The target cell of transformation is distinct from the leukemia stem cell in murine CALM/AF10 leukemia models

S Dutta, A Krause, S Vosberg, T Herold, B Ksienzyk, L Quintanilla-Martinez, B Tizazu, M Chopra, A Graf, S Krebs, H Blum, P A Greif, A Vetter, K Metzeler, M Rothenberg-Thurley, M R Schneider, M Dahlhoff, K Spiekermann, U Zimber-Strobl, E Wolf, S K Bohlander

Cite this article as: S Dutta, A Krause, S Vosberg, T Herold, B Ksienzyk, L Quintanilla-Martinez, B Tizazu, M Chopra, A Graf, S Krebs, H Blum, P A Greif, A Vetter, K Metzeler, M Rothenberg-Thurley, M R Schneider, M Dahlhoff, K Spiekermann, U Zimber-Strobl, E Wolf, S K Bohlander, The target cell of transformation is distinct from the leukemia stem cell in murine CALM/AF10 leukemia models, *Leukemia* accepted article preview 21 December 2015; doi: 10.1038/leu.2015.349.

Received 13 July 2015; revised 26 November 2015; accepted 3 December 2015; Accepted article preview online 21 December 2015

The target cell of transformation is distinct from the leukemia stem cell in murine CALM/AF10 leukemia models

Short title: CALM/AF10 in leukemogenesis

Sayantanee Dutta[1,2], Alexandre Krause[3], Sebastian Vosberg[1,2], Tobias Herold[1,2,4,5], Bianka Ksienzyk[1], Leticia Quintanilla-Martinez[6], Belay Tizazu[1,2], Martin Chopra[7], Alexander Graf[8], Stefan Krebs[8], Helmut Blum[8], Philipp A Greif[1,2,4,5], Anna Vetter[1,2,4,5], Klaus Metzeler[1,2], Maja Rothenberg-Thurley[1,2], Marlon R Schneider[9], Maik Dahlhoff[9], Karsten Spiekermann [1,2,4,5], Ursula Zimber-Strobl[10], Eckhard Wolf[9], Stefan K Bohlander[7]*

[1]Department of Medicine III, University Hospital Grosshadern, Ludwig Maximilians-University, Munich, Germany

[2]Clinical Cooperative Group Leukemia, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany

[3]Department of Small Animal Medicine, Federal University of Santa Maria, RS, Brazil

[4]German Cancer Consortium (DKTK), Heidelberg, Germany

[5]German Cancer Research Center (DKFZ), Heidelberg, Germany

[6]Institute for Pathology, University Hospital and Faculty of Medicine, University of Tübingen, Germany

[7]Department of Molecular Medicine and Pathology, Faculty of Medical and Health Sciences, The University of Auckland, New Zealand

[8]Laboratory for Functional Genome Analysis, Gene Center, Ludwig Maximilians-University, Munich, Germany

[9]Institute of Molecular Animal Breeding and Biotechnology, Ludwig Maximilians-University, Munich, Germany

[10]Department of Gene Vectors, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany

Corresponding author:

Prof. Dr. med. Stefan K. Bohlander

Department of Molecular Medicine and Pathology

Faculty of Medical and Health Sciences

The University of Auckland

New Zealand

Telephone: +64 (0)9 923 8348

Fax: +64 (0)9 364 7121

Email: s.bohlander@auckland.ac.nz

Word count text: 4297 words

Word count abstract: 199 words

Figure/table count: 6 figures, 1 table (13 Supplementary Figures, 9 Supplementary Tables)

Reference count: 52

Scientific category: Myeloid Neoplasia

Key points:

In unperturbed haematopoiesis, the cell of origin of leukemia (COL) in CALM/AF10 leukemia is a stem or early progenitor cell, while expression of CALM/AF10 in early B cells does not lead to leukemia. However, CALM/AF10-driven leukemia can be propagated by a leukemic stem cell (LSC) with B cell characteristics.

Abstract

The CALM/AF10 fusion gene is found in various hematological malignancies including AML, T-cell ALL, and malignant lymphoma. We have previously identified the leukemia stem cell (LSC) in a CALM/AF10-driven murine bone marrow transplant acute myeloid leukemia model as B220-positive lymphoid cells with B cell characteristics. To identify the target cell for leukemic transformation or 'cell of origin of leukemia' (COL) in non-disturbed steady-state hematopoiesis we inserted the CALM/AF10 fusions gene preceded by a loxP-flanked transcriptional stop cassette into the Rosa26 locus. Vav-Cre induced pan-hematopoietic expression of the CALM/AF10 fusion gene led to acute leukemia with a median latency of 12 months. Mice expressing CALM/AF10 in the B lymphoid compartment using Mb1-Cre or CD19-Cre inducer lines did not develop leukemia. Leukemias had a predominantly myeloid phenotype but showed co-expression of the B cell marker B220, and had clonal B-cell receptor rearrangements. Using whole-exome sequencing, we identified an average of 2 to 3 additional mutations per leukemia, including activating mutations in known oncogenes like *FLT3* and *PTPN11*. Our results show that the COL for CALM/AF10 is a stem or early progenitor cell and not a cell of B-cell lineage with a phenotype similar to that of the LSC in CALM/AF10-positive leukemia.

Introduction

The t(10;11)(p13;q14) is a rare but recurring chromosomal translocation leading to an in-frame fusion of the *AF10* (*ALL1 Fused gene from chromosome 10*, *MLLT10*) gene on chromosome 10 and the *CALM (Clathrin Assembly Lymphoid Myeloid, PICALM*) gene on chromosome 11[1]. The CALM/AF10 (C/A) fusion was initially cloned from the monocytic cell line U937, derived from a patient with histiocytic lymphoma[1]. The C/A fusion protein is found in various hematological malignancies including acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), and malignant lymphoma[2-9]. It is predominantly encountered in young patients and is associated with a poor prognosis[2,6,9-11]. In T-ALL with TCR $\gamma\delta$ rearrangement, the C/A fusion is the most frequent genetic abnormality with a prevalence of up to 30%[12].

CALM encodes a ubiquitously expressed protein involved in clathrin-mediated endocytosis and vesicular trafficking between the trans-golgi network and endosomes[13]. A nonsense point mutation in the murine *Calm* gene leads to hematopoietic abnormalities and a shortened life span in *fit1* mutant mice[14]. *AF10* is was initially identified as a fusion partner of the *Mixed Lineage Leukemia* (*MLL1*) gene[15]. AF10 has an N-terminal plant homeodomain (PHD) and a C-terminal octapeptide motif-leucine zipper domain (OM/LZ)[16,17]. The OM/LZ domain was is essential for AF10-mediated transformation[18] and also for mediating its interactions with the histone H3 lysine 79 (H3K79) methyltransferase hDOT1L[19] and the lymphoid regulator Ikaros (IKZF1)[20].

Panhematopoietic transgenic expression of C/A from a *Vav* promoter resulted in AML with a long latency of 12 months and a low penetrance of 40%[21]. Introducing additional genetic lesions in these mice by retroviral integration accelerated disease

induction and increased penetrance[22]. However, in a murine bone marrow transplantation (BMT) model, the retroviral expression of the C/A fusion protein led to an aggressive AML within 15 weeks in all mice[23]. The leukemia cells had a predominantly myeloid phenotype but were also positive for clonal B cell receptor rearrangements. In this C/A BMT model, we were able to identify the leukemia stem cell (LSC) as residing in a small subpopulation of leukemia cells that were myeloid marker negative but positive for the B cell marker B220. Importantly, human C/A-positive AML patient samples expressed CD45RA, the human homologue of murine B220, and showed IgH and TCR rearrangement[23].

It should be noted that both the transgenic Vav-C/A model by Aplan and colleagues and our BMT model have experimental limitations. In the transgenic model the integration of the Vav-C/A cassette into the genome is random resulting in variability of transgene expression. In our C/A BMT, recipient mice have their bone marrow ablated before transplantation and therefore hematopoiesis is severely perturbed with the retrovirally modified donor cells going through a phase of rapid proliferation after the transplant.

To determine the target cell or 'cell of origin of leukemia' (COL) for C/A mediated transformation in the setting of undisturbed, steady-state hematopoiesis, we generated a C/A knock-in mouse model (Rosa26 loxP-stop-loxP-CALM/AF10; R26LSLCA) with Cre-mediated expression of the C/A fusion from the Rosa26 locus. Expression of C/A in the entire hematopoietic compartment after Vav-Cre-mediated recombination caused AML with a median latency of 1 year and 100% penetrance. Most leukemias had clonal B cell receptor rearrangments and 40% of the leukemias also expressed the B cell marker B220. Interestingly, the C/A fusion did not induce

leukemia when it was expressed in cells of the B cell lineage with a similar phenotype to the C/A LSC found in the BMT model using Mb1-Cre or CD19-Cre-mediated recombination. Leukemic cells from Vav-Cre x R26LSLCA mice showed overexpression of *Hoxa* cluster genes and the *Hox* co-factor *Meis1* and had a gene expression profile that resembled human C/A-positive leukemias. Whole-exome sequencing revealed additional somatic mutations, some of them in known oncogenes like *FLT3* and *PTPN11*. In summary, our results show that the COL for C/A in unperturbed hematopoiesis is a HSC requiring additional genetic hits and that C/A is incapable of transforming cells of the B lineage even though the LSC in established C/A leukemia has B lineage characteristics.

Methods

Generation of the R26SLSCA knock-in mouse line

The C/A fusion cDNA was cloned into the pRosa26PA vector to generate the targeting construct Rosa26-loxP-Stop-loxP-CALM/AF10 (R26LSLCA) (Figure S1A) which was electroporated into 129X1/SvJ mouse embryonic stem (ES) cells. Recombinant ES cells were injected into C57BL/6 blastocysts that were transferred into foster mice to generate chimeric animals. A detailed description of the process is provided in supplementary methods.

Cre-induced expression of C/A

R26LSLCA mice were mated with Vav-Cre[24], Mb1-Cre (mb1Cre)[25], or CD19-Cre[26] mice (all C57BL/6-background) to achieve pan-hematopoietic or B cell–specific expression of the C/A fusion gene. Mice were bred and maintained under specific pathogen free (SPF) condition. All experiments were performed in compliance with the German Animal Welfare Act and were approved by the government of Upper Bavaria, Germany.

Analysis of leukemic mice

Mice expressing the C/A fusion were closely monitored and euthanized for *post mortem* analysis when moribund. Peripheral blood (PB), bone marrow (BM) and spleen (Spl) cells were analyzed by flow cytometry (FACS) using B220, CD19, CD24, CD4, CD43, CD8, cKit, Gr1, Mac1, Sca1, and Terr119 (BD Biosciences, Heidelberg, Germany), and F4/80 (Santa Cruz, Heidelberg, Germany) antibodies using

FACSCalibur (BD Biosciences). Results were analyzed with CELLQuest (BD Biosciences) or WinMDI 2.9 software.

PB smears, BM and Spl cytospins were assessed microscopically after Giemsa-May-Grünwald staining (Merck, Darmstadt, Germany). Paraformaldehyde fixed internal organs were prepared for standard hematoxylin and eosin staining, naphthol-AS-D-chloroacetate esterase (ASDCL) staining, and for immunohistochemistry using antibodies against myeloperoxidase (MPO), B220, CD3, and terminal desoxynucleotide transferase (TdT). The Bethesda proposals for the evaluation of leukemic tissues were followed[27,28]. Images were taken using an Axioskop 40 microscope (Carl Zeiss, Jena, Germany) and processed using Adobe photoshop CS3.

Gene expression profiling and microarray analyses

BM and B cells were sorted on a BD FACSVantage SE system using Gr1, Mac1 and B220 antibodies and total RNA and cDNA were prepared using TRIzol reagent (Life Technologies) and Ambion WT expression kit (Life Technologies, Darmstadt, Germany) and hybridized to Mouse Gene 1.0 ST transcript cluster arrays (Affymetrix, Santa Clara, USA). Data normalization was performed using the Robust Multichip Average (RMA) method as described[29]. The Linear Models for Microarray Data (LIMMA) package was used to compute differentially regulated probe sets. The R 2.12.2 and 3.0.1 software and routines from the Bioconductor biostatistics software repository were used for all statistical analyses[30]. Heatmaps were generated with the heatmap.2 function (r-package gplots) using the hclust function for hierarchical clustering.[31] All the expression data are publicly available on the Gene Expression Omnibus website (GSE58853).

Quantitative TaqMan real-time PCR

Expression levels of the C/A fusion transcript, *Hoxa5, Hoxa7, Hoxa9, Hoxa10, HoxB4, Meis1,* and *Bmi1* genes were determined by TaqMan real-time PCR (primers and probes Table S1). cDNA was prepared using the Thermoscript RT-PCR system (Life Technologies). Expression levels of the murine beta-actin and 18srRNA genes were used as internal controls (Life Technologies). Expression levels were calculated using the ΔC_T method.

Analysis of DJ and VDJ BCR rearrangments

Analysis of the DJ and VDJ rearrangements of the murine B cell receptor genes was performed using PCR based assay on genomic DNA of leukemic BM as described before[23].

Whole exome sequencing

Genomic DNA was prepared from unsorted BM cells of leukemic mice using the AllPrep DNA/RNA kit (Qiagen, Hilden, Germany). Tail snip genomic DNA from the same mice was used as germline control. 0.5-1 µg genomic DNA was fragmented to an average size of 250 bp using a Bioruptor sonicator (Diagenode, Seraing, Belgium). Paired-end sequencing library preparation and exon capturing were performed using the SureSelectXT Mouse All Exon kit (Agilent, Waldbronn, Germany). Exome libraries were sequenced with 76-bp paired-end reads on a Genome Analyzer IIx (Illumina). Sequences were mapped to the NCBI Mouse mm9 reference genome and somatic mutations were called by comparing leukemic samples to all germline control

samples. Variants annotated as polymorphisms in the dbSNP v128 database were removed from further analysis. Non-synonymous somatic variants detected in leukemic samples were verified by Sanger sequencing using a 3500xL Genetic Analyzer (Life Technologies).

Statistical analysis

Peripheral blood counts and percentages of immune cell subsets in peripheral blood, bone marrow, and spleen were compared between mice using two-tailed unpaired Student's t-tests.

Vav-Cre-mediated expression of the C/A fusion gene causes leukemia

To examine the requirements for C/A-induced leukemia in unperturbed hematopoiesis, we generated the R26LSLCA mouse model (Fig S1), allowing the tissue-specific expression of the C/A fusion from the Rosa26 promoter following Cre-mediated excision of a loxP site flanked transcriptional stop cassette (Fig S2). To achieve pan-hematopoietic expression of the C/A fusion, R26LSLCA mice were crossed with Vav-Cre mice to give rise to heterozygous CA+/Vav-Cre+ progeny. To achieve B cell-specific expression of C/A, R26LSLCA mice were crossed with Mb1-Cre and CD19-Cre mice which resulted in CA+/Mb1-Cre+ and CA+/CD19-Cre+ offspring, respectively.

All the CA+/Vav-Cre+ mice (n=23) developed leukemia with a median latency of 354 days (Figure 1A). Neither CA+/Mb1-Cre+ (n=25) nor CA+/CD19-Cre+ (n=20) mice developed any signs of leukemia after an observation period of 500 days. We were able to analyze 15 of 23 CA+/Vav-Cre+ mice *post mortem*. Eight animals were found dead in their cage, 6 of these mice showed splenomegaly, suggesting a death from leukemia. The remaining 2 mice showed autolysis. Leukemic mice exhibited splenomegaly (data not shown) and leukocytosis with a 7.5 fold increased white-blood cell count (WBC) in the PB over control animals (median of $94.2x10^6$ cells/ml (range 15-532 $x10^6$ cells/ml) vs. $12.5x10^6$ cells/ml (range 8-$22x10^6$ cells/ml), p=0.0193) (Figure 1B). Leukemic mice were anemic with a 2.8 fold reduced red-blood cell count (RBC) (median of $3.8x10^9$ cells/ml (range 0.6-9.2 $x10^9$ cells/ml) vs. $11x10^9$ cells/ml (range 7-$16x10^9$ cells/ml), p<0.0001) (Figure 1B). Leukemic blasts

accumulated in PB smears, BM and Spl cytospins (Figure 1C), and infiltrated multiple internal organs (Figures 1D and Figure 2A and B).

Leukemic blasts in CA+/Vav-Cre+ mice show co-expression of myeloid and lymphoid markers

Organ-infiltrating leukemic blasts of moribund CA+/Vav-Cre+ mice stained positive for the myeloid markers MPO and ASDCL (Figure 2A) and negative for lymphoid markers CD3 and TdT (Figure 2B). Flow cytometry of PB, BM, and Spl cells revealed increased percentages of myeloid cells in CA+/Vav-Cre+ leukemic mice compared to control mice (Figures 3). Mac1+ cells comprised a median of 70% (range 45-92%), 29% (range 7-71%), and 45% (range 26-64%) of total PB, BM, or Spl, respectively. This is a statistically significant increase over wild type mice ($p<0.0001$) (Figure 3B). The percentage of $CD3^+$ T cells was reduced in all organs of leukemic CA+/Vav-Cre+ mice.

In 6 out of 15 (40%) leukemic CA+/Vav-Cre+ mice, a median of 37% (range 35-47%), 32% (range 20-51%), and 43% (range 40-49%) cells in the PB, BM and Spl, respectively, co-stained for Mac1 and the B cell marker B220 (Figures 2B, 3). This biphenotypic population of cells was absent in the other 9 leukemic mice. We called the former 6 leukemias “biphenotypic” and the latter 9 leukemias “myeloid”. We did not detect pronounced changes in the expression of other surface markers like cKit, CD24, F4/80, and CD19 between the two groups of leukemia however Gr1 and cKIT were low in the biphenotypic group (Figure S3)

The CA+/Mb1-Cre+ and the CA+/CD19-Cre+ mice did not show significant changes in their hematopoietic organs (Figure S4) even though their B cells expressed C/A at comparable levels to the leukemic BM cells from CA+/Vav-Cre+ mice (Figure S5).

Both biphenotypic and myeloid CA leukemias have clonal B cell receptor DJ rearrangements

When we performed PCR assays to detect B cell receptor (BCR) DJ rearrangements in the genomic DNA of four myeloid and four biphenotypic leukemias, clonal DJ rearrangements were discovered in all 8 samples (Figure 4), and one of the biphenotypic leukemias (Leu57) also showed a clonal VJ rearrangement (Figure S6).

Pan-hematopoietic C/A expression causes a leukemia specific gene signature

To evaluate the impact of the C/A fusion on gene regulation, we compared gene expression profiles (GEPs) of BM cells from leukemic (n=8) and pre-leukemic (appr. 2 months old) (n=3) CA+/Vav-Cre+ mice to that of wild type mice (n=3) (Comparisons 2 and 3 in Figure S7 and Figure 5A). We identified 1930 probe sets (1524 with gene names) significantly (p value <0.001) deregulated between leukemic and wild type BM (up-regulated 807 (676 with gene names), down-regulated 1123 (848 with gene names) in leukemic BM) (Figure S8, Table S2), and 100 probe sets (38 with gene names) between preleukemic and wild type controls (up-regulated 36, down-regulated 64 in the preleukemic BM) (Figure S9, Table S3). 25 probesets of 38 with gene names were commonly deregulated in comparison 2 and 3 and included the up-regulation of 6 *Hoxa* cluster genes and the *Hox* cofactor *Meis1*, and the down-regulation of various genes belonging to the killer cell lectin-like receptor (*Klr*) gene family (Figure 5A).

The comparison of GEPs of B cells from preleukemic CA+/Vav-Cre+ mice (appr. 2 months old) to B cells from wild type mice (Comparison 4 Figure S7) yielded only 14

significantly deregulated probe sets (p value < 0.001). Importantly, these included the up-regulation of *Hoxa3*, *Hoxa5*, *Hoxa6*, and *Meis1* (Table S4).

Comparing GEPs of CALM/AF10 expressing B cells from non-leukemic CA+/Mb1-Cre+ mice to that of wild type B cells (Comparison 5 Figure S7), a total of 32 probe sets was found significantly deregulated (p value < 0.005; p value is less stringent). Eight annotated probe sets were up-regulated and 6 were down-regulated (Table S5). Again *Hoxa3*, *Hoxa5* and *Hoxa6* and *Meis1* were found up-regulated in the C/A expressing non-leukemic B cells.

The high expression levels of *Hoxa5*, *Hoxa7*, *Hoxa9*, *Hoxa10* and *Meis1* in leukemic bone marrow and spleen of CA+/Vav-Cre+ mice were confirmed by quantitative RT-PCR (Figure 5B). In contrast to the microarray, we did not see significantly higher expression of any *Hoxa* gene or *Meis1* in B cells from CA+/Mb1-Cre+ mice (Figure 5B) using realtime RT-PCR, probably due to a very moderate increase in expression.

Myeloid and biphenotypic C/A-driven leukemias show distinct GEPs

To gain further insight into the two phenotypically different CA leukemias, we sorted Gr1+/Mac1+ BM cells from myeloid leukemia (n=4) and B220+/Mac1+BM cells from biphenotypic leukemia (n=4) using FACS and compared their GEPs (Comparison 1 in Figure S7, Figure S10, Table S6). We detected 343 deregulated probe sets (p value <0.001) of which 251 had a gene annotation. Differentially regulated genes reflected the immunophenotypic differences between the two groups. Interestingly in biphenotypic leukemic we detected upregulation of T-lymphoid-specific genes like *Dpp4* (CD26), *Cd7*, and *Btla*.

Whole exome sequencing reveals additional somatic mutations in C/A-driven murine leukemia which can also be found in human C/A-positive leukemia

The long latency to leukemia onset in CA+/Vav-Cre+ mice together with the apparent necessity of the C/A fusion to be expressed in very early stages of hematopoiesis prompted us to search for additional mutations in the CA leukemias. To identify leukemia specific somatic variants, we performed whole exome sequencing (WES) and compared the exomes of 7 CA+/Vav-Cre+ leukemic samples with the corresponding germline controls. We identified between 1 and 6 somatic point mutations and indels per sample in the 5 exomes with the highest coverage and blast counts (Table 1, Figure S11). As expected, we observed a strong correlation between the number of mutations identified and the sequence coverage and blast percentages (Tables S7 and S8, Figure S12). Interestingly, we found two distinct point mutations in the same gene, 4930595M18Rik, in leukemic exomes Leu51 and Leu95. In addition, we found mutations in two known leukemia driver genes: in *Ptpn11* (exon 13) and in codon 838 (D838) in the tyrosine kinase domain (TKD) of the *Flt3* gene.

Interestingly, *PTPN11* mutations were also found in 2 of 7 CALM/AF10 positive leukemia patients and in the CALM/AF10 positive cell line U937, which were subjected to gene panel sequencing (10 genes) (Figure S13, Table S9).

Discussion

AML clonally originates from a single transformed cell[32] - often called the target of transformation or COL. In the established leukemia, only a small population of tumor cells, often referred to as the 'leukemia propagating cells' (LPCs), 'leukemia initiating cells' (LICs) or LSCs are clonogenic in culture[33] and able to propagate and maintain the disease *in vivo,*[34]. The target of transformation can either be hematopoietic stem cells (HSCs) or committed progenitors, depending on the oncogene and the assay system used[35,36]. MLL-AF9, for example, can confer self-renewal properties to common myeloid progenitors (CMP) and granulocyte-macrophage progenitors (GMP)[35], BCR/ABL can only transform LT-HSCs[37], and *MN1* can transform HSCs and CMPs but not GMPs[38]. To study the target of transformation of the CALM/AF10 fusion gene, we chose a knock-in approach, which does not suffer from some of the shortcomings of the BMT leukemia models like retroviral insertional mutagenesis, polyclonality and a highly perturbed hematopoietic system after the transplant.

The association of C/A fusion with ALL as well as AML suggests that an immature hematopoietic cell is the likely target for transformation by C/A. Using a murine BMT model, we have previously shown that in C/A-driven AML the LSC was characterized by the expression of the B220, DJ rearrangements in the BCR genes and the absence of myeloid markers[23]. In the transgenic model of Caudel and colleagues, the C/A fusion is driven from the Vav-promoter[21]. In this model, the leukemic phenotype is similar to our previous BMT model[23] and the CA+/Vav-Cre+ mice in the present study. The leukemias were either myeloid or biphenotypic with the expression of both Mac1 and B220. The CA+/Vav-Cre+ and the transgenic Vav-C/A[21] models have comparable latencies. However, the Vav-C/A transgenic mice showed a much lower leukemia penetrance (about 50%) compared to our model (100%). This difference

might be strain specific or due to the fact that CA+/Vav-Cre+ mice are expressing C/A from the Rosa26 promoter and not from a randomly integrated Vav-C/A transgenic construct.

The co-expression of B220, which is expressed from the pro-B cell stage onwards, and myeloid markers appears to be common in C/A-induced murine and human leukemia[21,23]. Since C/A-induced biphenotypic leukemic cells do not express the B cell markers CD79a, IgM or CD19, the aberrant expression of B220 might be a direct effect of C/A-mediated activation of the transcription of B220. In our previous BMT model the leukemic clones recurrently showed clonal D_H-J_H rearrangement, suggesting that malignant transformation might have occurred at or after the pro-B cell stage[23]. However, whether the B lymphoid phenotype in this model was really due to the transformation of a B cell progenitor or because the C/A fusion favors the differentiation along the B lymphoid lineage, was not known. To distinguish between these two possibilities, we developed mouse strains that would allow lineage specific as well as pan-hematopoietic C/A expression from a C/A allele knocked into the Rosa26 locus that can be activated by Cre-mediated excision of a transcriptional stop cassette. In CA+/Mb1-Cre+ mice the expression of the C/A fusion begins at the very early pro-B cell stage with the activity of the CD79a (Igα/Mb1) promoter and in CA+/CD19-Cre+ mice the expression of the C/A fusion starts from the pre-B cell stage. When Hobeika et al. tested the recombination efficiency of Mb1-Cre mice using Rosa-floxed EYFP reporter mouse line, along with the strong YFP signals in the B cells from the early Pro B cell stage onward they reported seeing low levels of YFP+ T cells in thymus, spleen and lymph nodes[25]. This strongly suggests that Mb1-Cre mediated recombination can take place in the very early progenitors that are not

irreversibly committed to the B cell fate, probably to some extent in CLPs. In a very recent work, Duque-Afonso and colleagures have conditionally expressed the E2A-PBX1 fusion protein using Mb1-Cre and CD19-Cre and achieved B-ALL in 59% and 7% of the mice, respectively. This shows that cells expressing Mb1 and even CD19 are still capable of malignant transformation[39].

Interestingly, expression of C/A after either Mb1-Cre or CD19-Cre-mediated recombination did not result in leukemia development. Pan-hematopoietic expression of C/A after Vav-Cre-mediated recombination led to leukemia with a median latency of about 1 year. These results show that C/A is only leukemogenic when expressed in the correct cellular context, most probably in a HSC or an early hematopoietic progenitor. In addition, these results clearly shows that early pro B cells, which are similar in differentiation stage to the LSC of CA leukemias identified in the BMT model, cannot be transformed into leukemic cells. Thus, at least in CA leukemias there is a clear distinction between the target of transformation or COL and the LSC. Further, and almost certain proof that the COL and LSC are distinct in CALM/AF10 leukemias would have to come from extensive serial and limiting dilution bone marrow transplantation experiments in the Vav-Cre/Rosa26-CA leukemia model. The B cell-like phenotype, including the presence of Igh-DJ rearrangements, of the C/A LSC would then support the hypothesis that the expression of the C/A fusion might drive or favor differentiation along the B lymphoid lineage (Figure 6).

In line with earlier studies[21,40,41], C/A expression led to increased expression of *Hoxa* cluster genes and the *Hox* co-factor *Meis1* in BM and spleen samples of leukemic mice. Up-regulated expression of these genes in the BM and B cells of young, pre-

leukemic CA+/Vav-Cre+ mice suggests this event to be an early step in C/A-induced transformation. A slightly higher expression of certain *Hoxa* cluster genes and *Meis1* was also observed in the B cells of CA+/Mb1-Cre+ mice. Our results, together with previous observations made by other groups, clearly indicate that the overexpression of AbdB like Hox proteins and Meis1 are important factors in C/A-mediated leukemogenesis. However, as overexpression of *Hoxa* cluster genes and *Meis1* was also observed in the preleukemic mice and in mice that would never develop leukemia it is clear that this overexpression of *Hoxa* cluster genes and *Meis1* is not sufficient for leukemia development.

The long latency of C/A-driven leukemia in our and previous mouse model strongly suggests the requirement for collaborating genetic events. Whole exome sequencing (WES) identified several somatic mutations in 5 out of 8 leukemic exomes in the setting of unperturbed hematopoiesis without the confounding effects of retroviral insertional mutagenesis. These mutations affected, among others, the *Flt3* and the *Ptpn11* genes. An association between *Flt3* and *Ptpn11* (Tyrosine-protein phosphatase non-receptor type 11) mutations and murine C/A-driven leukemias has already been reported[22,41]. We detected a C>G point mutation resulting in a D>H substitution at codon 838 of the murine *Flt3* gene. This codon corresponds to D835 in human *FLT3* and is the main site of tyrosine kinase domain (TKD) activating mutations of *FLT3* in human AML[42]. *PTPN11* is a negative regulator of the Ras pathway, and mutations in exon 3 and exon 13 of *PTPN11* are found in patients with AML, myelodysplastic syndrome and juvenile myelomonocytic leukemia (JMML)[43]. We found a missense mutation in exon 13, codon 507, which encodes the catalytic domain of *PTPN11*, leading to a G>V amino acid change, a mutation described at the

corresponding human codon G503 in JMML[43]. Interestingly, in a small series of 7 CALM/AF10 positve ALL and AML patients as well as the U937 cell line we found *PTPN11* exon3 mutations in two patient samples and the U937 cell line (3/8 = 38%), indicating that the activation of the RAS pathway might be an important factor in CA-driven leukemias as previously described by others[22] and that our CA leukemia mouse model not only recapitulates the immunophenotype but also collaborating genetic events of human CA-positive leukemias.

Other mutations affected genes encoding for the Rho GTPase activating protein Arhgap42, the DNA repair protein Ap5s1, the Golgi-associated protein Grasp65, which regulates spindle dynamics and is important for cell division[44], and Moesin (*Msn*), an ERM family protein that cross-links plasma membranes and actin skeletons, and plays a role in cancer progression/metastasis[45,46]. We identified one somatic mutation in *Ush2a*, a gene recently identified as recurrently mutated in splenic marginal zone B cell lymphoma[47]. Additionally, two different non-synonymous mutations in 4930595M18Rik in two mice were identified. This gene codes for an RNA recognition motif and a RING finger domain. It should be noted that there was a strong correlation between the number of somatic mutations detected in the leukemia exomes and the coverage and blast percentage of the individual exome in our WES experiments (Figure S12). One should bear in mind that WES is not able detect all leukemia-relevant genomic changes, like copy number changes or translocations, and a more complete genomic analysis (WES with a higher coverage, or whole genome sequencing), with all likelihood, would have uncovered additional genetic alterations and mutations in the two leukemias (Leu74 and Leu31) where no mutations were found using WES. It should also be noted that our very

small CALM/AF10 patient cohort was only sequenced using a limited gene panel (10 genes) so that a greater overlap between human and murine collaborating mutations cannot be expected.

In summary, our knock-in mouse model provides evidence that the target cell of transformation or the COL for C/A-mediated transformation is a HSC or a very early multipotent progenitor cell (MPP) and that this cell is distinct from the C/A LSC. Furthermore, our WES results identified additional mutations that are required to collaborate with C/A to drive full-blown leukemia. Earlier studies from our laboratory have shown C/A to reduce global histone H3 lysine 79 methylation, favoring chromosomal instability and possibly other mutational events[48]. Additionally studies from our group showed that the expression of C/A leads to the down-regulation of DNA repair genes and changes in cell cycle regulation[49]. We propose that the expression of the C/A fusion makes cells more prone to acquire additional mutations. These mutations will have more time to accumulate in HSCs rather than in more committed MPPs as their lifetime is limited. Furthermore, HSCs are more sensitive to mutagenesis than committed progenitors[50]. The development of both myeloid and biphenotypic leukemias in our model is in line with the `lineage promiscuity' model, which states that leukemias arising from HSCs can still differentiate into the different lineages[51,52], and further strengthens our assumption that such early cells are the target of transformation in C/A-induced leukemia.

Acknowledgments

We are thankful to Frank Constantini from Columbia University, USA, for kindly providing the pBigT and Rosa26PA vectors, to Dimitris Kioussis from MRC National Institute for Medical Research, UK, and Matthias Kieslinger from Helmholtz Zentrum München, Germany, for the Vav-Cre mouse line, to Elias Hobeika and Michael Reth, both from MPI Freiburg, Germany, for the mb1Cre mouse line (referred to as Mb1-Cre in this manuscript), and to Klaus Rajewsky from MDC Berlin, Germany, for the CD19-Cre mouse line. This work was supported by Deutsche Forschungsgemeinschaft grants to SKB, EW, KS and UZ (SFB 684 A6, A12, and A13). SKB and MC are supported by Leukaemia & Blood Cancer New Zealand and the family of Marijanna Kumerich.

Authorship Contributions:

SD and SKB designed the experiments. SD and BT performed experiments. AK, MRS, MD generated the knock-in R26LSLCA mouse line, and EW supervised this work. SV and PAG analyzed WES data. TH analyzed GEP data. BK performed the FACS sorting. LQM performed histology and IHC. SKB, AG and HB operated the sequencing platform. SK performed the sequencing experiments. KS contributed to the manuscript. UZS supervised the animal experiments. AV, KM and M R-T performed and analyzed the gene panel sequencing. SD, MC, and SKB analysed data and wrote the manuscript. SKB supervised the project.

Conflict of Interest Disclosures

Authors declare no conflict of interest.

References

1. Dreyling MH, Martinez-Climent JA, Zheng M, Mao J, Rowley JD, Bohlander SK. The t(10;11)(p13;q14) in the U937 cell line results in the fusion of the AF10 gene and CALM, encoding a new member of the AP-3 clathrin assembly protein family. *Proceedings of the National Academy of Sciences of the United States of America* 1996; **93**(10): 4804-9.
2. Dreyling MH, Schrader K, Fonatsch C, et al. MLL and CALM are fused to AF10 in morphologically distinct subsets of acute leukemia with translocation t(10;11): both rearrangements are associated with a poor prognosis. *Blood* 1998; **91**(12): 4662-7.
3. Carlson KM, Vignon C, Bohlander S, Martinez-Climent JA, Le Beau MM, Rowley JD. Identification and molecular characterization of CALM/AF10fusion products in T cell acute lymphoblastic leukemia and acute myeloid leukemia. *Leukemia* 2000; **14**(1): 100-4.
4. Jones LK, Chaplin T, Shankar A, et al. Identification and molecular characterisation of a CALM-AF10 fusion in acute megakaryoblastic leukaemia. *Leukemia* 2001; **15**(6): 910-4.
5. Abdou SM, Jadayel DM, Min T, et al. Incidence of MLL rearrangement in acute myeloid leukemia, and a CALM-AF10 fusion in M4 type acute myeloblastic leukemia. *Leukemia & lymphoma* 2002; **43**(1): 89-95.
6. Narita M, Shimizu K, Hayashi Y, et al. Consistent detection of CALM-AF10 chimaeric transcripts in haematological malignancies with t(10;11)(p13;q14) and identification of novel transcripts. *British journal of haematology* 1999; **105**(4): 928-37.
7. Nakamura F, Maki K, Arai Y, Nakamura Y, Mitani K. Monocytic leukemia with CALM/AF10 rearrangement showing mediastinal emphysema. *American journal of hematology* 2003; **72**(2): 138-42.
8. Salmon-Nguyen F, Busson M, Daniel M, Leblanc T, Bernard OA, Berger R. CALM-AF10 fusion gene in leukemias: simple and inversion-associated translocation (10;11). *Cancer genetics and cytogenetics* 2000; **122**(2): 137-40.
9. Kumon K, Kobayashi H, Maseki N, et al. Mixed-lineage leukemia with t(10;11)(p13;q21): an analysis of AF10-CALM and CALM-AF10 fusion mRNAs and clinical features. *Genes, chromosomes & cancer* 1999; **25**(1): 33-9.
10. Silliman CC, McGavran L, Wei Q, Miller LA, Li S, Hunger SP. Alternative splicing in wild-type AF10 and CALM cDNAs and in AF10-CALM and CALM-AF10 fusion cDNAs produced by the t(10;11)(p13-14;q14-q21) suggests a potential role for truncated AF10 polypeptides. *Leukemia* 1998; **12**(9): 1404-10.
11. Kobayashi H, Hosoda F, Maseki N, et al. Hematologic malignancies with the t(10;11)(p13;q21) have the same molecular event and a variety of morphologic or immunologic phenotypes. *Genes, chromosomes & cancer* 1997; **20**(3): 253-9.
12. Asnafi V, Radford-Weiss I, Dastugue N, et al. CALM-AF10 is a common fusion transcript in T-ALL and is specific to the TCRgammadelta lineage. *Blood* 2003; **102**(3): 1000-6.
13. Tebar F, Bohlander SK, Sorkin A. Clathrin assembly lymphoid myeloid leukemia (CALM) protein: localization in endocytic-coated pits, interactions with clathrin, and the impact of overexpression on clathrin-mediated traffic. *Molecular biology of the cell* 1999; **10**(8): 2687-702.
14. Klebig ML, Wall MD, Potter MD, Rowe EL, Carpenter DA, Rinchik EM. Mutations in the clathrin-assembly gene Picalm are responsible for the hematopoietic and iron metabolism abnormalities in fit1 mice. *Proceedings of the National Academy of Sciences of the United States of America* 2003; **100**(14): 8360-5.
15. Chaplin T, Ayton P, Bernard OA, et al. A novel class of zinc finger/leucine zipper genes identified from the molecular cloning of the t(10;11) translocation in acute leukemia. *Blood* 1995; **85**(6): 1435-41.
16. Aasland R, Gibson TJ, Stewart AF. The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends in biochemical sciences* 1995; **20**(2): 56-9.

17. Saha V, Chaplin T, Gregorini A, Ayton P, Young BD. The leukemia-associated-protein (LAP) domain, a cysteine-rich motif, is present in a wide range of proteins, including MLL, AF10, and MLLT6 proteins. *Proceedings of the National Academy of Sciences of the United States of America* 1995; **92**(21): 9737-41.
18. Deshpande AJ, Rouhi A, Lin Y, et al. The clathrin-binding domain of CALM and the OM-LZ domain of AF10 are sufficient to induce acute myeloid leukemia in mice. *Leukemia* 2011; **25**(11): 1718-27.
19. Okada Y, Jiang Q, Lemieux M, Jeannotte L, Su L, Zhang Y. Leukaemic transformation by CALM-AF10 involves upregulation of Hoxa5 by hDOT1L. *Nature cell biology* 2006; **8**(9): 1017-24.
20. Greif PA, Tizazu B, Krause A, Kremmer E, Bohlander SK. The leukemogenic CALM/AF10 fusion protein alters the subcellular localization of the lymphoid regulator Ikaros. *Oncogene* 2008; **27**(20): 2886-96.
21. Caudell D, Zhang Z, Chung YJ, Aplan PD. Expression of a CALM-AF10 fusion gene leads to Hoxa cluster overexpression and acute leukemia in transgenic mice. *Cancer research* 2007; **67**(17): 8022-31.
22. Caudell D, Harper DP, Novak RL, et al. Retroviral insertional mutagenesis identifies Zeb2 activation as a novel leukemogenic collaborating event in CALM-AF10 transgenic mice. *Blood* 2010; **115**(6): 1194-203.
23. Deshpande AJ, Cusan M, Rawat VP, et al. Acute myeloid leukemia is propagated by a leukemic stem cell with lymphoid characteristics in a mouse model of CALM/AF10-positive leukemia. *Cancer cell* 2006; **10**(5): 363-74.
24. Georgiades P, Ogilvy S, Duval H, et al. VavCre transgenic mice: a tool for mutagenesis in hematopoietic and endothelial lineages. *Genesis* 2002; **34**(4): 251-6.
25. Hobeika E, Thiemann S, Storch B, et al. Testing gene function early in the B cell lineage in mb1-cre mice. *Proceedings of the National Academy of Sciences of the United States of America* 2006; **103**(37): 13789-94.
26. Rickert RC, Roes J, Rajewsky K. B lymphocyte-specific, Cre-mediated mutagenesis in mice. *Nucleic acids research* 1997; **25**(6): 1317-8.
27. Morse HC, 3rd, Anver MR, Fredrickson TN, et al. Bethesda proposals for classification of lymphoid neoplasms in mice. *Blood* 2002; **100**(1): 246-58.
28. Kogan SC, Ward JM, Anver MR, et al. Bethesda proposals for classification of nonlymphoid hematopoietic neoplasms in mice. *Blood* 2002; **100**(1): 238-45.
29. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003; **4**(2): 249-64.
30. (2008) RDCT. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, www.R-project.org. 2008.
31. Gregory R. Warnes BB, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber Andy Liaw, Thomas Lumley, Martin Maechler, Arni Magnusson, Steffen Moeller, Marc Schwartz, Bill Venables. gplots: Various R programming tools for plotting data. . R package version 2.14.1. http://CRAN.R-project.org/package=gplots. 2014.
32. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000; **100**(1): 57-70.
33. Buick RN, Till JE, McCulloch EA. Colony assay for proliferative blast cells circulating in myeloblastic leukaemia. *Lancet* 1977; **1**(8016): 862-3.
34. Lapidot T, Sirard C, Vormoor J, et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* 1994; **367**(6464): 645-8.
35. Krivtsov AV, Twomey D, Feng Z, et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* 2006; **442**(7104): 818-22.
36. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature medicine* 1997; **3**(7): 730-7.
37. Kovacic B, Hoelbl A, Litos G, et al. Diverging fates of cells of origin in acute and chronic leukaemia. *EMBO molecular medicine* 2012; **4**(4): 283-97.

38. Heuser M, Yun H, Berg T, et al. Cell of origin in AML: susceptibility to MN1-induced transformation is regulated by the MEIS1/AbdB-like HOX protein complex. *Cancer cell* 2011; **20**(1): 39-52.
39. Duque-Afonso J, Feng J, Scherer F, et al. Comparative genomics reveals multistep pathogenesis of E2A-PBX1 acute lymphoblastic leukemia. *J Clin Invest* 2015.
40. Dik WA, Brahim W, Braun C, et al. CALM-AF10+ T-ALL expression profiles are characterized by overexpression of HOXA and BMI1 oncogenes. *Leukemia* 2005; **19**(11): 1948-57.
41. Novak RL, Harper DP, Caudell D, Slape C, Beachy SH, Aplan PD. Gene expression profiling and candidate gene resequencing identifies pathways and mutations important for malignant transformation caused by leukemogenic fusion genes. *Experimental hematology* 2012; **40**(12): 1016-27.
42. Small D. FLT3 mutations: biology and treatment. *Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program* 2006: 178-84.
43. Kratz CP, Niemeyer CM, Castleberry RP, et al. The mutational spectrum of PTPN11 in juvenile myelomonocytic leukemia and Noonan syndrome/myeloproliferative disease. *Blood* 2005; **106**(6): 2183-5.
44. Sutterlin C, Polishchuk R, Pecot M, Malhotra V. The Golgi-associated protein GRASP65 regulates spindle dynamics and is essential for cell division. *Molecular biology of the cell* 2005; **16**(7): 3211-22.
45. Zhu X, Morales FC, Agarwal NK, Dogruluk T, Gagea M, Georgescu MM. Moesin is a glioma progression marker that induces proliferation and Wnt/beta-catenin pathway activation via interaction with CD44. *Cancer research* 2013; **73**(3): 1142-55.
46. Clucas J, Valderrama F. ERM proteins in cancer progression. *Journal of cell science* 2014; **127**(Pt 2): 267-75.
47. Parry M, Rose-Zerilli MJ, Gibson J, et al. Whole exome sequencing identifies novel recurrently mutated genes in patients with splenic marginal zone lymphoma. *PloS one* 2013; **8**(12): e83244.
48. Lin YH, Kakadia PM, Chen Y, et al. Global reduction of the epigenetic H3K79 methylation mark and increased chromosomal instability in CALM-AF10-positive leukemias. *Blood* 2009; **114**(3): 651-8.
49. Mulaw MA, Krause A, Deshpande AJ, et al. CALM/AF10-positive leukemias show upregulation of genes involved in chromatin assembly and DNA repair processes and of genes adjacent to the breakpoint at 10p12. *Leukemia* 2012; **26**(5): 1012-9.
50. Mohrin M, Bourke E, Alexander D, et al. Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. *Cell stem cell* 2010; **7**(2): 174-85.
51. Greaves MF, Chan LC, Furley AJ, Watt SM, Molgaard HV. Lineage promiscuity in hemopoietic differentiation and leukemia. *Blood* 1986; **67**(1): 1-11.
52. So CW, Karsunky H, Passegue E, Cozzio A, Weissman IL, Cleary ML. MLL-GAS7 transforms multipotent hematopoietic progenitors and induces mixed lineage leukemias in mice. *Cancer cell* 2003; **3**(2): 161-71.

Figure legends

Figure 1: **CA+/Vav-Cre+ mice develop leukemia with 100% penetrance. A**: Kaplan-Meier survival plot of CA+/Vav-Cre+ mice (n=23), CA+/Mb1-Cre+ mice (n=25) and CA+/CD19-Cre+ mice (n=20). All CA+/Vav-Cre+ mice developed leukemia with a median latency of 354 days. Leukemic mice presented with hunched posture, ruffled fur, difficulties in breathing, and lethargy. Post mortem analysis was possible for 15 mice. The dots indicate 8 censored mice that could not be analyzed. None of CA+/Mb1-Cre+ or CA+/CD19-Cre+ mice developed leukemia. **B**: Dot plot of WBC and RBC counts in the peripheral blood (PB) of leukemic CA+/Vav-Cre+ mice and normal wild type mice. **C**: Giemsa-May-Grünwald staining of peripheral blood smear (PB Smear) and cytospins of bone marrow (BM Cytospin) and splenocytes (Spleen Cytospin). Presence of immature cells with large blastic chromatin and bluish cytoplasm is noted in all hematopoietic tissues. D: H&E staining of peripheral organs shows the presence of infiltrating leukemic blasts in spleen, thymus, lung, kidney, pancreas, salivary gland, and liver.

Figure 2: **Immunohistochemistry shows CA+/Vav-Cre+ mice develop myeloid and biphenotypic leukemia. A**: Immunohistochemical staining of thymus, spleen and liver sections from myeloid and biphenotypic leukemia for MPO (myeloperoxidase) reveals both groups of leukemic mice to have myeloid marker positive blasts infiltrating hematopoietic and non-hematopoietic organs. **B**: Immunohistochemical staining of the spleen from myeloid and biphenotypic leukemias for ASDCL (AS-D-chloracetate esterase), CD3, B220, and TdT (terminal desoxynucleotide transferase).

In both leukemias, CD3-, Tdt- ASDCL+ blasts are visible. In biphenotypic leukemia, B220+ blasts have infiltrated the red pulp of the spleen, whereas normal white pulp B cells are the only B220 positive cells in myeloid leukemia. Positive ASDCL staining indicates both leukemic samples to be of a mature granulocyte phenotype.

Figure 3: **Immunophenotype analysis of myeloid and biphenotypic leukemia using flow cytometry. A**: Representative flow cytometry scatter plots for the expression of Mac1 and B220 on cells from peripheral blood (PB), bone marrow (BM) and spleen (Spl) from wild type mice and CA+/Vav-Cre+ mice with myeloid or biphenotypic leukemia. **B**: Comparative immunophenotypic analysis of PB, BM and Spl cells from wild type mice and CA+/Vav-Cre+ mice with myeloid or biphenotypic leukemia. (Mean ± SD).

Figure 4: **D_HJ_H rearrangement of B cell receptor locus in myeloid and biphenotypic CA leukemias**: **A**. The murine IgH locus showing three members of the VH gene family (VH7183, VH558, VHQ52), 3 elements of the D_H gene family (DFL, DSP and DQ52) and the four members of JH (JH1, JH2, JH3, JH4). The arrows indicate the positions of two degenerate forward primers and one reverse primer used to detect the D_H-J_H rearrangements in the IgH locus. The diagram is not drawn to scale.

B: Clonal D_H-J_H rearrangement in the leukemic bone marrow cells isolated from 4 mice with myeloid (no 42, 52, 72, and 74) and biphenotypic leukemias (no 32, 57, 81 and 29). Genomic DNA from Mac1+ positive bone marrow cells and the 32D cell line served as IgH naive germline controls. Genomic DNA from B220+ spleen cells served as a control for polyclonally rearranged B cells.

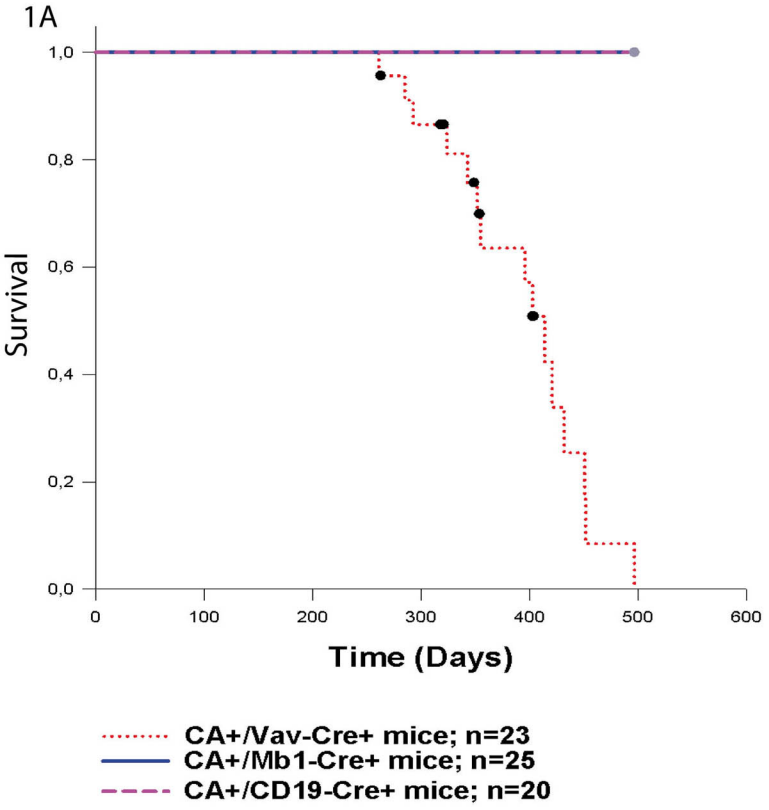
Figure 5: **Gene expression analysis. A**: Left panel: Venn diagram of genes significantly ($p \leq 0.001$) deregulated in preleukemic (n=3) and leukemic BM (n=8) in comparison to wildtype BM (n=3). The numbers refer to the probe sets for which a gene annotation was available. 25 genes were found deregulated in both comparisons, including 6 *Hoxa* cluster genes and the *Hox* co-factor *Meis1.* Right panel: Heatmap of these 25 deregulated genes in preleukemic and leukemic BM as compared to wild type BM. **B**: qRT-PCR of *Hoxa* cluster genes, *Meis1*, and *Bmi1* transcript levels. Expression of *Hoxb4* was assayed as a control outside the *Hoxa* cluster. Expression was normalized to β-actin and is presented as ΔC_T. * $p \leq 0.02$; ** $p \leq 0.002$. Expression levels were compared between leukemic CA+/Vav-Cre+ BM (n=4) and wild type BM (n=3), leukemic CA+/Vav-Cre+ spleen (n=4) and wild type spleen (n=3), and CA+/Mb1-Cre+ B cells (n=3) and wild type B cells (n=4). Note that *Hoxa7*, *Hoxa9* and *Hoxa10* transcripts were detected in leukemic spleen but not in wild type spleen. *Hoxa7*, *Hoxa9* and *Hoxa10* transcripts were not detected in either CA+/Mb1-Cre+ B cells or wild type B cells.

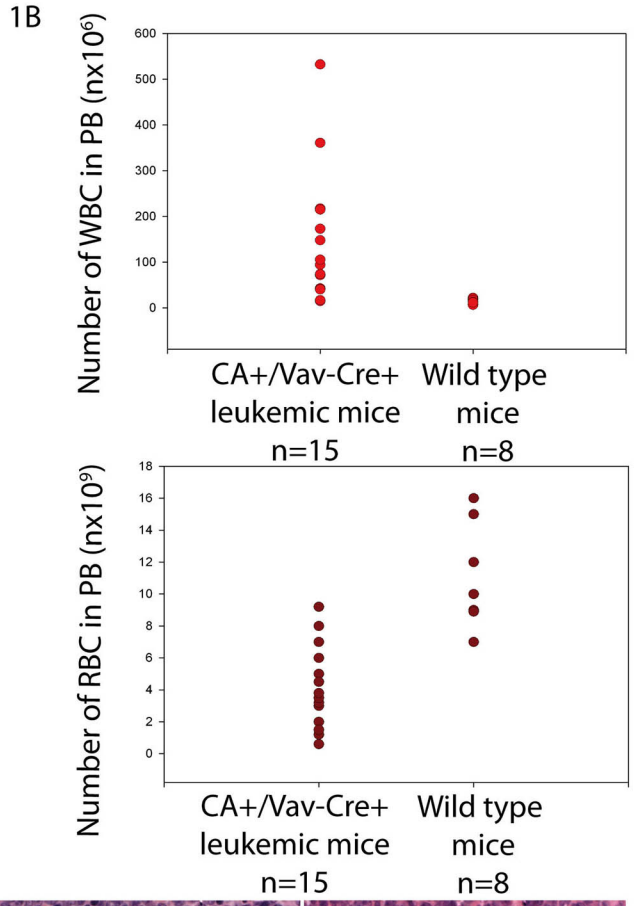
Figure 6: **Diagram showing the relationship between the CALM/AF10 target of transformation or cell of origin of leukemia (COL) and the leukemic stem cell (LCS) within the hematopoietic hierarchy. T**he genotype of the mice which develop leukemia (CA+/Vav-Cre+) is shown in red, the genotype of the mice that do not develop leukemia is shown in green (CA+/Mb1-Cre+, CA+/CD19-Cre+).

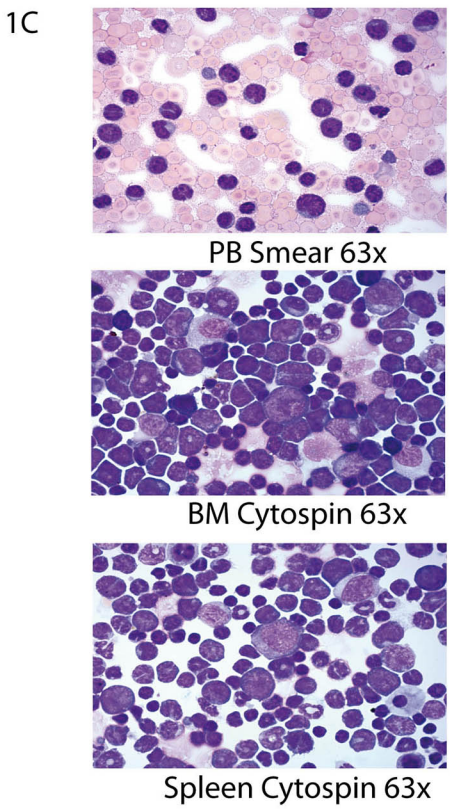
Mouse	Gene	Genomic position (mm9)	Reference allele	Variant allele	Amino acid change	Ensembl transcript
Leu 29	Nos1ap	chr1:172259273	G	A	T195M	ENSMUST00000160466
	Ush2a	chr1:190734992	C	T	P4224L	ENSMUST00000027905
	Gtpbp6	chr5:110534100	G	*/-GATGA	Q302Gfs	ENSMUST00000127628
	Flt3	chr5:148152827	C	G	D838H	ENSMUST00000110547
Leu 42	Ap5s1	chr2:131038711	C	T	R206C	ENSMUST00000110210
	Lrrc7	chr3:157823465	T	C	Q1201R	ENSMUST00000106044
	Reln	chr5:21733431	C	A	V224L	ENSMUST00000162622
	Arhgap42	chr9:9006619	G	A	P592L	ENSMUST00000093893
	Soga2	chr17:66728596	C	T	R885Q	ENSMUST00000086693
	Msn	chrX:93355514	C	T	R295C	ENSMUST00000117399
Leu 51	Zranb1	chr7:140174581	C	G	L632V	ENSMUST00000106157
	4930595M18Rik	chrX:78703310	T	C	D11G	ENSMUST00000080083
Leu 57	Ugt2b35	chr5:87430163	A	C	K83Q	ENSMUST00000031186
Leu 95	Ptpn11	chr5:121593103	C	A	G507V	ENSMUST00000100770
	Gorasp1	chr9:119841887	A	G	V66A	ENSMUST00000035099
	Csnka2ip	chr16:64478340	C	T	E52K	ENSMUST00000089279
	4930595M18Rik	chrX:78665692	A	T	S583T	ENSMUST00000080083

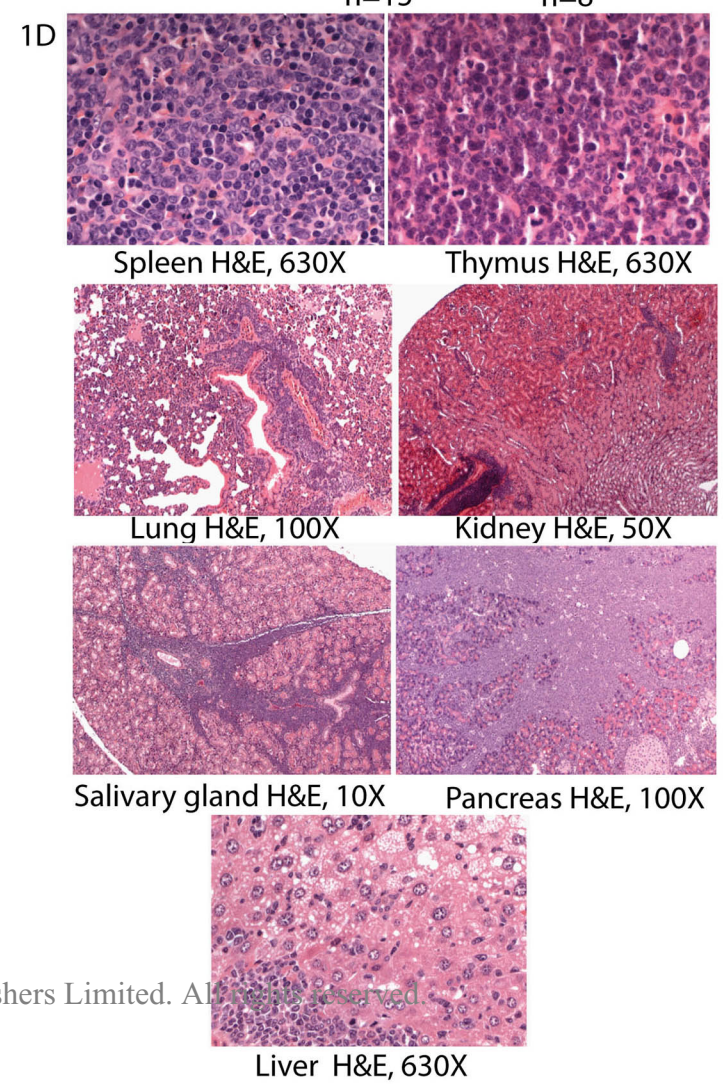
Table 1: **Somatic mutations identified in leukemic samples by WES.**

Read depth in leukemia	Variant read count	Variant frequency (%)	Mutated in hematopoietic and lymphoid malignancies (COSMIC) (%)	Leukemia Sub-type
22	7	31.82	0.08	Biphenotypic
39	8	20.51	Jan-62	
30	16	53.33	0	
23	7	30.43	21.14	
				Myeloid
12	9	75	0	
46	14	30.43	0.15	
14	4	28.57	Feb-15	
40	15	37.5	0	
29	10	34.48	0.08	
18	13	72.22	0	
				Myeloid
20	5	25	0	
44	12	27.27	0	
62	28	45.16	0	Biphenotypic
29	16	55.17	Jun-57	Biphenotypic
57	28	49.12	0	
28	13	46.43	0	
60	32	53.33	0	



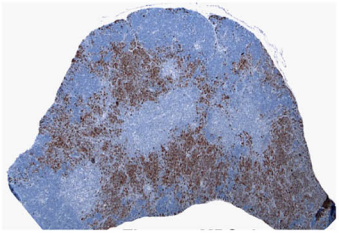




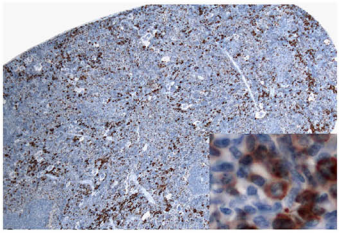


2A

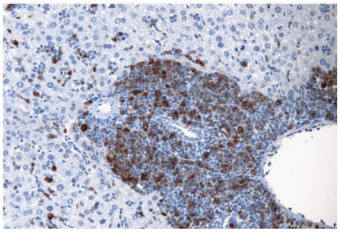
Myeloid leukemia



Thymus MPO 25x

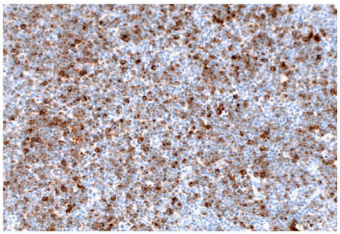


Spleen MPO 50x, Insert: 630x

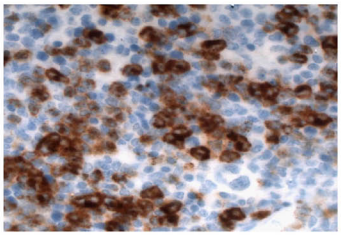


Liver MPO 200x

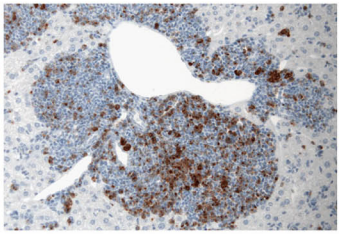
Biphenotypic Leukemia



Thymus MPO 200x



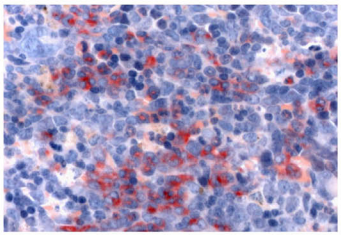
Spleen MPO 630x



Liver MPO 100x

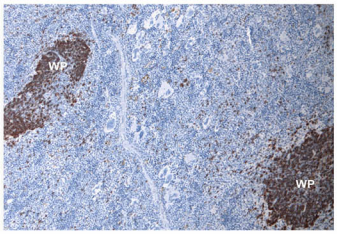
2B

Myeloid leukemia



Spleen ASDCL 630x

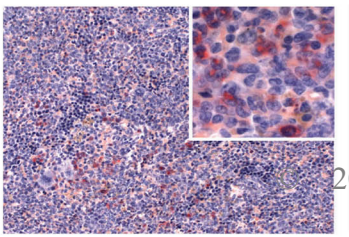
Spleen CD3 50x



Spleen B220 100x

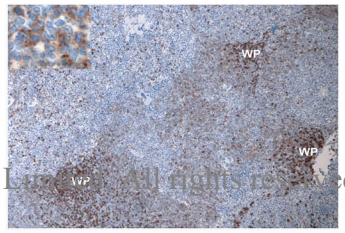
Spleen Tdt 50x

Biphenotypic Leukemia

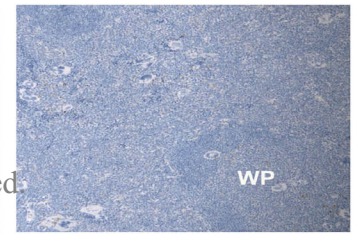


Spleen ASDCL 200x, Insert:630x

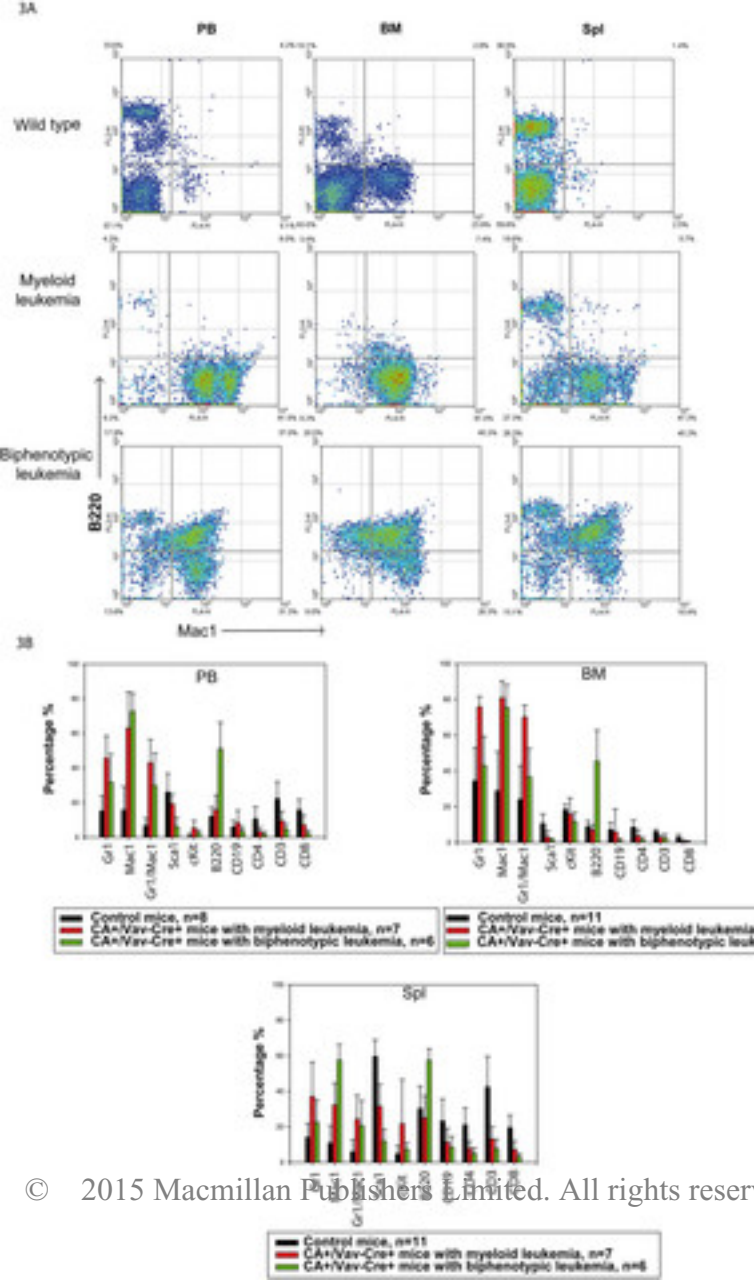
Spleen CD3 50x



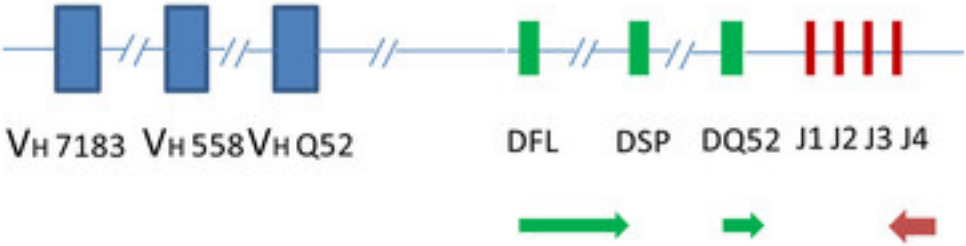
Spleen B220 100x, Insert: 630x



Spleen Tdt 100x



A



B