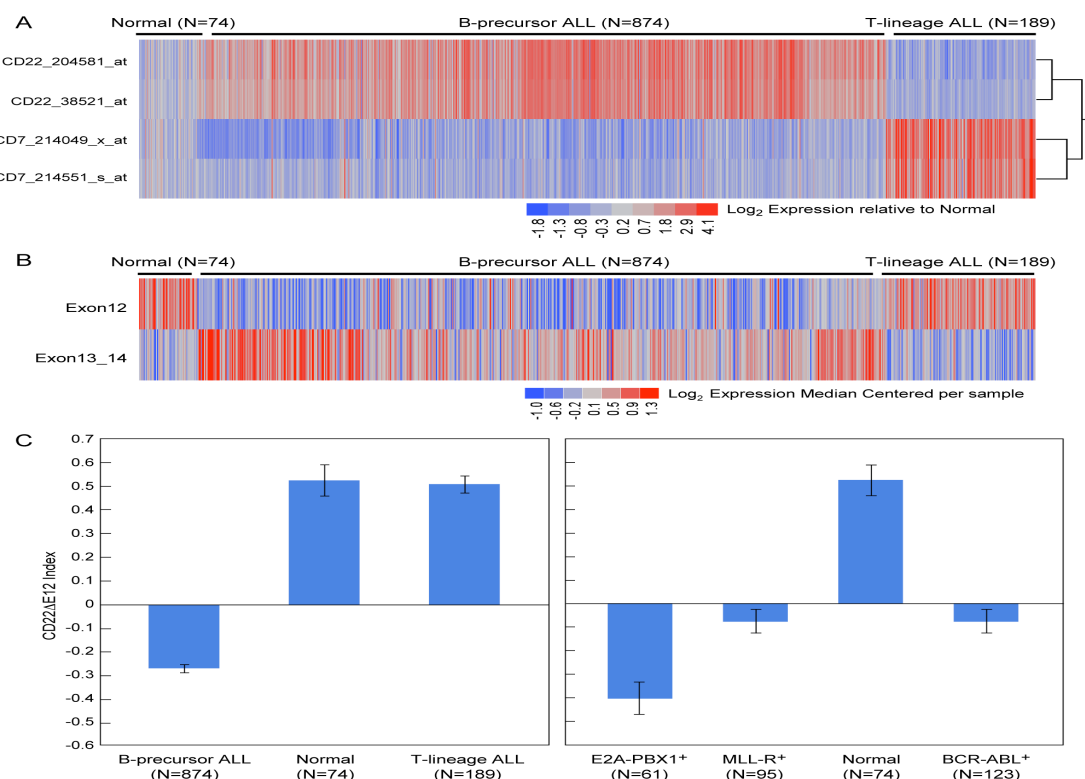**SUPPLEMENT**

Figure S1. Gene Expression Analysis Using Multiple Probesets to Examine CD22 expression and the Incidence of the CD22ΔE12 genetic defect in B-lineage lymphoid malignancies. [A] We compiled archived RMA normalized expression values for CD22 and CD7 probesets in newly diagnosed B-precursor ALL (BPL), newly diagnosed T-lineage ALL and normal hematopoietic cells. Heat map depicts the up- and down-regulated transcripts ranging from blue to red respectively and clustered according to average distance metric. Two probesets for CD22 were highly correlated exhibiting increased expression in BPL samples and the 2 probesets for CD7 were strongly co-regulated with increased expression in T-lineage ALL samples. Probeset expression values for BPL and T-lineage ALL were compared to normal hematopoietic cells using Students T-test (Unequal variance correction). CD22_204581_at, CD22_38521_at, CD7_214551_s_at and CD7_214049_x_at exhibited 3.90 (P = 7.2×10^{-46}), 3.50 (P = 4.2×10^{-50}), 0.78 (P = 1.1×10^{-7}) and 0.67 (P = 9.4×10^{-11})

fold difference in BPL compared to normal samples respectively. CD22_204581_at, CD22_38521_at, CD7_214551_s_at and CD7_214049_x_at exhibited 0.68 (P = 2.0 x 10^{-10}), 0.74 (P = 4.3 x 10^{-9}), 4.89 (P = 4.8 x 10^{-51}) and 6.21 (P = 1.8 x 10^{-61}) fold difference in T-lineage ALL compared to Normal samples respectively. **[B]** Probeset 217422_s_at was analyzed to compare differential expression of Exon 12 and Exon 13-14 in BPL cells, T-lineage ALL cells and normal hematopoietic cells. Significant differences were observed across patient subgroups for Exon 12 (ANOVA, $F_{2,1134}$ = 305.5, P<0.0001) and Exon 13-14 expression (ANOVA, $F_{2,1134}$ = 157.5, P<0.0001). Dunnett's post hoc tests revealed significant reduction in Exon 12 transcript levels in BPL samples (Normal = 0.586 ± 0.058 versus BPL = -0.215 ± 0.016, P<0.0001), but not normal versus T-lineage ALL (Mean = 0.579 ± 0.031, P = 0.99). Expression of Exons 13-14 exhibited a significant increase in BPL (Normal = -0.062 ± 0.039 versus BPL = 0.407 ± 0.017, P<0.0001) and a borderline significant decrease in T-lineage ALL (Mean = -0.200 ± 0.025, P = 0.045). **[C]** A CD22ΔE12 Index was calculated by subtracting the median centered expression values of the 7 probes for CD22 Exons 10-11 and CD22 Exons 13-14 in the 217422_s_at probeset from the median centered expression values of the 3 CD22 Exon 12 probes. Depicted are bar charts for the CD22ΔE12 index values comparing normal samples (N=74) versus leukemic samples from patients with BPL (N=874) versus T-lineage ALL (N=189) compiled from 3 studies, GSE11877, GSE13159 and GSE13351 (**left panel**), and 3 genetically defined high-risk BPL groups (E2A-PBX1[+], N=61; MLL-R[+], N=95; and BCR-ABL[+], N=123) (right panel) were compared to normal hematopoietic cells from non-leukemic control samples included in the same studies (N=74).
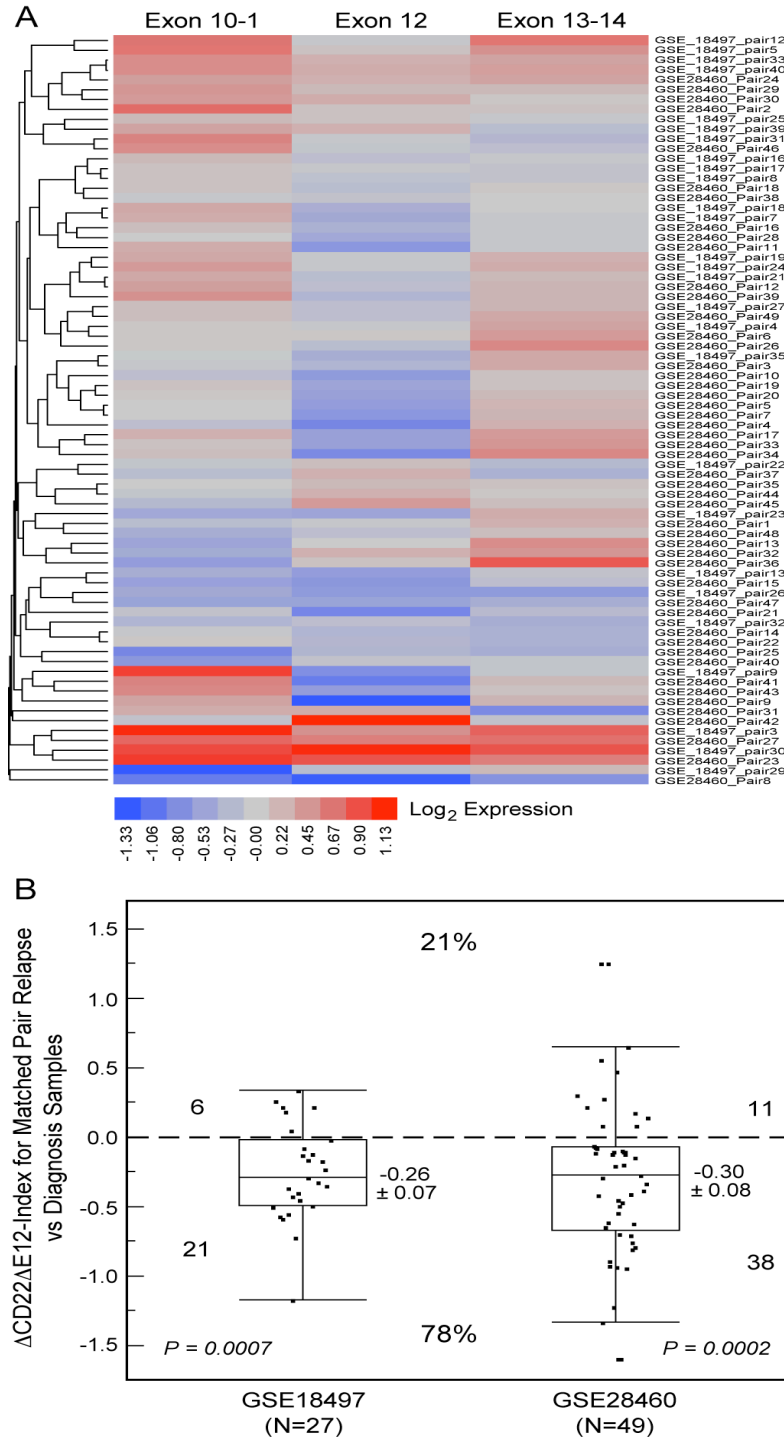
2

Figure S2. Gene Expression Analysis Using Multiple Probesets to Examine CD22 expression and the Incidence of the CD22ΔE12 genetic defect in relapsed BPL. **[A]** We compiled archived RMA normalized expression values for probe-specific CD22 expression levels for 217422_s_at probes in primary leukemia cells from 74 matched pair initial diagnosis vs. relapse bone marrow specimens

from 2 studies (GSE28460, 49 pairs and GSE18497, 27 pairs). For each sample, probe level log2 RMA normalized expression values were median centered to expression across 11 probes in the probeset deposited on the Affymetrix U133 plus 2 genechip for the CD22 receptor, 217422_s_at. Depicted are the mean levels of expression for each probe subtracting expression at Relapse from the expression at Diagnosis for each matched pair case across 2 studies (Probes 2-4 were specific to Exons 10-11, probes 5-6 were specific for Exon 12 and probes 8-11 were specific to Exons 13-14). Probes for Exon 12 demonstrated reduced expression compared to the probes for Exons 10-11 and Exons 13-14 consistent with a selective reduction of CD22 Exon 12 transcript levels at relapse for the matched pair samples. The heat map depicts the up- and down-regulated transcripts ranging from red to blue respectively and clustered according to average distance metric. Probe specific expression levels for 217422_s_at probes illustrated the reduced expression of CD22 exon 12 probes (probes 5-7) comparing expression at relapse versus diagnosis for 76 matched pair samples compiled from 2 studies (GSE28460, GSE18497), but not for probes that interrogated Exons 10-11 or 13-14. **[B]** A CD22ΔE12 Index was calculated by subtracting the median centered expression values of the 7 probes for CD22 Exons 10-11 and CD22 Exons 13-14 in the 217422_s_at probeset from the median centered expression values of the 3 CD22 Exon 12 probes. The distribution of the differences in the CD22ΔE12-index values between matched pair samples (i.e. CD22ΔE12-index at relapse minus the CD22ΔE12-index of the primary BPL cells from the same patients when they were initially diagnosed) taken at relapse minus diagnosis for each study are depicted as dot plots overlaid with Box-Whisker plots. In 59/76 cases (78%), the CD22ΔE12-index values were lower at relapse than at initial diagnosis (21/27 paired samples for GSE18497, Paired T-test, P = 0.0007, mean matched pair difference = -0.28 ± 0.07, Df = 26, T value = 3.86) ; 38/49 paired samples for GSE28460, Paired T-test, P = 0.0002, mean matched pair difference = -0.30 ± 0.08, Df = 48, T value = 4.03). Median and inter-quartile ranges for the matched pair differences for GSE18497 and GSE28460 were -0.289 (-0.50 - -0.02) and -0.276 (-0.67- -0.068) respectively.
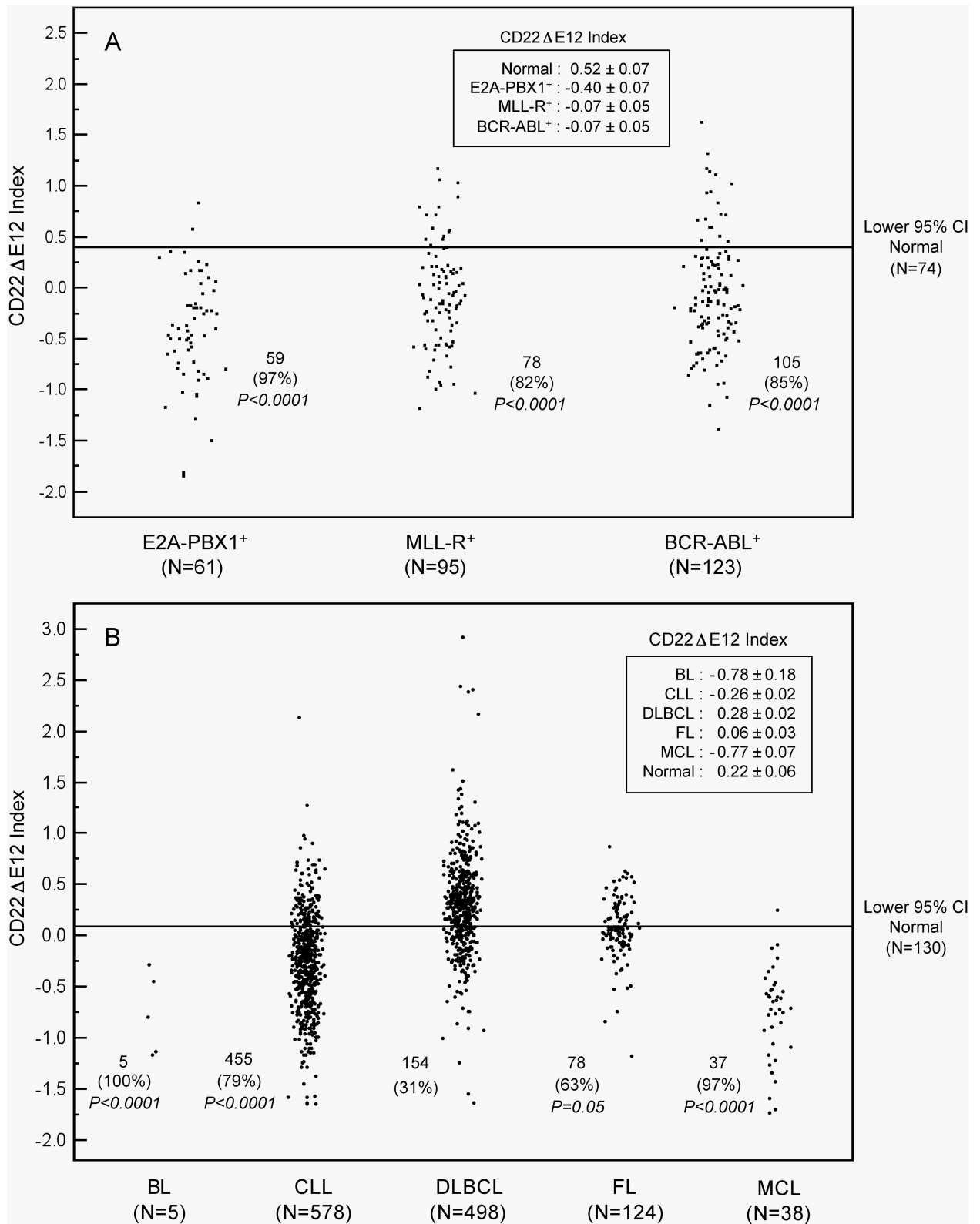
A

CD22 ΔE12 Index

Normal : 0.52 ± 0.07
E2A-PBX1⁺ : -0.40 ± 0.07
MLL-R⁺ : -0.07 ± 0.05
BCR-ABL⁺ : -0.07 ± 0.05

CD22 ΔE12 Index

Lower 95% CI
Normal
(N=74)

59
(97%)
P<0.0001

78
(82%)
P<0.0001

105
(85%)
P<0.0001

E2A-PBX1⁺
(N=61)

MLL-R⁺
(N=95)

BCR-ABL⁺
(N=123)

B

CD22 ΔE12 Index

BL : -0.78 ± 0.18
CLL : -0.26 ± 0.02
DLBCL : 0.28 ± 0.02
FL : 0.06 ± 0.03
MCL : -0.77 ± 0.07
Normal : 0.22 ± 0.06

CD22 ΔE12 Index

Lower 95% CI
Normal
(N=130)

5
(100%)
P<0.0001

455
(79%)
P<0.0001

154
(31%)

78
(63%)
P=0.05

37
(97%)
P<0.0001

BL
(N=5)

CLL
(N=578)

DLBCL
(N=498)

FL
(N=124)

MCL
(N=38)

5

Fig. S3. Incidence of CD22ΔE12 in Newly Diagnosed B-lineage lymphoid malignancies. BLAT analysis on CD22 probe sequences for probeset 217422_s_at (covering exons 10-14) deposited in the Affymetrix NetAffx™ Analysis Center (http://www.affymetrix.com/analysis/index.affx) were mapped onto specific CD22 exons and visualized using the UCSC genome browser (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start). An Exon 12 Index was calculated by subtracting the median centered expression values of the 7 probes for Exons 10-11 and Exons 13-14 from the median centered expression values of the 3 Exon 12 probes. **[A]** Depicted are dot plots of the sample level CD22ΔE12-index values in leukemic samples (compiled from GSE11877, GSE13159 and GSE13351) from 3 genetically defined subsets of newly diagnosed pediatric BPL patients referenced against the 95% lower confidence interval for the CD22ΔE12-index values for the 74 normal samples (95% CI = 0.394-0.655). Significant differences were observed comparing Normal with E2A-PBX1[+], MLL-R[+] and BCR-ABL[+] subsets (ANOVA, $F_{3,349}$ = 36.6, P<0.0001; Dunnett's post hoc P<0.0001 for all comparisons with normal). **[B]** Depicted are dot plots of the sample level CD22ΔE12-index values in leukemia/lymphoma samples (compiled from GSE13159, GSE39671 GSE10846, GSE12195, GSE12453, GSE12453, GSE53820, GSE36000) from adult patients with newly diagnosed B-lineage lymphoid malignancies referenced against the 95% lower confidence interval for the CD22ΔE12-index values for the 130 normal B-cell samples (95% CI = 0.094-0.349). Reductions in CD22 exon 12 expression levels were observed for 79% of CLL patients (N=578, P<0.0001) and 97% of MCL patients (N=38, P<0.0001). ANOVA of these patient groups also yielded significant differences between groups ($F_{4,1363}$ = 114.4, P<0.0001; Dunnett's post hoc P<0.0001 for CLL and MCL comparisons with normal, P=0.026 for FL versus normal and P=0.46 for DLBCL versus normal).
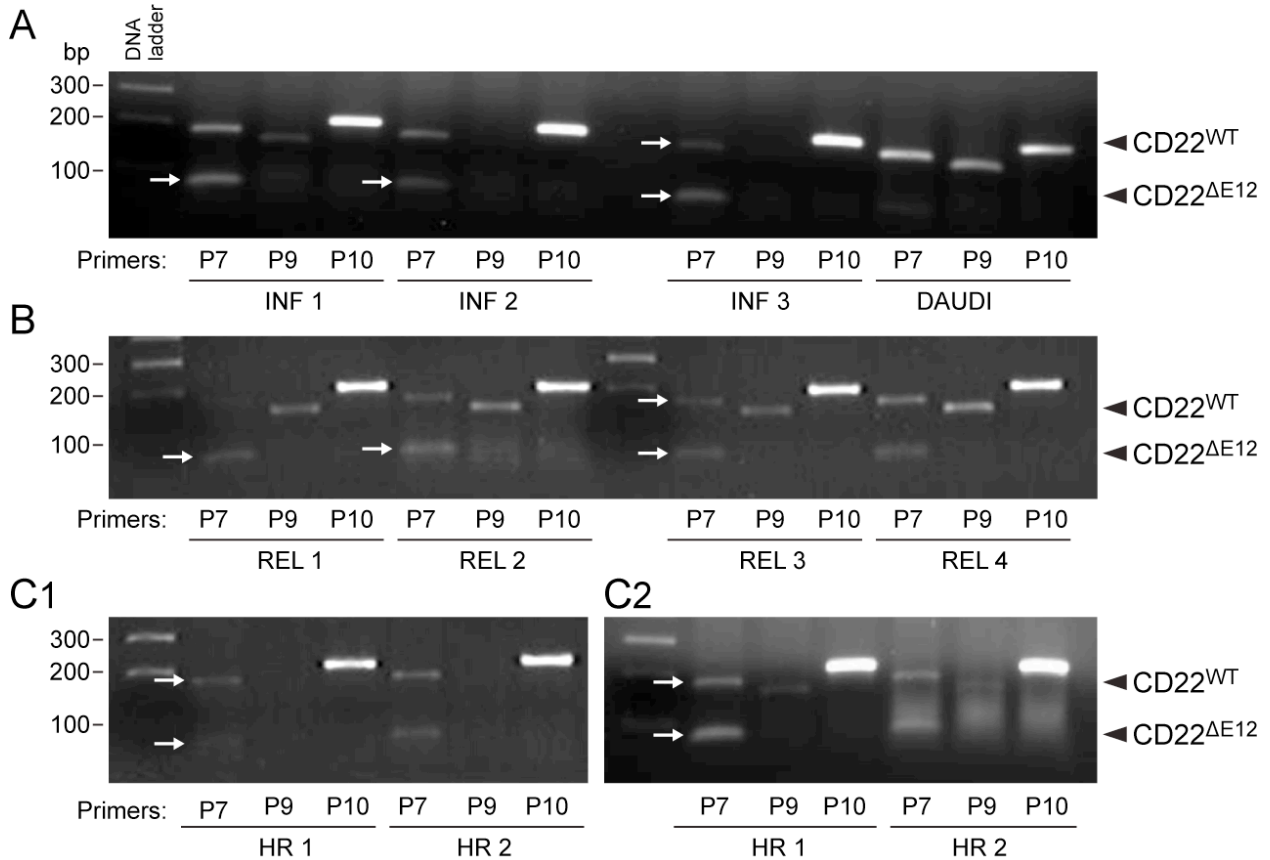
Figure S4. Primary Leukemic Cells in Initial Diagnostic Bone Marrow Specimens from Pediatric ALL Patients are CD22ΔE12-Positive. Total RNA was reversely transcribed and directly amplified in the same tube using the 3 specific primer sets P7, P9 and P10. The P10 primer set was used as a positive control of RNA integrity to amplify a 213-bp region (c.433–c.645) of Exon 4 of the CD22 cDNA present in both wildtype CD22 and CD22ΔE12 mRNA species. The P7 primer set was used to amplify a 182-bp region (c.2180–c.2361) of the CD22 cDNA extending from Exon 11 to Exon 13 and spanning the entire Exon 12. Deletion of Exon 12 results in a smaller PCR of product of 63-bp size with this primer set. The P9 primer set (Forward: CGAGATGAACATACCACGA, Reverse: TGAATCCCCTCATCTTCTGG) was used to amplify a 160-bp region (c.2304–c.2463) of the CD22 cDNA extending from Exon 12 to Exon 14 and spanning the entire Exon 13. Consequently, no PCR product can be obtained with this primer set from CD22 cDNA with the Exon 12 deletion. PCR products were separated on a 2% agarose gel containing ethidium bromide. Gel images were taken with an

UVP digital camera and UV light in an Epi Chemi II Darkroom using the LabWorks Analysis software (UVP, Upland, CA). Depicted are the agarose gel images of the PCR products of primary lekemic cells from (i) 3 infants (INF) with ALL as well as the CD22ΔE12-negative Burkitt's lymphoma/leukemia cell line DAUDI (Panel A), (ii) 4 pediatric patients with relapsed (REL) BPL (Panel B), and (iii) pediatric patients with newly diagnosed high-risk (HR) BPL (Panels C1 & C2).
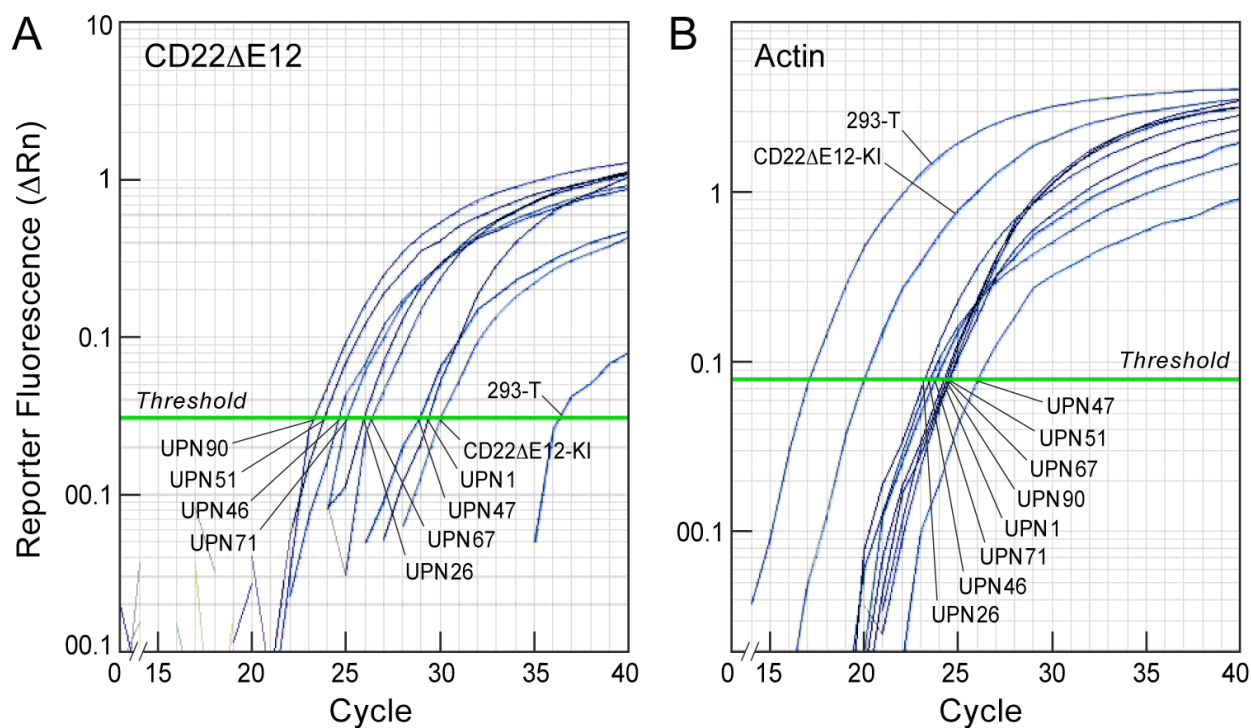
Figure S.5 Detection of CD22ΔE12 mRNA in ALL Cells Using Quantitative RT-PCR. We used the Applied Biosystems 7900HT Fast Real-Time PCR System. The PCR primer pair was designed to amplify a 113-bp fragment spanning from Exon 11 to Exon 13 of the human CD22 cDNA. The amplified fragment was then specifically annealed to a pre-mixed oligo DNA probe (5'-TGTGAGGAATAAAAAGAGATGCAGAGTCC-3') conjugated with 5' FAM reporter and 3'BHQ Quencher on the CD22ΔE12-specific unique junction region between Exon 11 and Exon 13. Actin RT-PCR was also performed in parallel. Depicted are representative cycle number vs. Log (ΔRn) (fluorescence signal) amplification plots obtained using the CD22ΔE12-specific test primers and Actin-specific reference primers for splenocytes from homozygous CD22ΔE12-KI mice (positive control), 293-T cells (negative control) and primary leukemic cells from 8 patients with high-risk BPL.
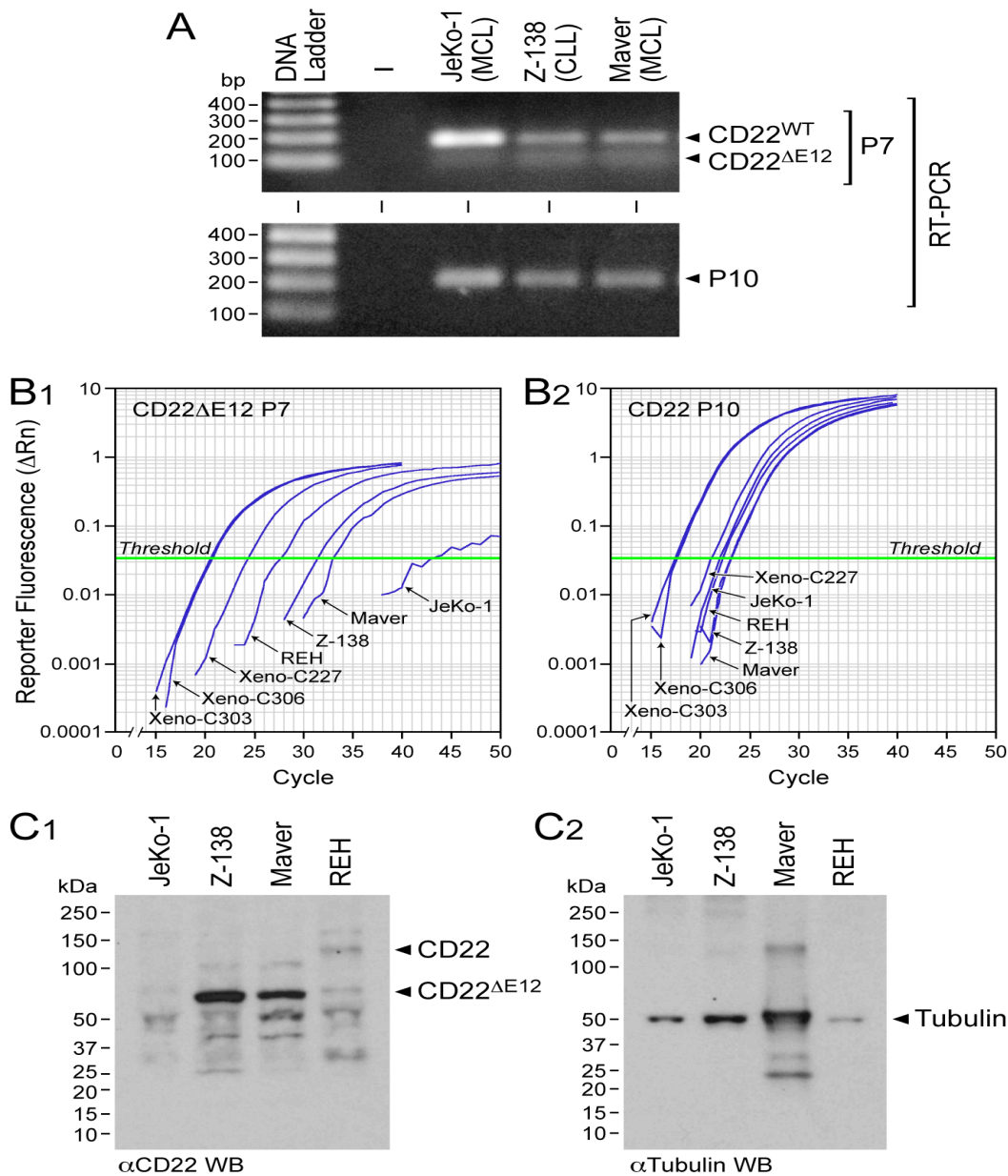
A

bp

DNA Ladder — JeKo-1 (MCL) — Z-138 (CLL) — Maver (MCL)

400
300
200
100

◄ CD22^WT
◄ CD22^ΔE12 } P7

400
300
200
100

◄ P10

} RT-PCR

B1 CD22ΔE12 P7

Reporter Fluorescence (ΔRn)

Threshold

Maver
JeKo-1
Z-138
REH
Xeno-C227
Xeno-C306
Xeno-C303

Cycle

B2 CD22 P10

Threshold

Xeno-C227
JeKo-1
REH
Z-138
Maver
Xeno-C306
Xeno-C303

Cycle

C1

kDa   JeKo-1 Z-138 Maver REH

250
150
100

◄ CD22
◄ CD22^ΔE12

50
37
25
20
15
10

αCD22 WB

C2

kDa   JeKo-1 Z-138 Maver REH

250
150
100

50   ◄ Tubulin
37
25
20
15
10

αTubulin WB

Figure S6. Identification of CD22ΔE12$^+$ CLL and MCL Cells. [A] RT-PCR detection of abnormal CD22 mRNA species with Exon 12 deletion. The P7 primer set was used to amplify a 182-bp region (c.2180–c.2361) of the CD22 cDNA extending from Exon 11 to Exon 13 and spanning the entire Exon

10

12. Deletion of Exon 12 results in a smaller CD22ΔE12-specific PCR product of 63-bp size with this primer set[4]. The P10 primer set was used as a positive control of RNA integrity to amplify a 213-bp region (c.433–c.645) of the CD22 cDNA present in both wildtype CD22 and CD22ΔE12 mRNA species. **[B] Detection of CD22ΔE12 mRNA in CLL and MCL Cells Using Quantitative RT-PCR.** We used the Applied Biosystems 7900HT Fast Real-Time PCR System. The PCR primer pair was designed to amplify a 113-bp fragment spanning from Exon 11 to Exon 13 of the human CD22 cDNA. The amplified fragment was then specifically annealed to a pre-mixed oligo DNA probe (5'-TGTGAGGAATAAAAAGAGATGCAGAGTCC-3') conjugated with 5' FAM reporter and 3'BHQ Quencher on the CD22ΔE12-specific unique junction region between Exon 11 and Exon 13. Actin RT-PCR was also performed in parallel. Depicted are representative cycle number vs. Log (ΔRn) (fluorescence signal) amplification plots obtained using the CD22ΔE12-specific test primers and Actin-specific reference primers for 2 MCL and one CLL cell line as well as 3 BPL xenograft clones and the BPL cell line REH that were included as positive controls. In control reactions we used the P10 primer set to confirm the RNA integrity by amplifying a 213-bp region (c.433–c.645) of the CD22 cDNA present in both wildtype CD22 and CD22ΔE12 mRNA species. Z-138 CLL cells as well as MAVER MCL cells expressed CD22ΔE12-mRNA at levels comparable to those of BPL xenograft clones and the BPL cell line REH. **[C] Western blot detection of a truncated (t) CD22 protein.** Whole cell lysates of MCL, CLL, and ALL cell lines were subjected to CD22 (Panel C.1) and Tubulin (Panel C.2) Western blot analysis. The positions of intact wild-type and truncated (t) mutant CD22 proteins are indicated with arrow heads. CLL cell line Z-138 and MCL cell line MAVER were characterized by abundant expression of the truncated CD22ΔE12 protein.

Table S1. Intronic Mutations of CD22 Gene in Primary BPL Cells

Intronic Mutations According to Standard HGVS Nomenclature (Corresponding SNP)

UPN	ΔCt	ΔΔCt	c.2327+128 A>G (rs4805119) HO	HE	c.2327+147 G>A (rs10406539) HO	HE	c.2327+203 C>G (rs10413500) HO	HE	c.2328-199 C>G (rs10413526) HO	HE	c.2328-195 A>G (rs4805120) HO	HE
42	5.7	-13.0	-	+	–	-	–	-	–	–	-	+
46	1.1	-17.6	+	–	+	-	+	-	–	–	+	–
66	1.7	-18.0	+	–	–	–	–	–	–	–	+	-
75	3.4	-15.3	–	+	–	-	–	-	–	–	+	-
90	-1.0	-19.7	–	+	–	–	–	–	–	-	+	-
94	4.9	-13.8	+	-	–	+	–	+	–	-	+	–
121	>12.3	-6.4	-	-	-	-	-	-	-	-	-	-
158	2.3	-16.4	-	+	-	+	-	+	-	+	+	-
168	0.2	-18.5	-	-	-	-	-	-	-	+	+	-
174	5.3	-13.4	-	+	-	+	-	+	-	+	-	+

PCR-based genomic sequence analysis of CD22 gene revealed mutations within a mutation-rich hotspot region of the intronic sequence between exons 12 and 13; HO = homozygous; HE = heterozygous; UPN: unique patient number. UPN 168 had high-risk BPL. All other patients had standard risk BPL.