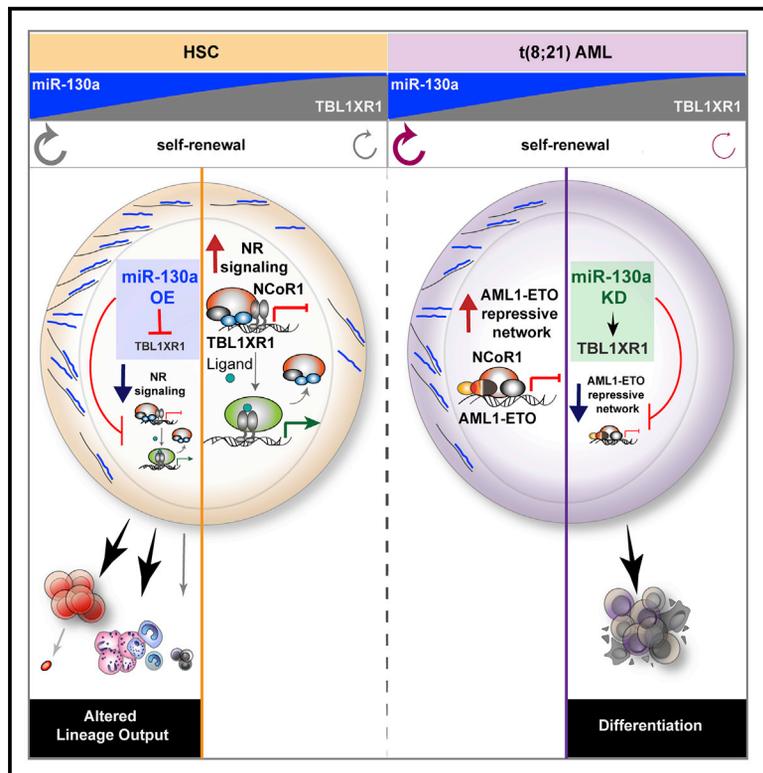


# Identification of the global miR-130a targetome reveals a role for TBL1XR1 in hematopoietic stem cell self-renewal and t(8;21) AML

## Graphical abstract



## Authors

Gabriela Krivdova, Veronique Voisin, Erwin M. Schoof, ..., Gene W. Yeo, John E. Dick, Eric R. Lechman

## Correspondence

john.dick@uhnresearch.ca (J.E.D.), eric.lechman@uhnresearch.ca (E.R.L.)

## In brief

Krivdova et al. identify TBL1XR1, a component of the NCoR1 repressor complex, as a direct target of miR-130a, which is involved in regulating HSC function and pathogenesis of t(8;21) AML. High levels of miR-130a are required for maintenance of this AML subtype by reinforcing the oncogenic molecular program of AML1-ETO.

## Highlights

- miR-130a is a regulator of HSC self-renewal and lineage commitment
- TBL1XR1 is a principal target of miR-130a
- TBL1XR1 loss of function in HSPC phenocopies enforced expression of miR-130a
- Elevated miR-130a levels maintain the AML1-ETO repressive program in t(8;21) AML



## Article

# Identification of the global miR-130a targetome reveals a role for TBL1XR1 in hematopoietic stem cell self-renewal and t(8;21) AML

Gabriela Krivdova,<sup>1,2</sup> Veronique Voisin,<sup>3</sup> Erwin M. Schoof,<sup>1</sup> Sajid A. Marhon,<sup>1</sup> Alex Murison,<sup>1</sup> Jessica L. McLeod,<sup>1</sup> Martino M. Gabra,<sup>4</sup> Andy G.X. Zeng,<sup>1,2</sup> Stefan Aigner,<sup>5</sup> Brian A. Yee,<sup>5</sup> Alexander A. Shishkin,<sup>5,10</sup> Eric L. Van Nostrand,<sup>5,11</sup> Karin G. Hermans,<sup>1,6</sup> Aaron C. Trotman-Grant,<sup>1,12</sup> Nathan Mbong,<sup>1</sup> James A. Kennedy,<sup>1,7</sup> Olga I. Gan,<sup>1</sup> Elvin Wagenblast,<sup>1</sup> Daniel D. De Carvalho,<sup>1,8</sup> Leonardo Salmena,<sup>1,4</sup> Mark D. Minden,<sup>1</sup> Gary D. Bader,<sup>1,2,3,9</sup> Gene W. Yeo,<sup>5</sup> John E. Dick,<sup>1,2,13,\*</sup> and Eric R. Lechman<sup>1,\*</sup>

<sup>1</sup>Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1L7, Canada

<sup>2</sup>Department of Molecular Genetics, University of Toronto, Toronto, ON M5S1A5, Canada

<sup>3</sup>The Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada

<sup>4</sup>Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON M5S 1A8, Canada

<sup>5</sup>Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92037, USA

<sup>6</sup>Program of Developmental & Stem Cell Biology, Peter Gilgan Centre for Research and Learning, The Hospital for Sick Children, Toronto, ON M5G0A4, Canada

<sup>7</sup>Division of Medical Oncology and Hematology, Sunnybrook Health Sciences Centre, Toronto, ON M4N3M5, Canada

<sup>8</sup>Department of Medical Biophysics, University of Toronto, Toronto, ON M5S 1A8, Canada

<sup>9</sup>Department of Computer Science, University of Toronto, Toronto, ON M5T 3A1, Canada

<sup>10</sup>Present address: Eclipse Bioinnovations, Inc., San Diego, CA 92121, USA

<sup>11</sup>Present address: Baylor College of Medicine, Houston, TX 77030, USA

<sup>12</sup>Present address: Immunology Graduate Program, Stanford University School of Medicine, Stanford, CA 94305, USA

<sup>13</sup>Lead contact

\*Correspondence: [john.dick@uhnresearch.ca](mailto:john.dick@uhnresearch.ca) (J.E.D.), [eric.lechman@uhnresearch.ca](mailto:eric.lechman@uhnresearch.ca) (E.R.L.)

<https://doi.org/10.1016/j.celrep.2022.110481>

## SUMMARY

Gene expression profiling and proteome analysis of normal and malignant hematopoietic stem cells (HSCs) point to shared core stemness properties. However, discordance between mRNA and protein signatures highlights an important role for post-transcriptional regulation by microRNAs (miRNAs) in governing this critical nexus. Here, we identify miR-130a as a regulator of HSC self-renewal and differentiation. Enforced expression of miR-130a impairs B lymphoid differentiation and expands long-term HSCs. Integration of protein mass spectrometry and chimeric AGO2 crosslinking and immunoprecipitation (CLIP) identifies TBL1XR1 as a primary miR-130a target, whose loss of function phenocopies miR-130a overexpression. Moreover, we report that miR-130a is highly expressed in t(8;21) acute myeloid leukemia (AML), where it is critical for maintaining the oncogenic molecular program mediated by the AML1-ETO complex. Our study establishes that identification of the comprehensive miRNA targetome within primary cells enables discovery of genes and molecular networks underpinning stemness properties of normal and leukemic cells.

## INTRODUCTION

The human blood system is a hierarchically organized continuum with the hematopoietic stem cell (HSC) residing at the apex (Bryder et al., 2006; Seita and Weissman, 2010; Laurenti and Göttgens, 2018). Acquisition of mutations and chromosomal abnormalities in HSCs and progenitor cells (HSPC) results in clonal dominance and transformation to acute myeloid leukemia (AML), with leukemic stem cells (LSCs) maintaining the malignant hierarchy and driving relapse (Kreso and Dick 2014; Shlush et al., 2014; Thomas and Majeti, 2017; Wang and Dick, 2005). Previous studies identified transcriptional and proteomic signatures un-

derpinning HSC and LSC states (Van Galen et al. 2019, Cabezas-Wallscheid et al., 2014; Velten et al., 2017; Ranzoni et al., 2021); however, little is known of the role that post-transcriptional regulation plays in governing the interrelationship between transcriptome and proteome profiles of different hematopoietic cell states in normal hematopoiesis and AML. MicroRNAs (miRNAs) represent a large class of non-coding RNAs that mediate repression of multiple target mRNAs at the post-transcriptional level (Bartel 2009, 2018). Although several miRNAs act as regulators of stemness and lineage commitment (Chung et al., 2011; Lechman et al., 2016), the mechanism of action and the comprehensive targetome they repress in human HSCs are



poorly defined. Moreover, how miRNA regulatory function is subverted in leukemogenesis and cooperates with specific genetic aberrations present in diverse AML subtypes remains largely unexplored.

In this study, we undertook a focused *in vivo* miRNA enforced expression screen in human HSPCs and identified miR-130a as a regulator of stemness and lineage commitment. Our mechanistic studies of the miR-130a targetome revealed repression of gene regulatory networks centered on transducin beta-like 1 X-linked receptor 1 (TBL1XR1). TBL1XR1 is a core component of the nuclear corepressor complexes NCoR1 and NCoR2, also known as NCoR and silencing mediator of retinoic acid and thyroid hormone (SMRT), respectively (Yoon et al., 2003). Both corepressors form multiprotein complexes consisting of HDAC3, G protein pathway suppressor 2 (GPS2), TBL1XR1, and its closely related homolog transducin beta-like protein 1X (TBL1X) (Guenther et al., 2001; Li et al., 2000; Wen et al., 2000; Yoon et al., 2003; Zhang et al., 2002). The NCoR1 and NCoR2 complexes function in repressing transcription governed by unliganded nuclear receptors (NRs) (Chen and Evans, 1995; Hörlein et al., 1995; Glass and Rosenfeld, 2000; Perissi et al., 2010; Watson et al., 2012). TBL1XR1 and TBL1X act as exchange factors to mediate the dismissal of the corepressors and subsequent transcriptional activation of the target genes (Perissi et al., 2004, 2008). Although deregulated function of the repressor complexes has been observed in many types of cancers (Mottis et al., 2013; Wong et al., 2014; Liang et al., 2019), their function in HSCs and LSCs is mostly unknown.

A common translocation in which the function of NCoR1 is dysregulated is t(8;21) AML, comprising up to 10% of all AML cases (Peterson and Zhang, 2004; Solh et al., 2014). The t(8;21) AML belongs to the core binding factor (CBF) leukemias, characterized by the cytogenic rearrangements that disrupt the function of RUNX1 (CBF or AML1) (Solh et al., 2014; de Bruijn and Speck, 2004). RUNX1 (AML1) is a heterodimeric transcription factor composed of one  $\alpha$  (RUNX1, RUNX2, or RUNX3) and one  $\beta$  (CBF $\beta$ ) subunit (Speck and Gilliland, 2002). RUNX1 functions as a transcriptional activator involved in different stages of adult hematopoiesis (de Bruijn and Speck, 2004; Guo et al., 2012; Ichikawa et al., 2004; Tober et al., 2013). The t(8;21) chromosomal translocation fuses the N-terminal DNA binding domain of the *RUNX1* gene on chromosome 21 to nearly the entire *ETO* (*RUNX1T1*) gene on chromosome 8 (Erickson et al., 1992, 1994; Miyoshi et al., 1991, 1993). The resulting AML1-ETO oncofusion protein recruits NCoR1 to repress transcription and therefore acts in a dominant negative fashion to inhibit the function of wild-type RUNX1 (Gelmetti et al., 1998; Lutterbach et al., 1998; Wang et al., 1998; Ptasinska et al., 2014). Although this AML subtype is generally associated with a favorable clinical outcome, only 40%–60% of adult patients achieve long-term survival (Appelbaum et al., 2006; Solh et al., 2014). Thus, a better understanding of the molecular mechanisms driving t(8;21) AML initiation, progression, and maintenance is critical for the development of targeted therapies and improved clinical outcomes.

The ability of miRNAs to repress multiple targets highlights the importance of uncovering their complete targetome to understand their mechanistic function in specific cell types. In this

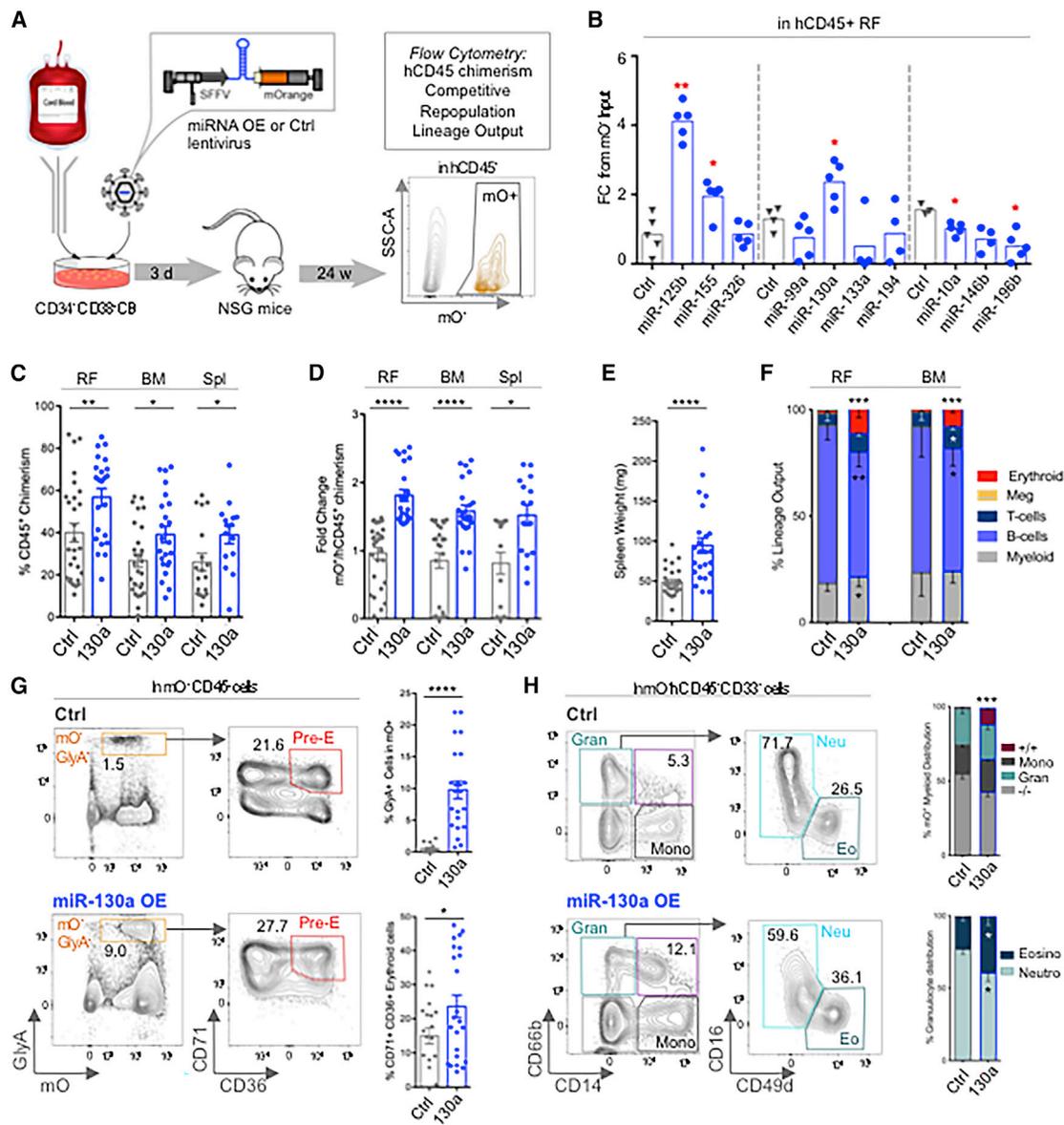
study, we integrate endogenous miRNA-target interactions from chimeric AGO2 enhanced crosslinking and immunoprecipitation (eCLIP) (Van Nostrand et al., 2016; Manakov et al., 2022) with proteomics and RNA-sequencing (RNA-seq) data to establish the mechanistic function of miR-130a in human HSPCs and t(8;21) AML. Our findings reveal that repression of TBL1XR1 by miR-130a in normal HSPCs impedes differentiation and expands the functional long-term (LT)-HSC population. Moreover, our study demonstrates that elevated levels of miR-130a in t(8;21) AML facilitate leukemogenesis by reinforcing the repressive AML1-ETO molecular program and preventing differentiation of leukemic cells.

## RESULTS

### *In vivo* microRNA screen identifies miR-130a as a regulator of HSC function

To investigate the functional role of candidate miRNAs previously found to be differentially expressed in normal and malignant human HSPCs (Lechman et al., 2016), a competitive repopulation screen was performed to determine the impact of overexpression (OE) of individual miRNAs on long-term repopulation potential. Lineage-negative (Lin<sup>−</sup>) CD34<sup>+</sup>CD38<sup>−</sup> cord blood (CB) cells were transduced with lentiviral constructs encoding selected miRNAs and an mOrange<sup>+</sup> (mO<sup>+</sup>) reporter gene and transplanted into immune-deficient NOD/Lt-scid/IL2R $\gamma$ <sup>null</sup> (NSG) mice (Figure 1A). Competitive repopulation was assessed by fold change in the percentage of transduced cells (mO<sup>+</sup>CD45<sup>+</sup>) in the injected right femur (RF) and non-injected bone marrow (BM) at 24 weeks following transplantation over mO<sup>+</sup> input levels. Enforced expression of miR-125b, miR-155, and miR-130a significantly increased long-term hematopoietic reconstitution, whereas miR-10a and miR-196b OE significantly reduced the repopulation capacity of HSPCs. Moreover, enforced expression of several miRNAs altered lineage distribution (Figures S1B and S1C). Based on the observed phenotype associated with miR-130a OE, we prioritized this miRNA for functional validation studies.

To investigate the impact of miR-130a OE on human HSC function, additional *in vivo* transplantation experiments were performed with Lin<sup>−</sup>CD34<sup>+</sup>CD38<sup>−</sup> CB cells. Increased expression of miR-130a in mO<sup>+</sup>CD45<sup>+</sup> cells from xenografts was confirmed by RT-qPCR (Figure S1D). Enforced expression of miR-130a resulted in greater human chimerism in the RF, BM, and spleen compared with controls at 24 weeks (Figure 1C). Notably, miR-130a OE conferred a competitive advantage and resulted in enlarged spleens in recipient mice (Figures 1D and 1E). Examination of lineage distribution revealed significant increases in erythroid, myeloid, and T cell outputs at the expense of B cell differentiation (Figures 1F and S1E–S1I). Enforced expression of miR-130a resulted in altered erythroid and myeloid differentiation programs, as evidenced by accumulation of immature CD71<sup>+</sup>CD36<sup>+</sup> erythroid progenitors, the presence of an aberrant CD33<sup>+</sup> cell population coexpressing granulocytic CD66b<sup>+</sup> and monocytic CD14<sup>+</sup> markers, a higher proportion of eosinophils, and reduced levels of neutrophils (Figures 1G and 1H). To further assess the functional significance of miR-130a in human HSPCs, we used a knockdown (KD) strategy utilizing an miRNA sponge

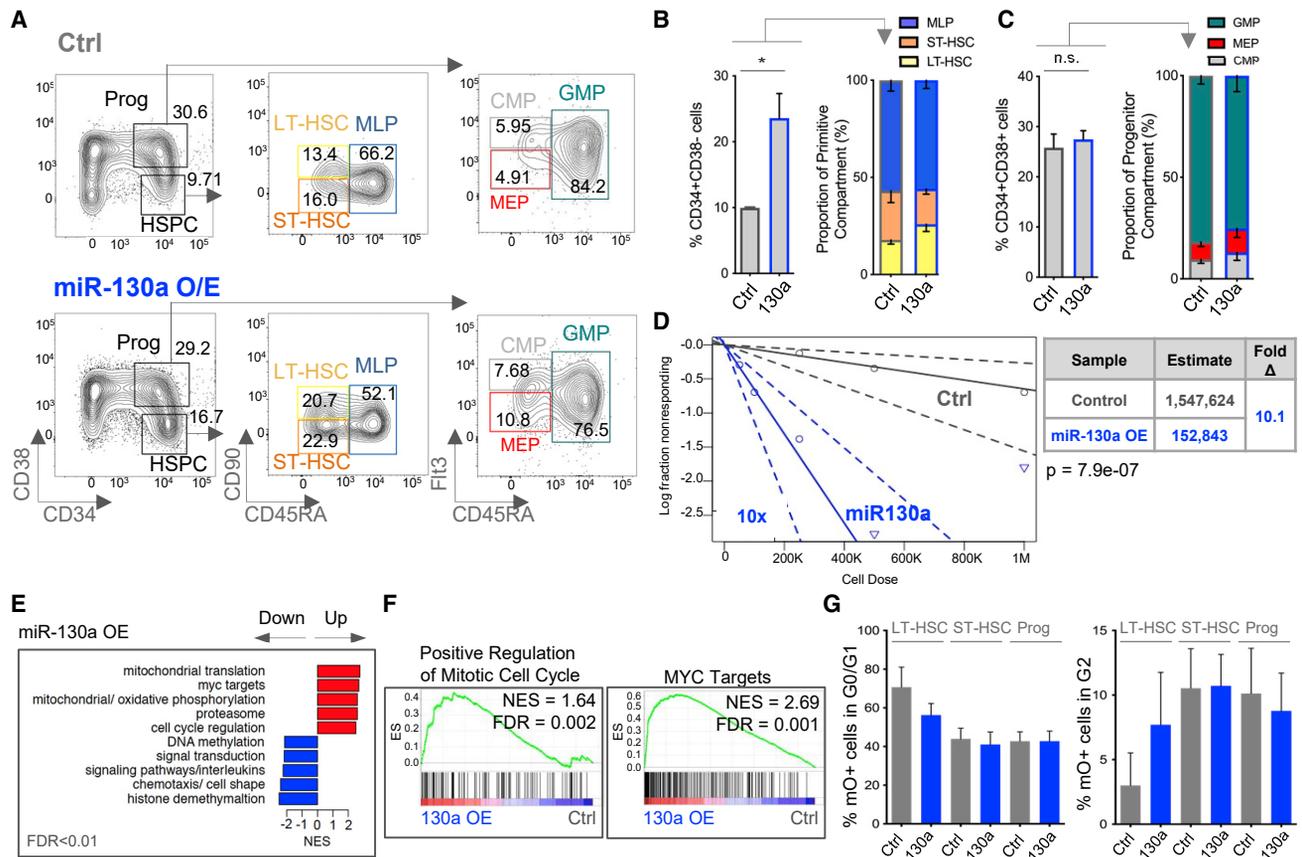


**Figure 1. Enforced expression of miR-130a enhances long-term repopulation capacity and alters lineage output of HSCs**

(A) Schematic of the *in vivo* miRNA overexpression (OE) repopulation assay.  
 (B) Fold change of  $mO^+CD45^+$  cells in the right femur (RF) at 24 weeks compared with input levels following enforced expression of individual miRNAs ( $n = 4-5$  mice/experimental group).  
 (C) Human  $CD45^+$  chimerism in the RF, bone marrow (BM), and spleen (Spl) ( $n = 3$  biological experiments, 7–10 mice/experimental group).  
 (D) Fold change of  $mO^+CD45^+$  cells at 24 weeks compared with input levels.  
 (E) Spleen weight of xenotransplanted mice.  
 (F) Lineage distribution of  $mO^+$  xenografts ( $n = 3$  independent biological experiments, 7–10 mice/experimental group).  
 (G) Representative flow cytometry plots of  $GlyA^+$  erythroid cells and  $CD36^+CD71^+$  progenitors (left) and graphs representing the proportions of these populations (right) ( $n = 13-22$  mice/experimental group).  
 (H) Representative flow cytometry plots of  $CD14^+$  monocytes,  $CD66b^+$  granulocytes,  $CD49d^+$  eosinophils, and  $CD16^+$  neutrophils (left) and graphs representing the proportions of these populations (right) ( $n = 8-9$  mice/experimental group).  
 (B–H) Mann-Whitney test, all error bars indicate  $\pm$ SEM, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ .

lentivector designed to specifically inhibit miR-130a function and GFP as a reporter gene. miR-130a KD caused substantial changes in lineage differentiation that were opposite the observed OE phenotype (Figures S1J–S1P). Collectively, our

data show that enforced expression of miR-130a increases the engraftment potential of human HSPCs, impairs B lymphoid lineage output, and alters erythroid and myeloid differentiation programs.



**Figure 2. Enforced expression of miR-130a expands HSCs by forcing them into cell cycle**

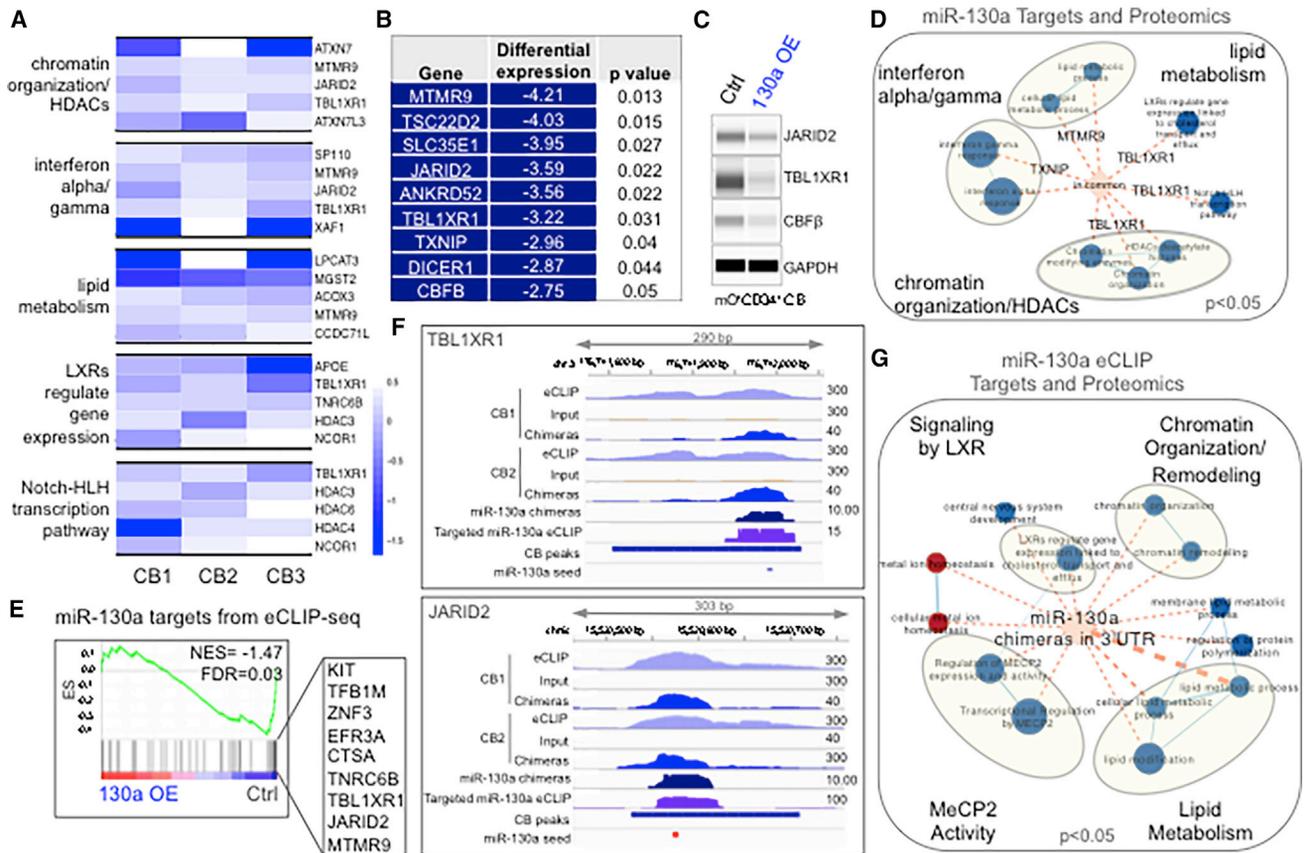
(A) Representative flow plots of HSPC populations from control and miR-130a OE xenografts at 12 weeks. (B) Proportion of CD34<sup>+</sup>CD38<sup>-</sup> compartment and frequency of LT-HSC, ST-HSC, and MLP (multi-lymphoid progenitor) cell populations (n = 4, each replicate contains pooled RF and BM from 2 to 4 individual mice). (C) Proportion of CD34<sup>+</sup>CD38<sup>+</sup> compartment and frequency of CMP (common myeloid progenitor), MEP (megakaryocyte-erythroid progenitor), and GMP (granulocyte-macrophage progenitor) cell populations (n = 4, each replicate contains pooled RF and BM from 2 to 4 individual mice). (D) Secondary transplantation of CD45<sup>+</sup>mO<sup>+</sup> from 12 week NSG mice at limiting dilution doses. Frequency of HSCs was evaluated from the engraftment in secondary mice (>0.05% CD45<sup>+</sup>mO<sup>+</sup> cells in BM, n = 19–31 mice from two independent experiments). (E) Normalized enrichment score (NES) of the gene sets significantly different between miR-130a OE and control-transduced HSCPs (n = 3). (F) GSEA plots of HSC cell-cycle-primed gene sets and MYC targets in the transcriptome profile following miR-130a OE in CD34<sup>+</sup> HSPCs. (G) Cell-cycle analysis of LT-HSCs, ST-HSCs, and CD34<sup>+</sup>CD38<sup>+</sup> cells transduced with miR-130a or control lentiviruses (n = 3). (B and C) Unpaired t test, all error bars indicate  $\pm$ SEM, \*p < 0.05.

### Enforced expression of miR-130a expands functional HSCs

We next sought to identify the HSPC populations in which miR-130a exerts its function. Combined BM and RF cells harvested from engrafted mice were depleted of murine and human lineage-committed cells and the remaining HSPCs were analyzed by flow cytometry. Enforced expression of miR-130a resulted in an approximately 2-fold increase in the proportion of CD34<sup>+</sup>CD38<sup>-</sup> cells and an increased proportion of immunophenotypic LT-HSCs (Figures 2A and 2B). The proportions of committed progenitors within the CD34<sup>+</sup>CD38<sup>+</sup> compartment were not altered (Figure 2C). The absolute number of CD34<sup>+</sup>CD38<sup>-</sup> and CD34<sup>+</sup>CD38<sup>+</sup> cells was increased 11- and 6-fold, respectively (Figure S2A). Secondary transplantation at limiting dilution demonstrated an approximately 10-fold higher frequency of LT-HSCs (1 in  $1.5 \times 10^5$  cells) in recipients trans-

planted with miR-130a OE HSPCs compared with controls (Figures 2D and S2B). In contrast, miR-130a KD caused notable reduction of the proportion of GFP<sup>+</sup> cells in the Lin<sup>-</sup> compartment and a decrease in the number of CD34<sup>+</sup>CD38<sup>-</sup> and CD34<sup>+</sup>CD38<sup>+</sup> cells (Figure S2C).

To identify the molecular pathways altered in miR-130a OE HSPCs, RNA-seq was performed on mO<sup>+</sup>CD34<sup>+</sup> HSPCs 3 days after transduction with control or miR-130a OE lentiviral constructs (Table S1). Gene set enrichment analysis (GSEA) revealed enrichment of mitochondrial translation, MYC signaling, proteasome function, and cell-cycle regulation pathways in genes upregulated in miR-130a OE cells (Figures 2E and S2D). In contrast, histone demethylation, interleukin signaling, and DNA methylation pathways were enriched in genes downregulated upon miR-130a OE (Figures 2E and S2D). As many of the upregulated gene sets were related to HSC programs, we performed GSEA using quiescent



**Figure 3. Mass spectrometry and chimeric AGO2 eCLIP identify miR-130a targetome in human HSPCs**

(A) Heatmap of downregulated pathways and proteins following miR-130a OE in CD34<sup>+</sup> CB cells; Wilcoxon one-sided test, n = 3, p < 0.05.  
 (B) List of miR-130a downregulated proteins that are predicted TarBase and mirDIP targets, limma t test.  
 (C) Western blot showing repression of miR-130a targets.  
 (D) Enrichment map showing gene sets containing miR-130a targets listed in (B); node size is proportional to NES; Wilcoxon one-sided test, p < 0.05.  
 (E) GSEA plots of miR-130a targets from chimeric AGO2 eCLIP in proteome profile following miR-130a OE in CD34<sup>+</sup> HSPCs.  
 (F) Genome browser tracks from chimeric AGO2 eCLIP-seq in CD34<sup>+</sup> HSPCs.  
 (G) Enrichment map of miR-130a targets from chimeric AGO2 eCLIP in downregulated protein sets following miR-130a OE; false discovery rate (FDR) < 0.05, node size is proportional to NES, Wilcoxon one-sided test; p < 0.05.

and activated HSC signatures (Forsberg et al., 2010) and observed enrichment of MYC targets and mitotic cell-cycle activation/proliferation gene sets and depletion of quiescent genes with miR-130a OE (Figures 2F and S2E). In addition, EdU (5-ethynyl-2'-deoxyuridine) incorporation assays were performed utilizing a fluorescently labeled thymidine analog to measure nascent DNA synthesis in sorted LT- and short-term (ST)-HSCs. Enforced expression of miR-130a in LT-HSCs, but not ST-HSCs, resulted in a decrease in the proportion of mO<sup>+</sup> cells in G0/G1 phase and an increase in G2 phase (Figures 2G and S2F). Thus, enforced expression of miR-130a induces HSC activation and cell-cycle progression while expanding the functional HSC pool.

### Repression of TBL1XR1 by miR-130a contributes to downregulation of chromatin remodeling and lipid metabolism pathways

To characterize potential miR-130a targets, low-cell-input, quantitative, label-free mass spectrometry (MS) was performed

in control or miR-130a OE CD34<sup>+</sup> CB cells (Table S2). GSEA revealed 145 upregulated and 128 downregulated pathways following miR-130a OE (Figure S3A). As miRNAs generally act as negative regulators of gene expression, our analysis focused on miR-130a targets within downregulated proteins. To identify which downregulated proteins are potential targets of miR-130a, experimentally validated TarBase targets were integrated into the subsequent analysis. TarBase miR-130a targets were strongly enriched among proteins downregulated upon miR-130a OE (Figures S3B and S3C). Enrichment mapping revealed that chromatin modification, interferon signaling, and lipid metabolism were significantly enriched in both downregulated proteins and miR-130a TarBase targets (Figures 3A and S3D). Examination of the overlap between predicted (mirDIP) and experimentally validated (TarBase) miR-130a targets identified nine genes also downregulated in our proteomics dataset, including TBL1XR1, CBFβ, and JARID2 (Figures 3B and S3E). Interestingly, JARID2 and CBFβ are components of the

well-characterized polycomb repressive complex 2 (PRC2) and RUNX1 transcription factor, respectively, which regulate chromatin organization and gene expression programs in adult hematopoiesis, and their function is frequently deregulated in leukemia (Cai et al., 2015; Ichikawa et al., 2004; Majewski et al., 2008; Radulović et al., 2013; Xie et al., 2014). Repression of TBL1XR1, CBF $\beta$ , and JARID2 by miR-130a OE was validated by western blot in human CD34<sup>+</sup> CB cells (Figures 3C and S3E). Enrichment mapping of the nine miR-130a targets with the downregulated gene sets revealed association of TBL1XR1 with multiple gene networks, including chromatin organization/modification, HDAC function, and lipid metabolism (Figure 3D).

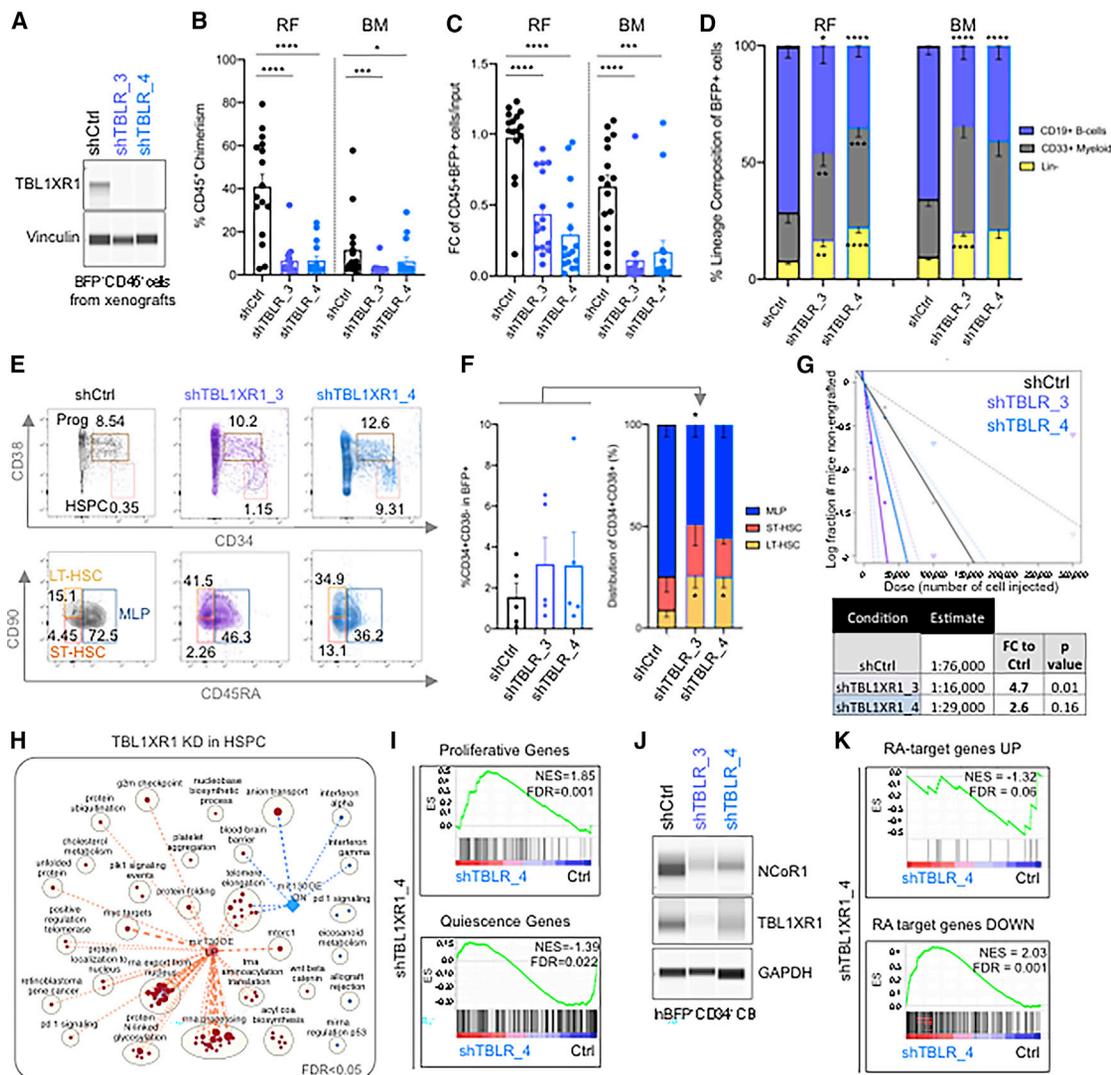
Currently, CLIP-seq is the only biochemical method that enables identification of global miRNA-target interactions (Chi et al., 2009; Darnell, 2010; Grosswendt et al., 2014; Helwak et al., 2013; Moore et al., 2015). To capture endogenous miR-130a targets in human HSPCs, we performed an enhanced (e) AGO2 CLIP-seq technique (Van Nostrand et al., 2016; Manakov et al., 2022) in CD34<sup>+</sup> HSPCs and analyzed chimeric reads containing miRNAs ligated to their targets (Table S3). In addition, due to the requirement of the eCLIP method for large input cell numbers, we used a targeted eCLIP-seq method that incorporates PCR amplification using miRNA-specific primers for more sensitive detection of miRNA targets (Manakov et al., 2022). Analysis of miR-130a targets revealed that 57% of miR-130a targets detected by eCLIP were present in TarBase (Figure S3G). As miRNA target sites located in the 3' UTR of mRNAs were previously shown to be associated with the greatest repression (Broughton et al., 2016; Grimson et al., 2007; Grosswendt et al., 2014; Helwak et al., 2013), we selected miR-130a targets chimeras containing 3' UTR target sites for further analysis. We focused on chimeras that were present in both targeted and transcriptome-wide eCLIP-seq experiments (n = 67) (Table S4). GSEA revealed significant enrichment of miR-130a-target chimeras in the list of proteins downregulated by miR-130a OE (Figure 3E). Notably, we identified *MTMR9*, *TBL1XR1*, and *JARID2* among the leading-edge genes enriched in the downregulated proteins (Figures 3E and S3H). Genome browser tracks containing miR-130a-target chimeras showed peaks corresponding to the binding of miR-130a to the 3' UTR of *TBL1XR1*, *JARID2*, and *CBF $\beta$* , providing evidence for the direct binding of miR-130a to these transcripts (Figures 3F and S3I). The intersection of miR-130a targets with the downregulated proteins showed enrichment of lipid metabolism and chromatin remodeling/organization and signaling by liver X receptors (LXRs), in line with our MS data (Figures 3G and S3J). Intriguingly, we noted that methyl CpG binding protein (MeCP) activity was also significantly downregulated following miR-130a OE. MeCP2 interacts directly with TBL1XR1, allowing its recruitment to heterochromatin and association with the NCoR1 and NCoR2 complexes (Kruusvee et al., 2017). Moreover, downregulation of LXR activity with concomitant decrease in the protein levels of NcoR1 and HDACs suggested that enforced expression of miR-130a impairs signaling via the NR/NCoR pathway (Figures 3A and S3K). Collectively, integration of MS and chimeric AGO2 eCLIP-seq identified TBL1XR1 as a direct miR-130a target that contributes to downregulation of genes involved in chromatin

organization/modification, interferon signaling, lipid metabolism, and NR signaling pathways.

### Repression of TBLXR1 by miR-130a impedes differentiation and expands LT-HSCs

Since TBL1XR1 is required for transcriptional activation by several NRs, including retinoic acid receptor (RAR) (Perissi et al., 2004), we interrogated the expression levels of RA-target genes following miR-130a OE. GSEA revealed depletion of genes activated by RA in miR-130a OE CD34<sup>+</sup> CB cells (Figure S4A). As the function of TBL1XR1 in hematopoiesis and HSCs is largely unknown, we investigated the impact of TBL1XR1 deficiency in human CD34<sup>+</sup>CD38<sup>-</sup> HSPCs by performing KD studies. We tested four TBL1XR1 short hairpin RNAs (shRNAs) for KD efficiency and selected two for *in vivo* studies (Figure S4B). Western blotting confirmed efficient KD of TBL1XR1 in transduced cells from xenografts (Figure 4A). TBL1XR1 KD significantly decreased human chimerism and the proportion of BFP<sup>+</sup> cells in the RF and BM of mice at 24 weeks post-transplantation (Figures 4B and 4C). Lineage analysis of TBL1XR1 KD xenografts revealed a significant decrease in mature CD19<sup>+</sup> B lymphoid cells and an increase in CD33<sup>+</sup> myeloid cells and CD45<sup>+</sup> human cells that did not express committed lineage markers (Lin<sup>-</sup>) (Figures 4D and S4C–S4G). Within the myeloid lineage, a significant decrease in mature granulocytes and an increase in monocytes were observed in TBL1XR1 KD xenografts (Figure S4H). Analysis of the HSPC compartment revealed an increase in CD34<sup>+</sup>CD38<sup>-</sup> cells and a significant expansion of immunophenotypic LT-HSC and ST-HSC populations with a concomitant decrease in the multi-lymphoid progenitor (MLP) population (Figures 4E and 4F). In addition, there was a significant expansion of CD7<sup>-</sup>CD10<sup>-</sup> progenitors within the CD34<sup>+</sup>CD38<sup>+</sup> population and an increase in the proportion of megakaryocyte-erythroid progenitor (MEP) in TBL1XR1 KD xenografts (Figure S4I). Serial transplantation at limiting dilution of BFP<sup>+</sup>CD45<sup>+</sup> cells from primary recipients into secondary NSG-GF recipients demonstrated a 4.7- and 2.6-fold increase in HSC frequency with TBL1XR1 shRNA3 and shRNA4, respectively (Figures 3G and S3J). In line with our findings from primary xenografts, TBL1XR1 KD significantly impaired B cell differentiation in secondary recipients, resulting in 80%–90% myeloid engraftment (Figure S4K).

To investigate the molecular mechanism of HSC expansion associated with TBL1XR1 KD, RNA-seq of BFP<sup>+</sup> progeny of transduced CD34<sup>+</sup>CD38<sup>-</sup> HSPCs cultured *in vitro* was performed (Table S1). Upon TBL1XR1 KD, there was a significant decrease in the expression of TBL1XR1 and a concomitant increase in its homolog TBL1X (Figure S4L). Despite the increased expression level of TBL1X, we observed significant gene expression changes upon TBL1XR1 KD (Table S1). Enrichment mapping revealed upregulation of genes in rRNA processing pathways and MYC targets in TBL1XR1 KD HSPCs, concordant with the observed enrichment of these gene sets following miR-130a OE (Figure 4H). GSEA showed enrichment of cell-cycle activation and proliferation genes and depletion of genes associated with cellular quiescence in TBL1XR1 KD cells (Figures 4I and S4M). As TBL1XR1 is a core component of the NcoR1 complex, we performed western blot analysis of

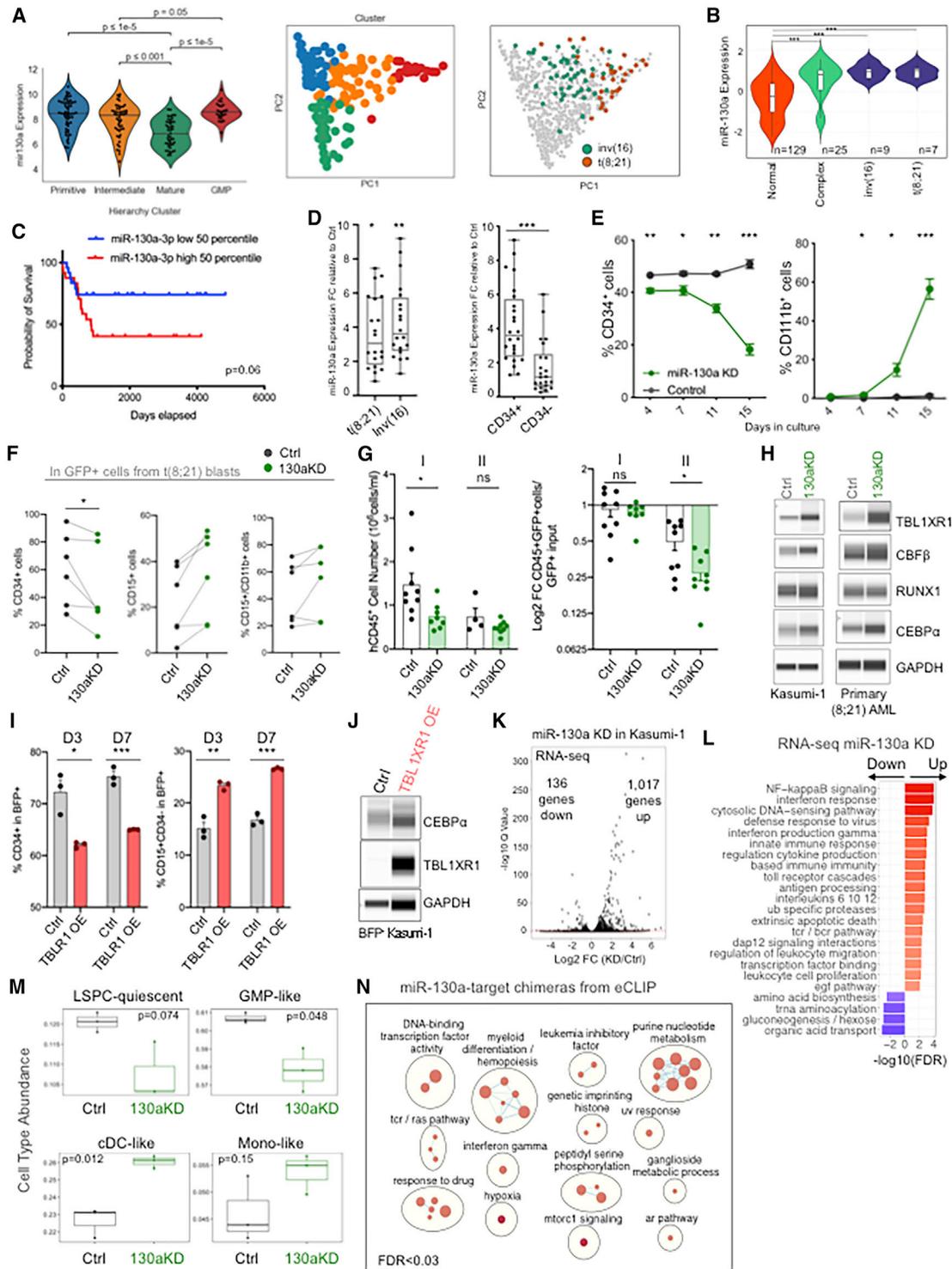


**Figure 4. Repression of TBL1XR1 impairs B lymphoid differentiation and expands LT-HSC**

(A) Western blot showing TBL1XR1 protein levels in xenografts.  
 (B) Human CD45<sup>+</sup> chimerism in the right femur (RF) and bone marrow (BM) at 24 weeks post-transplantation (n = 2 biological experiments, 6–8 mice/experimental group).  
 (C) Fold change in BFP<sup>+</sup>CD45<sup>+</sup> cells in RF and BM at 24 weeks post-transplantation compared with input levels.  
 (D) Lineage distribution of BFP<sup>+</sup> xenografts.  
 (E) Flow cytometry plots of HSPC populations, flow plots are representative of 3 samples overlaid together.  
 (F) Distribution of BFP<sup>+</sup>CD34<sup>+</sup>CD38<sup>-</sup> cell populations from 24 week xenografts (n = 5, each replicate contains pooled RF from 2 to 4 individual mice), unpaired t test, \*p < 0.05.  
 (G) Secondary transplantation of CD45<sup>+</sup>BFP<sup>+</sup> from 24 week NSG mice at limiting dilution doses. Frequency of HSC was evaluated from the engraftment in NSG-GF secondary mice (>0.05% CD45<sup>+</sup>BFP<sup>+</sup> cell in BM, n = 11–16 mice/experimental group).  
 (H) Enrichment map of upregulated (red nodes) and downregulated (blue nodes) pathways following TBL1XR1 KD intersected with miR-130a targets from eCLIP, hypergeometric t test, p < 0.05.  
 (I) GSEA plots of proliferative and quiescence genes in transcriptome profile following TBL1XR1 KD.  
 (J) Western blot showing protein levels of NCoR1 and TBL1XR1 following TBL1XR1 KD in CD34<sup>+</sup> CB cells.  
 (K) GSEA plots of genes upregulated and downregulated by RA in the transcriptome profile following TBL1XR1 KD.  
 (B–D) Unpaired Mann-Whitney U test, all error bars indicate  $\pm$ SEM, \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001, \*\*\*\*p < 0.0001.

TBL1XR1-deficient CD34<sup>+</sup> CB cells and confirmed decreased levels of NCoR1 (Figure 4J). Moreover, analysis of RA-target genes revealed strong enrichment of genes repressed by RA and depletion of genes upregulated by RA in TBL1XR1 KD cells,

indicating that a consequence of TBL1XR1 deficiency is abrogation of repression with concomitant activation of RA-target genes (Figures 4K and S4N). Overall, these findings demonstrate that repression of TBL1XR1 largely phenocopies the



**Figure 5. Loss of function of miR-130a and TBL1XR1 OE induces differentiation of t(8;21) AML**

(A) Clustering of TCGA AMLs into four groups based on scRNA-seq data (n = 173).

(B) Expression of miR-130a in AML subtypes from the TCGA dataset (n = 170).

(C) Kaplan-Meier curve showing correlation between miR-130a expression and CBF AML patient survival (n = 48).

(D) qRT-PCR showing expression levels of miR-130a in CD34<sup>+</sup> blasts from t(8;21) AML compared with healthy controls (n = 20) and CD34<sup>-</sup> cells from the same patient samples (n = 24).

(legend continued on next page)

functional effects and gene expression changes following the enforced expression of miR-130a.

### High miR-130a expression is required for maintenance of t(8;21) AML

As impaired myeloid differentiation and aberrant self-renewal programs are characteristic of leukemia, and several identified miR-130a targets are involved in oncogenic pathways in AML, we investigated whether expression of miR-130a is deregulated in some AML subtypes. Using the TCGA dataset, we identified four AML clusters based on miR-130a expression (Figure S5A). AML subtypes expressing high levels of miR-130a had higher LSC17 scores, a stemness signature associated with adverse risk in AML (Ng et al., 2016) (Figure S5B). Next, we mapped miR-130a expression onto TCGA, BEAT-AML, and Leucegene AMLs classified on the basis of leukemic hierarchy composition into granulocyte-monocyte progenitors (GMPs), intermediate, mature, and primitive groups using our new CIBERSORTx-based deconvolution approach involving single-cell (sc) RNA-seq signatures (Zeng et al., 2022). We observed high expression of miR-130a in primitive and GMP-like AMLs (Figure 5A). Similarly, gene set variation analysis (GSVA) using scRNA-seq data of normal hematopoietic populations revealed the association of miR-130a expression with cells exhibiting high HSC-progenitor and GMP-like signatures and low myeloid-like signatures (Figure S5C). Analysis of miR-130a expression across different AML subtypes from the TCGA dataset revealed high expression in cytogenetically complex AML (Figure 5B). In particular, we observed high expression of miR-130a in CBF AML, including t(8;21) and inv(16), which are highly enriched within the GMP-like cluster (Figures 5A, 5B, and S5D). Moreover, the majority of t(8;21) and inv(16) AML samples segregated into the clusters with high miR-130a expression (Figure S5E). To determine if miR-130a is of clinical importance, we examined miR-130a levels in CBF AML patient samples from the Princess Margaret Hospital Leukemia Tissue Bank (PM Bank) (Table S5) and observed that high miR-130a levels were associated with shorter overall survival and disease-free survival (Figures 5C and S5F). However, miR-130a expression was not prognostic when all cytogenetically normal or complex AML samples were considered in the analysis, suggesting that miR-130a plays a specific role in the pathogenesis of CBF AML (Figure S5G). RT-qPCR analysis of miR-130a expression levels in an independent cohort of CBF AML patients from the PM Bank (Table S5) revealed approximately 4-fold higher levels of miR-130a in CD34<sup>+</sup> blasts

compared with peripheral blood mononuclear cells from healthy controls or CD34<sup>+</sup> leukemic cells from the same patient samples (Figure 5D). In addition, profiling of miR-130a expression across different AML cell lines showed significantly higher expression in the Kasumi-1 cell line, which is derived from a t(8;21) AML and is characterized by an enrichment for the GMP-like hierarchy signature (Figures S5H and S5I). Collectively, these data demonstrate that miR-130a is highly expressed in CBF AML, including t(8;21) AML, and that higher miR-130a expression identifies a subset of CBF patients with inferior outcome.

To examine the functional significance of high miR-130a expression in t(8;21) AML, loss-of-function studies in Kasumi-1 cells and primary AML patient samples were performed. In Kasumi-1 cells, miR-130a KD induced a significant reduction in the CD34<sup>+</sup> cell population and an increase in differentiated CD11b<sup>+</sup>CD15<sup>+</sup> cells over a period of 2 weeks in culture (Figure 5E). Similarly, miR-130a KD in CD34<sup>+</sup> blasts from six t(8;21) AML patient samples followed by culture on the MS5 stromal cell line expressing human colony-stimulating factor 1 (hCSF1) resulted in a decreased proportion of CD34<sup>+</sup> cells and an increased percentage of differentiated CD11b<sup>+</sup>CD15<sup>+</sup> cells (Figures 5F and S5J). Transplantation of two t(8;21) patient samples transduced with miR-130a KD or control lentivectors into NSG-GF mice resulted in lower engraftment and a smaller proportion of GFP<sup>+</sup> cells compared with controls (Figure 5G). miR-130a KD in Kasumi-1 cells and primary t(8;21) AML cells increased levels of TBL1XR1 and CEBP $\alpha$  detected by Western blot, consistent with de-repression of these targets (Figure 5H). We also detected increased levels of the myeloid differentiation transcription factor CEBP $\alpha$ , in agreement with the differentiation immunophenotype observed with miR-130a KD. To examine whether de-repression of TBL1XR1 contributes to the differentiation phenotype, we overexpressed TBL1XR1 in Kasumi-1 cells (Figures 5I, 5J, and S5K). Notably, TBL1XR1 OE phenocopied the effects of miR-130a loss of function, including decreased proportion of CD34<sup>+</sup> cells, increased proportion of CD15<sup>+</sup> cells, and elevated levels of CEBP $\alpha$  (Figures 5I and 5J). Collectively, these results demonstrate that repression of TBL1XR1 by miR-130a is required for the maintenance of t(8;21) leukemia, and miR-130a loss of function or TBL1XR1 OE results in differentiation of leukemic cells.

To examine the effect of miR-130a on global gene expression in Kasumi-1 cells, we performed RNA-seq of Kasumi-1 cells following miR-130a KD (Table S1). We observed global activation of gene expression, including 1,017 transcripts upregulated and 136 transcripts downregulated following miR-130a KD

(E) Graphs representing quantitative changes in the immunophenotype of Kasumi-1 cells transduced with control and miR-130a KD lentiviruses (n = 3).

(F) Graphs representing quantitative changes in the immunophenotype of primary t(8;21) AML blasts (n = 6).

(G) Graphs representing changes in CD45<sup>+</sup> leukemic engraftment and log<sub>2</sub> fold change in GFP<sup>+</sup> cells from two primary t(8;21) samples in NSG-GF mice at 10 weeks. Unpaired Mann-Whitney U test.

(H) Western blots in t(8;21) AML following miR-130a KD.

(I) Graphs representing quantitative changes in the immunophenotype of Kasumi-1 cells transduced with control and TBL1XR1 OE lentiviruses (n = 3).

(J) Western blots showing changes in TBL1XR1 and CEBP $\alpha$  levels following TBL1XR1 OE in Kasumi-1 cells.

(K) Volcano plot showing DE genes following miR-130a KD in Kasumi-1 cells (n = 3).

(L) Bar graph representing upregulated and downregulated gene sets following miR-130a KD in Kasumi-1 cells, FDR < 0.01.

(M) Deconvolution of gene expression changes in control and miR-130a KD Kasumi-1 cells based on AML signatures from scRNA-seq data (n = 3).

(N) Enrichment map of upregulated gene sets following miR-130a KD in Kasumi-1 overlaid with miR-130a targets from chimeric AGO2 eCLIP, Mann-Whitney, p < 0.05.

(A–K) Unpaired t test unless indicated otherwise, all error bars indicate  $\pm$ SEM, \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001.

(Figure 5K). Genes in the interferon/immune response, ubiquitin processing/degradation, and apoptosis/cell proliferation pathways were strongly enriched following miR-130a KD (Figure 5L), in line with our functional data. Moreover, transcripts downregulated following miR-130a OE and TBL1XR1 KD in CD34<sup>+</sup> CB cells were enriched in miR-130a KD Kasumi-1 cells, providing evidence for de-repression of these gene networks (Figure S5L). Deconvolution of Kasumi-1 cells using the CIBERSORTx approach revealed loss of LSPC-quiescent and GMP-like signatures and a gain in cDC-like and mono-like signatures upon miR-130a KD (Figure 5M), consistent with the observed functional requirement of miR-130a in the differentiation block and maintenance of t(8;21) AML. To elucidate which of the upregulated transcripts are direct miR-130a targets, we performed chimeric AGO2 eCLIP-seq in Kasumi-1 cells and identified 100 and 890 miR-130a-target chimeras from the global and the targeted approaches, respectively (Table S4). miR-130a chimeric reads containing the 3' UTR of TBL1XR1 and CBF $\beta$  were captured by the more sensitive, targeted eCLIP-seq approach. By focusing on transcripts with target sites located in the 3' UTR and shared chimeras between the targeted and the transcriptome-wide approaches, we narrowed the list of miR-130a targets to 40 genes, including JARID2, TXNIP, TNRC6B, and KMT2C. Examination of the overlap between the miR-130a-target chimeras and the enrichment map from the RNA-seq revealed an association of these miR-130a targets with upregulated pathways, confirming global de-repression of miR-130a targets following its loss of function (Figure S5M). Myeloid differentiation and interferon signaling pathways were enriched in genes upregulated in miR-130-target chimeras (Figure 5N). Collectively, our results indicate that miR-130a KD in t(8;21) AML causes de-repression of targets centered on the myeloid differentiation and interferon pathways.

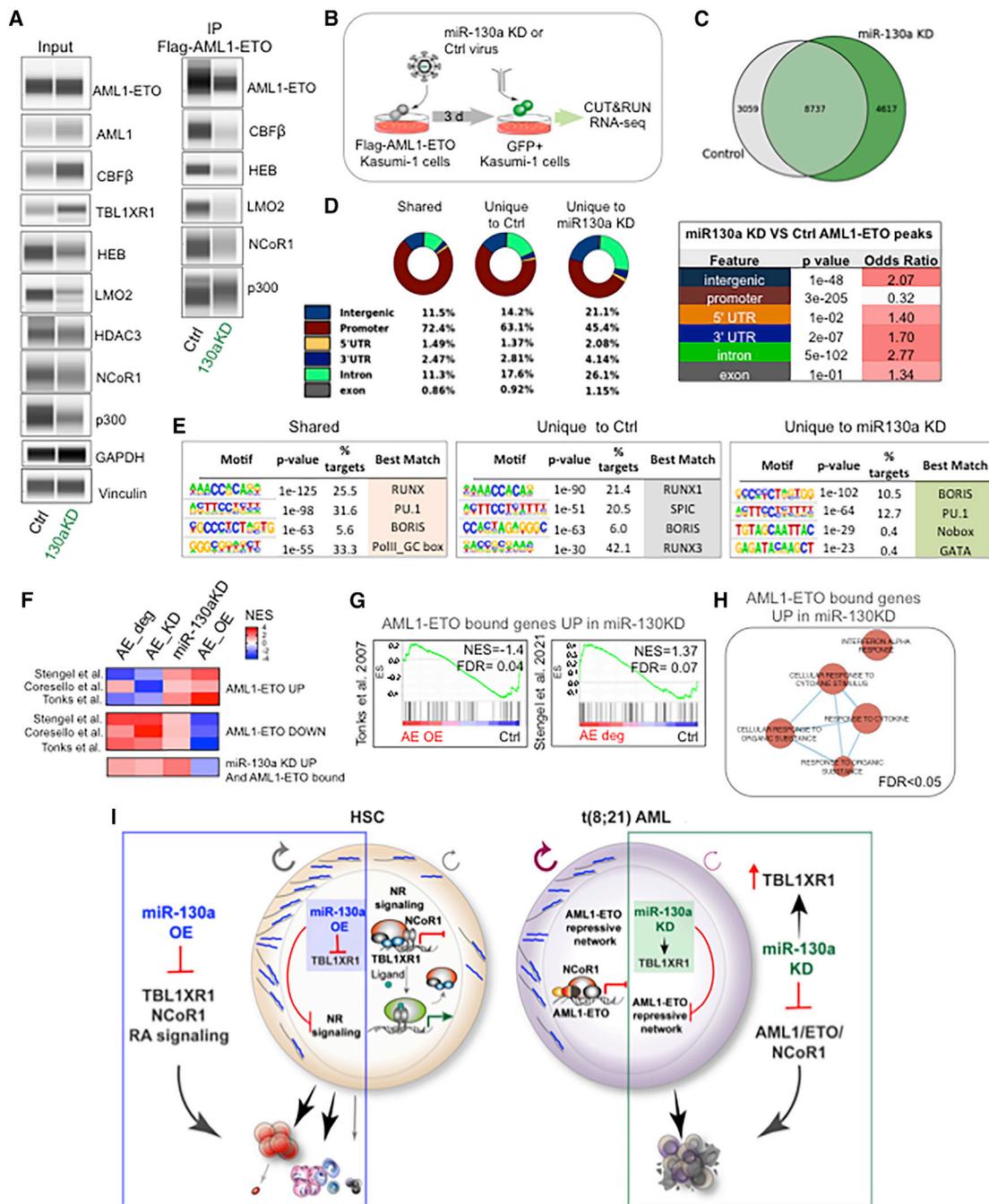
### miR-130a maintains the repressive AML1-ETO gene network

Apart from interactions with NCoR1 and HDAC3, AML1-ETO associates with several hematopoietic transcription (co)factors, including CBF $\beta$ , RUNX1, ERG, E proteins HEB and E2A, LMO2, and LYL1, which all contribute to leukemogenesis ((Mandoli et al., 2016); Sun et al., 2013). We performed immunoprecipitation (IP) of AML1-ETO in FLAG-tag knock-in AML1-ETO (FLAG-AE) Kasumi-1 cells to interrogate changes in the composition of the AML1-ETO complex following miR-130a KD. The specificity of the FLAG antibody and FLAG-AML1-ETO IP was confirmed with a wild-type (WT) Kasumi-1 cell line and a control IP sample with no addition of FLAG peptide for elution (Figures S6A and S6B). Western blot analysis revealed increased levels of CBF $\beta$  and TBL1XR1 and decreased levels of LMO1 and HEB in the input lysate in miR-130a KD Kasumi-1 cells. Surprisingly, we observed decreased association of AML1-ETO with CBF $\beta$  in the IP samples from miR-130a KD cells compared with control (Figure 6A). Moreover, western blot analysis revealed loss of association of AML1-ETO with NCoR1 and the transcription factors HEB and LMO2 (Figure 6A), suggesting that high miR-130a levels are required to maintain the composition of the AML1-ETO complex. KD of miR-130a did not affect the levels of AML1-ETO over the time course of 15 days, sug-

gesting that miR-130a KD does not induce degradation or affect the stability of AML1-ETO (Figure S6C).

To investigate whether changes in the composition of the AML1-ETO complex following miR-130a KD alter its binding occupancy, we performed CUT&RUN (cleavage under targets and release using nuclease) and RNA-seq (Figure 6B). CUT&RUN assays were performed with anti-FLAG and anti-ETO antibodies in FLAG-AML1-ETO Kasumi-1 cells, as the Kasumi-1 cells do not express ETO. In control cells, 12,260 of the 13,816 peaks (89%) detected with ETO antibody were also detected with anti-FLAG antibody, demonstrating a substantial overlap between the binding sites (Figure S6D). By contrast, overlap between the peaks detected by anti-FLAG and anti-ETO antibodies was less pronounced (62%) in miR-130a KD cells (Figure S6D). Importantly, the majority (~85%) of our annotated AML1-ETO binding sites overlapped with published CUT&RUN datasets of hemagglutinin (HA)-tagged AML1-ETO in Kasumi-1 cells (Stengel et al., 2021) (Figure S6E). Analysis of shared peaks captured with both anti-FLAG and anti-ETO antibodies, herein referred to as AML1-ETO peaks, revealed 4,617 unique peaks in miR-130a KD, 3,059 unique peaks in control, and 8,737 peaks shared between control and miR-130a KD cells (Figure 6C). Interestingly, AML1-ETO binding regions unique to miR-130a KD cells displayed significant loss of binding sites in promoter regions (45%) compared with peaks unique to control (63.1%) and shared peaks (72.4%) as well as a gain in intronic and intergenic regions at  $p < 0.001$  (Figure 6D). Motif enrichment analysis of peaks mapping to promoter regions of AML1-ETO-bound genes revealed the RUNX motif as the most highly enriched motif in both shared peaks and peaks unique to control (Figures 6E and S6F), supporting the specificity of the CUT&RUN approach. In contrast, RUNX motifs were not enriched in peaks unique to miR-130a KD cells (Figures 6E and S6F), indicating a shift in AML1-ETO binding with a resultant loss of specificity. K-means clustering based on the intensity of peaks unique to control identified two clusters, both with high enrichment of the RUNX motif (43% and 23%) (Figure S6G). By contrast, AML1-ETO peaks unique to miR-130a KD segregated to three clusters, none of which showed a strong enrichment of the RUNX motif (Figure S6G). Gene ontology (GO) analysis of promoter-bound AML1-ETO genes unique to control showed enrichment in cell-cycle and DNA-damage pathways (Figure S6H). In comparison, lipid metabolism, MAPK, and translation pathways were enriched in AML1-ETO promoter-bound regions unique to miR-130a KD cells. Overall, these findings show that loss of function of miR-130a in t(8;21) AML results in reorganization of the AML1-ETO repressor complex, including loss of association with CBF $\beta$  and NCoR1, with resultant altered binding of AML1-ETO across the genomic landscape.

Previous studies demonstrated that degradation of AML1-ETO in Kasumi cells results in de-repression of direct gene targets constituting a core AML1-ETO regulatory network, whereas AML1-ETO OE causes repression of these gene targets (Stengel et al., 2021; Tonks et al., 2007). To assess the impact of miR-130a KD on this known AML1-ETO gene network, we integrated RNA-seq analysis with AML1-ETO binding occupancy from the CUT&RUN analysis and interrogated several RNA-seq and chromatin immunoprecipitation (ChIP)-seq datasets, including



**Figure 6. miR-130a controls the composition and binding of the AML1-ETO complex**

- (A) Western blots of input lysates and IP of FLAG-AML1-ETO in Kasumi-1 cells.
- (B) Schematic depiction of the experimental design.
- (C) Venn diagram showing the number of shared and unique AML1-ETO peaks ( $n = 3$ ,  $q < 0.001$ ).
- (D) Genomic distribution of AML1-ETO peaks (left), Fisher t test was used to compare the distribution of peaks between control and miR-130a KD (right).
- (E) HOMER transcription factor binding site motif enrichment analysis.
- (F) Heatmap of gene expression signatures from AML1-ETO degradation, KD, and OE datasets compared with the transcriptome profile following miR-130a KD in Kasumi-1 cells.
- (G) GSEA plots showing enrichment of genes bound and upregulated by AML1-ETO in miR-130a KD Kasumi-1 cells in gene expression changes following AML1-ETO OE (left) or degradation (right).
- (H) Enrichment map showing upregulated, promoter-bound AML1-ETO genes unique to miR-130a KD (genes FDR < 0.05, pathways FDR < 0.05).
- (I) Mechanistic model of miR-130a function in HSC and t(8;21) AML.

AML1-ETO OE in HSPCs and AML1-ETO KD and degradation in Kasumi-1 cells (Corseello et al., 2009; Stengel et al., 2021; Tonks et al., 2007). AML1-ETO target genes identified from our CUT&RUN analysis as being promoter-bound and upregulated following miR-130a KD were enriched in the AML1-ETO degradation and KD gene sets (Figures 6F and 6G). The same genes were depleted in AML1-ETO OE samples, implying that these AML1-ETO targets are de-repressed following miR-130a KD (Figures 6F and 6G). This set of genes was strongly enriched in interferon alpha and cytokine response pathways (Figures 6H and S6I), and transcription factor enrichment analysis using iRegulon predicted STAT proteins as regulators of these pathways (Figures S6J–S6L). Collectively, our findings reveal a unique role for miR-130a in regulating HSC self-renewal and describe how elevated levels of miR-130a in t(8;21) AML contribute to the leukemogenesis of this AML subtype (Figure 6I).

## DISCUSSION

Our study establishes a fundamental mechanistic role for miR-130a in human HSCs. Our findings demonstrate that miR-130a is a regulator of HSC self-renewal and lineage commitment, with its enforced expression severely impairing B lymphoid differentiation and expanding LT-HSCs. TBL1XR1 was identified as a principal miR-130a target, whose repression results in downregulation of chromatin organization, lipid metabolism, and NR pathways controlled by NCoR1. TBL1XR1 loss of function largely phenocopied miR-130a enforced expression, establishing TBL1XR1 as a regulator of HSC self-renewal and lineage differentiation and highlighting the role of NR signaling pathways in human HSCs. Moreover, our study provides an insight into how elevated miR-130 levels maintain the AML1-ETO repressive transcriptional program in t(8;21) AML by governing the composition and binding occupancy of the AML1-ETO/NCoR1 repressor complex. Overall, these data provide evidence of the essential role that miRNA-mediated post-transcriptional regulation plays within both normal and malignant HSCs.

Our findings demonstrate that miR-130a is an important regulator of HSC self-renewal and differentiation. Although several studies have shown that miR-130a is highly expressed in HSCs and myeloid progenitors and that its expression declines during differentiation (Gentner et al., 2010; Georgantas et al., 2007; O'Connell et al., 2010; Petriv et al., 2010), the mechanistic function of miR-130a in human HSCs was previously unknown. Our analysis suggests that TBL1XR1 is associated with the preponderance of pathways that are downregulated in HSPCs following miR-130a OE, including chromatin organization, lipid metabolism, and NR signaling. Only a limited number of studies have addressed the contribution of NCoR1/SMRT complexes and diverse NR pathways in regulating HSC function. Conditional knockout of NCoR1 in murine LSK cells severely impaired B cell differentiation and increased frequency of LT-HSCs and ST-HSCs (Wan et al., 2019). Cell-cycle analysis of NCoR1-deficient LT-HSCs and murine HSPCs (LSK) revealed significant reduction in G0 cells with concomitant increases in G1 and S/G2/M cell-cycle phases, implicating NCoR1 in the regulation of HSC quiescence and proliferation (Wan et al., 2019). Our study corroborates the role of NCoR1 in controlling HSC properties and

identifies TBL1XR1, a core component of the NCoR1/SMRT repressor complexes, as a regulator of HSC self-renewal and cell-fate specification. Our data point to a convergent function of NCoR1 and TBL1XR1 in promoting lymphoid differentiation and HSC activation and further highlight the understudied role of NR pathways in regulating stemness properties of HSCs. Our results show that TBL1XR1 KD abrogates signaling by RA, demonstrated by de-repression of targets downregulated by RA and the inability to activate RA target genes in TBL1XR1-deficient cells. As RA signaling has been shown to regulate HSC maintenance/function (Cabezas-Wallscheid et al., 2017; Ghiaur et al., 2013) and TBL1XR1 is required for the transcriptional activation by NRs, we speculate that the inability to activate pathways downstream of NR including RA and impaired function of NCoR1 contributes to HSC activation and enhanced self-renewal of TBL1XR1-deficient HSC. Investigation of the precise role of TBL1XR1 and NCoR1-mediated repression and activation of signaling pathways governed by NRs in purified human HSC subsets will be challenging due to limited cell numbers, but clearly warranted.

Our data demonstrate that high miR-130a levels in t(8;21) AML are required for leukemia maintenance. Indeed, previous miRNA profiling studies reported elevated expression of miR-130a in CBF AML (Ding et al., 2018; Li et al., 2008). Moreover, miR-130a expression was shown to be elevated at relapse compared with diagnosis and associated with worse event-free survival of t(8;21) AML (Ding et al., 2018). Here, we show that miR-130a loss of function induces differentiation and apoptosis of t(8;21) AML cells accompanied by upregulation of genes enriched in myeloid differentiation and interferon signaling, suggesting that high miR-130a levels are required for t(8;21) AML maintenance. We show that TBL1XR1 OE in Kasumi-1 cells phenocopies these differentiation effects, implying that repression of TBL1XR1 by miR-130a is critical for t(8;21) AML maintenance. Interestingly, focal deletions and recurrent point mutations of TBL1XR1 occur in pediatric B-ALL and are enriched at relapse compared with diagnosis, implying that reduction in TBL1XR1 activity contributes to leukemogenesis (Mullighan et al., 2008; Parker et al., 2008; Zhang et al., 2011). Focal deletions of TBL1XR1 were reported in 10%–15% of B-ALL patients with the translocation t(12;21) characterized by the oncofusion protein ETV6-RUNX1, which also recruits NCoR1 (Parker et al., 2008). Decreased expression of TBL1XR1 in B-ALL was associated with resistance to glucocorticoid drugs (Jones et al., 2014). In addition, somatic mutations in TBL1XR1 also commonly occur in diffuse large B cell lymphomas and follicular lymphomas (~5%–10%), resulting in changes in the composition of the SMRT complex and impaired plasma cell differentiation (Venturutti et al., 2020). Moreover, recurrent fusions involving TBL1XR1 occur in acute promyelocytic leukemia and involve translocation between the *TBL1XR1*, *RARA*, and *RARB* genes, resulting in diminished transcriptional activity of RAR and inhibition of myeloid maturation (Chen et al., 2014; Osumi et al., 2018, 2019). Collectively, these studies and our findings demonstrate that TBL1XR1 plays a functional role in HSCs and imply that its loss of function contributes to diverse hematological malignancies.

Our mechanistic study of miR-130a function in t(8;21) provides insight into the complexity of interactions between the different

factors that constitute the AML1-ETO/NCoR1 repressive complex. Repression of CBF $\beta$  by miR-130a might also affect the function and composition of the AML1-ETO complex, as CBF $\beta$  is associated with both the RUNX1 (AML1) transcription factor and the AML1-ETO fusion protein. The mechanism driving upregulated expression of miR-130a in this AML subtype is unclear, but further understanding of how its expression is regulated could enable therapeutic manipulation of miR-130a expression levels to induce differentiation of t(8;21) AML. Moreover, the functional significance of elevated expression of miR-130a in inv(16) AML warrants further investigation as well. In summary, our comprehensive characterization of miR-130a function and identification of its targetome in normal HSPCs and t(8;21) AML have shed light on the transcriptional control of the pathways mediated by TBL1XR1 and NCoR1 in normal hematopoiesis and AML. Further investigation of the complex interactions between AML1-ETO, NCoR1, and TBL1XR1 will be beneficial for the design of future differentiation therapies targeting the repressor complex with the goal of developing better outcomes for patients with t(8;21) AML.

#### Limitations of the study

Although our findings show that the KD of TBL1XR1 phenocopies the effect of miR-130a OE, it does not definitively demonstrate that the phenotype is exclusively driven by repression of this single target. A lack of genetic rescue experiments in normal HSPCs and t(8;21) AML demonstrating the importance of TBL1XR1 in mediating the miR-130a phenotype is a minor shortcoming of this study. Furthermore, it is possible that miR-130a and TBL1XR1 have functional roles in other cytogenetically complex AML subtypes.

#### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Samples cryopreservation and thawing
  - Cord blood processing and HSPC enrichment
  - Cell culture
  - Mouse models
- **METHOD DETAILS**
  - Lentiviral vectors
  - Lentivirus production and titration
  - Lentiviral transduction
  - Cell sorting and flow cytometric analysis
  - Xenotransplantation
  - Lin<sup>-</sup> Cell depletions or CD34<sup>+</sup> enrichment
  - Flow cytometric analysis of xenografts
  - RNA isolation
  - RNA-seq processing and pathway enrichment analysis
  - Label free-mass spectrometry samples preparation

- Mass spectrometry acquisition
- Label-free quantitative proteomics analysis
- Pathway analysis and the enrichment map visualization
- Chimeric AGO2 eCLIP
- Chimeric AGO2 eCLIP data analysis
- Pathway analysis of miR-130a-target chimeras in CB and Kasumi-1 cells
- Cell cycle analysis
- Quantitative RT-PCR for expression levels of miR-130a in CBF AML
- miRNA microarray in CBF AML patient samples
- TCGA analysis of CBF AML patients
- GSVA and CIBERSORTx analysis and deconvolution of AML signatures from scRNAseq
- Western blot assay
- Immunoprecipitation of Flag-AML1-ETO
- CUT&RUN assay for Flag-AML1-ETO binding occupancy
  - CUT&RUN bioinformatic analysis
  - CUT&RUN and RNA-seq enrichment map
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

#### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2022.110481>

#### ACKNOWLEDGMENTS

We thank the obstetrics units at Trillium, William Osler, and Credit Valley hospitals for providing the cord blood samples; the Animal Resource Center (UHN) for support with the mouse work; SickKids UHN Flow and Princess Margaret Flow Facility for FACS; the Center for Applied Genomics and Princess Margaret Genomics Center for next-generation sequencing; and the laboratories of Steven Chan and Faiyaz Notta for sharing their equipment. We thank Scott W. Hiebert for supplying the FLAG-tagged AML1-ETO Kasumi-1 cell line, Steven Henikoff for Protein A-Micrococcal nuclease and spike-in yeast DNA, and M.D.M. for the human CSF1 MS5 stromal cell line. We thank Jean C.Y. Wang for detailed editing of the manuscript. We thank members of the Dick lab for assistance and critical feedback and Scott W. Hiebert, M.D.M., Courtney Jones, and Anastasia Tikhonova for comments on the manuscript. This work was supported by funds from the Princess Margaret Cancer Centre Foundation, Ontario Institute for Cancer Research through funding provided by the Government of Ontario, Canadian Institutes of Health Research (RN380110-409786), International Development Research Centre Ottawa Canada, Canadian Cancer Society (grants 703212 [end date 2019] and 706662 [end date 2025]), Terry Fox New Frontiers Program Project Grant (Project 1106), University of Toronto's Medicine by Design initiative with funding from the Canada First Research Excellence Fund, a Canada Research Chair, Princess Margaret Cancer Centre, The Princess Margaret Cancer Foundation, and Ontario Ministry of Health to J.E.D. G.K. is supported by a Canadian Institutes of Health Research (CIHR) Doctoral Award, Frederick Banting, and Charles Best Canada Graduate Scholarship (CGS-D). This work was supported by an NIH grant (HG009889) to G.W.Y.

#### AUTHOR CONTRIBUTIONS

G.K. conceived the study, performed *in vitro* and *in vivo* experiments, analyzed data, and wrote the manuscript. V.V. performed downstream analysis of RNA-seq, MS, and eCLIP-seq data, including enrichment mapping, GSEA, and analysis of TCGA AML. E.M.S. performed and analyzed mass spectrometry data. S.A.M. performed CUT&RUN analysis. A.M. performed RNA-seq

analysis. J.L.M., E.R.L., K.G.H., and A.G.T.G. assisted with mouse work. J.L.M. performed intrafemoral injections. M.G. performed analysis of the CBF-AML dataset. A.Z. performed CIBERSORTx analysis. S.A. and A.S. performed eCLIP-seq. B.Y. analyzed eCLIP-seq data. K.G.H., A.G.T.G., N.B., O.I.G., and E.W. assisted with *in vitro* experiments. V.V., O.I.G., E.W., E.R.L., and J.E.D. edited the manuscript. J.A.K. and M.D.M. coordinated AML patient consent and sample collection and provided patient information. E.R.L. conceived the project, analyzed data, and supervised the study. J.E.D. supervised research and secured funding for this study.

#### DECLARATION OF INTERESTS

G.W.Y. is a cofounder, a member of the board of directors, on the scientific advisory board, an equity holder, and a paid consultant for Locanabio and Eclipse Bioinnovations. G.W.Y. is a visiting professor at the National University of Singapore. G.W.Y.'s interests have been reviewed and approved by the University of California, San Diego, in accordance with its conflict-of-interest policies. E.L.V.N. is a co-founder, a member of the board of directors, on the SAB, an equity holder, and a paid consultant for Eclipse Bioinnovations. E.L.V.N.'s interests have been reviewed and approved by the University of California, San Diego, in accordance with its conflict of interest policies.

Received: August 9, 2021

Revised: December 3, 2021

Accepted: February 11, 2022

Published: March 8, 2022

#### REFERENCES

Amendola, M., Passerini, L., Pucci, F., Gentner, B., Bacchetta, R., and Naldini, L. (2009). Regulated and multiple miRNA and siRNA delivery into primary cells by a lentiviral platform. *Mol. Ther. J. Am. Soc. Gene Ther.* *17*, 1039–1052.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinforma. Oxf. Engl.* *31*, 166–169.

Appelbaum, F.R., Kopecky, K.J., Tallman, M.S., Slovak, M.L., Gundacker, H.M., Kim, H.T., Dewald, G.W., Kantarjian, H.M., Pierce, S.R., and Estey, E.H. (2006). The clinical spectrum of adult acute myeloid leukaemia associated with core binding factor translocations. *Br. J. Haematol.* *135*, 165–173.

Balmer, J.E., and Blomhoff, R. (2002). Gene expression regulation by retinoic acid. *J. Lipid Res.* *43*, 1773–1808.

Bartel, D.P. (2009). MicroRNA target recognition and regulatory functions. *Cell* *136*, 233–251.

Bartel, D.P. (2018). Metazoan MicroRNAs. *Cell* *173*, 20–51.

Broughton, J.P., Lovci, M.T., Huang, J.L., Yeo, G.W., and Pasquinelli, A.E. (2016). Pairing beyond the seed supports MicroRNA targeting specificity. *Mol. Cell* *64*, 320–333.

Bryder, D., Rossi, D.J., and Weissman, I.L. (2006). Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. *Am. J. Pathol.* *169*, 338–346.

Cabezas-Wallscheid, N., Klimmeck, D., Hansson, J., Lipka, D.B., Reyes, A., Wang, Q., Weichenhan, D., Lier, A., von Paleske, L., Renders, S., et al. (2014). Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell* *15*, 507–522.

Cabezas-Wallscheid, N., Buettner, F., Sommerkamp, P., Klimmeck, D., Ladel, L., Thalheimer, F.B., Pastor-Flores, D., Roma, L.P., Renders, S., Zeisberger, P., et al. (2017). Vitamin A-retinoic acid signaling regulates hematopoietic stem cell dormancy. *Cell* *169*, 807–823.e19.

Cai, X., Gao, L., Teng, L., Ge, J., Oo, Z.M., Kumar, A.R., Gilliland, D.G., Mason, P.J., Tan, K., and Speck, N.A. (2015). Runx1 deficiency decreases ribosome biogenesis and confers stress resistance to hematopoietic stem and progenitor cells. *Cell Stem Cell* *17*, 165–177.

Chen, J.D., and Evans, R.M. (1995). A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature* *377*, 454–457.

Chen, Y., Li, S., Zhou, C., Li, C., Ru, K., Rao, Q., Xing, H., Tian, Z., Tang, K., Mi, Y., et al. (2014). TBLR1 fuses to retinoid acid receptor  $\alpha$  in a variant t(3;17)(q26;q21) translocation of acute promyelocytic leukemia. *Blood* *124*, 936–945.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinforma. Oxf. Engl.* *34*, i884–i890.

Chi, S.W., Zang, J.B., Mele, A., and Darnell, R.B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* *460*, 479–486.

Chung, S.S., Hu, W., and Park, C.Y. (2011). The role of MicroRNAs in hematopoietic stem cell and leukemic stem cell function. *Ther. Adv. Hematol.* *2*, 317–334.

Corsello, S.M., Roti, G., Ross, K.N., Chow, K.T., Galinsky, I., DeAngelo, D.J., Stone, R.M., Kung, A.L., Golub, T.R., and Stegmaier, K. (2009). Identification of AML1-ETO modulators by chemical genomics. *Blood* *113*, 6193–6205.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* *26*, 1367–1372.

Darnell, R.B. (2010). HITS-CLIP: panoramic views of protein-RNA regulation in living cells. *Wiley Interdiscip. Rev. RNA* *1*, 266–286.

de Bruijn, M.F., and Speck, N.A. (2004). Core-binding factors in hematopoiesis and immune function. *Oncogene* *23*, 4238–4248.

Ding, C., Chen, S.-N., Macleod, R.A.F., Drexler, H.G., Nagel, S., Wu, D.-P., Sun, A.-N., and Dai, H.-P. (2018). MiR-130a is aberrantly overexpressed in adult acute myeloid leukemia with t(8;21) and its suppression induces AML cell death. *Ups. J. Med. Sci.* *123*, 19–27.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* *29*, 15–21.

Erickson, P., Gao, J., Chang, K.S., Look, T., Whisenant, E., Raimondi, S., Lasher, R., Trujillo, J., Rowley, J., and Drabkin, H. (1992). Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. *Blood* *80*, 1825–1831.

Erickson, P.F., Robinson, M., Owens, G., and Drabkin, H.A. (1994). The ETO portion of acute myeloid leukemia t(8;21) fusion transcript encodes a highly evolutionarily conserved, putative transcription factor. *Cancer Res.* *54*, 1782–1786.

Follenzi, A., Ailles, L.E., Bakovic, S., Geuna, M., and Naldini, L. (2000). Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat. Genet.* *25*, 217–222.

Forsberg, E.C., Passequé, E., Prohaska, S.S., Wagers, A.J., Koeva, M., Stuart, J.M., and Weissman, I.L. (2010). Molecular signatures of quiescent, mobilized and leukemia-initiating hematopoietic stem cells. *PLoS One* *5*, e8785.

Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* *47*, D766–D773.

Gelmetti, V., Zhang, J., Fanelli, M., Minucci, S., Pelicci, P.G., and Lazar, M.A. (1998). Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. *Mol. Cell. Biol.* *18*, 7185–7191.

Gentner, B., Schira, G., Giustacchini, A., Amendola, M., Brown, B.D., Ponzoni, M., and Naldini, L. (2009). Stable knockdown of microRNA *in vivo* by lentiviral vectors. *Nat. Methods* *6*, 63–66.

Gentner, B., Visigalli, I., Hiramatsu, H., Lechman, E., Ungari, S., Giustacchini, A., Schira, G., Amendola, M., Quattrini, A., Martino, S., et al. (2010). Identification of hematopoietic stem cell-specific miRNAs enables gene therapy of globoid cell leukodystrophy. *Sci. Transl. Med.* *2*, 58ra84.

Georgantans, R.W., Hildreth, R., Morisot, S., Alder, J., Liu, C., Heimfeld, S., Calin, G.A., Croce, C.M., and Civin, C.I. (2007). CD34+ hematopoietic stem-progenitor cell microRNA expression and function: a circuit diagram of differentiation control. *Proc. Natl. Acad. Sci. U S A* *104*, 2750–2755.

- Ghiaur, G., Yegnasubramanian, S., Perkins, B., Gucwa, J.L., Gerber, J.M., and Jones, R.J. (2013). Regulation of human hematopoietic stem cell self-renewal by the microenvironment's control of retinoic acid signaling. *Proc. Natl. Acad. Sci. U S A* *110*, 16121–16126.
- Glass, C.K., and Rosenfeld, M.G. (2000). The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev.* *14*, 121–141.
- Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A., and Enright, A.J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* *34*, D140–D144.
- Grimson, A., Farh, K.K.-H., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell* *27*, 91–105.
- Grosswendt, S., Filipchuk, A., Manzano, M., Klironomos, F., Schilling, M., Herzog, M., Gottwein, E., and Rajewsky, N. (2014). Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. *Mol. Cell* *54*, 1042–1054.
- Guenechea, G., Gan, O.I., Inamitsu, T., Dorrell, C., Pereira, D.S., Kelly, M., Naldini, L., and Dick, J.E. (2000). Transduction of human CD34+ CD38- bone marrow and cord blood-derived. *Mol. Ther. J. Am. Soc. Gene Ther.* *1*, 566–573.
- Guenther, M.G., Barak, O., and Lazar, M.A. (2001). The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. *Mol. Cell. Biol.* *21*, 6091–6101.
- Guo, H., Ma, O., Speck, N.A., and Friedman, A.D. (2012). Runx1 deletion or dominant inhibition reduces Cebpa transcription via conserved promoter and distal enhancer sites to favor monoopoiesis over granulopoiesis. *Blood* *119*, 4408–4418.
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* *38*, 576–589.
- Helwak, A., Kudla, G., Dudnakova, T., and Tollervey, D. (2013). Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell* *153*, 654–665.
- Hörlein, A.J., Näär, A.M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Söderström, M., and Glass, C.K. (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* *377*, 397–404.
- Hu, Y., and Smyth, G.K. (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* *347*, 70–78.
- Ichikawa, M., Asai, T., Saito, T., Seo, S., Yamazaki, I., Yamagata, T., Mitani, K., Chiba, S., Ogawa, S., Kurokawa, M., et al. (2004). AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. *Nat. Med.* *10*, 299–304.
- Jones, C.L., Bhatla, T., Blum, R., Wang, J., Paugh, S.W., Wen, X., Bourgeois, W., Bitterman, D.S., Raetz, E.A., Morrison, D.J., et al. (2014). Loss of TBL1XR1 disrupts glucocorticoid receptor recruitment to chromatin and results in glucocorticoid resistance in a B-lymphoblastic leukemia model. *J. Biol. Chem.* *289*, 20502–20515.
- Kaufmann, K.B., Garcia-Prat, L., Liu, Q., Ng, S.W.K., Takayanagi, S.I., Mitchell, A., Wienholds, E., van Galen, P., Cumba, C.A., Tsay, M.J., et al. (2019). A stemness screen reveals C3orf54/INKA1 as a promoter of human leukemia stem cell latency. *Blood* *133*, 2198–2211.
- Knott, S.R.V., Maceli, A., Erard, N., Chang, K., Marran, K., Zhou, X., Gordon, A., Demerdash, O.E., Wagenblast, E., Kim, S., et al. (2014). A computational algorithm to predict shRNA potency. *Mol. Cell* *56*, 796–807.
- Kreso, A., and Dick, J.E. (2014). Evolution of the cancer stem cell model. *Cell Stem Cell* *14*, 275–291.
- Kruusvee, V., Lyst, M.J., Taylor, C., Tarnauskaitė, Ž., Bird, A.P., and Cook, A.G. (2017). Structure of the MeCP2-TBLR1 complex reveals a molecular basis for Rett syndrome and related disorders. *Proc. Natl. Acad. Sci.* *114*, E3243.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* *9*, 357–359.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* *10*, R25.
- Laurenti, E., and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. *Nature* *553*, 418–426.
- Lechman, E.R., Gentner, B., van Galen, P., Giustacchini, A., Saini, M., Boccalatte, F.E., Hiramatsu, H., Restuccia, U., Bachi, A., Voisin, V., et al. (2012). Attenuation of miR-126 activity expands HSC in vivo without exhaustion. *Cell Stem Cell* *11*, 799–811.
- Lechman, E.R., Gentner, B., Ng, S.W., Schoof, E.M., van Galen, P., Kennedy, J.A., Nucera, S., Ciceri, F., Kaufmann, K.B., Takayama, N., et al. (2016). miR-126 regulates distinct self-renewal outcomes in normal and malignant hematopoietic stem cells. *Cancer Cell* *29*, 214–228.
- Li, J., Wang, J., Wang, J., Nawaz, Z., Liu, J.M., Qin, J., and Wong, J. (2000). Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. *EMBO J.* *19*, 4342–4350.
- Li, Z., Lu, J., Sun, M., Mi, S., Zhang, H., Luo, R.T., Chen, P., Wang, Y., Yan, M., Qian, Z., et al. (2008). Distinct microRNA expression profiles in acute myeloid leukemia with common translocations. *Proc. Natl. Acad. Sci. U S A* *105*, 15535–15540.
- Liang, N., Jakobsson, T., Fan, R., and Treuter, E. (2019). The nuclear receptor-Co-repressor complex in control of liver metabolism and disease. *Front. Endocrinol.* *10*, 411.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* *15*, 550.
- Lutterbach, B., Westendorf, J.J., Linggi, B., Patten, A., Moniwa, M., Davie, J.R., Huynh, K.D., Bardwell, V.J., Lavinsky, R.M., Rosenfeld, M.G., et al. (1998). ETO, a target of t(8;21) in acute leukemia, interacts with the N-CoR and mSin3 corepressors. *Mol. Cell. Biol.* *18*, 7176–7184.
- Majewski, I.J., Blewitt, M.E., de Graaf, C.A., McManus, E.J., Bahlo, M., Hilton, A.A., Hyland, C.D., Smyth, G.K., Corbin, J.E., Metcalf, D., et al. (2008). Polycomb repressive complex 2 (PRC2) restricts hematopoietic stem cell activity. *PLoS Biol.* *6*, e93.
- Manakov, A.S., Yee, B.A., Shen, K., Cox, D., Park, S., Shishkin, A.A., Yeo, G.W., and Van Nostrand, E.L. (2022). High depth profiling of miRNA targets with chimeric eCLIP. Preprint at bioRxiv. <https://doi.org/10.1101/2022.02.13.480296>.
- Mandoli, A., Singh, A.A., Prange, K.H.M., Tjichon, E., Oerlemans, M., Dirks, R., Ter Huurne, M., Wierenga, A.T.J., Janssen-Megens, E.M., Berentsen, K., et al. (2016). The Hematopoietic Transcription Factors RUNX1 and ERG Prevent AML1-ETO Oncogene Overexpression and Onset of the Apoptosis Program in t(8;21) AMLs. *Cell Rep* *17*, 2087–2100.
- Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albu, L.-P., Mushayamaha, T., and Thomas, P.D. (2021). PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* *49*, D394–D403.
- Miyoshi, H., Shimizu, K., Kozu, T., Maseki, N., Kaneko, Y., and Ohki, M. (1991). t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, Aml1. *Proc. Natl. Acad. Sci. U S A* *88*, 10431–10434.
- Miyoshi, H., Kozu, T., Shimizu, K., Enomoto, K., Maseki, N., Kaneko, Y., Kamada, N., and Ohki, M. (1993). The t(8;21) translocation in acute myeloid leukemia results in production of an. *EMBO J.* *12*, 2715–2721.
- Moore, M.J., Scheel, T.K.H., Luna, J.M., Park, C.Y., Fak, J.J., Nishiuchi, E., Rice, C.M., and Darnell, R.B. (2015). miRNA–target chimeras reveal miRNA 3'-end pairing as a major determinant of Argonaute target specificity. *Nat. Commun.* *6*, 8864.
- Mottis, A., Mouchiroud, L., and Auwerx, J. (2013). Emerging roles of the corepressors NCoR1 and SMRT in homeostasis. *Genes Dev.* *27*, 819–835.

- Mullighan, C.G., Phillips, L.A., Su, X., Ma, J., Miller, C.B., Shurtleff, S.A., and Downing, J.R. (2008). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**, 1377–1380.
- Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782.
- Ng, S.W.K., Mitchell, A., Kennedy, J.A., Chen, W.C., McLeod, J., Ibrahimova, N., Arruda, A., Popescu, A., Gupta, V., Schimmer, A.D., et al. (2016). A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature* **540**, 433–437.
- O’Connell, R.M., Chaudhuri, A.A., Rao, D.S., Gibson, W.S.J., Balazs, A.B., and Baltimore, D. (2010). MicroRNAs enriched in hematopoietic stem cells differentially regulate long-term hematopoietic output. *Proc. Natl. Acad. Sci. U S A* **107**, 14235–14240.
- Osumi, T., Tsujimoto, S., Tamura, M., Uchiyama, M., Nakabayashi, K., Okamura, K., Yoshida, M., Tomizawa, D., Watanabe, A., Takahashi, H., et al. (2018). Recurrent RARB translocations in acute promyelocytic leukemia lacking RARA translocation. *Cancer Res.* **78**, 4452.
- Osumi, T., Watanabe, A., Okamura, K., Nakabayashi, K., Yoshida, M., Tsujimoto, S.-I., Uchiyama, M., Takahashi, H., Tomizawa, D., Hata, K., et al. (2019). Acute promyelocytic leukemia with a cryptic insertion of RARA into TBL1XR1. *Genes. Chromosomes Cancer* **58**, 820–823.
- Parker, H., An, Q., Barber, K., Case, M., Davies, T., Konn, Z., Stewart, A., Wright, S., Griffiths, M., Ross, F.M., et al. (2008). The complex genomic profile of ETV6-RUNX1 positive acute lymphoblastic leukemia highlights a recurrent deletion of TBL1XR1. *Genes. Chromosomes Cancer* **47**, 1118–1125.
- Perissi, V., Aggarwal, A., Glass, C.K., Rose, D.W., and Rosenfeld, M.G. (2004). A corepressor/coactivator exchange complex required for transcriptional activation by nuclear receptors and other regulated transcription factors. *Cell* **116**, 511–526.
- Perissi, V., Scafoglio, C., Zhang, J., Ohgi, K.A., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (2008). TBL1 and TBLR1 phosphorylation on regulated gene promoters overcomes dual CtBP and NCoR/SMRT transcriptional repression checkpoints. *Mol. Cell* **29**, 755–766.
- Perissi, V., Jepsen, K., Glass, C.K., and Rosenfeld, M.G. (2010). Deconstructing repression: evolving models of co-repressor action. *Nat. Rev. Genet.* **11**, 109–123.
- Peterson, L.F., and Zhang, D.-E. (2004). The 8;21 translocation in leukemogenesis. *Oncogene* **23**, 4255–4262.
- Petriv, O.I., Kuchenbauer, F., Delaney, A.D., Lecault, V., White, A., Kent, D., Marmolejo, L., Heuser, M., Berg, T., Copley, M., et al. (2010). Comprehensive microRNA expression profiling of the hematopoietic hierarchy. *Proc. Natl. Acad. Sci. U S A* **107**, 15443–15448.
- Ptasinska, A., Assi, S.A., Martinez-Soria, N., Imperato, M.R., Piper, J., Cauchy, P., Pickin, A., James, S.R., Hoogenkamp, M., Williamson, D., et al. (2014). Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal. *Cell Rep.* **8**, 1974–1988.
- Radulović, V., de Haan, G., and Klauke, K. (2013). Polycomb-group proteins in hematopoietic stem cell regulation and hematopoietic neoplasms. *Leukemia* **27**, 523–533.
- Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165.
- Ranzoni, A.M., Tangherloni, A., Berest, I., Riva, S.G., Myers, B., Strzelecka, P.M., Xu, J., Panada, E., Mohorianu, I., Zaugg, J.B., et al. (2021). Integrative single-cell RNA-seq and ATAC-seq analysis of human developmental hematopoiesis. *Cell Stem Cell* **28**, 472–487.e7.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.* **26**, 139–140.
- Schoof, E.M., Lechman, E.R., and Dick, J.E. (2016). Global proteomics dataset of miR-126 overexpression in acute myeloid leukemia. *Data Brief* **9**, 57–61.
- Seita, J., and Weissman, I.L. (2010). Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2**, 640–653.
- Shlush, L.I., Zandi, S., Mitchell, A., Chen, W.C., Brandwein, J.M., Gupta, V., Kennedy, J.A., Schimmer, A.D., Schuh, A.C., Yee, K.W., et al. (2014). Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* **506**, 328–333.
- Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **6**, e21856.
- Skene, P.J., Henikoff, J.G., and Henikoff, S. (2018). Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019.
- Solh, M., Yohe, S., Weisdorf, D., and Ustun, C. (2014). Core-binding factor acute myeloid leukemia: heterogeneity, monitoring, and therapy. *Am. J. Hematol.* **89**, 1121–1131.
- Speck, N.A., and Gilliland, D.G. (2002). Core-binding factors in haematopoiesis and leukaemia. *Nat. Rev. Cancer* **2**, 502–513.
- Stengel, K.R., Ellis, J.D., Spielman, C.L., Bomber, M.L., and Hiebert, S.W. (2021). Definition of a small core transcriptional circuit regulated by AML1-ETO. *Mol. Cell* **81**, 530–545.e5.
- Sun, X.-J., Wang, Z., Wang, L., Jiang, Y., Kost, N., Soong, T.D., Chen, W.-Y., Tang, Z., Nakadai, T., Elemento, O., et al. (2013). A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis. *Nature* **500**, 93–97.
- Thomas, D., and Majeti, R. (2017). Biology and relevance of human acute myeloid leukemia stem cells. *Blood* **129**, 1577.
- Thomas, P.D., Campbell, M.J., Kejarival, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., and Narechania, A. (2003). PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* **13**, 2129–2141.
- Tober, J., Yzaguirre, A.D., Piwarzyk, E., and Speck, N.A. (2013). Distinct temporal requirements for Runx1 in hematopoietic progenitors and stem cells. *Dev. Camb. Engl.* **140**, 3765–3776.
- Tonks, A., Pearn, L., Musson, M., Gilkes, A., Mills, K.I., Burnett, A.K., and Darley, R.L. (2007). Transcriptional dysregulation mediated by RUNX1-RUNX1T1 in normal human progenitor cells and in acute myeloid leukaemia. *Leukemia* **21**, 2495–2505.
- Van Galen, P., Hovestadt, V., Wadsworth li, M.H., Hughes, T.K., Griffin, G.K., Battaglia, S., Verga, J.A., Stephansky, J., Pastika, T.J., Lombardi Story, J., et al. (2019). Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity. *Cell* **176**, 1265–1281.e24.
- Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514.
- Velten, L., Haas, S.F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B.P., Hirche, C., Lutz, C., Buss, E.C., Nowak, D., et al. (2017). Human hematopoietic stem cell lineage commitment is a continuous process. *Nat. Cell Biol.* **19**, 271–281.
- Venturutti, L., Teater, M., Zhai, A., Chadburn, A., Babiker, L., Kim, D., Béguelin, W., Lee, T.C., Kim, Y., Chin, C.R., et al. (2020). TBL1XR1 mutations drive extranodal lymphoma by inducing a pro-tumorigenic memory fate. *Cell* **182**, 297–316.e27.
- Wan, X., Liu, L., Zhou, P., Hui, X., He, Q., Yu, F., Zhang, W., Ding, X., Yuan, X., Zhang, N., et al. (2019). The nuclear receptor corepressor NCoR1 regulates hematopoiesis and leukemogenesis in vivo. *Blood Adv.* **3**, 644–657.
- Wang, J.C.Y., and Dick, J.E. (2005). Cancer stem cells: lessons from leukemia. *Trends Cell Biol.* **15**, 494–501.
- Wang, J., Hoshino, T., Redner, R.L., Kajigaya, S., and Liu, J.M. (1998). ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by

- interaction with the human N-CoR/mSin3/HDAC1 complex. *Proc. Natl. Acad. Sci. U S A* 95, 10860–10865.
- Watson, P.J., Fairall, L., and Schwabe, J.W.R. (2012). Nuclear hormone receptor co-repressors: structure and function. *Mol. Cell. Endocrinol.* 348, 440–449.
- Wen, Y.D., Perissi, V., Staszewski, L.M., Yang, W.M., Krones, A., Glass, C.K., Rosenfeld, M.G., and Seto, E. (2000). The histone deacetylase-3 complex contains nuclear receptor corepressors. *Proc. Natl. Acad. Sci. U S A* 97, 7202–7207.
- Wojtowicz, E.E., Lechman, E.R., Hermans, K.G., Schoof, E.M., Wienholds, E., Isserlin, R., van Veelen, P.A., Broekhuis, M.J.C., Janssen, G.M.C., Trotman-Grant, A., et al. (2016). Ectopic miR-125a expression induces long-term repopulating stem cell capacity in mouse and human hematopoietic progenitors. *Cell Stem Cell* 19, 383–396.
- Wong, M.M., Guo, C., and Zhang, J. (2014). Nuclear receptor corepressor complexes in cancer: mechanism, function and regulation. *Am. J. Clin. Exp. Urol.* 2, 169–187.
- Xie, H., Xu, J., Hsu, J.H., Nguyen, M., Fujiwara, Y., Peng, C., and Orkin, S.H. (2014). Polycomb repressive complex 2 regulates normal hematopoietic stem cell function in a developmental-stage-specific manner. *Cell Stem Cell* 14, 68–80.
- Yoon, H., Chan, D.W., Huang, Z., Li, J., Fondell, J.D., Qin, J., and Wong, J. (2003). Purification and functional characterization of the human NCoR complex: the roles of HDAC3, TBL1 and TBLR1. *EMBO J.* 22, 1336.
- Zeng, A.G.X., Bansal, S., Jin, L., Mitchell, A., Chen, W.C., Abbas, H.A., Chan-Seng-Yue, M., Voisin, V., van Galen, P., Tierens, A., et al. (2022). A cellular hierarchy framework for understanding heterogeneity and predicting drug response in AML. Preprint at bioRxiv. <https://doi.org/10.1101/2022.01.25.476266>.
- Zhang, J., Kalkum, M., Chait, B.T., and Roeder, R.G. (2002). The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Mol. Cell* 9, 611–623.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-seq (MACS). *Genome Biol.* 9, R137.
- Zhang, J., Mullighan, C.G., Harvey, R.C., Wu, G., Chen, X., Edmonson, M., Buetow, K.H., Carroll, W.L., Chen, I.-M., Devidas, M., et al. (2011). Key pathways are frequently mutated in high-risk childhood acute lymphoblastic leukemia: a report from the Children's Oncology Group. *Blood* 118, 3080–3087.
- Zhu, L.J., Gazin, C., Lawson, N.D., Pagès, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPpeakAnno: a bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* 11, 237.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
TBL1XR1	Cell Signaling	74499S; RRID:AB_2799857
JARID2	Cell Signaling	13594S; RRID:AB_2798269
AML1 (RUNX1)	Cell Signaling	4334S; RRID:AB_2184099
NCoR1	Cell signaling	5948S; RRID:AB_10834809
LMO2	Cell Signaling	87182S
HEB	Cell Signaling	11825S; RRID:AB_2797736
p300	Cell Signaling	86377S; RRID:AB_2800077
Vinculin	Cell Signaling	13901S; RRID:AB_2728768
GAPDH	Cell Signaling	14C10; RRID:AB_10693448
CEBPa	Cell Signaling	2295S; RRID:AB_10692506
HDAC3	Cell Signaling	3949S; RRID:AB_2118371
CBF $\beta$	Cell Signaling	62184S; RRID:AB_2722525
Flag M2	Sigma	F1804; RRID:AB_262044
ETO	Santa Cruz	sc-134335; RRID:AB_2184234
Mouse IgG Ctrl	ThermoFisher	10400C; RRID:AB_2532980
Rabbit IgG Ctrl	ThermoFisher	10500C; RRID:AB_2532981
Rabbit Anti-Mouse Secondary	ThermoFisher	31188; RRID:AB_228419
AGO2	Sino Biological	50683-R036; RRID:AB_2860516
AGO2 clone 4F9	Santa Cruz	sc-53521; RRID:AB_628697
Mouse monoclonal anti-CD45, PE-Cy5 (J33)	Beckman Coulter	Cat#IM2653U; RRID:AB_10641226
Mouse monoclonal anti-CD45, V500, clone HI30	BD	Cat#560777; RRID:AB_1937324
Mouse monoclonal anti-CD38, PC7, clone HB7	BD	Cat#335790; RRID:AB_399969
Mouse monoclonal anti-CD45RA, FITC, clone HI100	BD	Cat#555488; RRID:AB_395879
Mouse monoclonal anti-CD45RA, BV421, clone HI100	Biologend	Cat#304130; RRID:AB_10965547
Mouse monoclonal anti-CD90, APC (5E10)	BD	Cat#559869; RRID:AB_398677
Rat monoclonal anti-CD49f, PE-Cy5 (GoH3)	BD	Cat#551129; RRID:AB_394062
Mouse monoclonal anti-CD19, PE (4G7)	BD	Cat#349209; RRID:AB_400407
Mouse monoclonal anti-CD71, FITC (L01.1)	BD	Cat#347513; RRID:AB_400316
Mouse monoclonal anti-CD11b, PC5 (Bear1)	Beckman Coulter	Cat#IM3611; RRID:AB_131151
Mouse monoclonal anti-Ki67, FITC (B56)	BD	Cat#556026; RRID:AB_396302
Mouse monoclonal anti-CD3, FITC, clone SK7	BD	349201; RRID:AB_400405
Mouse monoclonal anti-CD7, AF700, clone M-T701	BD	Cat#561603; RRID:AB_10898348
Mouse monoclonal anti-CD19, V450, clone HIB19	BD	560353; RRID:AB_1645564
Mouse monoclonal anti-CD33, BV421, clone WM53	Biologend	303416; RRID:AB_2561690
Mouse monoclonal anti-CD33, APC, clone P67.6	BD	340474; RRID:AB_400518
Mouse monoclonal anti-CD34, APC-Cy7, clone 581	BD	624072
Mouse monoclonal anti-CD36, APC, clone CD38	BD	550956; RRID:AB_398480
Mouse monoclonal anti-CD41, APC-Cy7, clone HIP8	Biologend	303716; RRID:AB_10897646
Mouse monoclonal anti-FLT3, biotin, clone 4G8	BD	624008
Mouse monoclonal anti-GlyA, PC5, clone GA-R2	BD	559944; RRID:AB_397387
Mouse monoclonal anti-CD41, PC7, clone P2	Beckman Coulter	6607115; RRID:AB_2800448
Mouse monoclonal anti-CD45, APC Cy7, clone2D1	BD	348795; RRID:AB_400385
Streptavidin Conjugate FITC	BD	554060; RRID:AB_10053373

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse monoclonal anti-CD7, PeCy5	Beckman Coulter	IM3613U; RRID:AB_10643230
Mouse monoclonal anti-CD10, PeCy5, clone ALB1	Beckman Coulter	IM2721U; RRID:AB_131177
Mouse monoclonal anti-CD117, PeCy7, clone104D2	BD	339195; RRID:AB_647418
Mouse monoclonal anti-CD49d, PE-Cy5, clone 9F10	BD	559880; RRID:AB_397355
Mouse monoclonal anti-CD14, PC7, clone RMO52	Beckman Coulter	A22331; RRID:AB_10639528
Mouse monoclonal anti-CD66b, APC, clone G10F5	BD	561645; RRID:AB_10894001
Mouse monoclonal anti-CD16, AP- Cy7, clone 3G8	BD	557758; RRID:AB_396864
Mouse monoclonal anti-CD14, FITC, clone RMO52	Beckman Coulter	IM0645U; RRID:AB_130992
Mouse monoclonal anti-Flt3, BV711, clone 4G8	BD	563908; RRID:AB_2738479
Mouse monoclonal anti-CD10, AF700, clone HI10a	BD	563509; RRID:AB_2738247
Mouse monoclonal anti-CD15, V450, MMA	BD	642917; RRID:AB_1645751
Mouse monoclonal anti-annexinV, PE	BD	556422; RRID:AB_2869071
Mouse monoclonal anti-annexinV, APC	BD	550475; RRID:AB_2868885)

**Biological samples**

Human umbilical cord blood samples	Trillium, Credit Valley and Brampton Civic Hospital	N/A
Peripheral Blood from CBF AML patient samples	Princess Margaret Cancer Centre	N/A

**Chemicals, peptides and recombinant proteins**

FLT3 Ligand, human, premium grade	Miltenyi Biotec	130-096-479
G-CSF, human, premium grade	Miltenyi Biotec	130-093-861
TPO, human, premium grade	Miltenyi Biotec	130-095-752
IL-6, human, premium grade	Miltenyi Biotec	130-093-932
SCF, human, premium grade	Miltenyi Biotec	130-096-695
IL-3, human, premium grade	Miltenyi Biotec	130-095-069
IL-7, human, premium grade	Miltenyi Biotec	130-095-363
Propidium Iodide - 1.0 mg/mL Solution in Water	Thermo Fisher	P3566
AmpliTaq Gold 360 master mix	Thermo Fisher	4398881
DNase I	Roche	11284932001
Protease and Phosphatase Inhibitor Single-Use Cocktail	Thermo Fisher	78446
Ammonium Chloride	Stem Cell Technologies	07850
Bovine Serum Albumin Fraction V	Roche	10735086001
DMSO	Fisher Scientific	D128-500
Agarose	Thermo Fisher	16500500
EDTA	Sigma	E5134
1M Tris-HCl, pH 7.5	Thermo Fisher	15567027
5M NaCl solution	Sigma	S6546-1L
1M MgCl <sub>2</sub>	Thermo Fisher	AM9530G
IGEPAL CA-630	Sigma	I8896
Tween 20	Sigma	P9416
Sytox Blue	Thermo Fisher	S34857
Protein A-Micrococcal nuclease fusion protein	S. Henikoff's lab	N/A
BamHI-HF	NEB	Cat#R0136S
MluI-HF	Thermo Fisher	Cat#15432-016
Hoescht 33342	Thermo Fisher	Cat#H3570
FseI	NEB	R0588S
XhoI	NEB	R0146S
Pacl	NEB	R0547S
SphI-HF	NEB	R3182S

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
KOD Hot Start Polymerase	Sigma	71086
DpnI	NEB	R0176S
Sodium Butyrate	Sigma	B5887-1G
Trypsin MS grade	Promega	V5280
Guanidium Hydrochloride	Sigma	G3272
TCEP	Sigma	C4706
CAA	Sigma	C0267
Acetonitrile HPLC grade	Sigma	Merck 1.00020.1000
Trifluoroacetic Acid (TFA)	Sigma	302031
Formic Acid	Fluka	56302
Ammonium Acetate	Sigma	17836
Ezview Red ANTI-FLAG M2 Affinity Gel	Sigma	F2426
T4 PNK 3' phosphatase minus	NEB	M0236
Protease Inhibitor	Thermo Fisher	78437
3X FLAG peptide	Sigma	F4799
pA-Mnase	Henikoff's lab	N/A
Power SYBR Green Master Mix	Thermo Fisher	4367659
LifeTech Protease Inhibitor Cocktail III	LifeTech	539134-1SET
Proteinase K	NEB	P8107S
SuperScript IV Reverse Transcriptase	Thermo Fisher	18080044
Exo-SAP-IT	Affymetrix	78201
T4 PNK	NEB	M0201L
T4 RNA ligase 1 high conc	NEB	M0437M
RNase I	LifeTech	AM2295
Turbo Dnase	LifeTech	AM2239
<b>Critical commercial assays</b>		
CD34 MicroBead kit	Miltenyi Biotec	130-046-703?
Mouse cell depletion kit	Miltenyi Biotec	130-104-694
Click-iT EdU AF647 Flow Cytometry Assay Kit	Thermo Fisher	C10634
RNeasy Plus Micro kit	Qiagen	74034
MinElute PCR Purification kit	Qiagen	28004
Ampure XP beads	Beckman Coulter	A63880
CalPhos Mammalian Transfection Kit	Clontech	631312
SMART-Seq V4 Ultra Low Input RNA kit	Clontech	634891
Nextera XT DNA Library Preparation kit	Illumina	FC-131-1096
High Sensitivity DNA kit	Agilent	5067-4626
RNA 6000 Pico kit	Agilent	5067-1513
NEBNext Ultra II DNA Library Prep Kit for Illumina	NEB	E7645
NEBNext Multiplex Oligos for Illumina	NEB	E7500
PicoPure RNA Isolation Kit	Thermo Fisher	Cat#KIT0204
StemSep Human Hematopoietic Progenitor Cell Enrichment Kit	StemCell Technologies	Cat#14056
Anti-human CD41-TAC	StemCell Technologies	Cat# 14050
Quick Gel Extraction Kit	Qiagen	K210012
5X Rapid Ligation Kit	Thermo Fisher	K1422
Gateway LR Clonase Kit	Thermo Fisher	11791020
Mouse/Human Chimera Enrichment Kit	StemCell Technologies	28890
Zymo RNA Clean and Concentrator Kit	Zymo Research	R1013

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MirVana miRNA isolation kit	Thermo Fisher	AM1560
TaqMan microRNA reverse transcription kit	Thermo Fisher	4366596
Rneasy Mini Kit	Qiagen	74104
LUNA Universal qPCR 2x Master Mix	NEB	M3003
Q5 High-Fidelity 2x Master Mix	NEB	M0492
12-230 kDa Separation Module	ProteinSimple	SM-W004
66-440 kDa Separation Module	ProteinSimple	SM-W008
Anti-rabbit Detection Module	ProteinSimple	DM-001
Anti-mouse Detection Module	ProteinSimple	DM-002

**Deposited data**

RNA-seq data miR-130a OE in CD34+ CB; miR-130a KD in Flag-AML1-ETO Kasumi-1, TBL1XR1 in CD34+CD38- CB cells	JE Dick lab, Princess Margaret Cancer Centre	GSE181140
Total and miR-130a Targeted Chimeric AGO2 eCLIP in CD34+ HSPC from CB and Kasumi-1 cells	GW Yeo lab, UCSD	GSE181140
Protein Mass Spectrometry data miR-130a OE in CD34+ CB	JE Dick lab, Princess Margaret Cancer Centre	ProteomeXchange Consortium PXD027331
Microarray data miR-130a expression in CBF AML patients	L Salmena, Princess Margaret Cancer Centre	GSE181140
CUT&RUN data Flag-AML-ETO in Kasumi-1 with miR-130a KD	JE Dick lab/D De Carvalho lab Princess Margaret Cancer Centre	GSE181140
Raw RNA-seq and CUT&RUN data	JE Dick lab, Princess Margaret Cancer Centre	EGAD00001008412

**Experimental models: Cell lines**

Flag-AML1-ETO Kasumi-1	S Hiebert's lab, Vanderbilt U	N/A
Kasumi-1 t(8:21) (q22;q22) AML	ATCC	CRL-2724
HE293T	ATCC	CRL-3216
MS-5 CSF1	M. Minden lab, Princess Margaret Cancer Centre	N/A

**Experimental models: Organisms/strains**

NOD.Cg Prkdcscid <sup>l12rgtm1Wjl</sup> /SzJ (NSG)	The Jackson Laboratory	Cat#005557; RRID:IMSR_JAX:005557
NOD.Cg-Prkdc <sup>scid</sup> Il2rg <sup>tm1Wjl</sup> /Tg (CMV-IL3,CSF2,KITLG) 1Eav/MloySzJ (NSG-SGM3)	The Jackson Laboratory	Cat#013062; RRID:IMSR_JAX:013062

**Oligonucleotides**

See [Table S6](#)

**Recombinant DNA**

shRenilla	<a href="#">Kaufmann et al., 2019</a>	pLBC-BS lentiviral vector
shTBL1XR1_1	This manuscript	pLBC-BS lentiviral vector
shTBL1XR1_2	This manuscript	pLBC-BS lentiviral vector
shTBL1XR1_3	This manuscript	pLBC-BS lentiviral vector
shTBL1XR1_4	This manuscript	pLBC-BS lentiviral vector
LV.SFFV.intron.mOrange2 Control	<a href="#">Lechman et al., 2016</a>	LV.SFFV.intron.mOrange2
miR-130a_OE_mO	This manuscript	LV.SFFV.intron.mOrange2
LV.SFFV.GFP Control	<a href="#">Lechman et al., 2016</a>	LV.SFFV.GFP
miR-130a_KD_GFP	This manuscript	LV.SFFV.GFP
pLV-TagBFP2-SFFV>ORF_stuffer	VectorBuilder	pLV-TagBFP2-SFFV
pLV-TagBFP2-SFFV>hTBL1XR1	VectorBuilder	pLV-TagBFP2-SFFV

**Software and algorithms**

FACSDiva v. 8.0.1.1	BD	N/A
FlowJo v 9.96 and 10.6.2	FlowJo, LLC	N/A

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Prism 8.0.2	GraphPad Software	N/A
Extreme Limiting Dilution Analysis (ELDA)	Hu and Smyth, 2009	<a href="http://bioinf.wehi.edu.au/software/elda/">http://bioinf.wehi.edu.au/software/elda/</a>
GSEA v 4.1.0	N/A	<a href="http://software.broadinstitute.org/gsea/index.jsp">http://software.broadinstitute.org/gsea/index.jsp</a>
Cytoscape v 3.8.2	N/A	<a href="https://cytoscape.org">https://cytoscape.org</a>
R 3.5.0, R Studio 1.2	N/A	<a href="http://www.r-project.org/">http://www.r-project.org/</a> ; <a href="https://rstudio.com">https://rstudio.com</a>
STAR v.2.5.2b, 2.6.0c and 2.7.6a	Dobin et al., 2013	<a href="https://github.com/alexdobin/STAR">https://github.com/alexdobin/STAR</a>
HTSeq v.0.7.2	Anders et al., 2015	<a href="https://github.com/simon-anders/htseq">https://github.com/simon-anders/htseq</a>
picard v2.6.0	N/A	<a href="http://broadinstitute.github.io/picard/">http://broadinstitute.github.io/picard/</a>
DESeq2 v.1.22.2	Love et al., 2014	<a href="https://github.com/mikelove/DESeq2">https://github.com/mikelove/DESeq2</a>
Python 3.7.0	N/A	<a href="https://www.python.org/">https://www.python.org/</a>
BWA	Langmead et al., 2009	<a href="https://github.com/lh3/bwa">https://github.com/lh3/bwa</a>
GENCODE v32	Frankish et al., 2019	<a href="https://www.encodegenes.org/human/release_32.html">https://www.encodegenes.org/human/release_32.html</a>
MaxQuant 1.5.2.8	Cox and Mann, 2008	<a href="https://www.maxquant.org/">https://www.maxquant.org/</a>
mirDIP v1.0	N/A	<a href="https://ophid.utoronto.ca/mirDIP/">https://ophid.utoronto.ca/mirDIP/</a>
DIANA-TarBase v7.0	N/A	<a href="http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=tarbase/index">http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=tarbase/index</a>
Bowtie 2.2.6	Langmead and Salzberg, 2012	<a href="https://sourceforge.net/projects/bowtie-bio/files/bowtie2/2.2.6/">https://sourceforge.net/projects/bowtie-bio/files/bowtie2/2.2.6/</a>
miRBase release 22.1	Griffiths-Jones, 2006	<a href="http://www.mirbase.org/">http://www.mirbase.org/</a>
HOMER 4.9.1-6, 4.8	Heinz et al., 2010	<a href="http://homer.ucsd.edu/homer/">http://homer.ucsd.edu/homer/</a>
Compass for SW v 4.0.0	N/A	<a href="https://www.proteinsimple.com/compass/downloads/">https://www.proteinsimple.com/compass/downloads/</a>
CIBERSORTx	Newman et al., 2019	<a href="https://cibersortx.stanford.edu/">https://cibersortx.stanford.edu/</a>
PANTHER v. 15.0	Mi et al., 2021, Thomas et al., 2003	<a href="http://pantherdb.org/">http://pantherdb.org/</a>
<b>Other</b>		
IMDM media	Thermo Fisher	12440053
X-VIVO 10 hematopoietic serum-free culture media	Lonza	Cat#04-380Q
Fetal Bovine Serum (FBS)	Multicell	98150
Phosphate-Buffered Saline (PBS), pH 7.4	Multicell	311-010-LL
L-Glutamine	Multicell	Cat#609-065-EL
Penicillin-Streptomycin	Gibco	15140-122
autoMACS Running Buffer	Miltenyi Biotec	130-091-221
DEPC-Treated H <sub>2</sub> O	Thermo Fisher	AM9916
RIPA buffer	Thermo Fisher	89900
LS column	Miltenyi Biotec	130-042-401
Lymphocyte Separation medium	Wisent	305-010-CL
RPMI	Multicell	350-000-CL
BIT9500	Stem Cell Technologies	09500
4-12% SDS polyacrylamide gels	Thermo Fisher	NP0321
Nitrocellulose membrane	Thermo Fisher	IB23001
NuPAGE Transfer Buffer (20X)	Thermo Fisher	NP00061
NuPAGE MOPS SDS Running buffer (20X)	Thermo Fisher	LSNP0001
Affymetrix™ ExoSAP-IT™ For PCR Product Clean-Up	Affymetrix	501128967
Dynabeads MyOne Silane beads	Thermo Fisher	37002D
Dynabeads M-280 sheep anti-rabbit	Thermo Fisher	11203D
Concanavalin A Beads	Bangs Laboratories	BP531
AMPure XP beads	Beckman	A63880

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, John Dick ([john.dick@uhnresearch.ca](mailto:john.dick@uhnresearch.ca)).

### Materials availability

All unique/stable reagents generated in this study are available from the lead contact without restriction. Experimental cell lines, mouse strains, recombinant DNA and oligonucleotides are listed in the [Key Resources Table](#).

### Data and code availability

- CLIP-seq, RNA-seq, CUT&RUN and microarray data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the [key resources table](#). Mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository. The dataset identifier is listed in the [key resources table](#).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Samples cryopreservation and thawing

All samples were cryopreserved in 80% FBS (Multicell) and 20% DMSO (FisherScientific) solution added in 1:1 ratio. Samples were stored at  $-80^{\circ}\text{C}$  for short-term storage or  $-150^{\circ}\text{C}$  for long-term storage. All primary cells and cell lines were thawed by slow, drop-wise addition of thawing medium containing 50% X-VIVO 10 (Lonza), 50% FBS and DNaseI (100  $\mu\text{g}/\text{ml}$ ). Subsequently, cells were centrifuged at 1,450 rpm for 10 min at RT and resuspended in PBS (Multicell) + 2.5% FBS.

### Cord blood processing and HSPC enrichment

Human cord blood (CB) samples were obtained with informed consent from Trillium Health Center, Brampton Civic Hospital and Credit Valley Hospital, Ontario, Canada in accordance with guidelines approved by the University Health Network (UHN) Research Ethics Board. Cord blood samples were diluted 1:1 with PBS and mononuclear cells were enriched by density gradient centrifugation with lymphocyte separation medium (Wisent). Red blood cells were lysed with ammonium chloride solution (StemCell Technologies) and hematopoietic stem and progenitor cells were isolated with StemSep Human Hematopoietic Progenitor Enrichment Kit and Anti-Human CD41 TAC (StemCell Technologies) by negative selection according to manufacturer's instructions. Lineage depleted ( $\text{Lin}^{-}$ ) CB cells were resuspended in PBS +2% FBS and cryopreserved at  $-150^{\circ}\text{C}$ .

### Cell culture

CB cells were cultured in low cytokine conditions using X-VIVO 10 medium with 1% BSA, 1X penicillin-streptomycin (Gibco), 1X L-glutamine (Multicell) supplemented with the following cytokines: Flt3L (100 ng/ml), SCF (100 ng/ml), TPO (50 ng/ml) and IL-7 (10 ng/ml). All cell lines listed in the Key Resources table were grown at  $37^{\circ}\text{C}$ , 5%  $\text{CO}_2$ . Kasumi-1 and Flag-AML1-ETO Kasumi-1 cells were cultured in RPMI medium +20% FBS, 1% Pen/Strep. Flag-AML1-ETO Kasumi-1 cell line was generously provided by S.W. Hiebert's lab. Primary t(8;21)  $\text{CD}34^{+}$  blasts from AML patient samples were cultured in X-VIVO 10 medium supplemented with 20% BIT9500 (StemCell Technologies) and the following cytokines: IL-3 (10 ng/mL), SCF (50 ng/mL), IL-6 (10 ng/mL), Flt3L (50 ng/mL), G-CSF (10 ng/mL) and TPO (25 ng/mL) on MS5 stroma expressing human soluble CSF1. MS5 stroma cell line expressing human CSF1 was generously provided by M.D. Minden's lab. Details of AML patient samples are outlined in [Table S5](#). MS5-CSF1 stroma cells were grown in 6-well tissue culture treated plates seeded at a density of  $1-1.5 \times 10^5$  MS5 stroma cells/well. MS5-CSF1 stroma cells and HEK293T cells were cultured in IMEM medium +10% FBS, 1% Pen/Strep. Cells were propagated in T75 tissue culture treated flasks and passaged every 2-4 days for up to 7-8 passages.

### Mouse models

All mouse experiments were performed in accordance with the institutional guidelines approved by the University Health Network Animal Care Committee. All mouse strains used in this study are listed in the Key Resource table. All xenotransplantation experiments were performed with 8-12 week-old male and female NSG (NOD.Cg-Prkdc<sup>scid</sup> Il2rg<sup>tm1Wjl</sup>/Szj) and 8-12 week-old male and female NSG-SGM3 (NOD.Cg-Prkdc<sup>scid</sup> Il2rg<sup>tm1Wj</sup> Tg(CMV-IL3, CSF2, KITLG)1Eav/MloySzj) (JAX) mice. All mice were housed at the animal facility (ARC) at Princess Margaret Cancer Centre in a room designated only for immunocompromised mice with individually ventilated racks equipped with complete sterile micro-isolator caging (IVC), on corn-cob bedding and supplied with environmental enrichment in the form of a red house/tube and a cotton nestlet. Cages are changed every <7 days under a biological safety cabinet. Health status is monitored using a combination of soiled bedding sentinels and environmental monitoring. Mice were sublethally irradiated

(225 cGy) 24–48 h before transplantation. Intrafemoral injections were performed on anesthetized mice by injecting 30  $\mu$ l of cells resuspended in PBS into the right femur.

## METHOD DETAILS

### Lentiviral vectors

Lentiviral vectors for ectopic miRNA expression and stable knockdown have been described previously (Amendola et al., 2009; Gentner et al., 2009; Lechman et al., 2012, 2016). Briefly, pre-miRNA sequences were PCR amplified using sequence specific primers with Fse1 and Xho1 restriction enzyme sites (Table S6) and genomic DNA from CD34<sup>+</sup> CB cells. The PCR product was subsequently purified using the MinElute PCR Purification kit (Qiagen) and digested with Fse1 and Xho1. Digested DNA fragments encoding human pre-miRNA sequences (~350 nt in length) were cloned into the intron of LV.SFFV.intron.mOrange2 vector backbone. Empty vector backbone was used as an mO<sup>+</sup> control (Ctrl). For stable knockdown of miR-130a, DNA fragment containing 8 tandem copies of imperfectly complementary sequence to mature miR-130a and Pac1 and Sph1 restriction enzyme sites (Table S6) was ordered through GeneArt Gene Synthesis (ThermoFisher). Following restriction enzyme digest and gel electrophoresis, the DNA fragment was cloned into the LV.SFFV.GFP backbone using Rapid Ligation Kit (Thermo Fisher). Empty vector backbone was used as a negative control (Ctrl).

shRNA sequences for TBL1XR1 shRNA vectors were predicted using Sherwood algorithm (Knott et al., 2014) and ordered as Ultra-mer DNA oligos (IDT). The negative control for knockdown lentivirus (shCtrl) contains a sequence directed against Renilla luciferase in the pLBC2 backbone with BFP fluorescent reporter gene. Four different shRNA sequences for TBL1XR1 were PCR amplified using AmpliTaq Gold 360 Polymerase (ThermoFisher) with sequence specific primers containing BamH1 and Mlu1 restriction enzymes sites on the forward and reverse primers, respectively (Table S6). The PCR product was subsequently purified using the MinElute PCR Purification kit and digested with BamH1 and Mlu1. Digested PCR product and vector backbone were resolved by gel electrophoresis and extracted from the gel with Quick Gel Extraction Kit (Qiagen). Isolated PCR product was cloned into an Ultramir cassette (miR30) within a pRRL-based vector downstream of an SFFV promoter. Knock-down efficiency was assessed by qRT-PCR using TaqMan assay for TBL1XR1 (Thermo Fisher Cat#4331182; Hs00226564\_m1) and GAPDH (Hs02758991\_g1) and/or western blot assay using primer/probe sets and antibodies listed in the Table S6.

The TBL1XR1 overexpression vector (OE) was purchased from VectorBuilder as a codon-optimized pLV-TagBFP2-SFFV > hTBL1XR1 lentiviral vector. As a negative control, pLV-TagBFP2-SFFV > ORF\_stuffer lentiviral vector encoding E.coli beta-galactosidase ORF was used

### Lentivirus production and titration

VSV-G pseudotyped lentiviral vector particles were produced in 293T cells using 2<sup>nd</sup> and 3<sup>rd</sup> generation packaging protocols described previously (Follenzi et al., 2000; Guenechea et al., 2000). 293T cells were seeded in 15 mm plates with approximately 9–10  $\times$  10<sup>6</sup> cells per plate one day before transfection to achieve ~80% confluency at the time of transfection. Subsequently, cells were transfected using CalPhos Mammalian Transfection Kit (Takara) according to the manufacturer's protocol. Fresh AMEM medium +1%FBS, 1X L-glutamine, 1X penicillin-streptomycin and sodium butyrate (1mM) was added to transfected plates ~16 h post-transfection. Virus-containing medium was collected 48 h post-transfection and cell debris was pelleted by centrifugation at 1,800 rpm for 15 min at RT. Subsequently, supernatant was filtered using 0.45  $\mu$ m filter, virus particles were concentrated using ultracentrifugation at 22,000 rpm for 2.5 h at 4°C and stored at –80°C. Lentiviruses were titrated on 293T cells.

### Lentiviral transduction

Transduction of Lin<sup>–</sup> CB CD34<sup>+</sup>CD38<sup>–</sup> cells and other HSPC populations was performed in low cytokine conditions described above. Sorted CD34<sup>+</sup>CD38<sup>–</sup> CB cells were transduced with lentiviruses at a density of 5  $\times$  10<sup>5</sup> cell/ml. Fresh medium was added 24 h post-transduction and cells were harvested 72 h post-transduction and subsequently transplanted into mice or collected for other experimental procedures. Transduction efficiency was determined by removing 10% of the transduced cells and measuring the percentage of mOrange<sup>+</sup>, GFP<sup>+</sup> or BFP<sup>+</sup> cells by flow cytometry. Transduction of Kasumi-1 and Flag-AML1-ETO Kasumi-1 cells was performed at a density of 2  $\times$  10<sup>6</sup> cells/ml in RPMI medium with 20% FBS, 1% Pen/Strep. Cells were transduced in 24-well suspension plates and centrifuged at 1,400 rpm for 60 min at 32°C and fresh medium was added on the next day. For xenotransplantation of primary t(8;21) AML blasts, cells were enriched for CD34<sup>+</sup> blasts using human CD34 Microbead Kit (Miltenyi Biotec) according to manufacturer's instructions. Enriched CD34<sup>+</sup> blasts were transduced at a density of 1.5–9 $\times$ 10<sup>6</sup> cells/ml in X-VIVO + 20% BIT 9500 supplemented with the following cytokines: IL-3 (10 ng/ml), SCF (50 ng/ml), IL-6 (10 ng/ml), Flt3L (50 ng/ml), G-CSF (10 ng/ml) and TPO (25 ng/ml) and fresh media was added on the following day. Transduction efficiency was determined 72 h post-infection by flow cytometry. For AML I, 3  $\times$  10<sup>5</sup> CD34<sup>+</sup> blasts were transplanted per mouse. For AML II, 1  $\times$  10<sup>6</sup> CD34<sup>+</sup> blasts were transplanted per mouse. For *in vitro* experiments, CD34<sup>+</sup> blasts were transduced as described and transferred to MS5 CSF1 stroma 24 h post-transduction. Blasts were sorted for GFP<sup>+</sup> cells 72 h post-transduction, cultured for an additional 5 days and then analyzed by flow cytometry.

### Cell sorting and flow cytometric analysis

Cord blood cells, primary AML cells and cell lines were resuspended at 10<sup>6</sup> cells/ml in PBS +2% FBS for surface antibody staining and antibodies were used as indicated (Key Resources Table). Cells were stained for 20 min at room temperature and washed with

PBS. Cell sorting was performed at FACSaria III (BD Biosciences), SH800 (Sony) or MoFlo (Beckman Coulter) sorters or analyzed on Canto II or Celesta instruments (BD). Antibodies are listed in the [Key Resource Table](#).

### Xenotransplantation

For CB transplantation, progeny of  $\sim 1 \times 10^4$  CD34<sup>+</sup>CD38<sup>-</sup> CB cells were injected per mouse 72 h post-transduction. The miRNA OE screen was performed with 3 independent cord blood pools each transduced with 3 different miRNA OE and 1 control lentivirus and transplanted into 4–5 mice per experimental condition. For miR-130a OE repopulation assays, 2–4 independent CB pools were transduced for each experiment and transplanted into 8–10 mice per experimental condition. Mice were sacrificed after 12 or 24 weeks of repopulation and right femur (RF) and bone marrow (left femur and tibia) were flushed with 1 ml of PBS +2.5% FBS. Spleens were weighted, crushed and collected using cell strainer caps. All cells were centrifuged at 1,450 rpm for 10 min at RT and resuspended in 500  $\mu$ l of PBS +2.5% FBS. Cells were counted in ammonium chloride using Vicell XR (Beckman Coulter) and used for flow cytometric analysis or other assays.

Secondary transplantations for miR-130a OE and TBL1XR1 KD studies were performed by combining cells from 2–6 primary recipient mice and depleting murine cells using the Mouse Cell Depletion Kit (Miltenyi Biotec) according to manufacturer's instructions. Murine-depleted cells were sorted for CD45<sup>+</sup>mO<sup>+</sup> or CD45<sup>+</sup>BFP<sup>+</sup> cells and transplanted at defined doses into NSG or NSG-SGM3 mice using intrafemoral injections. Secondary recipient mice were sacrificed at week 8–12, right and left femur were flushed and analyzed by flow cytometry. Engraftment was scored positive if the percentage of CD45<sup>+</sup> BFP<sup>+</sup>/mO<sup>+</sup> cells was >0.05% in the bone marrow. LT-HSC frequency was estimated by linear regression analysis and Poisson statistics using the publicly available online software ELDA ([Key Resources Table](#)).

### Lin<sup>-</sup> Cell depletions or CD34<sup>+</sup> enrichment

For flow cytometric analysis of HSPC populations from miR-130a OE and miR-130a KD xenografted mice, cells from RF and BM of 2–4 mice were enriched for human HSPC with Mouse/Human Chimera Enrichment Kit (STEMCELL Technologies), StemSep Human Hematopoietic Progenitor Enrichment Kit (STEMCELL Technologies) and Anti-Human CD41 TAC (STEMCELL Technologies, 14050). For enrichment of CD34<sup>+</sup> blasts from peripheral blood of CBF-AML primary patient samples, Direct CD34<sup>+</sup> Progenitor Isolation Kit (Miltenyi Biotec) was used according to manufacturer's instructions. For secondary transplantation experiments and analysis of HSPC from TBL1XR1 KD xenografts, Mouse Cell Depletion Kit (Miltenyi Biotec) was used to deplete murine cells and isolated human cells were subsequently analyzed by flow cytometry or sorted using FACS for secondary transplantations or RNA-seq. Cells from 2–4 individual mice were combined prior to depletion of murine cells.

### Flow cytometric analysis of xenografts

Cells from RF, BM, or spleen of xenografted mice (50  $\mu$ l) were stained in 4 ml polystyrene tubes (Falcon, 352058) or 96 well U bottom plates (Corning, 351177) for 30 min at 4°C. Following antibodies ([Key Resources Table](#)) were used: 1. For the Pan Lin stain, CD3 FITC (1:100), GlyA PE Cy5 (1:100) or GlyA PE (1:100), CD41 PE Cy7 or CD41 PE Cy5 (1:100), CD33 APC, CD45 APC Cy7 (1:100) or CD45 V500(1:50), CD19 V450 (1:100) or CD19 PE C7 (1:100); 2. For the erythroid stain, CD71 FITC (1:100), GlyA PE Cy5 (1:100), CD117 PE cy7 (1:100), CD36 APC (1:50), CD34 APC Cy7 (1:200); 3. For the myeloid stain, CD45 FITC (1:100), CD49d PE Cy5 (1:50), CD14 PE Cy7 (1:100), CD66b APC (1:100), CD16 APC Cy7 (1:50), CD33 V450 (1:100). For the full stem and progenitor hierarchy stain, murine-depleted cells from TBL1XR1 KD and Ctrl xenografts were stained with the following antibodies: CD45RA FITC (1:50), CD45 V500 (1:50), Flt3 BV711 (1:50), CD7/CD10/CD19 AF700 (1:100), CD38 PE-Cy7 (1:100), CD90 APC (1:50), CD34 APC Cy7 (1:200) and CD49f PE-Cy5. For the miR-130a OE and KD, xenografts depleted of murine and lineage committed cells were stained with the following antibodies: Flt3-biotin (1:50), STV-FITC (1:100), CD7/10 PE-Cy5 (1:100), CD38 PE-Cy7 (1:100), CD90 APC (1:50), CD34 APC Cy7 (1:200), CD45RA BV421 or CD45RA FITC (1:100).

### RNA isolation

Sorted populations from CB or Kasumi-1 cells ( $\sim 5 \times 10^3$  –  $5 \times 10^4$  cells) were washed in PBS and centrifuged at 1,450 rpm for 10 min and RT. Cell pellets were frozen at  $-80^\circ\text{C}$  and total RNA was extracted using RNeasy Plus Micro Kit (Qiagen) or PicoPure RNA Isolation Kit (ThermoFisher) according to manufacturer's instructions. RNA integrity was assessed using Agilent Bioanalyzer (RNA Pico chip Agilent Technologies) and next-generation sequencing libraries were prepared at the Center of Applied Genomics, SickKids Hospital. SMART-Seq V4 Ultra Low Input RNA kit for sequencing (Clontech) was used to generate the cDNA samples. Subsequently, next-generation sequencing libraries were prepared with Nextera XT DNA Library Preparation Kit (Illumina). Equimolar quantities of libraries were pooled and sequenced on the Illumina HiSeq 2500 platform generating 125-bp paired-end sequencing reads to achieve depth of approximately 40–50M reads per sample. All RNAseq experiments were performed *in vitro* in triplicates using three individually transduced samples.

### RNA-seq processing and pathway enrichment analysis

Raw sequencing reads were trimmed to remove the adaptors and aligned against hg38 human genome with STAR v2.5.2b ([Dobin et al., 2013](#)) using GENCODE ([Frankish et al., 2019](#)). Read counts were generated using HTSeq counts v0.7.2 ([Anders et al., 2015](#)) and general statistics were obtained from picard/2.6.0 using Default parameters. Subsequently, unstranded counts were normalized

using normalization factor and annotated using BioMart\_ENSG\_hg38. For TBL1XR1 KD in CB, STAR v2.7.2 and HTSeq was used to count reads over GENCODE v32 features using default parameters. EdgeR\_3.28.1 (Robinson et al., 2010) was used to fit a glmQ model to generate estimation of differential expression between miR-130a OE in CB and miR-130a KD in Kasumi and TBL1XR1 KD in CB conditions and their respective control samples. A score was calculated using the formula  $\text{sign}(\log\text{FC}) \cdot \log_{10}(\text{p value})$  to rank all genes from top upregulated to downregulated and was subsequently used in pathway enrichment analysis.

#### Pathway enrichment analysis of miR-130a OE in CB

The ranked gene lists were used in GSEA\_4.1.0 using the Baderlab gene set file containing pathways from multiple databases (June 01 2020, <http://baderlab.org/GeneSets/>), using as parameters 2000 permutations and gene set size between 10 and 500. EnrichmentMap 3.3.2 in Cytoscape 3.8.2 was used to visualize the GSEA enrichment results and AutoAnnotate 1.3.3 was used to cluster and label functional modules. Significant overlap at a p value of 0.05 between the downregulated proteins in miR-130a OE VS Control and DIANA-TarBase v7.0 miR-130a targets was visualized in Cytoscape/EnrichmentMap using the post-analysis function. GSEA was also run on the following selected gene sets. Cell cycle gene sets used in the miR-130a OE GSEA including G1 S DNA damage, Myc targets and positive regulation of mitotic cell cycle were obtained from the GO biological process and MSig C2 databases. Proliferative and quiescent gene sets were obtained from Forsberg et al. comparing gene expression of quiescent and cytokine-induced mobilized HSC in mice (Forsberg et al., 2010). RA-target genes were extracted from several publicly available data (Balmer and Blomhoff, 2002; Cabezas-Wallscheid et al., 2017; Ghiaur et al., 2013). For the RA-target gene dataset from Ghiaur et al., (2013), DE genes between CD34<sup>+</sup>CD38<sup>-</sup> and CD34<sup>+</sup>CD38<sup>+</sup> cells from bone marrow donors and predicted RA targets (Table S1) were used. RA-target genes from Balmer and Blomhoff (2002) dataset were extracted, genes were converted to human orthologs using g:convert (<https://biit.cs.ut.ee/gprofiler/convert>) and separated into upregulated and downregulated by RA.

#### Pathway enrichment analysis of miR-130a KD in kasumi

The ranked gene lists were used in GSEA\_4.1.0 using the Baderlab gene set file containing pathways from multiple databases (June 01 2020, <http://baderlab.org/GeneSets/>), using as parameters 2000 permutations and gene set size between 10 and 500. EnrichmentMap 3.3.2 in Cytoscape 3.8.2 was used to visualize the GSEA enrichment results and AutoAnnotate 1.3.3 was used to cluster and label functional modules. Overlap between the enriched pathways and DIANA-TarBase v7.0 miR-130a targets was assessed using Fisher's exact test and significant overlap at p value 0.05 was visualized using the post-analysis function.

#### Pathway enrichment analysis of TBL1XR1 KD in CB

The ranked gene lists were used in GSEA\_4.1.0 using the Baderlab gene set file containing pathways from multiple databases (June 01 2020, <http://baderlab.org/GeneSets/>), using as parameters 2000 permutations and gene set size between 10 and 500. EnrichmentMap 3.3.2 in Cytoscape 3.8.2 was used to visualize the GSEA enrichment results and AutoAnnotate 1.3.3 was used to cluster and label functional modules. Overlap between the enriched pathways and eCLIP miR-130a targets was assessed using Fisher's exact test and significant overlap at p value 0.05 was visualized using the post-analysis function. The same gene sets used in the miR-130a OE analysis were used in the GSEA of TBL1XR1 KD differentially expressed (DE) genes including the proliferative and quiescent gene sets (Forsberg et al., 2010) and MYC targets and positive regulation of mitotic cell cycle obtained from the GO biological process and MSig C2 databases. RA-upregulated and downregulated genes were obtained from Cabezas-Wallscheid et al. (2017) comparing *in vitro* ATRA treatment of murine HSC to untreated HSC (LSK CD150 + CD48 – CD34 –) and RA-target genes from Balmer and Blomhoff (2002).

#### Label free-mass spectrometry samples preparation

Human Lin<sup>-</sup>CB cells from 3 independent cord blood pools were transduced individually with miR-130a OE and control lentiviruses and collected for FACS 72 h post-transduction. Equal number of sorted CD34<sup>+</sup>mO<sup>+</sup> cells ( $1 \times 10^5$ ) were washed twice with ice-cold PBS and resulting samples were subjected to sample preparation similar to Schoof et al. (2016) (Schoof et al., 2016). Cells were lysed using 20  $\mu$ l of lysis buffer (6M Guanidinium Hydrochloride, 10 mM TCEP, 40 mM CAA, 100 mM Tris pH8.5). Samples were boiled for 5 min 95°C, after which they were sonicated on high for  $3 \times 10$  s in a Bioruptor sonication water bath (Diagenode) at 4°C. Samples were diluted 1:3 with 10% Acetonitrile, 25 mM Tris pH 8.5, LysC (MS grade, Wako) was added in a 1:50 (enzyme to protein) ratio, and samples were incubated at 37°C for 4 h. Samples were further diluted to 1:10 with 10% Acetonitrile, 25 mM Tris pH 8.5, trypsin (MS grade, Promega) was added in a 1:100 (enzyme to protein) ratio and samples were incubated overnight at 37°C. Enzyme activity was quenched by adding 2% trifluoroacetic acid (TFA) to a final concentration of 1%. Prior to mass spectrometry analysis, the peptides were desalted and fractionated on in-house packed SCX Stagetips. For each sample, 3 discs of SCX material (3M Empore) were packed in a 200  $\mu$ l tip, and the SCX material activated with 80  $\mu$ l of 100% Acetonitrile (HPLC grade, Sigma). The tips were equilibrated with 80  $\mu$ l of 0.2% TFA, after which the samples were loaded using centrifugation at 4,000x rpm. After washing the tips twice with 100  $\mu$ l of 0.2% TFA, two initial fractions were eluted into clean 500  $\mu$ l Eppendorf tubes using 75 mM and 300 mM ammonium acetate in 20% Acetonitrile, 0.5% formic acid respectively. The final fraction was eluted using 5% ammonium hydroxide, 80% Acetonitrile. The eluted peptides were frozen on dry ice and concentrated in an Eppendorf Speedvac, and reconstituted in 1% TFA, 2% Acetonitrile for Mass Spectrometry (MS) analysis.

#### Mass spectrometry acquisition

For each sample, peptides were loaded onto a 2cm C18 trap column (ThermoFisher 164705), connected in-line to a 50 cm C18 reverse-phase analytical column (Thermo EasySpray ES803) using 100% Buffer A (0.1% Formic acid in water) at 750 bar, using

the Thermo EasyLC 1000 HPLC system in a single-column setup and the column oven operating at 45°C. Peptides were eluted over a 200 min gradient ranging from 5 to 48% of 100% acetonitrile, 0.1% formic acid at 250 nl/min, and the Orbitrap Fusion (Thermo Fisher Scientific) was run in a 3 s MS-OT, ddMS2-IT-HCD top speed method. Full MS spectra were collected at a resolution of 120,000, with an AGC target of  $4 \times 10^5$  or maximum injection time of 50ms and a scan range of 400–1500 m/z. Ions were isolated in a 1.6m/z window, with an AGC target of  $1 \times 10^4$  or maximum injection time of 35ms, fragmented with a normalized collision energy of 30 and the resulting MS2 spectra were obtained in the ion trap. Dynamic exclusion was set to 60 s, and ions with a charge state  $<2$ ,  $>7$  or unknown were excluded. MS performance was verified for consistency by running complex cell lysate quality control standards, and chromatography was monitored to check for reproducibility. The mass spectrometry data have been deposited to the ProteomeX-change Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository with the dataset identifier PXD027331.

### Label-free quantitative proteomics analysis

The raw files were analyzed using MaxQuant version 1.5.2.8 (Cox and Mann, 2008) and standard settings. Briefly, label-free quantitation (LFQ) was enabled with a requirement of 2 unique peptides per protein, and iBAQ quantitation was also enabled during the search. Variable modifications were set as Oxidation (M), Acetyl (protein N-term), Gln-  $\rightarrow$  pyro-Glu and Glu-  $\rightarrow$  pyro-Glu. Fixed modifications were set as Carbamidomethyl (C), false discovery rate was set to 1% and “match between runs” was enabled. The resulting protein groups file was processed with an in-house developed tool (PINT) (Wojtowicz et al., 2016), which imputes missing LFQ values with adjusted iBAQ values. Briefly, the distributions of iBAQ intensities for each sample are adjusted to overlap with the LFQ intensity distributions using median-based adjustment, enabling the direct imputation of missing LFQ values with adjusted iBAQ values for those proteins that did not have LFQ values across all the samples. Final list of 6735 proteins identified and quantified in all samples was generated by simultaneously filtering for reverse hits, contaminants and only those proteins observed in 3 biological replicates ( $n = 3$ ) in at least one group. Significance of protein ratios between control and miR-130a OE samples were estimated for each biological repeat and subjected to a statistical analysis in R using *limma\_3.42.2* with Benjamini-Hochberg adjustment. The 6735 detected proteins were ranked from top upregulated to downregulated in miR-130a OE compared to control using the *t* statistics values from the *limma* paired *t* test and this rank file was subsequently used in the pathway analysis (Table S2).

### Pathway analysis and the enrichment map visualization

GSEA\_4.10 was performed using the ranked list from the proteomics data and Baderlab gene set file containing pathways from multiple databases (June 1, 2020, <http://baderlab.org/GeneSets/>) using as parameters of 2000 permutations and gene set size between 10 and 500. EnrichmentMap 3.3.2/Cytoscape 3.8.2 was used to visualize 145 and 128 gene sets that were significantly enriched at *p* value 0.05 in genes upregulated and downregulated in miR-130a OE, respectively. To identify direct miR-130a targets, a proteomics analysis was combined with computationally predicted and experimentally validated miR-130a targets from the mirDIP v1.0 (microRNA data integration portal) and DIANA-TarBase v7.0 databases, respectively. To estimate the enrichment of miR-130a targets in downregulated proteins, a Wilcoxon rank sum test was applied between the ranked list of DE proteins between miR-130a OE and control and Tarbase predicted targets. The specificity of the enrichment of miR-130a targets in downregulated proteins was confirmed by comparing it to the ones obtained for the other 6 other members of the miR-130 family, 7 randomly chosen miRNAs, and 2000 random gene lists by plotting the observed and random *p* values on a histogram in R. GSEA was also used to confirm this enrichment by using a gene set of all Tarbase miR130a targets present in more than 2 independent studies. Significant overlap at a *p* value of 0.05 between the downregulated DE proteins between miR-130a OE and Control and Tarbase predicted targets was visualized in Cytoscape/EnrichmentMap using the post-analysis function. Furthermore, 9 miR-130a targets overlapping between both mirDIP and Tarbase databases with downregulated proteins at *p* value  $< 0.05$  were further visualized using the heatmap.2 function in R and Cytoscape/EnrichmentMap. GSEA was also used to confirm the enrichment of the 9 targets in downregulated proteins.

### Chimeric AGO2 eCLIP

Chimeric eCLIP for miRNA-mRNA chimeras (Manakov et al., 2022) was performed to identify global miRNA-target interactions in CD34<sup>+</sup> CB (replicates of  $\sim 5 \times 10^6$ ) and Kasumi-1 ( $\sim 2.5 \times 10^6$ ) cells. Briefly, the chimeric eCLIP method utilizes the enhanced crosslinking immunoprecipitation protocol (Van Nostrand et al., 2016; Manakov et al., 2022) with the following changes: after initial washes of AGO2 immunoprecipitates, samples were treated with T4 PNK 3' phosphatase minus (NEB) to phosphorylate RNA fragment 5' ends, followed by T4 RNA ligation without adapter to encourage chimeric ligation products. After standard eCLIP adapter ligation and gel electrophoresis, reverse transcription was performed under modified  $Mn^{2+}$  buffer conditions (Van Nostrand et al., 2016) to generate cDNA libraries, and PCR amplified as per the eCLIP protocol. To generate miR130a-targeted AGO2 eCLIP libraries, the content of miR130a chimeras in 0.25  $\mu$ l of cDNA library was quantified by qPCR (NEB LUNA Universal qPCR 2x Master Mix) using primer q7c and primer mir130a-enrich (Table S6). The remainder of the libraries were amplified (NEB Q5 High-Fidelity 2x Master Mix) with primer Q23C and primer miR-130a-enrich, using the following cycle parameters: 98°C, 30s; (98°C, 15s; 65°C, 30s; 72°C, 40s)  $\times$  6 cycles; (98°C, 15s; 72°C, 45s)  $\times$  7 cycles; 72°C, 1 min. PCR products were purified (AMPure XP beads, Beckman Coulter; 1.8  $\times$  bead to sample ratio), and amplified in a second PCR reaction using standard Illumina D5X and D7X index primers to generate sequencing libraries, using the following cycle parameters: 98°C, 30s; (98°C, 15s; 72°C, 1 min)  $\times$  12 cycles; 72°C, 1 min. PCR products were purified (AMPure XP beads; 1.5  $\times$  bead to sample ratio). All libraries were sequenced in PE100 mode on the Illumina NovaSeq-6000 platform.

### Chimeric AGO2 eCLIP data analysis

Processing and bioinformatics analysis was based on previously described methods (Moore et al., 2015), with the following modifications: AGO2 peaks were assigned using the eCLIP processing pipeline available at (<https://github.com/yeolab/eclip>). Briefly, unique molecular identifiers (UMIs) were assigned to each read using `umi_tools` (1.0.0) and trimmed of adapters with `cutadapt` (1.14.0). Reads were then mapped to repeat elements (RepBase release 18.05) using STAR (2.7.6a). Reads that did not align to repeat elements were then aligned to hg19 with STAR and UMI-collapsed with `umi_tools`. These PCR-deduped reads were then used to identify local enrichments with CLIPper (3.0) and normalized above SMIInput using custom scripts `overlap_peakfi_with_bam.pl` and `compress_l2foldenpeakfi_for_replicate_overlapping_bedformat.pl`.

To identify chimeras within total eCLIP datasets, UMI-tagged reads were adapter trimmed, sorted and collapsed, leaving only unique sequences. These sequences were then indexed with Bowtie2 (2.2.6) and used to reverse-map miRNA sequences (miRBase release 22.1). Each miR-mapped read was filtered to select one miR per unique sequence, prioritizing positive stranded reads and minimizing mismatches or indels. To identify chimeras, miR-mapped reads were then uncollapsed and those which contained an mRNA portion of at least 18nt were mapped to hg19 with STAR (2.7.6a). Chimeric reads were then PCR-deduplicated and overlapped with identified AGO2 peaks.

Targeted miR-130a-specific eCLIP reads were UMI-tagged and adapter trimmed in similar fashion as what was done with total datasets. Chimeras from these targeted libraries were identified as hg19-mapped reads that contain expected primer sequences. These reads were then PCR-deduplicated and used to identify enriched regions with Clipper. For both total and targeted chimeric datasets, peaks were annotated (Gencode v19) to identify the fractions of bound genic regions (ie. CDS, UTR, intron). De novo motif analysis using HOMER (4.9.1-6) was also performed to ensure that the most enriched motif corresponds to the expected miRNA seed. Sequencing and genome mapping statistics are listed in Table S3.

### Pathway analysis of miR-130a-target chimeras in CB and Kasumi-1 cells

Pathway enrichment analysis was performed with the web based tool g:Profiler (<https://biit.cs.ut.ee/gprofiler/gost>) using the Baderlab gene set file (June 1, 2020, <http://baderlab.org/GeneSets/>) and visualized using EnrichmentMap 3.3.2/Cytoscape 3.8.2 and Auto-Annotate 1.3.3. Significant overlap at a p value of 0.05 between the downregulated DE proteins between miR-130a OE and Control and the list of miR-130a-target chimeras was calculated using a Wilcoxon rank sum test and visualized in Cytoscape/EnrichmentMap using the post-analysis function. The same analysis was applied to DE transcripts between miR-130a KD and Control and the list of miR-130a-target chimeras from Kasumi-1 cells.

### Cell cycle analysis

Cell cycle analysis was performed with the Click-iT EdU AF647 Flow Cytometry Assay kit according to manufacturer's instructions. Lin<sup>-</sup>CB was used to sort LT-HSC, ST-HSC and CD34 + CD38 + progenitors using the following antibodies (BD): CD34 APC-Cy7 (1:200), CD38 PE-Cy7 (1:100), CD90 APC (1:50), CD45RA FITC (1:50) and CD49f PE-Cy5 (1:50) and PI. Sorted subpopulations (approximately  $7 \times 10^3$ - $1 \times 10^4$ ) were cultured in low cytokine conditions overnight and transduced with miR-130a OE or Control viruses 24 h later. EdU was added into the medium at 10 uM concentration and cells were pulsed for 1.5-2 h. Subsequently, cells were fixed and permeabilized as described in the protocol. Cells were stained with Ki67-FITC antibody (1:30, BD) in PermWash solution overnight at 4°C. The next day, cells were stained with DAPI (1:5000) to label DNA and analyzed by flow cytometry.

### Quantitative RT-PCR for expression levels of miR-130a in CBF AML

Peripheral blood samples from t(8;21) and inv(16) patients (Table S5) were thawed as described and CD34<sup>+</sup> blasts were enriched with Direct CD34<sup>+</sup> Progenitor Isolation Kit (Miltenyi) or sorted by FACS. Following column enrichment, RNA was extracted from approximately  $5 \times 10^5$  CD34<sup>+</sup> or CD34<sup>-</sup> cells with MirVana miRNA Isolation Kit (ThermoFisher) and RNA was reverse transcribed with TaqMan MicroRNA reverse transcription kit using miR-130a-3p (Thermo Fisher Cat#4427975; ID 000454) and RNU48 (ID001006) specific primers purchased from TaqMan Small RNA assay (ThermoFisher). Subsequently, qPCR was performed on the ABI-SDS7900HT instrument using the corresponding Taqman primer/probe sets. Relative quantification was performed using the  $\Delta\Delta C_t$  method normalized to the levels of RNU48 and miR-130a expression in PBMC from 3 healthy volunteers was used as a control. The following AML cell lines were used to measure the expression level of miR-130a: OCI-AML3, OCI-AML2, MOLM13, HL60, NB-4, Kasumi-1, Flag-AE Kasumi-1 and U937. Cells ( $\sim 5 \times 10^5$ ) were collected in triplicates, washed in PBS and centrifuged at 1,450 rpm for 10 min at 4°C. RNA was extracted with RNeasy Plus Mini Kit (Qiagen), RNA integrity was assessed using Agilent Bioanalyzer (RNA Pico chip Agilent Technologies) and 10 ng of RNA was used in the RT reaction. RT-qPCR was performed using the TaqMan Small RNA Assays as described above. Expression levels of miR-130a in AML cell lines were represented relative to the expression levels in Kasumi-1 cells.

### miRNA microarray in CBF AML patient samples

RNA was extracted from peripheral blood cells ( $1 \times 10^6$ - $1 \times 10^7$ ) using the RNeasy Mini kit (Qiagen). Briefly, samples were lysed, and total RNA was collected by column extraction according to the manufacturer's instructions. Samples were then examined by Nanostring, a probe-based assay that detects 827 common miRNAs (<https://www.nanostring.com/products/ncounter-assays-panels/immunology/mirna/>). The output of the assay was analyzed by nSolver 4.0 where the mean of the negative spike-in control was

used as the threshold of microRNA detection. With the expression profile table generated, patients were then stratified by miR-130a expression using a median split and a Kaplan-Meier curve was drawn by GraphPad Prism 7.

### TCGA analysis of CBF AML patients

Publicly available data from the TCGA-LAML database for which miRNA expression data was available was used to study the miR-130a expression pattern. TMM normalized count per million counts (CPM) were clustered using the heatmap.2 function available from the gplots R package (gplots\_3.1.1). The 4 main AML clusters separating the samples in very high, medium, low and very low miR-130a expression groups were retrieved from the heatmap cluster results using the R stats cutree function. A boxplot was constructed using CPM values for the 4 AML groups. One Way Analysis of Variance (ANOVA) was applied to the data to test the significance of the differences in the mean values between the groups. The TCGA-LAML bulk RNAseq dataset was then used to determine the association of miR-130a expression with different AML subtypes. Total of 170 AML patients including cytogenetically normal (129) and complex AML: 9 with inv(16), 7 with t(8;21) and 25 with other cytogenetic abnormalities, were included in the analysis. Expression levels of miR-130a in inv(16) and t(8;21) were compared to the other subtypes using normalized log2 CPM counts. LSC17 score was calculated on RPKM data of each TCGA-LAML patient sample data using regression coefficients of the 17 signature genes from Ng. et al. (Ng et al., 2016). A scatter and a boxplot plot were used to visualize the association of LSC17 score with miR-130a expression and clusters respectively. ANOVA was applied to the data to test the significance of the differences in the mean values between the 4 clusters.

### GSVA and CIBERSORTx analysis and deconvolution of AML signatures from scRNAseq

Deconvoluted AML data was obtained from Zeng et al. (Zeng et al., 2022). Briefly, a signature matrix was generated using scRNA-seq data from van Galen et al. (Van Galen et al., 2019) using expression profiles of seven malignant cell types: Leukemia Stem and Progenitor Cell (LSPC)-Quiescent, LSPC-Primed, LSPC-Cycle, Granulocyte-Monocyte-Progenitor (GMP)-like, ProMono-like, Mono-like, conventional Dendritic Cell (cDC)-like, as well as seven non-malignant cell types: T, Cytotoxic T Lymphocyte (CTL), Natural Killer (NK), B, Plasma, Monocyte, and cDC. LSPC populations were re-annotated in Zeng et al. Deconvolution using this signature matrix performed using CIBERSORTx with S-mode batch correction in absolute mode, and applied on TPM-normalized gene expression data from 173 TCGA AML patients and Kasumi-1 cells. Deconvolution analysis was also performed using RNAseq data from Kasumi-1 transduced with miR-130a KD or lentivirus. Inferred abundances of the seven malignant cell types were normalized to a total of 1, such that the score for a given population represents the proportion of total leukemic cells belonging to that population. For TCGA analysis, principal component analysis was performed based on malignant cell composition and clusters were labeled based on Zeng et al.

Single sample GSEA was run on bulk TCGA-LAML RNAseq RPKM data by using the gsva function available from GSVA\_1.34.0 using tumor-derived HSC/progenitor-like, GMP and myeloid signature gene-sets derived from AML scRNA set (Van Galen et al. 2019). Z scores were calculated by using the formula  $Z = (x - \mu) / \sigma$  (where  $\mu$  is the mean of all x values and  $\sigma$  is the standard deviation of all x values). Each sample was classified as highly enriched in HSC/progenitor-like, GMP or myeloid gene-sets when the gsva score was greater than 1 standard deviation. A boxplot was constructed by plotting the miR-130a expression values for each tumor derived signature classified as highly or lowly enriched samples. A Student's t test was applied to test the significance of the differences in the mean values between each 2 groups.

### Western blot assay

Sorted CD34<sup>+</sup>CD38<sup>-</sup> cord blood cells or Kasumi-1 cells ( $\sim 5 \times 10^4 - 2 \times 10^5$ ) were centrifuged at 1,450 rpm for 10 min at RT and washed with PBS. Washed cell pellets were lysed in RIPA buffer (ThermoFisher) containing protease and phosphatase inhibitors (ThermoFisher). Subsequently, samples were centrifuged at 16,000xg for 15 min at 4°C and supernatants were utilized for Western blot assay. Western blot assay was performed on the automated Simple Western capillary platform (Wes, ProteinSimple) using 12-230 kDa or 66-440 kDa capillary cartridges according to manufacturer's protocol. All antibodies listed in the [Key Resources Table](#) were titrated prior to use on lysates from CD34<sup>+</sup> cord blood cells or Kasumi-1 cells and subsequently used at the listed dilutions.

### Immunoprecipitation of Flag-AML1-ETO

Immunoprecipitation experiments were performed in Flag-AML1-ETO Kasumi-1 cell line which was generously provided by Scott Hiebert's lab. Flag-tagged AML1-ETO Kasumi-1 cells were transduced with miR-130a KD and control lentiviruses as described and sorted for GFP<sup>+</sup> cells 72 h post-transduction. Sorted cells ( $7.5 \times 10^5 - 1.5 \times 10^6$ ) were centrifuged at 1,450 rpm for 10 min at RT and washed with PBS. Cell pellets were frozen at -80°C until IP experiments were performed. IP experiments were performed using EZview Red ANTI-FLAG M2 Affinity Gel (Sigma) according to manufacturer's instructions. Briefly, 20  $\mu$ l of gel suspension was used per reaction. Gel beads were washed two times with 500  $\mu$ l of TBS buffer (50mM Tris HCl, 150 mM NaCl, pH 7.4) by centrifugation at 8,200 x g for 30 s at RT. Cells were lysed in 500  $\mu$ l of lysis buffer (50 mM Tris HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% TRITON X-100 and 1X protease inhibitors). Cell lysates were sonicated with Sonic Dismembrator Model 500 (Fisher Scientific) using 10% amplitude and 3 cycles with 10 s on and 20 s off pulses. Subsequently, samples were centrifuged at 20,000 x g for 20 min and 4°C to pellet the debris and supernatant was collected. A fraction of each lysate (50  $\mu$ l) was saved as an input and the remainder was added to the washed beads. Samples were incubated by gentle rotation overnight at 4°C. Next, beads were centrifuged at 8,200 x g

for 30 s and supernatant was discarded. Beads were washed 3 times with 500  $\mu$ l of TBS and Flag-AML1-ETO was eluted under native conditions by competition with 3XFLAG peptide. For elution, 50  $\mu$ l of TBS containing 150 ng/ $\mu$ l of FLAG-peptide was added to the washed beads. Samples were incubated by gentle rotation for 30 min at 4°C, followed by centrifugation at 8,200 $\times$ g for 30 s. Supernatant containing the eluted Flag-AML1-ETO complex was transferred to fresh tubes, stored at –20°C or used right away in Western blot assay.

### CUT&RUN assay for Flag-AML1-ETO binding occupancy

Flag-AML1-ETO Kasumi-1 cells were sorted 72 h post-transduction and subsequently  $2 \times 10^5$  cells per condition were used in CUT&RUN assay. Cells were washed in PBS and CUT&RUN assay was performed as previously described (Skene and Henikoff, 2017; Skene et al., 2018). Protein A-Micrococcal nuclease (pA-MNase) fusion protein and yeast spike-in DNA were kindly provided by S. Henikoff's lab. Briefly, cells were washed twice with a wash buffer and activated Concanavalin A Beads (Bangs Laboratories) were added dropwise while vortexing the samples. Wash buffer was removed by beads separation on a magnet and antibody buffer containing 0.0125% digitonin and Anti-Flag, anti-ETO or mouse IgG control antibody (Key Resources Table) were added to the beads. Samples were incubated on a rotator overnight at 4°C, followed by incubation with rabbit anti-mouse secondary antibody for 1 h at RT. Addition and activation of pA-MNase and isolation of soluble DNA was performed as previously described (Skene et al., 2018). DNA was extracted with the MinElute PCR Purification kit (Qiagen) and DNA libraries were prepared with NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB) and NEBNext Multiplex Oligos for Illumina (NEB) according to manufacturer's instructions. The number of PCR cycles for library preparation was determined from qPCR based on the Ct value. For each sample, 1  $\mu$ l of DNA, 5  $\mu$ l of Power SYBR Green PCR Master Mix (ThermoFisher), 0.4  $\mu$ l of each forward (10  $\mu$ M) and reverse primer (10  $\mu$ M) and 3.2  $\mu$ l of H<sub>2</sub>O were combined and qPCR was performed according to manufacturer's instruction. AMPure XP beads (Beckman) were used for cleanup of PCR reactions as described in the manufacturer's protocol. Subsequently, libraries were size selected for 150-400 bp range with Pippin HT and size verified with Bioanalyzer. Samples were sequenced on Illumina NextSeq500 using 75bp paired-end reads to achieve sequencing depth of approximately 40M reads/sample.

### CUT&RUN bioinformatic analysis

CUT&RUN paired-end data was trimmed using fastp v. 0.19.5 (Chen et al., 2018) to remove adapters and low quality base pairs (base pair quality score <30 and read length <35 bp). Trimmed reads were aligned to the hg38 human reference genome using bowtie2 v. 2.3.5 (Langmead and Salzberg, 2012) with the same alignment setting as described previously (Skene et al., 2018). Unaligned reads and discordantly aligned reads were eliminated and only primary aligned loci were kept. Duplicated reads were kept for the downstream analysis. The trimmed reads were also aligned to the *S. cerevisiae* yeast genome (sacCer3) yeast genome to evaluate the yeast spike-in reads for the spike-in normalization using the same bowtie2 parameters as described previously (Skene et al., 2018).

Control and miR-130a KD samples with triplicates were pooled and the peaks were called from the pooled bam files using MACS2 v. 2.2.5 (Zhang et al., 2008) with normalizing the samples using the yeast spike-in reads quantification. Calling peaks was performed by comparing the control and miR-130a KD CUT&RUN samples with the IgG and the peaks were filtered with the q-value cutoff <0.001. Peaks from different samples were intersected in a pairwise manner to determine the shared and unique peaks between two samples. Only shared peaks present with ETO and Flag antibodies in control and miR-130a KD samples were used for the downstream analysis. ChipPeakAnnot library (Zhu et al., 2010) in the R package v.3.6.1 was used to determine the overlap between peak sets and determine the shared and unique peaks represented in the Venn diagrams. Minimum peak overlap was set to 100 bp to determine the intersection peaks between the two samples. Peaks were annotated to determine the distribution across the genomic features such as promoters, 5'UTR, 3'UTR, introns, exons and intergenic regions using the Refseq genome annotation. Promoters were defined as the genomic regions starting 2Kb upstream of the transcriptional start site until 500 bp downstream of the transcriptional start site. Genomic distribution of the bound regions was represented in donut plots. p values and odds ratios have been calculated for the counts of the peaks that were annotated for genomic regions for the miR-130a KD specific peaks compared with their counterparts in the shared peaks between control and miR-130a KD using Fisher's exact test.

Motif analysis was performed on called peaks using homer v. 4.8 (Heinz et al., 2010). Normalized enrichment score (NES) for each motif was calculated as the fold change of the target to the background percentage. If the target percentage is less than 5, the pseudo count 1 was added to the target and the background percentages before calculating the fold change to attenuate the fold change of motifs with low target percentages. Volcano plots for the log<sub>10</sub>(p value) vs the NES were plotted for each motif analysis. Only motives with NES>1.5 were displayed. Unique peaks to control and to miR-130a KD were clustered based on their signal intensity scores using K-means clustering algorithm. Peaks unique to Control were clustered into two clusters and peaks unique to miR-130a KD were clustered into three clusters. Motif analysis was performed on the clusters of the peaks using Homer.

Gene ontology analysis was performed on the list of genes that have AML1-ETO peaks bound to their promoters. PANTHER web server v. 15.0 (Thomas et al., 2003) was used for gene ontology analysis. Whole genome coverage tracks in bigwig file format were created from the bam files using the bamCoverage command in the deepTools package v.3.5.0 (Ramírez et al., 2016). The heatmaps of the signal in the Control and miR-130a KD samples for the peaks have been plotted from the coverage tracks using the computeMatrix and the plotHeatmap commands in the deepTools package.

Shared peaks which are common in ETO and Flag in the Control condition (Ctrl\_ETO\_Flag) have been compared with the previously published CUT&RUN data in Stengel et al., (2021). The following samples from the GSE153279 repository were

processed: Kasumi-1\_(-)dTAG\_anti-HA-1 and Kasumi-1\_(-)dTAG\_anti-HA-2. The CUT&RUN samples have been downloaded, trimmed by fastp, and aligned to hg38 genome using bowtie2 aligner. Since there is no IgG sample provided in the repository for these samples, we called peaks using MACS2 without using control/background. We determined the intersection between our Ctrl\_ETO\_Flag peaks and the peaks of these two public data samples by allowing a minimum overlap of 100bp to determine the intersection between the peak sets.

### CUT&RUN and RNA-seq enrichment map

Thirty-one genes that have AML1-ETO peaks bound to their promoters and that are upregulated following miR-130a KD in Kasumi cells at FDR <0.05 have been subjected to pathway enrichment analysis testing against the GO biological process and MSigDB C2 databases using g:Profiler. Significant enrichment results under FDR 0.05 were visualized using Cytoscape/EnrichmentMap. For AML1-ETO regulated genes, publicly available datasets including AML1-ETO OE in human CD34<sup>+</sup> HSPC (Tonks et al., 2007), AML1-ETO KD in Kasumi (Corsello et al., 2009) and dTAG AML1-ETO-FKBP12F36V degradation in Kasumi cells (Stengel et al., 2021). For AML1-ETO OE, genes upregulated and downregulated 3 days following AML1-ETO OE were used in the analysis (E-MEXP-583). For AML1-ETO KD, differentially expressed genes following nucleofection of AML1-ETO-directed siRNAs were used in the analysis. For AML1-ETO degradation, a high confidence AML1-ETO repression signature consisting of 59 genes was used in the analysis (GSE153264).

### QUANTIFICATION AND STATISTICAL ANALYSIS

GraphPad Prism was used for all statistical analysis except RNA-seq, CUT&RUN, mass spectrometry, microarray and chimeric eCLIP-seq datasets. Unless otherwise noted, all individual values and/or mean  $\pm$  SEM values are shown in the graphs. Statistical significance and p values were calculated with Mann-Whitney U-test or unpaired student's t-test. Stars are used to indicate significance in figures (\*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001, \*\*\*\*p < 0.0001). Statistical details of each experiment can be found in the figure legends and STAR Methods sections.