

## **Supplemental Date**

### **Supplemental Methods**

#### **Patient samples**

AML samples for immunoblot analysis were described in previous study by Chou FS et al.<sup>1</sup> t(8;21) samples used in FOXO1 inhibitor experiment were described in previous study by Link KA et al.<sup>2</sup>

#### **Viral production and transduction**

Retrovirus was produced in 293T cells transfected with retroviral vectors, together with envelope RD114 and the gag-pol M57 constructs. CD34<sup>+</sup>CB cells were stimulated IMDM with 20% BIT, SCF, FLT3L, and TPO (100 ng/mL) for 24 h before transduction. Retronectin-coated plates (Takara) were preloaded with retroviral supernatant once by centrifuge at 2200 rpm and 4°C for 30 min. Stimulated cells were then seeded into loaded plates with additional retroviral supernatant on for 1 day. CB cells were co-transduced with tTRi-AE and MSCV-tTA viruses in order to establish AE-tet-off systems. To produce lentivirus, 293T cells were transfected with lentiviral shRNA vectors with lentiviral envelope and gag-pol constructs. For shRNA transduction, cells were mixed with lentiviral supernatant and transduced on retronectin-coated plates.

#### **Flow cytometry and cell sorting**

Antibodies (all BD unless noted) used were APC- and PE-CD33 (WM53), APC-CD11b (ICRF44), V450-human CD45 (HI30), APCCy7-mouse CD45 (30-F11). Cells from mouse tissues were incubated with nonspecific binding blocker (anti-mouse/human CD16/CD32 Fc g receptor; BD) before staining. Cells were analyzed on FACSCanto flow cytometer (BD) or sorted on FACS Aria (BD) or MoFlo XDP (Beckman Coulter), and the data was analyzed with FloJo software (TreeStar).

#### **Immunostaining**

Cytospin slides of AE cells were fixed in 4% paraformaldehyde, permeabilized with PBS containing 0.5% Triton X-100, and blocked with 10% FBS in PBS. Then cells were incubated with primary antibody against FOXO1 (Cell Signaling Technology #2880) at 4°C overnight, and then with goat anti-rabbit IgG (H+L) secondary antibody, Alexa Fluor 568 (Molecular Probes) for 1 h at room temperature. The images were acquired using Leica DMI6000 fluorescent microscope (Leica Microsystems) with ORCAER C4742-95 camera (Hamamatsu) driven by Openlab software. The Wright-Giemsa-staining images were obtained using Motic BA310 microscope and X40/0.65 objective with Moticom 2500 camera (Motic).

#### **Western blotting**

Whole cell lysate were obtained by directly lysing cells in 2x Laemmli Sample Buffer (Bio-Rad, #1610737). Nuclear lysates were obtained using NE-PER nuclear extraction kit (Thermo Scientific). Protein samples were run on 6% or 8% polyacrylamide gels and transferred to a nitrocellulose membrane. The primary antibodies used were anti-FOXO1 (Cell Signaling Technology #2880), anti-MPL (Millipore 06944), anti-UBASH3B (Abcam ab34781), anti-SOX4 (Sigma AV38234), anti-HA (Cell Signaling Technology #3724), anti-HDAC1 (Cell Signaling Technology #5363), anti-ACTIN (Sigma A3854) and anti-TUBULIN (Sigma T0198). Anti-ETO antibody was a gift from Dr. Inge Olsson. The secondary antibodies were HRP-linked goat anti-rabbit IgG and anti-mouse IgG at 1:1000 (Cell Signaling Technology #7074, #7076)

## **Protein immunoprecipitation**

Whole cell lysate of AE cells expressing shRNA were prepared using cell lysis buffer (Cell Signaling Technology; #9803) containing protease inhibitor cocktail (GBiosciences) and 1mM PMSF. Dynabeads Protein G (Life Technologies) were incubated with anti-FOXO1 (Cell Signaling Technology #2880) or anti-HA (Cell Signaling Technology #3724) antibodies in PBST (PBS with 0.1% Tween-20) with rotation for 30 min at room temperature. The antibody-bound Protein G were then incubated with cell lysate with rotation for 6 h at 4°C. The precipitates were separated by magnet, washed 3 times with cell lysis buffer, and then resolved in 2x Laemmli Sample Buffer (Bio-Rad, #1610737).

## **WST-1 assays**

Cells were cultured with titrating doses of FOXO1 inhibitor (0-1.6  $\mu$ M) in triplicate for 6 days, then were incubated with WST-1 cell proliferation assay premix (Takara MK 400) for 3 h. The proliferation index was measured as absorbance at 450 nm. Data were normalized to 0  $\mu$ M group. IC50s were calculated using Prism (GraphPad Software).

## **RT-PCR**

RNA was reversed transcribed using MuLV Reverse Transcriptase and random hexamers (Applied Biosystems). The cDNA was then subject to qPCR using SYBR Green technology (Roche). Expression level was calculated by  $\Delta\Delta$ Ct method, normalized to PGK1. Primers are:

PGK1 F: GGGAAAAGATGCTTCTGGGAA; PGK1 R: TTGGAAAGTGAAGCTCGGAAA;  
FOXO1 F: TCGGCGGGCTGGAAGAATTCAA; FOXO1 R: TTTCCCGCTCTTGCCACCCTC;  
FOXO3 F: TGGGCAAAGCAGACCCTCAA; FOXO3 R: GCGGTGGGATTCACAAAGGTG;  
FOXO4 F: AAGCGACTGACACTTGCCCA; FOXO4 R: ACAGGTTGTGGCGGATCGAG;  
AE F: CACCTACCACAGAGCCATCAAA; AE R: ATCCACAGGTGAGTCTGGCATT.

## **Analysis of RNA sequencing data**

RNAseq data was processed with Tophat/Cufflinks V using the UCSC ref 38 gene/transcript model. FPKM values less than 0.1 were adjusted to 0.1, and normalized per sample to the 60<sup>th</sup> %ile per sample, and then to the median of the 3 MIT empty vector treated samples or to 1, whichever was higher. Differentially expressed genes were identified by filtering for genes that were expressed at a level greater than 2 FPKM in at least two samples and that differed between MIT, AE, FOXO1-WT, and FOXO1-DB with  $p < 0.05$  using students t-test ANOVA. The normalized expression values were then subjected to hierarchical clustering using Pearson Correlation. Supplemental Table 1 lists each gene in the order in which it appears in the heatmap.

## **Analysis FOXO1-activating signature in patient data**

The genes more than 2-fold higher in FOXO1 WT compared to MIG were selected as FOXO1-activating signature (Supplemental Table 2). AML patient dataset were imported into Qluore Omics Explorer 3.1 software (Qluore) for analysis. Expression values were  $\log_2$  transformed. The probesets with a low variation across all samples ( $< 0.2$ ) were filtered out. If there are multiple probesets matched to one gene, these probesets were collapsed, and the average of them were used as expression value of the gene. The  $\log_2$  expression values were then centered to zero mean. Heatmaps were generated by Qluore based on the normalized values of FOXO1-activating signature genes. The average of the mean-centered  $\log_2$

expression values of all FOXO1-activating genes was calculated for each patient sample, which was used to as an indicator of the expression level of the whole FOXO1-activating signature in one patient sample. The bean-plots were generated by BoxPlotR (<http://boxplot.tyerslab.com/>).

### **Pathway enrichment analysis**

The pathway enrichment analysis was performed for FOXO1 WT-upregulated genes using the LRpath methodology with the gene lists from the MSigDB database.<sup>3</sup>

### **Chromatin immunoprecipitation (ChIP)**

Antibodies used were: anti-FOXO1 (Abcam ab39670) for FOXO1 pull-down; anti-HA (Sigma, H6908) and anti-ETO (Santa Cruz SC-9737) for AE pull-down in AE cells and t(8;21) cell lines, respectively. Cells were first washed with PBS and then cross-linked with 0.83 mg/ml Di(N-succinimidyl) glutarate (DSG) (Sigma) for 45 minutes at room temperature with rotation. Cells were washed four times with PBS and further crosslinked in PBS with 1% formaldehyde (~0.34 M) for 10 minutes at room temperature. All crosslinking reactions were quenched by adding 4 volumes of PBS and 0.125 M glycine, cells were washed twice in PBS and incubated with Buffer A (10 mM HEPES pH 8.0, 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100, protease inhibitor cocktail (Roche UK, Burgess Hill, UK) and 0.1 mM PMSF), incubated for 10 min at 4 °C with rotation, and centrifuged 5 min at 500 x g at 4 °C. The pellet was resuspended in 10 ml of ice-cold Buffer B (10 mM HEPES pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.01% Triton X-100, protease inhibitor cocktail and 0.1 mM PMSF), incubated for 10 min at 4 °C with rotation and centrifuged for 5 min at 500 x g at 4 °C. Cells were resuspended in 600 µl of ice-cold ChIP lysis buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.25% SDS, protease inhibitor cocktail and 0.1 mM PMSF), incubated 10 min on ice and sonicated at 5 °C using a Bioruptor™ (Diagenode) to generate fragments an average length of 500 bp (10 min with 30 s “ON” and “OFF” cycles, power setting high). The lysates were centrifuged for 5 min at 16,000 x g at 4 °C and the supernatants were diluted with two volumes of ice-cold ChIP dilution buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 7.5% glycerol, protease inhibitor cocktail and 0.1 mM PMSF). For each IP, 15 µl of Dynabeads® protein G (Dynal) were pre-incubated with 50 µg BSA and 1-2 µg antibody for 2 h at 4 °C with rotation. The blocked antibody-bound protein G mix was added to 20–25 µg chromatin in a total volume of 500 µl diluted ChIP lysis buffer and incubated for 2 h at 4 °C with rotation. After magnetic separation the beads were washed once with 1 ml wash buffer 1 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), twice with 1 ml wash buffer 2 (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), once with 1 ml LiCl buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% NP-40, 0.5% Na-deoxycholate) and twice with 1 ml TE/NaCl buffer (10 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM EDTA). For each wash the beads were mixed with ice-cold washing buffers for 10 min at 4 °C. The immunoprecipitated DNA was eluted two times with 50 µl ChIP elution buffer (100 mM NaHCO<sub>3</sub>, 1% SDS) for 15 min at RT with shaking. At this step the input control (1% of the starting material) was included in the experimental procedure after first adjusting the final volume to 100 µl with ChIP elution buffer. The eluted DNA was incubated overnight at 65 °C in the presence of 50 µg proteinase K. The DNA was finally purified using Agencourt AMPure (Beckman Coulter) magnetic beads according to the manufacturer’s instructions, eluted with 50 µl TE.

### **ChIP-qPCR**

ChIP-qPCR was performed using SYBR Green technology (Applied Biosystems). The chromatin enrichment of each gene locus was calculated by standard curve method and normalized to 1% input. IVL

or RPL30 gene loci were used as a negative control region. RPL30 primers were from Cell Signaling Technology (#7014). Other primers are:

IVL F: GCCGTGCTTTGGAGTTCTTA; IVL R: CCTCTGCTGCTGCCACTT;

FOXO1 F: TCTGGTGAGGAAACCACAGG; FOXO1 R: CGGTTGGGAGGAGTTGACC;

LAT2 F: AAACCCAGAACAACCCAGGC; LAT2 R: ATGAGGAAGGATGTGTGTGCGG;

BAIAP3 F: AGAGCAACCTGTCACCCTTG; BAIAP3 R: ACTCGTAGCTCCACCCACTC;

JUP F: GAAGGACCCCCTAAGCCTAA; JUP R: GCAGCTTTATCTGGGGGTTT;

PTGS2 F: CTGCTGAGGAGTTCCTGGAC; PTGS2 R: GTTCTCTCGGTTAGCGACCA;

SOX4 F: GAGCTGCGGCTCCAAAGT; SOX4 R: CCTGCCAGGATGAGCTTG;

UBASH3B F: ACCACCAGCAAATGACTTCC; UBASH3B R: GCAAGGACTGGCCTAGTGTC;

MPL F: GAACCTCCCTACAACCCGTC; MPL R: TGCAGGAAATCACCCCTTCC.

### **ChIP library preparation**

DNA libraries for sequencing were prepared from approximately 10 ng DNA from ChIP samples using the KAPA Hyper (KR096100) library preparation kit according to the manufacturer's instructions (Kapa Biosystems).

### **ChIP data analysis**

#### **-Alignment**

Sequences reads in fastq format were mapped onto the reference human genome version hg38, Genome Reference Consortium GRCh38. The Illumina reads were aligned to the human genome using Bowtie2.<sup>4</sup> Reads that were uniquely aligned to chromosomal positions were retained and duplicate reads were removed from the aligned data using Picard tools (<http://broadinstitute.github.io/picard/>). The filtered aligned reads were used to generate density profiles using “genomeCoverageBed” function from bedtools (<http://bedtools.readthedocs.org/en/latest/>). These tag densities were displayed using the UCSC Genome Browser.<sup>5</sup>

#### **-Peak calling**

Regions of enrichment (peaks) of ChIP sequencing data were identified using DFilter software<sup>6</sup> with recommended parameters (-bs=100 -ks=50 -refine). Peak overlaps, gene annotations were performed using in-house scripts. Peaks were allocated to genes if located in either their promoters or within the region of 500 bp downstream and 2000 bp upstream of the transcription start sites (TSS), as intragenic if not in the promoter but within the gene body region, or if intergenic, to the nearest gene located within 100 kb. Overlaps between ChIP-seq peaks were defined by requiring the summits of two peaks to lie within +/-200 pb.

#### **-Motif analysis**

De novo motif analysis was performed on peaks using HOMER.<sup>7</sup> Motif lengths of 6, 8, 10, and 12 bp were identified in within  $\pm$  200 bp from the peak summit and a random background sequence option was used. The motif matrices generated by HOMER were scanned against JASPAR with the use of STAMP

to identify similarity to known transcription factor binding sites. The top enriched motifs with a significant p value score were recorded. The annotatePeaks function in HOMER was used to find occurrences of motifs in peaks. In this case we used known motif position weight matrices (PWM) from HOMER database. Percentages of motifs are displayed as a heatmap.

#### **-Correlation of gene expression and ChIP-seq data**

Genes with at least two fold-changes in expression (either up or down) that changed expression in AE versus MIG were selected and correlated with FOXO1 and AE ChIP-Seq co-localized genes. p values were calculated by hypergeometric test using phyper function in R.

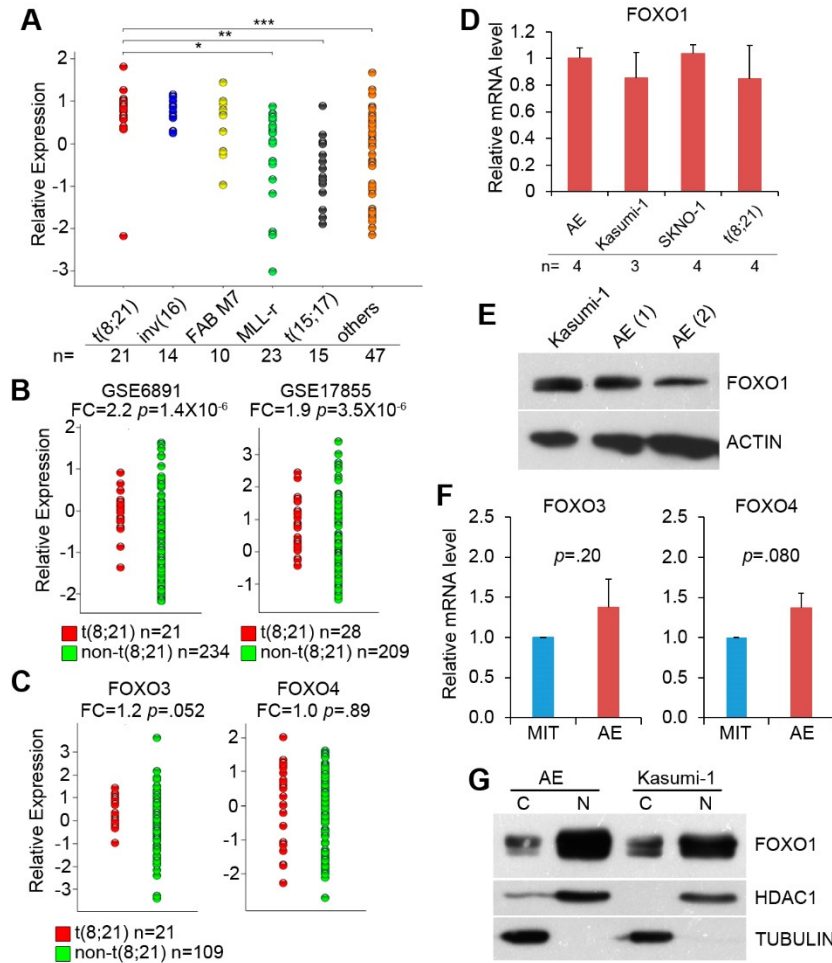
#### **Supplemental Tables**

**Supplemental Table 1. Differential expressed genes of CD34+HSPCs expressing AE or FOXO1 compared to control. Provided as Excel file.**

**Supplemental Table 2. FOXO1-activating gene signature. Provided as Excel file.**

## Supplemental Figures

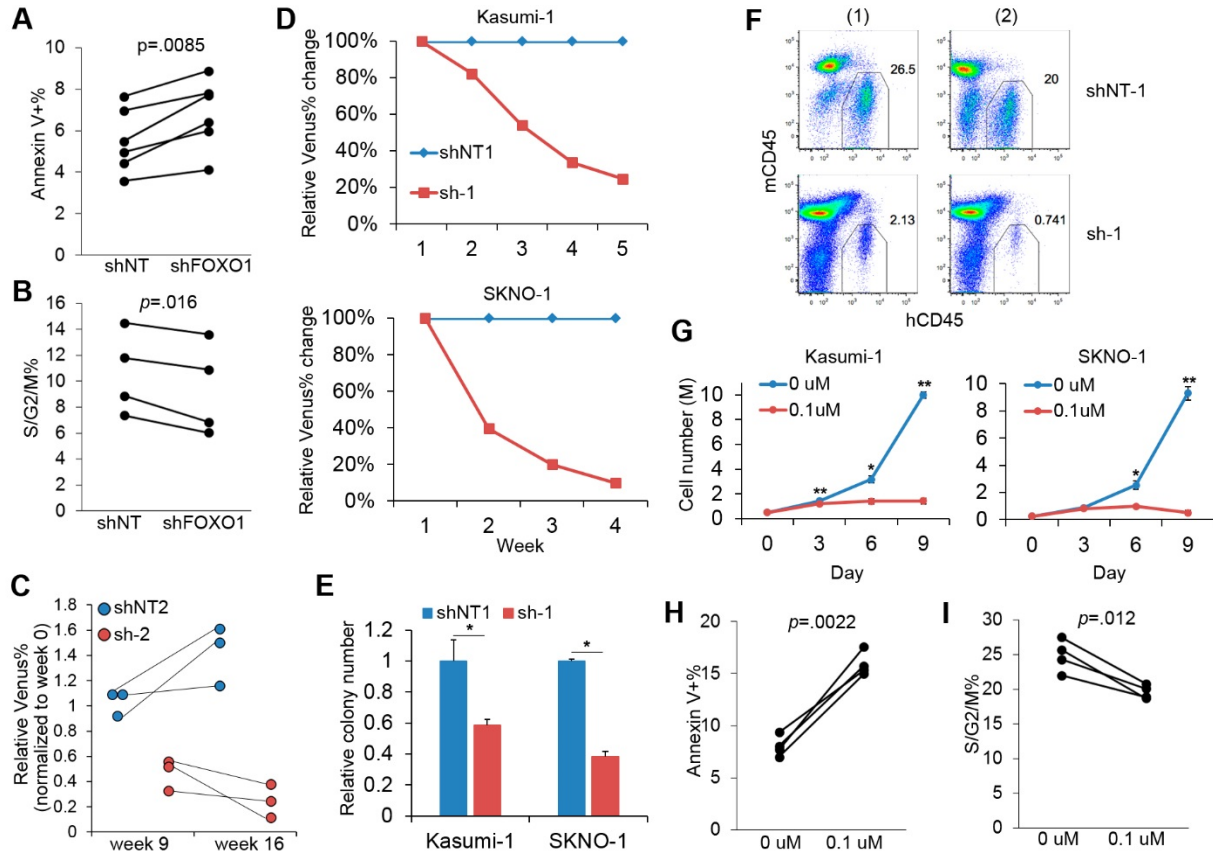
### Supplemental Figure 1



### Supplemental Figure 1. t(8;21) AML is associated with increased FOXO1.

(A) FOXO1 expression in AML patients with different cytogenetics. Dataset from Ross ME et al. \* $p=0.0042$ , \*\* $p=6.6 \times 10^{-6}$ , \*\*\* $p=0.00014$ . (B) FOXO1 expression in t(8;21) AML compared to non-t(8;21) AML, from datasets GSE6891 and GSE17855. (C) FOXO3 and FOXO4 transcript levels in t(8;21) AML compared to non-t(8;21) AML patient samples from Ross ME dataset. (D and E) Expression levels of FOXO1 mRNA (D, determined by qPCR) and protein (E) in AE pre-leukemia cells compared with t(8;21) AML cell lines (Kasumi-1 and SKNO-1) and t(8;21) primary patient samples. (F) qPCR analysis of FOXO3 and FOXO4 mRNA levels in AE- or empty vector (MIT)-transduced CD34<sup>+</sup> HSPCs. (G) Cellular fractionation immunoblot analysis to show the subcellular localization of FOXO1 in AE cells and Kasumi-1 cells. All p values were calculated by unpaired two-tailed t-test.

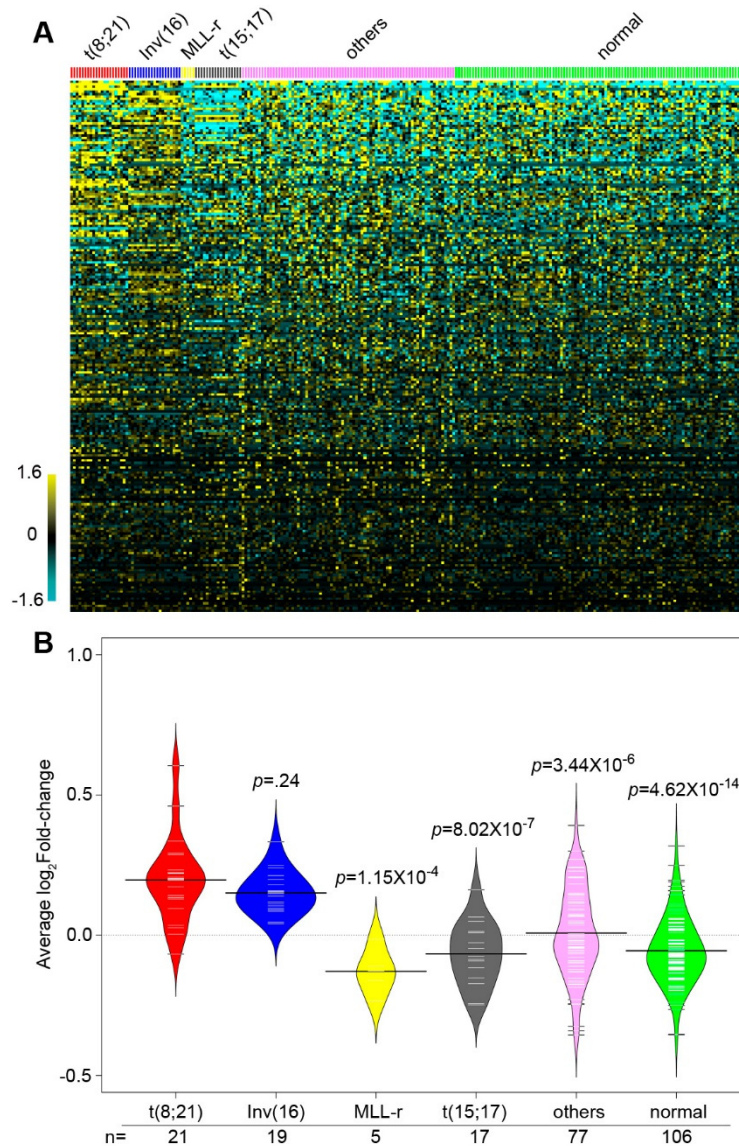
## Supplemental Figure 2



## Supplemental Figure 2. FOXO1 is required for maintenance of AE pre-leukemia and leukemia.

(A and B) Apoptosis (A) and cell-cycle status (B) of AE cells expressing non-targeting shRNA or FOXO1 shRNA at 1-2 weeks post-transduction. Each line represents one independent experiment. (C) shRNA-transduced AE cells were transplanted into immunodeficient mice. Normalized Venus<sup>+</sup> percentage of human engrafted cells in bone marrow was examined at different time points. The week 9 data is from the same experiment shown in Figure 3C. Data points from the same individual mouse are linked by lines. (D) Relative change of Venus<sup>+</sup> percentage of shRNA-transduced Kasumi-1 and SKNO-1 cells in liquid culture. One representative experiment of 2 replicates is shown. (E) CFU assay of sorted shRNA-transduced Kasumi-1 and SKNO-1 cells.  $n=3$ , results represent mean  $\pm$  SD. \* $p<0.01$ , unpaired two-tailed t-test. (F) Representative flow cytometry analysis of bone marrow from immunodeficient mice transplanted with shRNA-transduced AE leukemic clone. Bone marrow was examined when mice in shNT group got sick, leukemia burden was evaluated as percentage of human CD45<sup>+</sup> (hCD45) cells. Two individual mice from each group are shown. mCD45, murine CD45. (G) Cell growth of Kasumi-1 and SKNO-1 cells treated with 0.1  $\mu$ M FOXO1 inhibitor AS1842856 or DMSO.  $n=3$ , results represent mean  $\pm$  SD. \* $p<0.05$ , \*\* $p<0.01$ . (H and I) Apoptosis (H) and cell-cycle status (I) of AE cells treated with FOXO1 inhibitor or DMSO for 6 days. Each line represents one independent experiment. All p values were calculated by paired two-tailed t-test unless noted.

### Supplemental Figure 3

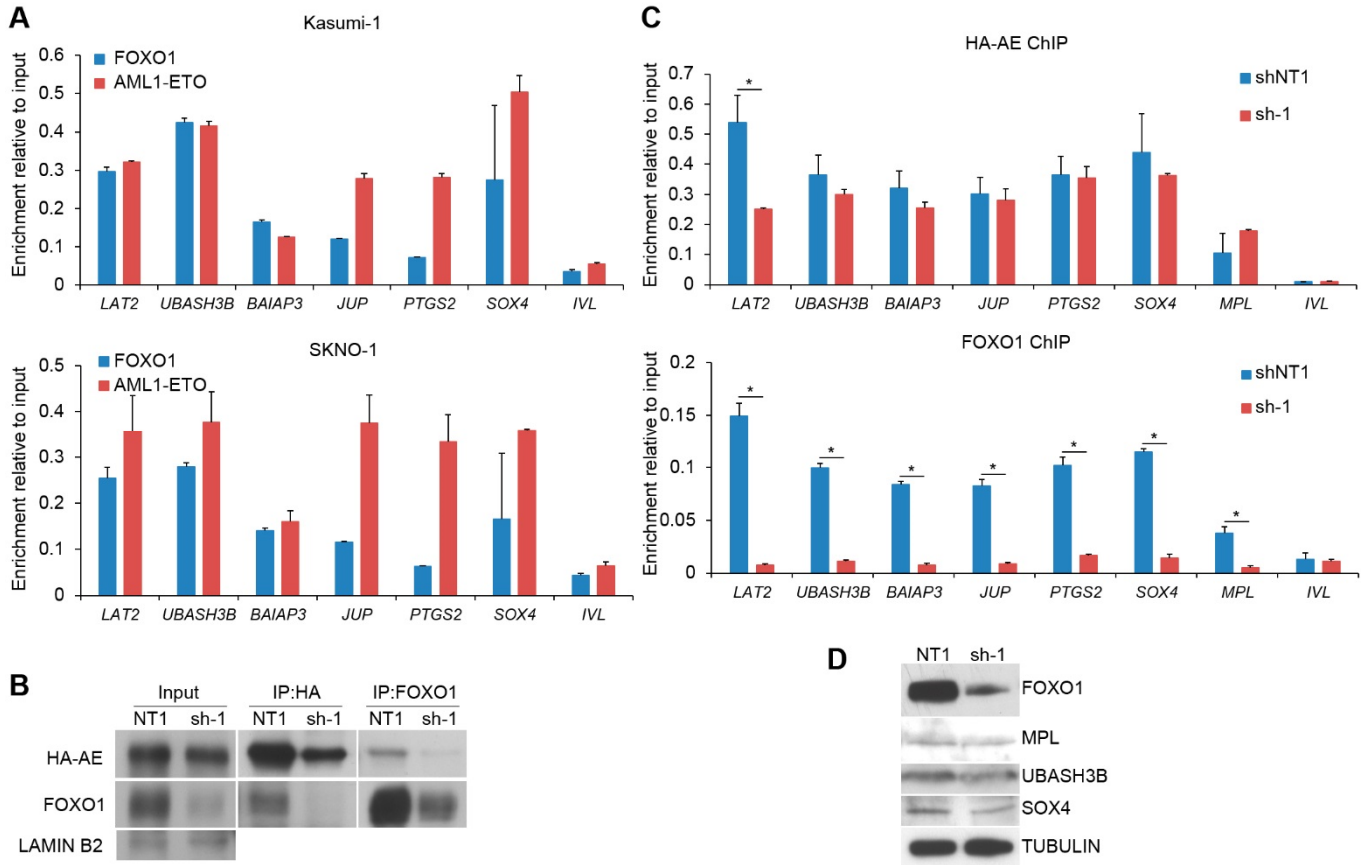


### Supplemental Figure 3. FOXO1-activating gene signature is overrepresented in t(8;21) AML patients.

The dataset was from GSE6891. (A) Heatmap showing the expression level of FOXO1-activating gene signatures in AML patient samples with different cytogenetics. (B) Bean-plot depicting the average of the mean-centered  $\log_2$  expression values of FOXO1-activating genes in different subtypes of AML patient samples. Black lines show the means of subgroups; each white line represents value of individual samples; polygons represent the estimated density of the data. p values were calculated by unpaired two-tailed t-test.



## Supplemental Figure 4



### Supplemental Figure 4. FOXO1 is involved in AE molecular network.

(A) ChIP-qPCR analysis of chromatin occupancy of AE and FOXO1 at selected target gene loci in Kasumi-1 and SKNO-1 cells. IVL locus was used as negative control. The result represents mean  $\pm$  SD,  $n=3$ . The enrichment level was normalized to input. (B) Co-immunoprecipitation assay of AE and FOXO1 in AE cells expressing control- or FOXO1- shRNA. (C) ChIP-qPCR analysis of chromatin occupancy of AE (HA-tagged) and FOXO1 at selected target gene loci in AE cells expressing control- or FOXO1- shRNA. The result represents mean  $\pm$  SD,  $n=3$ . The enrichment level was normalized to input. \* $p<0.05$ , unpaired two-tailed t-test. (D) Immunoblot to determine protein levels of MPL, UBASH3B and SOX4 upon FOXO1 knockdown in Kasumi-1 cells.

## Supplemental References

1. Chou FS, Griesinger A, Wunderlich M, et al. The thrombopoietin/MPL/Bcl-xL pathway is essential for survival and self-renewal in human preleukemia induced by AML1-ETO. *Blood*. 2012;120(4):709-719.
2. Link KA, Lin S, Shrestha M, et al. Supraphysiologic levels of the AML1-ETO isoform AE9a are essential for transformation. *Proc Natl Acad Sci U S A*. 2016;113(32):9075-9080.
3. Sartor MA, Leikauf GD, Medvedovic M. LRpath: A logistic regression approach for identifying enriched biological groups in gene expression data. *Bioinformatics*. 2009;25:211-217.
4. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357-359.
5. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12(6):996-1006.
6. Kumar V, Muratani M, Rayan NA, et al. Uniform, optimal signal processing of mapped deep-sequencing data. *Nat Biotechnol*. 2013;31(7):615-622.
7. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38(4):576-589.