RUNX1/ETO and mutant KIT both contribute to programming the transcriptional and chromatin landscape in t(8;21) AML

Chin, Paulynn Suyin; Assi, Salam; Ptasinska, Anetta; Imperato, Maria; Cockerill, Peter; Bonifer, Conny

DOI:

[10.1016/j.exphem.2020.10.005](https://doi.org/10.1016/j.exphem.2020.10.005)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version

Peer reviewed version

Citation for published version (Harvard):

Chin, PS, Assi, S, Ptasinska, A, Imperato, M, Cockerill, P & Bonifer, C 2020, 'RUNX1/ETO and mutant KIT both contribute to programming the transcriptional and chromatin landscape in t(8;21) AML', *Experimental Hematology*, vol. 92, pp. 62-74. <https://doi.org/10.1016/j.exphem.2020.10.005>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

RUNX1/ETO and mutant KIT both contribute to programming the transcriptional and chromatin landscape in t(8;21) AML

Paulynn Suyin Chin[&], Salam A. Assi[&], Anetta Ptasinska, Maria Rosaria Imperato, Peter N. Cockerill⁺, Constanze Bonifer^{}**

[&]Equal contribution

⁺Senior authors

Institute of Cancer and Genomic Sciences, College of Medical and Dental Sciences,
University of Birmingham, Birmingham, B15 2TT, UK.

*Corresponding Author: Prof. Constanze Bonifer, c.bonifer@bham.ac.uk, +44(0)121414 8881

Category: Malignant hematopoiesis

Word count: 4,808

Abstract

Acute myeloid leukemia development occurs in a step-wise fashion whereby an original driver mutation is followed by additional mutations. The first type of mutations tends to be in genes encoding members of the epigenetic/transcription regulatory machinery (i.e. RUNX1, DNMT3A, TET2), while the secondary mutations often involve genes encoding members of signalling pathways that cause uncontrolled growth of such cells such as the growth factor receptors c-KIT or FLT3. Patients usually present with both types of mutations, but it is currently unclear how both mutational events shape the epigenome in developing AML cells. To this end we generated an *in vitro* model of t(8;21) AML by expressing its driver oncoprotein RUNX1-ETO with or without a mutated (N822K) KIT protein. The expression of N822K-c-KIT strongly increases the self-renewal capacity of RUNX1-ETO expressing cells. Global analysis of gene expression changes and alterations in the epigenome show that N822K-c-KIT expression profoundly influences the open chromatin landscape and transcription factor binding. However, our experiments also show that double mutant cells still differ from their patient derived counterparts, highlighting the importance of studying patient cells to obtain a true picture of how gene regulatory networks have been reprogrammed during tumourigenesis.

Keywords

RUNX1/ETO, N822K c-KIT, t(8;21) AML, chromatin landscape, transcription factor

Introduction

The t(8;21) chromosomal translocation is the most common translocation found in acute myeloid leukemia (AML). In this translocation, the DNA binding domain of RUNX1 is maintained but the activation domain that binds a range of co-activators is replaced by the ETO protein (RUNX1T1). While RUNX1 plays an important role in definitive hematopoiesis, the RUNX1/ETO (RE) fusion is known to act as a transcriptional repressor by recruiting co-repressors, such as HDACs, NCoR/SMRT and mSin3a to their target genes¹. However, the RE knock-in, transgenic mice and inducible chromosomal translocation using Cre/loxP system fail to develop leukemia²⁻⁴. Therefore, additional mutations are required for leukemogenesis.

AML is characterised by recurrent mutations that affect both the epigenetic regulatory machinery and signalling molecules. Mutations affecting transcription factors are considered a leukemia-initiating event where the resulting chimeric transcription factor leads to

transformation of the hematopoietic stem cells as well as a block in differentiation at the myeloid progenitor stage. Another class of mutations involve mutations in growth factor receptors such as FLT3 which are rendered constitutively active. These mutations are thought to confer proliferative and/or survival advantage to hematopoietic stem cells. These receptors normally control the regulated growth and survival of myeloid progenitor cells, and function by activating the RAS/MAPK and JAK/STAT signalling pathways⁵. However, little is known about how aberrant signalling impact by these receptors on gene expression and the chromatin landscape in t(8;21) AML.

In patients with t(8;21) AML, activating mutations involving signal transduction pathways (c-KIT, FLT3 and N-RAS) are recurrently found^{6,7}, thus suggesting that RE requires further mutations for leukemogenesis. Specifically, up to 48% of t(8;21) AML have c-KIT mutations⁸ and they are associated with poor disease outcome and a high incidence of relapse after chemotherapy^{8,9}. c-KIT encodes a transmembrane glycoprotein, which is a type III receptor tyrosine kinase family¹⁰. The most common activating c-KIT mutations found in these t(8;21) AML reside in exon 17, the encoding regions around the A loop of the kinase domain, resulting in D816V or N822K mutations^{8,11}. Previous studies suggest that t(8;21) AML follow a stepwise model in leukemogenesis, where RE represents the first hit to initiate the disease, whereas activating c-KIT may be a secondary but crucial hit for a full-blown leukemia⁸.

Mulloy et al transduced RE into CD34⁺ cord blood cells and found that these cells maintained self-renewal and multipotential differentiation, however they do not cause leukemia in NOD/SCID mice¹², thus further supporting the stepwise model. Conversely, the introduction of the second hit- KIT (N822K), NRAS (G12D) or CBL mutant into RE CD34⁺ hematopoietic stem and progenitor cells (HSPC) still failed to initiate AML in immunodeficient mice, suggesting that other cooperating factors are required or the cells need to undergo a strong selection in culture¹³⁻¹⁵. Interestingly, Wang et al showed that activating c-KIT (N822K) alone induced leukemia development in mice and co-expression of both RE and activated c-KIT (N822K), led to a fatal AML¹⁶. Therefore, how RE and activating c-KIT cooperate to reprogram gene expression and cause leukemia remains unclear.

In this study, we want to determine the underlying molecular basis of how the c-KIT mutation alters the RE-driven gene expression program and how it reprograms the chromatin landscape in RE expressing cells. Previous work in the lab suggested that mutations in growth factor receptors not only regulate the proliferation of leukemic cells, but also have a substantial impact on gene expression and chromatin modification¹⁷. Therefore, we expressed RE and (N822K) mutant c-KIT in primary human CD34⁺ cells and conducted epigenetic profiling to examine how these two mutations cooperate in setting up an altered epigenotype in cells

without other leukemogenic mutations. Our results show that the activating c-KIT mutation imposes a motif signature that is specific for immature cells with increased binding of GATA2 and AP-1 transcription factors. However, our data also show that transduced cells differ from leukemic blast cells from t(8;21) patients.

Methods

Cell culture

Human umbilical cord blood cells were obtained from the Birmingham Women's NHS Foundation Trust, Birmingham, UK, from donors after informed consent and with the required ethical and protocol approval by the NHS National Research Ethics Committee of the UoB in accordance with the Declaration of Helsinki. CD34⁺ cells were purified using CD34⁺ magnetic antibodies (Miltenyi). RE/CB cells were established as previously described¹⁸. RE/WT and RE/N were also established in a similar way. Cells were maintained in a humidified incubator at 37 °C with 5% CO₂. RE/CB, RE/WT and RE/N cells were cultured in Iscove modified Dulbecco medium with 20% BIT Serum Substitute (Stemcell Technologies #9500), supplemented with 10 ng/ml stem cell factor (SCF), thrombopoietin (TPO), FLT3 ligand, interleukin-3 (IL-3) and IL-6.

Transfection of HEK293T cells for RE, WT c-KIT and N822K c-KIT virus production

MSCV-AE-IRES-GFP (RE plasmid) was a gift from James Mulloy, Cincinnati, USA. pRUFneo WT c-KIT and pRUFneo N822K c-KIT plasmids were gift from Richard Kahl, Newcastle, Australia. Each transgene expressing plasmid (24 µg) was mixed with Gag/Pol 20 µg and Env 6 µg (gift from James Mulloy). For each 15 cm² dish 3 ml of Optimem serum free media was mixed with 150 µl of TransIT-293 (Mirus, USA) and incubated at room temperature for 15 minutes. The TransIT-293– DNA mixture was added dropwise to the HEK293T plate. Viral supernatant was collected after 24 hours and subsequently every 12 hours for 36 hours. Viral supernatant was concentrated using a Centricon Plus-70 100 kDa filter (Millipore, USA), according to manufacturer's instructions.

Retroviral transduction with Retronectin

Non-tissue culture treated plates were incubated with 2 µg/ml RetroNectin (Takara, Japan) solution in PBS. Concentrated virus was coated onto the wells by centrifugation before CD34⁺ cord blood cells or RE/CB cells were added with polybrene at 8 µg/ml. The plate was left overnight at 37 °C, 5% CO₂ in a humidified incubator. Transduced CD34⁺ cord blood cells

were isolated by FACS using GFP labelling for RE/CB cells while RE/WT and RE/N cells were purified using CD117⁺ magnetic antibodies (Miltenyi).

Patient samples and PBSC cell processing

Human tissue was obtained with the required ethical approval from the NHS National Research Ethics Committee of the UoB in accordance with the Declaration of Helsinki. AML and PBSC samples used in this study were either surplus diagnostic samples, or were fresh samples obtained with informed consent from all patients. AML samples were obtained from either the Centre for Clinical Haematology, Queen Elizabeth Hospital Birmingham, Birmingham, UK or the West Midlands Regional Genetics Laboratory, Birmingham Women's NHS Foundation Trust, Birmingham, UK. Mononuclear cells were purified on the same day that they were received, and in most cases also directly further purified using either CD34 or CD117 (KIT) magnetic antibodies, as previously described¹⁹. For some samples with greater than 92% blast cells the column purification was not performed.

Western blotting, RNA extraction, cDNA synthesis and RT-PCR

RNA was extracted using the Machery-Nagel Nucleospin kit. The quality of RNA was assessed using a Total RNA PICO Bioanalyser chip (Agilent technologies, USA). Western blotting, cDNA synthesis, and RT-qPCR was performed as described previously²⁰. Antibodies and primers are listed in supplementary materials.

RNA Sequencing

RNA-seq libraries were prepared with a Total RNA Ribo-zero library preparation kit (with ribosomal RNA depletion) (Illumina, USA) as previously described²¹.

ATAC-seq

ATAC-seq was performed on 50,000 cells. ATAC-seq libraries were prepared essentially as described^{22,23}. Libraries were sequenced in a pool of 12 indexed libraries using a NextSeq 500/550 High Output Kit v2 (150 cycles) for paired end sequencing (Illumina) at the Genomics Birmingham sequencing facility.

Cross-linking, ChIP and library preparation for high through put sequencing

Cross-linking, Chromatin immunoprecipitation (ChIP) and library preparation for high throughput sequencing was performed as previously described^{19,21,24}. Full details with primer and antibody details can be found in the supplementary materials.

Colony Formation Assays

Colony formation assays were performed on sorted cells by seeding at 10,000 cells/ml in Methocult Express (Stem Cell Technologies). After 10 days colonies were counted and then the cells were collected for replating.

Data analysis

Full details of bioinformatics analysis can be found in the supplementary materials.

Results

Cells co-expressing RE and mutant c-KIT display features of oncogenic transformation

To understand the effect of RE and the co-operating effect of an activating c-KIT mutation in progenitor cells without the presence of other mutations, we generated CD34⁺ cord blood (CB) that stably express RE (RE/CB), RE with a wild type c-KIT (RE/WT) or RE with the activating N822K mutant c-KIT (RE/N). Purified CD34⁺ cells from CB were transduced with a retrovirus vector expressing the full length RE and GFP fluorescence marker. GFP expressing transduced cells were sorted (Supplementary Fig. 1a) and transduced with a second retroviral vector expressing either the WT or the N822K mutant c-KIT (Figure 1a). The presence of the N822K mutation was confirmed by sequencing (Supplementary Fig. 1b). The N822K mutation was selected for this study as this is one of the most common activating c-KIT mutations found in t(8;21) AML and it is also the same mutation that is found in Kasumi-1 cells. Cells that co-expressed both RE and c-KIT were sorted and we performed RNA-seq, ATAC-seq and ChIP-seq to profile their epigenome. Interestingly, the growth of RE/N cells was still cytokine-dependent (including SCF) as previously described¹³. In order to determine whether the c-KIT tyrosine kinase signalling pathway is constitutively activated in these cells, SCF was withdrawn for 24 hours prior to the isolation of cell extracts for western blotting. ERK1/2 is in the downstream signalling pathway of the MAPK families and is activated by the c-KIT receptor²⁵. We found increased levels of ERK1/2 phosphorylation in the RE/N cells as compared to the RE/CB cells and the RE/WT cells without the presence of SCF (Figure 1b). However, both levels of phosphorylation were similar between the RE/N cells cultured with and without SCF, suggesting that the tyrosine kinase signalling pathway is constitutively activated in these cells. RUNX1 and RE protein levels were found to be slightly higher as compared to the t(8;21) Kasumi-1 cell line but were expressed at similar levels in all transduced cells as compared to each other (Figure 1b, right panel). Morphological analysis of these cell populations showed that all cell populations resembled myeloid blast cells (Supplementary Fig. 1c).

Both RE/WT and RE/N cells displayed a growth advantage and enhanced cell expansion ability as compared to RE/CB cells (Figure 1c) demonstrating the transforming effect of RE. To confirm this finding, we examined the clonogenic potential of these cells using a methylcellulose colony re-plating assay. Cells were collected for replating every 7 days. In the first plating, RE/WT and RE/N cells generated increased number of colonies when compared to RE/CB (Figure 1d, Supplementary Fig. 1d). However, RE/N cells could be repeatedly replated whereas RE/CB and RE/WT cells stopped growing by the third plating, suggesting that overexpressing mutant KIT strongly enhances the self-renewal potential of these cells.

RE expressing cells display extensive differences in gene expression and the active chromatin landscape as compared to normal CD34⁺ progenitor cells

To compare RE cells with RE cells carrying either form of c-KIT on gene expression, we analysed gene expression profiles of these transduced cells using RNA-seq. We compared them to normal CD34⁺ progenitor cells isolated from CB cells, CD34⁺ progenitor cells purified from mobilised peripheral blood stem cells (PBSCs) and from three t(8;21) AML patients with one of them carrying the KIT mutation. Principal component analysis (PCA) of the RNA-seq data based on all expressed genes indicated that they segregated into four different clusters: CD34⁺ CB, CD34⁺ PBSC, t(8;21) AML and RE/CB, whereby the same cell types cluster together (Figure 2a). Scatter plots of the two biological replicates of each cell types (RE/CB, RE/WT and RE/N) showed that the replicates showed a good correlation (Supplementary Fig. 2a). However, the cells that co-express RE and KIT did not cluster with the cells from the t(8;21) patients. RE/CB cells showed extensive differences in gene expression as compared to CD34⁺ CB cells with 2,150 differentially upregulated and 2,349 differentially downregulated genes (Figure 2b and 2c). The analysis of 3-fold down-regulated genes by Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway analysis²⁶ showed a down-regulation of different signalling pathways and signalling genes such as *KIT* and *CSF3R* (Supplementary Fig. 2b, Supplemental Table S2). WT and mutant c-KIT caused only a limited change in gene expression as compared to RE/CB. Based on the RNA-seq data, *RUNX1* was expressed at lower levels as compared to PBSCs and t(8;21) patient cells (Figure 2d). Furthermore, overexpressing RE in the CD34⁺ CB cells downregulated *KIT* expression. Conversely, overexpressing KIT led to upregulation of genes involved in ERK, cAMP and WNT signalling (such as *LEF1*, *SOX13*, *ADCYAP1* and *ITGA1*²⁷) as well as the HIPPO signalling pathway (*TEAD2*) (Supplementary Fig. 2c).

While the overall direction of gene expression changes between RE/WT and RE/N (Fig 2b,c) was similar, we also found differences which shed light on the reason for the different behaviour of the two cell lines. This included *CSF1R* encoding for the CSF-1 receptor which

is absolutely required for macrophage differentiation²⁸ and was downregulated in cells with mutant KIT expression (RE/N) as compared to RE/WT cells (Supplemental Table S2), suggesting that the cells remained in a less mature state. Besides *CSF1R*, downregulated genes include HLA-DQ histocompatibility antigens which may impact on immune evasion. Up-regulated genes included *PAGE5* (CT16) which is involved in apoptosis regulation²⁹, the NOTCH regulator *DLK1* which is overexpressed in AML³⁰, as well as the Hox member *MEIS1* which is an oncogene involved in AML and regulates *DLK1*³¹. We also find upregulation of Calreticulin (*CALCRL*), an inhibitor of *CEBPA* expression (a driver of myeloid differentiation³²). Taken together, these results suggest that RE drives the majority of deregulation of gene expression changes while the addition of both forms of c-KIT affects fewer genes with many of them being involved in regulating signalling pathways. However, the expression of mutant KIT selectively changes the expression of genes involved in the control of differentiation and apoptosis.

In order to determine how the differences in gene expression correlated with chromatin accessibility, we performed an assay for transposase-accessible chromatin using sequencing (ATAC-seq) to map accessible regions of the chromatin in the transduced cells and compared them to the ATAC-seq from CD34⁺ CB cells. Hierarchical clustering of the ATAC-seq peaks again showed three distinct clusters consisting of a RE/CB/WT/N cluster, a CD34⁺ cluster and a t(8;21) patient cluster (Supplementary Fig. 3a), consistent with the PCA of the RNA-seq data. Interestingly, the *in vitro* generated cells clustered away from the t(8;21) patients including one carrying the KIT mutation, t(8;21)/KIT-2 AML³³. This result shows that although all cells both express RE and mutated c-KIT, they still differ from patient cells. RE/CB cells showed a large number of changes in the accessible chromatin landscape as compared to CD34⁺ CB cells, with 4,650 ATAC-seq peaks being specific for RE/CB and 5,071 ATAC-seq peaks being specific for CD34⁺ CB (Figure 3a). Furthermore, while ETS, RUNX1, and GATA motifs were shared between the DNA motifs present within distal ATAC-seq peaks specific for the RE/CB and CD34⁺ CB, we observed a specific enrichment of E-Box and HOX motifs in the distal CD34⁺ CB specific ATAC-seq peaks (Supplementary Fig. 3b). To correlate accessible chromatin sites with RUNX1-ETO binding, we performed chromatin immunoprecipitation assays. RE ChIP-seq in RE/CB cells showed that RE is binding to the ATAC-seq peaks that are shared between RE/CB and CD34⁺ CB. In addition, the ATAC-seq peaks gained when RE was overexpressed correlated with decreased gene expression (outmost right panel).

WT c-KIT and N822K c-KIT expression activate cis-regulatory elements with an AP-1 motif signature

We next examined the effect of overexpressing mutant KIT on the chromatin landscape. The presence of N822K c-KIT resulted a loss of 1,449 and a gain of 434 ATAC-seq peaks. We also determined the binding sites for R/E and AP-1/JUND in RE/CB using ChIP (Figure 3a,b). To investigate whether the open chromatin patterns in the transduced cells correlate with a specific stage of myeloid differentiation, we compared the ATAC-seq data from RE/CB, RE/WT and RE/N to published ATAC-seq data that define the open chromatin regions of normal stem and progenitor cells representing different developmental stages³⁴ (Figure 3b, Supplementary Fig. 3c). RE/CB open chromatin patterns correlated with those of CMP and to a lesser extent with GMP, MPP and MEP cells, suggesting a delay in hematopoietic differentiation at this stage and a presence of a mixed population of precursor cells. ATAC-seq peaks specific for RE/CB (bottom- group 3) showed increased ETS and PU.1 motif enrichment. Open chromatin patterns specific for RE/N (top- group 1) also correlated with those of CMP cells but were enriched in AP-1 and GATA motifs. We next performed Chromatin immunoprecipitation assays followed by next generation sequencing (ChIP-Seq) to determine AP-1/JUND binding in RE/N to investigate whether the AP-1 bound genes annotated to the RE/N specific (group 1) or the RE/CB specific (group 3) were correlated with up or down regulated genes using Gene Set Enrichment (GSEA) analysis (Supplementary Fig. S4a, b; Supplementary Fig. 3d, Supplemental Table S3). GSEA plots showed that the RE/N specific AP-1/JUND target genes correlated with increased gene expression in RE/N cells. In all the three cell types, GATA and RUNX motifs were evenly distributed in the ATAC-seq peaks with R/E binding in these chromatin accessible regions, consistent with the fact that all three cell types express RE.

To identify additional enriched transcription factor binding motifs within ATAC-seq peaks, we performed a more refined analysis. We compared each cell type of each other, then examined the enrichment of multiple motifs at peaks specific to each cell type and followed by clustering of the enrichment p-values (Figure 3c). This analysis shows how open chromatin regions specific for each cell type are enriched with a distinct cluster of motifs as compared to another. As compared to all other cell types, the open chromatin regions of CD34⁺ CB cells were enriched for TCF3, TAL1 and NF1 motifs, while ATAC-seq peaks in cells expressing RE were specifically enriched for WT1, STAT3, STAT5, ETS and MEF2 motifs. Cells expressing both the WT and N822K c-KIT, activated open chromatin regions enriched for AP-1, CREB, GATA and C/EBP motifs. Taken together, these results result strongly suggests that KIT overexpression activates the cAMP/AP-1 signalling pathway.

Newly gained RE sites in RE/N cells correlate with increased RUNX1 and GATA2 binding

The differential enrichment of GATA motifs prompted us to examine how the binding pattern of GATA2 and RUNX1 compared to the binding of AP-1 (JUND), and RE in N822K c-KIT expressing populations using ChIP-seq (Figure 4a). Overall, RE, GATA and JUND/AP-1 were bound predominantly to sites distal to the promoter and this general pattern were not affected by overexpression of KIT (Supplementary Fig. 4a). The binding of RE in RE/CB and RE/N cells showed a strong overlap with the RE binding sites from t(8;21) patients²⁰ which were entirely contained within this data-set (Figure 4a). However, the expression of mutant KIT together with RE led to a shift in binding for GATA, JUND/AP-1 and RUNX1 to different sites as compared to expressing RE alone (Supplementary Fig. 4b-d) supporting the notion from the ATAC-seq analysis (Figure 3b) that the cells had shifted identity. In addition, and in concordance with activated signalling, JUND/AP-1 peaks in N822K c-KIT expressing cells were enriched for AP-1 and CREB motifs. Moreover, also RE binding changed after the expression of mutant c-KIT (Supplementary Fig. 4e).

To test how the binding of RUNX1 and GATA in distal peaks related to distal RE peaks, we aligned RUNX1 and GATA2 binding sites in RE/CB and RE/N cells alongside the RE pattern. The analysis showed that 5,801 RE binding sites that were gained in RE/N cells, and 1,336 sites were lost (Figure 4b). The overall pattern of accessible chromatin sites at RE peaks was not changed, however, newly gained RE sites were enriched for RUNX1 and GATA motifs which correlated with increased RUNX1 and GATA2 binding, indicating that the cells entered an altered and most likely more immature differentiation state as previously demonstrated³⁵. This result was consistent with a reduced expression of the monocytic differentiation marker CD11b marker in KIT expressing cells, RE/WT and RE/N as compared to RE/CB cells (Supplementary Fig. 4f).

To further characterise enriched motifs at the binding sites of RE, RUNX1, GATA2 and JUND in chromatin, we examined the enrichment of multiple motifs at binding sites specific for each factor in either RE/CB or RE/N cells by performing pair-wise analyses and then clustering the enrichment scores (Figure 4c). The analysis showed that the JUND specific peaks in both cell types besides being enriched for their own motifs, were enriched for Sp1 and KLF binding sites. However, NFIL3 motifs were only enriched in the JUND sites of the KIT expressing cells. Similarly, the GATA2 specific peaks in both cell types were enriched for EVI1 motifs, but C/EBP, PU.1 and NFIL3 motifs were more enriched in the RE/CB cells (Figure 4c and Supplementary Fig. 4c). Interestingly, RUNX1 specific peaks in both cell types were each enriched for distinct set of motifs. RUNX1 specific peaks in RE/CB cells were enriched for KLF, SPI1 and SMAD motifs whereas RUNX1 specific peaks in RE/N cells were enriched for EVI1, FOXO, SOX, NFAT and MYB motifs, indicating a different binding site distribution of this factor in response to mutant KIT expression.

The open chromatin signature of RE + KIT *in vitro* generated cells differs from the t(8;21) AML patient signature

To investigate whether the open chromatin patterns from the *in vitro* generated double mutant cells were similar to those of t(8;21) AML patients that have the same KIT mutation, we compared the ATAC-seq peak profiles from RE/CB, RE/WT and RE/N cells to DNaseI-seq profiles from two t(8;21) AML patients³³ (Figure 5a-b). A large number of open chromatin regions (10,172) were specific to the t(8;21) AML patient cell carrying the KIT mutation and were enriched for C/EBP, E-box, NF-κB and POU4F1 motifs. We next compared these DHS peaks to the published ATAC-seq peaks from cells of different stages of hematopoiesis²³ and found the DHSs specific to the t(8;21) patients aligned closely to the GMP and monocytic cells. Similarly, the average profile of the t(8;21) patient specific peaks correlated with high GMP peaks coverage (Figure 5c) with about 50% of the peaks overlapping with GMP peaks and 40% of the peaks overlapping with those of monocytes (Figure 5d). The ATAC-seq peaks specific for *in vitro* generated RE/N cells aligned closely to the CMP open chromatin pattern by showing a high correlation and an about 60% overlap with CMP peaks (Figure 5c-d). RE/N specific ATAC-seq peaks were enriched for GATA motifs, again suggesting that the *in vitro* generated cells were more immature as compared to the patient cells. While both open chromatin regions from *in vitro* derived and patient cells showed enrichment of AP-1 motifs, the patient cells demonstrated an additional enrichment of the POU4F1 motif and motifs characteristic for an inflammatory signature (NF-κB) which were not observed in the *in vitro* cells. Our results therefore demonstrate a profound influence of the extracellular environment on the chromatin landscape of t(8;21) cells.

Discussion

Our work clearly demonstrates that the co-expression of RUNX1-ETO and N822K c-KIT has a profound effect on the growth behaviour of transduced cells with an enhanced re-plating ability and increased survival as shown previously¹³. Our study adds a global analysis to these findings by showing that the expression of N822K c-KIT in RE expressing cells is capable of further altering the accessible chromatin and transcription factor binding landscape. This result confirms our previous studies that chronic signalling processes in AML have the power to impose a specific chromatin signature on leukemic cells¹⁹. However, the major transforming agent is RUNX1-ETO whose expression in untransformed human progenitor cells correlated with an extensive reprogramming of gene expression. Deregulated genes consisted of those involved in multiple signalling pathways, with a down-regulation of genes involved in the MAPK and the inflammatory response pathways, including KIT itself. This finding could explain the

dependence of transduced cells on the continued presence of SCF even in the presence of an activated KIT. RE/CB cells may compensate for the lack of KIT expression by relying on off-target effects of SCF.

When compared to the impact of RE, the effect of N822K c-KIT on gene expression when compared to RE expressing cells was lower and we observed few differences between genes up-regulated by WT-c-KIT and mutant c-KIT. *KDR* which encodes for the VEGF receptor (VEGFR2) was up-regulated in RUNX1/ETO expressing cells (including patient cells) and the expression of both versions of KIT further up-regulated this gene. It is of note that all analysed cell types expressed VEGFA and B indicated the presence of a potential autoregulatory loop (Supplemental Table S2). However, there were notable differences: for example, the gene encoding for the adenylate cyclase activating protein (*ADCYAP1*) which activates cAMP driven gene expression was further enhanced only in the presence of mutant c-KIT. Such activation could explain the increased occurrence of a CREB/AP-1 motif signature in RE/WT and RE/N cells.

The expression of N822K c-KIT in RE/CB cells had a strong impact on the epigenome. Activation of this signalling pathway led to a shift in the active chromatin structure and affected the binding of several transcription factors, including RUNX1-ETO itself. The reason for this behaviour may be the activation of the AP-1 factor family which is known to cooperate with multiple other transcription factor to assist in the opening of novel chromatin regions, including other signalling responsive factors³⁶⁻³⁸. Our data are consistent with the idea that the enhanced re-plating activity of RE/N cells caused by the activation of KIT signalling is based on changes the identity of the cells towards a more immature state via a stable reprogramming of the epigenome. Further experiments will be required to test whether such signalling dependent epigenetic reprogramming is a priming event that is required to activate a gene expression program capable of generating leukemia *in vivo* as described in other blood cell types³⁹.

An important result from our study is our finding that in spite of a very similar RUNX1-ETO binding profile, *in vitro* generated cells still differ from the t(8;21) AML patient signature, as evidenced by a strong difference in the open chromatin profile with a different underlying transcription factor motif pattern. While both have the AP-1 motif signature, patient cells demonstrated an increased inflammatory signature in form of enriched NF-κB motifs which is not observed in the *in vitro* generated cells. We have previously shown that t(8;21) patients activate a large number of other signalling pathway genes³³. Patient cells are also more differentiated with a chromatin pattern closer to that of GMP/Monocytic cells. This could be due to the umbilical cord blood RE target cells being different from the mutant HSCs generating t(8;21) AML. However, from our correlation data (Supplementary Fig. 3c), the

CD34+ CB cells are highly similar to CD34+ PBSCs. We hypothesize that the difference between the *in vitro* generated cells to the t(8;21) AML is due to the fact that *in vitro* generated cells are grown in serum-free cultures containing a large number of cytokines (IL-3, IL-6, TPO and SCF) while blast cells from patients are exposed to the bone marrow niche and the cytokine mix present in a diseased organism which will shape the chromatin landscape in a different way. Our comparison therefore emphasizes the prominent role of signalling processes shaping the epigenome and thus the cellular identity of leukemic cells. While *in vitro* generated cells and cell culture are important tools to obtain mechanistic advances with the regards to the biochemistry of oncogene action, our study reinforces the importance of studying patient cells to obtain a true picture of how gene regulatory networks have been reprogrammed during the development of AML.

References

- 1 Amann, J. M. *et al.* ETO, a target of t (8; 21) in acute leukemia, makes distinct contacts with multiple histone deacetylases and binds mSin3A through its oligomerization domain. *Molecular and cellular biology* **21**, 6470-6483 (2001).
- 2 Higuchi, M. *et al.* Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t (8; 21) acute myeloid leukemia. *Cancer cell* **1**, 63-74 (2002).
- 3 Yuan, Y. *et al.* AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations. *Proceedings of the National Academy of Sciences* **98**, 10398-10403 (2001).
- 4 Buchholz, F., Refaeli, Y., Trumpp, A. & Bishop, J. M. Inducible chromosomal translocation of AML1 and ETO genes through Cre/loxP-mediated recombination in the mouse. *EMBO reports* **1**, 133-139 (2000).
- 5 Corces-Zimmerman, M. R., Hong, W.-J., Weissman, I. L., Medeiros, B. C. & Majeti, R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proceedings of the National Academy of Sciences* **111**, 2548-2553 (2014).
- 6 Beghini, A. *et al.* KIT activating mutations: incidence in adult and pediatric acute myeloid leukemia, and identification of an internal tandem duplication. *haematologica* **89**, 920-925 (2004).
- 7 Goemans, B. *et al.* Mutations in KIT and RAS are frequent events in pediatric core-binding factor acute myeloid leukemia. *Leukemia* **19**, 1536 (2005).
- 8 Wang, Y.-Y. *et al.* AML1-ETO and C-KIT mutation/overexpression in t (8; 21) leukemia: implication in stepwise leukemogenesis and response to Gleevec. *Proceedings of the National Academy of Sciences* **102**, 1104-1109 (2005).
- 9 Jiao, B. *et al.* AML1-ETO9a is correlated with C-KIT overexpression/mutations and indicates poor disease outcome in t (8; 21) acute myeloid leukemia-M2. *Leukemia* **23**, 1598 (2009).
- 10 Yarden, Y. *et al.* Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for an unidentified ligand. *The EMBO journal* **6**, 3341-3351 (1987).
- 11 Lennartsson, J. & Rönstrand, L. Stem cell factor receptor/c-Kit: from basic science to clinical implications. *Physiological reviews* **92**, 1619-1649 (2012).

- 12 Mulloy, J. C. *et al.* Maintaining the self-renewal and differentiation potential of human CD34+ hematopoietic cells using a single genetic element. *Blood* **102**, 4369-4376 (2003).
- 13 Wichmann, C. *et al.* Activating c-KIT mutations confer oncogenic cooperativity and rescue RUNX1/ETO-induced DNA damage and apoptosis in human primary CD34+ hematopoietic progenitors. *Leukemia* **29**, 279-289 (2015).
- 14 Chou, F.-S., Wunderlich, M., Griesinger, A. & Mulloy, J. C. N-RasG12D induces features of stepwise transformation in preleukemic human umbilical cord blood cultures expressing the AML1-ETO fusion gene. *Blood, The Journal of the American Society of Hematology* **117**, 2237-2240 (2011).
- 15 Goyama, S. *et al.* UBASH3B/Sts-1-CBL axis regulates myeloid proliferation in human preleukemia induced by AML1-ETO. *Leukemia* **30**, 728-739 (2016).
- 16 Wang, Y.-Y. *et al.* C-KIT mutation cooperates with full-length AML1-ETO to induce acute myeloid leukemia in mice. *Proceedings of the National Academy of Sciences* **108**, 2450-2455 (2011).
- 17 Ray, D. *et al.* Lineage-inappropriate PAX5 expression in t (8; 21) acute myeloid leukemia requires signaling-mediated abrogation of polycomb repression. *Blood* **122**, 759-769 (2013).
- 18 Wunderlich, M., Krejci, O., Wei, J. & Mulloy, J. C. Human CD34+ cells expressing the inv (16) fusion protein exhibit a myelomonocytic phenotype with greatly enhanced proliferative ability. *Blood* **108**, 1690-1697 (2006).
- 19 Cauchy, P. *et al.* Chronic FLT3-ITD Signaling in Acute Myeloid Leukemia Is Connected to a Specific Chromatin Signature. *Cell Rep* **12**, 821-836 (2015).
- 20 Ptasinska, A. *et al.* Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding. *Leukemia* **26**, 1829-1841 (2012).
- 21 Loke, J. *et al.* RUNX1-ETO and RUNX1-EVI1 Differentially Reprogram the Chromatin Landscape in t (8; 21) and t (3; 21) AML. *Cell reports* **19**, 1654-1668 (2017).
- 22 Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21.29.21-21.29.29 (2015).
- 23 Corces, M. R. *et al.* Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat Genet* **48**, 1193-1203, doi:10.1038/ng.3646 (2016).
- 24 Ptasinska, A. *et al.* Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal. *Cell Rep* **8**, 1974-1988 (2014).
- 25 Roux, P. P. & Blenis, J. ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. *Microbiol. Mol. Biol. Rev.* **68**, 320-344 (2004).
- 26 Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic acids research* **47**, D590-D595 (2019).
- 27 Li, H. *et al.* Integrin $\alpha 1$ promotes tumorigenicity and progressive capacity of colorectal cancer. *International Journal of Biological Sciences* **16**, 815 (2020).
- 28 Bonifer, C. & Hume, D. A. The transcriptional regulation of the Colony-Stimulating Factor 1 Receptor (csf1r) gene during hematopoiesis. *Front Biosci* **13**, 60 (2008).
- 29 Nylund, C. *et al.* Melanoma-associated cancer-testis antigen 16 (CT16) regulates the expression of apoptotic and antiapoptotic genes and promotes cell survival. *PLoS One* **7**, e45382 (2012).
- 30 Khoury, H., Suarez-Saiz, F., Wu, S. & Minden, M. D. An upstream insulator regulates DLK1 imprinting in AML. *Blood* **115**, 2260-2263 (2010).
- 31 Argiropoulos, B. *et al.* Linkage of Meis1 leukemogenic activity to multiple downstream effectors including Trib2 and Ccl3. *Exp Hematol* **36**, 845-859 (2008).
- 32 Helbling, D. *et al.* CBF β -SMMHC is correlated with increased calreticulin expression and suppresses the granulocytic differentiation factor CEBPA in AML with inv(16). *Blood* **106**, 1369-1375 (2005).

- 33 Assi, S. A. *et al.* Subtype-specific regulatory network rewiring in acute myeloid leukemia. *Nature genetics* **51**, 151-162 (2019).
- 34 Corces, M. R. *et al.* Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nature genetics* **48**, 1193 (2016).
- 35 de Boer, B. *et al.* Prospective isolation and characterization of genetically and functionally distinct AML subclones. *Cancer Cell* **34**, 674-689. e678 (2018).
- 36 Voss, T. C. *et al.* Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. *Cell* **146**, 544-554 (2011).
- 37 Jain, J. *et al.* The T-cell transcription factor NFAT p is a substrate for calcineurin and interacts with Fos and Jun. *Nature* **365**, 352-355 (1993).
- 38 Obier, N. *et al.* Cooperative binding of AP-1 and TEAD4 modulates the balance between vascular smooth muscle and hemogenic cell fate. *Development* **143**, 4324-4340 (2016).
- 39 Bonifer, C. & Cockerill, P. N. Chromatin priming of genes in development: concepts, mechanisms and consequences. *Experimental hematology* **49**, 1-8 (2017).

Acknowledgements

This work was funded by a programme grant from Bloodwise (15001) to C.B and P.N.C

Author contribution

P.S.C performed experiments and wrote the paper, A.P. and M.R.I. helped with experiments, S.A.A analyzed the data. P.N.C helped supervising the project, C.B. supervised the project and wrote the paper.

Declaration of Interests

The authors declare no competing financial interests.

Figures

Figure 1: The expression of mutated KIT leads to chronic ERK phosphorylation and increased growth and survival

- Experimental strategy. Approximately 1×10^6 CD34⁺ cord blood cells were retrovirally transduced with RE expressing plasmid. These cells were sorted as pure population, expanded for a week and then the RE/CB cells were transduced again with either the WT c-KIT expressing plasmid or the N822K c-KIT expressing plasmid along with the retroviral packaging vector Gag/Pol and envelope vectors. 10-14 days after transduction, pure population of c-KIT positive cells were sorted using the CD117 surface marker. The cells were then subjected to RNA-seq and ATAC-seq, ChIP-seq, colony formation assay (CFU) and western blot.
- Western blot showing phosphorylated ERK1/2 (p-ERK1/2), total ERK1/2, RE and RUNX1 protein from THP-1 untreated or treated with PMA and calcium ionophore for 6 hours or

untreated Kasumi-1, RE/CB, RE/WT in the presence or absence of SCF or SCF withdrawal for 24 hours and RE/N with SCF or SCF withdrawal for 24 hours. GAPDH served as loading control. Bar graphs showing quantification of p-ERK1/2 relative to GAPDH loading control. The respective values of biological replicates are indicated by the grey dots on the bar graphs; n=2.

- c. Growth curve of RE/CB, RE/WT (bulk population) and RE/N (bulk population) up to 18 days after c-KIT transduction. Error bars, SD; n=3. ** denotes $p < 0.01$ by paired t-test and * denotes $p < 0.05$ by paired t-test for RE/N vs. RE/CB and RE/WT vs. RE/CB.
- d. Colony replating assays for RE/CB, RE/WT and RE/N cells. Approximately 10,000 cells were seeded for each plating on methocellulose in triplicate. Colonies were counted 10 days after plating. Error bars, SD; n=3. ** denotes $p < 0.01$ by paired t-test and * denotes $p < 0.05$ by paired t-test.

a

b

c

d

Unchanged

Upregulated (RE/N versus RE/CB)

Down-regulated (RE/N versus RE/CB)

Figure 2: The additional expression of KIT in a RE expressing background activates a limited set of genes

- a. Principal component analysis of the gene expression showing that it segregates into four different clusters as indicated.
- b. Hierarchical clustering gene expression changes as determined by RNA-seq in the different cell types. Unsupervised clustering of expression values from two independent replicates of genes changing expression at least Log2 (1.5) with the heatmap colour showing the degree of differential expression (fold change) in RE/CB as compared to CD34⁺ CB, RE/WT as compared to RE/CB and RE/N as compared to RE/CB.
- c. Number of differentially expressed genes, respectively (indicated above bars) as measured by RNA-Seq. Differentially expressed genes are those with an at least Log2 (1.5) fold change in gene expression between RE/CB as compared to CD34⁺ CB, RE/WT as compared to RE/CB, RE/N as compared to RE/CB and RE/N as compared to RE/WT.
- d. Merged FPKM levels of c-KIT expression in CD34⁺ CB (n=1), RE/CB (n=2), RE/WT (n=2), RE/N (n=2), CD34⁺ PBSC (n=2) and t(8;21) patients (n=3). The respective values of independent replicates are indicated by the grey dots on the bar graphs.

Figure 3: Both mutant and WT c-KIT expression activate an AP-1 motif signature

- a. Comparison of distal ATAC-seq peaks in CD34⁺ CB and RE/CB cells. Peaks are ranked from top to bottom in order of increasing fold change relative to DNA sequence tag count for peaks identified in RE/CB relative to CD34⁺ CB cells. The bar alongside indicates the RE/CB specific ATAC-seq peaks (group 1), shared ATAC-seq peaks (group 2) and CD34⁺ CB specific ATAC-seq peaks (group 3). The motif density plots underlying the same coordinates are plotted along the same coordinates as the ATAC-seq peaks. The next panel shows RE ChIP binding in RE/CB cells plotted along the same coordinates as the ATAC-seq peaks. The heatmap on the outermost right depicts the relative expression of

genes nearest to each DHS and are ranked along the same coordinates as the ATAC-seq peaks.

- b. Comparison of distal ATAC-seq peaks in RE/CB and RE/N cells. Peaks are ranked from top to bottom in order of increasing fold change relative to DNA sequence tag count for peaks identified in RE/N relative to RE/CB. The distal ATAC-seq peaks in RE/WT cells are plotted alongside. The bar alongside indicates RE/N specific ATAC-seq peaks (group 1), shared ATAC-seq peaks (group 2) and RE/CB specific ATAC-seq peaks (group 3). The motif density plots underlying the same coordinates are plotted along the same coordinates. ATAC-seq tag counts from distinct myeloid progenitor cell types²³ are ranked along the same coordinates as the RE/CB ATAC-seq peaks. RE and JUND/AP-1 ChIP binding in RE/CB cells are plotted along the same coordinates as the ATAC-seq peaks. Gene expression heatmap depicts the relative expression of genes nearest to each ATAC-seq peaks and are ranked along the same coordinates as the ATAC-seq peaks.
- c. Heatmap showing motif enrichment after hierarchical clustering of motif occupancy in each specific peak set. Enrichment score was calculated by the level of motif enrichment in the unique DHSs, as compared to union of ATAC-seq peaks in the pair of experiments. The heatmap is based on motif enrichment scores thus depicting the degree of motif enrichment allowing identification of enriched motifs unique to each cell type.

a RUNX1/ETO peak overlap

b

c

Figure 4: Newly gained RE binding sites in RE/N cells correlate with increased RUNX1 and GATA2 binding

- a. Venn diagram showing the RE peak overlap between RE/CB, RE/N and t(8;21) patient.
- b. Comparison of RE binding in RE/CB and RE/N cells. RE ChIP-seq peaks were ranked according to the fold difference of the normalised tag counts across 2 kb window. The bar on the outermost left indicates the RE/N specific sites (group 1), shared sites (group 2) and RE/CB specific sites (group 3). The ATAC-seq tag counts from RE/CB, RE/WT and RE/N are ranked alongside the RE binding sites. Next panels show RUNX1 and GATA2 ChIP binding plotted alongside. The right panels show the density maps of the indicated motifs aligned to these coordinates.
- c. Hierarchical clustering of enriched motifs discovered in a pairwise comparison between ChIP-seq peaks between RE/CB and RE/N cells identifying unique peaks for each cell type. The enrichment score was calculated by the level of motif enrichment in the unique peaks as compared to union of peaks in the pair of experiments. The heatmap below depicts the degree of motif enrichment. Specific sets of enriched motifs unique to RE/CB are coloured in blue while specific sets of enriched motifs unique to RE/N are coloured in red.

Figure 5: CD34⁺ cells from N822K c-KIT mutant t(8;21) patients display a different open chromatin pattern than *in vitro* RUNX1/ETO - N822K c-KIT expressing cells

- a. Left panel: ATAC-seq profile spanning 2 kb windows for RE/N (*in vitro*) and DNaseI-seq profile from a t(8;21) patient with KIT mutation³³. Peaks are ranked from top to bottom in order of increasing relative DNA sequence tag count for peaks identified in RE/N (*in vitro*) relative to t(8;21) patient with KIT mutation. The ATAC-seq peaks in RE/WT cells plotted alongside. The bar on the outermost left indicates the t(8;21) KIT mutant specific open chromatin regions (group 1), shared open chromatin regions (group 2) and RE/N specific open chromatin regions (group 3). Density maps of the indicated motifs underlying the same coordinates of the open chromatin regions are plotted alongside within +/-1 Kb windows. The next panels show the open chromatin profiles spanning 2 kb windows for RE/CB, t(8;21) patient 3³³ and t(8;21) patient 4³³. The far right panels show ATAC-seq tag counts from distinct myeloid progenitor cell types²³ are ranked along the same coordinates as the RE/N ATAC-seq peaks.
- b. De novo motif discovery in open chromatin regions specific for RE/N (left panel) and t(8;21) patient (right panel) with KIT mutation.

- c. Average coverage of the ATAC-Seq peaks from distinct myeloid progenitor cell types²³ compared to the RE/N specific (top) and t(8;21) KIT mut specific (bottom) open chromatin regions.
- d. Bar graphs showing the percentage of peak overlap of the ATAC-Seq peaks from distinct myeloid progenitor cell types²³ with the RE/N specific and t(8;21) KIT mut specific open chromatin regions.