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

Dissertation
zur Erlangung des akademischen Grades
Dr. med.
an der Medizinischen Fakultät
der Universität Leipzig

eingereicht von: Ann-Kathrin Eisfeld


angefertigt an:

Universität Leipzig
Department für Innere Medizin, Neurologie und Dermatologie
Selbständige Abteilung für Hämatologie, internistische Onkologie und Hämostaseologie

und dem

Comprehensive Cancer Center,
Division of Hematology, Department of Medicine
The Ohio State University
Columbus, OH, USA

Betreuer: Prof. Dr. med. Dr. h.c. Dietger Niederwieser

Beschluss über die Verleihung des Doktorgrades vom 29.04.2014
### Inhaltsverzeichnis / Table of Contents:

Bibliografische Beschreibung .................................................................................................................. 1  
Referat / Abstract ....................................................................................................................................... 2  
Publikation / Publication ............................................................................................................................. 6  
Zusammenfassung / Conclusion ................................................................................................................... 16  
Referenz der Publikation / Reference of the publication ........................................................................ 27  
Komplette Publikationsliste / Complete List of Publications .................................................................. 28  
Lebenslauf / Curriculum Vitae ................................................................................................................... 32  
Erklärung über die eigenständige Abfassung der Arbeit ........................................................................... 36  
Danksagung / Acknowledgements ............................................................................................................ 37
Bibliografische Beschreibung

Ann-Kathrin Eisfeld

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

Universität Leipzig, Dissertation

37 Seiten, 46 Literaturangaben, 3 Abbildungen
Referat / Abstract

High expression levels of the gene *BAALC* (brain and acute leukemia, cytoplasmic) are associated with poor prognosis in acute myeloid leukemia (AML) patients, but the underlying mechanisms are not yet understood. We evaluated the prognostic significance of expression levels of *miR-3151*, a newly discovered microRNA embedded in intron 1 of the *BAALC* gene, in a cohort of 179 older (≥60 years) cytogenetically normal AML (CN-AML) patients, in the context of established molecular markers and especially with regard to the possible interplay with its host gene *BAALC*. In multivariable analyses, high *miR-3151* was associated with shorter disease-free and overall survival (OS), while higher *BAALC* expression strongly predicted failure of complete remission attainment and OS. Patients exhibiting both high *miR-3151* and *BAALC* expression had worse outcome than patients expressing low levels of either one of the genes or both. Next, gene- and microRNA-expression profiles associated with *miR-3151* expression were derived using microarrays, and a pathway analysis of the *miR-3151* associated gene signature was performed using Ingenuity software. High *miR-3151* expressers showed downregulation of genes involved in transcriptional regulation, post-translational modifications and cell-cycle control. Two genes of the ubiquitination pathway, *FBXL20* and *USP40*, were experimentally validated as direct *miR-3151* targets. In summary, we identified high expression levels of the intronic *miR-3151* as a novel, independent prognosticator for poor outcome in CN-AML. Interestingly, *miR-3151* impacted differently on outcome than its host gene *BAALC*; and the combination of both markers identified a patient subset with the poorest outcome, suggesting that the microRNA and its host gene contribute to clinical and prognostic features of CN-AML independently and through distinct mechanisms. This is the first example of the interplay of an intronic miR and its host gene in leukemia. Its discovery may have important biologic implications for future targeted treatment strategies.
Acute Myeloid Leukemia (AML) is a complex disease, which is characterized by the uncontrolled proliferation of a leukemic clone in the bone marrow. In recent years, new insights into the biology have significantly improved our understanding of AML. But, despite this progress, the outcome of the majority of AML patients still remains poor and only a fraction of the patients achieves long term survival.\textsuperscript{1-3} The outcome is especially poor in patients aged 60 years or older, of whom only $\sim$15\% achieve a long term survival.\textsuperscript{1-3}

Very soon it became clear, that the cytogenetical and molecular heterogeneity may be the reason for the different outcomes of the patient, since each chromosomal and/or molecular abnormality in the leukemic clone may lead to a more or less aggressive disease course. Today, numerous clinical, cytogenetic and molecular variables have been found to be associated with AML outcome.\textsuperscript{4-16} Their discoveries raised hopes to better predict the individual chances of treatment response and potentially lead to more personalized, risk-adapted therapeutic strategies for AML patients. Among the prognostic markers, mutations in the (nucleolar phosphoprotein B23, numatin) \textit{(NPM1)}\textsuperscript{15} and CCAAT/enhancer binding protein C/EBP, alpha \textit{(CEBPA)}\textsuperscript{8} genes are associated with favorable outcome and have been included as provisional entities in the World Health Organization classification of AML.\textsuperscript{17} These two mutations as well as the presence of internal tandem duplications of the fms-related tyrosine kinase 3 gene \textit{(FLT3-ITD)}\textsuperscript{5} 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),\textsuperscript{18} and should be determined in every AML patient.

In addition to the presence of gene mutations, the differential expression of specific genes has also been proven to be of prognostic significance in AML. Examples for those gene expression markers are, among others, the v-ets erythroblastosis virus E26 oncogene homolog (avian) \textit{(ERG)},\textsuperscript{19} menigioma (disrupted in balanced translocation) 1 \textit{(MN1)},\textsuperscript{16} the dominant negative helix-loop-helix protein \textit{ID1},\textsuperscript{20} and the surface marker \textit{CD200}.\textsuperscript{7} It is important to know, that the underlying (leukemogenic) mechanisms for
some of these markers are known, while for others their connection to leukemia pathogenesis has not been discovered yet.

Another gene, whose high expression levels have been shown to associate with poor prognosis in AML patients is the \textit{BAALC} (brain and acute leukemia cytoplasmic) gene, which was identified by cDNA-based representational difference analysis in leukemia patients.\textsuperscript{21} The prognostic impact of \textit{BAALC} expression has been most extensively studied in AML patients with a normal karyotype (CN-AML),\textsuperscript{19,22-24} with high expression levels being associated with a lower complete remission (CR) rate, as well as shorter disease-free (DFS) and overall survival (OS) of the patients.\textsuperscript{19,22} Although recent data suggest that \textit{BAALC} contributes to leukemogenesis by interfering with normal patterns of myeloid differentiation,\textsuperscript{23} the function of this gene and the mechanism(s) through which its high expression levels impact negatively on clinical outcome remain(s) unknown. Taking together the known published data on \textit{BAALC}’s impact on outcome of AML patients, deciphering \textit{BAALC}’s function may be of high interest to the scientific community. But, since multiple attempts of various study groups failed so far, new avenues and research approaches may have to be taken to gain more insights into the leukemogenic properties of the \textit{BAALC} locus.

MicroRNAs (miRs) are noncoding RNAs that downregulate gene expression by inhibiting translation or promoting mRNA degradation.\textsuperscript{26} Their targets can be various and may vary by tissue- and cell type; potentially providing simultaneous regulation of multiple genes to regulate a specific pathway. They are not only involved in such biological processes as cellular differentiation, proliferation and survival but also play an essential role in the development of solid tumors and AML.\textsuperscript{26-30} miR genes can be found throughout the genome, but about one third of mammalian miRs are located within introns of host genes,\textsuperscript{31,32} and most of them are believed to be co-expressed and processed from the same precursor mRNA in which they reside,\textsuperscript{33-35} while approximately 26\% of intronic miRs have their own promoters.\textsuperscript{36,37} The functional relationship of miR genes and their hosts is vastly unknown. Most intronic miRs are thought to strengthen the function of its host gene, which may happen in a direct manner by targeting genes within the same cellular pathway,\textsuperscript{38,39} while other miRs have
been shown to silence genes which are acting antagonistic to their hosts. Additionally, it could be shown that some miRs may even be made accountable for effects which have been thought to be caused by their host gene and that they can be achieved independent of the expression of their host gene.

Recently, deep sequencing of melanoma and acute lymphoblastic leukemia samples led to the discovery of a novel miR, miR-3151, which was found to be embedded in intron 1 of BAALC.

We therefore hypothesized, that miR-3151 might also be aberrantly expressed in patients with high expression levels of its host gene BAALC and might also – dependently or independently - impact on outcome of AML patients. Furthermore, since the function of BAALC is still unknown, we hypothesized that miR-3151 may either be the crucial partner needed for an unknown leukemogenic function of BAALC, or even be the element predominantly responsible for the adverse outcome of patients with high BAALC expression levels. Thus, miR-3151 may be a novel oncomiR.

In our here presented study, we measured the expression levels of miR-3151 and its host gene BAALC in a set of 179 CN-AML patients aged 60 years and older by RT Real-Time PCR. To compare the prognostic impact of miR-3151 expression levels to the impact of other well-established molecular markers, we additionally analyzed all patients for the presence or absence of FLT3-ITD, FLT3 tyrosine kinase domain mutations (FLT3-TKD), partial tandem duplication of the myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila) gene (MLL-PTD), mutations in the NPM1, CEBPA, tet methylcytosine dioxygenase 2 (TET2), additional sex combs like 1 (Drosophila) (ASXL1), DNA (cytosine-5-)methyltransferase 3 alpha (DNMT3A), runt-related transcription factor 1 (RUNX1), Wilms tumor 1 (WT1), and isocitrate dehydrogenase 1 (NADP+), soluble (IDH1) and isocitrate dehydrogenase 2 (NADP+), mitochondrial (IDH2) genes, and expression levels of ERG and MN1. Additionally, gene- and microRNA expression profiling was performed with the aim to analyze the derived gene expression signatures for enrichments of specific pathways and/or biologic functions. Finally, we aimed to validate two of the most important genes of the signature as direct targets of miR-3151, thereby giving first insights into the downstream biology of miR-3151.
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 Nicole1,1,2 Heiko Becker,1 Krzysztof Mrózek,1 Susan P. Whitman,1 Klaas H. Metzeler,1 Jason H. Mendler,1 Yue-Zhong Wu,1 Sandya Liyanarachchi,1 Ravi Patel,1 María R. Baer,3 Bayard L. Powell,4 Thomas H. Carter,4 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

1The Ohio State University Comprehensive Cancer Center, Columbus, OH; 2Alliance for Clinical Trials in Oncology Statistics and Data Center, Mayo Clinic, Rochester, MN; 3Greenbaum Cancer Center, University of Maryland, Baltimore, MD; 4Comprehensive Cancer Center of Wake Forest University, Winston-Salem, NC; 5University of Iowa, Iowa City, IA; 6Duke University Medical Center, Durham, NC; 7Montefiore Cancer Center, Hofstra North Shore–Long Island Jewish School of Medicine, Lake Success, NY; and 8Roswell Park Cancer Institute, Buffalo, NY; and 9University of Chicago Medical Center, Chicago, IL

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 expressers 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.4,5 Among them, mutations in the nucleoplasm (nucleolar phosphoprotein 223, name: 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.6,7 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).8

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.9 The prognostic impact of BAALC expression has been most extensively studied in CN-AML patients.10,11 Although recent data suggest that BAALC contributes to leukemogenesis by interfering with normal patterns of myeloid differentiation,12 the function of this gene and the mechanisms through which its high expression levels affect clinical outcome negatively remain(s) unknown.

Recently, small RNA deep sequencing of melanoma13 and pediatric acute lymphoblastic leukemia samples14 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 miRNAs15 and have been implicated in
leukemogenesis. Furthermore, expression levels of miRs may be used to refine risk assessment in AML in CN-AML patients. 24,25

miR genes are located throughout the genome, but approximately one-third of mammalian miRs reside within the introns of host genes. 24,25 Some of these intronic miRs have been found to act in functional synergism with their host genes. 26 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. 27 The diagnosis of normal cytogenetics was based on the analysis of ≥20 BM metaphase cells and confirmed by central karyotype review. 28 All patients gave informed consent for the research use of their specimens in accordance with the Declaration of Helsinki.

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 mir-3151 and the respective housekeeping genes JSR and KNU44 (Life Technologies Corporation/Applied Biosystems) according to protocol instructions (see supplemental Methods).

Additional molecular markers were analyzed centrally in pretreatment BM or peripheral blood samples as described previously and included: FLL3-ITD, 29 FLT3 tyrosine kinase domain mutations (FLT3-ITD/TKD), 30 partial tandem duplication of the myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila) gene (MLL-PTD), 31 mutations in the NPM1, 27 CEBPA, 27 tet methylcytosine dioxygenase 2 (TET2), 32 additional sex combs 1 (Drosophila, ASXL1), 32 DNA (cytosine-5-) methyltransferase 3 alpha (DNMT3A), 33 run-related transcription factor 1 (RUNX1), 34 Wilms tumor 1 (WT1), 34 and isocitrate dehydrogenase 1 (NADP+), 35 soluble (IDH1) and isocitrate dehydrogenase 2 (NADP+), 35 mitochondrial (IDH2), 35 genes, and expression levels of the c-kit erythroblastosis virus E26 oncogene homolog (avian; ERG) 36 and meningeoma (disrupted in balanced translocation) 1 (MEN1) 37 genes.

GEP 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). 32,33 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 or mutated mir-3151-binding sequence, see supplemental Table 2 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 (ie, 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 $P = .001$ (GEP) 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-1 Genetic Group, 11 and have mutations of RUNXI ($P < .001$) 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 MNI ($P = .05$). 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></td>
<td></td>
<td>.59</td>
</tr>
<tr>
<td>Median</td>
<td>68</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>60-79</td>
<td>60-81</td>
<td></td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td>.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>.10</td>
</tr>
<tr>
<td>White</td>
<td>73 (85)</td>
<td>83 (93)</td>
<td></td>
</tr>
<tr>
<td>Nonwhite</td>
<td>11 (15)</td>
<td>4 (5)</td>
<td></td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td></td>
<td></td>
<td>.47</td>
</tr>
<tr>
<td>Median</td>
<td>9.3</td>
<td>9.6</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>.06</td>
</tr>
<tr>
<td>Median</td>
<td>60</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>4-491</td>
<td>11-850</td>
<td></td>
</tr>
<tr>
<td>WBC count, × 10^9/L</td>
<td></td>
<td></td>
<td>.10</td>
</tr>
<tr>
<td>Median</td>
<td>26.2</td>
<td>31.7</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>1.4-400.0</td>
<td>1.1-1244.1</td>
<td></td>
</tr>
<tr>
<td>Percentage of blood blasts</td>
<td></td>
<td></td>
<td>.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>.02</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></td>
<td></td>
<td>.73</td>
</tr>
<tr>
<td>Mutated</td>
<td>21 (24)</td>
<td>24 (28)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>79 (73)</td>
<td>43 (49)</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>.63</td>
</tr>
<tr>
<td>Mutated</td>
<td>12 (13)</td>
<td>13 (15)</td>
<td></td>
</tr>
<tr>
<td>Single mutated</td>
<td>6 (9)</td>
<td>3 (4)</td>
<td></td>
</tr>
<tr>
<td>Double mutated</td>
<td>6 (2)</td>
<td>9 (10)</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>.05</td>
</tr>
<tr>
<td>Favorable</td>
<td>50 (55)</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;.001</td>
</tr>
<tr>
<td>Mutated</td>
<td>4 (5)</td>
<td>19 (24)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>80 (90)</td>
<td>59 (70)</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>70 (82)</td>
<td></td>
</tr>
<tr>
<td>WT1, n (%)</td>
<td></td>
<td></td>
<td>.06</td>
</tr>
<tr>
<td>Mutated</td>
<td>33 (33)</td>
<td>9 (10)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>87 (90)</td>
<td>78 (90)</td>
<td></td>
</tr>
<tr>
<td>TET2, n (%)</td>
<td></td>
<td></td>
<td>.33</td>
</tr>
<tr>
<td>Mutated</td>
<td>24 (27)</td>
<td>30 (33)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>65 (73)</td>
<td>56 (63)</td>
<td></td>
</tr>
<tr>
<td>MLL-PTD, n (%)</td>
<td></td>
<td></td>
<td>.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>V711 mutated</td>
<td>0 (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>.71</td>
</tr>
<tr>
<td>R140 mutated</td>
<td>20 (22)</td>
<td>17 (20)</td>
<td></td>
</tr>
<tr>
<td>V172 mutated</td>
<td>30 (33)</td>
<td>14 (15)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>65 (73)</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.
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†</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXIN1, n (%)</td>
<td></td>
<td></td>
<td>.13</td>
</tr>
<tr>
<td>Mutated</td>
<td>9 (10)</td>
<td>16 (19)</td>
<td></td>
</tr>
<tr>
<td>Wild-type</td>
<td>79 (90)</td>
<td>69 (81)</td>
<td></td>
</tr>
<tr>
<td>DNMT3A, n (%)</td>
<td></td>
<td></td>
<td>.74 (mut vs wt)</td>
</tr>
<tr>
<td>Mutated</td>
<td>27 (31)</td>
<td>29 (35)</td>
<td></td>
</tr>
<tr>
<td>R882</td>
<td>15</td>
<td>19</td>
<td>.56 (R882 vs wt)</td>
</tr>
<tr>
<td>Non-R882</td>
<td>12</td>
<td>10</td>
<td>1.00 (non-R882 vs wt)</td>
</tr>
<tr>
<td>Wild-type</td>
<td>59 (69)</td>
<td>55 (65)</td>
<td></td>
</tr>
<tr>
<td>ERG expression group, n (%)‡</td>
<td>42 (56)</td>
<td>35 (50)</td>
<td>.40</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>31 (42)</td>
<td>35 (50)</td>
<td></td>
</tr>
<tr>
<td>BAALC expression group, n (%)‡</td>
<td>34 (39)</td>
<td>36 (66)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>54 (61)</td>
<td>29 (34)</td>
<td></td>
</tr>
<tr>
<td>MVI expression group, n (%)‡</td>
<td>18 (37)</td>
<td>31 (56)</td>
<td>.06</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></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 Genomic Group comprises patients with mutated CEBPA and those with mutated NPM1 without FLT3–ITD; the ELN Intermediate-I Genomic 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.

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 a significant adverse impact on CR (miR-3151, P < .001, OR = 4.47; 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 expression 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,
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. Moreover, recent studies have demonstrated that the addition of selected new molecular markers can further improve prognosis within the ELN Genetic Groups. 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 MINT, the dominant negative helix-loop-helix protein ID1, the miR-3151 host BAALC, and the surface marker CD200. 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, and PRKCE, which has been shown to be involved in apoptosis and cellular signaling pathways.

Among the most down-regulated genes in high miR-3151 expressers was the HOX cofactor MESP1, 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, and the tumor suppressor PDCD4, which has been shown to contribute to retinoic acid-induced granulocytic differentiation. 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>
<thead>
<tr>
<th>End point</th>
<th>Variables in final model(s)</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>BAALC, high vs low</td>
<td>0.23</td>
<td>0.10-0.52</td>
<td>&lt; .001</td>
<td></td>
</tr>
<tr>
<td>WBC: continuous, 50-unit increase</td>
<td>0.65</td>
<td>0.48-0.88</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>DFS</td>
<td>miR-3151, high vs low</td>
<td>2.30</td>
<td>1.47-3.65</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>FLT3-TKD, positive vs no TKD</td>
<td>0.28</td>
<td>0.11-0.73</td>
<td>.009</td>
<td></td>
</tr>
<tr>
<td>ERG, high vs low</td>
<td>2.41</td>
<td>1.50-3.85</td>
<td>&lt; .001</td>
<td></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>ERG, high vs low</td>
<td>1.71</td>
<td>1.16-2.53</td>
<td>.007</td>
<td></td>
</tr>
<tr>
<td>DNMT3A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBB2 vs wild-type</td>
<td>1.66</td>
<td>1.04-2.66</td>
<td>.04</td>
<td></td>
</tr>
<tr>
<td>Non-RBB2 vs wild-type</td>
<td>1.11</td>
<td>0.61-2.01</td>
<td>.74</td>
<td></td>
</tr>
<tr>
<td>BAALC, high vs low</td>
<td>1.93</td>
<td>1.32-2.62</td>
<td>&lt; .001</td>
<td></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 OR rate for the higher values of the continuous variables and the first category listed for the categorical variables; OR < 1.0, a lower OR rate for the higher values of the continuous variables and the first category listed for the categorical variables; HR > 1.0, a higher 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 univariate 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-TKD (positive vs no TKD), NPM1 (mutated vs wild-type), WTC1 (mutated vs wild-type), DLL3 (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 univariate 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-TKD (positive vs no TKD), NPM1 (mutated vs wild-type), DNMT3A (RBB2 vs Non-RBB2 vs wild-type), RUNX1 (mutated vs wild-type), WT1 (mutated vs wild-type), and WBC (continuous, 50-unit increase).

Variables considered in the model based on univariate 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-TKD (positive vs no TKD), NPM1 (mutated vs wild-type), ASXL1 (mutated vs wild-type), RUNX1 (mutated vs wild-type), WT1 (mutated vs wild-type), DNMT3A (RBB2 vs Non-RBB2 vs wild-type), platelets (continuous, 50-unit increase), and WBC (continuous, 50-unit increase).
predicted targets of miR-3151 (www.micromat.org; supplemental Table 5).

Pathway analysis of the miR-3151-associated expression signature showed an enrichment of genes involved in transcriptional regulation, posttranslational modification, cell-cycle control, cellular development, and cancer pathways, suggesting an impact of miR-3151 on basic regulatory functions (http://humanity.com; supplemental Table 6).

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 \( P < 0.001 \). 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 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 host gene A1 (ABCA1), an important regulator of high-density lipoprotein synthesis and reverse cholesterol transport, thereby acting in synergy 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.\textsuperscript{13,15,16} 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 expresser 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
interfere with the normal development of hematopoietic precursors. In addition, we observed a down-regulation of multiple genes encoding ZNF proteins, suggesting a role for miR-3151 in pretranscriptional regulation.

Interestingly, pathway analysis of the miR-3151–associated gene expression signature suggested an involvement of miR-3151 in general regulatory processes on both the transcriptional and posttranslational levels. Identification of direct target genes of miR-3151 belonging to these pathways may pinpoint a root cause for the downstream effects, including the pathophysiological consequences seen in high miR-3151–expressing CN-AML patients. Indeed, we were able to validate FBXL20 and USP40 as direct targets of miR-3151. Both genes are involved in the ubiquitination pathway, which is known to be important for cell-cycle control, cell growth, and a multitude of transcriptional and posttranscriptional regulatory processes. Even though FBXL20 and USP40 are likely only the first of several important miR-3151 targets, their involvement may initiate and direct future research into our understanding of the function of miR-3151.

The miR-3151–associated GEP signature shared some features with a recently described GEP signature associated with BAALC expression in older CN-AML patients, in which we also reported an up-regulation of MYF1 and CD200 and down-regulation of MEIS1. However, changes in the expression of other genes were unique to the GEP signature associated with miR-3151 (eg, up-regulation of ID1 and down-regulation of ZNFs). Moreover, key components of the BAALC signature (eg, up-regulation of CD34 and PROM1 and down-regulation of several HOX gene clusters) were not found in our miR-3151–associated signature, suggesting important differences in the biologic activities of miR-3151 and BAALC.

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 HMG2 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 expression 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 Bacci 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, CA33146, CA33661, CA16058, CA77658, and CA129577), 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. Mrzoke, S.M.T., A.d.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., 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


Zusammenfassung / Conclusion

Since the here presented study evolved out of the aim to better understand the biology and function of BAALC, three main conclusions may be drawn.

1) We could, for the first time, demonstrate an independent prognostic impact of BAALC’s intronic microRNA miR-3151. In addition, we could validate the important prognostic impact of BAALC expression in older CN-AML patients, especially on the achievement of CR and OS.

2) Both genes impacted on different outcome endpoints and showed also differences in their associated gene-expression signatures. Thus, neither does miR-3151 explain -in full - the impact of BAALC expression levels (BAALC is not only a meaningless cover gene), nor does BAALC account for the impact of miR-3151 expression levels.

3) BAALC and miR-3151 may have to be interpreted as an oncogenic locus, which interplays towards leukemogenesis.

Within the scope of this thesis we present a study that reported miR-3151 as an independent prognostic factor for outcome in older patients with CN-AML. The impact of miR-3151 expression on survival probability was independent from other, well-established molecular markers (eg, NPM1 mutations and FLT3-ITD) and notably also independent from the expression status of its host gene BAALC. Importantly, higher expression of miR-3151 impacted on different outcome endpoints than BAALC: The effect of BAALC was mainly on achievement of CR while that of miR-3151 was mainly on outcome- once CR had been achieved (DFS); naturally then both BAALC and miR-3151 affected OS in these older patients. Patients with high expression levels of both genes were identified as the patient subset with the poorest outcome. In general, patients who had high expression of both miR-3151 and BAALC did worst, patients with low expression of both genes did best, and patients with discordant expresser status had intermediate outcomes.

The publication of a standardized reporting system for genetic abnormalities suggested by an international expert panel on behalf of the European LeukemiaNet (ELN) enabled
clinicians with a tool to stratify AML patients based on both cytogenetic abnormalities and the presence or absence of the mutational status of the established molecular markers *NPM1*, *CEBPA* and *FLT3*-ITD. Therefore, it was important to test the impact of miR-3151 expression levels within those ELN Genetic Groups. Interestingly, miR-3151 expression was able to identify a subset with an especially poor outcome within the ELN Intermediate-I Genetic Group. Notably, the DFS and OS rates of patients with low miR-3151 expression levels were comparable to the survival rates of patients belonging to the ELN Favorable Genetic Group. In contrast, miR-3151 expression levels were without influence on the outcome of patients belonging to the ELN Favorable Genetic Group. Thus, determination of miR-3151 expression levels may be a useful tool to refine the risk stratification proposed in the guidelines of the ELN. But, even knowing about the important prognostic information which may be gained by the determination of miR-3151 and BAALC expression levels in the pre-treatment peripheral blood of AML patients, it has to be kept in mind that the clinical use of those expression markers is so far still limited by the difficulties to standardize the methods for gene- and microRNA expression determination for individual patients and by the lack of absolute cut-offs to undoubtedly define high and low expressing patients. Once those technical difficulties are overcome, the usage of gene expression markers may be similar or even better than the one gained by the mutational analysis, since the expression markers may reflect the combined downstream effects of a plethora of molecular features, such as different gene mutations or epigenetic events. Besides the clinical observations, our aim was to gain first insights into the downstream biology of miR-3151-associated leukemia. Therefore, we performed a global gene-expression profiling in our patient cohort and were able to derive a distinct gene-expression signature associated with high miR-3151 expression levels. Using ingenuity software, we next performed a detailed gene ontology analysis to identify genes and subsequent pathways which are enriched in high miR-3151 expressing patients. This gene ontology analysis revealed pathways regulating gene expression, cell-cycle control and post-transcriptional regulation to be primarily affected by high miR-3151 expression levels. As a next step, we aimed to validate two of the most downregulated genes of the miR-3151-associated signature as direct miR-3151 targets. For identification of these genes, we used a significance-based
algorithm. Interestingly, 2 out of the 6 genes which fulfilled our selection criteria belonged to the ubiquitination pathway\textsuperscript{44} - which goes in line with our previous observation that \textit{miR-3151} impacts on post-transcriptional regulation pathways. Both genes, \textit{FBXL20}\textsuperscript{45} and \textit{USP40}\textsuperscript{46} could be validated as direct \textit{miR-3151} target genes. FBXL20 belongs to the F-BOX proteins and is - as a part of the SCF (Skp, Cullin, F-box containing) - complex - responsible for the ubiquitination and consecutive degradation of its target genes.\textsuperscript{45} The SCF-complex is a multiprotein complex catalyzing the ubiquitination of proteins destined for degradation, in which the F-box proteins (eg, FBXL20) contribute to the specificity of the complex, since they are the part which is binding to the target proteins. This protein- “capture” then allows the ubiquitin ligase (eg, E2) to attach the ubiquitin label (Figure 1).

![Figure 1. The SCF-(Skp, Cullin, F-box containing) complex.](image)

Figure adapted from: Morgan, David “Protein Degradation in Cell-Cycle Control”, The Cell Cycle; Principles of Control 2007

So far, nothing is known about specific FBXL20 targets. However, the identification of such genes may be an interesting topic of future research to further understand the downstream biology of high \textit{miR-3151} expression levels and to potentially find new links in the complex networking of AML signaling pathways. Since we already have a successfully working antibody for FBXL20, which has been used for Western blotting in the publication, a next step may be the immunoprecipitation and consecutive mass-spectrometric analysis to identify the proteins bound to FBXL20.

The comparison of the \textit{miR-3151}-associated gene- and microRNA-expression signatures with the gene- and microRNA-expression signatures previously reported to be associated with high expression levels of the host gene \textit{BAALC} led to additional
important observations. Despite the strong association of the expression levels of both
genres, interesting differences in the differentially expressed genes in both signatures
could be observed. While high BAALC expression was associated with high expression
of the stemness-markers PROM1 and CD34, no such association could be observed in
high miR-3151 expressing patients. In contrast, downregulation of the let-7 family, which
was the key-feature of the miR-3151-associated microRNA-expression signature, has
not been observed in high BAALC expressing patients.
Interestingly, in one of the initial studies of BAALC, Baldus et al. suggested that high
BAALC expression may lose its prognostic importance in patients undergoing allogeneic
hematopoietic stem cell transplantation (HCT). Thus, as a pilot study, we measured
pre-treatment BAALC and miR-3151 expression levels in a total of 75 older AML
patients with mixed karyotypes, which underwent allogeneic HCT at the University of
Leipzig. In these patients we also assessed the NPM1 mutation status. Similarly to the
observations in the patients from the Cancer and Leukemia Group B (CALGB) study
(presented study) we observed an association of high BAALC and high miR-3151
expresser status – using a median cut for both genes to define high and low expressers
- with NPM1 wild-type (P=0.0024 and P=0.013, respectively). Furthermore, again
validating our findings of the CALGB patient set, we observed an association between
high BAALC and high miR-3151 expresser status in the Leipzig patient set (P=.04).
However, about 40% (n=30) of the patients had a discordant expresser status of the two
genres. For a subset of the investigated Leipzig patients (n=22) follow-up data for a
preliminary outcome analysis was available. Utilizing the Kaplan-Meier Method, we
analyzed the OS according to the BAALC and miR-3151 expresser status. In this
analysis, both miR-3151 and BAALC failed as a single marker to significantly separate
the patients according to their overall survival (Figure 2A and B).
However, when following the proposal to analyze the combined effect of the BAALC/miR-3151 locus, we observed a significantly shorter OS (P=.05; log-rank test) in patients with high expression of BAALC, miR-3151 or both when compared to patients expressing both genes on low levels (Figure 3).

![Figure 2](image1.png)

**Figure 2:** (A) OS according to BAALC expression, using the median cut to define high and low expressers treated at the University of Leipzig (n=22). (B) OS according to miR-3151 expression, using the median cut to define high and low expressers in AML patients treated at the University of Leipzig (n=22).

Even though the analyzed groups are very small, these data may give further support to the importance of the combined effect of the BAALC/miR-3151 locus.
In conclusion, the here presented manuscript is not only the first report of the prognostic importance of a novel microRNA miR-3151, but it is the first example of an interaction between an intronic miR and its host gene in AML. Since so far the impact of intronic miRs on the effects of their host genes has been vastly overlooked, this discovery may initiate new research in the pathways of other human malignancies in which genes with intronic miRs are involved. In the case of BAALC and miR-3151, we may conclude that high expression of both oncogenes is important to promote leukemogenesis, since both are impacting on different downstream pathways and ultimately different outcome endpoints.

Nothing is known about the downstream targets about BAALC yet, but since we could validate two members of the ubiquitination pathway as direct miR-3151 target genes, it may be tempting to speculate that BAALC may either also be involved in the ubiquitination pathway or that the repressed ubiquitination (and consecutive degradation) of the FBXL20 targets may directly or indirectly support the function of BAALC.

The further characterization of the interplay of miR-3151 and its host gene BAALC may unravel novel pathways, which may lead to a better understanding of the pathogenesis of AML.
References


34) Kim YK, Kim VN. Processing of intronic microRNAs. *EMBO* 2007; J 26:775–783


The original article is available at:
http://bloodjournal.hematologylibrary.org/content/120/2/249.full.pdf
Komplette Publikationsliste / Complete List of Publications

Peer-Reviewed Publications


- **Eisfeld AK**, Westerman M, Krahl R, Leiblein S, Liebert UG, Hehme M, Teupser D, Niederwieser D, Al-Ali HK. *Elevated serum hepcidin in patients with AML prior to and after allogeneic hematopoietic cell transplantation does not correlate with transfusional body iron, HFE genotype or Graf versus Host Disease and may protect from excessive parenchymal iron loading* (ASH Annual Meeting Abstracts 2009, Abstract P22974)


Lebenslauf / Curriculum Vitae

Ann-Kathrin Eisfeld
*29.03.1981 Heide-Holstein/Germany
Office: 898 Biomedical Research Tower
460 W. 12th Ave., Columbus, OH 43210
Home: 5323 York County Rd, 43221 Columbus/Ohio
Ann-Kathrin.Eisfeld@osumc.edu, cell: 614-4775667

CURRENT POSITION
Since 10/09 Postdoctoral Fellow of Drs. Clara D. Bloomfield and Albert de la Chapelle, The Ohio State University Columbus/Ohio, Comprehensive Cancer Center

CURRENT RESEARCH
Since 10/09 Hereditary changes in Acute Myeloid Leukemia
Since 11/10 MicroRNA miR-3151: Analysis of prognostic significance, upstream regulation and target gene validation in AML

RESIDENT WORK
Since 06/07 Department of Hematology and Oncology, University of Leipzig/Germany

GENERAL EDUCATION
08/00-05/07 Medical School University of Leipzig/ Germany (M.D.)
08/91-07/00 Hermann-Tast-Gymnasium, Husum/Nordsee
HONORS

08/2009 Fellowship of the Leukemia Clinical Research Foundation: “Hereditary changes predisposing to Acute Myeloid Leukemia”, 3 year funding

05/2011 Pelotonia Postdoctoral Fellowship Award “The genomic basis for BAALC overexpression in AML”, 2 year funding starting 09/2011

CONFERENCE PRESENTATIONS

03/2012 “MiR-3151, a Novel MicroRNA Embedded in BAALC, Is Only Weakly Co-Expressed with Its Host Gene and Independently Impacts on the Clinical Outcome of Older Patients (Pts) with De Novo Cytogenetically Normal Acute Myeloid Leukemia (CN-AML)” (Leukemia Correlative Science Symposium, ALLIANCE Meeting Chicago).

02/2011 “Heritable Polymorphism Predisposes to High Expression of BAALC in Cytogenetically Normal Acute Myeloid Leukemia (CN-AML)” (Leukemia Symposium, Annual Meeting of the Comprehensive Cancer Center, February 28th 2011).

03/2009 "Iron depletion pattern in patients with iron overload after allogeneic haematopoietic cell transplantation treated by phlebotomy” (EBMT Annual Meeting 2009)

03/2006 “Donor HFE genotype influences the rate of iron depletion by phlebotomy for iron overload after allogeneic HCT” (EBMT Annual Meeting 2006)
03/2005  “Iron overload in patients after allogeneic does not correlate with HFE genotype which is always of donor origin” (EBMT Annual Meeting 2005)

03/2005  “Iron overload is a frequent hidden complication after allogeneic HCT and treatable by phlebotomy” (EBMT Annual Meeting 2005)

MEDICAL TEXTBOOKS/CONTRIBUTIONS


ACTIVE RESEARCH SUPPORT
2009-present  Fellowship of the Leukemia Clinical Research Foundation: “Hereditary changes predisposing to Acute Myeloid Leukemia”, 3 year funding started 10/2009

34
05/2011 Pelotonia Postdoctoral Fellowship Award “The genomic basis for BAALC overexpression in AML”, 2 year funding starting 09/2011

COMPLETED RESEARCH SUPPORT
01/09 Novartis Young Investigator Foundation 02/01/09-06/01/09. Title: Clinical relevance of HFE gene mutations of donor origin on iron metabolism of patients after allogeneic PBSCT. PI: Haifa Katrin Al-Ali, MD, Co-I: Ann-Kathrin Eisfeld, MD
07/09 Novartis Young Investigators Foundation 07/01/09-12/01/09. Title: Clinical relevance of serum hepcidin levels and detection of a possible predictive value for the development of GvHD, infections and relapse after allogeneic PBSCT. PI: Haifa Katrin Al-Ali, MD, Co-I: Ann-Kathrin Eisfeld, MD

EXTRACURRICULAR ACTIVITIES
Since 08/09 Author for “Die Schwarze Reihe”/ Internal Medicine (Georg Thieme Verlag KG)
Since 10/05 Copyediting of „Die Schwarze Reihe“ (Georg Thieme Verlag KG, literature for 1.+ 2. Medical State Examination)
05/03-01/07 Journalist for online- and print version of „Via medici“ (Georg Thieme Verlag KG, journal for medical students)
Erklärung über die eigenständige Abfassung der Arbeit


Leipzig, den 10.10.2012

Datum

Unterschrift
Danksagung / Acknowledgements

The decision to starting Medical school, the decision to become a hematologist, the decision to do a post-doctoral fellowship at The Ohio State University - would all not have been possible and especially all not have been successful without the continuous support and love of the people who mean so much to me.

First, I would like to thank my parents and my sister that they always supported all of my decisions and therefore gave me the strengths to believe in my way over every distance. I never felt away from home.

I would also like to thank Prof. Niederwieser, who very early on gave me the opportunity to become a part of the Department of Hematology and Oncology in Leipzig, and also supported my wish to do research and to start a postdoctoral fellowship at OSU and who will continue the mentorship after my return.

Especially, I would like to express my love and gratitude to my mentors Prof. Albert de la Chapelle and Prof. Clara D. Bloomfield, who despite my inexperience gave me the opportunity to work in their research groups, who continuously gave me both support and critique and – moreover- who gave me a vision of what can be achieved if you are persistent, hard working and never stop following your dreams.

Finally, I would like to thank Sebastian, without whom my life in the past years would not have been the same.
Supplement 1

Supplemental Material of Eisfeld et al. miR-3151 interplays with its host gene BAALC and independently impacts on outcome of patients with cytogenetically normal acute myeloid leukemia. Blood. 2012 Jul 12;120(2):249-58,