SUPPLEMENTAL DATA FOR

Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia

Maria E. Figueroa,[1] Shann-Ching Chen,[2] Anna K. Andersson,[2] Letha A. Phillips,[2] Yushan Li,[3] Jason Sotzen,[1] Mondira Kundu,[2] James R. Downing,[2] Ari Melnick,[3] and Charles G. Mullighan[2]
[1] Department of Pathology, University of Michigan, Ann Arbor, Michigan, USA. [2]Department of Pathology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA. [3]Department of Medicine, Hematology Oncology Division, Weill Cornell Medical College, New York, New York, USA.

Key words: Acute lymphoblastic leukemia, childhood ALL, epigenetics, DNA methylation, Integrative analysis

*Correspondence:

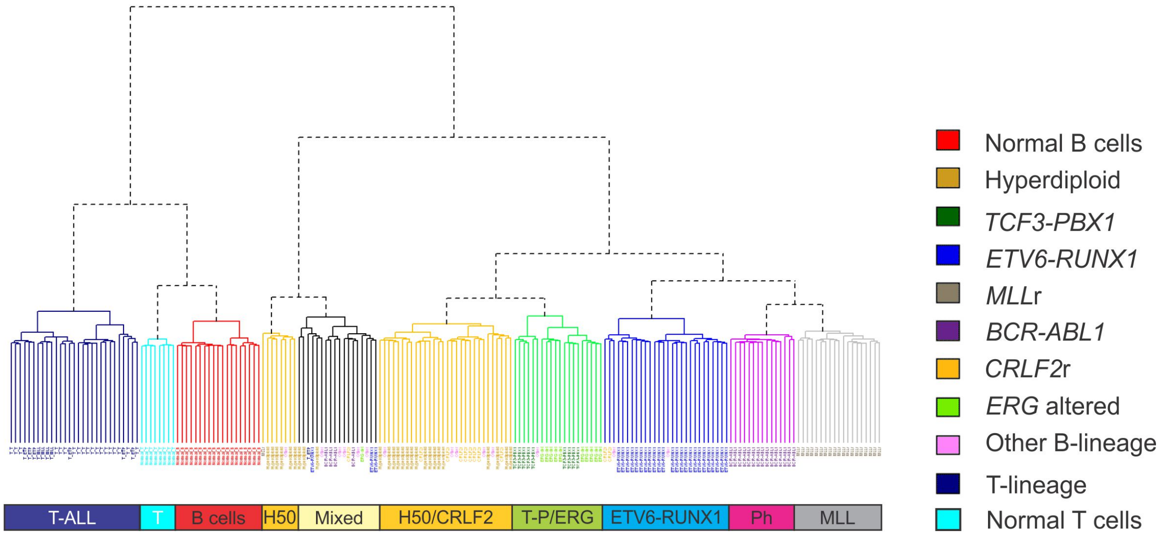
Charles G. Mullighan
Department of Pathology
MS 342, Room D4047E
St. Jude Children's Research Hospital
262 Danny Thomas Place
Memphis, TN 38105-3678
Phone: (901) 595-3387
FAX: (901) 595-5947
Email: charles.mullighan@stjude.org

Ari Melnick
Weill Cornell Medical College,
1300 York Ave, Room C620A,
New York, NY 10065
Phone: (212) 746-7643
Fax: (212) 746-8866
Email: amm2014@med.cornell.edu

SUPPLEMENTAL INFORMATION CONTENTS

Supplementary Figure 2
Clustering of HELP methylation data with data derived from normal CD3+ T cells. The normal T cell samples cluster distinctly from T-ALL and B-lineage samples. The normal T cell arrays were performed in a different batch from the remaining samples; to exclude temporal batch effects as

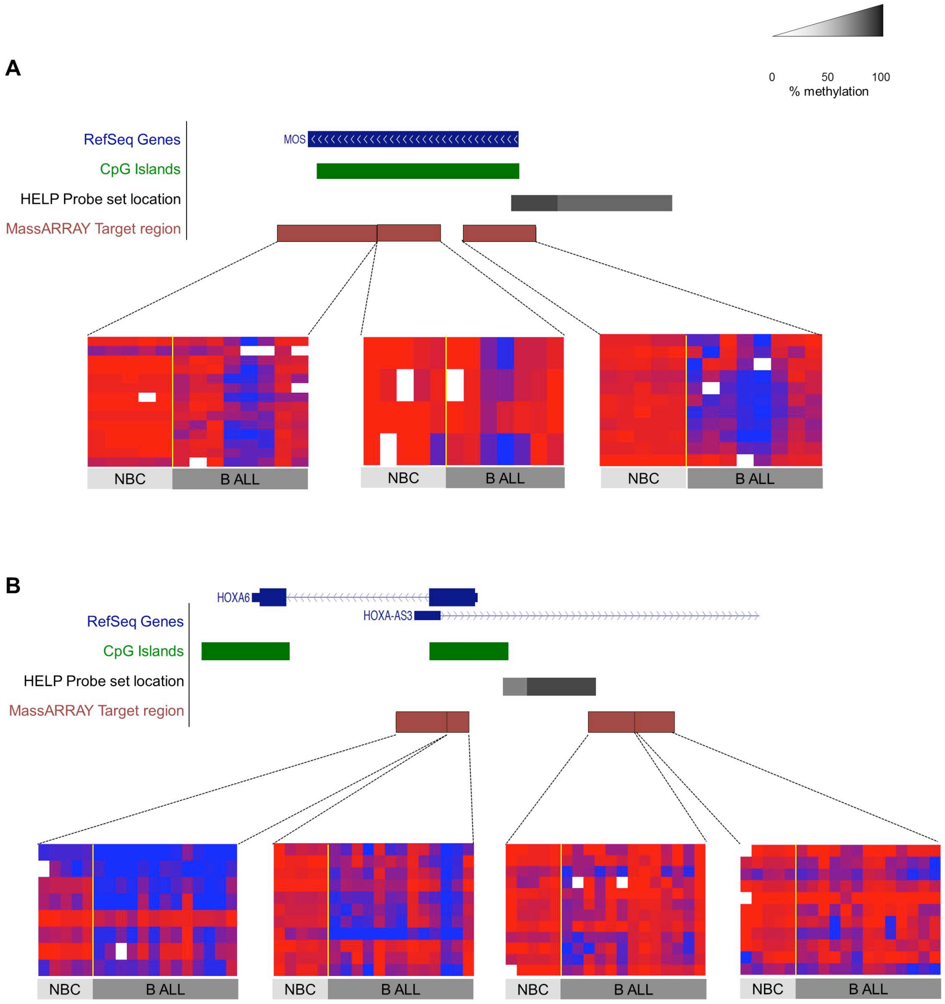
a cause of the clustering observed, 4 leukemic samples were rerun, and these clustered with the data from the sample samples ran previously. Clustering was performed with 5535 probesets (selected at SD>0.9, chosen using the optimal Rand index metric).



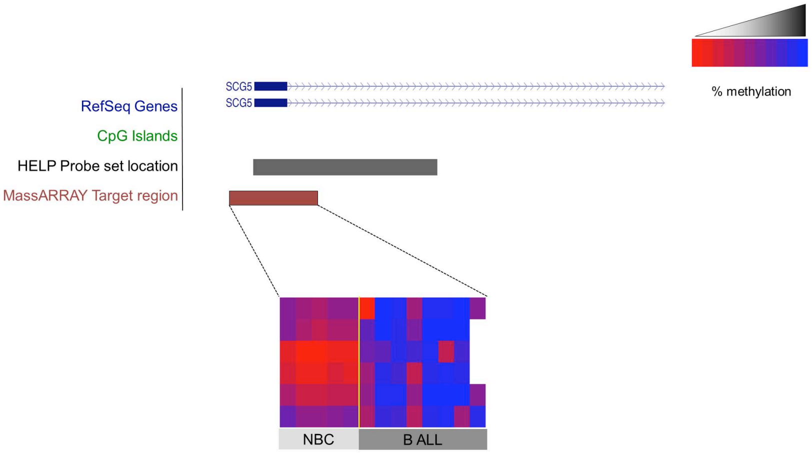
Supplementary Figure 4

Supplementary Figure 5

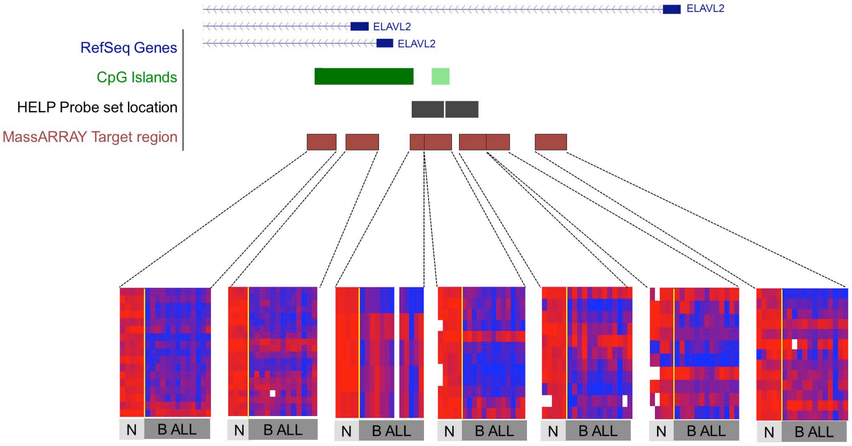
MassARRAY validation of common epigenetic signature genes. 5 genes were randomly selected from the common epigenetic signature for validation by MassARRAY EpiTYPER. For each of the 5 validated genes we depict (from top to bottom): Schematic representation of the RefSeq gene and where applicable its associated CpG island, along with the aligned location of the HELP probesets (grey) and the target regions validated by MassARRAY (brown). Below the schematic figure, heatmap representations depict the methylation status for normal B cells (NBM, left) and B ALL samples (B ALL, right) for each MassArray covered region. Each row represents a CpG site in the region, and each column represents a sample. CpG sites with missing vales for more than 5 cases were excluded from the analysis. (A) MOS, (B) HOXA6, (C) SCG5, (D) ELAVL2 and (E) ELF5.



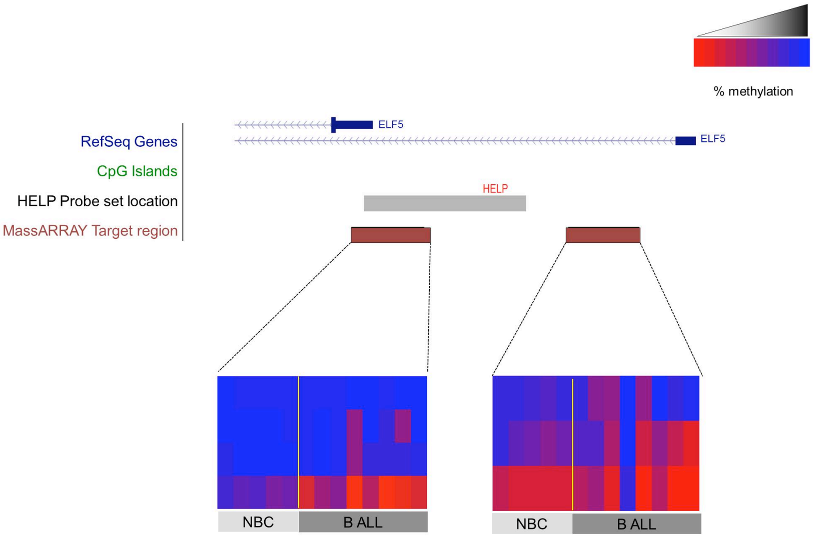
C


% methylation
SCG5
SCG5
RefSeq Genes
CpG Islands
HELP Probe set location
MassARRAY Target region
NBC
B ALL

D

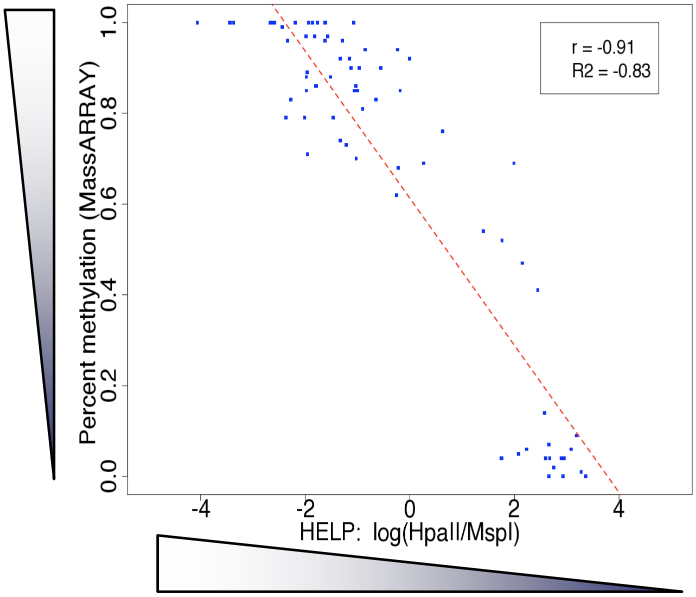

ELAVL2
ELAVL2
ELAVL2
RefSeq Genes
CpG Islands
HELP Probe set location
MassARRAY Target region
N
B ALL
N
B ALL
N
B ALL
N
B ALL
N
B ALL
N
B ALL
N
B ALL

E


% methylation
ELF5
ELF5
RefSeq Genes
CpG Islands
HELP Probe set location
MassARRAY Target region
HELP
NBC
B ALL
NBC
B ALL

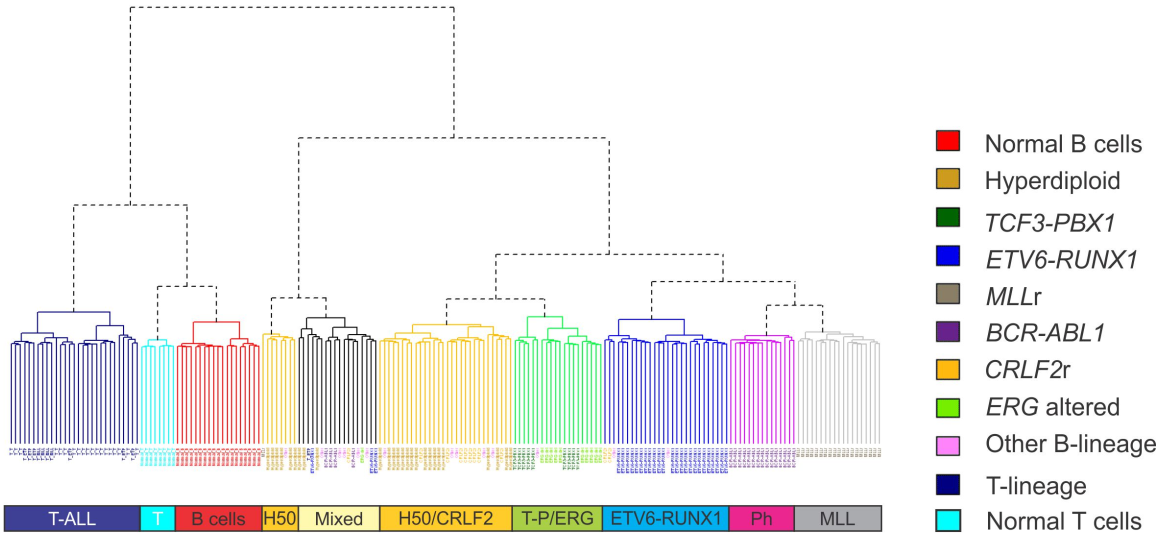
Supplementary Figure 1

Technical validation of the HELP microarray by quantitative bisulfate sequencing by MassARRAY. Correlation between HELP $\log_2$ ratio (x-axis) and percent methylation as measured by MassARRAY EpiTyping (y-axis), performed for 22 randomly selected HELP probe sets on 8 randomly selected cases. Pearson correlation coefficient: -0.91, p-value < 2.2e-16.



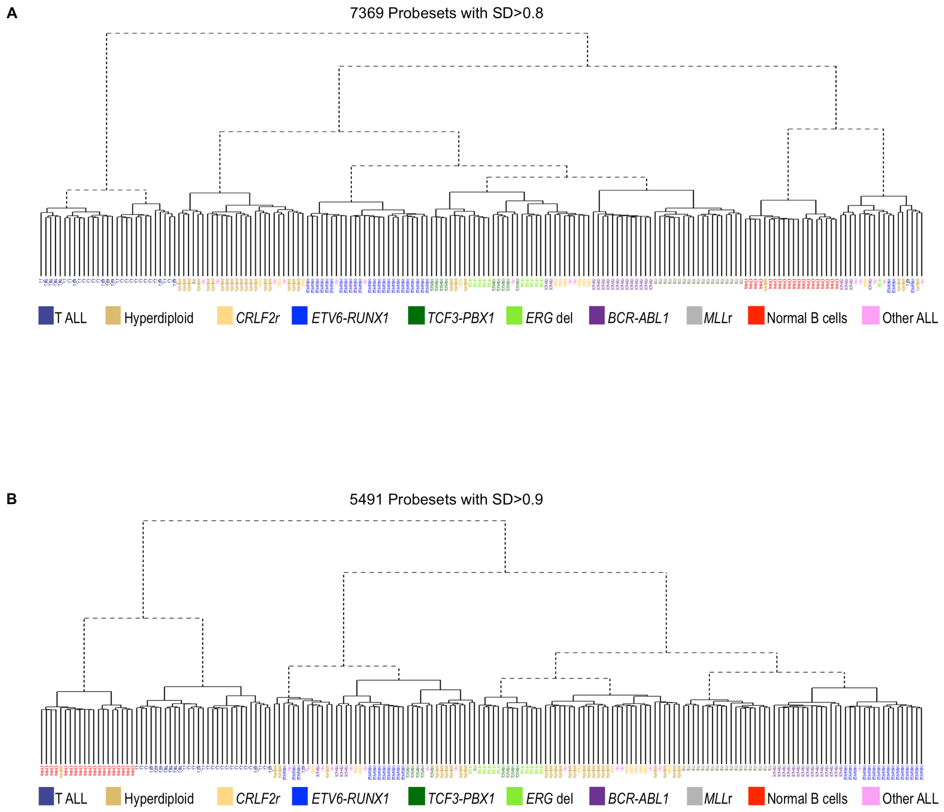
Supplementary Figure 2

Clustering of HELP methylation data with data derived from normal CD3+ T cells. The normal T cell samples cluster distinctly from T-ALL and B-lineage samples. The normal T cell arrays were performed in a different batch from the remaining samples; to exclude temporal batch effects as a cause of the clustering observed, 4 leukemic samples were rerun, and these clustered with the data from the sample samples ran previously. Clustering was performed with 5535 probesets (selected at SD>0.9, chosen using the optimal Rand index metric).

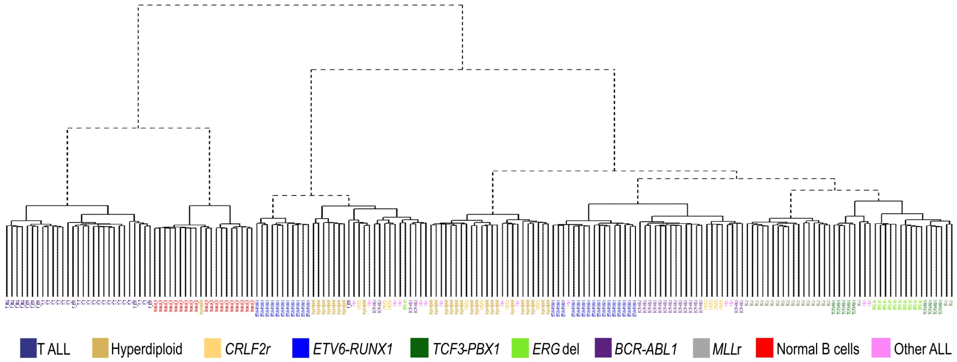


Supplementary Figure 3

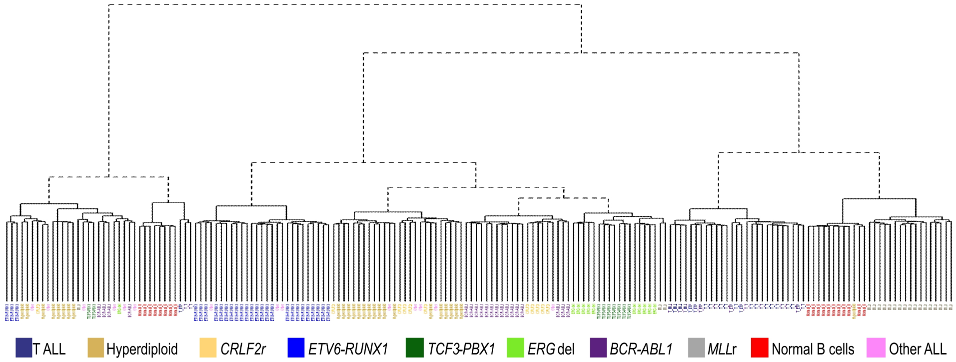
Hierarchical clustering of 167 childhood ALL cases and 19 normal B cells based on their DNA methylation profiles based on varying number of probe sets. Dendrogram representation of unsupervised analysis. Hierarchical clustering was performed on probe sets with (A) SD>0.8, (B) SD>0.9, (C) SD> 1.1 and (D) SD > 1.2 across all patients in to determine the natural segregation of childhood ALL cases and normal B cell controls based on their DNA methylation profiles.



C 2785 Probesets with SD>1.1

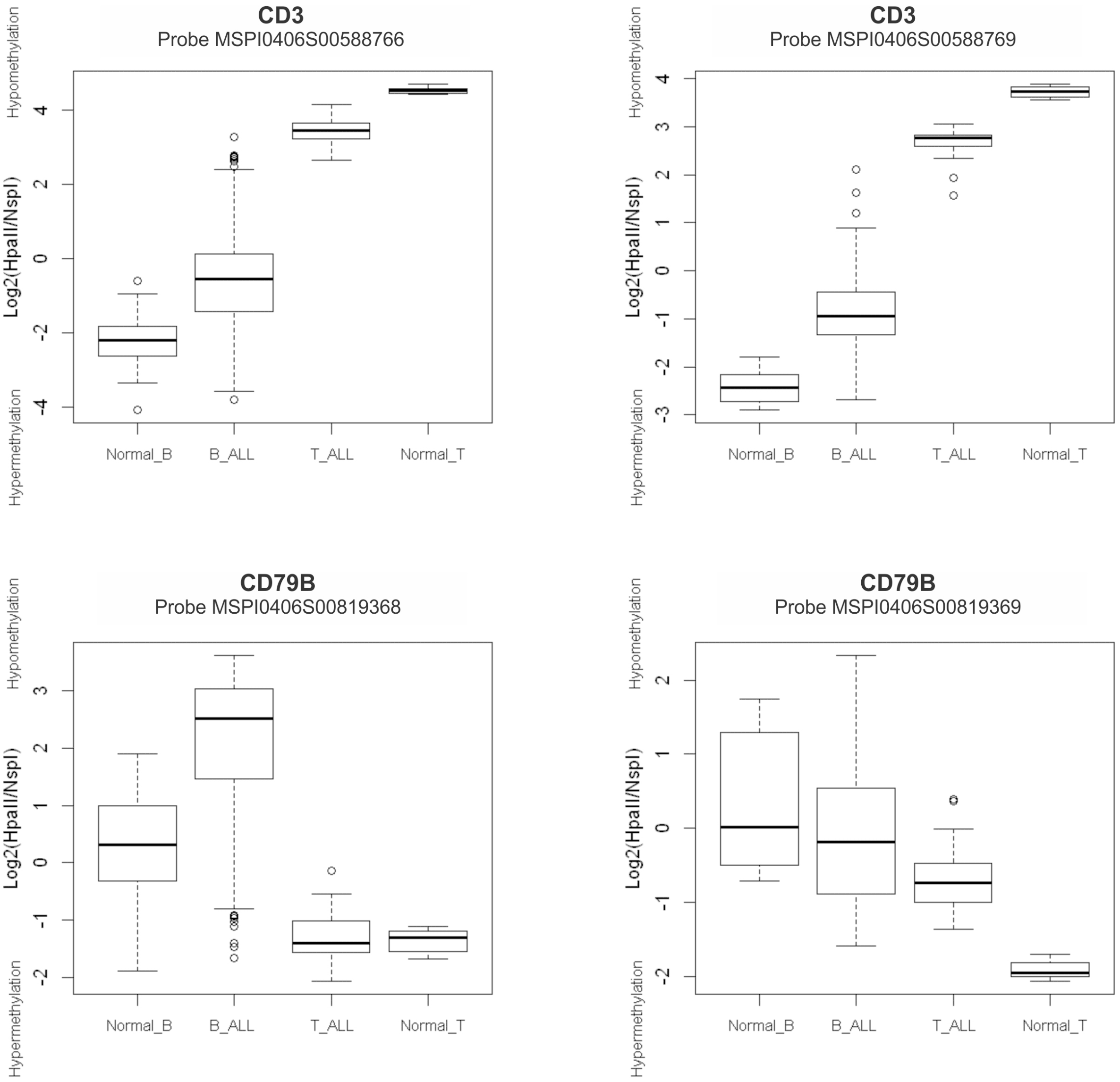


D 1866 Probesets with SD>1.2



Supplementary Figure 4

Patterns of methylation in the lymphoid signaling genes *CD3* and *CD79B* in normal and leukemic T and B cells



CD3
Probe MSPI0406S00588766

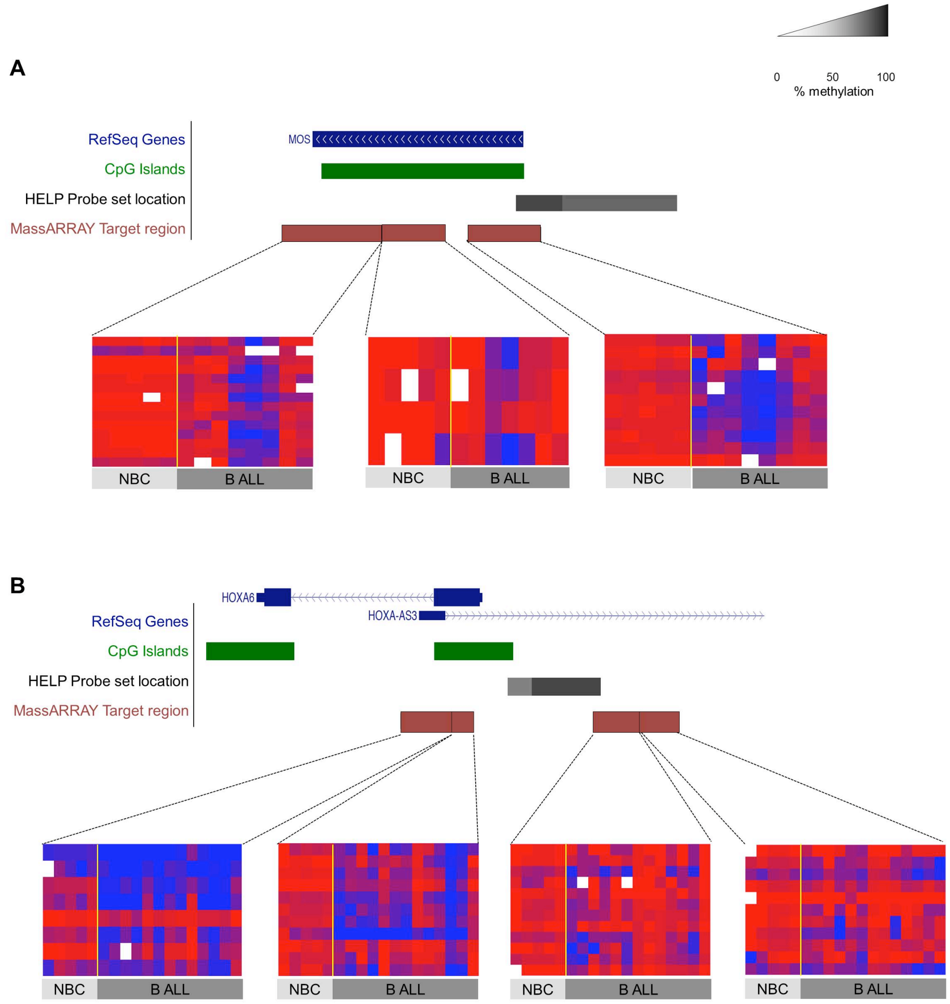
CD3
Probe MSPI0406S00588769

CD79B
Probe MSPI0406S00819368

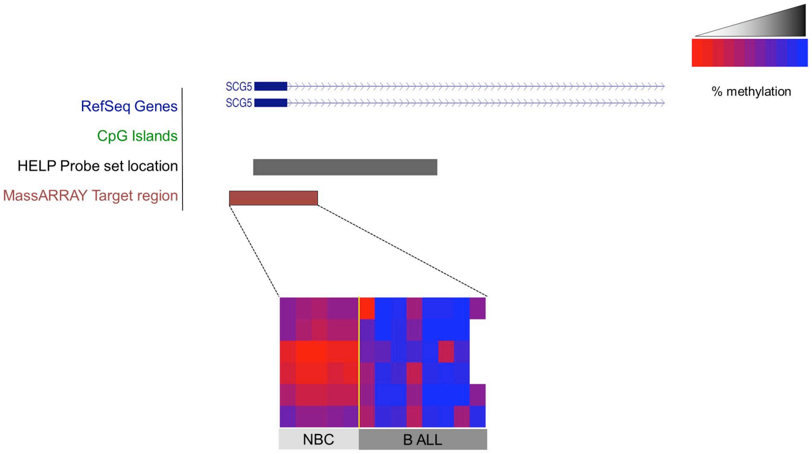
CD79B
Probe MSPI0406S00819369

Supplementary Figure 5

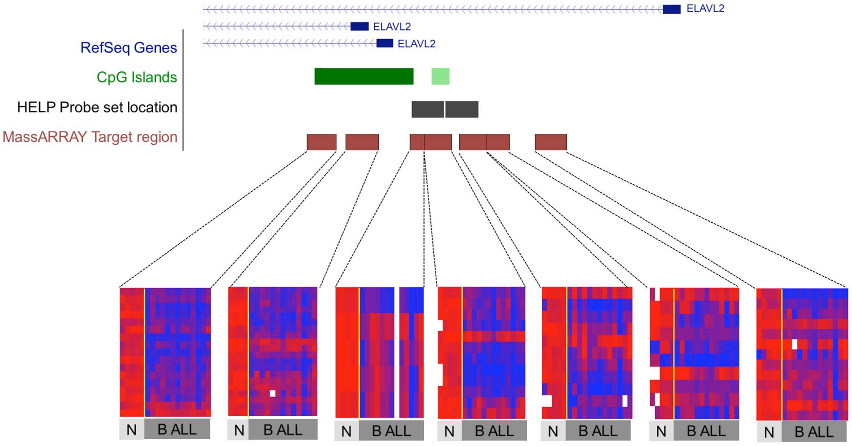
MassARRAY validation of common epigenetic signature genes. 5 genes were randomly selected from the common epigenetic signature for validation by MassARRAY EpiTYPER. For each of the 5 validated genes we depict (from top to bottom): Schematic representation of the RefSeq gene and where applicable its associated CpG island, along with the aligned location of the HELP probesets (grey) and the target regions validated by MassARRAY (brown). Below the schematic figure, heatmap representations depict the methylation status for normal B cells (NBM, left) and B ALL samples (B ALL, right) for each MassArray covered region. Each row represents a CpG site in the region, and each column represents a sample. CpG sites with missing vales for more than 5 cases were excluded from the analysis. (A) MOS, (B) HOXA6, (C) SCG5, (D) ELAVL2 and (E) ELF5.



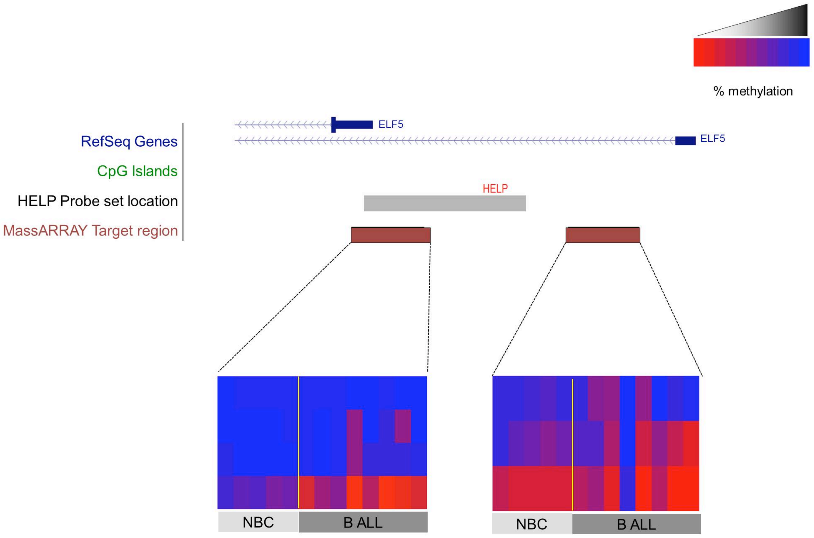
C


% methylation
SCG5
SCG5
RefSeq Genes
CpG Islands
HELP Probe set location
MassARRAY Target region
NBC
B ALL

D


ELAVL2
ELAVL2
ELAVL2
RefSeq Genes
CpG Islands
HELP Probe set location
MassARRAY Target region
N
B ALL
N
B ALL
N
B ALL
N
B ALL
N
B ALL
N
B ALL
N
B ALL

E


% methylation
ELF5
ELF5
RefSeq Genes
CpG Islands
HELP
HELP Probe set location
MassARRAY Target region
NBC
B ALL
NBC
B ALL

Supplementary Table 1

Listing of patient, sample and array characteristics. See separate excel workbook.

Supplementary Table 2:

Listing of differentially methylated regions for each subtype of ALL. See separate excel workbook.

Supplementary Table 3. Top networks enriched in each subtype-specific DNA methylation profile

Results of Ingenuity pathway analysis of the methylation signature of each subtype of ALL

Supplementary Table 3A. *ETV6-RUNX1* B-ALL

ID	Molecules in Network	Score	Focus Molecules	Top Functions
1	14-3-3, ATF7, BBC3, Cg, Creb, Cyclin A, DOK5, E2F6, E2f, EFEMP1, EIF2C1, EIF2C2, Estrogen Receptor, ESX1, FAIM, FUCA1, GFI1, IRAK3, KPNA4, KRAS, LIMD1, MAP2K1/2, MEFV, MIB1, MIRLET7, MLL2, MSX2, NFkB (complex), PDE3A, PELI3, RIN1, SAA1, SLC2A5, SLC35B1, USP6NL	47	26	Cell Death, Cancer, Cell Cycle
2	BRWD1, DBX2, ERK, FOXA2, FSH, GAB1, GLO1, Gpcr, Hsp70, Insulin, L-type Calcium Channel, Lh, NCKAP1, p85 (pik3r), PAX6, PDGF BB, PDK4, PIK3CG, Pka, Pka catalytic subunit, POU5F1, PPARD, Pro-inflammatory Cytokine, Proinsulin, RAB33B, RAD50, Ras, RASAL1, SCAP, STBD1, TFCP2L1, TNK1, Trk Receptor, ZNF143, ZNF462	32	20	Lipid Metabolism, Small Molecule Biochemistry, Cell Death
3	Actin, Akt, Alpha catenin, Ampa Receptor, Ap1, ARF1, ARF6, CALD1, CaMKII, DLGAP2, EIF6, ELOVL5, F Actin, Fcer1, GFAP, GRIA3, Laminin, MTORC1, Myosin, NACA, NEFL, NEFM, NFAT (complex), Nfat (family), OPHN1, PJA2, RAB8B, Rock, RRAGA, S100A9, TJP1 (includes EG:21872), TNFRSF9, TOPORS, UBE2J1, WIPF1	30	21	Cell Morphology, Cellular Assembly and Organization, Cellular Development
4	Alp, AMPK, BDNF, BGLAP, BMP7, Cbp/p300, CNR2, CXCR3, Dynamin, ERCC6, Growth hormone, HDL, Histone h4, Ifn gamma, ITSN2, KCNA3, LDL, MYO1E, N-type Calcium Channel, NADPH oxidase, NOS3, Nos, OGG1, PI3K (complex), PIAS2, Pias, PIK3C3, PIK3IP1, SCARF2, SERPINE1, SIRT2, Smad2/3, Tgf beta, UCP2, WWTR1	30	19	Cell Signaling, Small Molecule Biochemistry, Free Radical Scavenging
5	BCR, CD79B, CIITA, CLIP1, CTCF (includes EG:10664), DAXX, ERK1/2, FGFR4, HABP4, HLA-C, HLA-DQ, HLA-DQA1, HLA-DR, Ige, IgG1, IgG, Igm, IL12 (complex), Immunoglobulin, Interferon alpha, ISG20, Mek, MHC CLASS I (family), MHC Class II (complex), MTA3, NCK, PLC gamma, PTPRB, Rap1, SIGLEC7/SIGLEC9, Sos, SPOP, TUB, UCN2, VRK2	26	17	Cell Death, Inflammatory Response, Cellular Development

Supplementary Table 3B. *BCR-ABL1* ALL

ID	Molecules in Network	Score	Focus Molecules	Top Functions
1	Ap1, ATOH1, B3GNT5, BCL2L13, BCR, CARD9, CD22, CD38, CSDA, CSF1R, DAXX, DUSP5, FGF5, FGF6, Fgfr, GALP, GRB7, GUK1, IFN alpha/beta, IgG, Igm, IL1, IL12 (complex), IL2RA, Immunoglobulin, Interferon alpha, LILRB4, LIMD1, MHC Class II (complex), NCK, NFkB (complex), PLC gamma, Rsk, Sapk, TLR9	33	19	Cellular Growth and Proliferation, Hematological System Development and Function, Cell Death
2	ADCY, AMPK, BDNF, BRDT, CACNA1B, Calmodulin, CaMKII, Ck2, CLTB, CNTN2, Cofilin, Creb, CTXN1, Estrogen Receptor, FASN, Gsk3, Histone h4, Insulin, IQGAP2, ITIH3, KCNN2, KLC1, Lh, LRRTM1, Mlc, MSX2, PI3K (complex), PROC, RNA polymerase II, SFN, SIRT2, SSR1, STAT5a/b, TRIM13, TUBB2A	30	19	Lipid Metabolism, Small Molecule Biochemistry, Vitamin and Mineral Metabolism
3	ADAMTS12, Alp, Alpha catenin, ATP1A3 (includes EG:232975), Cadherin, CCL25, CDH5, CDH8, COL1A1, COL1A2, collagen, Collagen Alpha1, Collagen type I, Collagen type III, Collagen type IV, Collagen(s), CTNNA2, ERK1/2, Growth hormone, Integrin, ITGA2, KLF13, Laminin, MARCKSL1, Mmp, PCOLCE, PDGF BB, PROX1, Rock, Smad, Tgf beta, TINAG, TMEFF2, UCN2, ZEB2	29	17	Cellular Movement, Connective Tissue Development and Function, Gastrointestinal Disease
4	C11orf30, CCDC71, CCR7, CHEMOKINE, CLDN4, CXCL12 (includes EG:20315), ERK, Focal adhesion kinase, FOXQ1, GAB1, Gpcr, GPR6, GPR146, GPR153, Histone h3, IL12 (family), Jnk, LPAR2, Mapk, NKX2-1, p85 (pik3r), PDPN (includes EG:10630), Pka, Pkc(s), POU5F1, Rac, Ras, Ras homolog, SALL3, Sos, STAT, Trk Receptor, Ubiquitin, Vegf, VKORC1	26	16	Cellular Movement, Cellular Development, Cell Death
5	14-3-3, 26s Proteasome, Akt, ARID3A, BTG2, Caspase, CD3, E2f, Fcer1, GFAP, hCG, Hsp27, Ifn gamma, IL11, INPP5E, IP6K3, KLF6, LIF, MAP2K1/2, Mek, MUC4, Nfat (family), P38 MAPK, p70 S6k, Pro-inflammatory Cytokine, PRPH, SIGLEC7/SIGLEC9, SLC20A1, SP100, TCR, THEM4, Tnf, TNFRSF9, TRIB3, WIPF1	25	17	Cell Death, Gene Expression, Cell Signaling

Supplementary Table 3C. *TCF3-PBX1* ALL

ID	Molecules in Network	Score	Focus Molecules	Top Functions
1	ATP2B4, BCR, BLK, CD79B, CTCF (includes EG:10664), DEF6, ETS, Fc gamma receptor, FCGR2A, HISTONE, IDH2, IgD, IgG1, IgG, IgG2a, Igm, IL2RA, Immunoglobulin, LAT, MNAT1, MTA3, NCK2, NCK, NCKIPSD, NFkB (complex), PDCD1, PEA15, PLC gamma, PPP1R15B, PRDX1, SLAMF1, ST8SIA1, SYK/ZAP, TNFAIP8, WIPF1	37	21	Cell Death, Cell-To-Cell Signaling and Interaction, Hematological System Development and Function
2	ADIPOQ, AKR1B1, Alp, ASGR2, CCL25, CEP350, CLU, Collagen Alpha1, Collagen type III, Collagen type IV, Collagen(s), Cpla2, DUSP4, ERK1/2, Focal adhesion kinase, HGF, KLHL21, KRAS, LAIR1, LDL, Mek, NFAT (complex), p70 S6k, Pdgf (complex), PDGF BB, PPARD, Rock, SERPINE1, SORBS3, Sos, SOST, SOX9, STAB1, Tgf beta, VRK2	31	18	Cellular Growth and Proliferation, Connective Tissue Development and Function, Hepatic System Development and Function
3	ARRB1, ATP, CD164, CPNE8, CYSLTR1, DAZAP2, ERCC5, ERCC6, EWSR1, FAS, FBXO6, FOS, FPGS, FXYD2, IL13, IL13RA1, ITGB1, MAP3K7 (includes EG:172842), MT1A, NCLN, NUP107, PCNA, PIP4K2B, POLD2, POLD4, POLI, POLM, RAB37, RASL11B, ST6GAL1, TLN2, TSPAN3, TSPAN, VRK2, ZNF652	26	16	Cellular Function and Maintenance, Molecular Transport, Small Molecule Biochemistry
4	Akt, Ap1, ARHGEF6, BTG2, Cyclin A, DAXX, Dynamin, E2f, ERCC6, Estrogen Receptor, Fcer1, Gm-csf, GNA13, HABP4, hCG, HIST1H1C, IFN alpha/beta, Ikb, MAP2K1/2, Nfat (family), NFKB1, NGF, P38 MAPK, Pak, PDK1, POLD1, Rac, Rxr, Sapk, SOCS2, SPOCD1, SRGAP3, TERT, TFAP2C, TLR1	25	17	Cancer, Cell Death, Cellular Development
5	ADCY, ARHGAP24, ARL4C, CHEMOKINE, CRYM, CYSLTR1, ERK, FFAR1, FSH, G-protein beta, GAB1, GABBR1, Gpcr, GPR123, HBEGF, IKK (complex), Insulin, Lh, LHCGR, LPAR2, Mapk, Mmp, MND1 (includes EG:295160), p85 (pik3r), Pka, Pkc(s), PLC, PSMC3IP, RANBP9, Ras, Ras homolog, Shc, Trk Receptor, USP11, Vegf	24	15	Cell Signaling, Nucleic Acid Metabolism, Small Molecule Biochemistry

Supplementary Table 3D. *MLL*-rearranged B ALL

ID	Molecules in Network	Score	Focus Molecules	Top Functions
1	20s proteasome, 26s Proteasome, Alpha tubulin, ARRDC4, CLINT1, DAPK3, DAXX, ERK, Estrogen Receptor, GADD45B, GP9, HABP4, HISTONE, Hsp70, Hsp90, MAP1LC3A, MDM4, MEF2C, MLL2, MOV10, MYOD1, Nos, PAPOLA, POMP, PRX, PSMB5, RNA polymerase II, SFN, SIRT2, SLC27A2, SOX15, TERT, Ubiquitin, WDYHV1, ZFHX3	40	24	Post-Translational Modification, Protein Folding, Infectious Disease
2	ACSL1, Akt, Collagen type IV, DES, FAM189B, FLT3, GFAP, IP6K1, KCNK3, Laminin, LARGE, MAG, MCF2L, Mlc, MPP5, MUC4, MYL4, NCK2, NCK, NCKIPSD, NFAT (complex), NID1, Notch, Pak, Pdgf (complex), PHLDB1, PTPRE, PXK, Rock, SORBS3, Sos, TINAG, VAV, WIPF1, ZAP70	35	23	Cellular Assembly and Organization, Cellular Function and Maintenance, Cell Morphology
3	Ampa Receptor, BDNF, CACNB3, Calcineurin protein(s), Calpain, CaMKII, CNKSR2, CXCR5, DAPK1, DLGAP2, ERK1/2, FASN, G protein beta gamma, GHRL, GNB4, GRINA, GUCY1A3, ITIH3, ITPRIP, L-type Calcium Channel, MAP2K1/2, MAP3K, N-type Calcium Channel, NADPH oxidase, NMDA Receptor, NPR1 (includes EG:18160), PEA15, PHLPP1, PRKCH, Proinsulin, RAB11FIP1, RASGRF2, Sapk, SIK1, SORL1	34	21	Cell-To-Cell Signaling and Interaction, Nervous System Development and Function, Amino Acid Metabolism
4	B-cell receptor, C1R, CD8, CD22, CD37, CD180, CD79B, ETS, FAIM, FUCA1, HLA-B, HLA-C, IFN alpha/beta, IgD, IgG1, IgG, Igm, Immunoglobulin, IRAK3, KIR, MHC Class I (complex), NFkB (complex), NLRP12, PELI3, POU2AF1, RFTN1, RGS14, RHOH, SPIB, STAB1, SYK/ZAP, TARS, TLR7, Tlr, Tnf	32	20	Cell-To-Cell Signaling and Interaction, Hematological System Development and Function, Immune Cell Trafficking
5	AK1, Androgen-AR, APP, ARID1B, ATP2B3, Brd1, Calmodulin, Caspase, Ck2, Creb, EDEM1, FOXN3, Histone h3, Histone h4, IFITM2, IL12 (complex), Interferon alpha, KAT7, KCNC4, KDM4B, KDM5B, KPTN, LILRB4, LRRTM3, Mmp, P38 MAPK, PHF16, PIP4K2A, Pka, PRKCB, SSR1, STAT5a/b, SUV39H2, VKORC1, ZNRF3	26	18	Cell Death, Cell Cycle, Antigen Presentation

Supplementary Table 3E. *ERG*-deleted B-ALL

ID	Molecules in Network	Score	Focus Molecules	Top Functions
1	ACSBG1, ACSL3, ACSL, Alpha tubulin, Caspase, CCNI, CCNT1, CDKN1A, CDKN2C, CEP70, CLU, Cyclin A, Cytochrome c, EIF4G2, ERCC6, GMNN, HELZ, HIST1H1C, Histone H1, Histone h3, KAT7, MLL2, MTA2, Pkc(s), PPID, Rb, REPIN1, RFC4, RNA polymerase II, SECISBP2, SLC27A2, SPATA24, TGFBR2, TIMM13, TOMM40	39	25	Cellular Growth and Proliferation, Cancer, Dermatological Diseases and Conditions
2	ADCY1, CACNA1S, Calcineurin protein(s), CNTN2, CYTIP, ECM1 (includes EG:100332249), ERK1/2, Fgf, FGFR4, Fgfr, GLG1 (includes EG:20340), GLRA2, GNAZ, growth factor receptor, HS6ST1, HSPG2, Integrin, KLHL22, L-type Calcium Channel, L1CAM, LIN7A, PJA2, PLC gamma, PTPRB, RAB11FIP1, RGMA, Smad1/5/8, Sos, SQLE, SYVN1, UBE2J1, VAV, VRK2, ZMYM6, ZNF423	37	24	Cellular Movement, Cardiovascular System Development and Function, Embryonic Development
3	ATP2B4, BNIP3L, BRDT, CD33, CD38, CHRNA1, CHRNA2, CXCR3, EGLN3, ELP3, ERK, Fc gamma receptor, Histone h4, IFI30, Ifn gamma, Ikb, IL12 (complex), JAK, LFNG, LIF, MUC1, Pdgf (complex), PIAS2, RAD54L, RAMP1, RANBP9, SEMA4A, Smad, SNAI2, SOCS2, SOCS3, SOCS, STAT, STAT5a/b, Tgf beta	33	22	Cellular Development, Cellular Function and Maintenance, Cell-mediated Immune Response
4	ADCY, CCR7, CHEMOKINE, CHEMOKINE RECEPTOR, CNR2, CXCR1, CYSLTR1, Fibrinogen, G protein alphai, GAB1, GIPR, Gpcr, GPER, GPR44, GPR84, GPR123, GPR146, GPR148, GPR153, GPR155, Gs-coupled receptor, HCAR1, HTR7, LHCGR, Mapk, MC4R, PCNX, Pdgfr, Pik3r, Pld, Proinsulin, SHANK1, Shc, SSTR2, STK38	33	22	Cell Signaling, Nucleic Acid Metabolism, Small Molecule Biochemistry
5	ADAMTS9, B-cell receptor, CARD8, CARD14, CD180, CENPJ, ERAP2, FUCA1, G0S2, HSP, Ifn, IFN alpha/beta, IFN Beta, Iga, IgG1, Igg3, Igm, IRAK3, IRAK, MHC Class II (complex), MT3, NFkB (complex), NLRP12, PELI3, PNPT1, POU2AF1, RFTN1, SHIP, SLAMF1, SPIB, SYK/ZAP, TLR1, TLR7, Tlr, TNFAIP8	28	20	Cell-To-Cell Signaling and Interaction, Hematological System Development and Function, Immune Cell Trafficking

Supplementary Table 3F. High hyperdiploid B-ALL

ID	Molecules in Network	Score	Focus Molecules	Top Functions
1	Alpha catenin, B3GNT5, BCL3, CARD9, Cytokeratin, DGCR6L, Elastase, Estrogen Receptor, FAIM, FBN2 (includes EG:14119), GFI1, IL11RA, Kallikrein, KLK6, KPNA4, KRT1, KRT4, KRT15, KRT6A, LGALS7/LGALS7B, LIMD1, LOX, MEFV, MMP12, N-cor, NFkB (complex), Nuclear factor 1, RFTN1, RUNX1T1, SERPINA3, SERPINA4, TFAP2C, TNIP1, Trypsin, TXNDC11	44	26	Cell-To-Cell Signaling and Interaction, Connective Tissue Development and Function, Cancer
2	Actin, Akt, Ampa Receptor, APPL1, APPL2, CDH5, CLDN4, CNKSR2, Collagen type IV, DLGAP2, EPHA2, EPHB1, Fibrin, Fibrinogen, GEM, GFAP, GNAZ, GRIA3, INPP5E, Integrin, IP6K3, Laminin, MMP26, MTA2, NDUFS8, NEDD9, PCK1 (includes EG:18534), Pde, PDE11A, PHLDB1, PTGS1, Rap1, RELN, SERPINE1, SH2D3C	37	25	Cellular Movement, Cell Morphology, Cell-To-Cell Signaling and Interaction
3	ADCY, ADRA1D, ASGR1, Beta Arrestin, CCR6, CCR7, CSDA, DARC, EDNRA, EREG, ERK, FCGR2A, FSH, Gpcr, GPR39, GPR56, GPR84, GPR126, GPR162, GRM4, HCAR1, hCG, Lh, NMUR1, NPBWR2, NPW, Pak, PDLIM3, PLC, Proinsulin, PTH2, Sfk, SSTR1, SSTR2, STAT	37	23	Cell Signaling, Nucleic Acid Metabolism, Small Molecule Biochemistry
4	ARID3A, BBC3, Cbp/p300, Cdc2, CLSTN2, CNTN2, Collagen type I, Creb, CSDE1, Cyclin A, E2F2, E2f, ERCC6, FAT1, FHL2, HDAC1, Hdac, HISTONE, Histone h4, HOMER3, Jnk, KLC1, MAP3K8, MYOD1, NDN, POLD1, PRKCH, Rb, RYBP, Smad2/3-Smad4, SNAI3, SOX15, SP100, Thymidine Kinase, TNXB	33	22	Gene Expression, Cell Cycle, Cancer
5	ATP2A1, ATP2B3, ATP2B4, BCR, Ca2 ATPase, Cadherin, Calcineurin A, Calcineurin protein(s), CALD1, CaMKII, CD79B, ERK1/2, Fcer1, FES, KCNK18, KL, KLF6, LCK, MARCKSL1, MYH7, MYO1A, Myosin, NCK, NFAT (complex), Nfat (family), NFATC3, NPR1 (includes EG:18160), Pmca, RAB11FIP1, RAB4A, SCRIB, SIGLEC7/SIGLEC9, SYK/ZAP, TEAD1, UCN2	27	21	Cardiac Hypertrophy, Cardiovascular Disease, Developmental Disorder

Supplementary Table 3G. T-ALL v. B-ALL

ID	Molecules in Network	Score	Focus Molecules	Top Functions
1	ANK2, ARF1, ATL2, CD300C, CNKSR2, CNP, COP I, COPB1, COPG2, CYTIP, DLGAP2, DLK1, EMILIN1, ENaC, ERK1/2, Fcer1, FXYD2, GUCY, GUSB, H19, KLK4, LAT2, MDM4, MFN2, Na+, K+ -ATPase, NPR1 (includes EG:18160), PTPase, RAB11FIP1, SIK1, SNX27, SPRED2, TRPM2, TYROBP, UBASH3A, VRK2	40	28	Molecular Transport, Cellular Function and Maintenance, Cellular Assembly and Organization
2	AEBP1, ALT, AQP5, B3GNT5, CBR3, CENPJ, CFH, DMBT1, Ecm, ECM1 (includes EG:100332249), FAIM, Filamin, FOXJ1, GOT, HFE, HIVEP3, LARGE, MAP4K2, NFkB (complex), PARD6B, PELI3, peptidase, PPM1D, PRF1, PTP4A2, RCE1 (includes EG:19671), RFTN1, RIOK3, SRGN, Tnf receptor, TNFAIP8, TNFRSF9, TNIP1, TRAF1, TRAF	38	27	Developmental Disorder, Genetic Disorder, Hematological Disease
3	26s Proteasome, ADAM23, ATN1, Calpain, CAPN10, CASP12, Caspase, CSGALNACT1, DAXX, DNAJB6, DNAJC10, DNAJC19, Focal adhesion kinase, GAL, GIGYF1, Hsp70, Hsp90, Hsp22/Hsp40/Hsp90, HSPA1A/HSPA1B, IL2RA, MOV10, NAA10, PCNP, PPFIBP2, PRNP, SFN, THRA, TOP2B, UBE2, UBE2E3, UBE2J1, UBR4, UHRF2, ZMYND8, ZNRF4	38	27	Cell Death, Post-Translational Modification, Protein Folding
4	Akt, ALPP/ALPPL2, Ap1 gamma, Ap2 alpha, APPL1, APPL2, ASAP1, ASAP2, BIN1, CHCHD2, Clathrin, CLINT1, DNAH1, DNAH12, DNM3, Dynamin, Endophilin, GPHN, HSPA8, Inpp5, INPP5E, INPPL1, ITSN2, KIF23, MICALL1, MTM1, NUMB, P2RX7, PACSIN3, PFN2, Profilin, sGC, SYNJ2, SYNRG, WIPF1	35	26	Cellular Assembly and Organization, Cellular Movement, Nervous System Development and Function
5	ACSL3, ACSL5, ACSL, AMOTL2, ARID3A, ARID3B, CCT8, DAP, EGFL7, EGLN2, EIF1AY, ELF5, ELK3, ENG, ERK, ETS, GDF11, growth factor receptor, Il12 receptor, JAK, Lfa-1, LOXL1, LTBP3, MAL, PLEC, PTPRE, PURB, PXK, RANBP9, RGMA, SCAP, SH2B3, SLC27A3, Smad2/3, Sos	35	26	Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry

Supplementary Table 4

Complete listing of correlation data between differentially methylated probe sets and gene expression probe sets in Refseq centric analysis. Analyses of significance of correlation of methylation and expression are provided in Supplementary Table 5 (Fisher exact tests) and Supplementary Table 6 (Gene Set Enrichment Analysis)

See separate Excel workbook.

Supplementary Table 5

Fisher exact test results for correlation between differential expression and methylation.

Supplementary Table 5A shows correlation between downregulated and hypermethylated genes. The first 8 data rows of the table represent data for genes downregulated in each subgroup of ALL (e.g. CRLF2, ETV6-RUNX1 etc) compared to the comparison group (e.g. non-CRLF2r B-ALL, or non-ETV6-RUNX1 ALL). The second 8 rows show data for genes downregulated in the comparison group compared to each specific subtype. Significant Fisher P values are highlighted in pink

Downregulated, hypermethylated	(A) Downregulated FDR<0.1 and hypermethylated FDR<0.1	(B) Downregulated FDR>0.1 and hypermethylated FDR<0.1	(C) Downregulated FDR<0.1 and hypermethylated FDR>0.1	(D) Downregulated FDR>0.1 and hypermethylated FDR>0.1	Fisher P	Fraction of all hypermethylated genes that are downregulated (A/A+B)	Fraction of all downregulated genes that are hypermethylated (A/A+C)
B_CRLF2 v Non_B_CRLF2	0	52	17	12855	1	0.00	0.00
B_ETV6 v Non_B_ETV6	232	1336	1563	9793	0.28	0.15	0.13
B_H50 v Non_B_H50	1363	5146	968	5447	4.63E-18	0.21	0.58
B_MLLr v Non_B_MLLr	276	1656	881	10111	5.86E-17	0.14	0.24
B_ERGdel v Non_B_ERGdel	201	1204	1032	10487	1.10E-09	0.14	0.16
B_Ph v Non_B_Ph	163	1720	1001	10040	0.60	0.09	0.14
B_TCF3 v Non_B_TCF3	87	175	1999	10663	6.25E-12	0.33	0.04
T_all v B_all	629	2190	1248	8857	4.89E-37	0.22	0.34
Non_B_CRLF2 v B_CRLF2	1	57	14	12852	0.07	0.02	0.07
Non_B_ETV6 v B_ETV6	614	2432	1366	8512	1.53E-16	0.20	0.31
Non_B_H50 v B_H50	433	2383	1446	8662	0.16	0.15	0.23
Non_B_MLLr v B_MLLr	474	2957	1117	8376	0.002	0.14	0.30
Non_B_ERGdel v B_ERGdel	161	1140	1119	10504	0.002	0.12	0.13
Non_B_Ph v B_Ph	315	2408	791	9410	9.83E-10	0.12	0.28
Non_B_TCF3 v B_TCF3	25	162	1364	11373	0.23	0.13	0.02
B_all v T_all	1102	3779	1451	6592	4.78E-10	0.23	0.43

Supplementary Table 5B shows correlation between upregulated and hypermethylated genes. The first 8 data rows of the table represent data for genes downregulated in each subgroup of ALL (e.g. CRLF2, ETV6-RUNX1 etc) compared to the comparison group (e.g. non-CRLF2r B-ALL, or non-ETV6-RUNX1 ALL). The second 8 rows show data for genes upregulated in the comparison group compared to each specific subtype. Significant Fisher P values are highlighted in pink.

Upregulated, hypermethylated	(A) Upregulated FDR<0.1 and hypermethylated FDR<0.1	(B) Upregulated FDR>0.1 and hypermethylated FDR<0.1	(C) Upregulated FDR<0.1 and hypermethylated FDR>0.1	(D) Upregulated FDR>0.1 and hypermethylated FDR>0.1	Fisher P	Fraction of all hypermethylated genes that are upregulated (A/A+B)	Fraction of all upregulated genes that are hypermethylated (A/A+C)
B_CRLF2 v Non_B_CRLF2	0	52	15	12857	1	0.00	0.00
B_ETV6 v Non_B_ETV6	209	1359	1771	9585	0.02	0.13	0.11
B_H50 v Non_B_H50	817	5692	1062	5353	1.14E-10	0.13	0.43
B_MLLr v Non_B_MLLr	222	1710	1369	9623	0.24	0.11	0.14
B_ERGdel v Non_B_ERGdel	130	1275	1150	10369	0.42	0.09	0.10
B_Ph v Non_B_Ph	141	1742	965	10076	0.07	0.07	0.13
B_TCF3 v Non_B_TCF3	28	234	1361	11301	1	0.11	0.02
T_all v B_all	490	2329	2063	8042	0.003	0.17	0.19
Non_B_CRLF2 v B_CRLF2	0	58	17	12849	1	0.00	0.00
Non_B_ETV6 v B_ETV6	397	2649	1398	8480	0.12	0.13	0.22
Non_B_H50 v B_H50	455	2361	1876	8232	0.003	0.16	0.20
Non_B_MLLr v B_MLLr	291	3140	866	8627	0.26	0.08	0.25
Non_B_ERGdel v B_ERGdel	116	1185	1117	10506	0.46	0.09	0.09
Non_B_Ph v B_Ph	243	2480	921	9280	0.88	0.09	0.21
Non_B_TCF3 v B_TCF3	37	150	2049	10688	0.19	0.20	0.02
B_all v T_all	530	4351	1347	6696	8.81E-21	0.11	0.28

Supplementary Table 6. Summary of gene set enrichment analyses examining enrichment of hypo- and hypermethylated gene sets in gene expression signatures. G1 refers to the specific ALL subtype, G2 to the comparison group

Group comparison	Enrichment of G1 hypomethylated gene set in G1 vs. G2 gene signature	Enrichment of G1 hypermethylated gene set in G2 vs. G1 gene signature
*CRLF2*r vs non-*CRLF2*r B-ALL	No (FDR=0.89)	No (FDR 1.0)
ETV6-RUNX1 vs non-*ETV6-RUNX1* B-ALL	Yes (P<0.0001; FDR=0.18)	No (FDR 0.85)
High hyperdiploid vs non high hyperdiploid B-ALL	No (FDR=0.78)	No (FDR 0.96)
*MLL*r vs non-*MLL*r B-ALL	No (FDR=0.53)	Yes (P<0.0001, FDR 0.0016)
ERG B-ALL vs non-*ERG* B-ALL	No (FDR=0.65)	Yes (P<0.0001, FDR=0.07)
BCR-ABL1 B-ALL vs non-*BCR-ABL1* B-ALL	Yes (P<0.0001, FDR=0.0064)	No (FDR 0.87)
TCF3-PBX1 vs non-*TCF3-PBX1* B-ALL	Yes (P=0.003, FDR=0.11)	No (FDR 0.71)
T-ALL vs B-ALL	Yes (P<0.0001, FDR=0.013)	Yes (P<0.0001, FDR=0.00097)

Supplementary Table 7. Relationship between gene methylation and expression by chromosome in Hyperdiploid samples (provided as separate excel workbook).

Supplementary Table 8. Overlap of normal T- v B-cell and leukemic T- v B-ALL methylation signatures

Comparison	Signature 1 Probe Number	Signature 2 Probe Number	Overlap Probe Number	Signature 1 Residual Probe Number	Signature 2 Residual Probe Number	Signature 1 Residual/Probe Number	Overlap Ratio
Subtract Hyper signatures [Normal T v B] (FDR=0.1, Delta1) from [T- v B-ALL] (FDR=0.1, Delta1)	897	2463	413	484	2050	54.0%	39.7%
Subtract Hypo signatures [Normal T v B] (FDR=0.1, Delta1) from [T- v B-ALL] (FDR=0.1, Delta1)	422	1403	110	312	1293	73.9%	
Subtract Hyper signatures [Normal T v B] (FDR=1, Delta1) from [T- v B-ALL] (FDR=1, Delta1)	897	2545	413	484	2132	54.0%	39.7%
Subtract Hyper signatures [Normal T v B] (FDR=1, Delta1) from [T- v B-ALL] (FDR=1, Delta1)	422	1403	110	312	1293	73.9%	
Subtract Hyper signatures [Normal T v B] (FDR=1, Delta0.5) from [T- v B-ALL] (FDR=1, Delta0.5)	2214	5728	1086	1128	4642	50.9%	38.9%
Subtract Hyper signatures [Normal T v B] (FDR=1, Delta0.5) from [T- v B-ALL] (FDR=1, Delta0.5)	1923	6354	522	1401	5832	72.9%	
Subtract Hyper signatures [Normal T v B] (FDR=1, Delta2) from [T- v B-ALL] (FDR=1, Delta2)	181	345	59	122	286	67.4%	31.3%
Subtract Hyper signatures [Normal T v B] (FDR=1, Delta2) from [T- v B-ALL] (FDR=1, Delta2)	46	175	12	34	163	73.9%	

Supplementary Table 9. Genes in the common methylation signature of ALL

Methylation probeset	Chr	Start	End	Gene	mRNA Refseq ID	Status
MSPI0406S00033551	chr1	43765535	43765963	TIE1	NM_005424	Hypermethylated
MSPI0406S00047441	chr1	99729433	99730057	LPPR4	NM_001166252 / NM_014839	Hypermethylated
MSPI0406S00054375	chr1	150209809	150210213	ANP32E	NM_001136478 /NM_001136479 /NM_030920	Hypermethylated
MSPI0406S00114368	chr2	74758140	74758991	DQX1	NM_133637	Hypomethylated
MSPI0406S00114368	chr2	74758140	74758991	AUP1	NM_181575	Hypomethylated
MSPI0406S00114368	chr2	74758140	74758991	HTRA2	NM_013247 /NM_145074	Hypomethylated
MSPI0406S00136658	chr2	155554740	155554951	KCNJ3	NM_002239	Hypomethylated
MSPI0406S00163586	chr3	6902402	6902834	GRM7	NM_000844 / NM_181874	Hypermethylated
MSPI0406S00189293	chr3	101442882	101443442	CEP97	NM_024548	Hypermethylated
MSPI0406S00206872	chr3	180706718	180707350	DNAJC19	NM_145261	Hypermethylated
MSPI0406S00209099	chr3	185303158	185303485	SENP2	NM_021627	Hypermethylated
MSPI0406S00209527	chr3	186080357	186081059	DGKG	NM_001080744 / NM_001080745 / NM_001346	Hypermethylated
MSPI0406S00276902	chr5	89704526	89705315	CETN3	NM_004365	Hypermethylated
MSPI0406S00284502	chr5	134073213	134073675	CAMLG	NM_001745	Hypermethylated
MSPI0406S00287235	chr5	140242976	140243299	PCDHA11	NM_018902 /NM_031861	Hypermethylated
MSPI0406S00287329	chr5	140535915	140536611	PCDHB17	NR_001280	Hypermethylated
MSPI0406S00287350	chr5	140579044	140579962	PCDHB11	NM_018931	Hypermethylated
MSPI0406S00287351	chr5	140579962	140580502	PCDHB11	NM_018931	Hypermethylated
MSPI0406S00287436	chr5	140744981	140745428	PCDHGA5	NM_018918 / NM_032054	Hypermethylated
MSPI0406S00287436	chr5	140744981	140745428	PCDHGB3	NM_018924 / NM_032097	Hypermethylated
MSPI0406S00287891	chr5	141392900	141393794	GNPDA1	NM_005471	Hypomethylated
MSPI0406S00300233	chr5	177423738	177423997	PROP1	NM_006261	Hypermethylated
MSPI0406S00312956	chr6	25138315	25139939	CMAH	NR_002174 / NR_027626	Hypermethylated
MSPI0406S00313668	chr6	27101671	27102483	HIST1H2BJ	NM_021058	Hypomethylated
MSPI0406S00313668	chr6	27101671	27102483	HIST1H2AG	NM_021064	Hypomethylated
MSPI0406S00313668	chr6	27101671	27102483	HIST1H4I	NM_003495	Hypomethylated
MSPI0406S00327661	chr6	71376647	71377392	SMAP1	NM_001044305 / NM_021940	Hypermethylated
MSPI0406S00331617	chr6	96463499	96463874	FUT9	NM_006581	Hypermethylated
MSPI0406S00333064	chr6	106773679	106774170	ATG5	NM_004849	Hypermethylated
MSPI0406S00346373	chr6	160680891	160681505	SLC22A2	NM_003058	Hypermethylated
MSPI0406S00363123	chr7	27183696	27184157	HOXA5	NM_019102	Hypermethylated
MSPI0406S00363123	chr7	27183696	27184157	HOXA6	NM_024014	Hypermethylated
MSPI0406S00372326	chr7	56174388	56175223	CHCHD2	NM_016139	Hypermethylated
MSPI0406S00396886	chr7	134330441	134330836	BPGM	NM_001724 / NM_199186	Hypermethylated
MSPI0406S00425232	chr8	41434895	41435387	AGPAT6	NM_178819	Hypermethylated
MSPI0406S00429517	chr8	57026732	57027299	MOS	NM_005372	Hypermethylated
MSPI0406S00429579	chr8	57233598	57234355	SDR16C5	NM_138969	Hypermethylated
MSPI0406S00441641	chr8	124084236	124084728	WDR67	NM_001145088 / NM_145647	Hypermethylated
MSPI0406S00441831	chr8	124428126	124428638	WDYHV1	NM_018024	Hypermethylated
MSPI0406S00456924	chr9	23822613	23823089	ELAVL2	NM_004432	Hypermethylated
MSPI0406S00458933	chr9	34459232	34459602	C9orf25	NM_147202	Hypermethylated
MSPI0406S00458933	chr9	34459232	34459602	DNAI1	NM_012144	Hypermethylated
MSPI0406S00460735	chr9	37904664	37905009	MCART1	NR_024872 / NR_024873 / NM_033412	Hypermethylated
MSPI0406S00501433	chr10	7859702	7860348	TAF3	NM_031923	Hypermethylated
MSPI0406S00519670	chr10	70747999	70748453	KIAA1279	NM_015634	Hypermethylated
MSPI0406S00520388	chr10	71811620	71812475	H2AFY2	NM_018649	Hypermethylated
MSPI0406S00531953	chr10	102289691	102290274	NDUFB8	NM_005004	Hypermethylated
MSPI0406S00533558	chr10	104535364	104536033	C10orf26	NM_017787	Hypermethylated
MSPI0406S00540894	chr10	125425209	125425606	GPR26	NM_153442	Hypermethylated
MSPI0406S00558409	chr11	27384912	27385612	CCDC34	NM_030771 / NM_080654	Hypermethylated
MSPI0406S00560067	chr11	34533292	34534291	ELF5	NM_001422 / NM_198381	Hypomethylated
MSPI0406S00561064	chr11	40315539	40316098	LRRC4C	NM_020929	Hypermethylated

Methylation probeset	**Chr**	**Start**	**End**	**Gene**	**mRNA Refseq ID**	**Status**
MSPI0406S00586694	chr11	113928943	113929741	ZBTB16	NM_006006 /NM_001018011	Hypermethylated
MSPI0406S00592000	chr11	124542208	124542682	SIAE	NM_170601	Hypermethylated
MSPI0406S00592000	chr11	124542208	124542682	SPA17	NM_017425	Hypermethylated
MSPI0406S00606629	chr12	29935450	29936014	TMTC1	NM_175861	Hypermethylated
MSPI0406S00612910	chr12	52462629	52463347	C12orf44	NM_001098673 /NM_021934	Hypermethylated
MSPI0406S00644381	chr13	21751219	21751781	SKA3	NM_001166017 /NM_145061	Hypermethylated
MSPI0406S00644381	chr13	21751219	21751781	MRP63	NM_024026	Hypermethylated
MSPI0406S00653659	chr13	52378546	52379111	DHRS12	NM_024705 / NM_001031719	Hypermethylated
MSPI0406S00668515	chr14	23235280	23235901	OXA1L	NM_005015	Hypermethylated
MSPI0406S00668987	chr14	23789511	23790114	PABPN1	NM_004643	Hypermethylated
MSPI0406S00680432	chr14	70656074	70656983	SLC8A3	NM_033262 / NM_058240 / NM_182932 / NM_183002	Hypermethylated
MSPI0406S00704116	chr15	32933864	32934484	SCG5	NM_001144757 / NM_003020	Hypermethylated
MSPI0406S00755815	chr16	31482830	31483291	TGFB1I1	NM_001042454 / NM_001164719 / NM_015927	Hypermethylated
MSPI0406S00762508	chr16	58664013	58664819	CNOT1	NM_016284 / NM_206999	Hypermethylated
MSPI0406S00785511	chr17	4400482	4400882	SPNS2	NM_001124758	Hypomethylated
MSPI0406S00803475	chr17	35765935	35766548	ACACA	NM_198834 / NM_198839	Hypermethylated
MSPI0406S00803475	chr17	35765935	35766548	TADA2A	NM_001166105 / NM_001488 / NM_133439	Hypermethylated
MSPI0406S00808861	chr17	42144066	42144522	LSM12	NM_152344	Hypermethylated
MSPI0406S00808861	chr17	42144066	42144522	G6PC3	NM_138387 / NR_028581 / NR_028582	Hypermethylated
MSPI0406S00824789	chr17	72425973	72426791	GPRC5C	NM_022036 / NM_018653	Hypomethylated
MSPI0406S00844006	chr18	21978063	21978580	OSBPL1A	NM_080597	Hypermethylated
MSPI0406S00845949	chr18	33077934	33078448	INO80C	NM_001098817 / NM_194281	Hypermethylated
MSPI0406S00870600	chr19	9609540	9610004	ZNF560	NM_152476	Hypermethylated
MSPI0406S00908309	chr19	56905050	56905678	ZNF582	NM_144690	Hypermethylated
MSPI0406S00911634	chr20	3189309	3189933	DDRGK1	NM_023935	Hypermethylated
MSPI0406S00911634	chr20	3189309	3189933	ITPA	NM_033453 / NM_181493	Hypermethylated
MSPI0406S00925444	chr20	41819880	41820377	PTPRT	NM_007050 / NM_133170	Hypermethylated
MSPI0406S00939726	chr20	62507539	62508139	OK/SW-cl.69	uc002yhe.2	Hypomethylated
MSPI0406S00990104	chrX	20135218	20135929	MAP7D2	NM_152780	Hypermethylated
MSPI0406S00998736	chrX	51485701	51486446	GSPT2	NM_018094	Hypermethylated
MSPI0406S01006063	chrX	90690097	90690732	PABPC5	NM_080832	Hypermethylated
MSPI0406S01011804	chrX	119737171	119737856	MCTS1	NM_014060 / NM_001137554	Hypermethylated
MSPI0406S01014522	chrX	133680635	133681243	MGC16121	NR_024607	Hypermethylated

Supplementary Table 10: **Genes with recurrent genetic lesions in childhood ALL**

ABL1
ADAR
ADARB2
ADD3
ANL1
ARHGAP24
ARMC2
ARPP_21
ATM
ATP10A
ATP5O
BLNK
BTG1
BTLA
C13ORF21
C20orf94
CCDC26
CCNC
CDK6
CDKN2AB
CREBBP
CRLF2
CTCF
DLEU
DMD
DPF3
EBF1
ELF1
ERG
ETV6
FBXW7
FHIT
FLNB
GAB1
GRLF1
HBS1L
HIST1H2BE
IKZF1
IKZF3
IL1RAP
IL3RA
JAK1/2
KLF6
KRAS
LEF1
LMO2
LTK
MBNL1
MEF2C
MLLT3
MSRA
MYB
NF1
NR3C1
NR3C2
NRAS
NUP160
PAX5
PDE4B
PECAM1
PIK3AP1
PLCB3
PTEN
PTPN2
RAG
RB1
RUNX1
SPRED1
STIL
TBL1XR1
TCF3
THADA
TOX
VPREB1
X19PTEL
ZMYM5

Supplementary Table 11: MassARRAY primers

See separate excel workbook.