An FLT3 gene-expression signature predicts clinical outcome in normal karyotype AML

Lars Bullinger,1 Konstanze Döhner,1 Raphael Kranz,1 Christoph Stimer,1 Stefan Fröhling,1 Claudia Scholl,1 Young H. Kim,2 Richard F. Schlenk,1 Robert Tibshirani,3 *Hartmut Döhner,1 and *Jonathan R. Pollack2

1Department of Internal Medicine III, University of Ulm, Ulm, Germany; and Departments of 2Pathology and 3Statistics and Health Research & Policy, Stanford University, CA

Acute myeloid leukemia with normal karyotype (NK-AML) represents a cytogenetic group with intermediate prognosis but substantial molecular and clinical heterogeneity. Within this subgroup, presence of FLT3 (FMS-like tyrosine kinase 3) internal tandem duplication (ITD) mutation predicts less favorable outcome. The goal of our study was to discover gene-expression patterns correlated with FLT3-ITD mutation and to evaluate the utility of a FLT3 signature for prognostication. DNA microarrays were used to profile gene expression in a training set of 65 NK-AML cases, and supervised analysis, using the Prediction Analysis of Microarrays method, was applied to build a gene expression–based predictor of FLT3-ITD mutation status. The optimal predictor, composed of 20 genes, was then evaluated by classifying expression profiles from an independent test set of 72 NK-AML cases. The predictor exhibited modest performance (73% sensitivity; 85% specificity) in classifying FLT3-ITD status. Remarkably, however, the signature outperformed FLT3-ITD mutation status in predicting clinical outcome. The signature may better define clinically relevant FLT3 signaling and/or alternative changes that phenocopy FLT3-ITD, whereas the signature genes provide a starting point to dissect these pathways. Our findings support the potential clinical utility of a gene expression–based measure of FLT3 pathway activation in AML. (Blood. 2008; 111:4490-4495)

© 2008 by The American Society of Hematology

Introduction

In acute myeloid leukemia (AML), cytogenetics is used to stratify cases for appropriate risk-adapted therapy.1 The largest subset of patients, the 40% to 49% of AML cases with normal karyotype (NK-AML), comprises a heterogeneous group with intermediate prognosis.2 In the past few years, mutations in genes, such as FLT3, MLL, CEBPA, and NPM1, have been identified in NK-AML, and the presence of such mutations carries prognostic information.2,3 Despite these advances, there is still no clear consensus on how aggressively to treat patients with NK-AML.

Among these molecular alterations, FLT3 mutation has not only been linked to unfavorable outcome but is also a “drugable” target for molecularly directed therapy.4 FLT3 encodes a receptor tyrosine kinase that is preferentially expressed on hematopoietic progenitor cells and mediates stem cell differentiation and proliferation.5 In NK-AML, 28% to 33% of cases harbor an in-frame internal tandem duplication (ITD) mutation within the juxtamembrane (JM) domain of FLT3.6 ITD mutations disrupt the autoinhibitory conformation of the receptor7,8 and promote constitutive activation of FLT3 and downstream effectors, such as RAS and STAT5 (signal transducer and activator of transcription 5).9 The presence of FLT3-ITD mutation in NK-AML predicts shorter remission duration and overall survival.6 Within the group of FLT3-ITD mutations, the size of the ITD10 (but see Ponziani et al11), which varies from a few nucleotides to several hundred, and the ratio of FLT-ITD mutant to FLT3–wild-type (WT) alleles12,13 have each been positively correlated with less favorable outcome.

Approximately 5% to 15% of NK-AML cases carry an activating missense mutation of FLT3, mainly at Asp835 within the activation loop of the tyrosine kinase domain (TKD).14 The most commonly used diagnostic test (EcoRV digest) detects mutations of Asp835 and Ile836, although additional TKD mutations have been described (eg, residues 841 and 842).15,16 In comparison to FLT3-ITD, FLT3-TKD mutations are not strong activators of STAT517 and have not been definitively linked to poor prognosis.2,6,18,19

DNA microarray-based expression profiling also powerfully captures the molecular heterogeneity of cancers and has been applied to build classifiers and clinical-outcome predictors in AML.20-23 Regarding FLT3, prior studies have defined gene-expression patterns associated with FLT3 activation in cultured cells24,25 and in clinical specimens,26,27 but none has assessed the clinical relevance of identified signatures. Here, we profile a large set of clinically annotated NK-AML specimens, apply supervised analysis to define gene-expression patterns characterizing FLT3-ITD mutation, and evaluate the utility of a FLT3 signature in NK-AML classification and prognostication.

Methods

Specimens

NK-AML samples were provided by the German-Austrian AML Study Group (AMLSG) and comprised Ficoll gradient purified mononuclear cells (> 80%
Table 1. Clinicopathologic characteristics of sample sets

<table>
<thead>
<tr>
<th>Clinicopathologic parameter</th>
<th>Training set, n = 65</th>
<th>Test set, n = 72</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex, no. (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>28 (43)</td>
<td>27/72 (38)</td>
</tr>
<tr>
<td>Female</td>
<td>37 (57)</td>
<td>45/72 (63)</td>
</tr>
<tr>
<td><strong>Age,</strong> y [median [range]]</td>
<td>43 (16-60)</td>
<td>49 (27-60)</td>
</tr>
<tr>
<td><strong>WBC,</strong> f × 1000/µL [median [range]]</td>
<td>31 (0.5-238)</td>
<td>14 (0.8-156)</td>
</tr>
<tr>
<td><strong>LDH,</strong> U/L [median [range]]</td>
<td>553 (139-6676)</td>
<td>454 (133-1941)</td>
</tr>
<tr>
<td><strong>FAB n/N (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>6/64 (9)</td>
<td>5/65 (8)</td>
</tr>
<tr>
<td>M1</td>
<td>17/64 (27)</td>
<td>11/65 (17)</td>
</tr>
<tr>
<td>M2</td>
<td>15/64 (23)</td>
<td>21/65 (32)</td>
</tr>
<tr>
<td>M4</td>
<td>16/64 (25)</td>
<td>17/65 (26)</td>
</tr>
<tr>
<td>M5</td>
<td>10/64 (16)</td>
<td>9/65 (14)</td>
</tr>
<tr>
<td>M6</td>
<td>0/64 (0)</td>
<td>2/65 (3)</td>
</tr>
<tr>
<td><strong>FLT3-ITD mutation n/N (%)</strong></td>
<td>28/64 (44)</td>
<td>22/63 (35)</td>
</tr>
<tr>
<td><strong>FLT3-TKD mutation n/N (%)</strong></td>
<td>14/64 (22)</td>
<td>3/63 (5)</td>
</tr>
<tr>
<td><strong>MLL-PTD mutation, n/N (%)</strong></td>
<td>4/65 (6)</td>
<td>5/65 (8)</td>
</tr>
<tr>
<td><strong>CEBPA mutation n/N (%)</strong></td>
<td>6/56 (11)</td>
<td>10/60 (17)</td>
</tr>
<tr>
<td><strong>NPM1 mutation n/N (%)</strong></td>
<td>34/61 (56)</td>
<td>32/60 (53)</td>
</tr>
<tr>
<td><strong>Median follow-up, § d</strong></td>
<td>405</td>
<td>515</td>
</tr>
<tr>
<td><strong>Survivors, §</strong></td>
<td>1360</td>
<td>785</td>
</tr>
<tr>
<td><strong>Median survival time, § d</strong></td>
<td>405</td>
<td>675</td>
</tr>
<tr>
<td><strong>Median remission duration, § d</strong></td>
<td>920</td>
<td>920</td>
</tr>
</tbody>
</table>

*Significant differences between test and training sets:*
1. *P < .05 (Student t test).*
2. *P < .01 (Student t test).*
3. *P < .01 (Fisher exact test).*
5. *P = .01 (Log-rank test).*

---

**Results**

To discover gene-expression patterns associated with FLT3-ITD mutation, we used cDNA microarrays to profile expression across approximately 23,000 different genes in a training set of 65 NK-AML cases (Table 1) representing Ficoll-gradient purified blasts from bone marrow or peripheral blood. We then applied a supervised analysis method, PAM, to define a gene-expression classifier of FLT3-ITD status (hereafter, “FLT3 signature”). The optimal classifier comprised 25 cDNAs, representing 20 different genes (Figure 1A). As technical validation of the gene-expression differences, we carried out Q-RT-PCR of 4 representative signature genes (CFH, SOCS2, STON2, and MYO1B) on 20 samples exhibiting a range of expression levels for these genes. For each of the 4 genes, expression levels quantified by microarray and by Q-RT-PCR exhibited good correlation (*r = 0.55-0.82; Figure 2*).

To evaluate classification performance of the FLT3 signature, we profiled gene expression in an additional 72 cases of NK-AML (Table 1). In this independent test set, the FLT3 signature classified cases with FLT3-ITD mutations with an overall accuracy of 81%, with 73% sensitivity and 85% specificity (Figure 1B). None of the cases misclassified as FLT3-ITD carried FLT3-TKD mutations. For the subset of FLT3-ITD cases misclassified as not FLT3-ITD, there was no significant difference in the length of the ITD (*P = 1, Fisher exact test; combined training and test set, comparing proportion > vs ≤ median) or the FLT3-ITD/WT ratio (*P = .52*).

Expression levels for FLT3, measured by microarray (Figure 1), were also not significantly different in cases correctly or incorrectly classified as FLT3-ITD (*P = .58; Mann-Whitney U test), or as not FLT3-ITD (*P = .16*).

---

**Gene-expression profiling**

RNA was isolated from stored, frozen mononuclear AML-cell pellets using Trizol reagent and RNA quality assessed by gel electrophoresis. We hybridized Cy5-labeled total RNA from AML samples, along with Cy3-labeled universal reference RNA (pooled from 11 cell lines), on cDNA microarrays (manufactured by the Stanford Functional Genomics Facility) containing 39,632 nonredundant cDNA clones, representing 14,613 named genes, 4,925 additional UniGene clusters, and 3,068 ESTs not mapping to UniGene clusters. Hybridized arrays were imaged using an Axon GenePix 4000B scanner, fluorescence ratios extracted using GenePix software, and data entered into the Stanford Microarray Database for subsequent analysis. The complete microarray dataset is accessible at SMD[1] and at the Gene Expression Omnibus (GEO[2]; accession GSE8043).

---

**Quantitative RT-PCR**

To validate differential expression of selected FLT3-ITD classifier genes, we performed quantitative reverse transcription–PCR (Q-RT-PCR) analysis using ABI Assays-on-Demand primer pairs and reagents, run on an ABI Prism 7700 instrument. Expression values are reported normalized to GAPDH levels.

---

**Data analysis**

Fluorescence ratios were normalized by mean-centering genes for each array and then mean-centering each gene across all arrays within different print runs, to minimize potential print run-specific bias. For all subsequent analyses, we included only the 11,233 cDNAs on the microarray whose expression was both well measured and highly variable among samples. Well-measured genes were defined as those with signal intensity/background greater than 2 for either the Cy5-labeled AML sample or the Cy3-labeled reference sample, in at least 75% of AML samples hybridized. We defined variably expressed genes as those whose expression was higher or lower by a factor of at least 4 from the average expression of all AML samples in at least 2 AML samples. The FLT3-ITD gene expression-based classifier was built using the Prediction Analysis of Microarrays (PAM) method, based on nearest shrunken centroids. An optimal predictor set of 25 cDNAs (representing 20 different genes) was determined by minimizing the cross-validation error rate in the training set. The predictor was then evaluated by classifying expression profiles into FLT3-ITD present or absent, from an independent test set of 72 NK-AML cases. Prognostic value of the FLT3 signature was assessed by comparing the clinical outcomes (remission duration, disease-free and overall survival) of the predicted FLT3 groups using Kaplan-Meier survival analysis. Multivariate proportional-hazards analysis was done using the coxph function in the R statistical package.
Figure 1. Expression-based classification of FLT3-ITD status in NK-AML training and test sets. Heatmap representations of gene-expression levels of the PAM-derived 20-gene FLT3-ITD predictor, shown for the NK-AML (A) training and (B) independent test sets. Mean-centered log2 gene-expression ratios are depicted by ratio color scale, indicated. Gray denotes poorly measured data. Genes are ordered by fractional contribution to the PAM predictor, and samples by PAM probabilities of FLT3-ITD classification, also displayed graphically above. Molecular alterations are coded by shaded boxes: FLT3-ITD (■), not FLT3-ITD (□): FLT3-ITD length > 48 nt (median; ■); ≤ 48 nt (□); not FLT3-ITD (□); FLT3-ITD/WT ratio > 0.587 (median; ■); FLT3-ITD/WT ratio ≤ 0.587 (□); FLT3-TKD (□), not FLT3-TKD (□); not available (□). Misclassified samples are indicated by gray arrowhead (misclassified as not FLT3-ITD) or black arrowhead (misclassified as FLT3-ITD).

Although classification performance was modest, we nevertheless assessed the potential prognostic value of the FLT3 signature by Kaplan-Meier analysis. The FLT3 signature was a significant predictor of early relapse (P = .003, log-rank test; hazard ratio = 3.6; 1.5-8.8; Figure 3B) as well as unfavorable disease-free (P < .001, hazard ratio = 3.3; 1.6-6.6; Figure 3D) and overall survival (P = .001, hazard ratio = 3.9; 1.5-10.6; Figure 3F), and, remarkably, for all 3 outcome measures outperformed FLT3-ITD mutation status (Figure 3A,C,E). In multivariate analysis along with other prognostic factors, the FLT3 signature was the strongest prognosticator (Table 2).

Of note, none of the 6 cases misclassified as FLT3-ITD carried MLL-partial tandem duplication (PTD) (an alteration associated with poor prognosis), and cases misclassified (compared with correctly classified) as not FLT3-ITD were not enriched for CEBPA or NPM1 mutations (associated with good prognosis; P = 1, Fisher exact test). For misclassified cases, there were also no significant differences in patient age and serum lactate dehydrogenase levels (P values .08-.59, Mann-Whitney U test), nor in the proportion of patients treated by allogeneic stem cell transplant (P = 1, Fisher exact test). The only significant association found was between cases misclassified as FLT3-ITD and lower (ie, nonelevated) white blood cell (WBC) counts (P = .04, Mann-Whitney U test). Interestingly, the FLT3 signature prognostically stratified both the FLT3-ITD-positive and FLT3-ITD-negative subgroups, although only the latter reached statistical significance (Figure 4).

To assess the robustness of our findings, we performed 50 random training test splits of the combined 137 NK-AML samples. For each random split, an optimal PAM FLT3-ITD classifier was built using the training set and evaluated on the test set. On average, the FLT3 signature (median P = .001, hazard ratio = 7.4) outperformed FLT3-ITD mutation status (median P = .003, hazard ratio = 2.7) in predicting overall survival. The FLT3 signature also exhibited relative stability, with 8 of the 20 original signature genes occurring within the FLT3 classifier of
at least 50% of the 50 random training test splits and 14 of the original 20 genes found in at least 40% of the splits.

**Discussion**

By expression profiling and PAM analysis, we have defined a 20-gene FLT3-ITD classifier signature. Notably, despite modest FLT3-ITD classification accuracy, the 20-gene classifier outperformed FLT3-ITD mutation status in predicting disease relapse as well as disease-free and overall survival for patients with NK-AML, a finding not explainable by the presence of other known prognostic mutations (MLL-PTD, CEBP1, or NPM1), by other prognostic factors (age, lactate dehydrogenase, WBC) or by a spurious association with treatment. (Note, the one significant association found between cases misclassified as FLT3-ITD and lower WBC count (P = 0.04) would lose significance after correcting for multiple hypothesis testing; regardless, because lower WBC

![Figure 3](Figure3.png)

**Table 2. Multivariate proportional hazards analysis**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard ratio (95% CI)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted FLT3-ITD</td>
<td>3.68 (1.41-9.61)</td>
<td>0.01†</td>
</tr>
<tr>
<td>FLT3-ITD-/NPM1 mutation</td>
<td>1.53 (0.63-3.71)</td>
<td>0.35</td>
</tr>
<tr>
<td>CEBPA mutation</td>
<td>0.72 (0.20-2.61)</td>
<td>0.62</td>
</tr>
<tr>
<td>Age, per year</td>
<td>1.05 (1.00-1.10)</td>
<td>0.04†</td>
</tr>
<tr>
<td>WBC, per 1000/μL</td>
<td>1.00 (0.99-1.01)</td>
<td>0.81</td>
</tr>
<tr>
<td>LDH, per U/L</td>
<td>1.00 (1.00-1.00)</td>
<td>0.85</td>
</tr>
</tbody>
</table>

*Wald test.
†Significant P value.

![Figure 4](Figure4.png)

**Figure 4.** FLT3 signature prognostically stratifies both FLT3-ITD+ and FLT3-ITD− subgroups. Kaplan-Meier overall survival analysis of NK-AML independent test set for PAM-predicted FLT3-ITD mutation in (A) FLT3-ITD+ subgroup and (B) FLT3-ITD− subgroup. P values (log rank test) are indicated.
counts are associated with good prognosis any confounding effect would be in the wrong direction.) Rather, we speculate that the superior prognostic performance of the FLT3 signature reflects one or both of 2 possibilities. First, FLT3-ITD cases classified falsely by the signature as not having ITD might be associated with quantitatively or qualitatively less FLT3 pathway activation. Because systematic differences in the length of ITD, the ITD/WT ratio, and FLT3 mRNA levels were not observed, other alterations in FLT3 or downstream pathway genes would presumably account for the observed decreased FLT3 signature.

A second possibility is that cases classified falsely as FLT3-ITD might harbor alternative genetic or epigenetic alterations that phenocopy FLT3-ITD and have similar downstream gene expression read-outs and clinical impact. Because FLT3-TKD mutations were not observed among these cases and there was no significant difference in FLT3 mRNA levels, other alterations in FLT3 or other genes would have to account for the FLT3 pathway activation signature. Other alterations in FLT3 might include JM domain point mutations as well as a recently identified mutation in the N-terminal lobe of the TKD (K663Q), which have similar signaling characteristics, at least in vitro. Alterations in other genes might include recently described CBL-inactivating mutations, resulting in up-regulated wild-type FLT3 receptor signaling in AML. Because the FLT3 signature prognostically stratified both the FLT3-ITD-positive and -negative subgroups, probably both the above classes of possibility are operative. Regardless, our data suggest that a gene expression-based read-out might be a more accurate measure of clinically meaningful FLT3 pathway activation.

The genes included in the FLT3 signature may provide insight into FLT3-ITD-associated leukemogenesis. Some of the genes have known links to AML. For example, SOCS2 (Suppressor of cytokine signaling 2) was previously identified among genes up-regulated in 32Dc23 murine leukemia cells stably transfected with FLT3-ITD. SOCS2 is a known target gene induced by STAT5 and functions as a negative regulator of cytokine signaling. HOX genes, such as FLT3 signature genes HOXB2, HOXB5, and PBX, are frequently overexpressed in AML and have been linked to leukemogenesis. Other signature genes have no prior link to leukemia but have plausible leukemogenic roles based on known functions. For example, DPPA4 was recently identified in a shRNA screen for genes required for self-renewal in embryonic stem cells and may have a similar function in leukemic cells. LAPTM4b (lysosome-associated protein transmembrane 4 beta), up-regulated in several solid tumor types, has been shown to transform NIH3T3 cells. Other signature genes have no obvious link to FLT3 pathway activation but may suggest new areas of investigation. For example, CFH (complement factor H), one of the top FLT3-ITD discriminator genes, encodes a serum glycoprotein that regulates the alternative complement pathway but has not previously been linked to cancer pathogenesis. CFHR1 is a transcribed pseudogene of CFH.

Other prognostic signatures have been reported for NK-AML. By unsupervised analysis, Valk et al described clinically relevant AML subgroups with distinct gene-expression patterns, including 3 subgroups comprising mostly NK-AML cases. In our own prior studies, we used a semisupervised approach to train a clinical-outcome predictor in AML (including cases with cytogenetic aberration). The resultant 133-gene signature predicted outcome independent of FLT3-ITD and was also prognostic within the NK-AML subgroup. The 133-gene signature was recently validated in another NK-AML cohort but was found not independent of FLT3-ITD. We have applied a similar semisupervised approach (training on outcomes rather than FLT3-ITD) in the current NK-AML dataset, finding prognostic signatures (L.B., J.R.P., unpublished observation, December 2006) with substantial overlap with the FLT3 signature, suggesting that FLT3-ITD is a major component of prognostically relevant gene expression in NK-AML. Interestingly, in breast cancer, where multiple prognostic signatures have also been described, a recent study revealed concordance among individual patient outcome predictions, suggesting the different signatures are tracking a shared underlying biology. Future studies should clarify which of the proposed NK-AML outcome signatures might ultimately translate to clinical utility, using either microarray-based testing or other platforms.

Our finding that a gene expression–based FLT3-ITD mutation classifier is a better prognosticator than mutation status itself is reminiscent of a recent study of TP53 mutation. Miller et al defined a gene-expression signature of TP53 mutation in breast tumors. Although TP53 mutation is known to be prognostic, the TP53 signature was a stronger prognosticator and predictor of therapeutic response. Our study provides further support that gene-expression signatures might in general be more effective reporters of pathway status, by integrating downstream transcriptional responses of pathway activation.

Our study indicates a potential utility of the FLT3 signature for improved prognostication and risk-assessed treatment stratification for patients with NK-AML. Validation of prognostic performance in additional patient cohorts is warranted. In addition, and although speculative, the FLT3 signature might provide utility in guiding therapy with selective FLT3 inhibitors, several of which are now being evaluated in clinical trial. Specifically, the FLT3 signature might inform the selection of patients likely (or unlikely) to respond, where patients with the activated signature might benefit from treatment in the absence of FLT3 mutation and vice versa. Measuring the FLT3 signature might also be useful in monitoring the therapeutic response in patients treated with FLT3 inhibitors. Our findings provide a starting point to further dissect FLT3-ITD activation pathways and support the potential clinical utility of a gene expression–based measure of FLT3 pathway activation in patients with NK-AML.

Acknowledgments
The authors thank the Stanford Functional Genomics Facility for microarray manufacture, the Stanford Microarray Database for database support, the members of the German-Austrian AMLSG for their continuous support of the treatment protocols, and Keyan Salari for helpful discussion.

This work was supported by grants from the Leukemia & Lymphoma Society (6151-06; J.R.P., H.D.) and Else-Kröner-Fresenius-Stiftung (L.B., K.D.).

Authorship
Contribution: L.B. and J.R.P. designed and performed research, analyzed and interpreted data, and wrote the paper; K.D. designed the research, contributed vital new reagents, and analyzed and interpreted data; R.K., C. Stürner, C. Scholl, and Y.H.K. performed research and analyzed and interpreted data; S.F. performed research, analyzed and interpreted data, and wrote the paper; R.F.S. contributed vital new reagents, analyzed and interpreted data, and
wrote the paper; R.T. and H.D. designed the research, analyzed and interpreted data, and wrote the paper.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

References


17. Mead AJ, Linch DC, Hills RK, Wheatley K, Burnett AK, Gale RE. FLT3 tyrosine kinase domain mutations are biologically distinct from and have a significantly more favorable prognosis than FLT3 internal tandem duplications in patients with acute myeloid leukemia. Blood. 2007;110:1262-1270.


An FLT3 gene-expression signature predicts clinical outcome in normal karyotype AML

Lars Bullinger, Konstanze Döhner, Raphael Kranz, Christoph Stirner, Stefan Fröhling, Claudia Scholl, Young H. Kim, Richard F. Schlenk, Robert Tibshirani, Hartmut Döhner and Jonathan R. Pollack