Aus der Medizinischen Klinik und
Poliklinik III - Großhadern
der Ludwig-Maximilians-Universität München
Direktor: Prof. Dr. med. W. Hiddemann

Deciphering the genetic heterogeneity in Acute Myeloid Leukemia: Association of gene mutations with distinct chromosomal aberrations

Dissertation
zum Erwerb des Doktorgrades
der Humanbiologie
(Dr. rer. biol. hum.)

an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von
Luise Hartmann
aus Hannover

München, 2017

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter:	Prof. Dr. Karsten Spiekermann
Mitberichterstatter:	Prof. Dr. Elke Holinski-Feder
	Priv. Doz. Dr. Ursula Zimber-Strobl
	Priv. Doz. Dr. Michael Albert
Mitbetreuung durch den promovierten Mitarbeiter:	Dr. Philipp Greif
Dekan:	Prof. Dr. med. dent. Reinhard Hickel
Tag der mündlichen Prüfung:	18.04.2017

Eidesstattliche Versicherung

Hartmann, Luise

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Thema:

‚Deciphering the genetic heterogeneity in Acute Myeloid Leukemia: Association of gene mutations with distinct chromosomal aberrations'

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, den

_____

Luise Hartmann

Every snowflake that I caught was a miracle unlike any other.

-Alice Hoffman, *The Museum of Extraordinary Things*

Table of contents

I. Zusammenfassung

Das Hauptziel der vorliegenden Dissertation ist die genetische Charakterisierung von zytogenetischen Subgruppen der Akuten Myeloischen Leukämie (AML). Grundlage dieser kumulativen Dissertation sind die beiden aufgeführten Publikationen, die in renommierten Fachzeitschriften erschienen sind (Impact-factor von *Blood* in 2014: 10.452; aktueller Impact-factor von *Nature Communications*: 11.470):

- Herold, T., K. H. Metzeler, S. Vosberg, **L. Hartmann**, C. Röllig, F. Stölzel, S. Schneider, M. Hubmann, E. Zellmeier, B. Ksienzyk, V. Jurinovic, Z. Pasalic, P. M. Kakadia, A. Dufour, A. Graf, S. Krebs, H. Blum, M. C. Sauerland, T. Büchner, W. E. Berdel, B. J. Wörmann, M. Bornhäuser, G. Ehninger, U. Mansmann, W. Hiddemann, S. K. Bohlander, K. Spiekermann and P. A. Greif (2014). "Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis." Blood 124(8): 1304-1311.

- **Hartmann, L.**, S. Dutta, S. Opatz, S. Vosberg, K. Reiter, G. Leubolt, K. H. Metzeler, T. Herold, S. A. Bamopoulos, K. Bräundl, E. Zellmeier, B. Ksienzyk, N. P. Konstandin, S. Schneider, K. P. Hopfner, A. Graf, S. Krebs, H. Blum, J. M. Middeke, F. Stölzel, C. Thiede, S. Wolf, S. K. Bohlander, C. Preiss, L. Chen-Wichmann, C. Wichmann, M. C. Sauerland, T. Büchner, W. E. Berdel, B. J. Wörmann, J. Braess, W. Hiddemann, K. Spiekermann and P. A. Greif (2016). "ZBTB7A mutations in acute myeloid leukaemia with t(8;21) translocation." Nat Commun 7: 11733.

In beiden Arbeiten wurden Genmutationen identifiziert, die spezifisch bei AML Patienten mit bestimmten chromosomalen Veränderungen auftreten: *SRSF2* Mutationen bei Patienten mit Trisomie 13 und *ZBTB7A* Mutationen bei Patienten mit t(8;21) Translokation.

Es ist bekannt, dass die Entwicklung von AML als mehrstufiger Prozess abläuft, der von Veränderungen im Genom getrieben ist. Die spezifische Assoziation von bestimmten chromosomalen Veränderungen und Genmutationen, so wie in dieser Arbeit beschrieben, deutet auf eine definierte Kooperation der verschiedenen genetischen Veränderungen bei der Leukämogenese hin. Neue Einblicke in dieses Zusammenspiel können dazu beitragen, die Entstehung der AML besser zu verstehen und gezielte Therapieansätze zu entwickeln.

II. Summary

The main objective of this dissertation is the genetic characterization of cytogenetic subgroups of acute myeloid leukemia (AML). This cumulative dissertation is based on two articles that were published in leading scientific journals (impact factor of *Blood* in 2014: 10.452; recent impact factor of *Nature Communications*: 11.470):

- Herold, T., K. H. Metzeler, S. Vosberg, **L. Hartmann**, C. Röllig, F. Stölzel, S. Schneider, M. Hubmann, E. Zellmeier, B. Ksienzyk, V. Jurinovic, Z. Pasalic, P. M. Kakadia, A. Dufour, A. Graf, S. Krebs, H. Blum, M. C. Sauerland, T. Büchner, W. E. Berdel, B. J. Wörmann, M. Bornhäuser, G. Ehninger, U. Mansmann, W. Hiddemann, S. K. Bohlander, K. Spiekermann and P. A. Greif (2014). "Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis." Blood 124(8): 1304-1311.

- **Hartmann, L.**, S. Dutta, S. Opatz, S. Vosberg, K. Reiter, G. Leubolt, K. H. Metzeler, T. Herold, S. A. Bamopoulos, K. Bräundl, E. Zellmeier, B. Ksienzyk, N. P. Konstandin, S. Schneider, K. P. Hopfner, A. Graf, S. Krebs, H. Blum, J. M. Middeke, F. Stölzel, C. Thiede, S. Wolf, S. K. Bohlander, C. Preiss, L. Chen-Wichmann, C. Wichmann, M. C. Sauerland, T. Büchner, W. E. Berdel, B. J. Wörmann, J. Braess, W. Hiddemann, K. Spiekermann and P. A. Greif (2016). "ZBTB7A mutations in acute myeloid leukaemia with t(8;21) translocation." Nat Commun 7: 11733.

In both studies, gene mutations were found that occur specifically in AML patients with distinct chromosomal aberrations: *SRSF2* mutations in patients with trisomy 13 and *ZBTB7A* mutations in patients with t(8;21) translocation.

It is known that the development of AML is a multistep process driven by genomic alterations. The specific associations between certain chromosomal lesions and gene mutations, as described in this dissertation, point towards a defined leukemogenic cooperativity between the different kinds of genetic alterations. New insights into this interaction can contribute to a better understanding of the evolution of AML and to the development of targeted therapy approaches.

III. Abbreviations

2-DG	2-Deoxy-D-glucose
AML	Acute myeloid leukemia
CBF	Core binding factor
CLP	Common lymphoid progenitor
CML	Chronic myeloid leukemia
CMP	Common myeloid progenitor
CN-AML	Cytogenetically normal AML
ELN	European leukemia network
FAB	French-American-British
HSC	Hematopoietic stem cell
INDEL	Small insertion/deletion
ITD	Internal tandem duplication
MDS	Myelodysplastic syndrome
MPP	Multipotent progenitor
MRC	Medical Research Council
NGS	Next generation sequencing
PTD	Partial tandem duplication
SNV	Single nucleotide variant
TCGA	The cancer genome atlas
WHO	World Health Organization

IV. Tables and Figures

1. Introduction

1.1. Acute myeloid leukemia (AML)

Clinical characteristics

Acute myeloid leukemia (AML) is a hematopoietic malignancy characterized by excessive growth of clonal myeloid progenitor cells. The term '*leukemia*' was coined in the 19th century by Rudolf Virchow, based on his observations of 'white blood' (Kampen, 2012).

Common symptoms of AML include anemia, bleeding and frequent infections. The diagnosis is based on cytomorphological assessment of bone marrow and peripheral blood. AML is mostly a disease of the elderly, with a median age of >65 years at diagnosis (Juliusson et al, 2012; Wang, 2014). A combination of daunorubicin and cytarabine (the so-called '3+7' regimen) is the standard initial treatment for AML and results in remission, i.e. reduction of bone marrow blast counts to <5%, in 40-80% of patients (Burnett et al, 2011). However, a high proportion of patients will eventually relapse and become non-responsive to further therapy approaches. The five-year survival rate for adult AML can be as low as 10% (Burnett et al, 2011). Importantly, it was shown that remission and survival rates highly depend on clinical (e.g. age) and biological factors (e.g. karyotype, gene mutations), allowing for risk stratification and treatment adjustment such as consideration of allogeneic stem cell transplantation for suitable patients with high risk disease (Estey and Döhner, 2006; Döhner et al, 2010).

Initially, AML was classified based on cytomorphology. In 1976, the French-American-British (FAB) co-operative group proposed the so-called FAB classification which recognizes eight subtypes (M0- M7) with respect to cell type and differentiation (Bennett et al, 1976). Later, with better understanding of AML pathogenesis, a more refined classification established by the World Health Organization (WHO) also included biological and cytogenetic factors (Vardiman et al, 2009).

Table 1: WHO 2008 classification of acute myeloid leukemia (Vardiman et al, 2009)

Acute myeloid leukemia
Acute myeloid leukemia with recurrent genetic abnormalities Acute myeloid leukemia with myelodysplasia-related changes Therapy-related myeloid neoplasms Acute myeloid leukemia, not otherwise specified

Leukemogenesis

Normal hematopoiesis follows a tightly regulated hierarchy (Figure 1). Hematopoietic stem cells (HSC) reside in the bone marrow and have self-renewal capacities but can also differentiate into all blood cell types. Upon stimulation, HSCs differentiate to multipotent progenitors (MPP) which are still able to generate all kinds of mature blood cells but have lost self-renewal capacity (Fiedler and Brunner, 2012). The common lymphoid progenitors (CLP) and common myeloid progenitors (CMP) give rise to the mature cells of the lymphoid lineage (T-cells, B-cells, NK-cells) or the mature cells of myeloid lineage (erythrocytes, megakaryocytes, macrophages, granulocytes), respectively (Kondo et al, 1997; Akashi et al 2000). Differentiation and commitment to cell lineage fates have been demonstrated to highly depend on the expression of specific combinations of transcription factors (Tenen, 2003; Wilson et al, 2010; Pouzolles et al, 2016).

It was shown that AML derives from early progenitor cells (Bonnet and Dick, 1997). Ddifferentiation of myeloid progenitors is blocked and the cells proliferate unrestrictedly, leading to accumulation of clonal immature precursor cells in the bone marrow and consecutive suppression of normal hematopoiesis.

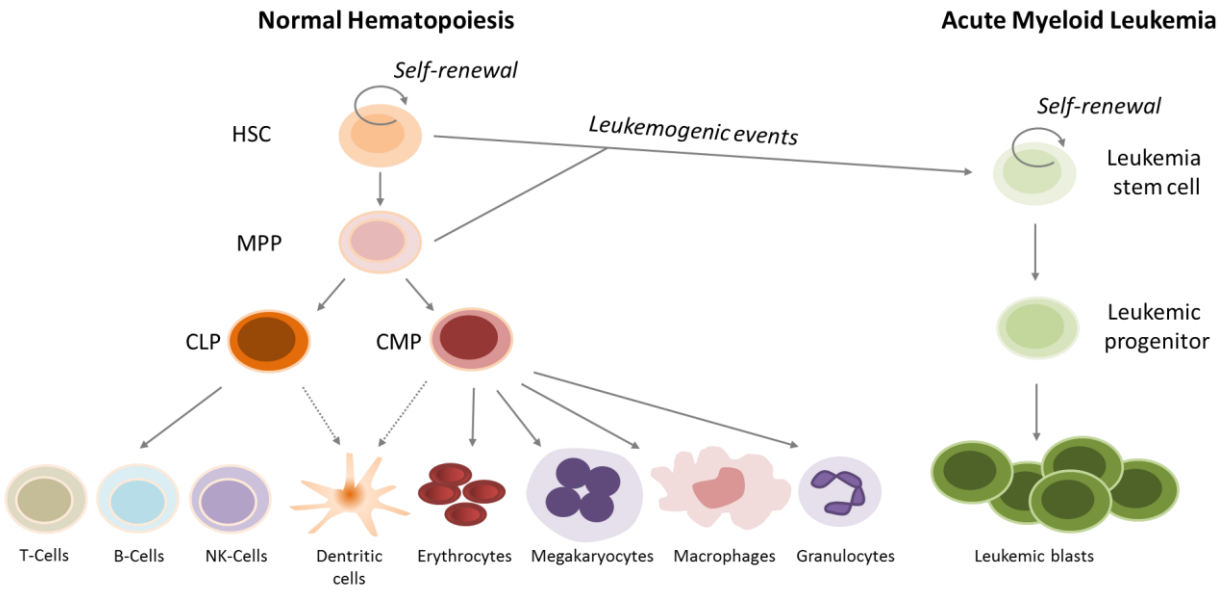


Figure 1: Normal hematopoiesis and acute myeloid leukemia (adapted from Tan et al, 2006). Blood cells derive from precursor cells that undergo multiple differentiation steps. In AML, differentiation of hematopoietic stem cells (HSC) or multipotent progenitors (MPP) is blocked, leading to accumulation of leukemic blasts. CLP= common lymphoid progenitor, CMP= common myeloid progenitor

The transformation of normal HSCs or MPPs to leukemic blasts is a multi-step process driven by sequential leukemogenic events (reviewed by Horton and Huntly, 2012). These events are commonly alterations of the genome. In consequence, characterization of genomic lesions in AML is essential to understand the pathogenesis of AML and ultimately to enable the development of tailored, more effective therapies.

1.2. Chromosomal alterations in AML

Recurrent cytogenetic alterations, i.e. structural or numerical chromosomal abnormalities, in AML were already described more than 40 years ago by pioneering work of Janet Rowley and others (reviewed by Freireich et al, 2014). The discovery of recurring balanced translocations between chromosomes 8 and 21, termed t(8;21)(q22;q22), in AML was the first translocation to be described in human cancers and is considered a milestone in our understanding of cancer genetics (Rowley, 1973). In approximately 50-60% of AML patients, abnormal karyotypes can be detected and as shown in Figure 2, the diversity of cytogenetic abnormalities is rather high.

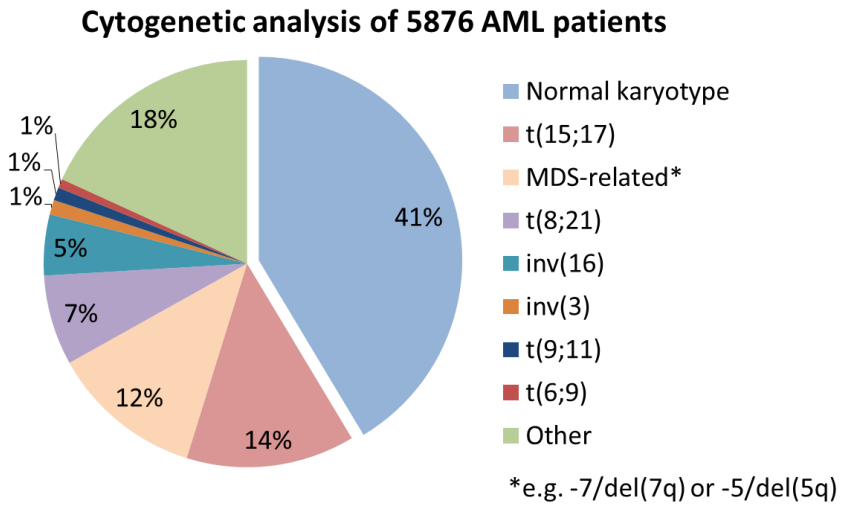


Figure 2: Cytogenetic results from the Medical Research Council (MRC) trials (Grimwade et al, 2010). A total of 5876 AML karyotypes were analyzed and abnormalities were identified in 59% of patients. Of note, these patients were <60 years old, and distribution of cytogenetic aberrations varies in different age groups. MDS= Myelodysplastic syndrome

Despite this complexity, the prognostic impact of the most common chromosomal abnormalities has been assessed through efforts of numerous study groups (overview in Burnett et al, 2011), leading to the widely used risk classification established by the European Leukemia Network (ELN) and Medical Research Council (MRC).

Table 2: MRC AML risk classification according to chromosomal aberrations (Grimwade et al, 2010)

Favorable Risk
t(15;17)(q22;q21) inv(16)(p13.1q22); t(16;16)(p13.1;q22) t(8;21)(q22;q22)
Intermediate Risk
Normal karyotype Cytogenetic abnormalities not classified as favorable or adverse
Adverse Risk
abnormal(3q), excluding t(3;5)(q21~25;q31~35) inv(3)(q21q26.2); t(3;3)(q21;q26.2) add(5q), del(5q), -5 -7, add(7q)/del(7q) t(6;11)(q27;q23) t(10;11)(p11~13;q23) t(11q23), excluding t(9;11)(p21~22;q23) and t(11;19)(q23;p13) t(9;22)(q34;q11) -17/abnormal(17p) complex karyotype*

* Defined as >4 independent chromosomal aberrations

Besides assessing their prognostic impact, understanding the underlying mechanisms how chromosome abnormalities arise and how they contribute to leukemogenesis is of great importance.

Aneuploidy, i.e. gain or loss of entire chromosomes, is the result of erroneous chromosome segregation during mitosis (Bakhoum and Compton, 2012). It is challenging to decipher the direct influence of numerical chromosomal aberrations on leukemogenesis since the aberrations affect numerous gene loci. However, gene dosage effects are believed to play an important role. For example, in a study of 80 patients with trisomy 8 (+8) as sole aberration, 452 genes were significantly upregulated and 329 downregulated in +8 AML compared to cytogenetically normal AML (Becker et al, 2014). Of the 452 upregulated genes, 189 (42%) were located on chromosome 8.

The precise molecular mechanism which causes chromosomal translocations remains elusive. Studies showed that homologous recombination, non-homologous end joining and chromosome fragile sites potentially trigger the formation of translocations (reviewed by Aplan, 2006). Moreover, it was shown that chromosome segregation errors during mitosis can lead to translocations as well (Janssen et al, 2011). In general, oncogenic translocations lead either to novel fusion genes (Hermans et al, 1987; de Thé et al, 1991) or juxtaposition of regulatory elements from one translocation partner to the other, resulting in aberrant gene expression (ar-Rushdi et al, 1983; Gröschel et al, 2014). The functional consequences of many chromosomal rearrangements have been subject to intensive studies. The recurrent translocation t(8;21)(q22;q22), for example, leads to the chimeric *RUNX1/RUNX1T1* gene (also known as *AML1-ETO*) (Erickson et al, 1992). RUNX1 is an important transcription factor for regulation of hematopoiesis (Tanaka et al, 1995; Okuda et al, 1996) and part of the so-called core binding factor (CBF) complex. Through fusion with RUNX1T1, normal function of RUNX1 in the CBF complex is disturbed, preventing transcription of CBF target genes important for myeloid differentiation, and thereby leading to disruption of normal hematopoiesis and inactivation of tumor suppressor genes (Westendorf et al, 1998; Goyoma and Mulloy, 2011).

Figure 3: The core binding factor (CBF) complex (adapted from Solh et al, 2014). (A) The CBF consists of 2 subunits. RUNX1 and CBFB form a complex known to initiate transcription of genes involved in myeloid differentiation. (B) The t(8;21) translocation leads to the RUNX1/RUNX1T1 fusion and, via recruitment of additional factors, to inactivation of CBF target genes.

However, *in vivo* models indicate the requirement of additional lesions, such as gene mutations, for leukemogenesis as the *RUNX1/RUNX1T1* fusion gene alone is not sufficient to induce leukemia in murine models (Rhoades et al, 2000; Yuan et al, 2001). Similarly, in children with t(8;21) positive AML, the *RUNX1/RUNX1T1* fusion could already be detected in neonatal blood samples but the full-blown leukemia was characterized by additional genomic aberrations (Wiemels et al, 2002).

1.3. The mutational landscape of AML

Besides microscopically detectable chromosomal alterations, gene mutations in AML have also been intensively investigated. Initially, gene mutations were identified based on candidate approaches or serendipitously. For example, AML samples were screened for *NRAS* mutations based on the observation that this oncogene is mutated in other types of cancer (Bos et al, 1985). *NPM1* mutations, which occur in approximately 25-35% of AML patients, were discovered after detection of aberrant cytoplasmic localization of the protein. It was shown that in most cases an insertion of 4 bases lead to a frame shift in the region encoding the C-terminus of NPM1, thereby truncating the protein and leading to loss of a nuclear localization signal and consequently abnormal sub-cellular localization (Falini et al, 2005).

With the introduction of next generation sequencing (NGS) technologies (reviewed by Welch and Link, 2011), the number of known recurrently mutated genes in AML has increased tremendously. In fact, the first human cancer genome to be completely sequenced was from a patient with AML (Ley et al, 2008). Shortly after, *DNMT3A* mutations were described by the same research group (Ley et al, 2010), followed by the discovery of several other novel gene mutations in AML such as *BCOR* (Grossmann et al, 2011), *GATA2* (Greif et al, 2012), *RAD21* (Dolnik et al, 2012) and *ASXL2* (Micol et al, 2014). Through high-throughput sequencing approaches, these and other mutations have been studied by several groups with regards to their frequency and prognostic significance (reviewed by Larsson et al, 2013; Meyer and Levine, 2014; Döhner et al, 2015). An overview of the most common recurrently mutated genes in AML is shown in Table 3.

Table 3: Recurrently mutated genes in AML (according to Döhner et al, 2015). ITD= Internal tandem duplication, PTD= Partial tandem duplication

Mutated gene	Frequency
NPM1	25-35%
FLT3-ITD	20%
DNMT3A	18-22%
NRAS	15%
TET2	7-25%
CEBPA	6-10%
RUNX1	5-15%
ASXL1	5-17%
IDH1; IDH2	7-14%; 8-19%
KIT	<5%
KMT2A-PTD	5%

Development of AML is believed to be a multistep process that requires the sequential acquisition of several mutations. Based on studies of CBF leukemia, it was proposed that these mutations would fall into two distinct categories (Speck and Gilliland, 2002). Class I mutations (for example in *FLT3*, *KIT* and *NRAS*) enhance proliferation and survival, predominantly through constitutively activated signaling pathways. In contrast, class II mutations result in impaired differentiation of hematopoietic progenitor cells and often affect transcription factors such as *RUNX1* or *GATA1/2*. Mutations of both classes are likely necessary to develop full-blown leukemia.

In the last years, with the discovery of numerous novel gene mutations, this model had to be revised. Functional analyses demonstrated that several mutations do not accurately fit in class I or II but can be categorized in other functional groups. *DNMT3A*, for example, encodes a DNA methyltransferase and *DNMT3A* mutations lead to global changes of the DNA methylation pattern (Russler-Germain et al, 2014). Likewise, *TET2* and *IDH1/2* mutations have also been associated with epigenetic changes (Figueroa et al, 2010). In consequence, new functional classifications of gene mutations in AML have been suggested as shown in Figure 4 (Thiede, 2012).

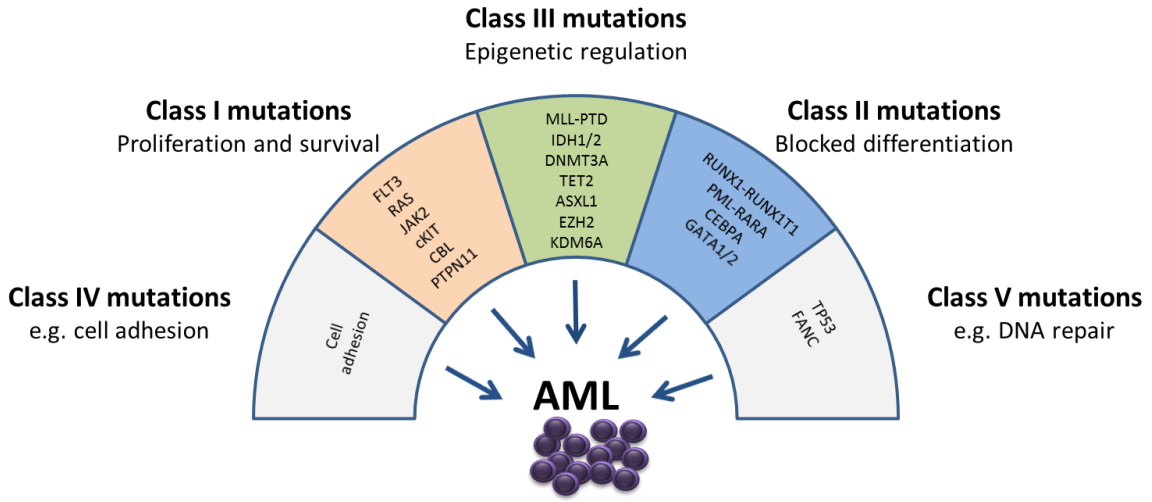


Figure 4: Molecular pathogenesis of AML (adapted from Thiede, 2012). Initially, mutations were only categorized in class I (affecting proliferation) and class II (affecting differentiation). This model was revised after discovery of gene mutations that affect further functional categories.

2. Specific aims and questions

AML is an exceedingly heterogeneous disease on the genetic level (Grimwade et al, 2016; Papaemmanuil et al, 2016; Metzeler et al, 2016). Probably, we will not identify two individuals with AML that are characterized by exactly the same genetic alterations. However, since associations between gene mutations and certain chromosomal aberrations have already been shown, e.g. *KIT* mutations in AML with t(8;21) or inv(16) (Beghini et al, 2000; Care et al, 2003) and *TP53* mutations in AML with complex karyotype (Haferlach et al, 2008), it is worth investigating cytogenetic subgroups of AML in order to identify further patterns of mutational co-occurrence and thereby decipher the genetic heterogeneity. Furthermore, it is of great interest to study the impact of these mutations on a clinical and functional level. Can we improve risk stratification if we include information about gene mutations? Are co-occurring gene mutations just bystanders or how do they contribute to the AML phenotype? This information might be particularly valuable for the design of novel targeted therapies.

The studies presented in this thesis aimed (I) to investigate the mutational landscape of selected cytogenetic subgroups and (II) to evaluate clinical and functional consequences of identified mutations.

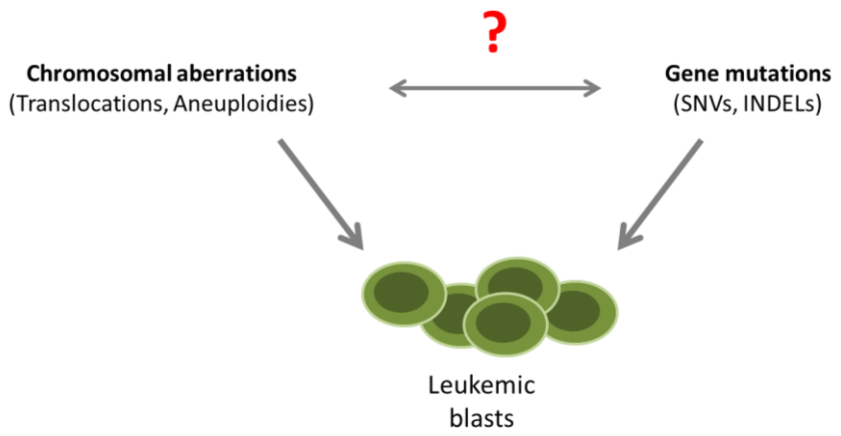


Figure 5: Contribution of chromosomal aberrations and gene mutations to leukemogenesis (adapted from Bochtler et al, 2015). Both types of genomic lesions can lead to leukemia. However, their synergism is not yet fully understood.

3. Summary of results

Paper I: Characterization of AML with trisomy 13

Herold T, Metzeler KH, Vosberg S, **Hartmann L**, Röllig C, Stölzel F, et al. Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis. *Blood*. 2014

Trisomy 13 (+13) as sole aberration is a rare cytogenetic finding in AML with an incidence of <1%. According to ELN and MRC risk stratification, patients with isolated +13 fall into the intermediate risk group. However, previous studies indicated adverse clinical outcome for AML patients with +13.

The aims of the presented study were (I) clinical characterization, (II) mutational profiling and (III) gene expression analysis of AML patients with +13.

Clinical data were available for 34 patients with isolated +13 and 850 patients with other cytogenetic findings that also fall into the same risk group. Patients with +13 were significantly older and had higher blast counts at diagnosis. Moreover, relapse-free survival and overall survival were inferior for the AML +13 group compared with the other intermediate-risk patients.

Exome sequencing of paired diagnostic and remission samples from two patients with +13 identified leukemia-specific mutations in 36 genes, including *RUNX1*, *ASXL1, BCOR, ZRSR2, NUP188* and *CEBPZ*. Next, targeted amplicon sequencing was performed on 16 AML +13 samples, revealing high frequencies of mutations in *RUNX1* (n=12, 75%) and the spliceosome complex (*SRSF2:* 81%, *SF3B1:* 6%, *SF1:* 6% and *ZRSR2:*13%). Moreover, novel mutations in *CEBPZ* were identified. The frequency of *SRSF2* mutations in AML +13 is the highest to be so far reported in any AML subgroup, pointing towards a joint contribution to cell transformation. Similarly, gene expression analysis identified genes that were significantly deregulated in AML +13, including *FLT3* (upregulation) and *SPRY2* (downregulation).

Contribution to this project as co-author:

Confirmation of *CEBPZ*, *ASXL1* and *SRSF2* mutations by Sanger sequencing (Tables S2 and S3, Figure S3), confirmation of somatic status (Figure S3), screening of cytogenetically normal AML (CN-AML) patients for *SRSF2* mutations, manuscript preparation and proof-reading.

Paper II: *ZBTB7A* mutations in t(8;21) positive AML

Hartmann L, Dutta S, Opatz S, Vosberg S, Reiter K, et al. ZBTB7A Mutations in Acute Myeloid Leukemia with t(8;21) Translocation, *Nat Commun*. 2016

The t(8;21) translocation is one of the most frequent chromosomal abnormalities in AML and leads to the fusion gene *RUNX1/RUNX1T1*. However, *in vivo* models indicate the requisite of additional lesions for leukemogenesis as *RUNX1/RUNX1T1* alone is not able to induce leukemia. Exome sequencing of matched diagnostic and remission samples of two patients with t(8;21) rearrangement identified leukemia-specific *ZBTB7A* mutations in both patients. ZBTB7A is a transcriptional repressor and plays a role in normal hematopoiesis. Previous studies indicated that ZBTB7A has both proto-oncogenic and tumor suppressor properties in a tissue-dependent fashion.

The aim of this study were to (I) assess the mutation frequency of *ZBTB7A* mutations in a large cohort of AML patients with t(8;21) translocation, (II) functionally characterize *ZBTB7A* mutations and (III) evaluate the clinical impact of *ZBTB7A* mutations and expression.

Using targeted amplicon sequencing, *ZBTB7A* mutations were identified in 13/56 (23%) of screened *RUNX1/RUNXT1* positive AML patients. Importantly, *ZBTB7A* mutations were not detected in 50 CN-AML patients. Two mutational hotspots (R402 and A175fs) were identified and further characterized on a functional level. The R402 mutations affect the zinc finger structure of ZBTB7A while the A175fs mutation leads to complete loss of the zinc finger domain. DNA pull-down assays and luciferase-based transcription reporter assays indicated that the analyzed *ZBTB7A* mutations lead to loss-of-function. Retroviral expression of wild-type *ZBTB7A* in a *RUNX1/RUNXT1* positive cell line as well as lineage negative murine bone marrow cells (co-expressing *RUNX1/RUNX1T1*) inhibited cell growth, whereas this anti-proliferative effect was lost or weakened upon expression of *ZBTB7A* mutants.

From a clinical perspective, *ZBTB7A* mutations showed no influence on patient outcome. However this evaluation was limited by the relatively small cohort size. Remarkably, in over 200 CN-AML patients treated on a clinical trial (NCT00266136), high expression of *ZBTB7A* was associated with a favorable outcome suggesting a relevance in AML beyond the t(8;21) subgroup.

4. Conclusion and outlook

The two studies presented in this thesis provided novel insides into the biology of acute myeloid leukemia:

-Isolated trisomy 13 is a rare cytogenetic finding but associated with inferior clinical outcome. Consequently, patients with this cytogenetic aberration should be stratified into the group of adverse risk.

-For the first time we have shown that trisomy 13 is associated with a high frequency of *SRSF2* mutations (13 of 16 patients, 81%). SRSF2 is a splicing factor and part of the spliceosome. It was shown that the common *SRSF2* P59H mutation leads to deregulated splicing because of altered RNA-binding affinities (Zhang et al, 2015). How this effect contributes to leukemogenesis and how mutated *SRSF2* and trisomy 13 may collaborate remains to be investigated.

-*ZBTB7A* mutations are a novel finding in AML. Just recently, another group also identified *ZBTB7A* mutations in 3/20 patients with t(8;21) translocation (Lavallée et al, 2016), independently confirming our data. Given the high frequency of these mutations, it is worth analyzing *ZBTB7A* mutations in a larger patient cohort to gain reliable information about the prognostic relevance of *ZBTB7A* mutations. This information can help to refine risk-stratification for t(8;21) positive patients.

-Our data indicates a specific association of *ZBTB7A* mutations and *RUNX1/RUNX1T1* suggesting oncogenic collaboration, however, the underlying mechanism remains elusive.

-ZBTB7A has been reported to act either as a tumor suppressor or oncogene, in a tissue-dependent fashion. The presented study indicates that ZBTB7A functions as a tumor suppressor in AML.

Ideally, therapy of AML could be improved by novel approaches that target one or more cooperating lesions. Since *ZBTB7A* mutations lead to loss of function in AML, therapies would either need to restore ZBTB7A function or reverse the consequences of insufficient ZBTB7A. It was shown that *ZBTB7A* mutations lead to higher glycolytic activity *in vitro* (Liu et al, 2015), thereby increasing tumor metabolism and promote cell proliferation. Consequently, it is attractive to explore if tumor metabolism could be

restricted in *ZBTB7A* mutated AML by treatment with glycolysis inhibitors such as 2-Deoxy-D-glucose (2-DG). For solid tumors, mouse transplantation assays already indicated that 2-DG treatment leads to reduced growth of *ZBTB7A*-knock down cells (Liu et al 2014). Importantly, clinical trials confirmed that the administration of 2-DG alone or combined with other anticancer therapies, such as chemotherapy and radiotherapy was safe and well tolerated by patients with solid tumors (Dwarakanath et al, 2009; Raez et al, 2013). It is therefore worthwhile investigating whether similar effects can also be observed in AML.

In 2013, the cancer genome atlas (TCGA) consortium published a series of 200 AML cases that were comprehensively characterized for gene mutations by either whole genome sequencing (n=50) or exome sequencing (n=150). The cohort comprised adult AML patients representing the major cytomorphologic and cytogenetic subtypes, including 7 patients that were *RUNX1*/*RUNX1T1* positive and a single patient with isolated trisomy 13. A total of 2315 somatic single nucleotide variants (SNV) and 270 small insertions or deletions (INDEL) in coding regions were identified. However, no *ZBTB7A* mutations and only a single *SRSF2* mutation were reported in this patient cohort (the *SRSF2* mutation was not found in the patient with isolated trisomy 13). This highlights that the genetic landscape of AML is still not fully understood and that focused analyses of cytogenetic subgroups is important for the discovery of novel mutations that might play an important role in leukemogenesis and provide the basis for tailored therapies that overcome the poor clinical outcome of patients with AML.

5. References

Akashi, K., D. Traver, T. Miyamoto and I. L. Weissman (2000). "A clonogenic common myeloid progenitor that gives rise to all myeloid lineages." Nature 404(6774): 193-197.

Aplan, P. D. (2006). "Causes of oncogenic chromosomal translocation." Trends Genet 22(1): 46-55.

ar-Rushdi, A., K. Nishikura, J. Erikson, R. Watt, G. Rovera and C. M. Croce (1983). "Differential expression of the translocated and the untranslocated c-myc oncogene in Burkitt lymphoma." Science 222(4622): 390-393.

Bakhoum, S. F. and D. A. Compton (2012). "Chromosomal instability and cancer: a complex relationship with therapeutic potential." J Clin Invest 122(4): 1138-1143.

Becker, H., K. Maharry, K. Mrozek, S. Volinia, A. K. Eisfeld, M. D. Radmacher, J. Kohlschmidt, K. H. Metzeler, S. Schwind, S. P. Whitman, J. H. Mendler, Y. Z. Wu, D. Nicolet, P. Paschka, B. L. Powell, T. H. Carter, M. Wetzler, J. E. Kolitz, A. J. Carroll, M. R. Baer, M. A. Caligiuri, R. M. Stone, G. Marcucci and C. D. Bloomfield (2014). "Prognostic gene mutations and distinct gene- and microRNA-expression signatures in acute myeloid leukemia with a sole trisomy 8." Leukemia 28(8): 1754-1758.

Beghini, A., P. Peterlongo, C. B. Ripamonti, L. Larizza, R. Cairoli, E. Morra and C. Mecucci (2000). "C-kit mutations in core binding factor leukemias." Blood 95(2): 726-727.

Bennett, J. M., D. Catovsky, M. T. Daniel, G. Flandrin, D. A. Galton, H. R. Gralnick and C. Sultan (1976). "Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group." Br J Haematol 33(4): 451-458.

Bochtler, T., S. Frohling and A. Kramer (2015). "Role of chromosomal aberrations in clonal diversity and progression of acute myeloid leukemia." Leukemia 29(6): 1243-1252.

Bonnet, D. and J. E. Dick (1997). "Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell." Nat Med 3(7): 730-737.

Bos, J. L., D. Toksoz, C. J. Marshall, M. Verlaan-de Vries, G. H. Veeneman, A. J. van der Eb, J. H. van Boom, J. W. Janssen and A. C. Steenvoorden (1985). "Amino-acid substitutions at codon 13 of the N-ras oncogene in human acute myeloid leukaemia." Nature 315(6022): 726-730.

Burnett, A., M. Wetzler and B. Lowenberg (2011). "Therapeutic advances in acute myeloid leukemia." J Clin Oncol 29(5): 487-494.

Care, R. S., P. J. Valk, A. C. Goodeve, F. M. Abu-Duhier, W. M. Geertsma-Kleinekoort, G. A. Wilson, M. A. Gari, I. R. Peake, B. Lowenberg and J. T. Reilly (2003). "Incidence and prognosis of c-KIT and FLT3 mutations in core binding factor (CBF) acute myeloid leukaemias." Br J Haematol 121(5): 775-777.

de The, H., C. Lavau, A. Marchio, C. Chomienne, L. Degos and A. Dejean (1991). "The PML-RAR alpha fusion mRNA generated by the t(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR." Cell 66(4): 675-684.

Döhner, H., E. H. Estey, S. Amadori, F. R. Appelbaum, T. Buchner, A. K. Burnett, H. Dombret, P. Fenaux, D. Grimwade, R. A. Larson, F. Lo-Coco, T. Naoe, D. Niederwieser, G. J. Ossenkoppele, M. A. Sanz, J. Sierra, M. S. Tallman, B. Lowenberg and C. D. Bloomfield (2010). "Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet." Blood 115(3): 453-474.

Döhner, H., D. J. Weisdorf and C. D. Bloomfield (2015). "Acute Myeloid Leukemia." N Engl J Med 373(12): 1136-1152.

Dolnik, A., J. C. Engelmann, M. Scharfenberger-Schmeer, J. Mauch, S. Kelkenberg-Schade, B. Haldemann, T. Fries, J. Kronke, M. W. Kuhn, P. Paschka, S. Kayser, S. Wolf, V. I. Gaidzik, R. F. Schlenk, F. G. Rucker, H. Dohner, C. Lottaz, K. Dohner and L. Bullinger (2012). "Commonly altered genomic regions in acute myeloid leukemia are enriched for somatic mutations involved in chromatin remodeling and splicing." Blood 120(18): e83-92.

Dwarakanath, B. S., D. Singh, A. K. Banerji, R. Sarin, N. K. Venkataramana, R. Jalali, P. N. Vishwanath, B. K. Mohanti, R. P. Tripathi, V. K. Kalia and V. Jain (2009). "Clinical studies for improving radiotherapy with 2-deoxy-D-glucose: present status and future prospects." J Cancer Res Ther 5 Suppl 1: S21-26.

Erickson, P., J. Gao, K. S. Chang, T. Look, E. Whisenant, S. Raimondi, R. Lasher, J. Trujillo, J. Rowley and H. Drabkin (1992). "Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt." Blood 80(7): 1825-1831.

Estey, E. and H. Dohner (2006). "Acute myeloid leukaemia." Lancet 368(9550): 1894-1907.

Falini, B., C. Mecucci, E. Tiacci, M. Alcalay, R. Rosati, L. Pasqualucci, R. La Starza, D. Diverio, E. Colombo, A. Santucci, B. Bigerna, R. Pacini, A. Pucciarini, A. Liso, M. Vignetti, P. Fazi, N. Meani, V. Pettirossi, G. Saglio, F. Mandelli, F. Lo-Coco, P. G. Pelicci and M. F. Martelli (2005). "Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype." N Engl J Med 352(3): 254-266.

Fiedler, K. and Brunner, C (2012). "Mechanisms Controlling Hematopoiesis, Hematology - Science and Practice.", Dr. Charles Lawrie (Ed.), ISBN: 978-953-51-0174-1, InTech

Figueroa, M. E., O. Abdel-Wahab, C. Lu, P. S. Ward, J. Patel, A. Shih, Y. Li, N. Bhagwat, A. Vasanthakumar, H. F. Fernandez, M. S. Tallman, Z. Sun, K. Wolniak, J. K. Peeters, W. Liu, S. E. Choe, V. R. Fantin, E. Paietta, B. Lowenberg, J. D. Licht, L. A. Godley, R. Delwel, P. J. Valk, C. B. Thompson, R. L. Levine and A. Melnick (2010). "Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation." Cancer Cell 18(6): 553-567.

Freireich, E. J., P. H. Wiernik and D. P. Steensma (2014). "The leukemias: a half-century of discovery." J Clin Oncol 32(31): 3463-3469.

Goyama, S. and J. C. Mulloy (2011). "Molecular pathogenesis of core binding factor leukemia: current knowledge and future prospects." Int J Hematol 94(2): 126-133.

Greif, P. A., A. Dufour, N. P. Konstandin, B. Ksienzyk, E. Zellmeier, B. Tizazu, J. Sturm, T. Benthaus, T. Herold, M. Yaghmaie, P. Dorge, K. P. Hopfner, A. Hauser, A. Graf, S. Krebs, H. Blum, P. M. Kakadia, S. Schneider, E. Hoster, F. Schneider, M. Stanulla, J. Braess, M. C. Sauerland, W. E. Berdel, T. Buchner, B. J. Woermann, W. Hiddemann, K. Spiekermann and S. K. Bohlander (2012). "GATA2 zinc finger 1 mutations associated with biallelic CEBPA mutations define a unique genetic entity of acute myeloid leukemia." Blood 120(2): 395-403.

Grimwade, D., R. K. Hills, A. V. Moorman, H. Walker, S. Chatters, A. H. Goldstone, K. Wheatley, C. J. Harrison and A. K. Burnett (2010). "Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials." Blood 116(3): 354-365.

Grimwade, D., A. Ivey and B. J. Huntly (2016). "Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance." Blood 127(1): 29-41.

Gröschel, S., M. A. Sanders, R. Hoogenboezem, E. de Wit, B. A. Bouwman, C. Erpelinck, V. H. van der Velden, M. Havermans, R. Avellino, K. van Lom, E. J. Rombouts, M. van Duin, K. Dohner, H. B. Beverloo, J. E. Bradner, H. Dohner, B. Lowenberg, P. J. Valk, E. M. Bindels, W. de Laat and R. Delwel (2014). "A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia." Cell 157(2): 369-381.

Grossmann, V., E. Tiacci, A. B. Holmes, A. Kohlmann, M. P. Martelli, W. Kern, A. Spanhol-Rosseto, H. U. Klein, M. Dugas, S. Schindela, V. Trifonov, S. Schnittger, C. Haferlach, R. Bassan, V. A. Wells, O. Spinelli, J. Chan, R. Rossi, S. Baldoni, L. De Carolis, K. Goetze, H. Serve, R. Peceny, K. A. Kreuzer, D. Oruzio, G. Specchia, F. Di Raimondo, F. Fabbiano, M. Sborgia, A. Liso, L. Farinelli, A. Rambaldi, L. Pasqualucci, R. Rabadan, T. Haferlach and B. Falini (2011). "Whole-exome sequencing identifies somatic mutations of BCOR in acute myeloid leukemia with normal karyotype." Blood 118(23): 6153-6163.

Haferlach, C., F. Dicker, H. Herholz, S. Schnittger, W. Kern and T. Haferlach (2008). "Mutations of the TP53 gene in acute myeloid leukemia are strongly associated with a complex aberrant karyotype." Leukemia 22(8): 1539-1541.

Hermans, A., N. Heisterkamp, M. von Linden, S. van Baal, D. Meijer, D. van der Plas, L. M. Wiedemann, J. Groffen, D. Bootsma and G. Grosveld (1987). "Unique fusion of bcr and c-abl genes in Philadelphia chromosome positive acute lymphoblastic leukemia." Cell 51(1): 33-40.

Horton, S. J. and B. J. Huntly (2012). "Recent advances in acute myeloid leukemia stem cell biology." Haematologica 97(7): 966-974.

Janssen, A., M. van der Burg, K. Szuhai, G. J. Kops and R. H. Medema (2011). "Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations." Science 333(6051): 1895-1898.

Juliusson, G., P. Antunovic, A. Derolf, S. Lehmann, L. Mollgard, D. Stockelberg, U. Tidefelt, A. Wahlin and M. Hoglund (2009). "Age and acute myeloid leukemia: real world data on decision to treat and outcomes from the Swedish Acute Leukemia Registry." Blood 113(18): 4179-4187.

Kampen, K. R. (2012). "The discovery and early understanding of leukemia." Leuk Res 36(1): 6-13.

Kondo, M., I. L. Weissman and K. Akashi (1997). "Identification of clonogenic common lymphoid progenitors in mouse bone marrow." Cell 91(5): 661-672.

Larsson, C. A., G. Cote and A. Quintas-Cardama (2013). "The changing mutational landscape of acute myeloid leukemia and myelodysplastic syndrome." Mol Cancer Res 11(8): 815-827.

Lavallee, V. P., S. Lemieux, G. Boucher, P. Gendron, I. Boivin, R. N. Armstrong, G. Sauvageau and J. Hebert (2016). "RNA-sequencing analysis of core binding factor AML identifies recurrent ZBTB7A mutations and defines RUNX1-CBFA2T3 fusion signature." Blood 127(20): 2498-2501.

Ley, T. J., L. Ding, M. J. Walter, M. D. McLellan, T. Lamprecht, D. E. Larson, C. Kandoth, J. E. Payton, J. Baty, J. Welch, C. C. Harris, C. F. Lichti, R. R. Townsend, R. S. Fulton, D. J. Dooling, D. C. Koboldt, H. Schmidt, Q. Zhang, J. R. Osborne, L. Lin, M. O'Laughlin, J. F. McMichael, K. D. Delehaunty, S. D. McGrath, L. A. Fulton, V. J. Magrini, T. L. Vickery, J. Hundal, L. L. Cook, J. J. Conyers, G. W. Swift, J. P. Reed, P. A. Alldredge, T. Wylie, J. Walker, J. Kalicki, M. A. Watson, S. Heath, W. D. Shannon, N. Varghese, R. Nagarajan, P. Westervelt, M. H. Tomasson, D. C. Link, T. A. Graubert, J. F. DiPersio, E. R. Mardis and R. K. Wilson (2010). "DNMT3A mutations in acute myeloid leukemia." N Engl J Med 363(25): 2424-2433.

Ley, T. J., E. R. Mardis, L. Ding, B. Fulton, M. D. McLellan, K. Chen, D. Dooling, B. H. Dunford-Shore, S. McGrath, M. Hickenbotham, L. Cook, R. Abbott, D. E. Larson, D. C. Koboldt, C. Pohl, S. Smith, A. Hawkins, S. Abbott, D. Locke, L. W. Hillier, T.

Miner, L. Fulton, V. Magrini, T. Wylie, J. Glasscock, J. Conyers, N. Sander, X. Shi, J. R. Osborne, P. Minx, D. Gordon, A. Chinwalla, Y. Zhao, R. E. Ries, J. E. Payton, P. Westervelt, M. H. Tomasson, M. Watson, J. Baty, J. Ivanovich, S. Heath, W. D. Shannon, R. Nagarajan, M. J. Walter, D. C. Link, T. A. Graubert, J. F. DiPersio and R. K. Wilson (2008). "DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome." Nature 456(7218): 66-72.

Liu, X. S., J. E. Haines, E. K. Mehanna, M. D. Genet, I. Ben-Sahra, J. M. Asara, B. D. Manning and Z. M. Yuan (2014). "ZBTB7A acts as a tumor suppressor through the transcriptional repression of glycolysis." Genes Dev 28(17): 1917-1928.

Liu, X. S., Z. Liu, C. Gerarduzzi, D. E. Choi, S. Ganapathy, P. P. Pandolfi and Z. M. Yuan (2015). "Somatic human ZBTB7A zinc finger mutations promote cancer progression." Oncogene.

Meyer, S. C. and R. L. Levine (2014). "Translational implications of somatic genomics in acute myeloid leukaemia." Lancet Oncol 15(9): e382-394.

Metzeler, K. H., T. Herold, M. Rothenberg-Thurley, S. Amler, M. C. Sauerland, D. Goerlich, S. Schneider, N. P. Konstandin, A. Dufour, K. Braundl, B. Ksienzyk, E. Zellmeier, L. Hartmann, P. A. Greif, M. Fiegl, M. Subklewe, S. K. Bohlander, U. Krug, A. Faldum, W. E. Berdel, B. Wormann, T. Buchner, W. Hiddemann, J. Braess and K. Spiekermann (2016). "Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia." Blood (in press).

Micol, J. B., N. Duployez, N. Boissel, A. Petit, S. Geffroy, O. Nibourel, C. Lacombe, H. Lapillonne, P. Etancelin, M. Figeac, A. Renneville, S. Castaigne, G. Leverger, N. Ifrah, H. Dombret, C. Preudhomme, O. Abdel-Wahab and E. Jourdan (2014). "Frequent ASXL2 mutations in acute myeloid leukemia patients with t(8;21)/RUNX1-RUNX1T1 chromosomal translocations." Blood 124(9): 1445-1449.

Okuda, T., J. van Deursen, S. W. Hiebert, G. Grosveld and J. R. Downing (1996). "AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis." Cell 84(2): 321-330.

Papaemmanuil, E., M. Gerstung, L. Bullinger, V. I. Gaidzik, P. Paschka, N. D. Roberts, N. E. Potter, M. Heuser, F. Thol, N. Bolli, G. Gundem, P. Van Loo, I. Martincorena, P. Ganly, L. Mudie, S. McLaren, S. O'Meara, K. Raine, D. R. Jones, J. W. Teague, A. P. Butler, M. F. Greaves, A. Ganser, K. Dohner, R. F. Schlenk, H. Dohner and P. J. Campbell (2016). "Genomic Classification and Prognosis in Acute Myeloid Leukemia." N Engl J Med 374(23): 2209-2221.

Pouzolles, M., L. Oburoglu, N. Taylor and V. S. Zimmermann (2016). "Hematopoietic stem cell lineage specification." Curr Opin Hematol 23(4): 311-317.

Raez, L. E., K. Papadopoulos, A. D. Ricart, E. G. Chiorean, R. S. Dipaola, M. N. Stein, C. M. Rocha Lima, J. J. Schlesselman, K. Tolba, V. K. Langmuir, S. Kroll, D. T. Jung, M. Kurtoglu, J. Rosenblatt and T. J. Lampidis (2013). "A phase I dose-

escalation trial of 2-deoxy-D-glucose alone or combined with docetaxel in patients with advanced solid tumors." Cancer Chemother Pharmacol 71(2): 523-530.

Rhoades, K. L., C. J. Hetherington, N. Harakawa, D. A. Yergeau, L. Zhou, L. Q. Liu, M. T. Little, D. G. Tenen and D. E. Zhang (2000). "Analysis of the role of AML1-ETO in leukemogenesis, using an inducible transgenic mouse model." Blood 96(6): 2108-2115.

Rowley, J. D. (1973). "Identificaton of a translocation with quinacrine fluorescence in a patient with acute leukemia." Ann Genet 16(2): 109-112.

Russler-Germain, D. A., D. H. Spencer, M. A. Young, T. L. Lamprecht, C. A. Miller, R. Fulton, M. R. Meyer, P. Erdmann-Gilmore, R. R. Townsend, R. K. Wilson and T. J. Ley (2014). "The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers." Cancer Cell 25(4): 442-454.

Solh, M., S. Yohe, D. Weisdorf and C. Ustun (2014). "Core-binding factor acute myeloid leukemia: Heterogeneity, monitoring, and therapy." Am J Hematol 89(12): 1121-1131.

Speck, N. A. and D. G. Gilliland (2002). "Core-binding factors in haematopoiesis and leukaemia." Nat Rev Cancer 2(7): 502-513.

Tan, B. T., C. Y. Park, L. E. Ailles and I. L. Weissman (2006). "The cancer stem cell hypothesis: a work in progress." Lab Invest 86(12): 1203-1207.

Tanaka, T., K. Tanaka, S. Ogawa, M. Kurokawa, K. Mitani, J. Nishida, Y. Shibata, Y. Yazaki and H. Hirai (1995). "An acute myeloid leukemia gene, AML1, regulates hemopoietic myeloid cell differentiation and transcriptional activation antagonistically by two alternative spliced forms." Embo j 14(2): 341-350.

Tenen, D. G. (2003). "Disruption of differentiation in human cancer: AML shows the way." Nat Rev Cancer 3(2): 89-101.

The Cancer Genome Atlas (2013). "Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia." N Engl J Med 368(22): 2059-2074.

Thiede, C. (2012). "Impact of mutational analysis in acute myeloid leukemia." Hematology Education: the education program for the annual congress of the European Hematology Association; 6:33-40

Vardiman, J. W., J. Thiele, D. A. Arber, R. D. Brunning, M. J. Borowitz, A. Porwit, N. L. Harris, M. M. Le Beau, E. Hellstrom-Lindberg, A. Tefferi and C. D. Bloomfield (2009). "The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes." Blood 114(5): 937-951.

Wang, E. S. (2014). "Treating acute myeloid leukemia in older adults." Hematology Am Soc Hematol Educ Program 2014(1): 14-20.

Welch, J. S. and D. C. Link (2011). "Genomics of AML: clinical applications of next-generation sequencing." Hematology Am Soc Hematol Educ Program 2011: 30-35.

Westendorf, J. J., C. M. Yamamoto, N. Lenny, J. R. Downing, M. E. Selsted and S. W. Hiebert (1998). "The t(8;21) fusion product, AML-1-ETO, associates with C/EBP-alpha, inhibits C/EBP-alpha-dependent transcription, and blocks granulocytic differentiation." Mol Cell Biol 18(1): 322-333.

Wiemels, J. L., Z. Xiao, P. A. Buffler, A. T. Maia, X. Ma, B. M. Dicks, M. T. Smith, L. Zhang, J. Feusner, J. Wiencke, K. Pritchard-Jones, H. Kempski and M. Greaves (2002). "In utero origin of t(8;21) AML1-ETO translocations in childhood acute myeloid leukemia." Blood 99(10): 3801-3805.

Wilson, N. K., S. D. Foster, X. Wang, K. Knezevic, J. Schutte, P. Kaimakis, P. M. Chilarska, S. Kinston, W. H. Ouwehand, E. Dzierzak, J. E. Pimanda, M. F. de Bruijn and B. Gottgens (2010). "Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators." Cell Stem Cell 7(4): 532-544.

Yuan, Y., L. Zhou, T. Miyamoto, H. Iwasaki, N. Harakawa, C. J. Hetherington, S. A. Burel, E. Lagasse, I. L. Weissman, K. Akashi and D. E. Zhang (2001). "AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations." Proc Natl Acad Sci U S A 98(18): 10398-10403.

Zhang, J., Y. K. Lieu, A. M. Ali, A. Penson, K. S. Reggio, R. Rabadan, A. Raza, S. Mukherjee and J. L. Manley (2015). "Disease-associated mutation in SRSF2 misregulates splicing by altering RNA-binding affinities." Proc Natl Acad Sci U S A 112(34): E4726-4734.

6. Acknowledgements

The acknowledgements have been removed in the publicly accessible version of this document.

7. Curriculum vitae – Luise Hartmann

1987	Born April 6th 1987 in Hannover (Germany)
2003	Mittlere Reife (Secondary school graduation) Johann-Winklhofer Realschule Landsberg am Lech
2005	Fachhochschulreife (Technical college entrance qualification) Staatliche Fachoberschule Landsberg am Lech
2005-2010	University of Applied Sciences Weihenstephan-Triesdorf Course: Biotechnology, degree: Diplom-Ingenieur (FH)
2010-2012	Technical University of Munich (TUM) Course: Biology, degree: Master of Science (passed with high distinction)
Since 2013	Doctoral student German Cancer Consortium (DKTK), partner site Munich Medizinischen Klinik und Poliklinik III - Großhadern Ludwig-Maximilians-Universität München Supervisors: Prof. Dr. med. Karsten Spiekermann and Dr. med. Philipp Greif

Publications

Metzeler, K. H., T. Herold, M. Rothenberg-Thurley, S. Amler, M. C. Sauerland, D. Goerlich, S. Schneider, N. P. Konstandin, A. Dufour, K. Braundl, B. Ksienzyk, E. Zellmeier, **L. Hartmann**, P. A. Greif, M. Fiegl, M. Subklewe, S. K. Bohlander, U. Krug, A. Faldum, W. E. Berdel, B. Wormann, T. Buchner, W. Hiddemann, J. Braess and K. Spiekermann (2016). "Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia." Blood (in press).

Hartmann, L., S. Dutta, S. Opatz, S. Vosberg, K. Reiter, G. Leubolt, K. H. Metzeler, T. Herold, S. A. Bamopoulos, K. Braundl, E. Zellmeier, B. Ksienzyk, N. P. Konstandin, S. Schneider, K. P. Hopfner, A. Graf, S. Krebs, H. Blum, J. M. Middeke, F. Stolzel, C. Thiede, S. Wolf, S. K. Bohlander, C. Preiss, L. Chen-Wichmann, C. Wichmann, M. C. Sauerland, T. Buchner, W. E. Berdel, B. J. Wormann, J. Braess, W. Hiddemann, K. Spiekermann and P. A. Greif (2016). "ZBTB7A mutations in acute myeloid leukaemia with t(8;21) translocation." Nat Commun 7: 11733.

Vosberg, S., T. Herold, **L. Hartmann**, M. Neumann, S. Opatz, K. H. Metzeler, S. Schneider, A. Graf, S. Krebs, H. Blum, C. D. Baldus, W. Hiddemann, K. Spiekermann, S. K. Bohlander, U. Mansmann and P. A. Greif (2016). "Close correlation of copy number aberrations detected by next-generation sequencing with results from routine cytogenetics in acute myeloid leukemia." Genes Chromosomes Cancer 55(7): 553-567.

Herold, T., K. H. Metzeler, S. Vosberg, **L. Hartmann**, C. Rollig, F. Stolzel, S. Schneider, M. Hubmann, E. Zellmeier, B. Ksienzyk, V. Jurinovic, Z. Pasalic, P. M. Kakadia, A. Dufour, A. Graf, S. Krebs, H. Blum, M. C. Sauerland, T. Buchner, W. E. Berdel, B. J. Woermann, M. Bornhauser, G. Ehninger, U. Mansmann, W. Hiddemann, S. K. Bohlander, K. Spiekermann and P. A. Greif (2014). "Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis." Blood 124(8): 1304-1311.

Hartmann, L., C. F. Stephenson, S. R. Verkamp, K. R. Johnson, B. Burnworth, K. Hammock, L. E. Brodersen, M. E. de Baca, D. A. Wells, M. R. Loken and B. K. Zehentner (2014). "Detection of clonal evolution in hematopoietic malignancies by combining comparative genomic hybridization and single nucleotide polymorphism arrays." Clin Chem 60(12): 1558-1568.

Zehentner, B. K., **L. Hartmann**, K. R. Johnson, C. F. Stephenson, D. B. Chapman, M. E. de Baca, D. A. Wells, M. R. Loken, B. Tirtorahardjo, S. R. Gunn and L. Lim (2012). "Array-based karyotyping in plasma cell neoplasia after plasma cell enrichment increases detection of genomic aberrations." Am J Clin Pathol 138(4): 579-589.

Hartmann, L., J. S. Biggerstaff, D. B. Chapman, J. M. Scott, K. R. Johnson, K. M. Ghirardelli, W. K. Fritschle, D. L. Martinez, R. K. Bennington, M. E. de Baca, D. A. Wells, M. R. Loken and B. K. Zehentner (2011). "Detection of genomic abnormalities in multiple myeloma: the application of FISH analysis in combination with various plasma cell enrichment techniques." Am J Clin Pathol 136(5): 712-720.

Awards

2014 ASH Abstract Achievement Award
Abstract #17 ‘Genetic Evolution of Cytogenetically Normal Acute Myeloid Leukemia (CN-AML) during Therapy and Relapse: An Exome Sequencing Study of 47 Cases’, selected for oral presentation.

2015 ASH Abstract Achievement Award
Abstract #690 ‘Mutations of Genes Linked to Epigenetic Regulation Are Frequently Gained in Relapsed Cytogenetically Normal Acute Myeloid Leukemia’, selected for oral presentation.

2016 EHA Travel Grant
Abstract #S119 ‘Frequent Recurring Mutations Disrupt the Anti-Proliferative Function of ZBTB7A in Acute Myeloid Leukemia with t(8;21) Translocation’, selected for oral presentation.

Appendix:

Paper I

Paper II

Regular Article

MYELOID NEOPLASIA

Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis

Tobias Herold,[1-4] Klaus H. Metzeler,[1-4] Sebastian Vosberg,[1-4] Luise Hartmann,[1-4] Christoph Röllig,[3-5] Friedrich Stölzel,[5] Stephanie Schneider,[1] Max Hubmann,[1,2] Evelyn Zellmeier,[1] Bianka Ksienzyk,[1] Vindi Jurinovic,[6] Zlatana Pasalic,[1] Purvi M. Kakadia,[7] Annika Dufour,[1] Alexander Graf,[8] Stefan Krebs,[8] Helmut Blum,[8] Maria Cristina Sauerland,[9] Thomas Büchner,[10] Wolfgang E. Berdel,[10] Bernhard J. Woermann,[11] Martin Bornhäuser,[3-5] Gerhard Ehninger,[3-5] Ulrich Mansmann,[3,4,6] Wolfgang Hiddemann,[1-4] Stefan K. Bohlander,[12] Karsten Spiekermann,[1-4] and Philipp A. Greif[1-4]

[1]Department of Internal Medicine 3, University Hospital Grosshadern, Ludwig-Maximilians-Universität, Munich, Germany; [2]Clinical Cooperative Group Leukemia, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany; [3]German Cancer Consortium (DKTK), Heidelberg, Germany; [4]German Cancer Research Center (DKFZ), Heidelberg, Germany; [5]Medizinische Klinik und Poliklinik I, Universitätsklinikum Dresden, Dresden, Germany; [6]Institute for Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany; [7]Center for Human Genetics, Philipps University, Marburg, Germany; [8]Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, Ludwig-Maximilians-Universität, Munich, Germany; [9]Institute of Biostatistics and Clinical Research, and [10]Department of Medicine A - Hematology, Oncology and Pneumology, University of Münster, Münster, Germany; [11]German Society of Hematology and Oncology, Berlin, Germany; and [12]Department of Molecular Medicine and Pathology, The University of Auckland, Auckland, New Zealand

Key Points

- AML patients with isolated trisomy 13 have a very poor clinical outcome
- Isolated trisomy 13 in AML is associated with a high frequency of mutations in *SRSF2* (81%) and *RUNX1* (75%)

In acute myeloid leukemia (AML), isolated trisomy 13 (AML+13) is a rare chromosomal abnormality whose prognostic relevance is poorly characterized. We analyzed the clinical course of 34 AML+13 patients enrolled in the German AMLCG-1999 and SAL trials and performed exome sequencing, targeted candidate gene sequencing and gene expression profiling. Relapse-free (RFS) and overall survival (OS) of AML+13 patients were inferior compared to other ELN Intermediate-II patients (n=855) (median RFS, 7.8 vs 14.1 months, *P* = .006; median OS 9.3 vs. 14.8 months, *P* = .004). Besides the known high frequency of *RUNX1* mutations (75%), we identified mutations in spliceosome components in 88%, including *SRSF2* codon 95 mutations in 81%. Recurring mutations were detected in *ASXL1* (44%) and *BCOR* (25%). Two patients carried mutations in *CEBPZ*, suggesting that *CEBPZ* is a novel recurrently mutated gene in AML. Gene expression analysis revealed a homogeneous expression profile including upregulation of *FOXO1* and *FLT3* and downregulation of *SPRY2*. This is the most comprehensive clinical and biological characterization of AML+13 to date, and reveals a striking clustering of lesions in a few genes, defining AML+13 as a genetically homogeneous subgroup with alterations in a few critical cellular pathways. Clinicaltrials.gov identifiers: AMLCG-1999: NCT00266136; AML96: NCT00180115; AML2003: NCT00180102; and AML60+: NCT00893373 (*Blood*. 2014;124(8):1304-1311)

Introduction

Acquired isolated trisomy 13 (+13) is a rare cytogenetic alteration in acute myeloid leukemia (AML). In a retrospective study of 22 856 AML patients from the Mayo Clinic, its incidence was 0.7%.[1] So far, the prognostic relevance of AML+13 has not been extensively studied, but assumed to be unfavorable based on small or heterogeneous patient cohorts.[2-4] However, according to the European LeukemiaNet (ELN) classification, AML+13 is currently classified in the Intermediate-II genetic group.[5] AML+13 is frequently associated with FAB M0 morphology and shows a high frequency (80% to 100%) of *RUNX1* mutations.[6,7] Overexpression of *FLT3* (located in band q12 on chromosome 13) due to a gene dosage effect was proposed as a potential mechanism of leukemogenesis in AML+13.[6,7] The possibility that AML+13 might be a marker for treatment response to lenalidomide has recently been raised.[8]

Constitutional aneuploidy is linked to increased cancer risk.[9] For example, Down syndrome (trisomy 21) predisposes to megakaryoblastic leukemia with a high frequency of acquired *GATA1* mutations.[10] Trisomy 13 (Patau syndrome) is a severe congenital disorder with cerebral, cardiac, and renal malformations.[11] An association of Patau syndrome and solid neoplasms including neuroblastoma and nephroblastoma was reported.[12] In the literature, we found a single case report of Patau syndrome with congenital myeloid leukemia.[13]

Submitted December 1, 2013; accepted May 28, 2014. Prepublished online as *Blood* First Edition paper, June 12, 2014; DOI 10.1182/blood-2013-12-540716.

Presented in abstract form at the 55th annual meeting of the American Society of Hematology, New Orleans, LA, December 7-10, 2013.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

Table 1. Patient characteristics

Variable	AML+13*	Control Group*	*P*
No. of patients	34	850	
Median age, years *(range)*	64 (43-80)	59 (17-84)	**.004**
Male sex, no. *(%)*	24 (70)	465 (55)	.08
WBC count, G/l, median *(range)*	10 (1-318)	11 (0.1-365)	.64
Hemoglobin, g/dl, median *(range)*	8.9 (4.6-12.8)	9.2 (2.9-17.2)	.2
Platelet count, G/l, median *(range)*	77 (1-399)	54 (1-1760)	.23
LDH (U/l), median *(range)*	269 (155-1011)	414 (115-11140)	**.009**
BM blasts, %, median *(range)*	80 (11-100)	68 (11-100)	**.02**
BM blasts at day 16, %, median *(range)*	5 (0-85)	9 (0-100)	.78
Performance status (ECOG) ≥ 2 *(%)*	8 (26)	263 (34)	.44
de novo AML *(%)*	26 (76)	646 (76)	1.0
Allogeneic transplantation, no. *(%)*	6 (18)	180 (21)	.83
CR, no. *(%)*	21 (62)	471 (55)	.49
Relapse, no. *(%)*	18 (86)	327 (69)	.14
Deceased, no. *(%)*	31 (91)	644 (76)	**.04**

Significant *P* values are indicated in bold.

*All patients were enrolled in the AMLCG-99 or SAL trials and received intensive induction treatment. All patients are classified as ELN Intermediate-II; AML+13: patients with isolated tri- or tetrasomy 13, additional aberrations of the sex chromosomes are allowed.

Considering that the vast majority of infants with Patau syndrome die before 1 year of age,[11] it remains unclear whether constitutional trisomy 13 predisposes to myeloid neoplasia.

We set out to characterize the clinical course of AML+13 patients and to elucidate the underlying spectrum of molecular genetic changes by exome sequencing, targeted sequencing, and gene expression profiling.

Materials and methods

Patients

In this analysis, a subgroup of patients enrolled in the German AML Cooperative Group (AMLCG) (NCT00266136) multicenter AMLCG-1999 trial, and the AML96, AML2003, and AML60+ trials of the Study Alliance Leukemia (SAL) was studied (for details, see supplemental Figure 1A-B on the *Blood* Web site).[14-17] All patients received intensive induction chemotherapy as described elsewhere.[14-17] The AMLCG and SAL clinical trials were approved by the local institutional review boards of all participating centers and informed consent was obtained from all patients in accordance with the Declaration of Helsinki.

Exome sequencing

To perform exome sequencing, genomic DNA of available paired diagnostic and remission samples was extracted from archived bone marrow (BM) samples and fragmented for library preparation as described previously.[18,19] Protein-coding regions were enriched using the SureSelect Human All Exon V4 Kit (Agilent), followed by multiplexed 80 bp paired-end sequencing on an Illumina Genome Analyzer IIx. In total, at least 3.2 Gb of raw sequence data were generated per sample (mean 3.5 Gb; quality metrics are summarized in supplemental Table 1). Raw sequence reads were filtered by Illumina's chastity filter and mapped to the NCBI human hg19 RefSeq reference genome using BWA mapper with default parameters.[20] Insufficiently mapped sequence reads (cutoff Q13, according to 95% confidence of correct mapping) and polymerase chain reaction (PCR) duplicate reads were removed using SAMtools[21]; realignment of mapped reads was performed using the Genome Analysis Toolkit to reduce false-positive single nucleotide variant calls.[22] Candidates for somatically acquired mutations were detected using VarScan with the following parameters: coverage ≥ 10×, variant allele frequency ≥ 20%, variant base calling quality ≥ Q13, and variant reads ≥ 3.[23] Positions with evidence for a variant in the corresponding remission sample or annotated polymorphism (as listed in dbSNP v135) were excluded.

Targeted amplicon sequencing

A selection of genes identified by exome sequencing (n = 9) and a panel of genes recurringly mutated in AML (n = 42) were studied by targeted amplicon sequencing (Haloplex; Agilent) in all AMLCG AML+13 patients with available material (16 of 23). The resulting libraries were sequenced in a single run on a MiSeq instrument. Sequence data were aligned to the human reference genome (version hg19) using BWA.[20] Single nucleotide variants and short insertions or deletions were called using VarScan 2 and Pindel, respectively.[24,25]

In addition, Sanger sequencing of genomic DNA was performed for additional validation of selected mutations. Primer sequences and PCR conditions (for *SRSF2*) are shown in supplemental Tables 2 and 3). PCR products were purified using NucleoFast 96 PCR Clean-up Kit (Macherey Nagel, Düren, Germany) and bi-directional sequencing was performed on an ABI 3500xL Genetic Analyzer using the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Sequences were aligned and compared with the reference sequences (NCBI accession numbers: NC_000002.11 [CEBPZ], NG_027868.1 [ASXL1], and NG_032905.1 [SRSF2]) using the Sequencher software (Gene Codes Corporation, Ann Arbor, MI)

Gene expression analysis

To further characterize the AML+13 subgroup, we compared gene expression profiles of 9 patients with AML+13 to 509 AML patients with various genetic abnormalities (except for numerical alterations affecting chromosome 13). The gene expression data set was published previously and is publicly available through the Gene Expression Omnibus Web site (GSE37642).[26] Eight of 9 patients were also included in the genetic analysis. Details of sample preparation, hybridization, and image acquisition were described previously.[26] For probe-to-probe set summarization, we used custom chip definition files based on GeneAnnot version 2.0 (available at http://www.xlab.unimo.it/GA_CDF/) as reported before.[18] Only the 17 389 probe sets present on both the Affymetrix HG-U133A and B chips, and the HG-U133 plus 2.0 chips were included in the analysis. To eliminate the batch effect resulting from the use of different chip designs, we applied an empirical Bayesian method as described previously.[27]

Gene set enrichment analysis (GSEA) was performed with GSEA software (MIT) using the "c5_all" collection consisting of 1454 gene sets derived from the controlled vocabulary of the Gene Ontology project.[28]

The Linear Models for Microarray Data package was used to compute differentially regulated probe sets. Differential regional gene expression on chromosome 13 was analyzed using MACAT (MicroArray Chromosome Analysis Tool) as described previously.[29,30]

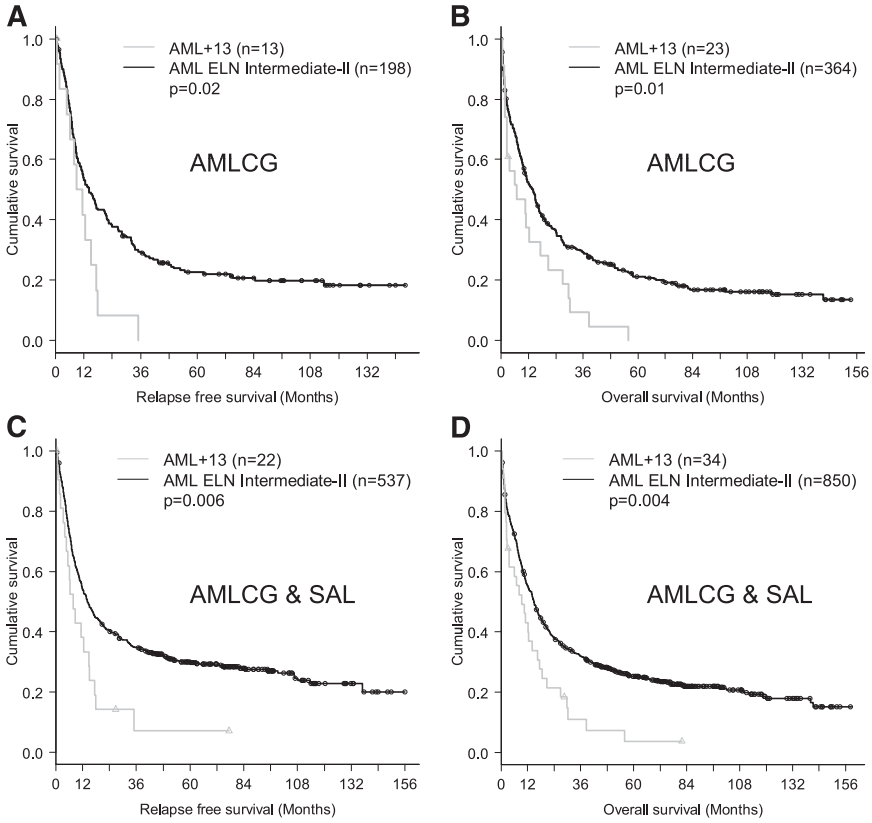


Figure 1. RFS and OS in AML patients. (A-B) AMLCG cohort. (C-D) Combined AMLCG and SAL cohort. Kaplan–Meier estimates of RFS and OS are significantly reduced for the AML+13 subgroup within the ELN Intermediate-II genetic group.

Statistical analyses

All statistical analyses were performed using the R 2.12.2 and 3.0.1 software[31] and routines from the biostatistics software repository Bioconductor, and SPSS version 21.0 (SPSS Inc., Chicago, IL). Two-sided Fisher's exact test was used to compare categorical variables, while Wilcoxon Mann-Whitney *U* test was applied for continuous variables. Adjustment for multiple hypothesis testing was performed using the Benjamini-Hochberg procedure.[32] Complete remission (CR) was defined as hematologic recovery with at least 1000 neutrophils per μL and at least 100 000 platelets per μL, and < 5% BM blasts in at least one measurement.[33] Relapse-free survival (RFS) was defined as time from the date of CR until relapse, or death. Overall survival (OS) was defined as time from study entry until death from any cause. Patients alive without an event were censored at the time of their last follow-up. The prognostic impact of AML+13 was evaluated according to the Kaplan-Meier method and the log-rank test. To adjust for other potential prognostic variables, we derived multivariate Cox models for RFS and OS. The following variables were included in the models, based on their role as potential confounders and availability of data: age (as a continuous parameter), sex, BM blasts at initial diagnosis and on day 16, Eastern Cooperative Oncology Group (ECOG) performance status, white blood cell (WBC) count, platelet count, hemoglobin, serum lactate dehydrogenase (LDH) level, de novo vs secondary AML, and presence of AML+13. No variable selection technique was applied, and all variables were retained in the final models. $P \leq .05$ was considered significant.

Results

Isolated trisomy 13 is associated with poor prognosis

We evaluated the cytogenetic reports of 6836 AML patients with available follow up data treated within the multicenter AMLCG-1999 and SAL trials for aneuploidy of chromosome 13. A total of 264 patients (3.9%) lacked sufficient cytogenetic data. Additional copies of chromosome 13 were reported in 99 of 6572 patients (incidence, 1.5%). Our analyses focused on patients with isolated trisomy (n = 33) or tetrasomy 13 (n = 1) (incidence, 0.5%). Patients with additional

Table 2. Multivariate analysis

	RFS*		OS†	
Variable‡	**HR (95% CI)**	***P***	**HR (95% CI)**	***P***
Age *(10 y increase)*	1.33 (1.21-1.46)	**<.001**	1.38 (1.27-1.5)	**<.001**
BM blasts on day 16 (*10% increase*)	1.04 (0.97-1.09)	.08	1.02 (1.02-1.09)	**.002**
WBC (*10 G/l increase*)	1.02 (0.99-1.05)	.15	1.02 (1-1.05)	**.04**
de novo vs secondary AML	1.02 (0.75-1.4)	.89	1.26 (1-1.59)	**.05**
AML+13	1.47 (0.82-2.62)	.2	1.65 (1.03-2.63)	**.04**

Significant P values are indicated in bold.

*n = 378, number of events = 275 (114 patients excluded due to missing covariables).

†n = 549, number of events = 410 (335 patients excluded due to missing covariables).

‡Only variables with $P \leq .05$ in either model are shown. The following variables were included in both models: sex, age (continuous variable), BM blasts at initial diagnosis and day 16, ECOG performance status, WBC count, platelet count, hemoglobin, serum LDH level, de novo vs secondary AML, and AML+13 status.

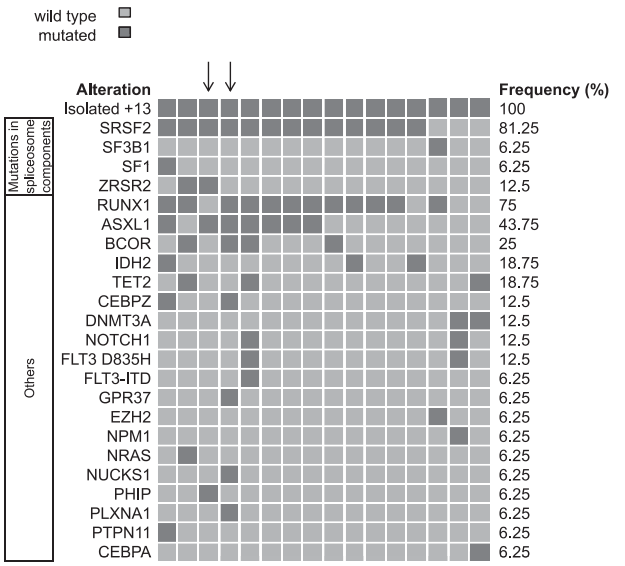


Figure 2. Frequency distribution of recurrently mutated genes in AML+13. Distribution of mutated genes in 16 patients with AML+13. Patients show a high frequency of mutations in spliceosome components and in *RUNX1*, *ASXL1*, and *BCOR*. Arrows highlight the 2 patients who were exome-sequenced.

numerical alterations of the sex chromosomes (n = 2) were included. These 34 patients (AML+13) were categorized into the Intermediate-II genetic category according to the ELN recommendations.[5] The remaining 65 patients had heterogeneous additional cytogenetic aberrations (aAML+13), frequently in the context of a complex karyotype, and were mostly classified as "adverse" according to ELN criteria (Favorable, n = 1; Intermediate-II, n = 20; Adverse, n = 44). AML+13 patients (n = 34 [AMLCG, n = 23; SAL, n = 11]) were compared with 850 ELN Intermediate-II genetic group patients without +13 enrolled in the same clinical trials. Detailed patient characteristics are given in Table 1 (and separated for the AMLCG and SAL subgroups in supplemental Table 4A-B). The study design is summarized in supplemental Figure 1A-B. In the combined data set, AML+13 patients were significantly older (P = .004) and had higher initial BM blast counts (P = .02), but significantly lower LDH levels (P = .009) than other patients in the ELN Intermediate-II genetic group. AML+13 and aAML+13 patients had similar baseline characteristics, except for significantly lower LDH levels and a higher CR rate in AML+13 and lower platelet counts than aAML+13 (supplemental Table 4C).

Twenty-one AML+13 patients (62%, 95% confidence interval [CI]: 44% to 77%) reached CR, compared with 471 (55%, 95% CI: 52% to 59%) of ELN Intermediate-II patients without +13 (P = .49). However, 18 of these 21 patients (86%, 95% CI: 63% to 96%) relapsed.

In the AMLCG trial, AML+13 was associated with inferior RFS and OS (median RFS = 8.7 vs 14.1 months, P = .02; median OS = 7 vs 13.9 months, P = .01; Figure 1A-B), whereas in the SAL cohort, the differences between AML+13 and other ELN Intermediate-II patients did not reach significance (RFS, P = .12; OS, P = .29; supplemental Figure 2A), possibly due to the small number of AML+13 cases (n = 11). RFS and OS in the combined SAL and AMLCG cohort were inferior for the AML+13 group compared with other ELN Intermediate-II patients (median RFS = 7.8 vs 14.1 months, P = .006; median OS = 9.3 vs 14.8 months, P = .004; Figure 1C-D).

In a multivariate analysis in the combined AMLCG and SAL cohorts that adjusted for other known prognostic markers, AML+13 remained a significant variable within the ELN Intermediate-II genetic group for OS, but not for RFS (Table 2).

There was no significant difference in RFS (P = .74) or OS (P = .82) between the AML+13 and aAML+13 subgroups, despite the high frequency of adverse cytogenetic alterations in the aAML+13 group (supplemental Figure 2B). We also compared the AMLCG AML+13 group (n = 23) to 463 patients treated on the AMLCG-1999 trial who had adverse cytogenetics. Baseline characteristics for these cohorts are shown in supplemental Table 4D. There was no significant difference regarding RFS (P = .78) or OS (P = .98) between both groups (supplemental Figure 2C).

High frequency of mutations affecting *SRSF2*, *RUNX1*, *ASXL1*, and *BCOR* in AML+13

To systematically identify somatic mutations associated with AML+13, we performed exome sequencing of paired diagnostic and remission samples from 2 patients with AML+13 (patients no. 8 and 11). We identified nonsynonymous leukemia-specific mutations affecting 36 genes, including *RUNX1*, *ASXL1*, *BCOR*, *ZRSR2*, *NUP188*, and *CEBPZ*. No recurring mutations were observed between the 2 patients. Nonsynonymous mutations in protein-coding transcripts are summarized in supplemental Table 5.

Targeted amplicon sequencing was performed on 16 AML+13 patient samples. Consistent with previous reports,[6,7] we found a high frequency of *RUNX1* mutations (n = 12, 75%). In addition, we detected mutations in spliceosome components in 14 AML+13 patients (88%), including *SRSF2* codon 95 mutations in 13 patients (81%) and an *SF3B1* mutation in 1 patient. The association of spliceosome component mutations (*SRSF2*, *SF3B1*, *SF1*, and *ZRSR2*) with *RUNX1* mutations was significant (P = .05). Additional recurring mutations affected *ASXL1* (n = 7, 44%) and *BCOR* (n = 4, 25%), and occurred with *RUNX1* and *SRSF2* mutations but these associations did not reach statistical significance (*ASXL1-SRSF2*, P = .21; *ASXL1-RUNX1*, P = .34; *BCOR-SRSF2*, P = .53; and *BCOR-RUNX1*, P = .53). The 2 patients without mutations in the splicing machinery had *DNMT3A* mutations, which were also mutually exclusive with mutations in *RUNX1* or *ASXL1*. Two patients carried mutations in *CEBPZ*, thus establishing *CEBPZ* as a novel recurrently mutated gene in AML. Details of all detected nonsynonymous variants are shown in Figure 2 and supplemental Table 6.

The mutations in *SRSF2* and *CEBPZ* were confirmed by Sanger sequencing (results summarized in supplemental Table 6). The correlation of the results from Sanger sequencing and targeted high throughput sequencing was 100% (for details, see supplemental Figure 3). In one of the patients with a *CEBPZ* mutation and an available remission sample, we could confirm the somatic nature of the mutation (supplemental Figure 3).

Both patients characterized by exome sequencing carried *SRSF2* mutations at codon 95, as identified by amplicon sequencing. However, these mutations were not detected by exome sequencing due to low coverage of this region in both samples. These results show that our targeted sequencing approach detects mutations in AML candidate genes with high sensitivity and specificity, including mutations in regions not covered by exome sequencing.

To further explore the association between *RUNX1* and *SRSF2* mutations, we analyzed the *SRSF2* gene in a cohort of 14 patients with a known *RUNX1* mutation and normal karyotype AML (CN-AML).[34] We found mutations in *SRSF2* in 3 of the 14 patients (21%).

Distinct gene expression pattern of AML+13

We identified 678 probe sets as significantly ($P \leq .05$ after adjustment for multiple testing) deregulated (upregulated, 492; downregulated, 186) in AML+13 patients (n = 9), when compared

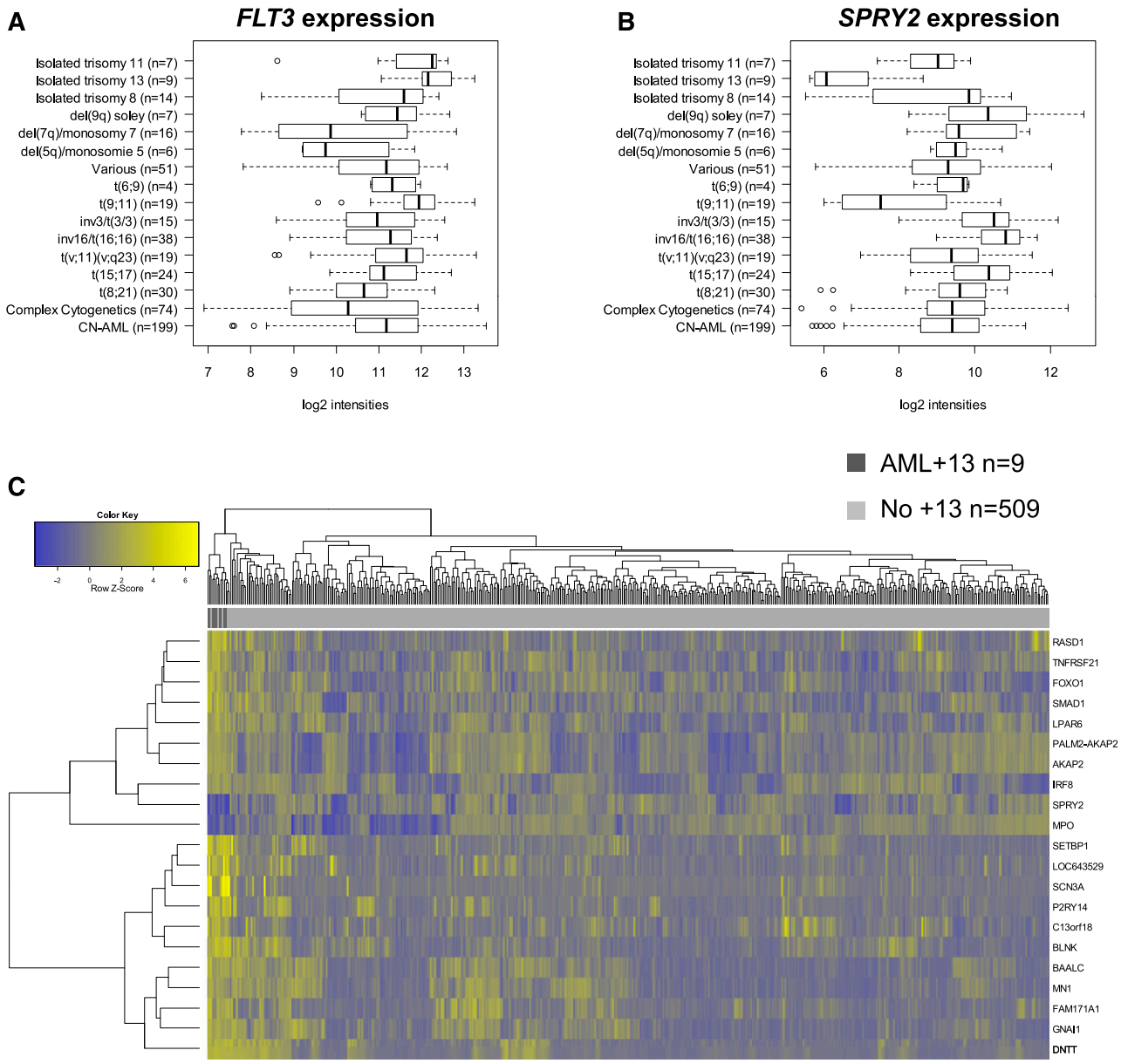
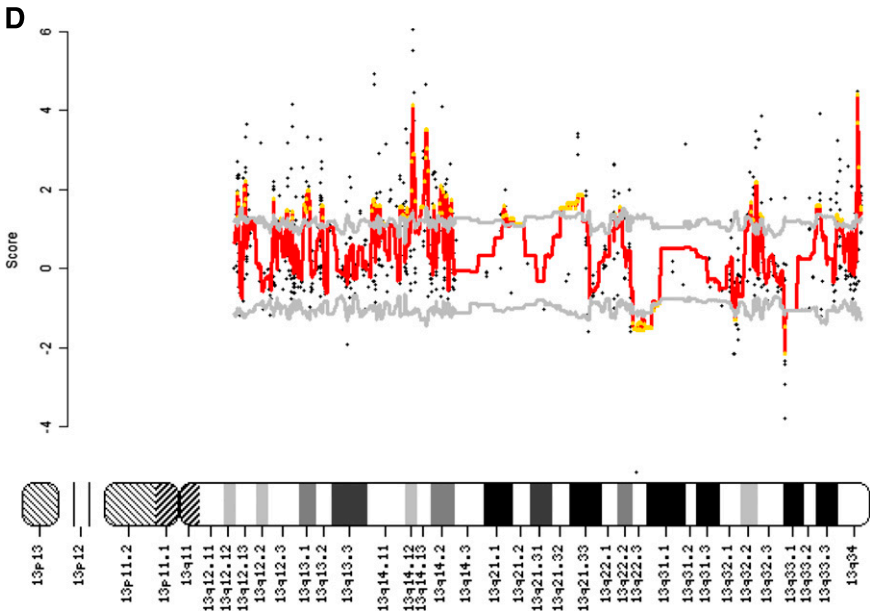


Figure 3. Gene expression profile of AML+13. (A-B) *FLT3* and *SPRY2* expression in AML subgroups. Boxplot showing *FLT3* (A) and *SPRY2* (B) expression levels in various cytogenetic AML subgroups. The boxes indicate the upper and lower quartiles. The band within the boxes represents the median. Outliers are plotted as individual points. *FLT3* expression is significantly higher in AML+13 compared with all other samples ($P = .04$). However, in several individual samples of various cytogenetic subgroups, *FLT3* was expressed at higher levels compared with AML+13. *SPRY2* expression is significantly lower in AML+13 ($P < .001$). (C) Clustering of AML+13 using 21 probe sets. Heatmap visualizing hierarchical clustering of AML+13 samples according to the 21 most differentially expressed probe sets (log-fold change ≥ 2 or ≤ -2 and adjusted *P*-value $< .001$) compared with AML with various other cytogenetic aberrations except for +13. All AML+13 samples cluster closely together, indicating a highly homogenous expression profile of this subgroup. (D) Regional gene expression on chromosome 13 in AML+13. Expression levels of probe sets located on chromosome 13 displayed by MACAT analysis in AML+13 patients (n = 9) compared with AML with various other cytogenetic abnormalities (except +13, n = 519). Scores for probe sets are shown as black dots. The sliding average of the 0.025 and 0.975 quantiles of the permuted scores are visualized as gray lines. The sliding average permuted scores (red line), and highlighted regions (yellow-dotted), where the score exceeds the quantiles, are plotted along chromosome 13. Despite the majority of probe sets showing elevated expression levels as expected, some regions were characterized by significantly lower expression levels.

to AML patients with various other cytogenetic abnormalities (n = 509). Detailed patient characteristics are given in supplemental Table 7. Only 59 (8.7%) of these probe sets were localized on chromosome 13, but of those, 55 were upregulated and only 4 were downregulated. Upregulated probe sets on chromosome 13 included *FOXO1*, *FLT3*, (Figure 3A) and *RB1*. The strongest downregulated probe set on chromosome 13 belonged to the tumor suppressor gene *SPRY2* (Figure 3B), which is a negative regulator of receptor tyrosine kinases. As described before, *FLT3* is significantly upregulated in AML+13, compared with all other AML samples in our gene expression data set ($P = .04$). However, as shown in Figure 3A, *FLT3* expression in AML shows a complex pattern with a wide range of expression levels, and AML+13 is not the only entity associated with high *FLT3* levels.

A total of 21 probe sets showed highly significant deregulation (log-fold change ≥ 2 or ≤ -2 and adjusted *P*-value $< .001$) and were therefore used for clustering (supplemental Table 8). The result of the clustering is shown in Figure 3C. Consistent with the results from our

Figure 3. Continued



genetic analysis, AML+13 shows a homogenous gene expression profile that is distinct from other AML subsets.

Surprisingly, some genes located on chromosome 13 showed significantly lower expression in AML+13 compared with patients with two copies of chromosome 13. The differential regional gene expression of AML+13 patient samples across chromosome 13 is visualized in Figure 3D (for details, see supplemental Table 9A-B). Despite the additional copy of chromosome 13, we identified several regions on chromosome 13 with significantly reduced gene expression levels compared with patients with two copies of chromosome 13.

By using GSEA, we see a potential deregulation of gene sets associated with cytoplasmatic and nuclear transport and the regulation of transcription. Details are given in supplemental Table 10. We could also observe that the expression levels of the transcription factor *FOXO1* correlated with higher expression levels of a predefined gene set consisting of target genes of this transcription factor (nominal *P*-value: .02; false discovery rate: .23). In summary, our gene expression studies reveal a complex picture of deregulated genes in AML+13 patients with a potential role in leukemogenesis. Some of these genes, such as *SPRY2* (Figure 3B) are downregulated despite their location on chromosome 13.

Finally, we compared the results of our gene expression analysis with data derived from the comparison of *RUNX1*-mutated and wild type AML with CN-AML.[34] This 85 gene *RUNX1* signature showed an overlap of 28 genes (33%) with differentially expressed genes in AML+13 (supplemental Table 11).

Discussion

Our study is the first to show that AML+13 patients have a significantly inferior RFS and OS compared with patients with other intermediate-risk cytogenetic abnormalities in a homogeneously treated cohort. Based on these findings, AML+13 should be considered as a subgroup associated with an extremely poor outcome. Furthermore, we provide evidence that AML+13 leukemia is genetically homogenous, not only on the cytogenetic but also on the molecular level. AML+13 is not only associated with a high frequency of *RUNX1* mutations, but also with mutations in *SRSF2*, *ASXL1*, and *BCOR*. To our knowledge, the incidence of mutations in *SRSF2* in AML+13 is the highest of any AML or myelodysplastic syndrome (MDS) subgroup reported so far.[35,36] An association between *SRSF2* and *RUNX1* mutations was already reported in patients with MDS.[35] We provide first evidence that an association between these mutations could also be observed in AML with *RUNX1* mutations. However, larger studies are necessary to verify this observation.

It is intriguing to speculate about functional interactions between mutations in these two genes and trisomy 13. It remains unclear whether mutations targeting *SRSF2* and *RUNX1*, and trisomy 13, affect a common pathway or different but complementary pathways on the way to leukemia. Although one of these lesions likely represents a near compulsory additional hit required by the initial event, the order of these events remains elusive. In light of the high prevalence of acquired *GATA1* mutations in AML of Down syndrome patients,[10] it is very likely that the chromosomal aneuploidy is the first event and determines the subsequent acquisition of mutations in precisely defined genes.

There is some, but limited overlap of recurrently mutated genes in AML and MDS. However, the high incidence of spliceosome gene mutations in both MDS and AML+13 is striking. A case report of 2 AML+13 patients who achieved sustained complete morphologic and cytogenetic remission while treated with high-dose, single-agent lenalidomide suggests a potential role of spliceosome gene mutations in the response to lenalidomide, which is also used in MDS therapy.[8] Otrock et al recently reported an association of lenalidome response with distinct mutation patterns.[37]

Of note, only one *SRSF2* mutation was found in 200 AML patients studied by whole exome or whole genome sequencing.[38] This *SRSF2*-mutated patient also had a *RUNX1* mutation. The study included a total of 19 *RUNX1*-mutated patients.[38] As is obvious from our study, it is likely that some *SRSF2* mutations in this study might have gone undetected, since exome sequencing may miss these mutations due to inefficient target enrichment.

It was proposed that overexpression of *FLT3,* which localizes to chromosome 13, could play a crucial role in AML+13.[6,7] Our

study confirms an elevated expression level of *FLT3* in the AML+13 subgroup. However, the levels are similar to other cytogenetic AML subgroups without additional chromosome 13, showing that high *FLT3* expression levels are not a defining feature of AML+13. Nevertheless, these findings do not rule out that high *FLT3* expression levels are an important leukemic driver in AML+13. High *FLT3* expression levels might be achieved by other mechanisms than an additional copy of chromosome 13 in other leukemias. Our gene expression analysis suggests several possible alternative or additional consequences of trisomy 13. *FOXO1* is overexpressed in AML+13, and GSEA revealed upregulated sets of *FOXO1* target genes. Recurrent mutations in *FOXO1* associated with poor survival were recently discovered in diffuse large B-cell lymphoma.[39] Furthermore, activation of *FOXO1* was observed in ~40% of AML patients.[40] Inhibition of *FOXO1* leads to reduced leukemic cell growth.[40] The tumor suppressor gene *SPRY2*, a negative regulator of receptor tyrosine kinases, had strikingly low expression levels even though it is located on chromosome 13 (Figure 3B). Downregulation of *SPRY2* was previously reported in a variety of solid tumors.[41-44] It is challenging to explain the underlying mechanism for this apparently contradictory result (ie, the downregulation despite an additional gene copy). Potential mechanisms for low *SPRY2* expression include epigenetic inactivation, submicroscopic deletions of *SPRY2*, or mutations in upstream regulators of *SPRY2*. These results again demonstrate the complexity of gene regulation and indicate that the concept of gene dosage is inadequate to explain all effects of an additional chromosome 13. Our gene expression data show a distinct gene expression profile of AML+13 partially overlapping with *RUNX1*- mutated CN-AML.

The striking association of mutations affecting only a few distinct genes in AML+13 suggests a strong synergism of these lesions during leukemogenesis. The fact that mutations in *RUNX1*, *ASXL1*, and upregulation of *FLT3* were previously reported as markers of poor prognosis in AML clearly suggests that the combination of these lesions is responsible for the extremely poor outcome of AML+13.

In summary, we discovered the highest incidence of *SRSF2* mutations in a specific AML subgroup reported so far. This rare, but genetically extremely homogenous group of AML+13 leukemia is characterized by concurrent mutations of *SRSF2* and *RUNX1*, as well as a specific gene expression profile. Consistent with other studies, our findings suggest a connection between mutations of *RUNX1* and *SRSF2* in myeloid leukemogenesis. AML+13 is associated with inferior survival despite intensive treatment. Therefore, new treatment strategies are highly warranted.

The discovery of rare, genetically homogenous AML subgroups indicates that the genetic complexity of AML is extremely high but mutations do not occur randomly. Despite the increasing number of comprehensively characterized AML cases, the understanding of oncogenic collaboration poses a challenge ahead.

Acknowledgments

The authors thank all participants and recruiting centers of the AMLCG and SAL trials.

This work was supported by a grant from the German Cancer Aid (109031) to P.A.G. and S.K.B., and start-up funding from the Ludwig-Maximilians-Universität to T.H. and K.H.M. (FöFoLe 798/774 and 783). S.K.B., K.S., and P.A.G. acknowledge support from the German Research Council (DFG) (Collaborative Research Center 684 Molecular Mechanisms of Normal and Malignant Hematopoesis, projects A6, A12, and start-up funding 2011).

Authorship

Contribution: T.H., K.H.M., and. P.A.G. conceived and designed the experiments; T.H., K.H.M., L.H., E.Z., B.K., and S.K. performed experiments; T.H., K.H.M., S.V., M.H., and V.J. analyzed data; S.V. and A.G. provided bioinformatics support; H.B. managed the Genome Analyzer IIx platform; B.K., A.D., E.Z., Z.P., P.M.K., S.S., S.K.B., and K.S. characterized patient samples; M.C.S., W.E.B., T.B., B.J.W., and W.H. coordinated the AMLCG clinical trial; P.A.G., U.M., K.S., and S.K.B. supervised the project; T.H., K.H.M., S.K.B., and P.A.G. wrote the manuscript; and C.R., F.S., M.B., and G.E. coordinated the SAL clinical trials, selected, contributed, and analyzed SAL data.

Conflict-of-interest disclosure: P.A.G. and S.K. received honoraria from Illumina. The remaining authors declare no competing financial interests.

Correspondence: Philipp A. Greif, University Hospital Grosshadern, Ludwig-Maximilians Universität, Marchioninistr. 15, 81377 Munich, Germany; e-mail: pgreif@med.uni-muenchen.de and p.greif@dkfz-heidelberg.de.

References

1. Mesa RA, Hanson CA, Ketterling RP, Schwager S, Knudson RA, Tefferi A. Trisomy 13: prevalence and clinicopathologic correlates of another potentially lenalidomide-sensitive cytogenetic abnormality. *Blood.* 2009;113(5):1200-1201.
2. Grimwade D, Hills RK, Moorman AV, et al; National Cancer Research Institute Adult Leukaemia Working Group. Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood.* 2010;116(3):354-365.
3. Baer MR, Bloomfield CD. Trisomy 13 in acute leukemia. *Leuk Lymphoma.* 1992;7(1-2):1-6.
4. Döhner H, Arthur DC, Ball ED, et al. Trisomy 13: a new recurring chromosome abnormality in acute leukemia. *Blood.* 1990;76(8):1614-1621.
5. Döhner H, Estey EH, Amadori S, et al; European LeukemiaNet. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood.* 2010;115(3):453-474.
6. Silva FP, Lind A, Brouwer-Mandema G, Valk PJ, Giphart-Gassler M. Trisomy 13 correlates with RUNX1 mutation and increased FLT3 expression in AML-M0 patients. *Haematologica.* 2007;92(8):1123-1126.
7. Dicker F, Haferlach C, Kern W, Haferlach T, Schnittger S. Trisomy 13 is strongly associated with AML1/RUNX1 mutations and increased FLT3 expression in acute myeloid leukemia. *Blood.* 2007;110(4):1308-1316.
8. Fehniger TA, Byrd JC, Marcucci G, et al. Single-agent lenalidomide induces complete remission of acute myeloid leukemia in patients with isolated trisomy 13. *Blood.* 2009;113(5):1002-1005.
9. Ganmore I, Smooha G, Izraeli S. Constitutional aneuploidy and cancer predisposition. *Hum Mol Genet.* 2009;18(R1):R84-R93.
10. Wechsler J, Greene M, McDevitt MA, et al. Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. *Nat Genet.* 2002;32(1):148-152.
11. Baty BJ, Blackburn BL, Carey JC. Natural history of trisomy 18 and trisomy 13: I. Growth, physical assessment, medical histories, survival, and recurrence risk. *Am J Med Genet.* 1994;49(2):175-188.
12. Satge D, Van Den Berghe H. Aspects of the neoplasms observed in patients with constitutional autosomal trisomy. *Cancer Genet Cytogenet.* 1996;87(1):63-70.

13. Schade H, Schoeller L, Schultze KW. [D-trisomy (Paetau syndrome) with congenital myeloid leukemia]. *Med Welt.* 1962;50:2690-2692.

14. Schaich M, Parmentier S, Kramer M, et al. High-dose cytarabine consolidation with or without additional amsacrine and mitoxantrone in acute myeloid leukemia: results of the prospective randomized AML2003 trial. *J Clin Oncol.* 2013; 31(17):2094-2102.

15. Schaich M, Röllig C, Soucek S, et al. Cytarabine dose of 36 g/m compared with 12 g/m within first consolidation in acute myeloid leukemia: results of patients enrolled onto the prospective randomized AML96 study. *J Clin Oncol.* 2011;29(19): 2696-2702.

16. Rollig C, Kramer M, Hanel M, et al. Induction treatment in elderly patients with acute myeloid leukemia (AML): randomized comparison of intermediate-dose cytarabine plus mitoxantrone (IMA) versus standard-dose cytarabine plus daunorubicin (DA) in 492 AML patients >60 years - Results from the SAL 60plus trial [abstract]. *Blood (ASH Annual Meeting Abstracts).* 2010;116(21). Abstract 334.

17. Büchner T, Berdel WE, Schoch C, et al. Double induction containing either two courses or one course of high-dose cytarabine plus mitoxantrone and postremission therapy by either autologous stem-cell transplantation or by prolonged maintenance for acute myeloid leukemia. *J Clin Oncol.* 2006;24(16):2480-2489.

18. Opatz S, Polzer H, Herold T, et al. Exome sequencing identifies recurring FLT3 N676K mutations in core-binding factor leukemia. *Blood.* 2013;122(10):1761-1769.

19. Greif PA, Dufour A, Konstandin NP, et al. GATA2 zinc finger 1 mutations associated with biallelic CEBPA mutations define a unique genetic entity of acute myeloid leukemia. *Blood.* 2012;120(2): 395-403.

20. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-1760.

21. Li H, Handsaker B, Wysoker A, et al; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-2079.

22. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9): 1297-1303.

23. Koboldt DC, Chen K, Wylie T, et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics.* 2009;25(17):2283-2285.

24. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics.* 2009;25(21):2865-2871.

25. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 2012;22(3):568-576.

26. Li Z, Herold T, He C, et al. Identification of a 24-gene prognostic signature that improves the European LeukemiaNet risk classification of acute myeloid leukemia: an international collaborative study. *J Clin Oncol.* 2013;31(9):1172-1181.

27. Herold T, Jurinovic V, Metzeler KH, et al. An eight-gene expression signature for the prediction of survival and time to treatment in chronic lymphocytic leukemia. *Leukemia.* 2011;25(10): 1639-1645.

28. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* 2005;102(43):15545-15550.

29. Herold T, Jurinovic V, Mulaw M, et al. Expression analysis of genes located in the minimally deleted regions of 13q14 and 11q22-23 in chronic lymphocytic leukemia-unexpected expression pattern of the RHO GTPase activator ARHGAP20. *Genes Chromosomes Cancer.* 2011;50(7):546-558.

30. Toedling J, Schmeier S, Heinig M, Georgi B, Roepcke S. MACAT—microarray chromosome analysis tool. *Bioinformatics.* 2005;21(9): 2112-2113.

31. R Development Core Team (2008). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, www.R-project.org.

32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc, B.* 1995;57(1): 289-300.

33. Cheson BD, Bennett JM, Kopecky KJ, et al; International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia [published correction appears in J Clin Oncol. 2004;22(3):576]. *J Clin Oncol.* 2003;21(24):4642-4649.

34. Greif PA, Konstandin NP, Metzeler KH, et al. RUNX1 mutations in cytogenetically normal acute myeloid leukemia are associated with a poor prognosis and up-regulation of lymphoid genes. *Haematologica.* 2012;97(12):1909-1915.

35. Thol F, Kade S, Schlarmann C, et al. Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. *Blood.* 2012; 119(15):3578-3584.

36. Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature.* 2011;478(7367):64-69.

37. Otrock ZK, Przychodzen BP, Husseinzadeh HD, et al. Molecular predictors of response to lenalidomide in myeloid malignancies. *Blood.* 2013;122(21):Abstract 2807.

38. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med.* 2013;368(22):2059-2074.

39. Trinh DL, Scott DW, Morin RD, et al. Analysis of FOXO1 mutations in diffuse large B-cell lymphoma. *Blood.* 2013;121(18):3666-3674.

40. Sykes SM, Lane SW, Bullinger L, et al. AKT/FOXO signaling enforces reversible differentiation blockade in myeloid leukemias [published correction appears in *Cell.* 2011;147(1):247]. *Cell.* 2011;146(5):697-708.

41. Kwak HJ, Kim YJ, Chun KR, et al. Downregulation of Spry2 by miR-21 triggers malignancy in human gliomas. *Oncogene.* 2011;30(21):2433-2442.

42. Fritzsche S, Kenzelmann M, Hoffmann MJ, et al. Concomitant down-regulation of SPRY1 and SPRY2 in prostate carcinoma. *Endocr Relat Cancer.* 2006;13(3):839-849.

43. Fong CW, Chua MS, McKie AB, et al. Sprouty 2, an inhibitor of mitogen-activated protein kinase signaling, is down-regulated in hepatocellular carcinoma. *Cancer Res.* 2006;66(4):2048-2058.

44. Lo TL, Yusoff P, Fong CW, et al. The ras/mitogen-activated protein kinase pathway inhibitor and likely tumor suppressor proteins, sprouty 1 and sprouty 2 are deregulated in breast cancer. *Cancer Res.* 2004;64(17):6127-6136.

blood®

2014 124: 1304-1311
doi:10.1182/blood-2013-12-540716 originally published online June 12, 2014

Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis

Tobias Herold, Klaus H. Metzeler, Sebastian Vosberg, Luise Hartmann, Christoph Röllig, Friedrich Stölzel, Stephanie Schneider, Max Hubmann, Evelyn Zellmeier, Bianka Ksienzyk, Vindi Jurinovic, Zlatana Pasalic, Purvi M. Kakadia, Annika Dufour, Alexander Graf, Stefan Krebs, Helmut Blum, Maria Cristina Sauerland, Thomas Büchner, Wolfgang E. Berdel, Bernhard J. Woermann, Martin Bornhäuser, Gerhard Ehninger, Ulrich Mansmann, Wolfgang Hiddemann, Stefan K. Bohlander, Karsten Spiekermann and Philipp A. Greif

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.

Supplemental Information

Isolated trisomy 13 defines a genetically homogenous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis.

Tobias Herold,[1,2,3,4] Klaus H. Metzeler,[1,2,3,4] Sebastian Vosberg,[1,2,3,4] Luise Hartmann,[1,2,3,4] Christoph Röllig,[3,4,5] Friedrich Stölzel,[5] Stephanie Schneider,[1] Max Hubmann,[1,2] Evelyn Zellmeier,[1] Bianka Ksienzyk,[1] Vindi Jurinovic,[6] Zlatana Pasalic,[1] Purvi M Kakadia,[7] Annika Dufour,[1] Alexander Graf,[8] Stefan Krebs,[8] Helmut Blum,[8] Maria Cristina Sauerland,[9] Thomas Büchner,[10] Wolfgang E. Berdel,[10] Bernhard J. Woermann,[11] Martin Bornhäuser,[3,4,5] Gerhard Ehninger,[3,4,5] Ulrich Mansmann, [3,4,6] Wolfgang Hiddemann,[1,2,3,4] Stefan K. Bohlander,[12] Karsten Spiekermann[1,2,3,4] and Philipp A. Greif[1,2,3,4]

[1]Department of Internal Medicine 3, University Hospital Grosshadern, Ludwig-Maximilians-Universität (LMU), München, Germany
[2]Clinical Cooperative Group Leukemia, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany
[3]German Cancer Consortium (DKTK), Heidelberg, Germany
[4]German Cancer Research Center (DKFZ), Heidelberg, Germany
[5]Medizinische Klinik und Poliklinik I, Universitätsklinikum Dresden, Dresden, Germany
[6]Institute for Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität (LMU), München, Germany
[7]Center for Human Genetics, Philipps University, Marburg, Germany
[8]Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, Ludwig-Maximilians-Universität (LMU), München, Germany
[9]Institute of Biostatistics and Clinical Research, and [10]Department of Medicine A-Hematology, Oncology and Pneumology, University of Münster, Münster, Germany
[11]German Society of Hematology and Oncology, Berlin, Germany
[12]Department of Molecular Medicine and Pathology, The University of Auckland, Auckland, New Zealand

Corresponding Author:
Philipp A. Greif, MD
Marchioninistr. 25
81377 München
Phone: +49 89 3187 1358
FAX: +49 89 3187 1358
Email: pgreif@med.uni-muenchen.de

Presented in abstract form at the 55th annual meeting of the American Society of Hematology, December 7-10, 2013 in New Orleans, LA (Publication Number: 608)

Running title: *SRSF2 mutations in AML +13*

Table S1: Quality metrics summary of exome sequencing data

	Patient 8 tumor	**Patient 8 control**	**Patient 11 tumor**	**Patient 11 control**
# total sequence reads	41,169,242	36,455,938	43,076,316	41,118,008
# total bases sequenced	3,595,441,280	3,170,797,760	3,766,165,920	3,579,766,560
% mapped reads	99.55	99.4721	99.4451	99.4387
% sequenced bases mapped	89.44	89.83	89.2	89.68
% target region sequenced*	97.65	97.52	97.67	97.71
% target region sequenced, minimum 10x coverage*	78.49	74.33	78.25	77.34
Mean coverage	33.91	28.83	33.83	31.62

*****NCBI human gemome (hg)19 protein coding region (34 Mb)

Table S2: *Primer Sequences*

Primer Name	**Sequence**
CEBPZ_1 forward	5'CAGCCTCAGGATGTTGTATCTAAG
CEBPZ_1 reverse	5'GCTTTTGTGGCAATTCTGTTC
CEBPZ_2 forward	5'AGCCCTTACCGTGGCTC
CEBPZ_2 reverse	5'GGGCACTGCTTGTGCTG
ASXL1 forward	5'AGTCCCTAGGTCAGATCACCC
ASXL1 reverse	5'CAACGGGGAGTTGGGAG
SRSF2_TO_fw	5'CAAGGTGGACAACCTGACCT
SRSF2_TO_rev	5'AGACGCCATTTCCCCAGT

Table S3: *PCR conditions*

Step	**Duration**	**Temperature**
Initial denaturation	3 min	94°C
Denaturation	0,5 min	94°C
Primer annealing	0,5 min	56°C
Extension	1 min	72°C
Final extension	10 min	72°C

Number of cycles: 35

Table S4 A: Patient characteristics AMLCG cohort

Variable	***AML+13****	**Control Group***	**P-value**
No. of patients	23	364	
Median age, years (range)	62 (45-80)	61 (18-82)	0.16
Male sex, no. (%)	16 (70)	200 (55)	0.2
White-cell count, G/l, median (range)	10.6 (0.7-318.1)	12 (0.6-341)	0.7
Hemoglobin, g/dl, median (range)	8.8 (4.6-12.8)	9.1 (3.8-16.9)	0.42
Platelet count, G/l ,median (range)	80 (1-283)	53.5 (1-1760)	0.19
LDH (U/l), median(range)	269 (155-869)	413 (115-11140)	**0.009**
Bone marrow blasts, %, median (range)	82 (11-100)	80 (11-100)	0.13
Bone marrow blasts at day 16, %, median (range)	5 (0-85)	5 (0-100)	0.88
Performance Status (ECOG) ≥ 2 (%)	5 (22)	114 (35)	0.26
de novo AML (%)	18 (78)	272 (75)	0.81
Allogeneic transplantation, no. (%)	5 (22)	87 (24)	1
Complete remission, no. (%)	13 (57)	198 (54)	1
Relapse, no. (%)	12 (92)	155 (78)	0.31
Deceased, no. (%)	22 (96)	286 (79)	0.06

*All patients were enrolled in the AMLCG-99 trial and received intensive induction treatment. All patients are classified as ELN Intermediate-II; AML+13: patients with isolated tri- or tetrasomy 13, additional aberrations of the sex chromosomes are allowed.

Table S4 B: Patient characteristics SAL cohort

Variable	AML+13*	Control Group*	P-value
No. of patients	11	486	
Median age, years (range)	66 (43-76)	56 (17-84)	**0.03**
Male sex, no. (%)	8 (72.7)	265 (55)	0.36
White-cell count G/l, median (range)	7 (1-237)	11 (0.4-365)	0.5
Hemoglobin, g/dl, median (range)	8.9 (5.3-12.1)	9.3 (2.9-17.2)	0.51
Platelet count, G/l, median (range)	74 (11-399)	56 (1-1043)	0.92
LDH, U/l, median (range)	371 (184-1011)	416 (122-5565)	0.4
Bone marrow blasts, %, median (range)	73 (28-92)	62 (11-99)	0.27
Bone marrow blasts at day 15, %, median (range)	5 (1-80)	10 (0-95)	0.78
Performance status (ECOG) >= 2, no., (%)	3 (37.5)	121 (29)	0.7
de novo AML, no. (%)	8 (80.0)	374 (77)	0.85
Allogeneic transplantation, no. (%)	1 (9.1)	93 (19)	0.7
Complete remission, no. (%)	8 (72.7)	273 (56)	0.36
Relapse, no. (%)	6 (75.0)	172 (63)	0.17
Deceased, no. (%)	9 (81.8)	358 (74)	0.74

*All patients were enrolled in SAL trials and received intensive induction treatment. All patients are classified as ELN Intermediate-II; AML+13: patients with isolated tri- or tetrasomy 13, additional aberrations of the sex chromosomes are allowed.

Table S4 C: Patient characteristics AML+13 versus aAML+13

Variable	***AML+13****	**aAML+13***	**P-value**
No. of patients	34	65	
Median age, years (range)	64 (43-80)	64 (32-86)	0.93
Male sex, no. (%)	24 (70)	41 (63)	0.51
White-cell count, G/l, median (range)	10 (1-318)	12 (0.1-269)	0.68
Hemoglobin, g/dl, median (range)	8.9 (4.6-12.8)	9.3 (4.4-14.8)	0.2
Platelet count, G/l ,median (range)	77 (1-399)	41 (1-592)	**0.03**
LDH (U/l), median(range)	269 (155-1011)	458 (104-7015)	**0.003**
Bone marrow blasts, %, median (range)	80 (11-100)	75 (12-100)	0.12
Bone marrow blasts at day 16, %, median (range)	5 (0-85)	5 (0-90)	0.91
Performance Status (ECOG) ≥ 2 (%)	8 (26)	19 (31)	0.64
de novo AML (%)	26 (76)	43 (66)	0.36
Allogeneic transplantation, no. (%)	6 (18)	13 (20)	1
Complete remission, no. (%)	21 (62)	25 (38)	**0.03**
Relapse, no. (%)	18 (86)	21 (84)	1
Deceased, no. (%)	31 (91)	56 (86)	0.54

*All patients were enrolled in the AMLCG-99 or SAL trials and received intensive induction treatment. AML+13: patients with isolated tri- or tetrasomy 13, additional aberrations of the sex chromosomes are allowed; aAML+13: patients with additional copies of chromosome 13 and further genetic aberrations not classified as AML+13.

Table S4 D: Patient characteristics of AML+13 versus and ELN Adverse

Variable	***AML+13****	**ELN Adverse Genetic Group***	**P-value**
No. of patients	23	463	
Median age, years (range)	62 (45-80)	62 (17-85)	0.44
Male sex, no. (%)	16 (70)	242 (52)	0.13
White-cell count, G/l, median (range)	10.6 (0.7-318.1)	4.5 (0.3-666)	0.2
Hemoglobin, g/dl, median (range)	8.8 (4.6-12.8)	8.8 (3.6-14.5)	0.94
Platelet count, G/l ,median (range)	80 (1-283)	52 (1-1110)	0.1
LDH (U/l), median(range)	269 (155-869)	342 (76-19624)	0.14
Bone marrow blasts, %, median (range)	82 (11-100)	60 (5-100)	**0.001**
Bone marrow blasts at day 16, %, median (range)	5 (0-85)	8 (0-100)	0.92
Performance Status (ECOG) ≥ 2 (%)	5 (22)	145 (34)	0.26
de novo AML (%)	18 (78)	113 (24)	**<0.001**
Allogeneic transplantation, no. (%)	5 (22)	27 (6)	**0.01**
Complete remission, no. (%)	13 (57)	146 (32)	**0.02**
Relapse, no. (%)	12 (92)	134 (92)	1
Deceased, no. (%)	22 (96)	393 (85)	0.23

*All patients were enrolled in the AMLCG-99 trial and received intensive induction treatment. AML+13: patients with isolated tri- or tetrasomy 13, additional aberrations of the sex chromosomes are allowed.

Table S5: Leukemia-specific variants identified by exome sequencing

Patient #	Gene	Transcript ID	Position	Reference reads	Variant reads	Variant allel frequency (%)	cDNA sequence change	protein equence change	HaloPlex	Sanger sequencing
8	*KIAA1549*	NM_020910	7:138597149	29	27	48	c.G2936A	p.R979K	N/A	N/A
8	*TG*	NM_003235	8:133899376	27	23	46	c.G1759C	p.V587L	N/A	N/A
8	*KDR*	NM_002253	4:55972889	23	22	49	c.A1501G	p.K501E	**not confirmed***	N/A
8	*TOMM70A*	NM_014820	3:100093990	9	12	57	c.G1099A	p.A367T	N/A	N/A
8	*PHIP*	NM_017934	6:79725406	11	12	52	c.A1330C	p.N444H	**confirmed**	N/A
8	*ZRSR2*	NM_005089	X:15822320	1	11	92	c.G399T	p.E133D	**confirmed**	N/A
8	*NUP188*	NM_015354	9:131765723	10	7	41	c.G4425A	p.R1475H	N/A	N/A
8	*FYB*	NM_001465	5:39153706	8	5	38	c.G1166A	p.S379N	N/A	N/A
8	*RTCA*	NM_001130841	1:100741174	10	3	23	c.C674G	p.A225G	N/A	N/A
8	*FOXF1*	NM_001451	16:86546689	9	3	25	c.T1138G	p.380stop_lost	N/A	N/A
8	*IHH*	NM_002181	2:219925092	9	3	25	c.T98G	p.V33G	N/A	N/A
8	*DDX17*	NM_006386	22:38895454	28	21	43	c.488_489insC	p.G163fs	N/A	N/A
11	*GPR37*	NM_005302	7:124386910	35	38	52	c.G1667A	p.C504Y	**confirmed**	N/A
11	*CEBPZ*	NM_005760	2:37455685	38	30	44	c.C651G	p.I217M	**confirmed**	N/A
11	*RUNX1*	NM_001754	21:36231791	3	21	84	c.A593T	p.D198V	**confirmed**	**confirmed**
11	*NUCKS1*	NM_022731	1:205687471	16	19	54	c.C669A	p.S223R	**confirmed**	N/A
11	*UBN1*	NM_016936	16:4925161	16	19	54	c.G2750A	p.S917N	N/A	N/A
11	*KIAA0319*	NM_014809	6:24563680	28	19	40	c.G2498A	p.R833Q	N/A	N/A
11	*USP21*	NM_012475	1:161131886	18	16	47	c.C659T	p.T220M	N/A	N/A
11	*PLXNA1*	NM_032242	3:126739168	15	12	44	c.C4019T	p.P1340L	**confirmed**	N/A
11	*DNAH5*	NM_001369	5:13845086	10	8	44	c.C5131T	p.R1711X	N/A	N/A
11	*FAM189B*	NM_006589	1:155217697	17	7	29	c.G1883A	p.R628H	N/A	N/A
11	*ASXL1*	NM_015338	20:31022592	15	7	32	c.C2077T	p.R693X	**confirmed**	N/A
11	*VIPR2*	NM_003382	7:158902599	10	7	41	c.G163A	p.V55I	N/A	N/A
11	*FMN2*	NM_020066	1:240370952	19	5	21	c.C2840T	p.P947L	N/A	N/A
11	*CHODL*	NM_024944	21:19628835	18	5	22	c.T89G	p.V30G	N/A	N/A
11	*SELENBP1*	NM_003944	1:151340744	6	4	40	c.A412C	p.T138P	N/A	N/A
11	*SELENBP1*	NM_003944	1:151340746	8	3	27	c.A410C	p.H137P	N/A	N/A
11	*ZMYM4*	NM_005095	1:35870649	9	3	25	c.T3554G	p.V1185G	N/A	N/A
11	*CACNB2*	NM_201596	10:18439813	10	3	23	c.C122T	p.S41L	N/A	N/A
11	*KCNK10*	NM_021161	14:88652357	11	3	21	c.T1139C	p.L380P	N/A	N/A
11	*TSHZ1*	NM_005786	18:72997846	8	3	27	c.A349C	p.T117P	N/A	N/A
11	*ZNF676*	NM_001001411	19:22363870	7	3	30	c.A349C	p.V117I	N/A	N/A
11	*NRXN2*	NM_015080	11:64457918	12	4	25	c.G808_809insC	p.A270fs	N/A	N/A
11	*ZNF676*	NM_001001411	19:22363810	10	3	23	c.707_708delT	p.R236fs	N/A	N/A
11	*ZNF676*	NM_001001411	19:22363814	11	3	21	c.704_705insG	p.N235fs	N/A	N/A
11	*BCOR*	NM_017745	X:39934065	10	3	23	c.533_534insG	p.P178fs	**confirmed**	N/A

*region was not covered by HaloPlex probe.

Table S6: Details of gene mutations found in 16 patients with AML+13 by targeted resequencing

SRSF2 (NM_003016)								
Patient	**Gene**	**Position**	**Reference reads**	**Variant reads**	**Variant allel frequency (%)**	**cDNA sequence change**	**Protein sequence change**	**Sanger sequencing**
1	SRSF2	17:74732959	165	108	40	c.C284G	p.P95R	**confirmed**
2	SRSF2	17:74732936	155	96	38	c.284_307del	p.95_103del	**confirmed**
4	SRSF2	17:74732959	215	115	35	c.C284A	p.P95H	**confirmed**
5	SRSF2	17:74732936	195	93	32	c.284_307del	p.95_103del	**confirmed**
6	SRSF2	17:74732959	271	159	37	c.283_284insGCC	p.P95delinsRP	**confirmed**
7	SRSF2	17:74732959	234	181	44	c.C284A	p.P95H	**confirmed**
8	SRSF2	17:74732959	152	138	48	c.C284A	p.P95H	**confirmed**
9	SRSF2	17:74732959	111	101	47	c.C284T	p.P95L	**confirmed**
10	SRSF2	17:74732936	146	62	30	c.284_307de	p.95_103del	**confirmed**
11	SRSF2	17:74732936	160	88	35	c.284_307del	p.95_103del	**confirmed**
14	SRSF2	17:74732959	271	107	28	c.C284T	p.P95L	**confirmed**
15	SRSF2	17:74732959	230	194	46	c.C284G	p.P95R	**confirmed**
16	SRSF2	17:74732936	118	72	38	c.284_307del	p.95_103del	**confirmed**
RUNX1 (NM_001754)								
Patient	**Gene**	**Position**	**Reference reads**	**Variant reads**	**Variant allel frequency (%)**	**cDNA sequence change**	**Protein sequence change**	**Sanger sequencing**
1	RUNX1	21:36206729	380	65	15	c.782_783insCC	p.P261fs	N/A
1	RUNX1	21:36252937	517	335	39	c.424_425insGGGCAGG	p.A142fs	N/A
2	RUNX1	21:36252940	388	245	39	c.C422T	p.S141L	**confirmed**
2	RUNX1	21:36164838	186	151	45	c.1036_1037insC	p.R346fs	N/A

3	RUNX1	21:36259204	522	317	38	c.A287G	p.N96S	N/A
4	RUNX1	21:36231783	125	439	78	c.C601T	p.R201X	N/A
5	RUNX1	21:36259156	606	346	36	c.T335C	p.L112P	N/A
5	RUNX1	21:36252940	498	429	46	c.C422A	p.S141X	N/A
6	RUNX1	21:36252920	605	265	30	c.441_442insGGCTGAGCTGAGAAATGCT	p.T148_A149delinsGX	**confirmed**
6	RUNX1	21:36252865	388	264	40	c.G497A	p.R166Q	N/A
7	RUNX1	21:36171728	2008	325	14	c.G837A	p.W279X	N/A
9	RUNX1	21:36252869	339	158	31	c.492_493insTAG	p.G165delinsX	N/A
9	RUNX1	21:36259181	274	221	45	c.309_310insCT	p.T104fs	N/A
10	RUNX1	21:36259173	54	705	93	c.G318C	p.W106C	**confirmed**
11	RUNX1	21:36231791	85	531	86	c.A593T	p.D198V	**confirmed**
15	RUNX1	21:36252961	103	422	80	c.C401A	p.A134D	N/A
16	RUNX1	21:36164745	473	393	45	c.1129_1130insT	p.Y377fs	N/A
CEBPZ (NM_005760)								
Patient	**Gene**	**Position**	**Reference reads**	**Variant reads**	**Variant allel frequency (%)**	**cDNA sequence change**	**Protein sequence change**	**Sanger sequencing**
10	CEBPZ	2:37455632	121	117	49	c.T704C	p.M235T	**confirmed**
11	CEBPZ	2:37455685	89	72	45	c.C651G	p.I217M	**confirmed**
BCOR (NM_017745)								
Patient	**Gene**	**Position**	**Reference reads**	**Variant reads**	**Variant allel frequency (%)**	**cDNA sequence change**	**Protein sequence change**	**Sanger sequencing**
5	BCOR	X:39933563	392	50	11	c.C1036G	p.P346A	N/A
5	BCOR	X:39914723	301	273	48	c.C4537T	p.R1513X	N/A
7	BCOR	X:39932330	237	32	12	c.2268_2269del	p.756_757del	N/A
9	BCOR	X:39911519	215	14	6	c.G5009A	p.S1670N	N/A

11	BCOR	X:39934065	132	62	32	c.533_534insC	p.P178fs	N/A
11	BCOR	X:39934069	136	65	32	c.G530C	p.S177T	N/A
ASXL1 (NM_015338)								
Patient	**Gene**	**Position**	**Reference reads**	**Variant reads**	**Variant allel frequency (%)**	**cDNA sequence change**	**Protein sequence change**	**Sanger sequencing**
1	ASXL1	20:31022928	394	315	44	c.2413delC	p.P805fs	N/A
4	ASXL1	20:31022441	389	113	23	c.1934dupG	p.G646fs	N/A
8	ASXL1	20:31022403	N/A	92	N/A	c.1900_1922del	p.E635fs	**confirmed**
9	ASXL1	20:31022441	332	113	25	c.1934dupG	p.G646fs	N/A
10	ASXL1	20:31022441	239	82	25	c.1934dupG	p.G646fs	N/A
11	ASXL1	20:31022592	299	218	42	c.C2077T	p.R693X	N/A
15	ASXL1	20:31022441	538	263	33	c.1926delA	p.G642fs	N/A
DNMT3A (NM_175629)								
Patient	**Gene**	**Position**	**Reference reads**	**Variant reads**	**Variant allel frequency (%)**	**cDNA sequence change**	**Protein sequence change**	**Sanger sequencing**
12	DNMT3A	2:25469616	477	206	30	c.1151_1152insT	p.F384fs	N/A
12	DNMT3A	2:25457185	476	397	45	c.T2702A	p.L901H	N/A
13	DNMT3A	2:25464508	250	272	52	c.T2005C	p.S669P	N/A
13	DNMT3A	2:25457158	470	532	53	c.C2729T	p.A910V	N/A

Table S7: Patient characteristics of the gene expression data set

Variable	***AML+13****	**All other (+13 excluded)**	**P-value**
No. of patients	9	509	
Median age, years (range)	64 (50-80)	57 (18-85)	0.06
Male sex, no. (%)	**8 (88.9)**	251 (49.3)	**0.04**
White-cell count, G/l, median (range)	10.6 (1.2-255)	19.5 (0.1-666)	0.69
Hemoglobin, g/dl, median (range)	8.7 (4.6-11.6)	9 (3.5-15.4)	0.85
Platelet count, G/l ,median (range)	84 (1-234)	49 (1-1760)	0.17
LDH (U/l), median(range)	**268 (166-459)**	465 (76-19624)	**0.004**
Bone marrow blasts, %, median (range)	**90 (80-100)**	80 (10-100)	**0.01**
Bone marrow blasts at day 16, %, median (range)	15 (0-95)	5 (0-100)	0.14
Performance Status (ECOG) ≥ 2 (%)	3 (33.3)	149 (31.2)	1
de novo AML (%)	9 (100)	399 (78.4)	0.22
Allogeneic transplantation, no.	0	37	1
Complete remission, no. (%)	5 (55.5)	282 (55.4)	1
Relapse, no. (%)	5 (100)	181 (64.2)	0.17
Deceased, no. (%)	9 (100)	354 (69.5)	0.06

All patients were enrolled in the AMLCG-99 trial and received intensive induction treatment.
* AML+13: patients with isolated tri- or tetrasomy 13, additional aberrations of the sex chromosomes are allowed; All other: patients with all kind of cytogenetic abnormalities expect of additional copies of chromosome 13.

Table S8: Top 21 differentially expressed genes in AML+13

Probe set	Gene	Description	Adjusted P-value	Log fold change	Chromosome
GC10P098054_at	DNTT	deoxynucleotidyltransferase, terminal	<0.001	5.12	10
GC10M097941_at	BLNK	B-cell linker	<0.001	4.02	10
GC07P079763_at	GNAI1	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	<0.001	3.16	7
GC03M150929_at	P2RY14	purinergic receptor P2Y, G-protein coupled, 14	<0.001	3.12	3
GC17M017397_at	RASD1	RAS, dexamethasone-induced 1	<0.001	2.88	17
GC13M048963_at	LPAR6	lysophosphatidic acid receptor 6	<0.001	2.72	**13**
GC08P104152_at	BAALC	brain and acute leukemia, cytoplasmic	<0.001	2.71	8
GC09P112542_at	PALM2-AKAP2	PALM2-AKAP2 readthrough	<0.001	2.54	9
GC06M047246_at	TNFRSF21	tumor necrosis factor receptor superfamily, member 21	<0.001	2.53	6
GC22M028144_at	MN1	meningioma (disrupted in balanced translocation) 1	0.001	2.52	22
GC18P042260_at	SETBP1	SET binding protein 1	<0.001	2.46	18
GC09P112810_at	AKAP2	A kinase (PRKA) anchor protein 2	<0.001	2.45	9
GC04P146402_at	SMAD1	SMAD family member 1	<0.001	2.35	4
GC13M041129_at	FOXO1	forkhead box O1	<0.001	2.32	**13**
GC10P091579_at	LOC643529	hCG2024094	<0.001	2.28	10
GC16P085932_at	IRF8	interferon regulatory factor 8	<0.001	2.25	16
GC13M046916_at	C13orf18	chromosome 13 open reading frame 18	<0.001	2.17	**13**
GC02M165908_at	SCN3A	sodium channel, voltage-gated, type III, alpha subunit	<0.001	2.17	2
GC10M015294_at	FAM171A1	family with sequence similarity 171, member A1	<0.001	2.14	10
GC13M080910_at	SPRY2	sprouty homolog 2 (Drosophila)	<0.001	-2.82	**13**
GC17M056347_at	MPO	myeloperoxidase	<0.001	-2.97	17

Top 21 significantly deregulated genes between AML+13 (n=9) and AML without an additional chromosome 13 (n=509) derived from the gene expression data set GSE37642. P-Value adjustment was done with the Benjamini Hochberg method. A positive value in log fold change means an over expression in the AML+13 subgroup, and a negative value a lower expression of this gene in the AML+13 subgroup.

Table S9 A: Genes within significant regions as identified by MACAT

ProbeSet ID	Cytoband	Gene Symbol	Gene Description	Score	p-Value
226724_s_at	13q34	GAS6	growth arrest-specific 6	-0.1	0.86
226574_at	13q12.2	RPL21	ribosomal protein L21	1.62	0.008
222612_at	13q12	HMGB1	high-mobility group box 1	2.7	<0.001
222611_s_at	13q12	HMGB1	high-mobility group box 1	1.48	0.016
218371_s_at	13q34	GAS6	growth arrest-specific 6	1.96	0.004
235620_x_at	13q13.1	N4BP2L2	NEDD4 binding protein 2-like 2	0.73	0.21
215948_x_at	13q13.1	N4BP2L2	NEDD4 binding protein 2-like 2	1.36	0.041
206744_s_at	13q33	ERCC5	excision repair cross-complementing rodent repair deficiency, complementation group 5	2.3	<0.001
206652_at	13q34	ARHGEF7	Rho guanine nucleotide exchange factor (GEF) 7	2.35	0.001
218479_s_at	13q34	ARHGEF7	Rho guanine nucleotide exchange factor (GEF) 7	0.9	0.163
222649_at	13q34	CDC16	cell division cycle 16 homolog (S. cerevisiae)	1.82	0.001
223379_s_at	13q14.1	FOXO1	forkhead box O1	0.86	0.042
223380_s_at	13q14.1	FOXO1	forkhead box O1	2.44	<0.001
227013_at	13q12.2-q13.3	SUCLA2	succinate-CoA ligase, ADP-forming, beta subunit	2.55	<0.001
230348_at	13q14.2	RB1	retinoblastoma 1	3.32	<0.001
214429_at	13q31.1	SPRY2	sprouty homolog 2 (Drosophila)	2.21	0.004
228789_at	13q12-q14	USPL1	ubiquitin specific peptidase like 1	2.09	0.008
204435_at	13q22	EDNRB	endothelin receptor type B	1.75	0.012
223984_s_at	13q22	EDNRB	endothelin receptor type B	0.5	0.292
225047_at	13q12.13	NUPL1	nucleoporin like 1	2.03	0.004
241425_at	13q12.3	PDS5B	PDS5, regulator of cohesion maintenance, homolog B (S. cerevisiae)	1.99	0.004
204831_at	13q14.3	RCBTB2	regulator of chromosome condensation (RCC1) and BTB (POZ) domain containing protein 2	0.98	0.138
200012_x_at	13q12	CDK8	cyclin-dependent kinase 8	1.15	0.014
238353_at	13q14	NUFIP1	nuclear fragile X mental retardation protein interacting protein 1	-0.3	0.49
243092_at	13q14	NUFIP1	nuclear fragile X mental retardation protein interacting protein 1	4.17	<0.001
225563_at	13q14	NUFIP1	nuclear fragile X mental retardation protein interacting protein 1	3.59	<0.001
233804_at	13q32.3	GPR183	G protein-coupled receptor 183	-0.6	0.111
227713_at	13q22	DACH1	dachshund homolog 1 (Drosophila)	2.22	0.002
223790_at	13q22	DACH1	dachshund homolog 1 (Drosophila)	1.46	0.01
224734_at	13q34	RASA3	RAS p21 protein activator 3	3.34	<0.001
224731_at	13q34	RASA3	RAS p21 protein activator 3	1.63	0.009
214938_x_at	13q33-q34	LIG4	ligase IV, DNA, ATP-dependent	1.24	0.01
200680_x_at	13q12	ZMYM5	zinc finger, MYM-type 5	1.55	0.002
200679_x_at	13q22	EDNRB	endothelin receptor type B	1.88	0.011
204190_at	13q12	ZMYM5	zinc finger, MYM-type 5	2.07	0.002
215105_at	13q34	UPF3A	UPF3 regulator of nonsense transcripts homolog A (yeast)	0.72	0.117
242576_x_at	13q34	UPF3A	UPF3 regulator of nonsense transcripts homolog A (yeast)	1.01	0.021
235547_at	13q12.3	PDS5B	PDS5, regulator of cohesion maintenance, homolog B (S. cerevisiae)	0.7	0.155
221899_at	13q21.2	TDRD3	tudor domain containing 3	1.73	0.013
214753_at	13q34	ING1	inhibitor of growth family, member 1	1.79	0.006
214748_at	13q31.2-q32.3	STK24	serine/threonine kinase 24	0.08	0.887
202259_s_at	13q31.2-q32.3	STK24	serine/threonine kinase 24	1.6	0.012
202258_s_at	13q14.3	LCP1	lymphocyte cytosolic protein 1 (L-plastin)	1.44	0.02
242302_at	13q34	CDC16	cell division cycle 16 homolog (S. cerevisiae)	0.04	0.895
215888_at	13q34	CDC16	cell division cycle 16 homolog (S. cerevisiae)	1.22	0.031
207956_x_at	13q34	ING1	inhibitor of growth family, member 1	2.52	0.001
204742_s_at	13q32	GPR18	G protein-coupled receptor 18	2.7	<0.001
233432_at	13q34	ING1	inhibitor of growth family, member 1	-0.3	0.439
228484_s_at	13q14.13	NEK3	NIMA (never in mitosis gene a)-related kinase 3	0.28	0.443
202724_s_at	13q14.2	RB1	retinoblastoma 1	4.93	<0.001
202723_s_at	13q14	TPT1	tumor protein, translationally-controlled 1	4.65	<0.001
212418_at	13q32.2	IPO5	importin 5	1.91	<0.001
212420_at	13q32.2	IPO5	importin 5	1.79	0.011
220656_at	13q32.2	IPO5	importin 5	0.22	0.404
219378_at	13q32.2	IPO5	importin 5	2.92	<0.001
205134_s_at	13q14	TPT1	tumor protein, translationally-controlled 1	1.08	0.047

205135_s_at	13q13	ELF1	E74-like factor 1 (ets domain transcription factor)	1.58	0.017
205136_s_at	13q13	ELF1	E74-like factor 1 (ets domain transcription factor)	1.37	0.008
229891_x_at	13q14	TPT1	tumor protein, translationally-controlled 1	2.18	0.005
229078_s_at	13q14.13	NEK3	NIMA (never in mitosis gene a)-related kinase 3	1.49	0.008
226429_at	13q33.1	C13orf27	chromosome 13 open reading frame 27	1.36	0.005
223606_x_at	13q21.2	TDRD3	tudor domain containing 3	1.63	0.035
220171_x_at	13q34	UPF3A	UPF3 regulator of nonsense transcripts homolog A (yeast)	0.34	0.593
216520_s_at	13q14	TPT1	tumor protein, translationally-controlled 1	0.68	0.132
214327_x_at	13q12	MTMR6	myotubularin related protein 6	0.14	0.65
212869_x_at	13q13.1	N4BP2L2	NEDD4 binding protein 2-like 2	-0.1	0.8
212284_x_at	13q13.1	N4BP2L2	NEDD4 binding protein 2-like 2	-0.1	0.799
211943_x_at	13q14.11	LRCH1	leucine-rich repeats and calponin homology (CH) domain containing 1	0.01	0.96
227709_at	13q12	HMGB1	high-mobility group box 1	1.71	0.005
227710_s_at	13q14.3	RNASEH2B	ribonuclease H2, subunit B	1.26	0.007
228913_at	13q12-q13	CG030	hypothetical CG030	0.32	0.503
238171_at	13q31.2-q32.3	STK24	serine/threonine kinase 24	0.51	0.23
226782_at	13q12.3	PDS5B	PDS5, regulator of cohesion maintenance, homolog B (S. cerevisiae)	3.53	<0.001
223450_s_at	13q12	ZMYM5	zinc finger, MYM-type 5	3.18	<0.001
238228_at	13q21	ATXN8OS	ATXN8 opposite strand (non-protein coding)	-0.2	0.827
208885_at	13q14	TPT1	tumor protein, translationally-controlled 1	0.32	0.608
219471_at	13q34	UPF3A	UPF3 regulator of nonsense transcripts homolog A (yeast)	5.51	<0.001
44790_s_at	13q14.3	ITM2B	integral membrane protein 2B	6.05	<0.001
235012_at	13q14.3	ITM2B	integral membrane protein 2B	3.74	<0.001
226795_at	13q14.2	MED4	mediator complex subunit 4	4.45	<0.001
214936_at	13q34	ANKRD10	ankyrin repeat domain 10	0.49	0.346
202930_s_at	13q14	RCBTB1	regulator of chromosome condensation (RCC1) and BTB (POZ) domain containing protein 1	0.68	0.382
219347_at	13q12.11	PSPC1	paraspeckle component 1	0.58	0.408
217843_s_at	13q11	XPO4	exportin 4	2.79	<0.001
222438_at	13q14	LPAR6	lysophosphatidic acid receptor 6	2.98	<0.001
217731_s_at	13q14.3	INTS6	integrator complex subunit 6	1.24	0.023
217732_s_at	13q14.3	RNASEH2B	ribonuclease H2, subunit B	0.95	0.079
203132_at	13q14.2	NUDT15	nudix (nucleoside diphosphate linked moiety X)-type motif 15	2.13	0.006
211540_s_at	13q14.11	NAA16	N(alpha)-acetyltransferase 16, NatA auxiliary subunit	3.21	<0.001
218589_at	13q14.13	C13orf18	chromosome 13 open reading frame 18	4.66	<0.001
204759_at	13q33	KDELC1	KDEL (Lys-Asp-Glu-Leu) containing 1	1.72	0.026
220813_at	13q13-q14	KIAA1704	KIAA1704	-0.7	0.17
218352_at	13q14.3	GUCY1B2	guanylate cyclase 1, soluble, beta 2	2.43	0.001
237417_at	13q14.11	NAA16	N(alpha)-acetyltransferase 16, NatA auxiliary subunit	0.76	0.05
223306_at	13q14.2	CYSLTR2	cysteinyl leukotriene receptor 2	0.95	0.2
221503_s_at	13q14.3	KPNA3	karyopherin alpha 3 (importin alpha 4)	2.45	<0.001
221502_at	13q14.3	KPNA3	karyopherin alpha 3 (importin alpha 4)	2.26	<0.001
233277_at	13q13.1	N4BP2L2	NEDD4 binding protein 2-like 2	-0.7	0.088
233156_at	13q14.3	INTS6	integrator complex subunit 6	4.09	<0.001
229210_at	13q14.2	MED4	mediator complex subunit 4	0.95	0.203
219056_at	13q12.11	PSPC1	paraspeckle component 1	2.04	0.002
215040_at	13q12.11	PSPC1	paraspeckle component 1	2.08	0.001
220506_at	13q11	XPO4	exportin 4	-0.5	0.239
218819_at	13q33.1	BIVM	basic, immunoglobulin-like variable motif containing	1.93	0.007
222239_s_at	13q34	ANKRD10	ankyrin repeat domain 10	2.14	0.001
235283_at	13q12-q13	EBPL	emopamil binding protein-like	1.38	0.013
213116_at	13q11-q12	LATS2	LATS, large tumor suppressor, homolog 2 (Drosophila)	2.66	0.001
211089_s_at	13q11-q12	LATS2	LATS, large tumor suppressor, homolog 2 (Drosophila)	1.38	0.012
208089_s_at	13q14.13	COG3	component of oligomeric golgi complex 3	1.99	0.003
214028_x_at	13q13-q14	KIAA1704	KIAA1704	1.57	0.001
232054_at	13q12.3	KATNAL1	katanin p60 subunit A-like 1	-0.6	0.197
223810_at	13q21	KLHL1	kelch-like 1 (Drosophila)	-0.3	0.433
216404_at	13q32.3	FKSG29	FKSG29	-0.1	0.742
228915_at	13q12.13	NUPL1	nucleoporin like 1	3.32	<0.001
205472_s_at	13q32.3	UBAC2	UBA domain containing 2	2.89	0.005
205471_s_at	13q12	HMGB1	high-mobility group box 1	3.4	0.002
225619_at	13q12	HMGB1	high-mobility group box 1	1.9	0.009
206701_x_at	13q12.13	NUPL1	nucleoporin like 1	1.28	0.03
204273_at	13q34	RASA3	RAS p21 protein activator 3	1.36	0.008
204271_s_at	13q12.2	PAN3	PAN3 poly(A) specific ribonuclease subunit homolog (S. cerevisiae)	1.42	0.022

204011_at	13q22.3	SLAIN1	SLAIN motif family, member 1	-5.1	<0.001
236734_at	13q34	ZNF828	zinc finger protein 828	0.01	0.977
236906_x_at	13q13-q14	KIAA1704	KIAA1704	0.6	0.196
211955_at	13q12.11	PSPC1	paraspeckle component 1	1.86	0.002
211954_s_at	13q34	ANKRD10	ankyrin repeat domain 10	2.06	0.001
211953_s_at	13q12.11	PSPC1	paraspeckle component 1	2.31	0.002
211952_at	13q14.13	SLC25A30	solute carrier family 25, member 30	0.41	0.55
215188_at	13q14.11	LRCH1	leucine-rich repeats and calponin homology (CH) domain containing 1	0.02	0.958
208855_s_at	13q11-q12	LATS2	LATS, large tumor suppressor, homolog 2 (Drosophila)	3.06	<0.001
208854_s_at	13q34	ANKRD10	ankyrin repeat domain 10	3.47	<0.001
224298_s_at	13q14.13	LOC100190939	hypothetical LOC100190939	1.22	0.046
210279_at	13q14.13	LOC100190939	hypothetical LOC100190939	3.88	<0.001
205419_at	13q12.3	KATNAL1	katanin p60 subunit A-like 1	1.67	0.047
223896_at	13q33-q34	LIG4	ligase IV, DNA, ATP-dependent	-0.5	0.239
213346_at	13q14.1	FOXO1	forkhead box O1	-2.3	0.002
219479_at	13q12	MTMR6	myotubularin related protein 6	-2.4	0.005
222761_at	13q14.13	LOC100190939	hypothetical LOC100190939	-3.8	<0.001
229478_x_at	13q22	DACH1	dachshund homolog 1 (Drosophila)	-0.3	0.436
229589_x_at	13q13-q14	KIAA1704	KIAA1704	-1.2	0.016
233255_s_at	13q14.3	RNASEH2B	ribonuclease H2, subunit B	-2.9	<0.001
202414_at	13q33.1	BIVM	basic, immunoglobulin-like variable motif containing	1.79	0.003
227766_at	13q33.1	BIVM	basic, immunoglobulin-like variable motif containing	1.92	0.011
206235_at	13q34	ARHGEF7	Rho guanine nucleotide exchange factor (GEF) 7	1.26	0.028
234993_at	13q13-q14	KIAA1704	KIAA1704	1.74	0.012
235348_at	13q11-q12	LATS2	LATS, large tumor suppressor, homolog 2 (Drosophila)	3.93	<0.001
208415_x_at	13q21	PCDH20	protocadherin 20	1.11	0.025
209808_x_at	13q14.3	RNASEH2B	ribonuclease H2, subunit B	1.68	0.004
210350_x_at	13q33.1	BIVM	basic, immunoglobulin-like variable motif containing	1.22	0.029
241414_at	13q14.3	DLEU7	deleted in lymphocytic leukemia, 7	0.74	0.034
239116_at	13q14.11	LOC646982	twelve-thirteen translocation leukemia gene	-0.3	0.477
227260_at	13q12.3	LOC440131	hypothetical LOC440131	0.56	0.386
226663_at	13q33.3	ABHD13	abhydrolase domain containing 13	1.07	0.123
223251_s_at	13q14.11	LRCH1	leucine-rich repeats and calponin homology (CH) domain containing 1	2.2	0.002
218093_s_at	13q14.3	INTS6	integrator complex subunit 6	3.24	<0.001
242999_at	13q33.3	ABHD13	abhydrolase domain containing 13	0.8	0.202
239397_at	13q34	ARHGEF7	Rho guanine nucleotide exchange factor (GEF) 7	0.01	0.99
236416_at	13q13.1	N4BP2L2	NEDD4 binding protein 2-like 2	0.07	0.812
235412_at	13q12	ZMYM5	zinc finger, MYM-type 5	2.75	<0.001
229642_at	13q34	ARHGEF7	Rho guanine nucleotide exchange factor (GEF) 7	0.57	0.318
202548_s_at	13q31.1	SLITRK1	SLIT and NTRK-like family, member 1	2.48	<0.001
202547_s_at	13q32.2	IPO5	importin 5	1.51	0.017
238226_at	13q14	RCBTB1	regulator of chromosome condensation (RCC1) and BTB (POZ) domain containing protein 1	0.2	0.667
1598_g_at	13q14.13	SLC25A30	solute carrier family 25, member 30	4.32	<0.001
202177_at	13q34	FAM70B	family with sequence similarity 70, member B	4.47	<0.001
225562_at	13q14.13	COG3	component of oligomeric golgi complex 3	1.92	0.004
206221_at	13q12.2	RASL11A	RAS-like, family 11, member A	0.5	0.218
206220_s_at	13q34	ANKRD10	ankyrin repeat domain 10	-0.3	0.482
202717_s_at	13q34	ARHGEF7	Rho guanine nucleotide exchange factor (GEF) 7	2.07	0.001
209658_at	13q34	ANKRD10	ankyrin repeat domain 10	1.47	0.02
209659_s_at	13q12.13	NUPL1	nucleoporin like 1	1.72	0.008
206958_s_at	13q12.3	PDS5B	PDS5, regulator of cohesion maintenance, homolog B (S. cerevisiae)	1.9	0.004
206959_s_at	13q13.1	N4BP2L2	NEDD4 binding protein 2-like 2	0.94	0.127
214323_s_at	13q34	ARHGEF7	Rho guanine nucleotide exchange factor (GEF) 7	1.6	0.026
217596_at	13q12.2	LOC100288730	hypothetical LOC100288730	0.9	0.099
226194_at	13q14.13	C13orf18	chromosome 13 open reading frame 18	1.84	0.007

Table S9 B: Differentially expressed genes located on chromosome 13 (Limma)

Probe set	Gene	Description	Adjusted P-value	Log fold change
GC13M048963_at	LPAR6	lysophosphatidic acid receptor 6	<0.001	2.72
GC13M041129_at	FOXO1	forkhead box O1	<0.001	2.32
GC13M046916_at	C13orf18	chromosome 13 open reading frame 18	<0.001	2.17
GC13M028710_at	LOC100288730	hypothetical LOC100288730	<0.001	1.68
GC13M099906_at	GPR18	G protein-coupled receptor 18	<0.001	1.68
GC13M072012_at	DACH1	dachshund homolog 1 (Drosophila)	0.001	1.67
GC13M023902_at	SACS	spastic ataxia of Charlevoix-Saguenay (sacsin)	0.001	1.35
GC13M028577_at	FLT3	fms-related tyrosine kinase 3	0.041	1.33
GC13P024734_at	SPATA13	spermatogenesis associated 13	0.001	1.24
GC13P088324_at	SLITRK5	SLIT and NTRK-like family, member 5	0.004	1.20
GC13P028713_at	PAN3	PAN3 poly(A) specific ribonuclease subunit homolog (S. cerevisiae)	<0.001	1.13
GC13M114523_at	GAS6	growth arrest-specific 6	<0.001	1.09
GC13M050106_at	RCBTB1	regulator of chromosome condensation (RCC1) and BTB (POZ) domain containing protein 1	0.037	0.97
GC13P041885_at	NAA16	N(alpha)-acetyltransferase 16, NatA auxiliary subunit	<0.001	0.95
GC13P042846_at	AKAP11	A kinase (PRKA) anchor protein 11	<0.001	0.91
GC13P046039_at	COG3	component of oligomeric golgi complex 3	<0.001	0.90
GC13P048877_at	RB1	retinoblastoma 1	0.001	0.88
GC13M099103_at	STK24	serine/threonine kinase 24	<0.001	0.85
GC13M030338_at	UBL3	ubiquitin-like 3	0.019	0.77
GC13M111530_at	ANKRD10	ankyrin repeat domain 10	<0.001	0.77
GC13M021547_at	LATS2	LATS, large tumor suppressor, homolog 2 (Drosophila)	0.009	0.76
GC13M048627_at	MED4	mediator complex subunit 4	<0.001	0.75
GC13M050273_at	KPNA3	karyopherin alpha 3 (importin alpha 4)	0.003	0.70
GC13P047127_at	LRCH1	leucine-rich repeats and calponin homology (CH) domain containing 1	<0.001	0.70
GC13P108870_at	ABHD13	abhydrolase domain containing 13	0.008	0.70
GC13P037572_at	EXOSC8	exosome component 8	0.045	0.69
GC13M051928_at	INTS6	integrator complex subunit 6	0.004	0.69
GC13P098086_at	RAP2A	RAP2A, member of RAS oncogene family	0.049	0.68
GC13P043597_at	DNAJC15	DnaJ (Hsp40) homolog, subfamily C, member 15	0.041	0.68
GC13M025820_at	MTMR6	myotubularin related protein 6	0.025	0.67
GC13P031191_at	USPL1	ubiquitin specific peptidase like 1	0.026	0.67
GC13P050069_at	PHF11	PHD finger protein 11	0.006	0.67
GC13P033160_at	PDS5B	PDS5, regulator of cohesion maintenance, homolog B (S. cerevisiae)	<0.001	0.65
GC13P021276_at	IL17D	interleukin 17D	0.005	0.62
GC13M113139_at	TUBGCP3	tubulin, gamma complex associated protein 3	<0.001	0.59
GC13P052158_at	WDFY2	WD repeat and FYVE domain containing 2	0.001	0.59

GC13P020532_at	ZMYM2	zinc finger, MYM-type 2	0.003	0.58
GC13P037393_at	RFXAP	regulatory factor X-associated protein	0.028	0.57
GC13M052706_at	NEK3	NIMA (never in mitosis gene a)-related kinase 3	0.002	0.57
GC13M041506_at	ELF1	E74-like factor 1 (ets domain transcription factor)	0.044	0.56
GC13P021714_at	SAP18	Sin3A-associated protein, 18kDa	0.001	0.56
GC13P027998_at	GTF3A	general transcription factor IIIA	0.005	0.56
GC13P098605_at	IPO5	importin 5	0.029	0.55
GC13P051483_at	RNASEH2B	ribonuclease H2, subunit B	0.032	0.54
GC13M030776_at	KATNAL1	katanin p60 subunit A-like 1	0.007	0.53
GC13M045967_at	SLC25A30	solute carrier family 25, member 30	<0.001	0.52
GC13M020249_at	PSPC1	paraspeckle component 1	0.008	0.49
GC13P025875_at	NUPL1	nucleoporin like 1	0.022	0.48
GC13M031032_at	HMGB1	high-mobility group box 1	0.001	0.48
GC13M073329_at	DIS3	DIS3 mitotic control homolog (S. cerevisiae)	0.038	0.46
GC13P060970_at	TDRD3	tudor domain containing 3	0.016	0.46
GC13M021950_at	ZDHHC20	zinc finger, DHHC-type containing 20	0.027	0.43
GC13M107194_at	ARGLU1	arginine and glutamate rich 1	0.013	0.38
GC13M020397_at	ZMYM5	zinc finger, MYM-type 5	0.015	0.35
GC13M079888_at	RBM26	RNA binding motif protein 26	0.033	0.30
GC13M096453_at	UGGT2	UDP-glucose glycoprotein glucosyltransferase 2	0.026	-0.31
GC13M103418_at	C13orf27	chromosome 13 open reading frame 27	0.026	-0.90
GC13P103451_at	BIVM	basic, immunoglobulin-like variable motif containing	0.001	-1.13
GC13M080910_at	SPRY2	sprouty homolog 2 (Drosophila)	<0.001	-2.82

Table S9 B displays all significantly deregulated genes located on chromosome 13 (Limma). P-Value adjustment was done with the Benjamini Hochberg method. A positive value in log fold change means an over expression in the AML+13 subgroup, and a negative value a lower expression of this gene in the AML+13 subgroup.

Table S10: GSEA results

NAME	ES	NES	NOM p-val	FDR q-val
NUCLEOCYTOPLASMIC_TRANSPORT	0.494	1.747	<0.001	0.117
PROTEIN_POLYMERIZATION	0.685	1.869	<0.001	0.118
N_ACETYLTRANSFERASE_ACTIVITY	0.605	1.773	0.010	0.123
PROTEIN_IMPORT_INTO_NUCLEUS	0.567	1.782	0.004	0.127
NUCLEAR_TRANSPORT	0.491	1.747	<0.001	0.128
REGULATION_OF_ORGANELLE_ORGANIZATION_AND_BIOGENESIS	0.525	1.789	0.002	0.133
PROTEIN_BINDING__BRIDGING	0.523	1.759	0.004	0.134
PROTEIN_IMPORT	0.534	1.799	<0.001	0.137
N_ACYLTRANSFERASE_ACTIVITY	0.608	1.748	0.008	0.139
NEGATIVE_REGULATION_OF_RNA_METABOLIC_PROCESS	0.492	1.719	<0.001	0.146
PROTEIN_MODIFICATION_BY_SMALL_PROTEIN_CONJUGATION	0.531	1.706	0.017	0.149
ACETYLTRANSFERASE_ACTIVITY	0.568	1.709	0.010	0.154
NEGATIVE_REGULATION_OF_TRANSCRIPTION__DNA_DEPENDENT	0.492	1.719	<0.001	0.158
NUCLEAR_IMPORT	0.575	1.804	0.002	0.159
TRANSCRIPTION_ACTIVATOR_ACTIVITY	0.424	1.694	0.008	0.165
PROTEIN_UBIQUITINATION	0.528	1.684	0.017	0.170
RNA_POLYMERASE_II_TRANSCRIPTION_FACTOR_ACTIVITY	0.393	1.686	0.004	0.175
MOLECULAR_ADAPTOR_ACTIVITY	0.597	1.808	0.002	0.199
HISTONE_ACETYLTRANSFERASE_ACTIVITY	0.623	1.656	0.040	0.222
SH3_SH2_ADAPTOR_ACTIVITY	0.627	1.870	<0.001	0.228
POSITIVE_REGULATION_OF_TRANSCRIPTION	0.402	1.645	0.010	0.239

Comparison of patients with AML+13 and the control group using the “c5all” gene sets implemented in GSEA. ES: enrichment score; NES; nominal enrichment score; NOM p-val: nominal p-value; FDR q-val: false discovery rate. Only gene sets with an FDR of <0.25 are displayed.

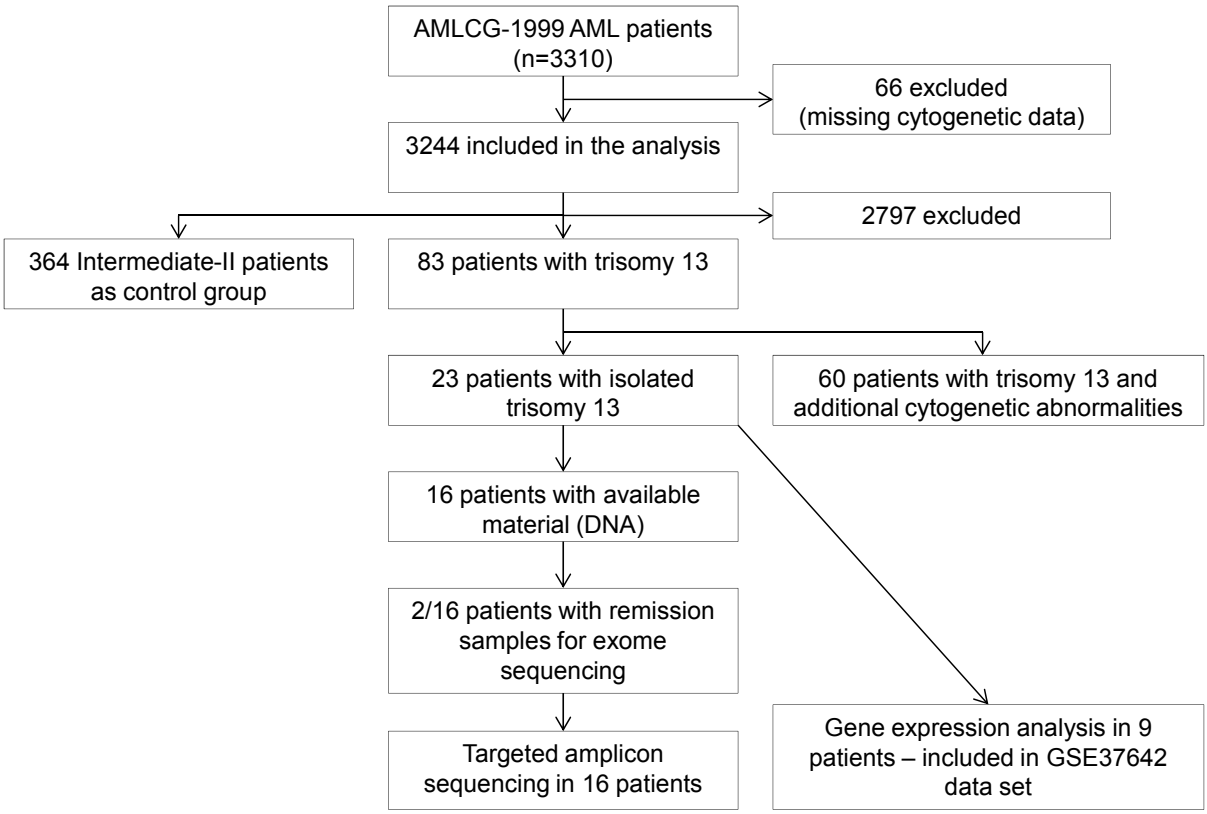
Table S11: Overlap of genes differentially expressed genes in RUNX1 mutated AML with normal karyotyp and AML+13

Microarray probe set	Gene symbol	Gene name	Differnetially expressed in AML+13
GC03M015700_at	ANKRD28	ankyrin repeat domain 28	no
GC04M122868_at	ANXA5	annexin A5	no
GC08P104222_at	BAALC	brain and acute leukemia, cytoplasmic	yes
GC02M060589_at	BCL11A	B-cell CLL/lymphoma 11A (zinc finger protein)	no
GC10M097941_at	BLNK	B-cell linker	yes
GC07P043764_at	BLVRA	biliverdin reductase A	no
GC13M102219_at	C13orf27	chromosome 13 open reading frame 27	yes
GC14M094943_at	C14orf139	chromosome 14 open reading frame 139	no
GC04P113286_at	C4orf32	chromosome 4 open reading frame 32	no
GC01P221966_at	CAPN2	calpain 2, (m/II) large subunit	no
GC11P034417_at	CAT	catalase	yes
GC16M087468_at	CBFA2T3	core-binding factor, runt domain, alpha subunit 2; translocated to, 3	no
GC04P110700_at	CCDC109B	coiled-coil domain containing 109B	no
GC01P026516_at	CD52	CD52 molecule	no
GC03P112743_at	CD96	CD96 molecule	no
GC0XM109724_at	CHRDL1	chordin-like 1	yes
GC05M149413_at	CSF1R	colony stimulating factor 1 receptor	no
GC14M024112_at	CTSG	cathepsin G	yes
GC11P065405_at	CTSW	cathepsin W	no
GC02P237143_at	CXCR7	chemokine (C-X-C motif) receptor 7	yes
GC15P020444_at	CYFIP1	cytoplasmic FMR1 interacting protein 1	no
GC10P098054_at	DNTT	deoxynucleotidyltransferase, terminal	yes
GC08P026491_at	DPYSL2	dihydropyrimidinase-like 2	no
GC06P116708_at	DSE	dermatan sulfate epimerase	no
GC18P027332_at	DSG2	desmoglein 2	no
GC02P047425_at	EPCAM	epithelial cell adhesion molecule	no
GC13M042358_at	EPSTI1	epithelial stromal interaction 1 (breast)	no
GC06M006089_at	F13A1	coagulation factor XIII, A1 polypeptide	no
GC10M015294_at	FAM171A1	family with sequence similarity 171, member A1	yes
GC01P117860_at	FAM46C	family with sequence similarity 46, member C	no
GC0XP135057_at	FHL1	four and a half LIM domains 1	yes
GC01M089290_at	GBP1	guanylate binding protein 1, interferon-inducible, 67kDa	no
GC01M089345_at	GBP2	guanylate binding protein 2, interferon-inducible	no
GC16M019422_at	GDE1	glycerophosphodiester phosphodiesterase 1	no
GC07P150015_at	GIMAP2	GTPase, IMAP family member 2	no
GC07M149953_at	GIMAP6	GTPase, IMAP family member 6	no
GC07P149842_at	GIMAP7	GTPase, IMAP family member 7	no
GC07P079602_at	GNAI1	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	yes
GC07P093388_at	GNG11	guanine nucleotide binding protein (G protein), gamma 11	yes
GC04P156807_at	GUCY1A3	guanylate cyclase 1, soluble, alpha 3	no
GC06P032649_at	HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	no
GC04M057210_at	HOPX	HOP homeobox	yes
GC01P078858_at	IFI44L	interferon-induced protein 44-like	no
GC11P000303_at	IFITM1	interferon induced transmembrane protein 1 (9-27)	no
GC14M105389_at	IGHM	immunoglobulin heavy constant mu	no
GC22M022239_at	IGLL1	immunoglobulin lambda-like polypeptide 1	no
GC11M000602_at	IRF7	interferon regulatory factor 7	yes
GC15P086983_at	ISG20	interferon stimulated exonuclease gene 20kDa	yes
GC07M150272_at	KCNH2	potassium voltage-gated channel, subfamily H (eag-related), member 2	no
GC14P105461_at	KIAA0125	KIAA0125	no
GC04M025425_at	KIAA0746	KIAA0746 protein	no
GC10U900364_at	LOC283070	hypothetical LOC283070	no
GC05M088051_at	MEF2C	myocyte enhancer factor 2C	no
GC17M053702_at	MPO	myeloperoxidase	yes
GC21P041720_at	MX1	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)	no
GC21P041655_at	MX2	myxovirus (influenza virus) resistance 2 (mouse)	no
GC09M138039_at	NACC2	NACC family member 2, BEN and BTB (POZ) domain containing	no
GC12M000543_at	NINJ2	ninjurin 2	yes
GC11P017255_at	NUCB2	nucleobindin 2	no

GC03M152412_at	P2RY14	purinergic receptor P2Y, G-protein coupled, 14	yes
GC07M139370_at	PARP12	poly (ADP-ribose) polymerase family, member 12	no
GC10M119033_at	PDZD8	PDZ domain containing 8	no
GC07M076779_at	PION	pigeon homolog (Drosophila)	no
GC02M037389_at	PRKD3	protein kinase D3	no
GC08M141737_at	PTK2	PTK2 protein tyrosine kinase 2	yes
GC02M001606_at	PXDN	peroxidasin homolog (Drosophila)	no
GC01M152220_at	RAB13	RAB13, member RAS oncogene family	no
GC08P030361_at	RBPMS	RNA binding protein with multiple splicing	no
GC14P020429_at	RNASE3	ribonuclease, RNase A family, 3 (eosinophil cationic protein)	yes
GC02M165652_at	SCN3A	sodium channel, voltage-gated, type III, alpha subunit	yes
GC18P040535_at	SETBP1	SET binding protein 1	yes
GC04M140646_at	SETD7	SET domain containing (lysine methyltransferase) 7	yes
GC22P049402_at	SHANK3	SH3 and multiple ankyrin repeat domains 3	no
GC12M044867_at	SLC38A1	solute carrier family 38, member 1	yes
GC12P092466_at	SOCS2	suppressor of cytokine signaling 2	yes
GC18M051045_at	TCF4	transcription factor 4	no
GC10P114700_at	TCF7L2	transcription factor 7-like 2 (T-cell specific, HMG-box)	yes
GC12M115961_at	TESC	tescalcin	no
GC05P135392_at	TGFBI	transforming growth factor, beta-induced, 68kDa	yes
GC01P012161_at	TNFRSF1B	tumor necrosis factor receptor superfamily, member 1B	no
GC07M047281_at	TNS3	tensin 3	no
GC22M026704_at	TTC28	tetratricopeptide repeat domain 28	yes
GC09M025668_at	TUSC1	tumor suppressor candidate 1	no
GC07M149094_at	ZNF467	zinc finger protein 467	no
GC18M020895_at	ZNF521	zinc finger protein 521	no

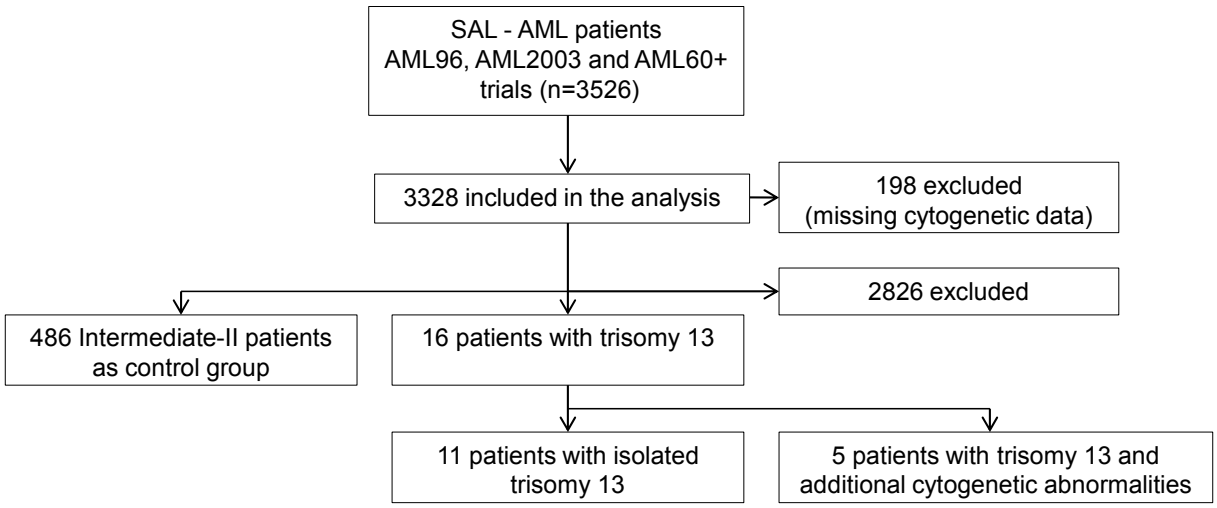
Differentially expressed genes in RUNX1-mut (n=15) vs. RUNX1-wt (n=26) patients. The analysis was restricted to NPM1-wt patients. Overlap with differentially expressed genes in AML+13 is indicated.

Figure S1 A: Study design AMLCG cohort



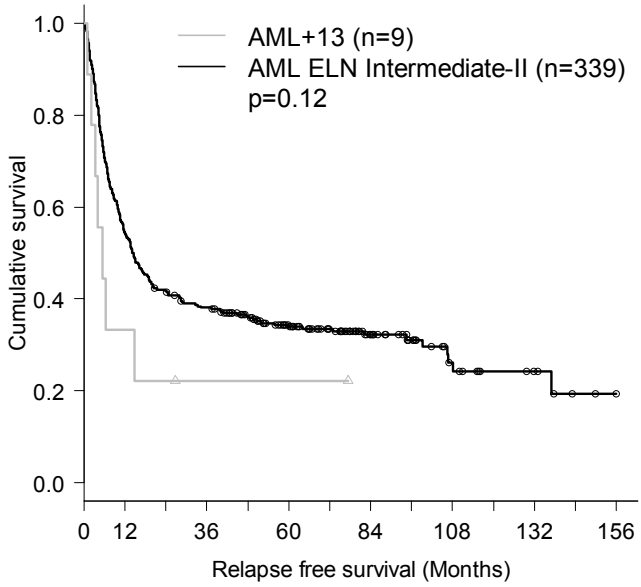
Definition of isolated trisomy 13: Isolated trisomy (n=22) or tetrasomy 13 (n=1) in absence of further cytogenetic aberrations except for numerical alterations of the sex chromosomes (n=2).

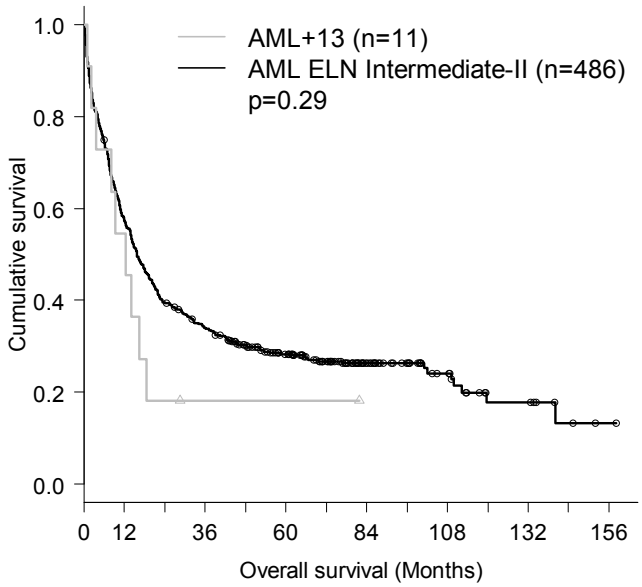
Figure S1 B: Study design SAL cohort



Definition of isolated trisomy 13: Isolated trisomy (n=11) or tetrasomy 13 (n=0) in absence of further cytogenetic aberrations except for numerical alterations of the sex chromosomes (n=0).

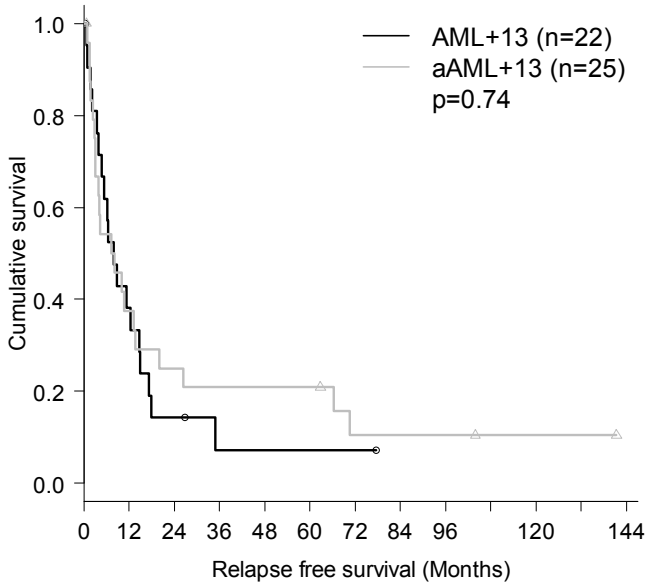
Figure S2 A: Relapse free and overall survival in AML patients (only SAL cohort)

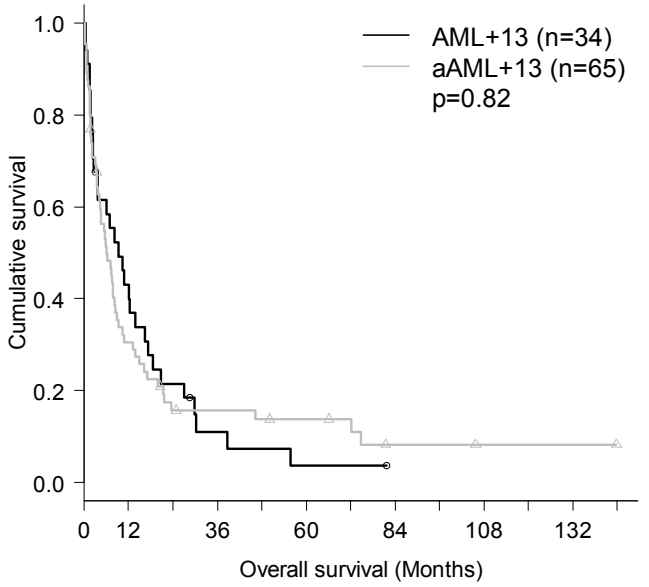




Kaplan–Meier estimates of SAL patients with isolated trisomy 13 (AML+13) and ELN Intermediate-II patients without amplifications of chromosome 13. The differences did not reach significance possibly due to the small number of AML+13 cases.

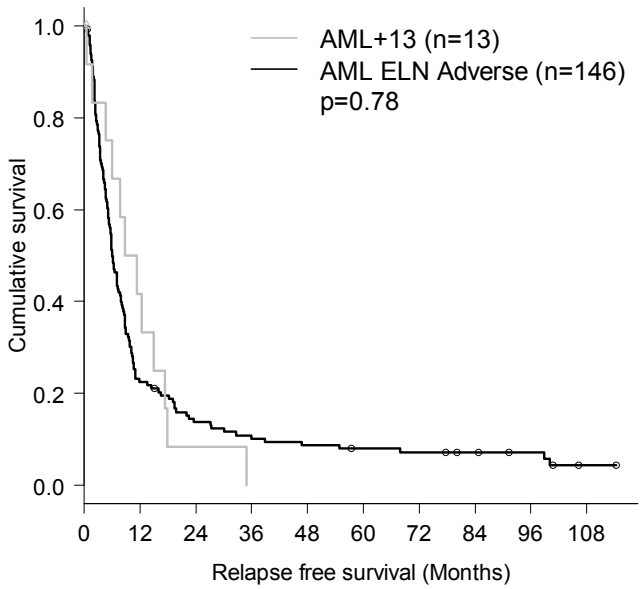
Figure S2 B: Relapse free and overall survival in AML+13 and aAML+13

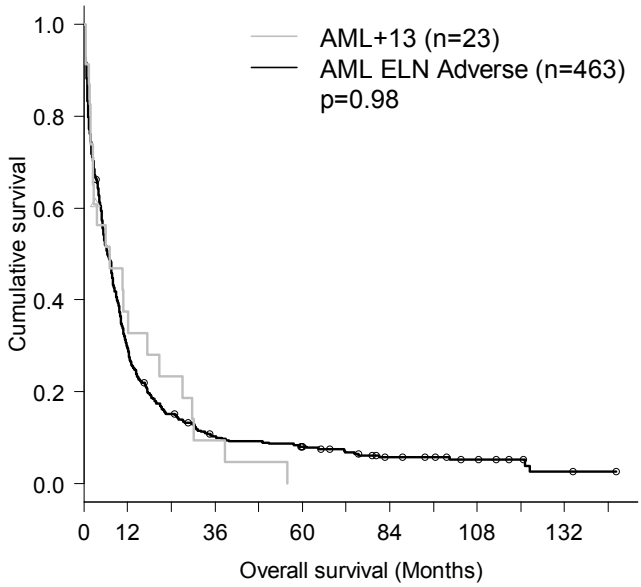




Kaplan–Meier estimates of AMLCG and SAL patients with isolated trisomy 13 (AML+13) and trisomy 13 and heterogeneous additional cytogenetic aberrations (aAML+13). There is no difference between AML+13 and the aAML+13 group regarding RFS and OS despite the high frequency of high risk aberrations in the aAML+13 group.

Figure S2 C: Relapse free and overall survival in AML+13 and ELN Adverse



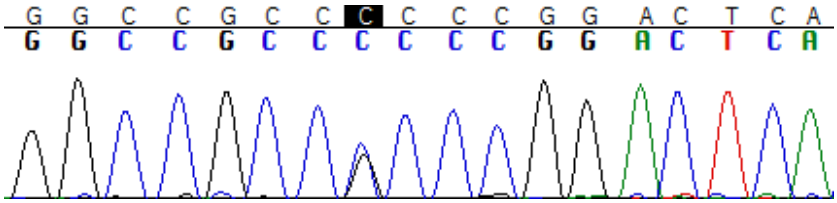


Kaplan–Meier estimates of AML in the ELN Adverse control group and patients with isolated trisomy 13 (AML+13). Only patients enrolled in the AMLCG trials are shown. There is no significant difference between the groups regarding RFS and OS

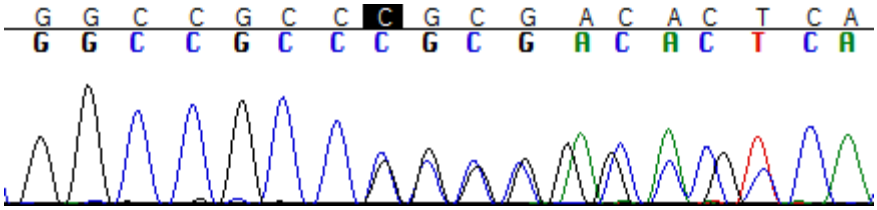
Figure S3: Results of Sanger sequencing

SRSF2

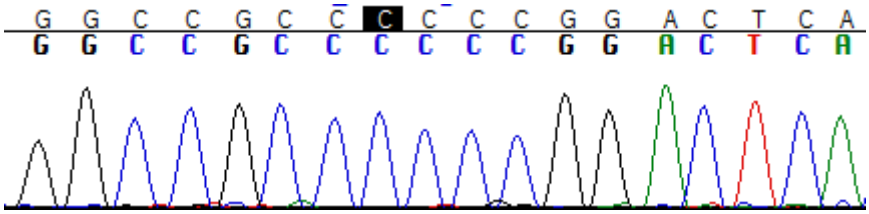
#1, SNV at position 17:74,732,959; C>G



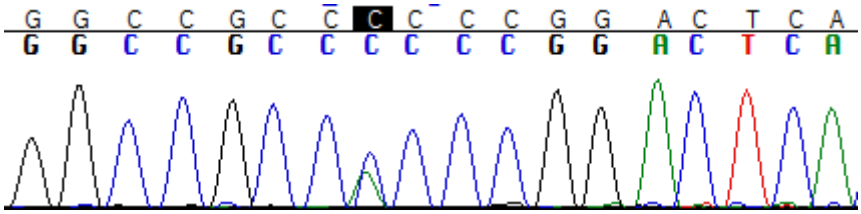
#2, in frame deletion at position 17:74,732,936-17:74,732,959



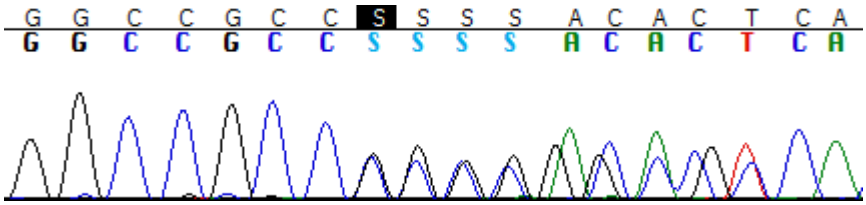
#3, wild type



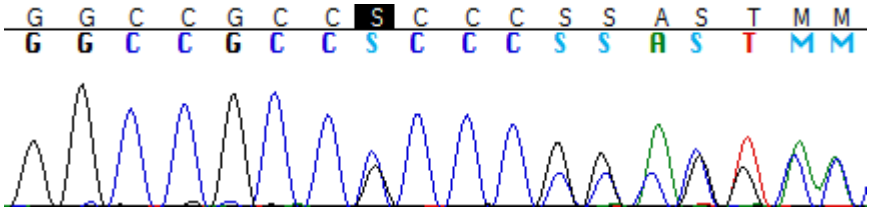
#4, SNV at position 17:74,732,959; C>A



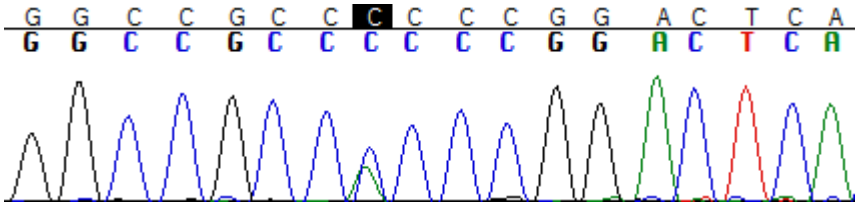
#5, in frame deletion at position 17:74,732,936-17:74,732,959



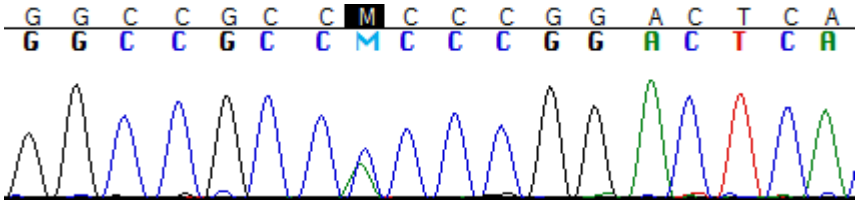
#6, in frame insertion at position 17:74,732,959



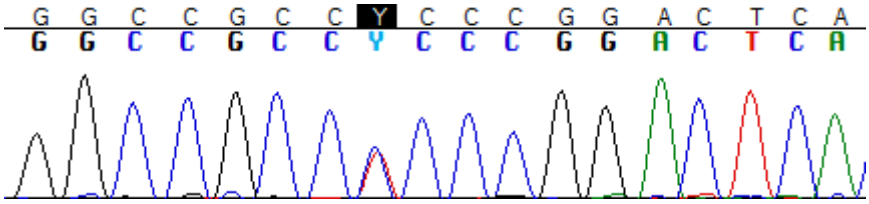
#7, SNV at position 17:74,732,959; C>A



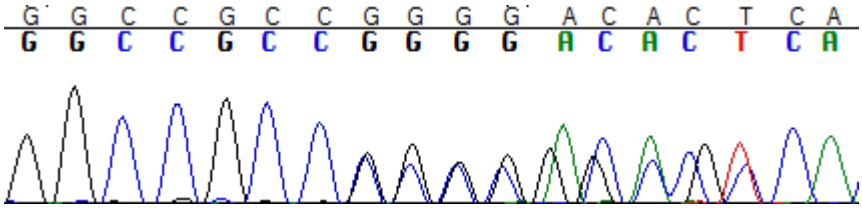
#8, SNV at position 17:74,732,959; C>A



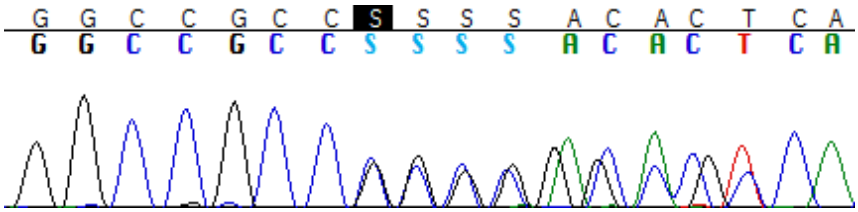
#9, SNV at position 17:74,732,959; C>T



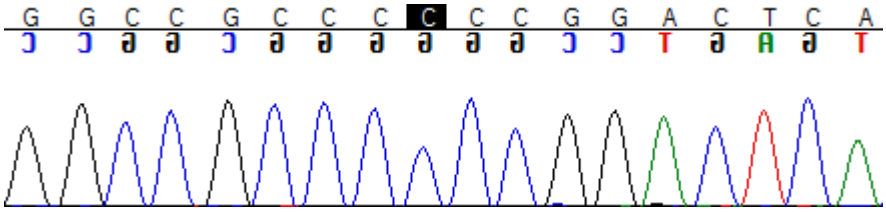
#10, in frame deletion at position 17:74,732,936-17:74,732,959



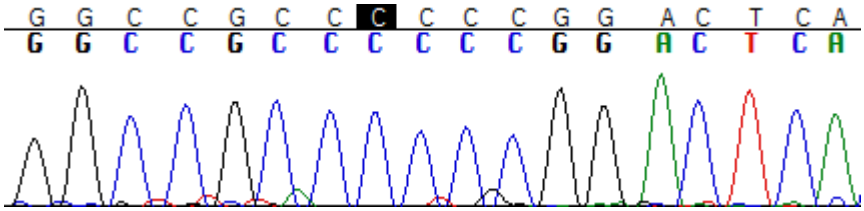
#11, in frame deletion at position 17:74,732,936-17:74,732,959



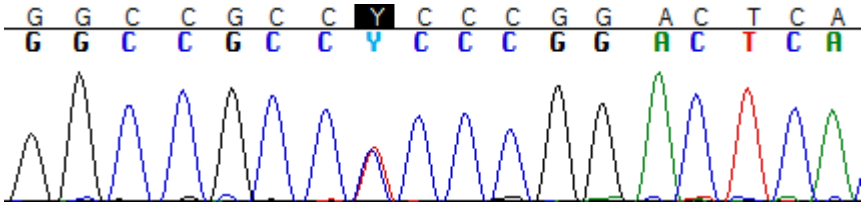
#11, Loss of SRSF2 mutation at complete remission



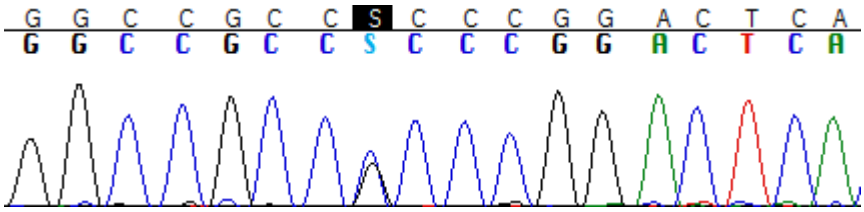
#13, wild type



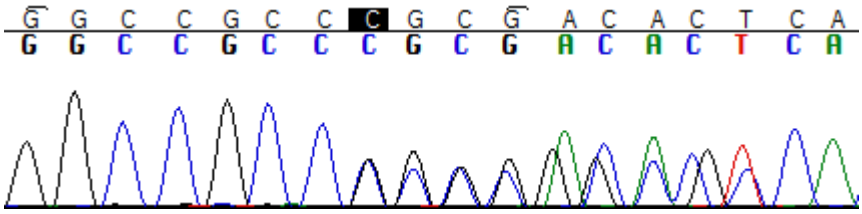
#14, SNV at position 17:74,732,959; C>T



#15, SNV at position 17:74,732,959; C>G

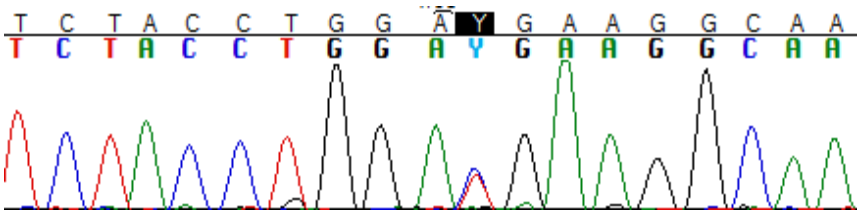


#16, in frame deletion at position 17:74,732,936-17:74,732,959

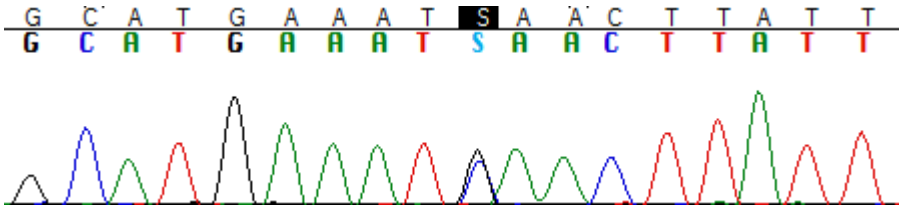


CEBPZ

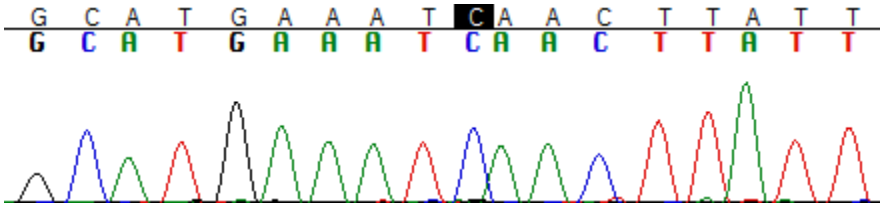
#10, SNV at position 2:37,455,632; T>C



#11, SNV at position 2:37,455,685; C>G

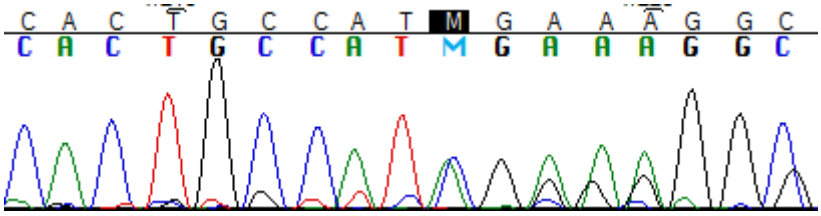


#11 Loss of *CEBPZ* mutation at complete remission



ASXL1

#8, frameshift deletion at position 20:31,022,415-20:31,022,437



nature COMMUNICATIONS

ARTICLE

Received 20 Oct 2015 | Accepted 26 Apr 2016 | Published 2 Jun 2016

DOI: 10.1038/ncomms11733 OPEN

ZBTB7A mutations in acute myeloid leukaemia with t(8;21) translocation

Luise Hartmann[1,2,3,4], Sayantanee Dutta[1,2,3,4], Sabrina Opatz[1,2,3,4], Sebastian Vosberg[1,2,3,4], Katrin Reiter[1,2,3,4], Georg Leubolt[1,2,3,4], Klaus H. Metzeler[1,2,3,4], Tobias Herold[1,2,3,4], Stefanos A. Bamopoulos[1], Kathrin Bräundl[1,2,3,4], Evelyn Zellmeier[1], Bianka Ksienzyk[1], Nikola P. Konstandin[1], Stephanie Schneider[1], Karl-Peter Hopfner[5], Alexander Graf[6], Stefan Krebs[6], Helmut Blum[3,4,6], Jan Moritz Middeke[3,4,7], Friedrich Stölzel[3,4,7], Christian Thiede[3,4,7], Stephan Wolf[4], Stefan K. Bohlander[8], Caroline Preiss[9], Linping Chen-Wichmann[9], Christian Wichmann[9], Maria Cristina Sauerland[10], Thomas Büchner[11], Wolfgang E. Berdel[11], Bernhard J. Wörmann[12], Jan Braess[13], Wolfgang Hiddemann[1,2,3,4], Karsten Spiekermann[1,2,3,4] & Philipp A. Greif[1,2,3,4]

The t(8;21) translocation is one of the most frequent cytogenetic abnormalities in acute myeloid leukaemia (AML) and results in the *RUNX1/RUNX1T1* rearrangement. Despite the causative role of the *RUNX1/RUNX1T1* fusion gene in leukaemia initiation, additional genetic lesions are required for disease development. Here we identify recurring *ZBTB7A* mutations in 23% (13/56) of AML t(8;21) patients, including missense and truncating mutations resulting in alteration or loss of the C-terminal zinc-finger domain of ZBTB7A. The transcription factor ZBTB7A is important for haematopoietic lineage fate decisions and for regulation of glycolysis. On a functional level, we show that *ZBTB7A* mutations disrupt the transcriptional repressor potential and the anti-proliferative effect of ZBTB7A. The specific association of *ZBTB7A* mutations with t(8;21) rearranged AML points towards leukaemogenic cooperativity between mutant ZBTB7A and the RUNX1/RUNX1T1 fusion.

[1] Department of Internal Medicine 3, University Hospital, Ludwig-Maximilians-Universität (LMU) München, 81377 München, Germany. [2] Clinical Cooperative Group Leukemia, Helmholtz Zentrum München, German Research Center for Environmental Health, 81377 München, Germany. [3] German Cancer Consortium (DKTK), 69121 Heidelberg, Germany. [4] German Cancer Research Center (DKFZ), 69121 Heidelberg, Germany. [5] Department of Biochemistry, Ludwig-Maximilians-Universität (LMU) München, 81377 München, Germany. [6] Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, Ludwig-Maximilians-Universität (LMU) München, 81377 München, Germany. [7] Medizinische Klinik und Poliklinik I, Universitätsklinikum Dresden, 01307 Dresden, Germany. [8] Department of Molecular Medicine and Pathology, The University of Auckland, Auckland 1142, New Zealand. [9] Department of Transfusion Medicine, Cell Therapeutics and Hemostasis, University Hospital, Ludwig-Maximilians-Universität (LMU) München, 81377 München, Germany. [10] Institute of Biostatistics and Clinical Research, University of Münster, 48149 Münster, Germany. [11] Department of Medicine A, Hematology, Oncology and Pneumology, University of Münster, 48149 Münster, Germany. [12] Department of Hematology, Oncology and Tumor Immunology, Charité University Medicine, Campus Virchow, 13353 Berlin, Germany. [13] Oncology and Hematology, St. John-of-God Hospital, 93049 Regensburg, Germany. Correspondence and requests for materials should be addressed to P.A.G. (email: pgreif@med.uni-muenchen.de and p.greif@dkfz-heidelberg.de).

Block of myeloid differentiation is one of the hallmarks of acute myeloid leukaemia (AML). First insights into this key mechanism were gained by the discovery of the t(8;21)(q22;q22) translocation, which was the first balanced translocation described in a tumour and results in the *RUNX1/RUNX1T1* fusion gene (also known as AML1/ETO)[1,2]. The *RUNX1/RUNX1T1* rearrangement is one of the most frequent chromosomal aberrations in AML and defines an important clinical entity with favourable prognosis according to the World Health Organization classification[3]. The RUNX1/RUNX1T1 fusion protein disrupts the core-binding factor complex, and thereby blocks myeloid differentiation. However, *in vivo* models indicate the requirement of additional lesions, such as of *KIT* or *FLT3* mutations, for leukaemogenesis as the *RUNX1/RUNX1T1* fusion gene alone is not sufficient to induce leukaemia[4–8]. In the present study, we set out to identify additional mutations in AML t(8;21) and discovered frequent mutations of *ZBTB7A*—encoding a transcription factor important for the regulation of haematopoietic development[9] and tumour metabolism[10]. It is very likely that *ZBTB7A* mutations are one of the important missing links in *RUNX1/RUNX1T1*-driven leukaemogenesis.

Results

***ZBTB7A* is frequently mutated in AML t(8;21).** To identify additional cooperating mutations, we performed exome sequencing of matched diagnostic and remission samples from two AML patients with t(8;21) translocation and detected 11 and 12 somatic variants, respectively (Supplementary Table 1). *ZBTB7A* was the only mutated gene identified in both patients. ZBTB7A (also known as LRF, Pokemon and FBI-1) is a member of the POZ/BTB and Krüppel (POK) transcription factor family[9], which is characterized by an N-terminal POZ/BTB protein–protein interaction domain and C-terminal C_2H_2 zinc fingers[11]. The first patient carried a homozygous missense mutation resulting in the amino-acid change R402H (NM_015898:exon2:c.1205G > A:p.R402H) affecting the highly conserved zinc-finger domain, while a heterozygous frameshift insertion (NM_015898:exon2:c.522dupC:p.A175fs) resulting in loss of the zinc-finger domain was identified in the second patient. Both mutations were validated by Sanger sequencing (Supplementary Fig. 1; Supplementary Table 2). Using targeted amplicon sequencing of *ZBTB7A* and 45 leukaemia relevant genes, we screened 56 diagnostic AML t(8;21) samples, including one of the two samples analysed by exome sequencing (UPN 1), whereas for the other one (UPN 2) availability of material was insufficient. *ZBTB7A* mutations were identified in 13 of 56 patients (23%; Fig. 1a,b; Supplementary Table 3). Patient characteristics are summarized in Supplementary Table 4. Two recurring mutational hotspots (A175fs and R402) in exon 2 were identified altering or resulting in loss of the zinc-finger domain (Fig. 1a). It was previously shown that the zinc-finger domain of *ZBTB7A* is essential for DNA binding[12]. Structural modelling revealed that arginine 402 binds into the major groove of the DNA double helix and likely contributes to the affinity or sequence specificity of the DNA interaction of the zinc-finger domain of ZBTB7A (Fig. 2a). We confirmed that both ZBTB7A mutants A175fs and R402H fail to bind DNA (Fig. 2b,c).

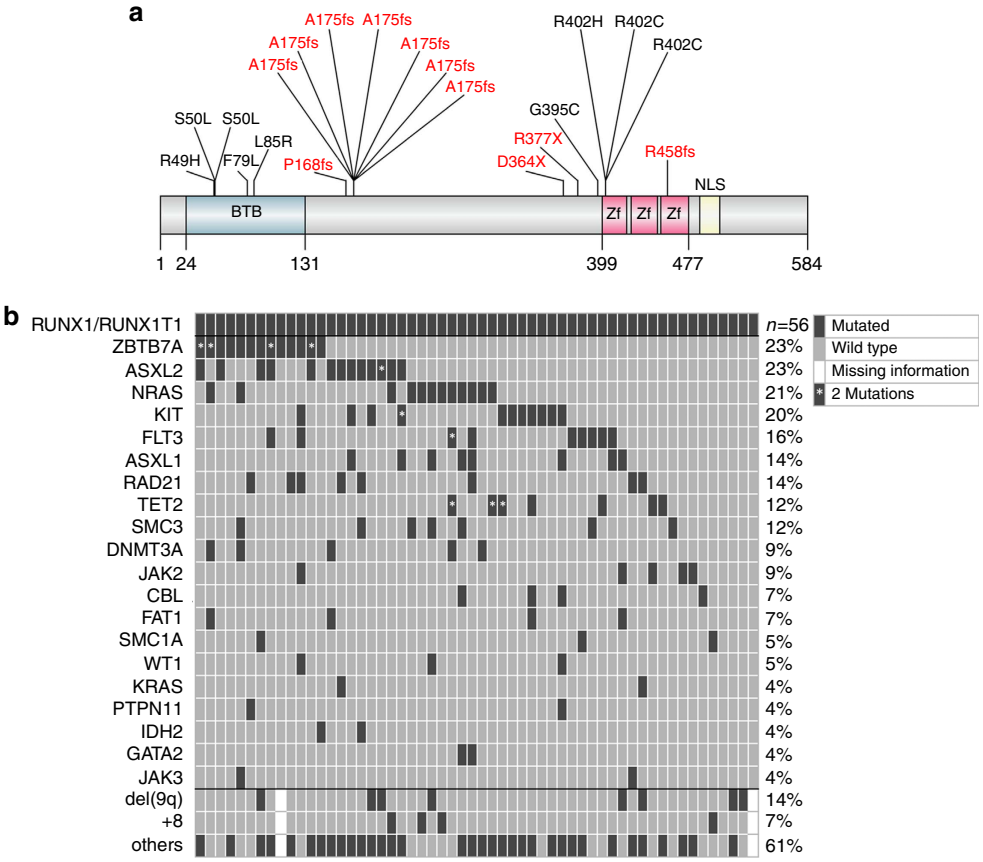


Figure 1 | *ZBTB7A* mutations in AML t(8;21). (**a**) ZBTB7A protein (NP_056982.1) and identified mutations (red = truncating; black = missense) illustrated using IBS software[31]. Amino-acid positions are indicated below the graph. BTB, BR-C ttk and bab; NLS, nuclear localization sequence; Zf, zinc finger. (**b**) Mutational landscape of 56 diagnostic AML samples with t(8;21) translocation. Each column represents one patient, each line one of the analysed genes or cytogenetic markers.

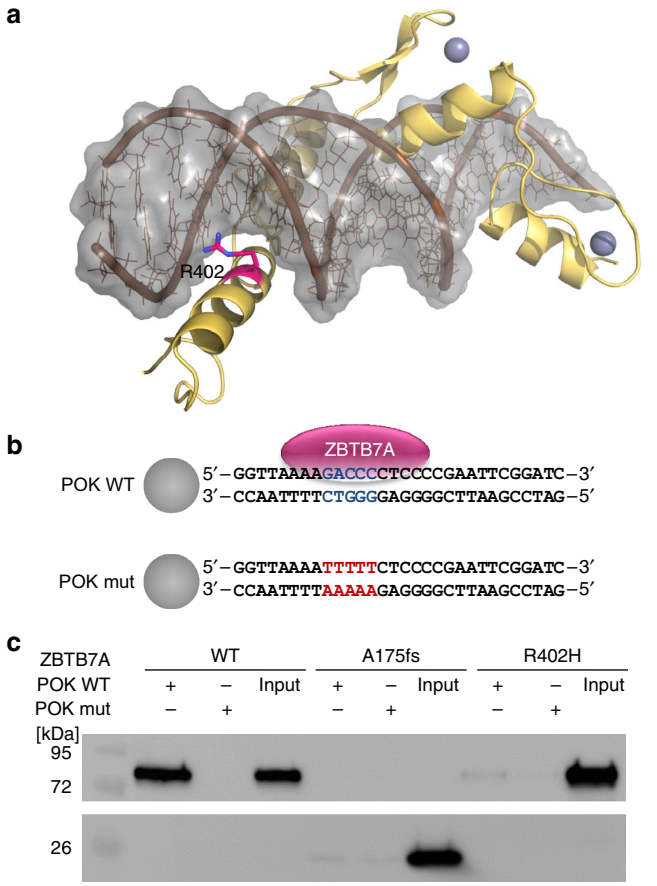


Figure 2 | Impact of ZBTB7A mutations on DNA binding. (**a**) Model for the C-terminal zinc-finger domain of ZBTB7A comprising residues 382–488. The model is depicted as yellow ribbon with highlighted secondary structure. Zinc ions are shown as grey spheres. DNA is shown in brown with a grey molecular surface. R402 (purple) binds into the major groove and likely contributes to the affinity or sequence specificity of the DNA interaction of the zinc-finger domain. (**b**) Biotinylated oligonucleotides containing the ZBTB7A (alias: Pokemon) consensus binding motif (POK WT) or a mutant thereof (POK mut)[14] used in DNA pull-down experiments. Spheres illustrate streptavidin-coated beads. (**c**) DNA pull-down using protein lysates from HEK293T cells expressing wild-type or mutant ZBTB7A. Western blot analysis shows that A175fs and R402H fail to bind oligonutides with a ZBTB7A-binding site (POK WT). Oligonucleotides with a mutated binding site (POK mut) were used as negative control. Input lanes were loaded with 10% of the protein lysate used for each binding reaction.

Variant allele frequency ranged from 5.4 to 76.2% (cut-off 2%) and 4 of 13 patients (31%) harboured two mutations of *ZBTB7A*. Fourteen of 17 mutations (82%) were validated by Sanger sequencing (Supplementary Fig. 1). Somatic status was confirmed in a total of three patients with available remission samples. Thirty-two additional samples of t(8;21)-positive AML with inadequate sample availability for gene panel sequencing were analysed by Sanger sequencing of exon 2 (encoding amino acids 1–421) resulting in the identification of two *ZBTB7A* mutations (2/32; 6%). This lower mutation frequency might be due to the lower sensitivity of Sanger sequencing and incomplete coverage of the coding exons of *ZBTB7A* (we were not able to reliably amplify exon 3 encoding amino acids 422–584). To evaluate the consequences of truncating *ZBTB7A* mutations on the protein level, we performed western blot analysis for one patient with available material and detected a shorter form of the ZBTB7A protein resulting from the R377X mutation (Supplementary Fig. 2).

Recently, frequent *ASXL2* mutations were identified in t(8;21) AML[13]. In our cohort, *ZBTB7A* and *ASXL2* mutations occurred at similar frequencies (Fig. 1b) and 5 of 13 patients carried mutations in both genes; however, there was no significant association of mutated *ZBTB7A* and mutations in *ASXL2* (Fisher's exact test, $P=0.12$) or any other recurrently mutated gene. Alterations of *ASXL1* were mutually exclusive with genetic lesions of *ZBTB7A* suggesting alternative routes of leukaemogenesis. Similarly, mutations of *ZBTB7A* and *KIT* were exclusive in all, but one patient. In the exome data of 22 patients with inversion inv(16) (another rearrangement disrupting the core-binding factor complex in AML), we found a single *ZBTB7A* mutation (A211V). Of note, we did not find any *ZBTB7A* mutations by exome sequencing of 50 patients with cytogenetically normal AML (CN-AML) or 14 AML patients with chromosomal aberrations other than t(8;21) or inv(16). These results point towards a specific association between *ZBTB7A* alterations and the *RUNX1/RUNX1T1* fusion.

Mutations disrupt the anti-proliferative function of ZBTB7A. To assess the functional consequences of the identified *ZBTB7A* mutations, we performed luciferase reporter gene assays. It is known that ZBTB7A represses the expression of *ARF* (alternate open reading frame of *CDKN2A*)[14]. In contrast to wild-type ZBTB7A, the R402H, R402C, A175fs or R377X mutants failed to repress a luciferase reporter containing ZBTB7A-binding elements derived from the *ARF* promoter (Fig. 3a). Expression of ZBTB7A constructs was confirmed by western blot (Fig. 3b).

In light of recent reports about the negative regulation of glycolysis by ZBTB7A[10], we assessed the expression of glycolytic genes (*SLC2A3*, *PFKP* and *PKM*) in the RNA-sequencing data from our AML t(8;21) patients (Supplementary Fig. 3). In *ZBTB7A*-mutated patients ($n=5$), we found a significantly higher expression of *PFKP* (Student's *t*-test, $P=0.03$) compared with patients without any detectable *ZBTB7A* mutation ($n=11$). On average, *PKM* and *SLC2A3* also showed higher expression levels in patients with *ZBTB7A* mutations, but did not reach statistical significance (Student's *t*-test, $P=0.17$ and $P=0.54$, respectively). In the latter case, the difference in the mean values can be attributed mainly to an outlier in the *ZBTB7A*-mutated group with very high *SLC2A3* expression. Expression levels of *ZBTB7A* were similar in both the patient groups, compatible with inactivation of *ZBTB7A* on the genetic level rather than on the transcriptional level.

The C-terminal part of ZBTB7A is important for nuclear localization[15]. Because some mutations result in loss of the C-terminal zinc-finger domain and nuclear localization signal, we evaluated the cellular localization of mutant ZBTB7A. Whereas wild-type ZBTB7A was detected in the nucleus, immunofluorescence staining of the A175fs and R377X mutants showed an altered cytoplasmic localization (Fig. 3c). In contrast, mutants R402H and R402C exhibited a variable cellular localization with cytoplasmic protein detectable only in a minor subset of cells (Supplementary Fig. 4a,b). Amino-acid substitutions of R402 showed a smaller increase in cytoplasmic protein fraction compared with truncation mutants as analysed by western blot (Supplementary Fig. 4c). Ultimately, the observed effect of mutations on ZBTB7A localization remains to be confirmed in appropriate primary patient material, which was not available in our study.

In the t(8;21) translocation-positive AML cell line Kasumi-1, retroviral expression of wild-type *ZBTB7A* inhibited cell growth, whereas this anti-proliferative effect was not observed upon

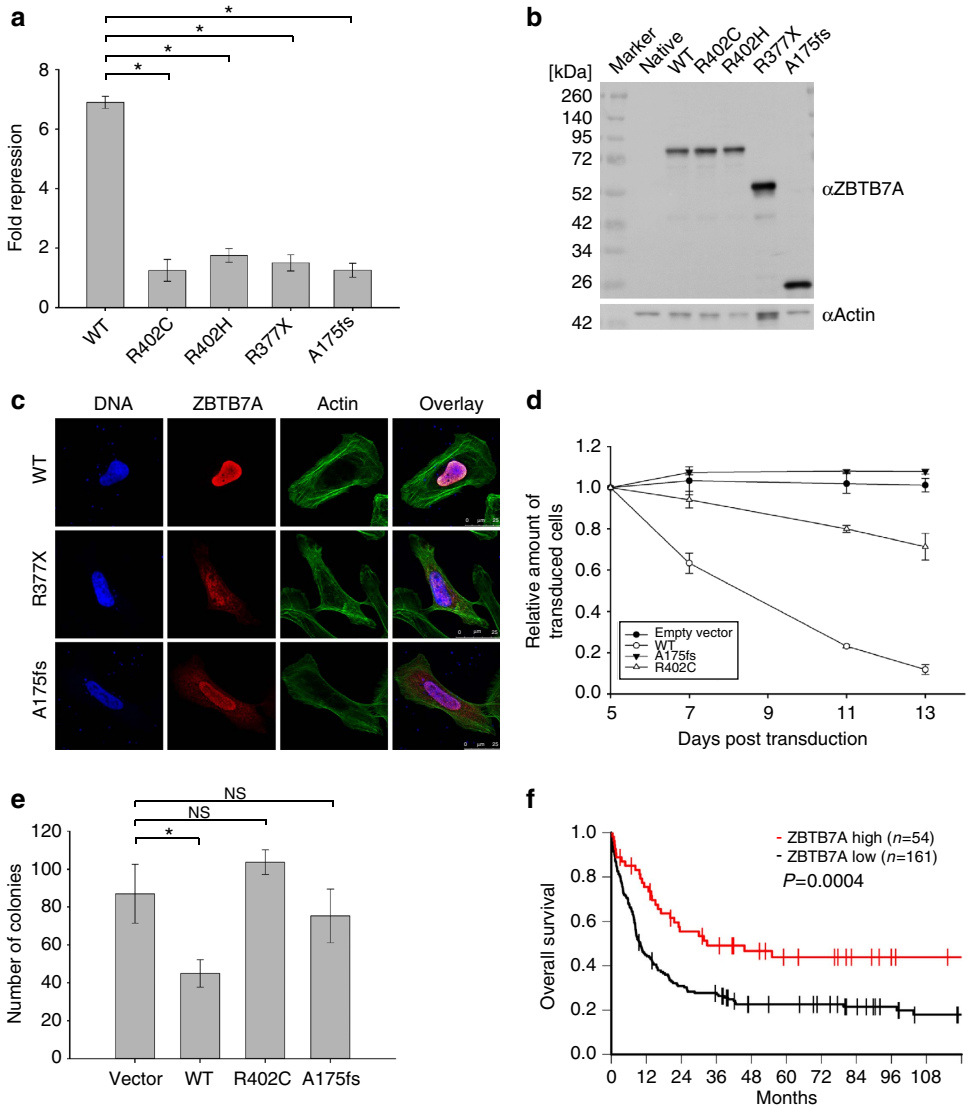


Figure 3 | Functional consequences of *ZBTB7A* mutations and clinical relevance of *ZBTB7A* expression. (**a**) Luciferase assay in transiently transfected HEK293T cells using the pGL2-p19ARF-Luc reporter combined with expression constructs for wild-type and mutant ZBTB7A. (**b**) Western blot of ZBTB7A constructs expressed in HEK293T cells. (**c**) Sub-cellular localization of ZBTB7A wild type, R377X and A175fs in transiently transfected U2OS cells. Scale bar, 25 μm. (**d**) Growth of Kasumi-1 cells stably expressing ZBTB7A wild type or mutants. (**e**) CFC assay of murine bone marrow lineage-negative cells co-expressing RUNX1/RUNX1T1 and wild-type or mutant ZBTB7A. (**f**) Overall survival of patients with CN-AML according to *ZBTB7A* expression (log-rank test, $P=0.0004$). *Two-tailed, unpaired Student's *t*-test, $P<0.05$; NS, not significant. Bar graphs or growth curves represent mean ± s.d. of three independent experiments.

expression of the A175fs *ZBTB7A* mutant (Fig. 3d). The R402C mutant expressing Kasumi-1 cells showed a trend towards reduced cell growth, suggesting residual activity. On the basis of this observation, we expressed *ZBTB7A* wild type or its mutants together with the *RUNX1/RUNX1T1* fusion in lineage-negative murine bone marrow cells and performed colony-forming cell (CFC) assays. *ZBTB7A* expression led to a significant decrease in the number of colonies in primary CFC (87 ± 12.6 versus 45 ± 5.8, Student's *t*-test, $P<0.0001$), while this effect was lost for both mutants tested (Fig. 3e). These findings support an oncogenic cooperativity between *RUNX1/RUNX1T1* and *ZBTB7A* mutations.

Prognostic relevance of *ZBTB7A* expression in CN-AML. The identification of a novel recurrently mutated gene demands the evaluation of its clinical relevance. We did not find a significant difference in overall or relapse-free survival between t(8;21)-positive AML patients with wild-type or mutant *ZBTB7A* (Supplementary Fig. 5). However, this evaluation was limited by the relatively small cohort size. Considering the potential role of *ZBTB7A* as tumour suppressor in AML and its anti-proliferative properties, we correlated *ZBTB7A* expression with clinical outcome in a larger cohort of AML patients (GSE37642). There was no significant difference in *ZBTB7A* expression levels between cytogenetic subgroups of AML (Supplementary Fig. 6). Remarkably, in over 200 CN-AML patients treated on clinical trial (NCT00266136), high expression of *ZBTB7A* was associated with a favourable outcome (Fig. 3f; Supplementary Fig. 7), suggesting a relevance in AML beyond the t(8;21) subgroup. The favourable prognostic impact of high *ZBTB7A* transcript levels was most obvious in elderly patients (age >60 years) and high *ZBTB7A* expression was associated with a 'low molecular risk genotype' (mutated *NPM1* without *FLT3*-ITD; Supplementary Fig. 7; Supplementary Table 5). We validated the association of

high *ZBTB7A* expression with favourable outcome in an independent CN-AML patient cohort[16,17] (Supplementary Fig. 8).

Discussion

In summary, we have identified *ZBTB7A* as one of the most frequently mutated genes in t(8;21)-positive AML. Consistent with our findings, *ZBTB7A* mutations in 3 of 20 (15%) AML t(8;21) patients and 1 of 395 AML inv(16) patients were reported[18] during the revision of the present manuscript. Our functional analyses indicate that *ZBTB7A* mutations result in loss of function, due to alteration or loss of the zinc-finger motives. Beyond DNA binding, the zinc-finger domain of ZBTB7A is also known to interact with TP53 and BCL6 (ref. 9). Thus, multiple pathways might be influenced by alteration or loss of the ZBTB7A zinc-finger domain. The N-terminal missense mutations in the BTB domain may result in failure of co-repressor recruitment. Considering that 4 of 13 of patients had more than one *ZBTB7A* mutation, our finding that overexpression of wild-type *ZBTB7A* leads to reduced proliferation of Kasumi-1 cells and a decreased number of CFCs of murine bone marrow cells, we suggest that *ZBTB7A* acts as a tumour suppressor in t(8;21)-positive AML. Initial studies characterized *ZBTB7A* as proto-oncogene in various tissues[14,19]. For example, Maeda *et al.* demonstrated that transgenic mice with *Zbtb7a* overexpression in the immature T- and B-lymphoid lineage develop precursor T-cell lymphoma/leukaemia[14]. In contrast, it was more recently shown that *ZBTB7A* can also act as a tumour suppressor. Overexpression of *Zbtb*7a in murine prostate epithelium did not result in neoplastic transformation; unexpectedly, *Zbtb7a* inactivation lead to the acceleration of Pten-driven prostate tumorigenesis[20]. Recently, somatic zinc-finger mutations of *ZBTB7A* were found at low frequencies ($<5\%$) in a variety of solid cancers suggesting a common mechanism across tumour entities[21]. In fact, the de-repression of glycolytic genes upon deletion or mutation of *ZBTB7A*[10,21] might underlie the loss of anti-proliferative properties that we observed for ZBTB7A mutants A175fs and R402C in the present study. Any inactivating alteration of *ZBTB7A* will likely increase glycolysis, and, thus, helps the tumour cells to produce more energy. Besides tumour metabolism, it is known that ZBTB7A also plays an important role in haematopoietic lineage fate decisions[9]. During lymphopoiesis *ZBTB7A* regulates B-cell development[22], whereas in the myeloid lineage it is essential for erythroid differentiation[23]. Thus, *ZBTB7A* mutations may contribute to the block of differentiation in AML t(8;21).

The favourable prognostic relevance of high *ZBTB7A* expression in CN-AML, which accounts for half of all AML patients, may point towards a more general tumour suppressor role of *ZBTB7A* in myeloid leukaemia. In particular, the anti-proliferative properties of ZBTB7A may slow down disease progression. High *ZBTB7A* expression as a favourable prognostic marker has been reported also in colorectal cancer[10], consistent with a clinicobiological role of *ZBTB7A* across malignancies of multiple tissue origins. Given that somatic mutations of *ZBTB7A* seem to be absent or rare in CN-AML, other mechanisms, including epigenetic changes or alterations of upstream regulators, may lead to inactivation or downregulation of *ZBTB7A*.

Our discovery of frequent *ZBTB7A* mutations in AML with t(8;21) translocation, one of the most common translocations in AML and the first balanced translocation identified in leukaemia[1], demonstrates that the mutational landscape of AML is still not fully understood. Further studies will be required to unravel the mechanism underlying leukaemogenic cooperativity between mutated *ZBTB7A* and the *RUNX1/RUNX1T1* fusion gene.

Methods

Patients. AML samples were collected within the German Cancer Consortium (DKTK) at the partner sites Munich and Dresden. Patients were treated according to the protocols of Acute Myeloid Leukemia Cooperative Group (AMLCG) or Study Alliance Leukemia (SAL) multicentre clinical trials. Study protocols were approved by the Institutional Review Boards of the participating centres. Informed consent was received in accordance with the Declaration of Helsinki.

Sequencing. Exome sequencing (mean coverage: 87x; range 80–90x) was performed on a HiSeq 2000 Instrument (Illumina), using the SureSelect Human All Exon V5 kit (Agilent). Pretreatment blood or bone marrow specimens from 56 AML patients with t(8;21) translocation were sequenced using Haloplex custom amplicons (Agilent) and a HiSeq 1500 instrument (Illumina). Target sequence included the entire open-reading frame of *ZBTB7A* in addition to 45 leukaemia-related genes or mutational hotspots (Supplementary Table 3). Variant calling was performed as described previously[24]. Sanger sequencing of PCR-amplified genomic DNA was carried out using a 3500xL Genetic Analyzer (Applied Biosystems). Primer sequences are provided in Supplementary Table 2. Sequencing of messenger RNA was performed using the TruSeq RNA Sample Preparation protocol, followed by sequencing on a HiSeq 2000 Instrument (Illumina). RNA sequence reads were aligned to the human genome (hg19) using STAR[25] (version 2.4.1b). Reads per gene were counted using HTseq[26] (version 0.6.1) with intersection-strict mode and normalized for the total number of reads per sample.

Structural modelling. Suitable templates for the modelling were searched with HHPRED[27], using the zinc-finger domain of ZBTB7A as input sequence. The highest scoring homologue, for which a structure of a DNA complex is available, was the Wilms tumour suppressor protein[28] (PDB accession code 2J9P, *E*-value 4.8E–29, *P*-value 1.3E–30). The model for ZBTB7A was generated on the basis of 2J9P using MODELLER[29]. Importantly, 2J9P also contains an arginine at the equivalent position of ZBTB7A's R402, allowing us to model the function of R402 as major groove binder with confidence.

Plasmids. The pcDNA3.1-His-ZBTB7A expression construct was a gift from Takahiro Maeda (Boston). ZBTB7A A175fs, R377X, R402C and R402H mutant plasmids were generated using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent) and confirmed by Sanger sequencing. *ZBTB7A* wild type and mutants were subcloned into pMSCV-IRES-YFP (pMIY), using the In-Fusion HD cloning kit (Clontech) and EcoRI restriction sites. The pMSCV-IRES-GFP(pMIG)-RUNX1/RUNX1T1 plasmid was provided by Christian Buske (Ulm).

DNA pull-down. HEK293T cells (DSMZ no.: ACC 635) were transfected with pcDNA3.1 His-Xpress-ZBTB7A (wild type or mutant). After 24 h, protein was extracted using lysis buffer (50 mM Tris HCl, pH 8.5, 150 mM NaCl, 1% Triton X-100, cOmplete Protease Inhibitor Cocktail). For each reaction, 20 µl protein lysate was incubated in binding buffer (PBS supplemented with 150 mM NaCl resulting in a total salt concentration of nearly 300 mM, 0.1% NP40, 1 mM ETDA) with 10 pM biotinylated double-stranded oligonucleotides that contain either the ZBTB7A consensus binding motif (POK WT; 5′-GGTTAAAAGACCCCTCCCCG AATTCGGATC-3′) or a mutant thereof (POK mut; 5′-GGTTAAAATTTTTCTCC CCGAATTCGGATC-3′). After 1 h of incubation at 4 °C, 10 µl streptavidin agarose beads (Sigma Aldrich) was added to each reaction and incubated for 30 min at 4 °C. Beads were washed three times with binding buffer and resupended in 10 µl Laemmli buffer for subsequent western blot analysis. ZBTB7A protein was detected using an antibody against the Xpress tag (1:5,000 dilution, clone R910-25; Life Technologies) and secondary goat anti-mouse IgG-HRP (1:10,000 dilution, clone sc-2060; Santa Cruz). The uncropped western blot scan underlying Fig. 2c is shown in Supplementary Fig. 9.

Reporter gene assay. HEK293T cells (DSMZ no.: ACC 635) were co-transfected with pcDNA3.1-His-ZBTB7A (wild type or mutant), pGL2-p19ARF-Luc (gift from Takahiro Maeda, Boston) as well as pRL-CMV (Renilla luciferase; Promega) using Lipofectamine 2000 (ThermoFischer). After 24 h, cells were lysed; Firefly and Renilla luciferase activity was measured with the dual-luciferase reporter assay system (Promega) according to the manufacturer's instructions. Three independent experiments were each performed in triplicates.

Western blot. HEK293T cells (DSMZ no.: ACC 635) were transfected using Lipofectamine 2000 (ThermoFischer) with pcDNA3.1-His-ZBTB7A (wild type or mutant). After 24 h, protein was either extracted by multiple freeze–thaw cycles in lysis buffer (600 mM KCl, 20 mM Tris-Cl pH 7.8, 20% Glycerol, cOmplete Protease Inhibitor Cocktail) or using the Qproteome Nuclear Protein Kit (Qiagen) for the analysis of nuclear and cytoplasmic protein fractions. From archived patient bone marrow samples, protein was isolated using the AllPrep DNA/RNA/Protein Mini Kit (Qiagen) according to the manufacturer's instructions. Following SDS–polyacrylamide gel electrophoresis and protein transfer to polyvinylidene difluoride membrane (Hybond PTM, Amersham Pharmacia biotech),

immunoblots were blocked with 5% nonfat dried milk, probed with anti-human Pokemon (ZBTB7A) purified antibody (1:5,000 dilution, clone: 13E9; eBioscience) and secondary anti-Armenian hamster IgG-HRP (1:10,000 dilution, clone: sc-2443; Santa Cruz). As loading control immunoblots were incubated with rabbit anti-actin (1:5,000 dilution, clone: sc-1616- R; Santa Cruz) and secondary goat anti-rabbit IgG-HRP (1:10,000 dilution, clone: sc-2030; Santa Cruz). For analysis of the nuclear and cytoplasmic ZBTB7A protein fractions, we used mouse anti-Xpress tag (1:5,000 dilution, clone R910-25, Life Technologies) and secondary goat anti-mouse IgG-HRP (1:10,000 dilution, clone: sc-2060; Santa Cruz). Mouse anti-GAPDH (1:10,000 dilution, clone: sc-32233; Santa Cruz) served as loading control for the cytoplasmic protein fraction. Proteins were detected with enhanced chemiluminescence (ECL, Amersham, GE Healthcare).

Immunofluorescence staining. U2OS human osteosarcoma cells (ATTC no.: HTB-96) were grown on coverslips and transiently transfected with pcDNA3.1-His-ZBTB7A wild type and mutant constructs using PoliFect (Qiagen) according to the manufacturer's guidelines. Cells were fixed 48 h post transfection using PBS 2% formaldehyde (37% stock solution; Merck Schuchardt) for 10 min, permeabilized with PBS 0.5% Triton X-100 (Carl Roth) for 10 min and blocked for 1 h with PBS 2% bovine serum albumin (Albumin Fraction V, AppliChem). Cells were then incubated with polyclonal rabbit His-probe (H-15) antibody (1:50 dilution; Santa Cruz) for 1 h. After extensive washing with PBS 0.1% Tween 20 (Carl Roth), secondary antibody incubation was performed for 1 h with goat anti-rabbit IgG (H + L), F(ab′)2 fragment Alexa Fluor 594 conjugate (1:500 dilution; Cell Signaling Technology). Counterstaining was performed using NucBlue Reagent and ActinGreen 488 ReadyProbes Reagent (Life Technologies; 2 drops per ml) at room temperature for 20 min. Coverslips were mounted using fluorescence mounting medium (DAKO). Specimens were analysed using a confocal fluorescence laser scanning system (TCS SP5 II; Leica). For image acquisition and processing, the LAS AF Lite Software (Leica) was used.

Retroviral transduction. Retroviral transduction of Kasumi-1 cells (DSMZ no.: ACC 220) was accomplished as outlined previously[30]. In brief, HEK293T cells were co-transfected with pMSCV-IRES-YFP (pMIY) vectors containing either wild-type or mutant (A175fs, R402C) ZBTB7A and packaging plasmids. After 48 h, the cell culture supernatant was collected, sterile filtered and used for viral loading of RetroNectin (Takara Clontech)-coated plates. A total of 3×10^5 Kasumi-1 cells were transduced per well. The percentage of YFP-positive cells was assessed on a FACSCalibur flow cytometer (BD Biosciences). Three independent experiments were each performed in duplicates.

Colony-forming cell assay. For *in vitro* CFC assays, bone marrow cells were collected from the femur and pelvic girdle of wild-type mice (C57BL/6X129/J). Lineage-negative haematopoietic progenitors were isolated using magnetic separation (MACS, murine lineage depletion kit, Miltenyi biotech). Retrovirally transduced cells were sorted for GFP/YFP and were plated in 1% myeloid-conditioned methylcellulose containing Iscove's modified Dulbecco medium-based Methocult (Methocult M3434; StemCell Technologies) at a concentration of 500 cells per ml. Single-cell suspensions of colonies were serially replated at the same concentration until the exhaustion of cell growth. Three independent experiments were each performed in duplicates.

Analysis of clinical and gene expression data. Clinical relevance of *ZBTB7A* mutations or expression levels was evaluated using the Kaplan–Meier method and the log-rank test. Fisher's exact test was used to compare categorical variables, while Wilcoxon Mann–Whitney *U*-test was applied for continuous variables. All patients included in this analysis were treated intensively with curative intent according to the AMLCG protocols. Gene expression profiling was performed on 215 adult patients with cytogenetically normal AML, using Affymetrix Human Genome (HG) U133A/B ($n = 155$) and HG U133Plus2.0 microarrays ($n = 60$). The RMA method was used for data normalization, and probe set summarization utilized custom chip definition files based on the GeneAnnot database (version 2.2.0). Probe set GC19M004001_at was used to determine *ZBTB7A* expression levels. High *ZBTB7A* expression was defined as the highest (4th) quartile of expression values observed in CN-AML patients. Patients with *ZBTB7A* expression levels in the 1st to 3rd quartile were classified as having low expression. The patients analysed here represent a subset of the previously published data set GSE37642. Validation of the results was done using data sets from the Haemato Oncology Foundation for Adults in the Netherlands (HOVON) study group (GSE14468 and GSE1159)[16,17].

Data availability. Data referenced in this study are available in the Gene Expression Omnibus database with the accession codes GSE37642, GSE14468 and GSE1159. The next-generation sequencing data that support the findings of this study are available on request from the corresponding author (P.A.G). The data are not publicly available due to them containing information that could compromise research participant privacy or consent. Explicit consent to deposit raw-sequencing data was not obtained from the patients, many samples were collected >10 years ago. Thus, the vast majority of patients cannot be asked to provide their consent for deposit of their comprehensive genetic data.

References

1. Rowley, J. D. Identificaton of a translocation with quinacrine fluorescence in a patient with acute leukemia. *Ann. Genet.* **16,** 109–112 (1973).
2. Erickson, P. *et al.* Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. *Blood* **80,** 1825–1831 (1992).
3. Vardiman, J. W. *et al.* The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood* **114,** 937–951 (2009).
4. Rhoades, K. L. *et al.* Analysis of the role of AML1-ETO in leukemogenesis, using an inducible transgenic mouse model. *Blood* **96,** 2108–2115 (2000).
5. Schessl, C. *et al.* The AML1-ETO fusion gene and the FLT3 length mutation collaborate in inducing acute leukemia in mice. *J. Clin. Invest.* **115,** 2159–2168 (2005).
6. Schwieger, M. *et al.* AML1-ETO inhibits maturation of multiple lymphohematopoietic lineages and induces myeloblast transformation in synergy with ICSBP deficiency. *J. Exp. Med.* **196,** 1227–1240 (2002).
7. Yuan, Y. *et al.* AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations. *Proc. Natl Acad. Sci. USA* **98,** 10398–10403 (2001).
8. Higuchi, M. *et al.* Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. *Cancer Cell* **1,** 63–74 (2002).
9. Lunardi, A., Guarnerio, J., Wang, G., Maeda, T. & Pandolfi, P. P. Role of LRF/Pokemon in lineage fate decisions. *Blood* **121,** 2845–2853 (2013).
10. Liu, X. S. *et al.* ZBTB7A acts as a tumor suppressor through the transcriptional repression of glycolysis. *Genes Dev.* **28,** 1917–1928 (2014).
11. Costoya, J. A. Functional analysis of the role of POK transcriptional repressors. *Brief Funct. Genomic Proteomic* **6,** 8–18 (2007).
12. Morrison, D. J. *et al.* FBI-1, a factor that binds to the HIV-1 inducer of short transcripts (IST), is a POZ domain protein. *Nucleic Acids Res.* **27,** 1251–1262 (1999).
13. Micol, J. B. *et al.* Frequent ASXL2 mutations in acute myeloid leukemia patients with t(8;21)/RUNX1-RUNX1T1 chromosomal translocations. *Blood* **124,** 1445–1449 (2014).
14. Maeda, T. *et al.* Role of the proto-oncogene Pokemon in cellular transformation and ARF repression. *Nature* **433,** 278–285 (2005).
15. Pendergrast, P. S., Wang, C., Hernandez, N. & Huang, S. FBI-1 can stimulate HIV-1 Tat activity and is targeted to a novel subnuclear domain that includes the Tat-P-TEFb-containing nuclear speckles. *Mol. Biol. Cell* **13,** 915–929 (2002).
16. Valk, P. J. *et al.* Prognostically useful gene-expression profiles in acute myeloid leukemia. *N. Engl. J. Med.* **350,** 1617–1628 (2004).
17. Wouters, B. J. *et al.* Double CEBPA mutations, but not single CEBPA mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. *Blood* **113,** 3088–3091 (2009).
18. Lavallée, V.-P. *et al.* RNA-sequencing analysis of core binding factor AML identifies recurrent ZBTB7A mutations and defines RUNX1-CBFA2T3 fusion signature. *Blood*, pii: blood-2016-03-703868 (2016).
19. Jeon, B. N. *et al.* Proto-oncogene FBI-1 (Pokemon/ZBTB7A) represses transcription of the tumor suppressor Rb gene via binding competition with Sp1 and recruitment of co-repressors. *J. Biol. Chem.* **283,** 33199–33210 (2008).
20. Wang, G. *et al.* Zbtb7a suppresses prostate cancer through repression of a Sox9-dependent pathway for cellular senescence bypass and tumor invasion. *Nat. Genet.* **45,** 739–746 (2013).
21. Liu, X. S. *et al.* Somatic human ZBTB7A zinc finger mutations promote cancer progression. *Oncogene,* doi:10.1038/onc.2015.371 (2015).
22. Maeda, T. *et al.* Regulation of B versus T lymphoid lineage fate decision by the proto-oncogene LRF. *Science* **316,** 860–866 (2007).
23. Maeda, T. *et al.* LRF is an essential downstream target of GATA1 in erythroid development and regulates BIM-dependent apoptosis. *Dev. Cell* **17,** 527–540 (2009).
24. Herold, T. *et al.* Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis. *Blood* **124,** 1304–1311 (2014).
25. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).
26. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).
27. Hildebrand, A., Remmert, M., Biegert, A. & Soding, J. Fast and accurate automatic structure prediction with HHpred. *Proteins* **77**(Suppl 9): 128–132 (2009).
28. Stoll, R. *et al.* Structure of the Wilms tumor suppressor protein zinc finger domain bound to DNA. *J. Mol. Biol.* **372,** 1227–1245 (2007).
29. Webb, B. & Sali, A. Protein structure modeling with MODELLER. *Methods Mol. Biol.* **1137,** 1–15 (2014).

30. Wichmann, C. *et al.* Activating c-KIT mutations confer oncogenic cooperativity and rescue RUNX1/ETO-induced DNA damage and apoptosis in human primary CD34+ hematopoietic progenitors. *Leukaemia* **29,** 279–289 (2015).
31. Liu, W. *et al.* IBS: an illustrator for the presentation and visualization of biological sequences. *Bioinformatics* **31,** 3359–3361 (2015).

Acknowledgements

We thank all participants and recruiting centres of the AMLCG and SAL trials. This work was supported by a grant to P.A.G. and H.B. from the Wilhelm-Sander-Stiftung (2014.162.1). K.H.M, K.S., P.A.G., W.H., K.-P.H. and H.B. acknowledge support from the German Research Council (DFG) within the Collaborative Research Centre (SFB) 1243 'Cancer Evolution' (projects A06, A07, A08, A10 and Z02). We thank Christina Schreck and Robert A.J. Oostendorp for providing murine bone marrow.

Author contributions

L.H. and. P.A.G. conceived and designed the experiments. L.H., S.D., S.O., G.L., K.R., K.B., C.P., L.C.-W. and S.K. performed the experiments. L.H., S.O., K.R., T.H., S.A.B., K.H.M. and S.V. analysed the data. S.V. and A.G. provided the bioinformatics support. H.B. and S.W. managed the sequencing platforms. K.B., E.Z., N.P.K., S.S., J.B., S.K.B., K.S., J.M.M., F.S. and C.T. characterized the patient samples. M.C.S., J.B., W.E.B., T.B., B.J.W. and W.H. coordinated the AMLCG clinical trials. K.-P.H. performed the structural modelling. P.A.G., C.W. and K.S. supervised the project. L.H. and P.A.G. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Hartmann, L. *et al.* ZBTB7A mutations in acute myeloid leukaemia with t(8;21) translocation. *Nat. Commun.* 7:11733 doi: 10.1038/ncomms11733 (2016).

Supplementary Information

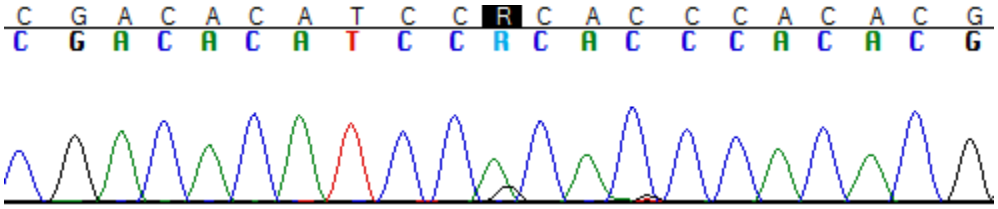
ZBTB7A Mutations in Acute Myeloid Leukemia with t(8;21) Translocation

Luise Hartmann[1,2,3,4], Sayantanee Dutta[1,2,3,4], Sabrina Opatz[1,2,3,4], Sebastian Vosberg[1,2,3,4], Katrin Reiter[1,2,3,4], Georg Leubolt[1,2,3,4], Klaus H. Metzeler[1,2,3,4], Tobias Herold[1,2,3,4], Stefanos A. Bamopoulos[1], Kathrin Bräundl[1,2,3,4], Evelyn Zellmeier[1], Bianka Ksienzyk[1], Nikola P. Konstandin[1], Stephanie Schneider[1], Karl-Peter Hopfner[5], Alexander Graf[6], Stefan Krebs[6], Helmut Blum,[6] Jan Moritz Middeke[3,4,7], Friedrich Stölzel[3,4,7], Christian Thiede[3,4,7], Stephan Wolf[4], Stefan K. Bohlander[8], Caroline Preiss[9], Linping Chen-Wichmann[9], Christian Wichmann[9], Maria Cristina Sauerland[10], Thomas Büchner[11], Wolfgang E. Berdel[11], Bernhard J. Wörmann[12], Jan Braess[13], Wolfgang Hiddemann[1,2,3,4], Karsten Spiekermann[1,2,3,4] and Philipp A. Greif[1,2,3,4]

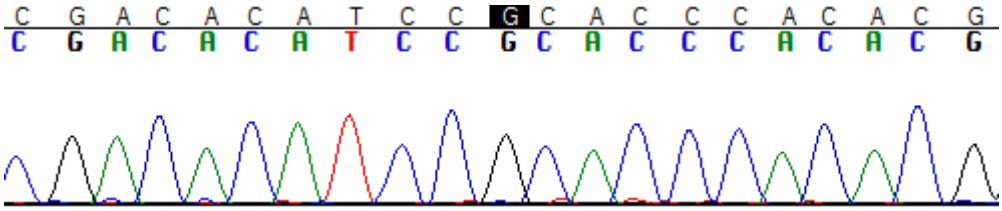
(1) Department of Internal Medicine 3, University Hospital, Ludwig-Maximilians-Universität (LMU) München, München, Germany
(2) Clinical Cooperative Group Leukemia, Helmholtz Zentrum München, German Research Center for Environmental Health, München, Germany
(3) German Cancer Consortium (DKTK), Heidelberg, Germany
(4) German Cancer Research Center (DKFZ), Heidelberg, Germany
(5) Department of Biochemistry, and (6) Laboratory for Functional Genome Analysis (LAFUGA), at the Gene Center, Ludwig-Maximilians-Universität (LMU) München, München, Germany
(7) Medizinische Klinik und Poliklinik I, Universitätsklinikum Dresden, Dresden, Germany
(8) Department of Molecular Medicine and Pathology, The University of Auckland, Auckland, New Zealand
(9) Department of Transfusion Medicine, Cell Therapeutics and Hemostasis, University Hospital Grosshadern, Ludwig-Maximilians-Universität (LMU) München, München, Germany
(10) Institute of Biostatistics and Clinical Research, and (11) Department of Medicine A - Hematology, Oncology and Pneumology, University of Münster, Münster, Germany
(12) German Society of Hematology and Oncology, Berlin, Germany
(13) Oncology and Hematology, St. John-of-God Hospital, Regensburg, Germany

Correspondence:
Philipp A. Greif, MD
Marchioninistr. 15
81377 München
Phone: +49 89 4400 43982
FAX: +49 89 4400 43970
pgreif@med.uni-muenchen.de
p.greif@dkfz-heidelberg.de

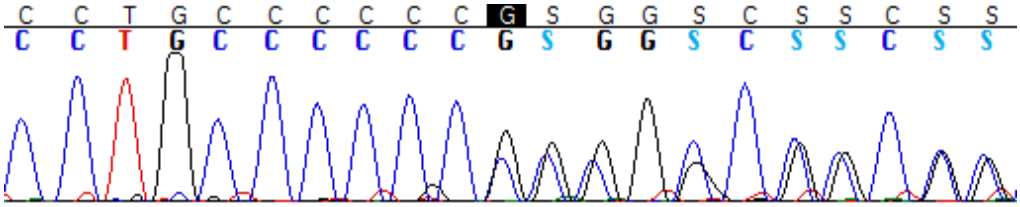
UPN#1-Diagnosis: NM_015898:exon2: c.1205G>A:p.R402H



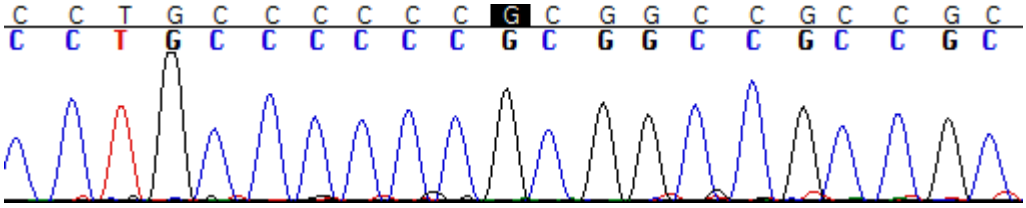
UPN#1-Remission: Wild-type



UPN#2-Diagnosis: NM_015898:exon2: c.522dupC:p.A175fs

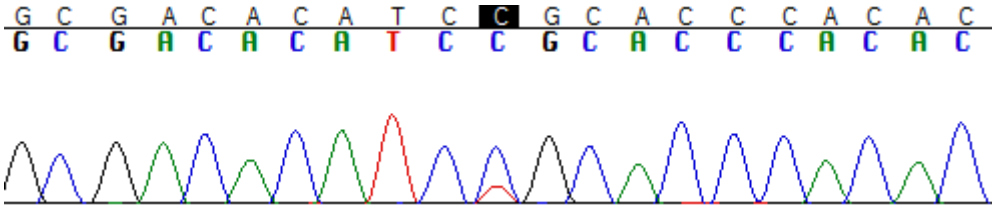


UPN#2-Remission: Wild-type

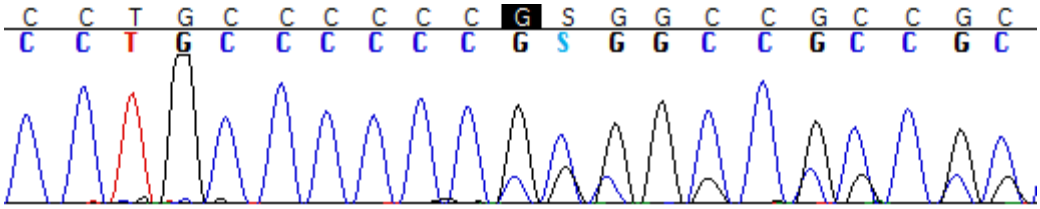


Supplementary Figure 1. Sanger sequencing confirms *ZBTB7A* mutations.

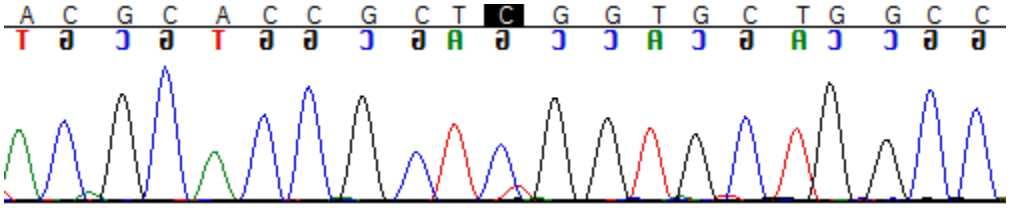
UPN#3-Diagnosis: NM_015898:exon2: c.1204C>T:p.R402C



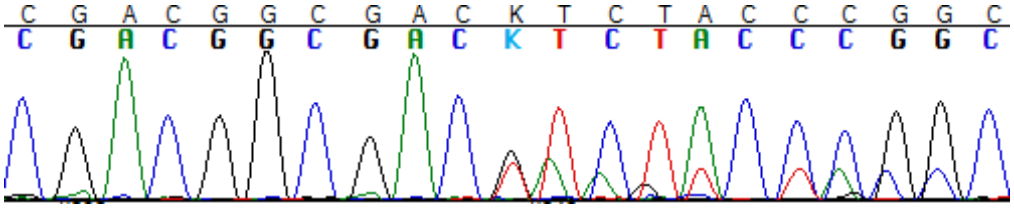
UPN#3-Diagnosis: NM_015898:exon2: c.522dupC:p.A175fs



UPN#4-Diagnosis: NM_015898:exon2: c.149C>T:p.S50L

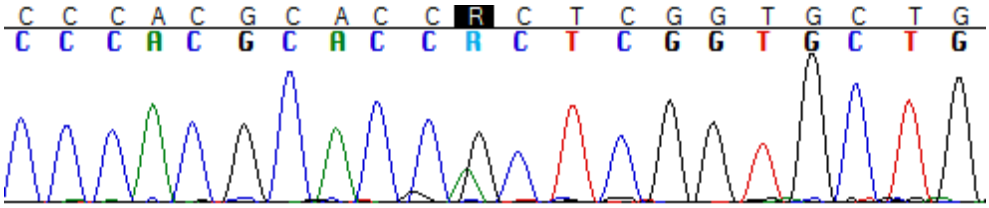


UPN#5-Diagnosis: NM_015898:exon2: c.1089_1090insTAA: p.V364delinsX

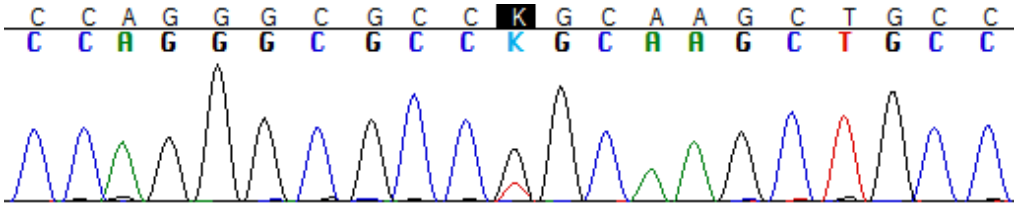


Supplementary Figure 1 (continued). Sanger sequencing confirms *ZBTB7A* mutations.

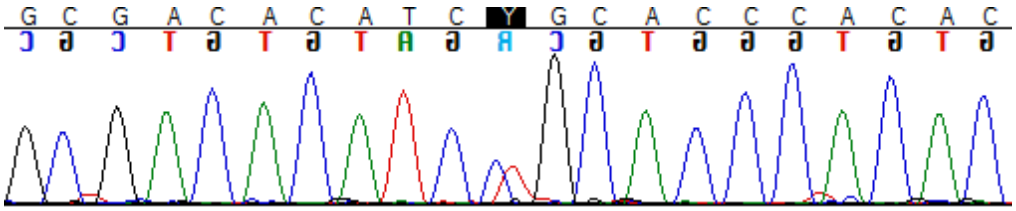
UPN#5-Diagnosis: NM_015898:exon2: c.146G>A:p.R49H



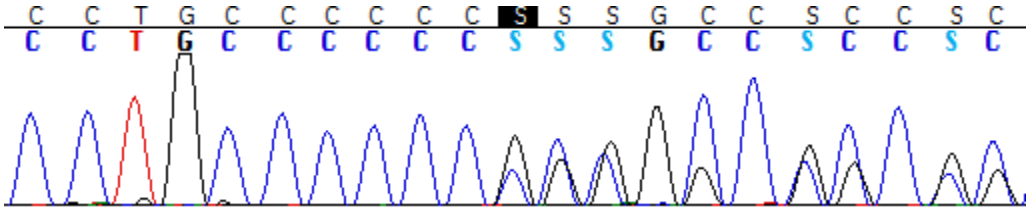
UPN#6-Diagnosis: NM_015898:exon2: c.1183G>T:p.G395C



UPN#7-Diagnosis: NM_015898:exon2: c.1204C>T:p.R402C

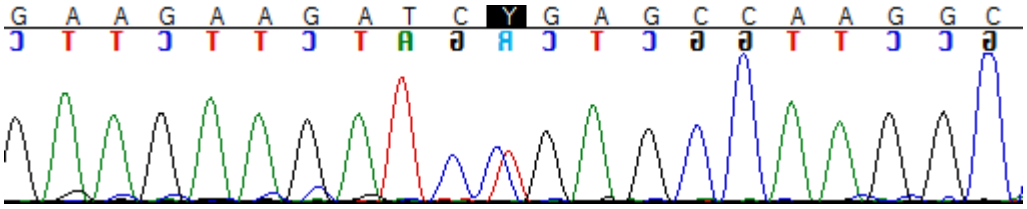


UPN#8-Diagnosis: NM_015898:exon2: c.522dupC:p.A175fs

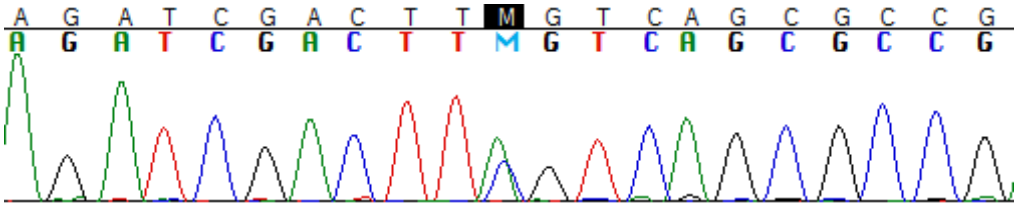


Supplementary Figure 1 (continued). Sanger sequencing confirms *ZBTB7A* mutations.

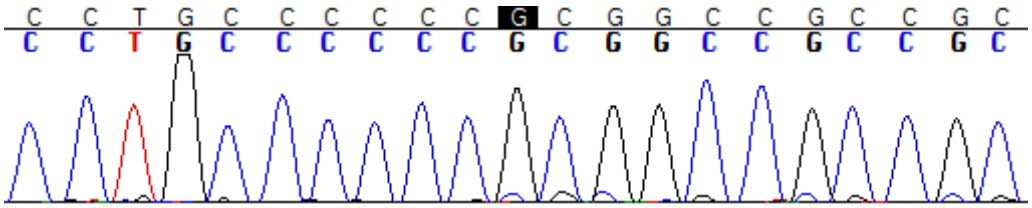
UPN#9-Diagnosis: NM_015898:exon2: c.1129C>T:p.R377X



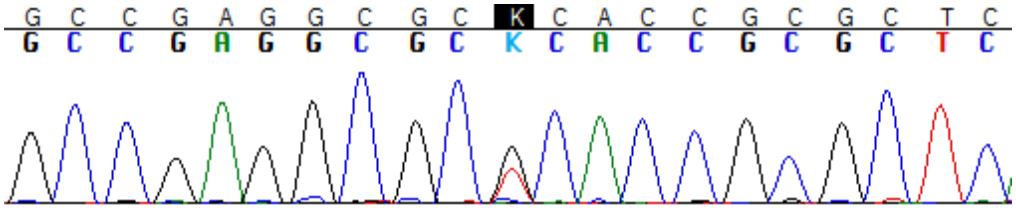
UPN#11-Diagnosis: NM_015898:exon2: c.237C>A:p.F79L



UPN#12-Diagnosis: NM_015898:exon2: c.522dupC:p.A175fs

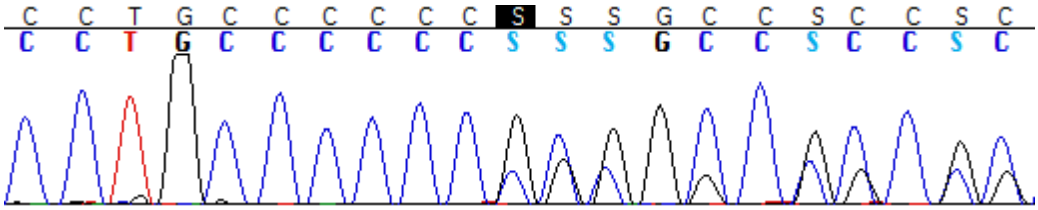


UPN#13-Diagnosis: NM_015898:exon2: c.254T>G:p.L85R

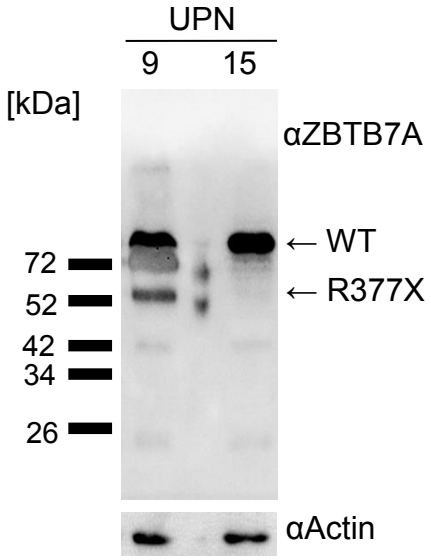


Supplementary Figure 1 (continued). Sanger sequencing confirms *ZBTB7A* mutations.

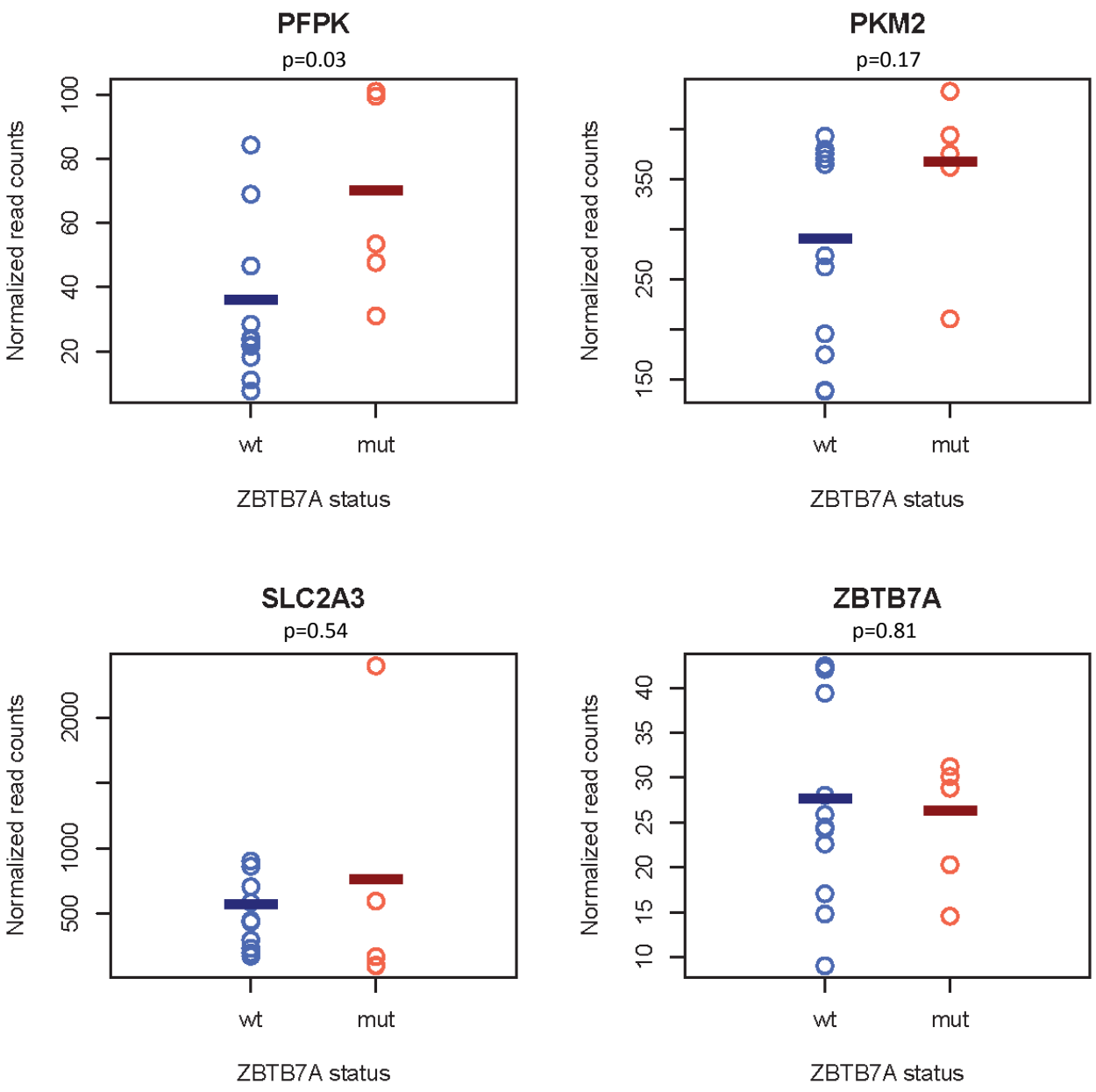
UPN#14-Diagnosis: NM_015898:exon2: c.522dupC:p.A175fs



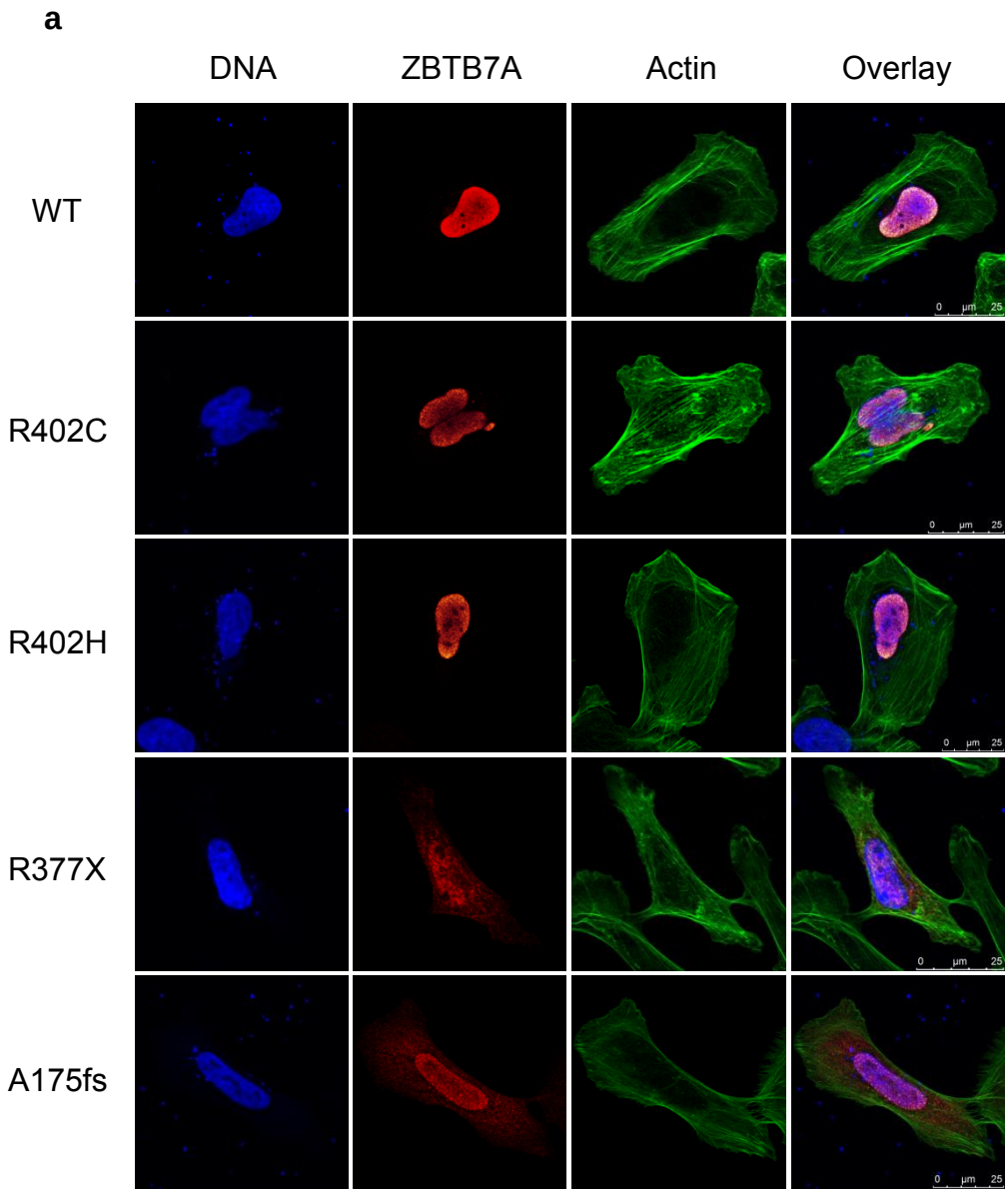
Supplementary Figure 1 (continued). Sanger sequencing confirms *ZBTB7A* mutations.



Supplementary Figure 2. Western blot analysis of a patient with the truncating R377X mutation (UPN9) and another patient with wild-type ZBTB7A (UPN15).

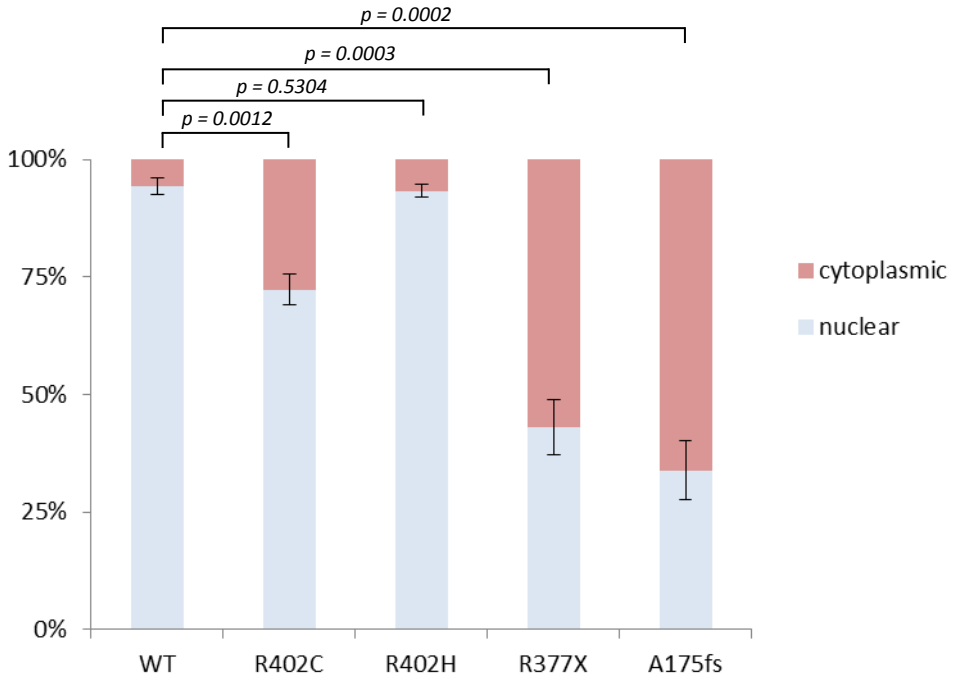


Supplementary Figure 3. Expression of glycolytic genes and *ZBTB7A* in AML t(8;21) patients according to *ZBTB7A* mutation status. Circles indicate mRNA sequence read counts from individual patients. Horizontal bars show mean values of the two patient groups (mutated n=5; wild-type n=11). Differences between groups were assessed using a two-tailed unpaired Student's t-test.



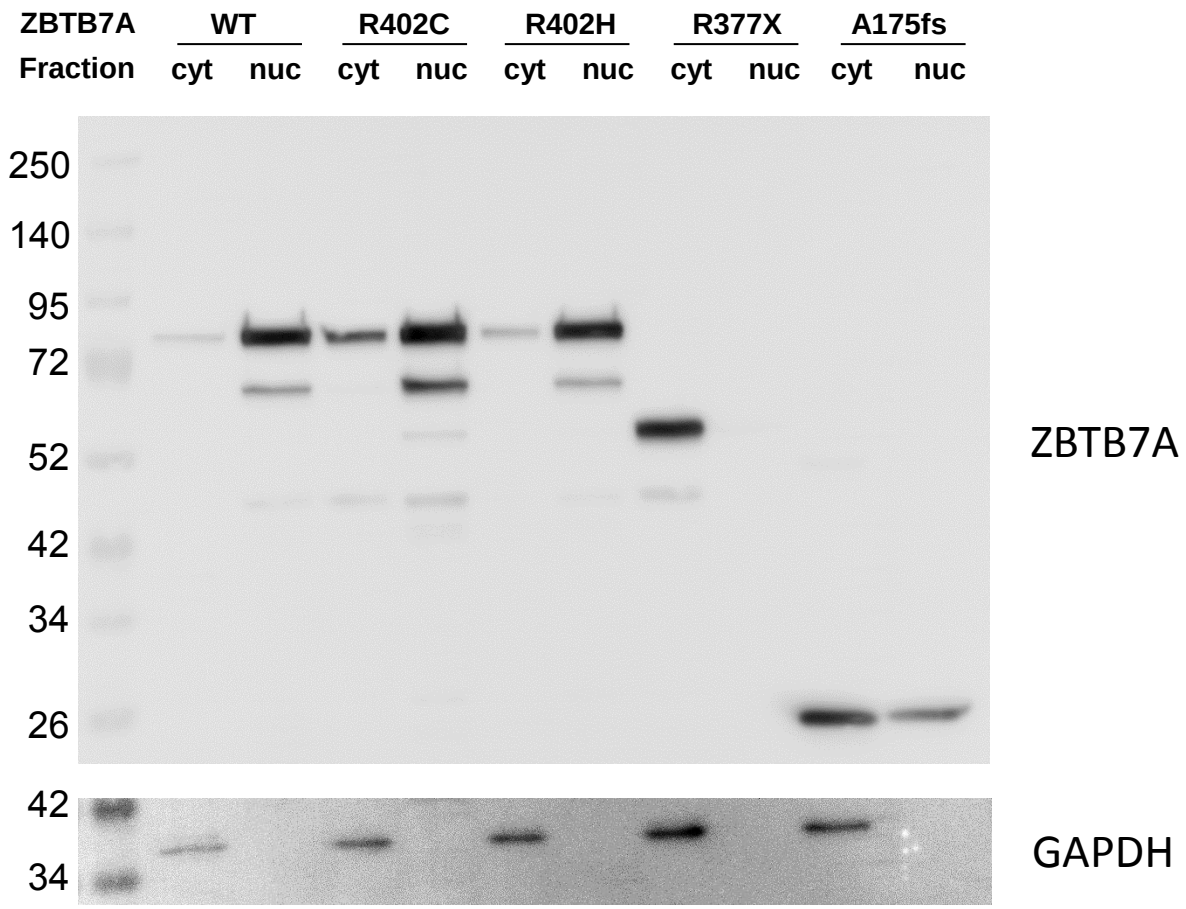
Supplementary Figure 4. Subcellular localization of ZBTB7A wild-type and mutants.
(a) Representative confocal laser scans of transiently transfected U2OS cells show the predominant protein distribution observed for each construct. Scale bar corresponds to 25 μm.

b

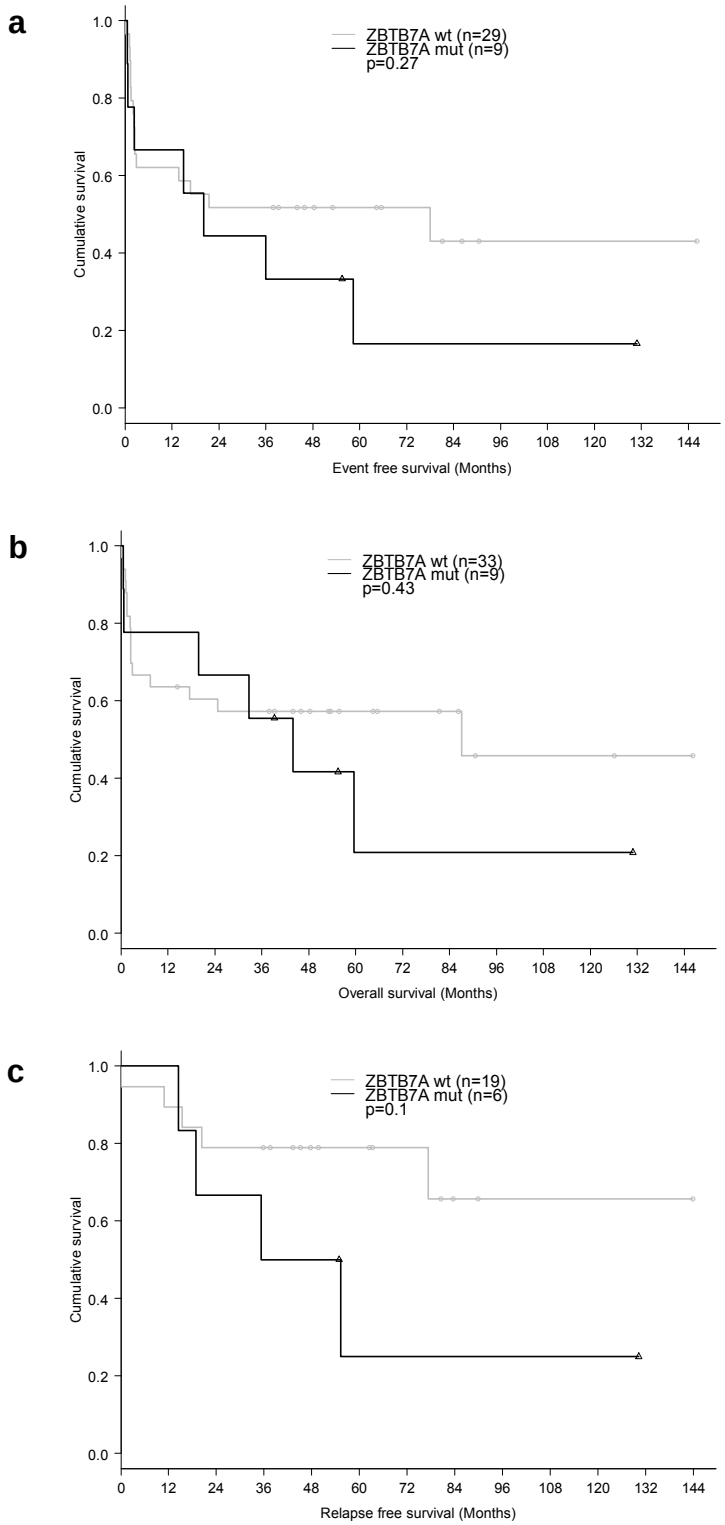


Supplementary Figure 4 (continued). (b) Cell counts after immunofluorescence staining of ZBTB7A wild-type and mutants in transiently transfected U2OS cells. Bar graph shows mean values ± standard deviation of 3 independent experiments with evaluation of 124 cells per construct (representing the minimum number of cells available for evaluation in each experiment). Statistical difference was assessed using a two-tailed unpaired Student's t-test. Nuclear localization was defined as detection of ZBTB7A exclusively in the cell nucleus, whereas cytoplasmic localization indicates ZBTB7A protein detected both in the nucleus and the cytoplasm.

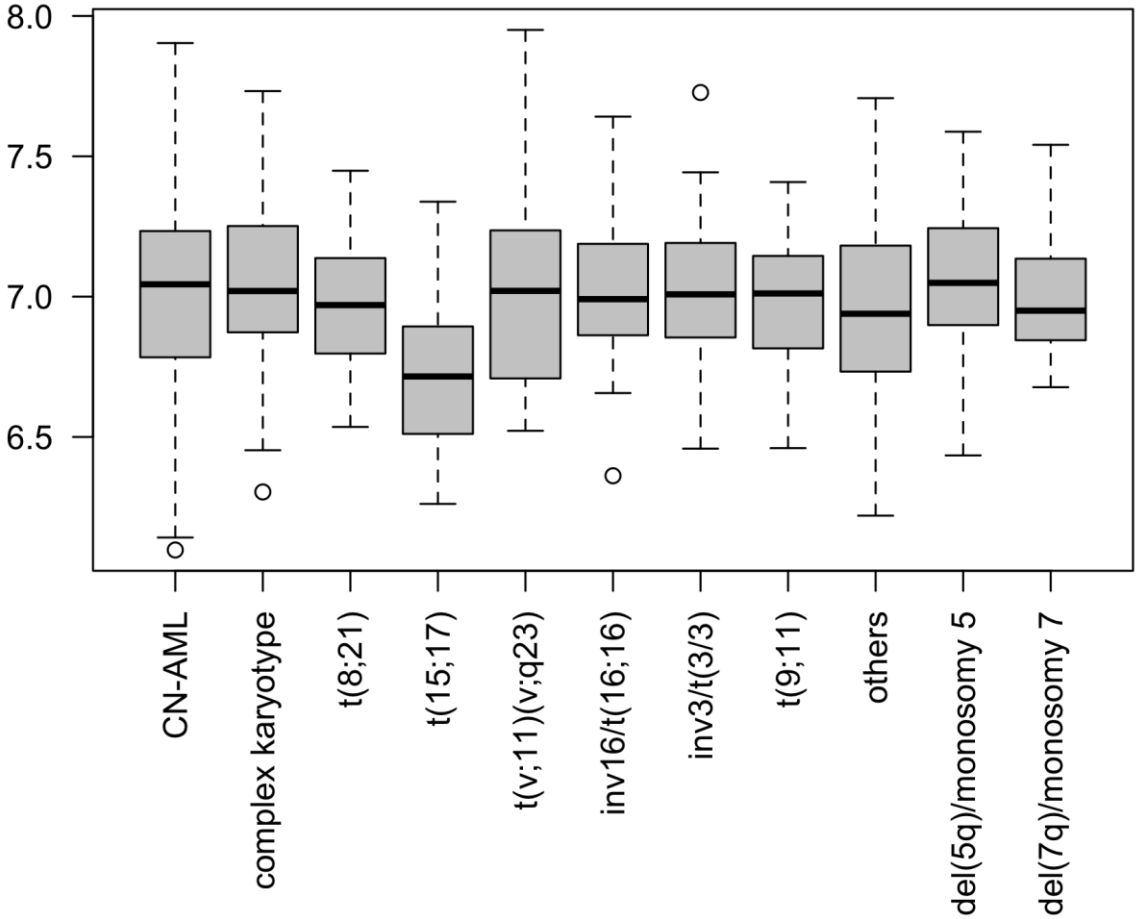
c



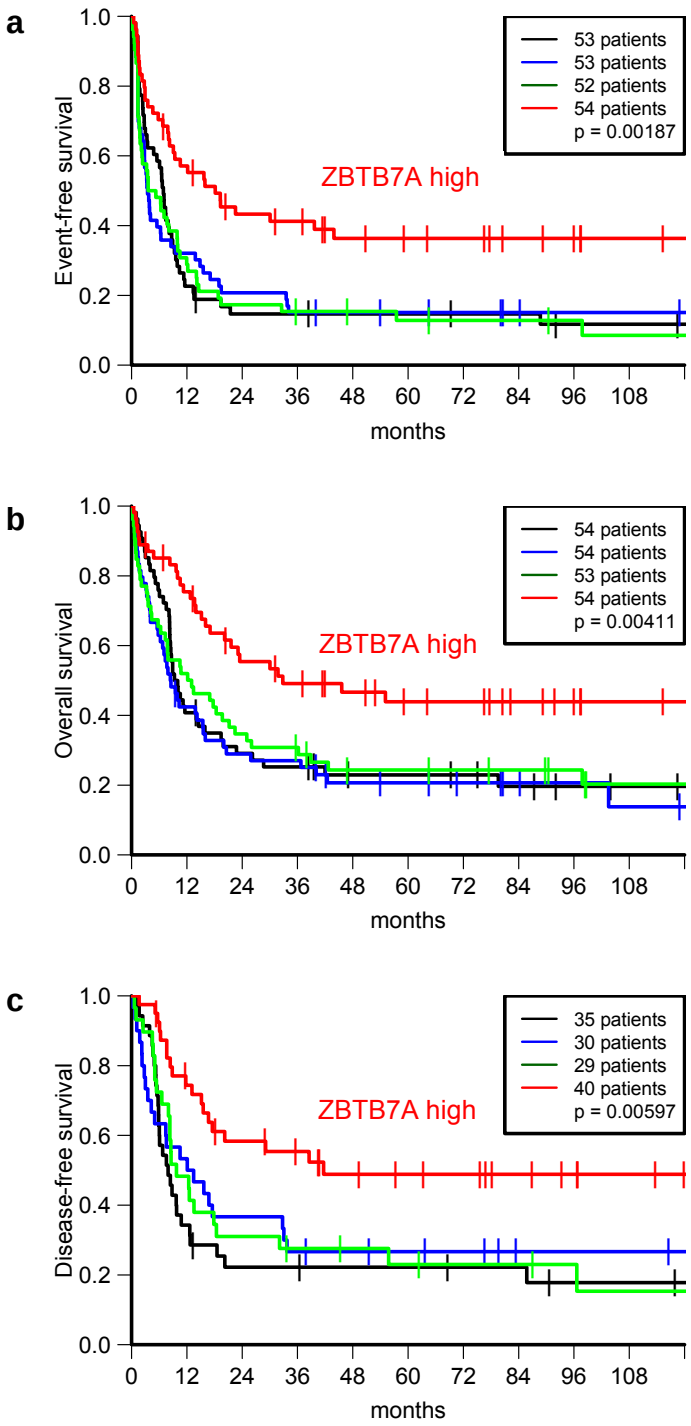
Supplementary Figure 4 (continued). (c) Western blot analysis of cytoplasmic (cyt) and nuclear (nuc) protein fractions extracted from HEK293T cells expressing ZBTB7A wild-type or mutants.



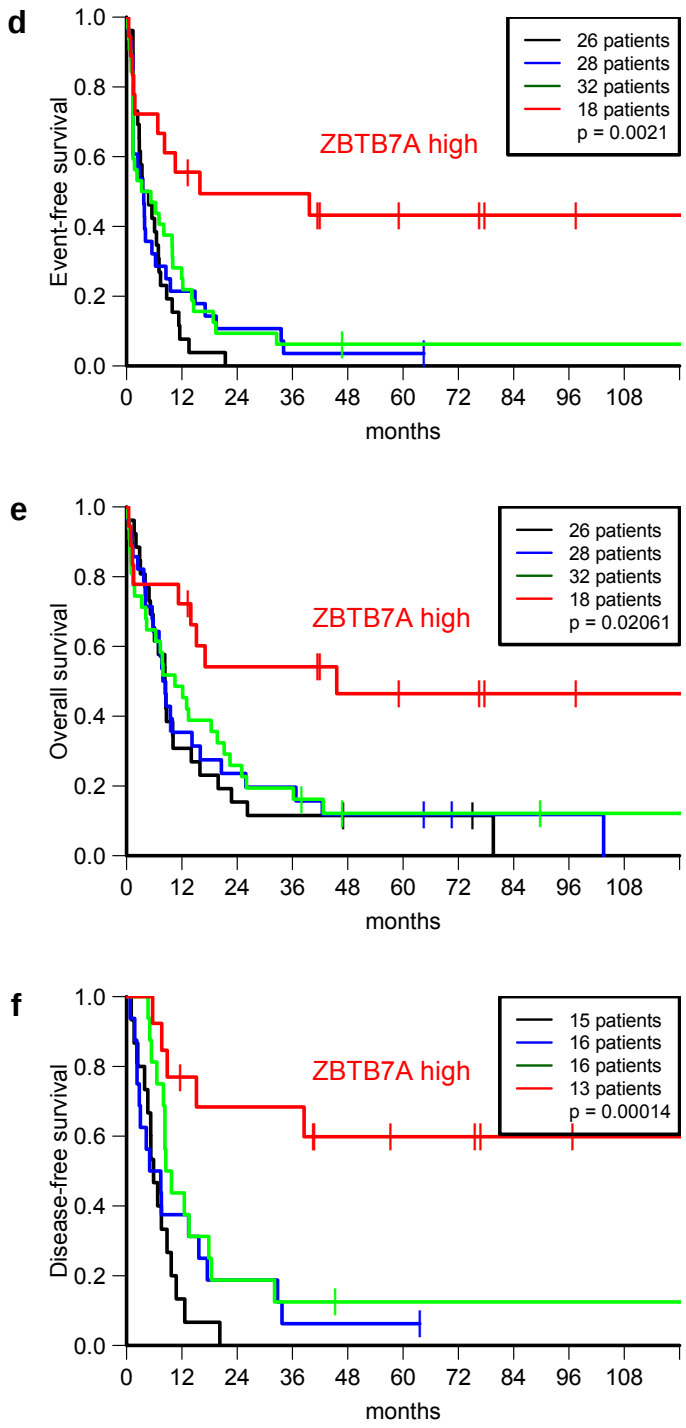
Supplementary Figure 5. Survival of t(8;21) positive AML patients according to *ZBTB7A* mutation status. P values were calculated by the log-rank test. (a) Event free survival, (b) Overall survival and (c) Relapse-free survival.



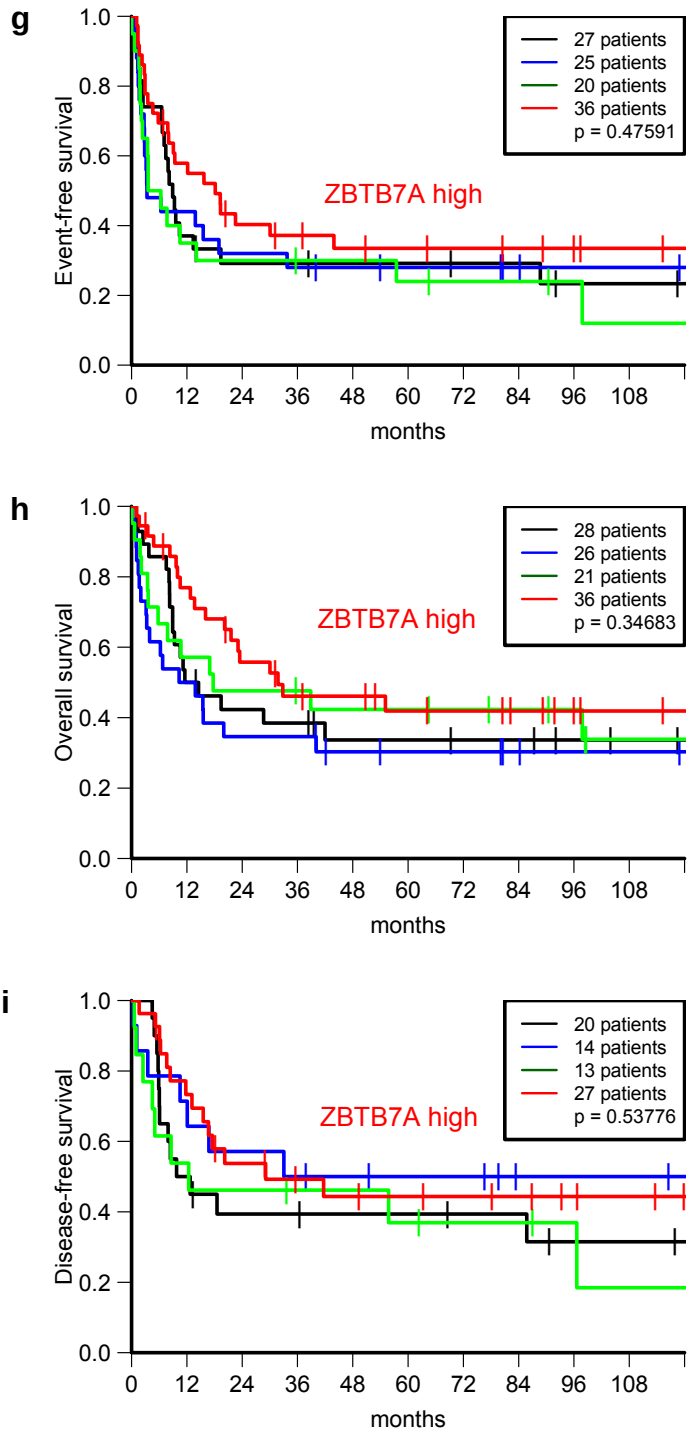
Supplementary Figure 6. *ZBTB7A* expression in cytogenetic subgroups of AML.



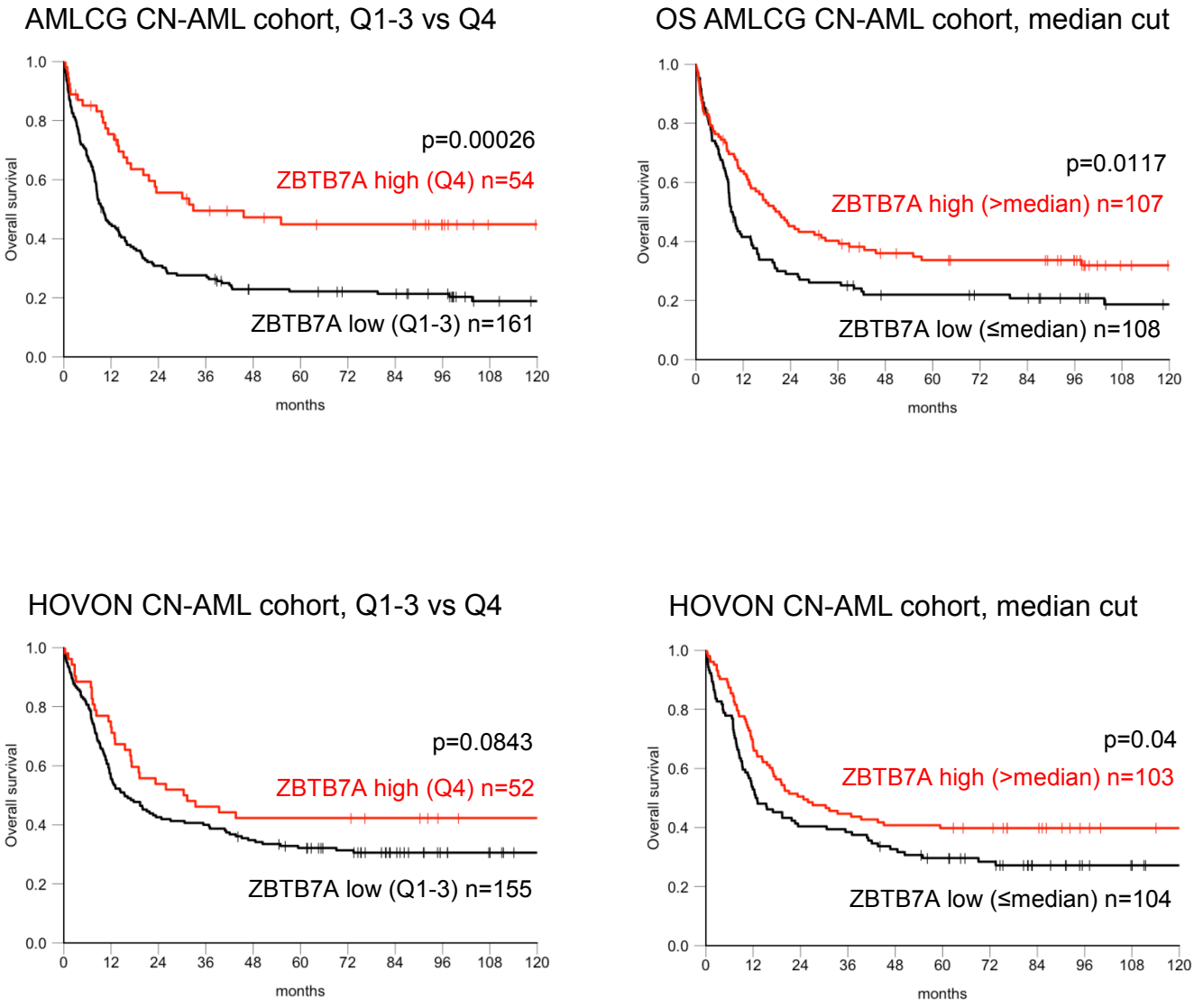
Supplementary Figure 7. Survival of patients with cytogenetically normal (CN-)AML according to *ZBTB7A* expression (GSE37642). High *ZBTB7A* expression (red) was defined as the highest (4th) quartile of expression values observed in CN-AML patients. Patients with *ZBTB7A* expression levels in the 1st to 3rd quartile were classified as having low expression. P values were calculated by the log-rank test. (a) Event-free survival (b) Overall survival (c) Relapse-free survival.



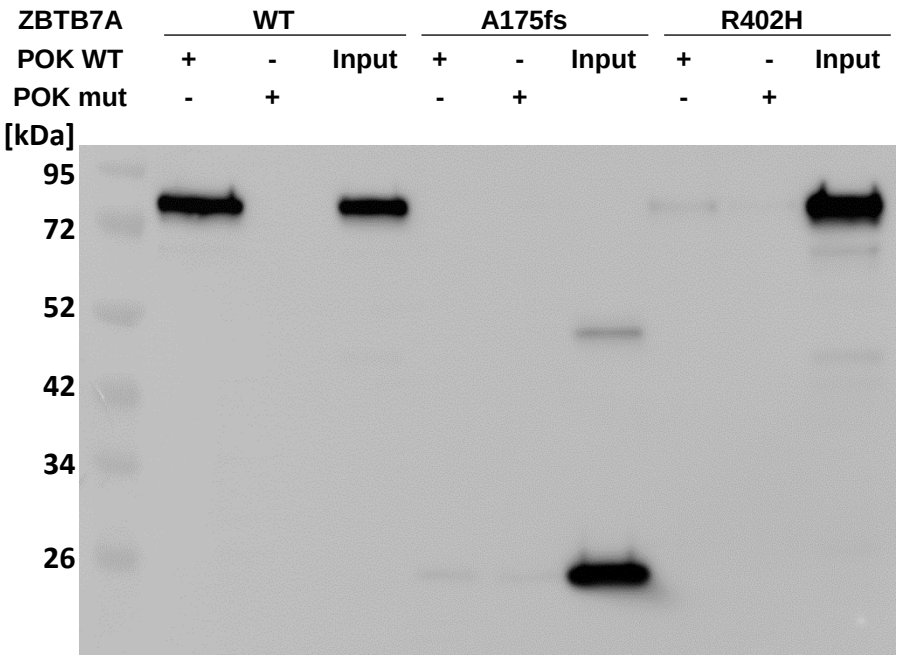
Supplementary Figure 7 (continued). Survival of **patients ≥ 60 years** with CN-AML according to *ZBTB7A* expression (GSE37642). (d) Event-free survival patients (e) Overall survival (f) Relapse-free survival.



Supplementary Figure 7 (continued). Survival of **patients < 60 years** with CN-AML according to *ZBTB7A* expression (GSE37642). (g) Event-free survival patients (h) Overall survival (i) Relapse-free survival.



Supplementary Figure 8. Overall survival of patients with CN-AML according to *ZBTB7A* expression in the AMLCG-cohort (GSE37642) and the HOVON cohort (GSE14468 and GSE1159). P values were calculated by the log-rank test.



Supplementary Figure 9. Uncropped Western blot scan underlying Figure 2c.

Supplementary Table 1. Somatic variants from exome sequencing of two AML t(8;21) patients.

UPN	Chr	Position (hg 19)	Gene	Ref	Var	dbSNP	VarFreq (%)	Type	AA Change
1	7	138437432	ATP6V0A4	C	T		52.9	nonsynonymous SNV	NM_020632:c.967G>A:p.A323T
1	19	16040374	CYP4F11	T	A		87	nonsynonymous SNV	NM_021187:c.236A>T:p.Q79L
1	3	183823156	HTR3E	A	-GCAAG		50.5	frameshift deletion	NM_001256613:c.662_666delGCAAG: p.K222fs
1	5	38869211	OSMR	A	G		37.0	nonsynonymous SNV	NM_001168355:c.65A>G:p.Q22R
1	16	67695894	PARD6A	C	A		41.8	nonsynonymous SNV	NM_016948:c.385C>A:p.P129T
1	6	150569915	PPP1R14C	G	A		26.2	nonsynonymous SNV	NM_030949:c.457G>A:p.G153S
1	14	61186589	SIX4	G	A		45.1	stopgain SNV	NM_017420: c.1438C>T:p.Q480X
1	15	62994266	TLN2	C	G		47.8	nonsynonymous SNV	NM_015059:c.1772C>G:p.S591C
1	15	81627077	TMC3	C	G	rs376889456	35.4	nonsynonymous SNV	NM_001080532:c.2443G>C:p. E815Q
1	19	4054026	ZBTB7A	C	T		75.2	nonsynonymous SNV	NM_015898:c.1205G>A:p.R402H
1	19	24288837	ZNF254	G	A		43.9	nonsynonymous SNV	NM_203282:c.126G>A:p:M42I
2	12	70724077	CNOT2	C	T		44	nonsynonymous SNV	NM_014515:c.397C>T:p.P133S
2	1	22903142	EPHA8	C	T		34	nonsynonymous SNV	NM_020526:c.592C>T:p.R198C
2	16	30495266	ITGAL	C	T		29.7	nonsynonymous SNV	NM_002209:c.841C>T:p.R281C
2	4	55599321	KIT	A	T	rs121913507	28.3	nonsynonymous SNV	NM_000222:c.2447A>T:p.D816V
2	18	30321954	KLHL14	G	A		45.6	nonsynonymous SNV	NM_020805:c.1006C>T:p.R336W
2	5	140348603	PCDHAC2	G	T	position of rs143196630	37.2	nonsynonymous SNV	NM_018899:c.2252G>T:p.R751M
2	6	42890875	PTCRA	C	T		40.7	nonsynonymous SNV	NM_138296:c.169C>T:p.L57F
2	19	46198897	QPCTL	C	T		51.3	nonsynonymous SNV	NM_017659:c.554C>T:p.T185M
2	6	28540575	SCAND3	A	C		46.8	nonsynonymous SNV	NM_052923:c.3091T>G:p.S1031A
2	3	36534709	STAC	C	T		29.5	nonsynonymous SNV	NM_003149:c.754C>T:p. R252C
2	X	104464034	TEX13A	G	A		44.9	nonsynonymous SNV	NM_031274:c.842C>T:p.T281M
2	19	4054708	ZBTB7A	C	+G		45.7	frameshift insertion	NM_015898:c.522dupC:p.A175fs

Supplementary Table 2. Primer sequences for Sanger sequencing of *ZBTB7A* exon 2.

Region	PCR-amplification primers		Sequencing primers	
Exon2_1	fwd	GGGTGGAACGCTGCTTCT	fwd	CTTGTCAGTGGGCACAGGAA
	rev	GTTCATGGGGTTGCTCTGGA	rev	CTGAGGATGTCACCCACGTT
Exon2_2	fwd	GCTCATGGACTTCGCCTAC	fwd	ACAGCCAACGTGGGTGAC
	rev	GGTAGTAGTCCATGACGCCC	rev	CTCCCGACAGGAAGCCC
Exon2_3	fwd	CCAGAGCGGGATGAGGAC	fwd	ACTCTCCGGGCTTCCTGTC
	rev	GTGTGCACGTGCGTGTATG	rev	GTATGTGTGCGTCTGCGTG

Supplementary Table 3. *ZBTB7A* mutations from gene panel* analysis of 56 AML t(8;21) cases.

UPN	Chr	Position (hg 19)	Gene	Ref	Var	Length	Ref Count	Var Count	VarFreq (%)	Type	AA Change	Sanger validated
1	19	4054026	ZBTB7A	C	T	1	298	901	75.2	nonsynonymous SNV	NM_015898:exon2: c.1205G>A:p.R402H	Yes
3	19	4054027	ZBTB7A	G	A	1	564	136	19.4	nonsynonymous SNV	NM_015898:exon2: c.1204C>T:p.R402C	Yes
3	19	4054708	ZBTB7A	-	G	1	446	156	25.9	frameshift insertion	NM_015898:exon2: c.522dupC:p.A175fs	Yes
4	19	4054727	ZBTB7A	G	-	1	2387	290	10.8	frameshift deletion	NM_015898:exon2: c.504delC:p.P168fs	No
4	19	4055082	ZBTB7A	G	A	1	888	438	33.0	nonsynonymous SNV	NM_015898:exon2: c.149C>T:p.S50L	Yes
5	19	4054141	ZBTB7A	-	TTA	3	242	89	26.9	stopgain insertion	NM_015898:exon2: c.1089_1090insTAA: p.V364delinsX	Yes
5	19	4055085	ZBTB7A	C	T	1	305	211	40.9	nonsynonymous SNV	NM_015898:exon2: c.146G>A:p.R49H	Yes
6	19	4054048	ZBTB7A	C	A	1	522	77	12.9	nonsynonymous SNV	NM_015898:exon2: c.1183G>T:p.G395C	Yes
7	19	4054027	ZBTB7A	G	A	1	2129	1117	34.4	nonsynonymous SNV	NM_015898:exon2: c.1204C>T:p.R402C	Yes
8	19	4054708	ZBTB7A	-	G	1	459	231	33.4	frameshift insertion	NM_015898:exon2: c.522dupC:p.A175fs	Yes
9	19	4054102	ZBTB7A	G	A	1	235	167	41.5	stopgain SNV	NM_015898:exon2: c.1129C>T:p.R377X	Yes
10	19	4048131	ZBTB7A	G	-	1	328	35	9.6	frameshift deletion	NM_015898:exon3: c.1374delC:p.R458fs	No
10	19	4054708	ZBTB7A	-	G	1	208	12	5.5	frameshift insertion	NM_015898:exon2: c.522dupC:p.A175fs	No
11	19	4054994	ZBTB7A	G	T	1	462	174	27.4	nonsynonymous SNV	NM_015898:exon2: c.237C>A:p.F79L	Yes
12	19	4054708	ZBTB7A	-	G	1	7326	552	7.0	frameshift insertion	NM_015898:exon2: c.522dupC:p.A175fs	Yes
13	19	4054977	ZBTB7A	A	C	1	197	629	76.2	nonsynonymous SNV	NM_015898:exon2: c.254T>G:p.L85R	Yes
14	19	4054708	ZBTB7A	-	G	1	872	362	29.3	frameshift insertion	NM_015898:exon2: c.522dupC:p.A175fs	Yes

*JAK1, NRAS, GATA3, PTEN, SMC3, WT1, SF1, CBL, ETV6, KRAS, PTPN11, FLT3, IDH2, TP53, SRSF2, JAK3, CEBPA, U2AF2, DNMT3A, SF3B1, IDH1, ASXL1, RUNX1, U2AF1, SF3A1, MYD88, GATA2, KIT, TET2, FBXW7, IL7R, NPM1, BRAF, EZH2, RAD21, JAK2, NOTCH1, ZRSR2, BCOR, GATA1, SMC1A, STAG2, PHF6, ZBTB7A, ASXL2, FAT1

Supplementary Table 4. Patient characteristics of AML t(8;21) gene panel sequencing cohort.

Variable	Wild-type *ZBTB7A*	Mutated *ZBTB7A*	P value*
No. of patients	43	13	
Median Age, years (range)	55 (23-79)	53 (16-66)	0.148
Male gender, no. (%)	29 (67)	10 (77)	0.7331
White blood cell count G/l, median (range)	9 (1.9-210)	8.3 (3.5-245)	0.9689
Bone marrow blasts %, median (range)	70 (4-95)	55 (14-90)	0.1141
French-American-British (FAB) classification, no. (%)	M1: 7 (20)	M1: 1 (3)	0.6593
	M2: 28 (80)	M2: 10 (83)	1.0000
		M4: 1 (3)	0.2553
Secondary AML (%)	7	8	1.0000
Allogeneic transplantation, no. (%)	4 (12)	2 (22)	0.5928
Complete Remission, no. (%)	18 (55)	6 (67)	0.7083
Relapse, no. (%)	5 (28)	4 (67)	0.1501
Deceased, no. (%)	15 (45)	6 (67)	0.4537

*Two-tailed Fisher's exact test was used to compare categorical variables, while Wilcoxon Mann-Whitney U test was applied for continuous variables

Supplementary Table 5. *ZBTB7A* expression in molecular and age subgroups of CN-AML.

	All CN-AML N=218			CN-AML <60 years N=112			CN-AML >=60 years N=106		
	ZBTB7A^{Q4} N=55	ZBTB7A^{Q1-3} N=163	P	ZBTB7A^{Q4} N=37	ZBTB7A^{Q1-3} N=75	P	ZBTB7A^{Q4} N=18	ZBTB7A^{Q1-3} N=88	P
FLT3-ITD	13/54	70/163	.015	11/36	36/75	.10	2/18	34/88	.03
NPM1	31/53	83/158	.52	20/36	47/74	.53	11/17	36/84	.11
LMR	24/53	34/159	.001	14/36	20/75	.20	10/17	14/84	<.001

ITD, Internal tandem duplication; LMR, low molecular risk genotype; mutated *NPM1* without *FLT3*-ITD, Q4, quartile of patients with highest expression levels of *ZBTB7A,* Q1-3, quartiles of patients with lower expression levels of *ZBTB7A.* P Values were calculated by two-tailed Fisher's exact test.