Gene expression profiles predictive of outcome and age in infant acute lymphoblastic leukemia: a Children’s Oncology Group study

Huining Kang, PhD¹, Carla S. Wilson, MD, PhD¹, Richard C. Harvey, PhD¹, I-Ming Chen, DVM, MS¹,², Maurice H. Murphy, MS¹, Susan R. Atlas, PhD¹, Edward J. Bedrick, PhD¹, Meenakshi Devidas, PhD²,³, Andrew J. Carroll, PhD²,⁴, Blaine W. Robinson, PhD⁵, Ronald W. Stam, PhD⁶, Maria G. Valsecchi, PhD⁷, Rob Pieters, PhD⁶, Nyla A. Heerema, PhD²,⁸, Joanne M. Hilden, MD²,⁹, Carolyn A. Felix, MD²,⁵,¹⁰, Gregory H. Reaman, MD²,¹¹, Bruce Camitta, MD²,¹², Naomi Winick, MD²,¹³, William L. Carroll, MD²,¹⁴, ZoAnn E. Dreyer, MD²,¹⁵, Stephen P. Hunger, MD²,⁹ and Cheryl L. Willman, MD¹,²,*

¹University of New Mexico Cancer Center and Departments of Internal Medicine, Pathology, and Physics and Astronomy, University of New Mexico, Albuquerque, NM; ²Children’s Oncology Group, Arcadia, CA; ³Department of Biostatistics, College of Medicine, University of Florida, Gainesville, FL; ⁴Department of Genetics, University of Alabama at Birmingham, Birmingham, AL; ⁵Division of Oncology, Children’s Hospital of Philadelphia, Philadelphia, PA; ⁶Department of Pediatric Oncology/Hematology, Erasmus MC-Sophia Children’s Hospital, Rotterdam, The Netherlands; ⁷Department of Clinical Medicine and Prevention, University of Milano-Bicocca, Monza, Italy; ⁸Department of Pathology, The Ohio State University, Columbus, OH; ⁹Children’s Hospital Colorado and the Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO; ¹⁰Dept of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA; ¹¹School of Medicine and Health Sciences, George Washington University Children's National Medical Center, Washington, DC; ¹²Pediatric Hematology/Oncology/BMT, Medical College of Wisconsin, Wauwatosa, WI; ¹³Pediatric Hematology-Oncology, University of Texas Southwestern Medical Center, Dallas,
TX; 14. New York University Medical Center/NYU Cancer Institute, New York, NY; 15. Texas Children's Hospital, Houston, TX.

*Correspondence: cwillman@salud.unm.edu
Abstract

Gene expression profiling was performed on 97 cases of infant ALL from Children’s Oncology Group Trial P9407. Statistical modeling of an outcome predictor revealed three genes highly predictive of event-free survival (EFS), beyond age and MLL status: FLT3, IRX2 and TACC2. Low FLT3 expression was found in a group of infants with excellent outcome (n=11; 5-year EFS of 100%), while differential expression of IRX2 and TACC2 partitioned the remaining infants into two groups with significantly different survivals (5-year EFS of 16% versus 64%; \( P < .001 \)). When infants with MLL-AFF1 were analyzed separately, a seven-gene classifier was developed that split them into two distinct groups with significantly different outcomes (5-year EFS of 20% versus 65%; \( P < .001 \)). In this classifier, elevated expression of NEGR1 was associated with better EFS while IRX2, EPS8 and TPD52 expression were correlated with worse outcome. This classifier also predicted EFS in an independent infant ALL cohort from the Interfant-99 trial. When evaluating expression profiles as a continuous variable relative to patient age, we further identified striking differences in profiles in infants ≤90 and those >90 days of age. These age-related patterns suggest different mechanisms of leukemogenesis and may underlie the differential outcomes historically seen in these age groups.
Introduction

Acute lymphoblastic leukemia (ALL) arising in infants less than one year of age is an aggressive disease, with >50% of cases relapsing within 5 years of diagnosis.\textsuperscript{1,2} Established prognostic factors include patient age and the presence of genomic rearrangements of the \textit{Mixed Lineage Leukemia (MLL)} gene on chromosome 11q23, with children ≤90 days of age at diagnosis and those with \textit{MLL} rearrangements experiencing a particularly poor outcome.\textsuperscript{1,2} \textit{MLL} rearrangements occur in nearly 80% of infants with ALL, resulting predominantly in translocations of \textit{MLL} with a variety of partner genes including \textit{AFF1(AF4), MLLT1(ENL)} and \textit{MLLT3(AF9)}.\textsuperscript{3,4}

The mechanism of leukemogenesis in infants, in both \textit{MLL}-rearranged (\textit{MLL-R}) and \textit{MLL}-germline (\textit{MLL-G}) cases, is a subject of intensive investigation and may differ from leukemogenesis in older children.\textsuperscript{4-10} Studies of identical twins and retrospective analyses of neonatal blood spots have demonstrated that in many children, leukemia-initiating events may be acquired prenatally.\textsuperscript{5} While several of the frequently recurring translocations (\textit{ETV6-RUNX1} and \textit{RUNX1-MTG8}) associated with leukemia in older children have been demonstrated to arise \textit{in utero}, they appear to be insufficient for leukemogenesis, which requires the acquisition of additional cooperating genetic lesions. This multistep process may account for the longer latency period before development of clinically overt leukemia. In contrast, in infants, latency is very short and \textit{MLL-R} alone or in concert with only a few cooperating lesions may be sufficient for leukemogenesis. In support of this working hypothesis, whole genome studies of copy number variations or CGH arrays in \textit{MLL-R} infant ALL cases have demonstrated a very limited number of DNA copy-number alterations,\textsuperscript{4-6} in contrast to ALL arising in older children.\textsuperscript{11}
MLL belongs to the Trithorax Group (TRX-G) of proteins that regulate gene expression through chromatin association and modification. MLL binds to thousands of loci in the genome that are important for the regulation of hematopoiesis, cell signaling, and transcription. Important downstream targets of MLL include genes within the Homeobox A (HOXA) cluster and MEIS1. These and other genes involved in transcriptional activation are partially regulated through histone H3K4 methylation by the SET domain of wild type MLL. In contrast, leukemogenic MLL fusion proteins lack a SET domain, but retain the ability to bind and methylate H3K4, causing persistent activation of target genes that contribute to a distinctive pattern of gene expression characteristic of MLL-R leukemias. Thus, perturbed epigenetic regulation is likely to play a critical role in MLL-mediated leukemogenesis and further insights into this mechanism may provide new insights for therapeutic intervention.

The etiology and pathogenesis of infant ALL is complex and multifactorial, with critical in utero exposures, potential genetic susceptibility, MLL and other acquired genetic abnormalities, epigenetic dysregulation, and modifying influences all affecting the development and phenotype of the leukemia. As patterns of gene expression may be reflective of these combined factors, we performed gene expression profiling in a cohort of 97 infant ALL cases accrued to COG Trial P9407, the largest infant cohort examined to date. We wished to determine if we could identify genes that might improve risk classification and outcome prediction in infant ALL, beyond the well-established factors of MLL status and patient age. We further wished to determine whether these profiles might provide new insights into this disease, reflect cooperating genetic lesions or pathways, and serve as potential diagnostic and therapeutic targets.
Materials and Methods

Patient Selection and Characteristics

A total of 212 infants <365 days of age with ALL were enrolled onto COG P9407 study (NCT00002756, ClinicalTrials.gov) in 3 consecutive cohorts. A subset of 70 infants was enrolled between 1996 and 2000 on Cohorts 1 and 2 and another 142 infants were enrolled between 2001 and 2006 on Cohort 3. Infants in Cohorts 1 and 2 were observed to have unacceptably high induction death rates leading to modification of the treatment regimen for Cohort 3. Pre-treatment leukemia specimens were available from 97 of 212 of the cases accrued to this trial, predominantly from cohort 3. The pre-clinical and outcome characteristics of this cohort of 97 cases did not differ appreciably from the total 212 cases accrued (data not shown). Treatment protocols were approved by the National Cancer Institute (NCI) and participating institutions through their Institutional Review Boards. MLL rearrangements and MLL partner genes were characterized using cytogenetic and molecular methods. Informed consent for participation in these research studies was obtained from patients or authorized representatives. An independent validation cohort consisting of 57 cases of infant ALL with a variety of MLL translocations, that had available U133_Plus_2 CEL files and outcome data, was provided by Stam et al. The criteria for selection of validation cases were the same as for COG P9407 samples.

Gene Expression Profiling

RNA was isolated from pre-treatment leukemic cell suspensions obtained from peripheral blood or bone marrow using TRIzol (Invitrogen, Carlsbad, CA). All samples had >70% leukemic blasts, with a median of 90% blasts. cDNA labeling, hybridization to U133_Plus_2 arrays (Affymetrix, Santa Clara, CA) and scanning were performed as previously described.
The default RMA and MAS5 algorithms of Expression Console (v 1.1, Affymetrix, Santa Clara, CA) were used to generate and normalize signal intensities. Array experimental quality parameters for inclusion into the study included: scale factor <40, \textit{GAPDH} M33197 3’ intensity >15000 and \textit{GAPDH} M33197 3’/5’ ratio <3. Microarray gene expression data were available from an initial 54,504 probe sets after filtering to remove probe sets associated with sex-related genes, globins and controls, as previously described. In order to minimize the impact of set effects in comparative analyses with the validation cohort, the 97 P9407 CEL files were combined with 73 infant cases from the validation cohort study (including cases for which clinical data were not available) and 21 \textit{MLL}-R cases from a previously reported pediatric study prior to performing RMA. The final RMA-normalized dataset contained 191 cases from three separate studies. The P9407 gene expression data may be accessed via the National Cancer Institute caArray site (https://array.nci.nih.gov/caarray/project/EXP-520).

\textbf{Statistical Analyses}

Event-free survival (EFS) was calculated from the date of trial enrollment to either the date of first event (induction failure, relapse, second malignancy, or death) or last follow-up. Outcome analysis was performed only on cases where a death did not occur within 7 days of trial enrollment. This eliminated two cases from P9407 and four from the validation cohort. All cases were retained for studies correlating gene expression profiles with age as a continuous variable, as outcome associations were not required for these analyses.

Three gene filtering methods based on the coefficient of variation (CV), the standard deviation (SD), and cancer outlier profile analysis (COPA) were used to pre-select probe sets. The union of the probes selected by these three methods was used for the further analyses. The
Significance Analysis of Microarray (SAM)²⁰ was then used to identify the probe sets that were significantly associated with EFS and those that were significantly associated with age. Cox regression and logrank tests further determined the significance of a variable (clinical covariate, gene expression value, predicted risk score, or risk group by a model) in predicting EFS, with and without adjusting for the effects of other variables. This provided a means to assess the predictive power of various genes or probe sets relative to established prognostic factors such as age and \textit{MLL} status. Supervised principal component analysis (SPCA)²¹ and regression tree analysis²² were used to build outcome prediction models and to predict a score that positively correlated with event risk (higher scores correlated with increased risk of an event). Binary risk classification for event prediction was determined by a threshold \(\tau\) (defined as \(\alpha\%\) of the predicted score) with values below the threshold predicting low risk and above the threshold predicting high risk, where \(\alpha\%\) is the 2 year EFS probability of the original data. All analyses were performed using statistical software R (http://www.R-project.org, version 2.14.0, with basic, samr, superpc, rpart and survival packages) and Stata (version 11, StataCorp, College Station, TX). Pathway analyses and methods are described in the Supplement.

Hierarchical Clustering and Heat Maps

Gene expression data (54,504 probe sets) from the 48 cases with \textit{MLL-AFF1} rearrangements in the infant ALL cohort from COG P9407 were sorted based on SD and the top 100 probe sets were selected for clustering. Hierarchical clustering was performed in \textit{MATLAB} (MathWorks, Inc., Natick, MA) using Euclidean distance and average linkage. For the heat map of age-associated probe sets (treating patient age as a continuous variable), probe sets were ordered by their modified correlation coefficients as calculated with SAM. Expression intensity colors were displayed for each probe set relative to the average of the mean value of cases \(\leq 90\)
days old and the mean value of cases >90 days old. Cases included all 97 P9407 infants in addition to 21 ALL samples from a cohort of older pediatric ALL with molecularly confirmed *MLL-AFF1* translocations.\(^\text{17}\)

### Results

#### Clinical Features

The clinical and biologic features of the cohort of 97 infant ALL cases studied from COG Trial P9407\(^\text{15}\) are provided in Table 1. The overall outcome of the infants in this cohort was very poor, with a 5-year EFS of 41% ± 5 (Table 1). Three variables including patient age, white blood cell count (WBC) at disease presentation, and *MLL* status were significantly correlated with EFS (Table 1), while sex, race/ethnicity, and CNS status were not (data not shown). When WBC count was evaluated as a categorical variable (with a cutoff of 50 x10\(^3\)/µL), patients with high WBC fared worse than those with low WBC.

Infants with *MLL* rearrangements, and specifically *MLL-AFF1*, had significantly poorer outcomes than patients lacking MLL rearrangements (*MLL-G*) (Figures 1A and 1B; Table 1). Applying an age threshold of 90 days, the youngest 22 patients had significantly worse EFS than older infants (Figure 1C). Since patient age and *MLL* status were both observed to be significant factors in predicting EFS, we explored their combined impact. The Kaplan-Meier survival curve in Figure 1D displays the behavior of each age/*MLL* category. Not surprisingly, the 18 patients who were in the highest risk category for both variables (*MLL-R* and ≤90 days) had the poorest overall outcome (Figure 1D).

#### Gene Expression Profiling and Modeling Outcome Genes in the Full Cohort
To identify genes from the expression profiles that were significantly correlated with EFS and to model an outcome predictor in the full cohort of 97 infant ALL cases, we performed SAM analysis on the 430 potentially informative probe sets first determined by high CV, standard deviation, and COPA. The results of this analysis are provided in Table 2. Applying a false discovery rate (FDR) cutoff of 1%, only 5 probe sets (derived from 4 genes) were retained in the analysis: TACC2, IRX2, IL1R2 and FLT3 (Table 2). Each of these genes retained prognostic significance for outcome prediction beyond or independent of patient age, and, all but FLT3 were significant for outcome prediction after adjusting for MLL status (Table 2). For all four genes, elevated expression was correlated with a poorer outcome. Interestingly, in pathways analysis (see Supplement), 3 (TACC2, IL1R2, FLT3) of these 4 genes are linked together with MEIS1, the Rab effector MYRIP, and the EGFR pathway gene EPS8, in a primary pathway centering on the GRB2 adapter signaling protein. A regression tree analysis of EFS was then developed, which resulted in a final model in which only 3 of the 4 genes were retained (FLT3, TACC2, and IRX2), as IL1R2 did not retain independent prognostic significance in the context of the other three. This model for EFS is shown in Figure 2A. At each decision branch the highest expressing cases fall into the high risk category. Low FLT3 expression was found in a group of infants with excellent outcome (n=11; 5-year EFS of 100%) termed the “low risk group,” while differential expression of IRX2 and TACC2 partitioned the remaining infants into two groups (termed “intermediate risk” or “high risk”) (Figure 2A and 2B). The net result of this predictive modeling defined and distinguished three outcome groups in infant ALL with significantly different EFS including: 11 low, 31 intermediate, and 53 high risk cases (Figure 2B). The intermediate versus high risk groups had 5 year EFS of 16% versus 64% (P<.001), respectively (29% versus 71% for OS, Supplement Figure S2A). Similar findings were seen among the 23
infants ≤90 days old: 15 were high risk (12 died; 80%), 4 were intermediate risk (3 died; 75%) and 3 were low risk (none died).

The majority of “low risk” cases (10/11) defined by our predictive model using FLT3 expression were MLL-G, with the remaining case an MLL-AFF1. In order to rule out the possibility that low FLT3 expression was simply identifying all of the MLL-G cases, we analyzed all 17 MLL-G cases in this cohort to see how they performed in our predictive model (Figure 2C). Although 10 of the MLL-G cases were assigned to our “low risk” group, the remaining 7 were divided among “intermediate” (1) and “high risk” (6) with highly significant different outcomes (P<.001). These analyses reveal that low FLT3 expression is not simply a reflection of MLL-G status, and conversely, that some patients with MLL-G have high FLT3 expression and a poorer outcome, as predicted by our refined model (P=.009 for OS, Supplement Figure S2B).

As a final step in the assessment of the predictive model, we analyzed the performance of this predictor on 22 MLL-AFF1 cases from an independent validation cohort of infants with ALL accrued to the Interfant-99 Infant ALL Study (Figure 2D, Supplement Figure S2C). These cases were chosen because they represented the largest subset in that cohort and shared similar overall survivals to the infants in COG P9407. No low FLT3-expressing cases were present among these patients, so only intermediate and high risk cases were examined. The model partitioned the 22 cases into 5 intermediate and 17 high risk infants. Although the outcome differences are large (HR= 4.31), the small sample size resulted in P-values for EFS and OS that did not reach levels of statistical significance (P=.064 and P=.119, respectively). However, when relapse-free survival (RFS) was modeled instead of EFS, the single event in the intermediate risk group disappeared and the performance reached statistical significance (P=0.018).

Gene Expression Profiling and Modeling Outcome Genes in MLL-AFF1 Cases
MLL-AFF1 (MLL-AF4) cases represent the single largest subset of patients within the COG P9407 infant ALL cohort, comprising 48 of the 97 (49%) cases. Despite sharing an initiating translocation (MLL-AFF1), unsupervised hierarchical clustering (using the top 100 genes ranked according to standard deviation) of the 48 MLL-AFF1 cases showed a surprising degree of heterogeneity with two predominant cluster groups (Figure 3). Similar to what has been reported recently by others,\(^8,23\) we observed a clear separation of the MLL-AFF1 cases into two distinctive clusters dominated by the expression of homeobox genes: HOXA3, HOXA5, HOXA9, HOXA10 (referred to as the HOXA pattern), or, by IRX1 and IRX2 (referred to as the IRX pattern). Several other genes were differentially expressed within these two distinct clusters, but the expression of either elevated HOXA or IRX genes was the dominant feature (Figure 3). Many of the genes that defined these distinct unsupervised clusters had associations with EFS and it was interesting to note that many of these cluster-associated probe sets (including IRX1, IRX2, NEGR1, TPD52 and VAT1L) were identified as genes predictive of outcome in supervised learning, as we built the model for outcome prediction for the MLL-AFF1 cases detailed below.

We next used supervised analysis to build a model for outcome prediction (EFS) in the infants with MLL-AFF1. As with the entire cohort, a SAM analysis was performed on the 214 potentially informative genes initially identified by CV, standard deviation and COPA. Applying a FDR cutoff of 15%, 29 probe sets (representing 22 genes) were retained. Unlike the model built on the entire cohort, elevated expression of two of these probe sets correlated with a lower risk or a good outcome (Table 3A) and expression of 27 probe sets correlated with a higher risk or poorer outcome (Table 3B). Among these 22 genes, five (EPS8, IL1R2, IRX2, MYRIP and TACC2) were previously identified when we built the predictive model for outcome in the entire infant ALL cohort.
Using these 29 probe sets, we built an SPCA model which significantly separated 47 MLL-AFF1 cases into a low and a high risk group; this model was also predictive of outcome when tested against the independent validation cohort of 22 MLL-AFF1 cases from the Interfant-99 study (see details in the Supplement). Due to the known impact of patient age on outcome in infant ALL, we rebuilt the SPCA model with only those probe sets that retained significance after adjusting for age (the 9 probe sets derived from seven genes as noted in bold font, Table 3): \textit{NEGR1, VAT1L, IRX2, EPS8, TPD52, HLA-DQB1, and HLA-DRB1}. This seven-gene predictor separated the 47 MLL-AFF1 cases from COG P9407 into 17 “low risk” (5-year EFS of 65% ± 12) and 30 “high risk” cases (5-year EFS of 20% ±7) (Figure 4A); HR= 4.35, logrank \( P < .001 \). The results for OS were comparable (\( P < .001 \); HR=5.61; Supplement Figure S4A). Figure 4B illustrates the performance of the model on the 22 MLL-AFF1 cases in the independent validation cohort, showing a clear separation of cases by EFS (HR = 4.17; \( P = .069 \)) and OS (HR=3.16; \( P = .124 \)). The small number of cases in the independent validation cohort and their limited risk distribution decreased the statistical power.

We further examined the power of this predictive model relative to patient age in the MLL-AFF1 cases (using a 90-day cutoff). As shown in Figure 4C, the seven gene classifier effectively stratified the 36 infants >90 days of age into low versus high risk for EFS (\( P < .001 \), HR = 7.12), and for OS (\( P < .001 \), HR=11.2; Supplement Figure S4C). All 11 infants \( \leq 90 \) days of age died, including 8 identified by the model as high risk and 3 as low risk.

While similar genes were identified through unsupervised clustering of MLL-AFF1 cases (Figure 3), the seven-gene predictive model built through supervised methods was more precise for outcome prediction and effectively stratified risk in both the IRX-clustered and HOXA-clustered cases (Figure 3; Figure 4D; Supplement Figure S4D). The 17 “low risk” cases include
15 HOXA cluster and 2 IRX cluster cases, with 5 and 1 event, respectively. Within the 30 “high risk” cases were 7 HOXA and 23 IRX cases, with 7 and 18 events, respectively. Even though the HOXA cluster phenotype (Figure 3) was generally associated with better outcome, the supervised EFS model distinguished 7 “high risk” cases, all of whom had events (contrasted with the 33% rate among the 15 HOXA cluster-type “low risk” cases).

**Gene Expression Profiles Associated with Patient Age**

Given the historical significance of patient age in infant leukemia and the striking association of outcome with patient age in our infant cohort, we sought to determine whether there was a significant correlation of gene expression and patient age in our cohort of *MLL-AFF1* infant ALL cases. By treating leukemia cell gene expression as a continuous variable related to patient age in these infants, we found a dramatic change in gene expression patterns around 90 days of age, with a striking difference in gene expression patterns in infants ≤90 days versus those greater than 90 days in age. Figure 5 shows the top 43 probe sets (24 with increased expression associated with lower age and 19 with increased expression associated with higher age). Additionally, when we included 21 older children with ALL containing *MLL-AFF1* translocations from our COG P9906 studies in this analysis, they clustered with the older infants from COG P9407. This showed that infants >90 days of age have a gene expression signature more similar to older children than to infants ≤90 days of age.

Several genes associated with antibody production and B-cell maturation (e.g., *IGJ*, *IGH*, *MS4A1*) were expressed at higher levels in the leukemic cells from infants >90 days old, raising the possibility that *MLL-AFF1* leukemias in older children might be arising in a more committed B cell progenitor than in younger infants. These older infants also had elevated expression of genes linked with NK cells and antigen stimulation (*KLRB1*, *GNLY*). In contrast, the younger
patients had elevated expression of interleukin-related genes (IL1R2, IL8, IRAK3), heat-shock proteins (HSPA1A) and HLA genes (HLA-DRB4), that are linked through IL13 and IL1A to define a pattern associated with inflammatory responses, cellular movement, and immune cell trafficking (See Supplement, Fig. S6B).

Discussion

Our study of 97 infants with ALL (80/97 with MLL-R), accrued to COG Infant ALL Trial P9407, represents the largest cohort of infants reported to date to undergo gene expression profiling. The 5-year EFS among these infants was very poor (41%), with superior survivals seen among infants with MLL-G, age >90 days, and low WBC counts at disease presentation. Expression profiling initially identified a number of genes that were significantly associated with EFS in the infant cohort (EPS8, TACC2, FLT3, MEIS1, IL1R2), including genes known to play a role in MLL-mediated leukemogenesis (MEIS1), tumor progression (STAB1), and therapeutic resistance in T-cell malignancies (KCNK12).24,25 Pathways analyses further demonstrated complex interaction patterns among these genes, all converging on the adaptor protein GRB2 which plays a critical role in tyrosine kinase and Ras cell signaling pathways (See detailed analysis in the Supplement; Fig. S5).

Our final model predictive of outcome in the entire infant ALL cohort included three key genes (FLT3, TACC2 and IRX2) that refined risk classification and outcome prediction beyond age and MLL status. Differential FLT3 expression was uniquely powerful, as low FLT3 expression identified a group of infants with a very favorable outcome. This group of infants, with a superior 5 year EFS of 100% (including 9 MLL-G and a one MLL-AFF1 case) had not been previously identified using any other classification algorithm. Conversely, the remaining
infants who were \textit{MLL-G} (7/17) but had high \textit{FLT3} expression had an outcome similar to those infants with \textit{MLL-R} (5-year EFS of 29\% ± 17 versus 34\% ± 5, respectively). High \textit{FLT3} expression is common in infants and children with \textit{MLL-R} or high hyperdiploidy (>50 chromosomes) ALL, even in the absence of activating \textit{FLT3} mutations.\textsuperscript{26} High expression of wild type \textit{FLT3} may act similarly to \textit{FLT3} mutants or internal tandem duplications by activating Ras/Raf/MAPK signaling pathways.\textsuperscript{27} As leukemogenesis is a multistep process, high \textit{FLT3} activity may be a second “hit”, in concert with \textit{MLL}, to induce leukemogenesis in \textit{MLL-R} infants.

In the development of the outcome predictor for the full cohort, our regression tree analysis further identified \textit{TACC2} and \textit{IRX2} as two important genes that significantly impacted infants with high \textit{FLT3} expression. Infants with high expression of either \textit{TACC2} or \textit{IRX2} had a significantly worse 5 year EFS of 15\% as compared to 64\% in infants with low expression of both genes. How the expression of these genes might be altered by underlying \textit{MLL} rearrangements or other epigenetic changes in infant ALL is of interest. \textit{MLL} functions as an epigenetic regulator during normal development by incorporating into a COMPAS-like complex that facilitates H3K4 methylation which is required for the transcriptional activation of developmental regulatory genes, including homeobox genes and target genes of the WNT signaling pathway.\textsuperscript{7} When \textit{MLL} is translocated, domains involved in methylation such as SET are lost and MLL fusion proteins recruit alternative histone methyltransferases, resulting in inappropriate histone modification and dysregulation of gene expression.\textsuperscript{3} In addition to the potential activation of \textit{FLT3}, this mechanism may underlie the activation of \textit{TACC2}, a component of the chromatin modification machinery whose expression is regulated through binding of SMYD2 to the \textit{TACC2} promoter.\textsuperscript{28} The SET domain on SMYD2 acts as a
methyltransferase and specifically methylates H3K4, similar to MLL. Methylation of H3K4 is generally associated with activation of gene expression, and in addition to TACC2 activation, SMYD2 upregulates proteins in the SWI/SNF chromatin remodeling complex and methylates p53. This complex interaction between TACC2 and the SWI/SNF complex suggests that perturbed chromatin remodeling may promote malignant transformation. The SWI/SNF complex has also been implicated in other cancers through a variety of mechanisms.

The iroquois (IRO/IRX) genes encode transcriptional regulators that belong to the TALE (tree amino acid loop extension) class of homeobox genes. IRX2 controls developmental processes via the WNT pathway, while WNT signaling itself regulates homeobox genes and other targets of hematopoietic differentiation. IRX2 is located in a chromosomal region, 5p15.33, linked to a number of cancer types. A related family member, IRX1, also located at 5p15.33, appears to function as a tumor suppressor in solid tumors. An additional gene localized to 5p15.33 that may be affected by abnormal IRX2 expression is TERT, which maintains genomic integrity by regulating telomerase activity to allow restoration of telomere length. Maintaining TERT expression and preventing senescence may be a mechanism whereby fusion proteins, such as MLL-AFF1, mediated by HOXA genes, support self-renewal in leukemic cells.

Unsupervised hierarchical clustering of the MLL-AFF1 infants in our cohort yielded two cluster groups with distinct gene expression signatures dominated by homeobox genes (HOXA or IRX). The cluster with high HOXA expression is characteristic of MLL-R ALL, while low HOXA (and high IRX) expression is unique among infants with MLL-AFF1 and has been previously shown to be associated with a higher rate of relapse. When we used supervised learning to model an outcome predictor in these MLL-AFF1 cases, a 29 probe set (22 gene) expression
classifier better predicted EFS than the unsupervised hierarchical clustering. This predictive model was further reduced to 7 genes (NEGR1, VAT1L, IRX2, EPS8, TPD52, HLA-DQB1, HLA-DRB1) after removal of genes associated with patient age. Some of the highly expressed, and age-independent genes, that best predicted for poor EFS in MLL-AFF1 infants also predicted for poor EFS in the entire study cohort (IRX2, EPS8). EPS8 is an oncoprotein that participates in v-Src-induced cellular transformation with overexpression enhancing cell proliferation, migration, and tumorigenicity. Tumor protein D52 (TPD52) plays a role in Ca(2+)-dependent membrane trafficking in proliferating cancer cells and is also a B-cell differentiation marker overexpressed in subsets of pediatric acute leukemia. Genes associated with a poorer outcome in this predictor also included human leukocyte antigen class II genes (HLA-DQB1, multiple HLA-DRB) involved in modulating immune responses. In the MLL-AFF1 infant ALL cases, NEGR1 was the most significant gene predictor of EFS; high expression was associated with a superior EFS. NEGR1 is an opioid-binding cell adhesion molecule that maps to chromosome 1p31.1. Although NEGR1 is not known to be a tumor suppressor gene, studies of similar genes in ovarian cancer suggest a role for gene inactivation or silencing through DNA hypermethylation.

When the gene expression profiles derived from the infant ALL cases were treated as a continuous variable related to patient age (without knowledge of MLL status, outcome, or other clinical features), a striking change was observed around 90 days of age. This biologic observation is highly interesting, as infants ≤90 days versus those >90 days of age have historically been observed to have significantly different treatment outcomes. Our studies are the first to show a striking transition in leukemia cell expression signatures at this time point.
age-related patterns suggest different mechanisms of leukemogenesis and may underlie the differential outcomes historically seen in these age groups.

The expression profiles in infants >90 days of age are more similar to older children with ALL than to infants ≤90 days old.\textsuperscript{43} MLL-AFF1 leukemias in older infants and children may arise hypothetically from a more committed B cell progenitor than in younger infants, given the elevated expression of genes linked with NK cells and antigen stimulation (\textit{KLRB1, GNLY}) and their associations with cellular assembly and organizational pathways (See Supplement, Fig. S6A). In contrast, younger infants had elevated expression of interleukin-related genes (\textit{IL1R2, IL8, IRAK3}), heat-shock proteins (\textit{HSPA1A, HSPA1B}), and HLA genes (\textit{HLA-DRB4, HLA-DQ}), associated with the regulation of inflammatory responses. Whether this profile is reflective of the transformed cell of origin, a different marrow microenvironment, or different immune status in younger infants remains to be determined. These youngest infants may be more susceptible to the development of acute leukemia in utero, as specific HLA-DRB and HLA-DRQ haplotypes have been associated with different forms of leukemia and may be linked to polymorphisms in heat shock protein genes (\textit{HSPA1A, HSPA1B}) located in the HLA class III region.\textsuperscript{41,44} Polymorphisms of \textit{HSPA1A, HSPA1B}, and other inflammatory or immunomodulatory genes, such as \textit{IL-8} or \textit{VCAN}, have also been reported to influence susceptibility to cancer.\textsuperscript{45,46}

In conclusion, we have developed gene expression classifiers that improve outcome prediction in infant ALL, beyond the well-established prognostic factors of patient age and \textit{MLL} gene rearrangement status. Assessment of the expression levels of this relatively small number of genes that improve outcome prediction in the full cohort (\textit{FLT3, TACC2} and \textit{IRX2}) or in the subset of \textit{MLL-AFF1} cases (\textit{NEGR1, VAT1L, IRX2, EPS8, TPD52, HLA-DQB1, HLA-DRB1}) can easily be transitioned to a clinically useful diagnostic assay. Importantly, we can define those
infants who today have excellent responses to current therapies versus those who require the development of new approaches for cure, and, we can distinguish outcomes within the most predominant subset of infant ALL cases with MLL-AFF1 translocations. We have also found that higher expression of FLT3 identifies a group of MLL-G cases that has a poorer outcome than expected. Future inclusion of these infants in clinical trials testing FLT3 inhibitors, such as the COG Trial AALL0631 (NCT00557193) which randomizes infants with MLL-R to receive chemotherapy with or without the FLT3 inhibitor Lestaurtinib may improve their outcomes. The genes and pathways that we have identified not only serve as potential diagnostic targets, but also as potential therapeutic targets. As proposed by Stumpel et al, gene expression patterns that predict for poorer EFS in infant ALL, reflective of the aberrant epigenetic regulation in this disease, such as those defined in this study, suggest a role for demethylating agents in the treatment of infant ALL. Our assessment of the role of these genes in pathways analysis has also identified key nodes that might be considered for targeted therapy, such as GRB2, that are centrally regulated by the high expression of multiple genes associated with a poorer outcome in infant ALL.
Acknowledgements

This work was supported by NCI U01 SPECS: CA114762 (C.L.W.), Leukemia & Lymphoma Society SCOR 7372-07 (C.A.F.), LLS SCOR 7388-06 (C.L.W.), NCI U10 CA098543 (COG, Chair’s Award), U10 CA98413 supporting COG Statistical Center, U24 CA114766 supporting Human specimen banking in NCI supported cancer trials, NIH R01 CA080175 (C.A.F.) and NCI P30 CA118100 supporting UNM Cancer Center Shared Resources. S.P.H. is the Ergen Family Chair in Pediatric Cancer.

Authorship Contributions

H.K. performed statistical modeling and analyses, and prepared the manuscript; C.S.W. performed data analysis and review and prepared the manuscript; R.C.H. performed microarray studies, statistical modeling and analyses, and prepared the manuscript; I.-M.C. performed leukemia sample processing, microarray studies; M.H.M. conducted data management and data analyses; S.R.A. conducted pathway analyses, data analysis and review; E.J.B. performed statistical modeling and analyses; M.D. conducted COG clinical and statistical analyses, data review, and analysis; A.J.C. and N.A.H. performed cytogenetic analysis; B.W.R. conducted MLL partner gene classification; R.W.S. Designed validation data set; M.G.V. collected and processed patient information for the validation data set; R.P. designed validation data set; J.M.H., C.A.F., G.H.R., B.C., N.W., W.L.C. Z.E.D. and S.P.H. designed COG Studies, performed data analysis and reviewed the manuscript; and C.L.W. oversaw all aspects of this project, performed study design, data analysis and review, and prepared the manuscript.
Disclosure of Conflicts of Interest

Carolyn A. Felix owns the following patent: Methods and Kits for Analysis of Chromosomal Rearrangements Associated with Leukemia-US Patent #6,368, 791 (issued April 9, 2002); the panhandle PCR method that is subject of the patent and derivative methods were used to classify \( MLL \) rearrangements in this study.

References


### Tables

#### Table 1. Clinical Characteristics and Their Outcome Association (EFS)

<table>
<thead>
<tr>
<th></th>
<th>Entire Cohort</th>
<th></th>
<th></th>
<th>MLL-AFF1 Patients</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value¹</td>
<td>HR (SE)²</td>
<td>5-yr EFS (SE)³</td>
<td>P⁴</td>
<td>Value¹</td>
<td>HR (SE)²</td>
</tr>
<tr>
<td>Overall</td>
<td>97 (100%)</td>
<td>0.41 (.05)</td>
<td></td>
<td>48 (100%)</td>
<td>0.36 (.07)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.009</td>
<td></td>
<td></td>
<td>.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 90 days</td>
<td>24 (25%)</td>
<td>2.13 (.62)</td>
<td>.19 (.09)</td>
<td>12 (25%)</td>
<td>3.25 (1.30)</td>
<td>0 (NA)</td>
</tr>
<tr>
<td>&gt; 90 days</td>
<td>73 (75%)</td>
<td>1</td>
<td>.48 (.06)</td>
<td>36 (75%)</td>
<td>1</td>
<td>.47 (.08)</td>
</tr>
<tr>
<td>Sex</td>
<td>.745</td>
<td></td>
<td></td>
<td>.099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>54 (56%)</td>
<td>1.09 (.30)</td>
<td>.40 (.07)</td>
<td>26 (54%)</td>
<td>1.82 (.68)</td>
<td>.28 (.09)</td>
</tr>
<tr>
<td>Female</td>
<td>43 (44%)</td>
<td>1</td>
<td>.42 (.08)</td>
<td>22 (46%)</td>
<td>1</td>
<td>.45 (.11)</td>
</tr>
<tr>
<td>WBC</td>
<td>.006</td>
<td></td>
<td></td>
<td>.228</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 50x10⁹/µL</td>
<td>83 (86%)</td>
<td>5.70 (4.11)</td>
<td>.35 (.05)</td>
<td>43 (91%)</td>
<td>3.20 (3.26)</td>
<td>.33 (.07)</td>
</tr>
<tr>
<td>&lt;50x10⁹/µL</td>
<td>13 (14%)</td>
<td>1</td>
<td>.84 (.10)</td>
<td>4 (9%)</td>
<td>1</td>
<td>.75 (.22)</td>
</tr>
<tr>
<td>MLL Status</td>
<td>.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rearranged</td>
<td>80 (82%)</td>
<td>3.23 (1.52)</td>
<td>.35 (.05)</td>
<td>48 (100%)</td>
<td></td>
<td>.36 (.07)</td>
</tr>
<tr>
<td>Germline</td>
<td>17 (18%)</td>
<td>1</td>
<td>.69 (.12)</td>
<td>0 (0%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹Value includes number and percentage of infants in parentheses.
²HR is Hazard ratio relative to the reference category (with HR = 1). SE is the standard error for each hazard ratio.
³Five year EFS survival probability and standard error
⁴P-value of logrank test
NA = not applicable
Table 2. Gene Expression Probe Sets Associated with Outcome (EFS) in the Entire Infant ALL Cohort

<table>
<thead>
<tr>
<th>Rank</th>
<th>Probe set ID</th>
<th>Gene Symbol</th>
<th>Elevated in¹</th>
<th>SAM Analysis</th>
<th>Cox Regression Analysis</th>
<th>P (adjusted for MLL)³</th>
<th>P (adjusted for age)⁴</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Score</td>
<td>FDR (%)</td>
<td>HR</td>
<td>P²</td>
</tr>
<tr>
<td>1</td>
<td>211382_s_at</td>
<td>TACC2</td>
<td>High Risk</td>
<td>2.50</td>
<td>&lt; 1</td>
<td>1.51</td>
<td>0.001</td>
</tr>
<tr>
<td>2</td>
<td>228462_at</td>
<td>IRX2</td>
<td>High Risk</td>
<td>2.23</td>
<td>&lt; 1</td>
<td>1.45</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>202289_s_at</td>
<td>TACC2</td>
<td>High Risk</td>
<td>2.15</td>
<td>&lt; 1</td>
<td>1.42</td>
<td>0.003</td>
</tr>
<tr>
<td>4</td>
<td>205403_at</td>
<td>IL1R2</td>
<td>High Risk</td>
<td>1.99</td>
<td>&lt; 1</td>
<td>1.48</td>
<td>0.001</td>
</tr>
<tr>
<td>5</td>
<td>206674_at</td>
<td>FLT3</td>
<td>High Risk</td>
<td>1.94</td>
<td>&lt; 1</td>
<td>1.54</td>
<td>0.009</td>
</tr>
<tr>
<td>6</td>
<td>202609_at</td>
<td>EPS8</td>
<td>High Risk</td>
<td>1.78</td>
<td>9.77</td>
<td>1.38</td>
<td>0.004</td>
</tr>
<tr>
<td>7</td>
<td>220448_at</td>
<td>KCNK12</td>
<td>High Risk</td>
<td>1.77</td>
<td>9.77</td>
<td>1.39</td>
<td>0.019</td>
</tr>
<tr>
<td>8</td>
<td>204150_at</td>
<td>STAB1</td>
<td>High Risk</td>
<td>1.77</td>
<td>9.77</td>
<td>1.42</td>
<td>0.003</td>
</tr>
<tr>
<td>9</td>
<td>214156_at</td>
<td>MYRIP</td>
<td>High Risk</td>
<td>1.76</td>
<td>9.77</td>
<td>1.49</td>
<td>0.001</td>
</tr>
<tr>
<td>10</td>
<td>204069_at</td>
<td>MEIS1</td>
<td>High Risk</td>
<td>1.72</td>
<td>9.77</td>
<td>1.37</td>
<td>0.027</td>
</tr>
</tbody>
</table>

¹Indicates risk group with which elevated expression is correlated
²P-value of Wald test for the hazard ratio indicating significance of the association between the expression of each individual gene and treatment outcome.
³ and ⁴P-value indicating significance of the association between the expression of each gene and treatment outcome adjusting for the MLL status or age effect.

Bold denotes probe sets that are used for building the regression tree model.
### Table 3. Gene Expression Probe Sets Associated with Outcome (EFS) in Patients with MLL-AFF1

#### A

<table>
<thead>
<tr>
<th>Rank</th>
<th>Probe set ID</th>
<th>Gene Symbol</th>
<th>Elevated in[^1]</th>
<th>SAM Analysis</th>
<th>Cox Regression</th>
<th>P (adjusted for age)[^3]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Score</td>
<td>FDR (%)</td>
<td>HR</td>
</tr>
<tr>
<td>1</td>
<td>229461_x_at</td>
<td>NEGR1</td>
<td>Low Risk</td>
<td>-2.35</td>
<td>0.54</td>
<td>0.005 [^2]</td>
</tr>
<tr>
<td>2</td>
<td>226415_at</td>
<td>VAT1L</td>
<td>Low Risk</td>
<td>-2.19</td>
<td>0.59</td>
<td>0.007 [^2]</td>
</tr>
</tbody>
</table>

#### B

<table>
<thead>
<tr>
<th>Rank</th>
<th>Probe set ID</th>
<th>Gene Symbol</th>
<th>Elevated in[^1]</th>
<th>SAM Analysis</th>
<th>Cox Regression</th>
<th>P (adjusted for age)[^3]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Score</td>
<td>FDR (%)</td>
<td>HR</td>
</tr>
<tr>
<td>1</td>
<td>228462_at</td>
<td>IRX2</td>
<td>High Risk</td>
<td>2.08</td>
<td>1.53</td>
<td>0.010 [^2]</td>
</tr>
<tr>
<td>2</td>
<td>230472_at</td>
<td>IRX1</td>
<td>High Risk</td>
<td>1.99</td>
<td>1.51</td>
<td>0.023 [^2]</td>
</tr>
<tr>
<td>3</td>
<td>202609_at</td>
<td>EPNII</td>
<td>High Risk</td>
<td>1.99</td>
<td>1.52</td>
<td>0.006 [^2]</td>
</tr>
<tr>
<td>4</td>
<td>201688_s_at</td>
<td>TPDS2</td>
<td>High Risk</td>
<td>1.97</td>
<td>1.63</td>
<td>0.008 [^2]</td>
</tr>
<tr>
<td>5</td>
<td>214150_at</td>
<td>MYRIP</td>
<td>High Risk</td>
<td>1.92</td>
<td>1.64</td>
<td>0.005 [^2]</td>
</tr>
<tr>
<td>6</td>
<td>209480_at</td>
<td>HLA-DQB1</td>
<td>High Risk</td>
<td>1.89</td>
<td>1.49</td>
<td>0.023 [^2]</td>
</tr>
<tr>
<td>7</td>
<td>205403_at</td>
<td>IL1R2</td>
<td>High Risk</td>
<td>1.84</td>
<td>1.54</td>
<td>0.015 [^2]</td>
</tr>
<tr>
<td>8</td>
<td>212192_at</td>
<td>KCTD12</td>
<td>High Risk</td>
<td>1.83</td>
<td>1.56</td>
<td>0.021 [^2]</td>
</tr>
<tr>
<td>9</td>
<td>201743_at</td>
<td>CD14</td>
<td>High Risk</td>
<td>1.73</td>
<td>1.50</td>
<td>0.017 [^2]</td>
</tr>
<tr>
<td>10</td>
<td>1555745_a_at</td>
<td>LYZ</td>
<td>High Risk</td>
<td>1.73</td>
<td>1.53</td>
<td>0.031 [^2]</td>
</tr>
<tr>
<td>11</td>
<td>203555_at</td>
<td>SI100A9</td>
<td>High Risk</td>
<td>1.70</td>
<td>1.51</td>
<td>0.030 [^2]</td>
</tr>
<tr>
<td>12</td>
<td>204971_at</td>
<td>C6ST</td>
<td>High Risk</td>
<td>1.67</td>
<td>1.49</td>
<td>0.029 [^2]</td>
</tr>
<tr>
<td>13</td>
<td>204971_at</td>
<td>C6ST</td>
<td>High Risk</td>
<td>1.65</td>
<td>1.51</td>
<td>0.032 [^2]</td>
</tr>
<tr>
<td>14</td>
<td>204971_at</td>
<td>C6ST</td>
<td>High Risk</td>
<td>1.64</td>
<td>1.65</td>
<td>0.008 [^2]</td>
</tr>
<tr>
<td>15</td>
<td>211571_s_at</td>
<td>PLBD1</td>
<td>High Risk</td>
<td>1.61</td>
<td>1.55</td>
<td>0.015 [^2]</td>
</tr>
<tr>
<td>16</td>
<td>218192_at</td>
<td>TACCC2</td>
<td>High Risk</td>
<td>1.59</td>
<td>1.44</td>
<td>0.046 [^2]</td>
</tr>
<tr>
<td>17</td>
<td>228434_at</td>
<td>BTNL9</td>
<td>High Risk</td>
<td>1.56</td>
<td>1.58</td>
<td>0.021 [^2]</td>
</tr>
<tr>
<td>18</td>
<td>202917_s_at</td>
<td>SI100A8</td>
<td>High Risk</td>
<td>1.55</td>
<td>1.48</td>
<td>0.054 [^2]</td>
</tr>
<tr>
<td>19</td>
<td>204620_s_at</td>
<td>VCAN</td>
<td>High Risk</td>
<td>1.54</td>
<td>1.43</td>
<td>0.058 [^2]</td>
</tr>
<tr>
<td>20</td>
<td>205835_at</td>
<td>SI100A12</td>
<td>High Risk</td>
<td>1.52</td>
<td>1.43</td>
<td>0.043 [^2]</td>
</tr>
<tr>
<td>21</td>
<td>201690_s_at</td>
<td>TPDS2</td>
<td>High Risk</td>
<td>1.50</td>
<td>1.49</td>
<td>0.047 [^2]</td>
</tr>
<tr>
<td>22</td>
<td>204619_s_at</td>
<td>VCAN</td>
<td>High Risk</td>
<td>1.48</td>
<td>1.40</td>
<td>0.058 [^2]</td>
</tr>
<tr>
<td>23</td>
<td>215646_s_at</td>
<td>VCAN</td>
<td>High Risk</td>
<td>1.45</td>
<td>1.39</td>
<td>0.070 [^2]</td>
</tr>
<tr>
<td>24</td>
<td>221731_x_at</td>
<td>VCAN</td>
<td>High Risk</td>
<td>1.45</td>
<td>1.41</td>
<td>0.076 [^2]</td>
</tr>
<tr>
<td>25</td>
<td>204971_at</td>
<td>CSTA</td>
<td>High Risk</td>
<td>1.40</td>
<td>1.44</td>
<td>0.071 [^2]</td>
</tr>
</tbody>
</table>

[^1]: Indicates risk group with which elevated expression is correlated
[^2]: P-value of Wald test for the hazard ratio indicating significance of the association between the expression of each individual gene and treatment outcome.
[^3]: P-value indicating significance of the association between the expression of each gene and treatment outcome adjusting for age effect.

Bold denotes probe sets that are used for building the 7-gene (9-probe set) SPCA model.
Figure Legends

Figure 1. Impact of Age and MLL Status on EFS. Kaplan-Meier survival curves show the impact of infant age and MLL status on EFS. (A) Patients with MLL-R have significantly shorter EFS than those without rearrangements (MLL-G) [P=.008, logrank test; Hazard ratio (HR)=3.23]. (B) MLL-AFF1 cases have a nearly identical outcome pattern to overall MLL-R (P=.010, HR=3.26). (C) Younger infants (≤90 days old) compared to older infants (>90 days old) have significantly worse EFS (P=.009 logrank test; HR=2.13). (D) Infants with MLL-G have the best EFS while infants ≤90 days of age with MLL-R have the worst EFS. Note: two patients who died within 7 days of diagnosis were excluded from these analyses.

Figure 2. Performance of the 3-gene Regression Tree Model of EFS in All Cases. A 3-gene model was developed for prediction of EFS in the entire study cohort. (A) FLT3 expression separates infants into low versus intermediate/high risk disease. Infants with high FLT3 expression are further divided into intermediate and high risk disease categories based on IRX2 and TACC2 expression. The oval/boxes contain the relative risk followed by # events/ # cases. (B) Kaplan-Meier survival curves show significant differences in EFS among the infants in the low, intermediate and high risk categories. (C) The model significantly separates MLL-G infants into 2 groups with significantly different EFS. NA indicates HR is not applicable due to absence of failures in one group. (D) The validation cohort of 22 MLL-AFF1 cases is also separated into 2 groups with different EFS. No infants with MLL-AFF1 and low FLT3 expression (low risk) were present.

Figure 3. Hierarchical Clustering of MLL-AFF1 cases using Top 100 Standard Deviation Probe Sets. The top 100 probe sets, ranked by standard deviation, were used to cluster the 48 MLL-AFF1 cases. The yellow bar indicates the branch point defining the two major cluster
patterns. These patterns are named after the family of homeobox genes most commonly and highly expressed by its members (IRX or HOXA). Samples are shown in columns and probe sets are in rows. Captions across the right indicate the positions of some of the more conserved genes across a cluster. Increasing (red) or decreasing (green) gene expression is shown relative to the median (black) for each gene.

**Figure 4. Performance of the 7-Gene (9-Probe Set) SPCA Model of EFS in MLL-AFF1 Cases.** Kaplan-Meier survival curves showing prediction of EFS and RFS in our cohort and the independent validation cohort. (A) The 7-gene model separates the 47 MLL-AFF1 cases into low and high risk groups with significantly different EFS. (B) The validation cohort of 22 MLL-AFF1 cases are similarly separated into two groups with significantly different EFS based on our model. (C) The 7 gene model also significantly separates the 36 older (>90 days) MLL-AFF1 infants into low and high risk groups. (D) The model overlaps significantly with the P9407 MLL-AFF1 unsupervised clusters, however it also adds significant predictive risk information, particularly in the HOXA pattern cases.

**Figure 5. Heat Map of Probe Sets Associated with Patient Age.** The top 43 probe sets associated with age (as a continuous variable) at the significance level FDR = 15% were used to generate a heat map. Patients are ordered from left to right by ascending age. In addition to the 97 infant ALL patients, 21 pediatric MLL cases are included. Vertical white lines indicate the positions of age landmarks, while the horizontal line separates between the probe sets whose expressions are positively (upper part) and negatively (lower part) correlated with age. Age of patients is indicated across the top.
Figure 1. Impact of Age and MLL Status on EFS

A. MLL-G (n = 17) vs. MLL-R (n = 78)

B. MLL-G (n = 17) vs. MLL-AFF1 (n = 47)

C. Age ≤ 90 days (n = 22) vs. Age > 90 days (n = 73)

D. Age ≤ 90 days/MLL-G (n = 4) vs. Age ≤ 90 days/MLL-R (n = 18)
Figure 2. Performance of the 3-gene Regression Tree Model for Outcome (EFS) Prediction

A

Low Risk

1
56/95

High FLT3
(≥ 7.265)

Low IRX2
(< 3.865)

Low TACC2
(< 5.520)

0.094
0/11

0.831
25/47

0.505
11/31

High Risk

High IRX2
(≥ 3.865)

High TACC2
(≥ 5.520)

1.205
56/84

1.880
31/37

1.760
14/16

Intermediate Risk

B

Low Risk (n=11)

Intermediate Risk (n=31)

High Risk (n=53)

P<.001

0 2 4 6 8
YEARS

% EFS

C

Low Risk (n=10)

High/Intermediate Risk
(n=7)

P<.001

HR=NA

0 2 4 6 8
YEARS

% EFS

D

Intermediate Risk (n=5)

High Risk (n=17)

P=.064 (one-sided)

HR=4.31

0 1 2 3 4 5 6 7
YEARS

% EFS
Figure 3. Hierarchical clustering of *MLL-AFF1* Patients Using Top 100 Standard Deviation Probe Sets
Figure 4. Performance of the 7-Gene (9-Probe Set) SPCA Model for Outcome Prediction (EFS) in MLL-AFF1 Cases

A

B

C

D

For personal use only.on October 24, 2017. By guest
Figure 5. Heat Map of Gene Expression Probe Sets Associated with Patient Age
Gene expression profiles predictive of outcome and age in infant acute lymphoblastic leukemia: a Children’s Oncology Group study