Deciphering the genetic diversity of pediatric AML using Next Generation Sequencing tools

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

Dr. rer. nat.

der Fakultät für

Biologie

an der

Universität Duisburg-Essen

vorgelegt von

Christiane Walter

aus Lemgo

08/2019

Angaben zur Prüfung

Die der vorliegenden Arbeit zugrunde liegenden Experimente wurden in der Klinik für Kinderheilkunde III in der Arbeitsgruppe Pädiatrische Hämatologie und Onkologie des Universitätsklinikums Essen durchgeführt.

1. Gutachter: Prof. Dr. Dirk Reinhardt

2. Gutachter: Prof. Dr. Matthias Gunzer

Vorsitzender des Prüfungsausschusses: Prof. Dr. Daniel Hoffmann

Tag der mündlichen Prüfung: 14.02.2020

DOI: 10.17185/duepublico/71940
URN: urn:nbn:de:hbz:464-20200729-115459-4

Table of Contents

1 Zusammenfassung

Die pädiatrische akute myeloische Leukämie (AML) ist eine heterogene Erkrankung, bei der neben der morphologischen und immunphänotypischen Charakterisierung vor allem genetische Veränderungen eine große Rolle spielen. Dabei kann die genaue Kenntnis der genetischen Signatur leukämischer Blasten und Stammzellen (LSC) genutzt werden, um die klonale Struktur der Erkrankung zu untersuchen. Die Kenntnis dieser Signatur könnte zu einer Verbesserung klonspezifischer Therapien und nachfolgend zu einer Verringerung der durch die Therapie verursachten Toxizität führen. Dabei kann insbesondere die Analyse auf Einzelzellebene weiterführende Informationen liefern.

In der vorliegenden Dissertation wurden verschiedene Next Generation Sequencing (NGS) Verfahren etabliert. Diese ermöglichen eine detaillierte genetische Analyse der pädiatrischen AML, auch auf Einzelzellebene. Durch Korrelation spezifischer genetischer Aberrationen mit den klinischen Daten konnten wir bei den pädiatrischen AML-Patienten Subgruppen identifizieren, die ein schlechteres Überleben im Vergleich zur gesamten Gruppe zeigen. Wir konnten zeigen, dass Patienten, bei denen eine Kombination aus zwei von drei Faktoren (*WT1* Mutation, *FLT3* ITD und *NUP98-NSD1* Fusion) vorliegt, eine signifikant schlechtere Prognose haben. Für diese Risikogruppen ist die Verfügbarkeit von neuen Therapien dringend notwendig.

Wir wendeten außerdem NGS-Technologien für die genetische Analyse einzelner LSC-angereicherter Blasten von pädiatrischen AML-Patienten mit bekannter *NPM1* Mutation und/oder *FLT3* ITD an. Für die Sequenzanalyse der DNA der leukämischen Blasten wurde Knochenmark von pädiatrischen AML Patienten mittels genspezifischer amplikonbasierter Sequenzierung (TS) sowie Exomsequenzierung (WES) untersucht. Parallel wurde DNA einzelner Zellen mit stammzellnahem Immunphänotyp extrahiert, das Genom amplifiziert (WGA) und anschließend mittels vollständiger Exomsequenzierung (WES) analysiert.

Wir konnten zeigen, dass die Qualität der Genomamplifikation aus Einzelzellen für die Generierung von auswertbaren NGS Daten geeignet war. Die statistische Auswertung der Sequenzdaten zeigte außerdem eine gute Vergleichbarkeit zwischen den Einzelzell- und den aus den aus dem nicht angereicherten Knochenmark generierten Daten. Der direkte Vergleich der verschiedenen NGS-Methoden ergab,

dass die mit WES erreichte Lesetiefen zwischen 50-72X nicht zur Identifizierung von seltenen AML-Subklonen ausreichend sind.

2 Summary

Pediatric acute myeloid leukemia (AML) is characterized by various genetic changes that can be used for classification in addition to morphological and immunophenotypic criteria. To determine the genetic signature of this disease might help to improve clone-specific therapies and to reduce the toxicity caused by therapy. In order to analyze the genetic profile even of leukemic stem cells, which are described to be the origin of the disease, analysis at the individual cell level in particular can provide more accurate information about clonal evolution in pediatric AML.

In this thesis, different NGS tools have been established that enabled detailed and in-depth genetic analysis of pediatric AML subpopulations on both bulk and single cell level. Linking the presence of combinations of specific genetic aberrations with the clinical data of the AML patients has revealed the existence of subgroups of children and adolescents with devastatingly poor outcome that urgently need novel therapies.

We used NGS technologies for reliable genetic analysis of single LSC-enriched blasts from pediatric AML patients with known *NPM1* mutation and/or *FLT3* ITD. Two NGS techniques, targeted amplicon-based sequencing (TS) and whole exome sequencing (WES) was applied to bone marrow bulk cells from pediatric AML patients at diagnosis. In parallel, single AML cells were sorted from a LSC-enriched compartment and DNA was extracted, whole genome-amplified (WGA) and subsequently subjected to whole exome sequencing (WES). We showed that the quality of the amplification achieved with a commercially available WGA kit was sufficient to generate evaluable sequence data and that coverage statistics were comparable between single cell and bulk WES data. Direct comparison between the different NGS methods also revealed that the initial screening of the bulk DNA by WES with read counts between 50–72× failed to detect rare AML subclones.

Additionally, we applied targeted NGS for 54 genes associated with myeloid malignancies, which are marker predicting a poor outcome in a contemporary pediatric AML cohort. Two of these genes further analyzed in this thesis were *FLT3* and *WT1.* By using qPCR, we identified another co-occuring prognostic indicator in a subgroup of patients with material for retrospective analysis available, namely *NUP98-NSD1* fusions. We could show that co-occurrence of any combination of *WT1* mutation, *FLT3* ITD and *NUP98-NSD1* fusion is a predictor of dismal response for pediatric AML.

3 Introduction

3.1 Biology of hematopoietic and leukemic stem cells

Hematopoietic stem cells (HSCs) are rare cells within the human bone marrow (BM), which are capable of both self-renewal and differentiation with the balance between these two stages being strictly regulated[105]. In healthy hematopoiesis, HSCs guarantee a constant production of multilineage precursor cells which differentiate in all types of mature hematopoietic cells such as red blood cells, megakaryocytes, myeloid cells (monocytes/macrophages, neutrophils) and lymphocytes (**Figure 1**)[110]. This continuous production of cells is necessary, as mature blood cells have relatively short lifespans[110].

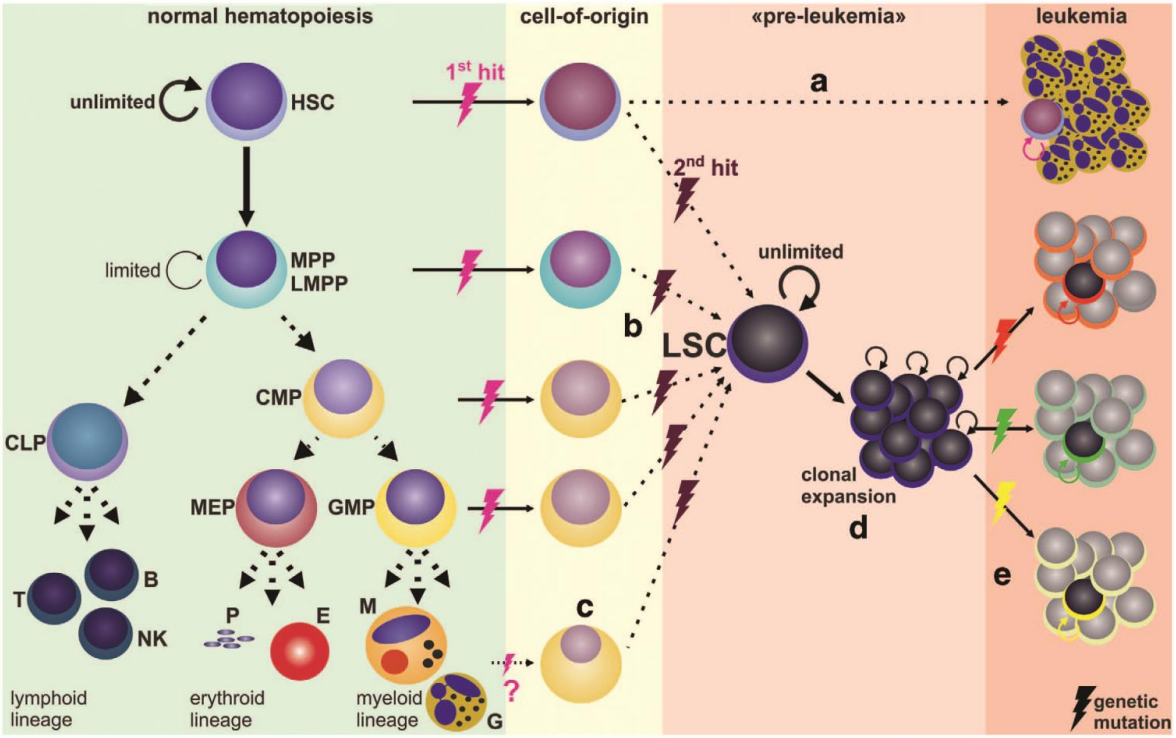


Figure 1: The hematopoietic and leukemic stem cell model. Hematopoietic stem cells (HSCs) give rise to multipotent progenitors (MPPs). MPPs further differentiate toward oligopotent lineage-restricted progenitors (LRPs), such as common lymphoid and myeloid progenitors (CLPs, CMPs) and granulocyte-macrophage progenitors (GMPs). LRPs proliferate intensely and produce all mature blood cell types required. From Riether *et al.* (2015)[125]

The concept that HSCs reside within the hematopoietic niche has first been described in the 1970s[132]. The hematopoietic niche is a complex system composed of a large number of cell types with distinct functions providing both chemical signals as well as physical interactions[83,99,105,128]. HSCs are thought to be in direct and also indirect contact with several cell types residing in the niche including macrophages,

adipocytes, mesenchymal stromal cells, osteoclasts, or megakaryocytes[83]. Soluble factors like cytokines and growth factors also participate in the regulation of stem cell maintenance and differentiation[83].

In contrast to normal hematopoiesis, in hematopoietic malignancies including leukemias remodeling of this microenvironment could be an important event in the disease development[128]. The modified microenvironment is on the other hand thought to be involved in controlling the maintenance and activity of leukemic stem cells (LSCs)[128]. LSCs are thought to originate from HSCs, but often are considered to be slightly more differentiated precursor/progenitor cells due to somatic mutations that affect proliferation and differentiation[20,48,77] as shown in **Figure 1**. Using CD34, CD38 and lineage-specific markers, LSCs have been first described according to their surface antigen expression to reside in the CD34+CD38- fraction[1,18] as these cells reconstituted human AML in immunodeficient mice[1,17,20]. More recent studies demonstrated that LSCs can also be present in the CD34+CD38+ or even in the CD34- fraction[85]. Several candidate antigens specifically distinguishing normal HSCs and LSCs were investigated during the past years, for example CD123 (IL-3R), TIM3 or CD96[64,103,153]. Since LSCs cannot be grown *in vitro* in long-term cell culture, different surrogate assays such as engraftment in immunodeficient mice, colony forming capacity, differentiation ability or long-term culture initiating activity have been used to functionally describe the LSC phenotype[48].

During therapy, LSCs can persist in the bone marrow of mice and patients and are able to regrow the leukemia as a relapse of the primary disease[56,77,159]. To overcome this, it is necessary to specifically target and eliminate these primitive cells for long-term cure of patients[56,77,159].

3.2 Clinical features and diagnosis of pediatric AML

Frequently occurring malignant diseases of the hematopoietic system in children and also adults are myeloproliferative neoplasms (MPNs), myelodysplastic syndrome (MDS), chronic myeloid leukemia (CML), acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL)[5]. As a chronic disease, pediatric CML is associated with preleukemic genetic events within progenitor cells resulting in advantages in survival and proliferation of these cells while terminally differentiated hematopoietic cells can

still be produced[73] (**Figure 2**). Additional changes in these altered progenitor cells leading to a block in differentiation then manifest as blast crisis. MDS are clonal disorders characterized by ineffective hematopoiesis and subsequent frequent development of AML[116]. Acute diseases such as AML and ALL are characterized by the malignant transformation of HSCs or progenitor cells that results in the accumulation of immature precursor cells and disruption of the normal hematopoiesis[73] (**Figure 2**).

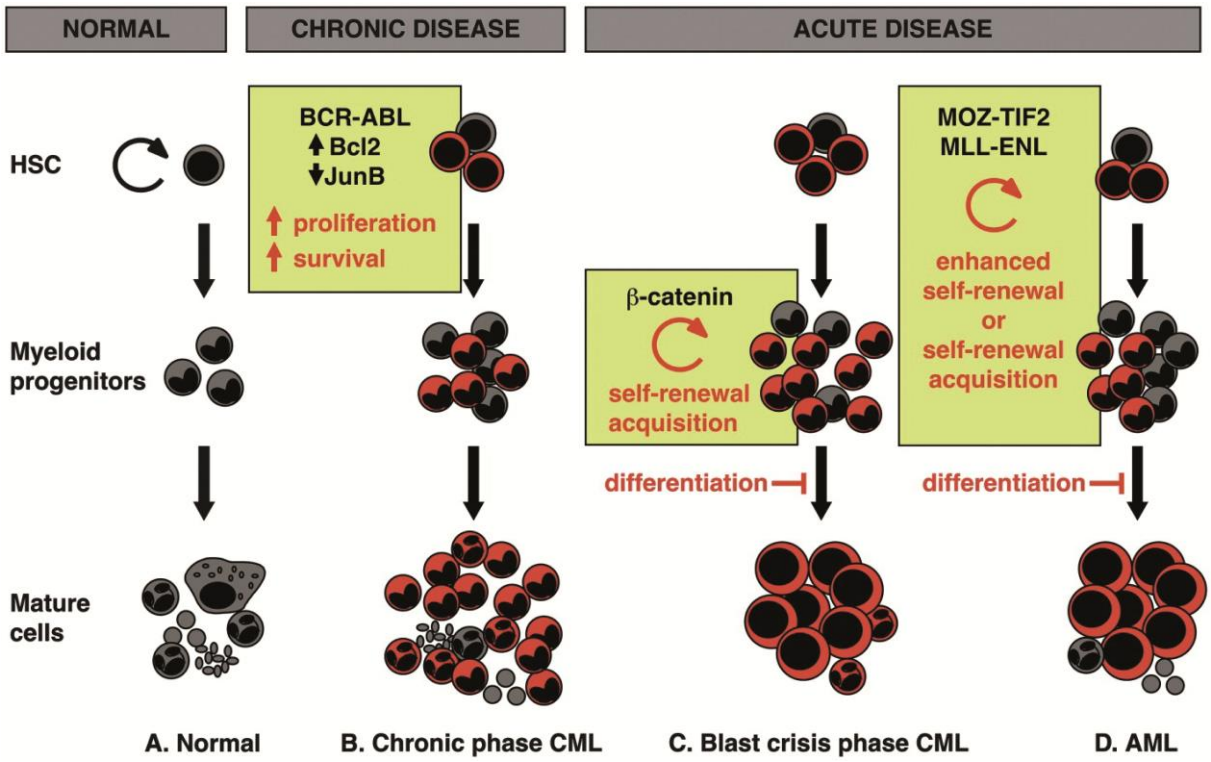


Figure 2: Pathogenesis of leukemias. Hematopoietic stem cells (HSCs) differenciate into committed myeloid progenitors and their respective terminally differentiated progeny (A). In chronic phase CML, preleukemic events that result in increased survival and proliferation occur within the stem and myeloid progenitor populations. Continued production of terminally differentiated progeny is still maintained (B). Acute diseases such as blast crisis CML and AML are characterized by acquisition of self-renewal capacity by progenitors or by enhanced self-renewal in HSC. Together with a subsequent block in differentiation, these leukemogenic events result in the accumulation of immature blasts and development of AML (C and D). From Jamieson *et al.* (2004)[73]

AML is the second most common form of pediatric acute leukemia with an incidence of 7 cases per million children younger than 15 years[32]. It is characterized by uncontrolled proliferation and impaired differentiation of malignant myeloid precursor cells, also called leukemic blasts[32,162]. Children suffering from AML may show clinical symptoms at diagnosis such as fever (40%), fatigue (19%), cutanous or mucosal bleeding (33%), bone pain (18%), pallor (25%) and weight loss (22%)[26]. The

diagnosis is made from bone marrow aspiration, in which more than 20% leukemic blasts have to be present[151]. Some of these blasts may contain Auer Rods, which are clumps of azurophilic granular material containing procoagulants[26]. Leukemic blasts are found mainly in the BM, but accumulation in the peripheral blood (PB) and organs such as spleen or liver are common features of AML[37]. Although the central nervous system is not often affected in AML, the examination of the cerebrospinal fluid for the presence of AML blasts is mandatory[26,32].

Table 1: WHO classification of myeloid neoplasms and acute leukemia From Arber *et al.* (2016) [5]

Acute myeloid leukemia (AML) and related neoplasms
AML with recurrent genetic abnormalities AML with t(8;21)(q22;q22.1);*RUNX1-RUNX1T1* AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22);*CBFB-MYH11* APL with *PML-RARA* AML with t(9;11)(p21.3;q23.3);*MLLT3-KMT2A* AML with t(6;9)(p23;q34.1);*DEK-NUP214* AML with inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2); *GATA2, MECOM* AML (megakaryoblastic) with t(1;22)(p13.3;q13.3);*RBM15-MKL1* Provisional entity: AML with *BCR-ABL1* AML with mutated *NPM1* AML with biallelic mutations of *CEBPA* Provisional entity: AML with mutated *RUNX1*
AML with myelodysplasia-related changes
Therapy-related myeloid neoplasms
AML, not otherwise specified AML with minimal differentiation AML without maturation AML with maturation Acute myelomonocytic leukemia Acute monoblastic/monocytic leukemia Pure erythroid leukemia Acute megakaryoblastic leukemia Acute basophilic leukemia Acute panmyelosis with myelofibrosis
Myeloid sarcoma
Myeloid proliferations related to Down syndrome Transient abnormal myelopoiesis (TAM) Myeloid leukemia associated with Down syndrome

Since AML is a very heterogeneous disease, several attempts to define subgroups have been made during the past decades. Two important proposals to classify AML into subtypes that became accepted are the French-American-British (FAB) classification and the newer World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia. AML has been divided into eight different subgroups (AML FAB M0-M7) based on the recognition of granulocytic (M1, M2 and M3), granulocytic-monocytic (M4), monocytic (M5a, M5b), erythroid (M6) and megakaryocytic (M7) types of cells by experts from France, America and Britain in the late 70s[12,13]. This classification has been widely accepted and is still used to morphologically, cytochemically and partly immunophenotypically define the eight subtypes of AML.

According to the latest revision of the WHO classification from 2016, AML can also be differentiated into four main subgroups using different criteria as shown in **Table 1**[144]. In recent years, genetic aberrations have become more important for the classification of AML and the largest subgroup in the WHO classification is defined as *AML with recurrent genetic abnormalities*[5]. The remaining subgroups contain AML with myelodysplasias-related changes, therapy-related myeloid neoplasms, myeloid proliferations related to Down syndrome and myelosarcomas[5], which represent an extramedullary tumor originating from granulocytic cells[3]. All other types fall into the group of not otherwise specified AML.

Additionally, the following technics are used for characterization of the leukemic blasts: **Immunophenotyping** defines the immunological phenotype of AML blasts including aberrant or asynchronous expression of surface antigens as well as cytoplasmatic factors. **Karyotyping** is a part of cytogenetics analysis and is used to examine whether the chromosome structure shows specific abnormalities. In addition, **Fluorescent In Situ Hybridization (FISH)** can reveal chromosomal aberrations sometimes even at the level of specific genes[50]. Molecular genetic analysis such as **Sanger or Next Generation Sequencing** as well as **fragment analysis** completes the diagnosis and reveals the presence of mutations.

Pediatric patients diagnosed with AML are stratified into different **risk groups** in the current AML-Berlin-Frankfurt-Münster (BFM) study group therapy protocols**: Standard risk** (SR), **intermediate risk** (IR) **and high risk** (HR)[32,124]. Risk stratification is based on different prognostic factors including clinical characteristics of the patients, the presence of chromosomal and genetic aberrations which are associated with outcome and also treatment response[32,162] (**Table 2**). Since the drugs

used in the therapy of AML are very toxic and too toxic side effects have to be avoided while at the same time leukemic blasts have to be eradicated effectively, the complex treatment is adapted to the three risk groups[31,32] with their different clinical presentation, as shown in **Table 2**.

Table 2: Current AML-BFM risk stratification and treatment recommendations. From Reinhardt *et al.* (2019)[124]

Standard-risk group	
Definition	All patients with the following evidence: • inv(16)(p13.1q22) • t(16;16)(p13;q22) • t(8;21)(q22;q22) • t(1;11) (q21;q23) • Normal karyotype and *NPM1* mutation • Normal karyotype and *CEBPA* (double) mutation
Treatment	Patients with inv(16) or t(16;16) are treated without the intensive chemotherapy block HAM (HD-cytarabine (3g/m^2) and mitoxantrone). These patients receive the following intensive therapy elements: ADxE (cytarabine, liposomal daunorubicin and etoposide) or AIE (cytarabine, idarubicin and etoposide) (induction 1), AI (cytarabine and idarubicin) (course 2), hAM (HD-cytarabine (1g/m^2) and mitoxantrone)(course 3), and HAE (course 4). Patients in the standard-risk group with t(8;21), t(1;11), normal karyotype and *NPM1* mutation or *CEBPA* (double) mutation receive the following intensive therapy elements: ADxE or AIE (induction 1), HAM (induction 2), AI (course 3), hAM (course 4), and HAE (HD-cytarabine and etoposide) (course 5). Maintenance therapy is administered for a year for all patients of the standard-risk group.
Re-stratification	Patients with nonresponse after first induction or who are *FLT3* ITD or *FLT3* TKD mutation positive are re-stratified into the intermediate-risk group. Patients with nonresponse after second induction are re-stratified into the high-risk group. Patients within the high-risk group have an indication for allogenic stem cell transplantation after the 4th course of chemotherapy, if they achieve CR. If re-stratified into the high-risk group and refractory disease persists, a salvage treatment with Dx-FLA or (Ida)-FLA and FLA should be performed, if the clinical conditions of the patient allow. Experimental “off study” therapy may be considered prior to SCT as an attempt to reach CR.

Intermediate-risk group	
Definition	All patients with de-novo AML, who do not belong to the standard-risk group (favorable prognosis) or to the high-risk group (unfavorable prognosis).
Treatment	ADxE or AIE (induction 1), HAM (induction 2), AI (course 3), hAM (course 4), and HAE (course 5). Maintenance therapy is administered for a year. In *FLT3* ITD or *FLT3* TKD mutation positive AML a treatment with Sorafenib may be considered on a compassionate-use basis.
Re-Stratification	Patients with nonresponse after first or second induction or individual disease characteristics are be re-stratified into the high-risk group, receiving an allogenic stem cell transplantation after the 4th course of chemotherapy, dependent on CR achievement.
High-risk group	
Definition	All patients with the following genetic evidence: • abnormalities in chromosome 12p/ t(2;12) • monosomy 5/5q- • *WT1* mutation and *FLT3* ITD • monosomy 7 (not in combination with favorable / KMT2A aberrations) • t(4;11)(q21;q23); *KMT2A/AF4* • t(5;11)(q35.3;p15); *NUP98/NSD1* • t(6;11)(q27;q23)); *KMT2A/AF6* • t(10;11)(p12;q23); *KMT2A/AF10* • t(6;9)(p23;q34) • t(7;12)(q36;p13) • t(9;22)(q34;q11) • complex karyotype (three or more aberrations, including at least one structural aberration, without favorable genetics and without *KMT2A*-rearrangement.) • inv(3)(q21q26.2)/t(3;3)(q21;q26.2) • t(16;21)(p11;q22); *FUS/ERG* • inv(16)(p13.3q24.3) *CBFA2T3-GLIS2*
Treatment	ADxE or AIE (induction 1), HAM (induction 2), AI (course 3), and hAM (course 4). In 1st remission, an allogenic stem cell transplantation is indicated. In *FLT3* ITD or *FLT3* TKD mutation positive AML a treatment with Sorafenib may be considered on a compassionate-use basis.
Re-Stratification	If nonresponse after 2nd Induction occurs, patients may be treated—if the conditions of the patients allow—with Dx-FLA or (Ida)-FLA (course 3) and FLA (course 4), to achieve CR before transplant. Experimental “off study” therapy may be considered prior SCT to reach CR.

3.3 Genetics and clonal evolution of AML

To evaluate and understand the genetic basis of leukemia, positional cloning of chromosomal translocations associated with human leukemias led to the identification of fusion genes that are implicated in the disease pathogenesis[48,78]. With more than 100 fusions identified so far, more different fusion genes than leukemic phenotypes are currently distinguished. This strongly implies that some of the fusion genes must activate similar signal transduction and/or transcriptional activation pathways in the leukemic blasts[78].

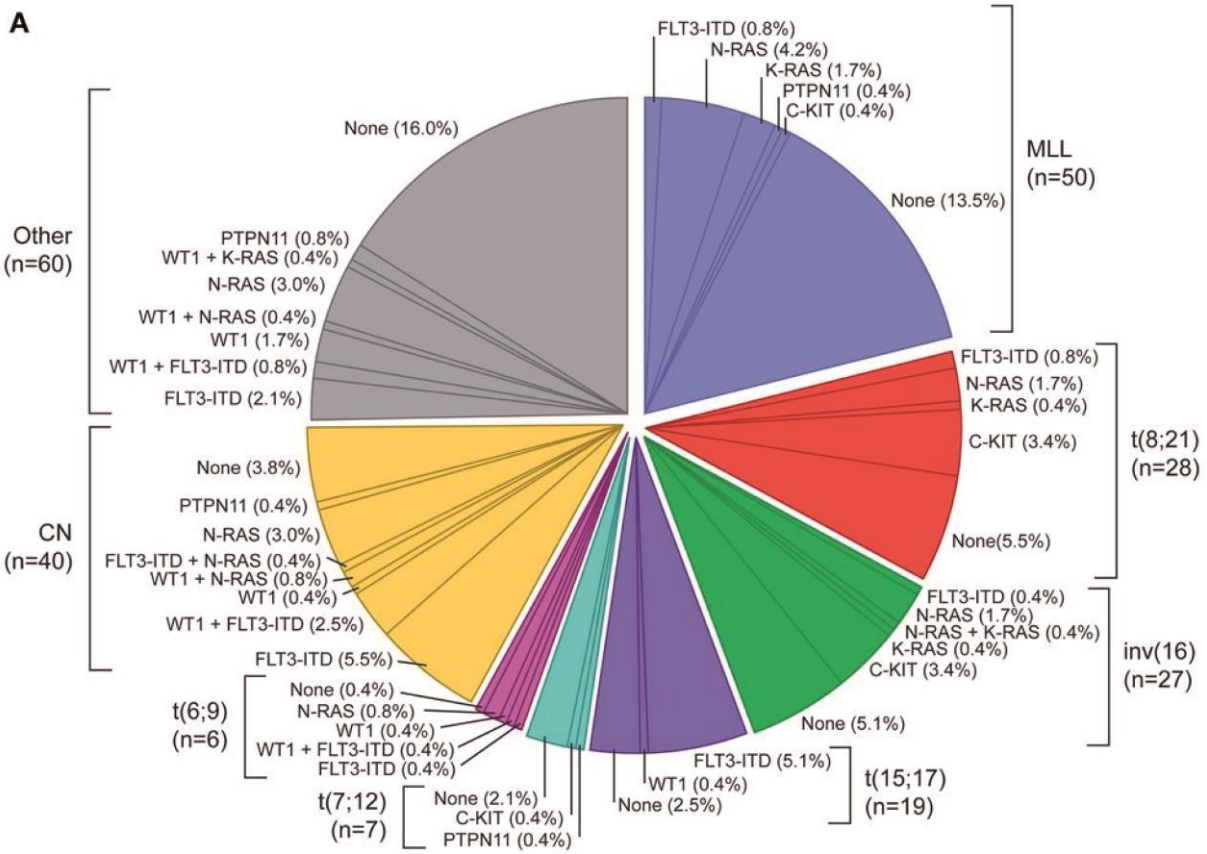


Figure *3:* Molecular and genetic aberrations as well as their nonrandom associations in pediatric AML. Integrative analysis of recurrent cytogenetic aberrations (*MLL*-rearrangements, t(8;21), inv(16), t(15;17), t(7;12), and t(6;9), normal karyotype (CN) and other nonspecific cytogenetic subgroups and molecular aberrations (in *FLT3*-ITD, *WT1*, *N-RAS*, *K-RAS*, *PTPN11*, and *c-KIT*) based on 237 unselected de novo pediatric AMLpatients. Each sector indicates the percentage of patients harboring 1 or more of the aforementioned mutations. From Creutzig *et al.* (2012)[32]

Apart from these gene fusions, genetic aberrations in the blasts of both adults and children include also somatic mutations[6,8,9,33,53,130] as summarized in **Figure 3** for pediatric AML. It could be shown, that neither loss of function mutations in hematopoietic transcription factors nor such gene fusions alone do result in malignant transformation but the cooperation between at least two broad classes of aberrations

is necessary for leukemogenesis[78]. This resulted in the postulation of a two-hit model of pathogenesis: **Class I mutations** or gene rearrangements lead to proliferation or survival advantage of the blasts, whereas **class II mutations** primarily lead to impaired hematopoietic differentiation[78]. Murine models demonstrated that cooperation of these two types of mutations can be sufficient for leukemogenesis[78]. Studies reporting that leukemia-associated class II aberrations can be detected in a proportion of healthy individuals [10,141] support this observation that a single genetic aberration alone does not appear to be sufficient for leukemogenesis[78]. However, it is hypothesized that these aberrations often are a predisposition and if another genetic event arises at approximately the same time, malignant transformation can occur[141]. Recent research also highlighted the presence of epigenetic modifications in AML due to altered DNA methylation or histone modifications as reviewed by Ungerstedt *et al*. (2018)[148].So far unknown mutations in genes related to epigenetic control of the genome and modifications to the epigenome itself neither belonging to class I nor class II mutations suggest that the two hit model has to be extended[136]. Additionally, recent studies suggest a temporal component to leukemogenesis; for leukemic transformation, it seems to be required that mutations occur at a specific point in cell development in a particular order[106,122,141].

Different studies also showed that the blast population in both adult and pediatric AML consists of several subclones characterized by the acquisition of additional mutations[41,44,98] Therefore, AML blasts obviously undergo clonal evolution with either the main clone or a genetic subclone of the primary disease changing the genetic phenotype and thereby surviving therapy, and then expanding in the patients clinically manifesting as relapse[41]. To understand this clonal evolution, it might be essential to identify driver mutations that occur early and/or late in leukemogenesis and persist or disappear during the course of the disease.

3.4 The main genetic aberrations in pediatric AML

The presence of certain genetic aberrations is an important part in current risk group stratification in pediatric AML treatment protocols[32] as shown in **Table 2**. **Standard risk** AML includes patients with core binding factor (CBF) mutations/dysregulation

such as inv(16), t(16;16), t(8;21)[162] and with mutations in *nucleophosmin1* gene (*NPM1*)[60,120] and *CCAAT/Enhancer Binding Protein Alpha gene* (*CEBPA*); all of these mutations are associated with a favorable prognosis[57,60,62]. Patients with mutations in the *Fms Related Tyrosine Kinase 3* (*FLT3*) gene are shown to have a poor prognosis[100,161]. However, if they occur together with *NPM*1 or *CBF* mutations, adult patients experience a better survival, while this effect could not be clearly shown in children[120]. Genetic aberrations defining the group of patients with **high risk** AML are monosomy 7 and monosomy 5[162]. Translocations involving *Histone-Lysine N-Methyltransferase 2A* (*KMT2A,* previously called *MLL)* occur in 15-20% of children with AML[118] and can be associated with poor outcome[8]. Interestingly, it is the translocation partner that predicts the clinical outcome[8]. *KMT2A* rearrangements with *Transcription Factor 7 Cofactor* (*MLLT11*) resulting in translocation t(1;11)(q21;q23) are associated with a good outcome whereas patients with translocations t(6;11)(q27;q23), t(10;11)(p11.2;q23) and t(10;11)(p12;q23) generally have a poor prognosis[8]. Other important poor prognostic factors in pediatric AML are loss of function mutations in *Wilms tumor gene 1* (*WT1*)[59], especially in combination with *FLT3* internal tandem duplications (ITD)[19]. Recent studies in both Europe and North America have additionally identified *NUP98/NSD1* fusions as poor prognostic factor occurring in 15% of cytogenetically normal pediatric AML associated with *FLT3* ITDs[61,111]. Patients who do not harbor any of the markers associated with either standard or high risk features are stratified into the **intermediate risk** group[124].

Further common genetic lesions are mutations in the RAS pathway occurring in 20% of pediatric AML patients with an association to *KMT2A* rearrangements, inv(16) and/or *NPM1* mutations[9]. In pediatric patients, the most common aberrations are *NRAS Proto-Oncogene* (*NRAS)* und *KRAS Proto-Oncogene* (*KRAS)* mutations which are not associated with a specific prognostic influence or treatment response[9].

3.5 Next Generation Sequencing (NGS)

Conventional Sanger sequencing was the gold standard for sequencing and the detection of mutations for the past decades[82,102]. This technology has recently been replaced by high-throughput approaches collectively described as Next Generation

Sequencing (NGS). Using different target enrichment strategies and clonal amplifications of DNA, it is now possible to sequence millions of DNA bases and strands in parallel, thereby dramatically reducing both time and costs of sequencing[34,102,126].

However, a reference genome was needed to asses the pathogenicity and relevance of genetic variants detected. The Human Genome project provided this in 2003, after 15 years of conducting Sanger sequencing to determine the sequence of

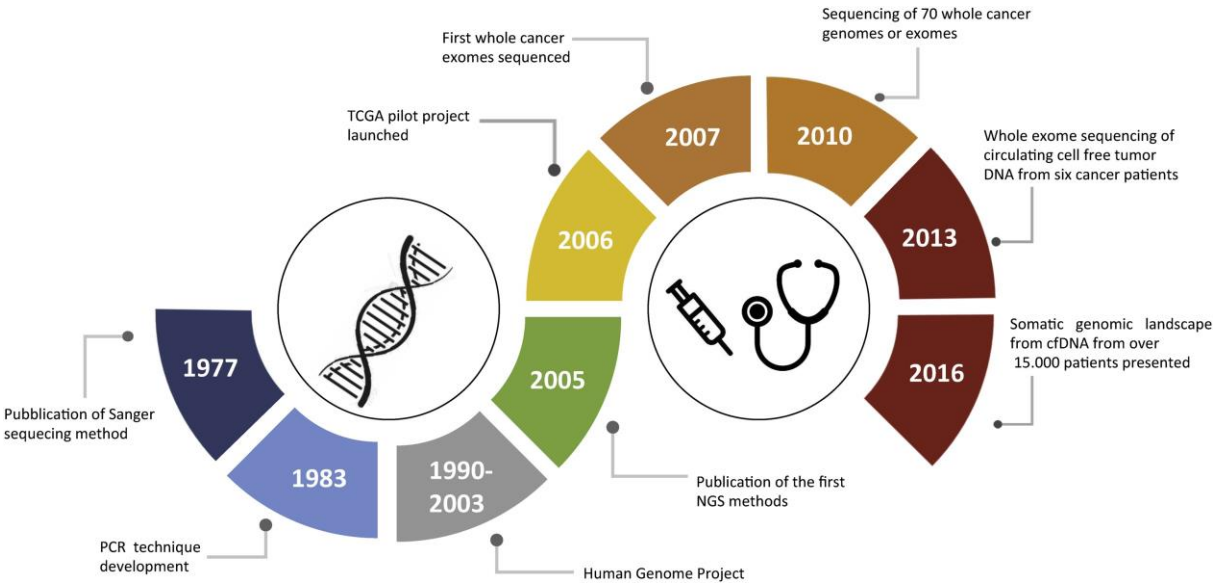


Figure 4: Timeline of major achievements in sequencing technologies. From Morganti *et al*. (2019)[104]

the human genome[52]. In parallel, several researchers identified new oncogenes and tumorsuppressor genes and could validate pathogenic alterations[104]. The evolution of sequencing was further promoted by large-scale cancer sequencing projects such as The Cancer Genome Atlas (TCGA)[104].

In 2005, Roche Diagnostics launched the 454 Life Sciences machine platform as the first NGS sequencing system[95], but other companies such as Illumina and Thermo Fisher Scientific introduced their own NGS platforms soon thereafter. The 454 pyrosequencing is based on adapter-ligated DNA fragments bound to small DNA-capture beads kept in a water-oil-emulsion[95]. Each bead is subjected to a small well with a mix of enzymes and fluorescence-labelled deoxyribonucleotide triphosphates (DNTPs) are added sequential while complementary nucleotides are embedded and a lightsignal is emitted due to the release of a pyrophosphate[95].

The Ion Torrent platform uses the ion semiconductor sequencing strategy and was launched by Thermo Fisher Scientific. Sequencing is based on the detection of

hydrogen ions that are released during the polymerization of DNA[4]. A microwell containing a template DNA strand to be sequenced is flooded with a single species of dNTP[4]. If the introduced dNTP is complementary to the leading template nucleotide, it is incorporated into the growing complementary strand[4]. This causes the release of a hydrogen ion that triggers an ion-sensitive field-effect transistor (ISFET) sensor[4].

Illumina launched several benchtop sequencer which all use a sequencing strategy that is based on reversible termination (RT) following bridge PCR on a flow cell. A mixture of RT nucleotides is added to the flow cell and a DNA polymerase incorporates them into the DNA strand[14]. Three steps are repeated in each sequencing cycle: The complementary RT nucleotide gets incorporated, the emitted fluorescence signal specific for each base gets detected, cleavage of the terminating and reporter molecule to restore a free 3′OH group for the next sequencing cycle[14].

With the several different NGS platforms available on the market, it has become readily possible to sequence all coding exomes of a genome (Whole Exome Sequencing) or even a complete genome (Whole Genome Sequencing) at a less expensive price and in a relatively short time. Analysis of these huge amount of NGS data is an indispensable need which usually consists of several steps: base calling, read alignment, variant identification and variant annotation[104]. After performing these analysis, a dataset of detected variants is available and can then be used for research or diagnostic approaches[34,90].

Recently, Bolouri *et al.* analyzed the molecular landscape of pediatric AML in a large cohort treated within Children's Oncology Group (COG) AML trials[19]. They extensively characterized the mutational, transcriptional and epigenetic landscape of a large well-described cohort of children, adolescent and young adult AML patients using a combination of different NGS approaches such as whole genome sequencing (WGS), targeted capture-based NGS, mRNA and microRNA (miRNA) sequencing and also methylation studies[19]. Comparison of different age groups revealed, that structural variants including gene fusions were more prevalent in young individuals whereas mutations in common genes in adult patients including *DNA methyl transferase 3 Alpha* (*DNMT3A*) and *tumor protein 53* gene (*TP53*) were extremely rare in pediatric patients[19]. So far, undescribed mutations in *GATA binding protein 2* (*GATA2*), *FLT3* and *Cbl Proto-Oncogene* (*CBL*) and recurrent mutations in the *MYC protooncogene* (*MYC*), *NRAS*, *KRAS* and *WT1* could be detected at high frequencies in pediatric AML patients[19]. Pathways such as Wnt signaling, Polycomb repression, innate immune cell interactions and a cluster of zinc finger encoding genes which are

associated with *KMT2A* rearrangements were impacted by different genetic aberrations as well as by epigenetic events such as promotor DNA hypermethylation[19]. The authors showed that the mutational burden in AML patients increases with age and that gene fusions and focal copy number aberrations were more common in younger subjects[19]. Additionally, it was described that the number and location of mutations in genes such as *FLT3* and *WT1*, which are also recurrently mutated in adults, is different in pediatric patients. Clonal *WT1* mutations occurred more often in children and with a higher variability[19]. This study demonstrates the wide range of applications of NGS technologies to extensively study the genetic and epigenetic diversity in pediatric AML. Nevertheless, further work is needed to decipher the genetic complexity of AML.

3.6 Minimal Residual Disease in AML

The identification and description of disease-specific markers is an initial step towards the encryption of leukemogenesis. However, monitoring the response to treatment in the patient over time is another important approach to assess the success of the different therapeutic regimens and to understand the biology of the disease on the other hand. Minimal Residual Disease (MRD) was initially described as the detection of leukemic blasts below the morphologic limit of detection (5%), during and after therapy[96]. Current methods for MRD detection are PCR and flow cytometry-based methods[30,45,54,55,86,101,150]. Especially in adult AML, MRD monitoring has become a widely used and accepted prognostic indicator important for risk stratification and treatment planning[76,133]. With respect to MRD-directed therapy in pediatric AML however, robust clinical data are still lacking[112]. For example, multiparameter flow cytometry, which is a very quick and relatively cheap method to determine MRD using the aberrant expression of antigens on the blast cells, is more straightforward in ALL than in AML[112]. This is due to the fact, that AML patients have a higher probability to present with multiple clones and every clone could contribute to MRD and thus requires monitoring[55,86].

Another widely used technique to monitor MRD in addition to qualitative PCR is the quantitative Real Time PCR (qPCR). This is a powerful tool using primer and probes with which the presence and amount of mutated genes or fusion genes can specifically be targeted[11,45,154]. For adult leukemia, an international group from ten countries already collaborated and established standardized protocols for qPCR analysis of fusion genes that are used to determine the MRD level[11,45]. Presently this kind of study has not yet been performed for pediatric AML patients.

Supplementing or maybe even replacing the current MRD monitoring platforms is another direction for NGS and several studies already suggested its applicability[38,113,147,152]. Ultra deep sequencing with close to 1 000 000 reads per marker per patient may enable sensitive and reliable MRD monitoring for mutations initially identified by NGS in patients lacking markers used by conventional methods such as qPCR[147].

4 Aims of the thesis

I established various NGS approaches for the Illumina platform in my thesis to address two important topics in current pediatric AML research.

In the first part of this thesis, an approach that allows to genetically characterize AML blasts and specifically AML LSCs was established. Previous studies used somatic changes as markers for clonality and have revealed that AML is a complex mosaic of different cell populations[35,89,97,98]. Therefore, we aimed to work on a workflow using a commercial WGA procedure to analyze the heterogeneity of pediatric AML even on a single-cell level, allowing the analysis of clonality and clonal evolution of this disease.

In the second part of this thesis, we retrospectively analyzed a cohort of pediatric AML patients treated on current protocols of the AML-BFM study group in Germany. We applied NGS and quantitative Real-Time PCR (qPCR) to identify mutations and gene fusions to analyze the correlation between the presence or absence of genetic aberrations and survival and treatment response. We thereby focused on and confirmed the data about the impact of *WT1* and *FLT3* mutational status on patient outcome using targeted amplicon-based NGS (TS) and further analyzed patient cells for the presence of *NUP98-NSD1* fusiontranscripts.

5 Results

5.1 Publications

Single-cell whole exome and targeted sequencing in NPM1/FLT3 positive pediatric acute myeloid leukemia

Christiane Walter, Christian Pozzorini, Katarina Reinhardt, Robert Geffers, Zhenyu Xu, Dirk Reinhardt, Nils von Neuhoff*, Helmut Hanenberg*

* Both authors contributed equally

Pediatric Blood & Cancer. 2017;e26848.

https://doi.org/10.1002/

Mutated WT1, FLT3-ITD, and NUP98-NSD1 Fusion in Various Combinations Define a Poor Prognostic Group in Pediatric Acute Myeloid Leukemia

Naghmeh Niktoreh, Christiane Walter, Martin Zimmermann, Christine von Neuhoff, Nils von Neuhoff, Mareike Rasche, Katharina Waack, Ursula Creutzig, Helmut Hanenberg*, and Dirk Reinhardt*

*Both authors contributed equally

Hindawi; Journal of Oncology; Volume 2019, Article ID 1609128

https://doi.org/10.1155/2019/1609128

The treatment outcome of AML in children and adolescents has drastically improved over the last decades, however the overall survival is still only around 70%[32,119,162]. Understanding the complex mechanisms of leukemogenesis and identifying biomarkers predicting initial treatment response and relapse will lead to more precise definition of specific risk groups, which is a prerequisite for more specifically adapting the treatment according to the individual risk and thereby possibly increasing the overall survival further. In this thesis, a workflow for reliable genetic analysis of single AML blasts from pediatric AML patients with known mutations in *NPM1* and/or *FLT3* ITD as genetic markers was established. To this end, DNA from bone marrow of pediatric patients with at least 80% blasts at initial diagnosis was isolated and then underwent both targeted amplicon-based sequencing (TS) as well as whole exome sequencing. Additionally, genomic DNA was extracted from single AML blasts sorted for a leukemic stem cell (LSC)-enriched immunophenotype and then subjected to whole genome amplification (WGA). The amplified genomic DNA was analyzed by whole exome sequencing (WES) and the findings compared to the results achieved on the bulk material. Importantly, this work demonstrated that the quality of WGA was sufficient to facilitate WES of single cells and that coverage statistics were comparable between single cell and bulk WES data.

In the second study we tested the hypothesis, whether mutations in *WT1* and *FLT3* that co-occur in AML blasts from children and adolescents at diagnosis are still predictors of poor outcome in contemporary treatment protocols. 353 patients (<18 years) treated in Germany in the BFM-AML studies since 2004 were analyzed. *WT1* mutational status was confirmed via targeted amplicon-based NGS using the TruSight myeloid panel (Illumina) and the presence of *FLT3* ITDs was additionally confirmed in 52 patients by fragment length analysis. The presence of *NUP98-NSD1* fusion, another genetic factor that is frequently associated with *FLT3* ITD, was determined retrospectively in our cohort from frozen material obtained at diagnosis. Comparing the clinical data/treatment outcome of different genetically defined risk groups demonstrated that co-occurrence of *WT1* mutation, *NUP98-NSD1* fusion and *FLT3* ITD as triple or double factors still predicts dismal response to contemporary first- and second-line treatment for pediatric AML.

5.2 Single-cell whole exome and targeted sequencing in NPM1/FLT3 positive pediatric acute myeloid leukemia

Contribution to present publication:

- Conception: 40%
- Experimental work: 80%
    - preparation of cells for sorting
    - sorting of cells
    - Establishment of WGA from sorted cell lines and patient cells
    - WGA and subsequent quality control
    - extraction of DNA from bulk cells
    - preparation of DNA libraries for targeted NGS
- Data analysis: 60%
- Writing the manuscript: 60%
- Revising the manuscript: 60%

The above listed contributions of Christiane Walter to the publication are correct.

Essen, den _____ _____

Unterschrift der Doktorandin

Essen, den _____ _____

Unterschrift des wissenschftl. Betreuers/Mitglieds der Universität Duisburg-Essen

Single-cell whole exome and targeted sequencing in NPM1/FLT3 positive pediatric acute myeloid leukemia

Christiane Walter[1], Christian Pozzorini[2], Katarina Reinhardt[1], Robert Geffers[3], Zhenyu Xu[2], Dirk Reinhardt[1,] Nils von Neuhoff[1,5,] *, and Helmut Hanenberg[1,4,5,] *

[1] Pediatrics III, University Children's Hospital Essen, University of Duisburg-Essen, 45147 Essen, Germany.

[2] Sophia Genetics Inc., 1025 Saint-Sulpice, Switzerland

[3] Genome Analytics, Helmholtz Centre for Infection Research, 38124 Braunschweig, Germany

[4] Otorhinolaryngology and Head/Neck Surgery (ENT), Heinrich Heine University, 40225 Düsseldorf, Germany

[5] Both authors contributed equally to this work

***Correspondence:**

Helmut Hanenberg, MD,

University Children's Hospital Essen, Virchowstr. 171, 45147 Essen, Germany

Helmut.Hanenberg@uk-essen.de,

Tel. +49-201-723-85185

Nils von Neuhoff, PhD,

University Children's Hospital Essen, Virchowstr. 171, 45147 Essen, Germany

Nils.vonNeuhoff@uk-essen.de

Tel. +49-201-723-1052

Keywords: pediatric AML; single-cell analysis, whole genome amplification, whole exome sequencing, targeted sequencing

Short title: Single cell sequencing in pediatric AML

Word count: abstract: 196; main text: 3470

Figure and table count: 1 table; 5 figures, 7 supplemental figures, 9 supplemental tables

abbreviation	full term
LSC	Leucemic stem cell
AML	Acute myeloid leukemia
NPM1	Nucleophosmin
FLT3	fms like tyrosine kinase 3
WES	whole exome sequencing
InDel	Insertion or Deletion
WGA	whole genome amplification
MDA	multiple displacement amplification
TS	targeted sequencing
AML-BFM	name of the german pediatric AML study group
BM	bone marrow
FAB classification	French–American–British (FAB) classification system for hematologic diseases
n.a.	not available
HR	high risk
SR	standard risk
WBC	white blood cell count
Hb	hemoglobin
PI	propidium iodide
SNV	single nucleotide variant
TSM panel	TruSight myeloid Panel

Abstract

Background

The small portion of leukemic stem cells (LSCs) in acute myeloid leukemia (AML) present in children and adolescents is often masked by the high background of AML blasts and normal hematopoietic cells. The aim of the current study was to establish a simple workflow for reliable genetic analysis of single LSC-enriched blasts from pediatric patients.

Procedure

For three AMLs with mutations in NPM1 and/or FLT3, we performed whole genome amplification on sorted single-cell DNA followed by whole exome sequencing (WES). The corresponding bulk bone marrow DNAs were also analysed by WES and additionally by targeted sequencing that included 54 genes associated with myeloid malignancies.

Results

Analysis revealed that read coverage statistics were comparable between single-cell and bulk WES data, indicating high-quality WGA. From 102 single-cell variants, 72 SNVs and InDels (70%) were consistently found in the two bulk DNA analyses. Variants reliably detected in single cells were also present in targeted sequencing. However, WES with read counts between 50-72X failed to detect rare AML subclones in the bulk DNAs.

Conclusions

In summary, our study demonstrated that single-cell WES combined with bulk DNA targeted sequencing is a promising tool for detecting AML subclones and possibly LSCs.

Introduction

Childhood acute myeloid leukemia (AML) is a rare hematological disorder caused by uncontrolled proliferation and impaired differentiation of malignant myeloid precursor cells (1, 2). Although the event-free and overall survival of children and

adolescents with AML has steadily increased over the last three decades (3-5) and small molecularly and cytogenetically defined subgroups with improved outcome exist (6, 7), overall survival after the very intense therapy is still only about 70% (1).

In order to improve the prognosis of pediatric AML, numerous approaches to understand the complex interplay of molecular and cellular factors regulating the development and growth of different AML subtypes have been pursued (8-12). Major findings include intrinsic or acquired changes in the microenvironment of the bone marrow niche (13, 14) (reviewed in (15)) as well as intrinsic factors in the AML blasts, such as epigenetic changes, predisposing germ-line drivers or somatic mutations, which all lead to altered differentiation, survival, dormancy and regenerative capacities (reviewed in (15, 16)). Contributing factors in the manifestation of AML may also comprise deficiencies in innate or acquired immune responses (reviewed in (17)), as AML blasts are known to express antigen(s) on their surface that can trigger immune responses (18) and can be recognized by donor lymphocyte infusions after allogeneic stem cell transplantation (19, 20).

More than 20 years ago, studies already suggested that similar to normal hematopoiesis, AML may be organized in a hierarchical manner (reviewed in (15)). Using CD34, CD38 and lineage specific markers, CD34+CD38- leukemic blasts were identified in AML patients that possessed long-term culture initiating activity *in vitro* (21, 22) and were capable of engraftment in immunodeficient mice (23-25). Later studies demonstrated that these leukemic stem cells (LSCs) can also reside in the CD34+/CD38+ fraction or even within CD34- cells (26). In some patients, more than one LSC co-exists within the CD34+ cell fraction (26). Importantly, the fact that LSCs persist in the bone marrow as a small population of cells capable of recapitulating the disease (27-29) implies that long-term cure can only be achieved by eliminating these primitive cells, which can be more resistant to common chemotherapy agents (30).

During the process of malignant transformation, LSCs undergo continuous genetic and epigenetic changes as part of their clonal evolution (31). Extensive studies using acquired changes as markers for clonality have indeed revealed that AML in individual patients is a complex mosaic of oligo- to polyclonal cell populations (32, 33). Chemotherapy exposure of AML LSCs and blasts leads to additional evolutionary stress that induces further alterations in the AML cells as adaptation to the treatment. This forced evolution can ultimately result in altered signatures of the LSCs and blasts at relapse (34). As LSCs with different profiles are present at very low percentages

(27-29), it remains a major technical challenge to genetically characterize few LSC cells against a high background of diverse AML blasts and also normal hematopoietic cells. Here, enrichment procedures are required to specifically analyse AML subpopulations and ideally single cells enriched for LSCs.

Next generation sequencing (NGS) has emerged as a technology that provides high-quality genetic data, especially in case of heterogeneous cell populations (11, 32, 35). However, the limited amount of DNA obtainable from single cells currently constitutes a major technical challenge, as the characterization of single cells with NGS technologies requires reliable and unbiased methods for single-cell whole genome amplification (WGA). Several WGA techniques are available with the most important ones being based on PCR and/or multiple displacement amplification (MDA) methods (36, 37).

Here, we established a procedure for WGA of DNA from single CD34+/CD38- cells, which is suitable for genomic characterization of these single cells via whole exome sequencing (WES). In order to validate this workflow, we performed WGA on 24 single cells from three pediatric AML patients and showed that unique markers of the patient blasts were still present in 16 out of 24 cells. NGS analysis of 5 single malignant cells demonstrated the high quality of the WGA procedures as the majority of variants identified in single CD34+CD38- blasts were consistently detected in the bone marrow bulk populations by both WES and targeted sequencing (TS).

Material and Methods

Patients and samples

Specimen were collected within the AML-BFM Trials and used with informed consent of the patients and legal guardians, approved by the Ethics Committees. Bone marrow (BM) samples were collected at diagnosis and mononuclear cells were stored in liquid nitrogen. The characteristic profile of the blasts, including their *NPM1* and *FLT3* status (**Table 1**), was assessed as previously described (38). The OCI-AML3 cell line was purchased from the DSMZ (Braunschweig).

Flow cytometry and cell sorting

Cryopreserved BM cells from each patient were thawed and stained with propidium iodide (PI), CD34-PE-Cy7 and CD38-FITC antibodies (Beckman Coulter). Single CD34+CD38-PI- cells were sorted on the MoFloXDP cell sorter (Beckman Coulter) into single wells of a 96-well plate for microscopic control. Wells with a single cell were subsequently transferred into 0.5ml low binding tubes (Biozym). Sorted PI- single cells from the OCI-AML3 cell line cells were processed as described above for initial establishment of the workflow.

Whole genome amplification (WGA) and quality control

Sorted single-cell DNAs were amplified using the MDA-based *REPLI-g* single-cell kit (Qiagen). The amplified DNA was purified using the Gel and PCR clean-up kit (Macherey-Nagel). To control the amplification quality, fragments of *GATA1*, *Albumin* and *GAPDH* located on different chromosomes were detected by Taqman qPCR (**Supplemental Table S2**).

Targeted Sequencing with the TruSight Myeloid panel

For targeted sequencing with the TruSight Myeloid (TSM) panel (Illumina, **Supplemental Table S3**), 50ng of bulk DNA derived from each of the three patients was used. The libraries were prepared according to the manufacturer`s protocol and sequenced on Illumina MiSeqDx using the MiSeq V2 Reagent kit (2x150 cycles, paired end run) with a total of 141kb per patient.

Whole exome sequencing (WES)

WES was performed on the amplified DNA derived from sorted single cells as well as from the bulk bone marrow cells from the three patients. Fragmentation of 100ng DNA was performed on a Covaris S2 and the size of the fragments was controlled with Agilent Technologies 2100 Bioanalyzer.

The DNA sequencing library was generated from 100ng of fragmented bulk and single-cell DNAs using Agilent SureSelectXT Reagent Kits v5_3´ UTR (75Mb)

according to the manufacturer´s protocols. The libraries were sequenced on Illumina HiSeq2500 or HiSeq2000, respectively, using TruSeq SBS Kit v3-HS (200 cycles, paired end run) with an average of 12.5 x10^6 reads per single exome.

WES and TSM data analysis

Raw reads (fastq format) from the WES samples (n=8) and the TSM targeted sequencing (n=3) were aligned to build 37 of the human genome (hg19) using the BWA MEM (http://bio-bwa.sourceforge.net/) or Bowtie2 (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) algorithms with default settings. Single-nucleotide variants (SNV) and Indels were called and annotated using a commercial software (Sophia DDM, Sophia Genetics, Switzerland). Variant calling was performed on each sample independently. For WES, genomic regions covered by <5 reads were not considered. For targeted sequencing, genomic regions that were only covered by <30 reads were not included. Variants with variant fraction <4% were filtered out. Variant analysis was only performed on genomic regions targeted by both the WES and TSM approaches.

Results

Patient selection

For WGA of single CD34+CD38- blasts, we selected three patients with mutations that can be readily detected by PCR followed by fragment analysis of the PCR amplification products (39-43). As shown in **Table 1** and **Supplemental Figure S1**, AML blasts from **patient 1** had a somatic insertion of 4bp pairs, c.860_863dupTCTG, in the *NPM1* gene and an internal tandem duplication (ITD), c.1812_1813ins36, in the *FLT3* gene. In contrast, the AML blasts from **patient 2** had wild-type *NPM1* sequences but were positive for the FLT3 ITD, c.1734_1837+2dup. The blasts of **patient 3** had the same *NPM1* mutation as patient 1 but had *FLT3* wild-type alleles. All three patients had elevated leukocyte counts in the peripheral blood with typical antigen expression profiles of myeloid blasts.

Quality control of the WGA DNAs and sequencing

The workflow was first established using cultivated CD34+CD38-PI- OCI-AML3 cells. To quickly check whether error-free amplification of defined areas of the genome had occurred, we analyzed fragments of *albumin*, *GATA1* and *GAPDH* located on different chromosomes by Taqman qPCR. With this readout, we achieved correct WGA in 14 out of 23 (61%) sorted single cells (data not shown).

Next, we thawed cryopreserved cells of the three patients from diagnosis and sorted single CD34+CD38-PI- cells followed by WGA. From a total of 32 single cells, correct amplification of *albumin*, *GATA1* and *GAPDH* was achieved in 18 (56%) single cell DNAs (**Supplemental Table S1**). Amplified DNAs from two WGA single AML cells from patients #1 and #2 and one cell from patient #3 (**Supplemental Table S4**) was then subjected to subsequent WES analysis. WES was also performed on the bulk DNA isolated from the mononuclear BM cells from each patient. As an independent high-quality sequencing tool, we analysed the bulk DNA from the BM cells with the TSM panel that specifically determines the presence of genetic variants in 54 genes typically associated with myeloid malignancies and reaches much higher coverage compared to WES.

Coverage of the genetic information in bulk versus single-cell DNA

We initially compared the coverage between the three sequencing approaches, WES of WGA single cell DNA, WES and targeted sequencing of the bulk DNA. The total number of aligned reads varied between 51,224,023 - 68,229,043 (~98% of all reads) for bulk DNA (n=3) and 32,826,454 - 83,090,381 (~98% of all reads) for single-cell DNA (n=5) (**Supplemental Table S5**). For the three bulk DNAs and three of the five WGA single-cell DNAs, >95% of the target regions was covered with at least 10 reads and 50% of the regions was covered more than 60 times (**Figure 2A**). In contrast, the WGA DNAs from two single cells, 73-KM from patient 1 and 92-UE from patient 2, had a median coverage of 0 (**Supplemental Table S6**) due to the facts that 90% of the target regions were not covered and that the fraction of the target regions covered with at least 50X was <10% (**Figure 2A**). Excluding the latter two single-cell DNAs, further analysis revealed that the median coverage per exome varied between 42 and 63 (mean 50, n=3) in the bulk DNAs and between 23 and 78 (mean 55.6, n=3) in the single-cell DNAs (**Figure 2B**).

The total number of reads that mapped to the target region of the TSM Panel (n=3) ranged between 1,551,921 and 2,291,044 (97-99% of all reads) and the minimum coverage observed on 90% of the target region was comprised between 806x and 1346x (data not shown). As shown graphically in **Figure 3**, the coverage for the bulk DNAs was comparable at each genomic position targeted by the TSM panel and the WES approach. In contrast, we observed a slightly less uniform coverage in the amplified single-cell DNAs, as more regions were present in the amplified single cell DNAs that were either over- or underrepresented when using 5-times or one fifth of the median coverage as cut-off.

When we restricted the comparison to the regions targeted by the TSM panel, the uniformity of the coverage in the bulk and the single-cell DNAs WES was comparable and the median coverage even higher for the single-cell DNAs, 44 vs. 78, as shown for patient 1 in **Figure 4**.

Somatic variants detected by WES or targeted TSM sequencing

In order to determine the quality of the single-cell amplifications, we analysed the somatic variants (SNV, InDels) detected by all three sequencing approaches. This analysis was possible in 50 out of 54 genes that were sufficiently covered in the analyses. As shown in **Figure 5a**, the majority of variants detected in the WES bulk data was found at comparable frequencies in the TSM panel data. The variant frequencies in the single-cell data from WES were more variable and more variants from the single-cell DNAs were not found in the TSM panel data (**Figure 5b**).

With all approaches, a total of 170 somatic variants were detected in the three bulk and the three single-cell DNA samples. 53 variants only detected by TSM panel were visually inspected using IGV and 49 were finally classified as false positives and therefore excluded from further analysis. The origin of the remaining 121 variants was visualized in **Figure 5c**. Importantly, none of the sequencing approaches was able to detect all 121 variants. However, as many as 79 out of 121 variants (65.3%) were present in all three data sets. WES readily identified 103 (85.1%) and 93 (76.9%) of the 121 variants in the bulk and single-cell DNAs, respectively. Only four (3.3%) variants were not detected by WES.

Twelve (9.9%) variants were exclusively found in the single-cell DNAs. Eleven of these were classified as artefacts generated during the DNA amplification process, as the variant frequencies were <25% and the variants could not be called in the TSM panel data (**Supplemental Table S7**), despite a high read coverage in these areas. A synonymous variant in the *CUX1* gene (c.603G>A; p.Leu201Leu) was detected at a variant frequency of 60% in the single cell from patient #2 (**Supplemental Table S7**).

Out of 28 variants (23.1%) detected in WES bulk data and/or TSM panel data but not in the single-cell WES data (**Figure 6**), 23 were missed in the single-cell data because of insufficient or absent coverage at the corresponding regions. Four variants were present in the single-cell data but not called by our pipeline due to too low read frequencies (**Supplemental Table S8**). Despite sufficient coverage, one missense variant in the *TET2* gene (c.5285A>G, p.Ile1762Val) was not present at all in the single-cell WES data (**Supplemental Table S8**), probably as the single cell used for DNA amplification did not harbor this variant.

We identified two variants (1.7%) in the single-cell WES and the TSM panel data, which were not detected in the bulk WES data (**Figure 6**). In depth analysis revealed that one variant was not detected due to insufficient coverage. The nonsense variant in the *KDM6A* gene (c.2581G>T; p.Glu861) present in the single cell of patient #2 with a variant frequency of 90% was found in the TSM panel data, albeit with a variant frequency of only 4.6%.

Retracing the FLT3 ITD and NPM1 mutations in the single-cell DNA WES

Finally, we checked whether the *NPM1* and *FLT3* alleles known for the three patients as characteristics of their AML blasts were also present in the data obtained from the different sequencing approaches. As shown in **Supplemental Table S9**, the TSM panel was an excellent tool for identifying the mutations and wild-type sequences in *NPM1* and *FLT3* in the bulk DNAs of all patients.

The WES approach was equally suited to detect the wild-type *FLT3* sequences and the FLT3 ITD c.1812_1813ins36 (**Supplemental Table S8**). However, the FLT3 ITD c.1734_1737+2dup from patient #2 was only detected by WES analysis with our improved algorithm. The NPM1 mutation in patient #1 and #3 was found in approximately 50 and 25% of reads by WES, respectively.

Discussion

Seminal work in adults demonstrated that the myeloid disorders MDS and AML exhibit clonal heterogeneity that evolves during disease progression and can be altered at relapse because of treatment (44). The initial publications have used NGS of bulk samples (32, 33, 45, 46), but recent studies also described the functional heterogeneity among AML subpopulations by combined single-cell and bulk analysis (47). In pediatric AML/MDS, we are often severely limited by the amount of material that is available for the diagnostic and experimental procedures. The aim of this work therefore was to evaluate whether a commercial WGA procedure allows to analyze the heterogeneity of AML in cryopreserved material on single-cell level.

For the study, we chose material from three pediatric patients at diagnosis with defined mutations in *NPM1* and/or *FLT3*. Five out of 18 successfully amplified single CD34+CD38- cell DNAs underwent WES and were compared to WES and also targeted sequencing with an amplicon-based approach of the bulk DNAs from the patients. Our results demonstrated that the quality of DNA amplified using the MDA technology is sufficient for WES of single cells, as also shown by others (48, 49). Importantly, the alignment of reads derived from the single-cell DNA was robust and highly specific, since consistently >98% of reads were unequivocally mapped to the corresponding sites in the reference genome. The median coverage between single-cell and bulk DNA was comparable. In addition, the coverage uniformity was slightly lower in DNA amplified from single cells, which is concordant with findings described previously (48, 49).

In order to access the error rate in variant detection introduced by the WGA-MDA technology, the called variants from the single-cell WES data were compared to those identified the bulk DNA WES data. As WES generally produces limited coverage, we also analyzed the bulk DNA with the TSM panel that readily achieves coverage of 1000X or more. For a meaningful comparison, we restricted the variant analysis to common target regions covered by all three approaches. Not surprisingly, the coverage in these target regions was slightly lower in single cells, but still comparable to bulk DNA.

From 170 somatic variants detected, 49 were artifacts from the PCR-based DNA amplification in the TSM panel. Importantly, 79 of the validated 121 variants (65.5%) were present in all three approaches and thus classified as real variants reflecting the genetic profile of three CD34+CD38- single cells derived from the patients. Forty-two variants were detected only in one or two of the data sets. Out of the twelve variants exclusively found in the single-cell WES data, eleven were classified as artifacts as their variant frequencies were below 20% and as they were not retraceable in the targeted sequencing analysis despite the high coverage achieved. One synonymous variant detected in the *CUX1* gene in the single-cell WES derived from patient #2 had a variant allele frequency of 60% and was classified as a heterozygous SNP from a small CD34+CD38-subclone, as it was present in the TSM data set but not the bulk WES.

We found 28 variants in the bulk DNAs from WES and the TSM panel, that were absent in the single-cell WES data; 23 of them were not called in the single-cell data due to insufficient coverage at these positions. A *TET2* missense variant found in the bulk sample (39% in WES and 53% variant allele frequency in targeted sequencing) was not be detected in the single-cell data, despite sufficient coverage at this position as the single cell simply did not harbor this variant. Three SNVs detected by the TSM panel in the bulk DNAs were not called by our pipeline due to insufficient coverage in WES single-cell and bulk DNA data. Similarly, a 107bp long insertion detected by the TSM panel in the bulk DNA was not called in WES bulk and single-cell data. Nevertheless, visual inspection by IGV suggested the presence of this variant in WES single-cell DNA, but not in WES bulk DNA.

Only two variants that were found in the single-cell and the TSM data sets were missing in the bulk WES data. One was not called in the bulk WES due to low coverage at this position. The other variant was a nonsense mutation detected in the *KDM6A* gene in the single-cell and the targeted sequencing data from patient #2. This variant was identified with a frequency of 4.6% in the TSM panel and therefore only present in a small subpopulation for which the WES approach on the bulk DNA did not have sufficient coverage to detect.

Our study clearly demonstrated that WGA followed by WES can reliably detect at least two third of the somatic changes that are present in single CD34+CD38- AML cells. For single cell analysis, the usual coverage of WES (about 50X) is sufficient as

variants will be present either in heterozygous or homozygous state in a largely diploid genome (except for areas that were amplified or deleted). WES of bulk DNA is helpful for screening approaches but does not add much more information when compared to WES of amplified single cells. WES however is not ideal to analyze small subpopulations of clones, simply due to the low coverage associated with this approach. Here, targeted sequencing e.g. by the TSM panel is clearly more powerful and also uses less resources. However, the main limitation for targeted sequencing is that the genes of interest have to be defined already prior to the sequencing. Targeted sequencing of bulk DNA nevertheless is a powerful and also economic research tool when used for confirmation of WES data obtained from amplified single-cell DNA, as even small subpopulation can readily be detected due to the high read counts. Here, the TSM panel as PCR-based amplicon technology will be replaced by capture-based assays. For our pediatric AML patients, WES of WGA single-cell DNA in combination with targeted bulk DNA analysis will allow to cost-effectively and reliably perform clonal analysis at diagnosis, after passage in immunodeficient mice and/or at relapse after therapy.

Conflict of interest Statement

CP and ZX are employed by Sophia Genetics.

Acknowledgments

We would like to thank the patients and their families for supporting our research efforts. We kindly acknowledge the excellent technical expertise of Chantal Daszkowski, Ellen Mahlow, Denise Kondryn, Christel Katerkamp, Andrea Drothler, Lisa Kellerstrass, Nicole Schuster and Carolin Augsburg.

Author contributions

The study was designed by CW, CP, ZX, DR, NvN and HH. All experiments and/or analyses were performed by CW, CP, KR, RG, ZX, DR, NvN and HH. The manuscript was written by CW, CP, ZX, NvN and HH with essential help from the other authors.

References

1. Creutzig U, van den Heuvel-Eibrink MM, Gibson B, Dworzak MN, Adachi S, de Bont E, 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.

2. Zwaan CM, Kolb EA, Reinhardt D, Abrahamsson J, Adachi S, Aplenc R, et al. Collaborative Efforts Driving Progress in Pediatric Acute Myeloid Leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(27):2949-62.

3. Creutzig U, Zimmermann M, Dworzak MN, Ritter J, Schellong G, Reinhardt D. Development of a curative treatment within the AML-BFM studies. Klin Padiatr. 2013;225 Suppl 1:S79-86.

4. Creutzig U, Reinhardt D, Zimmermann M, Group A-BS. Prognostic relevance of risk groups in the pediatric AML-BFM trials 93 and 98. Ann Hematol. 2004;83 Suppl 1:S112-6.

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

6. Balgobind BV, Raimondi SC, Harbott J, Zimmermann M, Alonzo TA, Auvrignon A, et al. Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. Blood. 2009;114(12):2489-96.

7. Creutzig U, Reinhardt D, Diekamp S, Dworzak M, Stary J, Zimmermann M. AML patients with Down syndrome have a high cure rate with AML-BFM therapy with reduced dose intensity. Leukemia. 2005;19(8):1355-60.

8. Li S, Mason CE, Melnick A. Genetic and epigenetic heterogeneity in acute myeloid leukemia. Curr Opin Genet Dev. 2016;36:100-6.

9. Tzelepis K, Koike-Yusa H, De Braekeleer E, Li Y, Metzakopian E, Dovey OM, et al. A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. Cell Rep. 2016;17(4):1193-205.

10. Lindsley RC, Mar BG, Mazzola E, Grauman PV, Shareef S, Allen SL, et al. Acute myeloid leukemia ontogeny is defined by distinct somatic mutations. Blood. 2015;125(9):1367-76.

11. Bacher U, Schnittger S, Haferlach T. Molecular genetics in acute myeloid leukemia. Current opinion in oncology. 2010;22(6):646-55.

12. Schlenk RF, Dohner K, Krauter J, Frohling S, Corbacioglu A, Bullinger L, et al. Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. The New England journal of medicine. 2008;358(18):1909-18.

13. Wei J, Wunderlich M, Fox C, Alvarez S, Cigudosa JC, Wilhelm JS, et al. Microenvironment determines lineage fate in a human model of MLL-AF9 leukemia. Cancer Cell. 2008;13(6):483-95.

14. Medyouf H, Mossner M, Jann JC, Nolte F, Raffel S, Herrmann C, et al. Myelodysplastic cells in patients reprogram mesenchymal stromal cells to establish a transplantable stem cell niche disease unit. Cell Stem Cell. 2014;14(6):824-37.

15. Reinisch A, Chan SM, Thomas D, Majeti R. Biology and clinical relevance of acute myeloid leukemia stem cells. Semin Hematol. 2015;52(3):150-64.

16. Saultz JN, Garzon R. Acute Myeloid Leukemia: A Concise Review. J Clin Med. 2016;5(3).

17. Anguille S, Van Tendeloo VF, Berneman ZN. Leukemia-associated antigens and their relevance to the immunotherapy of acute myeloid leukemia. Leukemia. 2012;26(10):2186-96.

18. Greiner J, Ringhoffer M, Taniguchi M, Hauser T, Schmitt A, Dohner H, et al. Characterization of several leukemia-associated antigens inducing humoral immune responses in acute and chronic myeloid leukemia. Int J Cancer. 2003;106(2):224-31.

19. Kolb HJ, Schattenberg A, Goldman JM, Hertenstein B, Jacobsen N, Arcese W, et al. Graft-versus-leukemia effect of donor lymphocyte transfusions in marrow grafted patients. Blood. 1995;86(5):2041-50.

20. Schmid C, Labopin M, Nagler A, Bornhauser M, Finke J, Fassas A, et al. Donor lymphocyte infusion in the treatment of first hematological relapse after allogeneic stem-cell transplantation in adults with acute myeloid leukemia: a retrospective risk factors analysis and comparison with other strategies by the EBMT Acute Leukemia Working Party. J Clin Oncol. 2007;25(31):4938-45.

21. Ailles LE, Gerhard B, Hogge DE. Detection and characterization of primitive malignant and normal progenitors in patients with acute myelogenous leukemia using long-term coculture with supportive feeder layers and cytokines. Blood. 1997;90(7):2555-64.

22. Blair A, Hogge DE, Ailles LE, Lansdorp PM, Sutherland HJ. Lack of expression of Thy-1 (CD90) on acute myeloid leukemia cells with long-term proliferative ability in vitro and in vivo. Blood. 1997;89(9):3104-12.

23. Lapidot T, Sirard C, Vormoor J, Murdoch B, Hoang T, Caceres-Cortes J, et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature. 1994;367(6464):645-8.

24. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nature medicine. 1997;3(7):730-7.

25. Blair A, Ailles LE, Hogge DE, Sutherland HJ. Phenotype of primitive AML cells which engraft NOD/SCID mice. Experimental Hematology. 1996;24(9):558-.

26. Kreso A, Dick JE. Evolution of the Cancer Stem Cell Model. Cell Stem Cell. 2014;14(3):275-91.

27. Guan Y, Gerhard B, Hogge DE. Detection, isolation, and stimulation of quiescent primitive leukemic progenitor cells from patients with acute myeloid leukemia (AML). Blood. 2003;101(8):3142-9.

28. Jordan CT. The leukemic stem cell. Best practice & research Clinical haematology. 2007;20(1):13-8.

29. Clarke MF, Dick JE, Dirks PB, Eaves CJ, Jamieson CH, Jones DL, et al. Cancer stem cells--perspectives on current status and future directions: AACR Workshop on cancer stem cells. Cancer research. 2006;66(19):9339-44.

30. Ye H, Adane B, Khan N, Sullivan T, Minhajuddin M, Gasparetto M, et al. Leukemic Stem Cells Evade Chemotherapy by Metabolic Adaptation to an Adipose Tissue Niche. Cell Stem Cell. 2016;19(1):23-37.

31. Jan M, Snyder TM, Corces-Zimmerman MR, Vyas P, Weissman IL, Quake SR, et al. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. Science translational medicine. 2012;4(149):149ra18.

32. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature. 2012;481(7382):506-10.

33. Masetti R, Castelli I, Astolfi A, Bertuccio SN, Indio V, Togni M, et al. Genomic complexity and dynamics of clonal evolution in childhood acute myeloid leukemia studied with whole-exome sequencing. Oncotarget. 2016.

34. Ho TC, LaMere M, Stevens BM, Ashton JM, Myers JR, O'Dwyer KM, et al. Evolution of acute myelogenous leukemia stem cell properties after treatment and progression. Blood. 2016;128(13):1671-8.

35. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, Chen K, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009;361(11):1058-66.

36. Babayan A, Alawi M, Gormley M, Muller V, Wikman H, McMullin RP, et al. Comparative study of whole genome amplification and next generation sequencing performance of single cancer cells. Oncotarget. 2016.

37. Gawad C, Koh W, Quake SR. Single-cell genome sequencing: current state of the science. Nat Rev Genet. 2016;17(3):175-88.

38. Hollink IH, Zwaan CM, Zimmermann M, Arentsen-Peters TC, Pieters R, Cloos J, et al. Favorable prognostic impact of NPM1 gene mutations in childhood acute myeloid leukemia, with emphasis on cytogenetically normal AML. Leukemia. 2009;23(2):262-70.

39. Falini B, Nicoletti I, Martelli MF, Mecucci C. Acute myeloid leukemia carrying cytoplasmic/mutated nucleophosmin (NPMc+ AML): biologic and clinical features. Blood. 2007;109(3):874-85.

40. Thiede C, Creutzig E, Reinhardt D, Ehninger G, Creutzig U. Different types of NPM1 mutations in children and adults: evidence for an effect of patient age on the prevalence of the TCTG-tandem duplication in NPM1-exon 12. Leukemia. 2007;21(2):366-7.

41. Brown P, McIntyre E, Rau R, Meshinchi S, Lacayo N, Dahl G, et al. The incidence and clinical significance of nucleophosmin mutations in childhood AML. Blood. 2007;110(3):979-85.

42. Gilliland DG, Griffin JD. The roles of FLT3 in hematopoiesis and leukemia. Blood. 2002;100(5):1532-42.

43. Cloos J, Goemans BF, Hess CJ, van Oostveen JW, Waisfisz Q, Corthals S, et al. Stability and prognostic influence of FLT3 mutations in paired initial and relapsed AML samples. Leukemia. 2006;20(7):1217-20.

44. Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, Koboldt DC, et al. The origin and evolution of mutations in acute myeloid leukemia. Cell. 2012;150(2):264-78.

45. Jan M, Majeti R. Clonal evolution of acute leukemia genomes. Oncogene. 2013;32(2):135-40.

46. Dvorakova D, Racil Z, Borsky M, Robesova B, Jeziskova I, Razga F, et al. Clonal heterogeneity in patients with cytogenetically normal acute myeloid leukemia with NPM1 mutations. Leukemia & lymphoma. 2013;54(5):1056-60.

47. Klco JM, Spencer DH, Miller CA, Griffith M, Lamprecht TL, O'Laughlin M, et al. Functional heterogeneity of genetically defined subclones in acute myeloid leukemia. Cancer Cell. 2014;25(3):379-92.

48. de Bourcy CF, De Vlaminck I, Kanbar JN, Wang J, Gawad C, Quake SR. A quantitative comparison of single-cell whole genome amplification methods. PLoS One. 2014;9(8):e105585.

49. Hou Y, Wu K, Shi X, Li F, Song L, Wu H, et al. Comparison of variations detection between whole-genome amplification methods used in single-cell resequencing. Gigascience. 2015;4:37.

Tables

Table 1. Patient characteristics.

	patient 1	patient 2	patient 3
sex	female	female	female
Age at diagnosis [years]	15	5	7
FAB classification	AML M2	AML M1	AML M2
karyotype	46,XX, inv(9)(p11q12)c [25]	46,XX,t(6;9)(p23;q34)[2]/47,idem,+8[18]	n.a.
genetics	NPM1 c.860_863dupTCTG (p.Trp288Cysfs*?) FLT3 ITD c.1812_1813ins36 (p.Glu604_Phe605ins12)	NPM1 WT FLT3 c.1734_1837+2dup (p.?)	NPM1 c.860_863dupTCTG (p.Trp288Cysfs*?), FLT3 WT
Flow cytometry	CD33, CD13, part CD34, cyMPO, CD7	CD33, part. CD34, CD117, part CD7, CD13, CD45, cyMPO	CD33, CD13, CD38, CD123, CD114low, cyMPO, part CD117
Risk group	HR	HR	SR
% blasts	86	78	90
WBC [10^3/µl]	256.2	46.9	85.1
Hb [g/dl]	7.9	8.8	9.8

Figures

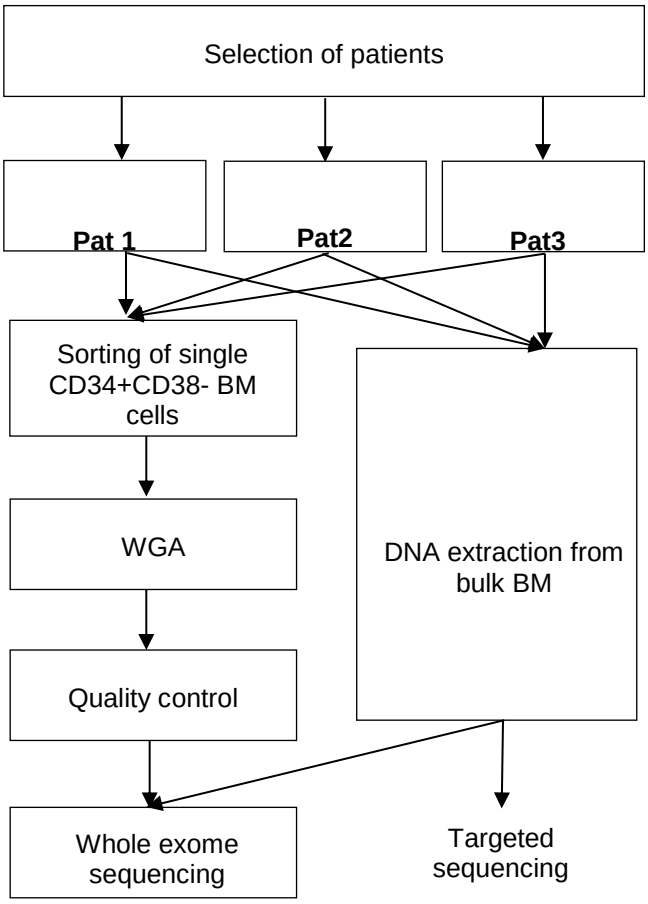


Figure 1. The workflow for three pediatric AML patients with defined mutations in FLT3 and/or NPM1. DNA derived from the bulk BM sample as well as WGA DNA from sorted CD34+/CD38- single cells of each patient was subjected to whole exome sequencing (WES). The DNA derived from the bulk samples was additionally analysed using targeted sequencing of 54 genes associated with myeloid malignancies.

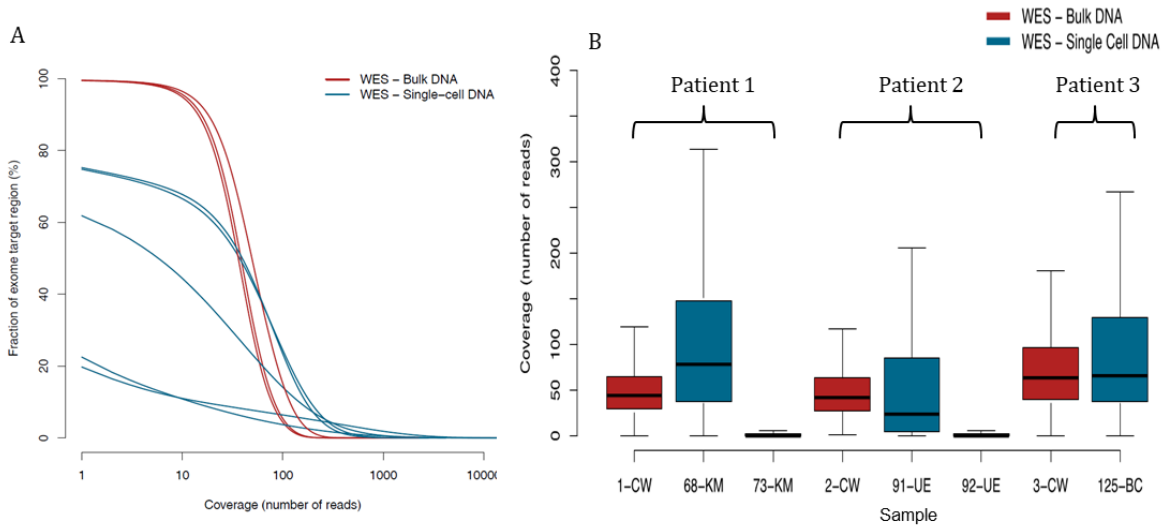


Figure 2. Fraction of the target region (total size ~75M bp) of WES covered with at least a certain number of reads (shown on the x-axis) (A). Coverage on the target region of the TSM panel achieved by whole exome sequencing in bulk samples (red) and single cells (blue) (B).

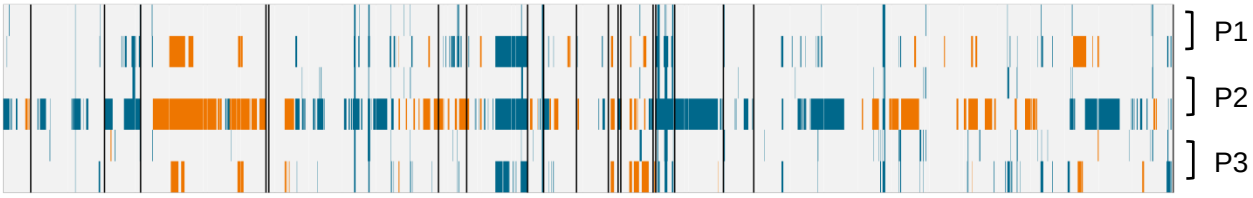


Figure 3. Coverage uniformity for bulk DNA and sorted single cells from the three patients (P1-3) displayed for each genomic position targeted with both the TSM panel and the WES approach. Orange: observed coverage is larger than 5-times the median coverage. Blue: observed coverage is smaller than one fifth of the median coverage. Gray: otherwise (i.e. observed coverage is not uniform). The upper line represents the single cell and the lower line represents the bulk sample for each patient respectively.

Figure 4. Coverage profile of the overlapping target regions from WES and TSM approach for the bulk and one single-cell DNA derived from patient 1.

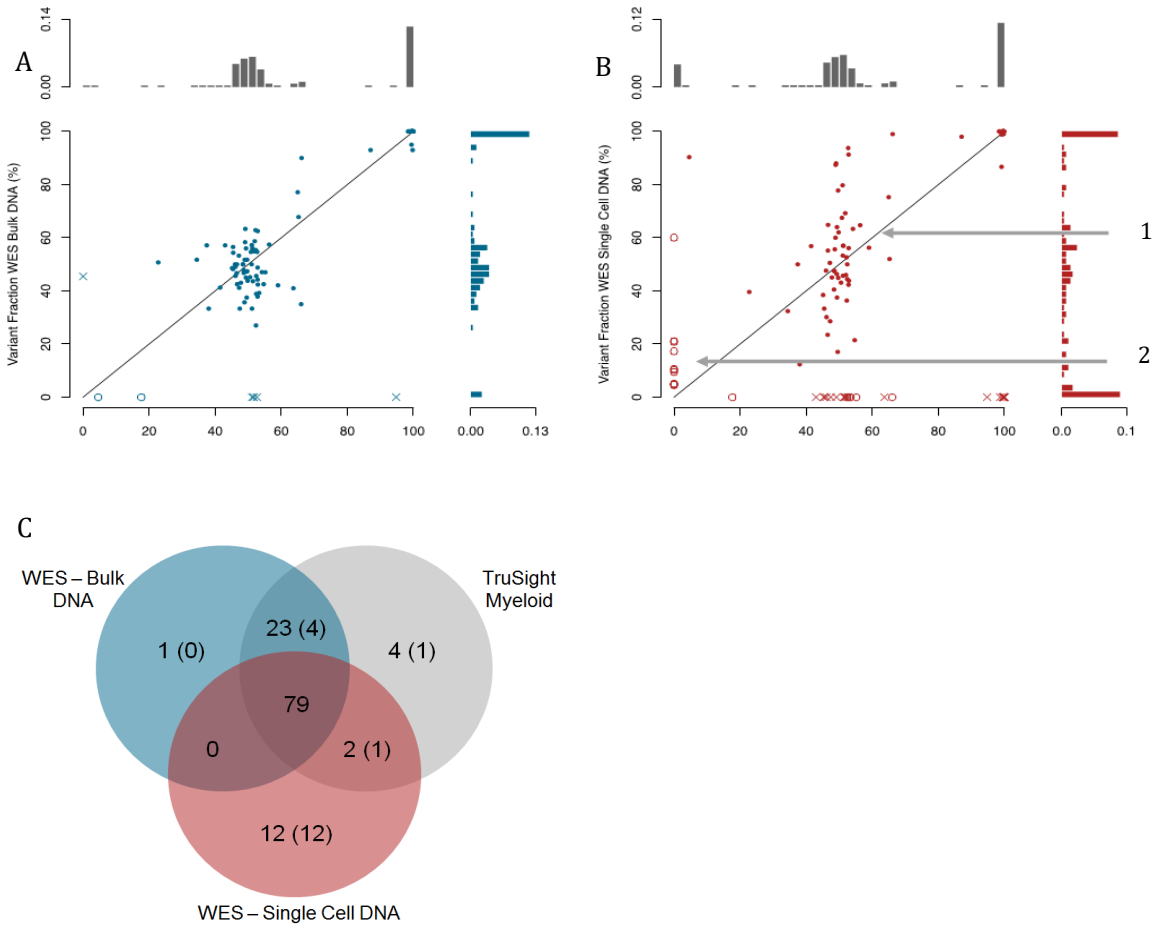


Figure 5. Comparison of the variant fractions found with the TSM panel to those identified by WES in bulk DNA (A) and in single-cell DNA (B). Filled circles indicate variants consistently found in both technologies; open circles and crosses show variants were not found by one technology because of low coverage (crosses) or for other reasons (open circles). 1: Variant fractions estimated from single-cell data were noisy (data dispersed along the y-axis). 2: All but one variant found in Single Cell DNA had low variant fraction. C: Number of variants called with at least one of the 3 approaches. Variants excluded due to insufficient coverage are not reported in the brackets.

5.3 Mutated WT1, FLT3-ITD, and NUP98-NSD1 Fusion in Various Combinations Define a Poor Prognostic Group in Pediatric Acute Myeloid Leukemia

Contribution to present publication:

- Conception: 20%
- Experimental work: 50%
    - preparation of DNA and RNA from cryopreserved BM or PB cells from pediatric AML patients
    - preparation of libraries for targeted NGS
    - cDNA synthesis
    - quantitative PCR and detection of NUP98-NSD1 fusion transcripts
- Data analysis: 50%
    - Analysis and interpretation of targeted NGS data
    - Analysis and interpretation of quantitative PCR data
- Writing the manuscript: 30%
- Revising the manuscript: 30%

The above listed contributions of Christiane Walter to the publication are correct.

Essen, den _____ _____

Unterschrift der Doktorandin

Essen, den _____ _____

Unterschrift des wissenschftl. Betreuers/Mitglieds
der Universität Duisburg-Essen

Hindawi
Journal of Oncology
Volume 2019, Article ID 1609128, 15 pages
https://doi.org/10.1155/2019/1609128

Research Article

Mutated *WT1*, *FLT3-ITD*, and *NUP98-NSD1* Fusion in Various Combinations Define a Poor Prognostic Group in Pediatric Acute Myeloid Leukemia

Naghmeh Niktoreh,[1] **Christiane Walter**,[1] **Martin Zimmermann**,[2] **Christine von Neuhoff**,[3] **Nils von Neuhoff**,[1] **Mareike Rasche**,[1] **Katharina Waack**,[4] **Ursula Creutzig**,[2] **Helmut Hanenberg**,[1,5] **and Dirk Reinhardt**[1]

[1]*Department of Pediatrics III, University Children's Hospital Essen, University of Duisburg-Essen, 45122 Essen, Germany*
[2]*Department of Pediatric Hematology and Oncology, Hannover Medical School, 30625 Hannover, Germany*
[3]*Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, LVR Klinikum Essen, University Hospital Essen, University of Duisburg-Essen, Germany*
[4]*Centre for Research Acceleration in Pediatrics GmbH, Hannover, Germany*
[5]*Department of Otorhinolaryngology & Head/Neck Surgery, Heinrich Heine University, 40225 Duesseldorf, Germany*

Correspondence should be addressed to Helmut Hanenberg; helmut.hanenberg@uk-essen.de
and Dirk Reinhardt; dirk.reinhardt@uk-essen.de

Received 26 April 2019; Accepted 24 June 2019; Published 30 July 2019

Guest Editor: Annalisa Lonetti

Acute myeloid leukemia is a life-threatening malignancy in children and adolescents treated predominantly by risk-adapted intensive chemotherapy that is partly supported by allogeneic stem cell transplantation. Mutations in the *WT1* gene and *NUP98-NSD1* fusion are predictors of poor survival outcome/prognosis that frequently occur in combination with internal tandem duplications of the juxta-membrane domain of *FLT3* (*FLT3-ITD*). To re-evaluate the effect of these factors in contemporary protocols, 353 patients (<18 years) treated in Germany with AML-BFM treatment protocols between 2004 and 2017 were included. Presence of mutated *WT1* and *FLT3-ITD* in blasts (n=19) resulted in low 3-year event-free survival of 29% and overall survival of 33% compared to rates of 45-63% and 67-87% in patients with only one (only *FLT3-ITD; n=33,* only *WT1* mutation; n=29) or none of these mutations (n=272). Including *NUP98-NSD1* and high allelic ratio (AR) of *FLT3-ITD* (AR ≥0.4) in the analysis revealed very poor outcomes for patients with co-occurrence of all three factors or any of double combinations. All these patients (n=15) experienced events and the probability of overall survival was low (27%). We conclude that co-occurrence of *WT1* mutation, *NUP98-NSD1,* and *FLT3-ITD* with an AR ≥0.4 as triple or double mutations still predicts dismal response to contemporary first- and second-line treatment for pediatric acute myeloid leukemia.

1. Introduction

Pediatric acute myeloid leukemia (AML) is a rare and heterogeneous disorder, for which continuous improvement of risk-adapted treatment approaches over the last 30 years has led to overall survival rates of approximately 70% [1, 2]. In current pediatric AML treatment protocols, cytogenetic abnormalities of the leukemic blasts at initial diagnosis are important indicators for risk group stratification and treatment assignment [1, 2]. Approximately, 25% of pediatric patients have AML blasts with a normal karyotype, but even these cases often harbor somatic mutations in genes such as *WILMS TUMOR 1* (*WT1*), *NPM1*, *NRAS*, *KRAS*, *Fms-like tyrosine kinase 3* (*FLT3*), and/or *c-KIT/CDII7* [1, 2].

The *WT1* gene is located on chromosome 11, has ten exons and four zinc finger domains, and functions as a transcription factor and master regulator of tissue development [3]. Within normal hematopoiesis, *WT1* has two distinct roles: in early

stages, it mediates quiescence of primitive progenitor cells, and later, WT1 expression is important for differentiation towards the myeloid lineage [4]. In AML, *WT1* mutations are present in approximately 10% of patients and predominantly located in exons 7 and 9, which contain the DNA-binding zinc finger domains of the protein. The majority of these mutations are out-of-frame deletion/insertions or premature termination codons that will lead to truncated proteins with altered functional consequences for the cells [5]. If these truncated proteins are stable, they might have dominant negative effects by partially blocking the wild-type WT1 protein; if unstable, the diminished WT1 protein levels may lead to haploinsufficiency [5]. Nevertheless, it has been clearly established that the occurrence of *WT1* mutations in AML blasts with normal karyotypes is associated with adverse clinical outcomes in adult [6–9] as well as pediatric patients [10, 11].

Somatic *WT1* mutations in AML blasts often co-occur with other genetic aberrations, most frequently with an internal tandem duplication in the juxta-membrane domain of the tyrosine kinase receptor *FLT3* (*FLT3-ITD*) [5]. Classified as type-I or proliferating mutation, *FLT3-ITDs* are present in 10-15% of pediatric AML cases and lead to poor clinical outcomes [12–14]. We previously demonstrated in a cohort of 298 pediatric patients with *de novo* AML treated before 2004 on AML-BFM protocols that the combination of *FLT3-ITD* and mutated *WT1* is associated with even worse survival [10]. Comparably, an independent study from the Children's Oncology Group (COG) in a cohort of 842 children with *de novo* AML showed that the poor prognostic impact of *WT1* mutations depends on the *FLT3-ITD* status [11]. These two pediatric studies confirmed earlier findings in adults that first established the adverse prognostic impact of both *WT1* and *FLT3-ITD* mutations [15, 16].

Two additional prognostic indicators in *FLT3-ITD*-positive AML cases established in the last few years are the mutational burden in each patient defined as the ratio between mutant and wild-type *FLT3-ITD* alleles (allelic ratio, AR) [12, 17, 18] and the co-occurrence of *FLT3-ITD* with a cytogenetically cryptic translocation of chromosomes 5 and 11 or t(5;11)(q35;p15) [19]. This translocation leads to fusion of the *nucleoporin* (*NUP98*) gene on chromosome 11 and the gene for nuclear receptor binding SET-domain protein 1 (*NSD1*) of chromosome 5 (*NUP98-NSD1*). As the breakpoints for the NUP98 gene are often not detected by classical cytogenetic due to its terminal localization at 11p15, it has been described in AML cases with a "normal" karyotype [20]. Importantly, this rare recurrent aberration is mutually exclusive with other recurrent translocations and more prevalent in pediatric AML, in which it is associated with the presence of *FLT3-ITD* and poor survival outcomes [21, 22].

In the present study, we re-evaluated the role of mutations in *WT1*, *FLT3-ITD*, and the *NUP98-NSD1* translocation as prognostic factors in two contemporary pediatric treatment protocols by analyzing their association with co-occurring genetic and cytogenetic aberrations and by determining their clinical significance and influence on treatment outcome. Thereby, we were able to define a group of high-risk patients for which the efforts for salvage/second line treatment largely failed.

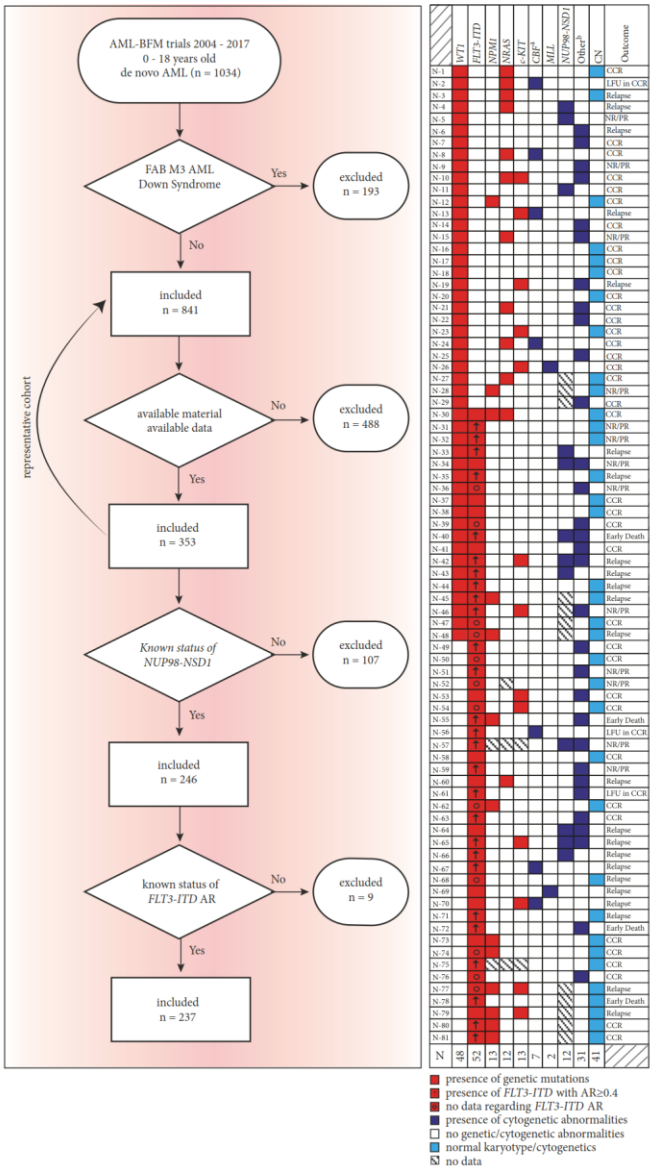
2. Materials and Methods

From April 2004 to May 2017, 841 patients aged 0–18 years with *de novo* AML (excluding FAB M3 and Down Syndrome) were treated in Germany according to the AML-BFM 04 trial (ClinicalTrials.gov Identifier: NCT00111345) or the AML-BFM 2012 registry and trial (EudraCT number: 2013-000018-39) (Figure 1(a)). Both trials were approved by the ethical committees and institutional review boards of university hospitals of Münster and Hannover and an informed consent was obtained from each patient or their legal guardians before the beginning of treatment. Standard procedures for the diagnosis of AML were carried out by the German AML-BFM reference laboratory as previously described [23–25]. This included mutation analysis in *WT1*, *FLT3-ITD*, *NPM1*, *NRAS*, and *c-KIT* by Sanger and/or next-generation sequencing or GeneScan analysis. In 353 patients (42%), sufficient material and clinical data were available for further analysis. As a confirmation, material from *WT1* and/or *FLT3-ITD* positive and negative cases was re-analyzed by next-generation sequencing (NGS) using the TruSight Myeloid Panel (Illumina)[26] with median read counts for *WT1* and *FLT3-ITD* of around 4,200 and 6,000 reads, respectively, as we described previously [27]. In addition, the allelic ratio of *FLT3-ITD* to *FLT3* wild-type was calculated via GeneScan analysis [13] and the expression of *NUP98-NSD1* was analyzed in 246 out of 353 patients with available material by real-time quantitative PCR using previously described primers [19]. Initial analysis demonstrated that the selected cohort was representative for all patients treated between 2004 and 2017 on the AML-BFM protocols for features such as gender, age, AML subtype, initial cytogenetics, and preliminary, early response to treatment (data not shown).

Clinical end-points were defined as previously described [28, 29] and survival rates were calculated via Kaplan-Meier analysis and compared by log-rank test. Multivariate analysis was performed using Cox regression model evaluating the hazard ratio (HR) of each covariate with 95% confidence interval (CI). Stem cell transplantation was included in the Cox regression model as a time-dependent variable. Differences with a p value less than 0.05 were considered as significant. Data were analyzed using the Statistical Analysis System software version 9.4 (SAS Institute, Cary, NC). Data acquisition was stopped at June 30, 2018, with a median follow-up of 3.6 years.

3. Results

3.1. Study Cohort and Patient Characteristics. In this study, we included 353 patients treated on either the AML-BFM 2004 or AML-BFM 2012 protocol for whom sufficient material and information were available (Figure 1(a)). As shown in Table 1, 48 (14%) patients had *WT1* and 52 (15%) *FLT3-ITD* mutations in their leukemic blasts at diagnosis. Mutations in *NPM1*, *NRAS*, and *c-KIT* were present in the blasts of 9%, 17%, and 12% of patients, respectively. Most patients with mutated *WT1* (n=35, 73%) harbored at least one co-occurring mutation in the AML blasts, with the most common being *FLT3-ITD* (n=19, 40%) followed by *NRAS* mutations (n=11,



(a) Study flowchart (b) Features

Figure 1: *Study flowchart and patient characteristics.* (a) Study flowchart outlining the process of patient recruitment in the data analysis. (b) *WT1* mutations often co-occurred with *FLT3-ITD* and other genetic aberrations. AML-BFM, acute myeloid leukemia-Berlin-Frankfurt-Muenster; n, number; *WT1*, *Wilms Tumor 1*; *FLT3-ITD*, fms-related tyrosine kinase 3-internal tandem duplication; *NPM1*, *nucleophosmin 1*; *NRAS*, *neuroblastoma RAS viral oncogene homolog*; *c-KIT*, *KIT proto-oncogene*; CBF, core binding factor; MLL, rearrangements of *MLL* gene; *NUP98-NSD1*, *Nucleoporin-Nuclear Receptor Binding SET Domain Protein 1* fusion gene; CN, cytogenetic-normal AML; AR, allelic ratio; CCR, continued complete remission; LFU, lost to followup; NR, non-response; PR, partial remission. [a]CBF aberrations include translocation of chromosomes 8 and 21 and inversion or translocation of chromosome 16. [b]Other cytogenetic aberrations such as trisomy 8, various chromosomal translocations, and complex karyotype alterations.

Table 1: Patient Characteristics.

FEATURES		All patients n	%	*WT1* wild-type n	%	*WT1* mutated n	%	p*	*FLT3-ITD* neg. n	%	*FLT3-ITD* pos. n	%	p*
Study Population													
Number (%)		353	100	305	100	48	100		301	100	52	100	
Age (years), median (range)		9.09 (0 -18)		7.97 (0 - 18)		10.68 (0.8 - 17.8)		0.03	7.8 (0 -18)		12.95 (2.7 - 17.9)		0.0001
Gender													
male		183	52%	159	52%	24	50%	0.783	153	51%	30	58%	0.36
female		170	48%	146	48%	24	50%		148	49%	22	42%	
WBC count at diagnosis													
median x 10^9 cells/L (range)		24.5 (0.019 - 475)		24 (0.23 - 475)		42.85 (0.019 - 324)		0.1	20.1 (0.019-475)		73.5 (1.8 - 324)		0.0001
Morphological Classification													
FAB	M0	9	3%	7	2%	2	4%	0.008	5	2%	4	8%	0.0001
	M1/M2	134	38%	107	35%	27	56%		100	33%	34	65%	
	M4Eo+	31	9%	27	9%	4	8%		29	10%	2	4%	
	M4Eo-/M5	132	37%	125	41%	7	15%		121	40%	11	21%	
	M6	3	1%	2	1%	1	2%		3	1%	0	0%	
	M7	22	6%	21	7%	1	2%		22	7%	0	0%	
	AUL/other	22	6%	16	5%	6	13%		21	7%	1	2%	
Cytogenetics													
	t(8;21)	38	11%	38	12%	4	8%	<0.0001	36	12%	2	4%	0.0001
	inv(16)	32	9%	28	9%	0	0%		31	10%	1	2%	
	MLL rearr.	61	17%	60	20%	1	2%		60	20%	1	2%	
	others	117	33%	99	32%	18	38%		97	32%	20	38%	
	normal	97	27%	72	24%	25	52%		69	23%	28	54%	
	no data	8	2%	8	3%	0	0%		8	3%	0	0%	
NUP98-NSD1													
	positive	15	4%	7	2%	8	17%	<0.0001	6	2%	9	17%	<0.0001
	negative	231	65%	198	65%	33	69%		197	65%	34	65%	
	no data	107	30%	100	33%	7	15%		98	33%	9	17%	
Co-mutations													
FLT3-ITD	negative	301	85%	272	89%	29	60%	<0.0001					
	positive	52	15%	33	11%	19	40%						
WT1	wildtype	305	86%						272	90%	33	63%	0.0001
	mutated	48	14%						29	10%	19	37%	
NPM1	wildtype	316	90%	273	90%	43	90%	0.69	277	92%	39	75%	0.0004
	mutated	31	9%	26	9%	5	10%		20	7%	11	21%	
	no data	6	2%	6	2%	0	0%		4	1%	2	4%	
nRAS	wildtype	282	80%	245	80%	37	77%	0.26	235	78%	47	90%	0.008
	mutated	59	17%	48	16%	11	23%		57	19%	2	4%	
	no data	12	3%	12	4%	0	0%		9	3%	3	6%	
KIT	wildtype	308	87%	267	88%	41	85%	0.51	266	88%	42	81%	0.31
	mutated	41	12%	34	11%	7	15%		33	11%	8	15%	
	no data	4	1%	4	1%	0	0%		2	1%	2	4%	
CEBPA	wildtype	163	46%	121	40%	42	88%	<0.0001	122	41%	41	79%	<0.0001
	single	6	2%	4	1%	2	4%		6	2%	0	0%	
	double	10	3%	9	3%	1	2%		8	3%	2	4%	
	no data	174	49%	174	57%	0	0%		165	55%	9	17%	
HSCT													
HSCT in Its CR		64	18%	56	18%	8	17%	0.78	53	18%	11	21%	0.54
Chemotherapy only		289	82%	249	82%	40	83%		248	82%	41	79%	
Patient Status													
	alive	251	71%	221	72%	30	63%	0.14	225	75%	26	50%	0.0001
	deceased	73	21%	58	19%	15	31%		51	17%	22	42%	
	LFU	29	8%	26	9%	3	6%		25	8%	4	8%	

n, number; *WT1*, *Wilms tumor 1*; *FLT3-ITD*, *fms related tyrosine kinase 3*-internal tandem duplication; WBC, white blood cell; FAB, French-American-British; M4Eo+, AML M4 subtype with the presence of atypical eosinophils; M4Eo-, AML M4 subtype without the presence of atypical eosinophils; AUL, acute undifferentiated leukemia; t, translocation; inv, inversion; MLL rear., rearrangement of MLL gene; *NUP98-NSD1*, *Nucleoporin-Nuclear Receptor Binding SET Domain Protein 1* fusion gene; *NPM1*, *nucleophosmin 1*; *NRAS*, *neuroblastoma RAS viral oncogene homolog*; *c-KIT*, *KIT* proto-oncogene; *CEBPA*, *CCAAT/enhancer binding protein (C/EBP) alpha*; HSCT, hematopoietic stem cell transplantation; CR, complete remission; HSCT, hematopoietic stem cell transplantation; LFU, lost to follow-up. *p-values derived from Pearson's Chi-squares test.

23%, Table 1 and Figure 1(b)). Comparably, the majority of patients with *FLT3-ITD* had additional mutations in other genes (n=32, 62%), most commonly in *WT1* (n=19, 37%) and *NPM1* (n=11, 21%). Patients with mutated *WT1* or *FLT3-ITD* were older compared to the rest of the study cohort, and AML FAB M1/M2 was the most common morphologic subtype in both groups (Table 1). In addition, the AML blasts of more than half of patients with *WT1* (n=25/48, 52%) and *FLT3-ITD* (n=28/52, 54%) mutations had a normal karyotype at diagnosis; these percentages were significantly higher than those in patients without mutations in each of the two genes ($p<0.0001$, Table 1).

3.2. Characteristics of WT1 Mutations. We identified 64 different *WT1* sequence alterations in 48 patients (Table 2). These alterations were frequently located in exon 7 (n=55, 86%) and predominantly resulted in frameshifts producing premature termination codons (PTCs). In total, nine single nucleotide variants (SNVs) were found, mostly in exon 9 (n=7, 78%). Only three of the nine SNVs were not previously reported as pathogenic (Table 2). Using NGS, we characterized multiple distinct *WT1* mutations with highly diverse variant allele frequencies in 13 patients (11 patients had two and 2 patients, three distinct mutations). We then analyzed the heterozygosity of these mutations via the integrative genomic viewer (Broad Institute, MA, USA) and determined that they were all located on individual/different alleles/reads (Table 2).

3.3. Survival Significance of the Genomic Aberrations. Next, we analyzed the impact of each mutation on the clinical outcomes. Our analysis identified *WT1* and *FLT3-ITD,* but not *NRAS*, *NPM1,* or *c-KIT* mutations as single factors that significantly increased the chance of relapse or treatment failure and reduced the probability of 3-year overall survival (OS) in our patient cohort (Figures 2(a), 2(b), and 3). In addition, *FLT3-ITD* but not *WT1* mutations significantly decreased the 3-year probability of event-free survival (EFS, Figure 2(b)). When we grouped the two mutations together, the survival analysis revealed a 3-year EFS of 29±11% for patients with both *WT1* and *FLT3-ITD* mutations compared to 63±3% for patients with none of these mutations ($p=0.0004$) and 61±11% or 45±9% for patients with only *WT1* mutation ($p=0.016$) or *FLT3-ITD* ($p=0.16$), respectively (Figure 2(c)). Corresponding to this low EFS, co-occurrence of these two mutations was associated with an increased cumulative incidence of relapse (CIR) of 65±12% compared to 32±12% for patients with none of these mutations ($p=0.002$) and 39±11% or 46±9% for patients with only *WT1* mutation ($p=0.05$) or *FLT3-ITD* ($p=0.08$), respectively (Figure 2(c)). Furthermore, we identified a low 3-year OS probability of 33±12% in patients with co-occurrence of *WT1* and *FLT3-ITD*, which was significantly lower than those of patients without these mutations (81±3%, $p<0.0001$), patients with only mutated *WT1* (87±7%, $p=0.0007$), and patients with only *FLT3-ITD* (67±9%, $p=0.017$, Figure 2(c)). Comparing the curves for EFS and OS clearly demonstrated that our second line treatment was not able to rescue any patient with co-occurrence of *WT1* and *FLT3-ITD* mutations, while the OS rates increased by more than 20% for the other three subgroups (Figure 2(c)).

3.4. Impact of NUP98-NSD1 Fusion. To further characterize the prognostic significance of *WT1* and *FLT3-ITD* mutations, we analyzed the expression of *NUP98-NSD1* fusion in our patient cohort (Figure 1(a)). From 246 patients with available material for this retrospective real-time quantitative PCR analysis, 15 (6%) of them were identified to have the *NUP98-NSD1* translocation. Most of these patients (12/15, 80%) harbored additional *WT1* or *FLT3-ITD* mutations: 3 patients carried both *WT1* and *NUP98-NSD1*, 4 had a co-occurrence of *FLT3-ITD* and *NUP98-NSD1*, and 5 patients carried all three genetic alterations (Figure 1(b)). Only 1 of these 15 patients had a previous known status of *NUP98-NSD1* by conventional karyotyping: 2 others were previously diagnosed with deletion of chromosome 5, 1 carried an inversion of chromosome 16 (no other mutations and still in continuous complete remission), 4 carried complex karyotypes or rare aberrations, and 7 had no other cytogenetic abnormalities (data not shown).

We then analyzed the prognostic significance of *NUP98-NSD1* in the cohort of 246 patients with the known status of this fusion gene (Figure 1(a)). As a single factor, the presence of *NUP98-NSD1* in AML blasts of patients at diagnosis was associated with a significant increase in CIR (81%) in addition to decreased probabilities of 3-year EFS and OS (Figure 4(a)). Combining *NUP98-NSD1* with *WT1* and *FLT3-ITD* mutations in our multifactor survival analysis revealed that patients with all three or either two of these mutations had worse survival outcomes. These patients had a higher CIR of 73±11% compared to the CIR of 30±4% for patients with none of these aberrations or *NUP98-NSD1* alone ($p<0.0001$) and the CIR of 37±13% or 38±10% for patients with only mutated *WT1* ($p=0.0078$) or *FLT3-ITD* ($p=0.013$), respectively (Figures 4(a) and 4(b)). The increased CIR translated into a lower 3-year EFS probability of 23±10% for patients with triple or double mutations compared to the EFS of 62±4% for patients with none of these mutations or only *NUP98-NSD1* ($p<0.0001$) and the EFS of 63±13% or 54±10% for patients with only *WT1* ($p=0.003$) or *FLT3-ITD* ($p=0.036$) mutations, respectively (Figure 4(b)). Moreover, co-occurrence of all three or any double mutations resulted in a significantly lower 3-year OS probability of 42±12% compared to 80±8% for patients with none of the mutations or only *NUP98-NSD1* ($p=0.0003$) and 88±8% or 73±10% for patients with only *WT1* ($p=0.0007$) or *FLT3-ITD* ($p=0.049$) mutations, respectively (Figure 4(b)).

3.5. Survival Significance of the FLT3-ITD Allelic Ratio. We have previously established the prognostic significance of an *FLT3-ITD* allelic ratio of ≥0.4 in pediatric AML [12]. Therefore, to determine the impact of the mutational burden of *FLT3-ITD* on treatment outcomes in the present cohort, we calculated the *FLT3-ITD* AR in patients with available data/material. As indicated in Figure 1(b), 27 patients had an AR ≥0.4 at diagnosis. Analyzing the survival impact of

Table 2: Characteristics of *WT1* Variants.

UPN	exon	seq. read	mutation sequence[a]	amino acid alteration	VF (%)	dbSNP or COSMIC ID	published	previously reported sample	outcome
missense substitutions									
8	9		c.1333C>T	p.Arg445Trp	19.1	rs121907900, COSM21417	Yes	WT	CCR
15	9		c.1345C>A	p.Leu449Met	5.49		No		CCR
20	9		c.1385G>A	p.Arg462Gln	47.21	rs121907903, COSM4191067	Yes	AML, colon cancer, adenocarcinoma	CCR
21	9		c.1343A>G	p.His448Arg	33.12	COSM7335365	Yes	AML, mesothelioma	CCR
23	9		c.1333C>T	p.Arg445Trp	72.42	rs121907900	Yes	WT, DDS	CCR
26	7		c.1097C>G	p.Ser366Cys	2.57		No		CCR
35	9	different	c.1334G>A	p.Arg445Gln	3.12	rs121907903, COSM4191067	Yes	AML, colon cancer, adenocarcinoma	Relapse
	9	different	c.1307G>A	p.Cys436Tyr	44.21	COSM21438	Yes	AML	
nonsense substitutions/insertions, deletions or duplications									
1	7		c.1090_1093dupTC	p.Ala365Valfs*4	43	COSM5487332	Yes	AML	CCR
2	7		c.1048-4_1056dupGCAGGATGTGCGA	p.Arg353Alafs*19	30.25		No		LFU in CCR
3	7		c.1087_1161dup74	p.Lys387Asnfs*44	n.d.		No		Relapse
4	7	different	c.1087_1091dupCGGTC	p.Ala365Glyfs*69	5.08	COSM28954	Yes	AML, T-ALL	Relapse
4	7	different	c.1091C>A	p.Ser364*	28.38	COSM27307	Yes	AML, WT	
5	7		c.1083_1098delTGTACGGTCGGCATCT	p.Val362Argfs*65	46.82		No		NR/PR
6	7		c.1059dupT	p.Val354Cysfs*14	35.9	COSM1317324	Yes	AML	Relapse
7	7		c.1179dupG	p.His394Alafs*8	25		No		CCR
9	7	different	c.1078_1079insGCCGA	p.Thr360Serfs*74	38.7		No		NR/PR
	7	different	c.1084_1085insGC	p.Val362Glyfs*71	52.9		No		
10	7		c.1074_1077dupCCCG	p.Thr360Profs*9	9.9		No		CCR
11	7		c.1079_1090delCTCTTGTACGGTinsTGGG	p.Thr360Metfs*5	55.23		No		CCR
12	7		c.1058_1059insGA	p.Val354Metfs*5	31.6		No		CCR
13	7	different	c.1058_1059insGGTG	p.Pro355Cysfs*14	5.6		No		Relapse
	7	different	c.1078_1084dupACTCTTG	p.Val362Aspfs*8	8.3	COSM5879281	Yes	AML	
14	7		c.1090_1093dupTCGG	p.Ala365Valfs*4	22.81	COSM21392	Yes	AML	CCR
16	7		c.1054_1084dup	p.Val362Alafs*16	7.3		No		CCR
17	7		c.1087delCinsGGG	p.Arg363Glyfs*70	24.3		No		CCR
18	7		c.1054_1055insT	p.Arg352Leufs*16	67.2	COSM5751511	Yes	T-ALL	CCR
19	7		c.1077_1078insTGTTTCTTCCGCCCAG	p.Thr360Cysfs*13	36.95		No		Relapse
22	7		c.1087delCinsGG	p.Arg363Glyfs*5	41.88		Yes	AML	CCR
24	7		c.1083_1090dupTGTACGGT	p.Ser364Leufs*71	3.8	COSM27309	Yes	AML	CCR

Table 2: Continued.

UPN	exon	seq. read	mutation sequence[a]	amino acid alteration	VF (%)	dbSNP or COSMIC ID	published	previously reported sample	outcome
25	9	different	c.1323_1338dupAAAGTTCTCCCGGTCC	p.Asp447Lysfs*18	40.1		No		CCR
	9	different	c.1322_1332dupGAAAGTTCTCC	p.Arg445Glufs*9	40.5		No		
27	7	different	c.1077_1078insGTTG	p.Thr360Valfs*9	43.71		No		CCR
	7	different	c.1089dupG	p.Ser364Valfs*4	49.24	COSM28966	Yes	AML	
28	7	different	c.1058delGinsCCA	p.Arg353Profs*6	19.72		No		NR/PR
	7	different	c.1054_1055insAAAAAGATT	p.Arg352delins4	19.55		No		
29	7		c.1179dupG	p.His394Alafs*8	25		No		CCR
30	7		c.1048_1057delGATGTGCGACinsAAGG	p.Asp350_Arg353	46.34		No		CCR
31	7		c.1093dupG	p.Ala365Glyfs*3	44.49		Yes	AML	NR/PR
32	7		c.1048-8_1055dupGCCTGCAGGATGTGCG	p.Arg353Profs*20	2.5		No		NR/PR
33	7		c.1090_1091dupTC	p.Ala365Argfs*68	44.25	COSM28955	Yes	AML	Relapse
	7	different	c.1087delCinsGA	p.Arg363Glufs*5	4.1		No		
34	7	different	c.1086dupA	p.Arg363Thrfs*5	5.33	COSM1166631	Yes	AML	NR/PR
	7	different	c.1090_1093dupTC	p.Ala365Valfs*4	36.29	COSM5487332	Yes	AML	
36	7	different	c.1090_1093dupTCGG	p.Ala365Valfs*4	6.94	COSM5487332	Yes	AML	NR/PR
	7	different	c.1091dupC	p.Ala365Glyfs*3	39.42	COSM27304	Yes	AML	
37	7	different	c.1057delCinsGG	p.Arg353Glyfs*15	42.78		Yes	AML	CCR
	7	different	c.1087delCinsGGG	p.Arg363Glyfs*70	52.11		No		
38	7		c.1068_1076delAGTAGCCCCinsGACGGTCGTTATTA	p.Val357Thrfs*77	42.14		No		CCR
39	7		c.1087delCinsGG	p.Arg363Glyfs*5	47.54		Yes	AML	CCR
40	7	different	c.1058_1059insGGTGCCGCTCG	p.Gly356Leufs*6	48.49		No		Early Death
	7	different	c.1082_1091dupTTGTACGGTC	p.Ala365Cysfs*6	41.83	COSM27303	Yes	AML	
41	7	different	c.1123dupA	p.Met375Asnfs*9	44.5		Yes	AML	CCR
	7	different	c.1057_1058insTA	p.Arg353Leufs*6	45.8		No		
42	7		c.1051_1055dupGTGCG	p.Arg353Cysfs*7	34.73		No		Relapse
43	7		c.1058delGinsCC	p.Arg353Profs*15	44.78	COSM28946	Yes	AML	Relapse
44	7	different	c.1079_1101delinsGAA	p.Thr360Argfs*4	20.37		No		Relapse
	7	different	c.1088_1089insCTCGG	p.Ala365Glyfs*69	10.69		No		
45	7		c.1090_1091insAGGT	p.Ser364*fs*1	42.97		No		Relapse
46	7		c.1058delGinsCC	p.Arg353Profs*15	51.08	COSM28946	Yes	AML, T-ALL	NR/PR
47	7	different	c.1048-3_1055dupCAGGATGTGCG	p.Val354Metfs*8	2.49		No		CCR
	7	different	c.1053dupG	p.Arg352Alafs*16	3.83	COSM28980	Yes	AML	
	7	different	c.1054delCinsGG	p.Arg352Glyfs*16	35.86	COSM28970	Yes	AML, T-ALL	
48	7		c.1089_1090insGGCCTCTTGTACGG	p.Ser364Glyfs*73	40.49		No		Relapse

UPN, unique patient number; Seq. read, sequence read; VF, variant allele frequency; dup, duplication; ins, insertion; indel, Insertion-deletion; fs, frame-shift; *termination codon; WT, Wilms tumor; DDS, Denis-Drash syndrome; T-ALL, T-cell acute lymphoblastic leukemia; CCR, continued complete remission, NR, non-response; PR, partial response; LFU, lost to follow-up.
[a]Transcript ID: NM_000378 was used to describe all alterations.

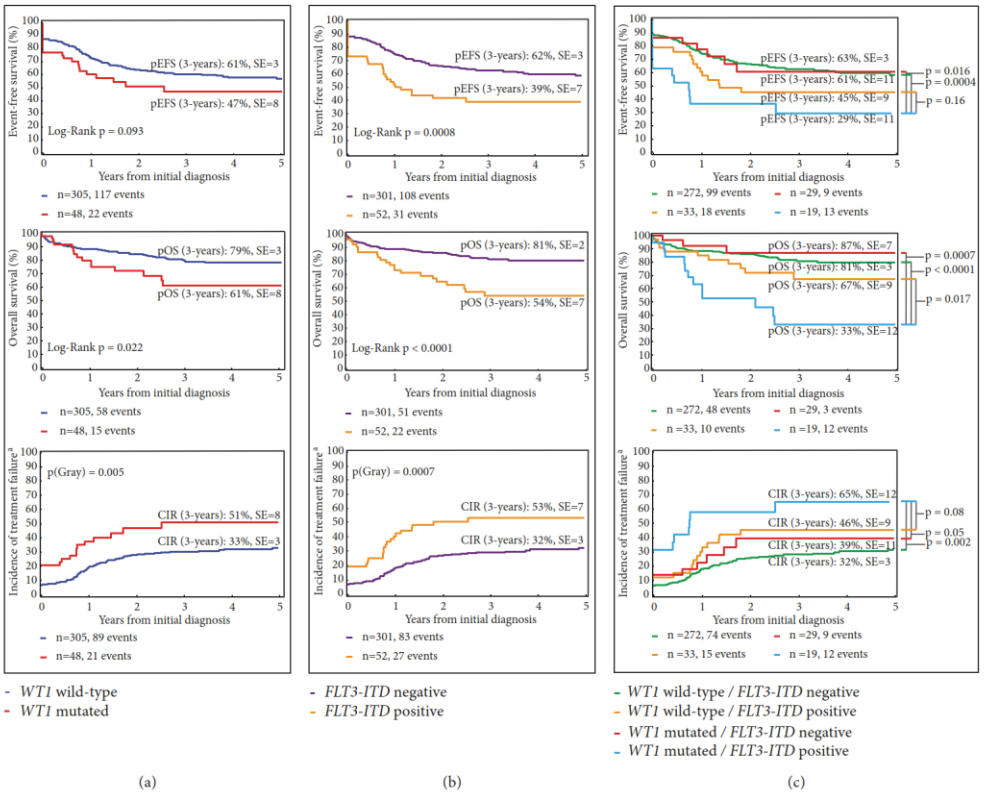


FIGURE 2: *Co-occurrence of WT1 and FLT3-ITD mutations at initial diagnosis of pediatric AML predicts poor survival outcomes.* (a) *WT1* mutation as single factor increased the incidence of relapse, reducing the probability of survival. (b) The presence of *FLT3-ITD*, individually, leads to an increased chance of relapse and decreased patient survival. (c) Clinical consequences of *WT1* mutations and *FLT3-ITD* were dependent on each other. *WT1*, *Wilms Tumor 1*; *FLT3-ITD*, fms related tyrosine kinase 3-internal tandem duplication; pEFS, probability of event-free survival; pOS, probability of overall survival; CIR, cumulative incidence of relapse; SE, standard error; n, number. [a]No response to treatment was considered as the occurrence of an event at time zero.

the *FLT3-ITD* AR ≥0.4 revealed that as a single factor, it was associated with an EFS of only 25±8% and an OS of only 47±10%, respectively (Figure 5(a)). Remarkably, the co-occurrence of *FLT3-ITD* AR ≥0.4, *WT1,* and *NUP98-NSD1* as triple or double mutations significantly increased the CIR to 93±15% compared to the CIR of 31±4% for patients with no mutations or only *NUP98-NSD1* or *FLT3-ITD* AR <0.4 (p<0.0001) and to the CIR of 31±11% or 36±15% in patients with only *WT1* (p<0.0001) or *FLT3-ITD* AR ≥0.4 (p=0.001) mutations, respectively (Figure 5(b)). The probability of 3-year EFS was zero in patients with double or triple *WT1*, *FLT3-ITD* AR ≥0.4, and *NUP98-NSD1* mutations as opposed to 61±4% in patients with no mutations or only *NUP98-NSD1* or *FLT3-ITD* AR <0.4 (p<0.0001) and 69±11% or 45±15% for patients with only mutated *WT1* (p<0.0001) or *FLT3-ITD* AR ≥0.4 (p=0.019), respectively (Figure 5(b)). Finally, the co-occurrence of double or triple mutations resulted in a 3-year OS probability of 27±13%, which was significantly lower than the 3-year OS of 79±3% in patients with no mutations or only *NUP98-NSD1* or *FLT3-ITD* AR <0.4 (p<0.0001) and 90±7% or 73±13% in patients with only *WT1* (p=0.0003) or *FLT3-ITD* AR ≥0.4 (p=0.06) mutations, respectively (Figure 5(b)). By multivariate analysis including *WT1* mutation, *FLT3-ITD* AR ≥0.4, core-binding factor aberrations, early bone marrow response to treatment, and stem cell transplantation as covariables, we confirmed that the interaction of these three factors, and not each of the aberrations individually, was a significant predictor of poor prognosis for EFS (p=0.008, HR: 3.88, 95% CI: 1.42 – 10.6) and OS (p=0.042, HR: 3.42, 95% CI: 1.04 – 11.21, Table 3). Importantly, none of the patients with triple mutations survived and the only patients who could be rescued harbored

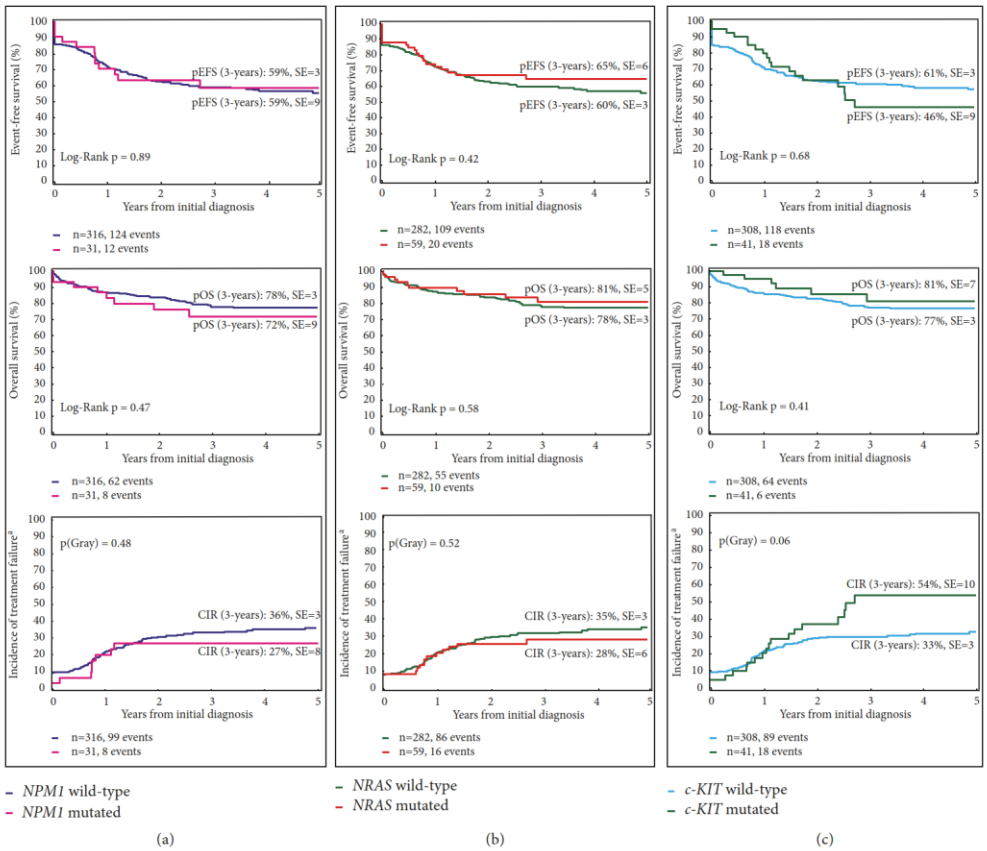


FIGURE 3: *Mutations in NPM1, NRAS, and c-KIT had no impact on survival.* (a) Prognostic impact of mutated *NPM1* on EFS, OS, and CIR. (b) Prognostic impact of mutated *NRAS* on EFS, OS, and CIR. (c) Prognostic impact of *c-KIT* mutation on EFS, OS, and CIR. *NPM1*, *nucleophosmin 1*; *NRAS*, *neuroblastoma RAS viral oncogene homolog*; *c-KIT*, *KIT* protooncogene; pEFS, probability of event-free survival; pOS, probability of overall survival; CIR, cumulative incidence of relapse; SE, standard error; n, number. [a]No response to treatment was considered as the occurrence of an event at time zero.

double *NUP98-NSD1* and *WT1* or *NUP98-NSD1* and *FLT3-ITD* mutations (Figure 1(b)), thus resulting in an OS of 27±13% (Figure 5(b)).

4. Discussion

Treatment of pediatric AML has significantly improved over the past three decades due to the development of intensified first-line treatments, efficient second-line therapies, and optimized supportive care [2, 30]. The success is, at least partly, achieved by more efficient risk group stratification using factors such as somatic mutations and cytogenetic aberrations of AML blasts at diagnosis as well as considering the primary response to treatment to optimize the allocation of patients to standard or enhanced treatment options [1]. In the present study, we analyzed the influence of three parameters, mutations in *WT1* and *FLT3* and the translocation of *NUP98-NSD1*, on the outcome of pediatric patients in the German AML-BFM 2004 and 2012 protocols. Although all three parameters have been established by us and others as important prognostic factors in both pediatric and adult patients [8–14, 20–22], their combined utility to identify high-risk patients likely to experience dismal treatment results has not yet been reported in a contemporary pediatric AML trial.

In a cohort of 237 patients treated within the AML-BFM 2004 and 2012 protocols and with sufficient material for re-analysis, we observed favorable outcomes for 3-year EFS of 61% and 69% and OS of 79% and 90% in patients

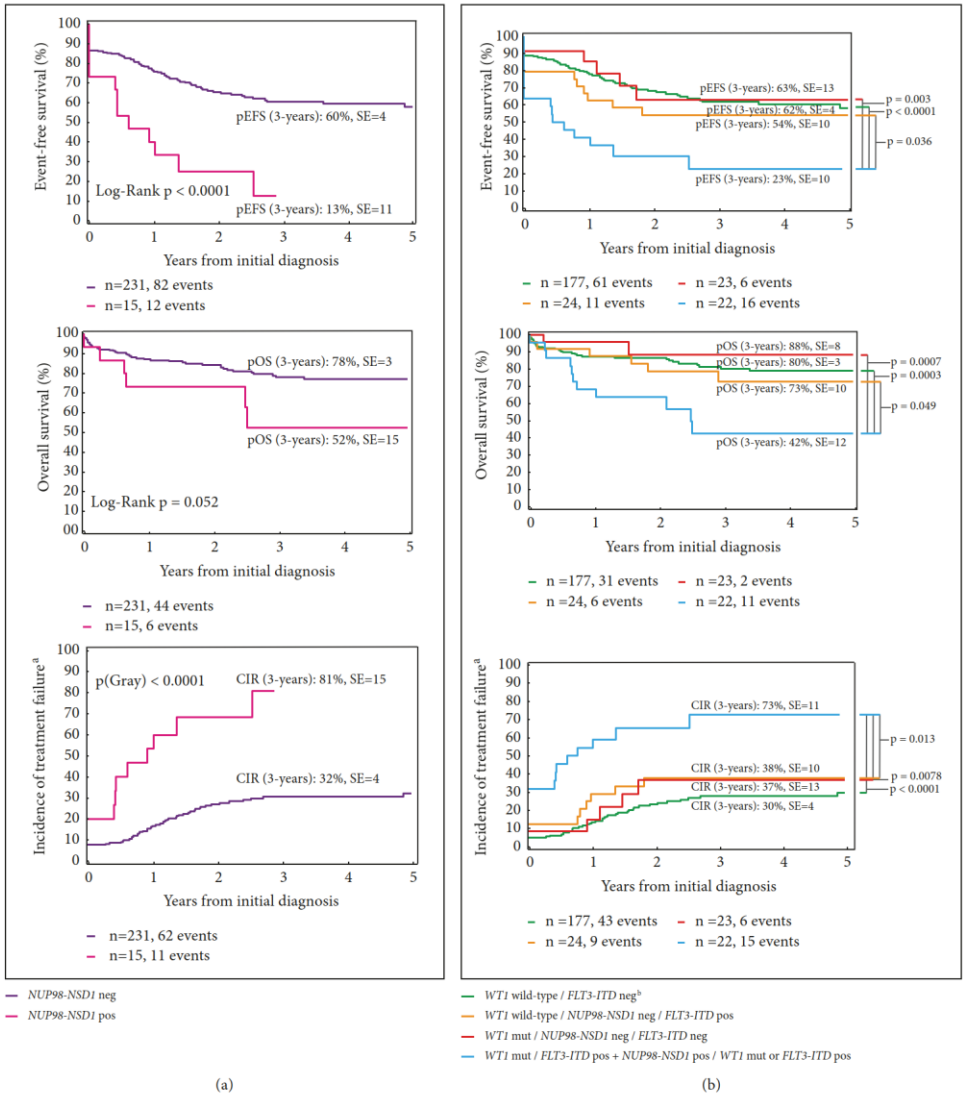


FIGURE 4: *Prognostic significance of NUP98-NSD1 fusion.* (a) *NUP98-NSD1* as single factor predicted poor outcomes. (b) Inclusion of *NUP98-NSD1* as poor prognostic factor with *WT1* mutation and *FLT3-ITD*, predicted poor outcomes for patients harboring all three factors in addition to patients with *NUP98-NSD1* and *WT1* mutation or *FLT3-ITD*. Patients with unknown status of *NUP98-NSD1* fusion were excluded from this analysis. WT1, *Wilms Tumor 1*; *FLT3-ITD*, fms related tyrosine kinase 3-internal tandem duplication; *NUP98-NSD1*, *Nucleoporin-Nuclear Receptor Binding SET Domain Protein 1* fusion gene; pEFS, probability of event-free survival; pOS, probability of overall survival; CIR, cumulative incidence of relapse; SE, standard error; mut, mutated; pos, positive; neg, negative. [a]No response to treatment was considered as the occurrence of an event at time zero. [b]Three patients with *NUP98-NSD1* are included in this group.

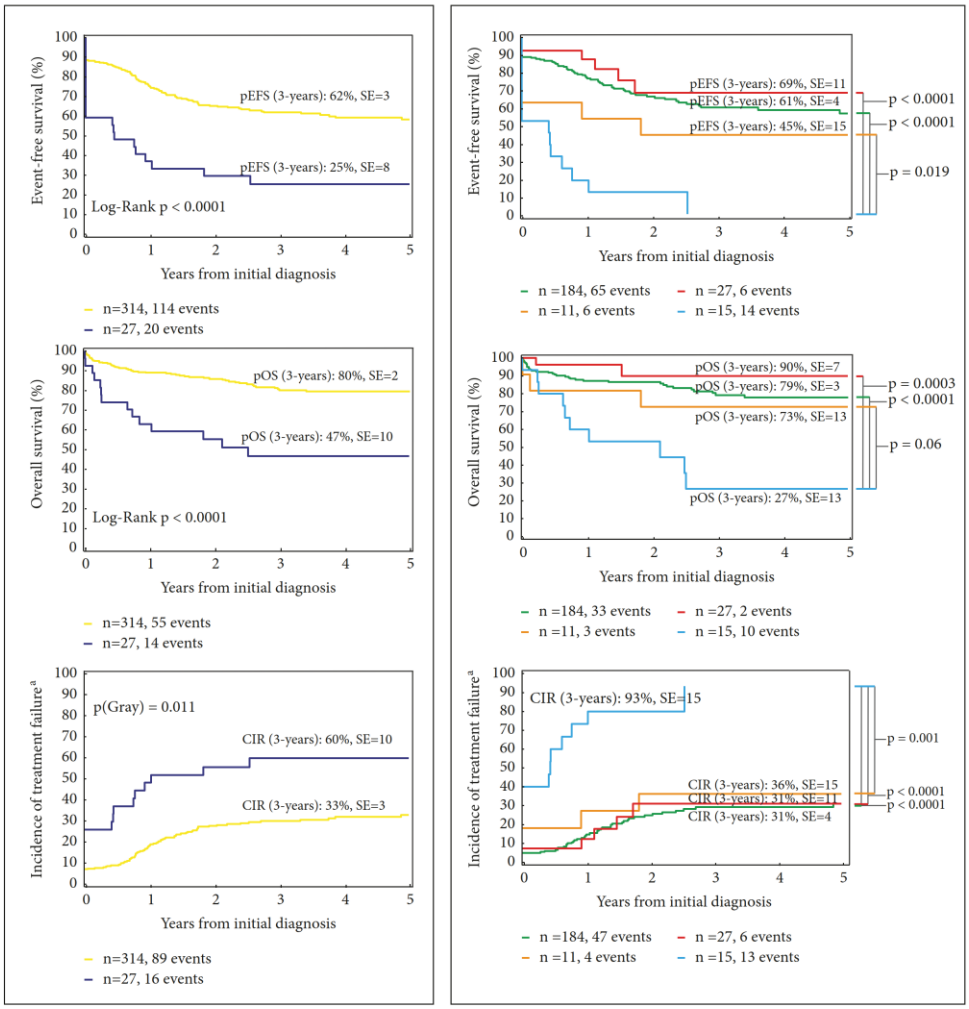


FIGURE 5: *Prognostic significance of mutational burden of FLT3-ITD.* (a) *FLT3-ITD* with an allelic ratio ≥0.4 as a single factor predicted poor outcomes. (b) High mutational burden of *FLT3-ITD* was another predictor of poor prognosis when it occurred with *WT1* and/or *NUP98-NSD1*. Patients with an unknown *FLT3-ITD* AR were excluded from this analysis. *NUP98-NSD1*, *Nucleoporin-Nuclear Receptor Binding SET Domain Protein 1* fusion gene; *FLT3-ITD*, fms related tyrosine kinase 3-internal tandem duplication; pEFS, probability of event-free survival; pOS, probability of overall survival; CIR, cumulative incidence of relapse; AR, allelic ratio; SE, standard error; n, number. [a]No response to treatment was considered as the occurrence of an event at time zero. [b]Three patients with *NUP98-NSD1* are included in this group.

Table 3: Multivariate analysis.

Cox regression analysis - Event-free survival				
Parameters	Hazard ratio	95% confidence interval		p value
		Lower limit	Upper limit	
WT1 mutation	0.79	0.41	1.53	0.479
FLT3-ITD AR ≥ 0.4	1.55	0.69	3.51	0.288
WT1 mutation, *FLT3-ITD* ≥ 0.4 and *NUP98-NSD1* interaction	3.88	1.42	10.66	0.008
t(8;21) and/or inv(16)	0.51	0.27	0.96	0.037
Unsatisfactory early response to treatment[a]	1.31	0.79	2.18	0.294
HSCT[b]	0.25	0.1	0.64	0.004
Cox regression analysis - Overall survival				
WT1 mutation	0.84	0.35	2.06	0.710
FLT3-ITD ≥ 0.4	1.51	0.57	4.02	0.404
WT1 mutation, *FLT3-ITD* ≥ *0.4* and *NUP98-NSD1* interaction	3.42	1.04	11.21	0.042
t(8;21) and/or inv(16)	0.45	0.16	1.31	0.143
Unsatisfactory early response to treatment[a]	1.21	0.61	2.42	0.589
HSCT[b]	1.18	0.51	2.73	0.700

WT1, Wilms tumor 1; *FLT3-ITD, fms related tyrosine kinase 3*-internal tandem duplication; *NUP98-NSD1, Nucleoporin-Nuclear Receptor Binding SET Domain Protein 1* fusion gen; t, translocation; inv, inversion; HSCT, hematopoietic stem cell transplantation.

[a]Unsatisfactory early response to treatment was defined as persistence of >5% blasts in bone marrow at day 15 and/or 28 after treatment. [b]hematopoietic stem cell transplantation events at first complete remission or after no-response to other treatments were included in the multivariate analysis as a time-dependent variable.

without *WT1* mutations or *NUP98-NSD1* fusion or with only one of these factors. Patients with leukemic blasts that were *FLT3-ITD* positive but negative for *WT1* and *NUP98-NSD1* mutations and that had an *FLT3-ITD* AR ≥0.4 still achieved an EFS of 45% and an OS of 73%. Surprisingly, our data therefore suggests that without *WT1* and *NUP98-NSD1* mutations, the negative impact of *FLT3-ITD* even with an AR≥0.4 might not be as severe as previously published [12, 17]. However, all patients positive for at least two of the three risk factors and with an *FLT3-ITD* AR ≥0.4 had events within the first three years and only 27% could be rescued by our salvage therapies. These unfavorable results in our double or triple mutated group unequivocally demonstrate that our current first-line treatment strategies for these patients are still insufficient/inadequate and urgently need improvement.

Of the three risk factors, currently only the *FLT3-ITD* mutation can be specifically targeted with inhibitors [31]. Although the first generations of these drugs only achieved limited and often transient efficacy due to intrinsic and extrinsic adaptations in the AML blasts and/or the environment [31], combination therapies of newer tyrosine kinase inhibitors such as Quizartinib with standard chemotherapy seem to be relatively well tolerated and in initial studies have demonstrated survival improvement in relapsed or refractory AML patients [32–34]. Due to the important role of FLT3 pathway activation in AML, numerous combinations of FLT3 inhibitors with other drugs are currently being tested. Whether these results will also be helpful for the treatment of pediatric AML will need to be carefully determined in future studies, especially considering the clonal heterogeneity of *FLT3-ITD* and the additional survival burden that it causes by increasing drug resistance through clonal evolution or selection and further expansion of resistant AML clones [35, 36]. Nevertheless, it is tempting to speculate that the simple addition of a newer FLT3 inhibitor to our standard therapy might be a feasible, well-tolerated, and effective approach for all patients with blasts that are positive for the *FLT3-ITD* mutation, regardless of the status of alterations in *WT1* or *NUP98*.

The role of WT1 in patients with AML is still controversial [4]. Although *WT1* is overexpressed in the majority of leukemias and can be used as a marker for minimal residual disease and maybe even vaccination attempts, the prognostic and therapeutic relevance of high or absent WT1 expression levels is not unequivocally accepted [37–39]. In contrast, mutations in *WT1* are clearly identified as determinants of poor prognosis and, as we showed here, confer a dismal prognosis especially in combination with *FLT3-ITD* or *NUP98-NSD1* fusion. In the present study, we identified 64 monoallelic *WT1* sequence alterations in exon 7 or exon 9 in the leukemic blasts of 48 patients. The majority of these alterations leads to frameshifts and/or premature terminations codons and thus shortened proteins. These mutant proteins can act in a dominant negative manner [40], which may contribute to a myeloid differentiation block present in AML blasts [41]. However, similar mutations have also been described in the context of Wilms tumors as gain-of-function mutations promoting proliferation [42]. Here, we show a favorable prognosis for patients with single *WT1* mutations, with 26 out of 29 cases reaching continued complete remission (CCR) (Figure 1(b)). Therefore, based on a 3-year EFS of 69% and an OS of 90%, the development of

new treatment approaches is not as urgently needed for these patients with *WT1* mutated blasts that do not harbor *FLT3-ITD* or *NUP98-NSD1* mutations.

Among the 31 different fusion gene partners of *NUP98* identified so far, the *NUP98-NSD1* t(5:11) translocation is the most frequent and present in 4-7% of patients in pediatric AML patients [20–22]. Importantly, the *NUP98* translocations that occur in AML all share the N-terminus of the protein and are thought to initially lead to epigenetic dysregulation of different leukemia-associated genes including *HOXA7*, *HOXA9*, and *HOXA10* in myeloid precursor cells [20]. Additional somatic mutations in other genes occur as secondary events and promote malignant transformation and uncontrolled cell growth [20]. As also shown in our patient data set, these secondary alterations often include activating mutations in *FLT3 (FLT3-ITD)* or truncating mutations *in WT1* [21]. Strikingly, only three patients in our study had a *NUP98-NSD1* translocation without mutations in *FLT3* or *WT1*; two of these patients achieved and remained in first CCR at the end of data acquisition. The third patient had no other genetic risk factors but a very high initial white blood cell count of almost 400,000 cells/μl. Complete remission induction was delayed, and the patient relapsed a year later but was successfully treated by allogeneic stem cell transplantation with a follow-up of 10 years. Therefore, as also described previously [21], our patients with *NUP98*-rearranged blasts with *WT1* and/or *FLT3-ITD* mutations had a poor prognosis, especially in contrast to patients with only *WT1* and *FLT3-ITD* mutations, who could at least partially be rescued by allogeneic transplantation. However, due to the high risk of failure of the first-line treatment, stem cell transplantation already in first CCR seems to be an attractive option for cases of *NUP98*-rearranged AML [21, 22]. Nevertheless, it should be noted that even allogeneic stem cell transplantation is not always effective in improving the treatment outcome in patients with a high probability of treatment failure based on risk stratification. Thus, introducing novel treatment approaches such as the use of small inhibitors, e.g., venetoclax and isadanutlin [43] or cellular therapies with allogeneic NK-cells or engineered T-cells with chimeric antigen receptors (CARs) [44] targeting leukemic blasts harboring *NUP98* rearrangement or *WT1* mutations should be taken into consideration in future clinical studies.

Recent analysis from a collaborative study between the American and Dutch children oncology groups (COG and DCOG) included patients from three clinical COG/DCOG trials and also young adults less than 39 years of age in the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) AML initiative [45]. Analysis of the different cohorts revealed similarly unfavorable outcomes with an EFS of 14-25% and an OS of 15-40% for patients with *FLT3-ITD* and *WT1* mutations and/or the *NUP98-NSD1* translocation [45]. In contrast to our findings however, the authors reported an EFS range of 15-35% in patients with *FLT3-ITD* only, which is lower than that achieved with current protocols, for which an EFS of 45% and an OS of 73% were found for patients with *FLT3-ITD* only. Notably, in the American-Dutch study, patients with co-occurrence of *NPM1* mutations and *FLT3-ITD* (and without *WT1* and *NUP98-NSD1*) were separated from patients with *FLT3-ITD* only and had a slightly increased, albeit probably not statistically significant, survival. Similarly, we have previously observed favorable outcomes for patients with *NPM1* mutations in their AML blasts with normal karyotype and proved this impact was not affected by the presence of *FLT3-ITD* [46]. In the current cohort, five patients were positive for mutations in *FLT3-ITD* and *NPM1* and negative for *WT1* and *NUP98* alterations. At present, four patients with a normal karyotype are still in first CCR, and the fifth patient with a complex karyotype and an *FLT3-ITD* AR >11 experienced early death. In summary, the principle findings of this American-Dutch study and the present study are very similar. However, the treatment outcomes for our patient groups are superior, most likely due to the fact that we included only patients between 0 and 18 years of age treated in Germany according to two contemporary protocols from the AML BFM study group.

5. Conclusion

Despite the fact that our study was partly based on data collected prospectively since 2004 and partly on data assessed *de novo* on stored material by either NGS or PCR, we can safely conclude that co-occurrence of the three factors, mutated *WT1* and *FLT3-ITD and/or NUP98-NSD1* translocation, still defines a subgroup of AML patients with devastating EFS and OS outcome, even with our current treatment protocols. Although the number of pediatric AML patients available for analysis of these three risk factors was limited and therefore not all interesting factors could be assessed in multivariate analysis, it is obvious that patients with double or triple mutations benefitted very little from the improved EFS and OS in our AML-BFM studies in recent years. Thus, for these pediatric patients, new and more targeted approaches are urgently needed for both first- and second-line treatments.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest regarding the current work. Dirk Reinhardt has consulting or advisory roles for Roche, Celgene, Hexal, Pfizer, Novartis, Boehringer and receives research funding from Celgene. Dirk Reinhardt received travel grants from Jazz Pharmaceuticals and Griffols. Naghmeh Niktoreh and Christine von Neuhoff received travel grants from Jazz Pharmaceuticals. The other authors have nothing to declare.

Authors' Contributions

Helmut Hanenberg and Dirk Reinhardt contributed equally. Naghmeh Niktoreh and Christiane Walter collected and assembled data; Martin Zimmermann, Naghmeh Niktoreh,

Christiane Walter, Helmut Hanenberg, and Dirk Reinhardt analyzed and interpreted data; Naghmeh Niktoreh and Helmut Hanenberg wrote the manuscript; and all authors gave final approval of manuscript.

Acknowledgments

We thank all patients and their families who participated in the corresponding AML-BFM protocols. We acknowledge the work of the physicians, nurses, and study personnel involved in patient care and documentation of the study results. We are indebted to the expert technical assistance of Carolin Augsburg, Chantal Daszkowski, Andrea Drothler, Fabienne Hülse, Christel Katerkamp, Lisa Kellerstrass, Denise Kondryn, and Ellen Mahlow. Work in our laboratory was supported, in part, by the Deutsche Krebshilfe e.V., the Deutsche José Carreras Leukämie-Stiftung e.V., the Deutsche Forschungsgemeinschaft e.V., the Essener Elterninitiative zur Unterstützung krebskranker Kinder e. V., and the Verein für krebskranke Kinder Hannover e. V.

References

[1] U. Creutzig, M. M. van den Heuvel-Eibrink, B. Gibson et al., "Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel," *Blood*, vol. 120, no. 16, pp. 3187–3205, 2012.

[2] C. M. Zwaan, E. A. Kolb, D. Reinhardt et al., "Collaborative efforts driving progress in pediatric acute myeloid leukemia," *Journal of Clinical Oncology*, vol. 33, no. 27, pp. 2949–2962, 2015.

[3] E. Toska and S. G. E. Roberts, "Mechanisms of transcriptional regulation by WT1 (Wilms' tumour 1)," *Biochemical Journal*, vol. 461, no. 1, pp. 15–32, 2014.

[4] V. Huff, "Wilms' tumours: about tumour suppressor genes, an oncogene and a chameleon gene," *Nature Reviews Cancer*, vol. 11, no. 2, pp. 111–121, 2011.

[5] C. Owen, J. Fitzgibbon, and P. Paschka, "The clinical relevance of Wilms Tumour 1 (WT1) gene mutations in acute leukaemia," *Hematological Oncology*, vol. 28, no. 1, pp. 13–19, 2009.

[6] P. Paschka, G. Marcucci, A. S. Ruppert et al., "Wilms' tumor 1 gene mutations independently predict poor outcome in adults with cytogenetically normal acute myeloid leukemia: a cancer and leukemia group B study," *Journal of Clinical Oncology*, vol. 26, no. 28, pp. 4595–4602, 2008.

[7] P. Virappane, R. Gale, R. Hills et al., "Mutation of the Wilms' tumor 1 gene is a poor prognostic factor associated with chemotherapy resistance in normal karyotype acute myeloid leukemia: The United Kingdom medical research council adult leukaemia working party," *Journal of Clinical Oncology*, vol. 26, no. 33, pp. 5429–5435, 2008.

[8] H. A. Hou, T. C. Huang, L. I. Lin et al., "WT1 mutation in 470 adult patients with acute myeloid leukemia: stability during disease evolution and implication of its incorporation into a survival scoring system," *Blood*, vol. 115, no. 25, pp. 5222–5231, 2010.

[9] M. Krauth, T. Alpermann, U. Bacher et al., "WT1 mutations are secondary events in AML, show varying frequencies and impact on prognosis between genetic subgroups," *Leukemia*, vol. 29, no. 3, pp. 660–667, 2015.

[10] I. H. Hollink, M. M. van den Heuvel-Eibrink, M. Zimmermann et al., "Clinical relevance of Wilms tumor 1 gene mutations in childhood acute myeloid leukemia," *Blood*, vol. 113, no. 23, pp. 5951–5960, 2009.

[11] P. A. Ho, R. Zeng, T. A. Alonzo et al., "Prevalence and prognostic implications of WT1 mutations in pediatric acute myeloid leukemia (AML): a report from the Children's Oncology Group," *Blood*, vol. 116, no. 5, pp. 702–710, 2010.

[12] S. Meshinchi, T. A. Alonzo, D. L. Stirewalt et al., "Clinical implications of FLT3 mutations in pediatric AML," *Blood*, vol. 108, no. 12, pp. 3654–3661, 2006.

[13] C. M. Zwaan, S. Meshinchi, J. P. Radich et al., "FLT3 internal tandem duplication in 234 children with acute myeloid leukemia: prognostic significance and relation to cellular drug resistance," *Blood*, vol. 102, no. 7, pp. 2387–2394, 2003.

[14] E. Manara, G. Basso, M. Zampini et al., "Characterization of children with FLT3-ITD acute myeloid leukemia: a report from the AIEOP AML-2002 study group," *Leukemia*, vol. 31, no. 1, pp. 18–25, 2017.

[15] V. I. Gaidzik, R. F. Schlenk, S. Moschny et al., " Prognostic impact of *WT1* mutations in cytogenetically normal acute myeloid leukemia: a study of the German-Austrian AML Study Group ," *Blood*, vol. 113, no. 19, pp. 4505–4511, 2009.

[16] A. Renneville, N. Boissel, V. Zurawski et al., "Wilms tumor 1 gene mutations are associated with a higher risk of recurrence in young adults with acute myeloid leukemia," *Cancer*, vol. 115, no. 16, pp. 3719–3727, 2009.

[17] R. F. Schlenk, S. Kayser, L. Bullinger et al., "Differential impact of allelic ratio and insertion site in FLT3-ITD-positive AML with respect to allogeneic transplantation," *Blood*, vol. 124, no. 23, pp. 3441–3449, 2014.

[18] S. B. Liu, H. J. Dong, X. B. Bao et al., "Impact of FLT3-ITD length on prognosis of acute myeloid leukemia," *Haematologica*, 2018.

[19] S. Akiki, S. A. Dyer, D. Grimwade et al., "NUP98-NSD1 fusion in association with FLT3-ITD mutation identifies a prognostically relevant subgroup of pediatric acute myeloid leukemia patients suitable for monitoring by real time quantitative PCR," *Genes, Chromosomes & Cancer*, vol. 52, no. 11, pp. 1053–1064, 2013.

[20] I. H. Hollink, M. M. van den Heuvel-Eibrink, S. T. Arentsen-Peters et al., "NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid leukemia with a distinct HOX gene expression pattern," *Blood*, vol. 118, no. 13, pp. 3645–3656, 2011.

[21] S. Struski, S. Lagarde, P. Bories et al., "NUP98 is rearranged in 3.8% of pediatric AML forming a clinical and molecular homogenous group with a poor prognosis," *Leukemia*, vol. 31, no. 3, pp. 565–572, 2017.

[22] F. Ostronoff, M. Othus, R. B. Gerbing et al., "NUP98/NSD1 and FLT3/ITD coexpression is more prevalent in younger AML patients and leads to induction failure: a COG and SWOG report," *Blood*, vol. 124, no. 15, pp. 2400–2407, 2014.

[23] U. Creutzig, M. Zimmermann, J.-P. Bourquin et al., "Randomized trial comparing liposomal daunorubicin with idarubicin as induction for pediatric acute myeloid leukemia: Results from study AML-BFM 2004," *Blood*, vol. 122, no. 1, pp. 37–43, 2013.

[24] J. Schweitzer, M. Zimmermann, M. Rasche et al., "Improved outcome of pediatric patients with acute megakaryoblastic leukemia in the AML-BFM 04 trial," *Annals of Hematology*, vol. 94, no. 8, pp. 1327–1336, 2015.

[25] C. Bachas, G. J. Schuurhuis, I. H. Hollink et al., "High-frequency type I/II mutational shifts between diagnosis and relapse are associated with outcome in pediatric AML: implications for personalized medicine," *Blood*, vol. 116, no. 15, pp. 2752–2758, 2010.

[26] R. Luthra, K. P. Patel, N. G. Reddy et al., "Next-generation sequencing-based multigene mutational screening for acute myeloid leukemia using MiSeq: applicability for diagnostics and disease monitoring," *Haematologica*, vol. 99, no. 3, pp. 465–473, 2014.

[27] C. Walter, C. Pozzorini, K. Reinhardt et al., "Single-cell whole exome and targeted sequencing in NPM1/FLT3 positive pediatric acute myeloid leukemia," *Pediatric Blood & Cancer*, vol. 65, no. 2, 2018.

[28] U. Creutzig, M. Zimmermann, J. Ritter et al., "Treatment strategies and long-term results in paediatric patients treated in four consecutive AML-BFM trials," *Leukemia*, vol. 19, no. 12, pp. 2030–2042, 2005.

[29] A. Sander, M. Zimmermann, M. Dworzak et al., "Consequent and intensified relapse therapy improved survival in pediatric AML: results of relapse treatment in 379 patients of three consecutive AML-BFM trials," *Leukemia*, vol. 24, no. 8, pp. 1422–1428, 2010.

[30] M. Rasche, M. Zimmermann, L. Borschel et al., "Successes and challenges in the treatment of pediatric acute myeloid leukemia: a retrospective analysis of the AML-BFM trials from 1987 to 2012," *Leukemia*, vol. 32, no. 10, pp. 2167–2177, 2018.

[31] M. Larrosa-Garcia and M. R. Baer, "FLT3 inhibitors in acute myeloid leukemia: current status and future directions," *Molecular Cancer Therapeutics*, vol. 16, no. 6, pp. 991–1001, 2017.

[32] T. M. Cooper, J. Cassar, E. Eckroth et al., "A phase I study of quizartinib combined with chemotherapy in relapsed childhood leukemia: a therapeutic advances in childhood leukemia & lymphoma (TACL) study," *Clinical Cancer Research*, vol. 22, no. 16, pp. 4014–4022, 2016.

[33] J. Cortes, A. E. Perl, H. Döhner et al., "Quizartinib, an FLT3 inhibitor, as monotherapy in patients with relapsed or refractory acute myeloid leukaemia: an open-label, multicentre, single-arm, phase 2 trial," *The Lancet Oncology*, vol. 19, no. 7, pp. 889–903, 2018.

[34] J. E. Cortes, M. S. Tallman, G. J. Schiller et al., "Phase 2b study of 2 dosing regimens of quizartinib monotherapy in FLT3-ITD–mutated, relapsed or refractory AML," *Blood*, vol. 132, no. 6, pp. 598–607, 2018.

[35] L. Y. Shih, C. F. Huang, J. H. Wu et al., "Internal tandem duplication of FLT3 in relapsed acute myeloid leukemia: a comparative analysis of bone marrow samples from 108 adult patients at diagnosis and relapse," *Blood*, vol. 100, no. 7, pp. 2387–2392.

[36] K. Schranz, M. Hubmann, E. Harin et al., "Clonal heterogeneity of FLT3-ITD detected by high-throughput amplicon sequencing correlates with adverse prognosis in acute myeloid leukemia," *Oncotarget*, vol. 9, no. 53, pp. 30128–30145, 2018.

[37] C. Rautenberg, S. Pechtel, B. Hildebrandt et al., "Wilms' tumor 1 gene expression using a standardized european leukemianet-certified assay compared to other methods for detection of minimal residual disease in myelodysplastic syndrome and acute myelogenous leukemia after allogeneic blood stem cell transplantation," *Biology of Blood and Marrow Transplantation*, vol. 24, no. 11, pp. 2337–2343, 2018.

[38] J. Nakata, Y. Nakae, M. Kawakami et al., "Wilms tumour 1 peptide vaccine as a cure-oriented post-chemotherapy strategy for patients with acute myeloid leukaemia at high risk of relapse," *British Journal of Haematology*, vol. 182, no. 2, pp. 287–290, 2018.

[39] L. A. Rein and N. J. Chao, "WT1 vaccination in acute myeloid leukemia: new methods of implementing adoptive immunotherapy," *Expert Opinion on Investigational Drugs*, vol. 23, no. 3, pp. 417–426, 2013.

[40] J. C. Reddy, J. C. Morris, J. Wang et al., "WT1-mediated transcriptional activation is inhibited by dominant negative mutant proteins," *The Journal of Biological Chemistry*, vol. 270, no. 18, pp. 10878–10884, 1995.

[41] S. Sinha, D. Thomas, L. Yu et al., "Mutant WT1 is associated with DNA hypermethylation of PRC2 targets in AML and responds to EZH2 inhibition," *Blood*, vol. 125, no. 2, pp. 316–326, 2015.

[42] M. Busch, H. Schwindt, A. Brandt et al., "Classification of a frameshift/extended and a stop mutation in WT1 as gain-of-function mutations that activate cell cycle genes and promote Wilms tumour cell proliferation," *Human Molecular Genetics*, vol. 23, no. 15, pp. 3958–3974, 2014.

[43] R. Pan, V. Ruvolo, H. Mu et al., "Synthetic lethality of combined Bcl-2 inhibition and p53 activation in AML: mechanisms and superior antileukemic efficacy," *Cancer Cell*, vol. 32, no. 6, pp. 748–760.e6, 2017.

[44] S. Hofmann, M. L. Schubert, L. Wang et al., "Chimeric antigen receptor (CAR) T cell therapy in acute myeloid leukemia (AML)," *Journal of Clinical Medicine*, vol. 8, no. 2, 2019.

[45] H. Bolouri, J. E. Farrar, T. Triche et al., "The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions," *Nature Medicine*, vol. 24, no. 1, pp. 103–112, 2018.

[46] I. H. Hollink, C. M. Zwaan, M. Zimmermann et al., "Favorable prognostic impact of NPM1 gene mutations in childhood acute myeloid leukemia, with emphasis on cytogenetically normal AML," *Leukemia*, vol. 23, no. 2, pp. 262–270, 2009.

6 Discussion

The concept of the hierarchic organization of AML with leukemic stem cells (LSCs) as the origin has first been described already decades ago[20]. During leukemogenesis, LSCs undergo further clonal evolution gaining additional genetic and/or epigenetic aberrations and give rise to different clones of leukemic blasts[35,39,40]. The LSCs are thought to replenish the leukemic blasts constantly leading to both uncontrolled proliferation, as it is seen clinically, and maintenance of the disease[20]. Characteristically for this clonal evolution especially in adult AML is, that preleukemic mutations can already be found in healthy people[47,157]. This results in a predisposition for hematologic malignancies, which is termed clonal hematopoiesis of indeterminate potential (CHIP)[142] which has not yet been described for pediatric patients.

Several approaches have aimed to unravel the genetic complexity of AML and have already increased our knowledge [8,9,19,44,98,135], all of them using bulk samples. Recent studies, especially in the field of adult AML, also accomplished to describe the functional heterogeneity among AML subpopulations by combined single-cell and bulk analysis of the AML blasts[74,81,109,135].

In this thesis, one aim was to establish a workflow using a commercial whole genome amplification (WGA) procedure to analyse the heterogeneity of pediatric AML on single-cell level, thereby better allowing to analyze the clonality and clonal evolution in this disease. To this end, bulk BM cells from three pediatric AML patients were analyzed using targeted NGS (TS) and also whole exome sequencing (WES). Additionally, sorted single BM cells were analysed via WES and the quality and results obtained were compared.

In the second part of the study, we retrospectively analyzed the outcome of a cohort of pediatric AML patients with sufficient material available treated on current protocols of the AML-BFM study group in Germany. We used NGS and quantitative Real-Time PCR (qPCR) to identify mutations in *WT1, FLT3* and other myeloid genes and retrospectively assessed the presence of a specific gene fusion, *NUP98-NSD1,* by qPCR. Using this data, we then investigated the correlation between the presence or absence of genetic aberrations for survival and treatment response/outcome in this cohort of pediatric AML patients.

6.1 Single cell Whole Genome Amplification

Standard methods rely on the sequence information obtained from DNA from multiclonal cell populations. The genetic information of individual cells is often lost and *de novo* mutations in a low percentage of AML blasts such as subclones or LSCs are concealed within the bulk signal in AML samples. Therefore, the need to establish methods that facilitate analysis of the genetic profile of single cells has emerged during the past years.

Almost all human diploid cells contain only ~6 pg of DNA per cell[49]. However, for the analysis of the genome, a few hundred ng to µg of genomic DNA is needed for library preparation for next generation sequencing. Thus, to analyze the genome of single cells, it is mandatory to first amplify the genome before using it in further complex genetic analyses[36]. Several whole genome amplification (WGA) methods are already commercially available, and their use is described in different studies[36,67,108]. WGA usually is either PCR-based or uses isothermal amplification methods[16]. Using degenerate primers, degenerate oligonucleotide-primed PCR (DOP-PCR) allows dense priming and is typically run in two PCR steps; extension of the hybridized oligos and replication of the amplicon[145]. A similar method is MALBAC[160]. However, the sequence of the constant regions of the primers is differently designed and in combination with thermal cycling during the initial reaction enables a quasi-linear amplification[160].

For our study, we chose a well-established assay based on multiple displacement-amplification (MDA), which works with the φ29 DNA polymerase and random hexamer primers in an isothermal reaction. After the hexamer primer annealed to the template DNA, the polymerase initiates replication at multiple sites simultaneously[108]. Complementary DNA strands are displaced und thus new single stranded DNA is generated[108]. Subsequently, priming and strand displacement replication ultimately results in the formation of amplified double stranded DNA[108]. The amplified DNA can subsequently be used for library preparation and NGS approaches. As already described by Hou *et al.* (2015)[67], this MDA-based WGA approach is sufficient to generate sequence reads with a higher and more uniform coverage compared to other WGA-methods. The authors isolated 29 single cells from a cell line and applied seven different commercial WGA kits. Methods used were DOP-PCR, MDA and MALBAC and different sequencing approaches were performed to compare the performance of the different WGA kits[67]. The MDA-based WGA approach showed

a lower duplication ratio, a higher mapping ratio and a higher genome recovery compared to PCR-based methods[67]. Although all WGA strategies tested by Hou *et al.* introduced some amplification bias, the MDA kit showed the highest covered sequencing depth and thus should be best suited for variant calling[67].

Our single cell sequencing results confirmed these findings. The alignment of reads derived from the single-cell DNA was robust and highly specific. Consistently, >98% of reads were mapped to the reference genome, which is very concordant with the findings reported by Hou *et al.* (2015)[67]. The median coverage between single-cell and bulk DNA was comparable. The coverage uniformity was slightly lower in DNA amplified from single cells, which was also described previously[36,67].

Since we were interested in the error rate in variant detection introduced by the WGA-MDA technology in our approach, we also compared the called variants from the single-cell WES data to those identified the bulk DNA WES data. We therefore additionally analyzed the bulk DNA with TS that readily achieves coverage of 1000X or more. The analysis was then restricted to the regions sufficiently covered by all three approaches in order to get a meaningful comparison. From 170 somatic variants detected, 49 were classified as artifacts from PCR errors or unbalanced distribution of PCR products due to uneven amplification. While 42 out of the 121 validated variants were detected in only one or two of the data sets, 79 were present in all three approaches and therefore classified as real variants reflecting the genetic profile of three CD34+CD38- single patient derived cells. We identified eleven out of twelve variants exclusively found in the single cell data set as artifacts since the variant frequencies were below 20% and they were not retraceable in the TS data despite high coverage. Introduction of false positives during the amplification is a well described effect of MDA-WGA[36,67]. Primer extension in WGA is dependent on a DNA polymerase and the MDA based approach used in this study is based on a φ29 polymerase with an error rate between 2×10^{-4} and 4×10^{-6} [43] explaining the detection of false positive variants. On the other hand, also allelic dropout (ADO) and selection based bias can be introduced during WGA and is described to be the main cause for false negatives and variants with allele frequencies lower than expected[68]

It is already known, that the bulk of leukemic blasts consists of several subclones that gain additional mutations[41,44,51,98]. Chemotherapy itself may promote this evolution even more and clones, which are more fit than others, may survive and give rise to relapse[98]. To characterize subclones within the AML blasts therefore is an essential step to understand clonal evolution. Additionally, it might be essential to

identify driver mutations that occur early in leukemogenesis and persist also during the course of the disease. This may then allow more precise disease monitoring and even enable more specific targeting the origin of the AML LSCs[98].

In our study, one synonymous variant only found in a single cell remained after filtering and was classified as a heterogenous SNP. This SNP was also present in the TS data but not in the bulk WES approach. Missing subclones in the bulk data is a problem that can be caused by insufficient coverage and due to the contamination by non-malignant cells. The other way around, we identified some variants in the bulk analyses which were not detectable in the single cell data. The main cause for this discrepancy was the low coverage at the specific regions. Only one missense variant detected in the bulk sample could clearly be excluded from the single cell data since the coverage at the specific region was sufficient. The single cell was therefore identified as belonging to a subclone that did not harbor this variant. By identifying a variant clearly present in the single cell data, present at a low variant frequency in the TS data and absent in the bulk WES data, we concluded that the bulk WES approach is not sufficient to detect somatic variants of rare subclones due to the low coverage that it achieves. Additionally, due to PCR duplicates generated during library preparation, the allele frequencies of variants detected in bulk samples can be influenced and biased[140]. To determine the presence of subclones even in the bulk analysis, the integration of unique molecular identifiers (UMIs) in the library preparation step could be another approach apart from single cell sequencing. During library preparation, each molecule gets an identifying sequence and therefore copies derived from each molecule can be identified in the sequencing data and PCR duplicates and false positives can be removed as already described[63,80,137].

Our study clearly showed that WGA followed by WES can readily be used for genetic analysis of single cells. Coverage achieved by WES in our study is sufficient to detect somatic variants in single cells, as variants will be present in homozygous or heterozygous state in a diploid genome (except for amplified or deleted areas). WES of bulk cells did not add much more information to this, since it is not ideal to analyze multiclonal subpopulations where a higher coverage is needed to detect the variants. For this purpose, the introduction of UMIs could be an appropriate next step to gain the information output of the bulk analysis and reveal the clonal structure. Nevertheless, limitations of WES such as unbalanced coverage and the introduction of biases due to increase of sequencing reads have also to be considered. A higher

coverage compared to WES will readily be achieved by TS. However, the fixed gene content can be a limitation; areas of interest must be defined prior to the analysis. TS nevertheless is a powerful and useful tool for confirmation of single cell WES data. Even small subpopulations can be detected easily due to the high read counts. Since this approach is an expensive research platform, the implementation in diagnostic workflows is still debatable.

The clonal structure of the AML and the importance of subclones in pediatric AML is not yet used for diagnostic approaches and clinical decisions and must be further investigated. Additionally, it has to be further evaluated how many single cells have to be used to differentiate subclones certainly in a bigger analysis. However, WES of WGA single-cell DNA in combination with targeted bulk DNA analysis will be a promising tool to perform research on clonality of pediatric AML at diagnosis, after passage in immunodeficient mice and/or at relapse after therapy.

6.2 Next Generation Sequencing in pediatric AML

The gold-standard for detection of mutations, Sanger sequencing, has gradually been replaced by NGS approaches. At this moment, both time and costs even for genome sequencing are thereby tremendously reduced[34,102,126] and these approaches are already validated for clinical and diagnostic applications[34,90]. In pediatric AML, this technology enables us, as partly conducted in this thesis, to analyze the genetic variation of AML blasts and LSCs even on a single cell but also on a cell population-based level. Additionally, applying NGS for MRD monitoring using somatic mutations seems to be a suitable approach and several proof-of-principle studies have already been published[38,147].

By establishing panels for TS of genes that are frequently mutated in AML, the number of genes analyzed simultaneously have drastically increased and the chance to identify at least one mutation per patient therefore increased as well[24]. Although this development in the diagnostic possibilities harbors a lot of chances to refine the detailed clinical evaluation and may lead to the use of targeted treatment strategies, one challenge for the identification of *'true'* mutations in AML blasts is the presence of variants with unknown significance (VUS)[7]. Several databases already take this into

account by trying to depict population allele frequencies and to enable a filtering procedure for polymorphisms that do not influence the pathogenesis of AML and other diseases[7]. Nevertheless, functional studies are often needed to clarify the role of these VUS and their pathogenicity. Such functional assays can be a paramount tool used to support clinical annotation of VUS[156].

For adult AML patients, the genetic diversity has already been studied in detail using NGS technologies[24,70,94]. However, the mutational spectrum in pediatric patients is different and needs to be considered carefully. In the present study, the introduction of a targeted NGS panel was successfully carried out and helped to identify mutations in a very reliable and sensitive manner. Here, the targeted panel was successfully introduced as a validation tool for variants detected by whole exome sequencing of either WGA single cells or bulk populations.

As an amplicon-based sequencing panel, the TSM panel nevertheless does have some technical limitations as already described by others due to false positive variants introduced by PCR amplification[2,127,129]. Sequencing of GC-rich genes is also a limitation for amplicon-based sequencing with loci with high GC content often being underrepresented or absent[2,127]. Capture-based library preparation may overcome these limitations by the use of hybridization probes that are longer than normal PCR primers and that can tolerate mismatches and enable enrichment of neighboring regions that cannot easily be captured[127].

6.3 Impact of genetic biomarkers in pediatric AML

Identification of biomarkers can be utilized by NGS to identify somatic mutations as described above. Mutations in *WT1* and *FLT3* and also the fusion of *NUP98-NSD1* have already been identified as poor prognostic factors in pediatric and adult AML [58,59,61,66,84,93,100,111,143,161], but their co-occurrence and association to high-risk patients have not yet been reported in a contemporary pediatric AML trial. With a 3-year overall survival (OS) probability of 67±9% achieved in our patients with *FLT3* ITD, we can not confirm the poor prognostic impact described by others[100,131]. However, we could show that co-occurrence of all three or any double mutations resulted in a significantly lower 3-year overall survival (OS) probability of 42±12% compared to 80±8% for patients with none of the mutations or only *NUP98-NSD1* (p=0.0003) and 88±8% or

73±10% for patients with only *WT1* mutation (p=0.0007) or *FLT3* ITD (p=0.049). The 3-year event-free survival (EFS) probability of 23±10% for patients with triple or double mutations was significantly lower compared to the EFS of 62±4% for patients with none of these mutations or only *NUP98-NSD1* (p<0.0001) and the EFS of 63±13% or 54±10% for patients with only *WT1* mutation (p=0.003) or *FLT3* ITD (p=0.036). In other words, devastatingly all patients enrolled experience an event within the first three years and our second line regimens - including stem cell transplantation - largely fail to rescue these patients.

We showed here, that detection of *WT1* mutations and *FLT3* ITD*s* using the targeted NGS approach is robust and sensitive. Both SNVs and InDels were detected and called reliably using our analysis and our NGS results were concordant with the state-of-the-art detection of *FLT3* ITD*s* - PCR amplification followed by capillary electrophoresis.

Although *WT1* mutations were identified as an important prognostic factor and, as shown here, in combination with *FLT3* ITDs *or NUP98-NSD1* are associated with a poor prognosis, the role of *WT1* in AML is still controversial[69]. Although *WT1* overexpression is described in the majority of acute leukemias and can also be used as a MRD marker and even in vaccination studies[65,123],the relevance of high or absent *WT1* expression is not unequivocally accepted[65,121,123]. The *WT1* mutations described in this study are monoallelic mutations mainly located in exon 7. They mostly lead to frameshifts and/or premature termination codons resulting in shortened protein. Although these mutant proteins could act as dominant negative factor and induce a myeloid differentiation block in AML blasts[139], such alterations have also been described as gain-of-function mutations promoting proliferation in the context of Wilms tumors[22]. Therefore, more functional studies exploring the functional consequences of *WT1* mutations are needed to clarify their meaning in leukemogenesis and analyze their potential as targetable markers. Nevertheless, the prognostic impact of *WT1* mutations without *FLT3* ITD or *NUP98-NSD1* seems to be favourable with 19 out of 29 cases reaching continued complete remission (CCR).

6.4 Implication of genetic biomarkers for treatment

Targeted treatment strategies aiming specifically for dysregulated signaling pathways are already being tested and implemented in recent clinical trials in adult AML[107,114,134,149,155]. However, in pediatric AML patients outside of initial phase I/II treatment studies, only *FLT3* ITDs are currently targeted via tyrosine kinase inhibitors (TKIs), albeit without final outcome on a larger cohort being reported yet[87]. Combination therapies of newer TKIs such as Quizartinib with standard chemotherapy seem to be relatively well tolerated in adults and initial studies have demonstrated survival improvement in relapsed or refractory AML patients[27-29]. Here, studies in children and adolescents are forthcoming.

Another genetic marker and possible target for new treatment strategies addressed in the present study is the fusion of *NUP98* and *NSD1*. This particular fusion is the most frequent among the 31 different fusion gene partners of *NUP98* identified so far and occurs in 4-7% of pediatric AML patients[61,111,143]. All *NUP98* fusions occurring in AML share the N-terminus[61]. Interactions of several *NUP98* fusions with the *Lysine Methyltransferase 2A* (*KMT2A*) and Histone-Modifying complexes such as the non-specific lethal (NSL) histone acetyltransferase (HAT) complex have been described recently[158]. As *KMT2A* is frequently involved in pediatric and also adult AML, a current working hypothesis is, that fusion oncoproteins can hijack histone-modifying complexes to drive leukemogenesis[15,25,158]. Inhibiting the function of *KMT2A* or indirect disruption of the KMT2A complex has therefore be proposed as a possible target for AML patients harboring *NUP98* fusions[158]. *NUP98* fusions are also frequently accompanied by additional somatic mutations and, as also shown in our data set, these mutations often include activating mutations in *FLT3* or truncating mutations in *WT1*[143]. A poor prognosis of these combinations was also shown in our data set, especially in combination with mutated *WT1* mutations or *FLT3* ITD.

Since its already known that specific myeloid genes are mutated more frequently in pediatric AML, we also evaluated the prognostic impact of *NRAS*, *NPM1*, *CEBPA* and *KIT* mutations. Despite a relatively small cohort, our current data suggest - in concordance with published results - that mutations in none of these genes are prognostically relevant[9,62]. These mutations as single factors neither significantly increased the chance of relapse or treatment failure nor reduced the probability of 3-year overall survival (OS). Nevertheless, these alterations are common genetic lesions in pediatric AML[19,32,162] and involve signal transduction pathways that have already

been targeted in several studies with ambiguous success[21,146], as reviewed in[75]. Nevertheless, combinations of mutations in *NRAS*, *NPM1* and double mutations in *CEBPA* can still be useful for risk group stratification.

Although not being an independent prognostic factor, mutations in the RAS pathway are hallmarks of the leukemic blasts and therefore offer a targetable aim for treatment strategies. Attempts to directly target NRAS via small molecules has not been successful, e.g. in melanoma patients[46]. Inhibition of farnesyltransferase as a RAS target with tipifarnib on the other hand did show responses in a single phase I trial in AML[79]. However, the mutational status of patients who did respond was unclear[79]. Inhibition of Mitogen-activated protein kinase kinase (MEK) as another downstream mediator of RAS signaling is another possible mechanism to target the RAS pathway. Selumetinib for example was used in a phase II trial in AML, where none of the patients harbouring a *NRAS* mutation in the cohort responded to the drug[72]. However, preclinical data support the administration of MEK inhibitors to treat *NRAS* mutated leukemia[88,91] and other MEK inhibitors are used in recent clinical trials[92,117].

In the future, we will be able to identify more and more genetic and epigenetic alterations in malignant cells. It is tempting to speculate that targeted treatment based on the identification and interaction of these alterations and affected signalling pathways will become clinically important and ultimately will allow to tailor the treatment better to each individual patient. To understand and target the interplay of these genetic alterations will therefore be paramount to address in future clinical trials and treatment protocols.

7 Outlook

To further decipher the underlying genetic mechanisms and improve treatment strategies in both pediatric and adult AML patients, several aspects can be addressed in the future. The workflow established here for single cell analysis - in combination with bulk analysis - should be applied to an expanded cohort of patients and different subpopulations of leukemic blasts e.g. based on phenotypic differences or specific clinical courses.

This approach can also be used in an *in vivo* xenotransplantation mouse model in which the sorted populations are injected into e.g. NOD/scid/*gamma* (NSG)[71,138] or MI(S)TRG mice[42], which lack most of the innate and adaptive immune system. Engraftment potential of defined subgroups even on a single cell level can thus be evaluated. These xenotransplantation models would also be applicable as a test system for specific inhibitors *in vivo*. Leukemic blasts harbouring specific genetic aberrations engrafted in immunodeficient mice could be treated *in vivo* and the efficacy of potential drugs on inhibiting the leukemic cell growth analyzed.

Additionally, the descriptive analysis of the mutational landscape of pediatric AML must be further expanded using additional NGS approaches to get better insights into the genetic basis of the disease and to refine the risk groups. Due to limitations observed by using amplicon-based targeted sequencing, establishment of a custom capture-based panel with a refined gene content would be an important future step. These improved sequencing techniques should be complemented by extensive screening for gene fusions via RNA-based NGS that would gain profound knowledge about cooperating fusion genes. RNA sequencing may reduce the need for cytogenetic analysis using fluorescence-in-situ hybridisation (FISH) which suffers from limitations such as the requirement for an adequate number of mitotic cells, a limited resolution and limitations in detecting cryptic fusions[115].

Epigenetic analyses will additionally help to decipher the genetic diversity of pediatric AML. Important epigenetic regulators are DNA methylation, posttranslational modifications of histones and miRNA regulation[23]. Mutations in genes involved in these processes have already be detected in AML[136] and to further investigate their impact, methylome sequencing and miRNA sequencing via NGS can be applied.

8 References

1. Ailles, L.E., Gerhard, B. & Hogge, D.E. Detection and characterization of primitive malignant and normal progenitors in patients with acute myelogenous leukemia using long-term coculture with supportive feeder layers and cytokines. *Blood* **90**, 2555-2564 (1997).
2. Aird, D., Ross, M.G., Chen, W.S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D.B., Nusbaum, C. & Gnirke, A. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol* **12**, R18 (2011).
3. Almond, L.M., Charalampakis, M., Ford, S.J., Gourevitch, D. & Desai, A. Myeloid Sarcoma: Presentation, Diagnosis, and Treatment. *Clin Lymphoma Myeloma Leuk* **17**, 263-267 (2017).
4. Ambardar, S., Gupta, R., Trakroo, D., Lal, R. & Vakhlu, J. High Throughput Sequencing: An Overview of Sequencing Chemistry. *Indian J Microbiol* **56**, 394-404 (2016).
5. Arber, D.A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M.J., Le Beau, M.M., Bloomfield, C.D., Cazzola, M. & Vardiman, J.W. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* **127**, 2391-2405 (2016).
6. Bacher, U., Schnittger, S. & Haferlach, T. Molecular genetics in acute myeloid leukemia. *Current opinion in oncology* **22**, 646-655 (2010).
7. Bacher, U., Shumilov, E., Flach, J., Porret, N., Joncourt, R., Wiedemann, G., Fiedler, M., Novak, U., Amstutz, U. & Pabst, T. Challenges in the introduction of next-generation sequencing (NGS) for diagnostics of myeloid malignancies into clinical routine use. *Blood Cancer J* **8**, 113 (2018).
8. Balgobind, B.V., Raimondi, S.C., Harbott, J., Zimmermann, M., Alonzo, T.A., Auvrignon, A., Beverloo, H.B., Chang, M., Creutzig, U., Dworzak, M.N., Forestier, E., Gibson, B., Hasle, H., Harrison, C.J., Heerema, N.A., Kaspers, G.J., Leszl, A., Litvinko, N., Nigro, L.L., Morimoto, A., Perot, C., Pieters, R., Reinhardt, D., Rubnitz, J.E., Smith, F.O., Stary, J., Stasevich, I., Strehl, S., Taga, T., Tomizawa, D., Webb, D., Zemanova, Z., Zwaan, C.M. & van den Heuvel-Eibrink, M.M. Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. *Blood* **114**, 2489-2496 (2009).
9. Balgobind, B.V., Hollink, I.H., Arentsen-Peters, S.T., Zimmermann, M., Harbott, J., Beverloo, H.B., von Bergh, A.R., Cloos, J., Kaspers, G.J., de Haas, V., Zemanova, Z., Stary, J., Cayuela, J.M., Baruchel, A., Creutzig, U., Reinhardt, D., Pieters, R., Zwaan, C.M. & van den Heuvel-Eibrink, M.M. Integrative analysis of type-I and type-II aberrations underscores the genetic heterogeneity of pediatric acute myeloid leukemia. *Haematologica* **96**, 1478-1487 (2011).
10. Basecke, J., Cepek, L., Mannhalter, C., Krauter, J., Hildenhagen, S., Brittinger, G., Trumper, L. & Griesinger, F. Transcription of AML1/ETO in bone marrow and cord blood of individuals without acute myelogenous leukemia. *Blood* **100**, 2267-2268 (2002).
11. Beillard, E., Pallisgaard, N., van der Velden, V.H., Bi, W., Dee, R., van der Schoot, E., Delabesse, E., Macintyre, E., Gottardi, E., Saglio, G., Watzinger, F., Lion, T., van Dongen, J.J., Hokland, P. & Gabert, J. Evaluation of candidate control genes for diagnosis and residual disease detection in leukemic patients using 'real-time' quantitative reverse-transcriptase polymerase chain reaction (RQ-PCR) - a Europe against cancer program. *Leukemia* **17**, 2474-2486 (2003).
12. Bennett, J.M., Catovsky, D., Daniel, M.T., Flandrin, G., Galton, D.A., Gralnick, H.R. & Sultan, C. Proposals for the classification of the acute leukaemias.

French-American-British (FAB) co-operative group. *British journal of haematology* **33**, 451-458 (1976).
13. Bennett, J.M., Catovsky, D., Daniel, M.T., Flandrin, G., Galton, D.A., Gralnick, H.R. & Sultan, C. Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. *Ann Intern Med* **103**, 620-625 (1985).
14. Berglund, E.C., Kiialainen, A. & Syvanen, A.C. Next-generation sequencing technologies and applications for human genetic history and forensics. *Investig Genet* **2**, 23 (2011).
15. Bernt, K.M., Zhu, N., Sinha, A.U., Vempati, S., Faber, J., Krivtsov, A.V., Feng, Z., Punt, N., Daigle, A., Bullinger, L., Pollock, R.M., Richon, V.M., Kung, A.L. & Armstrong, S.A. MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer cell* **20**, 66-78 (2011).
16. Blainey, P.C. The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol Rev* **37**, 407-427 (2013).
17. Blair, A., Ailles, L.E., Hogge, D.E. & Sutherland, H.J. Phenotype of primitive AML cells which engraft NOD/SCID mice. *Experimental Hematology* **24**, 558-558 (1996).
18. Blair, A., Hogge, D.E., Ailles, L.E., Lansdorp, P.M. & Sutherland, H.J. Lack of expression of Thy-1 (CD90) on acute myeloid leukemia cells with long-term proliferative ability in vitro and in vivo. *Blood* **89**, 3104-3112 (1997).
19. Bolouri, H., Farrar, J.E., Triche, T., Jr., Ries, R.E., Lim, E.L., Alonzo, T.A., Ma, Y., Moore, R., Mungall, A.J., Marra, M.A., Zhang, J., Ma, X., Liu, Y., Liu, Y., Auvil, J.M.G., Davidsen, T.M., Gesuwan, P., Hermida, L.C., Salhia, B., Capone, S., Ramsingh, G., Zwaan, C.M., Noort, S., Piccolo, S.R., Kolb, E.A., Gamis, A.S., Smith, M.A., Gerhard, D.S. & Meshinchi, S. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nature medicine* **24**, 103-112 (2018).
20. Bonnet, D. & Dick, J.E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature medicine* **3**, 730-737 (1997).
21. Burgess, M.R., Hwang, E., Firestone, A.J., Huang, T., Xu, J., Zuber, J., Bohin, N., Wen, T., Kogan, S.C., Haigis, K.M., Sampath, D., Lowe, S., Shannon, K. & Li, Q. Preclinical efficacy of MEK inhibition in Nras-mutant AML. *Blood* **124**, 3947-3955 (2014).
22. Busch, M., Schwindt, H., Brandt, A., Beier, M., Gorldt, N., Romaniuk, P., Toska, E., Roberts, S., Royer, H.D. & Royer-Pokora, B. Classification of a frameshift/extended and a stop mutation in WT1 as gain-of-function mutations that activate cell cycle genes and promote Wilms tumour cell proliferation. *Hum Mol Genet* **23**, 3958-3974 (2014).
23. Callinan, P.A. & Feinberg, A.P. The emerging science of epigenomics. *Hum Mol Genet* **15 Spec No 1**, R95-101 (2006).
24. Cancer Genome Atlas Research, N., Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A., Hoadley, K., Triche, T.J., Jr., Laird, P.W., Baty, J.D., Fulton, L.L., Fulton, R., Heath, S.E., Kalicki-Veizer, J., Kandoth, C., Klco, J.M., Koboldt, D.C., Kanchi, K.L., Kulkarni, S., Lamprecht, T.L., Larson, D.E., Lin, L., Lu, C., McLellan, M.D., McMichael, J.F., Payton, J., Schmidt, H., Spencer, D.H., Tomasson, M.H., Wallis, J.W., Wartman, L.D., Watson, M.A., Welch, J., Wendl, M.C., Ally, A., Balasundaram, M., Birol, I., Butterfield, Y., Chiu, R., Chu, A., Chuah, E., Chun, H.J., Corbett, R., Dhalla, N., Guin, R., He, A., Hirst, C., Hirst, M., Holt, R.A., Jones, S., Karsan, A., Lee, D., Li, H.I., Marra, M.A., Mayo, M., Moore, R.A., Mungall, K., Parker, J., Pleasance, E., Plettner,

P., Schein, J., Stoll, D., Swanson, L., Tam, A., Thiessen, N., Varhol, R., Wye, N., Zhao, Y., Gabriel, S., Getz, G., Sougnez, C., Zou, L., Leiserson, M.D., Vandin, F., Wu, H.T., Applebaum, F., Baylin, S.B., Akbani, R., Broom, B.M., Chen, K., Motter, T.C., Nguyen, K., Weinstein, J.N., Zhang, N., Ferguson, M.L., Adams, C., Black, A., Bowen, J., Gastier-Foster, J., Grossman, T., Lichtenberg, T., Wise, L., Davidsen, T., Demchok, J.A., Shaw, K.R., Sheth, M., Sofia, H.J., Yang, L., Downing, J.R. & Eley, G. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *The New England journal of medicine* **368**, 2059-2074 (2013).

25. Cao, F., Townsend, E.C., Karatas, H., Xu, J., Li, L., Lee, S., Liu, L., Chen, Y., Ouillette, P., Zhu, J., Hess, J.L., Atadja, P., Lei, M., Qin, Z.S., Malek, S., Wang, S. & Dou, Y. Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. *Mol Cell* **53**, 247-261 (2014).
26. Cheng, J. & Sakamot, K.M. Topics in pediatric leukemia--acute myeloid leukemia. *MedGenMed* **7**, 20 (2005).
27. Cooper, T.M., Cassar, J., Eckroth, E., Malvar, J., Sposto, R., Gaynon, P., Chang, B.H., Gore, L., August, K., Pollard, J.A., DuBois, S.G., Silverman, L.B., Oesterheld, J., Gammon, G., Magoon, D., Annesley, C. & Brown, P.A. A Phase I Study of Quizartinib Combined with Chemotherapy in Relapsed Childhood Leukemia: A Therapeutic Advances in Childhood Leukemia & Lymphoma (TACL) Study. *Clin Cancer Res* **22**, 4014-4022 (2016).
28. Cortes, J., Perl, A.E., Dohner, H., Kantarjian, H., Martinelli, G., Kovacsovics, T., Rousselot, P., Steffen, B., Dombret, H., Estey, E., Strickland, S., Altman, J.K., Baldus, C.D., Burnett, A., Kramer, A., Russell, N., Shah, N.P., Smith, C.C., Wang, E.S., Ifrah, N., Gammon, G., Trone, D., Lazzaretto, D. & Levis, M. Quizartinib, an FLT3 inhibitor, as monotherapy in patients with relapsed or refractory acute myeloid leukaemia: an open-label, multicentre, single-arm, phase 2 trial. *Lancet Oncol* **19**, 889-903 (2018).
29. Cortes, J.E., Tallman, M.S., Schiller, G.J., Trone, D., Gammon, G., Goldberg, S.L., Perl, A.E., Marie, J.P., Martinelli, G., Kantarjian, H.M. & Levis, M.J. Phase 2b study of 2 dosing regimens of quizartinib monotherapy in FLT3-ITD-mutated, relapsed or refractory AML. *Blood* **132**, 598-607 (2018).
30. Coustan-Smith, E., Ribeiro, R.C., Rubnitz, J.E., Razzouk, B.I., Pui, C.H., Pounds, S., Andreansky, M., Behm, F.G., Raimondi, S.C., Shurtleff, S.A., Downing, J.R. & Campana, D. Clinical significance of residual disease during treatment in childhood acute myeloid leukaemia. *British journal of haematology* **123**, 243-252 (2003).
31. Creutzig, U., Zimmermann, M., Ritter, J., Reinhardt, D., Hermann, J., Henze, G., Jurgens, H., Kabisch, H., Reiter, A., Riehm, H., Gadner, H. & Schellong, G. Treatment strategies and long-term results in paediatric patients treated in four consecutive AML-BFM trials. *Leukemia* **19**, 2030-2042 (2005).
32. Creutzig, U., van den Heuvel-Eibrink, M.M., Gibson, B., Dworzak, M.N., Adachi, S., de Bont, E., Harbott, J., Hasle, H., Johnston, D., Kinoshita, A., Lehrnbecher, T., Leverger, G., Mejstrikova, E., Meshinchi, S., Pession, A., Raimondi, S.C., Sung, L., Stary, J., Zwaan, C.M., Kaspers, G.J., Reinhardt, D. & Group, A.M.L.C.o.t.I.B.S. Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. *Blood* **120**, 3187-3205 (2012).
33. Creutzig, U., Zimmermann, M., Reinhardt, D., Rasche, M., von Neuhoff, C., Alpermann, T., Dworzak, M., Perglerova, K., Zemanova, Z., Tchinda, J., Bradtke, J., Thiede, C. & Haferlach, C. Changes in cytogenetics and molecular

genetics in acute myeloid leukemia from childhood to adult age groups. *Cancer* (2016).
34. Dahl, A., Mertes, F., Timmermann, B. & Lehrach, H. The application of massively parallel sequencing technologies in diagnostics. *F1000 Biol Rep* **2**, 59 (2010).
35. de Boer, B., Prick, J., Pruis, M.G., Keane, P., Imperato, M.R., Jaques, J., Brouwers-Vos, A.Z., Hogeling, S.M., Woolthuis, C.M., Nijk, M.T., Diepstra, A., Wandinger, S., Versele, M., Attar, R.M., Cockerill, P.N., Huls, G., Vellenga, E., Mulder, A.B., Bonifer, C. & Schuringa, J.J. Prospective Isolation and Characterization of Genetically and Functionally Distinct AML Subclones. *Cancer cell* **34**, 674-689 e678 (2018).
36. de Bourcy, C.F., De Vlaminck, I., Kanbar, J.N., Wang, J., Gawad, C. & Quake, S.R. A quantitative comparison of single-cell whole genome amplification methods. *PLoS One* **9**, e105585 (2014).
37. De Kouchkovsky, I. & Abdul-Hay, M. 'Acute myeloid leukemia: a comprehensive review and 2016 update'. *Blood Cancer J* **6**, e441 (2016).
38. Delsing Malmberg, E., Johansson Alm, S., Nicklasson, M., Lazarevic, V., Stahlman, S., Samuelsson, T., Lenhoff, S., Asp, J., Ehinger, M., Palmqvist, L., Brune, M. & Fogelstrand, L. Minimal residual disease assessed with deep sequencing of NPM1 mutations predicts relapse after allogeneic stem cell transplant in AML. *Leukemia & lymphoma*, 1-9 (2018).
39. Dick, J.E. Acute myeloid leukemia stem cells. *Ann N Y Acad Sci* **1044**, 1-5 (2005).
40. Dick, J.E. & Lapidot, T. Biology of normal and acute myeloid leukemia stem cells. *Int J Hematol* **82**, 389-396 (2005).
41. Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., McLellan, M.D., McMichael, J.F., Wallis, J.W., Lu, C., Shen, D., Harris, C.C., Dooling, D.J., Fulton, R.S., Fulton, L.L., Chen, K., Schmidt, H., Kalicki-Veizer, J., Magrini, V.J., Cook, L., McGrath, S.D., Vickery, T.L., Wendl, M.C., Heath, S., Watson, M.A., Link, D.C., Tomasson, M.H., Shannon, W.D., Payton, J.E., Kulkarni, S., Westervelt, P., Walter, M.J., Graubert, T.A., Mardis, E.R., Wilson, R.K. & DiPersio, J.F. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506-510 (2012).
42. Ellegast, J.M., Rauch, P.J., Kovtonyuk, L.V., Muller, R., Wagner, U., Saito, Y., Wildner-Verhey van Wijk, N., Fritz, C., Rafiei, A., Lysenko, V., Dudkiewicz, E., Theocharides, A.P., Soldini, D., Goede, J.S., Flavell, R.A. & Manz, M.G. INV(16) and NPM1mut AML engraft human cytokine knock-in mice. *Blood* (2016).
43. Esteban, J.A., Salas, M. & Blanco, L. Fidelity of phi 29 DNA polymerase. Comparison between protein-primed initiation and DNA polymerization. *J Biol Chem* **268**, 2719-2726 (1993).
44. Farrar, J.E., Schuback, H.L., Ries, R.E., Wai, D., Hampton, O.A., Trevino, L.R., Alonzo, T.A., Guidry Auvil, J.M., Davidsen, T.M., Gesuwan, P., Hermida, L., Muzny, D.M., Dewal, N., Rustagi, N., Lewis, L.R., Gamis, A.S., Wheeler, D.A., Smith, M.A., Gerhard, D.S. & Meshinchi, S. Genomic Profiling of Pediatric Acute Myeloid Leukemia Reveals a Changing Mutational Landscape from Disease Diagnosis to Relapse. *Cancer research* **76**, 2197-2205 (2016).
45. Gabert, J., Beillard, E., van der Velden, V.H., Bi, W., Grimwade, D., Pallisgaard, N., Barbany, G., Cazzaniga, G., Cayuela, J.M., Cave, H., Pane, F., Aerts, J.L., De Micheli, D., Thirion, X., Pradel, V., Gonzalez, M., Viehmann, S., Malec, M., Saglio, G. & van Dongen, J.J. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion

gene transcripts for residual disease detection in leukemia - a Europe Against Cancer program. *Leukemia* **17**, 2318-2357 (2003).
46. Gajewski, T.F., Salama, A.K., Niedzwiecki, D., Johnson, J., Linette, G., Bucher, C., Blaskovich, M.A., Sebti, S.M., Haluska, F., Cancer & Leukemia Group, B. Phase II study of the farnesyltransferase inhibitor R115777 in advanced melanoma (CALGB 500104). *J Transl Med* **10**, 246 (2012).
47. Genovese, G., Kahler, A.K., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakhoum, S.F., Chambert, K., Mick, E., Neale, B.M., Fromer, M., Purcell, S.M., Svantesson, O., Landen, M., Hoglund, M., Lehmann, S., Gabriel, S.B., Moran, J.L., Lander, E.S., Sullivan, P.F., Sklar, P., Gronberg, H., Hultman, C.M. & McCarroll, S.A. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *The New England journal of medicine* **371**, 2477-2487 (2014).
48. Gilliland, D.G., Jordan, C.T. & Felix, C.A. The molecular basis of leukemia. *Hematology Am Soc Hematol Educ Program*, 80-97 (2004).
49. Gillooly, J.F., Hein, A. & Damiani, R. Nuclear DNA Content Varies with Cell Size across Human Cell Types. *Cold Spring Harb Perspect Biol* **7**, a019091 (2015).
50. Gonzales, P.R. & Mikhail, F.M. Diagnostic and Prognostic Utility of Fluorescence In situ Hybridization (FISH) Analysis in Acute Myeloid Leukemia. *Curr Hematol Malig Rep* **12**, 568-573 (2017).
51. Greaves, M. Darwin and evolutionary tales in leukemia. The Ham-Wasserman Lecture. *Hematology Am Soc Hematol Educ Program*, 3-12 (2009).
52. Green, E.D., Watson, J.D. & Collins, F.S. Human Genome Project: Twenty-five years of big biology. *Nature* **526**, 29-31 (2015).
53. Grimwade, D., Moorman, A., Hills, R., Wheatley, K., Walker, H., Harrison, G., Harrison, C., Goldstone, A., Burnett, A. & Party, N.A.L.W. Impact of karyotype on treatment outcome in acute myeloid leukemia. *Ann Hematol* **83 Suppl 1**, S45-48 (2004).
54. Grimwade, D. & Hills, R.K. Independent prognostic factors for AML outcome. *Hematology Am Soc Hematol Educ Program*, 385-395 (2009).
55. Group, M.-A.-B.S., Langebrake, C., Creutzig, U., Dworzak, M., Hrusak, O., Mejstrikova, E., Griesinger, F., Zimmermann, M. & Reinhardt, D. Residual disease monitoring in childhood acute myeloid leukemia by multiparameter flow cytometry: the MRD-AML-BFM Study Group. *J Clin Oncol* **24**, 3686-3692 (2006).
56. Guan, Y., Gerhard, B. & Hogge, D.E. Detection, isolation, and stimulation of quiescent primitive leukemic progenitor cells from patients with acute myeloid leukemia (AML). *Blood* **101**, 3142-3149 (2003).
57. Ho, P.A., Alonzo, T.A., Gerbing, R.B., Pollard, J., Stirewalt, D.L., Hurwitz, C., Heerema, N.A., Hirsch, B., Raimondi, S.C., Lange, B., Franklin, J.L., Radich, J.P. & Meshinchi, S. Prevalence and prognostic implications of CEBPA mutations in pediatric acute myeloid leukemia (AML): a report from the Children's Oncology Group. *Blood* **113**, 6558-6566 (2009).
58. Ho, P.A., Zeng, R., Alonzo, T.A., Gerbing, R.B., Miller, K.L., Pollard, J.A., Stirewalt, D.L., Heerema, N.A., Raimondi, S.C., Hirsch, B., Franklin, J.L., Lange, B. & Meshinchi, S. Prevalence and prognostic implications of WT1 mutations in pediatric acute myeloid leukemia (AML): a report from the Children's Oncology Group. *Blood* **116**, 702-710 (2010).
59. Hollink, I.H., van den Heuvel-Eibrink, M.M., Zimmermann, M., Balgobind, B.V., Arentsen-Peters, S.T., Alders, M., Willasch, A., Kaspers, G.J., Trka, J., Baruchel, A., de Graaf, S.S., Creutzig, U., Pieters, R., Reinhardt, D. & Zwaan, C.M. Clinical relevance of Wilms tumor 1 gene mutations in childhood acute myeloid leukemia. *Blood* **113**, 5951-5960 (2009).

60. Hollink, I.H., Zwaan, C.M., Zimmermann, M., Arentsen-Peters, T.C., Pieters, R., Cloos, J., Kaspers, G.J., de Graaf, S.S., Harbott, J., Creutzig, U., Reinhardt, D., van den Heuvel-Eibrink, M.M. & Thiede, C. Favorable prognostic impact of NPM1 gene mutations in childhood acute myeloid leukemia, with emphasis on cytogenetically normal AML. *Leukemia* **23**, 262-270 (2009).
61. Hollink, I.H., van den Heuvel-Eibrink, M.M., Arentsen-Peters, S.T., Pratcorona, M., Abbas, S., Kuipers, J.E., van Galen, J.F., Beverloo, H.B., Sonneveld, E., Kaspers, G.J., Trka, J., Baruchel, A., Zimmermann, M., Creutzig, U., Reinhardt, D., Pieters, R., Valk, P.J. & Zwaan, C.M. NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid leukemia with a distinct HOX gene expression pattern. *Blood* **118**, 3645-3656 (2011).
62. Hollink, I.H., van den Heuvel-Eibrink, M.M., Arentsen-Peters, S.T., Zimmermann, M., Peeters, J.K., Valk, P.J., Balgobind, B.V., Sonneveld, E., Kaspers, G.J., de Bont, E.S., Trka, J., Baruchel, A., Creutzig, U., Pieters, R., Reinhardt, D. & Zwaan, C.M. Characterization of CEBPA mutations and promoter hypermethylation in pediatric acute myeloid leukemia. *Haematologica* **96**, 384-392 (2011).
63. Hong, J. & Gresham, D. Incorporation of unique molecular identifiers in TruSeq adapters improves the accuracy of quantitative sequencing. *Biotechniques* **63**, 221-226 (2017).
64. Hosen, N., Park, C.Y., Tatsumi, N., Oji, Y., Sugiyama, H., Gramatzki, M., Krensky, A.M. & Weissman, I.L. CD96 is a leukemic stem cell-specific marker in human acute myeloid leukemia. *Proc Natl Acad Sci U S A* **104**, 11008-11013 (2007).
65. Hosen, N., Maeda, T., Hashii, Y., Tsuboi, A., Nishida, S., Nakata, J., Oji, Y., Oka, Y. & Sugiyama, H. Wilms tumor 1 peptide vaccination after hematopoietic stem cell transplant in leukemia patients. *Stem Cell Investig* **3**, 90 (2016).
66. Hou, H.A., Huang, T.C., Lin, L.I., Liu, C.Y., Chen, C.Y., Chou, W.C., Tang, J.L., Tseng, M.H., Huang, C.F., Chiang, Y.C., Lee, F.Y., Liu, M.C., Yao, M., Huang, S.Y., Ko, B.S., Hsu, S.C., Wu, S.J., Tsay, W., Chen, Y.C. & Tien, H.F. WT1 mutation in 470 adult patients with acute myeloid leukemia: stability during disease evolution and implication of its incorporation into a survival scoring system. *Blood* **115**, 5222-5231 (2010).
67. Hou, Y., Wu, K., Shi, X., Li, F., Song, L., Wu, H., Dean, M., Li, G., Tsang, S., Jiang, R., Zhang, X., Li, B., Liu, G., Bedekar, N., Lu, N., Xie, G., Liang, H., Chang, L., Wang, T., Chen, J., Li, Y., Zhang, X., Yang, H., Xu, X., Wang, L. & Wang, J. Comparison of variations detection between whole-genome amplification methods used in single-cell resequencing. *Gigascience* **4**, 37 (2015).
68. Huang, L., Ma, F., Chapman, A., Lu, S. & Xie, X.S. Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. *Annu Rev Genomics Hum Genet* **16**, 79-102 (2015).
69. Huff, V. Wilms' tumours: about tumour suppressor genes, an oncogene and a chameleon gene. *Nat Rev Cancer* **11**, 111-121 (2011).
70. Hussaini, M.O., Mirza, A.S., Komrokji, R., Lancet, J., Padron, E. & Song, J. Genetic Landscape of Acute Myeloid Leukemia Interrogated by Next-generation Sequencing: A Large Cancer Center Experience. *Cancer Genomics Proteomics* **15**, 121-126 (2018).
71. Ishikawa, F., Yasukawa, M., Lyons, B., Yoshida, S., Miyamoto, T., Yoshimoto, G., Watanabe, T., Akashi, K., Shultz, L.D. & Harada, M. Development of functional human blood and immune systems in NOD/SCID/IL2 receptor {gamma} chain(null) mice. *Blood* **106**, 1565-1573 (2005).

72. Jain, N., Curran, E., Iyengar, N.M., Diaz-Flores, E., Kunnavakkam, R., Popplewell, L., Kirschbaum, M.H., Karrison, T., Erba, H.P., Green, M., Poire, X., Koval, G., Shannon, K., Reddy, P.L., Joseph, L., Atallah, E.L., Dy, P., Thomas, S.P., Smith, S.E., Doyle, L.A., Stadler, W.M., Larson, R.A., Stock, W. & Odenike, O. Phase II study of the oral MEK inhibitor selumetinib in advanced acute myelogenous leukemia: a University of Chicago phase II consortium trial. *Clin Cancer Res* **20**, 490-498 (2014).
73. Jamieson, C.H., Weissman, I.L. & Passegue, E. Chronic versus acute myelogenous leukemia: a question of self-renewal. *Cancer cell* **6**, 531-533 (2004).
74. Jan, M., Snyder, T.M., Corces-Zimmerman, M.R., Vyas, P., Weissman, I.L., Quake, S.R. & Majeti, R. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Science translational medicine* **4**, 149ra118 (2012).
75. Johnson, D.B., Smalley, K.S. & Sosman, J.A. Molecular pathways: targeting NRAS in melanoma and acute myelogenous leukemia. *Clin Cancer Res* **20**, 4186-4192 (2014).
76. Jongen-Lavrencic, M., Grob, T., Hanekamp, D., Kavelaars, F.G., Al Hinai, A., Zeilemaker, A., Erpelinck-Verschueren, C.A.J., Gradowska, P.L., Meijer, R., Cloos, J., Biemond, B.J., Graux, C., van Marwijk Kooy, M., Manz, M.G., Pabst, T., Passweg, J.R., Havelange, V., Ossenkoppele, G.J., Sanders, M.A., Schuurhuis, G.J., Lowenberg, B. & Valk, P.J.M. Molecular Minimal Residual Disease in Acute Myeloid Leukemia. *The New England journal of medicine* **378**, 1189-1199 (2018).
77. Jordan, C.T. The leukemic stem cell. *Best practice & research. Clinical haematology* **20**, 13-18 (2007).
78. Kelly, L.M. & Gilliland, D.G. Genetics of myeloid leukemias. *Annu Rev Genomics Hum Genet* **3**, 179-198 (2002).
79. Kirschbaum, M.H., Synold, T., Stein, A.S., Tuscano, J., Zain, J.M., Popplewell, L., Karanes, C., O'Donnell, M.R., Pulone, B., Rincon, A., Wright, J., Frankel, P., Forman, S.J. & Newman, E.M. A phase 1 trial dose-escalation study of tipifarnib on a week-on, week-off schedule in relapsed, refractory or high-risk myeloid leukemia. *Leukemia* **25**, 1543-1547 (2011).
80. Kivioja, T., Vaharautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S. & Taipale, J. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods* **9**, 72-74 (2011).
81. Klco, J.M., Spencer, D.H., Miller, C.A., Griffith, M., Lamprecht, T.L., O'Laughlin, M., Fronick, C., Magrini, V., Demeter, R.T., Fulton, R.S., Eades, W.C., Link, D.C., Graubert, T.A., Walter, M.J., Mardis, E.R., Dipersio, J.F., Wilson, R.K. & Ley, T.J. Functional heterogeneity of genetically defined subclones in acute myeloid leukemia. *Cancer Cell* **25**, 379-392 (2014).
82. Kohlmann, A., Bacher, U., Schnittger, S. & Haferlach, T. Perspective on how to approach molecular diagnostics in acute myeloid leukemia and myelodysplastic syndromes in the era of next-generation sequencing. *Leukemia & lymphoma* **55**, 1725-1734 (2014).
83. Krause, D.S. & Scadden, D.T. A hostel for the hostile: the bone marrow niche in hematologic neoplasms. *Haematologica* **100**, 1376-1387 (2015).
84. Krauth, M.T., Alpermann, T., Bacher, U., Eder, C., Dicker, F., Ulke, M., Kuznia, S., Nadarajah, N., Kern, W., Haferlach, C., Haferlach, T. & Schnittger, S. WT1 mutations are secondary events in AML, show varying frequencies and impact on prognosis between genetic subgroups. *Leukemia* **29**, 660-667 (2015).

85. Kreso, A. & Dick, J.E. Evolution of the Cancer Stem Cell Model. *Cell Stem Cell* **14**, 275-291 (2014).
86. Langebrake, C., Brinkmann, I., Teigler-Schlegel, A., Creutzig, U., Griesinger, F., Puhlmann, U. & Reinhardt, D. Immunophenotypic differences between diagnosis and relapse in childhood AML: Implications for MRD monitoring. *Cytometry B Clin Cytom* **63**, 1-9 (2005).
87. Larrosa-Garcia, M. & Baer, M.R. FLT3 Inhibitors in Acute Myeloid Leukemia: Current Status and Future Directions. *Mol Cancer Ther* **16**, 991-1001 (2017).
88. Lauchle, J.O., Kim, D., Le, D.T., Akagi, K., Crone, M., Krisman, K., Warner, K., Bonifas, J.M., Li, Q., Coakley, K.M., Diaz-Flores, E., Gorman, M., Przybranowski, S., Tran, M., Kogan, S.C., Roose, J.P., Copeland, N.G., Jenkins, N.A., Parada, L., Wolff, L., Sebolt-Leopold, J. & Shannon, K. Response and resistance to MEK inhibition in leukaemias initiated by hyperactive Ras. *Nature* **461**, 411-414 (2009).
89. Li, S., Mason, C.E. & Melnick, A. Genetic and epigenetic heterogeneity in acute myeloid leukemia. *Curr Opin Genet Dev* **36**, 100-106 (2016).
90. Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J. & Pallen, M.J. Performance comparison of benchtop high-throughput sequencing platforms. *Nature biotechnology* **30**, 434-439 (2012).
91. Lyubynska, N., Gorman, M.F., Lauchle, J.O., Hong, W.X., Akutagawa, J.K., Shannon, K. & Braun, B.S. A MEK inhibitor abrogates myeloproliferative disease in Kras mutant mice. *Science translational medicine* **3**, 76ra27 (2011).
92. Maiti, A., Naqvi, K., Kadia, T.M., Borthakur, G., Takahashi, K., Bose, P., Daver, N.G., Patel, A., Alvarado, Y., Ohanian, M., DiNardo, C.D., Cortes, J.E., Jabbour, E.J., Garcia-Manero, G., Kantarjian, H.M. & Ravandi, F. Phase II Trial of MEK Inhibitor Binimetinib (MEK162) in RAS-mutant Acute Myeloid Leukemia. *Clin Lymphoma Myeloma Leuk* **19**, 142-148 e141 (2019).
93. Manara, E., Basso, G., Zampini, M., Buldini, B., Tregnago, C., Rondelli, R., Masetti, R., Bisio, V., Frison, M., Polato, K., Cazzaniga, G., Menna, G., Fagioli, F., Merli, P., Biondi, A., Pession, A., Locatelli, F. & Pigazzi, M. Characterization of children with FLT3-ITD acute myeloid leukemia: a report from the AIEOP AML-2002 study group. *Leukemia* (2016).
94. Mardis, E.R., Ding, L., Dooling, D.J., Larson, D.E., McLellan, M.D., Chen, K., Koboldt, D.C., Fulton, R.S., Delehaunty, K.D., McGrath, S.D., Fulton, L.A., Locke, D.P., Magrini, V.J., Abbott, R.M., Vickery, T.L., Reed, J.S., Robinson, J.S., Wylie, T., Smith, S.M., Carmichael, L., Eldred, J.M., Harris, C.C., Walker, J., Peck, J.B., Du, F., Dukes, A.F., Sanderson, G.E., Brummett, A.M., Clark, E., McMichael, J.F., Meyer, R.J., Schindler, J.K., Pohl, C.S., Wallis, J.W., Shi, X., Lin, L., Schmidt, H., Tang, Y., Haipek, C., Wiechert, M.E., Ivy, J.V., Kalicki, J., Elliott, G., Ries, R.E., Payton, J.E., Westervelt, P., Tomasson, M.H., Watson, M.A., Baty, J., Heath, S., Shannon, W.D., Nagarajan, R., Link, D.C., Walter, M.J., Graubert, T.A., DiPersio, J.F., Wilson, R.K. & Ley, T.J. Recurring mutations found by sequencing an acute myeloid leukemia genome. *The New England journal of medicine* **361**, 1058-1066 (2009).
95. Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A.,

Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F. & Rothberg, J.M. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376-380 (2005).
96. Martens, A.C., van Bekkum, D.W. & Hagenbeek, A. Minimal residual disease in leukemia: studies in an animal model for acute myelocytic leukemia (BNML). *Int J Cell Cloning* **8**, 27-38 (1990).
97. Martignoles, J.A., Delhommeau, F. & Hirsch, P. Genetic Hierarchy of Acute Myeloid Leukemia: From Clonal Hematopoiesis to Molecular Residual Disease. *Int J Mol Sci* **19**(2018).
98. Masetti, R., Castelli, I., Astolfi, A., Bertuccio, S.N., Indio, V., Togni, M., Belotti, T., Serravalle, S., Tarantino, G., Zecca, M., Pigazzi, M., Basso, G., Pession, A. & Locatelli, F. Genomic complexity and dynamics of clonal evolution in childhood acute myeloid leukemia studied with whole-exome sequencing. *Oncotarget* (2016).
99. Medyouf, H. The microenvironment in human myeloid malignancies: emerging concepts and therapeutic implications. *Blood* (2017).
100. Meshinchi, S., Alonzo, T.A., Stirewalt, D.L., Zwaan, M., Zimmerman, M., Reinhardt, D., Kaspers, G.J., Heerema, N.A., Gerbing, R., Lange, B.J. & Radich, J.P. Clinical implications of FLT3 mutations in pediatric AML. *Blood* **108**, 3654-3661 (2006).
101. Meshinchi, S. & Arceci, R.J. Prognostic factors and risk-based therapy in pediatric acute myeloid leukemia. *Oncologist* **12**, 341-355 (2007).
102. Metzker, M.L. Sequencing technologies - the next generation. *Nat Rev Genet* **11**, 31-46 (2010).
103. Monney, L., Sabatos, C.A., Gaglia, J.L., Ryu, A., Waldner, H., Chernova, T., Manning, S., Greenfield, E.A., Coyle, A.J., Sobel, R.A., Freeman, G.J. & Kuchroo, V.K. Th1-specific cell surface protein Tim-3 regulates macrophage activation and severity of an autoimmune disease. *Nature* **415**, 536-541 (2002).
104. Morganti, S., Tarantino, P., Ferraro, E., D'Amico, P., Viale, G., Trapani, D., Duso, B.A. & Curigliano, G. Complexity of genome sequencing and reporting: Next generation sequencing (NGS) technologies and implementation of precision medicine in real life. *Crit Rev Oncol Hematol* **133**, 171-182 (2019).
105. Morrison, S.J. & Scadden, D.T. The bone marrow niche for haematopoietic stem cells. *Nature* **505**, 327-334 (2014).
106. Murati, A., Brecqueville, M., Devillier, R., Mozziconacci, M.J., Gelsi-Boyer, V. & Birnbaum, D. Myeloid malignancies: mutations, models and management. *BMC cancer* **12**, 304 (2012).
107. Naqvi, K., Konopleva, M. & Ravandi, F. Targeted therapies in Acute Myeloid Leukemia: a focus on FLT-3 inhibitors and ABT199. *Expert Rev Hematol* **10**, 863-874 (2017).
108. Nelson, J.R. Random-primed, Phi29 DNA polymerase-based whole genome amplification. *Curr Protoc Mol Biol* **105**, Unit 15 13 (2014).
109. Niemoller, C., Renz, N., Bleul, S., Blagitko-Dorfs, N., Greil, C., Yoshida, K., Pfeifer, D., Follo, M., Duyster, J., Claus, R., Ogawa, S., Lubbert, M. & Becker, H. Single cell genotyping of exome sequencing-identified mutations to characterize the clonal composition and evolution of inv(16) AML in a CBL mutated clonal hematopoiesis. *Leukemia research* **47**, 41-46 (2016).
110. Orkin, S.H. & Zon, L.I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631-644 (2008).
111. Ostronoff, F., Othus, M., Gerbing, R.B., Loken, M.R., Raimondi, S.C., Hirsch, B.A., Lange, B.J., Petersdorf, S., Radich, J., Appelbaum, F.R., Gamis, A.S., Alonzo, T.A. & Meshinchi, S. NUP98/NSD1 and FLT3/ITD coexpression is more

prevalent in younger AML patients and leads to induction failure: a COG and SWOG report. *Blood* **124**, 2400-2407 (2014).
112. Paietta, E. Minimal Residual Disease in AML: Why Has It Lagged Behind Pediatric ALL? *Clin Lymphoma Myeloma Leuk* **15 Suppl**, S2-6 (2015).
113. Patkar, N., Kodgule, R., Kakirde, C., Raval, G., Bhanshe, P., Joshi, S., Chaudhary, S., Badrinath, Y., Ghoghale, S., Kadechkar, S., Khizer, S.H., Kannan, S., Shetty, D., Gokarn, A., Punatkar, S., Jain, H., Bagal, B., Menon, H., Sengar, M., Khattry, N., Tembhare, P., Subramanian, P. & Gujral, S. Clinical impact of measurable residual disease monitoring by ultradeep next generation sequencing in NPM1 mutated acute myeloid leukemia. *Oncotarget* **9**, 36613-36624 (2018).
114. Perl, A.E. The role of targeted therapy in the management of patients with AML. *Blood Adv* **1**, 2281-2294 (2017).
115. Piazza, R., Cecchetti, C., Pirola, A., Donandoni, C., Fontana, D., Mezzatesta, C., Magistroni, V., Bianchi, B., Borin, L., Fumagalli, M. & Gambacorti-Passerini, C. RNA-seq is a valuable complement of conventional diagnostic tools in newly diagnosed AML patients. *Am J Hematol* **90**, E227-228 (2015).
116. Pui, C.H., Schrappe, M., Ribeiro, R.C. & Niemeyer, C.M. Childhood and adolescent lymphoid and myeloid leukemia. *Hematology Am Soc Hematol Educ Program*, 118-145 (2004).
117. Ragon, B.K., Odenike, O., Baer, M.R., Stock, W., Borthakur, G., Patel, K., Han, L., Chen, H., Ma, H., Joseph, L., Zhao, Y., Baggerly, K., Konopleva, M. & Jain, N. Oral MEK 1/2 Inhibitor Trametinib in Combination With AKT Inhibitor GSK2141795 in Patients With Acute Myeloid Leukemia With RAS Mutations: A Phase II Study. *Clin Lymphoma Myeloma Leuk* (2019).
118. Raimondi, S.C., Chang, M.N., Ravindranath, Y., Behm, F.G., Gresik, M.V., Steuber, C.P., Weinstein, H.J. & Carroll, A.J. Chromosomal abnormalities in 478 children with acute myeloid leukemia: clinical characteristics and treatment outcome in a cooperative pediatric oncology group study-POG 8821. *Blood* **94**, 3707-3716 (1999).
119. Rasche, M., Zimmermann, M., Borschel, L., Bourquin, J.P., Dworzak, M., Klingebiel, T., Lehrnbecher, T., Creutzig, U., Klusmann, J.H. & Reinhardt, D. Successes and challenges in the treatment of pediatric acute myeloid leukemia: a retrospective analysis of the AML-BFM trials from 1987 to 2012. *Leukemia* (2018).
120. Rau, R. & Brown, P. Nucleophosmin (NPM1) mutations in adult and childhood acute myeloid leukaemia: towards definition of a new leukaemia entity. *Hematol Oncol* **27**, 171-181 (2009).
121. Rautenberg, C., Pechtel, S., Hildebrandt, B., Betz, B., Dienst, A., Nachtkamp, K., Kondakci, M., Geyh, S., Wieczorek, D., Haas, R., Germing, U., Kobbe, G. & Schroeder, T. Wilms' Tumor 1 Gene Expression Using a Standardized European LeukemiaNet-Certified Assay Compared to Other Methods for Detection of Minimal Residual Disease in Myelodysplastic Syndrome and Acute Myelogenous Leukemia after Allogeneic Blood Stem Cell Transplantation. *Biol Blood Marrow Transplant* (2018).
122. Reilly, J.T. Pathogenesis of acute myeloid leukaemia and inv(16)(p13;q22): a paradigm for understanding leukaemogenesis? *British journal of haematology* **128**, 18-34 (2005).
123. Rein, L.A. & Chao, N.J. WT1 vaccination in acute myeloid leukemia: new methods of implementing adoptive immunotherapy. *Expert Opin Investig Drugs* **23**, 417-426 (2014).

124. Reinhardt, D., Dworzak, M., Klusmann, J.H., Rasche, M., Creutzig, U. Recommendations for diagnostics, therapy and follow-up care of children and adolescents with Acute Myeloid Leukemia (AML). (2019).
125. Riether, C., Schurch, C.M. & Ochsenbein, A.F. Regulation of hematopoietic and leukemic stem cells by the immune system. *Cell Death Differ* **22**, 187-198 (2015).
126. Roukos, D.H. Next-generation sequencing and epigenome technologies: potential medical applications. *Expert Rev Med Devices* **7**, 723-726 (2010).
127. Samorodnitsky, E., Jewell, B.M., Hagopian, R., Miya, J., Wing, M.R., Lyon, E., Damodaran, S., Bhatt, D., Reeser, J.W., Datta, J. & Roychowdhury, S. Evaluation of Hybridization Capture Versus Amplicon-Based Methods for Whole-Exome Sequencing. *Hum Mutat* **36**, 903-914 (2015).
128. Schepers, K., Campbell, T.B. & Passegue, E. Normal and leukemic stem cell niches: insights and therapeutic opportunities. *Cell Stem Cell* **16**, 254-267 (2015).
129. Schirmer, M., Ijaz, U.Z., D'Amore, R., Hall, N., Sloan, W.T. & Quince, C. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic acids research* **43**, e37 (2015).
130. Schlenk, R.F., Dohner, K., Krauter, J., Frohling, S., Corbacioglu, A., Bullinger, L., Habdank, M., Spath, D., Morgan, M., Benner, A., Schlegelberger, B., Heil, G., Ganser, A., Dohner, H. & German-Austrian Acute Myeloid Leukemia Study, G. Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. *The New England journal of medicine* **358**, 1909-1918 (2008).
131. Schlenk, R.F., Kayser, S., Bullinger, L., Kobbe, G., Casper, J., Ringhoffer, M., Held, G., Brossart, P., Lubbert, M., Salih, H.R., Kindler, T., Horst, H.A., Wulf, G., Nachbaur, D., Gotze, K., Lamparter, A., Paschka, P., Gaidzik, V.I., Teleanu, V., Spath, D., Benner, A., Krauter, J., Ganser, A., Dohner, H., Dohner, K. & German-Austrian, A.M.L.S.G. Differential impact of allelic ratio and insertion site in FLT3-ITD-positive AML with respect to allogeneic transplantation. *Blood* **124**, 3441-3449 (2014).
132. Schofield, R. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells* **4**, 7-25 (1978).
133. Schuurhuis, G.J., Heuser, M., Freeman, S., Bene, M.C., Buccisano, F., Cloos, J., Grimwade, D., Haferlach, T., Hills, R.K., Hourigan, C.S., Jorgensen, J.L., Kern, W., Lacombe, F., Maurillo, L., Preudhomme, C., van der Reijden, B.A., Thiede, C., Venditti, A., Vyas, P., Wood, B.L., Walter, R.B., Dohner, K., Roboz, G.J. & Ossenkoppele, G.J. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. *Blood* **131**, 1275-1291 (2018).
134. Shafer, D. & Grant, S. Update on rational targeted therapy in AML. *Blood Rev* **30**, 275-283 (2016).
135. Shiba, N., Yoshida, K., Shiraishi, Y., Okuno, Y., Yamato, G., Hara, Y., Nagata, Y., Chiba, K., Tanaka, H., Terui, K., Kato, M., Park, M.J., Ohki, K., Shimada, A., Takita, J., Tomizawa, D., Kudo, K., Arakawa, H., Adachi, S., Taga, T., Tawa, A., Ito, E., Horibe, K., Sanada, M., Miyano, S., Ogawa, S. & Hayashi, Y. Whole-exome sequencing reveals the spectrum of gene mutations and the clonal evolution patterns in paediatric acute myeloid leukaemia. *British journal of haematology* **175**, 476-489 (2016).
136. Shih, A.H., Abdel-Wahab, O., Patel, J.P. & Levine, R.L. The role of mutations in epigenetic regulators in myeloid malignancies. *Nat Rev Cancer* **12**, 599-612 (2012).

137. Shugay, M., Britanova, O.V., Merzlyak, E.M., Turchaninova, M.A., Mamedov, I.Z., Tuganbaev, T.R., Bolotin, D.A., Staroverov, D.B., Putintseva, E.V., Plevova, K., Linnemann, C., Shagin, D., Pospisilova, S., Lukyanov, S., Schumacher, T.N. & Chudakov, D.M. Towards error-free profiling of immune repertoires. *Nat Methods* **11**, 653-655 (2014).
138. Shultz, L.D., Lyons, B.L., Burzenski, L.M., Gott, B., Chen, X., Chaleff, S., Kotb, M., Gillies, S.D., King, M., Mangada, J., Greiner, D.L. & Handgretinger, R. Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *J Immunol* **174**, 6477-6489 (2005).
139. Sinha, S., Thomas, D., Yu, L., Gentles, A.J., Jung, N., Corces-Zimmerman, M.R., Chan, S.M., Reinisch, A., Feinberg, A.P., Dill, D.L. & Majeti, R. Mutant WT1 is associated with DNA hypermethylation of PRC2 targets in AML and responds to EZH2 inhibition. *Blood* **125**, 316-326 (2015).
140. Smith, E.N., Jepsen, K., Khosroheidari, M., Rassenti, L.Z., D'Antonio, M., Ghia, E.M., Carson, D.A., Jamieson, C.H., Kipps, T.J. & Frazer, K.A. Biased estimates of clonal evolution and subclonal heterogeneity can arise from PCR duplicates in deep sequencing experiments. *Genome Biol* **15**, 420 (2014).
141. Song, J., Mercer, D., Hu, X., Liu, H. & Li, M.M. Common leukemia- and lymphoma-associated genetic aberrations in healthy individuals. *J Mol Diagn* **13**, 213-219 (2011).
142. Steensma, D.P., Bejar, R., Jaiswal, S., Lindsley, R.C., Sekeres, M.A., Hasserjian, R.P. & Ebert, B.L. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* **126**, 9-16 (2015).
143. Struski, S., Lagarde, S., Bories, P., Puiseux, C., Prade, N., Cuccuini, W., Pages, M.P., Bidet, A., Gervais, C., Lafage-Pochitaloff, M., Roche-Lestienne, C., Barin, C., Penther, D., Nadal, N., Radford-Weiss, I., Collonge-Rame, M.A., Gaillard, B., Mugneret, F., Lefebvre, C., Bart-Delabesse, E., Petit, A., Leverger, G., Broccardo, C., Luquet, I., Pasquet, M. & Delabesse, E. NUP98 is rearranged in 3.8% of pediatric AML forming a clinical and molecular homogenous group with a poor prognosis. *Leukemia* **31**, 565-572 (2017).
144. Swerdlow S.H., C.E., Harris N.L., Jaffe E.S., Pileri S.A., Stein H., Thiele J. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues.*, (IARC, Lyon, 2017).
145. Telenius, H., Carter, N.P., Bebb, C.E., Nordenskjold, M., Ponder, B.A. & Tunnacliffe, A. Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics* **13**, 718-725 (1992).
146. Teo, T., Lam, F., Yu, M., Yang, Y., Basnet, S.K., Albrecht, H., Sykes, M.J. & Wang, S. Pharmacologic Inhibition of MNKs in Acute Myeloid Leukemia. *Mol Pharmacol* **88**, 380-389 (2015).
147. Thol, F., Gabdoulline, R., Liebich, A., Klement, P., Schiller, J., Kandziora, C., Hambach, L., Stadler, M., Koenecke, C., Flintrop, M., Pankratz, M., Wichmann, M., Neziri, B., Buttner, K., Heida, B., Klesse, S., Chaturvedi, A., Kloos, A., Gohring, G., Schlegelberger, B., Gaidzik, V.I., Bullinger, L., Fiedler, W., Heim, A., Hamwi, I., Eder, M., Krauter, J., Schlenk, R.F., Paschka, P., Dohner, K., Dohner, H., Ganser, A. & Heuser, M. Measurable residual disease (MRD) monitoring by NGS before allogeneic hematopoietic cell transplantation in AML. *Blood* (2018).
148. Ungerstedt, J.S. Epigenetic Modifiers in Myeloid Malignancies: The Role of Histone Deacetylase Inhibitors. *Int J Mol Sci* **19**(2018).
149. Ungewickell, A. & Medeiros, B.C. Novel agents in acute myeloid leukemia. *Int J Hematol* **96**, 178-185 (2012).

150. van Dongen, J.J., Macintyre, E.A., Gabert, J.A., Delabesse, E., Rossi, V., Saglio, G., Gottardi, E., Rambaldi, A., Dotti, G., Griesinger, F., Parreira, A., Gameiro, P., Diaz, M.G., Malec, M., Langerak, A.W., San Miguel, J.F. & Biondi, A. Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. Report of the BIOMED-1 Concerted Action: investigation of minimal residual disease in acute leukemia. *Leukemia* **13**, 1901-1928 (1999).
151. Vardiman, J.W., Harris, N.L. & Brunning, R.D. The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood* **100**, 2292-2302 (2002).
152. Vedula, R.S. & Lindsley, R.C. Measurement of Residual Disease in Acute Myeloid Leukemia. *Curr Hematol Malig Rep* **12**, 574-581 (2017).
153. Vergez, F., Green, A.S., Tamburini, J., Sarry, J.E., Gaillard, B., Cornillet-Lefebvre, P., Pannetier, M., Neyret, A., Chapuis, N., Ifrah, N., Dreyfus, F., Manenti, S., Demur, C., Delabesse, E., Lacombe, C., Mayeux, P., Bouscary, D., Recher, C. & Bardet, V. High levels of CD34+CD38low/-CD123+ blasts are predictive of an adverse outcome in acute myeloid leukemia: a Groupe Ouest-Est des Leucemies Aigues et Maladies du Sang (GOELAMS) study. *Haematologica* **96**, 1792-1798 (2011).
154. Viehmann, S., Teigler-Schlegel, A., Bruch, J., Langebrake, C., Reinhardt, D. & Harbott, J. Monitoring of minimal residual disease (MRD) by real-time quantitative reverse transcription PCR (RQ-RT-PCR) in childhood acute myeloid leukemia with AML1/ETO rearrangement. *Leukemia* **17**, 1130-1136 (2003).
155. Wei, A.H. & Tiong, I.S. Midostaurin, enasidenib, CPX-351, gemtuzumab ozogamicin, and venetoclax bring new hope to AML. *Blood* **130**, 2469-2474 (2017).
156. Woods, N.T., Baskin, R., Golubeva, V., Jhuraney, A., De-Gregoriis, G., Vaclova, T., Goldgar, D.E., Couch, F.J., Carvalho, M.A., Iversen, E.S. & Monteiro, A.N. Functional assays provide a robust tool for the clinical annotation of genetic variants of uncertain significance. *NPJ Genom Med* **1**(2016).
157. Xie, M., Lu, C., Wang, J., McLellan, M.D., Johnson, K.J., Wendl, M.C., McMichael, J.F., Schmidt, H.K., Yellapantula, V., Miller, C.A., Ozenberger, B.A., Welch, J.S., Link, D.C., Walter, M.J., Mardis, E.R., Dipersio, J.F., Chen, F., Wilson, R.K., Ley, T.J. & Ding, L. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nature medicine* **20**, 1472-1478 (2014).
158. Xu, H., Valerio, D.G., Eisold, M.E., Sinha, A., Koche, R.P., Hu, W., Chen, C.W., Chu, S.H., Brien, G.L., Park, C.Y., Hsieh, J.J., Ernst, P. & Armstrong, S.A. NUP98 Fusion Proteins Interact with the NSL and MLL1 Complexes to Drive Leukemogenesis. *Cancer cell* **30**, 863-878 (2016).
159. Ye, H., Adane, B., Khan, N., Sullivan, T., Minhajuddin, M., Gasparetto, M., Stevens, B., Pei, S., Balys, M., Ashton, J.M., Klemm, D.J., Woolthuis, C.M., Stranahan, A.W., Park, C.Y. & Jordan, C.T. Leukemic Stem Cells Evade Chemotherapy by Metabolic Adaptation to an Adipose Tissue Niche. *Cell Stem Cell* **19**, 23-37 (2016).
160. Zong, C., Lu, S., Chapman, A.R. & Xie, X.S. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* **338**, 1622-1626 (2012).
161. Zwaan, C.M., Meshinchi, S., Radich, J.P., Veerman, A.J., Huismans, D.R., Munske, L., Podleschny, M., Hahlen, K., Pieters, R., Zimmermann, M., Reinhardt, D., Harbott, J., Creutzig, U., Kaspers, G.J. & Griesinger, F. FLT3 internal tandem duplication in 234 children with acute myeloid leukemia:

prognostic significance and relation to cellular drug resistance. *Blood* **102**, 2387-2394 (2003).
162. Zwaan, C.M., Kolb, E.A., Reinhardt, D., Abrahamsson, J., Adachi, S., Aplenc, R., De Bont, E.S., De Moerloose, B., Dworzak, M., Gibson, B.E., Hasle, H., Leverger, G., Locatelli, F., Ragu, C., Ribeiro, R.C., Rizzari, C., Rubnitz, J.E., Smith, O.P., Sung, L., Tomizawa, D., van den Heuvel-Eibrink, M.M., Creutzig, U. & Kaspers, G.J. Collaborative Efforts Driving Progress in Pediatric Acute Myeloid Leukemia. *J Clin Oncol* **33**, 2949-2962 (2015).

9 List of Abbreviations

ADO	Allelic dropout
ADxE	cytarabine, liposomal daunorubicin and etoposide
AI	cytarabine and idarubicin
AIE	Cytarabine, idarubicin and etoposide
ALL	Acute Lymphoblastic Leukemia
AML	Acute myeloid leukemia
AR	Allelic ratio
BFM	Berlin-Frankfurt-Münster
BM	Bone marrow
CBL	Cbl Proto-Oncogene
CCR	Continued complete remission
CD34	CD34 Molecule
CD38	CD38 Molecule
CD96	CD96 Molecule
CD123	Interleukin 3 Receptor Subunit Alpha – IL3R
CEBPA	CCAAT Enhancer Binding Protein Alpha
CFU	Colony forming unit
CIR	Cumulative incidence of relapse
COG	Children's Oncology Group
DNA	Desoxyrbionucleic acid
DNMT3A	DNA Methyltransferase 3 Alpha
dNTP	Desoxyribonucleosidtriphosphate.
EFS	Event-free survival
FAB	French-American-British
FLT3	fms related tyrosine kinase 3
FLT3 ITD	fms related tyrosine kinase 3-internal tandem deletion
GATA2	GATA Binding Protein 2
HAE	HD-cytarabine and etoposide
haM	HD-cytarabine (1g/m^2) and mitoxantrone
HAM	HD-cytarabine (3g/m^2) and mitoxantrone
HAT	histone acetyltransferase
HR	High risk

HSC	Hematopoietic stem cell
IFSET	ion-sensitive field-effect transistor
InDel	Insertion, Deletion
IR	Intermediate risk
KIT	KIT Proto-Oncogene Receptor Tyrosine Kinase
KMT2A	Lysine Methyltransferase 2A
KRAS	KRAS Proto-Oncogene
LSC	Leukemic stem cell
MDA	Multiple displacement amplification
MDS	Myelodysplastic syndrome
MEK	Mitogen-activated protein kinase kinase
miRNA	microRNA
MLLT11	Transcription Factor 7 Cofactor
MRD	Minimal residual disease
MYC	MYC Proto-Oncogene, BHLH Transcription Factor
NGS	Next Generation Sequencing
NRAS	NRAS Proto-Oncogene, GTPase
NSL HAT complex	Non-specific lethal histone acetyltransferase complex
NPM1	Nucleophosmin 1
NRAS	NRAS Proto-Oncogene
NSD1	Nuclear Receptor Binding SET Domain Protein 1
NUP98	Nucleoporin 98
OS	Overall survival
PB	Peripheral blood
PCR	Polymerase chain reaction
qPCR	quantitative Real Time PCR
RNA	Ribonucleic acid
RT	reversible termination
SCT	Stem cell transplantation
SNV	Single nucleotide variant
SR	Standard risk
TIM3	Hepatitis A Virus Cellular Receptor 2 - HAVCR2
TKI	tyrosine kinase inhibitor
TP53	Tumor Protein P53

TS	Targeted amplicon-based Next Generation Sequencing
UMI	Unique molecular identifiers
VUS	Variant of unknown significance
WES	Whole exome seguencing
WGA	Whole genome amplification
WGS	Whole genome sequencing
WHO	World Health Organization
WT1	Wilms Tumor 1 gene

10 List of Figures

11 List of Tables

12 Curriculum vitae

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

13 Danksagung

An dieser Stelle möchte ich all jenen danken, die mich während der Bearbeitung meiner Dissertation unterstützt haben.

Mein besonderer Dank gilt Prof. Dr. med. Dirk Reinhardt für die Überlassung des interessanten Themas und die Möglichkeit, diese Dissertation in seiner Arbeitsgruppe „Pädiatrische Hämatologie und Onkologie“ anzufertigen. Ich möchte mich außerdem für die engagierte Betreuung, sowie die fachliche und persönliche Unterstützung bedanken und freue mich sehr über die Möglichkeit, weiterhin in der pädiatrischen Hämatologie und Onkologie wissenschaftlich tätig zu sein.

Herrn Prof. Dr. rer. nat. Nils von Neuhoff danke ich für die wissenschaftliche Ausbildung, die erfolgreiche Zusammenarbeit im Labor, die persönliche Unterstützung sowie die gemeinsame Motivation für dieses Thema. Herrn Prof. Dr. med. Helmut Hanenberg gilt mein Dank für die engagierte Betreuung und wissenschaftliche Ausbildung sowie seinen fachlichen Rat.

Dem gesamten Team des Labors und der AML-BFM Studiengruppe gilt mein Dank für den wissenschaftlichen Austausch, die gegenseitige Hilfe und die konstruktive Atmosphäre. Es war eine schöne und spannende Zeit mit euch und ich freue mich, hier nicht nur beruflich Anschluss gefunden zu haben.

Ganz besonders danke ich meiner Familie, die mich jederzeit während des Studiums und der Bearbeitung der Dissertation unterstützt hat. Kristopher Gundlach danke ich für das Verständnis, die Geduld und die uneingeschränkte moralische Unterstützung. Abschließend danke ich meinen engsten Freunden für die kostbare gemeinsame Zeit. Ihr habt meine Begeisterung für die Forschung immer unterstützt und mich doch daran erinnert, dass es neben dem Labor auch viele andere schöne Seiten im Leben gibt.

14 Eidesstattliche Erklärungen

Erklärung:

Hiermit erkläre ich, gem. § 6 Abs. (2) g) der Promotionsordnung der Fakultät für Biologie zur Erlangung der Dr. rer. nat., dass ich das Arbeitsgebiet, dem das Thema „Deciphering the genetic diversity of pediatric AML using Next Generation Sequencing tools“ zuzuordnen ist, in Forschung und Lehre vertrete und den Antrag von Christiane Walter befürworte und die Betreuung auch im Falle eines Weggangs, wenn nicht wichtige Gründe dem entgegenstehen, weiterführen werde.

Essen, den _____ _____ _____

Name des wissenschaftl. Betreuers/Mitglieds der Universität Duisburg-Essen	Unterschrift des wissenschftl. Betreuers/Mitglieds der Universität Duisburg-Essen

Erklärung:

Hiermit erkläre ich, gem. § 7 Abs. (2) e) + g) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich keine anderen Promotionen bzw. Promotionsversuche in der Vergangenheit durchgeführt habe und dass diese Arbeit von keiner anderen Fakultät/Fachbereich abgelehnt worden ist.

Essen, den _____ _____

Unterschrift der Doktorandin

Erklärung:

Hiermit erkläre ich, gem. § 7 Abs. (2) d) + f) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich die vorliegende Dissertation selbständig verfasst und mich keiner anderen als der angegebenen Hilfsmittel bedient, bei der Abfassung der Dissertation nur die angegeben Hilfsmittel benutzt und alle wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet habe.

Essen, den _____ _____

Unterschrift der Doktorandin