ORIGINAL ARTICLE

# Prognostically Useful Gene-Expression Profiles in Acute Myeloid Leukemia

Peter J.M. Valk, Ph.D., Roel G.W. Verhaak, M.Sc., M. Antoinette Beijen,
Claudia A.J. Erpelinck, Sahar Barjesteh van Waalwijk van Doorn-Khosrovani, M.Sc.,
Judith M. Boer, Ph.D., H. Berna Beverloo, Ph.D., Michael J. Moorhouse, Ph.D.,
Peter J. van der Spek, Ph.D., Bob Löwenberg, M.D., Ph.D., and Ruud Delwel, Ph.D.

## ABSTRACT

### BACKGROUND

In patients with acute myeloid leukemia (AML) a combination of methods must be used to classify the disease, make therapeutic decisions, and determine the prognosis. However, this combined approach provides correct therapeutic and prognostic information in only 50 percent of cases.

### METHODS

We determined the gene-expression profiles in samples of peripheral blood or bone marrow from 285 patients with AML using Affymetrix U133A GeneChips containing approximately 13,000 unique genes or expression-signature tags. Data analyses were carried out with Omniviz, significance analysis of microarrays, and prediction analysis of microarrays software. Statistical analyses were performed to determine the prognostic significance of cases of AML with specific molecular signatures.

### RESULTS

Unsupervised cluster analyses identified 16 groups of patients with AML on the basis of molecular signatures. We identified the genes that defined these clusters and determined the minimal numbers of genes needed to identify prognostically important clusters with a high degree of accuracy. The clustering was driven by the presence of chromosomal lesions (e.g., t(8;21), t(15;17), and inv(16)), particular genetic mutations (*CEBPA*), and abnormal oncogene expression (*EVI1*). We identified several novel clusters, some consisting of specimens with normal karyotypes. A unique cluster with a distinctive gene-expression signature included cases of AML with a poor treatment outcome.

### CONCLUSIONS

Gene-expression profiling allows a comprehensive classification of AML that includes previously identified genetically defined subgroups and a novel cluster with an adverse prognosis.

ACUTE MYELOID LEUKEMIA (AML) IS NOT a single disease but a group of neoplasms with diverse genetic abnormalities and variable responses to treatment. Cytogenetics and molecular analyses can be used to identify subgroups of AML with different prognoses. For instance, the translocations inv(16), t(8;21), and t(15;17) herald a favorable prognosis, whereas other cytogenetic aberrations indicate poor-risk leukemia.[1-5] Abnormalities involving 11q23, t(6;9), or 7(q) are defined as poor-risk markers by some groups[2,3] and as intermediate-risk markers by others.[3-5] These inconsistencies and the absence of cytogenetic abnormalities in a considerable proportion of patients argue for refinement of the classification of AML.

Additional reasons for extending the molecular analyses of AML are exemplified by findings regarding the gene for fms-like tyrosine kinase 3 (*FLT3*), the gene encoding ectotropic viral integration 1 site (*EVI1*), and the gene for CCAAT/enhancer binding protein alpha (*CEBPA*). An internal tandem duplication in *FLT3*, a hematopoietic growth factor receptor, is the most common molecular abnormality in AML.[6,7] The presence of such mutations in *FLT3* and elevated expression of the transcription factor EVI1 confer a poor prognosis,[6-8] whereas mutations in *CEBPA* are associated with a good outcome.[9,10]

Molecular classification based on DNA-expression profiling offers a powerful way of distinguishing myeloid from lymphoid cancer and subclasses within these two diseases.[11-14] DNA-microarray analysis has the potential to identify distinct subgroups of AML with the use of one comprehensive assay, to classify cases that currently resist categorization by means of other methods, and to identify subgroups with favorable or unfavorable prognoses within genetically defined subclasses. The goals of this study of 285 adults with AML were to use gene-expression profiles to identify established and novel subclasses of AML and otherwise unrecognized cases of poor-risk AML.

## METHODS

### PATIENTS AND CELL SAMPLES

Eligible patients had received a diagnosis of primary AML, which had been confirmed by means of a cytologic examination of blood and bone marrow (Table 1). All patients were treated according to the protocols of the Dutch–Belgian Hematology–Oncology Cooperative group (available at www.hovon.nl).[15-17] All subjects provided written informed consent. A total of 285 patients provided bone marrow aspirates or peripheral-blood samples at the time of diagnosis and 8 healthy control subjects provided peripheral-blood samples or bone marrow aspirates. Blasts and mononuclear cells were purified by Ficoll–Hypaque (Nygaard) centrifugation and cryopreserved. CD34+ cells from three control subjects were sorted by means of a fluorescence-activated cell sorter. The AML samples contained 80 to 100 percent blast cells after thawing, regardless of the blast count at diagnosis.

### ISOLATION AND QUALITY CONTROL OF RNA

After thawing, cells were washed once with Hanks' balanced-salt solution. High-quality total RNA was extracted by lysis with guanidinium thiocyanate followed by cesium chloride–gradient purification.[18] RNA levels, quality, and purity were assessed with the use of the RNA 6000 Nano assay on the Agilent 2100 Bioanalyzer (Agilent). None of the samples showed RNA degradation (ratio of 28S ribosomal RNA to 18S ribosomal RNA of at least 2) or contamination by DNA.

### GENE PROFILING AND QUALITY CONTROL

Samples were analyzed with the use of Affymetrix U133A GeneChips. Each gene on this chip is represented by 10 to 20 oligonucleotides, termed a "probe set." The intensity of hybridization of labeled messenger RNA (mRNA) to these sets reflects the level of expression of a particular gene. The U133A GeneChip contains 22,283 probe sets, representing approximately 13,000 genes. We used 10 μg of total RNA to prepare antisense biotinylated RNA. Single-stranded complementary DNA (cDNA) and double-stranded cDNA were synthesized according to the manufacturer's protocol (Invitrogen Life Technologies) with the use of the T7-(deoxythymidine)24-primer (Genset). In vitro transcription was performed with biotin-11-cytidine triphosphate and biotin-16-uridine triphosphate (Perkin–Elmer) and the MEGAScript T7 labeling kit (Ambion). Double-stranded cDNA and complementary RNA (cRNA) were purified and fragmented with the GeneChip Sample Cleanup Module (Affymetrix). Biotinylated RNA was hybridized to the Affymetrix U133A GeneChip (45°C for 16 hours). Staining, washing, and scanning procedures were carried out as described in the GeneChip Expression Analysis technical manual (Affymetrix). All GeneChips were visually inspected for irregularities. The global method of scaling, or normaliza-

---

**Glossary**

**Centroid:** In a self-organizing topologic map of gene expression, the centroid corresponds to the center of a cluster.
**Chromosomal abnormalities**
   **t(8;21):** One of the commonest cytogenetic abnormalities in AML; produces a hybrid gene by fusing *AML1* on the long arm of chromosome 21 with *ETO* on the long arm of chromosome 8.
   **inv(16):** Inversion of a segment of chromosome 16 that produces the *CBFβ-MYH11* fusion.
   **t(15;17):** Reciprocal translocation of genetic material between the long arms of chromosomes 15 and 17 that produces the *PML-RARα* fusion gene, typical of acute promyelocytic leukemia.
   **11q23:** A chromosomal region that becomes rearranged with various partner chromosomal regions in diverse forms of leukemia, involving the *MLL* gene.
   **t(6;9):** A rare translocation often found in young patients and sometimes associated with basophilia.
   **−7(q):** Loss of the long arm of chromosome 7, on monosomy 7.
**French–American–British (FAB) classification:** An internationally agreed-on method of classifying acute leukemia by morphologic means. There are eight subtypes, ranging from M0 (myeloblasts) to M7 (megakaryoblasts).
**Gene-expression profiling:** Determination of the level of expression of thousands of genes through the use of microarrays. Messenger RNA extracted from the test tissue or cells and labeled with a fluorescent dye is tested for its ability to hybridize to the spotted nucleic acids.
**Microarray or GeneChip:** A robotically spotted array of thousands of complementary DNAs or oligonucleotides.
**Patient-clustering technique:** A method of grouping patients with similar patterns of gene expression.
**Pearson's correlation coefficient:** A statistical measure of the strength of the relationship between variables.
**Pearson's Correlation Visualization tool of Omniviz:** Omniviz is a commercial multifunctional statistical package used for analysis of microarray data. It allows the visual representation of gene-expression profiles of patients in a Pearson's Correlation View.
**Prediction analysis of microarrays (PAM):** A statistical technique that identifies a subgroup of genes that best characterizes a predefined class.
**Probe set:** A group of 10 to 20 oligonucleotides; each set corresponds to one gene.
**Significance analysis of microarrays (SAM):** A statistical method used in microarray analyses that identifies genes that are significantly differentially expressed between groups of patients on the basis of a change in the level of gene expression relative to the standard deviation of repeated measurements.
**Supervised analysis:** An analysis of the results of microarray profiling that takes external factors into account.
**Unsupervised analysis:** An analysis of the results of microarray profiling that does not take external factors such as survival or clinical signs into account.
**10-Fold cross-validation:** A validation method that works as follows: the model is fitted on 90 percent of the samples, and the class of the remaining 10 percent is then predicted. This procedure is repeated 10 times, with each part playing the role of the test samples and the error of all 10 parts added together to compute the overall error. The error within the validation set reflects the number of samples wrongly predicted to be in this set.

---

tion, was applied, and the mean (±SD) difference between the scaling, or normalization, factors of all GeneChips (293 samples; 285 from patients with AML, 5 from subjects with normal bone marrow, and 3 from subjects with CD34+ cell samples) was 0.70±0.26. All additional measures of quality — the percentage of genes present (50.6±3.8), the ratio of action 3′ to 5′ (1.24±0.19), and the ratio of *GAPDH* 3′ to 5′ (1.05±0.14) — indicated a high overall quality of the samples and assays. Detailed clinical, cytogenetic, and molecular cytogenetic information is available at the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo, accession number GSE1159).

### DATA NORMALIZATION, ANALYSIS, AND VISUALIZATION

All intensity values were scaled to an average value of 150 per GeneChip according to the method of glob-

al scaling, or normalization, provided in the Affymetrix Microarray Suite software, version 5.0 (MAS5.0). Since our methods reliably identify samples with an average intensity value of 30 or more but do not reliably discriminate values between 0 and 30, these values were set to 30. This procedure affected 31 percent of all intensity values, of which 64 percent were flagged as absent by the MAS5.0 software, 3 percent were flagged as marginal, and 33 percent were flagged as present according to the MAS5.0 software.

For each probe set, the geometric mean of the hybridization intensities of all samples from the patients was calculated. The level of expression of each probe set in every sample was determined relative to this geometric mean and logarithmically transformed (on a base 2 scale) to ascribe equal weight to gene-expression levels with similar relative distances to the geometric mean. Deviation from the

**Table 1.** Clinical and Molecular Characteristics of the 285 Patients with Newly Diagnosed AML.

| Characteristic | Value |
|---|---|
| Sex — no. (%) | |
|     Male | 137 (48) |
|     Female | 148 (52) |
| Age group — no. (%) | |
|     <35 yr | 76 (27) |
|     35–60 yr | 177 (62) |
|     ≥60 yr | 32 (11) |
| Age — yr | |
|     Median | 44 |
|     Range | 15–78 |
| White-cell count — $\times10^{-3}/mm^3$ | |
|     Median | 28 |
|     Range | 0.3–582 |
| Bone marrow blasts count — % | |
|     Median | 66 |
|     Range | 0–98 |
| Platelet count — $\times10^{-3}/mm^3$ | |
|     Median | 45 |
|     Range | 3–931 |
| French–American–British classification — no. (%) | |
|     M0 | 6 (2) |
|     M1 | 63 (22) |
|     M2 | 66 (23) |
|     M3 | 19 (7) |
|     M4 | 53 (19) |
|     M5 | 65 (23) |
|     M6 | 3 (1) |
|     Not determined | 10 (4) |
| Cytogenetic abnormalities — no. (%)* | |
|     t(15;17) | 18 (6) |
|     t(8;21) | 22 (8) |
|     inv(16)/t(16;16) | 19 (7) |
|     +8 | 26 (9) |
|     +11 | 7 (2) |
|     +21 | 2 (1) |
|     −5 | 3 (1) |
|     −5(q) | 1 (<1) |
|     −7 | 13 (5) |
|     −7(q) | 7 (2) |
|     3q | 6 (2) |
|     t(6;9) | 4 (1) |
|     t(9;22) | 2 (1) |
|     t(11q23) | 19 (7) |
|     Complex karyotype (>3 chromosomal abnormalities) | 11 (4) |
|     Other abnormal karyotypes | 60 (21) |
|     Normal karyotype | 119 (42) |
|     Not determined | 10 (4) |
| Molecular abnormalities — no. (%) | |
|     Mutation | |
|         *FLT3* internal tandem duplication | 78 (27) |
|         *FLT3* tyrosine kinase domain | 33 (12) |
|         N-*RAS* | 26 (9) |
|         K-*RAS* | 9 (3) |
|         *CEBPA* | 17 (6) |
|     Overexpression | |
|         *EVI1* | 23 (8) |

\* All patients with a specific cytogenetic abnormality were included in the analysis, irrespective of the presence of additional abnormalities. A summary of the frequencies and percentages of the cytogenetic and molecular abnormalities for each of the assigned clusters can be found in Table Q of Supplementary Appendix 1 (available with the full text of this article at www.nejm.org). Some samples had more than one abnormality.

geometric mean reflects differential gene expression. The transformed expression data were subsequently imported into Omniviz software, version 3.6 (Omniviz), significance analysis of microarrays (SAM) software, version 1.21, and prediction analysis of microarrays (PAM) software, version 1.12.

*Use of Pearson's Correlation and Visualization Tool*

The Omniviz package was used to perform and visualize the results of unsupervised cluster analysis (an analysis that does not take into account external information such as the morphologic subtype or karyotype). Genes (probe sets) whose level of expression differed from the geometric mean (reflecting up- or down-regulation) in at least one patient were selected for further analysis. The clustering of molecularly recognizable specific groups of patients was investigated with each of the selected probe sets with the use of the Pearson's Correlation and Visualization tool of Omniviz (provided in Fig. B, C, D, E, F, G, and H in Supplementary Appendix 1, available with the full text of this article at www.nejm.org).

*The SAM Method*

All supervised analyses were performed with the use of SAM software.[19] A supervised analysis correlates gene expression with an external variable such as the karyotype or the duration of survival. SAM calculates a score for each gene on the basis of the change in expression relative to the SD of all 285 measurements. The criteria for identifying the top 40 genes for an assigned cluster were a minimal difference in gene expression between the assigned cluster and the other AML samples by a factor of 2 and a q value of less than 2 percent. The q value for each gene represents the probability that it is falsely called significantly deregulated.

*The PAM Method*

All supervised class-prediction analyses were performed by applying PAM software in R (version 1.7.1).[20] The method of the nearest shrunken centroids identifies a subgroup of genes that best characterizes a predefined class. The prediction error was calculated by means of 10-fold cross validation (see the Glossary) within the training set (two thirds of the patients) followed by the use of a second validation set (one third of the patients). All genes identified by the SAM and PAM methods are listed in Supplementary Appendix 1 (Tables A1 to P1 and R).

### REVERSE-TRANSCRIPTASE POLYMERASE CHAIN REACTIONS AND SEQUENCE ANALYSES

Reverse-transcriptase–polymerase-chain-reaction (RT-PCR) assays and sequence analyses for internal tandem duplication and tyrosine kinase domain mutations in *FLT3* and mutations in N-*RAS*, K-*RAS*, and *CEBPA*, as well as real-time PCR for *EVI1* were performed as described previously.[8,9,21,22] AML samples of the clusters characterized by favorable cytogenetic characteristics (t(8;21), t(15;17), and inv(16)) were analyzed for the expression of fusion genes by real-time PCR (Supplementary Appendix 1).

### STATISTICAL ANALYSIS

Statistical analyses were performed with Stata Statistical Software, release 7.0. Actuarial probabilities of overall survival (with failure defined as death from any cause) and event-free survival (with failure defined as incomplete remission [set at day 1], relapse, or death during a first complete remission) were estimated according to the Kaplan–Meier method.

## RESULTS

### VISUAL CORRELATION OF GENE EXPRESSION

All specimens of AML were classified into subgroups with the use of unsupervised ordering (i.e., without taking into account hematologic, cytogenetic, or other external information). Optimal clustering of these specimens was reached with the use of 2856 probe sets (a probe set consists of 10 to 20 oligonucleotides); 2856 sets represent 2008 annotated genes and 146 expressed-sequence tags, which are short sequences of unknown genes (Fig. 1A and Table 2, and Fig. B, C, D, E, F, G, and H in Supplementary Appendix 1).

Sixteen distinct groups of patients with AML were identified on the basis of strong similarities in gene-expression profiles. Figure 1A, a Pearson's correlation view, shows these clusters as red squares along the diagonal. A red rectangle indicates positive pairwise correlations (equality in gene expression between clusters) and a blue rectangle indicates negative pairwise correlations (inequality in gene expression between clusters) (Fig. 1A, and Fig. A in Supplementary Appendix 1). The final Omniviz Correlation View was adapted so that cytologic, cytogenetic, and molecular features were plotted directly adjacent to the original diagonal. This arrangement allowed the visualization of groups of patients with

similar patterns of gene expression along with relevant clinical and genetic findings (Fig. 1B).

Distinct clusters of t(8;21), inv(16), and t(15;17) were readily identified with 1692 probe sets (Table 2). Identification of clusters with mutations in *FLT3*, monosomy 7, or overexpression of *EVI1* required 2856 probe sets (Table 2, and Fig. B, C, D, E, F, G, and H in Supplementary Appendix 1). When more genes were used, the compact pattern of clustering vanished (Table 2). When included in the Omniviz Correlation View analyses (2856 probe sets), all five samples of bone marrow and three CD34+ samples from control subjects gathered within clusters 8 and 10, respectively.

Genes characteristic of each of the 16 clusters were obtained by means of supervised analysis (distinctions on the basis of predefined classes), with the use of the SAM method. The expression profiles of the top 40 genes of each cluster are plotted in Figure 1B beside the correlation view. The SAM analyses identified 599 discriminating genes (Tables A1 to P1 in Supplementary Appendix 1); we were unable to identify a distinct gene profile for cluster 14.

### RECURRENT TRANSLOCATIONS

#### *CBFβ-MYH11*

All AML samples with inv(16), which causes the *CBFβ-MYH11* fusion gene, gathered within cluster 9 (Fig. 1B, and Table I in Supplementary Appendix 1). Four specimens within this cluster were not known to harbor an inv(16), but molecular analysis and Southern blotting revealed that their leukemic cells had the *CBFβ-MYH11* fusion gene (Table I and Fig. I in Supplementary Appendix 1). SAM analysis revealed that *MYH11* was the most discriminative gene for this cluster (Table I1 and Fig. J in Supplementary Appendix 1). Interestingly, a low level of expression of *CBFβ* was correlated with this cluster, perhaps because of the decreased expression or deletion of the *MYH11-CBFβ* alternate fusion gene or down-regulation of the normal *CBFβ* allele by the CBFβ-MYH11 fusion protein.

#### *PML-RARα*

Cluster 12 contained all cases of acute promyelocytic leukemia (APL) with t(15;17) (Fig. 1B, and Table L in Supplementary Appendix 1), including one patient (Patient 322) who had previously received a diagnosis of APL with *PML-RARα* on the basis of RT-PCR alone. SAM analyses revealed that genes for hepatocyte growth factor (*HGF*), macrophage-stim-

**Figure 1 (facing page). Correlation View of Specimens from 285 Patients with AML Involving 2856 Probe Sets (Panel A) and an Adapted Correlation View (2856 Probe Sets) (Right-Hand Side of Panel B), and the Levels of Expression of the Top 40 Genes That Characterized Each of the 16 Individual Clusters (Left-Hand Side of Panel B).**

In Panel A, the Correlation Visualization tool displays pairwise correlations between the samples. The colors of the cells relate to Pearson's correlation coefficient values, with deeper colors indicating higher positive (red) or negative (blue) correlations. One hundred percent negative correlation would indicate that genes with a high level of expression in one sample would always have a low level of expression in the other sample and vice versa. Box 1 indicates a positive correlation between clusters 5 and 9 and box 2 a negative correlation between clusters 5 and 12. The red diagonal line displays the intraindividual comparison of results for a patient with AML (i.e., 100 percent correlation). To reveal the patterns of correlation, we applied a matrix-ordering method to rearrange the samples. The ordering algorithm starts with the most highly correlated pair of samples and, through an iterative process, sorts all the samples into correlated blocks. Each sample is joined to a block in an ordered manner so that a correlation trend is formed within a block, with the most correlated samples at the center. The blocks are then positioned along the diagonal of the plot in a similar ordered manner. Panel B shows all 16 clusters identified on the basis of the Correlation View. The French–American–British (FAB) classification and karyotype based on cytogenetic analyses are depicted in the columns along the original diagonal of the Correlation View; FAB subtype M0 is indicated in black, subtype M1 in green, subtype M2 in purple, subtype M3 in orange, subtype M4 in yellow, subtype M5 in blue, and subtype M6 in gray; normal karyotypes are indicated in green, inv(16) abnormalities in yellow, t(8;21) abnormalities in purple, t(15;17) abnormalities in orange, 11q23 abnormalities in blue, 7(q) abnormalities in red, +8 aberrations in pink, complex karyotypes (those involving more than three chromosomal abnormalities) in black, and other abnormalities in gray. *FLT3* internal tandem duplication (ITD) mutations, *FLT3* mutations in the tyrosine kinase domain (TKD), N-*RAS,* K-*RAS,* and *CEBPA* mutations, and the overexpression of *EVI1* are depicted in the same set of columns: red indicates the presence of a given abnormality, and green its absence. The levels of expression of the top 40 genes identified by the significance analysis of microarrays of each of the 16 clusters as well as in normal bone marrow (NBM) and CD34+ cells are shown on the left side. The scale bar indicates an increase (red) or decrease (green) in the level of expression by a factor of at least 4 relative to the geometric mean of all samples. The percentages of the most common abnormalities (those present in more than 40 percent of specimens) and the percentages of specimens in each cluster with a normal karyotype are indicated.
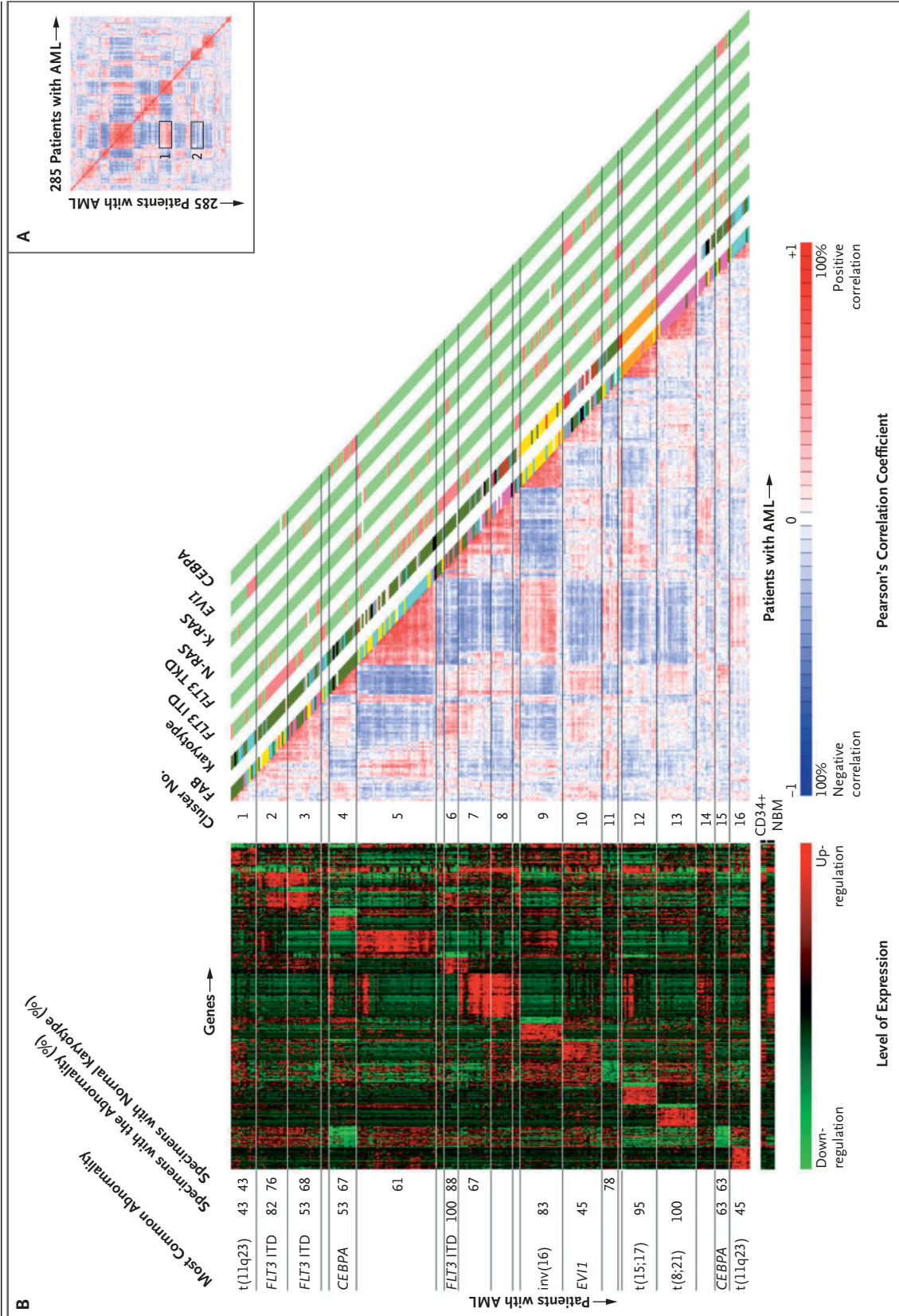
**Table 2.** Evaluation of the Omniviz Correlation View Results on the Basis of the Clustering of AML Specimens with Similar Molecular Abnormalities.*

| Variable | Distribution | | | | | | |
|---|---|---|---|---|---|---|---|
| No. of probe sets | 147 | 293 | 569 | 984 | 1692 | 2856 | 5071 |
| Factor increase or decrease in regulation† | >32 | >22.6 | >16 | >11.3 | >8 | >5.6 | >4 |
| **Chromosomal abnormalities** | | | | | | | |
| t(8;21) | ± | + | + | + | ++ | ++ | + |
| inv(16) | ± | ± | ± | + | ++ | ++ | ++ |
| t(15;17) | ± | + | ++ | ++ | ++ | ++ | + |
| 11q23 | ± | ± | ± | ± | + | + | ± |
| −7(q) | ± | ± | ± | ± | ± | + | ± |
| **Mutation** | | | | | | | |
| *FLT3* internal tandem duplication | ± | ± | ± | ± | ± | ± | ± |
| *FLT3* tyrosine kinase domain | − | − | − | − | − | − | − |
| *N-RAS* | − | − | − | − | − | − | − |
| *K-RAS* | − | − | − | − | − | − | − |
| *CEBPA* | − | ± | ± | + | + | + | + |
| **Overexpression** | | | | | | | |
| *EVI1* | − | − | − | − | ± | + | ± |

\* Two plus signs indicate that 100 percent of specimens were in a single cluster, a single plus sign that specimens were in no more than two recognizable clusters, a plus–minus sign that specimens were in more than two recognizable clusters, and a minus sign that no clustering occurred. Four patients with AML with abnormalities involving chromosome 5 were excluded.

† The factor increase or decrease in the regulation of gene expression is relative to the geometric mean by which the differentially expressed probe sets were selected.

ulating 1 growth factor (*MST1*), and fibroblast growth factor 13 (*FGF13*) were specific for this cluster. In addition, cluster 12 could be separated into two subgroups: one with a high and the other with a low white-cell count (Fig. K in Supplementary Appendix 1). This subdivision corresponds to the presence of *FLT3* internal tandem duplication mutations (Fig. 1B).

*AML1-ETO*

All specimens from patients with the t(8;21) that generates the *AML1-ETO* fusion gene grouped within cluster 13 (Fig. 1B, and Table M in Supplementary Appendix 1). SAM identified *ETO* as the most discriminative gene for this cluster (Table M1 and Fig. L in Supplementary Appendix 1).

**11q23 ABNORMALITIES**

Cases with 11q23 abnormalities were scattered among the 285 samples, although two subgroups were apparent: cluster 1 and cluster 16 (Fig. 1B, and

Tables A and P in Supplementary Appendix 1). Cluster 16, with 11 total cases, contained 4 cases of t(9;11) and 1 case of t(11;19). SAM analyses identified a strong signature of up-regulated genes in most cases in this cluster (Fig. 1B, and Table P1 in Supplementary Appendix 1). Although 6 of 14 cases within cluster 1 also had 11q23 abnormalities, this subgroup was more heterogeneous than cluster 16 (Fig. 1B).

***CEBPA* MUTATIONS**

Mutations in *CEBPA* occur in approximately 7 percent of patients with AML, most with a normal karyotype, and predict a favorable outcome.[9,10] Two clusters (4 and 15) had a high frequency of *CEBPA* mutations (Fig. 1B). The sets of up-regulated or down-regulated genes in cluster 4 discriminated the specimens it contained from those in cluster 15 (Table D1 in Supplementary Appendix 1). The up-regulated genes included the T-cell genes *CD7* and the T-cell receptor delta locus, which may be ex-

pressed by immature AML cells.[23,24] All but one of the top 40 genes of cluster 15 were down-regulated (Table O1 in Supplementary Appendix 1). These genes were also down-regulated in cluster 4 (Fig. 1B). The genes encoding alpha1-catenin (*CTNNA1*), tubulin beta-5 (*TUBB5*), and Nedd4 family interacting protein 1 (*NDFIP1*) were the only down-regulated genes among the top 40 in both cluster 4 and cluster 15.

### OVEREXPRESSION OF *EVI1*

High levels of expression of *EVI1*, which occur in approximately 10 percent of cases of AML, predict a poor outcome.[8] In cluster 10, 10 of 22 specimens (Table J in Supplementary Appendix 1) showed increased expression of *EVI1*, and 6 of these 10 specimens had chromosome 7 abnormalities. In cluster 8, 4 of 13 specimens also had chromosome 7 aberrations (Table H in Supplementary Appendix 1), but since its molecular signature differed from that of cluster 10 (Fig. 1B), the high level of expression of *EVI1* or *EVI1*-related proteins may have determined the molecular profile of cluster 10. In the heterogeneous cluster 1, 5 of 14 specimens also had increased *EVI1* expression. These specimens may have appeared outside cluster 10 because their molecular signatures were most likely the result of the overexpression of *EVI1* and an 11q23 abnormality.

### *FLT3* AND *RAS* MUTATIONS

Samples from most patients in clusters 2, 3, and 6 harbored a *FLT3* internal tandem duplication (Fig. 1B). Almost all these patients had a normal karyotype. The presence of *FLT3* internal tandem duplication seemed to divide clusters 3, 5, and 12 into two groups. Other individual specimens with a *FLT3* internal tandem duplication were dispersed over the entire series; mutations in the tyrosine kinase domain of *FLT3* were not clustered. Likewise, mutations in codon 12, 13, or 61 of the small GTPase RAS (N-*RAS* and K-*RAS*) had no apparent signatures and did not aggregate in the Correlation View (Fig. 1B).

### OTHER CLUSTERS

Specimens from patients with AML with a normal karyotype clustered into several subgroups within the assigned clusters (Fig. 1B). Most patients in cluster 11 had normal karyotypes and no consistent additional abnormality. Cluster 5 contained mainly specimens from patients with AML of subtype M4 or M5, according to the French–American–British (FAB) classification (Fig. 1B). Clusters 7, 8, 11, and 14 were not associated with a FAB subtype but had distinct gene-expression profiles.

### CLASS PREDICTION OF DISTINCT CLUSTERS

We used the PAM method to validate the cluster-specific genes identified by the SAM method and to determine the minimal number of genes that can be used to predict karyotypic or other genetic abnormalities with biologic significance in AML (Table 3). The 285 specimens were randomly divided into a training set (189 specimens) and a validation set (96 specimens). All patients in the validation set who had favorable cytogenetic findings were identified with 100 percent accuracy with the use of only a few genes (Table 3). As expected from the SAM analyses, *ETO* for t(8;21), *MYH11* for inv(16), and *HGF* for t(15;17) were among the best predictors of the cytogenetic abnormalities (Table R in Supplementary Appendix 1). Cluster 10 (which involved *EVI1* overexpression) was predicted with a high degree of accuracy, although with a higher 10-fold cross-validation error than that in the groups with favorable cytogenetic findings. In cluster 16 (involving 11q23 abnormalities), samples from 3 of 96 patients were wrongfully identified in the validation set. Since cluster 15 (involving *CEBPA* mutations) contained few samples, we combined both *CEBPA*-containing clusters. These combined clusters predicted the presence of *CEBPA* mutations within the validation set with 98 percent accuracy. We were unable to identify a signature that reliably identified *FLT3* internal tandem duplications.

### SURVIVAL ANALYSES

Overall survival, event-free survival, and relapse rates were determined among patients whose specimens were within clusters containing more than 20 specimens in the Correlation View (clusters 5, 9, 10, 12, and 13) (Fig. 2). The mean (±SE) actuarial probabilities of overall survival and event-free survival at 60 months were 59±10 percent and 55±11 percent, respectively, among patients with samples in cluster 13; 57±12 percent and 47±11 percent, respectively, among those with samples in cluster 12; and 72±10 percent and 52±10 percent, respectively, among those with samples in cluster 9. Patients with samples in cluster 5 had an intermediate rate of overall survival (32±8 percent) and event-free survival (27±8 percent), whereas survival among patients with samples in cluster 10 was poorer (the overall survival rate was 18±9 percent, and the event-

**Table 3.** Results of Class Prediction Analysis with the Use of Prediction Analysis of Microarrays.*

| Abnormality | Training Set (N=189) | Validation Set (N=96) | No. of Probe Sets Used | No. of Genes Represented |
|---|---|---|---|---|
| | *no. of errors* | | | |
| t(8;21), leading to *AML1-ETO* (cluster 13) | 0 | 0 | 3 | 2 |
| t(15;17), leading to *PML-RARα* (cluster 12) | 1 | 0 | 3 | 2 |
| inv(16), leading to *CBFβ-MYH11* (cluster 9) | 0 | 0 | 1 | 1 |
| 11q23 (cluster 16) | 3 | 3 | 31 | 25 |
| *EVI1* (cluster 10) | 16 | 0 | 28 | 25 |
| *CEBPA* (cluster 4) | 8 | 2 | 13 | 8 |
| *CEBPA* (cluster 15) | 17 | 6† | 36 | 32 |
| *CEBPA* (clusters 4 and 15) | 5 | 2 | 9 | 5 |
| *FLT3* internal tandem duplication | 27 | 21 | 56 | 41 |

* Prediction analysis of microarrays was performed to define the minimal numbers of genes that could predict whether a specimen from a particular patient belonged in one of the clusters (first column). The group of patients was randomly segregated into a training set (second column) and a validation set (third column). The 10-fold method of cross-validation, applied on the training set, works as follows: the model is fitted on 90 percent of the samples, and the class of the remaining 10 percent is then predicted. This procedure is repeated 10 times, with each part playing the role of the test samples and the error of all 10 parts added together to compute the overall error (second column). The minimal numbers of probe sets or genes (fourth and fifth columns, respectively) that were identified in the training were tested on the validation set (third column). The error within the validation set (third column) reflects the number of samples wrongly predicted in this set. The identities of the probe sets and genes are provided in Table R of Supplementary Appendix 1.
† After randomization none of the patients with *CEBPA* abnormalities in cluster 15 were included in the validation set.

free survival rate was 6±6 percent), mainly as a result of an increased incidence of relapse (Fig. 2C).

## DISCUSSION

In this study of 285 patients with AML that was characterized by cytogenetic analyses and extensive molecular analyses, we used gene-expression profiling to comprehensively classify the disorder. This method identified 16 groups on the basis of unsupervised analyses involving Pearson's correlation coefficient. Our results provide evidence that each of the assigned clusters represents true subgroups of AML with specific molecular signatures.

We were able to cluster all cases of AML with t(8;21), inv(16), or t(15;17), including those that had not been identified by cytogenetic examination, into three clusters with unique gene-expression profiles. Correlations between gene-expression profiles and prognostically favorable cytogenetic aberrations have been reported by others,[12,13] but we found that these cases can be recognized with a high degree of accuracy within a representative cohort of patients with AML.

The SAM and PAM methods were highly con-

cordant for the genes identified within the assigned clusters, indicating that these clusters contained discriminative genes. For instance, clusters 4 and 15, with overlapping signatures, both included specimens with normal karyotypes and mutations in *CEBPA*. Multiple genes appeared to be down-regulated in both clusters but were unaffected in any other subgroup of AML.

The discriminative genes identified by SAM and PAM may reveal functional pathways that are critical for the development of AML. These methods of statistical treatment of the data identified several genes that are implicated in specific subtypes of AML, such as the interleukin-5 receptor α (*IL5Rα*) gene in AML with t(8;21) abnormalities[25] and FLT3-STAT-5 targets — the gene for interleukin-2 receptor α (*IL2Rα*)[26] and the pim1 kinase gene (*PIM1*)[27] — in AML with *FLT3* internal tandem duplication mutations.

Five clusters (5, 9, 10, 12, and 13) with 20 or more specimens were evaluated in relation to outcome of disease. As expected, clusters 9 (involving *CBFβ-MYH11*), 12 (involving *PML-RARα*), and 13 (involving *AML1-ETO*) contained specimens with a relatively favorable prognosis.

Specimens in cluster 10 had a distinctly poor outcome. A randomly selected subgroup of patients with specimens in this cluster could be identified with a high degree of accuracy with the use of a minimal number of genes. The high frequency of poor prognostic markers in this cluster (−7(q), −5(q), t(9;22), or high levels of expression of *EVI1*) is in accord with the poor outcome of patients in this cluster. Since this cluster is heterogeneous with regard to both known poor-risk markers and the presence or absence of these markers, the molecular signature of this cluster may signify a biochemical pathway that causes a poor outcome. The fact that normal CD34+ cells segregate into this cluster suggests that the molecular signature of treatment resistance resembles that of normal hematopoietic stem cells.

The 44 patients with specimens in cluster 5 had an intermediate duration of survival. Since these specimens were of the FAB M4 or M5 subtype, it is possible that genes related to monocytes or macrophages were important in the clustering of these cases.

In three clusters more than 75 percent of specimens had a normal karyotype (clusters 2, 6, and 11). Most of the patients with specimens in clusters 2 and 6 had *FLT3* internal tandem duplication mutations, whereas patients with specimens in cluster 11, which had a discriminative molecular signature, did not have any consistent molecular abnormality.

Clusters 1 and 16 harbored 11q23 abnormalities, representing defects involving the mixed-lineage leukemia (*MLL*) gene. The different gene-expression profiles of these two clusters are most likely due to additional distinctive genetic defects. In cluster 1, this additional abnormality may be a high level of expression of the oncogene *EVI1*, which was not apparent in cluster 16. Similarly, distinctive additional genetic defects may explain the separation of clusters 4 and 15, both of which contained specimens with *CEBPA* mutations, clusters 1 and 10, both of which had high levels of *EVI1* expression, and clusters 8 and 10, both of which had a high frequency of monosomy 7.

Internal tandem duplications in *FLT3* adversely affect the clinical outcome.[6,7] The molecular signature associated with this abnormality is not distinctive; however, the clustering of specimens with these abnormalities within assigned clusters (e.g., cluster 12) suggests that these internal tandem duplications result in different biologic entities within the scope of AML.
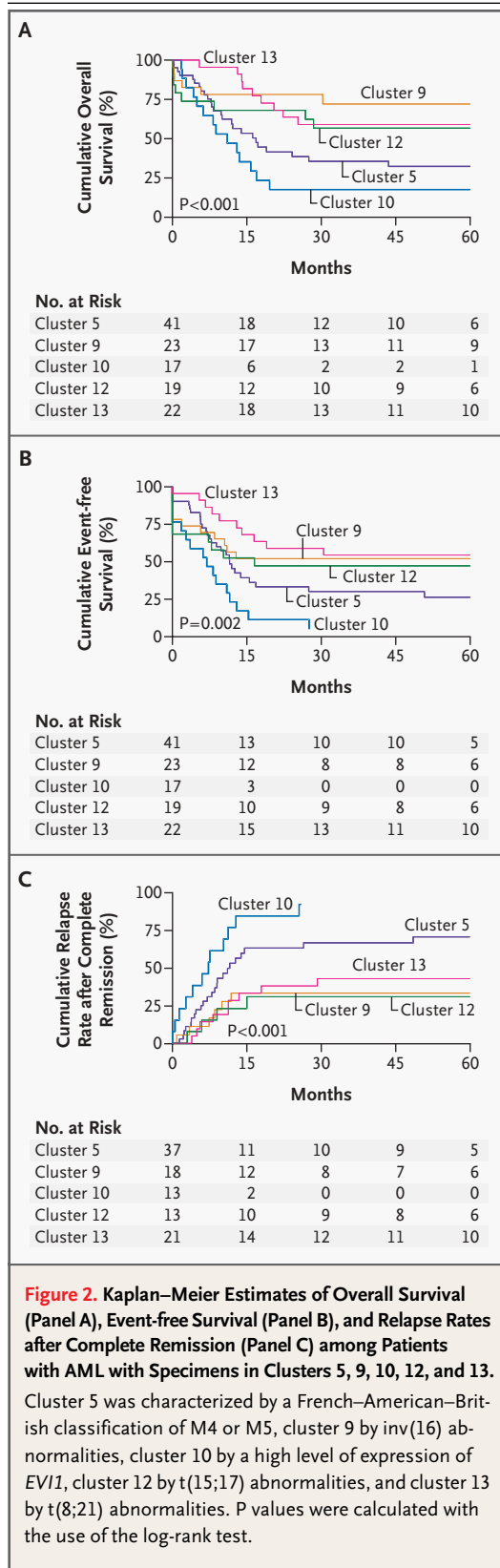
Our study demonstrates that cases of AML with



**Figure 2.** Kaplan–Meier Estimates of Overall Survival (Panel A), Event-free Survival (Panel B), and Relapse Rates after Complete Remission (Panel C) among Patients with AML with Specimens in Clusters 5, 9, 10, 12, and 13.

Cluster 5 was characterized by a French–American–British classification of M4 or M5, cluster 9 by inv(16) abnormalities, cluster 10 by a high level of expression of *EVI1*, cluster 12 by t(15;17) abnormalities, and cluster 13 by t(8;21) abnormalities. P values were calculated with the use of the log-rank test.

known cytogenetic abnormalities and new clusters of AML with characteristic gene-expression signatures can be identified with the use of a single assay. The applicability and performance of genome-wide analysis will advance with the availability of novel whole-genome arrays, improved sequence annotation, and the development of sophisticated protocols and software, allowing the analysis of subtle differences in gene expression and predictions of pathogenic pathways.

### REFERENCES

**1.** Lowenberg B, Downing JR, Burnett A. Acute myeloid leukemia. N Engl J Med 1999; 341:1051-62. [Erratum, N Engl J Med 1999; 341:1484.]

**2.** Slovak ML, Kopecky KJ, Cassileth PA, et al. Karyotypic analysis predicts outcome of preremission and postremission therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern Cooperative Oncology Group study. Blood 2000;96:4075-83.

**3.** Byrd JC, Mrozek K, Dodge RK, et al. Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). Blood 2002;100: 4325-36.

**4.** Grimwade D, Walker H, Oliver F, et al. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. Blood 1998;92:2322-33.

**5.** Grimwade D, Walker H, Harrison G, et al. The predictive value of hierarchical cytogenetic classification in older adults with acute myeloid leukemia (AML): analysis of 1065 patients entered into the United Kingdom Medical Research Council AML11 trial. Blood 2001;98:1312-20.

**6.** Kiyoi H, Naoe T, Nakano Y, et al. Prognostic implication of FLT3 and N-ras gene mutations in acute myeloid leukemia. Blood 1999;93:3074-80.

**7.** Gilliland DG, Griffin JD. The roles of FLT3 in hematopoiesis and leukemia. Blood 2002;100:1532-42.

**8.** Barjesteh van Waalwijk van Doorn-Khosrovani S, Erpelinck C, van Putten WL, et al. High EVI1 expression predicts poor survival in acute myeloid leukemia: a study of 319 de novo AML patients. Blood 2003;101:837-45.

**9.** van Waalwijk van Doorn-Khosrovani SB, Erpelinck C, Meijer J, et al. Biallelic mutations in the CEBPA gene and low CEBPA expression levels as prognostic markers in intermediate-risk AML. Hematol J 2003;4:31-40.

**10.** Preudhomme C, Sagot C, Boissel N, 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;100:2717-23.

**11.** Armstrong SA, Staunton JE, Silverman LB, et al. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat Genet 2002;30:41-7.

**12.** Debernardi S, Lillington DM, Chaplin T, 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;37:149-58.

**13.** Schoch C, Kohlmann A, Schnittger S, et al. Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. Proc Natl Acad Sci U S A 2002;99:10008-13.

**14.** Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531-7.

**15.** Lowenberg B, Boogaerts MA, Daenen SM, et al. Value of different modalities of granulocyte-macrophage colony-stimulating factor applied during or after induction therapy of acute myeloid leukemia. J Clin Oncol 1997;15:3496-506.

**16.** Löwenberg B, van Putten W, Theobald M, et al. Effect of priming with granulocyte colony-stimulating factor on the outcome of chemotherapy for acute myeloid leukemia. N Engl J Med 2003;349:743-52.

**17.** Ossenkoppele GJ, Graveland WJ, Sonneveld P, et al. The value of fludarabine in addition to ARA-C and G-CSF in the treatment of patients with high risk myelodysplastic syndromes and elderly AML. Blood (in press).

**18.** Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 1987;162:156-9.

**19.** Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001;98:5116-21. [Erratum, Proc Natl Acad Sci U S A 2001;98:10515.]

**20.** Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 2002;99:6567-72.

**21.** Valk PJM, Bowen DT, Frew ME, Goodeve AC, Löwenberg B, Reilly JT. Second hit mutations in the RTK/RAS signaling pathway in acute myeloid leukaemia and inv(16). Haematologica 2004;89:106.

**22.** Care RS, Valk PJ, Goodeve AC, et al. Incidence and prognosis of c-KIT and FLT3 mutations in core binding factor (CBF) acute myeloid leukaemias. Br J Haematol 2003; 121:775-7.

**23.** Lo Coco F, De Rossi G, Pasqualetti D, et al. CD7 positive acute myeloid leukaemia: a subtype associated with cell immaturity. Br J Haematol 1989;73:480-5.

**24.** Boeckx N, Willemse MJ, Szczepanski T, et al. Fusion gene transcripts and Ig/TCR gene rearrangements are complementary but infrequent targets for PCR-based detection of minimal residual disease in acute myeloid leukemia. Leukemia 2002;16:368-75.

**25.** Touw I, Donath J, Pouwels K, et al. Acute myeloid leukemias with chromosomal abnormalities involving the 21q22 region identified by their in vitro responsiveness to interleukin-5. Leukemia 1991;5:687-92.

**26.** Kim HP, Kelly J, Leonard WJ. The basis for IL-2-induced IL-2 receptor alpha chain gene regulation: importance of two widely separated IL-2 response elements. Immunity 2001;15:159-72.

**27.** Lilly M, Le T, Holland P, Hendrickson SL. Sustained expression of the pim-1 kinase is specifically induced in myeloid cells by cytokines whose receptors are structurally related. Oncogene 1992;7:727-32.

*Copyright © 2004 Massachusetts Medical Society.*