De novo activating mutations drive clonal evolution and enhance clonal fitness in KMT2A-rearranged leukemia

Axel Hyrenius-Wittsten, Mattias Pilheden, Helena Sturesson, Jenny Hansson, Michael P. Walsh, Guangchun Song, Julhash U. Kazi, Jian Liu, Ramprasad Ramakrishan, Cristian Garcia-Ruiz, Stephanie Nance, Pankaj Gupta, Jinghui Zhang, Lars Rönstrand, Anne Hultquist, James R. Downing, Karin Lindkvist-Petersson, Kajsa Paulsson, Marcus Järås, Tanja A. Gruber, Jing Ma & Anna K. Hagström-Andersson

Activating signaling mutations are common in acute leukemia with KMT2A (previously MLL) rearrangements (KMT2A-R). These mutations are often subclonal and their biological impact remains unclear. Using a retroviral acute myeloid mouse leukemia model, we demonstrate that FLT3ITD, FLT3N676K, and NRASG12D accelerate KMT2A-MLLT3 leukemia onset. Further, also subclonal FLT3N676K mutations accelerate disease, possibly by providing stimulatory factors. Herein, we show that one such factor, MIF, promotes survival of mouse KMT2A-MLLT3 leukemia initiating cells. We identify acquired de novo mutations in Braf, Cbl, Kras, and Ptpn11 in KMT2A-MLLT3 leukemia cells that favored clonal expansion. During clonal evolution, we observe serial genetic changes at the KrasG12D locus, consistent with a strong selective advantage of additional KrasG12D. KMT2A-MLLT3 leukemias with signaling mutations enforce Myc and Myb transcriptional modules. Our results provide new insight into the biology of KMT2A-R leukemia with subclonal signaling mutations and highlight the importance of activated signaling as a contributing driver.

1 Division of Clinical Genetics, Department of Laboratory Medicine, Lund University, 221 84 Lund, Sweden. 2 Division of Molecular Hematology, Department of Laboratory Medicine, Lund University, 221 84 Lund, Sweden. 3 Department of Pathology, St. Jude Children’s Research Hospital, Memphis, TN 38105, USA. 4 Division of Translational Cancer Research, Department of Laboratory Medicine, Lund University, 223 63 Lund, Sweden. 5 Department of Oncology, St. Jude Children’s Research Hospital, Memphis, TN 38105, USA. 6 Department of Computational Biology, St. Jude Children’s Research Hospital, Memphis, TN 38105, USA. 7 Lund Stem Cell Center, Department of Laboratory Medicine, Lund University, 221 84 Lund, Sweden. 8 Division of Oncology, Skane University Hospital, Lund University, 221 85 Lund, Sweden. 9 Department of Pathology, Skane University Hospital, Lund University, 221 85 Lund, Sweden. 10 Medical Structural Biology, Department of Experimental Medical Science, 221 84 Lund University, Lund, Sweden. Correspondence and requests for materials should be addressed to A.K.Höm-A. (email: Anna.Hagstrom@med.lu.se)
Genetic rearrangements of the Histone-lysine N-methyl-transferase 2A gene (KMT2A, previously MLL) is seen in approximately 10% of human acute leukemia. However, in acute lymphoblastic leukemia (ALL) occurring in infants less than 1 year of age, a leukemia subtype with a particularly poor prognosis, close to 80% have KMT2A-rearrangements (KMT2A-R)\(^2\). We have previously shown that infant KMT2A-R ALL has a strikingly low number of mutations with an average of 2.2 non-silent single-nucleotide variants per case, and 1.3 in the dominant leukemia clone\(^2\). About half of these mutations occurred in kinase/phosphoinositide 3-kinase (PI3K)/RAS signaling pathways, suggesting that they are important cooperating events\(^2\). Kinase/P3K/RAS pathway mutations are also common in KMT2A-R ALL and acute myeloid leukemia (AML) arising in older children and adults\(^2,4,6,7\).

In infant KMT2A-R ALL, most of the mutations that deregulate protein components of signal transduction networks are subclonal, as indicated by their mutant allele frequencies (MAF) \(<0.3\). In addition, some cases harbor multiple activating mutations, as indicated by their mutant allele frequencies (MAF). Further, analyses of paired diagnostic-relapse samples have shown that signaling mutations may be lost, maintained, or gained at relapse\(^4,8,9\). Thus, there is a delicate selection process that shapes clonal evolution during leukemia progression, treatment, and relapse. Mouse models have shown a potent genetic cooperativity between signaling mutations and KMT2A-Rs, with the most prominent pathophysiological combinatorial gain being acceleration of disease onset\(^10-17\). However, none of these models have recapitulated the biology of subclonal activating mutations.

The mechanisms by which signaling mutations affect leukemogenesis, whether present at high or low MAF, and the biological processes controlling and shaping the intricate selection of evolving leukemic clones, remain poorly understood. To gain a better understanding of these processes, we studied the functional importance of constitutively activated FLT3- and RAS-signaling in KMT2A-mediated leukemogenesis in mouse. We focused our efforts on FLT3\(^{ITD}\), FLT3\(^{N676K}\), and NRAS\(^{G12D}\), as FLT3 and NRAS are the most common targets of mutations that deregulate signal transduction in AML, and studied clonal evolution of leukemia cells carrying subclonal activating mutations over time.

Herein, we show that also subclonal activating mutations, as demonstrated by FLT3\(^{N676K}\), accelerate KMT2A-MLLT3 leukemia onset. Moreover, we identify acquired de novo activating mutations in known cancer-associated genes in KMT2A-MLLT3 leukemia cells that favored clonal expansion, indicating a strong

**Fig. 1** FLT3- and RAS-signaling mutations accelerate AML onset. a Kaplan–Meier survival curves of mice transplanted with bone marrow (BM) cells cotransduced with KMT2A-MLLT3 and either FLT3\(^{ITD}\), FLT3\(^{N676K}\), or NRAS\(^{G12D}\) (n = 7 for all groups), showing accelerated disease onset for mice receiving KMT2A-MLLT3 and either of the activating mutations (P = 0.0002 for KMT2A-MLLT3 + FLT3\(^{ITD}\), P = 0.0002 for KMT2A-MLLT3 + FLT3\(^{N676K}\), and P = 0.0004 for KMT2A-MLLT3 + NRAS\(^{G12D}\), as compared to KMT2A-MLLT3 + Empty-GFP, Mantel–Cox log-rank test). b Flow cytometric analysis of BM from sacrificed mice showed that a majority of cells were GFP\(^+\)mCherry\(^+\) in mice co-expressing KMT2A-MLLT3 and an activating mutation. c Hematoxylin-eosin stained sections from bone marrow, liver, and spleen (original magnification 200x, scale bar 0.1 mm, for bone marrow and 40x, scale bar 0.5 mm, for spleen and liver). The architecture of the spleen is effaced and the red pulp is expanded mainly due to expansion of immature myeloid cells. In the liver, periportal, perisinusoidal and intrasinusoidal extensive infiltrates of immature hematopoietic cells were noted. d Kaplan–Meier curves for secondary recipients transplanted with primary leukemic splenocytes showing that only KMT2A-MLLT3 + NRAS\(^{G12D}\) sustained a significant difference in disease latency when compared to KMT2A-MLLT3 + Empty-GFP (P < 0.0001, Mantel–Cox log-rank test). ***P ≤ 0.001, ****P ≤ 0.0001; ns, not significant
selective pressure for activated signaling as a cooperating event in KMT2A-R leukemogenesis.

**Results**

**FLT3- and RAS-signaling mutations accelerate AML onset.** We here studied constitutively activated FLT3- and RAS-signaling in KMT2A-R leukemogenesis. First, we focused on the impact of FLT3ITD, FLT3N676K, or NRASG12D when present as a dominant leukemia clone at disease manifestation. Mouse hematopoietic stem and progenitor cells (c-Kit+ cells) were co-transduced with retroviral vectors expressing KMT2A-MLLT3-mCherry and FLT3ITD-GFP, FLT3N676K-GFP, NRASG12D-GFP, or a GFP control vector (Empty-GFP) (Supplementary Fig. 1a, b). Expression of mutant NRAS and FLT3 in the mouse leukemia cell line Ba/F3 resulted in elevated phosphorylated p-AKT and p-ERK1/2, as well as p-STAT5 for FLT3ITD (Supplementary Fig. 1c).

Transduced cells were transplanted unfraccionated and at transplantation, only a minor fraction of cells co-expressed KMT2A-MLLT3 and NRASG12D, FLT3ITD, or FLT3N676K (Supplementary Fig. 1d). Despite this, all of them accelerated leukemia onset (median latency of 13, 23, 26 days, respectively, versus 50 days for KMT2A-MLLT3 alone; P = 0.0004, P = 0.0002, and P = 0.0002, respectively. Mantel–Cox log-rank test) (Fig. 1a), consistent with a competitive growth advantage for these cells. In agreement with this, a majority of cells in the bone marrow (BM) contained both a signaling mutation and KMT2A-MLLT3 (Fig. 1b, Supplementary Fig. 1e), supporting a prominent genetic cooperativity between the KMT2A-R and activating FLT3 and RAS mutations, including the recently identified FLT3N676K 4,18. Mice displayed leukocytosis, splenomegaly, expression of CD11b and Gr-1 in BM cells, and a similar burden of leukemic granulocyte-macrophage-like-progenitors10,19 (Supplementary Fig. 1f–i and Supplementary Fig. 2a–c). The BM, spleen, and liver showed infiltration of blasts and cytopenia, and peripheral blood (PB) an increase of blasts and progenitors (>20%) (Fig. 1c). Thus, mice succumbed to AML20. Leukemias with an activating mutation showed presence of p-ERK1/2, and also p-STAT5 for those with FLT3ITD. All leukemias displayed similar phosphorylation levels of P38 (Supplementary Fig. 3).

Leukemic cells isolated from the spleen of moribund mice gave rise to secondary malignancies identical to the primary disease in sublethally irradiated recipients, and with significantly reduced disease latency (Fig. 1d and Supplementary Fig. 4a–d). A continued enrichment for KMT2A-R leukemic cells carrying FLT3ITD, FLT3N676K, or NRASG12D was seen, consistent with a persistent competitive advantage for these cells (Supplementary Fig. 4e). Disease latency of secondary KMT2A-MLLT3 recipients was similar to those with an activating mutation, suggesting that they had adapted a comparable leukemic phenotype to those with activating mutations.

**Subclonal FLT3N676K accelerates AML onset.** In infant KMT2A-R ALL, a majority of the kinase/PI3K/RAS-pathway mutations are present in subclones4,22, raising the possibility that cells with

---

**Figure 2** Subclonal FLT3N676K accelerates AML onset. a Flow cytometric analysis on BM cells from primary KMT2A-MLLT3 + FLT3N676K recipients revealed the presence of either a dominant clone (>50% GFP+mCherry+, n = 17), or a subclone (<50% GFP +mCherry+, n = 7) of KMT2A-MLLT3-mCherry + FLT3N676K-GFP expressing cells. b Kaplan–Meier curves of mice transplanted with KMT2A-MLLT3 and FLT3N676K (n = 24) or with Empty-GFP vector control (n = 21) showing accelerated disease onset for both dominant and subclonal KMT2A-MLLT3 + FLT3N676K leukemias as compared to KMT2A-MLLT3 + Empty-GFP (both P < 0.0001. Mantel–Cox log-rank test). c Evolution of FLT3N676K-GFP+ cells within the mCherry+ leukemic population between primary (1°) spleen (SPL) and secondary (2°) BM showed a significant expansion of KMT2A-MLLT3-mCherry + FLT3N676K-GFP cells (n = 24, P < 0.0001. Paired t-test) but not of KMT2A-MLLT3-mCherry + Empty-GFP cells (n = 21, P = 0.1468. Paired t-test). d Cell cycle analysis of primary leukemia cells showed a higher cell cycle rate for KMT2A-MLLT3-mCherry harboring FLT3N676K-GFP (n = 7) as compared to KMT2A-MLLT3 + Empty-GFP (n = 7). Error bars are s.d. *P ≤ 0.01, ****P ≤ 0.0001; ns, not significant.
activating mutations could affect the growth or survival of other leukemia cells. We hypothesized that a low initial number of cells co-expressing an activating mutation and KMT2A-MLLT3 together with a several fold higher number of cells expressing only KMT2A-MLLT3 would allow for the formation of distinct subclones and performed three additional experiments accordingly (1:28, 1:41, and 1:156 for KMT2A-MLLT3 + FLT3*N676K, KMT2A-MLLT3 only, and KMT2A-MLLT3 + FLT3*N676K, KMT2A-MLLT3 alone, respectively) (Supplementary Fig. 5a). We focused our analyses on our series of infant and childhood leukemia cells. We hypothesized that a low initial number of cells co-expressing an activating mutation and KMT2A-MLLT3 together with a several fold higher number of cells expressing only KMT2A-MLLT3 would allow for the formation of distinct subclones and performed three additional experiments accordingly (1:28, 1:41, and 1:156 for KMT2A-MLLT3 + FLT3*N676K, KMT2A-MLLT3 only, and KMT2A-MLLT3 alone, respectively) (Supplementary Fig. 5a). We focused our analyses on FLT3*N676K, the most common FLT3 mutation in our series of infant and childhood KMT2A-R leukemia.

Combined analysis of the transplantations showed that despite the low number of cells that co-expressed KMT2A-MLLT3 and...
FLT3\textsuperscript{N676K} at transplantation, these mice displayed accelerated AML onset (median latency 34 versus 50 days for KMT2A-MLLT3 alone, P < 0.001). Mantel–Cox log-rank test (Supplementary Fig. 5b–f). Interestingly, the BM of KMT2A-MLLT3 + FLT3\textsuperscript{N676K} recipients were comprised of either (1) a dominant clone, >50% (17/24 mice), or (2) a subclone, <50% (7/24 mice) of KMT2A-MLLT3 + FLT3\textsuperscript{N676K} expressing cells, hereafter referred to as dominant clone and subclone, respectively (Fig. 2a). Importantly, when recipients were divided based on the fraction of FLT3\textsuperscript{N676K} expressing leukemic cells, mice with subclonal FLT3\textsuperscript{N676K} still succumbed to disease at an earlier onset (dominant clone, 32 days, P < 0.0001; subclone, 37 days P < 0.0001. Mantel–Cox log-rank test) (Fig. 2b). There was a weak but significant correlation between clone size and disease latency (r = -0.586, P = 0.0026). Spearman’s rank correlation coefficient (Supplementary Fig. 5g). Our data therefore suggest that subclonal activating mutations accelerate KMT2A-R leukemogenesis.

**Clonal evolution of FLT3\textsuperscript{N676K} containing leukemic cells.** We next investigated the evolution of cells co-expressing FLT3\textsuperscript{N676K} and KMT2A-MLLT3 in secondary subletally irradiated recipients. Primary splenocytes were used for these experiments and the fraction of cells co-expressing FLT3\textsuperscript{N676K} and KMT2A-MLLT3 was significantly smaller in spleen as compared to BM with 13/24 mice having a dominant clone and 11/24 a subclone (Supplementary Fig. 5h). Disease latency in secondary recipients was reduced and a small difference in survival was observed; median 16 days for KMT2A-MLLT3 + FLT3\textsuperscript{N676K} and 21 days for KMT2A-MLLT3 alone, P = 0.0293 (Mantel–Cox log-rank test) (Supplementary Fig. 5i). Secondary recipients were then divided based on the fraction of FLT3\textsuperscript{N676K} containing cells in the spleens of primary recipients, showing that size had a modest effect on disease latency (Supplementary Fig. 5j, k).

Clonal evolution of leukemia splenocytes from primary recipients in the BM of secondary recipients showed three distinct patterns. Most frequently, the KMT2A-MLLT3 + FLT3\textsuperscript{N676K} cells increased in size (21/24 mice) (Fig. 2c and Supplementary Fig. 5l). In concordance, FLT3\textsuperscript{N676K} expression induced a higher cell cycle rate (Fig. 2d). However, mice receiving cells from primary recipients with a very small KMT2A-MLLT3 + FLT3\textsuperscript{N676K} clone (≤10.1%, 3/24 mice) displayed a diverse evolution where one subclone (1.1%) was maintained in size, one (0.5%) disappeared, and one decreased in size from 10.1% to 0.1% (Fig. 2c and Supplementary Fig. 5m, n). Thus, cells containing an activating mutation typically had a persistent selective advantage also in established disease. However, in a few recipients, these cells failed to outgrow leukemia cells containing KMT2A-MLLT3 alone (see below).

**KMT2A-MLLT3 cells acquire de novo signaling mutations.** The failure of FLT3\textsuperscript{N676K} containing cells to expand in a subset of secondary recipients suggested that the KMT2A-MLLT3 only cells could have acquired de novo mutations. We therefore sequenced 41 selected genes that were either tyrosine kinases, within the PI3K/RAS pathways, epigenetic regulators, or recurrently mutated in KMT2A-R leukemia (see Methods and Supplementary Table 1) at a mean coverage of 3231 ± 1277 (s.d.) (Supplementary Data 1), in 62 primary (27 dominant and 7 subclonal KMT2A-MLLT3 and an activating mutation, and 28 KMT2A-MLLT3 alone) and 29 paired secondary recipients (derived from 15 dominant and 6 subclonal KMT2A-MLLT3 and an activating mutation, and 8 KMT2A-MLLT3 alone) (Supplementary Table 2). Notably, our analyses identified acquired de novo mutations in RAS pathway genes in 4/62 primary- and in 6/29 paired secondary recipients, all of which occurred in genes known to be mutated in human KMT2A-R leukemia (Ptpn11, Cbl, Braf, Kras) and all but one mutation (Cbl\textsuperscript{A308T}) occurred at amino acid positions known to be activating (Supplementary Fig. 6a–h and Supplementary Data 2). In every case, as determined by RNA sequencing, the acquired mutation was expressed (Supplementary Data 2).

Only 1/34 (3%) of the primary KMT2A-MLLT3 recipients that co-expressed FLT3\textsuperscript{ITD} + FLT3\textsuperscript{N676K}, or NRAS\textsuperscript{Q61D} had acquired a de novo mutation (Cbl\textsuperscript{A308T}) (Fig. 3a and Supplementary Data 2). Strikingly, Cbl\textsuperscript{A308T} was detected in the mouse in which the frequency of KMT2A-MLLT3 + FLT3\textsuperscript{N676K} cells decreased in size in the secondary recipient. The Cbl\textsuperscript{A308T} was subclonal in the primary (MAF 0.11), dominant in the secondary (MAF 0.37) and maintained in size in a tertiary recipient (Fig. 3a and Supplementary Data 2). Thus, in the primary recipient, KMT2A-MLLT3 cells acquired a Cbl\textsuperscript{A308T} that had a selective advantage over FLT3\textsuperscript{N676K} leukemia cells and gained clonal dominance in the secondary recipient. Cbl encodes a RING ubiquitin ligase that functions as a negative regulator of signaling pathways. The corresponding human residue of A308 (A310) is located in the highly conserved N-terminal SH2 domain within the tyrosine kinase-binding domain (Fig. 3b and Supplementary Fig. 6a, b), which mediates substrate specificity and is critical for substrate ubiquitination. Mutations occur over the whole region of CBL, with the RING domain being a hotspot; no A310 mutations have been reported in COSMIC\textsuperscript{25} (Supplementary Fig. 6b). The CBLG306E results in a complete loss of its ability to bind SYK and FLT3\textsuperscript{26,27}. Given the close proximity of A310 and G306 (Fig. 3b and Supplementary Fig. 6l), Cbl\textsuperscript{A308T} likely also interferes with binding and subsequent ubiquitination of target substrates, such as Syk and Flt3. Mutations in CBL cause factor independent growth of the IL3 dependent cell line Ba/F3 when overexpressed with certain receptor kinases such as FLT3\textsuperscript{28}. Using the same approach\textsuperscript{28–30}, we functionally demonstrated that expression of Cbl\textsuperscript{A308T} resulted in a significantly higher cell proliferation as compared to Cbl\textsuperscript{WT}, supporting that the A308T mutation interferes with the ability of Cbl to negatively regulate intracellular signaling (Supplementary Fig. 7a). Taken together, when the constitutively active mutation is present as a subclone, there is a selective benefit of acquiring activating mutations in leukemia cells lacking such a mutation.

Acquired mutations were more common in mice lacking a constitutively expressed activating mutation, with 3/28 (11%) primary and 5/8 secondary KMT2A-MLLT3 recipients containing such mutations. Mutations in Braf\textsuperscript{V637E}, Kras\textsuperscript{G12D}, and Ptpn11\textsuperscript{S506W} were identified in three primary and in their corresponding secondary recipient (Fig. 3c–f and Supplementary Data 2). In secondary transplants, the three KMT2A-MLLT3 leukemias with Braf\textsuperscript{V637E}, Kras\textsuperscript{G12D}, or Ptpn11\textsuperscript{S506W} were associated with accelerated disease (Fig. 3g). Further, a Ptpn11\textsuperscript{S506W} and a Ptpn11\textsuperscript{E69K} were identified in two additional secondary recipients, with no evidence of the mutant allele in their corresponding primary recipient (Fig. 3h, i and Supplementary Data 2). Using directed amplicon sequencing, additional secondary KMT2A-MLLT3 leukemias were screened for Braf\textsuperscript{V637E} (n = 14), Cbl\textsuperscript{A308T} (n = 14), Kras\textsuperscript{G12D} (n = 15), Ptpn11\textsuperscript{E69K} (n = 15), and Ptpn11\textsuperscript{S506W} (n = 13), but no mutations were identified. The Braf\textsuperscript{V637E} and Kras\textsuperscript{G12D} (human orthologues BRAF\textsuperscript{V600E} and KRAS\textsuperscript{G12D}) are well-described mutations that cause constitutively active signaling in humans and mice (Supplementary Fig. 6c–f)\textsuperscript{21,22}. The only recurrently mutated gene was the proto-oncogene tyrosine phosphatase Ptpn11 (encoding Shp-2), acquired in three separate leukemias, Ptpn11\textsuperscript{E69K} (n = 1) and Ptpn11\textsuperscript{S506W} (n = 2). These mutations are located in the C-terminal SH2 and protein tyrosine phosphorylation domain of Ptpn11, while the leukemogenic mutation (Ptpn11\textsuperscript{E69K}) is located in the N-terminal SH2 domain. The only recurrently mutated gene was the proto-oncogene tyrosine phosphatase Ptpn11 (encoding Shp-2), acquired in three separate leukemias, Ptpn11\textsuperscript{E69K} (n = 1) and Ptpn11\textsuperscript{S506W} (n = 2). These mutations are located in the C-terminal SH2 and protein tyrosine phosphorylation domain of Ptpn11, while the constitutively active mutation is present as a subclone, there is a selective benefit of acquiring activating mutations in leukemia cells lacking such a mutation.
phosphatase catalytic domains, known hotspots in human disease (E69 and S502) (Supplementary Fig. 6g, h). SHP-2 positively regulate RAS signaling and are seen in KMT2A-R acute leukemia of all ontogenies \(^{4,7}\). Mutations at E69 and S502 are known to increase SHP-2 phosphatase activity \(^{32,35}\), suggesting that also the mouse mutations would have similar effect. Co-expression of Ptpn11\(^{SS06W}\) and KMT2A-MLLT3 accelerated AML onset compared to control mice (median latency 27 vs. 44 days, \(P = 0.0005\). Mantel–Cox log-rank test) with a majority of KMT2A-MLLT3 cells containing Ptpn11\(^{SS06W}\) (Fig. 3) and Supplementary Fig. 7b–f). Thus, Ptpn11\(^{SS06W}\) is a cooperating mutation in KMT2A-R leukemogenesis.

The availability of a paired secondary leukemia allowed us to study the evolution of the acquired mutations. The Braf\(^{V637E}\) was subclonal at leukemia onset (MAF 0.20) and increased in size in the secondary recipient (MAF 0.30) (Fig. 3c and Supplementary Data 2). Similarly, Kras\(^{G12D}\) was subclonal (MAF 0.11) at leukemia manifestation and gained clonal dominance in the
Activating mutations enforce Myc and Myb modules. The transcriptional landscape of KMT2A-R leukemia has been comprehensively studied in mice\(^{19,32-35}\), however, insight into the effect of an activating mutation is limited\(^{33}\). We therefore performed RNA sequencing on 45 mouse leukemias (Supplementary Data 3) and also included data from purified normal immature hematopoietic subpopulations from the mouse (Supplementary Table 4)\(^{36}\). In addition, to study effects induced by activating mutations at the protein level, we performed mass spectrometry (MS)-based quantitative proteomic analysis by stable isotope labeling on 15 leukemias (Supplementary Data 3, 4).

Principal component analysis of purified normal hematopoietic subpopulations showed distinct clusters based on maturation stage and that the leukemias mostly resembled normal granulocyte–macrophage progenitors (GMPs) (Supplementary Fig. 8a)\(^{15}\). All leukemias had high expression of Hoxa9, Hoxa10, and Meis1, known KMT2A-R target genes, as compared to GMP cells (Supplementary Fig. 8b)\(^{19}\). Gene set enrichment analysis (GSEA)\(^{37}\) revealed enrichment of signatures from pediatric KMT2A-R AML (Supplementary Fig. 8c and Supplementary Data 5), supporting that mouse leukemias activate similar pathways as human KMT2A-R leukemia and is a representative model of this disease.

Despite a shared KMT2A-R leukemic signature, mutant FLT3 and NRAS expression also induced distinct gene expression profiles (GEPs) (Fig. 4a, Supplementary Fig. 8d–f and Supplementary Data 6). To link the GEPs induced by activating mutations in our mouse model to that seen in human disease, GSEA was performed using a gene signature associated with infants harboring KMT2A-AFF1 and an activating mutation\(^{4}\), which demonstrated enrichment of this signature in mice with activating mutations (Fig. 4b). This indicates that the transcriptional changes induced by activating mutations in human disease are preserved in the mouse. Notably, identified de novo mutations influenced the GEP for samples with mutations in the dominant leukemia clone, i.e., Kras and Ptpn11 (MAF 0.39–0.59), now clustering closely to those with a constitutively expressed mutation; this was not seen for mutations with MAF ≤ 0.30 (Fig. 4c).

Although FLT3\(^{1TD}\), FLT3\(^{N676K}\), and NRAS\(^{G12D}\) each had distinct GEPs, they induced similar biological processes at the gene and protein level, including a distinct enrichment of MYC signatures (Fig. 4d and Supplementary Data 7–14). The MYC regulatory network enhances embryonic stem cell (ESC) programs, but is separate from the pluripotency network\(^{38}\). Indeed, the signature evoked by activating mutations was specific to the MYC module\(^{38}\) and not linked to ESC self-renewal, given that no enrichment of the core and polyclonal modules\(^{39}\) was observed (Fig. 4e, Supplementary Fig. 8g, h, and Supplementary Data 15–23). In line with this, ribosomal and translational signatures, MYC influenced processes\(^{39}\), were enriched at the RNA and protein level (Supplementary Fig. 8i–k and Supplementary Data 24–26). The MYC signature likely reflect a worse prognosis\(^{38,40}\), in line with the rapid disease progression for these mice\(^{34,38,41}\) (Supplementary Fig. 9a and Supplementary Data 16–19). Myc is a direct target gene of MYB\(^{42}\), a transcription factor known to be a key regulator of hematopoietic stem and progenitor cells and to be an important mediator in KMT2A-R leukemia\(^{4,35}\). Further, Myb is associated with a leukemia maintenance signature in KMT2A-R leukemia models\(^{35}\) and this signature was enriched in leukemias with an activating mutation at the gene and protein level (Supplementary Fig. 9b, c and Supplementary Data 16–23). However, Myb was not part of the leading-edge genes, but was uniformly expressed in all leukemias, including those expressing KMT2A-MLLT3 alone Supplementary Fig. 9b–d, and Supplementary Data 17–19. The Myc and Myb signatures, mainly associated with committed progenitors as opposed to regenerative hematopoietic populations (Supplementary Fig. 9e, f). Enrichment of the Myc and Myb signatures was also observed in mouse leukemias with de novo Kras/Ptpn11 mutations (MAF 0.39–0.59) and in primary infant KMT2A-AFF1 ALL harboring activating mutations\(^{4}\) (Fig. 4f, Supplementary Fig. 9g–j, and Supplementary Data 27, 28).

To gain further insight into the transcriptional effect on downstream MEK/ERK signaling in leukemias with activating mutations, we selectively investigated if expression of FLT3\(^{1TD}\), FLT3\(^{N676K}\), and NRAS\(^{G12D}\) caused increased expression of known transcriptional output genes and negative feedback regulators of MEK/ERK signaling\(^{43}\). Indeed, enrichment of these genes were particularly strong in KMT2A-MLLT3 mouse leukemia with NRAS\(^{G12D}\) and in infant KMT2A-AFF1 ALL patients harboring activating mutations (Supplementary Fig. 9k and Supplementary Data 19, 28)\(^{4}\). However, enrichment was also seen for KMT2A-MLLT3 mouse leukemia with FLT3\(^{N676K}\) or Kras/Ptpn11 mutations (MAF 0.39–0.59) (Supplementary Fig. 9k and Supplementary Data 18, 27). This implicates an increased transcriptional activity of the MEK/ERK signaling pathway including negative feedback regulators in leukemia with signaling mutations, in particular those with FLT3\(^{N676K}\) and NRAS\(^{G12D}\), and that the mutant cells are insensitive to negative feedback control. The same enrichment was not observed for FLT3\(^{1TD}\) leukemias (Supplementary Fig. 9k and Supplementary Data 17).

KMT2A-MLLT3 recipients lacking an activating mutation had increased expression of genes involved in signal transduction (Fig. 4g) when compared to all leukemias that co-expressed KMT2A-MLLT3 and an activating mutation (Supplementary Fig. 8l). Consistent with this, GSEA revealed enrichment of intracellular signaling pathways including different MAPK pathways, both at gene- and protein level (Fig. 4h, Supplementary Fig. 8i–k and Supplementary Data 7–14, 24–26). This suggests that KMT2A-MLLT3 itself may cause sustained transcription and subsequent translation of genes involved in intracellular signaling.

MIF promotes survival of KMT2A-MLLT3 leukemia cells. Mice expressing subclonal FLT3\(^{N676K}\) succumb earlier to disease than those lacking such a mutation (Fig. 2b), raising the possibility that cells with an activating mutation, whether present at high or low MAF, provide factors that positively influence the growth or survival of leukemia cells. We therefore investigated the expression of 137 genes encoding cytokines or growth factors (Supplementary Data 29), revealing a differential expression of several
of these genes between leukemia cells with or without an activating mutation (Supplementary Fig. 10a). One cytokine commonly upregulated in KMT2A-MLLT3 leukemias co-expressing an activating mutation (FLT3ITD, FLT3N676K, or NRASG12D) as compared to those induced by KMT2A-MLLT3 alone, and normal healthy hematopoietic subpopulations, was the macrophage migration inhibitory factor (MIF), a lymphokine that stimulate pro-inflammatory processes (Fig. 5a and Supplementary Fig. 10b). Proteomic analyses confirmed a higher expression of MIF in a majority of KMT2A-MLLT3 leukemias with an activating mutation (Fig. 5b and Supplementary Fig. 11a). To establish the role of MIF in human KMT2A-R leukemia, we determined its expression in infant KMT2A-MLLT3 cells44 cultured ex vivo for 3 days with or without 500 ng/ml MIF. We therefore evaluated whether MIF stimulation maintained leukemia-initiating cells (LICs) ex vivo using a transplantation assay. For this purpose, 5000 serially transplanted and LIC-enriched KMT2A-MLLT3 c-Kit+ leukemia cells44 were cultured ex vivo for 3 days with or without MIF before transplantation into sublethally irradiated recipients (Supplementary Fig. 11b). Mice transplanted with MIF treated cells had significantly shorter disease latency as compared to no cytokine control mice (22.5 days versus 28 days, P ≤ 0.0001; ns, not significant).

Discussion
We lack a full understanding of the biological processes controlling and shaping the selection of evolving leukemia clones. Mouse models have proven powerful tools to uncover genetic interactions, and the spontaneous acquisition of mutations in genes that in humans promote leukemogenesis identified herein, highlights their remarkable capacity to recapitulate human disease. Our study demonstrates that an activating mutation present only in a small proportion of KMT2A-MLLT3 leukemia cells can
influence the rate of leukemogenesis and shorten time to leukemia onset. These data support a model whereby subclonal activating mutations are not bystanders, but influence leukemogenesis. This is the first time, to the best of our knowledge, that the functional importance of a defined leukemia subclone has been studied in the mouse. In addition, our data show that FLT3 \(^{\text{ITD}}\), the most frequent FLT3 mutation identified in our previous analysis of the mutational landscape of infant and childhood KMT2A-R leukemia, is a cooperating lesion in this disease.

The observation that subclonal activating mutations accelerate disease onset in the mouse, is consistent with our previous data showing that activating mutations, irrespective of their MAFs, associated with an average younger age at diagnosis in infants with KMT2A-AF7 ALL. Typically, subclonal FLT3/ITD cells had a continued proliferative advantage upon secondary transplantation. However, sometimes these cells failed to expand, and we identified spontaneous de novo signaling mutations in Ptpn11 \(^{E695/S656, Brk^{1637D}, Chk^{330T}, and Kras^{G12D}}\), in cells not carrying only KMT2A-MLLT3, thereby shifting clonal balance. Notably, further clonal evolution at the Kra loc region was evident upon secondary transplantation, consistent with duplication of the mutated allele and a strong selective advantage of additional Kras \(^{G12D} \text{D45F} \) (Fig. 3c). In human disease, data from large scale sequencing efforts have shown that patients may have more than one signaling mutation often present at different MAFs indicative of multiple leukemia clones. Our data are consistent with different activating mutations providing different fitness to evolving leukemia cells, thus influencing their selection potential. The lack of identified spontaneous mutations in cells co-expressing KMT2A–MLLT3 and FLT3/ITD, FLT3/ITD, or NRAS \(^{G12D} \) suggest that in the presence of a strong signaling mutation, the selective pressure for additional mutations within the pathway is not as prominent, at least not when the mutation is retrovirally overexpressed.

The fact that a subclonal activating mutation reduce leukemia latency raises the possibility that they secrete stimulating factors that influence other leukemia cells, as has been shown in solid tumor models with activating mutations. MIF, a pro-inflammatory cytokine, was highly expressed in KMT2A-MLLT3 leukemia with activating mutations both in mice and in primary infant ALL patients (Fig. 5c). MIF treatment ex vivo supported the survival of mouse KMT2A-MLLT3 LICs. Thus, MIF is likely one of many factors that mediate pro-leukemic signals in vivo. MIF can inactivate p53 and thereby suppress apoptosis and can also communicate with stromal cells, as recently described for primary AML cells, resulting in increased tumor cell survival. In addition, knock-out of Mif delayed development of chronic lymphocytic leukemia (CLL) in mice and, similar to our data, Mif extended survival of PLL cells in vitro.

It is well described that genetic aberrations shape GEPs in leukemia. Mutations in NRAS \(^{G12D} \), FLT3/ITD and FLT3/ITD further enforced Myc- and Myb-driven signatures previously associated with KMT2A-R leukemia. MYC levels are influenced by Ras and kinase signaling, in line with activating mutations extending this transcriptional program. Signalning mutations affecting the ERK pathway, in mice as well as in infant ALL, resulted in increased transcriptional output of the pathway including negative feedback regulators, suggesting that mutated cells are insensitive to negative regulation (Supplementary Fig. 9a). Interestingly, KMT2A-MLLT3 leukemias lacking a constitutively expressed signaling mutation, sustain or upregulate intracellular signaling pathways, such as RAC1 and MAPK, potentially as a contributing oncogenic event. Elevated RAC1 protein levels confer anti-apoptotic capacity in KMT2A-MLLT3 leukemia.

Taken together, these data provide a unique look into the mechanisms by which activating mutations affect leukemogenesis, also when present as a subclone. The data suggest that subclonal activating mutations, here demonstrated by FLT3/ITD, can influence leukemogenesis, possibly by providing pro-leukemic factors. In the absence of a constitutively active mutation, KMT2A-R leukemia cells may acquire spontaneous activating mutations, consistent with the importance of kinase/P13K/RAS pathways as a cooperative event in this disease. The acquired mutations typically conferred cells with a competitive advantage, altered clonal fitness, and favored clonal outgrowth. Our results provide insight into the mechanisms by which signaling mutations promote leukemogenesis and it will be important to determine whether interfering with intracellular signal transduction in human KMT2A-R leukemia, even in the absence of an activating mutation, may be a fruitful therapeutic approach.

**Methods**

**Vectors and virus production.** All vectors were based on the Murine Stem Cell Virus (MSCV) backbone (see schematics for the constructs). Retroviral vectors were made using the KMT2A-MLLT3-IRES-mCherry, MSCV-FLT3/ITD-IRES-GFP, MSCV-FLT3/ITD-IRES- GFP, MSCV-FLT3/ITD-IRES-GFP, MSCV-FLT3/ITD-IRES-GFP, MSCV-NRAS/G12D-IRES-eGFP, and MSCV-NRAS/G12D-IRES-eGFP. ORF cDNA clones for Chl (NM_007619) and Pmln1 (NM_011202) subcloned into MSCV-eGFP were constructed from GenScript (Piscataway, NJ, USA). To obtain Chl \(^{E695/S656, Brk^{1637D}, Chk^{330T}, and Kras^{G12D}}\) was performed using the QuickChange II XL Site Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer’s instructions using the following primers: Chl \(^{E695/S656} \) F5′- TGTTCACCATCGGATACAACTCACCAGAACC3′- Chl \(^{E695/S656} \) R5′- AGTGGGCTATTGGGTATGTTACTACCGATGGGAACA-3′. Pmln1 \(^{E695/S656} \), F5′- ACCATCCCCACCTCGGCAACGGC3′ and Pmln1 \(^{S656W} \), R5′- CGGTCCTCCAGGGTGATGT-3′ mutations were verified by Sanger sequencing. Empty vectors lacking the transgenes were used as controls. Retroviral supernatants were produced according to standard protocols using transient transfection of HEK 293T cells and viral containing medium was harvested 36 h later and stored at −80°C.

**Retroviral transduction.** Whole bone marrow (BM) cells were harvested from 8–12 weeks old C57Bl/6xB6S/J mice and CD117+ cells were enriched by magnetic activated cell sorting (Miltenyi Biotech, Bergisch Gladbach, Germany). Cells were pre-stimulated for 48 h in StemSpan SFEM (Stemcell Technologies, Grenoble, France) supplemented with 100 units/mL Penicillin and 100 μg/mL Streptomycin (Thermo Scientific, South Logan, UT, USA), 50 ng/mL mSCF, 50 ng/mL hIL6, 50 ng/mL hIL3 (all cytokines were purchased from Peprotech, Rocky Hill, NJ, USA). Retroviral co-transduction was performed by loading a defined mixture of retroviral supernatant on Retronectin (Takara, Otsu, Japan) pre-coated wells by centrifugation for 60 min at 1000×g at 4°C. Cells were then added, spun down at 3200×g for 5 min and incubated at 37°C for 15–18 h before bone marrow transplantation. A small aliquot of cells was saved in culture to determine the frequency of single and co-transduced cells by flow cytometry 48 h post transduction. IL3-dependent Ba/F3 cells (DSMZ, Braunschweig, Germany) were transduced in standard culture medium.

**In vivo bone marrow transplantation models.** C57Bl/6N Tac, C57Bl/6xB6S/JL mice were used throughout the experiments. C57Bl/6xB6S/JL, and B6S/JL mice were bred in-house and sex and age matched C57Bl/6N Tac were purchased from Harlan (Bar Harbor, ME). Eight bone marrow samples were harvested from the femurs of eight 12 weeks old and age matched female C57Bl/6N Tac mice were lethally irradiated with 900 cGy 18–20 h prior to transplantation and subsequently transplanted with 1 × 10⁶ unfractiacted co-transduced C57Bl/6xB6S/JL BM cells. For all primary transplantations, 400,000 freshly isolated B6S/JL whole BM cells were given as support, and recipient mice were given Ciprofloxacin (KRKA, Stockholm, Sweden) for 2 weeks post-irradiation in the drinking water. For secondary transplantations, 8–12 weeks old and age matched female C57Bl/6N Tac mice were sublethally irradiated with 450 cGy 6 h prior to transplantation and were subsequently transplanted with 1 × 10⁶ whole spleen cells from primary recipients. For all mice, when moribund, a PB sample was taken from vena saphena, and mice were then sacrificed by cervical dislocation. Upon sacrifice; one femur and a piece of the liver and spleen were saved for pathological evaluation. Remaining bones and spleen were made into single-cell suspension by manual trituration and BM and spleen cells were viably frozen. In addition, 1 × 10⁶ BM cells were lysed in RLT buffer (Qiagen, Hilden, Germany) supplemented with 1% β-mercaptoethanol (Sigma-Aldrich, St. Louis, MO, USA) and stored in −80°C before further DNA/RNA extraction. Complete blood cell count was performed on PB using Micros 60 CS

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-04180-1 | ARTICLE | (2018) 9:1770 | www.nature.com/naturecommunications
Histopathology. For immunohistochemistry, tissues were fixed in 4% formaldehyde (Merck, Darmstadt, Germany) and bones were subsequently decalciﬁed using formic acid (Histolab, Göteborg, Sweden) before parafﬁn embedment. For structural and morphological visualization, tissue sections were deparafﬁnized using xylene (WVR chemicals, Fontenay-sous-Bois, France) and decreasing concentrations of ethanol (95% and 70% Ronal, Findon, Sussex) and stained using Harris hematoxylin (Histolab) and erythrosine B (Merck) before pathological evaluation. Images were taken with a Nikon BS51 microscope (Nikon Instruments Inc., Melville, NY, USA).

Flow cytometry, cell cycle analysis, and cell sorting. Before flow cytometric analysis of primary mouse tissues (i.e., BM and spleen) red blood cells were lysed with ammonium chloride (Stemcell technologies). Dead cells were excluded using DragQ™ (1:1600, Biostatus, Shepshed, United Kingdom) or Fixable Viability Dye eFlour™ 780 (1:1000, Bioscience, San Diego, CA, USA). For cell cycle analysis, cells were ﬁxed and permeabilized using 1% formaldehyde (Merck) and 95% ethanol (CCS Healthcare AB) and stained with Ki-67 (1:40, clone REA183) (Miltenyi) and DAPI (1:50, Biolegend, San Diego, CA, USA). For immunophenotyping, the following antibodies were used: Ly-6G/Ly-6C/G-1 (1:800, clone R6C-8C5), CD11b (1:800, clone M1/70), CD45RA (1:800, clone RA3-62B), CD3e (1:800, clone 145-2C11), CD16/32 (1:200, clone 92), CD3e (1:200, clone MEC147), CD135 (1:200, clone AF210), Ly-6A/E-Ly-6C/ScA-1 (1:1000, clone D7C7), CD117 (1:100, clone ACK2), CD227 (78.25 clone A384), and a cocktail (1:50, clone 1:7A), Ly-6G/Ly-6C/G-1 (1:800, clone R6C-8C5); CD11b, clone M1/70; CD45RA/B220, clone RA3-62B; Ter-119, clone 119-Ter, all from Biolegend. Flow cytometric analysis and ﬂuorescence-activating cell sorting were performed on FACS Aria II Ulc or FACS Aria Fusion (BD, Franklin Lakes, NJ, USA). Data analysis was performed using FlowJo software (FlowJo, LLC, Ashland, OR, USA).

Ex vivo cultures. Establishment of mouse leukemic cell lines were done by culturing sorted GFP+mCherry™ (for KMT2A-MLLT3 + FLT3WT, KMT2A-MLLT3 + FLT3N676K, and KMT2A-MLLT3 + NRASQ61D) or mCherry™ (for KMT2A-MLLT3 + Empty-GFP) leukemic BM cells for 1 week in “C10” media, consisting of RPMI 1640 with 1-glutamine (Thermo Scientiﬁc), 10% fetal bovine serum (FBS, Thermo Scientiﬁc), 100 units/ml Penicillin and 100 µg/ml Streptomycin (Thermo Scientiﬁc), 55 µM L-mercaptoethanol (Sigma-Aldrich), 1 mM non-essential amino acids (Sigma-Aldrich), 1 mM sodium pyruvate (Thermo Scientiﬁc) and 10 mM HEPES (Thermo Scientiﬁc) supplemented with 100 µg/ml mSCF, 10 ng/ml hIL6 and 10 ng/ml mIL3 (Peprotech™), after which mSCF and hIL6 were withdrawn and mIL3 was decreased to 1 ng/ml for continuous culturing.

Validation of somatic mutations. Validation of each mutation was performed in matched primary, secondary recipient, and constitutional DNA, using PCR amplicon deep sequencing. Primers were designed using Primer-BLAST (sequences made to request). PCR was performed using Platinum® Ampliﬁer kit (Life Technologies) according to the manufacturer’s instructions and products were analyzed by gel electrophoresis. PCR amplicons were puriﬁed using AMPure XP beads (Beckman Coulter Inc., Brea, CA) and prepared for sequencing using the Nextera XT DNA Sample Preparation Kit and Index Kit (Illumina). 2 × 150 bp paired-end sequencing was performed using the Illumina NextSeq 500 (Illumina). Paired-end reads from the targeted amplicon deep sequencing were aligned to the mouse genome mm9 on BaseSpace (Illumina) using a banded Smith–Waterman algorithm and BedTools was used to calculate coverage. Mpileup ﬁles were created using SAMtools (1.3.1) and variant calling was performed using VarScan (2.4.1). Putative mutations were ﬁltered based on quality and resultant variants were manually reviewed using Bambino™. For calculations of mutant allele frequencies, the number of reads required for the mutant allele was ≥1000, otherwise data from the PCR ampliﬁcation validation was used. Fish plots depicting clonal evolution between primary and secondary samples were made for all targeted nucleotide variants by single nucleotide variants as determined by the targeted amplicon sequencing (Supplementary Data 2) and frequencies of cells displaying aberrant interphase FISH analysis signals (Supplementary Table 3).

Western blot analysis. Ba/F3 cells and Ba/F3 cells transduced with either FLT3WT-GFP, FLT3N676K-GFP, FLT3D835Y-GFP, FLT3ITD-GFP, and FLT3N676K-mCherry™ or mCherry™ (for KMT2A-MLLT3 + Empty-GFP) leukemic BM cells for 1 week in “C10” media, consisting of RPMI 1640 with 1-glutamine (Thermo Scientiﬁc), 10% FBS (Thermo Scientiﬁc) supplemented with 100 ng/ml mSCF, 10 ng/ml hIL6 and 10 ng/ml mIL3 (Peprotech™), after which mSCF and hIL6 were withdrawn and mIL3 was decreased to 1 ng/ml for continuous culturing.

DNA PCR validation of FLT3N676K subclones. PCR was performed on DNA extracted from BM for samples S016024 (minor clone: 0.77% FLT3N676K) and S016189 (minor clone: 0.38% FLT3N676K). S046281 (positive control: 62.7% FLT3N676K) and S047106 (negative control), using the AccuPrime™ Taq DNA Polymerase System (Life Technologies) according to the manufacturer’s instructions (94 °C, 2 min; and 35 cycles of: 94 °C, 0.5 min; 62 °C, 0.5 min; 68 °C, 1 min). Primers were designed using Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) and the following were used: Forward: 5′-GATGCCCCCGAGGA-3′ and 5′-IRES-5′: GAGAGGGCGGCAAAGATGCCC-3′. PCR products were analyzed by gel electrophoresis. PCR reactions were puriﬁed using ExoSA-P-IT (Affymetrix, Santa Clara, CA, USA) and sent to GATC Biotech (Constance, Germany) for Sanger sequencing using above primers.

Targeted amplicon deep sequencing and analysis. A custom targeted panel of 41 genes (Supplementary Table 1) was constructed based on two unbiased sequencing efforts of infant, childhood and adult KMT2A-R leukemias, and selected genes were either recurrently mutated in its respective study or part of a PI3K/RAS/kinase- or epigenetic pathway. An additional nine genes were manually added (Supplementary Table 1). A PCR ampliﬁcation library preparation was performed using PCR Low Input according to manufacturer’s instructions (v03). In short, 20 ng of mouse genomic DNA was hybridized with the custom oligo pool, bound oligos were extended and ligated after which the libraries were ampliﬁed. Library concentrations were determined by the Qubit Fluorometer (Life Technologies), libraries were pooled and 2 × 150 bp paired-end sequencing was performed using the Illumina NextSeq 500 (Illumina). Paired-end reads from the targeted amplicon deep sequencing were aligned to the mouse genome mm9 on BaseSpace (Illumina) using a banded Smith–Waterman algorithm and BedTools was used to calculate coverage. Mpileup ﬁles were created using SAMtools (1.3.1) and variant calling was performed using VarScan (2.4.1). Putative mutations were ﬁltered based on quality and resultant variants were manually reviewed using Bambino™. For calculations of mutant allele frequencies, the number of reads required for the mutant allele was ≥1000, otherwise data from the PCR ampliﬁcation validation was used. Fish plots depicting clonal evolution between primary and secondary recipients were made for all mutant allele frequencies for single nucleotide variants as determined by the targeted amplicon sequencing (Supplementary Data 2) and frequencies of cells displaying aberrant interphase FISH analysis signals (Supplementary Table 3).

Fluorescence in situ hybridization (FISH). FISH was performed according to standard methods on cultured BM cells from the primary (S016338) and secondary recipient (S046295) recipients with acquired KrasQ61D, utilizing BACs BMQ-415F20, BMQ-422H6 and BMQ-189B18 (Source BioScience, Nottingham, UK), overlapping the Kras locus, and a whole chromosome paint probe for chromosome 6 (MetaSystems Probes GmbH, Altlussheim, Germany). At least 100 interphase and 10 metaphase cells were scored for each sample.

Structural analyses. The three-dimensional structure of the N-terminal domain of human CBL (PDB ID: 2Y1M) was aligned with the complex structures of human CBL (ZAP-70 (PDB ID: 2Y1N), CBL/SKX (PDB ID: 3BUW) and CBL/EFGF (overexpressed 3BU)) applying the align command in the PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC. Only one CBL representation is clearly stated for reference.

Ba/F3 cell culture and cytokine independence assay. To analyze cytokine independence, Ba/F3 cells were transduced with the retroviral constructs CblA308T-mCherry™, eGFP + FLT3N676K-mCherry™, CblWT-mCherry™ + FLT3WT-mCherry™ or FLT3WT-mCherry™. mCherry, viable cells were sorted based on expression of GFP+mCherry™ (CblA308T-mCherry™, CblWT-mCherry™ + FLT3WT-mCherry™ or FLT3WT-mCherry™) and expanded in RPMI-1640 with 1-glutamine (Thermo Scientiﬁc), 10% FBS (Thermo Scientiﬁc), 100 units/ml Penicillin and 100
g/ml Streptomycin (Thermo Scientific) supplemented with 10 ng/ml Ml3 (PeproTech). For cytotoxic independence assay, cells were washed three times in PBS + 2% FBS and seeded in triplicates at a density of 0.5 x 10^6 cells in 1 ml in 12-well plates. Cells were counted continuously using CountBright beads (Life Technologies) on a FACS Fortessa (BD) for up to 19 days. Bf/F3 cells treated with 10 ng/ml Ml3 was included as a positive and Ba/F3 without Ml3 was used as a negative control. Cells were split when cell numbers reached 2 x 10^6 cells/ml or 0.2 million/ml.

**DNA/RNA extraction and RNA sequencing.** RNA and DNA were extracted using AllPrep DNA/RNA Mini Kit (Qiagen) according to the manufacturer’s instructions. The DNA quantity and quality was assessed by the NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and the Qubit Fluorometer (Life Technologies, Paisley, UK), respectively. The RNA quantity and quality was assessed by the NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific) and the Agilent 2100 Bioanalyzer (Agilent Technologies), respectively. Selection of poly(A) mRNA and subsequent cDNA synthesis was performed using the Illumina TrueSeq RNA sample preparation kit (Illumina, San Diego, CA, USA) according to the manufacturer’s instructions, but with a modified RNA fragmentation protocol and cDNA synthesis was performed at 94 °C from 8 min to 10 s. Quality analysis of the library was analyzed using the Agilent 2100 Bioanalyzer (Agilent Technologies) and 2 x 80 bp paired-end sequencing was performed on Illumina NextSeq 500 (Illumina).

**Analysis of RNA sequencing data.** Paired-end reads from RNA sequencing were aligned to the mouse genome mm9 using BWA (0.5.10) aligner and a bam file was then generated from which coverage was calculated using an in-house pipeline.

**Cuffdiff265 and log2 transformed.** A FPKM cutoff of per million mapped fragments (FPKM) and gene FPKM was calculated using the RNAseq data of expressed genes (6411 out of 6756). The RNAseq data have been deposited under the accession er PXD008213 to the Proteomics Database (PRIDE) for 3 days before transcription into sublethally irradiated (600 Gy) recipients. Blood sampling was performed 3 weeks post transplantation and at disease manifestation, BM cells from femurs and tibias were collected, and the percentage of dRed^2 leukemia cells were determined using flow cytometry.

**Proteome data analysis.** MS raw data were processed with Proteome Discoverer (version 2.1, Thermo Scientific). Enzyme set was to LysC and a maximum of two missed cleavages were allowed. TMT-K and TMT N-term were set as static modifications. Data were annotated using SEQUEST search engine against the Uniprot mouse database (downloaded 2016.05.29) containing 79,920 proteins to which 197 frequently observed contaminants had been added, as well as the trans and reporter genes used in this study. Maximal allowed precursor mass tolerance was set to 10 ppm. 6756 proteins were identified of which 5941 fulfilled an FDR < 0.05. Of identified proteins, 94.9% could be linked to their corresponding mRNA in the RNASeq data of expressed genes (6411 out of 6756).

**MIF ex vivo treatment assays.** For the cytokine assay, established mouse KMT2A-MLLT3 + Empty-GFP leukemia cells were washed once in cold PBS (GE Healthcare Life Sciences), resuspended in C10 medium supplemented with 25 ng/ml mSCF (PeproTech) and plated in 100 μl at a density of 50,000 cells per well in a 96-well U-bottom plate and 5 μl recombinant human MIF (PeproTech) was added at a final concentration of 10 or 500 ng/ml, and PBS (GE Healthcare Life Sciences) was added to an equal number of control cells. Cells were counted on day 6 after seeding using CountBright beads (Life Technologies, Eugene, OR, USA) on a FACS Fortessa (BD). Data analysis was performed using the FlowJo software (FlowJo, LLC).

**For the ex vivo transplantation assay, serially propagated dsRed^2 KMT2A-MLLT3 + Empty-GFP leukemia cells were used.** Briefly, 5000 freshly isolated CD117^+ quiescent transplant leukemia cells were treated ex vivo with or without 500 ng/ml MIF (PeproTech) in StemSpan™ SFEM (Stemcell Technologies) for 3 days before transplantation into sublethally irradiated (600 Gy) recipients. Blood sampling was performed 3 weeks post transplantation and at disease manifestation, BM cells from femurs and tibias were collected, and the percentage of dRed^2 leukemia cells were determined using flow cytometry.

**Statistical analysis.** Differences between groups were assessed by unpaired or paired (when noted) Students t-test. Correlation was assessed using Spearman’s rank correlation coefficient. Statistical analysis of survival curves was performed using Mantel-Cox log-rank test. All graphs show mean with all individual data points. All analyses were performed with Prism software version 6.0 (GraphPad software).

**Code Availability.** GSE106714 and PXD008213. Registered accession codes: SRP100345. Protein Data Bank: 211M, 2Y1N, 3BUW, 3BUQ, and 4ABB.

**Data availability.** RNA sequencing data have been deposited under the accession number GSE106714 in the Gene Expression Omnibus (GEO). Mass spectrometry proteome data have been deposited with the identifier PXD008213 to the ProteomeXchange Consortium via the Proteomics Identifications Database (PRIDE) partner repository. All other relevant data are available from the corresponding author upon request.
8. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
7. Lavallée, V.-P. et al. The transcriptomic landscape and directed chemical
9. Oshima, K. et al. Mutational landscape, clonal evolution patterns, and role of
10. Stubbs, M. C. et al. MLL-AF9 and FLT3 cooperation in acute myelogenous
6. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
11. Zuber, J. et al. Mouse models of human AML accurately predict chemotherapy
12. Moriya, K. et al. Development of a multi-step leukemogenesis model of MLL-
1. Muntean, A. G. & Hess, J. L. The pathogenesis of mixed-lineage leukemia.
20. Kogan, S. C. et al. Bethesda proposals for classi
23. LaRochelle, J. R. et al. Structural and functional consequences of three cancer-
24. Bentires-Alj, M. et al. Activating mutations of the noonan syndrome-
25. Forbes, S. A. et al. COSMIC: exploring the world
27. Sargin, B. et al. Flt3-dependent transformation by inactivating c-Cbl
919 (2011).
602 (2004).
8820 (2004).
28. Polzer, H., Janke, H., Schmid, D., Hiddemann, W. & Speikermann, K. Casitas B-lineage lymphoma mutants activate AKT to induce transformation in cooperation with class III receptor tyrosine kinases. Exp. Hematol. 41, 271–280 (2013).
29. Fernandes, M. S. et al. Novel oncogenic mutations of CBL in human acute
30. Fehn, C. et al. CBL exon 8/9 mutants activate the FLT3 pathway and cluster in core binding factor/11q deletion acute myeloid leukemia/myelodysplastic syndrome subtypes. Clin. Cancer Res. 15, 2238–2247 (2009).
31. Mohi, M. G. & Neel, B. G. The role of Shp2 (PTPN11) in cancer. Curr. Opin.
23, 23–30 (2007).
32. Chen, W. et al. Malignant transformation initiated by MLL-AF9: gene dosage
33. Sachs, Z. et al. NRASG12V oncogene facilitates self-renewal in a murine
34. Somervaille, T. C. P. et al. Hierarchical maintenance of MLL myeloid leukemia
35. Zuber, J. et al. An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-renewal program as essential for leukemia maintenance. Genes Dev. 25, 1628–1640 (2011).
36. Dang, J. et al. AMKL chimeric transcription factors are potent inducers of leukemia. Leukemia 31, 2228–2234 (2017).
37. Subramaniam, A. et al. Gene set enrichment analysis: a knowledge-based
38. Yagi, T. et al. Identification of a gene expression signature associated with
39. Zuber, J. et al. Mouse models of human AML accurately predict chemotherapy
40. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
41. Oshima, K. et al. Mutational landscape, clonal evolution patterns, and role of
42. Stubbs, M. C. et al. MLL-AF9 and FLT3 cooperation in acute myelogenous leukemia; development of a model for rapid therapeutic assessment. Leukemia 22, 66–70 (2008).
43. Sanger, B. et al. Blf3-dependent transformation by inactivating c-Cbl mutations in AML. Blood 110, 1004–1012 (2007).
44. Pui, C.-H. et al. Treating childhood acute lymphoblastic leukemia without cranial irradiation. N. Engl. J. Med. 360, 2730–2741 (2009).
45. Andersson, A. K. et al. The landscape of somatic mutations in infant MLL-
46. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
47. Oshima, K. et al. Mutational landscape, clonal evolution patterns, and role of
48. Stubbs, M. C. et al. MLL-AF9 and FLT3 cooperation in acute myelogenous leukemia: development of a model for rapid therapeutic assessment. Leukemia
49. Tamai, H. et al. Activated K-Ras protein accelerates human MLL/AF4-
50. Fujimoto, M. et al. The role of Shp2 (PTPN11) in cancer. Curr. Opin.
51. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
52. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
53. Ito, H. et al. Novel oncogenic mutations of CBL in human acute
54. Liljedahl, C. et al. CBL exon 8/9 mutants activate the FLT3 pathway and cluster in core binding factor/11q deletion acute myeloid leukemia/myelodysplastic syndrome subtypes. Clin. Cancer Res. 15, 2238–2247 (2009).
55. Mohi, M. G. & Neel, B. G. The role of Shp2 (PTPN11) in cancer. Curr. Opin.
56. Dang, J. et al. AMKL chimeric transcription factors are potent inducers of leukemia. Leukemia 31, 2228–2234 (2017).
57. Subramaniam, A. et al. Gene set enrichment analysis: a knowledge-based
58. Yagi, T. et al. Identification of a gene expression signature associated with
59. Zuber, J. et al. Mouse models of human AML accurately predict chemotherapy
60. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
61. Oshima, K. et al. Mutational landscape, clonal evolution patterns, and role of
62. Stubbs, M. C. et al. MLL-AF9 and FLT3 cooperation in acute myelogenous leukemia: development of a model for rapid therapeutic assessment. Leukemia
63. Tamai, H. et al. Activated K-Ras protein accelerates human MLL-AF4-
64. Sanger, B. et al. Blf3-dependent transformation by inactivating c-Cbl mutations in AML. Blood 110, 1004–1012 (2007).
65. Polzer, H., Janke, H., Schmid, D., Hiddemann, W. & Speikermann, K. Casitas B-lineage lymphoma mutants activate AKT to induce transformation in cooperation with class III receptor tyrosine kinases. Exp. Hematol. 41, 271–280 (2013).
66. Fernandes, M. S. et al. Novel oncogenic mutations of CBL in human acute
67. Lavallée, V.-P. et al. The transcriptomic landscape and directed chemical
68. Oshima, K. et al. Mutational landscape, clonal evolution patterns, and role of
69. Stubbs, M. C. et al. MLL-AF9 and FLT3 cooperation in acute myelogenous leukemia; development of a model for rapid therapeutic assessment. Leukemia
70. Tamai, H. et al. Activated K-Ras protein accelerates human MLL-AF4-
71. Zuber, J. et al. Mouse models of human AML accurately predict chemotherapy
72. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
73. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
74. Oshima, K. et al. Mutational landscape, clonal evolution patterns, and role of
75. Stubbs, M. C. et al. MLL-AF9 and FLT3 cooperation in acute myelogenous leukemia: development of a model for rapid therapeutic assessment. Leukemia
76. Tamai, H. et al. Activated K-Ras protein accelerates human MLL-AF4-
77. Zuber, J. et al. Mouse models of human AML accurately predict chemotherapy
78. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
79. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
80. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
81. Muntean, A. G. & Hess, J. L. The pathogenesis of mixed-lineage leukemia.
82. Pui, C.-H. et al. Treating childhood acute lymphoblastic leukemia without cranial irradiation. N. Engl. J. Med. 360, 2730–2741 (2009).
83. Andersson, A. K. et al. The landscape of somatic mutations in infant MLL-
84. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
85. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
86. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
87. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
88. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
89. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
90. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
91. Muntean, A. G. & Hess, J. L. The pathogenesis of mixed-lineage leukemia.
92. Pui, C.-H. et al. Treating childhood acute lymphoblastic leukemia without cranial irradiation. N. Engl. J. Med. 360, 2730–2741 (2009).
93. Andersson, A. K. et al. The landscape of somatic mutations in infant MLL-
94. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
95. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
96. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
97. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
98. Grossmann, V. et al. High incidence of RAS signalling pathway mutations in
99. Ma, X. et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-
Acknowledgements

We thank Andrea Biloglav for technical support and Kevin Shannon for critical reading of the manuscript. This work was supported by The Swedish Childhood Cancer Foundation, The Swedish Cancer Society, The Swedish Research Council, The Knut and Alice Wallenberg Foundation, BioCARE, The Gunnar Nilsson Cancer Foundation, Ellen Bachrachs Memorial Foundation, The Craaford Foundation, The Per-Eric and Ulla Schyberg Foundation, The Nilsson-Ehle Donations, The Wiberg Foundation, Governmental Funding of Clinical Research within the National Health Service. Support from the Swedish National Infrastructure for Biological Mass Spectrometry (BioMS) is gratefully acknowledged. A.H. is partly funded by The Georg Danielsson Foundation for Hematopoietic Disease.

Author contributions

A.K.H.A. designed the study; A.H.W. and A.K.H.A. designed experiments; A.H.W., H.S., J.H., J.U.K., J.L., R.R., C.G.R, and K.P. performed experiments; M.P., M.P.W., G.S., P.G., J.Z., and J.M performed computational data analyses; A.H.W and J.H. analyzed mass spectrometry data; S.N. and T.A.G. provided sequencing data of normal mouse progenitors; A.H.W., M.P., M.P.W., J.M., and A.K.H.A. analyzed sequencing data; K.L.P performed structural visualization; A.H. performed pathological evaluation; L.R., J.R.D., K.P., M.J., and T.A.G. performed critical reading of the manuscript; A.H.W. and A.K.H.A. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-04180-1.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.