ORIGINAL ARTICLE

MLL-AF9 and MLL-AF4 oncofusion proteins bind a distinct enhancer repertoire and target the RUNX1 program in 11q23 acute myeloid leukemia

KHM Prange1, A Mandoli1,3, T Kuznetsova1,3, S-Y Wang1, AM Sotoca1, AE Marneth2, BA van der Reijden2, HG Stunnenberg1 and JHA Martens1

INTRODUCTION

Mixed lineage leukemia (MLL), encoded by the lysine (K)-specific methyl transferase 2A (KMT2A) gene at 11q23, is a H3K4me3-depositing protein active during early development. MLL rearranged (MLLr) leukemias are responsible for about 10% of all acute lymphoblastic and myeloid leukemias (ALL/AML).1 The N-terminal CXXC-domain containing DNA-binding part of the MLL gene can, by genomic translocation, be fused to >60 different fusion partners.2 The two most common fusion partners, AF9 (MLLT3) and AF4 (AFF1), are found in 30% of MLLr AML and 66% of MLLr ALL cases, respectively.3

The most prevalent fusion partners (including AF9 and AF4) are present in the super elongation complex (SEC), therefore leading to a general MLLr mechanism of action despite being composed of different proteins.4,5 The SEC normally binds to RNA polymerase II and facilitates transcriptional elongation. In MLLr leukemias, the SEC is tethered to the DNA-binding domain of MLL via the fusion partner, leading to aberrant transcription of MLL target genes. Next to that, AF9 as well as other fusion partners such as ENL, AF10 and AF17 are present in the DOT1L complex (DOTCOM). This deposits H3K79me2 on actively transcribed genes, leading to aberrant H3K79me2 deposition in these subsets of MLLr-induced leukemias.4–6 As such, a promising avenue for treatment of MLLr-induced leukemia is inhibition of DOT1L.7,8

Furthermore, AMLs in general and MLLr leukemia in particular have been shown to be sensitive to inhibition of BET family proteins such as BRD4, which regulate transcription elongation via P-TEFb at promoters and enhancers.9–12 MLL fusion proteins have also been shown to interfere with RUNX113–15 and modulate PU.1 via its distal regulatory elements.16 Moreover, MLLr and other AMLs are sensitive to inhibition of mediator kinases,17 linking modulation of distal regulatory elements to execution of the MLLr leukemic program.

Thus far, genome-wide maps of MLL binding are only available in mouse models of AML,7 human ALL20–22 and of MLL-AF6 in the human AML cell line ML-2.23 No reports to date have described genome-wide MLL-AF9 and -AF4 binding in human AML.

Here we set out to investigate the molecular mechanisms and targets of MLLr-induced AML. For this, we characterized the genome-wide binding, epigenetic signature and gene expression program of wild-type (WT) MLL, MLL-AF9 and MLL-AF4 in human AML cells. We show that MLL fusion proteins bind in a ‘broad’ mode elongating over the gene body as well as in a ‘sharp’ mode stalled on the transcription start site (TSS), in addition to non-genic elements, such as distal enhancers. We show that MLL-AF9 and MLL-AF4 share only a subset of target genes, yet show enrichment for the same pathways in both the shared as well as the MLL-AF9- and MLL-AF4-specific gene sets. These target genes are marked by H3K79me2, H3K4me3 and H3K27ac enrichment as well as by RUNX1 occupancy and constitute a mixture of CD34+ and monocyte-expressed gene sets. Together these results suggest that, in MLLr AML, the RUNX1-mediated progenitor to monocyte differentiation program is deregulated.

RESULTS

MLL fusions and WT MLL show a broad and sharp binding mode

To investigate the binding sites of MLL-AF9, we used a substrain of THP-1 cells that express both WT MLL as well as MLL-AF9 but not fusion partners.2 The two most common fusion partners, AF9 (MLLT3) and AF4 (AFF1), are found in 30% of MLLr AML and 66% of MLLr ALL cases, respectively.3

The most prevalent fusion partners (including AF9 and AF4) are present in the super elongation complex (SEC), therefore leading to a general MLLr mechanism of action despite being composed of different proteins.4,5 The SEC normally binds to RNA polymerase II and facilitates transcriptional elongation. In MLLr leukemias, the SEC is tethered to the DNA-binding domain of MLL via the fusion partner, leading to aberrant transcription of MLL target genes. Next to that, AF9 as well as other fusion partners such as ENL, AF10 and AF17 are present in the DOT1L complex (DOTCOM). This deposits H3K79me2 on actively transcribed genes, leading to aberrant H3K79me2 deposition in these subsets of MLLr-induced leukemias.4–6 As such, a promising avenue for treatment of MLLr-induced leukemia is inhibition of DOT1L.7,8

Furthermore, AMLs in general and MLLr leukemia in particular have been shown to be sensitive to inhibition of BET family proteins such as BRD4, which regulate transcription elongation via P-TEFb at promoters and enhancers.9–12 MLL fusion proteins have also been shown to interfere with RUNX113–15 and modulate PU.1 via its distal regulatory elements.16 Moreover, MLLr and other AMLs are sensitive to inhibition of mediator kinases,17 linking modulation of distal regulatory elements to execution of the MLLr leukemic program.

Thus far, genome-wide maps of MLL binding are only available in mouse models of AML7 human ALL20–22 and of MLL-AF6 in the human AML cell line ML-2.23 No reports to date have described genome-wide MLL-AF9 and -AF4 binding in human AML.

Here we set out to investigate the molecular mechanisms and targets of MLLr-induced AML. For this, we characterized the genome-wide binding, epigenetic signature and gene expression program of wild-type (WT) MLL, MLL-AF9 and MLL-AF4 in human AML cells. We show that MLL fusion proteins bind in a ‘broad’ mode elongating over the gene body as well as in a ‘sharp’ mode stalled on the transcription start site (TSS), in addition to non-genic elements, such as distal enhancers. We show that MLL-AF9 and MLL-AF4 share only a subset of target genes, yet show enrichment for the same pathways in both the shared as well as the MLL-AF9- and MLL-AF4-specific gene sets. These target genes are marked by H3K79me2, H3K4me3 and H3K27ac enrichment as well as by RUNX1 occupancy and constitute a mixture of CD34+ and monocyte-expressed gene sets. Together these results suggest that, in MLLr AML, the RUNX1-mediated progenitor to monocyte differentiation program is deregulated.

RESULTS

MLL fusions and WT MLL show a broad and sharp binding mode

To investigate the binding sites of MLL-AF9, we used a substrain of THP-1 cells that express both WT MLL as well as MLL-AF9 but not fusion partners.2 The two most common fusion partners, AF9 (MLLT3) and AF4 (AFF1), are found in 30% of MLLr AML and 66% of MLLr ALL cases, respectively.3

The most prevalent fusion partners (including AF9 and AF4) are present in the super elongation complex (SEC), therefore leading to a general MLLr mechanism of action despite being composed of different proteins.4,5 The SEC normally binds to RNA polymerase II and facilitates transcriptional elongation. In MLLr leukemias, the SEC is tethered to the DNA-binding domain of MLL via the fusion partner, leading to aberrant transcription of MLL target genes. Next to that, AF9 as well as other fusion partners such as ENL, AF10 and AF17 are present in the DOT1L complex (DOTCOM). This deposits H3K79me2 on actively transcribed genes, leading to aberrant H3K79me2 deposition in these subsets of MLLr-induced leukemias.4–6 As such, a promising avenue for treatment of MLLr-induced leukemia is inhibition of DOT1L.7,8

Furthermore, AMLs in general and MLLr leukemia in particular have been shown to be sensitive to inhibition of BET family proteins such as BRD4, which regulate transcription elongation via P-TEFb at promoters and enhancers.9–12 MLL fusion proteins have also been shown to interfere with RUNX113–15 and modulate PU.1 via its distal regulatory elements.16 Moreover, MLLr and other AMLs are sensitive to inhibition of mediator kinases,17 linking modulation of distal regulatory elements to execution of the MLLr leukemic program.

Thus far, genome-wide maps of MLL binding are only available in mouse models of AML7 human ALL20–22 and of MLL-AF6 in the human AML cell line ML-2.23 No reports to date have described genome-wide MLL-AF9 and -AF4 binding in human AML.

Here we set out to investigate the molecular mechanisms and targets of MLLr-induced AML. For this, we characterized the genome-wide binding, epigenetic signature and gene expression program of wild-type (WT) MLL, MLL-AF9 and MLL-AF4 in human AML cells. We show that MLL fusion proteins bind in a ‘broad’ mode elongating over the gene body as well as in a ‘sharp’ mode stalled on the transcription start site (TSS), in addition to non-genic elements, such as distal enhancers. We show that MLL-AF9 and MLL-AF4 share only a subset of target genes, yet show enrichment for the same pathways in both the shared as well as the MLL-AF9- and MLL-AF4-specific gene sets. These target genes are marked by H3K79me2, H3K4me3 and H3K27ac enrichment as well as by RUNX1 occupancy and constitute a mixture of CD34+ and monocyte-expressed gene sets. Together these results suggest that, in MLLr AML, the RUNX1-mediated progenitor to monocyte differentiation program is deregulated.

RESULTS

MLL fusions and WT MLL show a broad and sharp binding mode

To investigate the binding sites of MLL-AF9, we used a substrain of THP-1 cells that express both WT MLL as well as MLL-AF9 but not
WT AF9 (Figures 1a and b, Supplementary Figure S1A). Using antibodies against the N-terminus of MLL-1 (ab1542/ab1547) and the C-terminus of AF9 (ab1327/ab1474) (Figure 1a) in chromatin immunoprecipitation (ChIP)-qPCR (quantitative PCR) showed enrichment for the canonical HOXA MLL fusion target region in THP-1 cells (Figure 1c), while HOXA7 and MEIS1 were, in contrast to expectations,24–26 not enriched. Western analysis of MLL and AF9 using our antibodies yielded no results (data not shown), suggesting that our antibodies were of ‘chIP-seq grade’ but not of ‘western grade’. Subsequently, we performed ChIP-seq experiments using MLL and AF9 antibodies in THP-1 MLLr cells that confirmed enrichment on the HOXA locus for both antibodies (Figure 1d, top). Interestingly, we observed that both MLL and AF9 show not only ‘broad mode’-enriched regions elongating over gene bodies (Figure 1d, middle), but also ‘sharp mode’-enriched regions on target promoters and enhancers (Figure 1d, bottom). Using MACS2 for defining sharp peaks and HOMER for defining broad regions (see Materials and methods), we identified 16,099 unfiltered MLL occupied regions in THP-1 cells (Supplementary Table S1). Of these, 8217 were of ‘broad mode’ (mean length ~12 kb) and 7882 of ‘sharp mode’ (mean length ~4.6 kb) (Figure 1e, left). Analysis of the genomic distribution revealed that the ‘broad mode’ peaks cover more TSS regions (Figure 1e, right), while the ‘sharp mode’ peaks seem to occur in intergenic regions more often.

As AF9 was not expressed from its endogenous locus in this substrain of THP-1 cells (Figure 1b, Supplementary Figure S1A), we defined MLL-AF9-binding peaks as those MLL-binding sites that show a high AF9 signal, and MLL WT-binding events as those that show a low AF9 signal (Figure 1f). This distilled our list of MLL-occupied regions down to 1613 high-confidence MLL-AF9-fusion-binding sites, including known AML and MLLr targets such as HOXA9, CDK6, MYB, MYC, JMJD1C, FOXP2, FLI1, RUNX1, PBX3, BCL2 and BRD4, as well as 439 high-confidence MLL WT-binding sites (Supplementary Table S1). In all, 84% of these MLL-AF9-binding sites are ‘broad mode’, versus 58% of WT MLL-binding sites (Figure 1g, left), and they occupy more TSS regions as compared with WT MLL (Figure 1g, right).

As MEIS1 occupancy and expression is near universal for all MLL-fusion-induced AML and ALL,27,28 we investigated the MEIS1 locus in THP-1 cells using our genome-wide data (Supplementary Figure S1B). This corroborated our ChIP-qPCR finding that MEIS1 is not bound by MLL-AF9 and not expressed in THP-1 cells, as it is marked only by H3K4me3 at its promoter but not by MLL, AF9, H3K27ac or H3K79me2. MEIS1 expression in MLL-fusion-induced AML has been shown to be especially important for the initial pathogenic signature of affected cells, compared with WT MLL (Figure 1h, Supplementary Table S2), while the NR family was depleted over background (Figure 2e, bottom), often involved in translocation events leading to myeloproliferative disorder.34 Motif analysis35 of MLL-AF9 targets revealed enrichment of the ETS, AP2 and C2H2-ZF families, while the POU family was depleted over background (Figure 2e, top left). WT MLL target genes were enriched for motifs recognized by C2H2-ZF and ETS families, while the NR family was depleted over background (Figure 2e, top right). Direct comparison of MLL-AF9 and MLL WT targets

Figure 1. Genome-wide binding patterns of MLL and AF9 in THP-1 cells. (a) Schematic representation of MLL, AF9 and MLL-AF9. Antibody binding locations are indicated with dotted lines, primer regions used in panel (b) with a filled line. (b) Reverse transcriptase–qPCR experiments (n = 5) in THP-1 cells with primers against the C and N termini of MLL and AF9 normalized to GAPDH. The N-terminus of AF9 is not expressed, indicating that there is no WT expression of AF9 in this cell line. ***P < 0.001 (Welch’s t-test). (c) ChIP-qPCR experiments using two anti-MLL-1 and two anti-AF9 antibodies in THP-1 cells and primers for HOXA7, 9, 10 and MEIS1. (d) ChIP-seq overview of MLL and AF9 binding at the HOXA, ZEB2 and CDKN2C loci in THP-1 cells. (e) Classification of MLL and AF9 binding events in ‘broad’ and ‘sharp’ modes. Left: boxplot showing dispersion of peak lengths. Right: barplot showing genomic distributions. (f) Classification of MLL-AF9 and MLL WT binding events. Average profiles showing ChIP-seq signal intensities for MLL-AF9 and MLL WT binding events in THP-1 cells. (g) Left: Distribution of MLL-AF9 and MLL WT binding events in the ‘broad’ and ‘sharp’ modes. Right: genomic distribution of MLL-AF9 and MLL WT binding events in the ‘broad’ and ‘sharp’ modes. (h) Venn diagram illustrating the overlap between our human (THP-1) MLL-AF9 AML targets, MLL-AF9 targets in a mouse LSC model (Bernt et al.) and human MLL-AF4 ALL targets (Guenther et al.).
revealed both gene sets as similar in terms of motif enrichment, except for the NR and AP-2 motif families, which were enriched in the MLL-AF9 target gene set (Figure 2e, bottom). This suggests that MLL-AF9-mediated deregulation of NR and AP-2 signaling might be involved in aberrant hematopoietic and immunological processes.
To identify common MLL fusion targets, we expanded our analysis by including an MLL-AF4-expressing AML. As both MLL-AF9 and MLL-AF4 are thought to bind to MLL target genes and are linked to aberrant elongation and transcription of their targets via the SEC, we set out to assess the subset of MLL target genes commonly bound by the fusion proteins. First, we created genome-wide binding profiles for MLL, AF4, H3K4me3, H3K79me2 and H3K27ac, as well as an RNA-seq expression profile in MV4-11 AML cells expressing the MLL-AF4 fusion protein (Figure 3a, Supplementary Figure S2A–B). As before, we divided the unfiltered MLL targets (28,656) into ‘broad mode’ (18,782) and ‘sharp mode’ (9874), and by rate of AF4 occupancy filtered them down to fusion (2560) and WT (828) binding events (Figure 3b, Supplementary Figures S2C and D, Supplementary Table S3), showing a similar distribution as in THP-1 cells (Supplementary Figure S2E). Expression of MLL-AF4 target genes (1722, identified by overlapping the high-confidence fusion-binding events with RefSeq hg19 genes) was significantly higher than MLL WT genes (308) (Figure 3c). Epigenetic signatures and pathway enrichments were also comparable to MLL-AF9 AMLs (Supplementary Figures S2F–H), suggesting that both MLL-AF9 and MLL-AF4 use similar...
molecular mechanisms to induce leukemia. We also found a similar overlap with the various other target sets as observed for MLL-AF9 (Supplementary Figure S2), Supplementary Table S2).

Comparing the MLL-AF9 and MLL-AF4 target gene sets (Supplementary Table S4) revealed that a significant (P = 1.98e-11) 29% (277) of MLL-AF9 target genes are also targeted by MLL-AF4 (Figure 3d, top), including known MLLr targets, such as BCL2, HOXA9, MYB and BRD4. In contrast, only 3% of WT MLL target genes found in THP-1 were targeted by WT MLL in MV4-11 (Figure 3d, bottom). Next, we analyzed the activity of these common MLL fusion and MLL WT target genes versus MLL-AF9 and -AF4-specific target genes. Common MLL-AF9 and MLL-AF4 target genes, as well as those specific for MLL-AF9 and -AF4 show a comparable level of H3K27ac and H3K4me3 (Figure 3e), while H3K79me2 is slightly lower in MLL-AF9 AMLs. The presence of H3K79me2 signal on MLL-AF4 target genes confirms the deposition of this histone tail mark also on MLL-AF4 targets in AML.

The gene expression levels as determined by RNA-seq are comparable for shared fusion targets in THP-1 and MV4-11, while the MLL-AF9- and -AF4-specific target genes are lower expressed in MV4-11 and THP-1, respectively, with median RPKM values of 23 and 17 for the MLL-AF9-specific target genes and 25 and 18 for the MLL-AF4-specific target genes, respectively (Figure 3f). Together, this indicates that the set of shared MLL fusion target genes might represent a ‘core’ set of targets important for driving the leukemic potential of both MLL-AF9 and -AF4.

Interplay of MLLr target genes with RUNX1 and CTCF

As RUNX1 is a known factor in several types of translocated AML and ALL and has been suggested to be involved in MLLr leukemias, we investigated RUNX1 DNA-binding in MLL-fusion-induced AML. Next to that, it was recently shown that MLL-translocated leukemias are affected by mediator kinase inhibition. The mediator complex is associated with regulation of RNA-polymerase II at promoters and distal regulatory elements. Cohesin, important for the establishment of promoter-enhancer interactions, co-localizes with CTCF and mediator. Therefore, we investigated CTCF binding at MLL fusion targets, as a proxy for mediator/cohesin binding. CTCF (GSM1335528) and RUNX1 show enrichment on MLL-AF9 target genes, while the signal on WT MLL genes is slightly lower for RUNX1 and almost absent for CTCF (Figure 4a, Supplementary Figure S3A). A total of 22% (215) of MLL-AF9 target gene promoters are occupied by CTCF, versus 11% for MLL WT target genes. For RUNX1, this overlap is 80% (767) and 61% (46), respectively (Figure 4b, left). Moreover, RUNX1 co-occupied MLL-AF9 target genes are higher expressed than MLL-AF9 target genes without RUNX1 co-occupancy (Figure 4b, right). As RUNX1 binding to MLL-AF4 target genes in MV4-11 cells follows a similar pattern as discussed for MLL-AF9 in THP-1 cells (Supplementary Figure S3B), together these results suggest that targeting the RUNX1 gene program is a common feature of MLLr AMLs.

MLL fusion binding at distal regulatory regions

MLLr leukemias have largely been described to function via aberrant elongation of MLL target genes. However, as we noticed a significant portion (~25%) of MLL-AF9 and MLL WT peaks occurring in distal regions (Figure 1g, right), we next set out to characterize these putative MLL-bound enhancers. Overlapping distal MLL peaks with H3K27ac peaks yielded 342 MLL-AF9-bound and 75 MLL WT-bound active enhancers with high H3K27ac and low H3K4me3 signal (Figure 4c, top and middle, Supplementary Table S5). Interestingly, while RUNX1 was detected on both MLL-AF9 and MLL WT enhancer peaks, CTCF and H3K79me2 signals were specific for MLL-AF9-bound enhancers (Figure 4c bottom, Supplementary Figure S4A).

Figure 3. Comparison of MLL-AF4 and MLL-AF9 target genes. (a) Overview of AF4, MLL, H3K79me2, H3K27ac and H3K4me3 binding and transcriptional activity at the ZEB2 locus in MV4-11 cells. (b) Genomic distribution of MLL-AF4 and MLL WT binding events in the ‘broad’ and ‘sharp’ modes. (c) Expression levels of MLL-AF4 and MLL WT target genes. ***P < 0.001 (Welch’s t-test) (d) Distribution of MLL-AF9-specific and MLL-AF9 and MLL-AF4 common target genes (top). Distribution of MLL WT target genes in THP-1 and MV4-11 (bottom). (e) Average signal of H3K4me3, H3K27ac and H3K79me2 on MLL-AF4 and MLL-AF9 and MLL-AF9 common target genes versus MLL-AF9 and MLL-AF4 specific target genes. (f) Expression level of MLL-AF9 target genes shared (+) or not shared (−) by MLL-AF4 in THP-1 and MV4-11 cells (left). Expression level of MLL-AF4 target genes shared (+) or not shared (−) by MLL-AF9 in THP-1 and MV4-11 cells (right). ***P < 0.001 (Welch’s t-test).
As MLL-AF9 and MLL WT enhancers occupy both intergenic as well as intronic regions (Figure 4d, top left), we further characterized only the set of MLL-bound intergenic enhancers to prevent mixing gene body and intronic enhancer chromatin signatures. A higher percentage of MLL-AF9-bound enhancers (39%) than MLL WT-bound enhancers (12%) was marked by H3K79me2 (Figure 4d, top right). This might reflect that aberrant deposition of the histone mark by DOT1L tethered to the MLL-AF9
fusion protein. RUNX1 differences were less striking, while a CTCF peak is present in 50% of all MLL-AF9-bound intergenic regions, versus 11% for MLL WT, pointing towards increased interaction with mediator for the MLL fusion-bound enhancers. Next to that, MLL-AF9-bound intergenic enhancer regions showed a significantly higher expression of enhancer RNA than MLL WT-bound regions (Figure 4d, bottom left), with MLL-AF9-bound intergenic regions co-occupied by H3K79me2 showing an even higher expression (Figure 4d, bottom right). Taken together, these findings indicate that MLL-AF9-bound enhancer regions are epigenetically more activated than their WT counterparts and might show aberrant H3K79me2 deposition and expression owing to the binding of the MLL fusion protein.

Next we wondered how distal binding sites identified in MV4-11 cells for the MLL-AF4 fusion would compare to the MLL-AF9-bound enhancers, as a difference in distal regulatory elements could potentially explain the difference in gene expression we found between the MLL-AF9- and -AF4-specific target genes (Figure 3f). MLL-AF4 active distal binding regions were similarly grouped based on H3K27ac (although average H3K27ac signal was lower as compared with MLL-AF9 distal regions), MLL and AF4 signals (Supplementary Figure S4B, Supplementary Table S5). Unlike the MLL fusion target genes, there was virtually no overlap (2%) between MLL-AF9 and MLL-AF4 enhancers, indicating that the core set of common target genes might be regulated by a variable set of regulatory regions in the different MLL fusions.

Interestingly, we observed the same decrease in co-occupancy of H3K27me2 on MLL-AF4 versus MLL WT intergenic enhancers, while RUNX1 co-occupancy was slightly higher in the WT set, and no significant difference in enhancer RNA expression between MLL-AF4 and WT MLL intergenic enhancers was observed (Supplementary Figure S4C). This could indicate that, as the epigenetic landscape of MLL-AF9- and MLL-AF4-bound active enhancers is similar except for a lower H3K27ac signal, transcriptional activity of these elements might be restricted to MLL-AF9-based leukemia, potentially either as a result of differences in complex presence (DOTCOM vs SEC) or H3K27ac occupancy.

Next we set out to determine the genes these enhancers are potentially interacting with. The majority of MLL-fusion-bound enhancer regions are located between 5 and 500 kb from the nearest TSS (Supplementary Figures S4D and E). However, as the TSS closest to an enhancer is not necessarily the one it acts upon, we refined our list of closest genes by comparison with topologically associating domain (TAD) data and by including active genes only. For this, we combined the CTCF-binding data in THP-1 with TADs in human monocytes as determined by HiC (TK, S-YW and HGS, in preparation) to get an approximate distribution of TADs. Subsequently, we linked the MLL-AF9-bound enhancers to the closest active (H3K27ac marked) gene within the same TAD, resulting in 247 genes putatively regulated by MLL-AF9-bound enhancers, including BCL2, PHLP1, Runx1 and SPI1 (Supplementary Table S6). We confirmed that our THP-1 TAD list includes the boundary at the 5' of the HOXA cluster, as described in THP-1 cells47 (Figure 4f), and performed further validation using 4C-seq experiments on the promoters of BCL2 and PHLP1, both MLL-AF9 target genes identified in this study, as bait. This allowed to confirm interactions formed by these promoters with MLL-AF9-bound active enhancers (Figure 4h).

Interestingly, of the 247 genes regulated by MLL-AF9-occupied enhancers, 25% (61) are MLL-AF9 promoter/body-occupied genes, which is significantly (P < 0.0001) more than in a random selection of genes. Next to that, GSEA (gene set enrichment analysis) pathway analysis of this set of 247 genes revealed a strong enrichment for cancer (q < 10^{-4}), MYC (q < 10^{-4}), AML (q < 10^{-4}) and immune system (q < 10^{-4}) pathways (Figure 4g), providing a strong indication that these genes and their putative enhancers are indeed implicated in MLL-AF9-mediated leukemogenesis and/or maintenance.

Figure 4. Characterization of MLL-AF9-bound distal regulatory elements. (a) Average signal of RUNX1 and CTCF on MLL-AF9 and MLL WT target genes. (b) Rate of co-occupancy of MLL-AF9 and MLL WT target genes by RUNX1 and CTCF (left). Expression level of MLL-AF9 target genes grouped by RUNX1 co-occupancy. **P < 0.001 (Welch's t-test) (right). (c) Average signal on MLL-AF9 (left) and MLL WT (right) bound enhancers for H3K4me3 and H3K27ac (top), MLL and AF9 (middle) and RUNX1, CTCF and H3K79me2 (bottom). (d) Genomic distribution of MLL-AF9 and MLL WT enhancers (top left). Co-occupancy of MLL-AF9 and MLL WT bound enhancers by CTCF, H3K79me2 and RUNX1 z(top right). Expression levels of MLL-AF9 and MLL WT intergenic enhancers (bottom left). Expression level of MLL-AF9 bound enhancers grouped by H3K79me2 co-occupancy. **P < 0.05 (Welch's t-test). (e) Motif family enrichment for MLL-AF9-bound enhancers. (f) Overview of HOXA locus in THP-1 cells showing alignment of AF9, MLL, H3K79me2, H3K4me3, H3K27ac, RUNX1 and RNA expression signal with the HOXA TAD boundary (gray box). (g) Pathway enrichments of active genes nearest to an MLL-AF9-bound enhancer within the same TAD. (h) Long range interactions from the BCL2 and PHLP1 promoters as measured by 4C-seq in THP-1 cells (black bars, q < 0.01) aligned with MLL, AF9 and H3K27ac ChIP-seq patterns on MLL-AF9-bound enhancers (gray boxes). Arrows highlight examples of interactions of the baited promoters with MLL-AF9-bound enhancers.
specificity, we compared the clustering results to the expression in other types of AML (Ley et al. 48). This identified three factors, ZNF521, ZNF433 and ZNF532, for which expression (Figure 5f) was increased in MLL-AF9-positive cells only, suggesting that these collaborate with MLL-AF9 in deregulating gene expression and driving leukemogenesis and several tumor-suppressing factors, such as ETV3, NR4A1 and EGR2, whose expression (Figure 5g) is downregulated in MLLr as well as all other types of AML included in the analysis. Deregulation of, for example, ZNF521, NR4A1 and EGR2, has indeed previously been implicated in AML. 49–51
In summary, these results show that MLL fusion target genes identified in this study can be divided into a group behaving more like CD34+ cells and a group behaving more like monocytes. Similarly, expression of TF family members whose motifs were enriched under MLL-AF9 and -AF4 target genes and enhancers can also be classified as CD34+ like, monocyte like or MLLr specific. Disturbance of the normal gene expression patterns of both direct MLLr targets and co-regulating TFs potentially produces the leukemogenic phenotype witnessed in MLL-translocated AML.

DISCUSSION

In this study, we investigated the genome-wide binding and epigenetic signature of MLL-AF9, MLL-AF4 and WT MLL in AML-derived cell lines carrying the respective MLL fusions. Enrichment of H3K79me2 and RUNX1 signal was high on both MLL-AF9 and MLL-AF4 target genes. Enrichment of H3K79me2 confirms that deposition of this histone modification on aberrantly activated MLL fusion target genes is also a feature of MLL-AF4-induced AML, as was shown for MLL-AF4 in murine and human ALL models.22,26.28 Deposition of H3K79me2 is possibly deregulated in all MLLr acute leukemias involving a component of the SEC via indirect association of the SEC with DOT1L via AF9 or ENL.22,34 However, as H3K79me2 is enriched on all activated genes in general and not just on aberrantly activated MLL fusion targets, inhibiting DOT1L function to non-specifically stop H3K79me2 deposition5,9 may introduce deleterious off-target effects.

Enrichment of H3K27ac on aberrantly activated MLL target gene promoters—and BRD4 being an MLLr target—corroborates the facilitating role of bromodomain proteins in transcription of MLLr targets, as evident from AML susceptibility to BET inhibition.56.28 This suggests a positive feedback loop where transcription of MLLr targets is facilitated by BRD4, and BDRD4 is transcribed because it is an MLLr target.

The large overlap between MLLr and RUNX1-binding sites and the identification of RUNX1 as an MLLr target gene suggests that MLLr AML deregulates (a subset of) the RUNX1 program, important for hematopoietic development.57 This is reminiscent of the way the CBFß-MYH11 oncofusion modulates the expression of RUNX1 targets59 and AML1-ETO can increase the expression of a subset of RUNX1 targets.60,59.66 Together with aberrant expression of genes such as MYC in most AML subtypes,69,66 this hints at the existence of a core set of genes, including a subset of the RUNX1 program, that is important for leukemic transformation and maintenance in translocation-induced AML.

In addition to MLL fusion binding to promoter regions as expected by the consensus of MLLr acting through aberrant activation of MLL target genes, we determined a significant number of MLL-fusion-binding sites at active distal regulatory elements enriched for H3K27ac, H3K79me2 and RUNX1 and in close proximity to genes enriched for pathways related to leukemia. In light of this, it seems likely that MLLr does not only act directly on their target genes but can also modulate the expression of target genes via distal regulatory elements. The MLL-AF4- and MLL-AF9-bound enhancers showed almost no overlap, indicating that each MLL subtype has a distinct enhancer repertoire. This is in line with several studies linking differences in acquired and innate resistance to BET inhibition in various AML cell lines and different clones to a dynamic or variable enhancer landscape.12,62

Moreover, MLL-AF9-bound enhancers are enriched for CTCF binding, which is in line with MLLr AML cells being responsive to treatment with mediator kinase inhibitors.19 This suggests an active role for MLL-AF9 in modulating the chromatin conformation to facilitate target gene expression via interference with the interplay between CTCF, RAD21 (cohesin) and the mediator complex.63,64

We extracted an extended set of core MLLr target genes including known targets such as MYC, RUNX1, BCL2 and CDK6, which can potentially be used to develop new strategies for combating MLLr leukemias, for instance, by fine-tuning the targeting of existing potential treatments, such as inhibition of DOT1L,79,95,98,101 or BET11 to MLL fusion target genes only. Next to that, the uncovered sets of MLL-AF9- and -AF4-specific target genes such as ZNF521 and CTK2NA, respectively, indicate that each specific fusion partner also has its own unique binding signature, which may potentially be exploited against MLLr blasts with resistance against a more general treatment, such as BET inhibition.35

Finally, we show that gene expression of MLL-AF9 target genes can be divided into CD34+ like and monocyte-like groups, thereby keeping the MLL-AF9-positive leukemic cells in a state between CD34+ progenitor cells and fully differentiated monocytes. Likewise, TFs from families with motifs enriched under MLL fusion targets can be divided into CD34+ like, monocyte-like and MLLr-specific groups, uncovering TFs such as ZNF521 and ZNF433 that are indirectly involved in the expression or selection of MLLr target genes and tumor-suppressing TFs such as ETF3 and NR4A1 that are downregulated in MLLr leukemias. Interestingly, ZNF521 is also an MLL-AF9 target, suggesting a feed-forward loop of ZNF521 and the MLL-AF9 leukemic program.

MATERIALS AND METHODS

Cell culture

THP-133 and MV4-1166 cells were routinely cultured in RPMI 1640 supplemented with 10% fetal calf serum and 1% penicillin/streptomycin at 37 °C in a humidified incubator with 5% CO2. Mycoplasma status was determined every 6 months.

Patient samples

Bone marrow samples from MLL-AF9-positive AML patients were collected at the Helsinki University Hospital, Helsinki and institutional guidelines and regulations (CMO 2013/064). Patient data are summarized in Table 1.

ChIP and ChIP-seq

Chromatin from cell lines was harvested as described.67 ChIPs were performed using antibodies against MLL-1, AF9, AF4, RUNX1, H3K4me3, H3K27ac and H3K79me2 and analyzed by qPCR or sequencing. Relative occupancy was calculated as fold over background, for which the promoter of the Myoglobin gene was used.
Reverse transcriptase–qPCR and RNA-seq
Total RNA was extracted with TRIzol (Invitrogen, Bleiswijk, The Netherlands) or RNeasy (QiaGen, Venlo, The Netherlands) and analyzed by treating reverse transcriptase–qPCR or strand-specific sequencing.

ChIP-seq and RNA-seq libraries were prepared according to the manufacturer’s instructions (Illumina, Eindhoven, The Netherlands). All data can be downloaded from the Gene Expression Omnibus GSE79899, or through the Blueprint DCC (http://dcc.blueprint-epigenome.eu/#/home).

**CONFLICT OF INTEREST**
The authors declare no conflict of interest.

**ACKNOWLEDGEMENTS**
We thank A Singh, A Brinkman, G Georgiou, S van Heeringen and K Francois for bioinformatic support. This work was supported by the Dutch Cancer Foundation (KWF KUN 2009-4527 and KUN 2011-4937) and Netherlands Organization for Scientific Research (NWO-Vidi to JHAM).

**AUTHOR CONTRIBUTIONS**
KHMP, AK and AMs designed and performed the wet lab experiments. KHMP, TK and S-YW analyzed the data. AEM and BAvdR provided patient blasts. KHMP, JHAM and HGS wrote the manuscript.

**REFERENCES**

34 Erben P, Gosen D, Muller MC, Reinhard J, Score J, del Valle F et al. Screening for diverse PDGFRα or PDGFRβ fusion genes is facilitated by generic quantitative reverse transcriptase polymerase chain reaction analysis. Haematologica 2010; 95: 739–744.


This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

© The Author(s) 2017

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