VOLUME 23 · NUMBER 36 · DECEMBER 20 2005

JOURNAL OF CLINICAL ONCOLOGY ORIGINAL REPORT

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

Guido Marcucci, Claudia D. Baldus, Amy S. Ruppert, Michael D. Radmacher, Krzysztof Mrózek, Susan P. Whitman, Jonathan E. Kolitz, Colin G. Edwards, James W. Vardiman, Bayard L. Powell, Maria R. Baer, Joseph O. Moore, Danilo Perrotti, Michael A. Caligiuri, Andrew J. Carroll, Richard A. Larson, Albert de la Chapelle, and Clara D. Bloomfield

From the Division of Hematology and Oncology, Department of Internal Medicine, and Division of Human Cancer Genetics, Department of Microbiology, Virology, Immunology, and Medical Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH; The Cancer and Leukemia Group B Statistical Center; Duke University Medical Center, Durham; Wake Forest University School of Medicine, Winston-Salem, NC; North Shore University Hospital, Manhasset; Roswell Park Cancer Institute, Buffalo, NY; University of Chicago, Chicago, IL; and University of Alabama at Birmingham, Birmingham, AL. C.D. Baldus is currently at Charité, Campus Benjamin Franklin, Medizinische Klinik III, Berlin, Germany.

Submitted July 28, 2005; accepted September 19, 2005.

Supported by National Cancer Institute (Bethesda, MD) Grants No. CA101140, CA77658, CA102031, CA31946, CA09512, CA16058, and CA90469 and The Coleman Leukemia Research Foundation.

Both G.M. and C.D. Baldus contributed equally to this work.

Authors' disclosures of potential conflicts of interest are found at the end of this article.

Address reprint requests to Guido Marcucci, MD, The Ohio State University, The Comprehensive Cancer Center, A433B Starling-Loving Hall, 320 W 10th Ave, Columbus OH 43210; e-mail: marcucci-1@medctr.osu.edu.

0732-183X/05/2336-9234/$20.00

DOI: 10.1200/JCO.2005.03.6137

ABSTRACT

Purpose
To test the prognostic significance of ETS-related gene (*ERG*) expression in cytogenetically normal primary acute myeloid leukemia (AML).

Patients and Methods
Pretreatment blood samples from 84 cytogenetically normal AML patients aged less than 60 years, who were characterized for *BAALC* expression, *FLT3* internal tandem duplication (ITD), and *MLL* partial tandem duplication (PTD) and uniformly treated on Cancer and Leukemia Group B 9621 protocol, were analyzed for *ERG* expression by real-time reverse transcriptase polymerase chain reaction. Patients were divided into quartiles according to *ERG* levels and were compared for clinical outcome. High-density oligonucleotide arrays were used to identify genes differentially expressed between high and low *ERG* expressers.

Results
With a median follow-up of 5.7 years, patients with the upper 25% of *ERG* expression values had a worse cumulative incidence of relapse (CIR; $P < .001$) and overall survival (OS; $P = .011$) than the remaining patients. In a multivariable analysis, high *ERG* expression ($P < .001$) and the presence of *MLL* PTD ($P = .027$) predicted worse CIR. With regard to OS, an interaction was observed between expression of *ERG* and *BAALC* ($P = .013$), with *ERG* overexpression predicting shorter survival only in low *BAALC* expressers ($P = .002$). *ERG* overexpression was an independent prognostic factor even when the unfavorable group of *FLT3* ITD patients lacking an *FLT3* wild-type allele was included. High *ERG* expression was associated with upregulation of 112 expressed-sequenced tags and named genes, many of which are involved in cell proliferation, differentiation, and apoptosis.

Conclusion
ERG overexpression in AML patients with normal cytogenetics predicts an adverse clinical outcome and seems to be associated with a specific molecular signature.

J Clin Oncol 23:9234-9242. © 2005 by American Society of Clinical Oncology

INTRODUCTION

Cytogenetic abnormalities detected at diagnosis have long been recognized as predictors for clinical outcome in acute myeloid leukemia (AML).[1] However, the largest cytogenetic subset of adult AML, approximately 45%, consists of patients with a normal karyotype.[1] In large studies of the clinical significance of cytogenetics in AML, these patients have been categorized in an intermediate-risk group, with 5-year survival rates varying

between 24% and 42%.[2-5] The difference in clinical outcome likely reflects molecular heterogeneity of this cytogenetic subset whose prognosis is influenced by submicroscopic gene mutations or overexpression.[6] The adverse prognostic impact of the partial tandem duplication (PTD) of *MLL*, internal tandem duplication (ITD) of *FLT3*, and overexpression of *BAALC* in karyotypically normal AML is now established,[7-17] as is the favorable prognostic significance of *CEBPA* gene mutations.[17,18] However, it is likely that, in addition to the aforementioned genetic abnormalities, others will be found to impact on the clinical outcome of cytogenetically normal AML. Given that intensive treatments such as allogeneic stem-cell transplantation (SCT), although potentially curative in patients with poor prognosis AML, are associated with high treatment-related mortality, novel molecular markers will likely be valuable to stratify karyotypically normal AML patients to risk-adapted therapies. Furthermore, because these markers are mutated or overexpressed genes encoding proteins with potentially pivotal roles in leukemogenesis, they could also serve as molecular targets for novel therapeutic approaches.[19]

We have recently shown that *ETS-related gene* (*ERG*), which is located at chromosome band 21q22, is frequently overexpressed in AML patients with complex karyotypes and cryptic amplification of chromosome 21.[20] *ERG* and other members of the *ETS* family are downstream effectors of mitogenic signaling transduction pathways and are involved in key steps regulating cell proliferation, differentiation, and apoptosis.[21-23] Although *ERG* rearrangements have been found in AML[24] and Ewing sarcoma[25] and its overexpression has been observed in prostate cancer,[26] little is known regarding how *ERG* contributes to malignant transformation.[27] In our previous report,[20] high *ERG* expression was not always associated with genomic amplification, thereby leaving *ERG* overexpression mechanistically unexplained. Nevertheless, the recurrent presence of *ERG* overexpression in AML with complex karyotypes, a prognostically unfavorable subgroup, suggests that *ERG* overexpression might not only be a nonrandom event in myeloid leukemogenesis, but also might contribute to an aggressive malignant phenotype.

To test this hypothesis, we analyzed karyotypically normal AML patients who were uniformly treated on the Cancer and Leukemia Group B (CALGB) 9621 protocol.[28] We show that the level of *ERG* expression varies among patients and that *ERG* overexpression constitutes an adverse prognostic factor in cytogenetically normal AML.

PATIENTS AND METHODS

Patients

ERG expression was analyzed in 84 adults aged less than 60 years with primary, untreated AML and normal cytogenetics confirmed by central morphologic and karyotype reviews. Eligible patients were enrolled onto the treatment trial CALGB 9621,[28] the prospective cytogenetic study CALGB 8461,[29] and the molecular study of *BAALC* expression (CALGB 9665).[15] Written institutional review board–approved informed consent was obtained from all patients.

Pretreatment cytogenetic analyses of bone marrow (BM) were performed as previously described.[4,30]At least 20 metaphases were analyzed, and the karyotype was normal in each case. Pretreatment *MLL* PTD and *FLT3* ITD status and *BAALC* levels were also determined centrally for each patient, as described previously.[7,10,15]

Treatment

Patients received induction chemotherapy with cytarabine, daunorubicin, and etoposide with valspodar (PSC-833) or without valspodar.[28] On achievement of complete remission (CR), patients received high-dose etoposide and cytarabine for stem-cell mobilization followed by myeloablative treatment with busulfan and etoposide supported by autologous peripheral-blood SCT (APBSCT). Patients unable to receive APBSCT received two additional cycles of high-dose cytarabine. After consolidation, patients received maintenance with interleukin-2.

RNA Extraction and Real-Time Reverse Transcriptase Polymerase Chain Reaction

ERG expression was measured in blood for consistency with our previous analysis of the same patients characterized for *BAALC* expression.[15] Mononuclear cells from pretreatment blood were enriched by Ficoll-Hypaque gradient and cryopreserved in liquid nitrogen. Samples were chosen based on the availability of procured material with adequate RNA quality in the CALGB Leukemia Tissue Bank. Total RNA was extracted using Trizol reagent (Invitrogen, Carlsbad, CA). cDNA synthesis and the real-time amplification reactions were performed as previously reported.[15]

The comparative cycle threshold (C_T) method was used to determine the relative expression levels of *ERG* to *GPI*, the internal control, in patients previously included in the *BAALC* expression study.[15] Of the original 86 patients, 84 were analyzed for *ERG/GPI* levels calculated using the mean of ΔC_T from two replicates and expressed as $2^{\mu(\Delta CT)}$. The results of real-time reverse transcriptase polymerase chain reaction (RT-PCR) were correlated with clinical end points.

To assess the impact of *ERG* levels on clinical outcome, we also adopted an alternative approach. Absolute *ERG* copy numbers were measured and normalized to the copy numbers of *ABL*, which was a different internal control validated by multicenter studies,[31,32] using standard curves constructed as reported previously.[33] We analyzed 73 samples comprising 48 samples from patients included in the set evaluated for *ERG/GPI* levels and for whom leftover material was available and 25 samples from additional, karyotypically normal AML patients enrolled onto CALGB 9621. Importantly, eight of the latter patients had the unfavorable $FLT3^{ITD/-}$ genotype,[10] the presence of which was a reason for exclusion from the previous *BAALC* study.[15]

Positive and negative controls were included in all assays. The reproducibility of the real-time RT-PCR assays was similar to what we reported previously.[15]

Gene Expression Profiling

Suitable RNA samples from 61 patients characterized for *ERG/ABL* expression were analyzed using Affymetrix U133 plus 2.0 GeneChips (Affymetrix, Santa Clara, CA). From 8 μg of total RNA, double-stranded cDNA was prepared (Invitrogen) with the

Table 1. Presenting Characteristics of Patients Divided Into Quartile Groups According to *ERG* Expression

Characteristic	Overall (N = 84)		*ERG* Expression Quartiles 1-3 (n = 63)		*ERG* Expression Quartile 4 (n = 21)		*P**
	No. of Patients	%	No. of Patients	%	No. of Patients	%	
Age, years							
Median	47		47		48		.616
Range	18-59		18-59		26-59		
Sex, male	45	54	35	56	10	48	.617
Race (n = 1 unknown)							.99
White	73	88	54	87	19	90	
Nonwhite	10	12	8	13	2	10	
FAB (n = 2 unknown)							.023
M0/M1	19	23	9	15	10	48	
M2	27	33	21	34	6	29	
M4	22	27	17	28	5	24	
M5	11	13	11	18	0	0	
M6	1	1	1	2	0	0	
AML, unclassified	2	2	2	3	0	0	
Hemoglobin, g/dL							
Median	8.9		9.1		8.4		.470
Range	4.6-12.9		4.6-12.8		7.1-12.9		
Platelets, $\times 10^9$/L							
Median	60.5		55		64		.473
Range	5-378		5-378		21-235		
WBC count, $\times 10^9$/L							
Median	22.8		19.9		24.9		.824
Range	0.8-295.0		0.8-295.0		1.6-118.4		
Percentage of BM blasts							
Median	57		53		66		.056
Range	28-90		28-88		28-90		
Percentage of PB blasts							
Median	50		46.5		58		.045
Range	0-97		0-97		20-95		
FLT3 status							.742
$FLT3^{WT/WT}$	70	83	53	84	17	81	
$FLT3^{ITD/WT}$	14	17	10	16	4	19	
MLL PTD (n = 1 unknown)							.411
Yes	8	10	5	8	3	14	
No	75	90	57	92	18	86	
BAALC expression†							.042
Low	42	50	36	57	6	29	
High	42	50	27	43	15	71	
Extramedullary involvement							
CNS	0	0	0	0	0	0	
Hepatomegaly	4	5	4	6	0	0	.568
Splenomegaly	5	6	4	6	1	5	.99
Lymphadenopathy	8	10	5	8	3	14	.406
Skin infiltrates	10	12	8	13	2	10	.99
Gum hypertrophy	16	19	15	24	1	5	.060
Induction regimen							.078
ADE	39	46	33	52	6	29	
ADEP	45	54	30	48	15	71	

Abbreviations: *ERG, ETS-related gene*; FAB, French-American-British classification; AML, acute myeloid leukemia; BM, bone marrow; PB, peripheral blood; *MLL* PTD, partial tandem duplication of the *MLL* gene; $FLT3^{WT/WT}$, patients with only wild-type *FLT3* genes; $FLT3^{ITD/WT}$ patients with internal tandem duplication of the *FLT3* gene and the wild-type *FLT3* allele; ADE, cytarabine, daunorubicin, and etoposide; ADEP, cytarabine, daunorubicin, etoposide, and valspodar.

*P compares differences in presenting characteristics between patients with the three lowest quartiles and the highest quartile of *ERG* expression.

†*BAALC* expression dichotomized at the median value.[15]

Table 2. Clinical Outcome of Patients Divided Into Quartile Groups According to *ERG* Expression

End Point	Overall (N = 84)	*ERG* Expression Quartiles 1-3 (n = 63)	*ERG* Expression Quartile 4 (n = 21)	*P**
CR				
No.	68	52	16	.532
%	81	83	76	
Relapse				
No.	30	17	13	.001
%	44	33	81	
Death in CR				
No.	7	6	1	.99
%	10	12	6	
CIR				
Median, years	Not reached	Not reached	0.7	< .001
CIR at 5 years				
%	44	33	81	
95% CI	32 to 56	20 to 46	60 to 100	
OS				
Median, years	3.1	Not reached	1.2	.011
Alive at 5 years				
%	43	51	19	
95% CI	32 to 53	38 to 62	6 to 38	

Abbreviations: *ERG, ETS-related gene*; CR, complete remission; CIR, cumulative incidence of relapse; OS, overall survival.
*P compares differences in outcome between patients with the three lowest quartiles and the highest quartile of *ERG* expression.

use of the T7-Oligo(dT) primer (Affymetrix). In vitro transcription was performed with the BioArray HighYield RNA Transcript Labeling Kit (T7) (Enzo Life Science, Farmingdale, NY). Twenty micrograms of fragmented, biotinylated RNA was hybridized to the U133 plus 2.0 GeneChip for 16 hours at 45°C. Scanned images were converted to CEL files using GCOS software (Affymetrix).

Definition of Clinical End Points

CR was defined as recovery of morphologically normal BM and blood counts (ie, neutrophils $\geq$ 1,500/μL and platelets $\geq$ 100,000/μL) and no circulating leukemic blasts or evidence of extramedullary leukemia. Relapse was defined by more than 5% blasts in marrow aspirates or the development of extramedullary leukemia in patients with previously documented CR, according to National Cancer Institute criteria.[34]

Cumulative incidence of relapse (CIR) was measured from the CR date to date of relapse, death, or date last known alive, where death in CR was considered a competing risk. Disease-free survival (DFS) was measured from the CR date until date of relapse or death (regardless of cause), censoring for patients alive at last follow-up. DFS was used only for the 73 patients analyzed for *ERG/ABL* because none of them died in CR, and therefore, DFS reflected the actual relapse risk. Overall survival (OS) was measured from the date the patient was enrolled onto the study until the date of death or date last known alive.

Statistical Methods

The main objective was to evaluate the impact of *ERG* expression on clinical outcome. A set of 84 patients was initially divided into quartile (Q) groups according to levels of *ERG/GPI* expression and subsequently dichotomized into groups including the three lower Qs (Q1, Q2, and Q3) and the upper Q (Q4) of *ERG/GPI* values. Q4 was chosen for the cut point because the relapse risk was significantly different for the Q4 group compared with the Q1 (P = .024), Q2 (P = .002), and Q3 (P = .009) groups. Similarly, the 73 patients characterized for *ERG/ABL* copy numbers in blood were dichotomized into Q1-3 and Q4 groups based on the number of normalized *ERG* copies. Pretreatment clinical features were compared between Q1-3 and Q4 groups using the Fisher's two-sided exact and Wilcoxon rank sum tests for categoric and continuous variables, respectively.

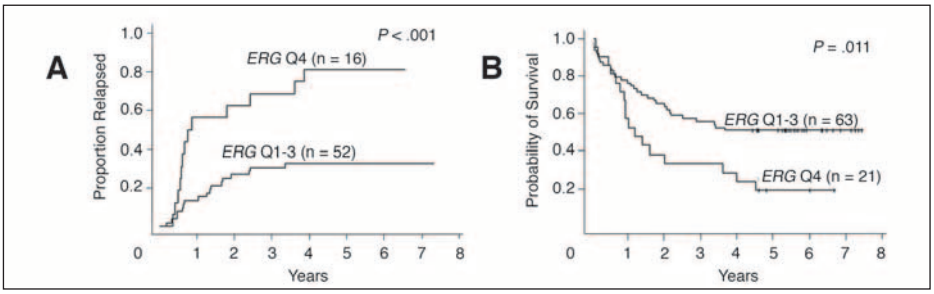

Fig 1. Clinical outcome of patients grouped by *ETS-related gene* (*ERG*) expression into quartile 4 (Q4), the uppermost quartile, and quartiles 1 to 3 (Q1-3), the lower quartiles. (A) Cumulative incidence of relapse. (B) Overall survival.

Table 3. Multivariable Analysis for Patients Divided Into Quartile Groups According to *ERG* Expression*

End Point	Variable	HR	95% CI	*P*
CIR	*ERG* expression, Q4 *v* Q1-3	3.71	1.88 to 7.31	< .001
	Present *v* absent *MLL* PTD	2.70	1.12 to 6.52	.027
OS	Interaction of *ERG* and *BAALC*			.013
	Patients with low *BAALC* expression;† *ERG* expression, Q4 *v* Q1-3	5.40	1.87 to 15.64	.002
	Patients with high *BAALC* expression;† *ERG* expression, Q4 *v* Q1-3	1.04	0.50 to 2.16	.922
	Log[WBC]	1.35	1.07 to 1.70	.012

NOTE. Hazard ratios greater than 1 indicate a higher risk of death or relapse for higher values of continuous variables and for the first category listed for categoric variables.

Abbreviations: *ERG*, *ETS-related gene*; HR, hazard ratio; CIR, cumulative incidence of relapse; Q, quartile; PTD, partial tandem duplication; OS, overall survival; ADE, cytarabine, daunorubicin, and etoposide; ADEP, cytarabine, daunorubicin, etoposide, and valspodar.

*Variables considered for model inclusion were *ERG* expression (Q4 *v* Q1-3), *BAALC* expression (high *v* low), hemoglobin, platelet count, log-transformed WBC count, percentage of peripheral-blood and bone marrow blasts, age, sex, race (white *v* nonwhite), induction treatment (ADE *v* ADEP), *FLT3* genotype ($FLT3^{WT/WT}$ *v* $FLT^{WT/ITD}$) and presence of *MLL* PTD.

†High and low *BAALC* groups defined by a cut point taken at the median *BAALC* expression value.[15]

Estimated probabilities of OS and DFS were calculated using the Kaplan-Meier method, and the log-rank test evaluated differences between survival distributions. Estimates of CIR were calculated, and Gray's k-sample test[35] evaluated differences in relapse rates. Proportional hazards models were constructed for OS and DFS,[36] whereas a multivariable model using Gray's method was constructed for CIR[37] using a limited backwards selection procedure. Variables remaining in the final models were significant at $\alpha = .05$. Adjusted survival curves were generated from the proportional hazards and Gray models using average covariate values.

For microarray data analysis, normalization and model-based expression index (MBEI) computations were performed using dChip version 1.3 (Harvard University, Cambridge, MA).[38-40] Only the perfect match probes were used in the computation of MBEIs, whereas mismatch probes were ignored. Log[MBEI] values were calculated and then exported to BRB-ArrayTools v3.2.3 (National Cancer Institute, Bethesda, MD) for further analysis. Probe sets with a variance in log[MBEI] values above the 80th percentile were retained for further analyses (n = 10,935). A comparison of gene expression between Q1-3 and Q4 was performed by two-sample *t* tests using $\alpha = .001$ as the significance level, which would result in approximately 11 expected false discoveries assuming no gene expression differences between the two groups. A permutation test was performed to validate the results of the parametric test. All analyses were performed by the CALGB Statistical Center.

RESULTS

Impact of ERG *Expression on Clinical Outcome*

No significant differences were observed in most pretreatment clinical characteristics between patients with the lowest 75% (Q1-3) and highest 25% (Q4) of *ERG* expression values. The two groups differed for *BAALC* expression levels ($P = .042$), circulating ($P = .045$) and BM ($P = .056$) blast percentages, gum hypertrophy ($P = .060$), and French-American-British subgroup distribution ($P = .023$; Table 1).

The CR rate was 81%, with no significant difference between patients in Q4 and Q1-3 ($P = .532$; Table 2). With a median follow-up of 5.7 years (range, 4.4 to 7.4 years), patients in Q4 had a worse CIR than patients in Q1-3 ($P < .001$; Fig 1A). The estimated 5-year relapse rate for Q4 patients was 81% compared with 33% for Q1-3 patients (Table 2). Furthermore, the OS was different between the two groups ($P = .011$; Fig 1B). Patients in Q4 had a median survival time of 1.2 years and an estimated 5-year survival rate of 19%; in contrast, median survival time for Q1-3 patients has not been reached, and their estimated 5-year survival rate was 51% (Table 2). When analysis was restricted to the set of patients who achieved CR and received APBSCT as a prescribed consolidation treatment (n = 49), *ERG* remained a significant adverse factor for outcome (data not shown).

On multivariable analysis, high *ERG* expression (ie, Q4) adversely impacted CIR ($P < .001$), as did the presence of *MLL* PTD ($P = .027$; Table 3, Fig 2A). Patients in Q4 had an estimated relapse risk almost four times higher than patients in Q1-3. For OS, an interaction between expression of *ERG* and *BAALC* ($P = .013$) was observed (Table 3). For low *BAALC* levels, patients in Q4 had a shorter survival than patients in Q1-3 ($P = .002$; Fig 2B, Table 3). However, the adverse impact of high *ERG* expression on OS was not observed in patients with high *BAALC* expression ($P = .922$; Fig 2C, Table 3). The only patients who maintained a long-term survival rate greater than 50% were patients with lower expression of both *ERG* and *BAALC*. Additionally, worse survival was associated with higher log[WBC] ($P = .012$; Table 3). Although Q4 patients had a higher percentage of circulating blasts at diagnosis (Table 1), this factor did not impact significantly on clinical outcome and, consequently, was not included in the multivariable models.

Similar results were obtained when we used a second analytic strategy uing specific standard curves to measure *ERG* copy numbers normalized to *ABL*.[32,33] To explore whether *ERG* expression predicts clinical outcome independently from other major unfavorable prognostic markers, we included samples with the most unfavorable *FLT3* genotype, $FLT3^{ITD/-}$. Normalized *ERG* copies ranged from 0.4 to

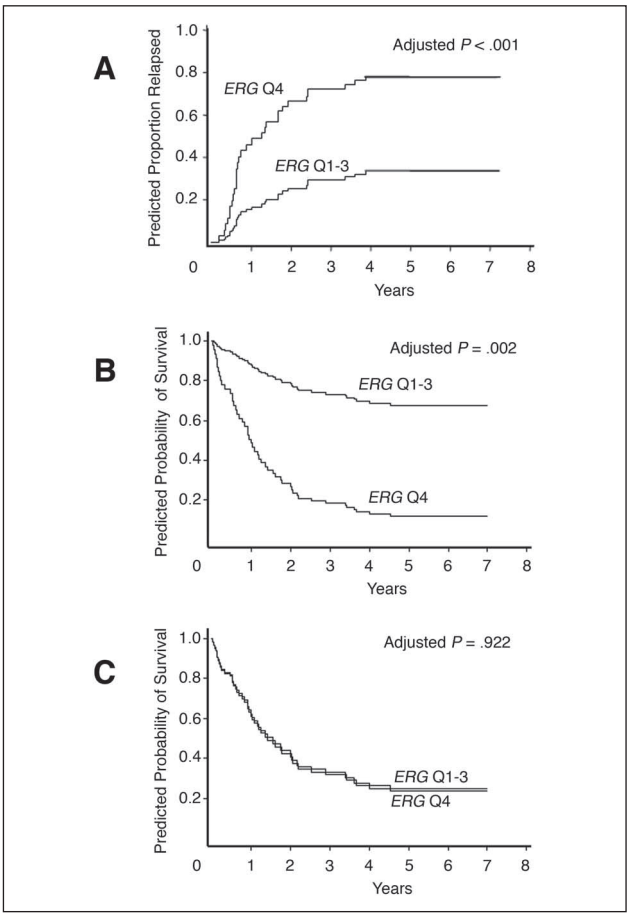

Fig 2. Predicted clinical outcome for the *ETS-related gene* (*ERG*) quartiles 1 to 3 (Q1-3) versus quartile 4 (Q4) groups. (A) Cumulative incidence of relapse. Curves are adjusted for *MLL* partial tandem duplication. (B and C) Overall survival for patients with (B) low and (C) high *BAALC* expression. Curves are adjusted for log[WBC] and *BAALC*. Results are based on 63 Q1-3 patients (36 with low and 27 with high *BAALC* expression) and 21 Q4 patients (six with low and 15 with high *BAALC* expression).

735.9 (median, 17.9 copies). Patients with the lowest 75% (Q1-3; n = 55) of *ERG* copies (median, 11.0 copies; range, 0.4 to 35.1 copies) were compared with patients with the highest 25% (Q4; n = 18) of *ERG* copies (median, 62.2 copies; range, 37.5 to 735.9 copies). There were no significant differences in *FLT3* genotype distribution, which included wild-type *FLT3* ($FLT3^{WT/WT}$), *FLT3* ITD expressing the wild-type allele ($FLT3^{ITD/WT}$), and *FLT3* ITD lacking the wild-type allele ($FLT3^{ITD/-}$; $P = .644$). Patients in Q4 had a worse DFS ($P = .092$) and OS ($P = .003$) than patients in Q1-3, with an estimated 5-year DFS of 25% (95% CI, 6% to 50%) *v* 53% (95% CI, 38% to 66%), respectively, and OS of 22% (95% CI, 7% to 43%) *v* 54% (95% CI, 40% to 67%), respectively. When the analysis was restricted to the subset of 25 patients not analyzed for *ERG*/*GPI*, *ERG* overexpression remained a significant adverse predictor for both OS and DFS (data not shown).

In a multivariable model, high *ERG* copy number (ie, Q4; $P = .023$; Fig 3A) and *FLT3* ITD ($P < .001$) independently predicted worse DFS. The estimated relapse risk for Q4 patients was more than twice the risk for Q1-3 patients (hazard ratio [HR] = 2.55; 95% CI, 1.14 to 5.69). Patients with $FLT3^{ITD/-}$ and $FLT3^{ITD/WT}$ had an estimated relapse risk eight times (HR = 8.37; 95% CI, 2.91 to 24.08) and three times (HR = 3.23; 95% CI, 1.44 to 7.23) the risk of patients with $FLT3^{WT/WT}$, respectively. In a multivariable model for OS, high *ERG* (HR = 3.05; 95% CI, 1.53 to 6.05; $P = .002$; Fig 3B), high log[WBC] (HR = 1.68; 95% CI, 1.14 to 2.46; $P = .009$), and *FLT3* ITD ($P = .002$) predicted shorter survival. Patients with $FLT3^{ITD/-}$ and $FLT3^{ITD/WT}$ had six times (HR = 6.21; 95% CI, 2.42 to 15.93) and two times (HR = 2.09; 95% CI, 0.98 to 4.47), respectively, the estimated risk of dying compared with $FLT3^{WT/WT}$ patients.

Gene Expression Profiling by Oligonucleotide Microarrays in High Versus Low ERG

Microarray gene expression profiling was conducted to assess whether *ERG* overexpression was associated with a specific signature suggestive of the gene's potential leukemogenic role. One hundred seventeen probes in the Affymetrix U133 plus 2.0 GeneChip were differentially expressed ($P < .001$) between the Q1-3 and Q4 groups, including 63 unique, named genes and 49 expressed-sequenced tags. Twenty-five probes corresponding to 14 named genes (Table 4) and eight expressed-sequenced tags had at least a two-fold difference in expression levels between the Q1-3 and Q4 groups (Fig 4). Supporting our real-time RT-PCR results, three probe sets for *ERG* were differentially expressed, with average expression 2.1 to 2.6 times higher in Q4 patients.

Of the 63 named genes overexpressed in Q4, functional characterization was available for 53. Twenty-three of the

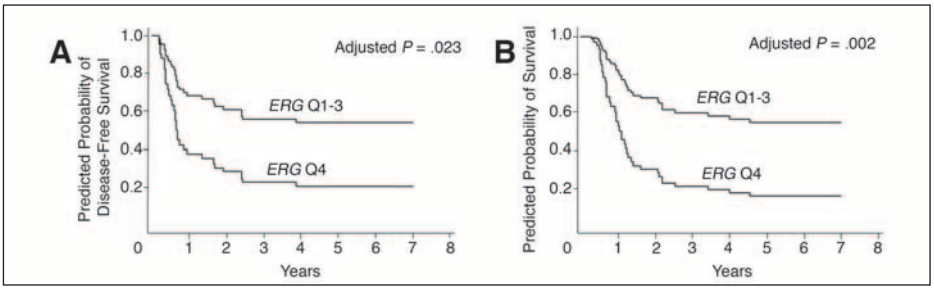

Fig 3. Predicted clinical outcome for patients grouped by *ETS-related gene* (*ERG*) copy number into the *ERG* quartiles 1 to 3 (Q1-3) and quartile 4 (Q4) groups. (A) Disease-free survival. Curves are adjusted for *FLT3* genotypes ($FLT3^{WT/WT}$ *v* $FLT3^{WT/ITD}$ and $FLT3^{WT/WT}$ *v* $FLT3^{ITD/-}$). (B) Overall survival. Curves are adjusted for log[WBC] and *FLT3* genotypes.

Table 4. Named, Differentially Expressed Genes That Were Expressed Two-Fold or More in the *ERG* Uppermost Quartile (quartile 4) Compared With Lower Quartiles (quartiles 1 to 3)

Gene Symbol	Name	Fold Change	*P*
HEMGN	Hemogen	2.69	.00073
ERG	V-ets erythroblastosis virus E26 oncogene like (avian)	2.60	.00039
IKIP	IKK interacting protein	2.48	.00049
BCL11A	B-cell CLL/lymphoma 11A (zinc finger protein)	2.46	.00002
DAPK1	Death-associated protein kinase 1	2.19	.00005
GRSP1	GRP1-binding protein GRSP1	2.16	.00066
GAS5	Growth arrest–specific 5	2.14	.00003
GUCY1A3	Guanylate cyclase 1, soluble, alpha 3	2.14	.00032
KLHDC1	Kelch domain containing 1	2.05	.00083
HIST2H4	Histone 2, H4	2.04	.00077
ATP6V1C1	ATPase, H+ transporting, lysosomal 42 kd, V1 subunit C, isoform 1	2.03	.00099
GTF2H2	General transcription factor IIH	2.03	.00002
RAB10	RAB10, member RAS oncogene family	2.01	.00049
ZNF638	Zinc finger protein 638	2.00	.00040

Abbreviations: *ERG, ETS-related gene*; IKK, I kappa B kinase; CLL, chronic lymphoid leukemia.

genes are involved in DNA and/or RNA binding and chromatin remodeling, including general transcription activators (ie, *GTF2H2*, *FAM48A*, *TCF12*, and *GCN5L2*) or repressors (*CTBP2*) and lineage-specific transcription regulators, such as *BCL11A*, which is involved in lymphoid cell development,[41] and *HEMGN*, which is involved in myeloid differentiation.[42-44] Other genes encode small GTPases, members of the RAS superfamily (*RALA* and *RAB10*) or their regulators (*RABGAP1* and *ARHGAP22*), or proteins involved in membrane-receptor signaling pathways (*GPR21*, *DPAGT1*, *PILRB*, *GPR89*, and *FRMD4B*), including activators of small GTPases. Surprisingly, we found overexpression of three proapoptotic genes (*PACAP*, *IKIP*, and *DAPK1*); *DAPK1* was recently reported to be silenced by methylation in AML.[45] Finally, the following three genes with unknown function but linked to hereditary diseases were upregulated: *ATXN2*, which is mutated in autosomal-dominant spinocerebellar ataxia-2,[46] *SHANK3*, which is involved in the 22q13 deletion syndrome,[47] and *OFD1*, which is mutated in orofaciodigital syndrome I.[48] To our knowledge, overexpression or mutations of these genes have hitherto not been reported in AML.

DISCUSSION

Although molecular mechanisms underlying diverse clinical outcome in cytogenetically normal AML are not fully

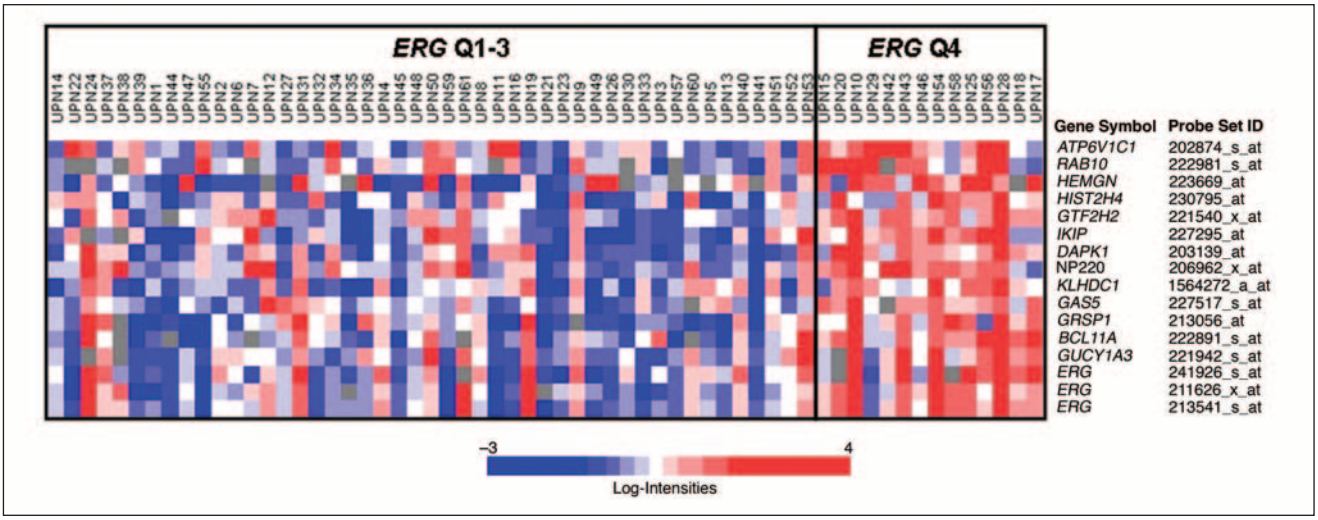

Fig 4. Heat map of two-fold or greater differentially expressed genes between *ETS-related gene* (*ERG*) quartiles 1 to 3 (Q1-3) and quartile 4 (Q4). Columns represent samples, and rows represent genes ordered by hierarchical cluster analysis. Shading indicates relative expression of each gene with respect to the gene median expression (white equal to, red above, and blue below the median value; gray, missing values as a result of unreliable measurement).

understood, recent studies have identified several biomarkers correlated with prognosis.[6-18] Here, we report for the first time that *ERG* overexpression predicted an increased relapse risk and short survival in AML patients with normal karyotype by both univariable and multivariable analyses. Despite a relatively small number of patients studied, our data support the addition of *ERG* overexpression to the emerging list of markers predictive for clinical outcome in cytogenetically normal AML.

With the number of prognostic markers growing, the relative contribution of each in predicting treatment outcome becomes important. In the current study, although high *ERG* levels correlated with an increased relapse risk regardless of the *BAALC* expression, a prognostic interaction between *ERG* and *BAALC* expression was observed for OS, with a shorter survival associated with high *ERG* levels only among patients with low *BAALC* expression. Likewise, interactions among prognostic markers, such as a greater prognostic importance of *FLT3* ITD compared with *BAALC* overexpression and of *BAALC* overexpression compared with *CEBPA* mutations, have been reported previously.[15,17,18] Because data on the impact of the concurrent presence of two or more molecular markers in cytogenetically normal AML are limited, further investigation of prognostic interactions is required by large prospective studies, with the goal of designing a prioritized, clinically relevant prognostic classification.

Although our data suggest that expression of *ERG* is useful for prognostic stratification of cytogenetically normal AML, the mechanism through which *ERG* overexpression contributes to myeloid leukemogenesis remains unknown. *ERG* is one of more than 30 members of the *ETS* gene family, most of which are downstream nuclear targets of signal transduction pathways regulating and promoting cell differentiation, proliferation, and tissue invasion.[21,22,27] Rearrangement of *ERG* was initially discovered in Ewing sarcomas harboring t(21;22)(q22;q12), which at the molecular level fuses *ERG* with *EWS*.[25] Other *ETS* family members also fuse with *EWS* in Ewing sarcomas.[49-51] In AML carrying t(16;21)(p11;q22), *ERG* was found rearranged with *FUS*, linking *ERG* with myeloid leukemogenesis for the first time.[24] Interestingly, like *EWS*, *FUS* is a member of the TET family of RNA-binding proteins,[52] supporting the notion that gene rearrangements involving *ETS* members are often characterized by a TET-related transactivation domain at the N terminus and ETS DNA binding and protein-protein interaction domains at the C terminus. This structure likely increases the oncogenic activity of the resulting chimeric transcription factors by redirecting them to specific targets. Interestingly, a role for *ERG* in endothelial cell differentiation and angiogenesis was recently suggested.[23,53]

Our microarray gene expression profiling analysis, using a supervised analysis, identified a molecular signature for patients in the highest Q (Q4) of *ERG* expression. Many genes upregulated in Q4, including *ERG*, encode proteins functioning as transcription factors or involved in chromatin remodeling and RNA processing; others encode proteins regulating cell differentiation, proliferation, and apoptosis. Interestingly, the recently described *HEMGN* gene (also known as *EDAG*) was the most differentially expressed gene between the two groups. This gene reportedly regulates proliferation, differentiation, and apoptosis of hematopoietic cells and seems to be significantly overexpressed in refractory AML patients but not in patients with chemotherapy-sensitive disease.[42-44]

Because the molecular signature for the *ERG* overexpressers was defined by patients' clustering based on previously identified prognostic groups (ie, Q4 *v* Q1-3), it was not surprising that this gene profile was associated with a worse clinical outcome (data not shown). Obviously, for *ERG* expression to become a molecular marker used routinely for risk stratification in AML with normal cytogenetics, our data require confirmation in an independent, large patient cohort.

Acknowledgment

We thank Marko I. Klisovic for his excellent technical support.

Appendix

The Appendix is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF (via Adobe® Acrobat Reader®) version.

Authors' Disclosures of Potential Conflicts of Interest

The authors indicated no potential conflicts of interest.

REFERENCES

1. Mrózek K, Heerema NA, Bloomfield CD: Cytogenetics in acute leukemia. Blood Rev 18: 115-136, 2004

2. 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 92:2322-2333, 1998

3. 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 96:4075-4083, 2000

4. Byrd JC, Mrózek 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 100:4325-4336, 2002

5. Farag SS, Ruppert AS, Mrózek K, et al: Outcome of induction and postremission therapy in younger adults with acute myeloid leukemia with normal karyotype: A Cancer and Leukemia Group B study. J Clin Oncol 23:482-493, 2005

6. Marcucci G, Mrózek K, Bloomfield CD: Molecular heterogeneity and prognostic biomarkers in adults with acute myeloid leukemia and normal cytogenetics. Curr Opin Hematol 12:68-75, 2005

7. Caligiuri MA, Strout MP, Lawrence D, et al: Rearrangement of *ALL1* (*MLL*) in acute myeloid leukemia with normal cytogenetics. Cancer Res 58:55-59, 1998

8. Schnittger S, Kinkelin U, Schoch C, et al: Screening for MLL tandem duplication in 387 unselected patients with AML identify a prognostically

unfavorable subset of AML. Leukemia 14:796-804, 2000

9. Döhner K, Tobis K, Ulrich R, et al: Prognostic significance of partial tandem duplications of the MLL gene in adult patients 16 to 60 years old with acute myeloid leukemia and normal cytogenetics: A study of the Acute Myeloid Leukemia Study Group Ulm. J Clin Oncol 20: 3254-3261, 2002

10. Whitman SP, Archer KJ, Feng L, et al: Absence of the wild-type allele predicts poor prognosis in adult *de novo* acute myeloid leukemia with normal cytogenetics and the internal tandem duplication of *FLT3*: A Cancer and Leukemia Group B study. Cancer Res 61:7233-7239, 2001

11. Kottaridis PD, Gale RE, Frew ME, et al: The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: Analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials. Blood 98:1752-1759, 2001

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 99:4326-4335, 2002

13. Schnittger S, Schoch C, Dugas M, et al: Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: Correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. Blood 100:59-66, 2002

14. Fröhling S, Schlenk RF, Breitruck J, et al: Prognostic significance of activating *FLT3* mutations in younger adults (16 to 60 years) with acute myeloid leukemia and normal cytogenetics: A study of the AML Study Group Ulm. Blood 100:4372-4380, 2002

15. 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 102:1613-1618, 2003

16. Baldus CD, Thiede C, Bloomfield CD, et al: High *BAALC* expression as an independent adverse risk factor in 309 patients with acute myeloid leukemia (AML) and normal cytogenetics: Results of the SHG'96 study. Blood 104: 554a, 2004 (abstr 2009)

17. Bienz M, Ludwig M, Oppliger Leibundgut E, et al: Risk assessment in patients with acute myeloid leukemia and a normal karyotype. Clin Cancer Res 11:1416-1424, 2005

18. 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 22:624-633, 2004

19. Whitman SP, Liu S, Vukosavljevic T, et al: The *MLL* partial tandem duplication: Evidence for recessive gain-of-function in acute myeloid leukemia identifies a novel patient subgroup for molecular-targeted therapy. Blood 106:345-352, 2005

20. 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 101: 3915-3920, 2004

21. Oikawa T: ETS transcription factors: Possible targets for cancer therapy. Cancer Sci 95: 626-633, 2004

22. Oikawa T, Yamada T: Molecular biology of the Ets family of transcription factors. Gene 303:11-34, 2003

23. McLaughlin F, Ludbrook VJ, Cox J, et al: Combined genomic and antisense analysis reveals that the transcription factor Erg is implicated in endothelial cell differentiation. Blood 98:3332-3339, 2001

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

25. Sorensen PHB, Lessnick SL, Lopez-Terrada D, et al: A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. Nat Genet 6:146-151, 1994

26. Petrovics G, Liu A, Shaheduzzaman S, et al: Frequent overexpression of *ETS*-related gene-1 (*ERG1*) in prostate cancer transcriptome. Oncogene 24:3847-3852, 2005

27. Hart AH, Corrick CM, Tymms MJ, et al: Human *ERG* is a proto-oncogene with mitogenic and transforming activity. Oncogene 10:1423-1430, 1995

28. Kolitz JE, George SL, Dodge RK, et al: Dose escalation studies of cytarabine, daunorubicin, and etoposide with and without multidrug resistance modulation with PSC- 833 in untreated adults with acute myeloid leukemia younger than 60 years: Final induction results of Cancer and Leukemia Group B Study 9621. J Clin Oncol 22:4290-4301, 2004

29. Döhner H, Arthur DC, Ball ED, et al: Trisomy 13: A new recurring chromosome abnormality in acute leukemia. Blood 76:1614-1621, 1990

30. Mitelman F (ed): ISCN (1995): An International System for Human Cytogenetic Nomenclature. Basel, Switzerland, Karger, 1995

31. Beillard E, Pallisgaard N, van der Velden VHJ, et al: Evaluation of candidate control genes for diagnosis and residual disease detection in leukemic patients using 'real-time' quantitative reverse-transcriptase polymerase chain reaction (RQ-PCR): A Europe Against Cancer program. Leukemia 17:2474-2486, 2003

32. Gabert J, Beillard E, van der Velden VHJ, et al: Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia: A Europe Against Cancer program. Leukemia 17: 2318-2357, 2003

33. Marcucci G, Caligiuri MA, Döhner H, et al: Quantification of *CBFbeta/MYH11* fusion transcript by real-time RT-PCR in patients with INV(16) acute myeloid leukemia. Leukemia 15: 1072-1080, 2001

34. Cheson BD, Cassileth PA, Head DR, et al: Report of the National Cancer Institute-sponsored workshop on definitions of diagnosis and response in acute myeloid leukemia. J Clin Oncol 8:813-819, 1990

35. Gray RJ: A class of k-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat 16:1141-1154, 1988

36. Cox DR: Regression models and life tables. J R Stat Soc B 34:187-202, 1972

37. Fine JP, Gray RJ: A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 94:496-509, 1999

38. Li C, Wong W: DNA-chip analyzer (dChip), in Parmigiani G, Garrett ES, Irizarry R, et al (eds): The Analysis of Gene Expression Data: Methods and Software. New York, NY, Springer-Verlag, 2003, pp 120-141

39. Schadt EE, Li C, Ellis B, et al: Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data. J Cell Biochem 37:120-125, 2001 (suppl)

40. Li C, Wong WH: Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection. Proc Natl Acad Sci U S A 98:31-36, 2001

41. Satterwhite E, Sonoki T, Willis TG, et al: The *BCL11* gene family: Involvement of *BCL11A* in lymphoid malignancies. Blood 98:3413-3420, 2001

42. Li CY, Zhan YQ, Xu CW, et al: EDAG regulates the proliferation and differentiation of hematopoietic cells and resists cell apoptosis through the activation of nuclear factor-kappaB. Cell Death Differ 11:1299-1308, 2004

43. Yang LV, Nicholson RH, Kaplan J, et al: *Hemogen* is a novel nuclear factor specifically expressed in mouse hematopoietic development and its human homologue EDAG maps to chromosome 9q22, a region containing breakpoints of hematological neoplasms. Mech Dev 104:105-111, 2001

44. An L-L, Li G, Wu K-F, et al: High expression of EDAG and its significance in AML. Leukemia 19:1499-1502, 2005

45. Voso MT, Scardocci A, Guidi F, et al: Aberrant methylation of DAP-kinase in therapy-related acute myeloid leukemia and myelodysplastic syndromes. Blood 103:698-700, 2004

46. Furtado S, Payami H, Lockhart PJ, et al: Profile of families with parkinsonism-predominant spinocerebellar ataxia type 2 (SCA2). Mov Disord 19:622-629, 2004

47. Wilson HL, Wong ACC, Shaw SR, et al: Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of *SHANK3/PROSAP2* in the major neurological symptoms. J Med Genet 40:575-584, 2003

48. Romio L, Wright V, Price K, et al: OFD1, the gene mutated in oral-facial-digital syndrome type 1, is expressed in the metanephros and in human embryonic renal mesenchymal cells. J Am Soc Nephrol 14:680-689, 2003

49. Delattre O, Zucman J, Plougastel B, et al: Gene fusion with an *ETS* DNA-binding domain caused by chromosome translocation in human tumours. Nature 359:162-165, 1992

50. Urano F, Umezawa A, Hong W, et al: A novel chimera gene between *EWS* and *E1A-F*, encoding the adenovirus E1A enhancer-binding protein, in extraosseous Ewing's sarcoma. Biochem Biophys Res Commun 219:608-612, 1996

51. Jeon I-S, Davis JN, Braun BS, et al: A variant Ewing's sarcoma translocation (7;22) fuses the *EWS* gene to the *ETS* gene *ETV1*. Oncogene 10:1229-1234, 1995

52. de Alava E, Pardo J: Ewing tumor: Tumor biology and clinical applications. Int J Surg Pathol 9:7-17, 2001

53. Baltzinger M, Mager-Heckel A-M, Remy P: Xl erg: Expression pattern and overexpression during development plead for a role in endothelial cell differentiation. Dev Dyn 216:420-433, 1999