neg” AML patients were historically classified in the intermediate-risk class. However, these patients had always showed a great variability in response to therapy and overall outcome, underlining the presence of unknown genetic changes, not detectable by conventional analyses, but relevant for pathogenesis, dynamics, and outcome of AML.

The results obtained from this study are expected to reveal novel biological insights into the rise of pediatric AML, to define novel prognostic markers and putative therapeutic targets, and to prospectively ensure, a correct risk stratification and risk-adapted therapy also for the “all-neg” AML subgroup.

CHAPTER II – Study design and Methods

2.1 Patients

Patients analyzed either in the massively parallel sequencing screening or in the validation cohort are children with newly diagnosed *de novo* AML other than acute promyelocytic leukemia enrolled in the AIEOP AML 2002/01 Protocol [72]. The initial diagnosis of AML was centrally established according to morphology-FAB classification and immunophenotypic analysis at the laboratory of Pediatric Hematology of the University-Hospital in Padova. Chromosome analysis was performed on BM using standard laboratory procedures. The karyotypes were reported according to the International System for Human Cytogenetic Nomenclature (ISCN 2005). For fluorescence in situ hybridization (FISH) an MLL locus specific (LSI) dual color probe for 11q23 (Abbot-Vysis, Downess Grove, IL) was employed. This analysis was performed on metaphases and/or nuclei, when necessary, according to the manufacturer's instructions. Patients were all negative for any known cytogenetic aberrancy and recurrent genetic abnormalities involving *MLL*, *CBFB*, and *FLT3* genes. Main clinical characteristics of the 19 AML patients analyzed by RNA-seq are reported in Table 4.

2.2 Whole-transcriptome sequencing

2.2.1 RNA-seq libraries preparation

Total RNA was extracted from BM leukemia cells of CN-AML patients by TRIzol, following the manufacturer's protocol (Invitrogen, Karlsruhe, Germany); 250 to 1000 ng of total RNA was used for the synthesis of cDNA libraries with TruSeq RNA Sample Prep Kit v2 (Illumina, San Diego, CA) according to the manufacturer's recommendations. Briefly, by the use of magnetic beads, PoliA+ RNA have been selected and purified from total RNA. Then, through incubation for 4 minutes at 94°C, the purified PoliA+ RNA have been fragmented (average length of 160bp) and sequentially retro-transcribed to cDNA by SuperScript II Reverse Transcriptase (Invitrogen)

and Random Examers Primers (Illumina), according to Illumina's libraries preparation protocol. cDNA were then purified by Agencourt AMPure XP (Beckman Coulter) magnetic beads purification system, and directed through the ends-adenilation step.

Table 4. Table 4. Clinical characteristics of the 19 AML patients analyzed by RNA-seq

#	ID	Age (y)	Gender	FAB	WBC, x 109/mL	BM blast, % at diagnosis	HSCT (type)	Relapse (site)	Status
1	CN-AML#1	0.3	F	M5A	18000	88	AUTO	-	Vivo
2	CN-AML#2	0.3	M	M4	82190	60	MFD	-	Vivo
3	CN-AML#3	0.5	F	M7	22800	40	-	-	Morto
4	CN-AML#4	0.8	M	M2	22400	52	MUD	-	Vivo
5	CN-AML#5	1.4	M	M7	22400	60	MUD	-	Vivo
6	CN-AML#6	1.4	M	M0	35000	86	MUD	-	Vivo
7	CN-AML#7	1.5	F	M7	13370	90	MMUD	-	Vivo
8	CN-AML#8	0.8	M	M7	188000	28	MFD	Yes	Morto
9	CN-AML#9	15.5	M	M5A	99420	90	AUTO	-	Vivo
10	CN-AML#10	6.5	F	M1	7870	70	AUTO	Yes	Vivo
11	CN-AML#11	8.1	M	M1	3500	70	AUTO	Yes	Vivo
12	CN-AML#12	4.6	F	M1	9980	90	MMFD	-	Vivo
13	CN-AML#13	9	M	M0	1220	70	MUD	-	Vivo
14	CN-AML#14	8.9	M	M5	96180	60	MUD	-	Vivo
15	CN-AML#15	3.4	M	M2	12010	70	AUTO	Yes	Vivo
16	CN-AML#16	0.1	F	M2	16000	41	MUD	-	Vivo
17	CN-AML#17	1.4	M	M0	8470	70	AUTO	-	Vivo
18	CN-AML#18	2.9	M	M1	187900	90	AUTO	Yes	Morto
19	CN-AML#19	0.5	M	M0	34770	21	-	-	Vivo

BM, bone marrow; HSCT, hematopoietic stem cell transplant; AUTO, autologous; MUD, matched unrelated donor; MMUD, mismatched unrelated donor; MFD, matched family donor; MMFD, mismatched family donor.

After a second purification, DNA ligase I was used to link the Illumina adapters at the ends of each cDNA fragment by incubation of 1 hour at 16°C. The ligation products were then submitted for a PCR amplification step with primers that map on the adapters (Illumina) which ensure the enrichment of the only cDNA fragments that have been correctly linked to the adapters. Finally, one last purification step through Agencourt AMPure XP (Beckman Coulter) magnetic beads is perform to obtain the final libraries ready to be validated, quantified and then loaded for the clonal amplification and sequencing.

2.2.2 RNA-seq libraries validation and quantification

All the libraries have been validated and quantified accordingly to Illumina's protocol. The average length distribution of the cDNA fragments of a library has been evaluated through analysis by Chip DNA 1000 on Bioanalyzer 2100 (Agilent Technologies). The concentration of each library has been evaluated through two different methods: i) quantitative-PCR (qPCR) using primer mapping on the Illumina adapters linked at the end of each cDNA fragments with KAPA SYBR® FAST qPCR Kits (KAPA Biosystems); ii) fluorimetric analysis by Quant-iT™ PicoGreen® dsDNA Assay Kit (Invitrogen). An average of the concentration obtained by these two assays were taken as library final concentration and used to dilute the each library to 10 ng/µl as indicated in Illumina protocol.

2.2.3 Bridge-Amplification and sequencing on HiScanSQ (Illumina)

After the quantification, 100ng total of each library was pooled with the other libraries according to the pooling strategy and the adaptors used in the library preparation step. Then, through the cBot cluster amplification system / TruSeq PE Cluster Kit v3-cBot-HS (Illumina) the libraries were load into 8 lane of glass slide (flow-cell). By complementary oligo-adapters linked on the flow-cell, each fragment of the cDNA library was anchored and, through a "bridge" solid-phase amplification, clonally amplified to generate a cluster of cDNA all derived from the same starting fragment of the library. Sequencing by synthesis was performed on HiScanSQ sequencer (Illumina) at 75bp in paired-end mode. An average of 6.6 Giga Base (GB) and 87.6 million of reads has been generated for each patient. The average coverage obtained was 34X.

2.3 Bioinformatic analysis

75x2 paired-end reads obtained by RNA-seq were trimmed at 74 nucleotides after checking the average of the base quality at each position. The reads were aligned to the reference human genome HG19/GRCh37 with TopHat2/BowTie2[73] with a percentage of mapping reads ranging from 50% to 84% and a mean depth of 35x. For each sample the mean depth was computed considering the observed portion of reference genome covered instead of the whole human transcriptome:

$$Mean\ Depth = \frac{N * 74}{L}$$

where N is the number of mapped reads and L is the total size of the regions covered by at least one read. By this we purpose to estimate a different transcriptome size for each sample keeping out the regions not expressed and including the putative unknown transcript. The presence of gene fusions in the samples was investigated with three different bioinformatic software packages suited to detect chimeric transcripts from RNA-seq data: Defuse,[74] ChimeraScan,[75] and FusionMap.[76] Only the putative chimeric transcripts detected from all the alghotihms were selected for further validations.

2.4 RNA isolation, RT-PCR and Sanger sequencing

Total RNA was extracted using Qiagen RNeasy mini kit (Qiagen) as per manifacturer's instructions. Then quantity and quality of the isolated RNA was assessed using the Nanodrop (Thermo Scientific).

Table 5. List of the primers used for RT-PCR and Sanger sequencing

Amplicon target	Primer Fw	Tm	%GC	Primer Rev	Tm	%GC	Amplicon lentgh (bp)
CBFA2T3-GLIS2	CATGTCGGAGCTG CAGAAAG	58	55	CAAAGAGCTGGT TACACTTGGCC	61	52	353
NUP98-PHF23	GGCCCCAGTAGCT TTGACAGAT	60	55	GCAGACGAGAGA AAGTGGACCT	61	57	216
NUP98-KDM5A	CCAGCAGCACATC AATAGTCTAACA	59	44	GCTCCTTTGATTT GTCTGCACC	60	50	278
PRDM16-SKI	GCTGCTTCTGGAC TCAAGGAGG	61	59	TATTGGAAGAGC CGGCCAAG	61	55	210
MYB-GATA1	AGCACAAAATGTC TCCAGTCATGTT	60	40	TCCTTCCGCATGG TCAGTG	59	58	351
MAML1-SQSTM1	CAGTACCAAGACC CGACACAAG	58	55	ACACAAGTCGTA GTCTGGGCAG	58	55	303
RUNX1-USP42	GGAGCTTGTCCTT TTCCGAGC	61	57	GCATGACATGTTT TGGAGTGTTC	58	43	364
DHH-RHEBL1	TCCACTACGAAGG CCGTGC	61	63	GGATCGTAGCCTT CCGAGAAC	59	57	205
TNIP1-PDGFRB	CCGGAATACACCT GGCGTCTAC	62	59	CCATCGGATCTCG TAACGTGG	61	57	216

Typically, 500-1000 ng of total RNA was retrotrascribed to cDNA using oligo-dT and random hexamers by Transcriptor First Strand cDNA Synthesis kit (Roche Diagnostics). RT-PCR reaction was performed according through the AmpliTaq GOLD kit (Applied Biosystems). PCR

amplicons were purified using a Microcon centrifugal filter (Millipore Corporation) and sequenced according to the BigDye terminator v 3.1 Cycle Sequencing kit (PE Applied Biosystems) on an Applied Biosystems 310 analyzer. BLAST (http://www.ncbi.nlm.nih.gov/BLAST) and CLC (CLC Bio - Quiagen) software were used for the analysis of sequence. A list of the primers used for RT-PCR and Sanger sequencing is reported in Table 5.

2.5 Gene expression analysis from RNA-seq data

The mapped reads obtained with TopHat2/BowTie2 pipeline were processed with SAMtools[77] in order to remove the potential optical or PCR duplicate (function "rmdup") and then the count of the mapped reads for each hg19 gene was performed by applying the Python package "htseq-count" (http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html). Gene annotations were derived from Ensembl Release 70 (January 2013). The differentially expressed genes were determined with edgeR, a R-bioconductor package suitable for analyzing RNA-seq data.[78] Three different comparisons were performed, corresponding to all the possible couples among the three groups of patients: 1) *DHH-RHEBL1*- positive and *CBFA2T3-GLIS2*-positive (N=2); 2) *DHH-RHEBL1*-negative and *CBFA2T3-GLIS2*-positive (N=2); 3) CN-AML (N=4). For each comparison, the complete set of genes, with the corresponding mapped reads count, was firstly reduced in order to consider in our analysis only the genes with count-per-million (CPM) > 3 in more than 2 samples. Then, adopting a statistical method based on the negative binomial distribution, the significance of the differences between the normalized reads count was determined for each gene. Differences with $P<0.05$ were considered to be statistically significant. The Multi Experiment Viewer (MeV) tool (http://www.tm4.org/mev.html) was used to visualize the expression data. Expression levels of *PRDM16* and *USP42* genes were obtained from the RNA-seq data according the same bioinformatics analysis described above. Graphpad Prims (GraphPad Software) was used to visualize the expression data.

2.7 Quantitative (q) PCR for *USP42* and *PRDM16* genes

Expression levels of *USP42* were measured by qPCR. Total RNA was obtained from BM blasts, and $CD34^+$ haematopoietic cells; 500ng total RNA was reverse-transcribed to single-stranded cDNA using the Transcriptor first strand cDNA synthesis kit (Roche Diagnostics) with oligo-dT primers (2.5 μM). Gene-specific primers amplifying *USP42* wild-type (wt) allele, and *USP42* wt plus fused variant (wt+fus) were designed with Primer Express 3.0 Software (Applied Biosystems) and qRT-PCR was performed using FastStart Sybr Green (Roche Diagnostics) on the LightCycler 480 apparatus (Roche Diagnostics). DDCt method was used to quantify gene product levels relative to two housekeeping genes, *GAPDH*, and *ATPS*. qPCR primer sequences were as follows: *USP42*-wt_Fw,5'-CTGGGGTGCTGTGTCTTCAT-3';*USP42*-wt_Rev,5'-GCTGCATTGGCAAAACAGGT-3';*USP42*-wt+fus_Fw,5'-TGCAAGGGCGTTTCAGATACT-3'; *USP42*-wt+fus_Rev,5'-GTTCCGGCTTCACAAACTGC-3'. All qPCR primer pairs spanned intron/exon boundaries. *USP42* alleles (wt and wt+fus) expression levels were analyzed in the t(7;21)-positive sample, in CN-AML cases (N=4), and in $CD34^+$ haematopoietic cells (N=1). q-PCR analysis was also performed in order to detect differential expression of the *PRDM16* gene in the t(7;21)(p22;q22) positive case (AML74), and in five paediatric CN-AML samples (CN21, CN23, CN25, CN65, CN68).

2.8 Fluorescence in situ hybridization (FISH) analysis for *NUP98-PHF23* fusion

Fluorescence in situ hybridization (FISH) analysis was performed on metaphase e interphase cells using two BlueFISH probes (BlueGnome Ltd., Cambridge) for chromosome 11 and 17 according to manufacturer's instructions. Specifically, BAC clones RP11-120E20 and RP11-348A20 (red) were used to probe the *NUP98* gene on chromosomes 11, while the BAC clone RP11-542C16 (green) was used to target the *PHF23* gene on chromosome 17.

2.9 Cloning and retroviral transduction of *TNIP1-PDGFRB* fusion

Full length *TNIP1-PDGFRB* chimeric transcript was amplified by PCR (Phusion® High-Fidelity DNA Polymerase, New England Bio Labs Inc.), cloned into pCR2.1®-TOPO® TA vector (Life Technologies) and then sub-cloned in murine stem cell virus (MSCV)-internal ribosome entry site (IRES)-green fluorescence protein (GFP) vector (MSCV-IRES-GFP retroviral vector). Retroviral supernatants containing MSCV-TNIP1-PDGFRB-IRES-GFP, as well as MSCV-IRES-GFP (empty vector control) were produced in 293T packaging cell line and used to infect interleukin (IL)-3-dependent hematopoietic murine cells (Ba/F3). After 3 days in culture, cells were harvested and sorted by flow-cytometry for GPF expression. The expression of the fusion gene was confirmed at mRNA and protein level by RT-PCR/Sanger sequencing and western blot, respectively.

2.10 *In-vitro* functional analyses for *TNIP1-PDGFRB* fusion

2.10.1 Cytokine-independent assay

To evaluate factor-independent growth, Ba/F3 cells were washed three times with PBS 1X, and then seeded in triplicate at the concentration of 0.2M cells/well in 24-wells plate with or without IL-3. Cell proliferation and viability were assessed every two days for a total of 14 days by trypan blue count using TC20™ Automated Cell Counter (Biorad).

2.10.2 TKIs treatment sensitivity assay

The potential effect of several TKIs was assessed on both mouse hematopoietic Ba/F3 cells expressing TNIP1-PDGFRB-fusion using CellTiter-Blue® Cell Viability Assay (Promega) accordingly the manufacturer's guidelines. A range of drug concentration was used (0.0025μM - 10μM) and the cell viability was assessed 72 hours (hrs) of treatment. IC50 for each TKI used was calculated using Graphpad Prims (GraphPad Software).

CHAPTER III – Results and Discussion

3.1 Gene fusions are a hallmark of pediatric CN-AML

Fusion genes are frequently detected in acute myeloid leukaemia as well as other haematopoietic cancers. In the present study, RNA-seq performed on diagnostic leukemic blasts of 19 pediatric CN-AML cases confirmed the considerable incidence of this type of genetic alteration in childhood CN-AML. The presence of gene fusions was investigated with three different bioinformatic software packages suited to detect chimeric transcripts from RNA-seq data (Defuse,[74] ChimeraScan,[75] and FusionMap,[76]) and only the putative chimeric transcripts detected from all the alghotihms were selected for further validations. Acconding to this quality selection cut-off, a total of 21 putative fusion genes have been identified (Table 6). Based on the incidence of the chimeric mRNA in analyzed cases, and the available information on partners genes involved in the fusion, the most relevant chimeric transcripts have been selected for further analysis as described in the following paragraphs (Figure 13).

3.2 Identification of *CBFA2T3-GLIS2* fusion transcript in children with CN-AML

The most recurrent (3 out of 19 cases, 16%) putative gene fusion revealed by RNA-seq was a chimeric transcript involving *CBFA2T3* and *GLIS2* genes. Resulting from a cryptic inversion of the telomeric region of chromosome 16, *CBFA2T3-GLIS2* fusion mRNA consists in the 5' portion of *CBFA2T3* gene fused *in-frame* to the 3' region of *GLIS2* gene. *CBFA2T3*, also known as *MTG16* or *ETO2*, is a member of the myeloid translocation gene family that is fused to AML1 in the t(16;21)(q24;q22) translocation that is identified in therapy-relatedAML.[79,80] *GLIS2* is a transcription factor of the GLI-similar Kruppel-like zinc finger protein family that acts as an inhibitor of the hedgehog signaling pathway[81]. Two patients showed the fusion between exon 11

of *CBFA2T3* and exon 3 of *GLIS2*, while the third showed *CBFA2T3* exon 10 fused to exon 2 of *GLIS2* (Figure 14A).

Table 6. Gene fusion identified by RNA-seq in pediatric CN-AML

Patient	Chromosome		5' gene	3' gene	Type	Frame	Function/Process
	5' gene	3' gene					
CN-AML#1	16	16	*CBFA2T3*	*GLIS2*	IC	orf	Trascriptional regulation
CN-AML#2	6	X	*MYB*	*GATA1*	Inter	orf	
CN-AML#3	16	16	*CBFA2T3*	*GLIS2*	IC	orf	
CN-AML#4	16	16	*CBFA2T3*	*GLIS2*	IC	orf	
CN-AML#4	5	5	*MAML1*	*SQSTM1*	Intra	orf	
CN-AML#19	6	X	*MYB*	*GATA1*	Inter	orf	
CN-AML#8	1	1	*PRDM16*	*SKI*	Intra	oof	
CN-AML#14	1	1	*PRDM16*	*SKI*	Intra	oof	
CN-AML#8	11	12	*NUP98*	*KDM5A*	Inter	orf	Chromatin-modification
CN-AML#13	11	17	*NUP98*	*PHF23*	Inter	orf	
CN-AML#18	11	17	*NUP98*	*PHF23*	Inter	orf	
CN-AML#11	5	5	*TNIP1*	*PDGFRB*	Intra	orf	Tyrosine-kinase
CN-AML#14	21	7	*RUNX1*	*USP42*	Inter	orf	Signaling
CN-AML#3	12	12	*DHH*	*RHEBL1*	RT	orf	
CN-AML#3	12	12	*DHH*	*RHEBL1*	RT	orf	
CN-AML#3	17	17	*METRNL*	*SLC16A3*	Intra	oof	Other
CN-AML#7	17	17	*METRNL*	*SLC16A3*	Intra	oof	
CN-AML#4	11	11	*LSP1*	*TNNT3*	RT	oof	
CN-AML#6	16	16	*GSE1*	*ATP6V0D1*	IC	orf	
CN-AML#7	2	2	*PTMA*	*CXCR4*	IC	orf	
CN-AML#17	13	13	*EPSTI1*	*ENOX1*	RT	oof	

IC, intrachromosomal complex; Inter, interchromosomal; RT, read through; Intra, intrachromosomal; orf, open-reading-frame; oof, out-of-frame.

These data were also confirmed by RT-PCR analysis and Sanger sequencing (Figure 14B). Interestingly, we identified 1 novel breakpoint for the *CBFA2T3-GLIS2* fusion transcript (*CBFA2T3*-ex10/ex2-*GLIS2*), which is different from breakpoints already described in the 2 recent reports on non-DS AMKL,[71,82] demonstrating that the cryptic inv(16) generates a *CBFA2T3-GLIS2* fusion gene that can be truncated at different positions.

3.2.1 *CBFA2T3-GLIS2* is recurrent in pediatric CN-AML

To assess the prevalence of *CBFA2T3-GLIS2* fusion, we then examined a validation cohort of 230 children with newly diagnosed *de novo* CN-AML, also negative for known recurrent genetic abnormalities involving *MLL*, *CBFB*, *NPM1*, and *FLT3* genes. Globally, the *CBFA2T3-GLIS2* rearrangement was detected in 20 of 237 cases (8.4%) with CN-AML. In the validation cohort

RT-PCR analysis and Sanger sequencing confirmed that all positive cases carried the *CBFA2T3* exon 11-*GLIS2* exon3 fusion. Fifty percent (N = 10, 50%) of positive patients belonged to the M7 FAB subgroup, while the remaining patients (N = 10, 50%) were distributed among the other FAB classes (Table 7). These results indicate that the *CBFA2T3-GLIS2* fusion transcript, recently described as a distinctive feature of pediatric non-DS AMKL,[71,82] should be more broadly considered as a genetic abnormality that is shared with other FAB subgroups of pediatric CN-AML.

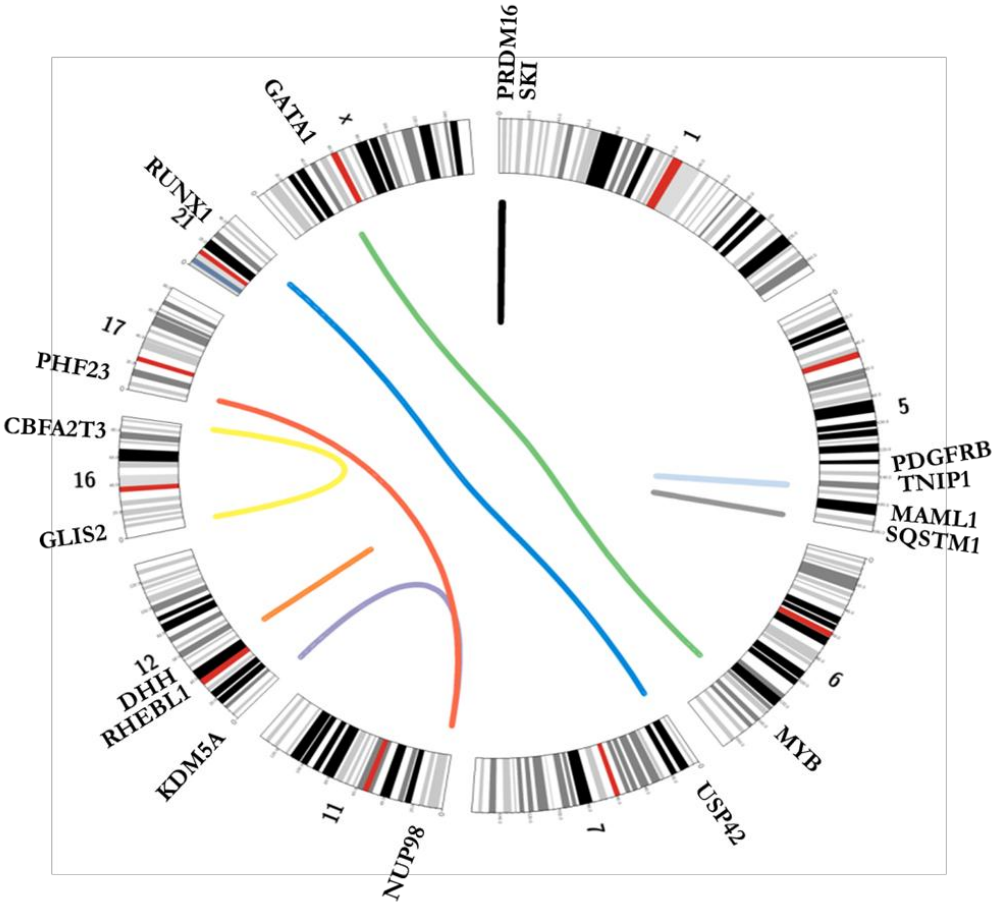
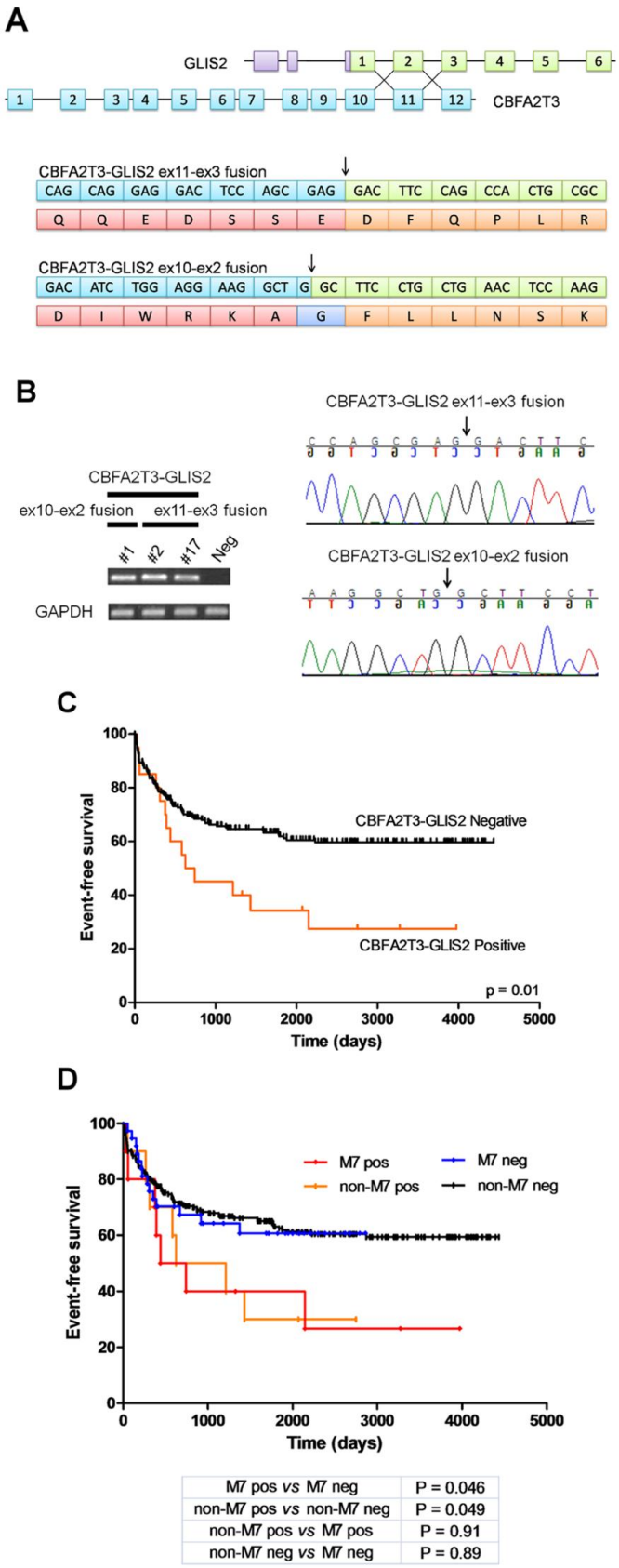


Figure 13. Most relevant fusion genes identified by RNA-seq in a pediatric cohort of 19 CN-AML patients. Circos representation (http://circos.ca/) reporting only the human chromosome found to be involved in rearrangements in the analyzed cases.



M7 pos *vs* M7 neg	P = 0.046
non-M7 pos *vs* non-M7 neg	P = 0.049
non-M7 pos *vs* M7 pos	P = 0.91
non-M7 neg *vs* M7 neg	P = 0.89

Figure 14. CBFA2T3-GLIS2 fusion transcript is a novel common feature of pediatric CN-AML, predicting poorer outcome. (A) Schematic representation of the fusion between *CBFA2T3* and *GLIS2* and predicted sequence of the fusion proteins found in CN-AML. The exon-intron gene structures are indicated. The purple blocks represent untranslated exons. Black arrows indicate the fusion breakpoint. **(B)** RT-PCR analysis and Sanger sequencing performed in order to validate the detection of the *CBFA2T3-GLIS2* fusion. Detection of the GAPDH transcript was used as an RNA quality control. A library negative for the *CBFA2T3-GLIS2* fusion transcript was used as negative control (Neg). Black arrows indicate the fusion breakpoint. **(C)** Probability of 5-year EFS in children with *CBFA2T3-GLIS2* fusion transcript in CN-AML. EFS of *CBFA2T3-GLIS2*–positive patients (27.4%, SE 10.5) *vs* *CBFA2T3-GLIS2*–negative patients (59.6%, SE 3.6; P 5 .01). **(D)** Probability of 5-year EFS in pediatric CN-AML with or without *CBFA2T3-GLIS2* fusion transcript stratified according to FAB subgroups (M7 vs non-M7): EFS of non-M7 CN-AML without *CBFA2T3-GLIS2* (59.4%, SE 3.5 vs EFS of non-M7 CN-AML with *CBFA2T3-GLIS2* (30.0%, SE 14.4 (P5.04). EFS of FAB-M7CN-AMLwithout *CBFA2T3-GLIS2* (60.7%, SE 8.3 vs EFS of FAB-M7 CN-AML with *CBFA2T3-GLIS2* (26.6%, SE 15.0 (P5.04). #1 is in this study CN-AML#1, #2 is in this study CN-AML#3, #17 is in this study CN-AML#4. From ***Masetti R, Pigazzi M, Togni M Astolfi A, Indio V, Manara E, Casadio R, Pession A, Basso G and Locatelli F. Blood 2013 121 (17): 3469-3473.***

Table 7. Clinical features of the CN-AML patients harboring the *CBFA2T3-GLIS2* fusion gene

ID	Age, y	Gender	WBC x $10^{9/L}$	FAB	Bone marrow blast percentage at diagnosis	Extra-medullary involvement	HSCT (Type)	CR after induction therapy	Relapse (site)	Disease free duration (months)	Survival duration (months)
#1*	0.8	M	13.63	M2	52	-	Yes (MFD)	Yes		46.7	+48.0
#2*	0.3	F	7.52	M5A	88	-	Yes (AUTO)	Yes		90.6	+91.4
#3	1.4	F	65.08	M0	98	-	Yes (AUTO)	Yes	Yes (BM)	9.3	13.0
#4	12.7	F	22.8	M1	61	-	Yes (AUTO)	Yes		70.1	+74.5
#5	17.0	F	0.91	M1	98	-	Yes (AUTO)	Yes	Yes (BM+skin)	38.7	50.4
#6	13.3	M	7.52	M5	90	-	No	No		-	2.6
#7	0.9	M	7.53	M5	60	-	Yes (MUD)	Yes	Yes (BM)	15.1	19.3
#8	12.1	M	20.82	M0	70	-	Yes (AUTO)	Yes	Yes (BM)	22.7	24
#9	17.2	M	35.7	M0	85	-	No	No		-	3.2
#10	16.4	F	26.4	M4	70	-	Yes (AUTO)	Yes	Yes (BM)	8.2	12.1
#11	3.2	F	24.2	M7	70	-	Yes (MUD)	Yes	Yes (BM)	12.2	24.7
#12	4.0	M	12.79	M7	74	-	Yes (MUD)	Yes	Yes (BM)	8.6	10.7
#13	0.7	M	13.6	M7	95	Yes (CNS)	Yes (MUD)	No		-	9.4
#14	0.7	F	9.6	M7	90	-	Yes (MUD)	No	Yes (BM)	-	21.1
#15	2.0	F	7.4	M7	95	-	Yes (MFD)	Yes		109.2	+111.8
#16	1.4	F	128.6	M7	50	-	No	No		-	2.1
#17*	0.5	F	13.25	M7	90	-	No	Yes	Yes (BM+CNS)	3.2	5.7
#18	3.0	F	13.63	M7	40	-	Yes (MUD)	Yes		37.2	+45.2
#19	1.6	F	115	M7	90	-	Yes (MFD)	Yes	Yes (BM)	17.7	24.7
#20	1.9	F	56.5	M7	30	-	Yes (MFD)	Yes		13.5	+15.6

*Indicates patients identified in the RNA-seq screening. MUD=Matched Unrelated Donor, MFD=matched family donor, AUTO=autologous. BM=Bone Marrow. CNS=Central Nervous System; +=patients alive and in CR. #1 is in this study CN-AML#1, #2 is in this study CN-AML#3, #17 is in this study CN-AML#4. From ***Masetti R, Pigazzi M, Togni M Astolfi A, Indio V, Manara E, Casadio R, Pession A, Basso G and Locatelli F. Blood 2013 121 (17): 3469-3473.***

3.2.2 *CBFA2T3-GLIS2* identifies a subset of childhood CN-AML with poor outcome

We evaluated whether the presence of *CBFA2T3-GLIS2* fusion product influences patients outcome. The 5-year event-free survival (EFS) of the 20 patients with *the CBFA2T3-GLIS2* fusion gene was significantly worse than that of the 217 pediatric CN-AML patients not harboring the translocation (27.4%, standard error [SE] 10.5 vs 59.6%, SE 3.6; P 5 .01; Figure 14C). We also stratified the patients with respect to FAB subgroups (M7 vs non-M7; Figure 14D). The 5- year EFS of FAB M7 patients with or without the *CBFA2T3-GLIS2* fusion gene was significantly different: 26.6% (SE 15.0) and 60.7% (SE 8.3; P5.04), respectively. Similar results were obtained in non-M7 patients; the 5-year EFS of patients assigned to other FAB categories with or without the *CBFA2T3-GLIS2* fusion transcript was 30.0% (SE 14.4) and

59.4% (SE 3.5; P 5 .04), respectively. No statistically significant difference in EFS of non-M7 and M7 patients harboring the *CBFA2T3-GLIS2* fusion transcript (30.0%, SE 14.4 vs 26.6%, SE 15.0; P 5 .91) was found, suggesting that FAB classification does not interact with the *CBFA2T3-GLIS2* fusion product in influencing outcome. Taken together, these data indicate that the *CBFA2T3-GLIS2* fusion transcript is a novel common feature in pediatric CN-AML that is not restricted to the FAB M7 subtype, predicting poor outcome.

3.3 Identification of a novel *DHH-RHEBL1* fusion transcript in pediatric *CBFA2T3-GLIS2*-positive AML patients

Despite the evidence that, in human AML, presence of *CBFA2T3-GLIS2* fusion predicts for a poor outcome, experimental evidence demonstrated that expression of this chimeric mRNA in mice is not sufficient to promote leukemogenesis, suggesting that the fusion protein *per se* may not be sufficient to promote leukemogenesis[71,83]. Starting from this observation, we reasoned that additional lesions can concur to leukemia development in children harboring *CBFA2T3-GLIS2* fusion transcript. Interestingly, further analysis on RNA-seq data revealed the presence of a novel fusion transcript in the *CBFA2T3-GLIS2*-positive patients (2 out of 3 *CBFA2T3-GLIS2*-positive patients initially identified by RNA-seq). In detail, this novel fusion transcript is the result of a read-through that combines at least part of one exon with each of two distinct (parent) genes that are adjacent on the same chromosome in the same orientation[84]. This transcript involves *DHH*, a member of the Hedgehog family,[85] and *RHEBL1*, a gene coding for a small GTPase of the Ras family. Both genes are contiguously localized on the reverse strand of chromosome 12 (Figure 15) and, although the mechanism that leads to generation of read-through fusion transcripts remains so far is obscure,[86] RT-PCR analysis and Sanger sequencing confirmed that all positive cases harbored the in-frame fusion between exon 2 of *DHH* and exon 2 of *RHEBL1* (Figure 15). Interestingly, both *DHH* and *RHEBL1* genes have been implicated in a variety of human diseases as well as cancer. *DHH* codes for a member of the Hedgehog (HH) signaling pathway, which, similar to other HH ligands, binds to its receptor Patched and leads to the signaling cascade of repressive interactions, culminating into effects on the transcription of target genes. The HH signaling, during embryogenesis, controls cell proliferation, differentiation and tissue morphogenesis[87]. However, it is also well known to have a role in tumors, and the role that HH signaling plays in the growth of tumors can be classified according to how the pathway

is activated[88]. These mechanisms include loss-of-function mutations in inhibitory proteins, such as Patched (PTC1), gain-of-function mutations in positive regulators, such as Smoothened (SMO), and overexpression of the HH ligands (Sonic, Indian and Desert Hedgehog), leading to either autocrine or paracrine activation of the pathway and renewal/propagation of cancer stem cells[87]. *RHEBL1* protein belongs to the Ras family of small GTPases and, similar to other Ras proteins, is a molecular switch that controls a wide variety of cellular functions including cell growth, differentiation and transformation[89]. Previous studies reported that *RHEBL1* could function as an activator of NF-kB[89] and mTOR[90] signaling, both of which are frequently altered in many solid tumors, as well as in leukemias and lymphomas[82,91–93].

3.3.1 *DHH-RHEBL1* fusion transcript is recurrent in pediatric *CBFA2T3-GLIS2*-positive AML

To assess the prevalence of *DHH-RHEBL1* fusion in pediatric AML, we then examined a validation cohort of 55 children with AML. The validation cohort included CN-AML patients (N=24), *CBFA2T3-GLIS2*-positive patients (N=16), patients harboring known cytogenetic/genetic abnormalities (alteration of *MLL*, *NPM1*, *FLT3*, t(8;22)(p11;q13), t(9;11)(p22;q23), inv(16) (p13;q22)) (N=12) and normal CD34+ hematopoietic stem cells (N=3). The *DHH-RHEBL1* fusion transcript was detected in 6 out of 16 patients carrying the *CBFA2T3-GLIS2* fusion, while it was never found in the other patients with AML, irrespectively of the mutational status, as well as in normal CD34+ cells. Thus, considering also the patients of the sequencing cohort, the *DHH-RHEBL1* fusion was globally present in 8 out of 20 (40%) of the *CBFA2T3-GLIS2*-positive patients, this demonstrating that this novel alteration is a common feature of this peculiar subset of childhood AML.

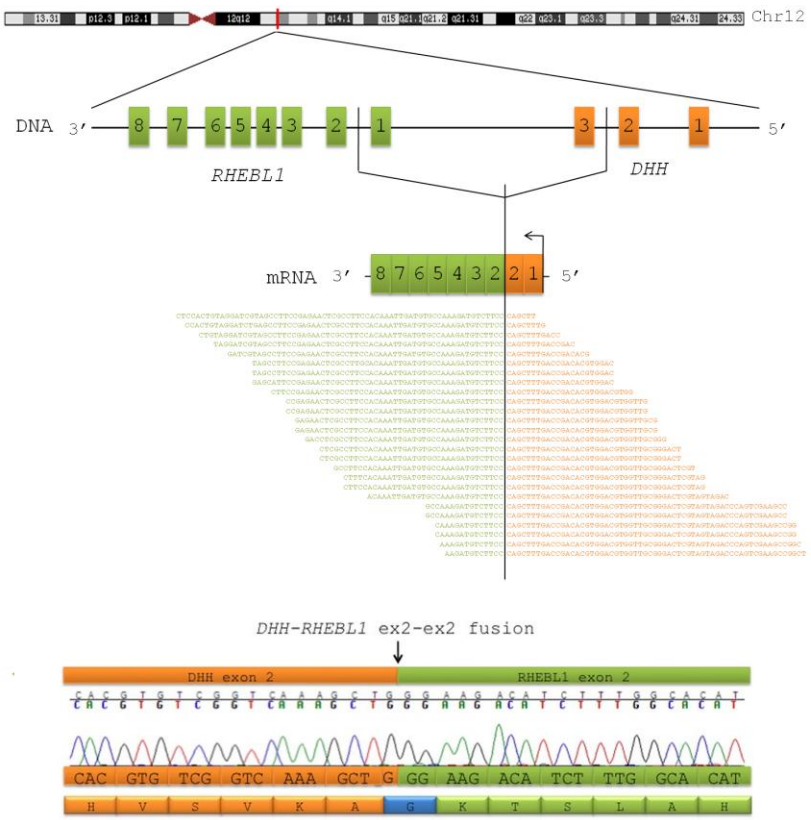


Figure 15. *DHH-RHEBL1* is a novel fusion transcript recurrent in pediatric *CBFA2T3-GLIS2* positive AML. Schematic representation of the fusion between *DHH* and *RHEBL1* identified by means of whole-transcriptome sequencing. The figure shows the position of *DHH* and *RHEBL1* on chromosome 12 and the fusion transcript detected by RNA-seq. The identification of this novel fusion transcript was supported by an average of 11 span and 21 split reads. Sanger sequencing performed in order to validate the detection of the *DHH-RHEBL1* fusion transcript. Electropherogram and predicted sequence of the fusion protein are shown. From ***Masetti R, Togni M, Astolfi A, Pigazzi M, Manara E, Indio V, Rizzari C, Rutella S, Basso G, Pession A and Locatelli F. Oncotarget. 2013 Oct;4(10):1712-20.***

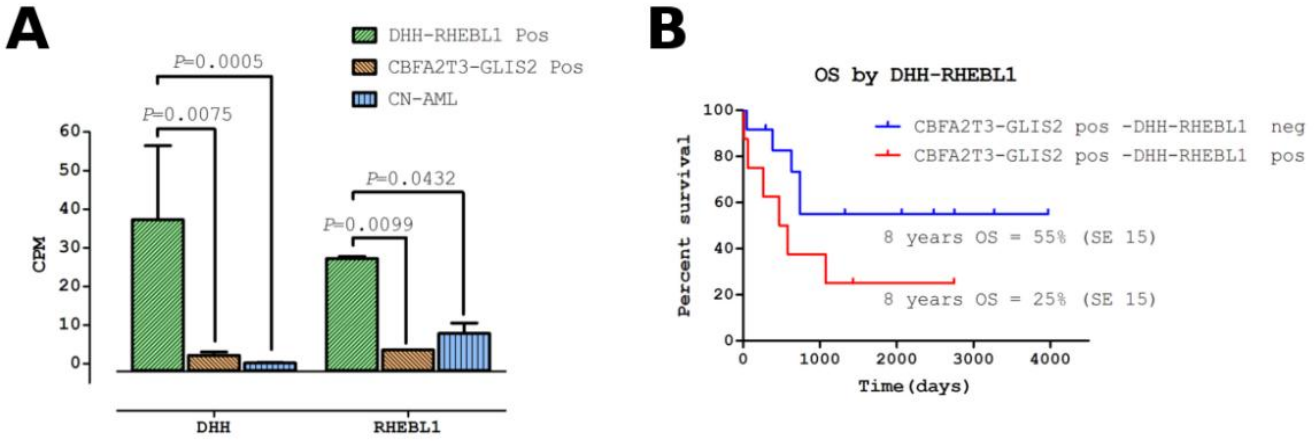


Figure 16. Implications of *DHH-RHEBL1* fusion transcript expression. (A) Expression levels of *DHH* and *RHEBL1* gene obtained from RNA-seq data in *DHH-RHEBL1* positive patients (N=2), in

CBFA2T3-GLIS2 positive patients (N=2) and in CN-AML patients (N=4). Abbreviations: CPM = count per million, Pos = positive. **(B)** Probability of 8-year overall survival (OS) in *CBFA2T3-GLIS2*- positive children who did or did not harbor the *DHH-RHEBL1* fusion transcript (25%, SE=15 vs 55%, SE=15) (P=0.1). From ***Masetti R, Togni M, Astolfi A, Pigazzi M, Manara E, Indio V, Rizzari C, Rutella S, Basso G, Pession A and Locatelli F. Oncotarget. 2013 Oct;4(10):1712-20.***

3.3.2 *DHH-RHEBL1*–positive patients exhibit a specific gene expression signature and an overexpression of both *DHH* and *RHEBL1*

To define the implications, if any, of *DHH-RHEBL1* fusion transcript expression, we performed a gene expression analysis on RNA-seq data. Firstly, we analyzed the expression level of the two genes involved in the fusion transcript and we found that the expression of both *DHH* and *RHEBL1* is significantly enhanced in the *DHH-RHEBL1*-positive patients as compared with patients harboring only *CBFA2T3-GLIS2* fusion (P=0.007 and P=0.009 respectively) and with the other CN-AML cases (P=0.0005 and P=0.043, respectively) (Figure 16A). Recently, with the identification of *CBFA2T3-GLIS2* fusion transcript in pediatric CN-AML, different studies[71,82,83] demonstrated that the presence of this fusion transcript leads to an aberrant activation of the HH signaling due to the ectopic expression of the *GLIS2* transcription factor. Notably, here, we demonstrate that patients harboring the *DHH-RHEBL1* fusion present an overexpression of *DHH* compared to both *CBFA2T3-GLIS2*-positive patients and CN-AML patients. Considering that overexpression of the HH ligands leads to activation of the HH pathway[87], it is tempting to speculate that overexpression of *DHH* could contribute to the aberrant activation of the HH pathway. Moreover, in view of *RHEBL1* over-expression in patients harboring the *DHH-RHEBL1* fusion transcript compared to those harboring only the *CBFA2T3-GLIS2* fusion transcript and to CN- AML children, it will be interesting to investigate more thoroughly its possible role in leukemogenesis. Additionally, *DHH-RHEBL1*-positive patients showed a distinctive gene expression signature both with respect to *CBFA2T3-GLIS2*-positive patients (518 differentially expressed genes; P<0.05), and CN-AML patients (596 differentially expressed

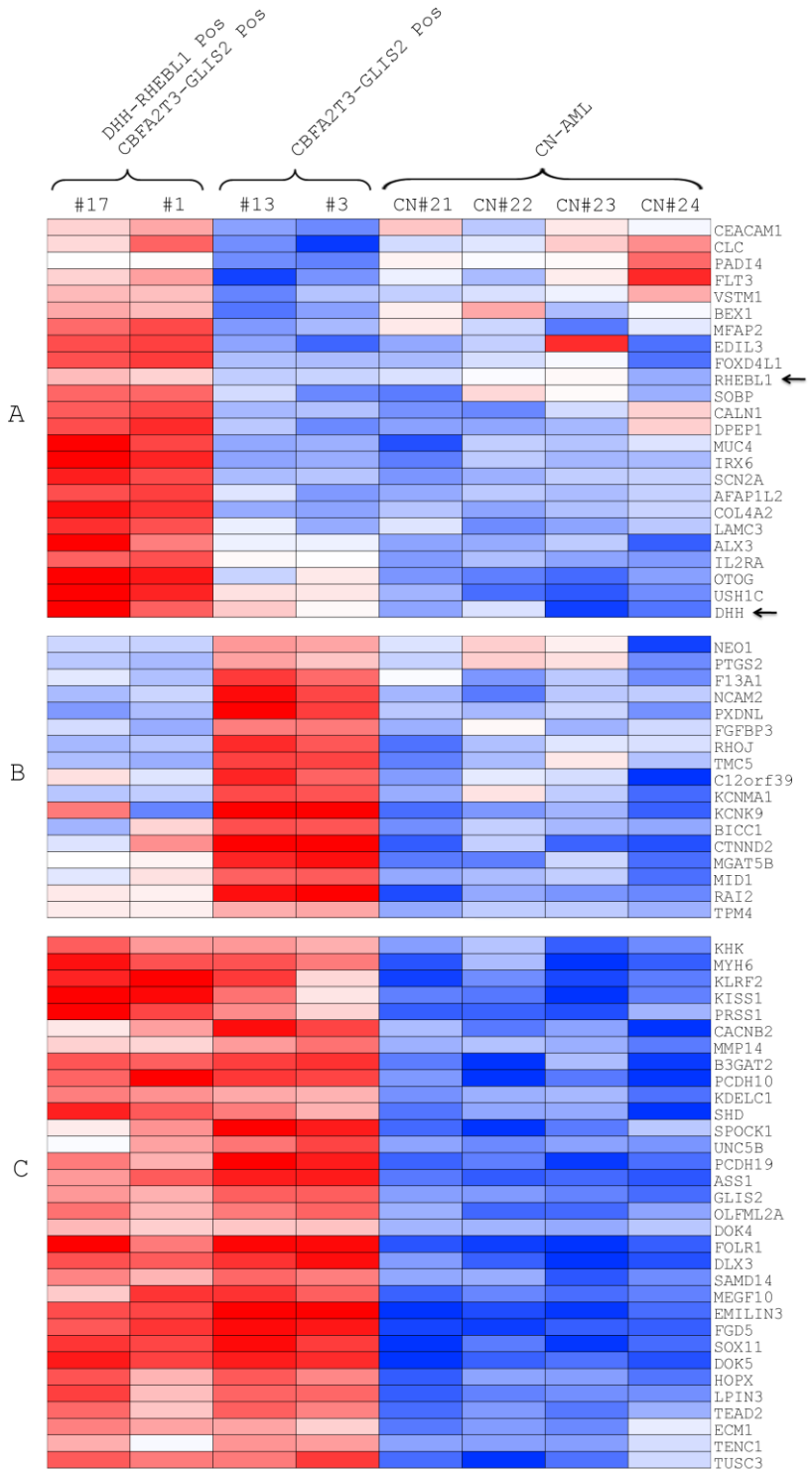


Figure 17. Analysis of gene expression profile of *DHH-RHEBL1*-positive patients. Heatmap of the top 30 differentially expressed genes in *DHH-RHEBL1*-positive patients compared with the patients harboring the *CBFA2T3-GLIS2* fusion only and CN- AML patients. **(A)** Gene expression signature of patients harboring both *CBFA2T3-GLIS2* and *DHH-RHEBL1* fusion transcript; **(B)** Gene expression

signature of *CBFA2T3-GLIS2*-positive patients; **(C)** Gene expression signature of children with CN-AML not harboring any detectable fusion transcript. Abbreviation: Pos = positive. From ***Masetti R, Togni M, Astolfi A, Pigazzi M, Manara E, Indio V, Rizzari C, Rutella S, Basso G, Pession A and Locatelli F. Oncotarget. 2013 Oct;4(10):1712-20.***

genes; P<0.05). Interestingly, *DHH-RHEBL1*-positive patients showed higher expression of several genes known to be associated with leukemia occurrence and/or tumor progression, such as *FLT3*,[94] *BEX1*,[95] *MUC4*[96] and *AFAP1L2*[97] (Figure 17). Finally, we also evaluated whether the presence of *DHH-RHEBL1* fusion transcript influences the outcome of *CBFA2T3-GLIS2*-positive patients. The 8-year overall survival of the 8 patients harboring the *DHH-RHEBL1* fusion transcript was worse than that of the 12 *CBFA2T3-GLIS2*-rearranged patients not harboring the *DHH-RHEBL1* fusion transcript (25% vs 55%). However, likely due to the small number of patients, this difference failed to achieve statistical significance (P=0.1) (Figure 16B).

3.4 Identification of *RUNX1-USP42* and *PRDM16-SKI* fusion transcripts

Among the CN-AML pediatric patients analyzed by the means of RNA-seq, an 8-year-old child diagnosed with AML French-American-British (FAB) type M5b showed an abnormal karyotype, 46, XY, del(5)(q14q34) with no recurrent genetic abnormality involving *MLL*, *CBFB*, *NPM1* and *FLT3* genes. However, due to the rarity of this particular alteration [del(5)(q14q34)] in *de novo* pediatric AML, and its association with a grim survival, the patient was included in the study cohort. Interestingly, giving value to the concordant results emerging from the use of three distinct algorithms, (Defuse,[74] ChimeraScan,[75] and FusionMap,[76]) two fusion transcripts, *RUNX1- USP42* and *PRDM16-SKI*, were identified in this patient. Perusal of the reads mapping these chimeric transcripts showed that the first was an *in-frame* fusion, while *PRDM16- SKI* was a novel *out-of-frame* fusion (Figure 18A). RT-PCR analysis and Sanger sequencing confirmed the presence of both chimeric transcripts (Figure 18B). To date, the cryptic t(7;21) leading to the fusion between *RUNX1* and *USP42* has been reported only once in children,[98] while seven adult cases with myeloid neoplasms harbouring the *RUNX1-USP42* transcript have been published (Table 8), suggesting that this genetic abnormality is a rare, but non-random, feature of myeloid malignancies in adults. To assess the recurrence of *RUNX1-USP42* fusion transcript in childhood

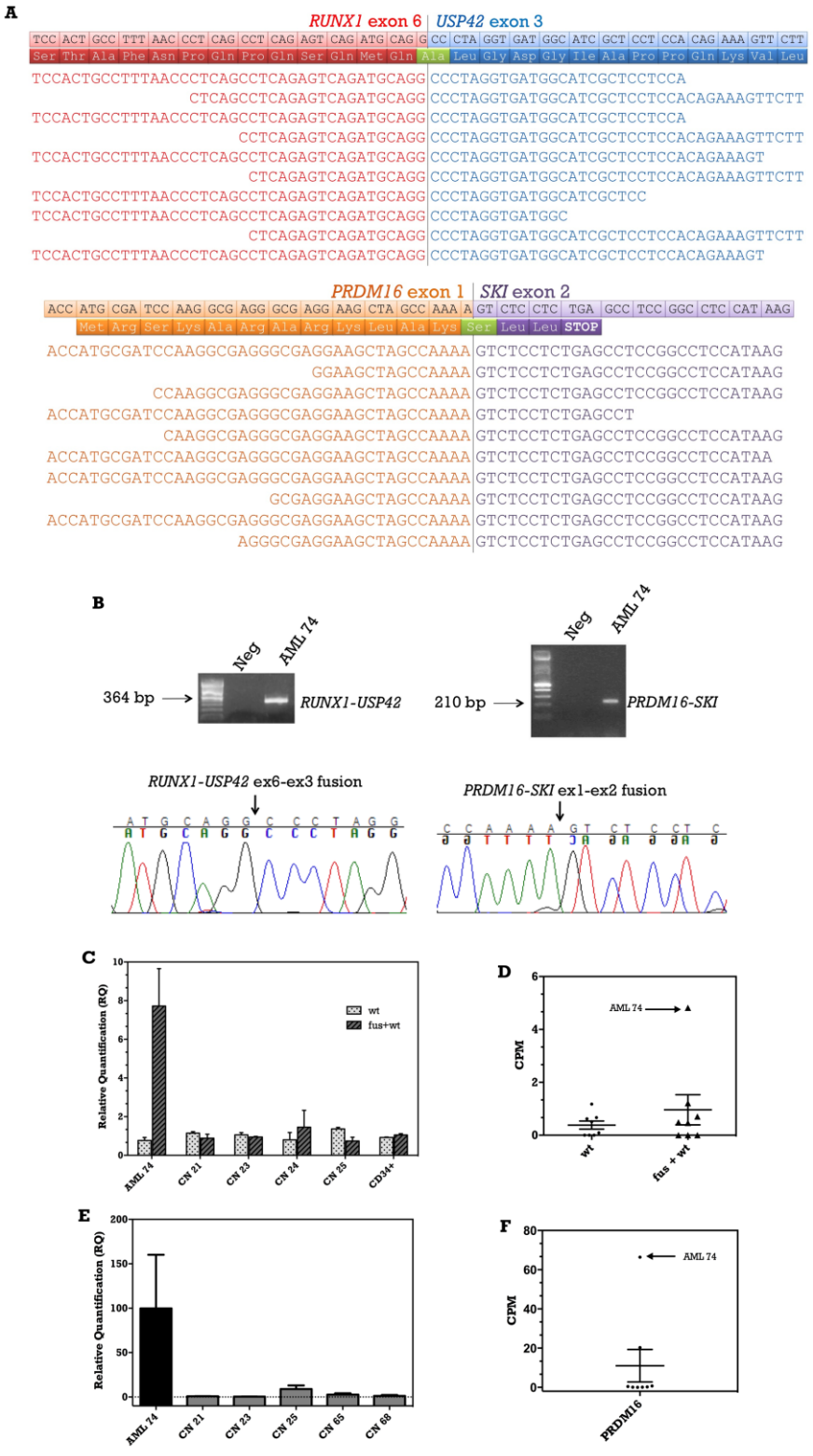


Figure 18. *RUNX1-USP42* and *PRDM16-SKI* fusion identified through RNA-seq in a child with acute myeloid leukaemia (AML) and del(5q). (A) Schematic representation of split read mapping the (Figure 18 legend continued…) *RUNX1-USP42* and *PRDM16-SKI* fusions obtained by whole-

transcriptome sequencing (WTS). Defuse, Chimerascan, and FusionMap packages were used to detect chimeric transcripts from RNA-seq data**.** Predicted sequences of the chimeric proteins are reported. **(B)** Reverse transcription polymerase chain reaction (RT-PCR) analysis and Sanger sequencing were performed in order to validate the detection of *RUNX1-USP42* and *PRDM16-SKI* fusions. Neg = negative. **(C)** Quantitative RT-PCR (qRT-PCR) analysis was performed in order to detect differential expression of the *USP42* gene. As a result of the t(7;21) translocation, *USP42* gene is placed under the control of the *RUNX1* promoter, and this could lead to over-expression of the rearranged *USP42* gene. To test this possibility, we designed a q-PCR with two different primer pairs (wild-type [wt]: primers mapping in exons 2–3of *USP42*; and wt+fusion [fus]: primers mapping in exons 7–8 of *USP42*) and we evaluated the expression of the wt and the wt+rearranged *USP42* in 4 cytogenetically normal AML (CN-AML) samples (CN21, CN23, CN24, CN25), the t(7;21)(p22; q22) positive case (AML74) and in CD34+ haematopoietic stem cells. **(D)** Expression levels of USP42 obtained from WTS data, in the t(7;21)(p22;q22) positive patient (AML74), and 7 CN-AML patients (CN21, CN22, CN23, CN24, CN25, CN65, CN68). CPM = count per million. **(E)** q-PCR analysis was performed in order to detect differential expression of the *PRDM16* gene in the t(7;21)(p22;q22) positive case (AML74), and in five paediatric CN-AML samples (CN21, CN23, CN25, CN65, CN68). **(F)** Expression levels of *PRDM16* obtained from WTS data in the t(7;21)(p22;q22) positive patient (AML74), and 7 CN-AML patients (CN21, CN22, CN23, CN24, CN25, CN65, CN68). AML74 is in this study CN-AML#14, CN21 is in this study CN-AML#5, CN22 is in this study CN-AML#6, CN23 is in this study CN-AML#7, CN24 is in this study CN-AML#9, CN25 is in this study CN-AML#10, CN65 is in this study CN-AML#11, CN68 is in this study CN-AML#12. From ***Masetti R, Togni M, Astolfi A, Pigazzi M, Indio V, Rivalta B, Manara E, Rutella S, Basso G, Pession A, Locatelli F. Br J Haematol. 2014 Aug;166(3):449-52.***

AML, we examined 132 children with *de novo* CN-AML enrolled in the AIEOP AML 2002/01 Protocol. No additional patient was found to carry this chimeric transcript. To the best of our knowledge, this is the second paediatric case in which the *RUNX1-USP42* fusion transcript has been detected and, unprecedentedly, it was observed at diagnosis together with del(5q), whereas in the case described by Paulsson et al (2006) the 5q- occurred later in the disease course.[98] In our case, the fusion occurred between exon 6 of *RUNX1* and exon 3 of *USP42* (Figure 18A-B). However, as previously reported, owing to different breakpoints and alternative splicing in *RUNX1*, several splice-variants of *RUNX1-USP42* chimeric transcript may be generated by the cryptic t(7;21) (Table 8) [99]. Noteworthy, all *RUNX1-USP42* isoforms harbour both the catalytic domain (UCH) of USP42 and the highly conserved Runt homology domain (RHD) of RUNX1, which mediates DNA binding and heterodimerization of RUNX1 with CBFb. Several mechanisms through which this chimeric transcript could contribute to the leukaemogenesis have been proposed: (i) dominant negative inhibition of wild-type RUNX1 transcription activation activity[98,100] (ii) USP42-mediated stabilization of RUNX1 from ubiquitin-proteasome degradation,[98] and iii) ectopic over-expression of USP42 and deregulation of TP53-dependent cell-cycle arrest[101]. In accordance with the last hypothesis, through q-PCR analysis, we demonstrated an over-expression of *USP42* in t(7;21)-positive cells when compared with other

CN-AML cells lacking this translocation, and CD34+ haematopoietic progenitors (P = 0.02) (Figure 18C). This finding was further validated through gene expression analysis performed on the RNA-seq data (Figure 18D). Taken together, these results confirm and extend previous reports,[100] enforcing the potential role of this proteinase in leukaemogenesis, and lending support to both the role of deubiquitinating proteins in tumourigenesis and the use of deubiquitinating enzyme inhibitors in cancer therapy. Despite this finding, animal models have shown that *RUNX1*-related translocations or haploinsufficiency of *RUNX1* are not sufficient for leukaemogenesis, indicating that additional genetic events are required. Interestingly, whole-transcriptome sequencing enabled the identification of a novel *out-of-frame* fusion transcript in which exon 1 of *PRDM16* (also known as *MEL1*) is fused with exon 2 of *SKI* (v-ski avian sarcoma viral oncogene homolog). Due to loss of the open-reading frame, no putative chimeric protein seems to be encoded by this fusion transcript. However, considering the relative position of the two genes on the forward strand of chromosome 1, we speculate that the juxtaposition of *SKI* downstream of exon 1 of *PRDM16* may lead to slipping of *PRDM16* and this, in turn, positions the gene under the control of ectopic promoter/enhancer elements (Figure 18A-B). To test this hypothesis, we evaluated the expression levels of *PRDM16* by q-PCR and gene expression analysis on RNA-seq data, and found that its expression was significantly enhanced in the t(7;21)-positive patient compared to other CN-AML patients (P = 0.01) (Figure 18E-F). Additionally, *SKI* expression was investigated through both q-PCR and RNA-seq gene expression analysis. However, no differential expression of this gene was detected (data not shown). Coding for a zinc-finger protein containing a DNA-binding PRDI-BF1/RIZ homologous (PR) domain, *PRDM16* belongs to the *EVI1* family, and is implicated in two translocations involving *RUNX1* and *RPN1* in myeloid malignancies. Interestingly, several reports indicate the link between overexpression of *PRDM16* and leukaemogenesis, strengthening its association with a worse outcome[102,103]. In summary, we reported, for the first time, the identification of a *RUNX1-USP42* fusion transcript detected at diagnosis in a child with AML and del(5). No other child was positive for this chimeric transcript in a lager validation cohort, demonstrating that this is a rare genetic lesion in childhood AML. Interestingly, RNA-seq also enabled the identification of a novel *out-of-frame PRDM16-SKI* fusion, and analysis of *PRDM16* revealed an overexpression of this gene. Notably, an additional CN-AML patient of our sequencing cohort was found to be positive for the *out-of-frame PRDM16-SKI* fusion and, preliminary results form

a screening in 230 children with newly diagnosed *de novo* CN-AML indicate both a considerable incidence of this genomic alteration in this subgroup of childhood leukemias and a consequent overexpression of *PRDM16* gene (data not show).

Table 8. Main characteristics of t(7;21)(p22; q22) positive patients

N	Age, years	Gender (M/F)	Disease	WBC, $x10^9$/L	BM blast percentage at diagnosis	del(5q) at diagnosis	Fusion (isoform)	HSCT (type)	Relapse	Ref
1	7	Male	AML M0	35.6	75-80%	No	RUNX1ex7(iso2) -USP42ex3(iso1)	Yes (SIB)	Yes	98
2	68	Female	MDS (RAEB-2)	NA	17%	No	RUNX1ex6(iso2) -USP42ex3(iso1)	No	No	99
3	68	Male	AML M5	NA	NA	No	RUNX1ex7(iso2) -USP42ex3(iso1)	No	No	99
4	32	Male	AML M1	12.1	90%	Yes	RUNX1ex7(iso2) -USP42ex3(iso1)	NA	NA	101
5	54	Male	AML M4/M5	17.6	73%	Yes	RUNX1ex7(iso2) -USP42ex3(iso1)	Yes (NA)	Yes	104
6	33	Male	AML M5a	5.4	84%	Yes	RUNX1ex7(iso2) -USP42ex3(iso1)	Yes (SIB)	No	104
7	39	Male	AML M4/M5	1.3	NA	Yes	RUNX1ex6(iso2) -USP42ex3(iso1)	No	NA	104
8	52	Female	AML M0	34	>70%	Yes	RUNX1ex7(iso2) -USP42ex3(iso1)	Yes (SIB)	No	100
9	8	Male	AML M5	96.18	60%	Yes	RUNX1ex6(iso2) -USP42ex3(iso1)	Yes (MUD)	No	Present Case

Sib = sibling; MUD = matched unrelated donor. NA= not available. From ***Masetti R, Togni M, Astolfi A, Pigazzi M, Indio V, Rivalta B, Manara E, Rutella S, Basso G, Pession A, Locatelli F. Br J Haematol. 2014 Aug;166(3):449-52.***

3.5 *NUP98-PHF23* is a recurrent fusion gene in pediatric CN-AML

RNA-seq performed on diagnostic leukemic cells was also able to reveal in 2 out of 19 analyzed patients a fusion involving the genes *NUP98* and *PHF23*, resulting from a cryptic translocation t(11;17)(p15;p13) (Figure 19A) (Table 9). Both patients showed an *in-frame* fusion between exon 13 of *NUP98* and exon 4 of *PHF23* and RT-PCR analysis and Sanger sequencing confirmed the fusion breakpoint (Figure 19B). To date, the identification of the cryptic translocation t(11;17)(p15;p13) has been described only once in an adult AML patient[105] and, to

our knowledge, this feature has never been reported in a pediatric AML cohort. In addition, differently from what previously reported by Reader and colleagues,[105] here the recurrent breakpoint in *PHF23* gene was in both cases identified at the start of exon 4 and not within it (Figure 17A-B). This finding indicates that *NUP98-PHF23* fusion gene can involve either the whole exon 4 of *PHF23* or only a fragment of it. To investigate the incidence of *NUP98-PHF23* fusion in pediatric CN-AML, we examined through RT-PCR analysis and Sanger sequencing a validation cohort of 152 AML children enrolled in the AIEOP AML 2002/01 Protocol (Pession, *et al* 2013): 135 patients were negative for the recurrent cytogenetic lesions or genetic abnormalities, involving *MLL*, *CBFB*, *NPM1* and *FLT3*, while the remaining 17 patients harbored *FLT3-ITD* mutation, and were chosen because we previously reported a strong association between *NUP98-NSD1* rearrangement and *FLT3-ITD*[106]. Overall, 2 out of 152 CN-AML patients were positive for the *NUP98-PHF23* fusion, demonstrating that this genomic aberrancy is not rare in pediatric CN-AML (Table 9). The presence of the cryptic chromosomal translocation t(7;11)(p15;p13), leading to the fusion between *NUP98* and *PHF23* genes was confirmed by fluorescence in situ hybridization (FISH) analysis in all cases (Figure 19C). Located on 11p15.5, *NUP98* encodes a 98-kD protein component of the nuclear pore complex, which facilitates mRNA export from the nucleus. So far many chromosomal rearrangements involving *NUP98* have been described to be associated with both *de novo* and therapy-related AML, but also with T-cell acute lymphoblastic leukemia (T-ALL). At present, more than 15 different partner genes have been found rearranged with *NUP98*[107]. Known fusion partners can be divided into three groups: homeobox (*HOX*) genes (*HOXA9, HOXA11, HOXA13, HOXC11, HOXC13, HOXD11* and *HOXD13*), nuclear non-homeotic genes (*LEDGF, NSD1, NSD3, DDX10, TOP1, TOP2B, FN1* and *C6orf80*) and two genes found exclusively fused with *NUP98* in T-ALLs (*RAP1GDS1* and *ADD3*). In childhood AML, recently, a novel fusion transcript involving *NUP98*, namely *NUP98/JARID1A* has been described to be a recurrent event in approximately 11% of pediatric patients with acute megakaryoblastic leukemia, with a distinct HOX gene expression pattern[108]. By contrast, chromosomal rearrangements and/or mutations of *PHF23* were never before described in children with AML. Located on the reverse strand of 17p13.1, *PHF23* gene encodes for a protein containing a plant homeodomain (PHD) finger[109]. PHD zinc fingers are structurally conserved motifs found in several nuclear proteins thought to be involved in epigenetic and chromatin-mediated transcriptional regulation.

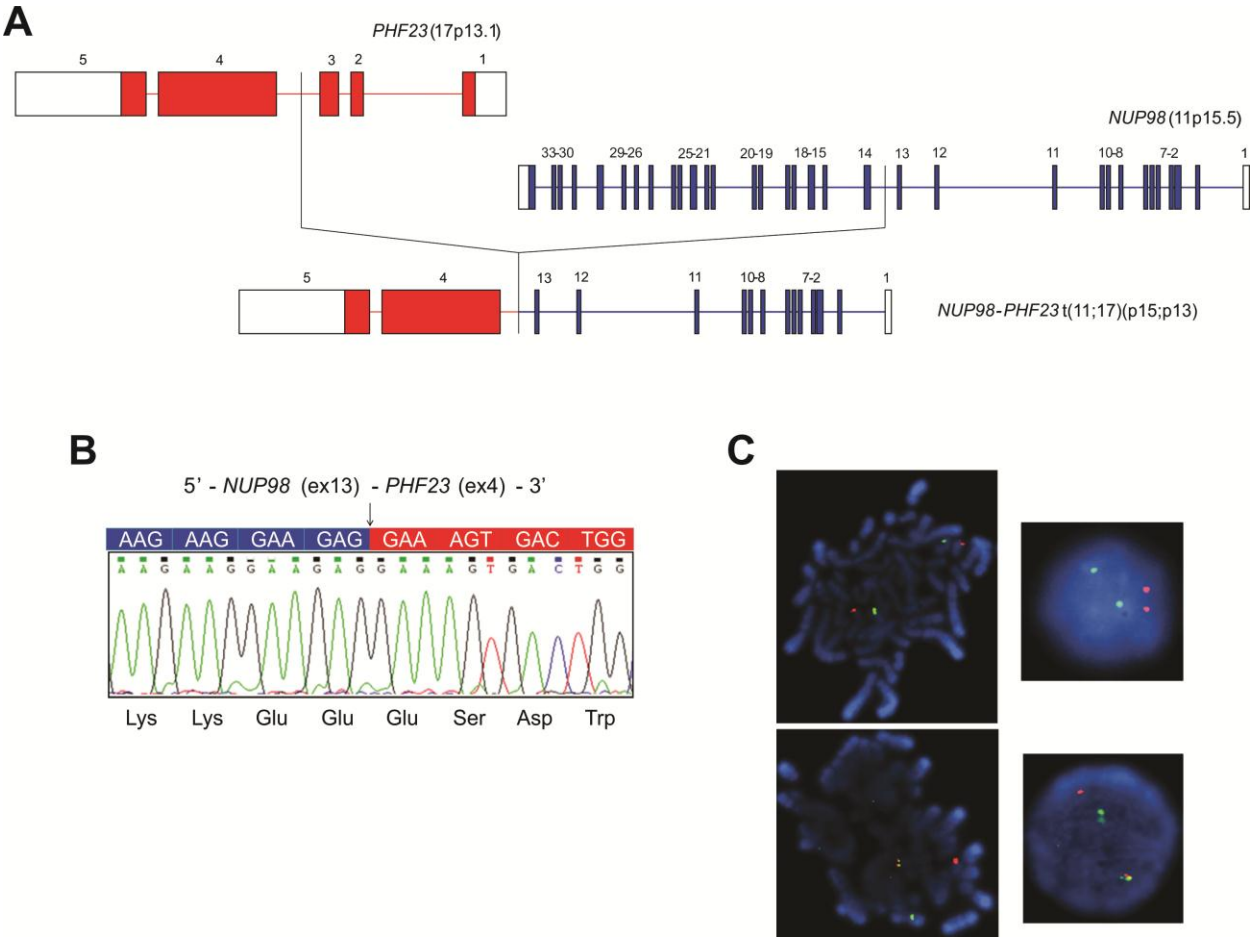


Figure 19. Identification of *NUP98-PHF23* in pediatric CN-AML. (A) Schematic representation of *NUP98-PHF23* fusion identified by RNA-seq in pediatric CN-AML. Fusion occurs between exon 13 of *NUP98* and exon 4 of *PHF23*. **(B)** Electropherogram from Sanger sequencing of the region surrounding the breakpoint confirmed the *in frame* fusion. Black arrow indicate the fusion breakpoint, predicted sequence of the fusion protein is showed. **(C)** FISH analysis was performed on metaphases and interphase cells using three BlueFISH probes (BlueGnome Ltd., Cambridge), according to manufacturer's instructions. BAC clones RP11-120E20 and RP11-348A20 (red) were used to probe the *NUP98* gene on chromosomes 11, while the BAC clone RP11-542C16 (green) was used to target the *PHF23* gene on chromosome 17. Normal metaphases (up left) and interphase nuclei (up right) showed two red signals representing normal copies of *NUP98* and two green signals representing normal copies of *PHF23*. Abnormal metaphases (down left) and interphase cells (down right) containing the *NUP98-PHF23* fusion gene showed one red (*NUP98*), one green (*PHF23*) and one yellow fusion signal which represents the juxtaposition of the translocated portions of the two genes.

Recent studies showed that these domains have a sophisticated histone sequence reading capacity, underscoring the functional versatility of PHD fingers as epigenome readers that

control gene expression through molecular recruitment of multi-protein complexes of chromatin regulators and transcription factors[105].

Table 9. Clinical features of pediatric CN-AML patients harboring the *NUP98-PHF23* fusion gene

Id	Age, (y)	Gender	WBC, x 10^9/L	FAB	BM (%)	ED	HSCT (type)	CR after induction therapy	Relapse (site)	DFD (m)	Sur (m)
CN-AML_54*	2.9	M	187	M1	90	No	AUTO	Yes	(BM)	5	† 30
CN-AML_66*	9.0	M	1.2	M0	70	No	MUD	Yes	-	65	66
AML_3	9.7	M	6.9	M4	40	No	MUD	Yes	-	40	41
AML_4	7.0	M	1.8	M5A	54	No	AUTO	Yes	-	103	104

, patients identified by RNA-seq; †, dead patients; AUTO, autologous; CR, complete remission; HSCT, hematopoietic stem cell transplantation; MUD, matched unrelated donor; WBC, white blood cells. y, years; BM, bone marrow blast % at diagnosis; EM, extramedullary involvement DFD, disease-free duration; Sur, survival duration: m, months. CN-AML_54 is in this study CN-AML#13, CN-AML_66* is in this study CN-AML#18.

The role of *NUP98-PHF23* in leukemogenesis has investigated by functional studies showing that the expression of this fusion protein impaired the differentiation of myeloid progenitor cells and is sufficient to promote leukemia development *in vitro* and *in vivo*[105,109,110]. Noteworthy, cells expressing NUP98-PHF23 are sensitive to low concentration of disulfiram, a U.S. Food and Drug Administration–approved drug, demonstrating the feasibility of targeting this oncoprotein as well as other PHD fusion proteins like NUP98-JARID1A[110]. In summary, here we reported the identification of a *NUP98-PHF23* chimeric transcript in a pediatric acute myeloid leukemia cohort, demonstrating that this genomic aberrancy is not rare (2.4%) in pediatric CN-AML. Taken together these findings strengthen the importance role of genes coding for epigenetic regulators in pediatric AML and suggest thethe putative use of novel epigenome-targeted therapies in this pathological setting.

3.6 Identification of *TNIP1-PDGFRB* fusion transcript in children with CN-AML

Although its presence was detected only in 1 out of 19 analyzed patients (5%), the identification of the *in-frame TNIP1-PDGFRB* fusion has been selected for further analyses due to the well-

known role in cancer of the platelet-derived growth factor receptor β (*PDGFRB*) gene.[111] Constitutive activation of the PDGFRB receptor tyrosine kinase has been reported in myeloid malignancies as a consequence of fusion to diverse partner genes (Table 10). PDGFRB overexpression is implicated in the pathogenesis of dermatofibrosarcoma protruberans (DFSP) and overexpression of PDGFRs and/or their ligands has been described in many solid neoplasms[112]. Furthermore, and more important, PDGFRB signaling is selectively inhibited by imatinib and other tyrosine kinase inhibitors (TKIs) already approved by FDA, and these selective drugs have been extensively used in several pathological settings in which PDGFR abnormalities have been implicated. Recently, analyzing a cohort of 15 BCR-ABL1-like acute lymphoblastic leukemia (Ph-like ALL) patients by RNA-seq and whole-genome sequencing, the laboratory leaded by Charles Mullighan showed that genetic alterations activating kinase or cytokine receptor signaling are a hallmark of this subtype of pediatric B-ALL, which lacks recurrent genetic alterations. Particularly, a novel gene fusion involving the early B-cell factor 1 (*EBF1*) gene and *PDGFRB* gene has been reported and a screening in an independent cohort of 231 high-risk B-progenitor ALL cases revealed the recurrence of this novel fusion gene in Ph-like ALL (up to 8%)[113]. Additionally, the effective clinical use of imatinib has been recently reported in two pediatric patients with *EBF1-PDGFRB*–positive B-ALL, demonstrating that TKIs treatment could be successfully used in this pathological setting[114,115]. By the means of RNA-seq, additional *PDGFRB*-fusions have been identified by our group and others in pediatric B-ALL. Particularly, Kobayashi K and collaborators have recently reported a novel *ATF7IP-PDGFRB* fusion in a 8-year-old boy with standard-risk B-ALL,[116] and Charles Mullighan's group have recently identified *ZEB2* and *TNIP1* genes as novel partner genes of *PDGFRB* in Ph-like ALL[113]. Interestingly, here I report the identification of *TNIP1-PDGFRB* fusion also in one patient of our CN-AML cohort, suggesting that this type of gene fusion could be a very early genetic alteration that occurs in the hematopoietic precursors before their commitment in the lymphoid or myeloid lineage. Specifically, the RNA-seq revealed an in-frame fusion involving the exon 17 of *TNIP1* gene and the exon 11 od *PDGFRB* gene leading to the generation of a chimeric protein in with both the dimerization domain of TNIP1 and the tyrosine-kinase domain of PDGFRB are maintained. Likewise the other PDGFRB-fusions, the suppose mechanism leading to the constitutive activation of this chimeric kinase is mediated by an increased ability

of the fusion protein to made dimer or oligomer, constitutively allowing the cross-activation of the tyrosine-kinase domain of PDGFRB (Figure 20).

Table 10. PDGFRB-fusions so far reported in myeloid malignancies

5' Gene Partner	Chromosomal rearrangement	Disease	Oligodimerization Domain	Ref
ETV6/TEL	t(5;12)(q33;p13)	CMML	SAM Pointed Domain #	[117]
CEV14/TRIP11	t(5;14)(q33;q32)	AML	Coiled Coil	[118]
HIP1	t(5;7)(q33;q11.2)	CMML	Coiled Coil	[119]
CCDC6/H4	t(5;10)(q33;q21)	MDS	Coiled Coil	[120]
RABAPTIN5	t(5;17)(q33;p13)	CMML	Coiled Coil	[121]
PDE4DIP/myomegalin	t(1;5)(q23;q33)	MDS	Coiled Coil	[122]
TP53BP1	t(5;15)(q33;q22)	MPN	NA	[123]
NIN	t(5;14)(q33;q24)	MPN	Coiled Coil	[124]
HCMOGT-1/CYTSB	t(5;17)(q33;p11.2)	JMML	Coiled Coil	[125]
KIAA1509/CCDC88C	t(5;14)(q31;q32)	MDS	Coiled Coil	[126]
TROPOMYOSIN TPM3	t(1;5)(q21;q33)	CEL	Coiled Coil	[127]
NDE1	t(5;16)(q33;p13)	CMML	Coiled Coil	[128]
GIT2	t(5;12)(q31-33;q24)	MDS	Ankyrin Repeat	[129]
GPIAP1/CAPRIN1	t(1;5;11)	CEL	Coiled Coil	[129]
PRKG2	t(4;5;5)(q23;q31;q33)	MPN	Coiled Coil #	[129,130]
ERC1	t(5;12)(q33;p13.3)	AML	Coiled Coil	[131]
SPTBN1	t(2;5)(p21;q33)	MPN	NA	[130]
Myosine/MYO18A	t(5;17)(q33-34;q11.2)	MPN	Coiled Coil	[132]
SART3	t(5;12)(q31–32;q23–24)	MPN	Coiled Coil	[133]
WDR48	t(5;3)(q33;p21)	MPN	NA	[134]
GOLGA4	t(5;3)(q33;p21)	MPN	Coiled Coil	[134]
BIN2	t(5;12)(q33;q13)	MPN	Coiled Coil	[134]
KANK1	t(5;9)(q32;p24)	Throm	Coiled Coil	[135]
EBF1	Del5(q33q33)	Ph-like ALL	Coiled Coil	[136]
CEP85L	t(5;6)(q33-34;q23)	T-ALL* -MPN	Coiled Coil	[137,138]
ATF7IP	t(5;12)(q33;p13.1)	Ph-like ALL	*Coiled Coil	[116]
DTD1	t(5;20)(q33;p11)	EosM/LN	Not present	[139]
CCDC88C	t(5;14)(q33;q32)	EosM/LN	Coiled Coil	[139]
TNIP1	Del5	Ph-like ALL	Coiled Coil	[113]
ZEB2	t(5;2)(q33;q22.3)	Ph-like ALL	NA	[113]

CMML, Chronic Myelomonocytic Leukemia; AML, Acute Myeloid Leukemia; MDS, Myelodysplastic Syndromes; MPN, Myeloproliferative Neoplasms; JMML, Juvenile Myelomonocytic Leukemia; CEL, Chronic Eosinophilic Leukemia; CMML, Chronic Myelomonocytic Leukemia;Throm, Thrombocythemia; Ph-like ALL, Ph-like Acute Lymphoblastic Leukemia; T-ALL, Precursor T Lymphoblastic Leukemia; EosM/LN, Eosinophilia-Associated Myeloid/Lymphoid Neoplasms; NA, not avaible.

3.6.1 TNIP1-PDGFRB sustains cytokine-independent growth *in-vitro*

To evaluate the transforming potential of TNIP1-PDGFRB *in-vitro*, we assessed the ability of murine interleukin-3 (IL-3)-dependent mouse hematopoietic Ba/F3 expressing TNIP1-PDGFRB to proliferate in the absence of exogenous cytokines.

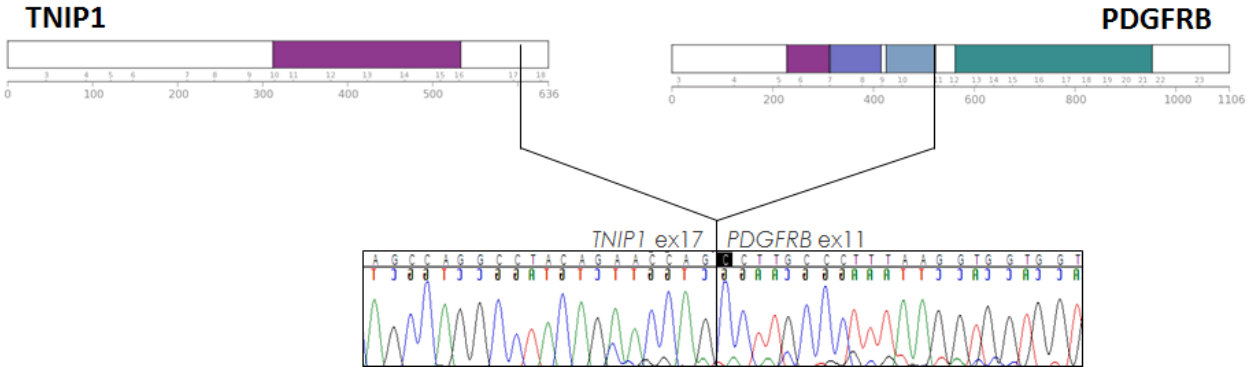


Figure 20. Identification of *TNIP1-PDGFRB* in pediatric CN-AML. Schematic representation of the two proteins and Sanger validation of the fusion breakpoint, confirming the juxtaposition of exon 11 of *PDGRFB* downstream the exon 17 of *TNIP1*.

As shown in Figure 21, preliminary data indicated that the expression of TNIP1-PDGFRB confers growth factor independence, suggesting that this chimeric protein could sustain an uncontrolled proliferation and could be important into leukemogenesis.

3.6.2 Ba/F3 cells expressing TNIP1-PDGFRB are sensitive to treatment with TKIs

Finally, the potential effect of several FDA-approved TKIs was investigated. Ba/F3 cells expressing TNIP1-PDGFRB were treated for 72h with increasing concentrations of imatinib, dasatinib and crenolanib (ABL1-kinase class inhibitors), and then analyzed using CellTiter-Blue® Cell Viability Assay. Noteworthy, all the drugs used showed a strong ability to inhibit the cytokine-independent proliferation as indicate in Figure 22

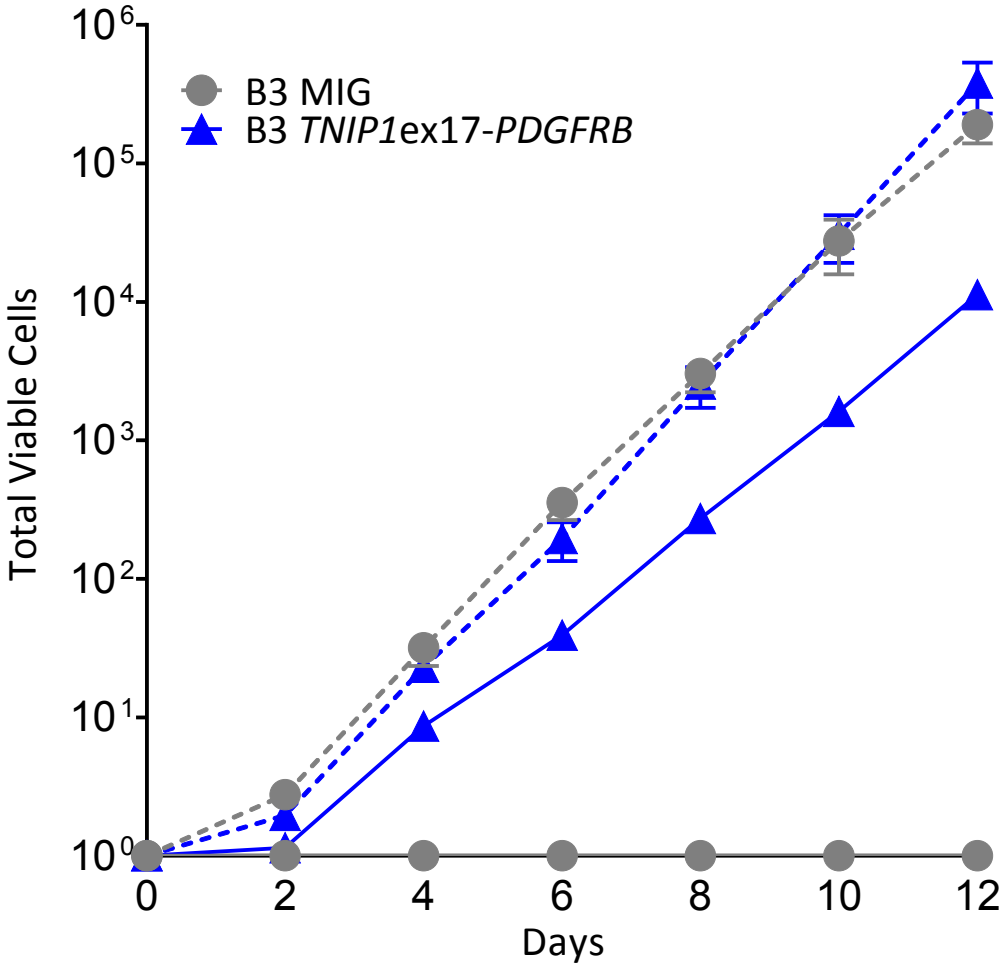
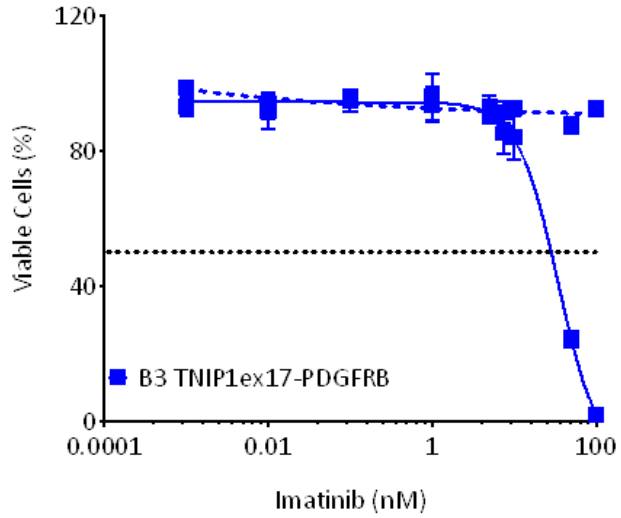
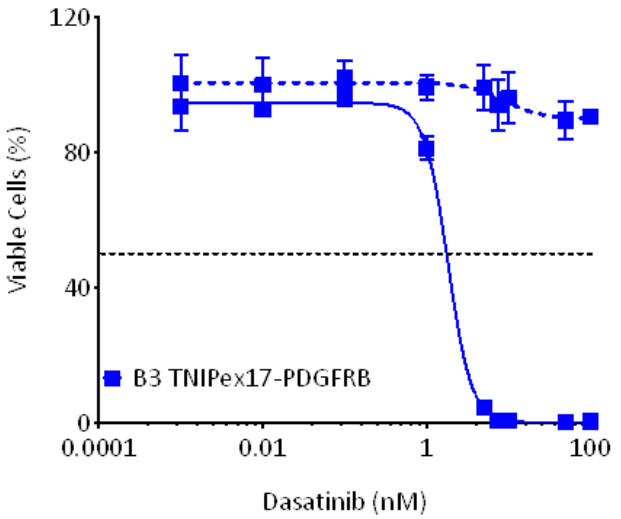
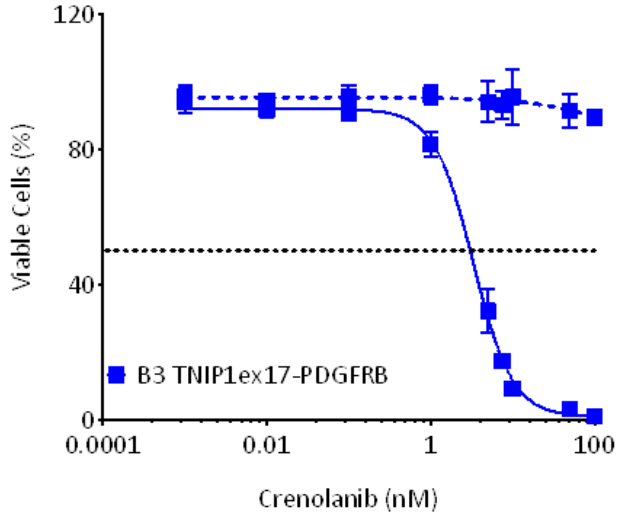


Figure 21. *TNIP1-PDGFRB* sustains cytokine-independent growth in-vitro. Mouse hematopoietic Ba/F3 cells were transduced with retroviral vector carrying MSCV-*TNIP1-PDGFRB*-IRES-GFP (blue curve) or MSCV-IRES-GFP (empty vector control) (gray curve) using RetroNectin (Takara Bio)-coated plates. After 3 days, cells were harvested and sorted by flow-cytometry for GPF expression. The expression of the fusion gene was checked at mRNA and protein level by RT-PCR/Sanger sequencing and western blot, respectively. To evaluate factor-independent growth, cells were washed three times with PBS 1X, and then seeded in triplicate at the concentration of 0.2M cells/well in 24-wells plate with (dotted-line) or without cytokine (solid-line). Cell proliferation and viability will be assessed every two days for a total of 12 days by trypan blue count using TC20™ Automated Cell Counter (Biorad). B3, Ba/F3 cells. Error bars represent mean ±SD of three independent experiments.







	Ba/F3 TNIP1-PDGFRB
Dasatinib IC50 (nM)	1.8
Imatinib IC50 (nM)	19.5
Crenolanib IC50 (nM)	5.2

Figure 22. Ba/F3 cells expressing TNIP1-PDGFRB are sensitive to treatment with TKIs. 72h treatment with TKIs in Ba/F3 cells expressing TNIP1-PDGFRB with (dotted-line) or without (solid-line) IL-3. B3, Ba/F3 cells. No cytotoxic effects were observed with cells maintained in factor indicating that imatinib specifically targets the activated PDGFRB and ABL1 kinases. Error bars represent mean ±SD of three independent experiments.

CHAPTER IV – Conclusions and future directions

Pediatric CN-AML patients have always showed a great variability in response to therapy and overall outcome, underlining the presence of unknown genetic changes, not detectable by conventional analyses, but relevant for pathogenesis, dynamics, and outcome of AML. Due to the lacking of any prognostic marker, "all-neg" CN-AML patients were historically classified in the intermediate-risk class. However, during the last decade, the development of novel genome-wide techniques such as next-generation sequencing, have tremendously improved our ability to interrogate the cancer genome, raising, for the first time, the possibility to have a complete picture of all the genetic alterations that could be present in a cancer cell. Along this line of research, the overall aim of this study was to investigate the mutational landscape of pediatric CN-AML patients negative for all the currently known somatic mutations reported in AML through whole-transcriptome sequencing.

Based on the results here obtained, the first consideration worth to be made is that RNA-seq revealed an high incidence of cryptic chromosomal rearrangements that generates fusion genes in pediatric CN-AML. Moreover, and more important, several of the fusion genes that were identified in this study are recurrent and could have a prognostic and/or therapeutic relevance. A paradigm of that is the *CBFA2T3-GLIS2* fusion, which has been demonstrate to be a novel common (8.4%) alteration in pediatric CN-AML, predicting poor outcome. Noteworthy, due to its prognostic relevance, the presence of this fusion gene will be included in the stratification of patients in the new AIEOP LAM 2012/01 protocol. Additionally, the identification in about 40% of *CBFA2T3-GLIS2*-positive patients of the *DHH-RHEBL1* read-through fusion transcript raise the possibility to define another subgroup of CN-AML with a very poor outcome, even if further studies are required to better characterized the prognostic significance of this alteration. Moving forward, whole-transcriptome sequencing enabled us to identify also a novel *out-of-frame* fusion transcript involving *PRDM16* and *SKI* genes. We demonstrated through q-PCR and gene

expression analysis base on RNA-seq data that *PRDM16-SKI*-positive patients showed a significant overexpression of *PRDM16* gene, suggesting that the juxtaposition of *SKI* downstream of exon 1 of *PRDM16* may lead to slipping of *PRDM16* and this, in turn, positions the gene under the control of ectopic promoter/enhancer elements. Interestingly, several reports indicate the link between overexpression of *PRDM16* and leukaemogenesis, strengthening its association with a worse outcome. Remarkably, preliminary results form a screening in 230 children with newly diagnosed *de novo* pediatric CN-AML indicate both a considerable incidence of this genomic alteration in this subgroup of childhood leukemias and a consequent overexpression of *PRDM16* gene (data not show), paving the way to further analysis of the prognostic value of *PRDM16-SKI* gene in this pathological setting.

Beyond the role of these fusion genes in defying the outcome of pediatric CN-AML patients, important findings have been obtained also regarding the possibility to develop targeted treatments against novel molecular alterations. On one side, the identification of *NUP98-PHF23* and *NUP98-JARID1A* fusions in 2.4% and 1.2% of pediatric CN-AML, respectively, raise the possibility to used disulfiram or other small molecules that inhibit the PHD binding to H3K4me3 also in paediatrics. On the other side, even if its presence was detected only in 1 out of 19 analyzed patients, the identification of the *in-frame TNIP1-PDGFRB* fusion is an important breakthrough in the use of tyrosine kinase inhibitors in AML. Through cloning and expression of *TNIP1-PDGFRB* in murine interleukin-3 (IL-3)-dependent mouse hematopoietic Ba/F3, here we demonstrated how the expression of this chimeric protein confers growth factor independence, suggesting that this chimeric protein have the ability to sustain an uncontrolled proliferation and could be important into leukemogenesis. Additionally, we also demonstrated that treatment with tyrosine kinase inhibitors blocks the cytokine-independent proliferation of Ba/F3 cells expressing TNIP1-PDGFRB and drive them to death. Taken together these data strongly indicate the possibility to use these targeted-drugs in *TNIP1-PDGFRB*-postive CN-AML pediatric patients and it is ongoing a screening in an independent and lager cohort of *de novo* pediatric CN-AML with the aim to investigate the recurrence of this alteration.

Despite the novel cryptic alteration identified so far in this study, several aspects need to be evaluated in the future to finally have a complete picture of the mutational landscape of pediatric CN-AML. Firstly, RNA-seq allows us to identify not only chromosomal rearrangements leading

to fusion genes, but also single-nucleotide-variants (SNVs), and insertion/deletion (InDel). However, in order to discriminate which SNVs are specifically present in the leukemic cells (somatic mutations), and which are wild-type variations (germ-line mutations) a diagnosis/germ-line paired analysis is required and the whole-exome sequencing of germ-line sample for the cases here analyzed is currently ongoing. Moreover, also the heterogeneity and the clonality of leukemic cells is an important aspect that we need to take in account: not all the leukemic cells show specific mutations and treatments can be more effective on specific clones rather than other. Base on that, the clonal dissection is a crucial to completely define the population of cells that could be present in a leukemic sample and it is now ongoing a paired whole-exome sequencing (diagnosis/remission/relapse) of three CN-AML pediatric patients of this study cohort in order to obtain data on the clonality of the identified mutation. The combination of the results here reported and the emerging data that will be obtained in the future by us and other important groups or consortia (e.g. TCGA, PCGP) are expected to be crucial to characterize the landscape of mutation in pediatric CN-AML, leading, ultimately, to make personalized medicine a reality in the near future.

LIST OF PAPERS

I. Masetti R, Pigazzi M, **Togni M**, Astolfi A, Indio V, Manara E, Casadio R, Pession A, Basso G, Locatelli F. CBFA2T3-GLIS2 fusion transcript is a novel common feature in pediatric, cytogenetically normal AML, not restricted to FAB M7 subtype. *Blood*. 2013 Apr 25;121(17):3469-72.

II. Masetti R*, **Togni M***, Astolfi A, Pigazzi M, Manara E, Indio V, Rizzari C, Rutella S, Basso G, Pession A, Locatelli F. DHH-RHEBL1 fusion transcript: a novel recurrent feature in the new landscape of pediatric CBFA2T3-GLIS2-positive acute myeloid leukemia. *Oncotarget*. 2013 Oct;4(10):1712-20. *** Co-first authors.**

III. Masetti R, **Togni M** $^{\delta}$, Astolfi A, Pigazzi M, Indio V, Rivalta B, Manara E, Rutella S, Basso G, Pession A, Locatelli F. Whole transcriptome sequencing of a paediatric case of de novo acute myeloid leukaemia with del(5q) reveals RUNX1-USP42 and PRDM16-SKI fusion transcripts. *Br J Haematol*. 2014 Aug;166(3):449-52. $^{\delta}$ **Corresponding author.**

IV. Masetti R, Rondelli R, Fagioli F, Mastronuzzi A, Pierani P, **Togni M**, Menna G, Pigazzi M, Putti MC, Basso G, Pession A, Locatelli F. Infants with acute myeloid leukemia treated according to the Associazione Italiana di Ematologia e Oncologia Pediatrica 2002/01 protocol have an outcome comparable to that of older children. *Haematologica*. 2014 Aug;99(8):e127-9.

V. **Togni M**, Masetti R, Pigazzi M, Astolfi A, Zama D, Indio V, Serravalle S, Manara E, Bisio V, Rutella S, Fagioli F, Basso, Pession A and Locatelli F. Identification of NUP98-PHF23 fusion gene in pediatric cytogenetically-normal acute myeloid leukemia by whole-transcriptome sequencing. *Br J Haematol*. 2015 Jan; (Submitted).

REFERENCES

1. Schuback HL, Arceci RJ, Meshinchi S. Somatic characterization of pediatric acute myeloid leukemia using next-generation sequencing. *Semin*. *Hematol*. 2013;50(4):325–32.

2. Tumori R. I TUMORI IN ITALIA - RAPPORTO 2012 I tumori dei bambini e degli adolescenti Cancer in children and adolescents. 2013;

3. Biondi A, Cimino G, Pieters R, Pui C. Biological and therapeutic aspects of infant leukemia. *Blood*. 2000;96(1):24–33.

4. Pui C-H, Carroll WL, Meshinchi S, Arceci RJ. Biology, risk stratification, and therapy of pediatric acute leukemias: an update. *J*. *Clin*. *Oncol*. 2011;29(5):551–65.

5. Kaspers GJL, Zwaan CM. Pediatric acute myeloid leukemia: towards high-quality cure of all patients. *Haematologica*. 2007;92(11):1519–32.

6. Rubnitz JE, Inaba H. Childhood acute myeloid leukaemia. *Br*. *J*. *Haematol*. 2012;159(3):259–76.

7. Creutzig U, Zimmermann M, Ritter J, et al. Treatment strategies and long-term results in paediatric patients treated in four consecutive AML-BFM trials. *Leukemia*. 2005;19(12):2030–42.

8. Gibson BES, Wheatley K, Hann IM, et al. Treatment strategy and long-term results in paediatric patients treated in consecutive UK AML trials. *Leukemia*. 2005;19(12):2130–8.

9. Abrahamsson J, Forestier E, Heldrup J, et al. Response-guided induction therapy in pediatric acute myeloid leukemia with excellent remission rate. *J*. *Clin*. *Oncol*. 2011;29(3):310–5.

10. Burnett AK, Hills RK, Milligan D, et al. Identification of patients with acute myeloblastic leukemia who benefit from the addition of gemtuzumab ozogamicin: results of the MRC AML15 trial. *J*. *Clin*. *Oncol*. 2011;29(4):369–77.

11. Cooper TM, Franklin J, Gerbing RB, et al. AAML03P1, a pilot study of the safety of gemtuzumab ozogamicin in combination with chemotherapy for newly diagnosed childhood acute myeloid leukemia: a report from the Children's Oncology Group. *Cancer*. 2012;118(3):761–9.

12. Pession A, Masetti R, Rizzari C, et al. Results of the AIEOP AML 2002/01 multicenter prospective trial for the treatment of children with acute myeloid leukemia. *Blood*. 2013;122(2):170–8.

13. Masetti R, Rondelli R, Fagioli F. with acute myeloid leukemia treated according to the Associazione Italiana di Ematologia e Oncologia Pediatrica 2002/01 protocol have an outcome. …. 2014;127–129.

14. Creutzig U, van den Heuvel-Eibrink MM, Gibson B, et al. Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. *Blood*. 2012;120(16):3187–205.

15. Tarlock K, Meshinchi S. Pediatric Acute Myeloid Leukemia: Biology and Therapeutic Implications of Genomic Variants. *Pediatr*. *Clin*. *North Am*. 2015;62(1):75–93.

16. Pui C-H. Actue myeloid leukemia in children treated with epipodophylloltoxins for acute lymphoblastic leukemia. *N*. *Engl*. *J*. *Med*. 1991;

17. Smith M, Rubinstein L. Secondary leukemia or myelodysplastic syndrome after treatment with epipodophyllotoxins. *J. Clin.* …. 1999;17(2):569–577.

18. Davies SM. Therapy-related leukemia associated with alkylating agents. *Med*. *Pediatr*. *Oncol*. 2001;36(5):536–40.

19. Ritchey AK, Land V, Steuber CP, Weinstein H. Acute Myeloid Leukemia (AML) in Down's Syndrome Is Highly Responsive to Chemotherapy: Experience on Pediatric Oncology Group AML Study 8498. 2014;80(9):2210–2214.

20. Shimamura A. Shwachman-Diamond syndrome. *Semin*. *Hematol*. 2006;43(3):178–88.

21. Valkhof MG, Sanders MA, Strien PMH Van, et al. Plenary paper Sequential gain of mutations in severe congenital neutropenia progressing to acute myeloid leukemia. 2014;119(22):5071–5078.

22. Tischkowitz M, Dokal I. Fanconi anaemia and leukaemia - clinical and molecular aspects. *Br*. *J*. *Haematol*. 2004;126(2):176–91.

23. Bennett JM, Catovsky D, Daniel MT, et al. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br*. *J*. *Haematol*. 1976;33(4):451–8.

24. Bennett JM, Catovsky D, Daniel MT, et al. Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. *Ann*. *Intern*. *Med*. 1985;103(4):620–5.

25. Vardiman JW, Thiele J, Arber D a, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009;114(5):937–51.

26. Gilliland DG, Jordan CT, Felix C a. The molecular basis of leukemia. *Hematology Am. Soc. Hematol. Educ. Program*. 2004;80–97.

27. Gilliland D, Griffin J. The roles of FLT3 in hematopoiesis and leukemia. *Blood*. 2002;100(5):1532–1543.

28. Nakao M, Janssen J, Erz D, Seriu T, Bartram C. Tandem duplication of the FLT3 gene in acute lymphoblastic leukemia: a marker for the monitoring of minimal residual disease. *Leukemia*. 2000;522–524.

29. Yamamoto Y, Kiyoi H, Nakano Y, et al. Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. 2014;97(8):2434–2440.

30. Dosil M, Wang S, Lemischka I. Mitogenic signalling and substrate specificity of the Flk2/Flt3 receptor tyrosine kinase in fibroblasts and interleukin 3-dependent hematopoietic cells. *Mol. Cell. Biol*. 1993;13(10):6572–6585.

31. Stirewalt DL, Radich JP. The role of FLT3 in haematopoietic malignancies. *Nat. Rev. Cancer*. 2003;3(9):650–65.

32. Schnittger S, Schoch C, Dugas M, et al. Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia : correlation to cytogenetics , FAB subtype , and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. 2014;100(1):59–67.

33. Kayser S, Levis MJ. FLT3 tyrosine kinase inhibitors in acute myeloid leukemia: clinical implications and limitations. *Leuk. Lymphoma*. 2014;55(2):243–55.

34. Paschka P, Marcucci G, Ruppert AS, 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. *J. Clin. Oncol*. 2006;24(24):3904–11.

35. Pollard J, Alonzo T. Prevalence and prognostic significance of KIT mutations in pediatric patients with core binding factor AML enrolled on serial pediatric cooperative trials for de novo. *Blood*. 2010;115(12):2372–2380.

36. Renneville a, Roumier C, Biggio V, et al. Cooperating gene mutations in acute myeloid leukemia: a review of the literature. *Leukemia*. 2008;22(5):915–31.

37. Tartaglia M, Niemeyer CM, Shannon KM, Loh ML. SHP-2 and myeloid malignancies. *Curr. Opin. Hematol*. 2004;11(1):44–50.

38. Tartaglia M, Mehler EL, Goldberg R, et al. Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat. Genet*. 2001;29(4):465–8.

39. Tartaglia M, Martinelli S, Iavarone I, et al. Somatic PTPN11 mutations in childhood acute myeloid leukaemia. *Br. J. Haematol*. 2005;129(3):333–9.

40. Rau R, Brown P. Nucleophosmin (NPM1) mutations in adult and childhood acute myeloid leukaemia: towards definition of a new leukaemia entity. *Hematol. Oncol*. 2009;(July):171–181.

41. Okuda M, Horn HF, Tarapore P, et al. Nucleophosmin/B23 Is a Target of CDK2/Cyclin E in Centrosome Duplication. *Cell*. 2000;103(1):127–140.

42. Schnittger S, Schoch C. Nucleophosmin gene mutations are predictors of favorable prognosis in acute myelogenous leukemia with a normal karyotype. …. 2005;106(12):3733–3740.

43. Verma A, Kambhampati S, Parmar S, Platanias LC. Jak family of kinases in cancer. *Cancer Metastasis Rev*. 2003;22(4):423–34.

44. Levine RL, Wadleigh M, Cools J, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell*. 2005;7(4):387–97.

45. Scott LM, Campbell PJ, East C, et al. Mechanisms of Disease Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. 2005;365:1054–1061.

46. Steensma DP, McClure RF, Karp JE, et al. JAK2 V617F is a rare finding in de novo acute myeloid leukemia, but STAT3 activation is common and remains unexplained. *Leukemia*. 2006;20(6):971–8.

47. Mrózek K, Marcucci G, Paschka P, Whitman SP, Bloomfield CD. Clinical relevance of mutations and gene-expression changes in adult acute myeloid leukemia with normal cytogenetics: are we ready for a prognostically prioritized molecular classification? *Blood*. 2007;109(2):431–48.

48. Manola KN. Cytogenetics of pediatric acute myeloid leukemia. *Eur. J. Haematol*. 2009;83(5):391–405.

49. Speck N a, Gilliland DG. Core-binding factors in haematopoiesis and leukaemia. *Nat. Rev. Cancer*. 2002;2(7):502–13.

50. Lowenberg B. Acute myeloid leukemia. *New Engl. J. ….* 1999;1051–1062.

51. Kawashima N, Shimada A, Taketani T, et al. Childhood acute myeloid leukemia with bone marrow eosinophilia caused by t(16;21)(q24;q22). *Int. J. Hematol*. 2012;95(5):577–80.

52. Hollink IHIM, van den Heuvel-Eibrink MM, Arentsen-Peters STCJM, et al. Characterization of CEBPA mutations and promoter hypermethylation in pediatric acute myeloid leukemia. *Haematologica*. 2011;96(3):384–92.

53. Fröhling S, Scholl C, Gilliland DG, Levine RL. Genetics of myeloid malignancies: pathogenetic and clinical implications. *J*. *Clin*. *Oncol*. 2005;23(26):6285–95.

54. Balgobind B V, Zwaan CM, Pieters R, Van den Heuvel-Eibrink MM. The heterogeneity of pediatric MLL-rearranged acute myeloid leukemia. *Leukemia*. 2011;25(8):1239–48.

55. Krivtsov A V, Armstrong S a. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat*. *Rev*. *Cancer*. 2007;7(11):823–33.

56. Mercher T, Raffel G. The OTT-MAL fusion oncogene activates RBPJ-mediated transcription and induces acute megakaryoblastic leukemia in a knockin mouse model. *J*. …. 2009;119(4):

57. Ley T, Ding L, Walter M. DNMT3A mutations in acute myeloid leukemia. … *Engl. J.* …. 2010;

58. Ho PA, Alonzo TA, Kopecky KJ, et al. Molecular alterations of the IDH1 gene in AML: a Children's Oncology Group and Southwest Oncology Group study. *Leukemia*. 2010;24(5):909–13.

59. Pigazzi M, Ferrari G, Masetti R, et al. Low prevalence of IDH1 gene mutation in childhood AML in Italy. *Leukemia*. 2011;25(1):173–4.

60. Campana D. Status of minimal residual disease testing in childhood haematological malignancies. *Br*. *J*. *Haematol*. 2008;143(4):481–9.

61. Meyer SC, Levine RL. Translational implications of somatic genomics in acute myeloid leukaemia. *Lancet*. *Oncol*. 2014;15(9):e382–94.

62. Mrózek K. Cytogenetic, molecular genetic, and clinical characteristics of acute myeloid leukemia with a complex karyotype. *Semin*. *Oncol*. 2008;35(4):365–77.

63. Ley TJ, Mardis ER, Ding L, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature*. 2008;456(7218):66–72.

64. Mardis EERE, Ding L, Dooling DJ, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. … *Engl. J.* …. 2009;361(11):1058–66.

65. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*. 2012;481(7382):506–10.

66. Welch J, Westervelt P, Ding L. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. *Jama*. 2011;305(15):1577–1584.

67. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. *Nat*. *Rev*. *Genet*. 2010;11(10):685–96.

68. Leukemia AM, Group TAG. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N*. *Engl*. *J*. *Med*. 2013;368(22):2059–74.

69. Ho P, Kutny M, Alonzo T. Leukemic Mutations in the Methylation-Associated Genes DNMT3A and IDH2 Are Rare Events in Pediatric AML : A Report From the Children ' s Oncology Group. *Pediatr. blood …*. 2011;(March):204–209.

70. Downing JR, Wilson RK, Zhang J, et al. The Pediatric Cancer Genome Project. *Nat*. *Genet*. 2012;44(6):619–22.

71. Gruber T a, Larson Gedman A, Zhang J, et al. An Inv(16)(p13.3q24.3)-encoded CBFA2T3-GLIS2 fusion protein defines an aggressive subtype of pediatric acute megakaryoblastic leukemia. *Cancer Cell*. 2012;22(5):683–97.

72. Pession A, Masetti R, Rizzari C, et al. Results of the AIEOP AML 2002/01 multicenter, prospective trial for treatment of children with acute myeloid leukemia. *Blood*. 2013;

73. Trapnell C, Pachter L, Salzberg SLS. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009;942–949.

74. McPherson A, Hormozdiari F, Zayed A, et al. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. *PLoS Comput*. *Biol*. 2011;7(5):e1001138.

75. Iyer MK, Chinnaiyan AM, Maher CA. ChimeraScan: a tool for identifying chimeric transcription in sequencing data. *Bioinformatics*. 2011;27(20):2903–4.

76. Ge H, Liu K, Juan T, et al. FusionMap: detecting fusion genes from next-generation sequencing data at base-pair resolution. *Bioinformatics*. 2011;27(14):1922–8.

77. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.

78. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25.

79. Berger R, Le Coniat M, Romana SP, Jonveaux P. Secondary acute myeloblastic leukemia with t(16;21) (q24;q22). involving the AML1 gene. *Hematol*. *Cell Ther*. 1996;38(2):183–6.

80. Gamou T, Kitamura E, Hosoda F, et al. The partner gene of AML1 in t(16;21) myeloid malignancies is a novel member of the MTG8(ETO) family. *Blood*. 1998;91(11):4028–37.

81. Li B, Rauhauser AA, Dai J, et al. Increased hedgehog signaling in postnatal kidney results in aberrant activation of nephron developmental programs. *Hum*. *Mol*. *Genet*. 2011;20(21):4155–66.

82. Thiollier C, Lopez CK, Gerby B, et al. Characterization of novel genomic alterations and therapeutic approaches using acute megakaryoblastic leukemia xenograft models. *J*. *Exp*. *Med*. 2012;209(11):2017–31.

83. Crispino JD, Le Beau MM. BMP meets AML: induction of BMP signaling by a novel fusion gene promotes pediatric acute leukemia. *Cancer Cell*. 2012;22(5):567–8.

84. Maher C a, Kumar-Sinha C, Cao X, et al. Transcriptome sequencing to detect gene fusions in cancer. *Nature*. 2009;458(7234):97–101.

85. Ingham PW, McMahon AP. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev*. 2001;15(23):3059–87.

86. Mitelman F, Johansson B, Mertens F. Fusion genes and rearranged genes as a linear function of chromosome aberrations in cancer. *Nat*. *Genet*. 2004;36(4):331–4.

87. Barakat MT, Humke EW, Scott MP. Learning from Jekyll to control Hyde: Hedgehog signaling in development and cancer. *Trends Mol*. *Med*. 2010;16(8):337–48.

88. Ng JMY, Curran T. The Hedgehog's tale: developing strategies for targeting cancer. *Nat*. *Rev*. *Cancer*. 2011;11(7):493–501.

89. Yuan J, Shan Y, Chen X, et al. Identification and characterization of RHEBL1, a novel member of Ras family, which activates transcriptional activities of NF-kappa B. *Mol*. *Biol*. *Rep*. 2005;32(4):205–14.

90. Tee AR, Blenis J, Proud CG. Analysis of mTOR signaling by the small G-proteins, Rheb and RhebL1. *FEBS Lett*. 2005;579(21):4763–8.

91. Baud V, Karin M. Is NF-kappaB a good target for cancer therapy? Hopes and pitfalls. *Nat*. *Rev*. *Drug Discov*. 2009;8(1):33–40.

92. Chappell WH, Steelman LS, Long JM, et al. Ras/Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR inhibitors: rationale and importance to inhibiting these pathways in human health. *Oncotarget*. 2011;2(3):135–64.

93. Zhou H, Huang S. Role of mTOR signaling in tumor cell motility, invasion and metastasis. *Curr*. *Protein Pept*. *Sci*. 2011;12(1):30–42.

94. Graf M, Hecht K, Reif S, et al. Expression and prognostic value of hemopoietic cytokine receptors in acute myeloid leukemia (AML): implications for future therapeutical strategies. *Eur. J. Haematol*. 2004;72(2):89–106.

95. Quentmeier H, Tonelli R, Geffers R, et al. Expression of BEX1 in acute myeloid leukemia with MLL rearrangements. *Leukemia*. 2005;19(8):1488–9.

96. Carraway KL, Perez A, Idris N, et al. Muc4/sialomucin complex, the intramembrane ErbB2 ligand, in cancer and epithelia: to protect and to survive. *Prog. Nucleic Acid Res. Mol. Biol*. 2002;71:149–85.

97. Emaduddin M, Edelmann MJ, Kessler BM, Feller SM. Odin (ANKS1A) is a Src family kinase target in colorectal cancer cells. *Cell Commun. Signal*. 2008;6:7.

98. Paulsson K, Békássy AN, Olofsson T, et al. A novel and cytogenetically cryptic t(7;21)(p22;q22) in acute myeloid leukemia results in fusion of RUNX1 with the ubiquitin-specific protease gene USP42. *Leukemia*. 2006;20(2):224–9.

99. Foster N, Paulsson K, Sales M, et al. Molecular characterisation of a recurrent, semi-cryptic RUNX1 translocation t(7;21) in myelodysplastic syndrome and acute myeloid leukaemia. *Br. J. Haematol*. 2010;148(6):938–43.

100. Panagopoulos I, Gorunova L, Brandal P, et al. Myeloid leukemia with t(7;21)(p22;q22) and 5q deletion. *Oncol. Rep*. 2013;30(4):1549–52.

101. Giguère A, Hébert J. Microhomologies and topoisomerase II consensus sequences identified near the breakpoint junctions of the recurrent t(7;21)(p22;q22) translocation in acute myeloid leukemia. *Genes. Chromosomes Cancer*. 2011;50(4):228–38.

102. Nishikata I, Sasaki H, Iga M, et al. A novel EVI1 gene family, MEL1, lacking a PR domain (MEL1S) is expressed mainly in t(1;3)(p36;q21)-positive AML and blocks G-CSF-induced myeloid differentiation. *Blood*. 2003;102(9):3323–32.

103. Duhoux FP, Ameye G, Montano-Almendras CP, et al. PRDM16 (1p36) translocations define a distinct entity of myeloid malignancies with poor prognosis but may also occur in lymphoid malignancies. *Br. J. Haematol*. 2012;156(1):76–88.

104. Jeandidier E, Gervais C, Radford-Weiss I, et al. A cytogenetic study of 397 consecutive acute myeloid leukemia cases identified three with a t(7;21) associated with 5q abnormalities and exhibiting similar clinical and biological features, suggesting a new, rare acute myeloid leukemia entity. *Cancer Genet*. 2012;205(7-8):365–72.

105. Reader JC, Meekins JS, Gojo I, Ning Y. A novel NUP98-PHF23 fusion resulting from a cryptic translocation t(11;17)(p15;p13) in acute myeloid leukemia. *Leukemia*. 2007;21(4):842–4.

106. Pigazzi M, Manara E, Bisio V, et al. Screening of novel genetic aberrations in pediatric acute myeloid leukemia: a report from the AIEOP AML-2002 study group. *Blood*. 2012;120(18):3860–2.

107. Romana SP, Radford-Weiss I, Ben Abdelali R, et al. NUP98 rearrangements in hematopoietic malignancies: a study of the Groupe Francophone de Cytogénétique Hématologique. *Leukemia*. 2006;20(4):696–706.

108. De Rooij JDE, Hollink IHIM, Arentsen-Peters STCJM, et al. NUP98/JARID1A is a novel recurrent abnormality in pediatric acute megakaryoblastic leukemia with a distinct HOX gene expression pattern. *Leukemia*. 2013;(March):1–9.

109. Reader JC, Leng Q, Rassool F V, Ning Y. Regulation of differentiation by a PHD domain in the NUP98-PHF23 fusion protein. *Leuk*. *Res*. 2010;34(8):1094–7.

110. Gough SM, Lee F, Yang F, et al. NUP98-PHF23 Is a Chromatin-Modifying Oncoprotein That Causes a Wide Array of Leukemias Sensitive to Inhibition of PHD Histone Reader Function. *Cancer Discov*. 2014;

111. Jones a V, Cross NCP. Oncogenic derivatives of platelet-derived growth factor receptors. *Cell*. *Mol*. *Life Sci*. 2004;61(23):2912–23.

112. Ostman A, Heldin C-H. PDGF receptors as targets in tumor treatment. *Adv*. *Cancer Res*. 2007;97(06):247–74.

113. Roberts KG, Li Y, Payne-Turner D, et al. Targetable Kinase-Activating Lesions in Ph-like Acute Lymphoblastic Leukemia. *N*. *Engl*. *J*. *Med*. 2014;371(11):1005–1015.

114. Lengline E, Beldjord K, Dombret H. Successful tyrosine kinase inhibitor therapy in a refractory B-cell precursor acute lymphoblastic leukemia with EBF1-PDGFRB fusion. …. 2013;146–148.

115. Weston B, Hayden M. Tyrosine Kinase Inhibitor Therapy Induces Remission in a Patient With Refractory EBF1-PDGFRB–Positive Acute Lymphoblastic Leukemia. *J. Clin.* .... 2013;31(25):

116. Kobayashi K, Mitsui K, Ichikawa H, et al. ATF7IP as a novel PDGFRB fusion partner in acute lymphoblastic leukaemia in children. *Br*. *J*. *Haematol*. 2014;1–6.

117. Golub TR, Barker GF, Lovett M, Gilliland DG. Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell*. 1994;77(2):307–16.

118. Abe A, Emi N, Tanimoto M, et al. Fusion of the platelet-derived growth factor receptor beta to a novel gene CEV14 in acute myelogenous leukemia after clonal evolution. *Blood*. 1997;90(11):4271–7.

119. Ross TS, Bernard OA, Berger R, Gilliland DG. Fusion of Huntingtin interacting protein 1 to platelet-derived growth factor beta receptor (PDGFbetaR) in chronic myelomonocytic leukemia with t(5;7)(q33;q11.2). *Blood*. 1998;91(12):4419–26.

120. Kulkarni S, Heath C, Parker S, et al. Fusion of H4/D10S170 to the platelet-derived growth factor receptor beta in BCR-ABL-negative myeloproliferative disorders with a t(5;10)(q33;q21). *Cancer Res*. 2000;60(13):3592–8.

121. Magnusson MK, Meade KE, Brown KE, et al. Rabaptin-5 is a novel fusion partner to platelet-derived growth factor beta receptor in chronic myelomonocytic leukemia. *Blood*. 2001;98(8):2518–25.

122. Wilkinson K, Velloso ERP, Lopes LF, et al. Cloning of the t(1;5)(q23;q33) in a myeloproliferative disorder associated with eosinophilia: involvement of PDGFRB and response to imatinib. *Blood*. 2003;102(12):4187–90.

123. Grand FH, Burgstaller S, Kühr T, et al. p53-Binding protein 1 is fused to the platelet-derived growth factor receptor beta in a patient with a t(5;15)(q33;q22) and an imatinib-responsive eosinophilic myeloproliferative disorder. *Cancer Res*. 2004;64(20):7216–9.

124. Vizmanos JL, Novo FJ, Román JP, et al. NIN, a gene encoding a CEP110-like centrosomal protein, is fused to PDGFRB in a patient with a t(5;14)(q33;q24) and an imatinib-responsive myeloproliferative disorder. *Cancer Res*. 2004;64(8):2673–6.

125. Morerio C, Acquila M, Rosanda C, et al. HCMOGT-1 is a novel fusion partner to PDGFRB in juvenile myelomonocytic leukemia with t(5;17)(q33;p11.2). *Cancer Res*. 2004;64(8):2649–51.

126. Levine RL, Wadleigh M, Sternberg DW, et al. KIAA1509 is a novel PDGFRB fusion partner in imatinib-responsive myeloproliferative disease associated with a t(5;14)(q33;q32). *Leukemia*. 2005;19(1):27–30.

127. Rosati R, La Starza R, Luciano L, et al. TPM3/PDGFRB fusion transcript and its reciprocal in chronic eosinophilic leukemia. *Leukemia*. 2006;20(9):1623–4.

128. La Starza R, Rosati R, Roti G, et al. A new NDE1/PDGFRB fusion transcript underlying chronic myelomonocytic leukaemia in Noonan Syndrome. *Leukemia*. 2007;21(4):830–3.

129. Walz C, Metzgeroth G, Haferlach C, et al. Characterization of three new imatinib-responsive fusion genes in chronic myeloproliferative disorders generated by disruption of the platelet-derived growth factor receptor beta gene. *Haematologica*. 2007;92(2):163–9.

130. Gallagher G, Horsman DE, Tsang P, Forrest DL. Fusion of PRKG2 and SPTBN1 to the platelet-derived growth factor receptor beta gene (PDGFRB) in imatinib-responsive atypical myeloproliferative disorders. *Cancer Genet. Cytogenet*. 2008;181(1):46–51.

131. Gorello P, La Starza R, Brandimarte L, et al. A PDGFRB-positive acute myeloid malignancy with a new t(5;12)(q33;p13.3) involving the ERC1 gene. *Leukemia*. 2008;22(1):216–8.

132. Walz C, Haferlach C, Hänel A, et al. Identification of a MYO18A-PDGFRB fusion gene in an eosinophilia-associated atypical myeloproliferative neoplasm with a t(5;17)(q33-34;q11.2). *Genes*. *Chromosomes Cancer*. 2009;48(2):179–83.

133. Erben P, Gosenca D, Müller MC, et al. Screening for diverse PDGFRA or PDGFRB fusion genes is facilitated by generic quantitative reverse transcriptase polymerase chain reaction analysis. *Haematologica*. 2010;95(5):738–44.

134. Hidalgo-Curtis C, Apperley JF, Stark A, et al. Fusion of PDGFRB to two distinct loci at 3p21 and a third at 12q13 in imatinib-responsive myeloproliferative neoplasms. *Br*. *J*. *Haematol*. 2010;148(2):268–73.

135. Medves S, Duhoux FP, Ferrant a, et al. KANK1, a candidate tumor suppressor gene, is fused to PDGFRB in an imatinib-responsive myeloid neoplasm with severe thrombocythemia. *Leukemia*. 2010;24(5):1052–5.

136. Roberts KG, Morin RD, Zhang J, et al. Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. *Cancer Cell*. 2012;22(2):153–66.

137. Chmielecki J, Peifer M, Viale A, et al. Systematic Screen for Tyrosine Kinase Rearrangements Identifies a Novel C6orf204-PDGFRB Fusion in a Patient with Recurrent T -ALL and an Associated Myeloproliferative Neoplasm. 2012;65(September 2011):54–65.

138. Winkelmann N, Hidalgo-Curtis C, Waghorn K, et al. Recurrent CEP85L-PDGFRB fusion in patient with t(5;6) and imatinib-responsive myeloproliferative neoplasm with eosinophilia. *Leuk*. *Lymphoma*. 2013;54(7):1527–31.

139. Gosenca D, Kellert B, Metzgeroth G, et al. Identification and functional characterization of imatinib-sensitive DTD1-PDGFRB and CCDC88C-PDGFRB fusion genes in eosinophilia-associated myeloid/lymphoid neoplasms. *Genes*. *Chromosomes Cancer*. 2014;53(5):411–21.

ACKNOWLEDGEMENTS

I thank Prof. Andrea Pession for being a dedicated and careful mentor in all these years, Dr. Riccardo Masetti and Dr. Annalisa Astolfi for their guide in the laboratory and for our continuous and productive scientific discussions. I thank the whole Pediatric Oncology and Hematology Laboratory "Lalla Seràgnoli" at S.Orsola-Malpighi Hospital, University of Bologna (Bologna, Italy).

I would like also to thank Dr Charles Mullighan for the great opportunity to perform part of my research in his laboratory at St. Jude Children's Research Hospital (Memphis, Tennesse, USA) and for his mentoring during this time.

This study has been supported by Department of Experimental, Diagnostic and Specialty Medicine (DIMES) – University of Bologna, Department of Medical and Surgical Science (DIMEC) – University of Bologna, Giorgio Prodi Cancer Research Center (CIRC) – University of Bologna, Associazione Italiana Ematologia Oncologia Pediatrica (AIEOP) and Fondazione Umberto Veronesi, Milan.