Published in final edited form as:
Leukemia. 2008 November ; 22(11): 2041–2047. doi:10.1038/leu.2008.198.

HOX expression patterns identify a common signature for favorable AML

M Andreeff[1], **V Ruvolo**[1,6], **S Gadgil**[2,6], **C Zeng**[3,6], **K Coombes**[4], **W Chen**[1], **S Kornblau**[1], **AE Barón**[3], and **HA Drabkin**[5]

1*Departments of Stem Cell Transplantation and Cellular Therapy, Section of Molecular Hematology and Therapy, The University of Texas MD Anderson Cancer Center, Houston, TX, USA*

2*Department of Medicine, Division of Hematology-Oncology, University of Colorado Health Sciences Center, Aurora, CO, USA*

3*Department of Preventive Medicine and Biometrics, University of Colorado Health Sciences Center, Aurora, CO, USA*

4*Bioinformatics and Applied Mathematics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA*

5*Department of Medicine, Division of Hematology-Oncology, Medical University of South Carolina, Charleston, SC, USA*

Abstract

Deregulated HOX expression, by chromosomal translocations and myeloid-lymphoid leukemia (MLL) rearrangements, is causal in some types of leukemia. Using real-time reverse transcription-PCR, we examined the expression of 43 clustered *HOX*, polycomb, *MLL* and *FLT3* genes in 119 newly diagnosed adult acute myeloid leukemias (AMLs) selected from all major cytogenetic groups. Downregulated *HOX* expression was a consistent feature of favorable AMLs and, among these cases, inv(16) cases had a distinct expression profile. Using a 17-gene predictor in 44 additional samples, we observed a 94.7% specificity for classifying favorable vs intermediate/unfavorable cytogenetic groups. Among other AMLs, *HOX* overexpression was associated with nucleophosmin (*NPM*) mutations and we also identified a phenotypically similar subset with wt-*NPM*. In many unfavorable and other intermediate cytogenetic AMLs, *HOX* levels resembled those in normal CD34+ cells, except that the homogeneity characteristic of normal samples was not present. We also observed that *HOXA9* levels were significantly inversely correlated with survival and that *BMI-1* was over-expressed in cases with 11q23 rearrangements, suggesting that $p19^{ARF}$ suppression may be involved in MLL-associated leukemia. These results underscore the close relationship between *HOX* expression patterns and certain forms of AML and emphasize the need to determine whether these differences play a role in the disease process.

Keywords

acute myeloid leukemia; homeobox protein HOXA9; myeloid-lymphoid leukemia gene; BMI-1 protein

Correspondence: Dr HA Drabkin, Division of Hematology-Oncology, Medical University of South Carolina, 96 Jonathan Lucas St, CSB 903, Charleston, SC 20425, USA. E-mail: E-mail: drabkin@musc.edu.

[6]These authors contributed equally to this work.

Supplementary Information accompanies the paper on the Leukemia website (http://www.nature.com/leu)

Introduction

Our understanding of acute myeloid leukemia (AML) has come from the elucidation of genes at recurrent chromosomal rearrangements. Mutations in growth factor receptors, signaling intermediates, transcription factors and the nucleophosmin (*NPM*) gene[1-9] have added to this knowledge. Chromosomal features define the major prognostic groups (favorable, intermediate and unfavorable), although some abnormalities are variably classified.[10,11] Within these broad groups are subsets with shared biological/clinical features. For example, t(8;21), inv(16)/t(16;16) and t(15;17) AMLs all share an excellent chance for long-term response and cure. The fact that both t(8;21) and inv(16) alterations affect *AML1* (*Runx1*) and a PML1 isoform reportedly interacts with AML1b[12] raises the possibility that these subtypes might share a set of deregulated genes. Although high-density microarrays have successfully identified individual subsets,[13,14] common expression signatures linking the favorable AMLs have not been reported.

Homeodomain genes, including the clustered *HOXA-D* loci, encode transcription factors regulating pattern formation, differentiation and proliferation. There is considerable evidence linking deregulated *HOX* expression with human acute leukemia. This includes translocations targeting specific *HOX* genes, particularly those located at 5′ in their clusters (for example, *HOXA9*),[15] or myeloid-lymphoid leukemia (*MLL*), which normally functions to maintain proper *HOX* levels.[16] In mice, retroviral integrations affecting *Hoxa7* and *Hoxa9* cause AML,[17] and integrations involving *MEIS1* facilitate AML development via the effects on HOX protein localization and DNA binding.[18] Golub *et al*.[19] reported that *HOXA9* was the single most informative gene distinguishing AML from acute lymphoblastic leukemia. In addition, although no expression signature correlated with treatment response or outcome, *HOXA9* was overexpressed in poor-outcome patients. However, no statistical correlates were provided and further studies were suggested. Overexpression of other *HOX* genes, that is *HOXA5*,[20] *HOXB3*[21] or *HOXB4*,[22,23] affects myeloid proliferation and differentiation without causing AML, suggesting that they may contribute to the leukemogenic process or influence the phenotype.

We first reported that favorable AMLs express low levels of *HOX* genes, although the number of cases ($n = 34$) was limited.[24] Similar conclusions were reached by Debernardi *et al*.[25] in 28 cases, of which half consisted of favorable AMLs, with the remainder having 11q23 abnormalities or normal karyotypes. Thus, although suggestive of a common expression pattern, the AMLs examined have been limited and biased. We also initially reported that the highest *HOX* expression occurred in a subset of intermediate cytogenetic cases with elevated FLT3 levels and *FLT3* mutation rates.[26] This subset was subsequently shown to contain *NPM* mutations.[2,27] *HOX* expression in other AMLs, except for MLL (11q23) rearrangements, has not been rigorously examined.

In the present study, we expanded the repertoire of genes to include most clustered *HOXA-D* genes, plus selected trithorax, polycomb and other homeodomain family members. We utilized real-time reverse transcription-PCR (RT-PCR) to obtain accurate assessments of gene expression. All SYBR-based assays were verified and applied to an independent set of AMLs ($n = 119$) plus sorted CD34+ cells from four normal marrow donors. Downregulation of *HOX* expression was a striking feature of favorable AMLs. Among these, inv(16) cases had a distinct profile. Using a 17-gene predictor in 44 additional samples, we observed a 94.7% specificity for classifying favorable vs intermediate or unfavorable cytogenetic groups. As anticipated, marked *HOX* overexpression was observed in intermediate AMLs with *NPM* mutations, although we also identified a subset of phenotypically similar cases without *NPM* mutations. In many unfavorable AMLs, we detected Hox levels that resembled those in normal CD34+ cells, except that the striking normal homogeneity was lost in the leukemias. Finally,

we found that *HOXA9* expression correlated with survival and response to therapy and that *BMI-1*, a polycomb gene known to repress $p19^{ARF}$[25], was overexpressed with *MEIS1* in 11q23 (MLL) rearrangements.

Methods

Patient samples

Specimens of bone marrow (BM) aspirates ($n = 60$), peripheral blood ($n = 41$) or peripheral white blood cells obtained by leukophoresis ($n = 18$) were from newly diagnosed AML patients under informed consent (MD Anderson IRB-approved protocol). Karyotypes were obtained by conventional methods as part of the diagnostic work-up, typically on BM samples. Except as indicated, samples were obtained from patients prior to treatment and, before processing, contained a median percentage blast count of 77.0±s.e.m. 2.8 or, in acute promyelocytic leukemia (APL), 84.5% ±s.e.m. 16.9 promyelocytes. Mononuclear cells were isolated by centrifugation through Ficoll-Hypaque (Sigma Chemical Co., St Louis, MO, USA). Cells were stored in liquid nitrogen or lysed with RNA Stat 60 (Tel-Test, Friendswood, TX, USA), as directed by the manufacturer, and the RNA stored at -80 °C. Samples from healthy BM donors were selected for CD34+ cells using a Mini-Macs magnetic-antibody column (Miltenyi Biotech, Auburn, CA, USA) as directed by the manufacturer.

Gene expression assays and mutation analysis

Preparation of RNA and cDNA and assessment of gene expression by quantitative (q) real-time PCR were performed as described.[26] SYBR Green-based qRT-PCR assays were developed to run under a standard set of conditions, and each product was initially verified by DNA sequencing from multiple cDNA and genomic DNA templates. Assays were validated in three independent labs using ABI Model 7700 and 5700 instruments. For the ABI 7000 and newer instruments, conditions included 9600 emulation. In all the samples analyzed, PCR products were shown to represent a single product by dissociation curves. Genes chosen for the present study were shown to be variably expressed among 10 AML, biphenotypic and chronic myelogenous leukemia blast crisis leukemia cell lines (not shown). Raw data were normalized against glyceraldehyde-3-phosphate dehydrogenase (GAPDH), that is, $\Delta C_t = C_{t\text{specific gene}} - C_{t\ \text{GAPDH}}$. Similar results were obtained using 18S rRNA (data not shown). Expression data are reported as the number of transcripts per thousand copies of glyceraldehyde-3-phosphate dehydrogenase; high HOX expression is defined as levels >10 and low expression as levels equal to or <0.1. A comparison of HOX expression in paired BM and peripheral blood samples from patients with high (>80%) blast counts did not show differences (not shown). Mutation analysis for *NPM*, *FLT3* and *C/EBP*α used 65-55 °C touchdown PCR on cDNA, followed by DNA sequence analysis (sequencing performed by the DNA Sequencing Core of the University of Colorado Cancer Center). Primer sequences are provided in supplementary material.

Statistical methods

Bootstrap-validated unsupervised hierarchical clustering used the R statistical software package (http://www.R-project.org).[28] Clustering used a distance metric based on the Pearson correlation coefficient and average linkage. We performed bootstrap validation by resampling rows (genes) for sample clustering or columns (patient samples) for gene clustering from the data matrix with replacement, then repeating the clustering.[29] We recorded for each pair of genes or patient samples the percentage of times, in 500 bootstrap samples, they were included as part of the same cluster. To examine whether any genes were differentially expressed between groups of AML patients based on cytogenetics, we performed a separate analysis of variance for each gene. The Cox proportional hazards model was used to determine whether any *HOX* gene was potentially useful for predicting survival. The two-sample *t*-test was used

to examine *HOX* genes predictive of response analyzing each gene in patients who achieved a complete remission (CR) to those who failed or were resistant to therapy. To examine the association between *HOX* expression and mutations of *FLT3, NPM* and *C/EBPα*, we performed two-sample *t*-tests for each gene, comparing expression in patients with and without mutations. To correct for multiple testing, we applied the stepwise permutation-based procedure of Korn *et al*.,[30] which controls the actual number of false discoveries. This approach contrasts with methods, like the false discovery rate of Benjamini and Hochberg,[31] which control the expected number of false discoveries, and is especially suited to the case in which there are dependencies among genes, such as we have previously reported among the genes in the *HOX* family.[24,26] Analyses were performed using BRB ArrayTools developed by Dr Richard Simon and Amy Peng Lam. For analysis of variance, after determining the genes that were significantly different, we used Tukey's test for honestly significant difference to determine the groups that were different for each gene. For correlations with clinical parameters, we computed Pearson correlation coefficients among *HOX* genes, as well as between HOX genes and other clinical variables, for example, BM blast percentage. The qRT-PCR gene expression data were log 2-transformed in the statistical analysis described above. To develop a multivariate predictor of genes for determining the cytogenetic groups to which a given sample belongs, diagonal linear discriminant analysis, implemented in the Class Prediction Tool in BRB ArrayTools, was used.[32] A separate test data set was used to evaluate the accuracy of the gene classifier for classifying new samples.[33]

Results

Hierarchical clustering based on HOX expression

One hundred and nineteen newly diagnosed cases of AML were studied initially, as well as normal CD34+ cells from four marrow donors. Samples were arbitrarily selected from favorable, intermediate and poor prognostic cytogenetic cases without consideration of response or outcome. Only samples with intact RNA, as judged by gel electrophoresis and the relative staining intensities of 28S/18S rRNA bands, were used. The median age was 60.2 years, and the percentage of favorable, intermediate and unfavorable disease was 14, 44 and 38%, respectively. Four patients had no analyzable metaphases. All received a high-dose Ara-C-containing regimen, unless they had APL, in which case they received all-*trans* retinoic acid, arsenic trioxide and/or idarubicin. Overall, 58% of patients achieved a CR. Survivals ranged from 0 to 452 weeks with a median of 38 weeks. Other clinical parameters are provided in Supplementary Table 1. Specimens were obtained from BM aspirates ($n = 60$), peripheral blood ($n = 41$) and peripheral white blood cells obtained by leukophoresis ($n = 18$). We have not detected differences in HOX expression between samples when both corresponding BM and peripheral blood were compared (not shown).

Forty-three genes were analyzed by qRT-PCR. These included 39 homeodomain and two polycomb genes (*EZH2* and *BMI-1*) plus *MLL* and *FLT-3*. Normalized expression levels for each are provided in Supplementary Table 2. To determine if any genes were reproducibly associated with cytogenetic groups, we applied hierarchical clustering with validation by bootstrap sampling, using the expression of 43 genes to cluster 115 AML and four normal CD34+ samples. These results are shown in Figure 1, in which the solid yellow and blue colors represent, respectively, complete concordance and the lack of clustering. Dendrograms for hierarchical clustering are shown on both edges, and the color bars indicate the cytogenetic group of each sample (blue, favorable; green, intermediate; red, unfavorable). As can be seen, cases from each major cytogenetic group were predominantly clustered with others from the same group. In addition, the normal CD34+ samples indicated in black form a tight subcluster reflecting a high degree of homogeneous expression. This uniformity was absent in the leukemic samples.

Low levels of HOX expression are characteristic of favorable cytogenetic AMLs

A striking feature observed in each favorable AML was the low level of *HOXA* and *HOXB* gene expression (Supplementary Table 2). Minimal expression was detected for genes in the *HOXC* and -*D* clusters, excluding alternative *HOX* gene usage. Among favorable AMLs, cases with inv(16) or t(16;16) had statistically significant higher *HOX* expression, particularly *HOXB2*, *HOXB3*, *HOXB4* and *MEIS1*. Shown in Table 1 are levels of 12 genes that distinguish inv(16) and t(16;16) cases with parametric *P*-values <0.05. On the basis of a permutation distribution of the *t*-test statistics for each gene, with 95% confidence, there were no false positives among the first eight genes. We also compared *HOX* levels in five t(15;17) patients to five patients with t(8;21) AML. Levels of *MEIS2* ($P = 0.004$) and *BMI-1* ($P = 0.03$) were significantly higher in t(8;21) AML. Among 44 additional AML cases, six had favorable subtypes. Again, we observed that *HOX* expression was uniformly low and that inv(16) cases had a distinguishable pattern (Supplementary Table 3). In fact, favorable AMLs could be distinguished from other cases with high sensitivity and specificity (see below).

HOX expression in other AML subtypes

Low *HOX* levels in favorable AMLs were contrasted by marked overexpression in cases with intermediate cytogenetics and *NPM* mutations (Supplementary Table 2). In our series, there were 27 *NPM* mutations among 50 evaluable intermediate cases; 35 mutations were identified among all 115 cases with cytogenetic data. Thirty genes were differentially expressed in samples with *NPM* mutations (parametric *P*-values <0.05). Using the permutation distribution of the *t*-test statistics for each gene, we identified 21 with no false positives at 95% confidence (Supplementary Table 4). It is worth noting that we identified four cases (nos. 18, 17, 35 and 22) with wild-type *NPM*, which had a very similar phenotype to those with *NPM* mutations. This included *HOX* overexpression, diploid or pseudodiploid karyotype, and reduced numbers of CD34+ cells. Also, this wild-type *NPM* subset had significantly higher *FLT3* mRNA levels compared to either unfavorable AMLs (Kruskal-Wallis overall χ^2 test, $P = 0.012$, Bonferroni-corrected $P = 0.015$) or favorable AMLs (Bonferroni-corrected $P = 0.0138$). Many unfavorable and other intermediate cytogenetic cases had *HOX* levels resembling those in normal CD34+ samples. However, compared to normals (notable for their homogeneity of expression), the leukemia samples were much more variable (Supplementary Table 2).

To determine if specific homeodomain genes could be used to distinguish major cytogenetic groups, we performed a separate one-way analysis of variance for each gene. Twenty-four genes (Supplementary Table 5) had parametric *P*-values <0.05. On the basis of permutation distribution of the F-ratio for each, with 95% confidence, there were no false positives among the top 20 genes. Using Tukey's test for honestly significant difference, all 20 differentially expressed genes, except *EZH2*, were able to distinguish favorable vs intermediate cytogenetic AMLs. Although most *HOXA* genes could distinguish favorable vs unfavorable cases, only the *HOXB* genes were effective at distinguishing intermediate vs unfavorable AMLs.

With these results as a training set, we developed a 17-gene multivariate predictor using diagonal linear discriminant analysis from BRB ArrayTools. The predictor was comprised mostly of *HOXA* and -*B* genes, along with *PBX3* and *MEIS1*. The sensitivity and specificity of the classifier for the favorable cytogenetics group was 88.2 and 89.8%, whereas for intermediate and unfavorable cases, the sensitivity/specificity percentages were 67.9/82.3 and 66.7/81.4, respectively. We then validated the multivariate predictor on the 44 additional cases. Although it was only 66% accurate (95% CI: 50-80%) in classifying the independent samples into major cytogenetic groups, it was 83.3% sensitive and 94.7% specific for classifying favorable vs intermediate or unfavorable cytogenetic groups combined, with an overall accuracy of 93% (95% CI: 81-97%). This supports the hypothesis that favorable cases of AML have a distinct and common homeodomain gene expression signature.

HOXA9 as a prognostic marker and BMI-1 in MLL rearrangements

Golub *et al*.[19] reported that *HOXA9* correlated with outcome in AML, although further studies were suggested. We examined this question by performing separate analyses using a Cox proportional hazards model on patients for whom survival data were available (n = 119). For overall survival, the levels of eight *HOXA* genes (that is, *HOXA1-10*) and *MEIS1* had parametric *P*-values less than 0.05 (Supplementary Table 6), with low expression favoring survival. However, based on a permutation distribution of the χ^2 statistics for each gene, we only have 90% confidence that there are no more than six false positives among these. The Kaplan-Meier curve for *HOXA9* expression, grouped in tertiles and analyzed by logrank test (low vs intermediate plus high), is shown in Figure 2. Similar analyses were performed for event-free survival. Only one gene, *HOXA9*, had unadjusted parametric *P*-values <0.05 (P = 0.0256). For response to therapy, we compared 67 patients who achieved CR to 49 patients who were resistant (excluding those with early death). Twelve genes (Supplementary Table 7) had parametric *P*-values <0.05, with *HOXA9* ranking predominantly. On the basis of permutation distribution, with 95% confidence there are no more than six false positives among the first 11 genes. Levels of *EZH2* were higher in patients achieving CR, whereas expression levels for the other genes were lower.

AMLs with 11q23 (MLL) rearrangements overexpress *MEIS1*.[34] We noted that expression of *BMI-1*, a polycomb gene reported to repress $p19^{ARF}$, was significantly correlated with *MEIS1* (Pearson r = 0.37, P<0.001). Compared to other unfavorable AMLs, *BMI-1* levels were significantly elevated in cases with 11q23 rearrangements (P = 0.038). Similar results were observed in the additional 44 AMLs, suggesting that $p19^{ARF}$ suppression by BMI-1 may play a role in the development of AML with MLL rearrangements. Other genes associated with 11q23 rearrangements are shown in Supplementary Table 8.

Discussion

Selective *HOX* deregulation by chromosomal translocations or retroviral integrations leads to acute leukemia. Our findings suggest that differences in homeodomain gene expression are related to the major cytogenetic groups. Strikingly, very low levels of homeodomain genes are characteristic of favorable AML. The results from 119 initial cases, of which 17 were favorable, were confirmed in 44 subsequent AMLs with six favorables. Using a 17-gene predictor in the 44 test samples, we observed a 94.7% specificity in classifying favorable vs intermediate or unfavorable cytogenetic groups. Moreover, cases of inv(16) and t(16;16) were readily discernable on the basis of *HOX* expression (Table 1).

We first reported that low *HOX* levels correlated with favorable AML, although the number of cases was limited.[24,26] Debernardi *et al*.[25] described similar results, although their series was also small and biased toward favorable cases. The current work involving 159 AMLs with cytogenetic data confirms this association, thus linking each of the favorable subtypes, that is, t(8;21), inv(16) and t(15;17), to a common gene expression signature. It is interesting to note that an isoform of PML1 reportedly interacts with AML1b,[12] raising the possibility that AML1 function could be affected in all favorable subtypes. Thus, it will be important to determine if AML1, AML1-ETO or alterations of Cbfβ (the β-subunit with AML1) affects *HOX* expression.

The identification of a common expression pattern, particularly when members of the gene set are known to be causally involved in leukemia, raises the question whether their expression levels are involved in the disease process or merely reflect the cell of origin. In favorable AMLs, it is possible that failure to upregulate certain *HOX* genes might impair differentiation, causing cells to persist in the BM with susceptibility to acquired mutations. Such a scenario has been suggested for cancer stem cells,[35] and AML1-ETO reportedly affects DNA repair.[36] Furthermore, Kuo *et al*.[37] proposed that the inv[16] Cbfβ-SMMHC fusion impairs

differentiation/proliferation and noted that Cbfβ-SMMHC was more leukemogenic than Runx1-Eto. It is worth noting that the higher levels of *HOXA2*, *HOXB2-B4* and *MEIS1* in inv (16) cases could be responsible for this difference.

Could downregulation of *HOX* genes be related to C/EBPα? Kirstetter *et al*.[38] reported that eliminating C/EBPα p42 results in AML by committed leukemia-initiating cells and that *HoxA9* and *Meis1* were downregulated. In this regard, AML1-ETO inhibits *C/EBPα*,[39,40] raising the question whether homeo-domain profiles in t(8;21) or other favorable AMLs resemble those in leukemias with *C/EBPα* mutations. However, our analyses indicate that *HOX* levels in favorable AMLs are far lower than those with *C/EBPα* mutations (Supplementary Table 2), suggesting that *C/EBPα* is not responsible, at least as a single factor.

Although homeodomain expression patterns in favorable AML are the most notable in terms of linking an entire cytogenetic group to a common expression signature, other AML subsets show distinct patterns of *HOX* expression. We first reported that the highest *HOX* levels occurred in a subset of intermediate cytogenetic cases with increased *FLT3* expression and *FLT3* mutations.[26] Falini et al.[2] and Verhaak et al.[27] subsequently discovered that this phenotype was associated with *NPM* mutations. In agreement with previous reports, we found that 54% (that is, 27/50) of evaluable intermediate cytogenetic cases had *NPM* mutations. Interestingly, we found four additional cases with similar levels of *HOX* genes and wild-type *NPM*. These cases shared other phenotypic features associated with NPM mutations, that is, diploid/pseudodiploid karyotype, reduced number of CD34+ cells and elevated *FLT3* mRNA levels.

Although *HOX* levels tended to be higher in intermediate cytogenetic AMLs with both *NPM* and *FLT3* mutations vs *NPM* mutation alone, the differences were not significant. On the other hand, *HOX* expression was significantly higher in cases with mutation of both *NPM* and *FLT3* compared to *FLT3* mutations alone. In related studies, we have found that FLT3 inhibition by PKC412 in MV4;11 cells, FLT3 stimulation by FLT ligand in RS4;11 cells and BCR/ABL inhibition by imatinib in K562 cells consistently affect only a few non-clustered homeo-domain genes (Starkova *et al*., to be presented elsewhere). Thus, *FLT3* mutations are probably not causally related to the high levels of *HOX* expression associated with *NPM* mutations. Although it is unknown whether *NPM* mutations directly affect *HOX* expression, the identification of a phenotypically similar subset with wt-NPM argues, in this instance, that high *HOX* expression relates to the cell of origin. Alternatively, if *NPM* mutations do affect *HOX* expression, then alterations in other pathway components might explain the apparent phenocopy.

Our results are not dissimilar to those obtained from genome-wide cDNA or oligo-nucleotide arrays. For example, the array analysis of Bullinger *et al*.[13] was able to distinguish major cytogenetic AML groups, although it tended to split subtypes with no common expression pattern identified for favorable AMLs. Valk *et al*.[14] reported that *C/EBPα* mutations clustered into two predominant groups, and for 11q23 rearrangements, both Bullinger *et al*.[13] and Valk *et al*.[14] presented evidence for molecular heterogeneity, consistent with clinical observations.[41] On the basis of homeodomain expression, we found similar heterogeneity for both *C/EBPα* and 11q23 rearrangements (Supplementary Table 2).

Interestingly, in cases with 11q23 rearrangements, which affect MLL and are associated with overexpression of *HOXA9* and MEIS1,[34] we found significantly higher levels of *BMI-1*. *MLL* and *BMI*-1 are antagonistic in some developmental contexts[16] and in experimental leukemia induced by MLL-ENL.[42] BMI-1 is a core component of the Polycomb (PcG) PRC1 complex, which ubiquitylates Lys119 of histone H2A. In knockout mice, Bmi-1 loss suppresses 3′-*Hox* genes, while upregulating 5′ genes (for example, *HoxA9*).[43] Targeting PRC1 to DNA-

binding sites is controlled through methylation of H3-K27 by EZH2, the catalytic component of a second PcG complex, ESC-E(Z). Why *BMI*-1 overexpression would occur in some MLL rearrangements with *HOXA9* overexpression is unclear. One hypothesis could be that *HOX* expression becomes *BMI*-1-independent as a result of some MLL fusions or secondary genetic changes and the selection for *BMI*-1 upregulation comes from its effects on other genes. In this regard, *BMI*-1 is overexpressed along with *HOXA9*, *HOXA10* and *MEIS1* in T-cell acute lymphoblastic leukemias with CALM-AF10 fusions, which lack $p16^{INK4A}/p19^{ARF}$ deletions common in other types of T-cell acute lymphoblastic leukemia.[44] These results suggest that in the context of 11q23/MLL rearrangements, *BMI*-1 overexpression may suppress $p16^{INK4A}/p19^{ARF}$ and play a role in the disease process.

Finally, among all cases, we found that lower *HOXA9* expression was the best predictor of overall/disease-free survival and response to therapy, extending the preliminary findings of Golub *et al*.[19] This appears to reflect the low level of *HOXA9* in favorable AMLs together with its elevated expression in other subtypes, including the unfavorable group. Other genes (Supplementary Table 7), including *HOXA* family members, also correlated with outcome. This reflects the high degree of coordinate regulation observed among members of the same *HOX* group, for example the Pearson correlation coefficient between *HOXA4* and *HOXA9* is 0.92 ($P<0.0001$). However, in a multivariate analysis, *HOX* expression had no additional impact when age and cytogenetics were considered (not shown). Interestingly, *EZH2* levels were higher in patients achieving CR. In the absence of *MLL* rearrangements, this is consistent with targeting PRC1 to HOX loci to suppress expression. In summary *HOX* expression patterns are closely linked to some major cytogenetic subgroups, particularly favorable AMLs. Because *HOX* genes play a pathogenetic role in leukemia development when deregulated by certain chromosomal translocations, these characteristic expression patterns suggest that they may have a role in the disease process or phenotype. Interestingly, pair-rule genes such as Runt (Runx1/AML1) are involved in a hierarchy with segment polarity genes, including homeotic genes, during normal development (Choe *et al*.[45] and references therein).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Drs Hagop Kantarjian and Elihu Estey from the Department of Leukemia (UT-MDA). This work is supported by grants from the National Institutes of Health (CA97710, CA55164 and CA16672) and the Paul and Mary Haas Chair in Genetics (MA).

References

1. Boissel N, Renneville A, Biggio V, Philippe N, Thomas X, Cayuela JM, et al. Prevalence, clinical profile, and prognosis of NPM mutations in AML with normal karyotype. Blood 2005;106:3618–3620. [PubMed: 16046528]
2. Falini B, Mecucci C, Tiacci E, Alcalay M, Rosati R, Pasqualucci L, et al. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. N Engl J Med 2005;352:254–266. [PubMed: 15659725]
3. Gari M, Goodeve A, Wilson G, Winship P, Langabeer S, Linch D, et al. c-kit proto-oncogene exon 8 in-frame deletion plus insertion mutations in acute myeloid leukaemia. Br J Haematol 1999;105:894–900. [PubMed: 10554798]
4. Goemans BF, Zwaan CM, Miller M, Zimmermann M, Harlow A, Meshinchi S, et al. Mutations in KIT and RAS are frequent events in pediatric core-binding factor acute myeloid leukemia. Leukemia 2005;19:1536–1542. [PubMed: 16015387]

5. Kiyoi H, Naoe T, Nakano Y, Yokota S, Minami S, Miyawaki S, et al. Prognostic implication of FLT3 and N-RAS gene mutations in acute myeloid leukemia. Blood 1999;93:3074–3080. [PubMed: 10216104]
6. Kottaridis PD, Gale RE, Frew ME, Harrison G, Langabeer SE, Belton AA, et al. The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials. Blood 2001;98:1752–1759. [PubMed: 11535508]
7. Preudhomme C, Sagot C, Boissel N, Cayuela JM, Tigaud I, de Botton S, et al. Favorable prognostic significance of CEBPA mutations in patients with de *novo* acute myeloid leukemia: a study from the Acute Leukemia French Association (ALFA). Blood 2002;100:2717–2723. [PubMed: 12351377]
8. Schnittger S, Schoch C, Dugas M, Kern W, Staib P, Wuchter C, et al. Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. Blood 2002;100:59–66. [PubMed: 12070009]
9. Yamamoto Y, Kiyoi H, Nakano Y, Suzuki R, Kodera Y, Miyawaki S, et al. Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. Blood 2001;97:2434–2439. [PubMed: 11290608]
10. Grimwade D, Walker H, Oliver F, Wheatley K, Harrison C, Harrison G, et al. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. Blood 1998;92:2322–2333. [PubMed: 9746770]
11. Mrozek K, Heerema NA, Bloomfield CD. Cytogenetics in acute leukemia. Blood Rev 2004;18:115–136. [PubMed: 15010150]
12. Nguyen LA, Pandolfi PP, Aikawa Y, Tagata Y, Ohki M, Kitabayashi I. Physical and functional link of the leukemia-associated factors AML1 and PML. Blood 2005;105:292–300. [PubMed: 15331439]
13. Bullinger L, Dohner K, Bair E, Frohling S, Schlenk RF, Tibshirani R, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med 2004;350:1605–1616. [PubMed: 15084693]
14. Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med 2004;350:1617–1628. [PubMed: 15084694]
15. Borrow J, Shearman AM, Stanton VP Jr, Becher R, Collins T, Williams AJ, et al. The t(7;11) (p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP98 and class I homeoprotein HOXA9. Nat Genet 1996;12:159–167. [PubMed: 8563754]
16. Hanson RD, Hess JL, Yu BD, Ernst P, van Lohuizen M, Berns A, et al. Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. Proc Natl Acad Sci USA 1999;96:14372–14377. [PubMed: 10588712]
17. Nakamura T, Largaespada DA, Shaughnessy JD Jr, Jenkins NA, Copeland NG. Cooperative activation of Hoxa and Pbx1-related genes in murine myeloid leukaemias. Nat Genet 1996;12:149–153. [PubMed: 8563752]
18. Shen WF, Rozenfeld S, Kwong A, Kom ves LG, Lawrence HJ, Largman C. HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. Mol Cell Biol 1999;19:3051–3061. [PubMed: 10082572]
19. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531–537. [PubMed: 10521349]
20. Crooks GM, Fuller J, Petersen D, Izadi P, Malik P, Pattengale PK, et al. Constitutive HOXA5 expression inhibits erythropoiesis and increases myelopoiesis from human hematopoietic progenitors. Blood 1999;94:519–528. [PubMed: 10397719]
21. Sauvageau G, Thorsteinsdottir U, Hough MR, Hugo P, Lawrence HJ, Largman C, et al. Overexpression of HOXB3 in hematopoietic cells causes defective lymphoid development and progressive myeloproliferation. Immunity 1997;6:13–22. [PubMed: 9052833]

22. Antonchuk J, Sauvageau G, Humphries RK. HOXB4 overexpression mediates very rapid stem cell regeneration and competitive hematopoietic repopulation. Exp Hematol 2001;29:1125–1134. [PubMed: 11532354]

23. Kyba M, Perlingeiro RC, Daley GQ. HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. Cell 2002;109:29–37. [PubMed: 11955444]

24. Drabkin HA, Parsy C, Ferguson K, Guilhot F, Lacotte L, Roy L, et al. Quantitative HOX expression in chromosomally defined subsets of acute myelogenous leukemia. Leukemia 2002;16:186–195. [PubMed: 11840284]

25. Debernardi S, Lillington DM, Chaplin T, Tomlinson S, Amess J, Rohatiner A, et al. Genome-wide analysis of acute myeloid leukemia with normal karyotype reveals a unique pattern of homeobox gene expression distinct from those with translocation-mediated fusion events. Genes Chromosomes Cancer 2003;37:149–158. [PubMed: 12696063]

26. Roche J, Zeng C, Baron A, Gadgil S, Gemmill RM, Tigaud I, et al. Hox expression in AML identifies a distinct subset of patients with intermediate cytogenetics. Leukemia 2004;18:1059–1063. [PubMed: 15085154]

27. Verhaak RG, Goudswaard CS, van Putten W, Bijl MA, Sanders MA, Hugens W, et al. Mutations in nucleophosmin (NPM1) in acute myeloid leukemia (AML): association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. Blood 2005;106:3747–3754. [PubMed: 16109776]

28. Team RDC. R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org. ISBN 3-900051-07-0

29. Kerr MK, Churchill GA. Bootstrapping cluster analysis: assessing the reliability of conclusions from microarray experiments. Proc Natl Acad Sci USA 2001;98:8961–8965. [PubMed: 11470909]

30. Korn EL TJ, McShane LM, Simon R. Controlling the number of false discoveries: Application to high-dimensional genomic data. J Stat Plan Inference 2004;124:379–398.

31. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 1995;57:289–300.

32. Dudoit S, Fridlyand J, Speed TP. Comparison of discrimination methods for the classification of tumors using gene expression data. J Am Stat Assoc 2002;97:77–87.

33. Dobbin KK, Simon RM. Sample size planning for developing classifiers using high-dimensional DNA microarray data. Biostatistics 2007;8:101–117. [PubMed: 16613833]

34. Rozovskaia T, Feinstein E, Mor O, Foa R, Blechman J, Nakamura T, et al. Upregulation of Meis1 and HoxA9 in acute lymphocytic leukemias with the t(4: 11) abnormality. Oncogene 2001;20:874–878. [PubMed: 11314021]

35. Rossi DJ, Jamieson CH, Weissman IL. Stems cells and the pathways to aging and cancer. Cell 2008;132:681–696. [PubMed: 18295583]

36. Xu M, Li D, Lu Y, Chen GQ. Leukemogenic AML1-ETO fusion protein increases carcinogen-DNA adduct formation with upregulated expression of cytochrome P450-1A1 gene. Exp hematol 2007;35:1249–1255. [PubMed: 17560011]

37. Kuo YH, Landrette SF, Heilman SA, Perrat PN, Garrett L, Liu PP, et al. Cbf beta-SMMHC induces distinct abnormal myeloid progenitors able to develop acute myeloid leukemia. Cancer Cell 2006;9:57–68. [PubMed: 16413472]

38. Kirstetter P, Schuster MB, Bereshchenko O, Moore S, Dvinge H, Kurz E, et al. Modeling of C/EBPalpha mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells. Cancer Cell 2008;13:299–310. [PubMed: 18394553]

39. Westendorf JJ, Yamamoto CM, Lenny N, Downing JR, Selsted ME, Hiebert SW. The t(8;21) fusion product, AML-1-ETO, associates with C/EBP-alpha, inhibits C/EBP-alpha-dependent transcription, and blocks granulocytic differentiation. Mol Cell Biol 1998;18:322–333. [PubMed: 9418879]

40. Pabst T, Mueller BU, Harakawa N, Schoch C, Haferlach T, Behre G, et al. AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia. Nat Med 2001;7:444–451. [PubMed: 11283671]

41. Schoch C, Schnittger S, Klaus M, Kern W, Hiddemann W, Haferlach T. AML with 11q23/MLL abnormalities as defined by the WHO classification: incidence, partner chromosomes, FAB subtype,

age distribution, and prognostic impact in an unselected series of 1897 cytogenetically analyzed AML cases. Blood 2003;102:2395–2402. [PubMed: 12805060]

42. Zeisig BB, Milne T, Garcia-Cuellar MP, Schreiner S, Martin ME, Fuchs U, et al. Hoxa9 and Meis1 are key targets for MLLENL-mediated cellular immortalization. Mol Cell Biol 2004;24:617–628. [PubMed: 14701735]
43. Cao R, Tsukada Y, Zhang Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol cell 2005;20:845–854. [PubMed: 16359901]
44. Dik WA, Brahim W, Braun C, Asnafi V, Dastugue N, Bernard OA, et al. CALM-AF10+ T-ALL expression profiles are characterized by overexpression of HOXA and BMI1 oncogenes. Leukemia 2005;19:1948–1957. [PubMed: 16107895]
45. Choe CP, Miller SC, Brown SJ. A pair-rule gene circuit defines segments sequentially in the short-germ insect Tribolium castaneum. Proc Natl Acad Sci USA 2006;103:6560–6564. [PubMed: 16611732]

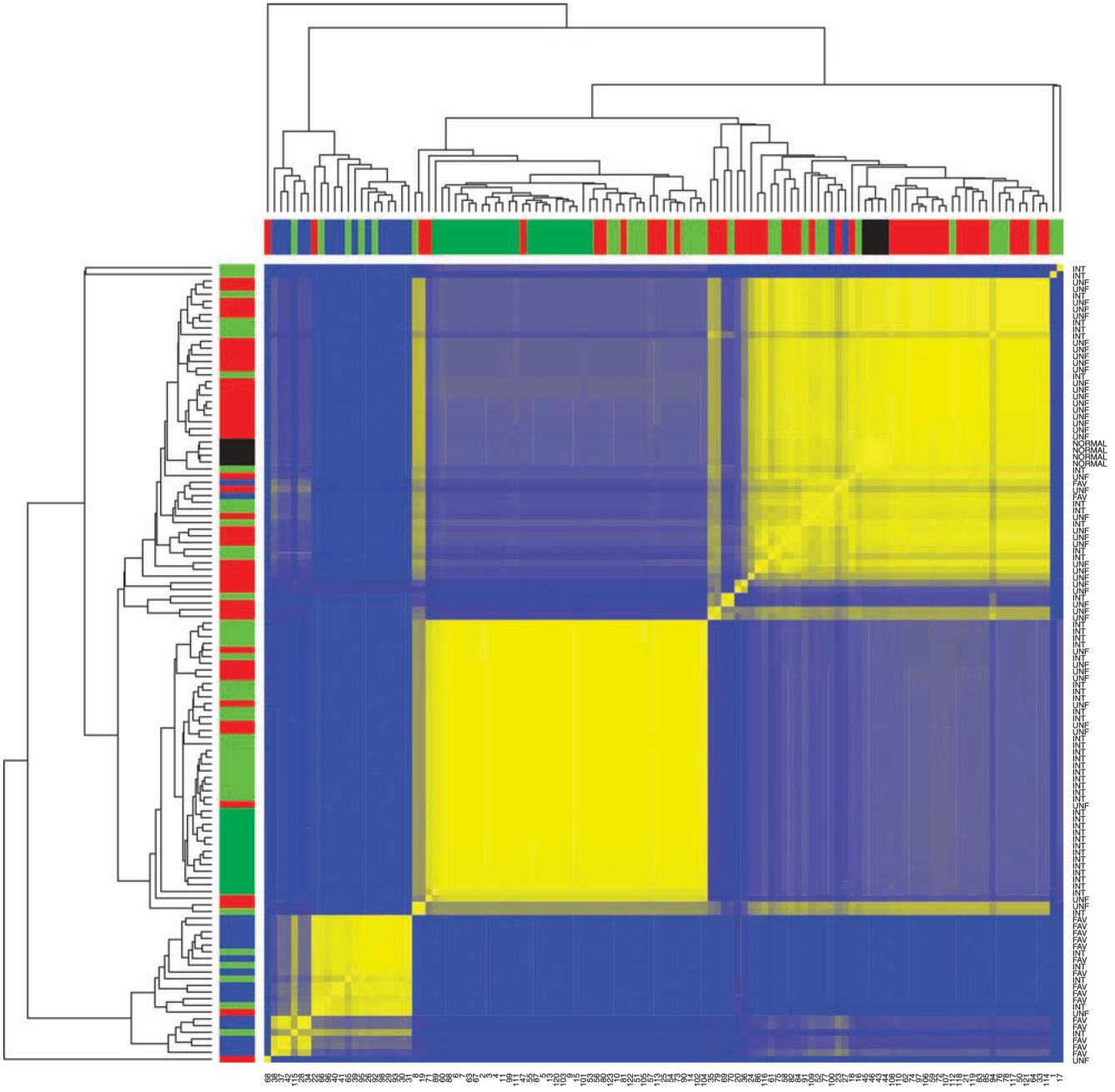

Figure 1.

Bootstrap-validated hierarchical clustering of samples. We applied bootstrap-validated unsupervised hierarchical clustering, using the expression of 43 genes to cluster 119 samples (excluding four NAM cytogenetics). Clustering used a distance metric based on the Pearson correlation coefficient and average linkage. We performed bootstrap validation by resampling genes from the data matrix with replacement and then repeating the clustering. We recorded for each pair of samples the percentage of times, in 500 bootstrap samples, that they were included as part of the same cluster (cutting the dendrogram each time to produce seven clusters). The percentage of times each pair of samples is clustered together is illustrated in the main part of the image, ranging from 0% (pure blue) to 100% (pure yellow). Dendrograms for hierarchical clustering by Pearson correlation with average linkage are shown on both edges.

Color bars indicate sample type, which includes normal (NORMAL) CD34+ controls (black), favorable (FAV) cytogenetics (blue), intermediate (INTERM) cytogenetics (green) and unfavorable (UNFAV) cytogenetics (red). Individual patient numbers are indicated along the bottom.

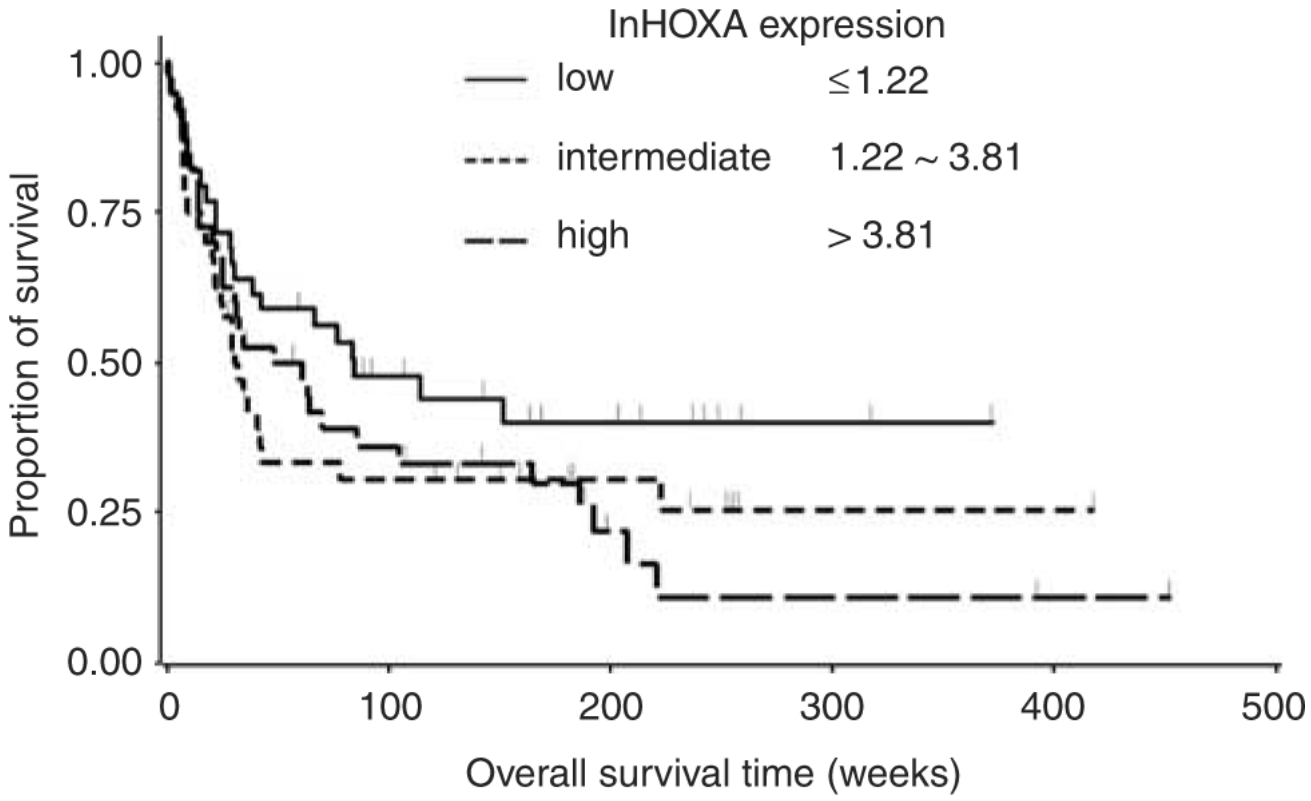

Figure 2.
Kaplan-Meier survival curve for HOXA9 gene expression grouped by its tertiles. Patients in the lowest tertile of expression have better survival than patients with moderate or high expression of HOXA9 ($P = 0.073$).

Table 1

Genes that discriminate inv(16) and t(16;16) cases from other favorable subtypes

Gene	*Parametric* P-*value*	Geometric mean expression	
		FAV t(15;17) or t(8;21) N = 10	*FAV inv(16) or t(16;16)* N = 7
HOXB4	<1e-07	0.0400	3.8773
MEIS1	<1e-07	0.0659	26.4084
HOXB2	<1e-07	0.2862	17.4490
HOXB3	<1e-07	0.0427	1.8071
HOXA2	4.80E-06	0.0174	0.2941
HOXA3	8.60E-06	0.0096	0.0943
HOXB6	1.31E-05	0.0057	0.1318
HOXA1	0.00039	0.0133	0.2323
HOXA4	0.0049521	0.0099	0.0936
HOXA5	0.0183021	0.0094	0.0241
HOXA10	0.022956	0.0291	0.2814
HOXB5	0.0404683	0.1189	0.4742

HOX genes that discriminate inv(16) and t(16;16) cases from other favorable prognosis subtypes. On the basis of a permutation distribution of the *t*-test statistics for each gene, with 95% confidence, there were no false-positives among the first eight genes. Shown are the geometric means derived from the log 2-transformed expression values and the parametric *P*-values for the *t*-statistic.