OVERVIEW

C-terminal BRE overexpression in 11q23-rearranged and t(8;16) acute myeloid leukemia is caused by intragenic transcription initiation

AE Marneth1, H K M Prange2, ASA Al Hinai3, SM Bergevoet1, N Tesi2, EM Janssen-Megens2, B Kim2, N Sharifi2, ML Yaspo4, J Kuster2, MA Sanders3, ECG Stoetman1, J Krijnenburg4, TCJM Arends-Peters5, CM Zwaan6, HG Stunnenberg2, MM van den Heuvel-Eibrink6,7, T Haferlach8, M Fornerod6, JH Jansen1, PJM Valk7, BA van der Reijden1,9 and JHA Martens2,9

Overexpression of the BRE (brain and reproductive organ-expressed) gene defines a distinct pediatric and adult acute myeloid leukemia (AML) subgroup. Here we identify a promoter enriched for active chromatin marks in BRE intron 4 causing strong biallelic expression of a previously unknown C-terminal BRE transcript. This transcript starts with BRE intron 4 sequences spliced to exon 5 and downstream sequences, and if translated might code for an N terminally truncated BRE protein. Remarkably, the new BRE transcript was highly expressed in over 50% of 11q23/KMT2A (lysine methyl transferase 2A)-rearranged and t(8;16)/KAT6A-CREBBP cases, while it was virtually absent from other AML subtypes and normal tissues. In gene reporter assays, the leukemia-specific fusion protein KMT2A-MLLT3 transactivated the intragenic BRE promoter. Further epigenome analyses revealed 97 additional intragenic promoter marks frequently bound by KMT2A in AML with C-terminal BRE expression. The corresponding genes may be part of a context-dependent KMT2A-MLLT3-driven oncogenic program, because they were higher expressed in this AML subtype compared with other groups. C-terminal BRE might be an important contributor to this program because in a case with relapsed AML, we observed an ins(11;2) fusing CHORDC1 to BRE at the region where intragenic transcription starts in KMT2A-rearranged and KAT6A-CREBBP AML.

Leukemia advance online publication, 24 October 2017; doi:10.1038/leu.2017.280

INTRODUCTION

Oncogene activation contributes to cancer onset and can be associated with treatment response. For instance, in acute myeloid leukemia (AML), high MECOM (MDS1-EVI1 complex) expression predicts poor outcome.1 Chromosome 3 rearrangements juxtaposing KMT2A, a frequent fusion partner of leukemia (AML), high MECOM associated with treatment response. For instance, in acute myeloid leukemia (AML), we reported that activation by KMT2A-MLLT3 may be context dependent. Earlier we lineage leukemia translocated to chromosome 3

…Continued on next page…

1Department of Laboratory Medicine, Laboratory of Hematology, Radboud Institute for Molecular Life Sciences (RIMLS), Radboud University Medical Center, Nijmegen, The Netherlands; 2Department of Molecular Biology, Faculty of Science, RIMLS, Radboud University, Nijmegen, The Netherlands; 3Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands; 4Max Planck Institute for Molecular Genetics, Berlin, Germany; 5Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands; 6Pediatric Oncology/Hematology, Erasmus University Medical Center-Sophia Children’s Hospital, Rotterdam, The Netherlands; 7Princess Maxima Center for Pediatric Oncology, Utrecht, The Netherlands and 8MLL Munich Leukemia Laboratory, Munich, Germany. Correspondence: Dr BA van der Reijden, Department of Laboratory Medicine, Laboratory of Hematology, Radboud Institute for Molecular Life Sciences (RIMLS), Radboud University Medical Center, Geert Grooteplein zuid 8, Nijmegen 6525 GA, The Netherlands; or Dr JHA Martens, Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences, Radboud University, Geert Grooteplein zuid 8, 6525 GA Nijmegen, The Netherlands. E-mail: Bert.vanderReijden@radboudumc.nl or J.Martens@ncmlls.ru.nl

9These authors contributed equally to this work.

Received 20 March 2017; revised 16 July 2017; accepted 10 August 2017; accepted article preview online 5 September 2017
RESULTS

The epigenome of human KMT2A-MLLT3 AML

To gain more insight into the epigenetic regulation of the MECOM and BRE loci in KMT2A-MLLT3 AML, we generated H3K4me1, H3K4me3 and H3K27ac ChIP-seq profiles from two primary KMT2A-MLLT3 AML samples (Figure 1a, one MECOM and one BRE overexpression (patients P1a and P2a; Table 1)). This revealed ~30 000 and ~26 000 active promoters (H3K27ac/H3K4me3) as well as ~40 500 and ~46 000 active enhancers (H2K27ac/H3K4me1) for P1a and P2a, respectively. The relative abundance and genomic distribution of active promoters and active enhancers was similar between P1a and P2a (Figure 1b and Supplementary Figure 1a). The common promoter and enhancer regions (>65% of active promoters/enhancers) showed comparable average ChIP-seq signal intensities (Figure 1b). In contrast, patient-specific active promoters and enhancers showed markedly higher average signals (Supplementary Figures 1B and C). Remarkably, for H3K4me3/H3K27ac peaks, only 50% overlapped with annotated (RefSeq hg19) promoters, suggesting that KMT2A-MLLT3 samples might harbor many alternate transcription start sites. Thus, the epigenomes of MECOM- and BRE-overexpressing KMT2A-MLLT3 AML samples are comparable, and also contain unique features.

Deposition of active histone marks at annotated promoters in MECOM-overexpressing KMT2A-MLLT3 AML

The sample with MECOM overexpression exhibited higher H3K4me3 signals than the sample with BRE overexpression at the two annotated MECOM promoter regions (Supplementary Figure 1D). Additionally, H3K27ac was only detected at the H3K4me3 peaks in P1a, while no signal was found in P2a, which is in line with the low MECOM expression in P2a. These observations were subsequently confirmed in additional primary KMT2A-MLLT3 AML samples: by H3K4me3 ChIP-seq \((n=2 \text{ MECOM, } n=5 \text{ BRE overexpression})\), RNA-seq \((n=2 \text{ MECOM, } n=2 \text{ BRE overexpression})\) and a H3K27ac ChIP-seq profile (BRE overexpression, P4; Supplementary Figure 1E). Taken together, these results show that MECOM overexpression is associated with increased deposition of active promoter marks at its two annotated promoters.

Intragenic promoter marks in BRE in KMT2A-MLLT3 AML

Subsequently, the BRE locus was examined. No differences were detected in active enhancers around the BRE locus between the two initial KMT2A-MLLT3 AML samples (not shown). Interestingly, apart from enrichment on the annotated promoter, the sample

<p>| Table 1. Characteristics of KMT2A-MLLT3 samples used for ChIP-seq and RNA-seq experiments |
|---------------------------------------------------------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Simple number</th>
<th>Study number</th>
<th>ChIP-seq/RNA-seq Study number</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1a</td>
<td>AML_MLLAF9_125</td>
<td>AML_MLLAF9_125</td>
</tr>
<tr>
<td>P2a</td>
<td>AML_MLLAF9_30</td>
<td>AML_MLLAF9_30</td>
</tr>
<tr>
<td>P3</td>
<td>5358</td>
<td>21</td>
</tr>
<tr>
<td>P4</td>
<td>11-030076</td>
<td>12-040773</td>
</tr>
<tr>
<td>P5</td>
<td>11-030076</td>
<td>12-040773</td>
</tr>
<tr>
<td>P6</td>
<td>11-030076</td>
<td>12-040773</td>
</tr>
<tr>
<td>P7</td>
<td>11-030076</td>
<td>12-040773</td>
</tr>
</tbody>
</table>

Abbreviations: BM, bone marrow; BRE, brain and reproductive organ-expressed; ChIP-seq, chromatin immunoprecipitation-sequencing; KMT2A, lysine methyl transferase 2A; MECOM, MDS1-EVI1 complex.
with BRE overexpression contained intragenic active promoter marks (H3K4me3/H3K27ac) in BRE intron 4 that were absent in the sample with normal BRE/high MECOM expression (Figure 1c).

These data were confirmed using ChIP-qPCRs on two KMT2A-MLLT3 cell lines and primary samples (Figure 1d and Supplementary Figure 2A). Analysis of all H3K4me3 ChIP-seq
profiles revealed that 5/6 samples with BRE overexpression harbored an intragenic H3K4me3 peak in BRE (the peak in P6 was not called, likely due to the higher background levels), whereas none of the three samples with MECOM overexpression did (Figure 1e and Supplementary Figures 2B and C). A second H3K27ac profile generated from a sample with BRE overexpression (P4) also showed a clear peak in the same region (not shown). These results show that the BRE intragenic active promoter marks were specific to KMT2A-MLLT3 samples with BRE overexpression.

Intragenic transcription activation causes high downstream BRE expression in KMT2A-MLLT3 AML

Next, we investigated whether the intragenic BRE promoter was accompanied by the expression of a specific transcript. Indeed, RNA-seq from sample P2a (BRE overexpression) showed expression of intronic sequences adjacent to the active promoter marks, ~3 kb upstream of exon 5 (Figure 1c). Downstream exons were expressed at a higher level compared with upstream exons. RNA-seq and RT-qPCR in a larger cohort showed that exons upstream of the intragenic promoter were equally expressed in all KMT2A-MLLT3 samples (Figures 1f and g). In contrast, downstream exons were five- to ninefold higher expressed in samples with BRE versus MECOM overexpression (Figure 1g). The start of the new BRE transcript and fusion of intragenic reads to exon 5 as determined by RNA-seq (Figure 1c) was confirmed by 5′-RACE (Figure 1h) followed by Sanger sequencing. The new transcript contained a ~100 bp region within intron 4 (exon 1b) fused to exon 5. As several samples harbored a heterozygous single-nucleotide polymorphism within exon 1b, which was expressed, we could determine that alternate BRE expression was biallelic (not shown). Besides exons 1b and 5, the novel BRE transcript contained exons 6–12 explaining high expression of these downstream exons (Figures 1c and 1f and not shown). Importantly, the new transcript was not detected by 5′-RACE in three KMT2A-MLLT3 samples with MECOM overexpression. Instead, the expected transcript containing BRE exons 1–7 was amplified (~870 bp; Figure 1h). Exon 5 codes for two methionines that potentially serve as translation start sites (Figure 1i), which, if used, might result in the formation of a 138 or 154 amino-acid N-terminal truncated BRE protein. In conclusion, these results show that BRE overexpression in KMT2A-MLLT3 AML is caused by biallelic intragenic BRE transcription activation from an alternate promoter in intron 4.

To study whether MECOM directly repressed the intragenic BRE promoter, we performed MECOM ChIP-qPCRs in a KMT2A-MLLT3 sample with high MECOM expression. The SPI1 enhancer region showed a clear enrichment for MECOM compared with control regions (Supplementary Figure 3A). MECOM binding was also found on the normal BRE promoter, yet no enrichment was detected at the intragenic BRE promoter (Supplementary Figure 3A). This suggests that MECOM does not directly repress the intragenic BRE promoter. Additionally, there is no indication for heterochromatin formation at the intragenic BRE promoter in samples with high MECOM expression, as the repressive promoter marks H3K27me3 and H3K9me3 were not detected (Supplementary Figure 3A).

Novel BRE transcript in KMT2A-rearranged and KAT6A-CREBBP AML

Our next goal was to study whether BRE intragenic transcription activation is unique to KMT2A-MLLT3 AML. We previously described samples with BRE overexpression without detectable rearrangements involving KMT2A.12 To determine whether BRE overexpression in these samples is caused by intragenic transcription activation, we performed 5′-RACE on two such samples (patients 2184 and 6948 from Valk et al21). Identical fragments were identified in these samples as in KMT2A-MLLT3 AML with BRE overexpression (Figure 1h), suggesting that expression of the new BRE transcript may occur independently of KMT2A rearrangements. We developed an RT-qPCR for detection of the new BRE transcript (Figure 1i) to determine its prevalence in AML in general (Figure 1j). The new BRE transcript was highly expressed in approximately half of the adult KMT2A-MLLT3 (13/24) as well as ~70% and ~50% of pediatric KMT2A-MLLT3 (13/19) and KMT2A-MLLT10 AMLs (7/15), respectively. As most KMT2A-rearranged AML samples have French–American–British M4 or M5 morphology, additional samples were screened for expression of the novel BRE transcript. Two of the 36 M4/M5 samples exhibited high expression of the new BRE transcript. Thus, morphological classification per se does not correlate with high expression of the new BRE transcript. The novel BRE transcript was only sporadically highly expressed in remaining AML samples (2/65) and was not found in other hematological malignancies (n = 16; Figure 1j).

As BRE overexpression was previously detected in one KAT6A-CREBBP AML sample,13 additional samples were screened for the new BRE transcript. Indeed, 3/7 adult and 10/13 pediatric KAT6A-CREBBP samples exhibited high new BRE transcript levels (Figure 1j). The new BRE transcript was hardly detected in normal primary hematopoietic cells (Figure 1j and 20 different human hematopoietic cell lines) and was not found in other hematological malignancies (n = 16; Figure 1j).

Figure 1. Intragenic transcription activation in BRE in primary KMT2A-MLLT3 samples with BRE overexpression. (a) Representation of the HOXA locus, a well-known KMT2A-MLLT3 target. H3K4me1, H3K27ac and H3K4me3 ChIP-seq profiles of two primary KMT2A-MLLT3 AML samples: P1a (MECOM overexpression) and P2a (BRE overexpression). (b) ChIP-seq signal intensity at active promoter loci (H3K4me3/H3K27ac, left) and active enhancer loci (H3K4me1/H3K27ac, right) for samples P1a and P2a. (c) ChIP-seq and RNA-seq results at the BRE locus showing an intragenic active promoter (H3K27ac/H3K4me3, gray box) and enhanced expression of adjacent and downstream sequences in P2a but not P1a. (d) H3K4me3/H3K27ac ChIP-qPCRs at the BRE intragenic active promoter (int AP) over background (Myoglobin) in P2b and P1b. Four different qPCRs were used. (e) Normalized H3K4me3 signal at the BRE intragenic active promoter in KMT2A-MLLT3 profiles (n = 6 BRE overexpression, n = 3 MECOM overexpression). (f) Expression of BRE downstream and upstream exons relative to the intragenic active promoter based on RNA-seq data from three samples with BRE overexpression (P2a, P2b and P5) and three samples with MECOM overexpression (P1a, P1b and P3). (g) Expression of BRE exons 2–3, 3–4, 5–6 and 10–11 (RT-qPCR) normalized to PBGD expression and a KMT2A-MLLT3 AML sample with MECOM overexpression (calibrator sample, P3). AML cell lines: NB4, OCI-AML3 and THP-1. (h) 5′-RACE using BRE exon 8 reverse primer for PCR. Samples 6955, 2184, 7416, 7178 and 6948 are from Valk et al.21 (i) Schematic overview of the most abundant BRE transcript in blood cells (NM_199191) and the new BRE transcript. The new BRE transcript contains two potential translation start sites in exon 5 that are in-frame with the protein encoded by transcript NM_199191. RT-qPCR primers and probe used in (j) are indicated. (j) Expression of the new BRE transcript (RT-qPCR) in indicated samples normalized to PBGD expression and a calibrator sample (P3). The novel BRE transcript was higher expressed in ‘Adult KMT2A-MLLT3’, ‘Pediatric KMT2A-MLLT3’, ‘Pediatric KMT2A-MLLT10’, ‘FAB M4 AML’, ‘Adult KAT6A-CREBBP’ and ‘Pediatric KAT6A-CREBBP’ samples compared with ‘remaining AML’ samples (P < 0.004). Mal, malignancies.
Figure 2. Genome-wide intragenic active promoter marks in BRE-overexpressing primary KMT2A-MLLT3 samples. (a) Overlap of intragenic active promoter peaks in KMT2A-MLLT3 samples with MECOM overexpression (P1a) and BRE overexpression (P2a, left). Overlap of intragenic H3K4me3 peaks including an additional five samples with high BRE expression and two samples with high MECOM expression (right). (b) Average H3K4me3 signal at 98 intragenic active promoter sites for nine KMT2A-MLLT3 AML samples (BRE overexpression n = 6, MECOM overexpression n = 3). (c) Expression of 90 genes containing 98 intragenic active promoter marks, in various subtypes of AML. (d) Reporter assays showing Renilla normalized Firefly luciferase signal from a control vector (pGL3 basic), from the Hoxa7 promoter and from three fragments of the new BRE promoter upon increasing concentrations of KMT2A-MLLT3. Mean ± s.d. (e) Average KMT2A signal at 98 intragenic promoter mark regions (b and Supplementary Figure S3B) in a KMT2A-MLLT3 sample with BRE overexpression (P4). (f) TF family motifs enriched under 98 intragenic active promoter marks. (g) Differential expression of TFs belonging to enriched motif families (AP-2, C2H2 and ETS, panel f) in BRE-overexpressing versus MECOM-overexpressing samples. Red dots: (n = 168) fold change > 2; orange dots: (n = 23) fold change > 2, Bonferroni-adjusted P-value < 0.05; green dots: (n = 35) fold change > 4, Bonferroni-adjusted P-value < 0.01. AP, active promoter; int BRE prom, intragenic BRE promoter; Sig., significant; TF, transcription factor.
Genome-wide intragenic promoter marks in BRE-overexpressing KMT2A-MLLT3 AML

To examine whether intragenic transcription activation is unique for BRE, we performed genome-wide analyses revealing ~7000 H3K4me3/H3K27ac intragenic peaks that were specific for cases with BRE overexpression (Figure 2a). To avoid false positives, we narrowed down the group of specific intragenic active promoters among others by including only those with a marked difference in the average H3K4me3 signal (Figure 2b and Supplementary Figures 3B and C). This analysis yielded intragenic active promoter marks that were unique to samples with BRE overexpression in 98 regions, residing in 90 genes (Figure 2b, Supplementary Figure 3C and Supplementary Table 1). For MECOM-overexpressing samples, similar filtering revealed eight (Supplementary Figure 3B and Supplementary Table 1) specific intragenic active promoter regions. In conclusion, BRE-overexpressing KMT2A-MLLT3 AML is characterized by specific genome-wide intragenic active promoters.

Genes with intragenic active promoter marks are activated in BRE-overexpressing KMT2A-MLLT3 AML

We analyzed the expression levels of the 90 genes with intragenic promoters in a large adult AML cohort.24 This showed significantly higher expression of these genes in BRE-overexpressing samples as opposed to 11q23/KMT2A-rearranged MECOM-overexpressing samples ($P = 0.015$) and a trend towards higher expression compared with other genetically defined patient subsets (Figure 2c). Remarkably, RNA-seq profiles described above indicated that only five (BRE, PAN3, SMYD3, MS4A7 and TGM5) of the 90 genes with intragenic promoters had $\geq 2$-fold increased expression. In conclusion, the presence of intragenic active promoters is characteristic of the AML subtype

Figure 3. A novel t(11;2) translocation fuses CHORDC1 exon 5 to BRE exon 5. (a) Sanger sequencing of the CHORDC1-BRE fusion transcript. (b) Karyotype of AML at the time of relapse: 46,XY;ins(11;2)(q14;p13p23)[10]. (c) Schematic representation of the full-length and predicted new BRE proteins, CHORDC1 and the CHORDC1-BRE fusion protein (UEV: ubiquitin E2 variant, CHORD: cysteine- and histidine-rich, CS: interaction module named after CHORD-containing proteins and SGT1). Amino-acid numbers refer to proteins encoded by BRE transcript 3 (NM_199191) and CHORDC1 transcript 1 (NM_012124).
expression of downstream versus upstream exons relative to their intragenic active promoter mark (not shown). These data suggest that intragenic active promoter marks are found in actively transcribed genes, but do not necessarily cause transcription initiation at that position. Alternatively, intragenic transcription initiation might occur, but the expression or stability of downstream sequences may be so low that it does not result in a significant increase in read count.

The high expression of downstream SMYD3, PAN3 and TGM5 sequences in KMT2A-MLLT3 samples with versus without BRE overexpression was confirmed by RT-qPCR, but this was not found for KAT6A-CREBBP AML samples (Supplementary Figure 3D). These results could imply that KAT6A-CREBBP and KMT2A-MLLT3 AML do not harbor a similar genome-wide intragenic active promoter signature.

KMT2A-MLLT3 transactivates the intragenic BRE promoter

To determine whether KMT2A-MLLT3 contributes to intragenic BRE transcription, gene reporter assays were performed. Both the Hoxa7 promoter (positive control) and BRE fragments containing sequences from +17 to +410 bp relative to the novel transcription initiation might occur, but the expression or stability of downstream sequences may be so low that it does not result in a significant increase in read count.

The expression of downstream SMYD3, PAN3 and TGM5 sequences in KMT2A-MLLT3 samples with versus without BRE overexpression was confirmed by RT-qPCR, but this was not found for KAT6A-CREBBP AML samples (Supplementary Figure 3D). These results could imply that KAT6A-CREBBP and KMT2A-MLLT3 AML do not harbor a similar genome-wide intragenic active promoter signature.

KMT2A-MLLT3 transactivates the intragenic BRE promoter

To determine whether KMT2A-MLLT3 contributes to intragenic BRE transcription, gene reporter assays were performed. Both the Hoxa7 promoter (positive control) and BRE fragments containing sequences from +17 to +410 bp relative to the novel transcription start site were transactivated in a dose-dependent manner by KMT2A-MLLT3, whereas a fragment containing the region −133 to +17 bp was not transactivated (Figure 2d). To determine whether KMT2A or KMT2A-fusion proteins regulate the 98 intragenic active promoter marks, we analyzed three KMT2A ChiP-seq profiles,25 one from a primary KMT2A-MLLT3 sample with BRE overexpression and two from the KMT2A-rearranged cell lines THP-1 and MV4-11 (Supplementary Figure 4A). KMT2A signals were enriched at the 98 intragenic active promoter marks in all three samples (Figure 2e and Supplementary Figure 4B). The level of KMT2A enrichment at the 98 intragenic promoter marks was similar to KMT2A enrichment in previously defined25 KMT2A-fusion binding sites in THP-1 and MV4-11 cells (Supplementary Figures 4C and D). These data suggest that KMT2A/KMT2A-fusion proteins bind to the intragenic active promoter marks to contribute to high expression of the respective genes.

AP-2, C2H2-Znf and ETS TF motif enrichment at intragenic promoter marks

As only part of the KMT2A-MLLT3 AML samples have BRE overexpression, other factors besides KMT2A-MLLT3 might be involved in expression of the new BRE transcript. To identify candidates that may bind intragenic active promoter marks, we performed transcription factor (TF) motif analysis. This showed enrichment of motifs for AP-2, C2H2-Znf and ETS TF families (Figure 2f). RNA-seq data from two BRE- and two MECOM-overexpressing samples revealed significant differential expression of 58 TFs belonging to these families (Figure 2g). Only two TFs were higher expressed in BRE compared with MECOM-overexpressing samples (PLAG1 and ZNF595). Possibly, PLAG1 and ZNF595 are involved in enhancing transcription of the 90 genes with intragenic promoters.

A novel ins(11;2) fuses CHORDC1 to BRE exon 5

As described above, the BRE gene is actively transcribed from an intragenic promoter preceding exon 5. To find other changes in BRE in AML, we screened an RNA-seq library (32 paired diagnosis-relapse samples) for alterations in the BRE gene. Interestingly, an in-frame fusion transcript between CHORDC1 exon 5 (chromosome 11) and BRE exon 5 (chromosome 2) was found in a relapsed AML case. This fusion was confirmed by RT-PCR followed by Sanger sequencing, and an insertion of chromosome 2 into 11 was confirmed by karyotyping (Figures 3a and b). Heterozygous CHORDC1 (Morgania) inactivation in mice results in spontaneous atypical chronic myeloid leukemia.26 The CHORDC1 fusion to BRE exon 5 is remarkable because the alternate BRE transcript starts with exon 1b also fused to exon 5. This means that if the CHORDC1-BRE and alternate BRE transcripts would be translated, they might contain a similar C-terminal part of the BRE protein (Figure 3c). BRE and CHORDC1 were equally expressed in the diagnosis and relapse samples (stranded RNA-seq data, not shown). The CHORDC1-BRE fusion was not detected at diagnosis nor in 200 randomly selected de novo AML cases (not shown), indicating that this fusion is rare in de novo AML and may be associated with relapsed AML. Thus, leukemia-specific expression of the C terminus of BRE may be caused by intragenic transcription activation or chromosomal insertions.

DISCUSSION

Here, we report abnormal expression of C-terminal BRE sequences specific for AML, caused by distinct, non-random events, by a chromosomal insertion and by an intragenic transcription initiation in 11q23/KMT2A-rearranged and t(8;16)/KAT6A-CREBBP AML. The intragenic transcriptional activation cannot be explained by recurrent chromosomal abnormalities affecting the BRE locus itself, because the alternate transcript was expressed from both alleles. Instead, the oncofusion protein KMT2A-MLLT3 could be involved in intragenic BRE activation. In KMT2A-MLLT3 AML harboring the intragenic BRE promoter, we observed active intragenic promoter marks in 89 additional genes. KMT2A/KMT2A-fusion protein enrichment was observed at these promoter marks and the corresponding genes were overexpressed in KMT2A-rearranged AML with high BRE expression. Moreover, KMT2A-MLLT3 activated the intragenic BRE promoter in reporter assays. These findings suggest that KMT2A-MLLT3 contributes to alternate transcription initiation in BRE. It should, however, be noted that especially in adult AML a significant number of samples with these two fusion oncoproteins lack high expression of the new BRE transcript. Thus, intragenic BRE activation may be dependent on additional factors. Motif analyses of KMT2A-MLLT3 AML showed an overrepresentation of AP-2, C2H2-Znf, and ETS TF binding sites at intragenic promoter marks with C2H2-Znf TFs (Supplementary Figure 4) and KAT6A-CREBBP and KMT2A-MLLT3 TFs (Figure 2g) being highly expressed in BRE overexpression cases. It would be interesting to test the contribution of especially PLAG1 to alternate BRE expression, as PLAG1 has been implicated in AML development.27,28

Earlier we reported that patients with BRE overexpression have a superior outcome compared with MECOM-overexpressing patients. The unique 5’ end of the novel BRE transcript allows for reliable detection by RT-qPCR enabling an optimal recognition of both subtypes. Importantly, the association of alternate BRE expression with good clinical outcome appears to be restricted to KMT2A-rearranged AML, as patients with t(8;16)/KAT6A-CREBBP have a very dismal outcome in general.

KMT2A-MLLT3 and KAT6A-CREBBP may contribute to AML pathogenesis by altering gene expression.29–34 We observed that the genes with active intragenic promoter marks were bound by KMT2A/KMT2A-fusion proteins, and, indeed, the respective genes showed highest expression compared with other AML subtypes. These genes might represent a newly identified oncogenic program. It is tempting to speculate that alternate BRE contributes to this oncogenic program for various reasons. Abnormal BRE expression is exclusively observed in AML and caused by distinct mechanisms. The novel BRE transcript contains two putative translation start sites in exon 5, which if used might possibly lead to N terminally truncated BRE isoforms. The in-frame CHORDC1-BRE fusion includes a similar C-terminal part of BRE and Chordc1 functions as a tumor suppressor in mice.26 BRE is required for DNA repair and cell cycle regulation.14,17,35–38 Another important constituent of these complexes is BRCC3. Mutations in BRCC3 and other components of these two
complexes have recently been observed in myeloid malignancies, including AML. BRE overexpression is also found in other malignancies and correlates with higher grade tumors. In addition to AML, BRE has prognostic impact in breast cancer and a translocation including downstream BRE sequences has been observed in adrenocortical carcinoma. We conclude that it will be interesting to determine whether leukemia-specific alternate BRE transcripts are translated into protein and contribute to disease pathogenesis.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Dr Robert Slayn for kindly sharing the pGL3-Hoxa7 promoter plasmid. This work was supported by the Radboud University Medical Center for Oncology, the Dutch Cancer Foundation (KUN 2011-4937) and the European Union’s Seventh Framework Programme (FP/2007-2013, 282510-BLUEPRINT).

AUTHOR CONTRIBUTIONS

AEM, KHMP, BAR and JHAM coordinated research and wrote the manuscript. AEM, KHMP and SMB conducted and analyzed most experiments. KHMP, EMJM, AEM, KHMP, BAR and JHAM coordinated research and wrote the manuscript. ASAAH, ECGS, JK, MAS and PJMV detected CHORDC1-BK, NS, MLY, JK, HGS and JHAM contributed to ChIP-seq and RNA-sequencing of AEM, KHMP and SMB conducted and analyzed most experiments. KHMP, EMJM, AEM, KHMP, BAR and JHAM coordinated research and wrote the manuscript.

REFERENCES


Supplementary Information accompanies this paper on the Leukemia website (http://www.nature.com/leu)