

## Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Mullighan CG, Su X, Zhang J, et al. Deletion of *IKZF1* and prognosis in acute lymphoblastic leukemia. *N Engl J Med* 2009;360. DOI: 10.1056/NEJMoa0808253.

## SUPPLEMENTARY APPENDIX

for

### DELETION OF *IKZF1* AND PROGNOSIS IN ACUTE LYMPHOBLASTIC LEUKEMIA

Charles G. Mullighan<sup>1</sup>, Xiaoping Su<sup>1</sup>, Jinghui Zhang<sup>8</sup>, Ina Radtke<sup>1</sup>, Letha A. A. Phillips<sup>1</sup>, Christopher B. Miller<sup>1</sup>, Jing Ma<sup>3</sup>, Wei Liu<sup>4</sup>, Cheng Cheng<sup>4</sup>, Brenda A. Schulman<sup>5,7</sup>, Richard C. Harvey<sup>12</sup>, I-Ming Chen<sup>12</sup>, Robert J. Clifford<sup>9</sup>, William L. Carroll<sup>13</sup>, Gregory Reaman<sup>14</sup>, W. Paul Bowman<sup>15</sup>, Meenakshi Devidas<sup>16</sup>, Daniela S. Gerhard<sup>10</sup>, Wenjian Yang<sup>2</sup>, Mary V. Relling<sup>2</sup>, Sheila A. Shurtleff<sup>1</sup>, Dario Campana<sup>1,6</sup>, Michael J. Borowitz<sup>17</sup>, Ching-Hon Pui<sup>6</sup>, Malcolm Smith<sup>11</sup>, Stephen P. Hunger<sup>18</sup>, Cheryl L. Willman<sup>12</sup>, James R. Downing<sup>1</sup>, and the Children's Oncology Group

<sup>1</sup>Pathology, <sup>2</sup>Pharmaceutical Sciences, <sup>3</sup>The Hartwell Center for Bioinformatics and Biotechnology, <sup>4</sup>Biostatistics, <sup>5</sup>Structural Biology and <sup>6</sup>Oncology, St Jude Children's Research Hospital, Memphis, TN; <sup>7</sup>Howard Hughes Medical Institute; <sup>8</sup>The Center for Biomedical Informatics and Information Technology, <sup>9</sup>Laboratory of Population Genetics; <sup>10</sup>Office of Cancer Genomics and <sup>11</sup>Cancer Therapy Evaluation Program, National Cancer Institute, National Institutes of Health; <sup>12</sup>University of New Mexico Cancer Research and Treatment Center, Albuquerque, NM; <sup>13</sup>New York University Cancer Institute, New York, NY; <sup>14</sup>Children's National Medical Center, Washington, DC; <sup>15</sup>Hematology Oncology, Cook Children's Medical Center, Fort Worth, TX; <sup>16</sup>Epidemiology and Health Policy Research, College of Medicine, University of Florida, Gainesville, FL; <sup>17</sup>Pathology, Johns Hopkins Medical Institute, Baltimore, MD; and <sup>18</sup>The University of Colorado Denver School of Medicine and The Children's Hospital, Aurora, CO; and the Children's Oncology Group.

## CONTENTS

<b>SUPPLEMENTARY METHODS</b> .....	<b>4</b>
Patients and samples.....	4
Single nucleotide polymorphism microarray analyses.....	5
<i>PAX5</i> , <i>IKZF1</i> and <i>EBF1</i> resequencing and mutation detection.....	6
Structural modeling of <i>PAX5</i> mutations .....	7
Analysis of associations between DNA copy number abnormalities and outcome .....	7
Gene expression profiling of high risk ALL .....	9
Genomic data access .....	10
<b>SUPPLEMENTARY RESULTS</b> .....	<b>23</b>
<b>REFERENCES</b> .....	<b>58</b>

## SUPPLEMENTARY TABLES

Supplementary Table 1. Samples studied from the Children’s Oncology Group P9906 cohort ...	12
Supplementary Table 2. 258 St Jude B-progenitor ALL cases examined. ....	17
Supplementary Table 3. DNA copy number abnormality frequency in high-risk pediatric ALL.....	23
Supplementary Table 4. Regions of recurring copy number alteration in the P9906 cohort.....	24
Supplementary Table 5. Regions of recurring copy number alteration in the St Jude cohort. ....	27
Supplementary Table 6. Results of <i>PAX5</i> genomic quantitative PCR.....	29
Supplementary Table 7. <i>PAX5</i> sequence mutations in the P9906 cohort. ....	30
Supplementary Table 8. Description all P9906 cases harboring <i>IKZF1</i> deletions and/or sequence mutations, and results of genomic quantitative PCR results for <i>IKZF1</i> deletions. ....	31
Supplementary Table 9. Description of B-cell pathway lesions observed in the P9906 cohort. ....	33
Supplementary Table 10. Variation in number of B cell pathway lesions between P9906 ALL subtypes.....	37
Supplementary Table 11. Genes with univariate Cox score exceeding threshold of $\pm 1.8$ in SPC analysis, P9906 cohort. ....	38
Supplementary Table 12. Genes with univariate Cox score exceeding threshold of $\pm 1.9$ in SPC analysis, St Jude cohort.....	39
Supplementary Table 13. Associations between of B cell pathway lesions, <i>IKZF1</i> alterations and hematologic relapse, P9906 cohort .....	40
Supplementary Table 14. Kaplan-Meier estimates of EFS by B cell pathway or <i>IKZF1</i> lesions, P9906 cohort.....	41

Supplementary Table 15. Hazard ratio estimates of B-cell pathway and <i>IKZF1</i> lesions on relapse and event free survival, P9906 cohort. ....	42
Supplementary Table 16. Associations between genomic abnormalities and day 8 MRD, P9906 cohort. ....	43
Supplementary Table 17. Associations between genetic lesions and day 29 MRD, P9906 cohort .....	44
Supplementary Table 18. Association of genetic lesions with day 29 MRD adjusted by age, presentation leukocyte count and genetic subtype in the P9906 cohort. ....	45
Supplementary Table 19. Multivariable analysis of associations between <i>IKZF1</i> deletion/mutation and outcome in the P9906 cohort adjusting for age, presentation leukocyte count, leukemia subtype, and day 8 or day 29 MRD level .....	46
Supplementary Table 20. Cumulative incidence of isolated or combined hematologic relapse by genetic lesions, St Jude cohort with MRD data. ....	47
Supplementary Table 21. Cumulative incidence of hematologic relapse by genetic lesions, St Jude cohort with MRD data, excluding BCR-ABL1 positive ALL.....	48
Supplementary Table 22. Associations of genetic lesions and day 19 MRD, St Jude cohort (all B-progenitor ALL cases with MRD data) .....	49
Supplementary Table 23. Associations of genetic lesions and day 46 MRD, St Jude cohort (all B-progenitor ALL cases with MRD data) .....	50
Supplementary Table 24. Associations of genetic lesions and day 19 MRD, St Jude cohort (B-progenitor ALL cases, excluding <i>BCR-ABL1</i> ALL) .....	51
Supplementary Table 25. Associations of genetic lesions and day 46 MRD, St Jude cohort (B-progenitor ALL cases, excluding <i>BCR-ABL1</i> ALL) .....	51
Supplementary Table 26. Genes driving positive enrichment of hematopoietic stem/progenitor gene set in P9906 high-risk ALL.....	52
Supplementary Table 27. Genes driving negative enrichment of the B-cell signal transduction gene set in P9906 high-risk ALL. ....	53

## **SUPPLEMENTARY FIGURES**

Supplementary Figure 1. <i>PAX5</i> deletions in P9906 ALL. ....	54
Supplementary Figure 2. <i>IKZF1</i> deletions in P9906 ALL.....	55
Supplementary Figure 3. Gene set enrichment analysis (GSEA) of poor outcome P9906 ALL, poor outcome St Jude ALL, and <i>BCR-ABL1</i> positive St Jude ALL. ....	56

## SUPPLEMENTARY METHODS

### *Patients and samples*

The P9906 cohort comprised 221 B-progenitor ALL cases treated on the Children's Oncology Group P9906 study with an augmented intensive regimen of post-induction chemotherapy<sup>1</sup> (Supplementary Table 1). All patients were high risk based on the presence of central nervous system or testicular disease, *MLL* rearrangement, or based on age, gender and presentation leukocyte count<sup>2</sup>. *BCR-ABL1* and hypodiploid ALL, and cases of primary induction failure were excluded. Hyperdiploid (as defined by trisomy of chromosomes 4 and 10 on cytogenetic analysis) and *ETV6-RUNX1* cases were excluded unless CNS or testicular involvement was present at diagnosis. Of 276 cases enrolled, 271 were eligible, and 221 had suitable material for genomic analysis. Twenty-five (11.3%) cases were *TCF3-PBX1* positive, 19 harbored *MLL*-rearrangements, four were hyperdiploid, and three were *ETV6-RUNX1* positive. One hundred seventy (76.9%) lacked a recurring chromosomal abnormality. The clinical protocol was approved by the National Cancer Institute and by the Institutional Review Board at each of the Children's Oncology Group institutions. Patients and/or a parent/guardian provided written informed consent to participate in the clinical trial and for future research using clinical specimens. The validation cohort comprised 258 B-progenitor ALL cases treated at St Jude Children's Research Hospital<sup>3,4</sup>, and included 44 high hyperdiploid (greater than 50 chromosomes), 10 hypodiploid, 17 *TCF3-PBX1* positive, 50 *ETV6-RUNX1* positive, 21 *BCR-ABL1* positive and 24 *MLL* rearranged B-progenitor ALL cases, and 92 cases with low hyperdiploid, pseudodiploid, normal or miscellaneous karyotypes. These cases were treated on St Jude Total XI (N=8), XII (N=13), XIII (N=105), XIV (N=4), XV (N=114) and Interfant-99 (infant; N=5) protocols<sup>5-9</sup>. Nine cases were treated off protocol. The study was approved by the

St Jude Children's Research Hospital Institutional Review Board, and informed consent was obtained from patients and/or parents/guardians.

### ***Single nucleotide polymorphism microarray analyses***

Cases in the P9906 cohort were genotyped using 250K Sty and Nsp arrays (Affymetrix, Santa Clara, CA), which together examine over 500,000 genomic loci. Thirty-six cases from the St Jude cohort were genotyped using SNP 6.0 arrays (Affymetrix) which examine over 1.87 million loci; 37 with 250K Sty and Nsp arrays, and 185 with both 250K and two 50K arrays that together examine over 615,000 markers. SNP array data preprocessing and inference of DNA copy number abnormalities (CNA) and loss-of-heterozygosity (LOH) was performed as previously described<sup>3,4</sup>. Briefly, SNP calls were generated using the DM or Birdseed algorithms in GTYPE 4.0 or Genotyping Console (Affymetrix). Summarization of probe level data was performed using the PM/MM (50K and 250K arrays) or PM-only (SNP 6.0 arrays) model-based expression algorithms in dChip ([www.dchip.org](http://www.dchip.org))<sup>10</sup>. Normalization of array signals was performed using a reference normalization algorithm that utilizes only those SNP probes from diploid regions of each array to guide normalization<sup>3,11</sup>. To identify all tumor-acquired regions of CNA for each sample, circular binary segmentation (CBS)<sup>12</sup> (implemented as the DNACopy package in R) was performed by directly comparing each tumor sample to the corresponding remission sample.

To ensure data from the SNP 50k and 250k arrays (the annotations for which are based on the hg17 genome build) were comparable to data obtained from the SNP 6.0 arrays (based on the hg18 genome build), genome coordinates of the segments from the SNP 6.0 arrays were converted to the corresponding hg17 genome coordinates using the command line *liftOver* tool available from UCSC genome browser (<http://hgdownload.cse.ucsc.edu/admin/exe/>). Curated

CBS segments were then re-extrapolated into a genome-wide matrix of 615,000 probe-set features for all cases.

***PAX5, IKZF1 and EBF1 resequencing and mutation detection.***

Genomic resequencing of all coding exons of *PAX5*, *IKZF1* and *EBF1* was performed for 197 P9906 samples with available DNA and gene expression data by Agencourt Biosciences (Beverly, MA). Genomic DNA was amplified in 384 well plates, with each PCR reaction containing 10 ng DNA, 1X HotStar buffer, 0.8 mM dNTPs, 1 mM MgCl<sub>2</sub>, 0.2U HotStar enzyme (Qiagen) and 0.2 μM forward and reverse primers in 10 μl reaction volumes. PCR cycling parameters were: one cycle of 95°C for 15 min, 35 cycles of 95°C for 20s, 60°C for 30s and 72°C for 1 min, followed by one cycle of 72°C for 3 min. PCR products were purified using proprietary large scale automated template purification systems using solid-phase reversible immobilization, and then sequenced using dye-terminator chemistry and ABI 3700/3730 machines (Applied Biosystems, Foster City, CA). Base calls and quality scores were determined using the program PHRED<sup>13,14</sup>.

Sequence variations including substitutions and insertion/deletions (indel) were analyzed using the SNPdetector<sup>15</sup> and the IndelDetector<sup>16</sup> software. A useable read was required to have at least one 30-bp window in which 90% of the bases have PHRED quality score of at least 30. Poor quality reads were filtered prior to variation detection. The minimum threshold of secondary to primary peak ratio for substitution and indel detection was set to be 20% and 10%, respectively. All sequence variations were annotated using a previously developed variation annotation pipeline<sup>17</sup>. Any variation that did not match a known polymorphism (defined as a dbSNP record that does not belong to OMIM SNP nor COSMIC somatic variation database<sup>18,19</sup>) and resulted in a non-silent amino acid change was considered a putative mutation.

All putative sequence mutations were confirmed by repeat genomic PCR and sequencing of both tumor and remission DNA. Where possible, expression of mutated *PAX5* and *IKZF1* alleles was confirmed by amplification and direct sequencing of full length *PAX5* and *IKZF1* cDNA as previously described<sup>3,4</sup>. Transcripts were then cloned into pGEM-T-Easy (Promega, Madison, WI) and multiple colonies sequenced. Confirmation of CNA involving *PAX5* and *IKZF1* by genomic quantitative PCR was performed as previously described<sup>3,4</sup>.

### ***Structural modeling of PAX5 mutations***

Missense substitutions were generated in the *PAX5* (residues 1-149)/ETS-1(residues 331-440)/DNA structure<sup>20</sup> and subjected to local refinement using the program O<sup>21</sup>. Structural representation was performed with the program PyMOL (Delano Scientific)<sup>22</sup>.

### ***Analysis of associations between DNA copy number abnormalities and outcome***

Supervised principal components (SPC) analysis<sup>23,24</sup> was used to examine associations between CNA and outcome of therapy in a genome-wide fashion. This method has previously been used to examine associations between gene expression profiling data and outcome in cancer<sup>23</sup>. To implement the SPC method for analysis of CNA and outcome, regions of somatic CNA for each sample were transformed into a matrix in which each column represented an individual case, each row represented an individual gene, and each cell represented copy number status for each gene targeted by CNA in at least one case. Using the P9906 cohort as the training set, a modified univariate Cox score was calculated for the association between copy number status of each gene and event-free survival, and genes whose Cox score exceeded a threshold that best predicted survival were used to carry out supervised principal components analysis. To determine the Cox threshold, the training set was split and principal components were derived from one half of the samples, and then used in a Cox model to predict survival in the other half. By varying the

threshold of Cox scores and using twofold cross-validation, this process was repeated ten times, and a threshold of  $\pm 1.8$  (averaged over ten separate repeats of this procedure) was used to generate the principal components subsequently used to predict outcome.

For each case, we used the first principal component in a regression model to calculate a SPC risk score that represents the sum of the weighted copy number levels for each gene found to be significantly associated with prognosis. To validate the SPC predictor, we computed risk scores for each of the 258 cases in the St Jude validation cohort using the model developed in the P9906 training set, and tested whether these scores were correlated with survival. To illustrate the performance of the SPC risk score in predicting survival, cases in the validation cohort were classified as being high or low risk according to the calculated SPC risk score, and cumulative incidence of hematologic relapse and any relapse in each SPC risk group analyzed using Gray's test<sup>25</sup>. To examine the role of individual genes in determining outcome, we computed importance scores for genes with Cox scores exceeding the threshold defined by cross validation. The importance score is equivalent to the correlation between each gene and the first supervised principal component. Associations between genes with the top importance scores and hematologic and any relapse were then analyzed using Gray's test<sup>25</sup>. Event-free survival (EFS) was defined as the time from diagnosis until the date of failure (relapse, death, or second malignancy) or until the last follow-up date for all event-free survivors. Associations between genetic variables (deletions  $\pm$  sequence mutations of individual genes, presence and number B-cell pathway lesions) and EFS were estimated by the methods of Kaplan and Meier. Standard errors were calculated by the methods of Peto *et al.*<sup>26</sup>. The Mantel–Haenszel test was used to compare EFS estimates for patients with and without lesion at each locus<sup>27</sup>. The proportional hazards model of Fine and Gray was used to adjust for age, presentation leukocyte count,

cytogenetic subtype and levels of minimal residual disease (MRD)<sup>28</sup>. Analyses were performed using R (www.r-project.org)<sup>29</sup>, SAS (SAS v9.1.2, SAS Institute, Cary, NC) and SPLUS (SPLUS 7.0, Insightful Corp., Palo Alto, CA)

To evaluate associations between genetic alterations and MRD, MRD data was converted into an ordinal variable ( $<0.01\%$ ,  $0.01 \leq \text{MRD} < 1\%$ , and  $\geq 1\%$ ) and association analyses performed using the Chi-Square test (FREQ procedure, SAS) with estimation of false discovery rate (MULTTEST, SAS). Significantly associated variables were then adjusted for age, presentation leukocyte count and genetic subtype using logistic regression.

### ***Gene expression profiling of high risk ALL***

Gene expression profiling was performed using U133 Plus 2 microarrays (Affymetrix) for 197 P9906 samples, and using U133A microarrays (Affymetrix) for 175 St Jude samples. Probe intensities were generated using the MAS 5.0 algorithm, probe sets called absent in all samples in each cohort were excluded, and expression data log-transformed. To define the gene expression signature of poor outcome ALL in each cohort, we used *limma* (Linear Models for Microarray Analysis)<sup>30</sup>, the empirical Bayes t-test implemented in Bioconductor<sup>31</sup> and the Benjamini-Hochberg method of false discovery rate (FDR) estimation<sup>32</sup> to identify probe sets differentially expressed between cases defined as high or low risk according to their SPC risk score. This approach was also used to define the gene expression signature of *BCR-ABL1* positive *de novo* pediatric ALL in the St Jude cohort.

To assess similarity between the high-risk gene expression signatures of the P9906 and St Jude cohorts, and between the high-risk signatures and the signature of *BCR-ABL1* positive ALL, gene set enrichment analysis (GSEA)<sup>33</sup> and direct comparison of the signatures was performed.

Gene sets of the top up- and down-regulated genes in the signatures of high risk P9906 and St Jude ALL, and *BCR-ABL1* positive ALL were created and added to the collection of curated gene sets available at the Molecular Signatures Database (<http://www.broad.mit.edu/gsea/msigdb/>). GSEA of high risk ALL was then performed for each cohort using this expanded collection of gene sets. In a complementary approach, we determined the fraction of the top 100 differentially expressed probe sets in P9906 high-risk ALL that were also differentially expressed in St Jude *BCR-ABL1* positive ALL (at an FDR threshold of 5%). The Gene Set Analysis (GSA) algorithm, a modification of GSEA that allows testing of associations between gene sets and time-dependent variables such as survival time<sup>34</sup>, was used to examine associations between gene sets and EFS in the P9906 cohort.

### ***Genomic data access***

P9906 primary SNP and expression profile data, as well as the computed cancer-specific copy number alterations and sequence mutations, are available through the Childhood Cancer TARGET data portal (<http://target.cancer.gov/data/>). The TARGET Open-Access Data tier includes publicly accessible data that cannot be aggregated to generate a dataset unique to an individual using TARGET data and other publicly available data. The Open-Access Data tier does not require user certification for data access. This tier allows open access by researchers to all transcriptome data, mutations identified in the tumor tissue, as well as processed data describing cancer-specific copy number alterations and loss-of-heterozygosity described in detail below. The Controlled-Access Data tier contains broader clinical data and molecular signatures, which while stripped of direct identifiers, are individually unique. The Controlled-Access Data tier is for research projects related to target identification and therapeutics development for which the Open-Access data are not sufficient. Access to this tier requires user certification and

approval by the Data Access Committee responsible for the TARGET project, as described at the TARGET data portal site. P9906 gene expression data has also been deposited at the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO)<sup>35,36</sup> (accession GSE11877). St Jude gene expression data has been deposited at GEO (accession GSE 12995). St Jude primary SNP array data is available to academic researchers from the authors upon request. All P9906 sequencing traces and sequencing primer information have been deposited with NCBI's trace archive. (<http://www.ncbi.nlm.nih.gov/Traces/trace.cgi?>).

**Supplementary Table 1. Samples studied from the Children’s Oncology Group P9906 cohort**

<b>Sample ID</b>	<b>Group</b>	<b>U133 Plus 2 data</b>
9906_001		Yes
9906_002	TCF3-PBX1	Yes
9906_003	TCF3-PBX1	Yes
9906_004		Yes
9906_005		Yes
9906_006	MLL	Yes
9906_007		Yes
9906_008		Yes
9906_009		Yes
9906_010		
9906_011		
9906_012		Yes
9906_013		Yes
9906_014		
9906_016		
9906_017	TCF3-PBX1	Yes
9906_018		Yes
9906_019		Yes
9906_020		Yes
9906_021		Yes
9906_022		Yes
9906_023	MLL	Yes
9906_024		Yes
9906_027		Yes
9906_028	TCF3-PBX1	Yes
9906_030		Yes
9906_031		Yes
9906_032	MLL	Yes
9906_033		Yes
9906_034		Yes
9906_036		Yes
9906_037		Yes
9906_038		Yes
9906_039		Yes
9906_040		
9906_041	MLL	Yes
9906_042		Yes
9906_043	TCF3-PBX1	Yes
9906_045		Yes
9906_046	TCF3-PBX1	Yes
9906_047		Yes
9906_048		Yes
9906_049		Yes
9906_050		Yes
9906_051	MLL	Yes
9906_052		Yes

<b>Sample ID</b>	<b>Group</b>	<b>U133 Plus 2 data</b>
9906_055		Yes
9906_057		
9906_058	TCF3-PBX1	Yes
9906_060		Yes
9906_061		Yes
9906_062		Yes
9906_063	TCF3-PBX1	Yes
9906_064		Yes
9906_065		Yes
9906_066		Yes
9906_069		Yes
9906_070		
9906_071	TCF3-PBX1	Yes
9906_073		Yes
9906_074	MLL	Yes
9906_075	TCF3-PBX1	
9906_076		Yes
9906_078		
9906_079	TCF3-PBX1	Yes
9906_080		Yes
9906_082		Yes
9906_083	ETV6-RUNX1	Yes
9906_084		Yes
9906_085		Yes
9906_086		Yes
9906_087		
9906_090		Yes
9906_092		Yes
9906_093		Yes
9906_094		Yes
9906_095	MLL	Yes
9906_096	TCF3-PBX1	Yes
9906_097	MLL	Yes
9906_098		Yes
9906_099		Yes
9906_100	TCF3-PBX1	
9906_101		Yes
9906_102		Yes
9906_106		Yes
9906_107		Yes
9906_108		Yes
9906_109		
9906_110		Yes
9906_111		Yes
9906_113		Yes
9906_114		Yes
9906_115	MLL	Yes
9906_116	MLL	Yes
9906_117		Yes

<b>Sample ID</b>	<b>Group</b>	<b>U133 Plus 2 data</b>
9906_118		Yes
9906_119		Yes
9906_120		Yes
9906_121		Yes
9906_122	Hyperdiploid	Yes
9906_123	MLL	Yes
9906_124		Yes
9906_126		Yes
9906_128		Yes
9906_129		Yes
9906_132		Yes
9906_133		Yes
9906_135		Yes
9906_136		Yes
9906_137	MLL	Yes
9906_138		Yes
9906_139	MLL	Yes
9906_141		Yes
9906_142	MLL	Yes
9906_143		Yes
9906_144		Yes
9906_145		Yes
9906_146		Yes
9906_147		Yes
9906_148		Yes
9906_149		
9906_150		Yes
9906_151		Yes
9906_152	TCF3-PBX1	Yes
9906_153		Yes
9906_154		
9906_155		Yes
9906_156	TCF3-PBX1	Yes
9906_157		Yes
9906_159	TCF3-PBX1	Yes
9906_160		Yes
9906_161		Yes
9906_163	TCF3-PBX1	Yes
9906_165		
9906_166	TCF3-PBX1	Yes
9906_167		Yes
9906_168		Yes
9906_170		Yes
9906_171	Hyperdiploid	Yes
9906_173		Yes
9906_174		Yes
9906_175		Yes
9906_176		Yes
9906_177		Yes

<b>Sample ID</b>	<b>Group</b>	<b>U133 Plus 2 data</b>
9906_179		Yes
9906_180		Yes
9906_182		
9906_183		Yes
9906_184		Yes
9906_185		Yes
9906_186		Yes
9906_187	TCF3-PBX1	Yes
9906_188		Yes
9906_189		Yes
9906_190		Yes
9906_192		Yes
9906_193		
9906_195		Yes
9906_196		Yes
9906_198	TCF3-PBX1	Yes
9906_199		Yes
9906_202	TCF3-PBX1	Yes
9906_203	TCF3-PBX1	
9906_206		Yes
9906_207		Yes
9906_209	Hyperdiploid	Yes
9906_210		Yes
9906_211		
9906_214		Yes
9906_215		Yes
9906_216		Yes
9906_217		Yes
9906_218	TCF3-PBX1	Yes
9906_219		Yes
9906_220		Yes
9906_221		Yes
9906_222		Yes
9906_224	ETV6-RUNX1	Yes
9906_225		Yes
9906_227	MLL	Yes
9906_228		Yes
9906_229	MLL	Yes
9906_230	MLL	Yes
9906_231		
9906_233		Yes
9906_234		Yes
9906_235		Yes
9906_236	TCF3-PBX1	Yes
9906_238		Yes
9906_239		Yes
9906_240		Yes
9906_241		Yes
9906_242		Yes

<b>Sample ID</b>	<b>Group</b>	<b>U133 Plus 2 data</b>
9906_243	ETV6-RUNX1	Yes
9906_244		Yes
9906_245	Hyperdiploid	Yes
9906_246		Yes
9906_247	MLL	Yes
9906_248		Yes
9906_249		Yes
9906_250		
9906_251		Yes
9906_252		
9906_253		Yes
9906_254		Yes
9906_255		Yes
9906_256		Yes
9906_257		Yes
9906_258		Yes
9906_259		Yes
9906_260		Yes
9906_261	MLL	Yes
9906_262		Yes
9906_263		Yes
9906_264		Yes
9906_265		Yes
9906_267		Yes
9906_268		
9906_269		
9906_271		Yes
9906_272	TCF3-PBX1	Yes

**Supplementary Table 2. 258 St Jude B-progenitor ALL cases examined.**

\*250K/50K refers to cases examined with 250K Sty and 250K Nsp SNP microarrays; all but 37 cases were also examined with the 50K Hind 240 and 50K Xba 240 arrays.

Sample ID	SNP platform*	U133A expression chip
Hyperdip>50-SNP-#01	250K/50K	JD-ALD485-v5-U133A
Hyperdip>50-SNP-#02	250K/50K	JD0070-ALL-v5-U133A
Hyperdip>50-SNP-#03	250K/50K	JD-ALD510-v5-U133A
Hyperdip>50-SNP-#04	250K/50K	JD0017-ALL-v5-U133A
Hyperdip>50-SNP-#05	250K/50K	JD0020-ALL-v5-U133A
Hyperdip>50-SNP-#06	250K/50K	JD0023-ALL-v5-U133A
Hyperdip>50-SNP-#07	250K/50K	JD-ALD611-v5-U133A
Hyperdip>50-SNP-#08	250K/50K	JD-ALD612-v5-U133A
Hyperdip>50-SNP-#09	250K/50K	JD0041-ALL-v5-U133A
Hyperdip>50-SNP-#10	250K/50K	JD0077-ALL-v5-U133A
Hyperdip>50-SNP-#11	250K/50K	JD0111-ALL-v5-U133A
Hyperdip>50-SNP-#12	250K/50K	JD0097-ALL-v5-U133A
Hyperdip>50-SNP-#13	250K/50K	JD0117-ALL-v5-U133A
Hyperdip>50-SNP-#14	250K/50K	JD0120-ALL-v5-U133A
Hyperdip>50-SNP-#15	250K/50K	JD0121-ALL-v5-U133A
Hyperdip>50-SNP-#16	250K/50K	JD0127-ALL-v5-U133A
Hyperdip>50-SNP-#17	250K/50K	JD0151-ALL-v5-U133A
Hyperdip>50-SNP-#18	250K/50K	JD0168-B-ALL-v5-U133A
Hyperdip>50-SNP-#19	250K/50K	JD0178-ALL-v5-U133A
Hyperdip>50-SNP-#20	250K/50K	JD0191-ALL-v5-U133A
Hyperdip>50-SNP-#21	250K/50K	JD0196-ALL-v5-U133A
Hyperdip>50-SNP-#22	250K/50K	JD0219-ALL-v5-U133A
Hyperdip>50-SNP-#23	250K/50K	JD0222-ALL-v5-U133A
Hyperdip>50-SNP-#24	250K/50K	JD-ALD085-v5-U133A
Hyperdip>50-SNP-#25	250K/50K	
Hyperdip>50-SNP-#26	250K/50K	
Hyperdip>50-SNP-#27	SNP 6.0	JD-ALD013-v5-U133A
Hyperdip>50-SNP-#28	250K/50K	
Hyperdip>50-SNP-#29	250K/50K	JD-ALD112-v5-U133A
Hyperdip>50-SNP-#30	250K/50K	JD-ALD163-v5-U133A
Hyperdip>50-SNP-#31	250K/50K	
Hyperdip>50-SNP-#32	250K/50K	
Hyperdip>50-SNP-#33	250K/50K	
Hyperdip>50-SNP-#34	250K/50K	
Hyperdip>50-SNP-#35	250K/50K	
Hyperdip>50-SNP-#36	250K/50K	
Hyperdip>50-SNP-#37	250K/50K	
Hyperdip>50-SNP-#38	250K/50K	
Hyperdip>50-SNP-#39	250K/50K	
Hyperdip50-SNP-#51	SNP 6.0	
Hyperdip50-SNP-#52	SNP 6.0	
Hyperdip50-SNP-#53	SNP 6.0	

Sample ID	SNP platform*	U133A expression chip
Hyperdip50-SNP-#54	SNP 6.0	
Hyperdip50-SNP-#55	SNP 6.0	
E2A-PBX1-SNP-#01	250K/50K	JD0004-ALL-v5-U133A
E2A-PBX1-SNP-#02	250K/50K	JD0015-ALL-v5-U133A
E2A-PBX1-SNP-#03	250K/50K	JD0036-ALL-v5-U133A
E2A-PBX1-SNP-#04	250K/50K	JD0042-ALL-v5-U133A
E2A-PBX1-SNP-#05	250K/50K	JD0083-ALL-v5-U133A
E2A-PBX1-SNP-#06	250K/50K	JD0099-ALL-v5-U133A
E2A-PBX1-SNP-#07	250K/50K	JD0104-ALL-v5-U133A
E2A-PBX1-SNP-#08	250K/50K	
E2A-PBX1-SNP-#09	250K/50K	JD0203-ALL-v5-U133A
E2A-PBX1-SNP-#10	250K/50K	JD-ALD019-v5-U133A
E2A-PBX1-SNP-#11	250K/50K	JD-ALD025-v5-U133A
E2A-PBX1-SNP-#12	SNP 6.0	JD-ALD437-v5-U133A
E2A-PBX1-SNP-#13	250K/50K	JD-ALD034-v5-U133A
E2A-PBX1-SNP-#14	250K/50K	JD-ALD041-v5-U133A
E2A-PBX1-SNP-#15	250K/50K	JD-ALD071-v5-U133A
E2A-PBX1-SNP-#16	250K/50K	JD-ALD073-v5-U133A
E2A-PBX1-SNP-#17	250K/50K	JD-ALD079-v5-U133A
TEL-AML1-SNP-#01	250K/50K	JD0002-ALL-v5-U133A
TEL-AML1-SNP-#02	250K/50K	JD0066-ALL-v5-U133A
TEL-AML1-SNP-#03	250K/50K	JD0056-ALL-v5-U133A
TEL-AML1-SNP-#04	250K/50K	JD-ALD493-v5-U133A
TEL-AML1-SNP-#05	250K/50K	JD0058-ALL-v5-U133A
TEL-AML1-SNP-#06	250K/50K	JD0059-ALL-v5-U133A
TEL-AML1-SNP-#07	250K/50K	JD0005-ALL-v5-U133A
TEL-AML1-SNP-#08	250K/50K	JD0009-ALL-v5-U133A
TEL-AML1-SNP-#09	250K/50K	JD0033-ALL-v5-U133A
TEL-AML1-SNP-#10	250K/50K	JD0014-ALL-v5-U133A
TEL-AML1-SNP-#11	250K/50K	JD0016-ALL-v5-U133A
TEL-AML1-SNP-#12	250K/50K	JD0018-ALL-v5-U133A
TEL-AML1-SNP-#13	250K/50K	JD0048-ALL-v5-U133A
TEL-AML1-SNP-#14	250K/50K	JD0085-ALL-v5-U133A
TEL-AML1-SNP-#15	250K/50K	JD0101-ALL-v5-U133A
TEL-AML1-SNP-#16	250K/50K	JD0118-ALL-v5-U133A
TEL-AML1-SNP-#17	250K/50K	JD0107-ALL-v5-U133A
TEL-AML1-SNP-#18	250K/50K	JD0109-ALL-v5-U133A
TEL-AML1-SNP-#19	250K/50K	JD0123-ALL-v5-U133A
TEL-AML1-SNP-#20	250K/50K	JD0139-ALL-v5-U133A
TEL-AML1-SNP-#21	250K/50K	JD0149-ALL-v5-U133A
TEL-AML1-SNP-#22	250K/50K	JD0170-ALL-v5-U133A
TEL-AML1-SNP-#23	250K/50K	
TEL-AML1-SNP-#24	250K/50K	JD0175-ALL-v5-U133A
TEL-AML1-SNP-#25	250K/50K	JD0193-ALL-v5-U133A
TEL-AML1-SNP-#26	250K/50K	JD0201-ALL-v5-U133A
TEL-AML1-SNP-#27	250K/50K	
TEL-AML1-SNP-#28	250K/50K	JD0212-ALL-v5-U133A

Sample ID	SNP platform*	U133A expression chip
TEL-AML1-SNP-#29	250K/50K	JD0221-ALL-v5-U133A
TEL-AML1-SNP-#30	250K/50K	
TEL-AML1-SNP-#31	250K/50K	JD-ALD004-v5-U133A
TEL-AML1-SNP-#32	250K/50K	JD-ALD005-v5-U133A
TEL-AML1-SNP-#33	250K/50K	JD-ALD006-v5-U133A
TEL-AML1-SNP-#34	250K/50K	JD-ALD096-v5-U133A
TEL-AML1-SNP-#35	250K/50K	
TEL-AML1-SNP-#36	250K/50K	JD-ALD108-v5-U133A
TEL-AML1-SNP-#37	250K/50K	JD-ALD109-v5-U133A
TEL-AML1-SNP-#38	250K/50K	
TEL-AML1-SNP-#39	250K/50K	
TEL-AML1-SNP-#40	250K/50K	
TEL-AML1-SNP-#41	250K/50K	
TEL-AML1-SNP-#42	250K/50K	
TEL-AML1-SNP-#43	250K/50K	
TEL-AML1-SNP-#44	250K/50K	JD-ALD054-v5-U133A
TEL-AML1-SNP-#45	250K/50K	
TEL-AML1-SNP-#46	250K/50K	
TEL-AML1-SNP-#47	250K/50K	
TEL-AML1-SNP-#48	SNP 6.0	
TEL-AML1-SNP-#49	SNP 6.0	
TEL-AML1-SNP-#50	SNP 6.0	
MLL-SNP-#01	250K/50K	JD0080-ALL-v5-U133A
MLL-SNP-#02	250K/50K	JD0084-ALL-v5-U133A
MLL-SNP-#03	250K/50K	
MLL-SNP-#04	250K/50K	JD0124-ALL-v5-U133A
MLL-SNP-#05	250K/50K	JD-ALD009-v5-U133A
MLL-SNP-#06	250K/50K	JD-ALD433-v5-U133A
MLL-SNP-#07	250K/50K	JD-ALD180-v5-U133A
MLL-SNP-#08	250K/50K	JD-ALD057-v5-U133A
MLL-SNP-#09	250K/50K	JD-ALD052-v5-U133A
MLL-SNP-#10	250K/50K	JD-ALD294-v5-U133A
MLL-SNP-#11	250K/50K	JD-ALD078-v5-U133A
MLL-SNP-#12	250K/50K	
MLL-SNP-#13	250K/50K	
MLL-SNP-#15	250K/50K	
MLL-SNP-#16	250K/50K	
MLL-SNP-#17	250K/50K	JD0284-ALL-v5-U133A
MLL-SNP-#18	250K/50K	JD-ALD232-v5-U133A
MLL-SNP-#19	250K/50K	
MLL-SNP-#20	250K/50K	
MLL-SNP-#21	250K/50K	
MLL-SNP-#22	250K/50K	
MLL-SNP-#23	250K/50K	JD-ALD385-v5-U133A
MLL-SNP-#24	SNP 6.0	
MLL-SNP-#25	SNP 6.0	
BCR-ABL-SNP-#01	250K/50K	JD-ALD494-v5-U133A

Sample ID	SNP platform*	U133A expression chip
BCR-ABL-SNP-#02	250K/50K	JD-ALD613-v5-U133A
BCR-ABL-SNP-#03	250K/50K	JD0102-ALL-v5-U133A
BCR-ABL-SNP-#04	250K/50K	JD0129-ALL-v5-U133A
BCR-ABL-SNP-#05	250K/50K	JD0154-ALL-v5-U133A
BCR-ABL-SNP-#06	250K/50K	JD0192-ALL-v5-U133A
BCR-ABL-SNP-#07	250K/50K	JD0206-ALL-v5-U133A
BCR-ABL-SNP-#08	250K/50K	JD-ALD008-v5-U133A
BCR-ABL-SNP-#09	250K/50K	JD-ALD035-v5-U133A
BCR-ABL-SNP-#10	250K/50K	JD-ALD386-v5-U133A
BCR-ABL-SNP-#11	SNP 6.0	JD-ALD387-v5-U133A
BCR-ABL-SNP-#12	250K/50K	JD-ALD388-v5-U133A
BCR-ABL-SNP-#13	250K/50K	JD-ALD389-v5-U133A
BCR-ABL-SNP-#14	250K/50K	JD-ALD390-v5-U133A
BCR-ABL-SNP-#15	SNP 6.0	JD-ALD233-v5-U133A
BCR-ABL-SNP-#16	SNP 6.0	
BCR-ABL-SNP-#17	250K/50K	JD-ALD428-v5-U133A
BCR-ABL-SNP-#18	SNP 6.0	JD-ALD264-v5-U133A
BCR-ABL-SNP-#19	250K/50K	JD-ALD171-v5-U133A
BCR-ABL-SNP-#20	250K/50K	JD-ALD039-v5-U133A
BCR-ABL-SNP-#21	250K/50K	JD-ALD391-v5-U133A
Hypodip-SNP-#01	250K/50K	JD0057-ALL-v5-U133A
Hypodip-SNP-#02	250K/50K	JD-ALD536-v5-U133A
Hypodip-SNP-#03	250K/50K	JD0025-ALL-v5-U133A
Hypodip-SNP-#04	250K/50K	JD0037-ALL-v5-U133A
Hypodip-SNP-#05	250K/50K	JD0087-ALL-v5-U133A
Hypodip-SNP-#06	250K/50K	JD0095-ALL-v5-U133A
Hypodip-SNP-#07	250K/50K	
Hypodip-SNP-#08	250K/50K	
Hypodip-SNP-#09	250K/50K	JD-ALD196-v5-U133A
Hypodip-SNP-#10	250K/50K	
Hyperdip>50-SNP-#40	250K/50K	JD-ALD280-v5-U133A
Hyperdip47-50-SNP-#01	250K/50K	JD0064-ALL-v5-U133A
Hyperdip47-50-SNP-#02	250K/50K	JD-ALD509-v5-U133A
Hyperdip47-50-SNP-#03	250K/50K	JD0062-ALL-v5-U133A
Hyperdip47-50-SNP-#04	SNP 6.0	JD-ALD554-v5-U133A
Hyperdip47-50-SNP-#05	250K/50K	JD0098-ALL-v5-U133A
Hyperdip47-50-SNP-#06	250K/50K	JD0112-ALL-v5-U133A
Hyperdip47-50-SNP-#07	250K/50K	JD0108-ALL-v5-U133A
Hyperdip47-50-SNP-#08	250K/50K	JD0132-ALL-v5-U133A
Hyperdip47-50-SNP-#09	250K/50K	JD0133-ALL-v5-U133A
Hyperdip47-50-SNP-#10	250K/50K	JD0137-ALL-v5-U133A
Hyperdip47-50-SNP-#11	250K/50K	JD0138-ALL-v5-U133A
Hyperdip47-50-SNP-#12	250K/50K	JD0150-ALL-v5-U133A
Hyperdip47-50-SNP-#13	250K/50K	JD0157-ALL-v5-U133A
Hyperdip47-50-SNP-#14	250K/50K	JD0181-ALL-v5-U133A
Hyperdip47-50-SNP-#15	250K/50K	JD0186B-ALL-v5-U133A
Hyperdip47-50-SNP-#16	250K/50K	

Sample ID	SNP platform*	U133A expression chip
Hyperdip47-50-SNP-#17	250K/50K	
Hyperdip47-50-SNP-#18	250K/50K	
Hyperdip47-50-SNP-#19	250K/50K	
Hyperdip47-50-SNP-#20	250K/50K	
Hyperdip47-50-SNP-#21	250K/50K	
Hyperdip47-50-SNP-#22	250K/50K	
Hyperdip47-50-SNP-#23	250K/50K	
Hyperdip47-50-SNP-#24	250K/50K	JD-ALD242-v5-U133A
Other-SNP-#01	250K/50K	JD0065-ALL-v5-U133A
Other-SNP-#02	250K/50K	JD0116-ALL-v5-U133A
Other-SNP-#03	250K/50K	JD0122-ALL-v5-U133A
Other-SNP-#04	250K/50K	JD0131-ALL-v5-U133A
Other-SNP-#05	250K/50K	JD0166-ALL-v5-U133A
Other-SNP-#06	250K/50K	JD0202-ALL-v5-U133A
Other-SNP-#07	250K/50K	JD0226-ALL-v5-U133A
Other-SNP-#08	250K/50K	JD-ALD340-v5-U133A
Other-SNP-#09	250K/50K	JD-ALD363-v5-U133A
Other-SNP-#10	250K/50K	
Other-SNP-#11	250K/50K	
Other-SNP-#12	250K/50K	JD-ALD279-v5-U133A
Other-SNP-#13	250K/50K	
Other-SNP-#14	250K/50K	JD-ALD194-v5-U133A
Other-SNP-#15	250K/50K	JD-ALD066-v5-U133A
Other-SNP-#16	250K/50K	
Other-SNP-#17	250K/50K	JD-ALD329-v5-U133A
Other-SNP-#18	250K/50K	JD-ALD115-v5-U133A
Other-SNP-#19	250K/50K	JD-ALD185-v5-U133A
Other-SNP-#20	250K/50K	JD-ALD297-v5-U133A
Other-SNP-#21	SNP 6.0	JD0021-ARD-v5-U133A
Other-SNP-#22	250K/50K	JD0031-ARD-v5-U133A
Other-SNP-#23	250K/50K	JD0025-ARD-v5-U133A
Other-SNP-#24	250K/50K	JD0003-ARD-v5-U133A
Other-SNP-#25	250K/50K	JD0018-ARD-v5-U133A
Other-SNP-#26	250K/50K	JD0014-ARD-v5-U133A
Other-SNP-#27	SNP 6.0	
Other-SNP-#28	SNP 6.0	
Other-SNP-#29	SNP 6.0	
Other-SNP-#30	SNP 6.0	
Other-SNP-#31	SNP 6.0	
Other-SNP-#32	SNP 6.0	
Other-SNP-#33	SNP 6.0	
Other-SNP-#34	SNP 6.0	
Other-SNP-#35	SNP 6.0	
Other-SNP-#36	SNP 6.0	
Other-SNP-#37	SNP 6.0	JD-ALD146-v5-U133A
Other-SNP-#38	SNP 6.0	JD-ALD420-v5-U133A
Other-SNP-#39	SNP 6.0	

Sample ID	SNP platform*	U133A expression chip
Other-SNP-#40	SNP 6.0	
Other-SNP-#41	SNP 6.0	
Other-SNP-#42	SNP 6.0	JD0019-ALL-v5-U133A
Other-SNP-#43	SNP 6.0	
Pseudodip-SNP-#01	250K/50K	JD0001-ALL-v5-U133A
Pseudodip-SNP-#02	250K/50K	JD0071-ALL-v5-U133A
Pseudodip-SNP-#03	250K/50K	JD0012-ALL-v5-U133A
Pseudodip-SNP-#04	250K/50K	JD0032-ALL-v5-U133A
Pseudodip-SNP-#05	250K/50K	JD0021-ALL-v5-U133A
Pseudodip-SNP-#06	250K/50K	JD-ALD610-v5-U133A
Pseudodip-SNP-#07	250K/50K	JD0103-ALL-v5-U133A
Pseudodip-SNP-#08	250K/50K	
Pseudodip-SNP-#09	250K/50K	JD0173-ALL-v5-U133A
Pseudodip-SNP-#10	250K/50K	JD0185B-ALL-v5-U133A
Pseudodip-SNP-#11	250K/50K	JD0188-ALL-v5-U133A
Pseudodip-SNP-#12	250K/50K	JD0225-ALL-v5-U133A
Pseudodip-SNP-#13	250K/50K	
Pseudodip-SNP-#14	250K/50K	
Pseudodip-SNP-#15	250K/50K	
Pseudodip-SNP-#16	250K/50K	JD-ALD164-v5-U133A
Pseudodip-SNP-#17	250K/50K	
Pseudodip-SNP-#18	SNP 6.0	
Pseudodip-SNP-#19	250K/50K	
Pseudodip-SNP-#20	250K/50K	
Pseudodip-SNP-#21	250K/50K	JD-ALD176-v5-U133A
Pseudodip-SNP-#22	250K/50K	JD0088-ALL-v5-U133A
Pseudodip-SNP-#23	250K/50K	JD-ALD136-v5-U133A
Pseudodip-SNP-#24	250K/50K	JD-ALD325-v5-U133A

## SUPPLEMENTARY RESULTS

**Supplementary Table 3. DNA copy number abnormality frequency in high-risk pediatric ALL**

Group	All lesions			Deletions			Gains		
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
<i>ETV6-RUNX1</i> N=3	9.00	10	1-16	8.67	9	1-16	.67	0	0-2
<i>TCF3-PBX1</i> N=25	3.52	4	0-9	2.44	2	0-8	1.08	1	0-4
<i>MLL</i> -rearranged N=19	1.84	1	0-11	1.26	1	0-10	.58	0	0-2
High hyperdiploid N=4	16.5	16.5	6-27	2.0	2	0-4	14.5	14.5	0-23
Other N=170	9.59	7	0-86	5.84	5	0-33	3.78	1	0-75
Total N=221	8.36	6	0-86	5.03	4	0-33	3.35	1	0-75
<i>P</i>	<0.0001			<0.0001			<0.0001		

**Supplementary Table 4. Regions of recurring copy number alteration in the P9906 cohort.**

Abnormalities are deletions unless otherwise indicated. \*Adjacent to gene. \*\*Deletion or gain \*\*\*Focal amplifications only. †Focal deletions and amplifications. ††Includes focal deletions, broad Xp deletions, and CNA immediately adjacent to the minimal region of focal deletion. †††B cell pathway lesions include deletions or sequence mutations involving *BCL11A* (N=1), *BLNK* (N=2), *EBF1* (N=17), *IKZF1* (N=67), *IKZF2* (N=1), *LEF1* (N=1), *MEF2C* (N=1), *PAX5* (N=81), *RAG1/2* (N=8), *SOX4* (N=1), *SPI1* (N=1) and *TCF3* (N=21); no lesions were found in *CD79A*, *GABPA*, *IKZF3*, *IL7RA*, *IRF4*, *IRF8*, *STAT3*, *STAT5A*, or *STAT5B*. *VPREB1* may be considered part of the “B cell pathway” but is located in the immunoglobulin lambda light chain locus at 22q11.22, and is commonly deleted upon rearrangement of this locus. Consequently the biologic significance of *VPREB1* deletions in B-ALL is unclear. Accordingly, the frequency of B cell pathway lesions is shown excluding and including *VPREB1* alterations. iAmp21, intrachromosomal amplification of chromosome 21.

Lesion	Location	All	%	Hyperdiploid	%	TCF3-PBX1	%	ETV6-RUNX1	%	MLL	%	Other	%
		221		4		25		3		19		170	
<i>PDE4B</i>	1p31.2	6	2.7	0	0	0	0	0	0	0	0	6	3.5
<i>NRAS</i>	1p13.1	4	1.8	0	0	0	0	0	0	0	0	4	2.4
<i>ADAR</i>	1q22	4	1.8	0	0	0	0	0	0	0	0	4	2.4
<i>ZNF528*</i>	1q44	6	2.7	0	0	0	0	0	0	0	0	6	3.5
1q gain	1q23.3-1qtel	22	10.0	1	25	16	64	0	0	0	0	5	2.9
<i>ARPP-21</i>	3p22.3	7	3.2	0	0	0	0	0	0	0	0	7	4.1
<i>FHIT</i>	3p14.2	1	0.5	0	0	0	0	0	0	0	0	1	0.6
<i>FLNB</i>	3p14.3	5	2.3	0	0	0	0	0	0	0	0	5	2.9
<i>BTLA/CD200</i>	3q13.2	13	5.9	0	0	0	0	0	0	0	0	13	7.6
<i>MBNL1</i>	3q25.1	8	3.6	0	0	0	0	1	33.3	0	0	7	4.1
<i>TBL1XR1</i>	3q26.32	7	3.2	0	0	0	0	0	0	0	0	7	4.1
<i>IL1RAP</i>	3q28	3	1.4	0	0	0	0	0	0	0	0	3	1.8
<i>ARHGAP24*</i>	4q21.23	4	1.8	0	0	0	0	0	0	0	0	4	2.4
<i>NR3C2*</i>	4q31.23	5	2.3	0	0	0	0	2	66.7	0	0	3	1.8
<i>FBXW7</i>	4q31.3	3	1.4	0	0	0	0	0	0	0	0	3	1.8
<i>EBF1</i>	5q33.3	17	7.7	0	0	0	0	1	33.3	0	0	16	9.4
Histone cluster	6p22.2	9	4.0	0	0	0	0	1	33.3	0	0	8	4.7
<i>GRIK2</i>	6q16	14	6.3	0	0	2	8	1	33.3	0	0	11	6.5
<i>ARMC2/SESNI</i>	6q21	15	6.8	0	0	2	8	1	33.3	0	0	12	7.1
<i>ARID1B*</i>	6q25.3	8	3.6	0	0	1	4	1	33.3	0	0	6	3.5

Lesion	Location	All	%	Hyperdiploid	%	TCF3- PBX1	%	ETV6- RUNX1	%	MLL	%	Other	%
		221		4		25		3		19		170	
<i>IKZF1</i>	7p13	63	28.6	0	0	0	0	0	0	1	5.3	61	35.9
<i>IKZF1</i> CNA or sequence mutation	7p13	67	30.3	0	0	0	0	0	0	2	10.5	65	38.2
<i>MSRA</i>	8p23	4	1.8	0	0	0	0	0	0	0	0	4	2.4
<i>TOX</i>	8q12.1	8	3.6	0	0	0	0	1	33.3	0	0	7	4.1
<i>CCDC26</i> **	8q24.21	23	10.4	2	50	3	12	0	0	2	10.5	16	9.4
<i>CDKN2A/B</i>	9p21.3	101	45.7	2	50	9	38	1	33.3	4	21.1	85	50
<i>PAX5</i> CNA	9p13.2	70	31.7	1	25	10	40	1	33.3	1	5.3	57	33.5
<i>PAX5</i> CNA or sequence mutation	9p13.2	81	36.7	1	25	11	44	1	33.3	1	5.3	67	39.4
<i>ABL1</i> ***	9q34.13	3	1.4	0	0	0	0	0	0	0	0	3	1.8
<i>ADARB2</i>	10p15.2	4	1.8	0	0	0	0	0	0	0	0	4	2.4
<i>COPEB/KLF6</i>	10p15	2	0.9	0	0	0	0	0	0	0	0	2	1.17
<i>ADD3</i>	10q25.2	18	8.1	0	0	1	4	1	33.3	0	0	16	9.4
<i>RAG1/2</i>	11p12	8	3.6	0	0	0	0	0	0	1	5.3	7	4.1
<i>NUP160/PTPRJ</i>	11p11.2	4	1.8	0	0	0	0	0	0	0	0	4	2.4
<i>ETV6</i>	12p13.2	28	12.7	0	0	0	0	1	33.3	0	0	27	15.8
<i>KRAS</i>	12p12.1	14	6.3	0	0	2	8	0	0	1	5.3	11	6.5
<i>BTG1</i>	12q21.33	23	10.4	0	0	0	0	0	0	0	0	23	13.5
<i>ZMYM5</i>	13q12.11	3	1.4	0	0	0	0	0	0	0	0	3	1.8
<i>ELF1</i>	13q14.11												
<i>C13orf21/TSC22D1</i>	13q14	20	9.1	0	0	5	20	0	0	0	0	15	8.8
<i>RBI</i>	13q14.2	25	11.3	0	0	5	20	0	0	0	0	20	11.8
<i>DLEU2/7/mir15/- 16a)</i>	13q14	21	9.5	0	0	5	20	0	0	0	0	16	9.4
<i>ATP10A</i>	15q12	6	2.7	0	0	0	0	0	0	0	0	6	3.5
<i>SPRED1 (5')</i>	15q14	0	0	0	0	0	0	0	0	0	0	0	0
<i>LTK</i>	15q15.1	0	0	0	0	0	0	0	0	0	0	0	0
<i>NFI</i> <sup>†</sup>	17q11.2	6	2.7	0	0	1	4	0	0	1	5.3	4	2.3
<i>TCF3</i>	19p13.3	21	9.5	0	0	15	60	0	0	0	0	6	3.5
<i>C20orf94</i>	20p12.2	19	8.6	0	0	0	0	0	0	0	0	19	11.2
<i>ERG</i>	21q22	11	5	0	0	0	0	0	0	0	0	11	6.5
iAmp21	21, varies	10	4.5	0	0	0	0	0	0	0	0	10	5.8
<i>VPREB1</i>	22q11.22	57	25.8	0	0	0	0	2	66.7	0	0	55	32.4
<i>IL3RA/CSF2RA</i> <sup>††</sup>	Xp22.33	18	6.8	0	0	0	0	0	0	0	0	18	10.6

Lesion	Location	All	%	Hyperdiploid	%	TCF3- PBX1	%	ETV6- RUNX1	%	MLL	%	Other	%
		221		4		25		3		19		170	
<i>DMD</i>	Xp21.1	15	6.8	0	0	5	20	0	0	0	0	10	5.8
B cell pathway <sup>†††</sup>		147	66.5	1	25	2	66.7	121	71.2	18	72	5	26.3
B cell pathway including <i>VPREB1</i> <sup>†††</sup>		154	69.7	1	25	2	66.7	128	75.3	18	72	5	26.3
B cell pathway lesion per case (mean, range)		1.3 (0-5)		0.3 (0-1)		1.0 (0-2)		1.7 (0-3)		0.3 (0-1)		1.5 (0-5)	

**Supplementary Table 5. Regions of recurring copy number alteration in the St Jude cohort.**

H50, high hyperdiploid, iAmp21, Intrachromosomal amplification of chromosome 21. \*Adjacent to *ZNF238*.

Lesion	Location	All	%	H50	%	TCF3- PBX1	%	ETV6- RUNX1	%	MLL	%	Ph	%	Hypo	%	Other	%
		258		44		17		50		24		21		10		92	
<i>PDE4B</i>	1p31.2	2	8	0	0	0	0	2	4	0	0	0	0	0	0	0	0
<i>NRAS</i>	1p13.1	1	4	0	0	0	0	0	0	0	0	0	0	0	0	1	1.1
<i>ADAR</i>	1q22	2	8	0	0	0	0	0	0	0	0	0	0	0	0	2	2.2
<i>LOC440742*</i>	1q44	2	8	0	0	0	0	0	0	0	0	0	0	0	0	2	2.2
1q gain	1q23.3- 1qtel	30	11.6	13	29.5	16	94.1	0	0	0	0	0	0	0	0	1	1.1
<i>ARPP-21</i>	3p22.3	8	3.1	1	2.3	0	0	2	4	0	0	1	4.8	2	20	2	2.2
<i>FHIT</i>	3p14.2	12	4.7	0	0	0	0	6	12	0	0	2	9.5	1	10	3	3.3
<i>FLNB</i>	3p14.3	7	2.7	1	2.3	0	0	1	2	0	0	1	4.8	1	10	3	3.3
<i>BTLA/CD200</i>	3q13.2	16	6.2	0	0	0	0	8	16	0	0	5	23.8	1	10	2	2.2
<i>MBNL1</i>	3q25.1	9	3.5	2	4.5	0	0	3	6	0	0	2	9.5	1	10	1	1.1
<i>TBL1XR1</i>	3q26.32	15	5.8	1	2.3	0	0	8	16	1	4.2	1	4.8	0	0	4	4.3
<i>ILIRAP</i>	3q28	3	1.2	0	0	0	0	1	2	0	0	1	4.8	1	10	0	0
<i>ARHGAP24</i>	4q21.23	2	8	0	0	0	0	0	0	0	0	0	0	1	10	1	1.1
<i>NR3C2</i>	4q31.23	10	3.9	0	0	0	0	6	12	0	0	0	0	1	10	3	3.3
<i>LEF1</i>	4q25	5	1.9	0	0	0	0	2	4.0	0	0	0	0	1	10	2	2.2
<i>FBXW7</i>	4q31.3	5	1.9	0	0	0	0	1	2	0	0	1	4.8	1	10	2	2.2
<i>EBF1</i>	5q33.3	12	4.7	1	2.3	0	0	5	10	0	0	3	14.3	1	10	2	2.2
Histone cluster	6p22.2	21	8.1	1	2.3	0	0	3	6	0	0	3	14.3	3	30	11	12
<i>GRIK2</i>	6q16	11	4.3	1	2.3	1	5.9	7	14	0	0	0	0	0	0	2	2.2
<i>ARMC2/SESNI</i>	6q21	13	5	0	0	0	0	8	16	0	0	0	0	0	0	5	5.4
4LOC389437	6q25.3	7	2.7	0	0	0	0	4	8	0	0	0	0	1	10	2	2/2
<i>IKZF1</i>	7p13	48	18.6	4	9.1	0	0	0	0	1	4.2	16	76.2	5	50	22	22.8
<i>CDK6</i>	7q21.2	8	3.1	1	2.3	0	0	0	0	0	0	2	9.5	3	30	2	2.2
<i>MSRA</i>	8p23	6	2.3	0	0	0	0	2	4.0	0	0	1	4.8	2	20	1	1.1
<i>TOX</i>	8q12.1	11	4.3	0	0	0	0	5	10	0	0	1	4.8	0	0	5	5.4
<i>CCDC26</i>	8q24.21	5	1.9	1	2.3	0	0	0	0	0	0	0	0	0	0	4	4.3
<i>CDKN2A/B</i>	9p21.3	87	33.7	9	20.5	6	35.3	15	30	4	16.7	11	52.4	10	100	32	34.8
<i>PAX5</i> CNA	9p13.2	79	30.6	4	9.1	7	41.2	17	34	4	16.7	11	52.4	10	100	26	28.3
<i>PAX5</i> CNA or sequence mutation	9p13.2	83	32.2	4	9.1	8	47.1	17	34	5	20.8	11	52.4	10	100	28	30.4

Lesion	Location	All	%	H50	%	TCF3- PBX1	%	ETV6- RUNX1	%	MLL	%	Ph	%	Hypo	%	Other	%
		258		44		17		50		24		21		10		92	
<i>ABL1</i>	9q34.13	5	1.9	0	0	0	0	0	0	0	0	4	19	1	10	0	0
<i>ADARB2</i>	10p15.2	1	4	0	0	0	0	0	0	0	0	1	4.8	0	0	0	0
<i>COPEB/KLF6</i>	10p15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>PTEN</i>	10q23.31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>BLNK</i>	10q24.1	3	1.2	0	0	0	0	2	4	0	0	0	0	0	0	1	1.1
<i>ADD3</i>	10q25.2	14	5.4	1	2.3	0	0	4	8	0	0	5	23.8	0	0	4	4.3
<i>RAG1/2</i>	11p12	15	5.8	0	0	0	0	8	16	1	4.2	0	0	0	0	6	6.5
<i>NUP160/PTPRJ</i>	11p11.2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	1	1.1
<i>ATM</i>	11q22.3	7	2.7	0	0	0	0	2	4	0	0	1	4.8	0	0	4	4.3
<i>ETV6</i>	12p13.2	63	24.4	5	11.5	0	0	34	68	2	8.3	2	9.5	2	20	18	19.6
<i>KRAS</i>	12p12.1	20	7.8	2	4.5	0	0	8	16	1	4.2	0	0	1	10	8	8.7
<i>BTG1</i>	12q21.33	18	7.0	0	0	0	0	7	14	0	0	4	19	1	10	6	6.5
<i>ZMYM5</i>	13q12.11	5	1.9	1	2.3	0	0	2	4	0	0	0	0	0	0	2	2.2
<i>ELF1</i>	13q14.11	12	4.7	2	4.5	2	11.8	4	8	1	4.2	0	0	1	10	2	2.2
<i>C13orf21/TSC22D1</i>	13q14	15	5.8	2	4.5	2	11.8	4	8	1	4.2	2	9.5	1	10	3	3.3
<i>RB1</i>	13q14.2	15	5.8	3	6.8	2	11.8	2	4	2	8.3	4	19	0	0	2	2.2
<i>DLEU2/7/mir15/-16a)</i>	13q14	16	6.2	5	11.4	2	11.8	3	6	3	12.5	1	4.8	0	0	2	2.2
<i>ATP10A</i>	15q12	5	1.9	0	0	0	0	1	2	0	0	1	4.8	1	10	2	2.2
<i>SPRED1 (5')</i>	15q14	6	2.3	0	0	0	0	0	0	0	0	1	4.8	1	10	4	4.3
<i>LTK</i>	15q15.1	6	2.3	0	0	0	0	3	6	0	0	0	0	1	10	2	2.2
<i>NF1</i>	17q11.2	8	3.1	1	2.3	0	0	2	4	0	0	0	0	1	10	4	4.3
<i>IKZF3 (AIOLOS)</i>	17q21.1	3	1.2	0	0	0	0	0	0	0	0	0	0	2	20	1	1.1
<i>TCF3</i>	19p13.3	17	6.6	1	2.3	16	94.1	0	0	0	0	0	0	0	0	0	0
<i>C20orf94</i>	20p12.2	20	7.8	2	4.5	0	0	7	14	0	0	7	33.3	0	0	4	4.3
<i>ERG</i>	21q22	14	5.4	0	0	0	0	0	0	0	0	0	0	0	0	14	15.2
iAmp21*	21, varies	11	4.3	0	0	0	0	5	10	0	0	0	0	0	0	6	6.5
<i>VPREB1</i>	22q11.22	80	31	7	15.9	1	5.9	35	70	1	4.2	7	33.3	3	30	26	28.3
<i>IL3RA</i>	Xp22.33	18	7.0	1	2.3	0	0	6	12	0	0	0	0	1	10	10	10.9
<i>DMD</i>	Xp21.1	11	4.3	1	2.3	0	0	4	8	0	0	0	0	0	0	6	6.5
B pathway		137	53.1	11	25	17	100	27	54	6	25	16	76.2	10	100	50	53.4
B pathway with VPREB1		169	65.5	16	36.4	17	100	42	64	6	25	16	76.2	10	100	62	67.4

**Supplementary Table 6. Results of *PAX5* genomic quantitative PCR.**

Results represent means of duplicate measurements, and are ratios of *PAX5* to control (RNase P). Values <0.75 represent hemizygous deletion, and <0.3 homozygous deletion.

Sample ID	Group	<i>PAX5</i> deletion region	Exon 3	Exon 6	Exon 8
9906_002	<i>TCF3-PBX1</i>	All gene	0.60	0.58	0.61
9906_004	Other	5' to distal	0.13	0.16	0.20
9906_009	Other	5' to distal	0.47	0.53	0.58
9906_013	Other	e3 - distal	0.54	0.31	0.33
9906_014	Other	e2 - e5	0.57	1.05	1.05
9906_028	<i>TCF3-PBX1</i>	All gene	0.44	0.50	0.44
9906_034	Other	All gene	0.31	0.30	0.29
9906_037	Other	e6 – distal	0.70	0.38	0.38
9906_040	Other	All gene	0.31	0.35	0.37
9906_045	Other	All gene	0.60	0.60	0.61
9906_046	<i>TCF3-PBX1</i>	All gene	0.42	0.41	0.39
9906_048	Other	5' - e7	0.39	0.12	0.65
9906_055	Other	e2 - e5	0.51	0.92	1.02
9906_063	<i>TCF3-PBX1</i>	e6	0.48	0.32	0.55
9906_065	Other	All gene	0.50	0.52	0.47
9906_070	Other	Promoter - e3	0.59	0.76	0.96
9906_080	Other	e7-distal	0.80	0.84	0.50
9906_098	Other	e9	0.95	0.95	0.94
9906_102	Other	Amplification e2-e5	2.77	1.21	0.90
9906_107	Other	e7-distal	1.15	1.16	0.63
9906_111	Other	e6-distal	1.09	0.60	0.57
9906_118	Other	Promoter - e5	0.72	1.16	0.97
9906_124	Other	e2-e4, e6	0.54	0.53	1.00
9906_141	Other	e2-e5	0.61	1.06	0.98
9906_154	Other	e2-e8	0.49	0.48	0.43
9906_157	Other	e2-e6	0.42	0.44	0.86
9906_160	Other	All gene	0.57	0.63	0.56
9906_161	Other	5' - e3	0.33	0.59	0.65
9906_163	<i>TCF3-PBX1</i>	5' - e7	0.03	0.03	0.04
9906_175	Other	e6-8	0.97	0.54	0.55
9906_180	Other	5' - e7	0.71	0.35	0.81
9906_192	Other	e8-9	0.78	0.83	0.43
9906_196	Other	e2-e7, homozygous e6-7	0.49	0.06	0.99
9906_218	<i>TCF3-PBX1</i>	All gene	0.44	0.46	0.50
9906_268	Other	e6-distal	1.13	0.64	0.76

**Supplementary Table 7. *PAX5* sequence mutations in the P9906 cohort.**

Five cases had two point mutations in *trans*, and 11 cases had deletions of one *PAX5* allele and point mutation of the second allele. e, exon; fs, frameshift

Sample ID	Group	<i>PAX5</i> deletion	<i>PAX5</i> deletion region	<i>PAX5</i> mutation description
9906_034	Other	Yes	All gene	P80R
9906_060	Other	No		V151I
9906_065	Other	Yes	All gene	G24R
9906_086	Other	No		G24R
9906_106	Other	Yes	All gene	P80R
9906_110	Other	No		Exon 3 splice; D53V
9906_113	Other	No		I139T
9906_121	Other	Yes	All gene	P80R
9906_156	<i>TCF3-PBX1</i>	No		I301T
9906_173	Other	Yes	All gene	T333fs
9906_179	Other	No		P80R; E201fs
9906_180	Other	Yes	Promoter-e7	S213L
9906_188	Other	Yes	All gene	P80R
9906_192	Other	Yes	e8-9	R59G
9906_195	Other	No		T75R; V336fs
9906_228	Other	No		P80R; E201fs
9906_233	Other	No		P80R; E7 splice
9906_234	Other	Yes	Focal promoter	E9 splice
9906_235	Other	Yes	All gene	E9 splice
9906_239	Other	No		V319FS
9906_256	Other	Yes	All gene	P80R
9906_258	Other	No		V319FS

**Supplementary Table 8. Description all P9906 cases harboring *IKZF1* deletions and/or sequence mutations, and results of genomic quantitative PCR results for *IKZF1* deletions.**

Results represent means of duplicate measurements, and are ratios of *IKZF1* to control (RNase P). Values <0.75 represent hemizygous deletion, and <0.3 homozygous deletion. e, exon; fs, frameshift (mutation); gqPCR, genomic quantitative PCR; \*, nonsense mutation

Sample ID	Group	<i>IKZF1</i> deletion	Region	gqPCR exon 1	gq PCR exon 3	gqPCR exon 5	gqPCR exon 6	<i>IKZF1</i> mutation
9906_261	MLL	Yes	e3 - e6	0.96		0.58	0.59	
9906_001	Other	Yes	e1 - e6					
9906_007	Other	Yes	e3 - e6	1.57	0.56		0.53	
9906_014	Other	Yes	e1 - 7	0.56	0.61		0.56	
9906_019	Other	Yes	All gene	0.67	0.66		0.59	G158S
9906_021	Other	Yes	All gene					
9906_024	Other	No						H224fs
9906_027	Other	Yes	e3 - e6	1.09	0.63		0.55	
9906_030	Other	Yes	e3 - e6	1.23			0.59	
9906_033	Other	Yes	e3 - e6	1.17	0.57		0.54	
9906_038	Other	Yes	e3 - distal					
9906_039	Other	Yes	All gene	0.54			0.53	
9906_040	Other	Yes	All gene	0.73			0.59	
9906_045	Other	Yes	e3 - e6	0.97		0.56	0.71	
9906_047	Other	Yes	e3 - e6	1.17		0.71	0.68	
9906_048	Other	Yes	e1 - e6	0.65	0.72		0.64	
9906_049	Other	Yes	All gene	0.71	0.75		0.71	
9906_055	Other	Yes	All gene					L117fs
9906_064	Other	Yes	5' - e1	0.59		0.99	1.02	
9906_065	Other	No						S402fs
9906_078	Other	Yes	5' - e1	0.52	1.23		1.08	
9906_082	Other	Yes	5' - e1	0.56		1.02	1.09	
9906_084	Other	Yes	e3 - e6	1.22		0.53	0.57	
9906_087	Other	Yes	All gene	0.64	0.70		0.68	
9906_090	Other	No						R111*
9906_093	Other	Yes	All gene					
9906_107	Other	Yes	e3 - e6	1.11		0.57	0.65	
9906_109	Other	Yes	5' - e1	0.57		1.05	1.20	
9906_113	Other	Yes	5' - e1	0.54		1.05	1.47	
9906_118	Other	Yes	All gene	0.59	0.61		0.61	
9906_120	Other	Yes	All gene					
9906_124	Other	Yes	e1 - e5	0.69		0.57	0.81	
9906_135	Other	Yes	All gene	0.57	0.57		0.57	
9906_138	Other	Yes	e1 - e5					
9906_141	Other	Yes	e3 - e6					
9906_146	Other	Yes	e3 - e6					
9906_151	Other	Yes	e3 - e6					
9906_153	Other	Yes	All gene					
9906_154	Other	Yes	e1 - e3	0.53		1.00	1.40	

Sample ID	Group	<i>IKZF1</i> deletion	Region	gqPCR exon 1	gq PCR exon 3	gqPCR exon 5	gqPCR exon 6	<i>IKZF1</i> mutation
9906_161	Other	Yes	e3 – distal	1.29	0.97		0.71	
9906_168	Other	Yes	5' - e1	0.67	1.07		1.05	
9906_170	Other	Yes	e1 - e4	0.49		0.91	1.23	
9906_173	Other	Yes	All gene	0.74		0.66	0.65	
9906_174	Other	Yes	e3 – distal	1.63		0.57	0.63	
9906_175	Other	Yes	All gene	0.61			0.62	
9906_179	Other	No						E504fs
9906_192	Other	Yes	e3 – distal	1.10	0.64		0.58	
9906_196	Other	Yes	e2 - e5	1.09		0.54	0.79	
9906_206	Other	Yes	e1 - e4	0.62		0.92	0.98	
9906_210	Other	Yes	e1 - e6	0.58	0.52		0.52	
9906_215	Other	Yes	e1 - e4	0.63		0.95	0.95	
9906_217	Other	Yes	5' - e6, homozygous 5' - e1	0.11		0.55	0.82	
9906_219	Other	Yes	All gene	0.63	0.75		0.52	
9906_222	Other	Yes	e3 - e6	0.99	0.65		0.63	
9906_225	Other	Yes	e3 - e6	1.21		0.43	0.58	
9906_231	Other	Yes	e3 - e6					
9906_234	Other	Yes	e1 - e6	0.62		0.66	0.58	
9906_240	Other	Yes	e1 - e6	0.59		0.63	0.69	
9906_242	Other	Yes	e3 - e6	0.99		0.49	0.48	
9906_244	Other	Yes	e3 - e6	1.06		0.60	0.54	
9906_250	Other	Yes	e1 - e6	0.45		0.50	0.49	
9906_252	Other	Yes	e3 - e6	1.04		0.60	0.57	
9906_253	Other	Yes	All gene	0.45		0.46	0.45	
9906_257	Other	Yes	e3 - e6	0.97	0.91		0.65	
9906_258	Other	Yes	e1 - e6	0.61		0.62	0.62	
9906_262	Other	Yes	e3 - e6	0.96		0.53	0.52	
9906_271	Other	Yes	e3 - distal					

**Supplementary Table 9. Description of B-cell pathway lesions observed in the P9906 cohort.**

In addition to *PAX5* and *IKZF1* abnormalities, lesions were also identified in *TCF3* (N=21), *EBF1* (N=17), *RAG1/2* (N=8), *BLNK* (N=2), *BCL11A*, *IKZF2* (encoding the IKAROS family member HELIOS), *LEF1*, *MEF2C*, *SOX4* and *SP11* (PU.1) (1 each). deln, deletion; FS, frameshift.

Sample ID	Group	Number of lesions	B cell pathway lesions
9906_001	Other	2	<i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_002	E2A	2	<i>PAX5</i> deln; <i>TCF3</i> deln
9906_003	E2A	1	<i>TCF3</i> deln
9906_004	Other	1	<i>PAX5</i> deln
9906_007	Other	3	<i>IKZF1</i> deln; <i>VPREB1</i> deln; <i>MEF2C</i> deln
9906_009	Other	2	<i>PAX5</i> deln; <i>VPREB1</i> deln
9906_010	Other	2	<i>EBF1</i> deln; <i>VPREB1</i> deln
9906_011	Other	1	<i>IKZF2</i> deln
9906_012	Other	1	<i>EBF1</i> deln
9906_013	Other	1	<i>PAX5</i> deln;
9906_014	Other	4	<i>EBF1</i> deln; <i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_016	Other	1	<i>EBF1</i> deln
9906_017	E2A	1	<i>TCF3</i> deln
9906_019	Other	3	<i>IKZF1</i> deln; G158S <i>IKZF1</i> mutation; <i>VPREB1</i> deln
9906_020	Other	2	<i>RAG1/2</i> deln; <i>VPREB1</i> deln
9906_021	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_024	Other	1	H224FS <i>IKZF1</i> mutation
9906_027	Other	3	<i>EBF1</i> deln; <i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_028	E2A	1	<i>PAX5</i> deln
9906_030	Other	1	<i>IKZF1</i> deln
9906_031	Other	1	<i>VPREB1</i> deln
9906_033	Other	1	<i>IKZF1</i> deln
9906_034	Other	4	<i>PAX5</i> deln; P80R <i>PAX5</i> mutation; <i>RAG1/2</i> deln; <i>LEF1</i> deln
9906_037	Other	2	<i>PAX5</i> deln; <i>RAG1/2</i> deln
9906_038	Other	1	<i>IKZF1</i> deln
9906_039	Other	1	<i>IKZF1</i> deln
9906_040	Other	2	<i>IKZF1</i> deln; <i>PAX5</i> deln
9906_045	Other	2	<i>IKZF1</i> deln; <i>PAX5</i> deln
9906_046	E2A	2	<i>PAX5</i> deln; <i>TCF3</i> deln
9906_047	Other	2	<i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_048	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln , homozygous
9906_049	Other	1	<i>IKZF1</i> deln
9906_052	Other	1	<i>VPREB1</i> deln
9906_055	Other	3	<i>IKZF1</i> deln; L117FS <i>IKZF1</i> mutation; <i>PAX5</i> deln
9906_057	Other	1	<i>RAG1/2</i> deln
9906_058	E2A	1	<i>TCF3</i> deln
9906_060	Other	2	V151I <i>PAX5</i> mutation; <i>VPREB1</i> deln
9906_063	E2A	2	<i>PAX5</i> deln; <i>TCF3</i> deln
9906_064	Other	2	<i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_065	Other	4	S402FS <i>IKZF1</i> mutation; <i>PAX5</i> deln; G24R <i>PAX5</i> mutation; <i>VPREB1</i> deln
9906_066	Other	1	<i>PAX5</i> deln
9906_070	Other	1	<i>PAX5</i> deln

Sample ID	Group	Number of lesions	B cell pathway lesions
9906_071	E2A	1	<i>PAX5</i> deln
9906_073	Other	2	<i>RAG1/2</i> deln; <i>TCF3</i> deln
9906_075	E2A	2	<i>PAX5</i> deln; <i>TCF3</i> deln
9906_078	Other	2	<i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_079	E2A	2	<i>PAX5</i> deln; <i>TCF3</i> deln
9906_080	Other	1	<i>PAX5</i> deln
9906_082	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_083	TEL	3	<i>EBF1</i> deln; <i>VPREB1</i> deln; <i>SOX4</i> deln
9906_084	Other	3	<i>EBF1</i> deln; <i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_085	Other	1	<i>RAG1/2</i> deln
9906_086	Other	1	G24R <i>PAX5</i> mutation
9906_087	Other	1	<i>IKZF1</i> deln
9906_090	Other	4	<i>EBF1</i> deln; R111* <i>IKZF1</i> mutation; <i>RAG1/2</i> deln; <i>VPREB1</i> deln
9906_092	Other	1	<i>VPREB1</i> deln
9906_093	Other	1	<i>IKZF1</i> deln
9906_094	Other	1	<i>BLNK</i> deln
9906_096	E2A	2	<i>PAX5</i> deln; <i>TCF3</i> deln
9906_097	MLL	1	<i>PAX5</i> deln
9906_098	Other	1	<i>PAX5</i> deln
9906_102	Other	2	focal internal <i>PAX5</i> amplification; <i>VPREB1</i> deln
9906_106	Other	2	<i>PAX5</i> deln; P80R <i>PAX5</i> mutation
9906_107	Other	2	<i>IKZF1</i> deln; <i>PAX5</i> deln
9906_108	Other	1	<i>PAX5</i> deln
9906_109	Other	3	<i>EBF1</i> deln; <i>IKZF1</i> deln; <i>TCF3</i> deln
9906_110	Other	3	D53V and E3 splice <i>PAX5</i> mutation; <i>VPREB1</i> deln
9906_111	Other	1	<i>PAX5</i> deln
9906_113	Other	3	<i>IKZF1</i> deln; I139T <i>PAX5</i> mutation; <i>VPREB1</i> deln
9906_114	Other	2	<i>VPREB1</i> deln; <i>BCL11A</i> deln
9906_117	Other	2	<i>EBF1</i> deln; <i>VPREB1</i> deln
9906_118	Other	2	<i>IKZF1</i> deln; <i>PAX5</i> deln
9906_120	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_121	Other	2	<i>PAX5</i> deln; P80R <i>PAX5</i> mutation
9906_124	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_128	MLL	1	<i>RAG1/2</i> deln
9906_135	Other	1	<i>IKZF1</i> deln
9906_137	MLL	1	<i>RAG1/2</i> deln
9906_138	Other	1	<i>IKZF1</i> deln
9906_141	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_144	Other	1	<i>VPREB1</i> deln
9906_145	Other	2	<i>PAX5</i> deln; <i>VPREB1</i> deln
9906_146	Other	1	<i>IKZF1</i> deln
9906_147	Other	2	<i>PAX5</i> deln; <i>VPREB1</i> deln
9906_148	Other	1	<i>PAX5</i> deln
9906_150	Other	2	<i>PAX5</i> deln; <i>VPREB1</i> deln
9906_151	Other	2	<i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_153	Other	2	<i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_154	Other	2	<i>IKZF1</i> deln; <i>PAX5</i> deln
9906_155	Other	1	<i>VPREB1</i> deln
9906_156	E2A	1	I301T <i>PAX5</i> mutation
9906_157	Other	2	<i>PAX5</i> deln; <i>VPREB1</i> deln
9906_159	E2A	1	<i>TCF3</i> deln
9906_160	Other	2	<i>PAX5</i> deln; <i>VPREB1</i> deln

Sample ID	Group	Number of lesions	B cell pathway lesions
9906_161	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_163	E2A	2	<i>PAX5</i> deln; <i>TCF3</i> deln
9906_166	E2A	1	<i>TCF3</i> deln
9906_168	Other	3	<i>EBF1</i> deln; <i>IKZF1</i> deln; <i>RAG1/2</i> deln
9906_170	Other	1	<i>IKZF1</i> deln
9906_171	T4_10	1	<i>PAX5</i> deln
9906_173	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; T333FS <i>PAX5</i> mutation
9906_174	Other	1	<i>IKZF1</i> deln
9906_175	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_176	Other	1	<i>PAX5</i> deln
9906_177	Other	1	<i>PAX5</i> deln
9906_179	Other	4	E504FS <i>IKZF1</i> mutation; P80R and E201FS <i>PAX5</i> mutation; <i>VPREB1</i> deln
9906_180	Other	2	<i>PAX5</i> deln; S213L <i>PAX5</i> mutation
9906_183	Other	2	<i>RAG1/2</i> deln; <i>SPI1</i> deln
9906_184	Other	1	<i>PAX5</i> deln
9906_185	Other	2	<i>VPREB1</i> deln; <i>BLNK</i> deln
9906_188	Other	2	<i>PAX5</i> deln; P80R <i>PAX5</i> mutation
9906_192	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; R59G <i>PAX5</i> mutation
9906_193	Other	1	<i>VPREB1</i> deln
9906_195	Other	2	T75R and V336FS <i>PAX5</i> mutation
9906_196	Other	4	<i>IKZF1</i> deln; <i>PAX5</i> deln, homozygous; <i>VPREB1</i> deln
9906_199	Other	1	<i>PAX5</i> deln
9906_202	E2A	1	<i>TCF3</i> deln
9906_206	Other	2	<i>EBF1</i> deln; <i>IKZF1</i> deln
9906_210	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_215	Other	3	<i>PAX5</i> deln; <i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_217	Other	3	<i>EBF1</i> deln; <i>IKZF1</i> deln, homozygous
9906_218	E2A	2	<i>PAX5</i> deln; <i>TCF3</i> deln
9906_219	Other	2	<i>IKZF1</i> deln; <i>TCF3</i> deln
9906_222	Other	2	<i>EBF1</i> deln; <i>IKZF1</i> deln
9906_225	Other	4	<i>IKZF1</i> deln; <i>PAX5</i> deln; <i>RAG1/2</i> deln; <i>VPREB1</i> deln
9906_227	MLL	1	M31V <i>IKZF1</i> mutation
9906_228	Other	2	P80R and E201FS <i>PAX5</i> mutation
9906_231	Other	2	<i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_233	Other	2	P80R and E7 splice <i>PAX5</i> mutation
9906_234	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; E9 splice <i>PAX5</i> mutation
9906_235	Other	2	<i>PAX5</i> deln; E9 splice <i>PAX5</i> mutation
9906_236	E2A	1	<i>TCF3</i> deln
9906_239	Other	2	V319FS <i>PAX5</i> mutation; <i>VPREB1</i> deln
9906_240	Other	1	<i>IKZF1</i> deln
9906_242	Other	1	<i>IKZF1</i> deln
9906_243	TEL	2	<i>PAX5</i> deln; <i>VPREB1</i> deln
9906_244	Other	1	<i>IKZF1</i> deln
9906_250	Other	4	<i>EBF1</i> deln; <i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_252	Other	1	<i>IKZF1</i> deln
9906_253	Other	1	<i>IKZF1</i> deln
9906_254	Other	1	<i>PAX5</i> deln
9906_255	Other	1	focal internal <i>PAX5</i> amplification
9906_256	Other	2	<i>PAX5</i> deln; P80R <i>PAX5</i> mutation
9906_257	Other	3	<i>IKZF1</i> deln; <i>PAX5</i> deln; <i>VPREB1</i> deln
9906_258	Other	5	<i>EBF1</i> deln; <i>IKZF1</i> deln; V319FS <i>PAX5</i> mutation; <i>RAG1/2</i> deln;

Sample ID	Group	Number of lesions	B cell pathway lesions
			<i>VPREB1</i> deln
9906_259	Other	1	<i>VPREB1</i> deln
9906_260	Other	1	<i>RAG1/2</i> deln
9906_261	MLL	1	<i>IKZF1</i> deln
9906_262	Other	2	<i>IKZF1</i> deln; <i>VPREB1</i> deln
9906_263	Other	2	<i>TCF3</i> deln; <i>VPREB1</i> deln
9906_264	Other	1	<i>PAX5</i> deln; <i>TCF3</i> deln
9906_265	Other	1	<i>TCF3</i> deln
9906_268	Other	1	<i>PAX5</i> deln
9906_271	Other	3	<i>EBF1</i> deln; <i>IKZF1</i> deln; <i>VPREB1</i> deln

**Supplementary Table 10. Variation in number of B cell pathway lesions between P9906 ALL subtypes**

Group	Mean number of B cell pathway lesions	Range
All	1.3 (0-5)	0-5
<i>TCF3-PBX1</i>	1.0 (0-2)	0-2
<i>MLL</i> -rearranged	0.3 (0-1)	0-1
Other	1.5 (0-5)	0-5
Hyperdiploid	0.3 (0-1)	0-1
<i>ETV6-RUNX1</i>	1.7 (0-1)	0-3
ANOVA P=0.0001		
Results of ANOVA post hoc Fisher's PLSD test		
	Mean Diff.	P
Hyperdiploid, <i>TCF3-PBX1</i>	-0.79	0.18
Hyperdiploid, <i>ETV6-RUNX1</i>	-1.417	0.093
Hyperdiploid, <i>MLL</i>	-0.013	0.98
Hyperdiploid, Other	-1.215	0.030
<i>TCF3-PBX1</i> , <i>ETV6-RUNX1</i>	-0.627	0.35
<i>TCF3-PBX1</i> , <i>MLL</i>	0.777	0.02
<i>TCF3-PBX1</i> , Other	-0.425	0.07
<i>ETV6-RUNX1</i> , <i>MLL</i>	1.404	0.041
<i>ETV6-RUNX1</i> , Other	0.202	0.75
<i>MLL</i> , Other	-1.202	<0.0001

**Supplementary Table 11. Genes with univariate Cox score exceeding threshold of  $\pm 1.8$  in SPC analysis, P9906 cohort.**

Raw score refers to the modified univariate Cox score calculated for each gene. Importance score is a measure of correlation between each gene and the first principal component derived from the SPC analysis.

Name	Raw score	Importance score
<i>IKZF1</i>	-4.002	-26.569
<i>MKKS/C20orf94</i>	-2.075	-7.928
<i>BTLA</i>	-2.137	-5.625
<i>EBF1</i>	-2.277	-5.316
<i>NRAS</i>	1.824	1.53
<i>FLNB</i>	-1.824	-1.469
<i>ZNF238</i>	-1.885	-1.159
<i>C1orf131</i>	-1.802	-0.878
<i>HAAO</i>	-1.861	0.787
<i>SRBD1</i>	-1.936	0.671
<i>SUSD3</i>	-1.84	-0.593
<i>FOXN2</i>	-1.904	0.57
<i>PRKCE</i>	-1.811	0.541
<i>C9orf71</i>	-1.907	0.491
<i>PPM1B</i>	-1.902	0.478
<i>FAM82A</i>	-1.872	0.359
<i>FXN</i>	-2.503	0.325
<i>SLC46A2</i>	-1.816	-0.254
<i>PRKACG</i>	-2.862	-0.123
<i>RHOQ</i>	-1.855	0.077

**Supplementary Table 12. Genes with univariate Cox score exceeding threshold of  $\pm 1.9$  in SPC analysis, St Jude cohort.**

Raw score refers to the modified univariate Cox score calculated for each gene. Importance score is a measure of correlation between each gene and the first principal component derived from the SPC analysis.

Name	Raw score	Importance score
<i>IKZF1</i>	-3.164	-18.607
<i>TAS2R5</i>	-2.056	-8.751
<i>LOC136242</i>	-2.034	-8.674
<i>SVOPL</i>	-1.951	-8.584
<i>C7orf34</i>	-1.901	-8.577
<i>FLJ36031</i>	-1.944	-8.177
<i>GPR37</i>	-1.931	-8.107

**Supplementary Table 13. Associations between of B cell pathway lesions, *IKZF1* alterations and hematologic relapse, P9906 cohort**

	N	Relapse N	Competing Risks N	Cumulative Incidence (SE)%	
				5 year	P
<b>Hematologic relapse</b>					
<b>B cell pathway</b>					
No	74	9	9	15.5 (5.7)	0.072
Yes	147	36	25	30.8 (4.9)	
<b>Number of B pathway lesions</b>					
N=0	67	9	7	17.3 (6.4)	0.026
N=1	67	12	8	23.1 (6.8)	
N=2	52	10	8	27.5 (8.9)	
N>=3	35	14	11	43.1 (9.4)	
<b><i>IKZF1</i> deletion</b>					
No	158	20	21	14.4 (3.2)	<0.0001
Yes	63	25	13	52.9 (8.5)	
<b><i>IKZF1</i> deletion or mutation</b>					
No	153	19	20	14.0 (3.1)	<0.0001
Yes	68	27	14	55.2 (8.6)	
<b>Any relapse</b>					
<b>B cell pathway lesions</b>					
No	74	17	1	26.6 (6.5)	0.041
Yes	147	58	3	46.3 (5.1)	
<b>Number of B pathway lesions</b>					
N=0	67	15	1	26.6 (7.0)	<0.0001
N=1	67	19	1	33.6 (7.3)	
N=2	52	17	1	41.4 (9.4)	
N>=3	35	24	1	72.1 (8.7)	
<b><i>IKZF1</i> deletion</b>					
No	158	38	3	26.0 (3.8)	<0.0001
Yes	63	37	1	73.4(8.0)	
<b><i>IKZF1</i> deletion or mutation</b>					
No	153	36	3	25.2 (3.8)	<0.0001
Yes	68	38	1	73.4 (8.0)	

**Supplementary Table 14. Kaplan-Meier estimates of EFS by B cell pathway or *IKZF1* lesions, P9906 cohort**

	N	Any Relapse or Death N	Event-free survival (SE)%	
			5-Year	P
<b>B cell pathway lesions</b>				
No	74	18	72.0 (13.5)	0.037
Yes	147	61	51.6 (8.5)	
<b>Number of B cell pathway lesions</b>				
N=0	67	16	71.9 (14.4)	<0.0001
N=1	67	20	64.9 (10.7)	
N=2	52	18	56.6 (16.7)	
N>=3	35	25	25 (12.5)	
<b><i>IKZF1</i> deletion</b>				
No	158	41	72.0 (8.3)	<0.0001
Yes	63	38	25.2 (9.7)	
<b><i>IKZF1</i> deletion or mutation</b>				
No	153	40	72.8 (8.3)	<0.0001
Yes	68	40	25.9 (10.0)	

**Supplementary Table 15. Hazard ratio estimates of B-cell pathway and *IKZF1* lesions on relapse and event free survival, P9906 cohort.**

Fine and Gray test, after adjustment for age, presentation leukocyte count and cytogenetic subtype.

	<b>Hazard Ratio (95% CI)</b>	<b>P</b>
<b>Hematologic relapse</b>		
B cell pathway lesions	1.86 (0.85-4.10)	0.12
Number of B cell pathway lesions	1.59 (1.12-2.26)	0.010
<i>IKZF1</i> deletion	4.05 (2.09-7.85)	<0.0001
<i>IKZF1</i> deletion or mutation	4.07 (2.08-7.97)	<0.0001
<b>Any relapse</b>		
B cell pathway lesion	2.08 (1.12-3.88)	0.021
Number of B cell pathway lesions	1.82 (1.41-2.34)	<0.0001
<i>IKZF1</i> deletion	2.96 (1.78-4.94)	<0.0001
<i>IKZF1</i> deletion or mutation	2.96 (1.74-4.81)	<0.0001
<b>Event free survival</b>		
B cell pathway lesion	1.85 (1.07-3.19)	0.029
Number of B cell pathway lesions	1.66 (1.32-2.10)	<0.0001
<i>IKZF1</i> deletion	2.78 (1.75-4.39)	<0.0001
<i>IKZF1</i> deletion or mutation	2.70 (1.71-4.26)	<0.0001

**Supplementary Table 16. Associations between genomic abnormalities and day 8 MRD, P9906 cohort.**

	Day 8 MRD N (%)			P- Value
	MRD≤0.01%	0.01%<MRD≤1.0%	MRD>1.0%	
<b><i>RBI</i></b>				
No	25 (14.79)	58 (34.32)	86 (50.89)	0.14
Yes	8 (29.63)	9 (33.33)	10 (37.04)	
<b><i>EBF1</i></b>				
No	32 (17.78)	66 (36.67)	82 (45.56)	0.0055
Yes	1 (6.25)	1 (6.25)	14 (87.50)	
<b><i>IKZF1</i> deletion</b>				
No	27 (19.71)	50 (36.50)	60 (43.80)	0.066
Yes	6 (10.17)	17 (28.81)	36 (61.02)	
<b><i>IKZF1</i> deletion or mutation</b>				
No	26 (19.55)	50 (37.59)	57 (42.86)	0.041
Yes	7 (11.11)	17 (26.98)	39 (61.90)	
<b><i>PAX5</i> deletion or mutation</b>				
No	14 (11.48)	39 (31.97)	69 (56.56)	0.0076
Yes	19 (25.68)	28 (37.84)	27 (36.49)	
<b>Age group</b>				
1<age≤10 years	14 (21.21)	22 (33.33)	30 (45.45)	0.50
Age>10 years	19 (14.62)	45 (34.62)	66 (50.77)	
<b>WBC group</b>				
WBC< 50K	17 (17.35)	38 (38.78)	43 (43.88)	0.32
WBC≥50k	16 (16.33)	29 (29.59)	53 (54.08)	
<b>Subtype</b>				
Hyperdiploid or <i>ETV6-RUNX1</i>	2 (40.00)	2 (40.00)	1 (20.00)	0.063
<i>TCF3-PBX1</i>	2 (9.52)	12 (57.14)	7 (33.33)	
<i>MLL</i> -rearranged	4 (23.53)	8 (47.06)	5 (29.41)	
Others	25 (16.34)	45 (29.41)	83 (54.25)	

**Supplementary Table 17. Associations between genetic lesions and day 29 MRD, P9906 cohort**

	Day 29 MRD; N (%)			P
	MRD≤0.01%	0.01%<MRD≤1.0%	MRD>1.0%	
<b>1q gain</b>				
No	113 (61.41)	46 (25.00)	25 (13.59)	
Yes	18 (90.00)	2 (10.00)	0 (0.00)	0.034
<b>ABL1</b>				
No	131 (65.17)	48 (23.88)	22 (10.95)	
Yes	0 (0.00)	0 (0.00)	3 (100.0)	<0.0001
<b>ADD3</b>				
No	124 (66.67)	44 (23.66)	18 (9.68)	
Yes	7 (38.89)	4 (22.22)	7 (38.89)	0.0012
<b>BTLA/CD200</b>				
No	127 (66.49)	44 (23.04)	20 (10.47)	
Yes	4 (30.77)	4 (30.77)	5 (38.46)	0.0052
<b>C20orf94</b>				
No	124 (66.31)	43 (22.99)	20 (10.70)	
Yes	7 (41.18)	5 (29.41)	5 (29.41)	0.044
<b>EBF1</b>				
No	129 (68.62)	40 (21.28)	19 (10.11)	
Yes	2 (12.50)	8 (50.00)	6 (37.50)	<0.0001
<b>IKZF1deletion</b>				
No	102 (71.83)	29 (20.42)	11 (7.75)	
Yes	29 (46.77)	19 (30.65)	14 (22.58)	0.0010
<b>IKZF1 deletion or mutation</b>				
No	100 (72.99)	28 (20.44)	9 (6.57)	
Yes	31 (46.27)	20 (29.85)	16 (23.88)	0.0001
<b>PAX5 deletion</b>				
No	80 (58.39)	39 (28.47)	18 (13.14)	
Yes	51 (76.12)	9 (13.43)	7 (10.45)	0.034
<b>PAX5 deletion or mutation</b>				
No	70 (55.56)	38 (30.16)	18 (14.29)	
Yes	61 (78.21)	10 (12.82)	7 (8.97)	0.0039
<b>RAG1/2</b>				
No	131 (66.16)	43 (21.72)	24 (12.12)	
Yes	0 (0.00)	5 (83.33)	1 (16.67)	0.0012
<b>Age</b>				
1<age≤10 years	48 (73.85)	14 (21.54)	3 (4.62)	
Age>10 years	83 (59.71)	34 (24.46)	22 (15.83)	0.048
<b>WCC</b>				
WBC< 50K	67 (64.42)	27 (25.96)	10 (9.62)	
WBC≥50k	64 (64.00)	21 (21.00)	15 (15.00)	0.42
<b>Subtype</b>				
Hyperdiploid or <i>ETV6-RUNX1</i>	5 (83.33)	0 (0.00)	1 (16.67)	
<i>TCF3-PBX1</i>	22 (100.0)	0 (0.00)	0 (0.00)	
<i>MLL</i> -rearranged	7 (43.75)	8 (50.00)	1 (6.25)	
Others	97 (60.63)	40 (25.00)	23 (14.38)	0.0015

**Supplementary Table 18. Association of genetic lesions with day 29 MRD adjusted by age, presentation leukocyte count and genetic subtype in the P9906 cohort. \*Other B-ALL (*ETV6-RUNX1*, *TCF3-PBX1*, hyperdiploid and normal/miscellaneous cases) v. *MLL*-rearranged ALL. WBC, presentation peripheral blood leukocyte count**

<b>Lesion</b>	<b>factor</b>	<b>Odds Ratio (95%CI)</b>	<b>P</b>
1q gain	WBC<50 vs. WBC≥50	0.68 (0.36-1.29)	0.24
	age≤10yr vs. age>10yr	0.36 (0.18-0.73)	0.0050
	Subtype*	0.61 (0.22-1.71)	0.35
	1q gain present v. absent	0.16 (0.03-0.71)	0.016
<i>ABL1</i>	WBC<50 vs. WBC≥50	0.75 (0.40-1.43)	0.39
	age≤10yr vs. age>10yr	0.44 (0.21-0.89)	0.022
	Subtype*	0.50 (0.18-1.37)	0.18
	<i>ABL1</i> deletion present v. absent	N/A	0.98
<i>ADD3</i>	WBC<50 vs. WBC≥50	0.77 (0.41-1.45)	0.42
	age≤10yr vs. age>10yr	0.41 (0.20-0.83)	0.013
	Subtype*	0.44 (0.16-1.24)	0.12
	<i>ADD3</i> deletion present v. absent	4.38 (1.73, 11.13)	0.0019
<i>BTLA/CD200</i>	WBC<50 vs. WBC≥50	0.78 (0.41-1.49)	0.45
	age≤10yr vs. age>10yr	0.42 (0.21-0.87)	0.019
	Subtype*	0.46 (0.16-1.27)	0.13
	<i>BTLA/CD200</i> deletion present v. absent	4.87 (1.64, 14.45)	0.0044
<i>C20orf94</i>	WBC<50 vs. WBC≥50	0.74 (0.39-1.39)	0.35
	age≤10yr vs. age>10yr	0.43 (0.21-0.87)	0.019
	Subtype*	0.48 (0.17-1.32)	0.16
	<i>C20orf94</i> deletion present v. absent	2.87 (1.11-7.44)	0.030
<i>EBF1</i>	WBC<50 vs. WBC≥50	0.61 (0.32-1.16)	0.13
	age≤10yr vs. age>10yr	0.37 (0.18-0.76)	0.0069
	Subtype*	0.43 (0.15-1.20)	0.11
	<i>EBF1</i> deletion present v. absent	8.98 (3.27, 24.61)	<0.0001
<i>IKZF1</i> deletion	WBC<50 vs. WBC≥50	0.72 (0.38-1.38)	0.33
	age≤10yr vs. age>10yr	0.40 (0.19-0.82)	0.013
	Subtype*	0.35 (0.12-1.00)	0.050
	<i>IKZF1</i> deletion present v. absent	3.41 (1.84-6.33)	0.0001
<i>IKZF1</i> deletion or mutation	WBC<50 vs. WBC≥50	0.72 (0.38-1.39)	0.33
	age≤10yr vs. age>10yr	0.41 (0.20-0.84)	0.016
	Subtype*	0.35 (0.12-1.00)	0.05
	<i>IKZF1</i> alteration present v. absent	3.71 (2.02-6.84)	<0.0001
<i>PAX5</i> deletion	WBC<50 vs. WBC≥50	0.70 (0.37-1.31)	0.26
	age≤10yr vs. age>10yr	0.42 (0.21-0.85)	0.016
	Subtype*	0.65 (0.23-1.83)	0.41
	<i>PAX5</i> deletion present v. absent	0.53 (0.28-1.03)	0.06
<i>PAX5</i> deletion or mutation	WBC<50 vs. WBC≥50	0.70 (0.37-1.31)	0.26
	age≤10yr vs. age>10yr	0.42 (0.21-0.85)	0.016
	Subtype*	0.65 (0.23-1.83)	0.41
	<i>PAX5</i> alteration present v. absent	0.53 (0.28-1.03)	0.060
<i>RAG1/2</i>	WBC<50 vs. WBC≥50	0.65 (0.35-1.23)	0.18
	age≤10yr vs. age>10yr	0.40 (0.19-0.81)	0.011
	Subtype*	0.51 (0.18-1.40)	0.19
	<i>RAG1/2</i> deletion present v. absent	5.51 (1.20, 25.39)	0.029

**Supplementary Table 19. Multivariable analysis of associations between *IKZF1* deletion/mutation and outcome (isolated or combined hematologic relapse, any relapse, or any event) in the P9906 cohort adjusting for age, presentation leukocyte count, leukemia subtype, and day 8 or day 29 MRD level**

\*Other B-ALL (*ETV6-RUNX1*, *TCF3-PBX1*, hyperdiploid, normal/miscellaneous ALL cases) v. *MLL*-rearranged. No association between sex or race (white, black, other), thus these variables were not included in the multivariable model. †As there was no event for day 8 MRD<0.01%, this analysis could not be performed for day 8 MRD and relapse.

<b>Outcomes/Factors</b>	<b>Hazard Ratio (95% CI)</b>	<b>P</b>
<b>Hematological Relapse (adjusting for day 29 MRD)†</b>		
age>10yr vs. age≤10yr	0.67 (0.25-1.72)	0.40
Subtype*	1.27 (0.29-5.59)	0.75
WBC≥50K vs. WBC<50K	0.79 (0.38-1.63)	0.52
0.01%<D29MRD≤1.0% vs. D29MRD≤0.01%	4.50 (2.05-9.86)	0.0002
D29MRD>1.0% vs. D29MRD≤0.01%	5.46 (2.39-12.50)	0.0001
<i>IKZF1</i> alteration present vs. absent	2.87 (1.35-6.09)	0.0062
<b>Any Relapse (adjusting for day 29 MRD)†</b>		
age>10yr vs. age≤10yr	0.79 (0.43-1.47)	0.46
Subtype*	1.10 (0.37-3.23)	0.86
WBC≥50K vs. WBC<50K	1.21 (0.72-2.04)	0.47
0.01%<D29MRD≤1.0% vs. D29MRD≤0.01%	2.33 (1.31-4.15)	0.0041
D29MRD>1.0% vs. D29MRD≤0.01%	2.55 (1.34-4.85)	0.0044
<i>IKZF1</i> alteration present vs. absent	2.40 (1.370-4.20)	0.0021
<b>Any Event (adjusting for day 8 MRD)</b>		
age>10yr vs. age≤10yr	1.05 (0.58-1.87)	0.88
Subtype*	1.34 (0.52-3.48)	0.54
WBC≥50K vs. WBC<50K	1.30 (0.75-2.26)	0.35
0.01%<D8MRD≤1.0% vs. D8MRD≤0.01%	1.33 (0.52-3.41)	0.55
D8MRD>1.0% vs. D8MRD≤0.01%	2.62 (1.09-6.26)	0.031
<i>IKZF1</i> alteration present vs. absent	2.71 (1.65-4.45)	<0.0001
<b>Any Event (adjusting for day 29 MRD)</b>		
age>10yr vs. age≤10yr	0.85 (0.48-1.50)	0.57
Subtype*	1.07 (0.43-2.64)	0.89
WBC≥50K vs. WBC<50K	1.34 (0.79-2.27)	0.27
0.01%<D29MRD≤1.0% vs. D29MRD≤0.01%	2.57 (1.49-4.43)	0.0007
D29MRD>1.0% vs. D29MRD≤0.01%	3.17 (1.66-6.05)	0.0005
<i>IKZF1</i> alteration present vs. absent	1.93 (1.17-3.18)	0.010

**Supplementary Table 20. Cumulative incidence of isolated or combined hematologic relapse by genetic lesions, St Jude cohort with MRD data (N=160).**

\*Stratified according to treatment protocol: Total XIII intermediate to high risk N=23, Total XIII low risk N=28, Total XIV and XV standard and high risk N=50, total XV low risk N=59.

Lesion	Sub N	Hematologic relapse N	Competing Risks N	Cumulative Incidence (SE)%		
				5-Year	P Unstratified Gray's Test	P Stratified Gray's Test*
<b><i>RAG1/2</i></b>						
No	150	17	6	10.4 (2.7)		
Yes	10	2	0	25.0 (17.2)	0.32	0.050
<b><i>ATM1</i></b>						
No	158	18	6	10.7 (2.7)		
Yes	2	1	0	NA	0.022	0.008
<b><i>KRAS</i></b>						
No	147	15	6	9.8 (2.6)		
Yes	13	4	0	30.8 (16.9)	0.015	0.013
<b><i>IKZF1</i></b>						
No	139	12	4	8.4 (2.6)		
Yes	21	7	2	29.4 (10.5)	0.001	0.039

**Supplementary Table 21. Cumulative incidence of hematologic relapse by genetic lesions, St Jude cohort with MRD data, excluding BCR-ABL1 positive ALL (N=151).**

\*Stratified according to treatment protocol: Total XIII intermediate to high risk N=20, Total XIII low risk N=28, Total XIV and XV standard and high risk N=44, total XV low risk N=59.

<b>Lesion</b>	<b>N</b>	<b>Hematologic relapse N</b>	<b>competing Risks N</b>	<b>Cumulative Incidence (SE)%</b>		
				<b>5-Year</b>	<b>P Unstratified Gray's Test</b>	<b>P Stratified Gray's Test</b>
<b><i>RAG1/2</i></b>						
No	141	12	5	7.1 (2.4)		
Yes	10	2	0	25.0 (17.2)	0.157	0.032
<b><i>KRAS</i></b>						
No	138	10	5	6.5 (2.3)		
Yes	13	4	0	30.8 (16.9)	0.002	0.002
<b><i>IKZF1</i></b>						
No	136	10	4	6.8 (2.4)		
Yes	15	4	1	20.6 (11.1)	0.022	0.11

**Supplementary Table 22. Associations of genetic lesions and day 19 MRD, St Jude cohort (all B-progenitor ALL cases)**

Lesion	Day 19 MRD N (%)			P
	MRD<0.01% 71 (44.1%)	0.01%≤MRD<1.0% 64 (39.8%)	MRD≥1.0% 26 (16.2%)	
<b><i>ATP10A</i></b>				
No	71 (44.94)	63 (39.87)	24 (15.19)	0.035
Yes	0 (0.00)	1 (33.33)	2 (66.67)	
<b><i>ARPP-21</i></b>				
No	70 (45.45)	62 (40.26)	22 (14.29)	0.01
Yes	1 (14.29)	2 (28.57)	4 (57.14)	
<b><i>GAB1</i></b>				
No	71 (45.22)	63 (40.13)	23 (14.65)	0.0064
Yes	0 (0.00)	1 (25.00)	3 (75.00)	
<b><i>HIST1H2BE</i></b>				
No	68 (46.26)	59 (40.14)	20 (13.61)	0.013
Yes	3 (21.43)	5 (35.71)	6 (42.86)	
<b><i>IKZF1</i></b>				
No	69 (49.29)	58 (41.43)	13 (9.29)	<0.0001
Yes	2 (9.52)	6 (28.57)	13 (61.90)	
<b><i>CDK6</i></b>				
No	71 (45.81)	63 (40.65)	21 (13.55)	0.0003
Yes	0 (0.00)	1 (16.67)	5 (83.33)	
<b><i>ABL1</i></b>				
No	71 (44.94)	64 (40.51)	23 (14.56)	0.0040
Yes	0 (0.00)	0 (0.00)	3 (100.0)	

**Supplementary Table 23. Associations of genetic lesions and day 46 MRD, St Jude cohort (all B-progenitor ALL cases with MRD data, N=160)**

Lesion Loci	Day 46 MRD N (%)			P
	MRD<0.01% 126 (78.8%)	0.01%≤MRD<1.0% 26 (16.3%)	MRD≥1.0% 8 (5%)	
<b><i>NFI</i></b>				
No	123 (80.39)	22 (14.38)	8 (5.23)	0.043
Yes	3 (42.86)	4 (57.14)	0 (0.00)	
<b><i>EBF1</i></b>				
No	122 (80.79)	23 (15.23)	6 (3.97)	0.021
Yes	4 (44.44)	3 (33.33)	2 (22.22)	
<b>6p22 Histone cluster</b>				
No	120 (82.19)	21 (14.38)	5 (3.42)	0.0030
Yes	6 (42.86)	5 (35.71)	3 (21.43)	
<b><i>HBS1L</i> (5' of <i>MYB</i>)</b>				
No	119 (78.81)	26 (17.22)	6 (3.97)	0.038
Yes	7 (77.78)	0 (0.00)	2 (22.22)	
<b><i>IKZF1</i></b>				
No	119 (85.61)	19 (13.67)	1 (0.72)	<0.0001
Yes	7 (33.33)	7 (33.33)	7 (33.33)	
<b><i>CDKN6</i></b>				
No	125 (81.17)	23 (14.94)	6 (3.90)	0.0034
Yes	1 (16.67)	3 (50.00)	2 (33.33)	
<b><i>ABL</i></b>				
No	126 (80.25)	25 (15.92)	6 (3.82)	0.0015
Yes	0 (0.00)	1 (33.33)	2 (66.67)	

**Supplementary Table 24. Associations of genetic lesions and day 19 MRD, St Jude cohort (B-progenitor ALL cases, excluding *BCR-ABL1* ALL)**

Lesion Loci	Day 19 MRD N (%)			P
	MRD<0.01% 71 (46.4%)	0.01%≤MRD<1.0% 61 (39.9%)	MRD≥1.0% 8 (38.1%)	
<b><i>ATP10A</i></b>				
No	71 (47.33)	60 (40.00)	19 (12.67)	0.023
Yes	0 (0.00)	1 (33.33)	2 (66.67)	
<b><i>ARPP-21</i></b>				
No	70 (47.62)	59 (40.14)	18 (12.24)	0.027
Yes	1 (16.67)	2 (33.33)	3 (50.00)	
<b><i>GABI</i></b>				
No	71 (47.65)	60 (40.27)	18 (12.08)	0.0041
Yes	0 (0.00)	1 (25.00)	3 (75.00)	
<b><i>IKZF1</i></b>				
No	69 (50.00)	56 (40.58)	13 (9.42)	0.0001
Yes	2 (13.33)	5 (33.33)	8 (53.33)	
<b><i>CDK6</i></b>				
No	71 (47.97)	60 (40.54)	17 (11.49)	0.0007
Yes	0 (0.00)	1 (20.00)	4 (80.00)	

**Supplementary Table 25. Associations of genetic lesions and day 46 MRD, St Jude cohort (B-progenitor ALL cases, excluding *BCR-ABL1* ALL)**

Lesion	Day 46 MRD N (%)			P
	MRD<0.01% 124 (82.2%)	0.01%≤MRD<1.0% 23 (15.2%)	MRD≥1.0% 4 (2.6%)	
<b><i>NFI</i></b>				
No	121 (84.03)	19 (13.19)	4 (2.78)	0.024
Yes	3 (42.86)	4 (57.14)	0 (0.00)	
<b><i>IKZF1</i></b>				
No	117 (86.03)	18 (13.24)	1 (0.74)	0.0001
Yes	7 (46.67)	5 (33.33)	3 (20.00)	
<b><i>CDK6</i></b>				
No	123 (84.25)	20 (13.70)	3 (2.05)	0.0096
Yes	1 (20.00)	3 (60.00)	1 (20.00)	
<b><i>CCDC26</i></b>				
No	124 (82.67)	23 (15.33)	3 (2.00)	0.030
Yes	0 (0.00)	0 (0.00)	1 (100.0)	

**Supplementary Table 26. Genes driving positive enrichment of hematopoietic stem/progenitor gene set in P9906 high-risk ALL.**

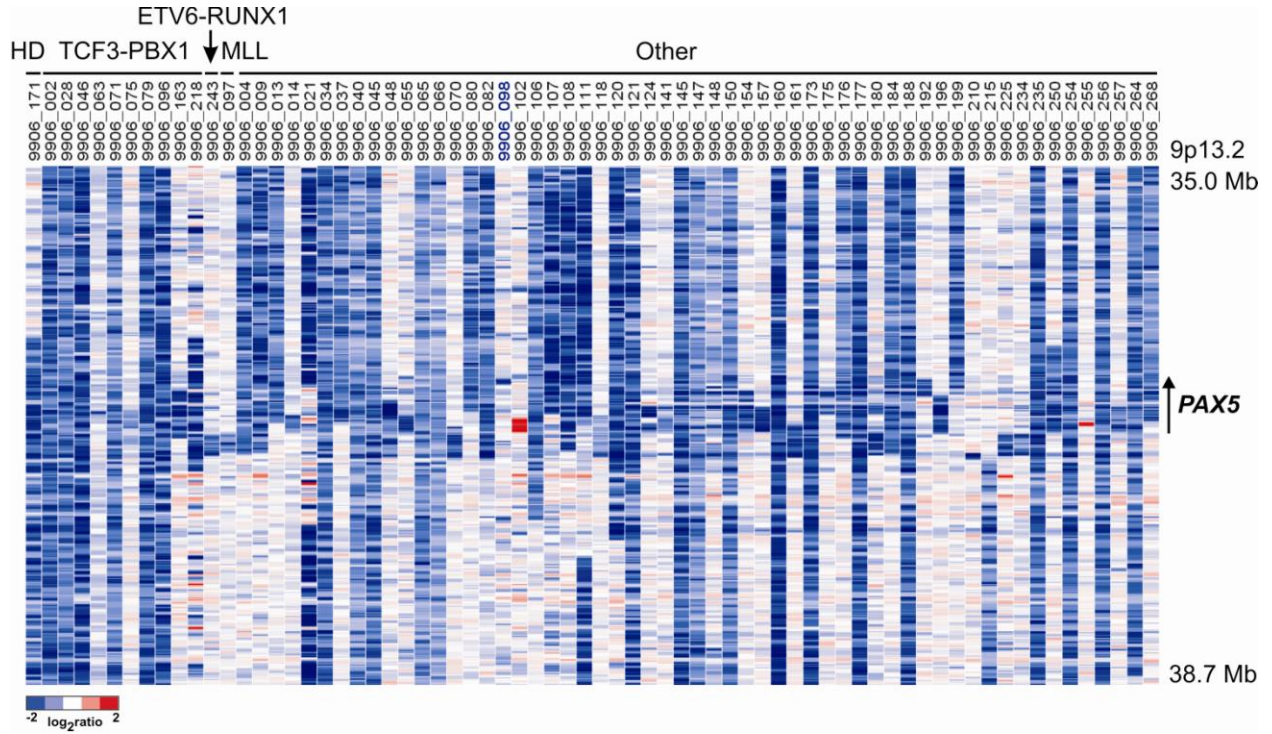
Gene	Running enrichment score	Core enrichment
<i>GIMAP6</i>	0.0992	YES
<i>CD52</i>	0.178	YES
<i>CLEC2B</i>	0.254	YES
<i>NRIP1</i>	0.305	YES
<i>CRHBP</i>	0.36	YES
<i>HIF0</i>	0.415	YES
<i>MMRN1</i>	0.448	YES
<i>ZFAND3</i>	0.478	YES
<i>TIPARP</i>	0.516	YES
<i>MLLT3</i>	0.53	YES
<i>SPINK2</i>	0.55	YES
<i>ROBO4</i>	0.572	YES
<i>GUCY1B3</i>	0.602	YES
<i>HOXB6</i>	0.61	YES
<i>BST2</i>	0.627	YES
<i>COX6B1</i>	0.577	NO
<i>HIST1H1C</i>	0.586	NO
<i>MPL</i>	0.561	NO
<i>GPC5</i>	0.555	NO
<i>CYBRD1</i>	0.516	NO
<i>GUCY1A3</i>	0.503	NO
<i>KLF2</i>	0.504	NO
<i>HIST1H2BG</i>	0.447	NO
<i>HIST1H2AE</i>	0.355	NO
<i>RHOQ</i>	0.345	NO
<i>NPR3</i>	0.335	NO
<i>HLA-E</i>	0.315	NO
<i>C5ORF23</i>	0.308	NO
<i>C4ORF18</i>	0.278	NO
<i>HOXB3</i>	0.278	NO
<i>TEAD2</i>	0.194	NO
<i>PGM5</i>	0.2	NO
<i>PCDH9</i>	0.148	NO
<i>BIRC3</i>	0.158	NO
<i>ATP8B4</i>	0.0388	NO

**Supplementary Table 27. Genes driving negative enrichment of the B-cell signal transduction gene set in P9906 high-risk ALL.**

Gene	Running enrichment score	Core enrichment
<i>AKT1</i>	0.0279	YES
<i>AKT2</i>	-0.021	YES
<i>LYN</i>	-0.0703	YES
<i>PIK3R1</i>	-0.115	YES
<i>BTK</i>	-0.15	YES
<i>ITPKB</i>	-0.192	YES
<i>SYK</i>	-0.234	YES
<i>CD19</i>	-0.268	YES
<i>NFKBIB</i>	-0.3	YES
<i>NFKB2</i>	-0.336	YES
<i>GRB2</i>	-0.372	YES
<i>PLCG2</i>	-0.406	YES
<i>SHC1</i>	-0.43	YES
<i>RAF1</i>	-0.46	YES
<i>NFKB1</i>	-0.466	YES
<i>SOS2</i>	-0.436	NO
<i>AKT3</i>	-0.438	NO
<i>DAG1</i>	-0.456	NO
<i>PPP1R13B</i>	-0.47	NO
<i>PIK3CD</i>	-0.451	NO
<i>SOS1</i>	-0.394	NO
<i>VAV1</i>	-0.36	NO
<i>NFKBIL2</i>	-0.362	NO
<i>MAPK1</i>	-0.274	NO
<i>NFKBIE</i>	-0.247	NO
<i>MAP2K1</i>	-0.25	NO
<i>EPHB2</i>	-0.184	NO
<i>PIK3CA</i>	-0.175	NO
<i>NFAT5</i>	-0.154	NO
<i>BCR</i>	-0.153	NO
<i>BAD</i>	-0.162	NO
<i>SERPINA4</i>	-0.1	NO
<i>PI3</i>	-0.0799	NO
<i>MAP2K2</i>	-0.00261	NO
<i>ITPKA</i>	-0.00194	NO
<i>NFKBIL1</i>	-0.0269	NO
<i>BLNK</i>	-0.0225	NO
<i>NFKBIA</i>	-0.0533	NO
<i>CSK</i>	-0.0845	NO

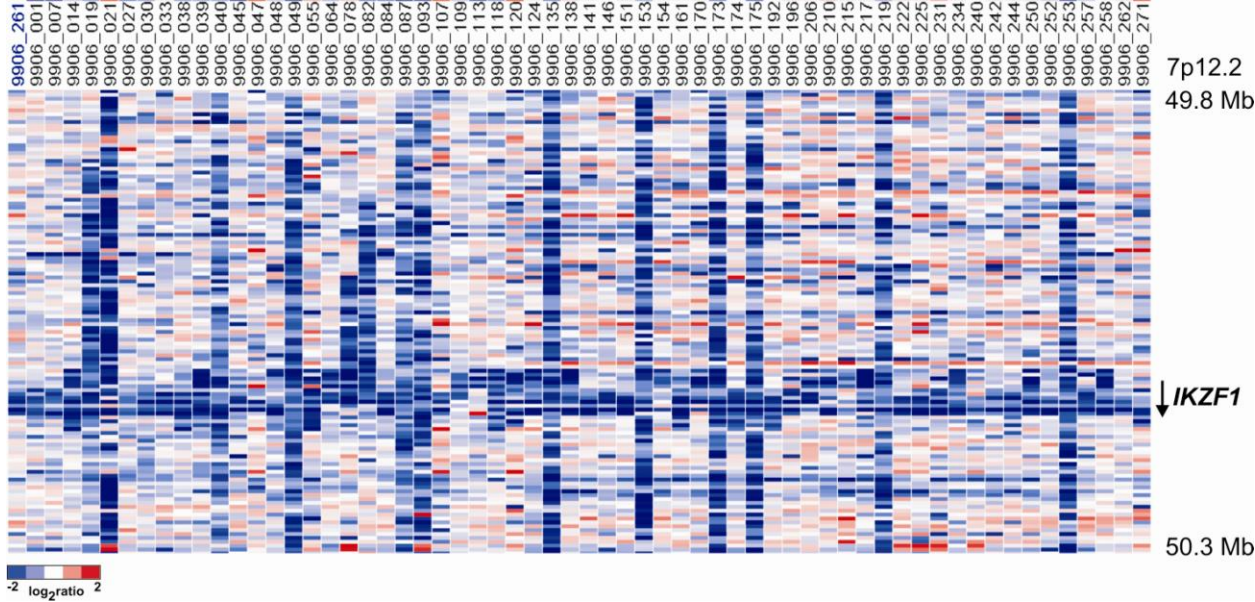
### Supplementary Figure 1. *PAX5* deletions in P9906 ALL.

Raw log ratio copy number at the *PAX5* locus is shown for all cases with an *IKZF1* CNA. Blue is deletion, and red gain. HD, hyperdiploid.



**Supplementary Figure 2. *IKZF1* deletions in P9906 ALL.**

Raw log ratio copy number at the *IKZF1* locus is shown for all cases with an *IKZF1* CNA.

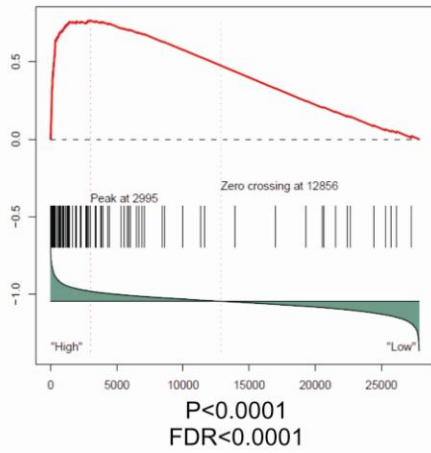


### Supplementary Figure 3. Gene set enrichment analysis (GSEA) of poor outcome P9906 ALL, poor outcome St Jude ALL, and *BCR-ABL1* positive St Jude ALL.

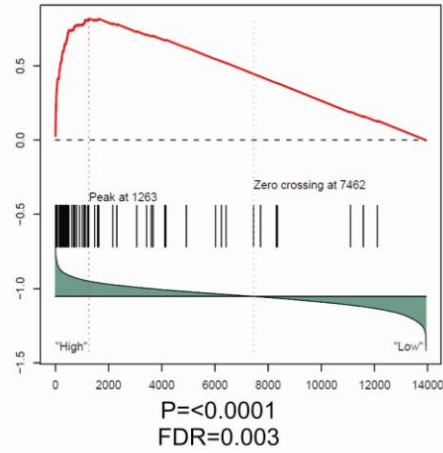
**A**, Genes are ranked (bottom of panel, green) based on correlation between expression and class distinction (here SPC predicted poor outcome v non-poor outcome). GSEA then determines if the members of a gene set (here a gene set of the top 100 upregulated genes in St Jude poor outcome ALL) are randomly distributed in the ranked gene list, or primarily found at the top or bottom. Occurrences of members of the gene set in the ranked gene list are shown as vertical black lines above the ranked signature. An enrichment score *ES* is calculated that reflects the degree to which a gene set is overrepresented at the top or bottom of the entire ranked list. The *ES* is a running sum, Kolmogorov-Smirnov-like statistic calculated by walking down list *L* and increasing the statistic when a gene in *S* is encountered, and decreasing it when it is not. The magnitude of the increment depends on the strength of association with phenotype, and the *ES* is the maximum deviation from zero encountered in the random walk, and is depicted as a red curve. The "leading edge" genes are those members of the gene set responsible for the observed enrichment, and are those hits occurring to the left of the vertical dotted red line. The significance level of *ES* is calculated by phenotype-based permutation testing, and when a database of gene sets are evaluated, as in this analysis, the significance level is adjusted for multiple hypothesis testing by calculation of a false discovery rate (FDR). Here there is highly significant enrichment of the St Jude poor outcome upregulated gene set in the P9906 poor outcome signature. **B**, enrichment of the P9906 poor outcome upregulated gene set in the St Jude poor outcome signature. These analyses demonstrate similarity between the signatures of P9906 and St Jude poor outcome ALL. **C**, enrichment of the P9906 poor outcome upregulated gene set in St Jude *BCR-ABL1* positive ALL, demonstrating similarity of P9906 poor outcome (*BCR-ABL1* negative) and St Jude *BCR-ABL1* positive signatures. **D**, heatmap of St Jude ALL and P9906 poor outcome upregulated genes, corresponding to the GSEA plot in C. B-A, *BCR-ABL1* positive; E-R, *ETV6-RUNX1* positive; H50, high hyperdiploid; Hypo, hypodiploid; T-P, *TCF3-PBX1*. Increased expression genes of the P9906 poor outcome gene set is seen in *BCR-ABL1* ALL; "leading edge" genes responsible for the enrichment are shown at the right of the panel. **E**, positive enrichment of genes upregulated in human hematopoietic stem cells (CD34+/CD38-/Lin- cells from bone marrow, umbilical cord blood and peripheral blood stem-prognitor cells compared to the stem cell-depleted populations)<sup>37</sup> in poor outcome P9906 ALL. Positive enrichment of this gene set in was also observed in poor outcome St Jude ALL (P=0.0023, FDR=0.03). **F**, negative enrichment of B cell antigen receptor/signal transduction genes ([http://stke.sciencemag.org/cgi/cm/stkecm;CMP\\_6909](http://stke.sciencemag.org/cgi/cm/stkecm;CMP_6909))<sup>38</sup> in P9906 poor outcome ALL. The ranked gene list compares high risk cases that relapse to low risk cases that do not relapse.

**A**

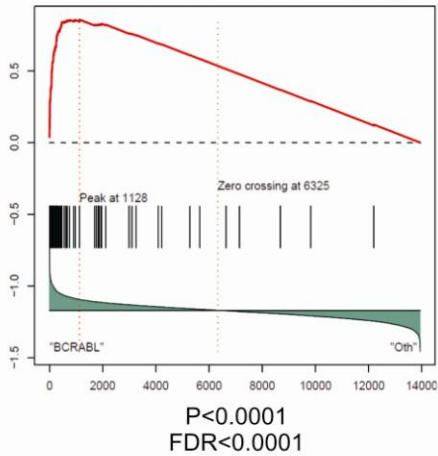
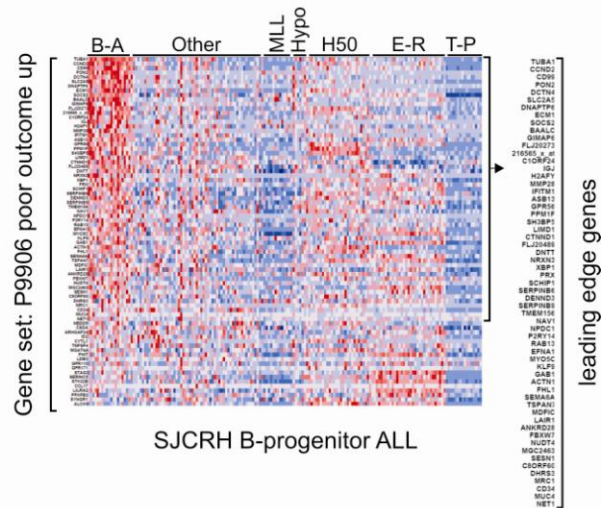
Ranked gene list: P9906 poor outcome  
Gene set: St Jude poor outcome upregulated

**B**

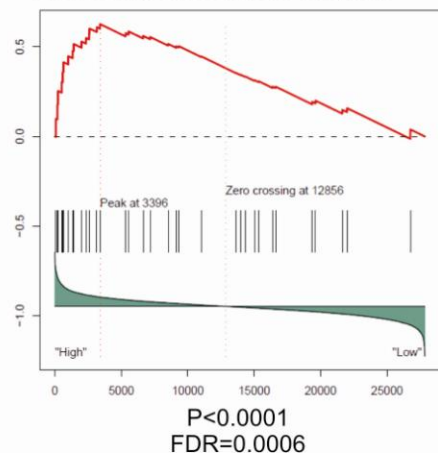
Ranked gene list: St Jude poor outcome  
Gene set: P9906 poor outcome upregulated

**C**

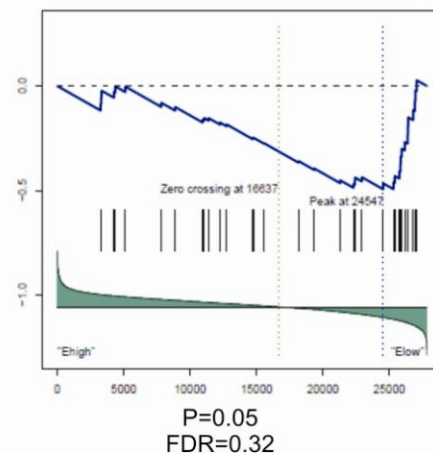
Ranked gene list: St Jude *BCR-ABL1*+ ALL  
Gene set: P9906 poor outcome upregulated

**D****E**

Ranked gene list: P9906 poor outcome  
Gene set: Hematopoietic stem cell

**F**

Ranked gene list: P9906 poor outcome  
Gene set: ST\_B\_cell\_antigen\_receptor



## REFERENCES

1. Nachman JB, Sather HN, Sensel MG, et al. Augmented post-induction therapy for children with high-risk acute lymphoblastic leukemia and a slow response to initial therapy. *N Engl J Med* 1998;338:1663-71.
2. Shuster JJ, Camitta BM, Pullen J, et al. Identification of newly diagnosed children with acute lymphocytic leukemia at high risk for relapse. *Cancer Res Ther and Control* 1999;9:101-7.
3. Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 2007;446:758-64.
4. Mullighan CG, Miller CB, Radtke I, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature* 2008;453:110-4.
5. Pui CH, Boyett JM, Rivera GK, et al. Long-term results of Total Therapy studies 11, 12 and 13A for childhood acute lymphoblastic leukemia at St Jude Children's Research Hospital. *Leukemia* 2000;14:2286-94.
6. Pui CH, Sandlund JT, Pei D, et al. Improved outcome for children with acute lymphoblastic leukemia: results of Total Therapy Study XIII B at St Jude Children's Research Hospital. *Blood* 2004;104:2690-6.
7. Kishi S, Griener J, Cheng C, et al. Homocysteine, pharmacogenetics, and neurotoxicity in children with leukemia. *J Clin Oncol* 2003;21:3084-91.
8. Pui CH, Relling MV, Sandlund JT, Downing JR, Campana D, Evans WE. Total Therapy study XV for newly diagnosed childhood acute lymphoblastic leukemia: study design and preliminary results. *Ann Hematol* 2006;85 Suppl 1:88-91.
9. Pieters R, Schrappe M, De Lorenzo P, et al. A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. *Lancet* 2007;370:240-50.
10. Lin M, Wei LJ, Sellers WR, Lieberfarb M, Wong WH, Li C. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* 2004;20:1233-40.
11. Pounds SB, Cheng C, Mullighan CG, Raimondi SC, Shurtleff SA, Downing JR. Reference Alignment of SNP Microarray Signals for Copy Number Analysis of Tumors. *Bioinformatics* 2008; doi:10.1093/bioinformatics/btn624.
12. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* 2007;23:657-63.
13. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 1998;8:175-85.
14. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 1998;8:186-94.
15. Zhang J, Wheeler DA, Yakub I, et al. SNPdetector: a software tool for sensitive and accurate SNP detection. *PLoS Comput Biol* 2005;1:e53.

16. Zhang J, Finney RP, Rowe W, et al. Systematic analysis of genetic alterations in tumors using Cancer Genome WorkBench (CGWB). *Genome Res* 2007;17:1111-7.
17. Zhang J, Rowe WL, Struewing JP, Buetow KH. HapScope: a software system for automated and visual analysis of functionally annotated haplotypes. *Nucleic Acids Res* 2002;30:5213-21.
18. Bamford S, Dawson E, Forbes S, et al. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer* 2004;91:355-8.
19. Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;29:308-11.
20. Garvie CW, Hagman J, Wolberger C. Structural studies of Ets-1/Pax5 complex formation on DNA. *Mol Cell* 2001;8:1267-76.
21. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 1991;47(Pt 2):110-9.
22. DeLano WL. The PyMOL Molecular Graphics System. <http://www.pymol.org>. 2002.
23. Bair E, Tibshirani R. Semi-supervised methods to predict patient survival from gene expression data. *PLoS Biol* 2004;2:E108.
24. Bair E, Hasle H, Debashis P, Tibshirani R. Prediction by Supervised Principal Components. *J Am Stat Assoc* 2006;101:119-37.
25. Gray RJ. A class of  $K$ -sample tests for comparing the cumulative incidence of a competing risk. *Annals Statistics* 1988;16:1141-54.
26. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br J Cancer* 1977;35:1-39.
27. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep* 1966;50:163-70.
28. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc* 1999;94:496-509.
29. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing 2006;Vienna:Austria.
30. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;3:Article3.
31. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;5:R80.
32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995;57:289-300.
33. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545-50.

34. Efron B, Tibshirani R. On testing the significance of sets of genes. *Ann Appl Stat* 2007;1:107-29.
35. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002;30:207-10.
36. Barrett T, Troup DB, Wilhite SE, et al. NCBI GEO: mining tens of millions of expression profiles--database and tools update. *Nucleic Acids Res* 2007;35:D760-5.
37. Georgantas RW, 3rd, Tanadve V, Malehorn M, et al. Microarray and serial analysis of gene expression analyses identify known and novel transcripts overexpressed in hematopoietic stem cells. *Cancer Res* 2004;64:4434-41.
38. Buhl AM, Nemazee D, Cambier JC, Rickert R, Hertz M. B-cell antigen receptor competence regulates B-lymphocyte selection and survival. *Immunol Rev* 2000;176:141-53.