Integrated genome-wide genotyping and gene expression profiling reveals *BCL11B* as a putative oncogene in acute myeloid leukemia with 14q32 aberrations

Saman Abbas,[1] Mathijs A. Sanders,[1] Annelieke Zeilemaker,[1] Wendy M.C. Geertsma-Kleinekoort,[1] Jasper E. Koenders,[1] Francois G. Kavelaars,[1] Zabiollah G. Abbas,[1] Souad Mahamoud,[1] Isabel W.T. Chu,[1] Remco Hoogenboezem,[1] Justine K. Peeters,[1] Ellen van Drunen,[2] Janneke van Galen,[2] H. Berna Beverloo,[2] Bob Löwenberg,[1] and Peter J.M. Valk[1]

[1]Department of Hematology, Erasmus University Medical Center, Rotterdam; and [2]Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, the Netherlands

ABSTRACT

Acute myeloid leukemia is a neoplasm characterized by recurrent molecular aberrations traditionally demonstrated by cytogenetic analyses. We used high density genome-wide genotyping and gene expression profiling to reveal acquired cryptic abnormalities in acute myeloid leukemia. By genome-wide genotyping of 137 cases of primary acute myeloid leukemia, we disclosed a recurrent focal amplification on chromosome 14q32, which included the genes *BCL11B*, *CCNK*, *C14orf177* and *SETD3*, in two cases. In the affected cases, the *BCL11B* gene showed consistently high mRNA expression, whereas the expression of the other genes was unperturbed. Fluorescence *in situ* hybridization on 40 cases of acute myeloid leukemia with high *BCL11B* mRNA expression [2.5-fold above median; 40 out of 530 cases (7.5%)] revealed 14q32 abnormalities in two additional cases. In the four *BCL11B*-rearranged cases the 14q32 locus was fused to different partner chromosomes. In fact, in two cases, we demonstrated that the focal 14q32 amplifications were integrated into transcriptionally active loci. The translocations involving *BCL11B* result in increased expression of full-length BCL11B protein. The *BCL11B*-rearranged acute myeloid leukemias expressed both myeloid and T-cell markers. These biphenotypic acute leukemias all carried *FLT3* internal tandem duplications, a characteristic marker of acute myeloid leukemia. *BCL11B* mRNA expression in acute myeloid leukemia appeared to be strongly associated with expression of other T-cell-specific genes. Myeloid 32D(GCSF-R) cells ectopically expressing Bcl11b showed decreased proliferation rate and less maturation. In conclusion, by an integrated approach involving high-throughput genome-wide genotyping and gene expression profiling we identified *BCL11B* as a candidate oncogene in acute myeloid leukemia.

Introduction

Acute myeloid leukemia (AML) is a heterogeneous clonal neoplasm characterized by accumulated genetic aberrations, which result in enhanced proliferation, blocked differentiation and increased survival of the leukemic blast cells and a variable response to therapy.[1,2] In the past decades a number of recurrent cytogenetic abnormalities have been identified in AML, such as the chromosomal aberrations t(8;21) and inv(16).[1,2] These recurrent molecular lesions result in expression of fusion proteins of which the leukemic potential, in combination with additional genetic events, has been demonstrated in both *in vitro* and *in vivo* models.[3] In addition to cytogenetic abnormalities, acquired mutations in disease genes, such as *FLT3*, *NPM1*, *RUNX1* and *CEBPA*, have recently been demonstrated to be involved in AML as well.[1,2] Several acquired molecular aberrations carry prognostic value and have been incorporated in routine molecular analyses of AML.[1,2]

Nowadays, various genome-wide approaches, such as gene expression profiling, genome-wide genotyping and next-generation sequencing, enable detailed analyses of hematologic malignancies to identify novel pathogenic genes.[2,4,5] Examples of aberrations in myeloid proliferative malignancies revealed with these novel technologies are mutations in the genes *IDH1*, *TET2*, *DNMT3A*, *ASXL1*, and *EZH2*.[6-11]

Besides balanced translocations, large chromosomal regions showing losses or gains of genetic material are apparent in the leukemic blasts of AML patients, e.g., those involving chromosome 5 and 7.[1] In the past two decades, attempts to identify the tumor suppressor genes located on these chromosomes have failed. Through genome-wide single nucleotide polymorphism genotyping it has become possible to simultaneously genotype hundreds of thousands of single nucleotide polymorphisms in a single assay. In addition, single nucleotide polymorphism platforms can also be conveniently used to determine chromosomal copy numbers, similarly to array comparative genomic hybridization. Genomic DNA can be examined with an inter-marker distance of several hundreds of base pairs, which makes it feasible to detect (micro) deletions and/or amplifications that are missed with conventional cytogenetics. The application of high-throughput single nucleotide polymorphism genotyping has been elegantly demonstrated to be powerful for the identification of disease genes, such as those for acute lymphoblastic leukemia[4,12,13] or AML.[14] Another major advantage of single

 doi:10.3324/haematol.2013.095604
The online version of this article has a Supplementary Appendix.
Manuscript received on August 18, 2013. Manuscript accepted on January 16, 2014.
Correspondence: p.valk@erasmusmc.nl

nucleotide polymorphism arrays is the fact that allele losses are directly recognizable as loss-of-heterozygosity. In fact, single nucleotide polymorphism arrays revealed that approximately 20% of cases of AML exhibit large non-random regions of homozygosity without changes in copy number as a result of segmental uniparental disomy, often indicating mutations in genes within these regions. These areas of uniparental disomy have been associated with mutations in *CEBPA*, *WT1*, *FLT3* and *RUNX1*.[15,16]

In addition, deletion, amplification and uniparental disomy may have other important consequences, such as those on gene expression. Juxtaposition of regulatory sequences may result in increased or decreased expression of affected genes. Thus, genome-wide analyses to detect copy number changes and loss of heterozygosity in the context of gene expression may also point towards pathogenic genes. We recently developed SNPExpress, an easily accessible software tool to analyze single nucleotide polymorphism genotype calls, copy number and gene expression in an accurate, efficient combinatorial way.[17]

In this study, we identified *BCL11B* as the candidate oncogene in AML through an integrated approach of genome-wide genotyping and gene expression profiling, followed by next-generation sequencing. *BCL11B* is a Kruppel family zinc finger family gene located at 14q32, associated with transcriptional co-repressor complexes in mammalian cells and a pivotal regulator of differentiation and survival of hematopoietic cells, especially T cells.[18] We demonstrated that *BCL11B* is involved in a number of cryptic 14q32 translocations in AML, in which *BCL11B* and T-cell associated genes expression levels are increased concomitantly. Overexpression of BCL11B in a murine myeloid cell line model inhibits proliferation.

Methods

Patients' samples

This study was approved by our local Medical Ethical Committee (MEC-2004-030 and MEC2007-364). After informed consent, bone marrow aspirates or peripheral blood samples were collected from a representative cohort of AML patients. Eligible patients had a diagnosis of primary AML, confirmed by cytological examination of blood and bone marrow. All patients were treated according to HOVON (Dutch-Belgian Hematology-Oncology Co-operative group) protocols (*http://www.hovon.nl*). Further details are available in the *Online Supplementary Appendix*.

Genome-wide genotyping and gene expression profiling

Genome-wide genotyping data sets of 48 patients with various subtypes of AML were generated using Affymetrix 500K *Nsp*I/*Sty*I DNA mapping arrays and of 89 patients with cytogenetically normal AML using Affymetrix 250K *Nsp*I or *Sty*I DNA Mapping arrays. The copy numbers of all AML samples were calculated using diploid references, i.e., 15 normal karyotype AML samples. Further details are available in the *Online Supplementary Appendix*.

Gene expression profiles of the same AML cases were generated using Affymetrix HG-U133 plus 2.0, as described elsewhere (GEO Series accession number GSE6891).[19] Pearson correlation analyses were performed as described previously.[20]

The genome-wide genotyping and gene expression profiling data sets were examined using SNPExpress.[17]

Fluorescence in situ hybridization

Dual color fluorescence *in situ* hybridization (FISH) was performed with BAC clones RP11-431B1, RP11-876E22, RP11-830F3, RP11-782I5, RP11-450C22, RP11-57E12, RP11-1069L3 and RP11-242A7 covering the *BCL11B* region and regions up- and downstream (BACPAC resources, Oakland, USA). Further details are available in the *Online Supplementary Appendix*.

Targeted sequencing of the 14q32 genomic region

Library preparation and targeted re-sequencing was performed following previously described protocols.[21] In brief, high molecular weight DNA from AML #2301 and #7073 was sheared using a Covaris E210 waterbath sonificator. The *BCL11B* 14q32 - tel genomic region [chr14:93930247-105928955 (hg19)] was captured with a Roche/Nimblegen SeqCap EZ Choice XL Library. The captured region was subsequently paired-end sequenced using Illumina HiSeq2000. The data were analyzed using an in-house pipeline, including visualization with IGV, which identifies single nucleotide variants, small and large indels and copy number variations. The chromosomal breakpoints were identified using Breakdancer[22] in the 14q32 region and the partner chromosome. The genomic fusions were subsequently confirmed by Sanger sequencing.

Western blot analyses

Western blot analyses were carried out using affinity-purified rabbit polyclonal anti-BCL11B antibody (Novus Biologicals, Littleton, USA). Immune complexes were detected by binding anti-mouse IgG conjugated to horseradish peroxidase (DAKO, Heverlee, Belgium) followed by an enhanced chemiluminescence assay (Amersham Bioscience, Piscataway, NJ, USA) and GAPDH was stained with primary affinity-purified rabbit polyclonal antibody (α-GAPDH FL-335) (Santa Cruz Biotechnology, California, USA). Further details are available in the *Online Supplementary Appendix*.

DNA constructs and generation of BCL11B-expressing 32D/GCSFR cells

A murine Bcl11b cDNA (kindly donated by Dorina Avram, Albany Medical Center, Albany, NY, USA) was subcloned into a pLXSN expression vector under the control of the 5' long terminal repeat of the Moloney murine sarcoma virus (Clontech, Mountainview, USA). Vector constructs were confirmed by nucleotide sequencing and retrovirally transfected into 32D cells that stably express human granulocyte colony-stimulating factor receptor (GCSF-R)[23] using Fugene transfection reagent (Roche, Indianapolis, USA). Cells were stimulated with interleukin-3 (25 ng/mL) or granulocyte colony-stimulating factor (25 ng/mL), counted and assessed for proliferation and granulocytic differentiation. Further details are available in the *Online Supplementary Appendix*.

Results

Genome-wide genotyping of cytogenetically abnormal and normal cases of acute myeloid leukemia

In total, DNA mapping array profiles of 137 patients with AML were generated (Figure 1). Initially a subgroup of 48 AML samples was selected based on previous gene expression profiling studies, i.e., 21 AML cases from gene expression profiling clusters #4 and #15 (100% *CEBPA* mutant or *CEBPA* silenced[24]), 13 AML cases from cluster #9 [100% inv(16)] and 14 AML cases from cluster #10 (adverse prognosis).[20] In addition, DNA mapping array profiles, i.e., Affymetrix 250K *Nsp*I or *Sty*I DNA mapping

array, of 89 AML cases with normal karyotypes were generated.

With the Affymetrix 500K *Nsp*I/*Sty*I DNA Mapping arrays, all known numerical cytogenetic aberrations, i.e., whole chromosome and interstitial deletions and amplifications that had been identified with cytogenetic banding analysis, were recognized in the 48 cytogenetically abnormal AML samples, as long as the abnormalities were present in over 30% of the AML cells.[17] Also, in approximately 25% of all cases large regions of segmental uniparental disomy were detected, often involving whole chromosome arms.[17]

In addition to the known cytogenetic aberrations, relatively low numbers of small interstitial deletions and amplifications were detected in the 137 AML cases. However, some of these were indicative of the presence of cryptic translocations, such as cryptic t(5;11), t(9;22) and t(4;11), which are known to encode chimeric fusion proteins essential for leukemogenesis. All fusion transcripts involved in these translocations, i.e., *NUP98-NSD1*,[25] *BCR-ABL* and *MLL-AF4*, were confirmed by reverse transcriptase polymerase chain reaction analysis. Thus, although relatively small numbers of aberrations were found, most being non-recurrent, they may reliably tag relevant leukemic lesions.

Integrated analysis of genome-wide genotypes and gene expression profiles

By an integrated approach using the genome-wide genotyping data and previously determined gene expression profiles of the primary AML samples (Figure 1),[20] we searched for genes aberrantly expressed as a result of numerical changes in the AML genome. Using SNPExpress,[17] we identified two AML cases with relatively small interstitial amplifications in the 14q32 region (#2301 amplification: 482 Kb, 3 copies and #7073 amplification: 460 Kb, 3 copies) (Figure 2A,B). The amplified region encompassed the genes *BCL11B*, *CCNK*, *C14orf177* and *SETD3*. Interestingly, *BCL11B* mRNA was highly expressed in the two AML cases with numerical changes, whereas the expression of *C14orf177*, *CCNK* and *SETD3* were not affected as compared to other AML cases (Figure 2A,B). In addition, *BCL11B* mRNA was highly expressed in AML #2301 and #7073, whereas expression was low or absent in control AML cases (Figure 2A). This could indicate that *BCL11B* becomes overexpressed as a result of a genomic rearrangement in these AML cases. The small interstitial amplifications in AML #2301 and #7073 may point towards cryptic translocations, as was seen in AML cases carrying the cryptic translocations t(5;11), t(9;22) and t(4;11).

Fluorescence in situ hybridization reveals translocations in acute myeloid leukemia #2301 and #7073 involving BCL11B

To confirm the amplifications in the *BCL11B* locus in the two AML cases, we performed FISH analysis with a probe covering the *BCL11B* gene (RP11-431B1) and a probe flanking this locus (RP11-74H1) (Figure 3A). On metaphase spreads of both AML cases an additional *BCL11B* allele was apparent (Figure 3B). This is in line with the expected copy number change for the *BCL11B* locus (n=3) as shown with SNPExpress (Figure 2A,B). In fact, through verification using chromosomal paints we showed that *BCL11B* was translocated to chromosome 6 in AML case #2301 and chromosome 8 in AML #7073 (*data not shown*).

Amplified 14q32 genomic regions are integrated into transcriptionally active loci

We performed 14q32 capture sequencing on AML cases #2301 and #7073. Paired-end sequencing of the captured region on the Illumina HiSeq2000 in AML case #2301 and case #7073 revealed the partner chromosomal regions on chromosome 6 and 8, respectively. Paired-end reads spanning 14q32 and chr6q25.3 (chr6:156717480 and chr14:99110325; chr6:156587275 and chr14:99748893) were identified in AML case #2301 and reads spanning 14q32 and 8q24.21 (chr8:130485869 and chr14:99179210) in case #7073. These breakpoints were confirmed by polymerase chain reaction analysis of genomic DNA, followed by Sanger sequencing. The *BCL11B* encoding-amplified

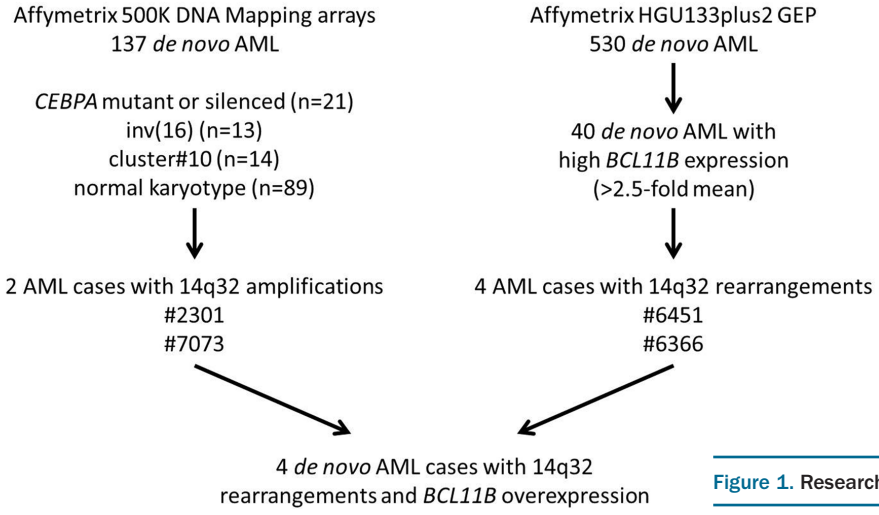

Figure 1. Research design. GEP: gene expression profiling.

DNA integrated into two transcriptionally active regions on chromosomes 6 and 8, i.e., on 6q25.3 into an expressed sequence tag sequence CB984582 and on 8q24.21 into the large non-coding RNA gene Coiled-Coil Domain Containing 26 (CCDC26). Both polyA genes were transcriptionally active in the two cases of AML, and various other AML cases, and are subjected to mRNA splicing as demonstrated by RNA sequencing on an Illumina HiSeq2500 (*data not shown*), indicating that these RNA are expressed in myeloid cells. No fusion transcripts between *BCL11B* and RNA encoded by the partner chromosomes could be detected by RNAseq, suggesting that regulatory sequences on chr6q25.3 and 8q24.21 may activate the *BCL11B* gene in the *BCL11B*-rearranged AML cases.

Acute myeloid leukemia case #2301 expresses full-length BCL11B

The translocations involving *BCL11B* could result in increased expression of either full-length BCL11B or a fusion protein involving BCL11B. In fact, *BCL11B* mRNA expression in AML#2301 and AML#7073, were, respectively, 12- and 8-fold over mean *BCL11B* mRNA expression in 530 AML cases (219528_s_at; 22895_s_at; 224310_s_at).[19] Next, we examined the expression profiles

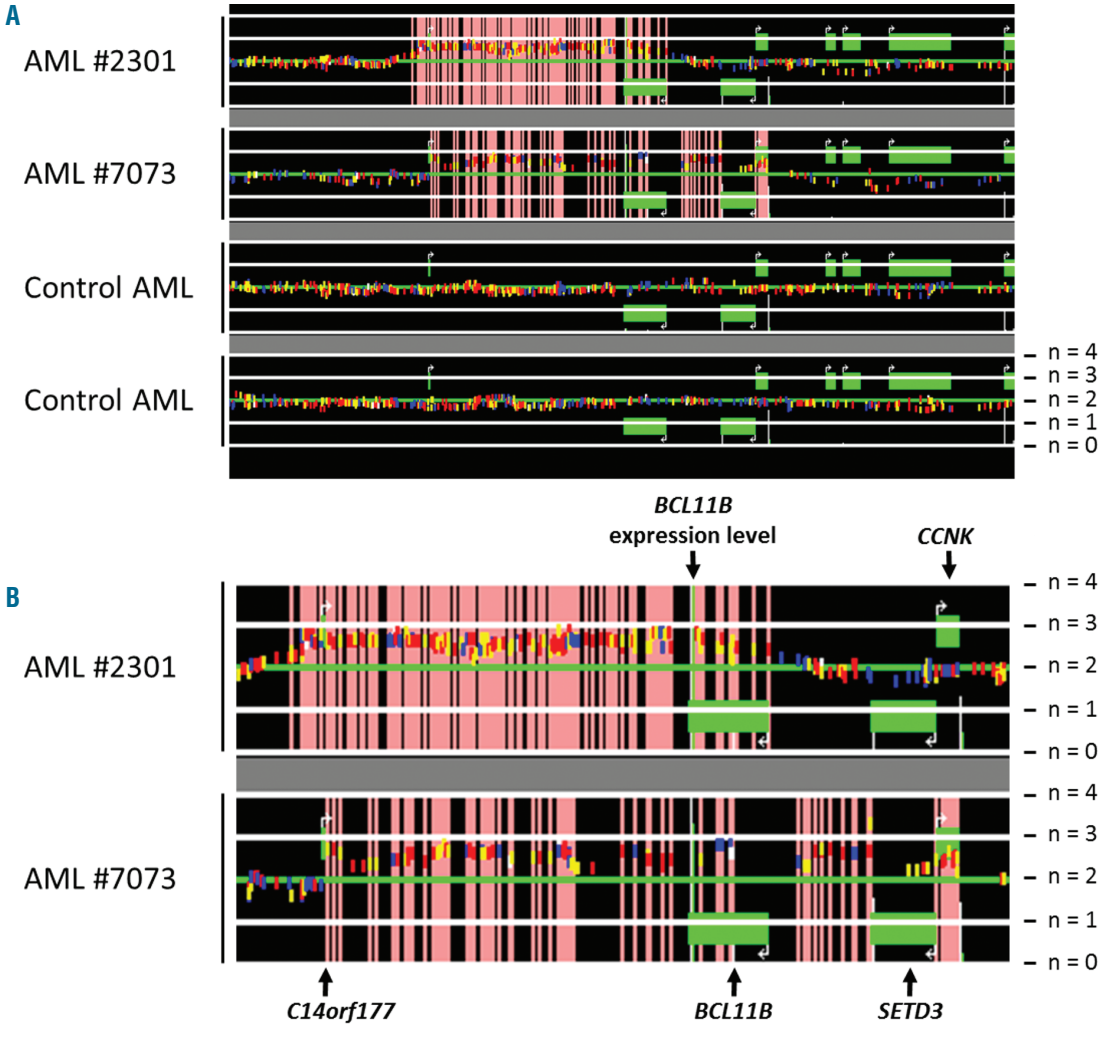

Figure 2. Identification of interstitial amplifications on 14q32.2 using SNPExpress. (A) Sequential alignment of the genotypes with copy numbers from the Affymetrix DNA mapping array of chromosome 14q32.2 of four AML samples.[17] The copy numbers are shown for each individual patient by horizontal lines (n=0, 1, 2, 3, 4). The single nucleotide polymorphism genotypes are sequentially aligned along the chromosome (AA: red; BB: yellow; AB: blue, no call: white). Gains (default n>2.5) are depicted as the pink background. Gene expression levels are visualized as vertical white bars at the chromosomal position of the gene-specific probe set. In the event that multiple probe sets span the same region in the chromosome-wide view the vertical gene expression bars are green and proportional to the highest expression value. The green boxes represent exons of the encoded genes, and the arrows indicate the orientation. In AML #2301 and AML #7073 clear amplifications are visible, whereas these aberrations are absent in the two control case of AML. (B) Snapshot of SNPExpress illustrating the amplified region in AML cases #2301 and #7073 from Figure 2A, showing the genes located within the amplified regions. *C14orf177* and *BCL11B* are amplified in both AML cases, whereas *SETD3* and *CCNK* are amplified only in AML #7073. BCL11B expression is increased in AML #2301 and #7073 as indicated by the green bar (multiple probe sets), whereas BCL11B expression is absent in control AML cases (Figure 2A).

obtained with Affymetrix Human Exon 1.0 ST Array for AML case #2301. This analysis showed that in AML #2301 all four exons of *BCL11B* were expressed at similarly high levels (*data not shown*). The fact that exon 1 of *BCL11B,* containing the ATG start codon, was expressed suggested that full-length BCL11B is expressed rather than a fusion protein in which parts of BCL11B are involved.

Protein lysates of AML case #2301 were available for western blot analyses. With western blot analyses of whole, cytoplasmic and nuclear cell lysates of AML #2301 we assessed both the size and localization of the BCL11B protein. Immunodetection with BCL11B antibodies confirmed the expression of full-length BCL11B protein (Figure 4). Moreover, BCL11B expression appeared to be restricted to the nucleus. Of note, full-length BCL11B was also highly expressed in AML case #2238, an AML without any known aberration involving *BCL11B*.

Fluorescence in situ *hybridization analyses of selected acute myeloid leukemia cases with high* BCL11B *mRNA expression reveals additional cases with* BCL11B *translocations*

The fact that FISH analysis of AML cases showing high expression of *BCL11B* mRNA revealed translocations involving *BCL11B* raised the possibility that other AML cases with aberrantly high *BCL11B* expression would also harbor translocations involving *BCL11B*. Gene expression profiling of 530 AML cases[19] showed variable expression of *BCL11B* mRNA in subsets of AML, including case #2301 and #7073 (Figure 5A). Interestingly, AML cases with high *BCL11B* expression appear to cluster in several distinct clusters (Figure 5A), suggesting that *BCL11B* overexpression is associated with distinct AML subtypes. We selected 40 AML cases with increased *BCL11B* mRNA expression, i.e., >2.5-fold above mean *BCL11B* expression in primary AML (Figure 5A), and performed FISH analysis on the *BCL11B* chromosomal region.

FISH analyses revealed two additional AML cases with a *BCL11B* translocation (AML #6366 and #6451) (Figure 5A,B). With specific chromosomal paints, we showed that in AML case #6451 the *BCL11B* locus was translocated to chromosome 7 (*data not shown*). Further FISH could not be carried out on AML #6366 due to a lack of material.

Immunophenotyping and molecular analyses of acute myeloid leukemia cases carrying BCL11B *aberrations*

Immunophenotyping the AML cases demonstrated that those with *BCL11B* translocations expressed not only myeloid markers, but also lymphoid markers such as CD2, CD3, and CD7 (Table 1). Cytoplasmic CD3 expression was present in case #2301, suggesting that it may better correspond to a T-acute lymphoblastic leukemia, but cytoplasmic CD3 was absent in the remaining cases. In fact, these AML cases appeared to have a biphenotypic signature, i.e., expressing (early) myeloid as well as T-cell-associated markers.

Well-known recurrent molecular abnormalities were searched for in the AML cases carrying *BCL11B* aberrations. These analyses demonstrated, with no exception, that these AML cases carried mutations in the *FLT3* gene, i.e., *FLT3* internal tandem duplications (ITD) in all cases and one case had a concurrent *FLT3* mutation in the tyro-

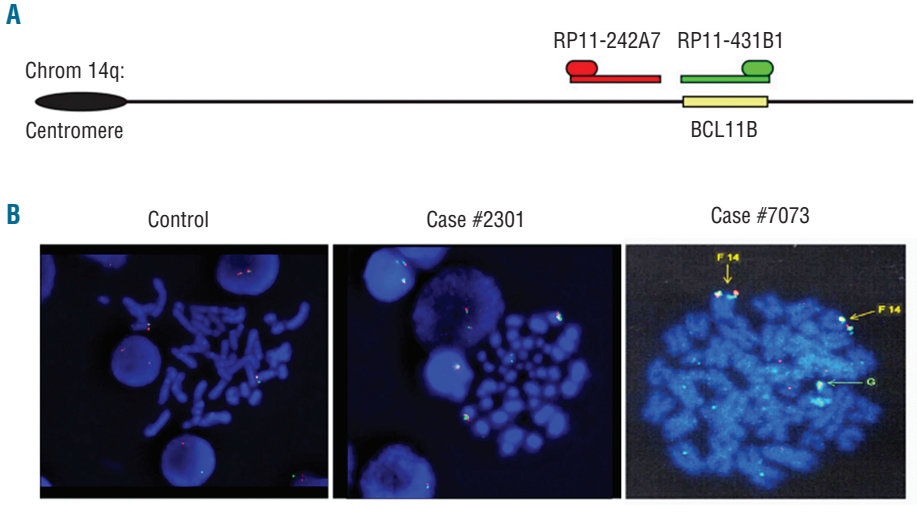

Figure 3. FISH analysis of AML cases #2301 and #7073 using probes specific for *BCL11B* and flanking *BCL11B*. (A) Schematic representation of the fluorescein isothiocyanate-labeled BAC probe (RP11 431B1) covering the *BCL11B* locus and Texas red-labeled BAC probe (RP11 242A7) covering the region adjacent to *BCL11B*. (B) Microscope images of FISH analysis performed on metaphase chromosomes of AML cases #2301 and #7073 showing additional green signal (RP11431B1) indicative of an extra copy of the *BCL11B* locus.

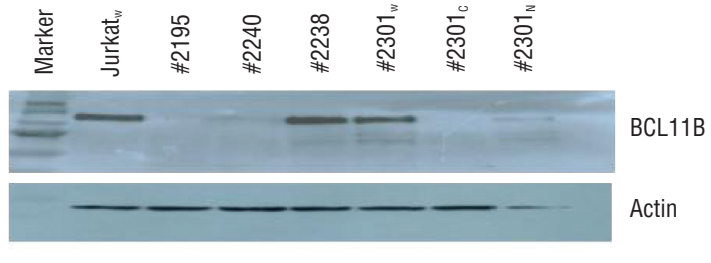

Figure 4. Western blot analysis of BCL11B in AML case #2301. Western blot analysis with a BCL11B-specific antibody demonstrates high expression of full-length BCL11B in AML case #2301, in the nuclear compartment (upper panel). Whole cell lysates from Jurkat, an acute T-cell leukemia cell line, and AML #2238 show high BCL11B expression. AML cases #2195 and #2240 with low *BCL11B* mRNA expression were used as negative controls (#2301$_w$:whole cell lysate; #2301$_c$: cytoplasmic lysate; #2301$_n$: nuclear lysate). β-actin was used as a loading control (lower panel).

sine kinase domain (TKD) (Table 1). We did not identify mutations in *K-RAS*, *N-RAS*, *c-KIT*, *IDH1*, *IDH2*, *ASXL1* or *CEBPA* in the AML cases carrying a *BCL11B* abnormality. Case #6366 also carried a *DNMT3A* mutation (Table 1).

We analyzed the immunophenotype of the *BCL11B* non-rearranged cases with *BCL11B* overexpression and did not find a specific pattern of T-cell specific markers. Several cases did express CD7, but this aberrant marker is relatively frequently present in AML (present in approximately 30% of cases). Moreover, we were unable to demonstrate a significant association between *BCL11B* overexpression and *FLT3*-ITD or -TKD mutations in *BCL11B* non-rearranged cases of AML.

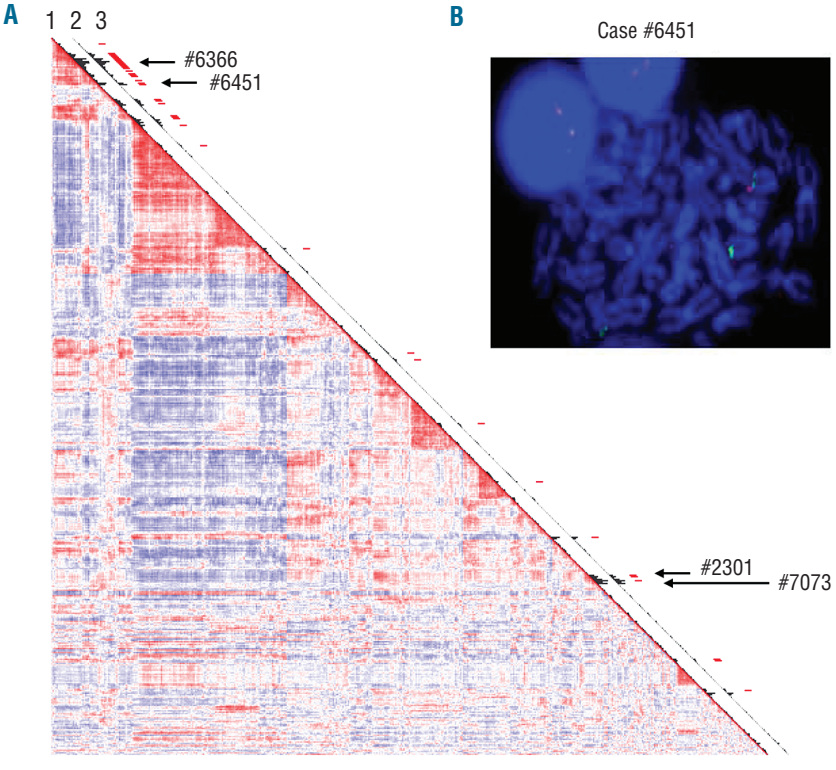

Figure 5. Correlation view based on gene expression profiling of 530 AML cases.[19] (A) Pearson correlation view of 530 AML cases showing gene expression correlation based on 2847 probe sets. The black bars indicate expression of *BCL11B* 1: *BCL11B* expression: 219528_s_at and 2: *BCL11B* expression: 222895_s_at, where the size of the bars is proportional to the levels of *BCL11B* expression; 3: selected AML with *BCL11B* overexpression (>2.5-fold mean). The *BCL11B*-rearranged cases #2301, #6451, #6366 and #7073 are indicated by arrows. (B) FISH analysis performed on metaphase spreads of AML case #6451 showing disassociation of the probe RP11242A7 (red) and RP11431B1 (green) indicating translocation of *BCL11B*.

Table 1. Clinical, molecular and immunophenotypic data of the AML cases with *BCL11B* translocations. The mutation status for *FLT3*, *NPM1*, *N-RAS*, *K-RAS*, *CEBPA*, *c-KIT*, *ASXL1*, *IDH1*, *IDH2* and *DNMT3A* was determined as described previously.[47-49] No mutations were present in *NPM1*, *N-RAS*, *K-RAS*, *CEBPA*, *KIT*, *ASXL1*, *IDH1*, and *IDH2* in the four AML cases (pos: mutant; neg: wild-type).

Patient number	AML #2301	AML #7073	AML #6451	AML #6366
FLT3 ITD	pos	pos	pos	pos
FLT3 TKD835	neg	neg	pos	neg
DNMT3A mutation	neg	neg	neg	pos
FAB subtype	M1	M4	M1	M2
WHO	1 WHO	1 WHO	2 WHO	0 WHO
Gender	male	male	female	female
Karyotype	46,XY[21]/ ?46,XY,inc[9]	46,XY[20]	46,XX,del(7)(q21q35)[5]/ 46,idem,add(13)(q3?4)[17]/ 46,idem,add(9)(q3?4)[2]/ 46,XX[15]	53,XX,+4,+8,+10,+13, +14,+15,+20[4]/ 46,XX[35]
Immunophenotype	CD45(+),HLA-DR-,CD34+, TdT+,MPOpartial+,CD1-, CD2+,CD3+,CD4-,CD5-	CD15partial+,CD33+,CD7+, CD36partial+,CD56-,CD65s-, CD117partial+,CD133+, CD4partial+	CD45(+),HLA-DR+,CD34+, TdTpartial+,MPOpartial+, CD11c-,CD13+,CD15-, CD15s partial+,CD33-, CD65s-,CD117+,CD133+, CD2+,CD7partial+	CD45(+),HLA-DR+, CD34partial+,TdT-, MPOpartial+,CD11c partial+, CD13partial+,CD15partial+, CD33+,CD36partial+, CD117 partial+, CD133+, CD4 partial+, CD7(+)

BCL11B *is aberrantly expressed in acute myeloid leukemia and associated with the T-cell gene expression signature*

To investigate whether other AML cases with elevated *BCL11B* mRNA expression show full-length BCL11B expression, we carried out western blot analyses on a number of AML cases with high *BCL11B* mRNA expression. All analyzed samples with high *BCL11B* mRNA showed full-length BCL11B protein expression at variable levels (Figure 6). Because of the lack of specimens, BCL11B protein expression analyses in non-rearranged *BCL11B* cases was limited to those shown in Figures 4 and 6.

To examine which genes are co-expressed with *BCL11B* in AML, we performed a Pearson correlation analysis using Affymetrix gene expression profiling data from 530 AML cases.[19] *BCL11B* co-regulated probe sets were calculated across all AML patients. The top 50 *BCL11B* correlating probe sets are highly associated with T cells and T-cell development (*Online Supplementary Table S1*). In fact, the majority of *BCL11B*-associated genes are T-cell specific genes, with *CD3*, *TRBV19*, *IL32*, *LCK*, *TCF7* and *CD2* being just a few examples among many others (*Online Supplementary Table S1*).

Increased Bcl11b *expression results in decreased proliferation of the myeloid cell line 32D(GCSF-R)*

To investigate the effect of *Bcl11b* expression on proliferation and differentiation, immortalized myeloblast-like murine bone marrow cells stably expressing human GCSF-R [32D(GCSF-R)] were transfected with full-length murine *Bcl11b* cDNA. Three 32D(GCSF-R) clones expressing *Bcl11b* were selected and incubated for 10 days in the presence of interleukin-3 or granulocyte colony-stimulating factor. Western blot analyses demonstrated that BCL11B was expressed at all time points (Figure 7A). BCL11B-expressing 32D(GCSF-R) clones showed a consis-

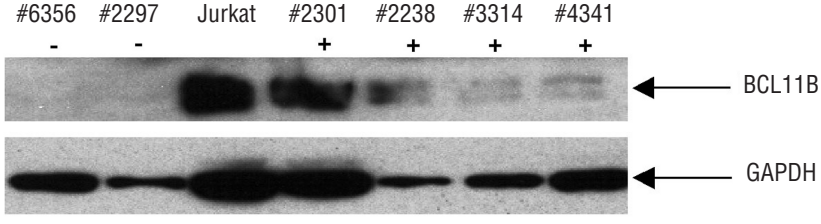

Figure 6. Western blot analyses for BCL11B primary AML. Immuno-detection of the BCL11B protein in AML cases with elevated levels of ***BCL11B*** mRNA (+) and cases with undetectable levels of ***BCL11B*** mRNA (-) (upper panel; Jurkat cell lysate as positive control). GAPDH was used as the loading control (lower panel).

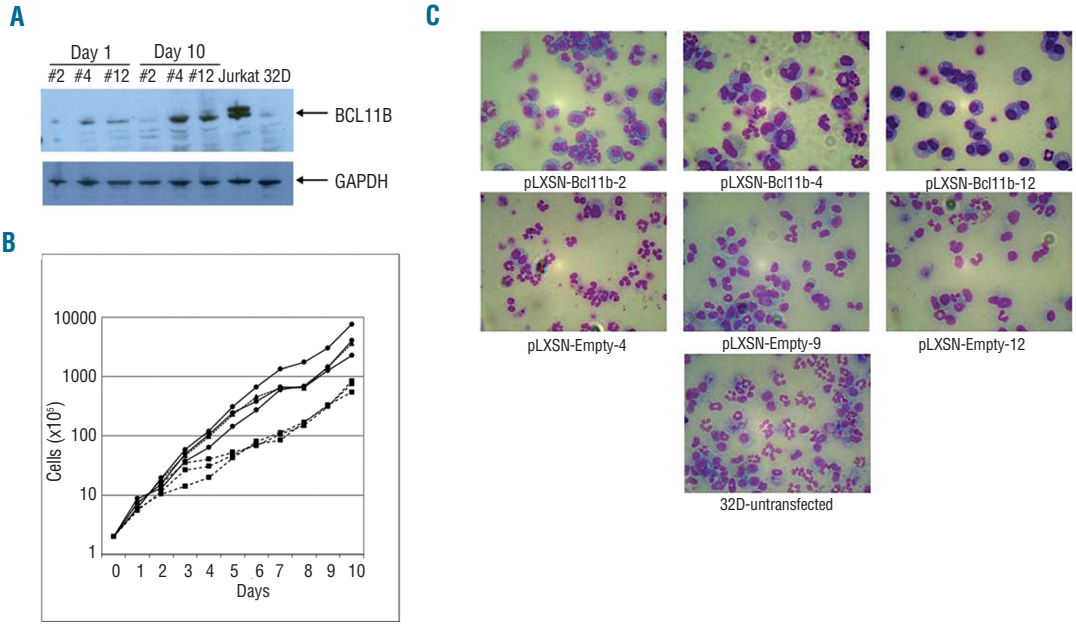

Figure 7. Effects of BCL11B overexpression in murine 32D(GCSF-R) cells. (A) Western blot analyses for BCL11B in 32D(GCSFR) cells. 32D(GCSFR) clones overexpressing BCL11B are indicated by #2, #4, and #12 (interleukin 3 for 1 and 10 days). Lysates obtained from these clones were immunostained for BCL11B at day 1 and day 10 [Jurkat cells: positive control; 32D: 32D(GCSF-R) cells]. GAPDH was used as the loading control (lower panel). (B) Growth curve of 32D(GCSFR) cells with (squares, dashed line) and without (round, solid line) BCL11B expression and parental 32D(GCSFR) cells (triangles, dotted line) incubated with interleukin 3. 32D cells were counted every 24 h for 10 days. (C) May-Grünwald-Giemsa-stained cytospins of 32D(GCSF-R) cells with (upper panel) and without (lower panel) BCL11B expression incubated with granulocyte colony-stimulating factor for 7 days. Granulocytic differentiation is monitored by the presence of cells with segmented nuclei.

tent decreased proliferation rate when cultured in the presence of interleukin in comparison to 32D(GCSF-R) clones containing a control, empty vector (Figure 7B).

We also evaluated the granulocytic differentiation abilities of the same 32D(GCSF-R) clones upon stimulation with granulocyte colony-stimulating factor. Morphological analyses of cytospins did not show consistent maturation defects in the BCL11B-expressing 32D(GCSF-R) clones. However, less maturation towards granulocytes in 32D(GCSF-R) cells expressing BCL11B and more undifferentiated blast cells compared to cells with empty vector were present in some 32D(GCSF-R) BCL11B- expressing clones (Figure 7C). This effect was most apparent in 32D(GCSF-R) BCL11B clones #4 and #12, the 32D(GCSF-R) clones with the highest expression levels of BCL11B (Figure 7A).

Discussion

Simultaneous analyses of genome-wide genotyping and copy number data with gene expression profiling enable the identification of pathogenic genes that, as a result of genomic imbalances, are aberrantly expressed. By an integrated approach involving genome-wide genotyping and gene expression profiling, we identified *BCL11B* as a novel oncogene in AML. Interstitial amplification of 14q32 was initially revealed in two AML cases (#2301 and #7073) by genome-wide genotyping, showing relatively small amplifications, including the *BCL11B* gene. Although other genes were affected by these numerical genetic changes, these 14q32 aberrations resulted in unique *BCL11B* mRNA and full-length BCL11B protein overexpression. By FISH of a selection of 40 AML cases with high *BCL11B* mRNA expression, we identified two additional AML cases bearing *BCL11B* translocations (AML #6451 and AML #6366).

The BCL11 family has two members, BCL11A and BCL11B.[18] Bcl11a was identified as a common retroviral insertion site (Evi9) in murine myeloid leukemias and is required for normal B-cell development.[26] Mice carrying biallelic inactivation of *Bcl11b* developed thymic lymphomas, indicating that loss-of-function mutations in *Bcl11b* contribute to mouse lymphomagenesis and, possibly, to human cancer development.[27] *BCL11B* is a four-exon gene located on 14q32, encoding a Kruppel family zinc finger transcription factor and a key regulator of differentiation and survival of thymocytes.[18]

BCL11B was first associated with hematologic malignancies due to its recurrent involvement with the homeobox transcription factor *TLX3* in a relatively high percentage of cases of pediatric and adult T-cell acute lymphoblastic leukemia carrying the cryptic t(5;14)(q35;q32).[28,29] Less frequently, T-cell acute lymphoblastic leukemia samples with an inv(14)(q11.2q32.31) carry an in-frame transcript of *BCL11B* and the T-cell receptor gene segment *TRDV1*. These cases of acute lymphoblastic leukemia do not express wild-type *BCL11B* transcripts, suggesting that *BCL11B* disruption may contribute to T-cell malignancies in humans.[30] Interestingly, a DNA copy number and sequencing analysis approach recently revealed monoallelic *BCL11B* deletions and missense mutations in 10-15% of cases of T-cell acute lymphoblastic leukemia.[31,32] Structural homology modeling showed that several of the BCL11B mutations disrupted the structure of zinc finger domains required for DNA binding.

A number of myeloid, mixed-lineage, and non-lymphocytic leukemias with 14q32 abnormalities have been reported, however, in these instances the affected genes were not identified.[33-37] The first evidence of *BCL11B* involvement in 14q32 translocations in AML was reported by Bezrookove *et al.*[34] These authors reported one case of t(6;14)(q25~q26;q32) in an adult with AML and used bacterial artificial chromosomes to demonstrate the involvement of *BCL11B* in that case.[34] Due to a lack of the patient's material, the investigators could neither establish the deregulation of *BCL11B* nor identify the partner genomic locus.[34] Of note, the breakpoint in this AML case appeared to be located upstream of the *BCL11B* gene. This resembles the AML cases described here and suggests that the breakpoints in AML are clustered upstream, whereas in acute lymphoblastic leukemia they are downstream, of *BCL11B*.

Specific chromosomal paints demonstrated that different partner chromosomes were involved in the AML cases with a *BCL11B* translocation. The fact that in two AML cases the BCL11B-containing amplified region integrated in transcriptionally active large non-coding RNA may suggest that different regulatory regions of large non-coding RNA are capable of activating the *BCL11B* oncogene in these cases of AML. The breakpoints in AML #2301 and #7073 are 10 kb and 600 kb away from the transcriptional start site of *BCL11B*. There is no obvious reason why *BCL11B* specifically, and not the other genes, would be activated. Interestingly, however, the rearranged *BCL11B* allele in both AML #2301 and AML #7073 is juxtaposed to recently described super-enhancers, which have been shown to act as key oncogenic drivers.[38,39] These putative super-enhancers are present in the cell line MOLM-1 and seem to be linked to *ARID1B* on chromosome 6 (#2301) and *GSDMC* on chromosome 8 (#7073) and may be responsible for increased *BCL11B* expression. The selective overexpression of *BCL11B* may give the cells a specific advantage, whereas the other genes located on 14q32 would not.

Interestingly, BCL11B protein appeared to be expressed in additional cases of primary AML that do not carry *BCL11B* translocations. In these AML cases other mutations may be present or *BCL11B* may be activated by other means.

The 14q32 region, including *BCL11B*, is subject to translocation in T-cell acute lymphoblastic leukemia and acute mixed lineage leukemia.[23,33-37] In fact, the involvement of 14q32 translocations and *BCL11B* in AML has been debated.[40] However, the immunophenotyping and molecular analyses of the AML samples with *BCL11B* translocations described here showed that these leukemias have a biphenotypic immunophenotype, but also all carry a common AML-associated *FLT3* ITD mutation. These leukemias do, therefore, share a characteristic genetic feature with AML.

BCL11B is expressed in T-lymphocytes and T-cell leukemias and is a pivotal regulator of number of genes related to T-cell proliferation and differentiation such as *IL2*, *NF-kappaB*, *TCRβ* and *p21*.[41-45] It was shown recently that the expression of *BCL11B* in T-cell lines resulted in markedly increased apoptosis resistance following treatment with radiomimetic drugs accompanied by a cell cycle delay caused by accumulation of cells at G1.[46] We exam-

ined the consequences of *Bcl11b* overexpression on proliferation and differentiation in a mouse myeloid 32D(GCSF-R) cell line model. The 32D(GCSF-R) cells, expressing full-length murine *Bcl11b* cDNA, had a consistently decreased proliferation rate compared to cells expressing the empty vector or to the parental untransfected cells. Upon stimulation with granulocyte colony-stimulating factor, 32D(GCSF-R) cells overexpressing Bcl11b showed less maturation towards granulocytes compared to cells expressing empty vector, giving supporting evidence that BCL11B is partially blocking or delaying differentiation in 32D(GCSF-R) cells. The decreased proliferation rate in BCL11B-expressing cells suggests that a proliferative mutation, such as a *FLT3*-ITD, may indeed be required for full-leukemic transformation.

In conclusion, we have shown that *BCL11B* is involved in 14q32 translocations with different putative chromosomal partners in well-characterized AML cases using high-throughput genome-wide genotyping, cytogenetics and gene expression profiling. In these translocations, full-length BCL11B is highly expressed concomitantly with T-cell specific markers. We speculate that due to the translocations, *BCL11B* expression is influenced by active transcriptional elements on the partner chromosomes resulting in high *BCL11B* expression and consequently T-cell associated genes. The murine cell line 32D(GCSF-R) overexpressing BCL11B shows decreased proliferation and partial delayed differentiation, which provides evidence that *BCL11B* may have suppressive and disruptive effects on cell proliferation and differentiation of myeloid cells. Altogether, these analyses indicate that *BCL11B* is a putative oncogene in AML with and possibly without aberrations involving 14q32.

Acknowledgments

We are indebted to the participants of the HOVON clinical trials, and our colleagues from the stem cell transplantation and molecular diagnostics laboratories who provided, collected and analyzed AML cell samples, respectively.

Funding

This work was supported by grants from the Dutch Cancer Society (Koningin Wilhelmina Fonds) and performed within the framework of CTMM, the Center for Translational Molecular Medicine (Leukemia BioCHIP project; grant 030-102).

Authorship and Disclosures

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Burnett A, Wetzler M, Lowenberg B. Therapeutic advances in acute myeloid leukemia. J Clin Oncol. 2011;29(5):487-94.
2. Marcucci G, Haferlach T, Dohner H. Molecular genetics of adult acute myeloid leukemia: prognostic and therapeutic implications. J Clin Oncol. 2011;29(5):475-86.
3. Goyama S, Mulloy JC. Molecular pathogenesis of core binding factor leukemia: current knowledge and future prospects. Int J Hematol. 2011;94(2):126-33.
4. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature. 2007;446(7137):758-64.
5. Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, Chen K, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature. 2008;456(7218):66-72.
6. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A, et al. Mutation in TET2 in myeloid cancers. N Engl J Med. 2009;360(22):2289-301.
7. Langemeijer SM, Kuiper RP, Berends M, Knops R, Aslanyan MG, Massop M, et al. Acquired mutations in TET2 are common in myelodysplastic syndromes. Nat Genet. 2009;41(7):838-42.
8. Gelsi-Boyer V, Trouplin V, Adelaide J, Bonansea J, Cervera N, Carbuccia N, et al. Mutations of polycomb-associated gene ASXL1 in myelodysplastic syndromes and chronic myelomonocytic leukaemia. Br J Haematol. 2009;145(6):788-800.
9. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, Chen K, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009;361(11):1058-66.
10. Ernst T, Chase AJ, Score J, Hidalgo-Curtis CE, Bryant C, Jones AV, et al. Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. Nat Genet. 2010;42(8):722-6.
11. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, et al. DNMT3A mutations in acute myeloid leukemia. N Engl J Med. 2010;363(25):2424-33.
12. Mullighan CG, Flotho C, Downing JR. Genomic assessment of pediatric acute leukemia. Cancer J. 2005;11(4):268-82.
13. Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. N Engl J Med. 2009;360(5):470-80.
14. Bullinger L, Frohling S. Array-based cytogenetic approaches in acute myeloid leukemia: clinical impact and biological insights. Semin Oncol. 2012;39(1):37-46.
15. Raghavan M, Lillington DM, Skoulakis S, Debernardi S, Chaplin T, Foot NJ, et al. Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. Cancer Res. 2005;65(2):375-8.
16. Gorletta TA, Gasparini P, D'Elios MM, Trubia M, Pelicci PG, Di Fiore PP. Frequent loss of heterozygosity without loss of genetic material in acute myeloid leukemia with a normal karyotype. Genes Chromosomes Cancer. 2005;44(3):334-7.
17. Sanders MA, Verhaak RG, Geertsma-Kleinekoort WM, Abbas S, Horsman S, van der Spek PJ, et al. SNPExpress: integrated visualization of genome-wide genotypes, copy numbers and gene expression levels. BMC Genomics. 2008.25;9:41.
18. Liu P, Li P, Burke S. Critical roles of Bcl11b in T-cell development and maintenance of T-cell identity. Immunol Rev. 2010;238(1):138-49.
19. Verhaak RG, Wouters BJ, Erpelinck CA, Abbas S, Beverloo HB, Lugthart S, et al. Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. Haematologica. 2009;94(1):131-4.
20. Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med. 2004;350(16):1617-28.
21. Beekman R, Valkhof MG, Sanders MA, van Strien PM, Haanstra JR, Broeders L, et al. Sequential gain of mutations in severe congenital neutropenia progressing to acute myeloid leukemia. Blood. 2012;119(22):5071-7.
22. Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nat Methods. 2009;6(9):677-81.
23. Dong F, van Buitenen C, Pouwels K, Hoefsloot LH, Lowenberg B, Touw IP. Distinct cytoplasmic regions of the human granulocyte colony-stimulating factor receptor involved in induction of proliferation and maturation. Mol Cell Biol. 1993;13(12):7774-81.
24. Wouters BJ, Jorda MA, Keeshan K, Louwers I, Erpelinck-Verschueren CA, Tielemans D, et al. Distinct gene expression profiles of acute myeloid/T-lymphoid leukemia with silenced CEBPA and mutations in NOTCH1. Blood. 2007;110(10):3706-14.
25. Hollink IH, van den Heuvel-Eibrink MM, Arentsen-Peters ST, Pratcorona M, Abbas S, Kuipers JE, et al. NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid leukemia with a distinct HOX gene expression pattern. Blood. 2011;118(13):3645-56.
26. Liu P, Keller JR, Ortiz M, Tessarollo L, Rachel RA, Nakamura T, et al. Bcl11a is essential for normal lymphoid development. Nat Immunol. 2003;4(6):525-32.
27. Wakabayashi Y, Inoue J, Takahashi Y, Matsuki A, Kosugi-Okano H, Shinbo T, et al. Homozygous deletions and point mutations

of the Rit1/Bcl11b gene in gamma-ray induced mouse thymic lymphomas. Biochem Biophys Res Commun. 2003;301 (2):598-603.
28. Bernard OA, Busson-LeConiat M, Ballerini P, Mauchauffe M, Della Valle V, Monni R, et al. A new recurrent and specific cryptic translocation, t(5;14)(q35;q32), is associated with expression of the Hox11L2 gene in T acute lymphoblastic leukemia. Leukemia. 2001; 15(10):1495-504.
29. Berger R, Dastugue N, Busson M, Van Den Akker J, Perot C, Ballerini P, et al. t(5;14)/HOX11L2-positive T-cell acute lymphoblastic leukemia. A collaborative study of the Groupe Francais de Cytogenetique Hematologique (GFCH). Leukemia. 2003;17(9):1851-7.
30. Przybylski GK, Dik WA, Wanzeck J, Grabarczyk P, Majunke S, Martin-Subero JI, et al. Disruption of the BCL11B gene through inv(14)(q11.2q32.31) results in the expression of BCL11B-TRDC fusion transcripts and is associated with the absence of wild-type BCL11B transcripts in T-ALL. Leukemia. 2005;19(2):201-8.
31. De Keersmaecker K, Real PJ, Gatta GD, Palomero T, Sulis ML, Tosello V, et al. The TLX1 oncogene drives aneuploidy in T cell transformation. Nat Med. 2010;16(11):1321-7.
32. Gutierrez A, Kentsis A, Sanda T, Holmfeldt L, Chen SC, Zhang J, et al. The BCL11B tumor suppressor is mutated across the major molecular subtypes of T-cell acute lymphoblastic leukemia. Blood. 2011;118 (15):4169-73.
33. Batanian JR, Dunphy CH, Gale G, Havlioglu N. Is t(6;14) a non-random translocation in childhood acute mixed lineage leukemia? Cancer Genet Cytogenet. 1996;90(1):29-32.
34. Bezrookove V, van Zelderen-Bhola SL, Brink A, Szuhai K, Raap AK, Barge R, et al. A novel t(6;14)(q25-q27;q32) in acute myelocytic leukemia involves the BCL11B gene. Cancer Genet Cytogenet. 2004;149(1):72-6.
35. Georgy M, Yonescu R, Griffin CA, Batista DA. Acute mixed lineage leukemia and a t(6;14)(q25;q32) in two adults. Cancer Genet Cytogenet. 2008;185(1):28-31.
36. Hayashi Y, Pui CH, Behm FG, Fuchs AH, Raimondi SC, Kitchingman GR, et al. 14q32 translocations are associated with mixed-lineage expression in childhood acute leukemia. Blood. 1990;76(1):150-6.
37. Raimondi SC, Kalwinsky DK, Hayashi Y, Behm FG, Mirro J Jr, Williams DL. Cytogenetics of childhood acute nonlymphocytic leukemia. Cancer Genet Cytogenet. 1989;40(1):13-27.
38. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, et al. Super-enhancers in the control of cell identity and disease. Cell. 2013;155(4):934-47.
39. Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell. 2013;153(2):320-34.
40. MacLeod RA, Nagel S, Drexler HG. BCL11B rearrangements probably target T-cell neoplasia rather than acute myelocytic leukemia. Cancer Genet Cytogenet. 2004; 153(1):88-9.
41. Cismasiu VB, Adamo K, Gecewicz J, Duque J, Lin Q, Avram D. BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter. Oncogene. 2005;24(45):6753-64.
42. Wakabayashi Y, Watanabe H, Inoue J, Takeda N, Sakata J, Mishima Y, et al. Bcl11b is required for differentiation and survival of alphabeta T lymphocytes. Nat Immunol. 2003;4(6):533-9.
43. Cherrier T, Suzanne S, Redel L, Calao M, Marban C, Samah B, et al. p21(WAF1) gene promoter is epigenetically silenced by CTIP2 and SUV39H1. Oncogene. 2009;28(38): 3380-9.
44. Cismasiu VB, Duque J, Paskaleva E, Califano D, Ghanta S, Young HA, et al. BCL11B enhances TCR/CD28-triggered NF-kappaB activation through up-regulation of Cot kinase gene expression in T-lymphocytes. Biochem J. 2009;417(2):457-66.
45. Cismasiu VB, Ghanta S, Duque J, Albu DI, Chen HM, Kasturi R, et al. BCL11B participates in the activation of IL2 gene expression in CD4+ T lymphocytes. Blood. 2006; 108(8):2695-2702.
46. Grabarczyk P, Nahse V, Delin M, Przybylski G, Depke M, Hildebrandt P, et al. Increased expression of bcl11b leads to chemoresistance accompanied by G1 accumulation. PLoS One. 2010;5(9). pii: e12532.
47. Rockova V, Abbas S, Wouters BJ, Erpelinck CA, Beverloo HB, Delwel R, et al. Risk-stratification of intermediate-risk acute myeloid leukemia: integrative analysis of a multitude of gene mutation and expression markers. Blood. 2011;118(4):1069-76.
48. Pratcorona M, Abbas S, Sanders MA, Koenders JE, Kavelaars FG, Erpelinck-Verschueren CA, et al. Acquired mutations in ASXL1 in acute myeloid leukemia: prevalence and prognostic value. Haematologica. 2012;97(3):388-92.
49. Ribeiro AF, Pratcorona M, Erpelinck-Verschueren C, Rockova V, Sanders M, Abbas S, et al. Mutant DNMT3A: a marker of poor prognosis in acute myeloid leukemia. Blood. 2012;119(24):5824-31.