miR-3151 interplays with its host gene BAALC and independently affects outcome of patients with cytogenetically normal acute myeloid leukemia

Ann-Kathrin Eisfeld,1 Guido Marcucci,1 Kati Maharry,1,2 Sebastian Schwind,1 Michael D. Radmacher,1,2 Deedra Nicolet,1,2 Heiko Becker,1 Krzysztof Mrózek,1 Susan P. Whitman,1 Klaus H. Metzeler,1 Jason H. Mendler,1 Yue-Zhong Wu,1 Sandya Liyanarachchi,1 Ravi Patel,1 Maria R. Baer,3 Bayard L. Powell,1 Thomas H. Carter,5 Joseph O. Moore,6 Jonathan E. Kolitz,7 Meir Wetzler,8 Michael A. Caligiuri,1 Richard A. Larson,9 Stephan M. Tanner,1 Albert de la Chapelle,1 and Clara D. Bloomfield1

High BAALC expression levels are associated with poor outcome in cytogenetically normal acute myeloid leukemia (CN-AML) patients. Recently, miR-3151 was discovered in intron 1 of BAALC. To evaluate the prognostic significance of miR-3151 expression levels and to gain insight into the biologic and prognostic interplay between miR-3151 and its host, miR-3151 and BAALC expression were measured in pretreatment blood of 179 CN-AML patients. Gene-expression profiling and miRNA-expression profiling were performed using microarrays. High miR-3151 expression was associated with shorter disease-free and overall survival, whereas high BAALC expression predicted failure of complete remission and shorter overall survival. Patients exhibiting high expression of both miR-3151 and BAALC had worse outcome than patients expressing low levels of either gene or both genes. In gene-expression profiling, high miR-3151 expression showed down-regulation of genes involved in transcriptional regulation, posttranslational modification, and cancer pathways. Two genes, FBXL20 and USP40, were validated as direct miR-3151 targets. The results of the present study show that high expression of miR-3151 is an independent prognosticator for poor outcome in CN-AML and affects different outcome endpoints than its host gene, BAALC. The combination of both markers identified a patient subset with the poorest outcome. This interplay between an intronic miR and its host may have important biologic implications. (Blood. 2012;120(2):249-258)

Introduction

Acute myeloid leukemia (AML) is a clinically, cytogenetically, and molecularly heterogeneous disease. Despite recent advances in our understanding of the mechanisms of leukemogenesis and the identification of markers that allow molecular-based stratification to risk-adapted therapies, the majority of patients with AML are not cured.1 The clinical outcome is particularly poor in older (≥60 years of age) patients, who have long-term survival rates of only 7%-15%.2

To date, the prognostic impact of molecular genetic markers has been studied most extensively in patients with cytogenetically normal AML (CN-AML).3 In this large cytogenetic subset, several molecular markers have been found to be associated with outcome.3,9 Among them, mutations in the nucleophosmin (nucleolar phosphoprotein B23, numatrin; NPM1) and CCAAT/enhancer binding protein (C/EBP) genes are associated with favorable outcome and have been included as provisional entities in the World Health Organization classification of AML.10 In addition, these 2 mutations and the presence of internal tandem duplications of the fms-related tyrosine kinase 3 gene (FLT3-ITD) that has been associated with unfavorable outcome have been incorporated into a standardized reporting system for genetic abnormalities suggested by an international expert panel on behalf of the European LeukemiaNet (ELN).11

Another strong prognostic factor associated with treatment resistance and poor outcome is high expression of the brain and acute leukemia cytoplasmic (BAALC) gene, which was identified by cDNA-based representational difference analysis in leukemia patients.12 The prognostic impact of BAALC expression has been most extensively studied in CN-AML patients.13–16 Although recent data suggest that BAALC contributes to leukemogenesis by interfering with normal patterns of myeloid differentiation,17 the function of this gene and the mechanism(s) through which its high expression levels affect clinical outcome negatively remain(s) unknown.

Recently, small RNA deep sequencing of melanoma18 and pediatric acute lymphoblastic leukemia samples19 identified a new miR, miR-3151, embedded in intron 1 of the BAALC gene. miRs are short, noncoding RNAs that hybridize to and regulate the expression of targeted mRNAs20 and have been implicated in...
leukemogenesis. Furthermore, expression levels of miRs may be used to refine risk assessment in AML in CN-AML patients.

miR genes are located throughout the genome, but approximately one-third of mammalian miRs reside within the introns of host genes. Some of these intronic miRs have been found to act in functional synergism with their host genes. These findings led us to hypothesize that miR-3151 could be overexpressed in patients with elevated BAALC levels and thus might contribute to the adverse prognostic impact of its host gene, either acting in concert with BAALC or perhaps even being the element predominantly responsible for the adverse outcome observed in BAALC-overexpressing AML patients.

In the present study, we measured miR-3151 and BAALC expression in a cohort of molecularly well-characterized de novo CN-AML patients to assess the impact of miR-3151 expression alone and in combination with the expression of its host BAALC on outcome and also explored some downstream effects of miR-3151 expression.

**Methods**

**Patients, treatment, and cytogenetic studies**

One-hundred-seventy-nine patients 60 years of age or older with de novo CN-AML who were treated with intensive cytarabine/daunorubicin-based regimens on Cancer and Leukemia Group B (CALGB) frontline clinical protocols (for details see supplemental Methods, available on the Blood Web site; see the Supplemental Materials link at the top of the online article), were included in this study. All study protocols received institutional review board approval at the participating institutions. Cytogenetic analyses were performed on pretreatment BM samples by CALGB-approved institutional cytogenetic laboratories as part of CALGB 8461, a prospective cytogenetic companion study. The diagnosis of normal cytogenetics was based on the analysis of ≥20 BM metaphase cells and confirmed by central karyotype review.

**Molecular analyses**

Pretreatment peripheral blood samples were analyzed for miR-3151 and BAALC expression levels by real-time RT-PCR. The TaqMan assays were carried out for each sample in triplicate using TaqMan Primer-Probe sets for BAALC and the respective housekeeping genes (RNU44). Additional molecular markers were analyzed centrally in pretreatment BM or peripheral blood samples as described previously and included: FLT3-ITD,

**GERM and MEP**

Gene-expression profiling (GEP) of samples was performed using the Affymetrix U133 plus 2.0 array (Affymetrix) and miR-expression profiling (MEP) was performed using The Ohio State University custom miR array (OSU CCC Version 4.0), as described previously (see supplemental Methods for details). The miR microarray data have been deposited in ArrayExpress under accession number E-MTAB-1074 and the microarray expression data under accession number E-MTAB-1075.

**Validation of FBXL20 and USP40 as direct miR-3151 targets**

For stable expression of miR-3151, the miR-3151 stem-loop was cloned into a lentiviral expression vector, as described in supplemental Methods, using lentiviral miR-scramble as the respective control for all experiments. To analyze the effects of forced miR-3151 expression on the predicted target genes, we assessed the expression of FBXL20 and USP40 on mRNA level, as described in the supplemental materials, and compared it with the effects of cells infected with scramble control using real-time RT-PCR. Western Blotting to test the effects of miR-3151 on protein level was performed as described in detail in the supplemental materials. The 3′-untranslated regions of FBXL20 and USP40 were cloned into a luciferase reporter vector (wild-type vs mutated miR-3151–binding sequence, see supplemental Table 3 for details) and luciferase activity was assessed as described in supplemental Methods.

**Definition of clinical end points and statistical analysis**

The main objective of this study was to evaluate the prognostic impact on clinical outcome in older CN-AML patients of miR-3151 expression alone and in combination with its host gene, BAALC. Median expression levels of miR-3151 and BAALC were used to define low and high miR-3151 and BAALC expressers, respectively, for all analyses (see supplemental Methods for details).

Definitions of clinical end points (i.e., complete remission [CR], disease-free [DFS], and overall survival [OS]) and details of statistical analyses, including variable selection for statistical modeling, are provided in supplemental Methods. Associations between patients with low and high expression of miR-3151 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 was used to evaluate differences between survival distributions. Multivariable logistical 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 miR-3151 expression by adjusting for other variables using a limited backward selection procedure. For achievement of CR, estimated odds ratios (ORs) and for survival end points, hazard ratios (HRs) with their corresponding 95% confidence intervals were obtained for each significant prognostic factor.

For the GEP and MEP, summary measures of gene and miR expression, respectively, were computed, normalized, and filtered (supplemental Methods). The profiles were derived by comparing gene expression between low and high miR-3151 expressers. Univariable significance levels of P = .001 for GEP (P = .005 for MEP) were used to determine the probe sets (probes) that comprised the signatures.

All clinical analyses were performed by the Alliance for Clinical Trials in Oncology Statistics and Data Center.

**Results**

**Associations of miR-3151 expression with clinical and molecular characteristics**

Patients with high miR-3151 expression had lower percentages of circulating blasts (P = .02), and were more likely to be NPM1 wild-type (P < .001), belong to the ELN Intermediate-I Genetic Group, and harbor mutations of RUNX1 (P < .001) than low expressers (P = .05; Table 1). In addition, high miR-3151 expression was associated with high expression levels of its host gene, BAALC (P < .001), and also of MN1 (P = .05). Approximately
Table 1. Clinical and molecular characteristics according to miR-3151 expression status in CN-AML patients 60 years of age or older

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Low miR-3151* (n = 90)</th>
<th>High miR-3151* (n = 89)</th>
<th>P†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>68.99</td>
<td>68.88</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>60-79</td>
<td>60-81</td>
<td></td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td>0.88</td>
</tr>
<tr>
<td>Male</td>
<td>48 (53)</td>
<td>49 (55)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>42 (47)</td>
<td>40 (45)</td>
<td></td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td>White</td>
<td>79 (88)</td>
<td>83 (95)</td>
<td></td>
</tr>
<tr>
<td>Nonwhite</td>
<td>11 (12)</td>
<td>4 (5)</td>
<td></td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td></td>
<td></td>
<td>0.47</td>
</tr>
<tr>
<td>Median</td>
<td>9.3</td>
<td>9.5</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>6.0-11.7</td>
<td>6.5-15.0</td>
<td></td>
</tr>
<tr>
<td>Platelet count, × 10^9/L</td>
<td></td>
<td></td>
<td>0.36</td>
</tr>
<tr>
<td>Median</td>
<td>60</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>4-481</td>
<td>11-850</td>
<td></td>
</tr>
<tr>
<td>WBC count, × 10^9/L</td>
<td></td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td>Median</td>
<td>26.2</td>
<td>38.7</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>1.4-450.0</td>
<td>1.1-434.1</td>
<td></td>
</tr>
<tr>
<td>Percentage of blood blasts</td>
<td></td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>Median</td>
<td>65</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>0-96</td>
<td>0-99</td>
<td></td>
</tr>
<tr>
<td>Percentage of BM blasts</td>
<td></td>
<td></td>
<td>0.52</td>
</tr>
<tr>
<td>Median</td>
<td>70</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>11-97</td>
<td>4-96</td>
<td></td>
</tr>
<tr>
<td>Extramedullary involvement, n (%)</td>
<td>21 (24)</td>
<td>24 (28)</td>
<td>0.73</td>
</tr>
<tr>
<td>NPM1, n (%)</td>
<td></td>
<td></td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Mutated</td>
<td>71 (79)</td>
<td>43 (49)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>19 (21)</td>
<td>45 (51)</td>
<td></td>
</tr>
<tr>
<td>FLT3-ITD, n (%)</td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Present</td>
<td>33 (37)</td>
<td>33 (38)</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>57 (63)</td>
<td>55 (63)</td>
<td></td>
</tr>
<tr>
<td>CEBPA, n (%)</td>
<td></td>
<td></td>
<td>0.83</td>
</tr>
<tr>
<td>Mutated</td>
<td>12 (13)</td>
<td>13 (15)</td>
<td></td>
</tr>
<tr>
<td>Single mutated</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Double mutated</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>78 (87)</td>
<td>76 (85)</td>
<td></td>
</tr>
<tr>
<td>ELN Genetic Group, n (%)‡</td>
<td></td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>Favorable</td>
<td>50 (56)</td>
<td>35 (40)</td>
<td></td>
</tr>
<tr>
<td>Intermediate-I</td>
<td>40 (44)</td>
<td>52 (60)</td>
<td></td>
</tr>
<tr>
<td>RUNX1, n (%)</td>
<td></td>
<td></td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Mutated</td>
<td>4 (5)</td>
<td>19 (24)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>80 (95)</td>
<td>59 (76)</td>
<td></td>
</tr>
<tr>
<td>FLT3-TKD, n (%)</td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Present</td>
<td>7 (8)</td>
<td>7 (8)</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>82 (92)</td>
<td>79 (92)</td>
<td></td>
</tr>
<tr>
<td>WT1, n (%)</td>
<td></td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>Mutated</td>
<td>3 (3)</td>
<td>9 (10)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>87 (97)</td>
<td>79 (90)</td>
<td></td>
</tr>
<tr>
<td>TET2, n (%)</td>
<td></td>
<td></td>
<td>0.33</td>
</tr>
<tr>
<td>Mutated</td>
<td>24 (27)</td>
<td>30 (35)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>65 (73)</td>
<td>56 (65)</td>
<td></td>
</tr>
<tr>
<td>MLL-PTD, n (%)</td>
<td></td>
<td></td>
<td>0.50</td>
</tr>
<tr>
<td>Present</td>
<td>3 (5)</td>
<td>6 (8)</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>63 (95)</td>
<td>67 (92)</td>
<td></td>
</tr>
<tr>
<td>IDH1, n (%)</td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>R132 mutated</td>
<td>12 (13)</td>
<td>10 (11)</td>
<td></td>
</tr>
<tr>
<td>V71I mutated</td>
<td>0</td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>77 (87)</td>
<td>76 (87)</td>
<td></td>
</tr>
<tr>
<td>IDH2, n (%)</td>
<td></td>
<td></td>
<td>0.71</td>
</tr>
<tr>
<td>R140 mutated</td>
<td>20 (22)</td>
<td>17 (20)</td>
<td></td>
</tr>
<tr>
<td>R172 mutated</td>
<td>20 (22)</td>
<td>15 (16)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>69 (78)</td>
<td>70 (80)</td>
<td></td>
</tr>
</tbody>
</table>

*The median expression value was used as the cutoff point.
†P values for categorical variables are from the Fisher exact test; P values for continuous variables are from the Wilcoxon rank-sum test.
‡The ELN Favorable Genetic Group comprises patients with mutated CEBPA and those with mutated NPM1 without FLT3-ITD; the ELN Intermediate-I Genetic Group includes patients with CEBPA wild-type who are FLT3-ITD–positive and NPM1–mutated, FLT3-ITD–negative and NPM1 wild-type, or FLT3-ITD–positive and NPM1 wild-type.
two-thirds of the patients had a concordant expresser status for miR-3151 and BAALC expression levels, whereas one-third of the patients exhibited up-regulation of only 1 of the 2 markers.

**Prognostic value of miR-3151 expression in CN-AML**

Patients with high miR-3151 expression had a lower CR rate ($P = .005$, 62% vs 81%) compared with low expressers. With a median follow-up time for living patients of 5.1 years (range, 4.1-11.6), high miR-3151 expressers had a shorter DFS ($P = .003$, HR = 1.76; Figure 1A). At 3 years after CR achievement, only 7% of high miR-3151-expressing patients were disease free compared with 26% of low expressers. High miR-3151 expressers also had a shorter OS ($P < .001$, HR = 1.86; Figure 1B). Three years after diagnosis, 10% of high miR-3151 expressers were still alive compared with 32% of low expressers.

Because miR-3151 expression levels were associated with the expression levels of its host gene, BAALC, we analyzed the impact on outcome end points of both genes using bivariable models. Analyses showed that high expression levels of either marker had significant adverse impact on CR (miR-3151, $P < .001$, OR = 0.47; BAALC, $P < .001$, OR = 0.3), DFS (miR-3151, $P = .01$, HR = 1.68; BAALC, $P = .003$, HR = 1.82), and OS (miR-3151, $P = .002$, HR = 1.68; BAALC, $P < .001$, HR = 2.01). Therefore, miR-3151 and BAALC independently added information for determination of all outcome end points (supplemental Table 1).

Despite the strong association of miR-3151 and BAALC expression levels, approximately one-third of patients were discordant in expresser status of the 2 markers (Table 1). Therefore, we next investigated whether their combination (miR-3151/BAALC, high/high, high/low, low/high, low/low) would reveal differences in impact on outcome end points. Patients who had high expression of both miR-3151 and BAALC demonstrated the lowest CR rates (50%), whereas patients who highly expressed only 1 of the 2 markers had intermediate CR rates (low miR-3151/high BAALC, 71%; high miR-3151/low BAALC, 79%), and low expressers of both markers had the highest CR rate (87%, $P < .001$). Patients with high expression of both miR-3151 and BAALC had significantly shorter DFS and OS than those expressing both markers at low levels ($P < .001$ for both DFS and OS; Figure 2A-B) or those who exhibited high expression of only one of the markers (DFS, $P = .03$, OS, $P = .01$). Of patients who expressed both the miR and its host gene at high levels, only 1 survived longer than 3 years after diagnosis.

In multivariable analyses (Table 2), after adjustment for BAALC expression status and WBC count, patients with high miR-3151 expression had a trend toward a lower CR rate ($P = .13$, OR = 0.56). High miR-3151 expressers had shorter DFS ($P < .001$,...

---

**Table 1. (continued)**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Low miR-3151* (n = 90)</th>
<th>High miR-3151* (n = 89)</th>
<th>$P^\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ASXL1, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutated</td>
<td>9 (10)</td>
<td>16 (19)</td>
<td>.13</td>
</tr>
<tr>
<td>Wild-type</td>
<td>79 (90)</td>
<td>69 (81)</td>
<td></td>
</tr>
<tr>
<td><strong>DNMT3A, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutated</td>
<td>27 (31)</td>
<td>29 (35)</td>
<td>.74</td>
</tr>
<tr>
<td>Non-R882</td>
<td>12</td>
<td>10</td>
<td>1.00</td>
</tr>
<tr>
<td>Wild-type</td>
<td>59 (69)</td>
<td>55 (65)</td>
<td></td>
</tr>
<tr>
<td><strong>ERG expression group, n (%)</strong></td>
<td></td>
<td></td>
<td>.40</td>
</tr>
<tr>
<td>High</td>
<td>42 (58)</td>
<td>35 (50)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>31 (42)</td>
<td>35 (50)</td>
<td></td>
</tr>
<tr>
<td><strong>BAALC expression group, n (%)</strong></td>
<td></td>
<td></td>
<td>&lt; .001</td>
</tr>
<tr>
<td>High</td>
<td>34 (39)</td>
<td>56 (66)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>54 (61)</td>
<td>29 (34)</td>
<td></td>
</tr>
<tr>
<td><strong>MVI expression group, n (%)</strong></td>
<td></td>
<td></td>
<td>.05</td>
</tr>
<tr>
<td>High</td>
<td>18 (37)</td>
<td>31 (56)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>31 (63)</td>
<td>24 (44)</td>
<td></td>
</tr>
</tbody>
</table>

*The median expression value was used as the cutoff point.
†$P$ values for categorical variables are from the Fisher exact test; $P$ values for continuous variables are from the Wilcoxon rank-sum test.

†The ELN Favorable Genetic Group comprises patients with mutated CEBPA and those with mutated NPM1 without FLT3-ITD; the ELN Intermediate-I Genetic Group includes patients with CEBPA wild-type who are FLT3-ITD–positive and NPM1-mutated, FLT3-ITD–negative and NPM1 wild-type, or FLT3-ITD–positive and NPM1 wild-type.
miR-3151 expression in the context of ELN Genetic Groups

The ELN Genetic Groups have been shown to be associated with outcome in older CN-AML patients.6 Moreover, recent studies have demonstrated that the addition of selected new molecular markers can further improve prognostication within the ELN Genetic Groups.5,7,34 Therefore, we investigated whether miR-3151 expression levels could improve outcome prediction within those ELN Genetic Groups that include CN-AML. Within the ELN Favorable Genetic Group, there were no differences in outcome between high and low miR-3151–expressing patients. However, in the ELN Intermediate-I Genetic Group, high miR-3151 expression levels identified patients with particularly poor prognosis for all 3 outcome end points. Only 52% of high miR-3151 expressers in this ELN group achieved a CR, compared with 78% of low miR-3151 expressers (P < .001). Likewise, patients with high miR-3151 expression levels had significantly shorter DFS (< .001; Figure 3A) and OS (P < .001; Figure 3B) compared with low miR-3151 expressers. For all 3 end points, the outcome of the low miR-3151 expressers classified in the ELN Intermediate-I Group was similar to that of both high and low miR-3151–expressing patients in the ELN Favorable Genetic Group (Figure 3A-B and supplemental Table 2).

Biologic insights

To gain biologic insights into miR-3151–associated leukemia, we derived a gene-expression signature comparing high versus low miR-3151 expressers. High miR-3151 expresser status was associated with the differential expression of 597 probe sets, representing 374 annotated genes. Of these, 192 probe sets (116 annotated genes) were up-regulated and 405 probe sets (258 annotated genes) were down-regulated (Figure 4A and supplemental Tables 4 and 5).

High miR-3151 expressers exhibited up-regulation of genes previously associated with worse outcome in CN-AML, including the transcriptional coregulator MN1,34 the dominant negative helix-loop-helix protein ID1,37 the miR-3151 host BAALC, and the surface marker CD200.38 In addition, we observed up-regulation of genes encoding several kinases, such as DDR1, which has been shown to be important for cell growth and differentiation by activation of NOTCH1,39 and PRKCE,40 which has been shown to be involved in apoptosis and several cellular signaling pathways.

Among the most down-regulated genes in high miR-3151 expressers was the HOX cofactor MEIS1, which is a key regulator in developmental processes and the absence of which is known to cause disturbances in the colony-forming ability of hematopoietic stem cells,41 and the tumor suppressor PDCD4, which has been shown to contribute to retinoic acid-induced granulocytic differentiation.42 Furthermore, down-regulation of 23 genes encoding different zinc finger proteins (ZNFs), which are known to be involved in transcriptional regulation, was found in high miR-3151 expressers. Of the 258 down-regulated genes, 73 were in silico

---

**Table 2. Multivariable analysis for outcome according to the miR-3151 expression status in older patients with CN-AML**

<table>
<thead>
<tr>
<th>End point</th>
<th>Variables in final models</th>
<th>OR/HR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR*</td>
<td>miR-3151, high vs low</td>
<td>0.56</td>
<td>0.27-1.19</td>
<td>.13</td>
</tr>
<tr>
<td></td>
<td>BAALC, high vs low</td>
<td>0.23</td>
<td>0.10-0.52</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>WBC, continuous, 50-unit increase</td>
<td>0.65</td>
<td>0.48-0.88</td>
<td>.01</td>
</tr>
<tr>
<td>DFS†</td>
<td>miR-3151, high vs low</td>
<td>2.38</td>
<td>1.47-3.85</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>FLT3-TKD, positive vs no TKD</td>
<td>0.28</td>
<td>0.11-0.73</td>
<td>.009</td>
</tr>
<tr>
<td></td>
<td>ERG, high vs low</td>
<td>2.41</td>
<td>1.50-3.86</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>OS‡</td>
<td>miR-3151, high vs low</td>
<td>1.69</td>
<td>1.14-2.50</td>
<td>.009</td>
</tr>
<tr>
<td></td>
<td>ERG, high vs low</td>
<td>1.71</td>
<td>1.16-2.53</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td>DNMT3A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R882 v wild-type</td>
<td>1.66</td>
<td>1.04-2.66</td>
<td>.04</td>
</tr>
<tr>
<td></td>
<td>Non-R882 v wild-type</td>
<td>1.11</td>
<td>0.61-2.01</td>
<td>.74</td>
</tr>
<tr>
<td></td>
<td>BAALC, high vs low</td>
<td>1.93</td>
<td>1.32-2.82</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*CR indicates complete remission; OR, odds ratio; HR, hazard ratio; OS, overall survival; DFS, disease-free survival; OR > 1.0, a higher CR rate for the higher values of the continuous variables and the first category listed for the categorical variables; CR < 1.0, a lower CR rate for the higher values of the continuous variables and the first category listed for the categorical variables; DDS ≥ 1.0, a higher risk for an event for the first category listed for the categorical variables; DDS < 1.0, a lower risk for an event for the first category listed for the categorical variables; and 95% CI, 95% confidence interval.

†Variables considered in the model based on univariable analyses were: miR-3151 expression (high vs low; median cut), BAALC expression (high vs low; median cut), ERG expression (high vs low; median cut), FLT3-ITD (positive vs no ITD), NPM1 (mutated vs wild-type), WT1 (mutated vs wild-type), MLL-PTD (present vs absent), ASXL1 (mutated vs wild-type), RUNX1 (mutated vs wild-type), WBC (continuous, 50-unit increase), platelets (continuous, 50-unit increase), and age (continuous, 10-year increase).

‡Variables considered in the model based on univariable analyses were: miR-3151 expression (high vs low; median cut), BAALC expression (high vs low; median cut), ERG expression (high vs low; median cut), FLT3-ITD (positive vs no ITD), FLT3-TKD (positive vs no TKD), NPM1 (mutated vs wild-type), DNMT3A (R882 vs non-R882 vs wild-type), RUNX1 (mutated vs wild-type), WT1 (mutated vs wild-type), MLL-PTD (present vs absent), ASXL1 (mutated vs wild-type), RUNX1 (mutated vs wild-type), WBC (continuous, 50-unit increase), and platelets (continuous, 50-unit increase).
let-7a, let-7b, and let-7c by participating in many cell-proliferation pathways and the miR-3151 up-regulated and 9 down-regulated in high compared with low (supplemental Table 6).

miR-3151 lar development, and cancer pathways, suggesting an impact of regulation, posttranslational modification, cell-cycle control, cellular function, and its host gene is an illustrative example. It has been shown that the intronic host gene, the established molecular prognostic marker, is a direct target of miR-3151 and is down-regulated in high compared with low.

To further investigate the networking processes of known miRs in which miR-3151 might be involved, we derived miR-expression signatures associated with miR-3151 expression status (Figure 4B). We found 15 differentially expressed probes, representing 14 miRs, 5 up-regulated and 9 down-regulated in high compared with low miR-3151 expressers. Among the down-regulated miRs were let-7a, let-7b, and let-7c, which are known to suppress tumorigenesis by participating in many cell-proliferation pathways and the down-regulation of which has been previously associated with AML leukemogenesis. We also observed down-regulation of miR-10a and miR-10b, which are miRs embedded in HOX gene clusters, and miR-99a and miR-100, the reduced expression of which has been implicated in tumor progression in cervical and prostate cancers.

To gain initial insights into the downstream effects of high miR-3151 expression, we sought to validate 1 or 2 of the down-regulated genes in the miR-3151-associated gene expression signature as a direct target of miR-3151. For the identification of potential candidate genes, we searched among the in silico predicted targets for probe sets of annotated genes that showed at least a 25% down-regulation with P < .0001. Only 6 of the 73 genes fulfilled these criteria (see supplemental Table 5 gray highlights). Among these, we selected as potential candidates those with a described function that was associated with the pathways shown to be preferentially involved (supplemental Table 6).

Strikingly, 2 of the 6 genes, the F-box and leucine-rich repeat protein 20 (FBXL20) and the ubiquitin-specific protease 40 (USP40), are involved in the ubiquitination pathway, suggesting an impact of miR-3151 on this important posttranslational regulatory process. FBXL20 is a member of the F-box gene family. As a part of the SCF (Skp, Cullin, F-box containing) ubiquitin ligase complex, it is responsible for the ubiquitination of proteins, thereby labeling them for consecutive proteasomal degradation. USP40 belongs to the family of cysteine proteases that function as deubiquitinating enzymes.

To validate FBXL20 and USP40 as direct targets of miR-3151, we stably overexpressed miR-3151 in KG1 cells using a lentiviral system. Forced expression of miR-3151 resulted in significant down-regulation of the FBXL20 and USP40 transcripts (FBXL20, 85% decrease, Figure 5A; USP40, 66% decrease, Figure 5B; both P < .001) compared with scramble control. MiR-3151 also down-regulated FBXL20 on the protein level (Figure 5D). However, none of the commercially available USP40 Abs showed a band at the predicted protein size of 140 kDa. To demonstrate that FBXL20 and USP40 are direct targets of miR-3151, the respective 3′-untranslated regions of both genes with the sequence containing the predicted miR-3151–binding sites were cloned into luciferase reporter vectors. The luciferase assays demonstrated a 54% and 33% decrease in luciferase activity for the FBXL20 and USP40 constructs, respectively, after addition of miR-3151 compared with scramble control (Figure 5E-F). The observed down-regulations were abrogated after mutation of the seed sequence of the predicted miR-3151–binding sites (Figure 5E-F). These results demonstrate that FBXL20 and USP40 are direct targets of miR-3151.

**Discussion**

In the present study, we report an intronic miR as an independent prognostic factor for outcome in older CN-AML patients. Furthermore, our results suggest that miR-3151 may act in concert with its host gene, the established molecular prognostic marker BAALC.

In considering previous findings of a similar nature, the interplay of miR-33 and its SREBP host gene is an illustrative example. It has been shown that the intronic miR-33 targets the ATP-binding cassette transporter A1 (ABCA1), an important regulator of high-density lipoprotein synthesis and reverse cholesterol transport, thereby acting in synergism with SREBP to control cholesterol homeostasis. Recently, cooperation between intronic miRs and their host genes has been also demonstrated in prostate cancer and hepatocellular carcinoma. However, to our knowledge, similar mechanisms have not been reported so far in AML. Even though it was not the aim of our study to determine a functional mechanism of miR-3151 and BAALC interaction, our data suggest that both markers contribute independently to poor outcome in CN-AML patients.

We also found that, like BAALC, higher expression levels of miR-3151 were associated with poor prognosis. However, the hosted miR and the hosting gene have an independent clinical significance and had different effects on specific outcome endpoints. Whereas miR-3151 expression did not remain an independent predictor in the multivariable model for achievement of CR, BAALC expression remained a strong prognostic factor for CR. This is consistent with the findings of our group and others that CR
rate is an outcome end point that is strongly affected by aberrant BAALC expression levels. Conversely, high miR-3151 expression and not BAALC expression was a strong predictor of DFS. Finally, both miR-3151 and BAALC expression levels remained important prognostic factors for OS.

We also found that patients overexpressing both miR-3151 and its host gene BAALC had particularly poor outcome for all outcome end points. In contrast, patients exhibiting up-regulation of only 1 of the 2 markers had a significantly better outcome; their DFS was comparable to that of low miR-3151/low BAALC-expressing patients and their OS was intermediate in comparison with OS of both groups of patients who had concordant miR-3151 and BAALC exprasser status. These findings suggest that miR-3151 and BAALC act independently to affect outcome, thereby possibly creating a synergism to support leukemogenesis.

In the present study, by deriving a GEP signature comparing high and low miR-3151–expressing patients, we were able to gain initial insights into the biology and possible downstream effects of miR-3151 in CN-AML patients. We showed that high miR-3151-expressing patients also had up-regulation of genes that are known prognosticators of worse outcome in AML patients, such as MN1, ID1, and CD200 and the miR-3151 host gene BAALC. Among the down-regulated genes associated with high miR-3151 expression, we found several that were implicated in hematopoietic differentiation, including MEIS1, the absence of which has been shown to
ences in the biologic activities of patients. Indeed, we were able to validate consequences seen in high for the downstream effects, including the pathophysiological and posttranscriptional regulatory processes. Even though cell-cycle control, cell growth, and a multitude of transcriptional ubiquitination pathway, which is known to be important for direct targets of miR-3151.

Regarding the derived MEP signature, we found 14 miRs to be differentially expressed between high and low miR-3151 expressers. Among these miRs, a key component was down-regulation of members of the let-7 family, which are known tumor suppressors and the down-regulation of which is found in various types of cancer, linking the let-7 family members to the HMGA2 and RAS pathways.

Comparing the derived MEP signature with the signature described to be associated with BAALC expression, we observed interesting similarities but also differences. High BAALC-expressing patients exhibited down-regulation of miR-99a, miR-100, and let-7b, whereas down-regulation of let-7a and let-7c seemed to be uniquely associated with high miR-3151 expression levels. Down-regulation of miR-9 was only observed in the signature associated with high BAALC expression levels. No up-regulated miR was shared by both MEP signatures.

The fact that we were able to identify miR-3151 as the second miR after miR-181a that independently affects outcome of CN-AML patients provides further support to the importance of aberrantly expressed miRs as prognostic factors in CN-AML.

It is noteworthy that miR-3151 is not simply a new molecular marker, but to our knowledge is the first example of how known factors (here BAALC) in AML leukemogenesis may be supported by miRs located in the locus itself. In the case of the BAALC gene, this finding is of special interest because the gene’s function and therefore the reasons for its strong prognostic impact in CN-AML patients are unknown.

In conclusion, we have shown in the present study that high expression of miR-3151 is an independent prognostic factor associated with poor outcome in older CN-AML patients.
partly discordant expresser status of miR-3151 and its host gene BAALC, the independent impact of the 2 genes on outcome, and the fact that patients overexpressing both miR-3151 and its host gene have significantly worse outcomes than those exhibiting up-regulation of only 1 of the genes suggest that the 2 genes contribute to the aggressiveness of the disease through different mechanisms. We conclude that determining the expression levels of miR-3151 at diagnosis will help to improve the risk stratification of older CN-AML patients. The development of therapies targeting miR-3151 up-regulation with synthetic inhibitors may provide novel, more effective strategies for personalized treatment of these patients.

Acknowledgments

The authors thank Donna Bucci of the CALGB Leukemia Tissue Bank at The Ohio State University Comprehensive Cancer Center (Columbus, OH) for sample processing and storage services; Lisa J. Sterling and Colin G. Edwards (CALGB Cytogenetics Data Management Center at The Ohio State University Comprehensive Cancer Center) for data management; and The Ohio State University Comprehensive Cancer Center’s Nucleic Acid and Microarray Shared Resources for technical support.

This work was supported in part by the National Cancer Institute (grants CA101140, CA114725, CA140158, CA31946, CA33601, CA16058, CA77658, and CA129657), the Coleman Leukemia Research Foundation, the Deutsche Krebshilfe–Dr Mildred Scheel Cancer Foundation (H.B.), the Pelotonia Fellowship Program (A.-K.E.), and the Conquer Cancer Foundation (J.H.M.).

Authorship

Contribution: A.-K.E., G.M., K. Mrózek, S.M.T., A.d.l.C., and C.D.B. designed the study, analyzed the data, and wrote the manuscript; A.-K.E., S.S., H.B., S.P.W., K.H.M., J.H.M., Y.-Z.W., and R.P. performed the laboratory-based research; K. Maharry, M.D.R., D.N., and S.L. performed the statistical analyses; G.M., M.R.B., B.L.P., T.H.C., J.O.M., J.E.K., M.W., M.A.C., R.A.L., and C.D.B. were involved directly or indirectly in the care of patients and/or sample procurement; and all authors read and agreed upon the final version of the manuscript.

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

For a complete list of participating Alliance institutions, principal investigators, and cytogeneticists, please see the supplemental Appendix.

Correspondence: Clara D. Bloomfield, MD, The Ohio State University Comprehensive Cancer Center, 1216 James Cancer Hospital, 300 West 10th Ave, Columbus, OH 43210; e-mail: clara.bloomfield@osumc.edu.

References


miR-3151 interplays with its host gene BAALC and independently affects outcome of patients with cytogenetically normal acute myeloid leukemia