Supplementary Information for:

Integrative epigenomic analysis identifies biomarkers and therapeutic targets in adult B-acute lymphoblastic leukemia

Huimin Geng[1,2,7], Sarah Brennan[1], Thomas A. Milne[4,9], Wei-Yi Chen[5], Yushan Li[1], Christian Hurtz[6,7,15], Soo-Mi Kweon[6], Lynette Zickl[8], Seyedmehdi Shojaee[6,7], Chuanxin Huang[1], Donna Neuberg[8], Debabrata Biswas[5], Yuan Xin[1], Janis Racevskis[3], Rhett P. Ketterling[10], Selina M. Luger[11], Hillard Lazarus[14], Martin S. Tallman[12], Jacob M. Rowe[13], Mark R. Litzow[10], Monica L. Guzman[1], C. David Allis[4], Robert G. Roeder[5], Markus Müschen[6,7], Elisabeth Paietta[3,*], Olivier Elemento[2,*] and Ari Melnick[1,*]

[1]Department of Medicine/Hematology-Oncology Division, [2]Institute for Computational Biomedicine, Weill Medical College of Cornell University, New York, New York
[3]Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, New York
[4]Laboratory of Chromatin Biology and Epigenetics, [5]Laboratory of Biochemistry and Molecular Biology, the Rockefeller University, New York, New York
[6]Children's Hospital Los Angeles, University of Southern California, Los Angeles, California
[7]Department of Laboratory Medicine, University of California San Francisco, San Francisco, California
[8]Dana Farber Cancer Institute, Boston, Massachusetts
[9]MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, Headington, Oxford, United Kingdom
[10]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota
[11]Abramson Cancer Center, University of Pennsylvania, Philadelphia, Pennsylvania
[12]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York
[13]Department of Hematology and BMT, Rambam Medical Center, Haifa, Israel
[14]Department of Medicine, Case Western Reserve University, Cleveland, Ohio
[15]Max-Planck Institute for Immunobiology, Freiburg, Germany

Running Title: Integrative epigenomics of adult B-ALL

***To whom Correspondence should be addressed:**

Ari Melnick, MD
Division of Hematology/Oncology, Department of Medicine
Weill Cornell Medical College
1300 York Ave, New York, NY 10065
email: amm2014@med.cornell.edu

Olivier Elemento, PhD
Institute for Computational Biomedicine
Weill Cornell Medical College
1305 York Ave, New York, NY 10065
email: ole2001@med.cornell.edu

Elisabeth Paietta, PhD
Montefiore Medical Center-North Division
Albert Einstein College of Medicine
600 East 233rd Street, Bronx, NY 10466
e-mail: epaietta@earthlink.net

SUPPLEMENTARY TABLES

Table S1: Patient characteristics at diagnosis

Table S2: Characteristics of the 215 E2993 B-ALL patients and the 12 normal bone marrow samples (**excel file**).

Table S3: MassArray primers used for technical validation of HELP assays.

Table S4: Genes included in *BCR-ABL1*, *E2A-PBX1* or MLLr DNA methylation signatures vs. normal pre-B cells (**excel file**).

Table S5: Lymphoid gene sets and gene ontology terms enriched in B-ALL DNA methylation signatures vs. normal pre-B cells.

Table S6: Detailed genes in the lymphoid gene sets and gene ontology terms (**excel file**).

Table S7: DNA methylation and gene expression signature of *BCR-ABL1*-positive B-ALL (**excel file**).

Table S8: MassArray primers for DNA methylation signature genes in *BCR-ABL1* and MLLr B-ALL.

Table S9: QPCR primers for gene expression signature genes in *BCR-ABL1* and MLLr B-ALL and in E2A-PBX1 and MLL-AF4 knockdown experiments.

Table S10: DNA methylation and gene expression signature of *E2A-PBX1*-positive B-ALL (**excel file**).

Table S11: Antibodies used for QChIP and ChIP-seq.

Table S12: Primers used for QChIP.

Table S13: DNA methylation and gene expression signature of MLLr ALL (**excel file**).

Table S14: Summary of number of reads and peaks detected in ChIP-seq experiments.

Table S15: Genomic regions enriched in MLL^N, $AF4^C$, H3K79me2 and the common of the three ChIP-seq in RS4;11 cells (**excel file**).

SUPPLEMENTARY FIGURES

Figure S1: Technical validation of HELP arrays by MassArray EpiTyper

Figure S2: QPCR and MassArray EpiTyper validation of core signature genes in *BCR-ABL1*-positive B-ALL

Figure S3: *IL2RA*(CD25) expression is associated with *BCR-ABL1*-positive B-ALL

Figure S4: *IL2RA*(CD25) is the top differentially expressed and methylated genes between CD25-positive and CD25-negative, *BCR-ABL1*-positive B-ALL

Figure S5: Whole-genome-wide binding patterns of E2A, PBX1 and p300.

Figure S6: QPCR and MassArray EpiTyper validation of core signature genes in MLLr ALL

Figure S7: Primer locations for MLL^N, $AF4^C$, H3K79me2 and Histone 3 QChIP experiments in RS4;11 and CCRF-CEM cells

Figure S8: QChIP validation for the MLL/AF4 fusion binding in SEM cells

Figure S9: MLL-fusion target genes are more hypomethylated and expressed than non-target genes in MLLr B-ALL

Figure S10: *BCL6* mRNA expression in MLLr B-ALL cell lines and primary samples

Figure S11: Colony formation assay in two non-MLLr B-ALL cell lines, REH and Nalm6

Figure S12: *BCL6* mRNA expression in MLLr B-ALL primary samples

Figure S13: Effect of RI-BPI on viability of MLLr and non-MLLr B-ALL specimens

Figure S14: Scenario - DNA hypomethylation signatures in MLL/AF4 ALL are linked with MLL fusion oncoprotein binding and H3K79me2, which induce overexpression of these genes.

SUPPLEMENTARY METHODS

Patient treatment

Patients received 2 phases of standard induction therapy and, if in remission, were assigned to myeloablative allogeneic stem cell transplantation (alloHSCT) if they had a compatible donor, matched sibling or unrelated donor, and/or were BCR-ABL1 positive. Most of BCR-ABL1-positive patients without a suitable donor were treated with chemotherapy alone rather than being subjected to randomization. Other patients were randomized to chemotherapy for 2.5 years versus an autologous transplant. For the most recent patients accrued to the trial, a protocol modification introduced imatinib (600mg daily) into intensification after the second block of induction. From late 2005, a final cohort within the trial was given imatinib earlier, in conjunction with the 2nd phase of induction. Both cohorts resumed the drug for a further 2 years following alloHSCT, if tolerated. If alloHSCT was not possible, imatinib was permitted for 2 years, with maintenance. Among our study population, 26 patients received imatinib as part of their treatment regimen. The subset of imatinib treated patients were excluded in our survival analyses. Full details of the protocol were described previously (1).

Total RNA extraction

Total RNA was extracted from 10^7 cells using the RNeasy mini kit from Qiagen (Valencia, CA), and eluted in RNAse-free water.

Genomic DNA extraction

High molecular weight DNA was extracted from 5–10×10^6 cells using the Qiagen Puregene Gentra cell kit (Qiagen, Valencia, CA). DNA was diluted in 10mM Tris-HCl pH 8.0 and the quality was assessed in 1% agarose gel.

Array-based methylation analysis using HELP

One microgram of high molecular weight DNA was digested overnight with isoschizomer enzymes HpaII and MspI respectively (NEB, Ipswich, MA). DNA fragments were purified using phenol/chloroform, resuspended in 10mM Tris-HCl pH 8.0, and used immediately to set up the ligation reaction with MspI/HpaII-compatible adapters and T4 DNA ligase. Ligation-mediated PCR was

performed with enrichment for the 200 to 2000 base pair (bp) products and then submitted for labeling and hybridization onto a human HG_17 promoter custom-designed oligonucleotide array covering 25,626 HpaII amplifiable fragments within the promoters of the genes.

Single locus quantitative DNA methylation assays

MassArray EpiTyper assays (Sequenom, CA) were performed on bisulfite-converted DNA, as previously described(2). For HELP array technical validations, MassArray primers were designed to cover the flanking HpaII sites of selected HpaII Amplifiable Fragments (HAF), as well as any other HpaII sites found up to 2,000 bp upstream of the downstream site and up to 2,000 bp downstream of the upstream site, in order to cover all possible alternative sites of digestion. For the biological validation of the signature genes, MassArray primers were designed to cover CpG dense areas of interest associated with the respective HAFs. The primers were designed using Sequenom EpiDesigner beta software (Sequenom, Inc) and data were analyzed using EpiTYPER software version 1.0 (Sequenom, Inc). Note that MassArray and QPCR validation studies were performed in different sets of randomly selected specimens with available DNA and RNA, from among the cohort of 215 profiled E2993 patients.

Gene expression array data analysis

Gene expression microarrays were performed using NimbleGen HG18 60mer expression 385K platform. Among the 215 B-ALL patient samples, 191 had enough cDNA for gene expression profiling. mRNA isolation, labeling, hybridization, and quality control was performed strictly according to the manufacturers protocol (Roche NimbleGen, Madison, WI). Raw pair files from the scanner were processed with the RMA algorithm in NimbleScan 2.5 software (Roche NimbleGen). Differentially expressed probesets between two groups were determined by Student's t-test and multiple testing correction with the BH method.

Quantitative real time PCR (QPCR)

After total RNA extraction, cDNA synthesis was done using the Superscript III First Strand Kit from Invitrogen (Carlsbad, CA). QPCR reactions was performed using TaqMan™ or SyberGreen™ on a 7900 ABI RT-PCR machine (Applied Biosystems, CA).

MLL-AF4 and E2A-PBX1 knockdown

Briefly, the MLL-AF4 siRNAs used in this study were siMARS (sense, 5'-ACUUUAAGCAGACCUACUCCA-3'; antisense, 5'-UGGAGUAGGUCUGCUUAAAGUCC-3'), targeting the MLL-AF4 fusion site specific in RS4;11 cells, and siMA6 (sense, 5'-AAGAAAAGCAGACCUACUCCA-3'; antisense, 5'-UGGAGUAGGUCUGCUUUUCUUUU-3'), targeting the MLL-AF4 fusion site specific in SEM cells. The mismatch control was siMM (sense, 5'-AAAAGCUGACCUUCUCCAAUG-3'; antisense, 5'-CAUUGGAGAAGGUCAGCUUUUCU-3').

Genomic DNA-fragment library

Genomic DNA-fragment libraries were prepared using the Illumina ChIP-seq Library preparation Kit following the manufacturer's instructions (Illumina, CA). Briefly 10ng of purified ChIP DNA was end repaired by conversion of overhangs into phosphorylated blunt ends with the use of T4 DNA polymerase and E. coli DNA polymerase I Klenow fragment. Illumina single-end adapters were ligated to the ends of the DNA fragments. Ligation products were purified on a 2% agarose gel with a size selection of 200-300bp. Fifteen PCR cycles were performed with Illumina genomic DNA primers that anneal to the ends of the adapters. The purified PCR-amplified fragment libraries were quantified with the use of the PicoGreen dsDNA Quantitation Assay with the Qubit Fluorometer (Invitrogen, CA). The size range of libraries was validated on the Agilent Technologies 2100 Bioanalyzer with the High Sensitivity DNA Kit (Agilent, CA).

ChIP sequencing (ChIPseq)

After library preparation, the protocols for the Illumina Single-Read Cluster Generation Kit were used for cluster generation on the cBOT (Illumina). The targeted samples were diluted to 10 nmol and denatured with sodium hydroxide. Seven pico moles of each target-enriched sample and Phix control were loaded into separate lanes of the same flow cell, hybridized, and isothermally amplified. After

linearization, blocking, and primer hybridization, sequencing was performed for 36 cycles on an Illumina GAIIx or HiSeq2000. Raw image data were converted into base calls via the Illumina pipeline CASAVA version 1.7 with default parameters. Rigorous quality control was performed with the use of data from reports generated by the Illumina pipeline. All 36-bp-long reads were mapped to the reference human genome sequence, hg18, using Illumina's ELAND aligner with the default parameters. Only reads mapping uniquely to the genome with not more than 2 mismatches were retained for further analysis. Clonal reads (i.e., reads mapping at the same genomic position and on the same strand) were collapsed into a single read. Peaks from ChIP-seq data were called using the ChIPseeqer program (3) with parameters indicated in **Table S14** and annotated to genes (-2kb to TSS and +1kb to TES) and/or promoters (+/- 2kb to TSS) based on hg18 refseq genes downloaded from the UCSC Genome Browser.

Pathway analysis

The gene sets for canonical pathways and Gene Ontology (GO) terms were downloaded from the Molecular Signatures Database (MSigDB) at http://www.broad.mit.edu/gsea/msigdb/ using C5 collection. The B-cell and lymphoid specific signatures were curated by Staudt laboratory and downloaded from http://lymphochip.nih.gov/signaturedb/. This database contains ~250 gene sets/signatures associated with normal lymphoid biology and lymphoid neoplasms(4). Fisher's Exact test was used to calculate enrichment p values for each of those gene sets and the BH method(5) was used for False Discovery Rate (FDR) control. The Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems, Inc. Redwood City, CA) was also used to identify deregulated gene networks.

SUPPLEMENTARY TABLES

Table S1: Patient characteristics at diagnosis

Characteristic	Value
Male/Female	120/95 (55.6%, 44.4%)
Age, year (1 unknown)	
<20	10 (4.7%)
20-29	44 (20.6%)
30-39	54 (25.2%)
40-49	51 (23.8%)
50-59	46 (21.5%)
>=60	9 (4.2%)
Median (range)	39 (17-63)
WBC X 109/L (8 unknown)	
<30	102 (49.3%)
>=30	105 (50.7%)
Median (range)	30.7 (0.8-438)
Immunophenotype	
Pro-B-ALL	34 (15.8%)
Pre-B-ALL	25 (11.6%)
Early-Pre-B-ALL	146 (67.9%)
Transitional-Pre-B-ALL	5 (2.3%)
Mature-B-ALL	5 (2.3%)
Cytogenetics	
BCR-ABL1	83 (38.6%)
E2A-PBX1	7 (3.3%)
MLLr	28 (13.0%)
Negative of the above	97 (45.1%)
BCR-ABL1 breakpoint (n=83)	
e1a2	58 (69.9%)
b2a2/b3a2	25 (30.1%)

Table S2: Characteristics of the 215 E2993 B-ALL patients (A) and the 12 normal bone marrow samples (B) (**excel file**).

Table S3: MassArray primers used for technical validation of HELP assays. Forward primers were indicated by "_F" and reverse primers by "_R".

Target Promoter	sequence (5' → 3')
B2M_MA1_F	aggaagagagGTGGGGTTAGTTTAGGGTTGGATTT
B2M_MA1_R	cagtaatacgactcactatagggagaaggctACCCCACATAAACCCAAAAATACTAAAA
B2M_MA2_F	aggaagagag TTTTGTTTGGTTGTTTTTAAGATGTATTGT
B2M_MA2_R	cagtaatacgactcactatagggagaaggctAAAACCTAAACTCAAAAAATCCTCCC
B2M_MA3_F	aggaagagagTTAGTATTTTTGGGGTTAGTTTGTAAAG
B2M_MA3_R	cagtaatacgactcactatagggagaaggctACCACCAAAAAAAACTTAAAAAAAA
CARD6-MA1_F	aggaagagagGGAGGTGATTTTGGGTTTGAGTTAT
CARD6-MA1_R	cagtaatacgactcactatagggagaaggctTTTTCTTCAAAAAATCTATAACATTCTCCA
CARD6-MA2_F	aggaagagagGTTTTTTAGTGAAATAAGAGGTGGTTGG
CARD6-MA2_R	cagtaatacgactcactatagggagaaggctAAAACCTCAAAACTTCATCCCTCTC
CEP70-MA1_F	aggaagagagGGGTTAGGTTAGGTTGGGATTAGTTA
CEP70-MA1_R	cagtaatacgactcactatagggagaaggctTCAACAAATATTTATTAAAATCCTCCAAAT
CEP70-MA2_F	aggaagagagTATTGGTTTTAAGAGTTTTTTATGGGTTTT
CEP70-MA2_R	cagtaatacgactcactatagggagaaggctTCTAACTCAACATCTAATTCAATCCCTAAT
CFHL5_MA1_F	aggaagagagTTTTTTGGTTGGGAGTGAATATATTAAGAT
CFHL5_MA1_R	cagtaatacgactcactatagggagaaggctCCTATAATCCCAACACTTTAAAAAACC
CFHL5_MA2_F	aggaagagagTTGGGATTATAGGTTTGAGTTATTG
CFHL5_MA2_R	cagtaatacgactcactatagggagaaggctAACAATCAAAAATAAACTATTTAAAAAAAA
CFHL5_MA3_F	aggaagagagTTTTTGAAATAGTTGAATTTTTGGTTAGAA
CFHL5_MA3_R	cagtaatacgactcactatagggagaaggctTCCCCCAACAATAAATACCCATAAA
CSF2_MA1_F	aggaagagagTAGATTGTTTAGGGAGGGTTGGAGAG
CSF2_MA1_R	cagtaatacgactcactatagggagaaggctCAAACCACAATACCCAAAAACAACA
CSF2_MA2_F	aggaagagagGGGTTGGGGGTAGTAGTAAAAAGGA
CSF2_MA2_R	cagtaatacgactcactatagggagaaggctACCCCCTAAAATCCTAAACAACCAC
CSF2_MA3_F	aggaagagagGGTTTTTGATTTTGGTTGTTATTGGTAG
CSF2_MA3_R	cagtaatacgactcactatagggagaaggctACTATAACCCACCCCTTAATCCCTC
CSF2_MA4_F	aggaagagagTGTTGTTTTTGGGTATTGTGGTTTG
CSF2_MA4_R	cagtaatacgactcactatagggagaaggctCAATTACTACCCCCTCCTCCTACAAC
HEL308-MA1_F	aggaagagagGTTATAATGGAGTATAGGATATGGAGATTT
HEL308-MA1_R	cagtaatacgactcactatagggagaaggctCTTCTACATCCACCTTAAAAAAAA
HEL308-MA2_F	aggaagagagTTTTGAGAGTTAATTTGGGGTTTTTTTT
HEL308-MA2_R	cagtaatacgactcactatagggagaaggctAATCTCCCACCCCTCTATCAATAAAC
HEL308-MA3_F	aggaagagagGGAGGTAGAGATGGGAGGATTATTTAAGT
HEL308-MA3_R	cagtaatacgactcactatagggagaaggctCACCCAAACTAAAACACAATAACTCCA
IL10-MA1_F	aggaagagagTTTTTAGATATTTGAAGAAGTTTTGATGTT
IL10-MA1_R	cagtaatacgactcactatagggagaaggctAAATAATACTCACCATAACCCCTACC
IL10-MA2_F	aggaagagagTTTGGGTATTTATTTTAGGTTGGGG
IL10-MA2_R	cagtaatacgactcactatagggagaaggctAACTCAAAAAAACCTCTTCATTCATTAAAA
IL3-MA1_F	aggaagagagTGAGTATGTAAGGAATAAATTTTGGGTT
IL3-MA1_R	cagtaatacgactcactatagggagaaggctAAACACAACTCAACACACAAAAAAAA
IL3-MA2_F	aggaagagagGTTTGGGGTTGAGTTAGGAGGGTATAT
IL3-MA2_R	cagtaatacgactcactatagggagaaggctATCCCAAAAAACAACCCCTAAAAAA
IL3-MA3_F	aggaagagagAAGATTGAGTTGTTAGTGGAGGTTATTATG
IL3-MA3_R	cagtaatacgactcactatagggagaaggctCTCAACCCCAAACTAAACCCTAAAC

IL3-MA4_F	aggaagagagGTTTGTGGTTTTTTATGGAGGTTTTATGT
IL3-MA4_R	cagtaatacgactcactatagggagaaggctAACAACCACCACCCCTCACTACTAC
IL3-MA5_F	aggaagagagTTAGAGGGATTGTTAGGGGTTGTAGT
IL3-MA5_R	cagtaatacgactcactatagggagaaggctACCCAATAAAATCTCATCTACCATCAC
P53AIP1_MA1_F	aggaagagagGGGATGTTGAGGTGGGTAGATTATTT
P53AIP1_MA1_R	cagtaatacgactcactatagggagaaggctAAACTTCCAACCTTCCCCTAACAAT
P53AIP1_MA2_F	aggaagagagGTGATTGTTTTTTGGTTGTTGGAG
P53AIP1_MA2_R	cagtaatacgactcactatagggagaaggctCAAATTCCTCAAAAACTACTTCTATTTTAA
P53AIP1_MA3_F	aggaagagagTGTTTGAAAGTTGGTAATATGTAAAAAGGT
P53AIP1_MA3_R	cagtaatacgactcactatagggagaaggctAAACAAAACCCCCTAAAATCCTAAC
P53AIP1_MA4_F	aggaagagag AGGTGAGAAAGATTAAAATTAGTTTTGTGT
P53AIP1_MA4_R	cagtaatacgactcactatagggagaaggctTCTCCTCACCTACCTCCTAAAAATC
P53AIP1_MA5_F	aggaagagagGATTGTGGTTGGTAAGTAGGGGAAG
P53AIP1_MA5_R	cagtaatacgactcactatagggagaaggctCCCCTAAACCCTATACCCATTTCTC
RIT1_MA1_F	aggaagagagGGTTAGGAGTTTGAGATTAGTTTGGTTAAT
RIT1_MA1_R	cagtaatacgactcactatagggagaaggctCTACTAAACTCTTACAACCCAAACATACCA
RIT1_MA2_F	aggaagagagTTGTTTTTATTGGTTAGTGGGGATTG
RIT1_MA2_R	cagtaatacgactcactatagggagaaggctAAAATACCCCTATCCTCCCTTCCTC
RIT1_MA3_F	aggaagagagAGAGGAAGGGAGGATAGGGGTATTT
RIT1_MA3_R	cagtaatacgactcactatagggagaaggctACCCAAAAAAACATTTAAAAAAAACAATAA
RIT1_MA4_F	aggaagagagTTGTTTTTTTTAAATGTTTTTTTGGGT
RIT1_MA4_R	cagtaatacgactcactatagggagaaggctAATAATTCTCCTACCTCAACCTCCC
S100A9_MA1_F	aggaagagagTTATTTATATTGGTTTTTAGGGGTGA
S100A9_MA1_R	cagtaatacgactcactatagggagaaggctAAACCCTCTTAAAATAACCCTACCTA
S100A9_MA2_F	aggaagagagAGGAATTAGAATAGTTGGGTTTTTTTT
S100A9_MA2_R	cagtaatacgactcactatagggagaaggctAAACACTTAACCCTTTAACCCTATC
S100A9_MA3_F	aggaagagagGGGGGTTTTTTTAATTTTAAGGGTTT
S100A9_MA3_R	cagtaatacgactcactatagggagaaggctCCACCCAACATCCTACCTACCTAAA
ST18-MA1_F	aggaagagagGTTTGTATGTTTGTAGTGGGGAGGG
ST18-MA1_R	cagtaatacgactcactatagggagaaggctAAAAAAAACAAAAATTCCCAAAACCA
ST18-MA2_F	aggaagagagAGGGGAGAGGGTTGTTTTTTGATAG
ST18-MA2_R	cagtaatacgactcactatagggagaaggctCCTCTAATATTCCCCTAAAATCACATTCA
ST18-MA3_F	aggaagagagGTATTGGGTTGAATTTTGAGAGAAAATATA
ST18-MA3_R	cagtaatacgactcactatagggagaaggctTCCTAAATAAAAAATAAATCAATCCCC
TNFRSF9_MA1_F	aggaagagagTATAATTGAGGTAGGAGAATGGGGAATAT
TNFRSF9_MA1_R	cagtaatacgactcactatagggagaaggctTATAATTGAGGTAGGAGAATGGGGAATAT
TNFRSF9_MA2_F	aggaagagagTTGTTGTGTTTTTTTTGTTTTTATTTTTTT
TNFRSF9_MA2_R	cagtaatacgactcactatagggagaaggctATCCCCTAACTCCATAATCACTATAATACA
TNFRSF9_MA3_F	aggaagagagTAGGGTTGTTATAGAGTTGTGGTGGG
TNFRSF9_MA3_R	cagtaatacgactcactatagggagaaggctAAAAACAAAAATTCCAATCTCACCC
TNFRSF9_MA4_F	aggaagagagATTTTTGTATTTTGGTAGAATATGATGGAA
TNFRSF9_MA4_R	cagtaatacgactcactatagggagaaggctTAAATTTTCATTTTCCTTTCCTTAAAAAAA
TREML2_MA1_F	aggaagagagTGGGTTTTTATATATTGAAGGAGTAGATGA
TREML2_MA1_R	cagtaatacgactcactatagggagaaggctACATACACATACACTCACACTCACAATCTC
TREML2_MA2_F	aggaagagagGGGGTTTTAGATTGGGTGGAGTAGT
TREML2_MA2_R	cagtaatacgactcactatagggagaaggctCCACAACAACAACAACAACAAAAAAA
TREML2_MA3_F	aggaagagagTTAGGAGTTTGGGAGAGAGAGAGAGG
TREML2_MA3_R	cagtaatacgactcactatagggagaaggctCAATATAAAACCCACTACCCCCAAC
TREML2_MA4_F	aggaagagagATTATGAGGAAAGGGGTTGTGTTTG

TREML2_MA4_R	cagtaatacgactcactatagggagaaggctAAATCTCCCCTTCAAAAAACCTCAC
TSPAN-MA1_F	aggaagagagGTTTTTGATTTAGGATTTTAAAGAAA
TSPAN-MA1_R	cagtaatacgactcactatagggagaaggctATTTTAAAACAAAATCTCCCTCTATC
TSPAN-MA2_F	aggaagagagTAAGTGAAAAGGTGGAGGTTTTTTTTATTT
TSPAN-MA2_R	cagtaatacgactcactatagggagaaggctCCCCTTCCCCCTATTCTACAATAAC
TSPAN-MA3_F	aggaagagagGGTTTTTAGAGTTGGGGAGATGGTT
TSPAN-MA3_R	cagtaatacgactcactatagggagaaggctAAAAAATAAACCCTCAAAATCATCCTATCA
TTRAP-MA1_F	aggaagagagGTTTAGAAGTTAGTGAAAGAAGGAGGATAT
TTRAP-MA1_R	cagtaatacgactcactatagggagaaggctAAACAAACTCCACACACAAAAATC
TTRAP-MA2_F	aggaagagagGATTTTTGTGTGTGGAGTTTGTTT
TTRAP-MA2_R	cagtaatacgactcactatagggagaaggctCCCTCATCACTTACTAAATCTTAAACTCAA
TTRAP-MA3_F	aggaagagagTTAAGATTTAGTAAGTGATGAGGGGAG
TTRAP-MA3_R	cagtaatacgactcactatagggagaaggctAACCAAATAATAAAAAACCTTATAAAACAA
TTRAP-MA4_F	aggaagagagTTTTTTGTAATAGGTTTTTAGAAAAGATG
TTRAP-MA4_R	cagtaatacgactcactatagggagaaggctTTCACTCCAAAATTAAAATTCCTAC

Table S4: Genes included in *BCR-ABL1*, *E2A-PBX1* or MLLr DNA methylation signatures vs. normal pre-B cells (**excel file**)

Table S5: Lymphoid gene sets (A) and gene ontology terms (B) enriched in B-ALL DNA methylation signatures vs. normal pre-B cells, and detailed information of the enriched lymphoid gene sets (C). P values were calculated by Fisher's Exact test and multi-testing adjusted by Benjamini and Hochberg (BH) method. Pathways with BH<0.05 in at least one subtype were shown.

(A)

Lymphoid gene sets	**BH *BCR-ABL1***	**BH *E2A-PBX1***	**BH MLLr**
Myc_ChIP_PET_2plus	0.0005	8.2e-12	8.7e-08
BCL6_target_ChIPseq	0.002	0.09	0.049
Dendritic_cell_CD123pos_blood	0.005	0.11	0.018
HRAS_overexpression_2x_up	0.007	0.29	0.18
p53_up_Xray	0.012	0.50	0.77
CD8_T_effectorDn_memoryIm_NaiveUp	0.013	0.09	1.00
IL10_OCILy3_Up	0.013	1.00	0.31
Cell_cycle_Liu	0.022	0.65	0.18
Lymph_node_DLBCL	0.030	0.35	0.51
Blimp_proliferation_repressed	0.031	0.81	0.92
Resting_monocyte_GNF	0.042	1.00	0.09
Notch_T-ALL_down_Palomero	0.043	0.42	1.00
Tcell_cytokine_induced_IL2_IL7_IL15only	0.049	0.42	0.31
SREBP1a&2_up_Scap_dep	0.53	0.016	0.39
CNS_PNS_Node1663	0.57	0.016	0.30
Glutamine_Glucose_starve_both_down	0.26	0.021	0.08
Tcell_cytokine_induced_IL4only	0.44	0.042	0.51

(B)

Gene Ontology (GO) terms	BH *BCR-ABL1*	BH *E2A-PBX1*	BH MLLr
Signal_Transduction	0.026	0.044	6E-06
Cell_Cell_Signaling	0.001	0.018	0.007
Negative_Regulation_of_Biological_Process	0.011	0.019	0.013
Cell_Proliferation	0.010	0.031	0.001
Negative_Regulation_of_Cellular_Process	0.014	0.051	0.013
Regulation_of_Cell_Proliferation	0.010	0.103	0.001
Positive_Regulation_of_Biological_Process	0.033	0.138	0.0002
Cell_Surface_Receptor_Linked_Signal_Transduction	0.027	0.378	0.007
Protein_Metabolic_Process	0.035	0.0003	0.106
Cellular_Protein_Metabolic_Process	0.050	0.001	0.126
Cellular_Macromolecule_Metabolic_Process	0.026	0.001	0.140
Transmission_of_Nerve_Impulse	0.021	0.033	0.089
Response_to_Ionizing_Radiation	0.024	0.040	0.192
Negative_Regulation_of_Cell_Proliferation	0.014	0.063	0.159
Regulation_of_Multicellular_Organismal_Process	0.012	0.087	0.185
Regulation_of_Developmental_Process	0.042	0.131	0.114
Synaptic_Transmission	0.013	0.131	0.123
Positive_Regulation_of_Response_To_Stimulus	0.050	0.139	0.534
Programmed_Cell_Death	0.026	0.219	0.114
Apoptosis	0.041	0.219	0.114
Negative_Regulation_of_Cytokine_Biosynthetic_Process	0.045	0.318	0.288
Regulation_of_Response_to_Stimulus	0.050	0.318	0.828
Actin_Polymerization_and_or_Depolymerization	0.044	0.710	0.126
Cytoskeleton_Organization_and_Biogenesis	0.026	0.819	0.553
System_Process	0.257	0.005	0.013
Response_to_Extracellular_Stimulus	0.104	0.033	0.013
Protein_Modification_Process	0.385	0.001	0.420
Biopolymer_Modification	0.480	0.003	0.492
Multicellular_Organismal_Development	0.104	0.003	0.114
System_Development	0.252	0.006	0.289
Anatomical_Structure_Development	0.543	0.008	0.126
Post_Translational_Protein_Modification	0.951	0.019	0.872
Biopolymer_Metabolic_Process	0.151	0.021	0.160
Nucleotide_Biosynthetic_Process	0.351	0.023	0.509
Regulation_of_Response_to_External_Stimulus	0.064	0.031	0.341
Purine_Nucleotide_Metabolic_Process	0.227	0.031	0.341
Nervous_System_Development	0.285	0.033	0.793
Biosynthetic_Process	0.104	0.035	0.114
Positive_Regulation_of_Protein_Modification_Process	0.835	0.040	0.145
Positive_Regulation_of_Phosphate_Metabolic_Process	0.851	0.045	0.151
Calcium_Ion_Transport	1.000	0.045	0.312
Positive_Regulation_of_Cellular_Process	0.104	0.251	0.0003
Cellular_Response_to_Extracellular_Stimulus	0.683	0.318	0.020

Intracellular_Signaling_Cascade	0.683	0.319	0.019
Positive_Regulation_of_Cell_Proliferation	0.162	0.718	0.020
Multi_Organism_Process	0.277	0.904	0.047
Positive_Regulation_of_Signal_Transduction	1.000	0.925	0.047
Receptor_Mediated_Endocytosis	0.935	1.000	0.040

(C)

Signature names	**Signature identifiers** (Shaffer et al, Immunol Rev. 2006)	**Detailed explaination** (Shaffer et al, Immunol Rev. 2006)	**Reference for the signatures**	**Selection criteria** (Shaffer et al, Immunol Rev. 2006)
MYC target genes	MYC_ChIP_PET_2plus	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=248	Zeller KI et. al. Global mapping of c-Myc binding sites and target gene networks in human B cells. Proc Natl Acad Sci U S A. 2006 Nov 21;103(47):17834-9.	Table 8. MYC targets in human B cells
BCL6 target genes	BCL6_target_ChIPseq		Katerina Chatzi et al, unpublished data (Melnick lab)	BCL6 targets in DLBCL cells identified by ChIPseq
CD123+ Dendritic cell overexpression genes	Dendritic_cell_CD123pos_blood	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=175	Lindstedt M et. al. Gene family clustering identifies functionally associated subsets of human in vivo blood and tonsillar dendritic cells. Lindstedt et al. J Immunol 175:4839 (2005)	Genes 3x higher in blood CD123+ DCs vs. blood BDCA1+ or BDCA3+ or CD16+ DCs (p<0.01) with signal > 7 in blood CD123+ DCs.
HRAS overexpression genes	HRAS_overexpression_2x_up	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=169	Bild AH et. al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Bild et al. Nature 439:353 (2006)	Supplemental Table 1; Genes upregulated by H-ras by 2 fold
p53 upregulated genes	p53_up_Xray	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=207	Rosenwald A et. al. Fludarabine treatment of patients with chronic lymphocytic leukemia induces a p53-dependent gene expression response. Rosenwald et al. Blood. 104:1428 (2004)	p53 wt cells:at least 2 time points induced by > 1log2; p53 mutant cells: no time point induced >0.5log2
Memory T cell signature genes	CD8_T_effectorDn_memoryIm_NaiveUp	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=151	Holmes S et. al. Memory T cells have gene expression patterns intermediate between naive and effector. Proc Natl Acad Sci U S A. 2005 Apr 12;102(15):5519-23.	Supplemental Table 2, EFFE (effector) down, MEM (memory) intermediary, NAI (naive) up
IL-10 upregulated genes in ABC DLBCL	IL10_OCILy3_Up	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=253	Lam LT et. al. Cooperative signaling through the signal transducer and activator of transcription 3 and nuclear factor-{kappa}B pathways in subtypes of diffuse large B-cell lymphoma. Blood. 2008 Apr 1;111(7):3701-13.	Fig. 2B, IL-10 upregulated genes in ABC DLBCL cell line OCILy3
Cell cycle genes	Cell_cycle_Liu	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=157	Liu D et. al. A random-periods model for expression of cell-cycle genes. Proc Natl Acad Sci U S A. 2004 May 11;101(19):7240-5.	Supporting Table 5, genes periodically expressed in the human cell cycle using a cancer cell line (HeLa)
Gene expression signatures in DLBCL	Lymph_node_DLBCL	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=8	Rosenwald A et. al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med. 2002 Jun 20;346(25):1937-47.	Supplementary appendix 1, Gene-Expression Signatures Defined by Hierarchical Clustering in DLBCL
Blimp-1 repressed genes	Blimp_proliferation_repressed	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=88	Shaffer AL et. al. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. Immunity. 2002 Jul;17(1):51-62.	Blimp-1 repressed genes in human mature B cell
Genes upregulated in resting monocyte	Resting_monocyte_GNF	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=41	Su AI et. al. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A. 2004 Apr 20;101(16):6062-7.	3x higher in CD14 monocytes v all other heme subsets (CD19, CD8, CD4, CD56, CD14, BDCA4, CD71, CD105)
Notch1 downregulated genes	Notch_T-ALL_down_Palomero	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=236	Palomero T et. al. NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. Proc Natl Acad Sci U S A. 2006 Nov 28;103(48):18261-6.	Supplemental Figure 10; genes upregulated by gamma secretase inhibitor and direct target of Notch1 ICN1 by CHIP-CHIP
IL2, IL7 and IL15 induced genes	Tcell_cytokine_induced_IL2_IL7_IL15only	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=72	target genes. Identification of dual-specificity phosphatase 5 (DUSP5) as a regulator of mitogen-activated protein kinase activity in interleukin-2 signaling. J Biol Chem. 2003 Feb 14;278(7):5205-13.	Fig. 1C, IL2, IL7 and IL15 induced genes
SREBP1a and SREBP2 upregulated genes	SREBP1a&2_up_Scap_dep	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=185	Horton JD et. al. Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. Horton et al. PNAS 100:12027 (2003)	Table 1; Genes up in SREBP1a and SREBP2 Tg mice and down in Scap -/- mice.
Co-expressed genes in human and mouse	CNS_PNS_Node1663	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=96	Su AI et. al. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A. 2004 Apr 20;101(16):6062-7.	Cluster with correlation coefficient of 0.75 on GNF U133A data
Glutamine Glucose starve downregulated genes	Glutamine_Glucose_starve_both_down	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=137	Peng T et. al. The immunosuppressant rapamycin mimics a starvation-like signal distinct from amino acid and glucose deprivation. Mol Cell Biol. 2002 Aug;22(15):5575-84.	List from Subramanian et al. PNAS 102:15545 (2005)
IL4 induced genes	Tcell_cytokine_induced_IL4only	http://lymphochip.nih.gov/cgi-bin/signaturedb/signatureDB_DisplayGenes.cgi?signatureID=73	Kovanen PE et. al. Analysis of gamma c-family cytokine target genes. Identification of dual-specificity phosphatase 5 (DUSP5) as a regulator of mitogen-activated protein kinase activity in interleukin-2 signaling. J Biol Chem. 2003	Fig. 1D, genes preferentially induced by IL-4

Table S6: Detailed genes enriched in the lymphoid gene sets and gene ontology terms shown in Table S5 (**excel file**)

Table S7: DNA methylation (A) and gene expression (B) signature of *BCR-ABL1*-positive B-ALL (**excel file**)

Table S8: MassArray primers used to validate differential DNA methylation of genes in *BCR-ABL1*-positive and MLLr B-ALL. Forward primers were indicated by "_F" and reverse primers by "_R".

Target Promoter	sequence (5' → 3')
ACSL1_MA1_F	aggaagagagGGGATGGTTTTGTAGGGTTATTAG
ACSL1_MA1_R	cagtaatacgactcactatagggagaaggctAACAACTCCTCTCCCTAAAAAACC
ACSL1_MA2_F	aggaagagagGGTTTTTTAGGGAGAGGAGTTGTT
ACSL1_MA2_R	aggaagagagGTGGTTGAGGGAATTTTTTTTAGA
ACSL1_MA3_F	cagtaatacgactcactatagggagaaggctCCCACATACCAAACCCTTTAATTC
ACSL1_MA3_R	aggaagagagTTAGTTTTAAAGTTAGGGTGGTGGG
BCL6_MA1_F	aggaagagagTTGTGTTTATTGTTAGTTGGGTTTTG
BCL6_MA1_R	cagtaatacgactcactatagggagaaggctAAATTTTAAACAAAACCTCTCACCC
CD200_MA1_F	aggaagagagTTTGTAATTTTAGTATTTTGGGAGG
CD200_MA1_R	cagtaatacgactcactatagggagaaggctTAAAAAATTTCCTTCCACTTACATT
CD200_MA2_F	aggaagagagTATTTTTGAAAAGGGAAAAATGTTTGA
CD200_MA2_R	cagtaatacgactcactatagggagaaggctAAACCACCCTTTAAAAAACAAAATC
CD38_MA1_F	aggaagagagTGGGTGATTATTATGTGTTAGGTATTG
CD38_MA1_R	cagtaatacgactcactatagggagaaggctAAAAAAACAACCACCTTAATCCTCT
CD38_MA2_F	aggaagagagGTGGGTGGGTTATTGTATAGGAGTT
CD38_MA2_R	cagtaatacgactcactatagggagaaggctCCAAAACAAAATTAAACTCTCCTAA
CD38_MA3_F	aggaagagagTTTTAGGAGAGTTTAATTTTGTTTTGG
CD38_MA3_R	cagtaatacgactcactatagggagaaggctAACTAAACTCAAAACACCTTACCCA
CD38_MA4_F	aggaagagagGTTTTTTGGGTAAGGTGTTTTGAGT
CD38_MA4_R	cagtaatacgactcactatagggagaaggctTAATTTAACCCTTCTACACCCTCCT
CNTNAP1_MA1_F	aggaagagagTGATTTAGTTTTTAGAGAAGATGGAGAA
CNTNAP1_MA1_R	cagtaatacgactcactatagggagaaggctCCTAACCTCAACCAACAAAAATAAA
CNTNAP1_MA2_F	aggaagagagGGGGTGTTAGTGGTTAGAGTTTTAGT
CNTNAP1_MA2_R	cagtaatacgactcactatagggagaaggctAAAAACTACCTCTATACCTTACCACATT
FLT3_MA1_F	cagtaatacgactcactatagggagaaggctTAAAAAATATTTCCCCACTCTCCAC
FLT3_MA1_R	cagtaatacgactcactatagggagaaggctTCCTACTTACTCACCTACCTACCCC
FLT3_MA2_F	cagtaatacgactcactatagggagaaggctTCAAACTCTACAAACACCAAACAAA
FLT3_MA2_R	aggaagagagTTTGTTTGGTGTTTGTAGAGTTTGA
FLT3_MA3_F	cagtaatacgactcactatagggagaaggctCTCTATACCCCAACTTCCTCTTCTA
FLT3_MA3_R	aggaagagagAGGGTTATGTTTGGGTTGTAGGTAG
FLT3_MA4_F	cagtaatacgactcactatagggagaaggctTCCAACAACCCTTAAAAACAATAAA
FLT3_MA4_R	aggaagagagTTTATTGTTTTTAAGGGTTGTTGGA
FUT4_MA1_F	cagtaatacgactcactatagggagaaggctAAAAAAAATTTACACACACACACCA
FUT4_MA1_R	aggaagagagTATTTTTTTTGTGGTGTGTGTGTGT
FUT4_MA2_F	cagtaatacgactcactatagggagaaggctAAAAAAAACTCAAATAATATCCCAA
FUT4_MA2_R	aggaagagagTTTAATAAATTGGTGAAGGGTTTTT
GAB1_MA1_F	aggaagagagAGGAGGGGATAGTTGAAGTTAGAAG
GAB1_MA1_R	cagtaatacgactcactatagggagaaggctCAAACAAACAATAACTAACAAAATTAACC
GAB1_MA2_F	aggaagagagGGAAAGAAATGATATTTTTGTGTTGAT
GAB1_MA2_R	cagtaatacgactcactatagggagaaggctCCCTACCTTTCTCTTTAAAACAACTAA

GAB1_MA3_F	aggaagagagAGATTTGTTTGATTGATTGGATTGT
GAB1_MA3_R	cagtaatacgactcactatagggagaaggctCAAATTCAAACATTTCTCCTACCTC
GIMAP5_MA1_F	aggaagagagTAGAAGTAGAGTTTGTGAGAAGGGG
GIMAP5_MA1_R	cagtaatacgactcactatagggagaaggctTTAAAAACAAAACCCTAAAAACCAA
GIMAP5_MA2_F	aggaagagagAATGGGATGGGGTATATTGTTTTAG
GIMAP5_MA2_R	cagtaatacgactcactatagggagaaggctACCCTATCCTCCAAAATAACAATCA
GIMAP5_MA3_F	aggaagagagTGATTGTTATTTTGGAGGATAGGGT
GIMAP5_MA3_R	cagtaatacgactcactatagggagaaggctCCTCTAAAAACTTCAATTCTCCCAT
GIMAP5_MA4_F	aggaagagagGGGAGTAAGGTTGAATTGATTTTTT
GIMAP5_MA4_R	cagtaatacgactcactatagggagaaggctCTACCCCAACCCATATATTTCCAA
HLA-DQA1_MA1_F	aggaagagagGGGATTGTTAGGGAGGGAAATTA
HLA-DQA1_MA1_R	cagtaatacgactcactatagggagaaggctCTAAAACTACAAACACCTACCACCA
HLA-DQA1_MA2_F	aggaagagagTGGTGGTAGGTGTTTGTAGTTTTAG
HLA-DQA1_MA2_R	cagtaatacgactcactatagggagaaggctAACCCTACCAAATCAAATTTTAACC
HLA-DQA1_MA3_F	aggaagagagTAAATAAAGTTTTGTTGTTGGGGG
HLA-DQA1_MA3_R	cagtaatacgactcactatagggagaaggctTAAAATAACCCCAAAATCTACCACC
IGF2BP2_MA1_F	cagtaatacgactcactatagggagaaggctTAAAAACTATTCCAACACCAACTCC
IGF2BP2_MA1_R	aggaagagagGGAGTTGGTGTTGGAATAGTTTTTA
IGF2BP2_MA2_F	cagtaatacgactcactatagggagaaggctTAAAATTAAAATTTCCCCCAAATTA
IGF2BP2_MA2_R	aggaagagagTTTTAATTTGATTTTAATTTGGGGG
IL2RA_MA1_F	aggaagagagTGGAATTATTTATTTGGGGATTTTT
IL2RA_MA1_R	cagtaatacgactcactatagggagaaggctCCAACCCAATACTTAAAAAAAACAA
IL2RA_MA2_F	aggaagagagGGGTTGTTATTTTTGTGGGTTTATT
IL2RA_MA2_R	cagtaatacgactcactatagggagaaggctAAAACTTTACTCCTTCATCCCAAAT
ITGAE_MA1_F	aggaagagagTATTGGGTTGGTAGGTAAAGGAAAT
ITGAE_MA1_R	aggaagagagGTTGTAGGTGGTTTTTTTAAGGATG
ITGAE_MA2_F	cagtaatacgactcactatagggagaaggctTAATTTATAACTCCCTACAATCCCC
ITGAE_MA2_R	cagtaatacgactcactatagggagaaggctAAAACTAACATCCAAAACTTTCCAA
ITGAE_MA3_F	aggaagagagTTGGAAAGTTTTGGATGTTAGTTTT
ITGAE_MA3_R	cagtaatacgactcactatagggagaaggctATACAAACATCTCTTCAAAACCCTA
ITGAE_MA4_F	aggaagagagGAGTAGAGTTTTGGAGATTGGTTGT
ITGAE_MA4_R	cagtaatacgactcactatagggagaaggctAAACAAAATCTCACTCTATCACCCA
ITGAE_MA5_F	aggaagagagTTGTTTTAAAGGAAATTGATTGGTT
ITGAE_MA5_R	cagtaatacgactcactatagggagaaggctCAATTCTCTACCTCAACCTCCTAAA
ITGAE_MA6_F	aggaagagagTTTAGGAGGTTGAGGTAGAGAATTG
ITGAE_MA6_R	cagtaatacgactcactatagggagaaggctAAACAACCTTCTAATACTTCCCACC
ITGAE_MA7_F	aggaagagagTGTGTGTGAATATTTTTTGTTTTTG
ITGAE_MA7_R	cagtaatacgactcactatagggagaaggctACTAATAACCCTACAAAACCATCCC
LTB_MA1_F	aggaagagagTTTAGTTTTTGTTGGGTTTGTGTTT
LTB_MA1_R	cagtaatacgactcactatagggagaaggctAATCTTTTACCCCCTCTAACTCAAC
LTB_MA2_F	aggaagagagGTTGAGTTAGAGGGGGTAAAAGATT
LTB_MA2_R	cagtaatacgactcactatagggagaaggctCCTTTAATTCCACCCTAAAAACCTA
LTB_MA3_F	aggaagagagGATTTGTATATTTGGTTGGGATTTT
LTB_MA3_R	cagtaatacgactcactatagggagaaggctCCTTTCTCACTCAACACAAAAAACT
LTB_MA4_F	aggaagagagGGATGGAAATGGAGTTTTATTTTTT
LTB_MA4_R	cagtaatacgactcactatagggagaaggctTAACTCACTACAACCTCTACCTCCC
MAP1A_MA1_F	cagtaatacgactcactatagggagaaggctAAACCAAAATCTAAAAAAAATTCCC
MAP1A_MA1_R	cagtaatacgactcactatagggagaaggctAAAAATTAAAACTCCTTACCCAAAA
MAP1A_MA2_F	cagtaatacgactcactatagggagaaggctAAAATAAAAAAACCCACTACAAACC
MAP1A_MA2_R	aggaagagagGGAATGTGTTTAATAGGAGGGTTTT
MAP1A_MA3_F	aggaagagagGGGGTTTTTGAATAGAGGAGTTAGT

MAP1A_MA3_R	aggaagagagTTTGGTTGGAGTTGAGGGTATAGTA
MAP7_MA1_F	cagtaatacgactcactatagggagaaggctTCAAAATCCTATTCAATCAATTTCC
MAP7_MA1_R	aggaagagagGGAAATTGATTGAATAGGATTTTGA
MAP7_MA2_F	cagtaatacgactcactatagggagaaggctTAACTCTCCCCAATACCCTTAAAAC
MAP7_MA2_R	aggaagagagAAGGGTATTGGGGAGAGTTATAGTG
MRPL33_MA1_F	aggaagagagTTTTTTTGTTAATGTAAGGTAAGGAAGG
MRPL33_MA1_R	cagtaatacgactcactatagggagaaggctCCTATAATCCCAACCCTTTAAAAAA
MRPL33_MA2_F	aggaagagagGTTTTAGTTTTTTAAAGGGTTGGGA
MRPL33_MA2_R	cagtaatacgactcactatagggagaaggctCAAAAACTCATACCTTTCCCAATAA
MRPL33_MA3_F	aggaagagagGGTTTTAATGTTTATTGGGAAAGGT
MRPL33_MA3_R	cagtaatacgactcactatagggagaaggctCCCTAACTCCTATTTCTAAATCAAA
NOV_MA1_F	aggaagagagTAAGAGTGGGGTTTAGGAATTTGTAT
NOV_MA1_R	cagtaatacgactcactatagggagaaggctAATCACCTCAAACCCCTTCTAAATAA
NOV_MA2_F	aggaagagagTATTTAGAAGGGGTTTGAGGTGATT
NOV_MA2_R	cagtaatacgactcactatagggagaaggctATCAAAACCTCCAACCAAAAAAA
NOV_MA3_F	aggaagagagTGTTTTAGGAGATGGAGAAGTAGGA
NOV_MA3_R	cagtaatacgactcactatagggagaaggctCCACCCTCTAAAAAAAACCAATC
NOV_MA4_F	aggaagagagGTTTTTTGGGAAGGGGAATATAAAA
NOV_MA4_R	cagtaatacgactcactatagggagaaggctAAAATACCCCCAAATTACTTTACCC
PARP8_MA1_F	cagtaatacgactcactatagggagaaggctCCCTCCTATAAAATTTCATTTCCTAA
PARP8_MA1_R	aggaagagagGAAGATTTTAAGTTTTGGATTTTTGG
PARP8_MA2_F	aggaagagagTTTTTTTGAGATGAAGTTTTTTTGTG
PARP8_MA2_R	cagtaatacgactcactatagggagaaggctAATTAAACAAAACTAACCTCACCCTA
PCDHGA5_MA1_F	cagtaatacgactcactatagggagaaggctAAACATTACTCAACAAAACCTAACCC
PCDHGA5_MA1_R	aggaagagagTGTGGTAAAATGTTTTTTGTTGTATT
PCDHGA5_MA2_F	cagtaatacgactcactatagggagaaggctCAAACTTTCCCAAATAACTTAAACCT
PCDHGA5_MA2_R	cagtaatacgactcactatagggagaaggctAAACAAAATCTTACTCTATCTCCCAAA
PCDHGA5_MA3_F	aggaagagagTAGTTGAGAGATTTATAGTTGTGGGTG
PCDHGA5_MA3_R	cagtaatacgactcactatagggagaaggctTCATCCCTAAAAAAATACAAAAAACTC
PCDHGA5_MA4_F	aggaagagagTTATGAAGTTTAATGGTTATTTGAGGG
PCDHGA5_MA4_R	aggaagagagTTATTTGAGGAGAGAGGTAGTAGATGG
PRKCH_MA1_F	cagtaatacgactcactatagggagaaggctCATTTTCTACACCCATAAAAATTATACA
PRKCH_MA1_R	aggaagagagGTTTGATTAATATGGAGAAATTTTGTTT
PRKCH_MA2_F	cagtaatacgactcactatagggagaaggctCCCTCAAATAACCATTAAACTTCATAAT
PRKCH_MA2_R	cagtaatacgactcactatagggagaaggctATTAATCCTATTTCCCTTAATCTCTACC
PRKCH_MA3_F	cagtaatacgactcactatagggagaaggctCACAATACTCAACCAATACTAAAATTAC
PRKCH_MA3_R	cagtaatacgactcactatagggagaaggctAAAATCTTAATACTACCTAAATCAACCAA
RBKS_MA1_F	aggaagagagTTAGTTTTTTGTTGTTTGATGAGGG
RBKS_MA1_R	cagtaatacgactcactatagggagaaggctTAAAACACCATCCAAATTCTTACAA
RBKS_MA2_F	aggaagagagATTGTAAGAATTTGGATGGTGTTTT
RBKS_MA2_R	cagtaatacgactcactatagggagaaggctAACTCCAAAATCAAAAATCACTTAAC
S100A9_MA1_F	aggaagagagGTGTTTTTATAGTGGGTAGGGAGGT
S100A9_MA1_R	cagtaatacgactcactatagggagaaggctTCTCAAAAATAAACAACCCTCTTACA
S100A9_MA2_F	aggaagagagAAAAGGAAGGGGTAGATTGTTTATG
S100A9_MA2_R	cagtaatacgactcactatagggagaaggctAAAACTAACAACTCACTTACCAAAACC
ZAP70_MA1_F	aggaagagagGGTTTTTAGAGGTTTTGGGGTATTA
ZAP70_MA1_R	cagtaatacgactcactatagggagaaggctAAACAACCTACCCTTACTCTCCAAT
ZAP70_MA2_F	aggaagagagGATGTTTGGTATAGAAGGGGTTTTT
ZAP70_MA2_R	cagtaatacgactcactatagggagaaggctCCAACACTACCTACCCAACTACAAA

Table S9: QPCR primers for gene expression signatures genes in *BCR-ABL1* and MLLr B-ALL and for the genes in the E2A-PBX1 and MLL-AF4 knockdown and BCL6 mRNA expression experiments.

Genes	Forward (5' → 3')	Reverse (5' → 3')
ACSL1	TTCGTGTTTGACTTGTCCATTC	GTCTGCAACATGAGGTGACTGT
CD200	AAAGCTGTAAGCCCAGAAAACA	CTCCAGGGTGATATTCCAGAAG
CD38	ATCACCACACCAAACCTCTCTT	ACATCATGCAACAAGGTCATTC
CNTNAP1	CTAAACCAATGCCCTTCTCATC	GCAGAAAGGACCTATGTCAACC
FLT3	CCTTTCAGCAGAGAGATGGATT	TGGAGGGATGAAGTCCTTAAAA
FUT4	ATTCACGTTTTTCTGGACCAAG	CTGTGAGGAATCCAAACAACAA
GAB1	CCCATGATGTTTTACCAAAGGT	CCCAGATGGGGTATTAAAAACA
GAPDH	CGACCACTTTGTCAAGCTCA	CCCTGTTGCTGTAGCCAAAT
GIMAP5	CATCAGGAAGGTGAAAGAGGTC	CAGTTGTCCGTGTTTGCTACAT
HLA-DQA1	CAGAGCTCACAGAGACTGTGGT	GTCTGGAAGCACCAACTGAAC
HPRT	TGTTGTAGGATATGCCCTTGA	AGGCTTTGTATTTTGCTTTTCC
IGF2BP2	GAACAAGCTTTACATCGGGAAC	GTAGTCCACGAAGGCGTAGC
IL2RA	AGTCATGAAGCCCAAGTGAAAT	TGTGACTTCAGAGCTTCCAAAA
ITGAE	CCAGCATCCTTTGCATTACTTT	ACTTTTTGCACAATGCACAGAC
LST1	GGATCCAAGAGCTGACTATGC	GTTTTTACTGGCTGGACCACAG
LTB	GAAGGAACAGGCGTTTCTGAC	TAGCCGACGAGACAGTAGAGGT
MAP1A	GTGCCCCTGTATATGTGGATCT	CACTCGACGGAAGAAGTCAAG
MAP7	CGTCCAAGTCTCTTCCTCATTT	CTTCTCAGGCTCCACTTTGACT
MRPL33	GATTCAGAAATCCTGTAGCGTGT	ACAATGGCCTTCAAATCACAG
NOV	GGCAAATAGTCAAGAAGCCAGT	TCCCTCTGGTAGTCTTCAGCTC
PARP8	CTGCAGCTGTTAAGTCAGAGGA	GCTTCAAGTTTGTGTTGTGCTC
PCDHGA5	CGACTACCGCCAGAATGTCTAC	CTTCTTGCCCGACTTCTTCTT
PRKCH	CTGCTGGACCCCTATCTGAC	GTACGTGGGTTTGTTGGTCTTC
RBKS	GCTGCATTAGTGCTCTTGAAAA	GGTATCCACAGCCTTGACTTTC
RHOH	TGGACTTCAGAGTAGGACAGCA	CTCGTACACTGTGGGCTTGTAG
S100A9	GAGCTTCGAGGAGTTCATCAT	ATCTTGGCCACTGTGGTCTTA
ZAP70	TCTTCTACGGCAGCATCTCG	AAAGTGGTGGAAGCGCACAT
MLL/AF4 (RS4;11)	TCAGCACTCTCTCCAATGGCAATAG	GGGGTTTGTTCACTGTCACTGTCC
HOXA9*	AAAACAATGCCGAGAATGAGAGCG	TGGTGTTTTGTATAGGGGGACC
BCL6	TGAGAAGCCCTATCCCTGTG	TGTGACGGAAATGCAGGTTA
E2A	ACGCAGCCCAAGAAGGTCCG	CGGTGGCATCCCTGCCGTAG
E2A-PBX1 (697)	CCAGCCTCATGCACAACCA	GGGCTCCTCGGATACTCAAAA
CALD1	GCTTGCTCTCGGCTGTGCTCC	AGATGTCTGAGCCTTCTGGTTGGTC
ARL4C	ACAGGGCAGGCCCCTGTTGA	CGCTGAACCGTCCCAGGCAG
EXTL3	CCAGGCCATCCGGGACATGGT	GACAGGGCCTGAGGGCATCCT
ST6GALNAC3	TACGTGACCACAGAGAAGCGCA	ACAGGCGTCCATGGCCAGAAG
Cox6b	AACTACAAGACCGCCCCTTT	GCAGCCAGTTCAGATCTTCC

* HOXA9 Taqman FAM/TAMRA probe (5' → 3') CCCCATCGATCCCAATAACCCAGC

Table S10: DNA methylation (A) and gene expression (B) signature of *E2A-PBX1* B-ALL (**excel file**)

Table S11: Antibodies used for QChIP and ChIP-seq.

ChIP antibody	Host	Amount/IP	Catalog Number	Provider
MLL/AF4:				
$MLL1^{N}$	Rabbit	2ug	A300-086A	Bethyl Laboratories
$AF4^{C}$	Rabbit	2ug	ab31812	Abcam
H3K79me2	Rabbit	2ug	04-835	Millipore
H3	Rabbit	2ug	ab1791	Abcam
E2A-PBX1:				
$E2A^{N}$ SDI	Rabbit	2 ug	-	Roeder Laboratory (unpublished)
$PBX1/2/3^{C}$	Rabbit	2 ug	sc-888	Santa Cruz
$PBX1^{N}$	Rabbit	2 ug	sc-889	Santa Cruz
P300	Rabbit	2 ug	sc-585	Santa Cruz

Table S12: Primers used for QChIP.

Target promoter	Forward primer (5'→ 3')	Reverse primer (5'→ 3')
MLL/AF4:		
Control	GGCTCCTGTAACCAACCACTACC	CCTCTGGGCTGGCTTCATTC
HOXA7	AGCCGACTTCTTGCTCCTTTGC	CCCGTAGCCGCTTCTCTGTGAG
HOXA10	CGCAACCACCCCAGCCAG	TTGTCCGCCGAGTCGTAGAGG
MEIS1	TGACAACCTCGCCTGTGATTGAC	CGCCCCTCAGACCCAACTACC
CTGF	CAACTCGCATCCCTGACGCTC	AGCACCATCTTTGGCTGGAGAAG
IGF2BP2/IMP-2	TTCGTGGACTACCCCGACCAG	GCGTGGAAGTGAGCGTGCG
C20orf103	CCTGATGTTGTTCCGTGAGTAGCG	CCCGTCCTCTCTCCCAGTCG
FLT3	GCCTCGGTCCCTGCCTCTG	CCCAGCCCTGCGATGCC
FUT4	AGACCGTGCCAACTACGAGCG	GAAGGAGGTGATGTGGACAGCG
RUNX2	TGCTTCATTCGCCTCACAAACAAC	CCAGAGACTTCTTGCTGTCCTCCTG
IGFBP7	CAGCGACGGCACCACCTACC	CAGGGTTGGAGAGGGAAGCG
ITGAE	CTGCGGCGGGAACCACTG	TGAACCTCTTGGCGGGTGC
VAT1L	CGCTTAGGGCACGGCTCTTG	TTGAATGTGGGATGCTCGGC
BCL6	GGAAATCCTTACGGCTTGGCAG	GGGCTGAGGGTCTCCATCATAGG
HOXA9*	GGTGCGCTCTCCTTCGC	GCATAGTCAGTCAGGGACAAAGTGT
E2A-PBX1 :		
Control	5-TGCCTTAACTCCACAAGTTTGA	5-ACAGAGGCTCAAAGGAAAACAA
CALD1	AAGCAGCGTGAGGTGAAATGAAGC	ACACACACACACACACACACACAC
ARL4C	CCCTGGAGGCCCAGCCTTTCTTG	TGCGGACGAAATGAGGTAGCCC
EXTL3	ACAAGTGAGGCCCCCAACTGTGA	GCTTGGGTCGTGTGTCCACCC
ST6GALNAC3	GTTAGGCCCTCTCCCCTGGGTG	TGAGGCCTGGTTACATGCTGCCC

* HOXA9 Taqman FAM/TAMRA probe (5'→ 3') TACCCTCCAGCCGGCCTTATGGC

Table S13: DNA methylation (A) and gene expression (B) signature of MLLr B-ALL (**excel file**)

Table S14: Summary of number of reads and peaks detected in each ChIP-seq experiment

ChIP-seq experiment**	Uniquely aligned reads* (Mbases)	Number of peaks detected	ChIPseeqer arameters
MLL^N RS4;11	686	3,312	-t 15 -f 2 -fraglen
INPUT MLL^N RS4;11	947		
H3K79me2 RS4;11	911	28,476	-t 15 -f 2 -fraglen
INPUT H3K79me2 RS4;11	951		
$AF4^C$ RS4;11	5,585	6,253	-t 15 -f 2 -fraglen
INPUT $AF4^C$ RS4;11	4,391		
$E2A^N$ 697	815	3,358	-t 15 -f 2 -fraglen
$PBX1^C$ 697	1,022	19,108	-t 15 -f 2 -fraglen
p300	1,109	15,501	-t 15 -f 2 -fraglen
INPUT 697	874		

*Aligned to hg18 using ELAND allowing up to 2 mismatches

** $AF4^C$ and INPUT $AF4^C$ were done with HiSeq2000, and the others with GAII, Illumina

Table S15: Genomic regions enriched in MLL^N (A), $AF4^C$ (B), H3K79me2 (C) and common in the three ChIP-seq (D) in RS4;11 cells (**excel file**)

SUPPLEMENTARY FIGURES

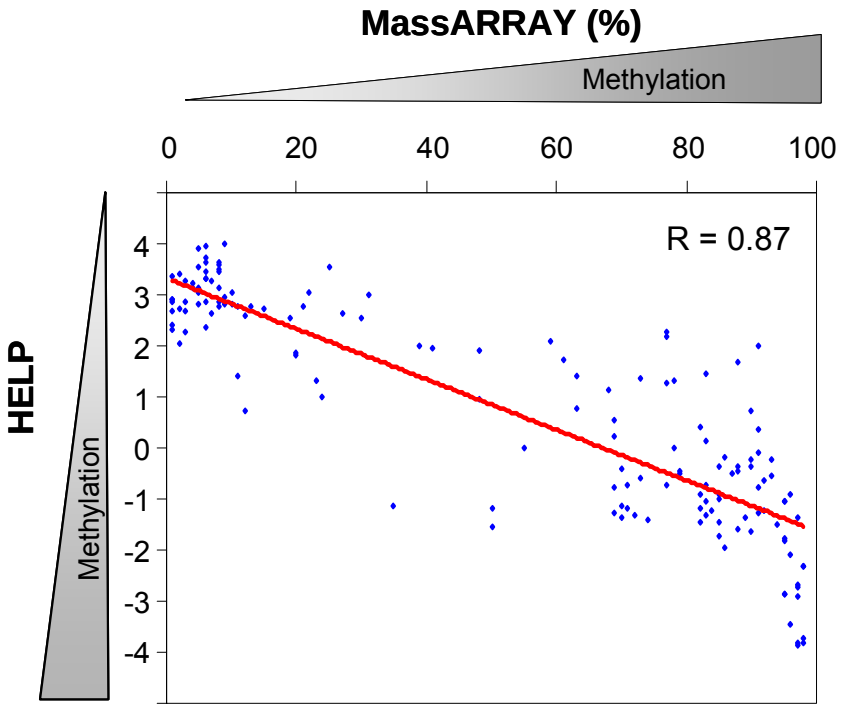

Figure S1: Technical validation of HELP arrays using MassArray EpiTyper. Correlation between HELP log ratio (y-axis) and percent methylation as measured by MassArray EpiTyper (x-axis), performed for 15 randomly selected probesets on 10 randomly selected E2993 primary B-ALL cases. Pearson correlation coefficient: -0.87, p-value < 1e-6. Methylation difference of 1 (dx>1) on HELP array corresponds to 20%, and dx>1.5 corresponds to 30%, methylation difference by MassArray (Equation for the trendline: y = - 4.9484x + 3.3232).

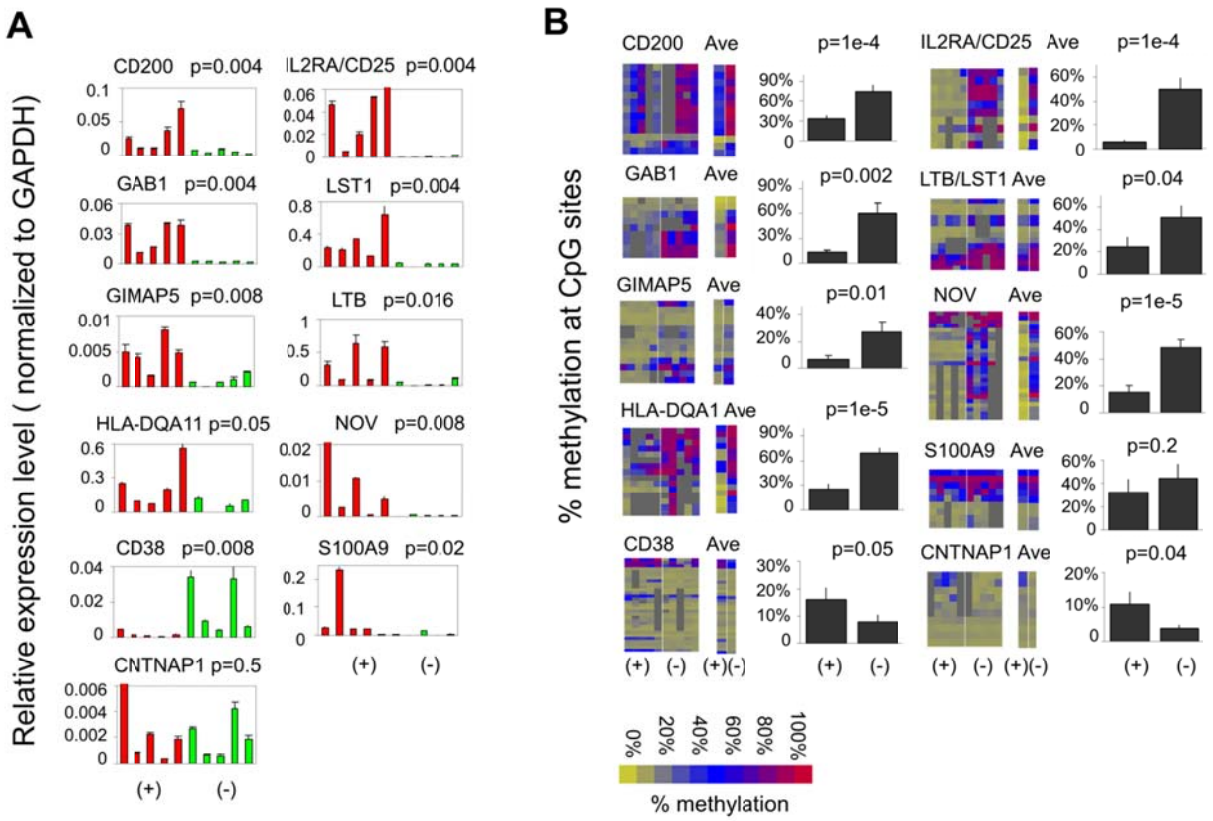

Figure S2: QPCR and MassArray EpiTyper validation of core signature genes in *BCR-ABL1*-positive B-ALL. **(a)** QPCR validation of 11 *BCR-ABL* core signature genes, performed in five randomly selected *BCR-ABL1*-positive (in red) and five *BCR-ABL1*-negative (in green) primary B-ALL cases. Transcript abundance was calculated by normalizing to the control gene GAPDH (Y axis). Error bars represent s.e.m. P values were derived by one-tailed t-test. **(b)** Heatmap representation of MassArray results for each of the 11 core signature genes performed in five *BCR-ABL1*-positive and five *BCR-ABL1*-negative B-ALL cases. Rows represent individual CpGs and columns represent individual cases. The color key represents percent methylation. The summarized heatmaps in the middle show the average methylation levels for each CpG in *BCR-ABL1*-positive and -negative patients respectively. The bar graphs show the average methylation level of all CpGs with error bars representing the s.e.m for all the CpG sites. P values were derived by one-tailed t-test.

	CD25 (+) (>0%)	CD25 (-) (=0%)
BCR/ABL (+)	69	44
BCR/ABL (-)	18	334

Fisher's Exact test p<1e-6

Figure S3: *IL2RA*(CD25) expression is associated with *BCR-ABL1*-positive B-ALL. 465 B-ALL patients enrolled in the E2993 clinical trial were studied (Imatinib-treated cases were excluded from the survival analysis). CD25 expression was measured by flow cytometry. Among the 113 *BCR-ABL1*-positive B-ALL patients, 69 (61.1%) samples are CD25 positive (CD25 expression > 0% by flow cytometry); while in the 352 *BCR-ABL1*-negative B-ALL, only 18 (5.1%) samples are CD25 positive. p<1e-6 by Fisher's exact test.

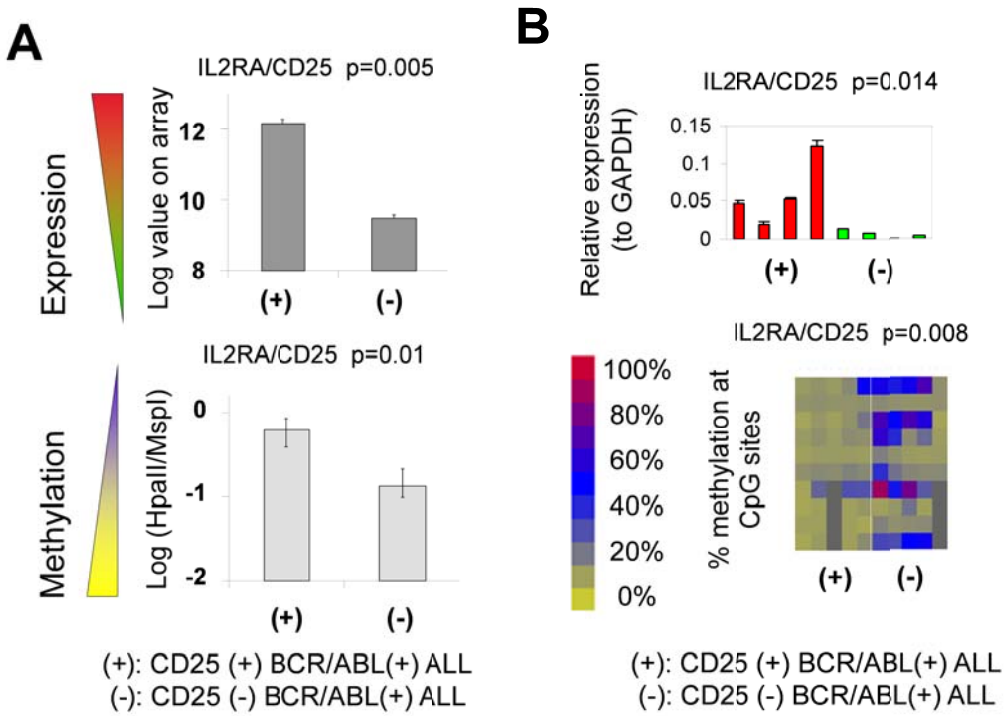

Figure S4: Differential expression and methylation of *IL2RA* in CD25+ vs. CD25- *BCR-ABL1*-positive B-ALL. **(a)** Average expression and methylation levels of *IL2RA* from the array data for the CD25-positive (CD25% >0%, n=77), and CD25-negative (CD25% =0%, n=6), *BCR-ABL1*-positive B-ALL samples. Error bars represent s.e.m. P values were shown above each panel by one-tailed t-test. **(b)** Top panel: *IL2RA* QPCR performed in four randomly selected CD25-positive (in red) and four CD25-negative (in green) *BCR-ABL1*-positive B-ALL primary cases. Transcript abundance was calculated by normalizing to GAPDH. Eror bars represent s.e.m. from PCR triplicates. Bottom: *IL2RA* MassArray performed in five randomly selected CD25-positive and five CD25-negative *BCR-ABL1*-positive B-ALL primary cases. Each row represents an individual CpGs whereas columns represent individual cases. The color key represents % methylation. P values were calculated by one-tailed t-test for the average methylation value for all CpG sites in the positive and negative groups.

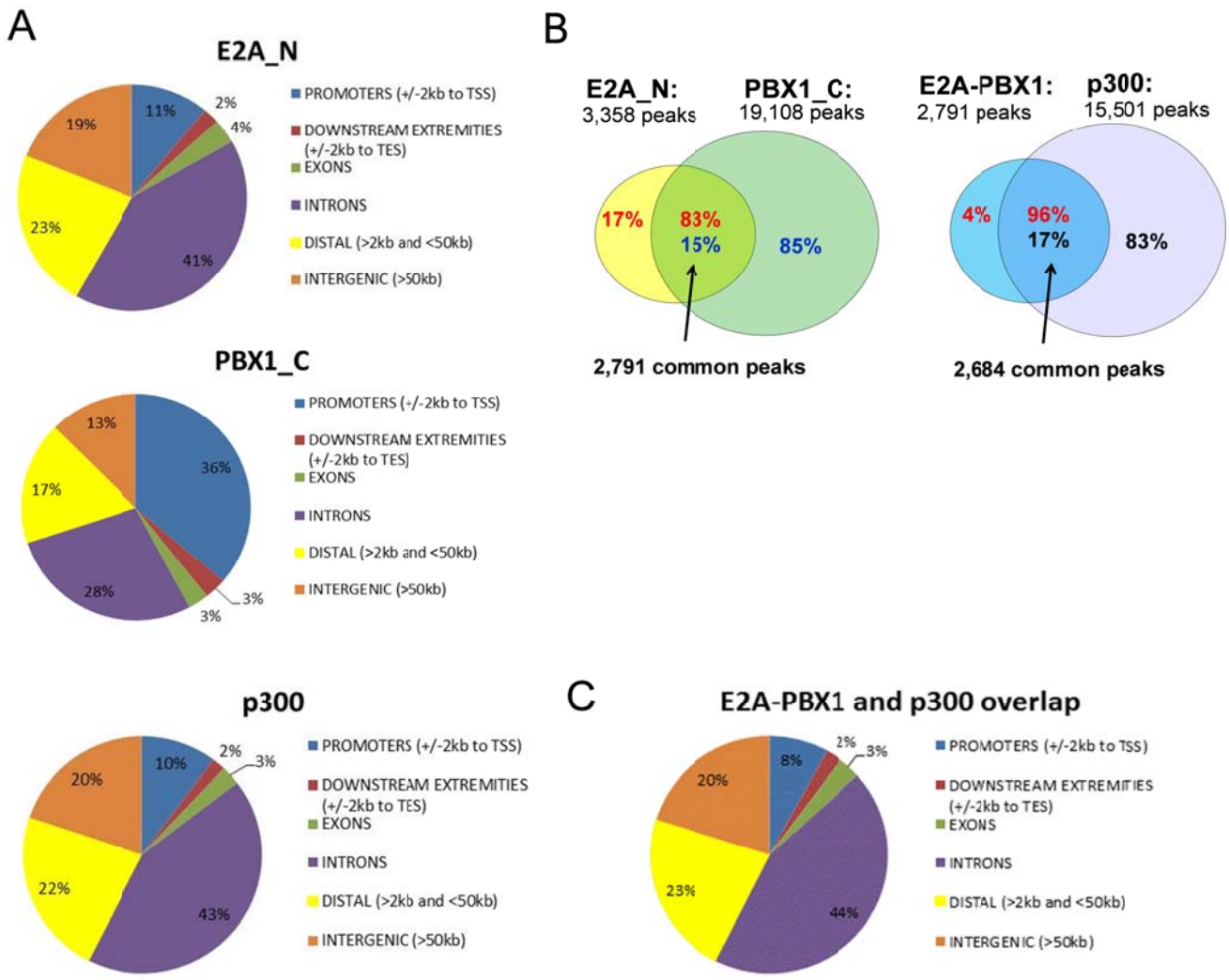

Figure S5: Whole-genome-wide binding patterns of E2A, PBX1 and p300. (**A**) E2A, PBX1 and p300 peaks were annotated to genomic regions including promoters (+/-2kb to TSS), downstream extremities (+/-2kb to TES), exons, introns, distal regions (between 2kb to 50 kb to TSS) and intergenic regions (>50kb to TSS). (**B**) Venn diagrams of overlap of peaks by E2AN and PBX1C, and the E2A-PBX1 and p300. (**C**) Annotation of genomic regions for the common peaks by E2A-PBX1 and p300.

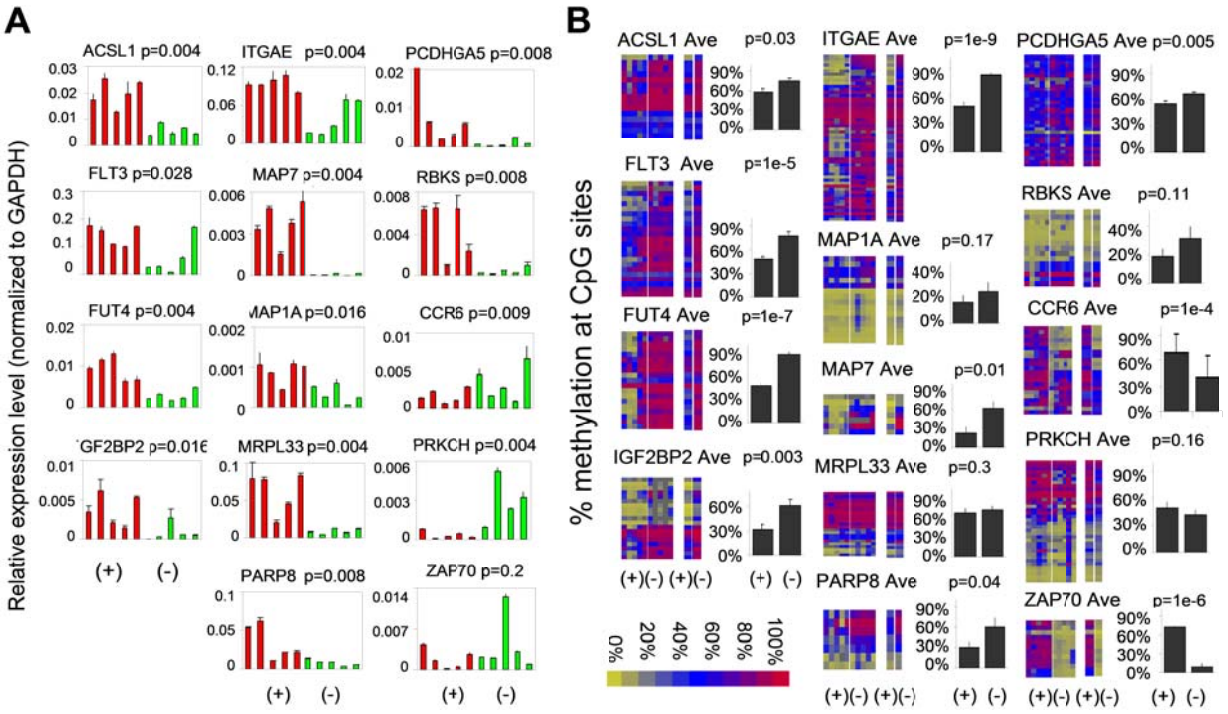

Figure S6: Validation of core MLLr signature genes. (a**)** QPCR was performed for fourteen core signature genes in five MLLr-positive (in red) and five MLLr-negative (in green) B-ALLs. Transcript abundance was calculated by normalizing to GAPDH. Error bars represent s.e.m. from PCR triplicates. P values were dervied by one-tailed t-test. **(b)** Heatmap representation of MassArray performed in five MLLr-positive and five MLLr-negative B-ALLs. Rows represent individual CpGs whereas columns represent individual cases. The color key indicates percent methylation. Panels in the middle show the averaged methylation levels for each CpG in the MLLr and non-MLLr groups. Bar graphs show average methylation level of all CpGs with error bars indicating s.e.m. for all the CpG sites. P values were derived by one-tailed t-test.

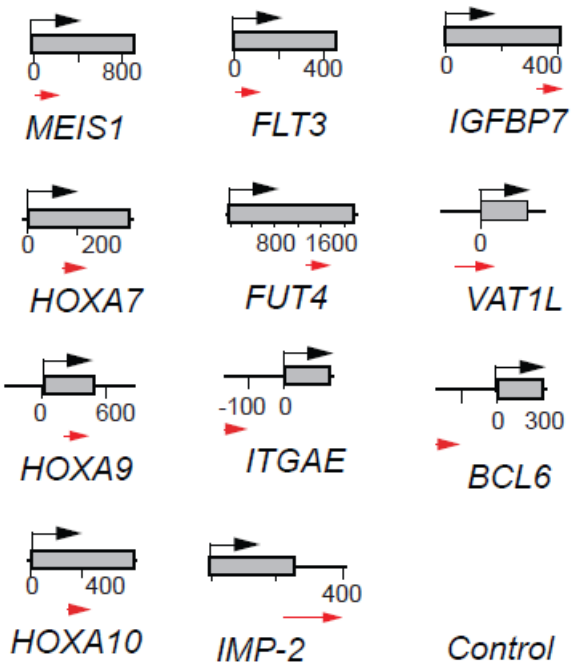

Figure S7: Primer locations for MLL^{N}, $AF4^{C}$, H3K79me2 and H3 QChIP in RS4;11 and CCRF-CEM cells. Red arrows show the primer locations and black arrow show the transcriptional start sites of the genes.

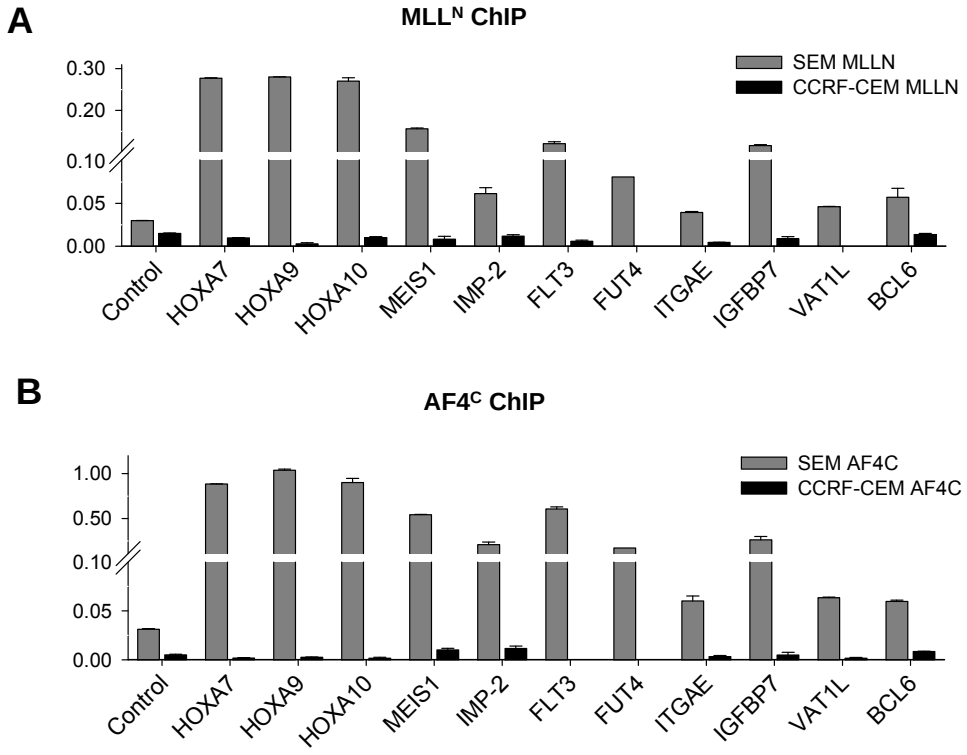

Figure S8: QChIP validation for the MLL/AF4 fusion binding in SEM cells. Each panel shows QChIP results for the four positive control, ten MLLr signature genes and one negative control, using MLL^N (**a**) and $AF4^C$ (**b**) in SEM and CCRF-CEM cells. ChIP signal was calculated as a % of input and error bars represent s.e.m. from three independent ChIP experiments.

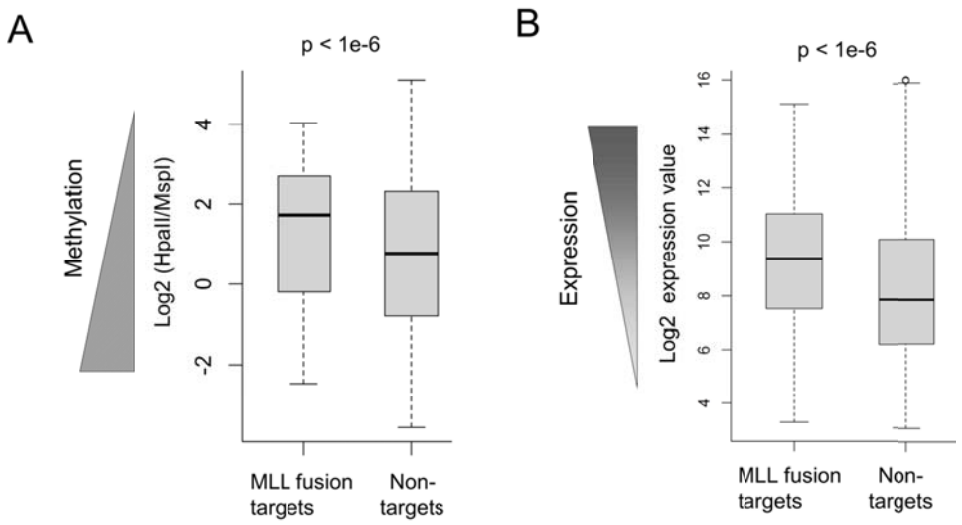

Figure S9: MLL-fusion target genes are more hypomethylated and expressed than the non-target genes in MLLr B-ALL. Box plot of methylation levels **(a)** and expression levels **(b)** of MLL-fusion target genes (n=603, identified by the overlap of MLL^N, $AF4^C$ and H3K79me2 ChIP-seq) vs. the rest non-target genes in the MLLr B-ALL from the array data. P values were calculated by one-tailed Wilcoxon test.

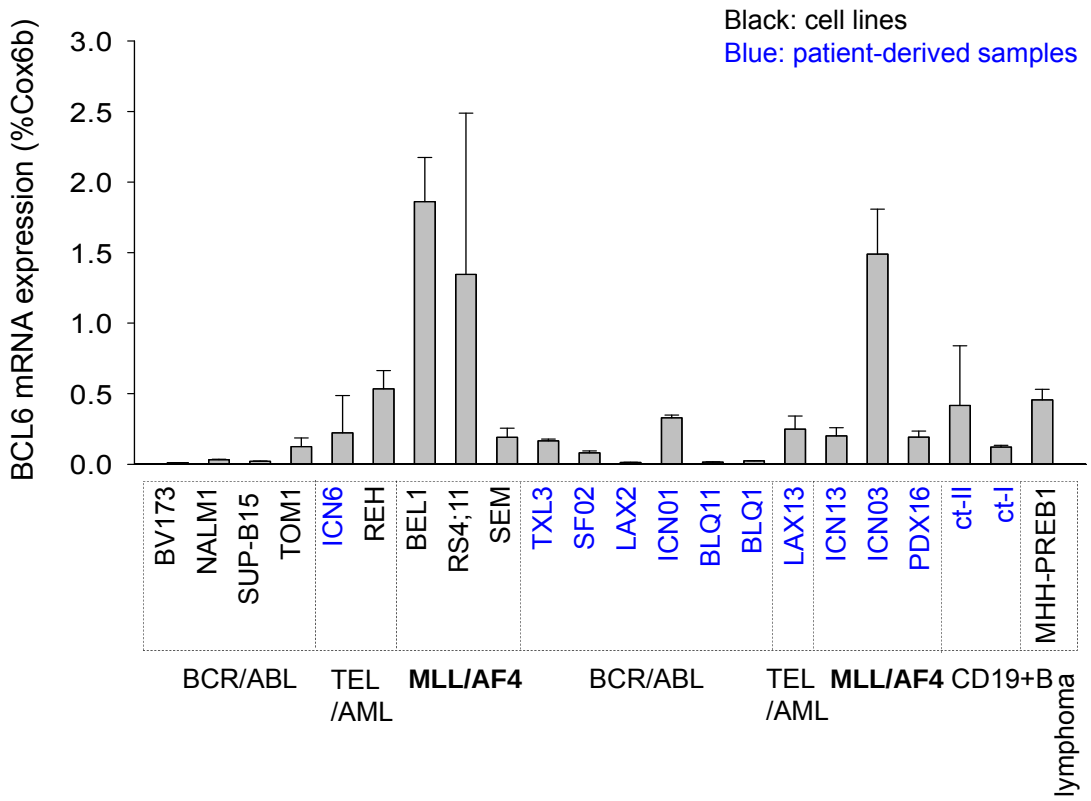

Figure S10: *BCL6* mRNA expression in MLLr B-ALL cell lines and primary samples. Figure 6D showed BCL6 western blots performed in B-ALL cell lines (in black labels) and primary specimens. This figure showed the mRNA expression of BCL6 in B-ALL cell lines (in black labels) and patient samples (in blue labels). One BCL6 expressing lymphoma cell line MHH-PREB1 was used as control.

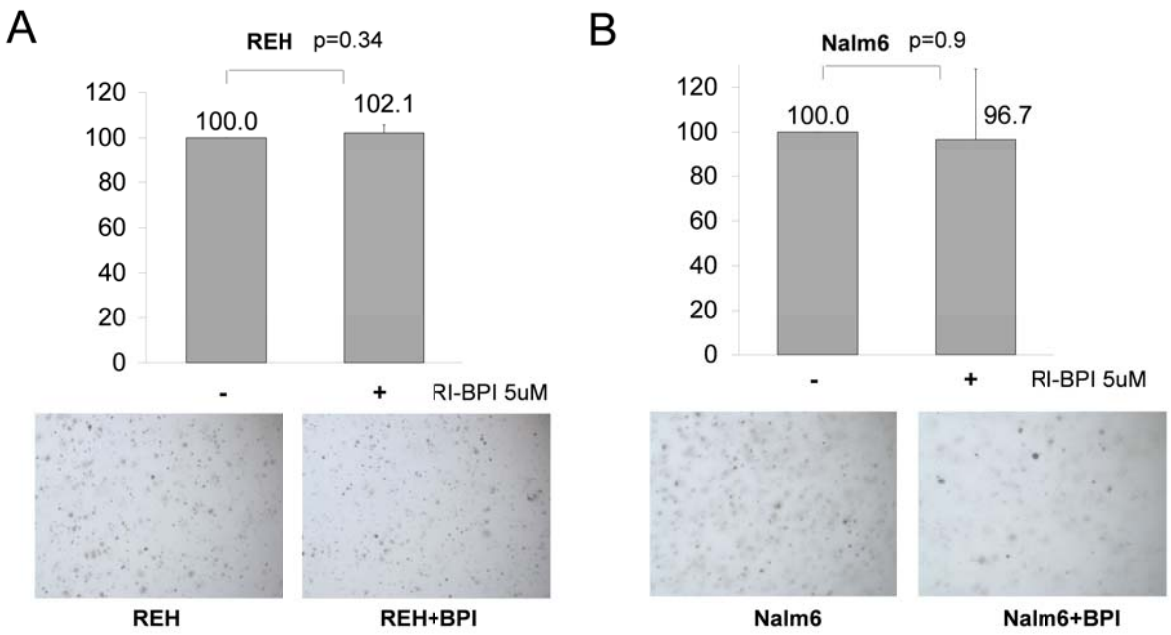

Figure S11: Colony formation assay in two non-MLLr B-ALL cell lines, REH (A) and Nalm6 (B). REH (AML/TEL) and Nalm6 cells and were exposed to 5 μM RI-BPI, or vehicle for 48 hrs, followed by colony formation assay. Y axis shows percent colony formation of RI-BPI treated cells vs. vehicle. Data represent means ± s.e.m. (n=3). Student's t-test was used to calculate p-values.

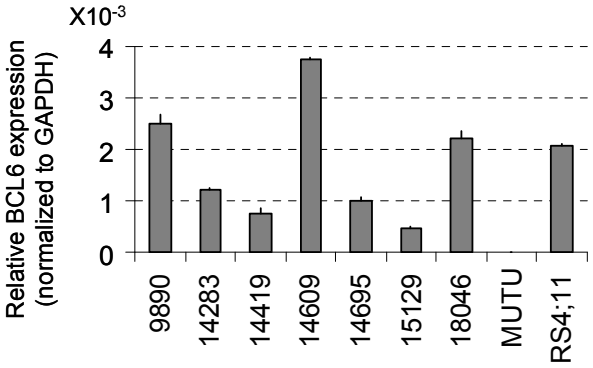

Figure S12: *BCL6* mRNA expression in MLLr B-ALL primary samples. QPCR was performed to measure BCL6 transcripts in the seven MLLr B-ALL samples treated in figure 7D. The MLL-AF4 cell line RS4;11 is shown as a positive control and the MUTU B-cell line which is known to not express BCL6 is the negative control. Transcript abundance was calculated by normalizing to GAPDH. Error bar represent the s.e.m. of PCR triplicates.

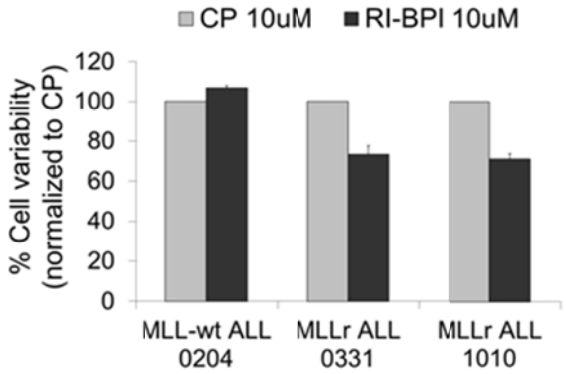

Figure S13: Effect of RI-BPI on viability of MLLr and non-MLLr B-ALL specimens. Three B-ALL specimens were obtained, two of which were MLLr and one was MLL wild type. These specimens were exposed to the indicated dose of RI-BPI or CP (control peptide). Viability was assessed by annexin V/7AAD flow cytometry and the percent viability of RI-BPI treated cells vs. CP was represented on the Y axis. Experiments were performed in triplicates and error bars represent s.e.m.

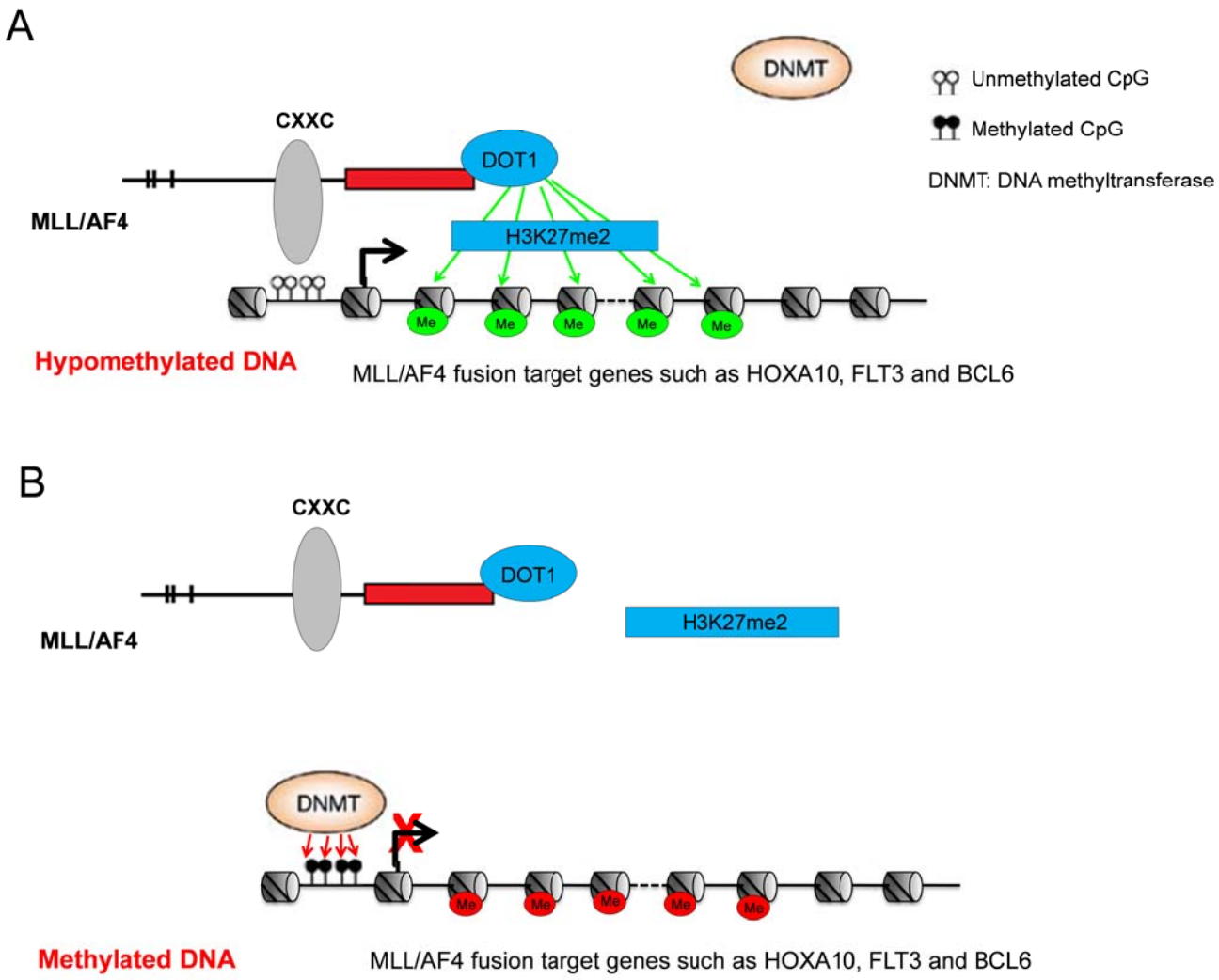

Figure S14: Scenario - DNA hypomethylation signatures in MLL/AF4 ALL are linked with MLL fusion oncoprotein binding and H3K79me2, which induce overexpression of these genes. **(A) (i),** CXXC domain in the MLL/AF4 fusion protein bind to unmethylated DNA (Cierpicki et al. Nat Struct Mol Biol 2010(6)); **(ii),** MLL-AF4 recruit DOT1, a H3K79 methyltransferase, which methylated H3K79 (Krivtsov et al, Cancer Cell 2008(7)); **(iii),** H3K79 methylation is associated with transcription activation (reviewed in Shilatifard A., Annu. Rev. BioChem. 2006(8)). **(B) (i),** CXXC domain in the MLL/AF4 fusion protein cannot bind to methylated DNA in a gene promoter; therefore **(ii),** the factors such as DOT1 recruited by MLL/AF4 fusion proteins cannot activate the transcription of the genes through H3K79 methylation **(iii)**.

Reference

1. Rowe JM, Buck G, Burnett AK, Chopra R, Wiernik PH, Richards SM, *et al.* Induction therapy for adults with acute lymphoblastic leukemia: results of more than 1500 patients from the international ALL trial: MRC UKALL XII/ECOG E2993. Blood 2005;106(12):3760-7.
2. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, *et al.* Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc Natl Acad Sci U S A 2005;102(44):15785-90.
3. Giannopoulou EG, Elemento O. An integrated ChIP-seq analysis platform with customizable workflows. BMC Bioinformatics 2011;12:277.
4. Schafer E, Irizarry R, Negi S, McIntyre E, Small D, Figueroa ME, *et al.* Promoter hypermethylation in MLL-r infant acute lymphoblastic leukemia: biology and therapeutic targeting. Blood 2010;115(23):4798-809.
5. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B (Methodological) 1995;57 (1):289–300.
6. Cierpicki T, Risner LE, Grembecka J, Lukasik SM, Popovic R, Omonkowska M, *et al.* Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. Nat Struct Mol Biol 2010;17(1):62-8.
7. Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU, *et al.* H3K79 methylation profiles define murine and human MLL-AF4 leukemias. Cancer Cell 2008;14(5):355-68.
8. Shilatifard A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu Rev Biochem 2006;75:243-69.