BAALC and *ERG* expression levels are associated with outcome and distinct gene and microRNA expression profiles in older patients with de novo cytogenetically normal acute myeloid leukemia: a Cancer and Leukemia Group B study

Sebastian Schwind,[1] Guido Marcucci,[1,2] Kati Maharry,[1,3] Michael D. Radmacher,[1,3] Krzysztof Mrózek,[1] Kelsi B. Holland,[1,3] Dean Margeson,[1,3] Heiko Becker,[1] Susan P. Whitman,[1] Yue-Zhong Wu,[1] Klaus H. Metzeler,[1] Bayard L. Powell,[4] Jonathan E. Kolitz,[5] Thomas H. Carter,[6] Joseph O. Moore,[7] Maria R. Baer,[8] Andrew J. Carroll,[9] Michael A. Caligiuri,[1,2] Richard A. Larson,[10] and Clara D. Bloomfield[1]

[1]Division of Hematology and Oncology, Department of Internal Medicine, Comprehensive Cancer Center, The Ohio State University, Columbus, OH; [2]Department of Microbiology, Virology, Immunology, and Medical Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH; [3]The Cancer and Leukemia Group B Statistical Center, Duke University Medical Center, Durham, NC; [4]The Comprehensive Cancer Center of Wake Forest University, Winston-Salem, NC; [5]North Shore University Hospital, Manhasset, NY; [6]University of Iowa, Iowa City, IA; [7]Duke University Medical Center, Durham, NC; [8]University of Maryland, Baltimore, MD; [9]University of Alabama at Birmingham, Birmingham, AL; and [10]University of Chicago, Chicago, IL

BAALC and *ERG* expression levels are prognostic markers in younger (< 60 years) cytogenetically normal acute myeloid leukemia (CN-AML) adults; their prognostic impact in older (≥ 60 years) patients requires further investigation. We evaluated pretreatment expression of *BAALC* and *ERG* in 158 de novo patients treated on cytarabine/daunorubicin-based protocols. The patients were also characterized for other established molecular prognosticators. Low *BAALC* and *ERG* expression levels were associated with better outcome in univariable and multivariable analyses. Expression levels of both *BAALC* and *ERG* were the only factors significantly associated with overall survival upon multivariable analysis. To gain biological insights, we derived gene expression signatures associated with *BAALC* and *ERG* expression in older CN-AML patients. Furthermore, we derived the first microRNA expression signatures associated with the expression of these 2 genes. In low *BAALC* expressers, genes associated with undifferentiated hematopoietic precursors and unfavorable outcome predictors were down-regulated, whereas *HOX* genes and *HOX*-gene–embedded microRNAs were up-regulated. Low *ERG* expressers presented with down-regulation of genes involved in the DNA-methylation machinery, and up-regulation of *miR-148a*, which targets *DNMT3B*. We conclude that in older CN-AML patients, low *BAALC* and *ERG* expression associates with better outcome and distinct gene and microRNA expression signatures that could aid in identifying new targets and novel therapeutic strategies for older patients. (*Blood*. 2010; 116(25):5660-5669)

Introduction

Acute myeloid leukemia (AML) is a cytogenetically and molecularly heterogeneous disease characterized by clonal proliferation of myeloid precursors and maturation arrest. Despite progress in our understanding of the biology of this disease and investigation of therapies targeting distinct clinical, cytogenetic, and/or molecular subsets, outcome remains poor for the majority of patients. This is especially true for patients aged 60 years or more, of whom only approximately 7%-15% achieve long-term survival.[1,2] The reasons for the poor outcome of this older patient population may not only relate to higher frequencies of secondary disease (ie, AML after antecedent hematologic disorders and/or therapy-related disease), high-risk cytogenetics, clinical comorbid conditions, and poor performance status, but also to the presence of specific molecular genetic alterations, including gene mutations and changes in gene expression.[3]

To date, the prognostic significance of molecular genetic alterations has been studied most extensively in younger (< 60 years) patients and found its maximum applicability in cytogenetically normal AML (CN-AML), which constitutes the largest AML subset.[4] In this cytogenetic subset, mutations in the *NPM1*[5] and *CEBPA* genes,[6,7] and lower

expression levels of the *BAALC*[8,9] and *ERG*[10,11] genes have been associated with favorable outcome, whereas internal tandem duplication of the *FLT3* gene (*FLT3*-ITD)[12] and *WT1* mutations[13] have been shown to confer adverse prognosis. Furthermore, because these molecular alterations are not mutually exclusive, combinations of 2 or more of them have been used to refine prognostication of CN-AML patients and are recommended by best practice guidelines for cytogenetic/molecular risk stratification of AML patients.[14]

CN-AML is also the largest cytogenetic subset among patients aged 60 years or older.[1,2] But, despite the relatively large number of patients presenting with this feature, few studies have investigated the prognostic significance of molecular markers in this age group. Recently, we reported a study demonstrating that *NPM1* mutations are associated with a more favorable outcome in older CN-AML patients.[15] However, to our knowledge, studies testing the prognostic impact of *BAALC* and *ERG* gene expression levels in relatively large cohorts of older CN-AML patients have not been reported.

The *BAALC* (brain and acute leukemia, cytoplasmic) gene, located at chromosome band 8q22.3, was cloned in the course of

---

research aimed at the identification of genes associated with a trisomy of chromosome 8 in AML.[16] High levels of *BAALC* expression were, indeed, found in AML patients with trisomy 8, but also in a subset of CN-AML patients.[8,16] The *ETS*-related gene, *ERG*, located at chromosome band 21q22.3, was first shown to be involved in leukemogenesis as a fusion partner with the *FUS* gene in the rare, but recurrent in AML, t(16;21)(p11;q22).[17] Moreover, *ERG* overexpression was demonstrated in AML patients with complex karyotypes with cryptic amplification of chromosome 21,[18] which was first discovered using spectral karyotyping,[19] and it was also found in a fraction of patients with CN-AML.[10] Our group was the first to report that high expression levels of the *BAALC* and *ERG* genes contribute to poor prognosis in younger CN-AML patients,[8,10] and these results have been corroborated by others.[20-24]

Herein, we sought to determine the prognostic impact of the expression of *BAALC* and *ERG* in older de novo CN-AML patients. We show, for the first time, that low levels of expression of these 2 genes are significantly associated with improved outcome in older CN-AML patients, even after adjustment for other prognostic clinical and molecular variables, including *NPM1* mutations. In addition, using genome-wide microarray profiling, we reveal changes in the expression of specific genes and microRNAs in low *BAALC* and *ERG* expressers that may contribute to the less aggressive disease in these patients. These biological features may potentially be exploited and lead to new treatment strategies.

Methods

Patients and treatment

A total of 158 patients aged 60 years or older with de novo CN-AML, who had pretreatment blood available, were analyzed for *BAALC* and *ERG* expression. The patients were treated with intensive cytarabine/daunorubicin-based regimens on Cancer and Leukemia Group B (CALGB) front-line clinical protocols 8525, 8923, 9420, 9720, or 10201 (see supplemental Methods for details, available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article). Patients with antecedent hematologic disorders or therapy-related AML, and those transplanted in first complete remission (CR), were excluded. The Ohio State University Institutional Review Board–approved, written informed consent for participation in these studies was obtained from all patients in accordance with the Declaration of Helsinki.

Cytogenetics and additional mutation markers

Pretreatment cytogenetic analyses of bone marrow (BM) were performed by CALGB-approved institutional cytogenetic laboratories as part of CALGB 8461, a prospective cytogenetic companion study, and the results were reviewed centrally.[25] To be considered cytogenetically normal, at least 20 metaphase cells had to be analyzed and the karyotype found to be normal.[25]

The presence or absence of *FLT3*-ITD, *FLT3* tyrosine kinase domain mutations (*FLT3*-TKD), *MLL* partial tandem duplication (*MLL*-PTD), and mutations in the *NPM1*, *CEBPA*, *WT1*, *IDH1*, and *IDH2* genes was also determined centrally in blood or marrow pretreatment samples, as previously described.[5,7,12,13,26,27]

RNA extraction and real-time RT-PCR to measure *BAALC* and *ERG* expression levels

Preparation of pretreatment blood samples and analysis of *BAALC* and *ERG* expression were performed as previously described.[8-11] Briefly, total RNA was extracted using the Trizol method and complementary DNA was synthesized from total RNA. Quantitative real-time reverse-transcription–polymerase chain reaction (RT-PCR) amplification of *BAALC*, *ERG*, and *ABL1* was performed using standard curves. *BAALC* and *ERG* expression levels are reported as copy numbers normalized to *ABL1* copy numbers.

Gene and microRNA expression profiling

For gene and microRNA expression profiling, total RNA was extracted from pretreatment BM or blood mononuclear cells. The gene and micro RNA expression profiling was performed using the Affymetrix U133 plus 2.0 array (Affymetrix) and The Ohio State University custom microRNA array (OSU_CCC Version 4.0), respectively, as previously reported[15,28] and detailed in supplemental Methods. The microarray data are available on ArrayExpress under accession numbers E-TABM-1071 and E-TABM-1072.

Definition of clinical endpoints and statistical analysis

The main objective of this study was to evaluate the prognostic impact of *BAALC* and *ERG* expression on clinical outcome in older CN-AML patients. The median *BAALC/ABL1* or *ERG/ABL1* copy number values, respectively, were used to define low and high *BAALC* or *ERG* expressers. This cutoff was based on the trends for overall survival (OS) of patients divided into quartiles by expression values, where patients in quartile 1 had better outcome than patients in quartile 2, followed by patients in quartiles 3 and 4 for *BAALC* and *ERG* expression ($P < .001$ and $P < .001$, test for trend[29]).

Definitions of clinical endpoints (eg, CR, disease-free survival [DFS] and OS) and details of statistical analyses, including variable selection for statistical modeling, are provided in the supplemental material. Associations between patients with low or high expression of *BAALC* or *ERG* for baseline demographic, clinical, and molecular features were compared using the Fisher exact and Wilcoxon rank-sum tests for categorical and continuous variables, respectively. Estimated probabilities of DFS and OS were calculated using the Kaplan-Meier method, and the log-rank test evaluated differences between survival distributions. Multivariable logistic regression models were constructed to analyze factors related to the probability of achieving CR using a limited backward selection procedure. Multivariable proportional hazards models were constructed for DFS and OS to evaluate the impact of *BAALC* and *ERG* expression (low/high) by adjusting for other variables using a limited backward selection procedure. For achievement of CR, estimated odds ratios, and for survival endpoints, hazard ratios with their corresponding 95% confidence intervals were obtained for each significant prognostic factor.

For the gene and microRNA expression profiling, summary measures of gene and microRNA expression were computed, normalized, and filtered (see supplemental material). The profiles were derived by comparing gene and microRNA expression between low and high *BAALC* expressers and between low and high *ERG* expressers. Univariable significance levels of .001 for gene and .005 for microRNA expression profiling were used to determine, respectively, the probe sets and microRNA probes that comprised the signatures.

All analyses were performed by the CALGB Statistical Center.

Results

Associations of *BAALC* and *ERG* expression with clinical and molecular characteristics

At diagnosis, low *BAALC* expression was associated with mutated *NPM1* ($P < .001$), wild-type *CEBPA* ($P = .02$), and low *ERG* expression levels ($P < .001$; Table 1).

Low *ERG* expression was associated with lower white blood count (WBC; $P = .005$), lower percentages of blood ($P < .001$) and BM ($P = .001$) blasts, the absence of *FLT3*-ITD ($P < .001$), and low *BAALC* expression ($P < .001$; Table 1).

Prognostic value of *BAALC* and *ERG* expression

Low *BAALC* expressers had a higher CR rate (86% vs. 54%; $P < .001$) and longer DFS ($P = .01$; Figure 1A) and OS ($P < .001$; Figure 1B) than high expressers (Table 2). Similarly, patients with low *ERG* expression had a trend for better CR rates (76% vs. 65%;

Table 1. Clinical and molecular characteristics at diagnosis according to _BAALC_ and _ERG_ expression status in older CN-AML patients

Characteristic	Low _BAALC_ (n = 79)	High _BAALC_ (n = 79)	P	Low _ERG_ (n = 79)	High _ERG_ (n = 79)	P
Age, y			.17			.95
Median	67	69		68	69	
Range	60-79	60-83		60-81	60-83	
Sex, no. (%)			.52			.11
Male	46 (58)	41 (52)		38 (48)	49 (62)	
Female	33 (42)	38 (48)		41 (52)	30 (38)	
Race, no. (%)			.40			1.0
White	73 (94)	68 (89)		70 (92)	71 (91)	
Nonwhite	5 (6)	8 (11)		6 (8)	7 (9)	
Hemoglobin, g/dL			.43			.76
Median	9.6	9.5		9.5	9.5	
Range	5.4-15.0	6.8-14.5		5.4-13.6	6.0-15.0	
Platelet count, ×10⁹/L			.31			.51
Median	55	67		69	55	
Range	5-481	4-850		4-271	11-850	
WBC count, ×10⁹/L			.78			.005
Median	27.8	33.5		20.6	38.0	
Range	1.4-450.0	1.0-173.1		1.1-234.5	1.0-450.0	
Blood blasts, %			.36			< .001
Median	52	59		34	68	
Range	0–97	0–99		0–93	0–99	
BM blasts, %			.80			.001
Median	70	67		59	75	
Range	11-97	7-97		7-97	23-97	
Extramedullary involvement, no. (%)	16 (21)	19 (25)	.70	16 (21)	19 (25)	.56
NPM1, no. (%)			< .001			.33
Mutated	64 (81)	33 (42)		45 (57)	52 (66)	
Wild-type	15 (19)	46 (58)		34 (43)	27 (34)	
FLT3-ITD, no. (%)			.14			< .001
Present	24 (30)	34 (43)		12 (15)	46 (58)	
Absent	55 (70)	45 (57)		67 (85)	33 (42)	
FLT3-TKD, no. (%)			.61			.61
Present	10 (13)	7 (9)		10 (13)	7 (9)	
Absent	69 (87)	72 (91)		69 (87)	72 (91)	
_WT1_, no. (%)			.21			.21
Mutated	3 (4)	8 (10)		3 (4)	8 (10)	
Wild-type	76 (96)	71 (90)		76 (96)	71 (90)	
CEBPA, no. (%)			.02			.65
Mutated	6 (8)	17 (22)		10 (13)	13 (16)	
Wild-type	73 (92)	62 (78)		69 (87)	66 (84)	
IDH1, no. (%)			.09			.40
Mutated	4 (6)	10 (17)		5 (8)	9 (15)	
Wild-type	59 (94)	50 (83)		59 (92)	53 (85)	
_IDH2_*, no. (%)			.30			.10
R140 _IDH2_	18 (29)	7 (12)		17 (28)	8 (13)	
R172 _IDH2_	0 (0)	5 (8)		2 (3)	3 (5)	
Wild-type	45 (71)	48 (80)		42 (69)	51 (82)	
MLL-PTD, no. (%)			.68			.42
Present	2 (4)	4 (7)		4 (7)	2 (3)	
Absent	55 (96)	55 (93)		51 (93)	59 (97)	
ERG expression†, no. (%)			< .001			
Low	51 (65)	28 (35)				
High	28 (35)	51 (65)				
BAALC expression†, no. (%)						< .001
Low				51 (65)	28 (35)	
High				28 (35)	51 (65)	

BM, bone marrow; _FLT3_-ITD, internal tandem duplication of the _FLT3_ gene; _FLT3_-TKD, tyrosine kinase domain mutation of the _FLT3_ gene; _MLL_-PTD, partial tandem duplication of the _MLL_ gene; WBC, white blood count.

*P value is for the comparison of _IDH2_ mutated (R140 or R172) versus _IDH2_ wild-type.

†The median expression value was used as a cut point.

$P = .12$) and longer DFS ($P = .03$; Figure 2A) and OS ($P = .003$; Figure 2B; Table 2). The median follow-up for patients alive was 3.8 years (range: 2.8-11.6 years).

In a multivariable model for CR attainment (see supplemental material for full model building description), _BAALC_ expression remained a significant prognosticator ($P < .001$) after adjustment

Figure 1. Outcome of cytogenetically normal older AML patients according to *BAALC* expression levels. (A) DFS. (B) OS.
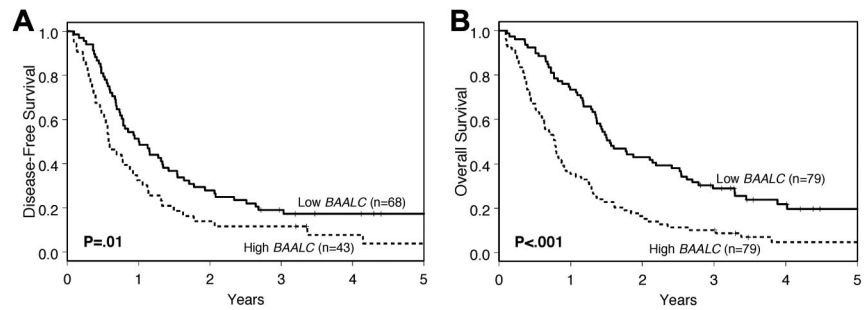
for *NPM1* mutation status (*P* = .04) and WBC (*P* = .02; Table 3). The odds for achieving a CR were more than 4 times higher for the low *BAALC*-expressing group than for high *BAALC* expressers. *BAALC* expression, but not *ERG* expression, was also a significant factor for DFS (*P* = .03), after adjustment for the *FLT3*-ITD (*P* < .001) and age (*P* = .05; Table 3). The risk of relapse or death was reduced by 36% for low *BAALC* expressers. Furthermore, expression levels of both *BAALC* and *ERG* were the only factors significantly associated with OS upon multivariable analysis (*P* < .001 and *P* = .03, respectively; Table 3). The risk of death was reduced by approximately one-half in the low *BAALC*-expressing group and by approximately one-third in the low *ERG*-expressing group.

Because both expression markers were significantly associated with OS, we investigated the combination of the 2 markers. Patients who had lower expression of *BAALC* or *ERG*, or both, had a significantly longer DFS and OS (*P* < .001 for both DFS and OS; Figure 3) than those expressing both markers at high levels. At 3 years, 26% of the patients with low expression of *BAALC* or *ERG*, or both, were alive, compared with only 6% of patients who expressed *BAALC* and *ERG* at high levels. However, there were no significant differences in outcome between patients with low expression of both *BAALC* and *ERG* and those who had low expression of only 1 of these 2 genes.

We recently reported a strong prognostic impact of *NPM1* mutations on CR rates, DFS, and OS in older CN-AML patients.[15] However, in the set of patients investigated here, *NPM1* mutation status was retained only in our model for CR. Therefore, we compared different multivariable models for CR attainment, DFS, and OS using the Akaike information criterion (AIC; supplemental Table 1A-C). We found that the multivariable model for DFS that included *BAALC* expresser status, *FLT3*-ITD mutation status, and age, excluding *NPM1* mutation status, appeared to be better than other evaluated models that comprised *NPM1* mutation status (supplemental Table 1B). For OS, we found that a model including only the *BAALC* and *ERG* expresser status was better than 2 other evaluated models that comprised *NPM1* mutation status (supplemental Table 1C).

We have also recently reported a trend toward a lower CR rate and shorter OS in older CN-AML patients harboring *WT1* mutations.[13] This report suggested that the impact of *WT1* mutation status might be obscured by the generally poor outcome of older CN-AML patients. In the current study, we only had 11 patients who harbored *WT1* mutations, and we did not see a significant impact of *WT1* mutation status on outcome upon univariable analyses. Furthermore, only 3 of the 11 *WT1*-mutated cases were low *BAALC* or low *ERG* expressers. Therefore, we were not able to develop multivariable regression models for outcome that would have evaluated *WT1* mutations in the context of low *BAALC* or *ERG* expression status.

Genome-wide gene expression profiling

To gain insights into the biology of older CN-AML patients, we derived gene expression signatures associated with *BAALC* and *ERG* expresser status.

We observed that *BAALC* expression was associated with the differential expression of 482 probe sets, representing 292 annotated genes. Of these, 283 probe sets, which represented 165 annotated genes, were up-regulated, and 199 probe sets, representing 127 annotated genes, were down-regulated in low *BAALC* expressers (Figure 4A; supplemental Table 2). Probe sets representing *BAALC* were among the most down-regulated probe sets in low *BAALC* expressers, confirming our real-time RT-PCR results. The gene expression signature of *BAALC* derived in older CN-AML patients was very similar to that we reported in patients younger than 60 years.[9] Low *BAALC* expression was associated with up-regulation of genes of the *HOXA* and *HOXB* clusters, as well as the *HOX* cofactor, *MEIS1*, that are important for developmental processes and hematopoietic stem cell function.[30] We also observed a higher expression of *CTSG* and *AZU1*—genes that are expressed in mature neutrophils.[31] As in younger patients and consistent with the more favorable prognosis associated with low *BAALC* expression, we found that low *BAALC* expressers had down-regulation

Table 2. Outcomes according to *BAALC* and *ERG* expression in older CN-AML patients

Outcome	Low *BAALC* (n = 79)	High *BAALC* (n = 79)	*P*	Low *ERG* (n = 79)	High *ERG* (n = 79)	*P*
Complete remission rate, no. (%)	68 (86)	43 (54)	< .001	60 (76)	51 (65)	.12
Disease-free survival			.01			.03
Median, y	1.0	0.6		1.0	0.6	
Disease-free at 3 y, % (95% CI)	19 (11-29)	12 (4-23)		18 (10-29)	14 (6-25)	
Overall survival			< .001			.003
Median, y	1.5	0.8		1.4	0.8	
Alive at 3 y, % (95% CI)	29 (19-39)	10 (5-18)		24 (15-34)	15 (8-24)	

CI indicates confidence interval.
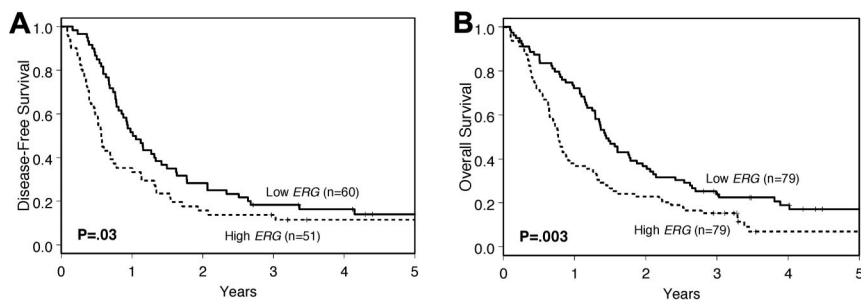
Figure 2. Outcome of cytogenetically normal older AML patients according to *ERG* expression levels. (A) DFS. (B) OS.

of genes previously associated with worse outcome in CN-AML, including the transcription coregulator, *MN1*,[32] the surface marker, *CD200*,[33] and the growth factor, *HGF*[34] (Figure 4A). Furthermore, we observed a down-regulation of genes found expressed in less differentiated precursors, such as *PROM1*, *CD34*, *JUP*, *C5orf23*, *FZD6*, *B4GALT6*, and *APP* (Figure 4A),[18,35,36] and *ABCB1* (*MDR1*), a gene encoding the multidrug resistance protein, whose high expression has been associated with worse outcome in older AML patients.[37]

The gene expression signature associated with *ERG* expression comprised 1554 differentially expressed probe sets, representing 1089 genes, of which 208 probe sets, representing 127 genes, were up-regulated, and 1346 probe sets, representing 962 genes, were down-regulated in low *ERG* expressers (Figure 5A; supplemental Table 3). Probe sets representing *ERG* were among the most down-regulated in low *ERG* expressers, again confirming our real-time RT-PCR results. We found previously identified putative *ERG* targets and binding partners, such as *ICAM2*, *SOCS2*, and *FLI1*,[38,39] down-regulated in low *ERG* expressers. Also down-regulated were *DNMT3A* and *DNMT3B*, genes associated with aberrant DNA hypermethylation in AML[40]; *SET*, which inhibits active DNA demethylation via chromatin modification and inhibits the activation of the tumor suppressor, PP2A phosphatase[41,42]; and the histone methyltransferases, *SMYD2* and *SMYD3*, that have a role in cell proliferation in cancer.[43,44] Among the overexpressed genes associated with low *ERG* levels were *TGFBR3*, which was shown to suppress breast cancer progression,[45] and *TOP1*, which has been associated with enhanced sensitivity to chemotherapy.[46] Furthermore,

when we compared the gene expression signature of *ERG* derived in older CN-AML patients to the signature reported in patients younger than 60 years, we found them to be very similar.[10] We found that genes such as *BCL11A*, *DAPK1*, *GUCY1A3*, and *KLHDC1*,[10] were also down-regulated in low *ERG* expressers in the current study.

Genome-wide microRNA expression profiling

To further investigate the biology related to low *BAALC* and *ERG* expression, we derived microRNA expression signatures associated with *BAALC* and *ERG* expresser status. For the first time, we were able to derive a microRNA expression signature associated with *BAALC* expression (Figure 4B). We found 22 differentially expressed probes, representing 18 microRNAs, with 10 up-regulated and 8 down-regulated in low *BAALC*-expressing patients. Consistent with the higher expression of *HOX* genes in low *BAALC* expressers, we observed up-regulation of microRNAs embedded in the *HOX* cluster, namely, *miR-10a*, *miR-10b*, and *miR-9*. Underexpressed was *miR-126*, which we previously found correlated with *MN1* expression in younger patients,[32] as well as *miR-222*, which is known to target *KIT* and has been linked to hematologic lineage differentiation.[47]

We also report here the first microRNA expression signature associated with *ERG* expression in CN-AML. We observed 11 differentially expressed probes, representing 11 microRNAs, associated with *ERG* expression (Figure 5B), with 5 up-regulated and 6 down-regulated in low *ERG* expressers. Among the up-regulated microRNAs in low *ERG*-expressing patients was *miR-107*, known

Table 3. Multivariable logistic regression analysis for achievement of CR, DFS, and OS in older CN-AML patients

Group	CR*			DFS†			OS‡		
	OR	95% CI	*P*	HR	95% CI	*P*	HR	95% CI	*P*
BAALC, low vs high	4.32	1.83-10.20	< .001	0.64	0.42-0.97	.03	0.49‖	0.34-0.71	<.001
ERG, low vs high							0.69	0.48-0.97	.03
NPM1, mutated vs wild-type	2.34	1.03-5.35	.04						
FLT3-ITD, present vs absent				1.79§	1.11-2.87	< .001			
WBC, each 50-unit increase	0.68	0.49-0.93	.02						
Age, each 10-year increase				0.65	0.42-1.00	.05			

Odds ratios greater than (less than) 1.0 mean higher (lower) CR rate for the higher values of the continuous variables and the first category listed for the categorical variables. Hazard ratios greater than (less than) 1.0 indicate higher (lower) risk for relapse or death (DFS) or death (OS) for the higher values of the continuous variables and the first category listed for the categorical variables.

*Variables considered in the model based on univariable analyses were *BAALC* expression (low vs high; median cut), *ERG* expression (low vs high; median cut), *NPM1* (mutated vs wild-type), WBC (in 50-unit increments), and age (in 10-year increments).

†Variables considered in the model based on univariable analyses were *BAALC* expression (low vs high; median cut), *ERG* expression (low vs high; median cut), *FLT3*-ITD (positive vs negative), *FLT3*-TKD (positive vs negative), and age (in 10-year increments).

‡Variables considered in the model based on univariable analyses were *BAALC* expression (low vs high; median cut), *ERG* expression (low vs high; median cut), *NPM1* (mutated vs wild-type), *FLT3*-ITD (positive vs negative), and platelet count.

‖Does not meet the proportional hazards assumption. Low *BAALC* was associated with better OS until approximately 2 years, and after that, did not seem to impact OS. The hazard ratio for *BAALC*, low vs high, is reported at 1 year.

§Does not meet the proportional hazards assumption. *FLT3*-ITD was associated with worse DFS until approximately 1 year, and after that, its adverse impact for DFS declined. The hazard ratio for *FLT3*-ITD is reported at 9 months.

Figure 3. Outcome of cytogenetically normal older AML patients according to a combination of *BAALC* and *ERG* expression levels. Patients with low expression of *BAALC* or *ERG*, or both, were compared with those having high expression of both *BAALC* and *ERG*. (A) DFS. (B) OS.
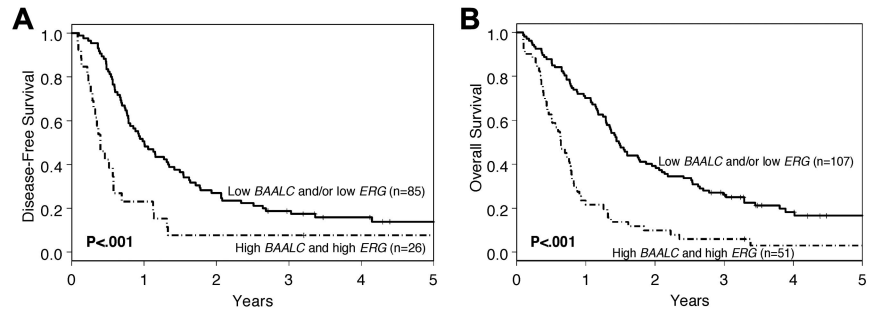
Discussion

Although some studies have shown an association of mutation markers and outcome in older CN-AML patients,[13,15,26] relatively little is known with regard to the prognostic value of gene expression markers in older patients with CN-AML. The aim of this study was to elucidate the prognostic impact of *BAALC* and *ERG* expression, in the context of other well-established molecular markers, in older de novo CN-AML patients similarly treated with intensive chemotherapy.

First, we demonstrate that low expression levels of both *BAALC* and *ERG* are associated with better outcome in older patients with CN-AML. As in younger CN-AML patients,[9] low *BAALC* expression was associated with a higher CR rate and longer DFS and OS in older CN-AML patients investigated here. Low *ERG* expression was also associated with longer DFS and OS and a trend toward a higher CR rate in older patients, which is similar to our findings in younger CN-AML patients, where low *ERG* expressers had higher CR rates and longer event-free survival (EFS) than high *ERG* expressers.[11] In multivariable analyses, *BAALC*, but not *ERG*, expression levels were significant factors for CR attainment and DFS. Expression levels of both *BAALC* and *ERG* were the only factors associated with OS in our cohort of older CN-AML patients.

To our knowledge, only 3 studies that analyzed the prognostic significance of *BAALC* expression and one study analyzing *ERG* expression in CN-AML included patients age 60 years or older.[20,21,24] Two of these studies that comprised, respectively, 67 and 98 adults with CN-AML, including an unspecified number of patients age 60 years or older, showed that low *BAALC* expression was associated with better outcome, but did not evaluate older patients separately. The third study[20] also included older patients and showed that low *BAALC* expression was associated with higher probability of CR achievement and longer EFS and OS. In their subgroup of 101 patients age 60 years or older, low *BAALC* expression showed a trend toward a higher CR rate (66% vs. 50%; *P* = .1) and a longer EFS (*P* = .03; Klaus H. Metzeler, Ludwig-Maximilians-Universität München, e-mail, April 26, 2010). On the other hand, low *ERG* expression seemed to be a stronger factor for outcome in the study of Metzeler et al,[20] where it was significantly associated with a higher CR rate and longer EFS and OS in all patients analyzed. Subset analyses revealed a significant impact of low *ERG* expression on OS in both the younger patients and those aged 60 years or older, but neither of these age groups was evaluated separately in multivariable models considering other molecular markers.

Second, because we recently reported a strong prognostic impact of *NPM1* mutations on CR rates, DFS, and OS in older CN-AML patients,[15] it was of particular interest to us to investigate the relationship of *NPM1* mutation status with *BAALC* and *ERG* expresser status. In multivariable models, *NPM1* mutation status remained a strong, favorable prognostic marker for achievement of CR, but did not remain in the models for DFS and OS. Interestingly, in our previously reported gene expression signature associated with *NPM1* mutations in older patients with CN-AML, expression of both the *BAALC* and *ERG* genes was low.[15] We believe that measuring *BAALC* and *ERG* expression at diagnosis seems to provide more prognostic information with regard to survival than *NPM1* mutational status alone. Indeed, patients with high *BAALC* expression had worse outcome irrespective of their *NPM1* mutational status. This is supported by the comparison of different multivariable models in our dataset. Thus, once reliable methods to measure the pretreatment expression levels of *BAALC* and *ERG* in individual patients are available, these markers should be included in diagnostic testing for a more accurate risk stratification of older CN-AML patients.

Finally, to gain insights into the biology of the disease, we derived gene and, for the first time, microRNA expression signatures associated with *BAALC* and *ERG* expression. Many genes in the *BAALC* signature for older patients reported here have also been found in the gene expression signature in younger patients.[9] This indicates that low *BAALC* expression defines a distinct subset of biologically similar patients, regardless of age. The characteristic features of this signature included overexpression of the *HOXA* and *HOXB* gene clusters, underexpression of known adverse outcome predictors, and down-regulation of genes associated with an undifferentiated state. Furthermore, in the first microRNA expression signature associated with *BAALC* expression, we found down-regulation of microRNAs embedded in the *HOX* cluster and up-regulation of *miR-222*, which is known to target the tyrosine kinase, *KIT*. The lower expression of *ABCB1* (*MDR1*) observed in older CN-AML patients with low *BAALC* expression may account, at least in part, for the observed higher CR rate and better survival.

Interestingly, patient subgroups characterized by mutated *NPM1* or low *BAALC* expression shared some common features in their gene and microRNA expression profiles. The second most down-regulated gene in *NPM1* mutated patients was *BAALC*.[15] Both the gene and microRNA expression profiles associated with mutated *NPM1*[15] and those associated with low *BAALC* expression show down-regulation of *HOX* genes and cofactors such as *MEIS1* and

A *BAALC* expression
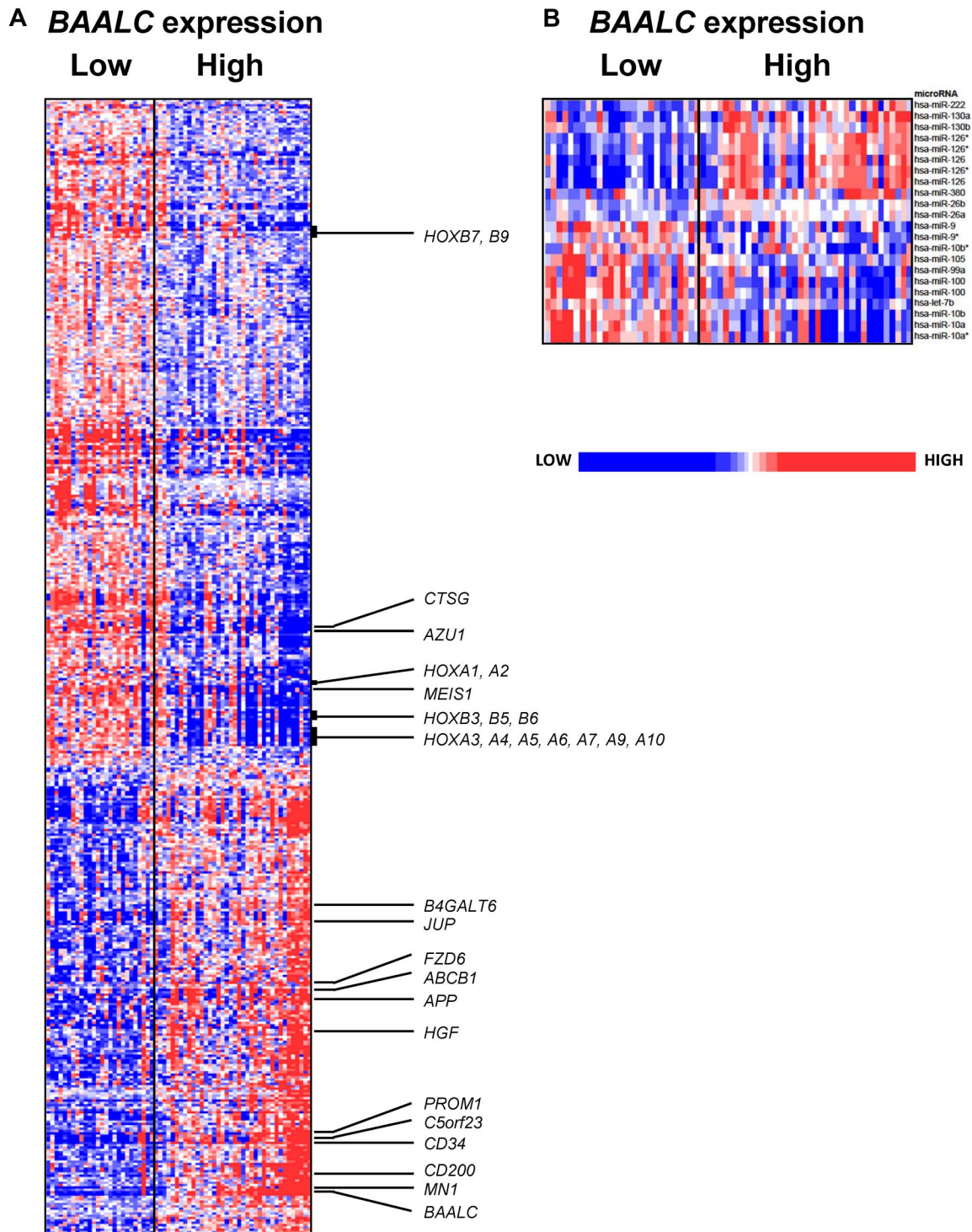
B *BAALC* expression

Figure 4. Heat maps of the *BAALC* signatures. Derived gene expression (A) and microRNA expression signatures (B) associated with *BAALC* expression in the group of CN-AML patients ≥ 60 years. Patients are ordered from left to right by increasing *BAALC* expression. Expression values of the gene probe sets (microRNA probes) are represented by color, with blue indicating expression less than and red indicating expression greater than the median value for the given gene probe set (microRNA probe). For the gene expression heat map, up- and down-regulated genes mentioned in the text are indicated.

up-regulation of *CTSG*, as well as overexpression of *miR-9*, *miR-10a*, *miR-10b*, and *miR-100*, and underexpression of *miR-126* and *miR-130a*. Thus, both low *BAALC* expression and mutated *NPM1* appear to share some biological features. To our knowledge, no molecular link between the 2 markers has been hitherto found. Further studies of possible biological links between those 2 markers are needed.

As with *BAALC*, the gene expression signature associated with *ERG* was similar to the signature we previously derived in younger patients.[10] Not surprisingly, we found lower expression of previously identified putative *ERG* targets and binding partners associated with low *ERG* expression. Up-regulation of topoisomerase 1 might, in part, explain outcome differences caused by enhanced sensitivity to chemotherapy of low *ERG* expressers. Furthermore,
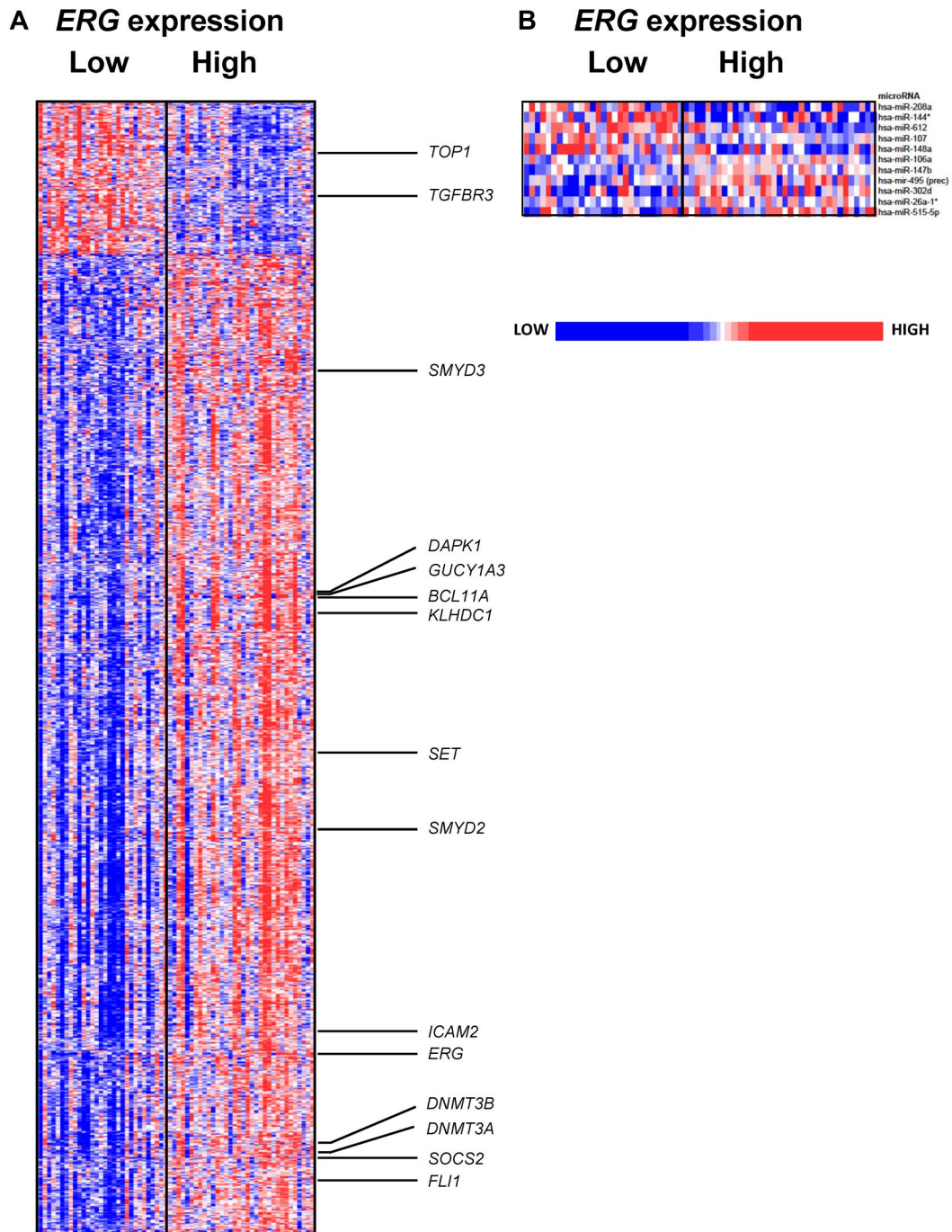
Figure 5. Heat maps of the *ERG* signatures. Derived gene expression (A) and microRNA expression signatures (B) associated with *ERG* expression in the group of CN-AML patients ≥ 60 years. Patients are ordered from left to right by increasing *ERG* expression. Expression values of the gene probe sets (microRNA probes) are represented by color, with blue indicating expression less than and red indicating expression greater than the median value for the given gene probe set (microRNA probe). For the gene expression heat map, up- and down-regulated genes mentioned in the text are indicated.

low *ERG* expressers also had down-regulation of DNA methyltransferases, histone methyltransferases, and an up-regulation of *miR-148a*, which targets *DNMT3B*.[49] We also observed a down-regulation of *SET*, which inhibits active DNA demethylation. This may indicate a higher activity of the DNA methylation machinery in high *ERG* expressers, which could contribute to their worse outcome. This should be further investigated and, if confirmed, might potentially lead to new treatment strategies in this subset of patients.

The similarity of the respective *BAALC* and *ERG* gene expression signatures between younger and older CN-AML patients, and the fact that these expression markers affect outcomes of both younger and older CN-AML patients, suggest that 60 years might not be an optimal age cutoff to separate younger and older patients' eligibility for treatment trials. Older patients with favorable molecular risk factors, such as low *BAALC* and/or *ERG* expression, if treated more intensively, might have outcomes comparable with those of younger patients with corresponding molecular features.

In conclusion, we show that low *BAALC* expression is an important factor for CR achievement and longer DFS; low *BAALC* and *ERG* expresser status are both significant factors associated with improved survival in older CN-AML patients. However, before the pretreatment expression of the 2 genes can be used for risk-stratification of older CN-AML patients, future studies should establish a standardized method of *BAALC* and *ERG* expression quantification and define absolute cutoff points. Once this obstacle is overcome, these important markers should be included in future risk-classification schemas. Furthermore, the derived gene and microRNA expression signatures shed light on the biology of this complex disease and identified new targets that might help in developing new therapeutic strategies for older CN-AML patients.

Authorship

Contribution: S.S., G.M., K. Maharry, M.D.R., K. Mrózek, and C.D.B. contributed to the design and analysis of this study and the writing of the manuscript, and all authors agreed on the final version; S.S., H.B., S.P.W., Y.-Z.W., and K.H.M. carried out laboratory-based research; K. Maharry, M.D.R., K.B.H., and D.M. performed statistical analyses; and G.M., B.L.P., J.E.K., T.H.C., J.O.M., M.R.B., M.A.C., R.A.L., and C.D.B. were involved directly or indirectly in the care of patients and/or sample procurement.

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

For a complete list of participating Cancer and Leukemia Group B institutions, principal investigators, and cytogeneticists, please see the supplemental Appendix.

Correspondence: Clara D. Bloomfield, The Ohio State University, James Cancer Hospital and Solove Research Institute, Rm 1216, 300 West 10th Ave, Columbus, OH 43210; e-mail: clara.bloomfield@osumc.edu; or Guido Marcucci, MD, The Ohio State University, Comprehensive Cancer Center, Biomedical Research Tower, 460 W 12th Ave, Columbus, OH 43210; e-mail: guido.marcucci@osumc.edu.

References

1. Farag SS, Archer KJ, Mrózek K, et al. Pretreatment cytogenetics add to other prognostic factors predicting complete remission and long-term outcome in patients 60 years of age or older with acute myeloid leukemia: results from Cancer and Leukemia Group B 8461. *Blood.* 2006;108(1):63-73.

2. Fröhling S, Schlenk RF, Kayser S, et al. Cytogenetics and age are major determinants of outcome in intensively treated acute myeloid leukemia patients older than 60 years: results from AMLSG trial AML HD98-B. *Blood.* 2006;108(10):3280-3288.

3. Estey E. Acute myeloid leukemia and myelodysplastic syndromes in older patients. *J Clin Oncol.* 2007;25(14):1908-1915.

4. Mrózek K, Marcucci G, Paschka P, Whitman SP, Bloomfield CD. Clinical relevance of mutations and gene-expression changes in adult acute myeloid leukemia with normal cytogenetics: are we ready for a prognostically prioritized molecular classification? *Blood.* 2007;109(2):431-448.

5. Döhner K, Schlenk RF, Habdank M, et al. Mutant nucleophosmin (NPM1) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics: interaction with other gene mutations. *Blood.* 2005;106(12):3740-3746.

6. Fröhling S, Schlenk RF, Stolze I, et al. *CEBPA* mutations in younger adults with acute myeloid leukemia and normal cytogenetics: prognostic relevance and analysis of cooperating mutations. *J Clin Oncol.* 2004;22(4):624-633.

7. Marcucci G, Maharry K, Radmacher MD, et al. Prognostic significance of, and gene and microRNA expression signatures associated with, CEBPA mutations in cytogenetically normal acute myeloid leukemia with high-risk molecular features: a Cancer and Leukemia Group B study. *J Clin Oncol.* 2008;26(31):5078-5087.

8. Baldus CD, Tanner SM, Ruppert AS, et al. BAALC expression predicts clinical outcome of de novo acute myeloid leukemia patients with normal cytogenetics: a Cancer and Leukemia Group B study. *Blood.* 2003;102(5):1613-1618.

9. Langer C, Radmacher MD, Ruppert AS, et al. High BAALC expression associates with other molecular prognostic markers, poor outcome, and a distinct gene expression signature in cytogenetically normal patients younger than 60 years with acute myeloid leukemia: a Cancer and Leukemia Group B (CALGB) study. *Blood.* 2008; 111(11):5371-5379.

10. Marcucci G, Baldus CD, Ruppert AS, et al. Overexpression of the ETS-related gene, ERG, predicts a worse outcome in acute myeloid leukemia with normal karyotype: a Cancer and Leukemia Group B study. *J Clin Oncol.* 2005;23(36):9234-9242.

11. Marcucci G, Maharry K, Whitman SP, et al. High expression levels of the ETS-related gene, ERG, predict adverse outcome and improve molecular risk-based classification of cytogenetically normal acute myeloid leukemia: a Cancer and Leukemia Group B study. *J Clin Oncol.* 2007;25(22):3337-3343.

12. Thiede C, Steudel C, Mohr B, et al. Analysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups with poor prognosis. *Blood.* 2002;99(12):4326-4335.

13. Becker H, Marcucci G, Maharry K, et al. Mutations of the Wilms tumor 1 gene (WT1) in older patients with primary cytogenetically normal acute myeloid leukemia: a Cancer and Leukemia Group B study. *Blood.* 2010;116(5):788-792.

14. Döhner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood.* 2010;115(3):453-474.

15. Becker H, Marcucci G, Maharry K, et al. Favorable prognostic impact of NPM1 mutations in older patients with cytogenetically normal de novo acute myeloid leukemia and associated gene- and microRNA-expression signatures: a Cancer and Leukemia Group B study. *J Clin Oncol.* 2010;28(4):596-604.

16. Tanner SM, Austin JL, Leone G, et al. BAALC, the human member of a novel mammalian neuroectoderm gene lineage, is implicated in hematopoiesis and acute leukemia. *Proc Natl Acad Sci U S A.* 2001;98(24):13901-13906.

17. Ichikawa H, Shimizu K, Hayashi Y, Ohki M. An RNA-binding protein gene, TLS/FUS, is fused to ERG in human myeloid leukemia with t(16;21) chromosomal translocation. *Cancer Res.* 1994; 54(11):2865-2868.

18. Baldus CD, Liyanarachchi S, Mrózek K, et al. Acute myeloid leukemia with complex karyotypes and abnormal chromosome 21: amplification discloses overexpression of APP, ETS2, and ERG genes. *Proc Natl Acad Sci U S A.* 2004;101(11):3915-3920.

19. Mrózek K, Heinonen K, Theil KS, Bloomfield CD. Spectral karyotyping in patients with acute myeloid leukemia and a complex karyotype shows hidden aberrations, including recurrent overrepresentation of 21q, 11q, and 22q. *Genes Chromosomes Cancer.* 2002;34(2):137-153.

20. Metzeler KH, Dufour A, Benthaus T, et al. ERG expression is an independent prognostic factor and allows refined risk stratification in cytogenetically normal acute myeloid leukemia: a comprehensive analysis of ERG, MN1, and BAALC transcript levels using oligonucleotide microarrays. *J Clin Oncol.* 2009;27(30):5031-5038.

21. Bienz M, Ludwig M, Leibundgut EO, et al. Risk assessment in patients with acute myeloid leukemia and a normal karyotype. *Clin Cancer Res.* 2005;11(4):1416-1424.

22. Baldus CD, Thiede C, Soucek S, Bloomfield CD, Thiel E, Ehninger G. BAALC expression and FLT3 internal tandem duplication mutations in acute myeloid leukemia patients with normal cytogenetics: prognostic implications. *J Clin Oncol.* 2006;24(5):790-797.

23. Eid MA, Attia M, Abdou S, et al. BAALC and ERG

expression in acute myeloid leukemia with normal karyotype: impact on prognosis. *Int J Lab Hematol.* 2010;32(2):197-205.

24. Santamaría C, Chillón MC, García-Sanz R, et al. BAALC is an important predictor of refractoriness to chemotherapy and poor survival in intermediate-risk acute myeloid leukemia (AML). *Ann Hematol.* 2010;89(5):453-458.

25. Mrózek K, Carroll AJ, Maharry K, et al. Central review of cytogenetics is necessary for cooperative group correlative and clinical studies of adult acute leukemia: the Cancer and Leukemia Group B experience. *Int J Oncol.* 2008;33(2):239-244.

26. Marcucci G, Maharry K, Wu Y-Z, et al. IDH1 and *IDH2* gene mutations identify novel molecular subsets within de novo cytogenetically normal acute myeloid leukemia: a Cancer and Leukemia Group B study. *J Clin Oncol.* 2010;28(14):2348-2355.

27. Caligiuri MA, Strout MP, Schichman SA, et al. Partial tandem duplication of *ALL1* as a recurrent molecular defect in acute myeloid leukemia with trisomy 11. *Cancer Res.* 1996;56(6):1418-1425.

28. Marcucci G, Radmacher MD, Maharry K, et al. MicroRNA expression in cytogenetically normal acute myeloid leukemia. *N Engl J Med.* 2008; 358(18):1919-1928.

29. Klein JP, Moeschberger ML. Survival Analysis: Techniques for Censored and Truncated Data. New York, NY: Springer-Verlag; 1997.

30. Argiropoulos B, Humphries RK. HOX genes in hematopoiesis and leukemogenesis. *Oncogene.* 2007;26(47):6766-6776.

31. Gullberg U, Andersson E, Garwicz D, Lindmark A, Olsson I. Biosynthesis, processing, and sorting of neutrophil proteins: insight into neutrophil granule development. *Eur J Haematol.* 1997;58(3):137-153.

32. Langer C, Marcucci G, Holland KB, et al. Prognostic importance of MN1 transcript levels, and biologic insights from MN1-associated gene and microRNA expression signatures in cytogenetically normal acute myeloid leukemia: a Cancer

and Leukemia Group B study. *J Clin Oncol.* 2009; 27(19):3198-3204.

33. Tonks A, Hills R, White P, et al. CD200 as a prognostic factor in acute myeloid leukaemia [letter]. *Leukemia.* 2007;21(3):566-568.

34. Kim JG, Sohn SK, Kim DH, et al. Clinical implications of angiogenic factors in patients with acute or chronic leukemia: hepatocyte growth factor levels have prognostic impact, especially in patients with acute myeloid leukemia. *Leuk Lymphoma.* 2005;46(6):885-891.

35. Toren A, Bielorai B, Jacob-Hirsch J, et al. CD133-positive hematopoietic stem cell "stemness" genes contain many genes mutated or abnormally expressed in leukemia. *Stem Cells.* 2005; 23(8):1142-1153.

36. Steidl U, Kronenwett R, Rohr U-P, et al. Gene expression profiling identifies significant differences between the molecular phenotypes of bone marrow–derived and circulating human CD34+ hematopoietic stem cells. *Blood.* 2002;99(6):2037-2044.

37. Leith CP, Kopecky KJ, Godwin J, et al. Acute myeloid leukemia in the elderly: assessment of multidrug resistance (MDR1) and cytogenetics distinguishes biologic subgroups with remarkably distinct responses to standard chemotherapy. A Southwest Oncology Group study. *Blood.* 1997; 89(9):3323-3329.

38. Forsberg EC, Prohaska SS, Katzman S, Heffner GC, Stuart JM, Weissman IL. Differential expression of novel potential regulators in hematopoietic stem cells. *PLoS Genet.* 2005;1(3):e28.

39. Kruse EA, Loughran SJ, Baldwin TM, et al. Dual requirement for the ETS transcription factors Fli-1 and Erg in hematopoietic stem cells and the megakaryocyte lineage. *Proc Natl Acad Sci U S A.* 2009;106(33):13814-13819.

40. Mizuno S, Chijiwa T, Okamura T, et al. Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. *Blood.* 2001;97(5):1172-1179.

41. Cervoni N, Detich N, Seo S, Chakravarti D, Szyf M. The oncoprotein Set/TAF-1beta, an inhibitor of histone acetyltransferase, inhibits active demethylation of DNA, integrating DNA methylation and transcriptional silencing. *J Biol Chem.* 2002;277(28):25026-25031.

42. Neviani P, Santhanam R, Trotta R, et al. The tumor suppressor, PP2A, is functionally inactivated in blast crisis CML through the inhibitory activity of the BCR/ABL-regulated SET protein. *Cancer Cell.* 2005;8(5):355-368.

43. Komatsu S, Imoto I, Tsuda H, et al. Overexpression of SMYD2 relates to tumor cell proliferation and malignant outcome of esophageal squamous cell carcinoma. *Carcinogenesis.* 2009;30(7):1139-1146.

44. Hamamoto R, Silva FP, Tsuge M, et al. Enhanced SMYD3 expression is essential for the growth of breast cancer cells. *Cancer Sci.* 2006;97(2):113-118.

45. Dong M, How T, Kirkbride KC, et al. The type III TGF-beta receptor suppresses breast cancer progression. *J Clin Invest.* 2007;117(1):206-217.

46. Burgess DJ, Doles J, Zender L, et al. Topoisomerase levels determine chemotherapy response in vitro and in vivo. *Proc Natl Acad Sci U S A.* 2008;105(26):9053-9058.

47. Felli N, Fontana L, Pelosi E, et al. MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation. *Proc Natl Acad Sci U S A.* 2005; 102(50):18081-18086.

48. Garzon R, Pichiorri F, Palumbo T, et al. MicroRNA gene expression during retinoic acid–induced differentiation of human acute promyelocytic leukemia. *Oncogene.* 2007;26(28):4148-4157.

49. Duursma AM, Kedde M, Schrier M, le Sage C, Agami R. miR-148 targets human DNMT3b protein coding region. *RNA.* 2008;14(5):872-877.

50. Barroso-del Jesus A, Lucena-Aguilar G, Menendez P. The miR-302-367 cluster as a potential stemness regulator in ESCs. *Cell Cycle.* 2009;8(3):394-398.