



MicroRNA Expression-Based Model Indicates Event-Free Survival in Pediatric Acute Myeloid Leukemia

Emilia L. Lim, Diane L. Trinh, Rhonda E. Ries, Jim Wang, Robert B. Gerbing, Yussanne Ma, James Topham, Maya Hughes, Erin Pleasance, Andrew J. Mungall, Richard Moore, Yongjun Zhao, Richard Aplenc, Lillian Sung, E. Anders Kolb, Alan Gamis, Malcolm Smith, Daniela S. Gerhard, Todd A. Alonzo, Soheil Meshinchi, and Marco A. Marra

Author affiliations and support information (if applicable) appear at the end of this article.

Published at jco.org on October 25, 2017.

S.M. and M.A.M. contributed equally to this work.

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

Corresponding author: Marco A. Marra, PhD, FRSC, Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, 675 W 10th Ave, Vancouver, BC, Canada V5Z 1L3; e-mail: mmarra@bcgsc.ca.

© 2017 by American Society of Clinical Oncology. Creative Commons Attribution Non-Commercial No Derivatives 4.0 License.



0732-183X/17/3535w-3964w/\$20.00

A B S T R A C T

Purpose

Children with acute myeloid leukemia (AML) whose disease is refractory to standard induction chemotherapy therapy or who experience relapse after initial response have dismal outcomes. We sought to comprehensively profile pediatric AML microRNA (miRNA) samples to identify dysregulated genes and assess the utility of miRNAs for improved outcome prediction.

Patients and Methods

To identify miRNA biomarkers that are associated with treatment failure, we performed a comprehensive sequence-based characterization of the pediatric AML miRNA landscape. miRNA sequencing was performed on 1,362 samples—1,303 primary, 22 refractory, and 37 relapse samples. One hundred sixty-four matched samples—127 primary and 37 relapse samples—were analyzed by using RNA sequencing.

Results

By using penalized lasso Cox proportional hazards regression, we identified 36 miRNAs the expression levels at diagnosis of which were highly associated with event-free survival. Combined expression of the 36 miRNAs was used to create a novel miRNA-based risk classification scheme (AMLmiR36). This new miRNA-based risk classifier identifies those patients who are at high risk (hazard ratio, 2.830; $P \leq .001$) or low risk (hazard ratio, 0.323; $P \leq .001$) of experiencing treatment failure, independent of conventional karyotype or mutation status. The performance of AMLmiR36 was independently assessed by using 878 patients from two different clinical trials (AAML0531 and AAML1031). Our analysis also revealed that miR-106a-363 was abundantly expressed in relapse and refractory samples, and several candidate targets of miR-106a-5p were involved in oxidative phosphorylation, a process that is suppressed in treatment-resistant leukemic cells.

Conclusion

To assess the utility of miRNAs for outcome prediction in patients with pediatric AML, we designed and validated a miRNA-based risk classification scheme. We also hypothesized that the abundant expression of miR-106a could increase treatment resistance via modulation of genes that are involved in oxidative phosphorylation.

J Clin Oncol 35:3964-3977. © 2017 by American Society of Clinical Oncology. Creative Commons Attribution Non-Commercial No Derivatives 4.0 License: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

INTRODUCTION

Acute myeloid leukemia (AML) comprises almost 25% of pediatric leukemias¹ and is characterized by genetic alterations that lead to impaired differentiation and clonal expansion.² Approximately 80% of patients achieve complete response after the induction chemotherapy, 40% of whom subsequently suffer from relapsed disease^{1,3} (Fig 1).

Pediatric AML patients separate into distinct risk categories on the basis of specific chromosomal alterations.^{1,2} Somatic mutations in genes—such as *FLT3*, *NPM1*, and *CEBPA*—have been associated with outcome and are used as prognostic markers¹; however, approximately 60% of children with AML who lack clinically informative mutation profiles are not amenable to risk-based therapy allocation solely on the basis of cytogenetic/molecular subgroups. In these patients, risk of relapse (RR) cannot be determined

ASSOCIATED CONTENT



Data Supplement
DOI: <https://doi.org/10.1200/JCO.2017.74.7451>

DOI: <https://doi.org/10.1200/JCO.2017.74.7451>

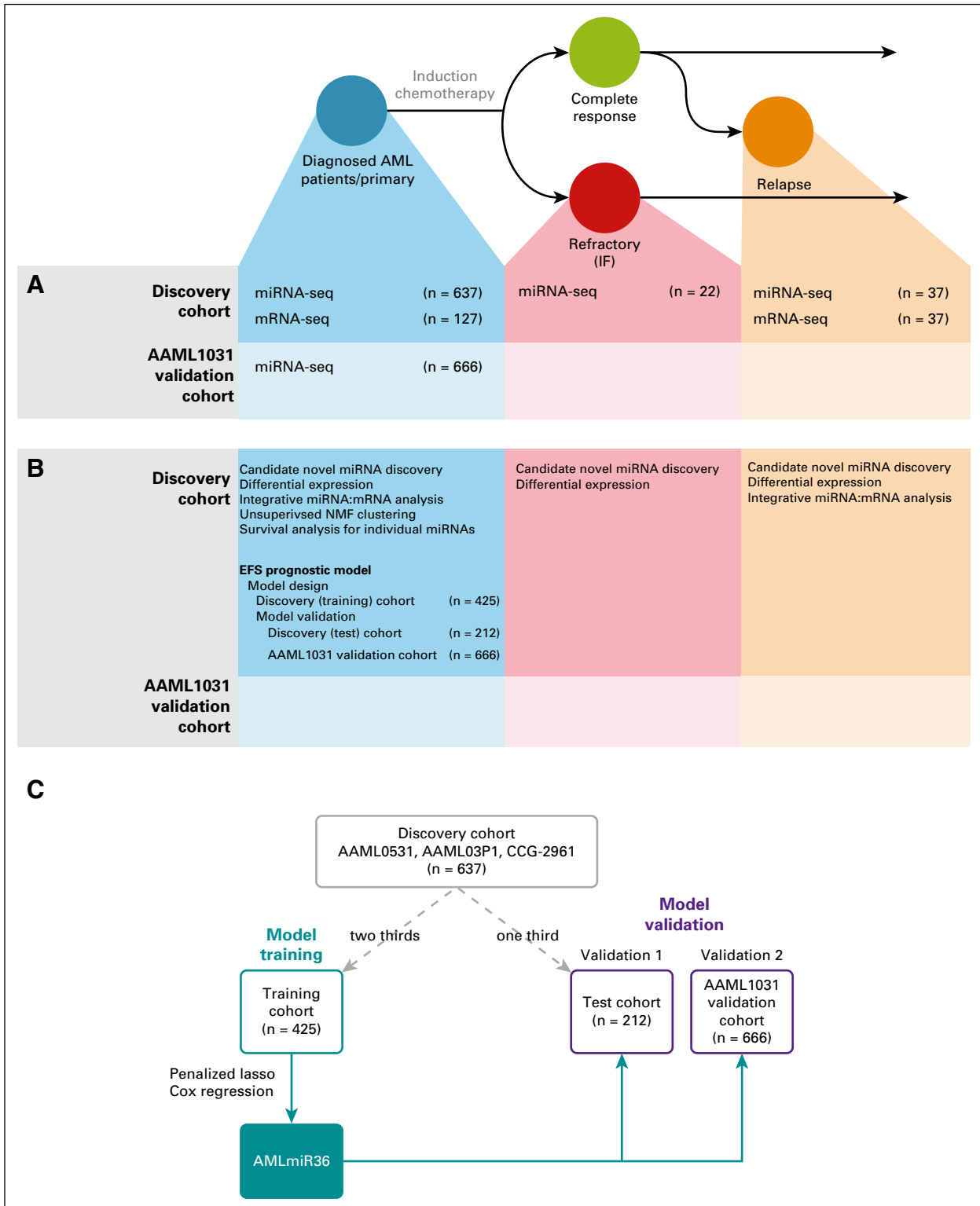


Fig 1. Transcriptome analysis of pediatric acute myeloid leukemia (AML). Schematic diagram of our experimental design and possible outcome trajectories of pediatric patients with AML. Data set consists of primary samples that are obtained at the time of diagnosis (blue), relapse samples (orange), and induction failure (IF)/refractory samples (red). (A) Sequence data (miRNA sequencing [miRNA-seq] and mRNA sequencing [mRNA-seq]) generated in our study. Samples were obtained in two batches: the discovery cohort consisted primarily of diagnostic samples from the AAML0531 trial (n = 528), but also included a few samples from the AAML03P1 (n = 71) and CCG-2961 (n = 38) trials; the AAML1031 validation cohort consisted of patients from the more recent AAML1031 trial (n = 666). (B) Analyses were performed for each sample and sequence data type. The bulk of the analyses were performed on primary (diagnostic) samples from the discovery cohort. (C) Study design for the training and validation of AMLmiR36. The discovery cohort (gray box) was randomly divided into a training cohort (two thirds; n = 425) and test cohort (one third; n = 212). AMLmiR36 (filled blue box) was trained on data from the training cohort (blue box) and validated on independent data from the test cohort and AAML1031 validation cohort (gold boxes). EFS, event-free survival; NMF, non-negative matrix factorization.

at diagnosis, as risk is assessed by using measurements of minimal residual disease obtained after the first round of induction chemotherapy¹; therefore, the identification of additional biomarkers and therapeutic targets may enhance existing risk-based therapy schemas whereby either those with high RR or those with appropriate targets for directed therapy can be allocated to an alternative therapy to optimize outcome and minimize toxicity.¹

Efforts to identify patients who are destined to experience treatment failure are ongoing,^{4,5} and evaluations of microRNA (miRNA) dysregulation to support this endeavor have been reported. For example, probe-based methods have been used to profile miRNA expression in pediatric AML⁶⁻¹¹ and have identified expression patterns that are characteristic of AML subtypes^{6,7} or inferior outcomes⁸⁻¹²; however, the full complement of miRNA expression has not been comprehensively assessed to produce a meaningful signature with which to predict patient outcomes.

miRNA sequencing provides unbiased digital read counts of miRNA expression and the opportunity to discover novel miRNA transcripts.¹³ By using this technology, The Cancer Genome Atlas has generated sequence profiles of 200 adult AML patients¹⁴; however, some genetic events observed in adults seem to be rare or absent in pediatric AML,¹⁵ which limits the applicability of data from adults to children and indicates a need to directly profile pediatric AML samples.

Here, we provide a detailed analysis of miRNA dysregulation in pediatric patients with AML, including expression patterns of previously annotated and novel miRNAs and identification of candidate miRNA:mRNA interactions. We designed AMLmiR36, a miRNA-based predictor of post-standard induction chemotherapy outcome in childhood AML. AMLmiR36 is the first mutation and translocation-independent schema to identify event-free survival (EFS) in children with AML and offers the potential for improved patient stratification and management.

PATIENTS AND METHODS

Patient Samples and Treatment Protocol

Our Discovery Cohort data set consisted of diagnostic, refractory, and relapse samples from 654 patients who were enrolled in one of the following pediatric AML studies: AAML0531 (n = 528),¹⁶ AAML03P1 (n = 71),¹⁷ or CCG-2961 (n = 38).¹⁸ The Medical Research Council–based therapy backbone¹⁶ was identical in AAML0531 and AAML03P1 trials. We also obtained 666 cases from the AAML1031 trial to validate our findings.

Patients who were enrolled in the AAML0531 trial who were deemed to be at high risk received hematopoietic stem cell transplantation (HSCT). Eighty-two patients received HSCT in the first round of induction chemotherapy compared with 401 patients who received chemotherapy consolidation. In AAML1031, a criterion for exclusion from analysis was high allelic ratio *FLT3*-ITD. Such patients were deemed to be high risk and received a different treatment regimen.

Analytic Platforms

We generated miRNA sequencing data from 1,303 primary, 22 refractory, and 37 relapse samples, as well as mRNA sequencing data from 127 primary and 37 relapse samples (Fig 1A and Data Supplement). Sequencing for all samples was performed by using the Illumina HiSeq2500 (Illumina, San Diego, CA). miRNA sequencing data were aligned to hg19 and annotated on the basis of miRBase (version 21).¹⁹ All

data sets and sequences are available online (dbGaP accession no.: phs000465; SRA accession no.: SRP012000).^{20,21}

Statistical Methods

Patient baseline characteristics were compared between groups by using χ^2 tests. The miRNA expression threshold—10 or more reads per million mapped reads (RPM) in 10 or more miRNA sequencing libraries—was based on miRBase criteria for high-confidence miRNAs.¹⁹ Differentially expressed (DE) miRNAs were identified by using two-tailed Wilcoxon tests, where significantly DE miRNAs were those with Benjamini-Hochberg²² multiple-test corrected *P* values of < .05. Unsupervised clustering was performed by using non-negative matrix factorization (NMF).²³

Univariable analyses were performed by using Cox proportional hazards (PH) regression models²⁴ to assess the association between expression levels of individual miRNAs and EFS or overall survival (OS). Expression level analyses were performed by using low/high expression groups or continuous expression values (log₂ RPM). Low/high expression designations were assigned by using X-tile cohort separation²⁵ on EFS data. By using the sample function in R (version 3.3.2), we divided the discovery cohort into training and test cohorts, which consisted of two thirds and one third, respectively, of patients in the discovery cohort. The miRNA-based EFS predictive model that was established in the discovery (training) cohort was tested in the discovery (test) cohort and AAML1031 validation cohort. The model was estimated in the discovery (training) cohort by using penalized lasso Cox PH regression (GLMnet R Package),²⁶ where coefficients that were estimated for each miRNA feature in the training cohort were carried over to the discovery (test) cohort and AAML1031 validation cohort. Coefficients were not re-estimated in either validation cohort. Integrative miRNA:mRNA expression analysis was performed as previously described.¹³

RESULTS

miRNA Expression in Childhood AML

Filtering the miRNA sequencing data from the discovery cohort (n = 696) against annotated miRNAs revealed 61 candidate novel miRNA species (122 miRNAs; Data Supplement). miRNA expression profiling revealed that 529 miRNAs, including 22 candidate novel miRNAs, were expressed at 10 or more RPM in 10 or more miRNA sequencing libraries (Data Supplement). These miRNAs were included in subsequent analyses.

To determine the extent of heterogeneity in miRNA expression across our discovery cohort, we performed NMF clustering by using the miRNA expression profiles of 637 primary samples. This analysis identified four subgroups (Fig 2A and Data Supplement) that were characterized by distinct miRNA expression patterns (Fig 2B and Data Supplement), enriched for cytogenetic and molecular alterations (Fig 2C), and correlated with outcome (Fig 2D). This result was consistent with reports that associated genomic variants with miRNA expression.²⁷

Our analysis revealed 20 alterations that were associated with miRNA expression (*q* < 0.05, Wilcoxon test; Data Supplement). Specifically, we confirmed reported associations that included that of *t*(8;21) with abundant miR-181b, miR-146b, miR-181a, miR-146a, and miR-126, as well as with reduced miR-133; that of *NPM1* mutation with abundant miR-10a/b and miR-196b and reduced miR-128, miR126, miR-130a, and miR-451; that of *FLT3*-ITD with abundant miR-155; that of *CEBPA* mutation with abundant miR-181a and miR-335; and that of *MLL* rearrangements with reduced

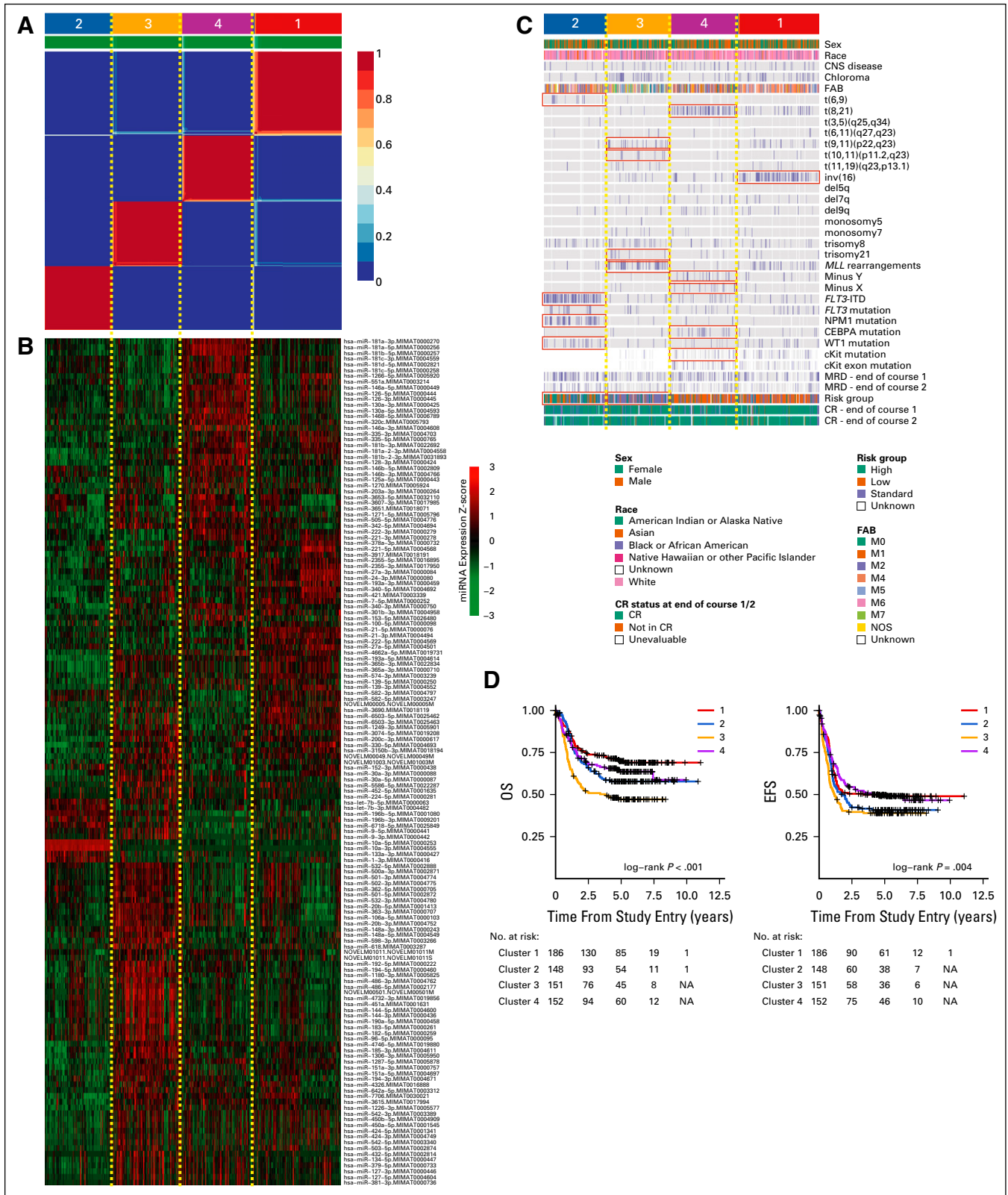


Fig 2. Unsupervised non-negative matrix factorization (NMF) clustering of miRNA expression profiles of primary samples. (A) NMF consensus map for k = 4 subgroups. Deep red blocks numbered 1 to 4 indicate four subgroups that were identified by using NMF. (B) Heat map displaying the expression of miRNAs that are significantly differentially expressed between subgroups ($P < .05$, Wilcoxon test proportional hazards regression adjusted; \log_2 fold change > 1). (C) Covariate tracks displaying the clinical attributes of each patient. Red boxes indicate clinical attributes that enrich one cluster over the others ($P < .01$, Fisher's exact test). (D) Kaplan-Meier plots illustrating overall survival (OS) and event-free survival (EFS) status of patients in each subgroup. CR, completed response; FAB, French-American-British; ITD, internal tandem duplication; MRD, minimal residual disease; NA, not applicable.

expression of miR-126, miR-125b, miR-146a, miR-181a, miR-146b, and let-7c.²⁷ A full list of associations of miRNA expression with genomic variants is provided in the Data Supplement.

miRNAs Associated With Patient Survival

To identify miRNAs whose expression correlated with outcome, we performed Cox PH regression analyses on low/high expression groups of each expressed miRNA. In the discovery cohort, this analysis revealed that there were 380 miRNAs whose expression was associated with EFS ($q < 0.05$, univariable Cox PH regression; hazard ratio (HR), 0.509 to 0.795/1.26 to 2.05) and 185 miRNAs that were associated with OS (HR, 0.482 to 0.736/1.36 to 2.32). Two hundred sixteen and 80 miRNAs were associated with EFS (HR, 0.509 to 0.795/1.27 to 2.04) and OS (HR, 0.482 to 0.730/1.36 to 2.32), respectively, independent of conventional risk factors (cytogenetic risk group and WBC cell count; $P < .05$, multivariable Cox PH regression; Data Supplement). We repeated survival analyses by using data from the AAML1031 validation cohort, which confirmed that 48 and 36 miRNAs were associated with EFS ($P < .05$, univariable Cox PH regression; HR, 0.003 to 0.798/1.28 to 3.30) and OS (HR, 0.003 to 0.766/1.45 to 2.52), respectively, and that 12 and 5 miRNAs were associated with EFS (HR, 0.003 to 0.789/1.46 to 3.30) and OS (HR, 1.58 to 2.52), respectively, independent of conventional risk factors ($P < .05$, multivariable Cox PH regression; Data Supplement).

We then performed Cox PH regression analysis on continuous expression values—measured in RPM—of each miRNA. This demonstrated that 66 and 55 miRNAs had linear associations with EFS ($q < 0.05$, Cox PH regression; HR, 0.730 to 0.946/1.05 to 1.38) and OS ($q < 0.05$, Cox PH regression; HR, 0.646 to 0.904/1.07 to 1.36), respectively. The same analysis using data from the AAML1031 validation cohort confirmed that 39 and 46 miRNAs had linear associations with EFS ($q < 0.05$, Cox PH regression; HR, 0.829 to 0.943/1.08 to 1.30) and OS ($q < 0.05$, Cox PH regression; HR, 0.748 to 0.904/1.07 to 1.42; Data Supplement), respectively. The 31 miRNAs that had linear associations with both OS and EFS in both cohorts are presented in the Data Supplement.

miRNA Expression–Based EFS Predictive Model (AMLmiR36)

Having observed many individual miRNAs with significant associations with outcome, we next developed a miRNA expression–based predictor of EFS by combining multiple miRNAs. We divided our discovery cohort ($n = 637$; Table 1) into a discovery (training) cohort (two thirds; $n = 425$) to generate the model, then tested its performance twice, first in the discovery (test) cohort (one third; $n = 212$), then subsequently in the AAML1031 validation cohort (Fig 1C). The discovery (training) and discovery (test) cohorts were derived by random selection, and there were no significant differences between cohorts with respect to patient characteristics ($P > .01$, χ^2 test; Table 1). Only previously annotated miRNAs were considered for this analysis.

The model was estimated in the discovery (training) cohort by using the penalized lasso Cox PH regression (GLMnet R Package).²⁶ The resulting model was composed of 36 miRNAs (Fig 3A and Data Supplement), of which 16 were overexpressed and 20 were underexpressed in patients who experienced an event (Fig 3B). Only four of 36 miRNAs—miR-155,¹¹ miR-335,¹⁰ miR-139,²⁸ and miR-375⁹—had

previously been associated with survival in pediatric AML. We named our model, AMLmiR36.

To demonstrate the potential clinical utility of AMLmiR36, we created a miRNA-based risk stratification scheme in which we determined score thresholds in the discovery (training) cohort to separate patients into low, intermediate, and high AMLmiR36 groups (Fig 3C). The AMLmiR36-based risk stratification scheme was significantly associated with EFS (Fig 3E; $P < .001$). Patients with high AMLmiR36 scores had inferior outcomes (HR, 3.659; 95% CI, 2.77 to 4.83; 5-year EFS, 9.26%; 5-year OS, 29.6%; $P < .001$), whereas cases with low AMLmiR36 scores had superior outcomes (HR, 0.265; 95% CI, 0.16 to 0.43; 5-year EFS, 84.4%; 5-year OS, 90.3%; $P < .001$). This observation was validated in both the discovery (test) cohort and the AAML1031 validation cohort (Table 2 and Fig 3E; $P < .001$).

The High AMLmiR36 Group Is Independent of Established Indicators of Outcome

We observed an enrichment for patients who were characterized by poor prognostic indicators—that is, high-risk cytogenetics, *FLT3*-ITD positive²⁹—in the high AMLmiR36 group ($P < .05$, Fisher's exact test; Fig 3D and Data Supplement). Despite this, multivariable analyses of the discovery (training), discovery (test), and AAML1031 validation cohorts indicated that the high AMLmiR36 risk group was independent of cytogenetic risk group, WBC count, *FLT3*-ITD status, and HSCT status ($P < .001$, multivariable Cox PH regression; Table 2), which are established indicators of outcome. Multivariable analyses of AMLmiR36, along with expression values of the four miRNAs—miR-155,¹¹ miR-335,¹⁰ miR-139,²⁸ and miR-375⁹—that had previously individually associated with survival, indicated that AMLmiR36 is a more robust predictor of EFS than the individual miRNAs ($P > .05$; Data Supplement). Moreover, with patients for which disease-free survival information was available, we demonstrated the AMLmiR36 was independent of minimal residual disease and RR ($P > .05$; Data Supplement).

AMLmiR36 Improves the Identification of High-Risk Patients

Patients who were enrolled in AAML0531 and who were deemed to be at high risk received HSCT instead of consolidation chemotherapy. Despite this, AMLmiR36 defined three distinct outcome cohorts in both patients receiving consolidation chemotherapy and HSCT ($P < .001$, log rank; Data Supplement), which indicated that AMLmiR36 can identify high-risk patients regardless of consolidation therapy received.

The AAML1031 validation cohort was obtained from a clinical trial that was different than that from which the discovery cohort was obtained (Table 1). In AAML1031, as patients with high allelic ratio *FLT3*-ITD were excluded from analysis, the representation of high-risk patients and the definition of cytogenetic/molecular risk groups were different from the discovery cohort. In particular, the AAML1031 validation cohort had fewer patients with chloroma, inv(16), *FLT3*-ITD, high WBC, and high cytogenetic/molecular high-risk cases compared with the discovery cohort ($P > .05$, χ^2 test; Table 1). Despite this difference, we addressed whether the AMLmiR36 signature could identify a group of AMLmiR36

Pediatric AML miRNA-Based Event-Free Survival Predictor

Table 1. Patient Characteristics

| Characteristic | Discovery Cohort (n = 637) | | | | AAML1031 Validation Cohort (n = 666) | |
|---|----------------------------|---------------------------------------|-----------------------------------|--|--------------------------------------|--|
| | All Patients (n = 637) | Discovery (training) Cohort (n = 425) | Discovery (test) Cohort (n = 212) | Training v Test Comparison, <i>P</i> * | All Patients (n = 666) | Discovery v AAML1031 Validation Comparison, <i>P</i> * |
| Study | | | | | | |
| AAML03P1 | 71 (11.146) | 44 (10.353) | 27 (12.736) | .368 | 0 (N/A) | < .001 |
| AAML0531 | 528 (82.889) | 361 (84.941) | 167 (78.774) | .051 | 0 (N/A) | < .001 |
| CCG-2961 | 38 (5.965) | 20 (4.706) | 18 (8.491) | .057 | 0 (N/A) | < .001 |
| AAML1031 | 0 (0) | 0 (0) | 0 (0) | N/A | 666 (100.000) | < .001 |
| Sex | | | | | | |
| Male | 322 (50.549) | 217 (51.059) | 105 (49.528) | .716 | 356 (53.453) | .321 |
| Female | 315 (49.451) | 208 (48.941) | 107 (50.472) | .716 | 310 (46.547) | .321 |
| Race | | | | | | |
| American Indian or Alaska Native | 3 (0.471) | 1 (0.235) | 2 (0.943) | .219 | 6 (0.901) | .547 |
| Asian | 26 (4.082) | 17 (4.0) | 9 (4.245) | .883 | 31 (4.655) | .711 |
| Native Hawaiian or other Pacific Islander | 3 (0.471) | 1 (0.235) | 2 (0.943) | .219 | 2 (0.300) | .96 |
| Black or African American | 68 (10.675) | 47 (11.059) | 21 (9.906) | .657 | 81 (12.162) | .45 |
| White | 473 (74.254) | 315 (74.118) | 158 (74.528) | .911 | 476 (71.471) | .286 |
| Other | 29 (4.553) | 21 (4.941) | 8 (3.774) | .505 | 0 (0.000) | N/A |
| Unknown | 35 (5.495) | 23 (5.412) | 12 (5.66) | .897 | 70 (10.511) | .00127 |
| Ethnicity | | | | | | |
| Hispanic or Latino | 122 (19.152) | 80 (18.824) | 42 (19.811) | .765 | 125 (18.769) | .916 |
| Not Hispanic or Latino | 497 (78.022) | 331 (77.882) | 166 (78.302) | .904 | 525 (78.829) | .774 |
| Liver | | | | | | |
| Normal | 444 (69.702) | 294 (69.176) | 150 (70.755) | .683 | N/A | N/A |
| Enlarged | 176 (27.63) | 118 (27.765) | 58 (27.358) | .914 | N/A | N/A |
| Spleen | | | | | | |
| Normal | 439 (68.917) | 293 (68.941) | 146 (68.868) | .985 | N/A | N/A |
| Enlarged | 181 (28.414) | 120 (28.235) | 61 (28.774) | .887 | N/A | N/A |
| CNS disease | | | | | | |
| No | 588 (92.308) | 392 (92.235) | 196 (92.453) | .923 | N/A | N/A |
| Yes | 49 (7.692) | 33 (7.765) | 16 (7.547) | .923 | N/A | N/A |
| Choloroma | | | | | | |
| No | 554 (86.97) | 371 (87.294) | 183 (86.321) | .731 | 639 (95.946) | < .001 |
| Yes | 82 (12.873) | 53 (12.471) | 29 (13.679) | .668 | 26 (3.904) | < .001 |
| FAB | | | | | | |
| M0 | 16 (2.512) | 12 (2.824) | 4 (1.887) | .477 | N/A | N/A |
| M1 | 65 (10.204) | 40 (9.412) | 25 (11.792) | .35 | N/A | N/A |
| M2 | 136 (21.35) | 90 (21.176) | 46 (21.698) | .88 | N/A | N/A |
| M4 | 155 (24.333) | 109 (25.647) | 46 (21.698) | .274 | N/A | N/A |
| M5 | 115 (18.053) | 78 (18.353) | 37 (17.453) | .781 | N/A | N/A |
| M6 | 10 (1.57) | 6 (1.412) | 4 (1.887) | .649 | N/A | N/A |
| M7 | 26 (4.082) | 19 (4.471) | 7 (3.302) | .482 | N/A | N/A |
| Other | 36 (5.651) | 20 (4.706) | 16 (7.547) | .143 | N/A | N/A |
| Unknown | 78 (12.245) | 51 (12.0) | 27 (12.736) | .789 | N/A | N/A |
| Cytogenetics | | | | | | |
| Normal | 150 (23.548) | 92 (21.647) | 58 (27.358) | .109 | 144 (21.622) | .444 |
| t(8;21) | 89 (13.972) | 59 (13.882) | 30 (14.151) | .927 | 107 (16.066) | .327 |
| inv(16) | 93 (14.6) | 65 (15.294) | 28 (13.208) | .482 | 67 (10.060) | .0159 |
| t(9;11)/11q23 | 104 (16.327) | 75 (17.647) | 29 (13.679) | .202 | 163 (24.474) | < .001 |
| t(6;9) | 13 (2.041) | 7 (1.647) | 6 (2.83) | .32 | 7 (1.051) | .22 |
| Monosomy 7 | 12 (1.884) | 8 (1.882) | 4 (1.887) | .997 | 10 (1.502) | .749 |
| Del(7q) | 4 (0.628) | 2 (0.471) | 2 (0.943) | .477 | 12 (1.802) | .0946 |
| -5/5q- | 6 (0.942) | 2 (0.471) | 4 (1.887) | .081 | 7 (1.051) | 1.0 |
| Plus 8 | 38 (5.965) | 23 (5.412) | 15 (7.075) | .403 | 33 (4.955) | .496 |
| Other | 99 (15.542) | 72 (16.941) | 27 (12.736) | .167 | 113 (16.967) | .534 |
| Missing | 29 (4.553) | 20 (4.706) | 9 (4.245) | .793 | 3 (0.450) | < .001 |
| FLT3-ITD | | | | | | |
| No | 514 (80.691) | 343 (80.706) | 171 (80.66) | .989 | 621 (93.243) | < .001 |
| Yes | 113 (17.739) | 73 (17.176) | 40 (18.868) | .598 | 45 (6.757) | < .001 |
| FLT3-ITD point mutation | | | | | | |
| No | 579 (90.895) | 392 (92.235) | 187 (88.208) | .096 | N/A | N/A |
| Yes | 46 (7.221) | 23 (5.412) | 23 (10.849) | .012 | N/A | N/A |
| NPM mutation | | | | | | |
| No | 563 (88.383) | 378 (88.941) | 185 (87.264) | .534 | 608 (91.291) | .0995 |
| Yes | 53 (8.32) | 31 (7.294) | 22 (10.377) | .184 | 58 (8.709) | .879 |

(continued on following page)

Table 1. Patient Characteristics (continued)

| Characteristic | Discovery Cohort (n = 637) | | | Training v Test Comparison, <i>P</i> * | AAML1031 Validation Cohort (n = 666) | |
|-----------------------------------|----------------------------|---------------------------------------|-----------------------------------|--|--------------------------------------|--|
| | All Patients (n = 637) | Discovery (training) Cohort (n = 425) | Discovery (test) Cohort (n = 212) | | All Patients (n = 666) | Discovery v AAML1031 Validation Comparison, <i>P</i> * |
| <i>CEBPA</i> mutation | | | | | | |
| No | 588 (92.308) | 394 (92.706) | 194 (91.509) | .593 | 626 (93.994) | .273 |
| Yes | 32 (5.024) | 20 (4.706) | 12 (5.66) | .603 | 40 (6.006) | .513 |
| <i>WT1</i> mutation | | | | | | |
| No | 571 (89.639) | 380 (89.412) | 191 (90.094) | .79 | N/A | N/A |
| Yes | 47 (7.378) | 30 (7.059) | 17 (8.019) | .662 | N/A | N/A |
| WBCs | | | | | | |
| < 100,000 | 501 (78.65) | 329 (77.412) | 172 (81.132) | .28 | 555 (83.333) | .037 |
| ≥ 100,000 | 136 (21.35) | 96 (22.588) | 40 (18.868) | .28 | 110 (16.517) | .0309 |
| Cytogenetic/molecular risk group† | | | | | | |
| Standard | 269 (42.229) | 186 (43.765) | 83 (39.151) | .267 | 375 (56.306) | < .001 |
| Low | 252 (39.56) | 164 (38.588) | 88 (41.509) | .477 | 272 (40.841) | .678 |
| High | 94 (14.757) | 58 (13.647) | 36 (16.981) | .264 | 16 (2.402) | < .001 |
| Unknown | 22 (3.454) | 17 (4.0) | 5 (2.358) | .285 | 3 (0.450) | < .001 |
| CR1 status | | | | | | |
| CR | 483 (75.824) | 320 (75.294) | 163 (76.887) | .658 | 514 (77.177) | .61 |
| Not in CR | 132 (20.722) | 88 (20.706) | 44 (20.755) | .989 | 129 (19.369) | .589 |
| Death | 14 (2.198) | 10 (2.353) | 4 (1.887) | .705 | 10 (1.502) | .466 |
| Unevaluable | 8 (1.256) | 7 (1.647) | 1 (0.472) | .209 | 13 (1.952) | .437 |
| CR2 status | | | | | | |
| CR | 536 (84.144) | 359 (84.471) | 177 (83.491) | .75 | 566 (84.985) | .731 |
| Not in CR | 62 (9.733) | 38 (8.941) | 24 (11.321) | .34 | 55 (8.258) | .404 |
| Death | 19 (2.983) | 15 (3.529) | 4 (1.887) | .251 | 10 (1.502) | .104 |
| Unevaluable | 17 (2.669) | 10 (2.353) | 7 (3.302) | .484 | 35 (5.255) | .0249 |
| MRD1 status | | | | | | |
| No | 350 (54.945) | 242 (56.941) | 108 (50.943) | .152 | 474 (71.171) | < .001 |
| Yes | 145 (22.763) | 89 (20.941) | 56 (26.415) | .121 | 159 (23.874) | .683 |
| MRD2 status | | | | | | |
| No | 361 (56.672) | 248 (58.353) | 113 (53.302) | .225 | N/A | N/A |
| Yes | 64 (10.047) | 42 (9.882) | 22 (10.377) | .845 | N/A | N/A |

Abbreviations: CR, complete response; FAB, French-American-British subtype classification system; ITD, internal tandem duplication; MRD, minimal residual disease; N/A, not applicable.

* χ^2 tests were performed to compare values between cohorts.

†Cytogenetic/molecular risk group is defined differently in the AAML1031 cohort than in the discovery cohort (Data Supplement).

high-risk patients despite having excluded the known high-risk patients from the cohort. Our analysis revealed 88 AMLmiR36 high-risk patients who had a 13.6% 3-year EFS rate compared with intermediate patients who had a 42.7% 3-year EFS rate (Table 2 and Fig 3E; $P < .001$). Moreover, by using the conventional cytogenetic/molecular risk group scheme, 82 (93%) of these 88 AMLmiR36 high-risk patients were classified as only low or standard conventional risk.

We then assessed the performance of AMLmiR36 within the standard conventional risk patients across three different clinical trials (Data Supplement). We noted consistent differences in 3-year EFS rates between AMLmiR36 high and AMLmiR36 intermediate patients (AAML03P1: 18% v 38%; AAML0531: 9% v 41%; 12% v 32%). These evaluations demonstrated that AMLmiR36 could identify a group of high-risk patients from within conventionally assigned risk groups.

miRNA Expression in Treatment-Resistant Samples

We next conducted DE analyses to identify miRNAs that were associated with induction failure (IF) or relapse. Comparing the

diagnostic miRNA expression profiles of patients with and without IF, we revealed 41 miRNAs that were DE between patients who responded and those who did not ($q < 0.05$, Wilcoxon test; Fig 4A and Data Supplement). We repeated this analysis in the AAML1031 cohort and confirmed that five miRNAs—miR-466, miR-5683, miR-106a-3p, miR-20b-3p, and miR-106a-5p—were consistently more abundantly expressed in patients with IF, whereas three miRNAs—miR-365a-3p, miR-199b-3p, and miR-199a-3p—were consistently less abundantly expressed in patients with IF ($P < .05$, Wilcoxon test; Data Supplement).

In patients with IF, 209 miRNAs were DE between diagnostic and refractory samples ($q < 0.05$, Wilcoxon test; Fig 4B and Data Supplement), and 156 miRNAs were DE between diagnostic and postinduction relapse samples ($q < 0.05$, Wilcoxon test; Fig 4C and Data Supplement).

Our DE analyses indicated that two miRNAs—miR-106a-3p and miR-106a-5p—were consistently more abundantly expressed in treatment-resistant contexts—that is, refractory and relapse samples and diagnostic samples of patients with IF ($q < 0.05$, Wilcoxon test; Figs 4A-4C and Data Supplement). This result is compatible with the notion that these miRNAs could be markers of

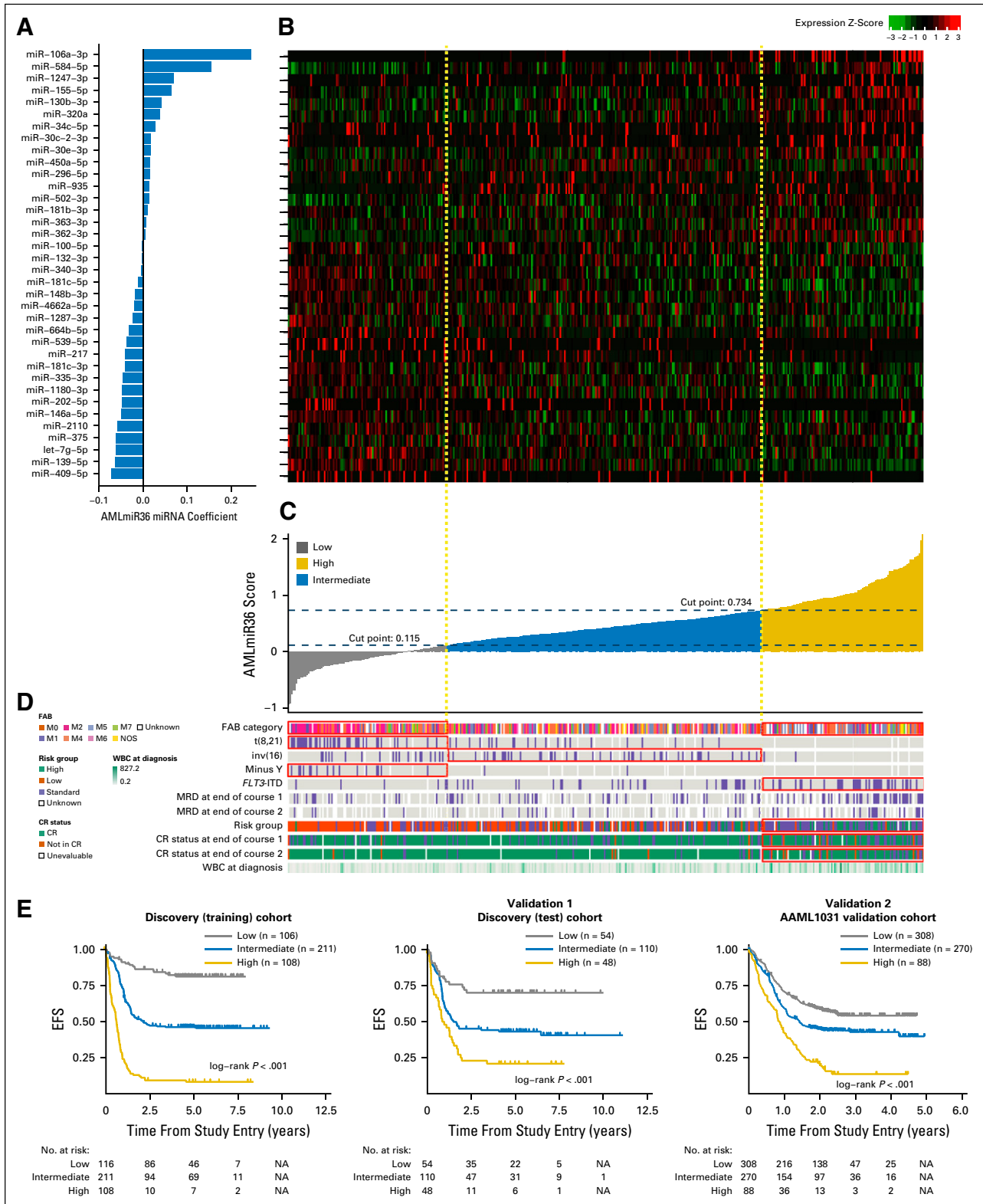


Fig 3. miRNA-based event-free survival (EFS) predictive model. (A) Predictor equation coefficients of the 36 miRNA features in the EFS prognostic model. (B) Heat map of relative expression levels of miRNA features across samples in the discovery (training) cohort (n = 425). (C) Model scores of each patient in the discovery (training) cohort derived using the EFS prognostic model. (D) Covariate tracks displaying the clinical attributes of each patient. Red boxes indicate that the model score group is enriched for the indicated attribute ($P < .01$, Fisher's exact test). (E) Kaplan-Meier plots displaying EFS differences between patients in various—low, intermediate, and high—model score groups within the discovery (training) cohort (n = 425), the discovery (test) cohort (n = 212), and the AAML1031 validation cohort (n = 666). Patients with high model scores had the poorest outcomes, whereas patients with low model scores had superior outcomes. Cox proportional hazards regression ratios are listed in Table 2. CR, completed response; FAB, French-American-British; ITD, internal tandem duplication; MRD, minimal residual disease; NA, not applicable.

Table 2. Univariable and Multivariable Cox Proportional Hazards Regression Analysis of AMLmiR36

| Category | No. | EFS From Study Entry | | |
|--|-----|----------------------|--------------|--------|
| | | HR | 95% CI | P |
| Univariable analysis | | | | |
| Discovery (training) cohort (n = 425) | | | | |
| AMLmiR36 miRNA group | | | | |
| Intermediate | 211 | 1.0 | | |
| High | 108 | 3.659 | 2.77 to 4.83 | < .001 |
| Low | 106 | 0.265 | 0.16 to 0.43 | < .001 |
| Discovery (test) cohort (n = 212) | | | | |
| AMLmiR36 miRNA group | | | | |
| Intermediate | 110 | 1.0 | | |
| High | 48 | 1.79 | 1.19 to 2.68 | .00497 |
| Low | 54 | 0.43 | 0.25 to 0.75 | .00292 |
| AAML1031 validation cohort (n = 666) | | | | |
| AMLmiR36 miRNA group | | | | |
| Intermediate | 270 | 1.0 | | |
| High | 88 | 1.92 | 1.45 to 2.55 | < .001 |
| Low | 308 | 0.68 | 0.54 to 0.86 | .001 |
| Multivariable analysis | | | | |
| Discovery (training) cohort (n = 400)* | | | | |
| AMLmiR36 miRNA group | | | | |
| Intermediate | 198 | 1.0 | | |
| High | 100 | 3.395 | 2.48 to 4.64 | < .001 |
| Low | 102 | 0.286 | 0.17 to 0.48 | < .001 |
| Risk group | | | | |
| Standard | 181 | 1.0 | | |
| High | 58 | 1.047 | 0.64 to 1.71 | .854 |
| Low | 161 | 0.656 | 0.46 to 0.93 | .018 |
| WBCs at diagnosis | | | | |
| < 100,000 | 309 | 1.0 | | |
| ≥ 100,000 | 91 | 1.159 | 0.84 to 1.6 | .366 |
| <i>FLT3</i> -ITD | | | | |
| No | 328 | 1.0 | | |
| Yes | 72 | 1.063 | 0.68 to 1.66 | .788 |
| HSCT | | | | |
| No | 343 | 1.0 | | |
| Yes | 57 | 0.464 | 0.29 to 0.74 | .001 |
| Discovery (test) cohort (n = 206)* | | | | |
| AMLmiR36 miRNA group | | | | |
| Intermediate | 105 | 1.0 | | |
| High | 48 | 1.625 | 1.05 to 2.51 | .029 |
| Low | 53 | 0.443 | 0.25 to 0.79 | .006 |
| Risk group | | | | |
| Standard | 82 | 1.0 | | |
| High | 36 | 1.826 | 0.83 to 4.01 | .133 |
| Low | 88 | 0.535 | 0.34 to 0.85 | .008 |
| WBCs at diagnosis | | | | |
| < 100,000 | 169 | 1.0 | | |
| ≥ 100,000 | 37 | 1.177 | 0.73 to 1.9 | .508 |
| <i>FLT3</i> -ITD | | | | |
| No | 166 | 1.0 | | |
| Yes | 40 | 0.597 | 0.28 to 1.28 | .185 |
| HSCT | | | | |
| No | 171 | 1.0 | | |
| Yes | 35 | 0.317 | 0.16 to 0.62 | .001 |
| AAML1031 validation cohort (n = 666)* | | | | |
| AMLmiR36 miRNA group | | | | |
| Intermediate | 269 | 1.0 | | |
| High | 87 | 1.41 | 1.05 to 1.88 | .021 |
| Low | 306 | 0.90 | 0.71 to 1.16 | .420 |
| Risk group† | | | | |
| Standard | 374 | 1.0 | | |
| High | 16 | 1.34 | 0.76 to 2.36 | .305 |
| Low | 272 | 0.38 | 0.29 to 0.49 | < .001 |
| WBCs at diagnosis | | | | |
| < 100,000 | 553 | 1.0 | | |
| ≥ 100,000 | 109 | 1.54 | 1.18 to 2.00 | .001 |

(continued on following page)

Table 2. Univariable and Multivariable Cox Proportional Hazards Regression Analysis of AMLmiR36 (continued)

| Category | EFS From Study Entry | | | |
|------------------|----------------------|------|--------------|------|
| | No. | HR | 95% CI | P |
| <i>FLT3</i> -ITD | | | | |
| No | 617 | 1.0 | | |
| Yes | 45 | 0.93 | 0.61 to 1.44 | .747 |
| HSCT | | | | |
| No | 603 | 1.0 | | |
| Yes | 59 | 0.88 | 0.60 to 1.30 | .882 |

Abbreviations: EFS, event-free survival; HR, hazard ratio; HSCT, hematopoietic stem cell transplantation; ITD, internal tandem duplication.

*Patients without risk group, *FLT3*-ITD, or HSCT information available were excluded from analysis.

†The definition of the cytogenetic/molecular risk group is different for the AAML1031 trial (Data Supplement).

IF and relapse, and that a study of their expression and function could reveal biology that is important to treatment resistance.

Putative miRNA:mRNA Interactions

To study candidate targets of miRNAs in pediatric AML, we performed an integrative miRNA:mRNA analysis¹³ by using matched miRNA sequencing and mRNA sequencing data for each patient (Fig 4D and Data Supplement). This analysis revealed 396 miRNAs with at least one putative target (mean, 526; range, 1 to 3,334). Pathways that were significantly enriched by both the putative targets of miRNAs that were abundant in refractory samples and miRNAs that were abundant in relapse samples are displayed in Figure 4D. Of these, only oxidative phosphorylation was among the top five most enriched pathways in the both analyses, which suggests that abundant expression of these miRNAs could suppress oxidative phosphorylation target genes in treatment-resistant disease. A full list of significantly enriched pathways is provided in the Data Supplement ($q < 0.05$, Fisher's exact test).

Targets of miR-106a Involved in Oxidative Phosphorylation

miR-106a-3p and miR-106a-5p were consistently abundantly expressed in treatment-resistant contexts (Figs 4A-4C) and were significantly associated with inferior OS and EFS ($q < 0.05$, univariable Cox PH regression; Data Supplement). Because oxidative phosphorylation was predicted to be consistently dysregulated (Fig 4D), we investigated the targets of miR-106a-5p that were involved in this pathway. To test whether miR-106a-5p can act on predicted miRNA binding sites of genes, constructs that contained the predicted binding sites of oxidative phosphorylation genes—*ATP5J2*, *PTCD1*, *ATP5S*, *NDUFA10*, and *NDUFC2/UQCRB*—were generated (Fig 4E) and luciferase reporter assays were performed. Overexpression of miR-106a-5p inhibited the luciferase activity of these constructs in a dose-dependent manner (Fig 4F), which indicated that miR-106a-5p expression could modulate the expression of mRNA targets that harbored these predicted binding sites.

DISCUSSION

We present AMLmiR36, a novel mutation and translocation-independent miRNA-based risk stratification scheme that can

identify at the time of diagnosis patients with adverse outcomes. We identified 61 candidate novel miRNAs that were expressed across our discovery cohort. Of these, six were DE between the NMF subgroups, which indicated that they may be important in defining biologic differences between pediatric AML subgroups. Moreover, we revealed nine candidate miRNAs that were DE between primary and relapse samples and three candidate miRNAs that were DE between primary and refractory samples, which indicated that they may be associated with treatment resistance.

We used miRNA expression profiles to identify four patient subgroups that were enriched for cytogenetic aberrations. Similar enrichment was observed in miRNA adult AML.²⁷ We also identified additional associations of miRNA expression with specific cytogenetic alterations (Data Supplement), including those that involved established indicators of outcome—for example, t(8;12), *CEBPA* mutation, and *FLT3*-ITD.

Ideally, a robust prognostic biomarker of treatment outcomes would identify a group of patients who were at sufficiently high risk of relapse and treatment resistance early enough in the course of treatment to justify consideration of alternate therapies. Here, we present AMLmiR36, a miRNA expression-based predictor of outcome that is independent of known cytogenetic and molecular risk classifiers. AMLmiR36 was developed and validated on a large cohort (N = 1,303) of primary patient samples. Of importance, AMLmiR36 was able to identify groups of high-risk patients from different clinical trials, independently of established indicators of outcome.

Our integrative miRNA:mRNA analysis provides a global view of candidate miRNA:mRNA interactions and pathways that are dysregulated in pediatric AML. In particular, we show that miRNAs that are abundantly expressed in treatment-resistant contexts seem to target genes that are involved in RNA processing, cellular signaling, and energy metabolism, and their reduced expression in relapse and refractory samples may support a role for quiescence in treatment resistance.³⁰

We show that abundant miR-106a expression was robustly associated with outcome, and its targets genes may be involved in oxidative phosphorylation, a process that is reduced in treatment-resistant cell line populations and consequently results in a quiescent cell state.³⁰ Such quiescent cells may evade therapies that are selective for rapidly dividing and proliferating cells.³¹ Thus, we speculate that abundant expression of miR-106a-363 could contribute to treatment resistance by repressing oxidative phosphorylation.

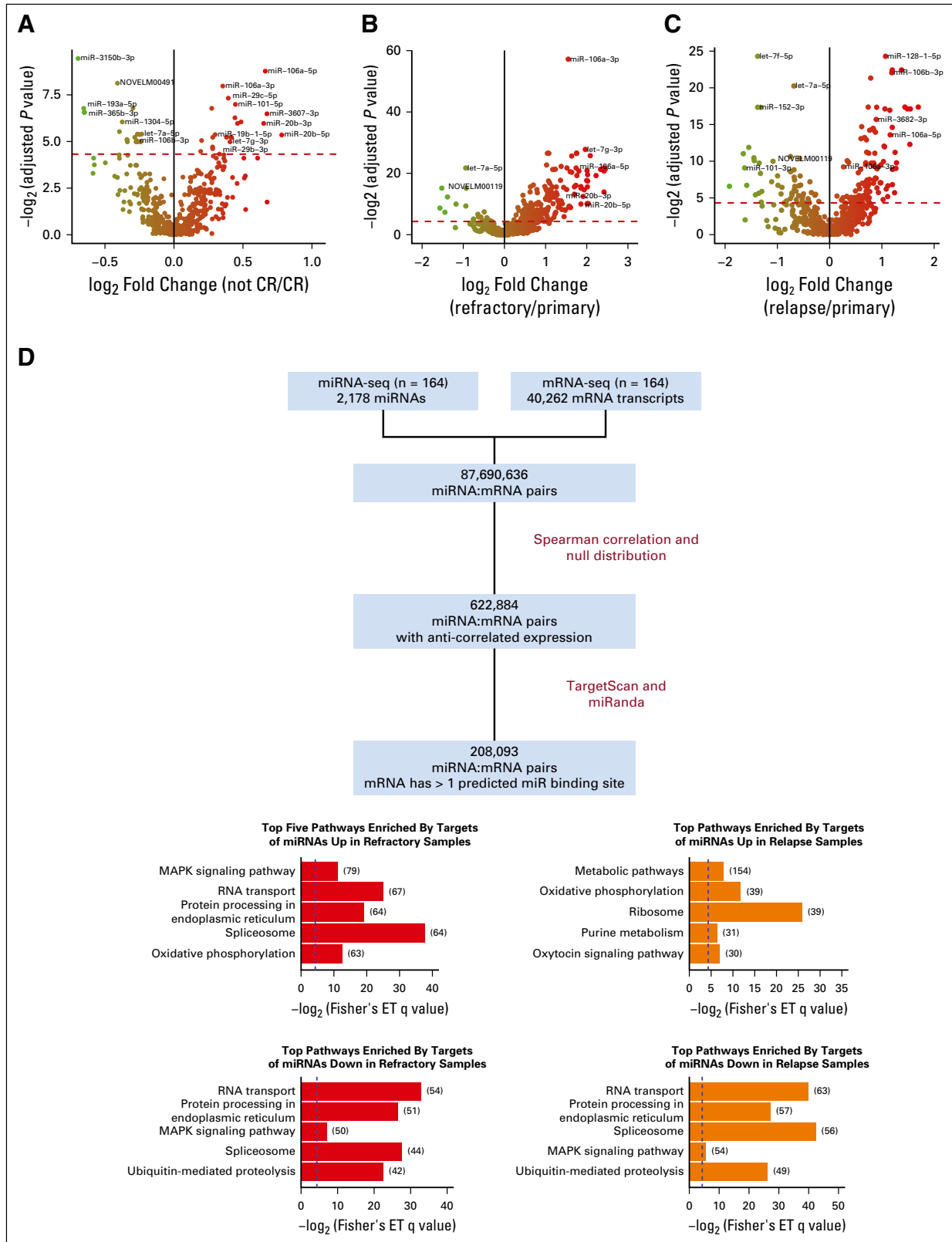


Fig 4. Integrative miRNA:mRNA analysis to identify putative targets of miRNAs. (A-C) Volcano plots displaying differentially expressed miRNAs (A) between primary samples from patients with refractory disease and primary samples from patients who achieved complete response; (B) between refractory samples and primary samples; and (C) between relapse and primary samples. Red dotted line represents the significant differential expression threshold ($q < 0.05$). (D) Workflow for miRNA:mRNA integrative analysis. Putative miRNA targets are those mRNA with expression patterns that are not correlated with the expression of their targeting miRNA and that harbor at least one predicted binding site for the targeting miRNA (left). Top five KEGG pathways enriched by targets of miRNAs that are significantly differentially expressed (right). (continued on next page)

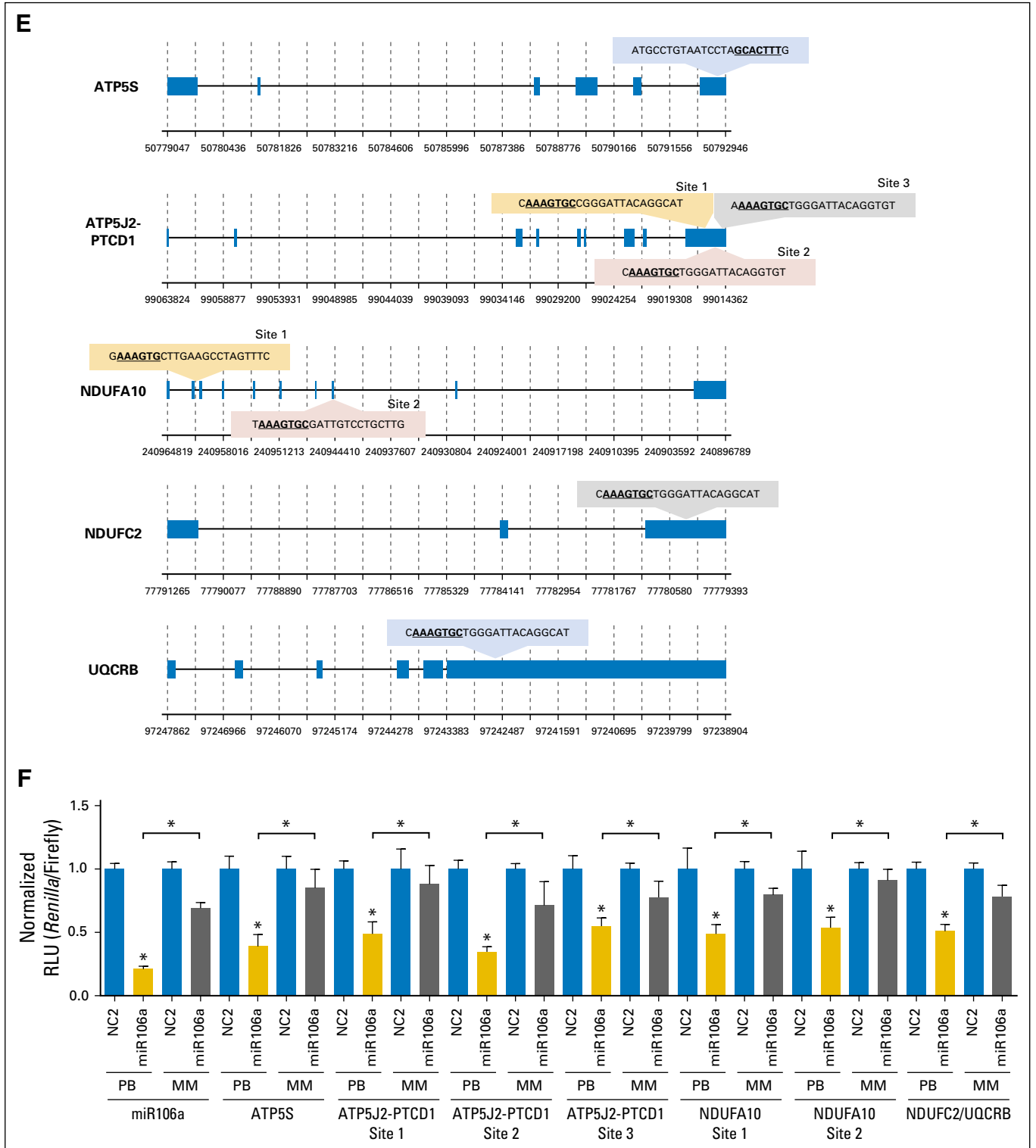


Fig 4. (Continued) The numbers of target genes that fall into each pathway are indicated in brackets. Dotted lines indicate the q value significance threshold, set at 0.05. (E) miR-106a-5p binding sites predicted by both TargetScan and miRanda on candidate target genes that are involved in oxidative phosphorylation: *ATP5S*, *ATP5J2-PTCD1*, *NDUFA10*, *NDUFC2*, and *UQCRB*. (F) miR-106a-5p activity in HEK-293 cells was assessed by using a psiCHECK2 dual luciferase reporter construct that contained each of the putative *ATP5S*, *ATP5J2-PTCD1*, *NDUFA10*, or *NDUFC2/UQCRB* binding sites. Activity is measured as *Renilla* luminescence normalized to Firefly luminescence to control for transfection efficiencies. Data are shown as normalized relative luciferase units (RLU) with respect to the corresponding dose of the control mimic and are representative of three independent experiments (means \pm SEM). Statistically significant comparisons between the cotransfected miR-106a-5p miRNA (20 pmol) and the NC2 control for the perfect binding reporter vector are noted over the solid colored bars. Statistical significance between perfect binding and mismatch constructs are indicated above the comparisons. * $P < .05$. Blue bars: NC2 negative control mimics; gold bars: miR-106a-5p mimics on perfect binding (PB) sites; gray bars: miR-106a-5p mimics on mismatched (MM) sites. ET, Fisher's exact test; MAPK, mitogen-activated protein kinase; seq, sequencing.

Overall, to our knowledge, we provide the first comprehensive sequence-based miRNA expression profiles of primary relapse and refractory pediatric AML, thus cataloguing the full repertoire of known and novel miRNAs. We also report a miRNA-based predictive model and identify miR-106a-363 as a miRNA that may be useful in future explorations of the biology that underpins treatment-resistant pediatric AML.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at jco.org.

REFERENCES

1. Tarlock K, Meshinchi S: Pediatric acute myeloid leukemia: Biology and therapeutic implications of genomic variants. *Pediatr Clin North Am* 62:75-93, 2015
2. Schuback HL, Arceci RJ, Meshinchi S: Somatic characterization of pediatric acute myeloid leukemia using next-generation sequencing. *Semin Hematol* 50:325-332, 2013
3. Pui C-H, Carroll WL, Meshinchi S, et al: Biology, risk stratification, and therapy of pediatric acute leukemias: An update. *J Clin Oncol* 29:551-565, 2011
4. Lacayo NJ, Alonzo TA, Gayko U, et al: Development and validation of a single-cell network profiling assay-based classifier to predict response to induction therapy in paediatric patients with de novo acute myeloid leukaemia: A report from the Children's Oncology Group. *Br J Haematol* 162:250-262, 2013
5. Ostronoff F, Othus M, Gerbing RB, et al: NUP98/NSD1 and FLT3/ITD coexpression is more prevalent in younger AML patients and leads to induction failure: A COG and SWOG report. *Blood* 124:2400-2407, 2014
6. Daschkey S, Röttgers S, Giri A, et al: MicroRNAs distinguish cytogenetic subgroups in pediatric AML and contribute to complex regulatory networks in AML-relevant pathways. *PLoS One* 8: e56334, 2013
7. Zhang H, Luo X-Q, Zhang P, et al: MicroRNA patterns associated with clinical prognostic parameters and CNS relapse prediction in pediatric acute leukemia. *PLoS One* 4:e7826, 2009
8. Zhu C, Wang Y, Kuai W, et al: Prognostic value of miR-29a expression in pediatric acute myeloid leukemia. *Clin Biochem* 46:49-53, 2013
9. Wang Z, Hong Z, Gao F, et al: Upregulation of microRNA-375 is associated with poor prognosis in pediatric acute myeloid leukemia. *Mol Cell Biochem* 383:59-65, 2013
10. Lin X, Wang Z, Zhang R, et al: High serum microRNA-335 level predicts aggressive tumor

progression and unfavorable prognosis in pediatric acute myeloid leukemia. *Clin Transl Oncol* 17: 358-364, 2015

11. Xu L-H, Guo Y, Cen J-N, et al: Overexpressed miR-155 is associated with initial presentation and poor outcome in Chinese pediatric acute myeloid leukemia. *Eur Rev Med Pharmacol Sci* 19:4841-4850, 2015
12. Ramamurthy R, Hughes M, Morris V, et al: miR-155 expression and correlation with clinical outcome in pediatric AML: A report from Children's Oncology Group. *Pediatr Blood Cancer* 63: 2096-2103, 2016
13. Lim EL, Trinh DL, Scott DW, et al: Comprehensive miRNA sequence analysis reveals survival differences in diffuse large B-cell lymphoma patients. *Genome Biol* 16:18, 2015
14. Cancer Genome Atlas Research Network; Ley TJ, Miller C, Ding L, et al: Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 368:2059-2074, 2013
15. Ho PA, Kutny MA, Alonzo TA, et al: Leukemic mutations in the methylation-associated genes DNMT3A and IDH2 are rare events in pediatric AML: A report from the Children's Oncology Group. *Pediatr Blood Cancer* 57:204-209, 2011
16. Gamis AS, Alonzo TA, Meshinchi S, et al: Gemtuzumab ozogamicin in children and adolescents with de novo acute myeloid leukemia improves event-free survival by reducing relapse risk: Results from the randomized phase III Children's Oncology Group trial AAML0531. *J Clin Oncol* 32:3021-3032, 2014
17. Cooper TM, Franklin J, Gerbing RB, et al: AAML03P1, a pilot study of the safety of gemtuzumab ozogamicin in combination with chemotherapy for newly diagnosed childhood acute myeloid leukemia: A report from the Children's Oncology Group. *Cancer* 118:761-769, 2012
18. Lange BJ, Smith FO, Feusner J, et al: Outcomes in CCG-2961, a Children's Oncology Group phase 3 trial for untreated pediatric acute myeloid leukemia: A report from the children's oncology group. *Blood* 111:1044-1053, 2008

AUTHOR CONTRIBUTIONS

Conception and design: Emilia L. Lim, Richard Aplenc, Lillian Sung, E. Anders Kolb, Alan Gamis, Malcolm Smith, Daniela S. Gerhard, Soheil Meshinchi, Marco A. Marra

Collection and assembly of data: Rhonda E. Ries, Maya Hughes, Andrew J. Mungall, Richard Moore, Yongjun Zhao, Todd A. Alonzo

Data analysis and interpretation: Emilia L. Lim, Diane L. Trinh, Jim Wang, Robert B. Gerbing, Yussanne Ma, James Topham, Erin Pleasance, Todd A. Alonzo, Marco A. Marra

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

19. Kozomara A, Griffiths-Jones S: miRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 42:D68-D73, 2014

20. National Cancer Institute: TARGET Data Matrix. <https://target.nci.nih.gov/dataMatrix/>

21. National Cancer Institute: The Next Generation Cancer Knowledge Network. <https://gdc.cancer.gov/>

22. Benjamini Y, Hochberg Y: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B (Method)* 57:289-300, 1995

23. Gaujoux R, Seoighe C: A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* 11:367, 2010

24. Therneau TM, Grambsch PM: *Modeling Survival Data: Extending the Cox Model*. New York, NY, Springer, 2000 doi:10.1007/978-1-4757-3294-8

25. Camp RL, Dolled-Filhart M, Rimm DL: X-tile: A new bioinformatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* 10:7252-7259, 2004

26. Friedman J, Hastie T, Tibshirani R: Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 33:1-22, 2010

27. Marcucci G, Mrózek K, Radmacher MD, et al: The prognostic and functional role of microRNAs in acute myeloid leukemia. *Blood* 117:1121-1129, 2011

28. Emmrich S, Engeland F, El-Khatib M, et al: miR-139-5p controls translation in myeloid leukemia through EIF4G2. *Oncogene* 35:1822-1831, 2016

29. Levis M, Small D: FLT3: ITDoes matter in leukemia. *Leukemia* 17:1738-1752, 2003

30. Lagadinou ED, Sach A, Callahan K, et al: BCL-2 inhibition targets oxidative phosphorylation and selectively eradicates quiescent human leukemia stem cells. *Cell Stem Cell* 12:329-341, 2013

31. Braess J, Jahns-Streubel G, Schoch C, et al: Proliferative activity of leukaemic blasts and cytosine arabinoside pharmacodynamics are associated with cytogenetically defined prognostic subgroups in acute myeloid leukaemia. *Br J Haematol* 113: 975-982, 2001

Affiliations

Emilia L. Lim, Diane L. Trinh, Yussanne Ma, James Topham, Erin Pleasance, Andrew J. Mungall, Richard Moore, Yongjun Zhao, and Marco A. Marra, Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency; Marco A. Marra, University of British Columbia, Vancouver, British Columbia; Lillian Sung, The Hospital for Sick Children, Toronto, Ontario, Canada; Rhonda E. Ries, Maya Hughes, and Soheil Meshinchi, Fred Hutchinson Cancer Research Center; Rhonda E. Ries, Maya Hughes, and Soheil Meshinchi, University of Washington, Seattle, WA; Jim Wang, Robert B. Gerbing, E. Anders Kolb, Alan Gamis, and Todd A. Alonzo, Children's Oncology Group, Monrovia; Todd A. Alonzo, University of Southern California, Los Angeles, CA; Richard Aplenc, The Children's

Hospital of Philadelphia, Philadelphia, PA; **Malcolm Smith and Daniela S. Gerhard**, Office of Cancer Genomics, National Cancer Institute, Bethesda, MD; and **Robert J. Arceci**, Phoenix Children's Hospital, Phoenix, AZ.

Support

Funded in whole or in part by the National Institutes of Health, National Cancer Institute (Grant No. HHSN261200800001E), the Children's Oncology Group Statistics and Data Center (Grants No. U10CA180899 and U10CA098413), the National Clinical Trials Network Operations Center (Grant No. U10CA180886), and from Kimberley Beare and Nathaniel Stoffelsma through the BC Cancer Foundation, and the St Baldrick's Foundation Consortium Grant. E.L.L. is supported by a Canadian Institutes of Health Research Doctoral Award and a University of British Columbia Four-Year Fellowship. M.A.M. holds the University of British Columbia Canada Research Chair in Genome Science and is supported by the BC Cancer Foundation, Genome Canada, Genome British Columbia, the Cancer Research Society, the Leukemia and Lymphoma Society of Canada, and the Canadian Institutes of Health Research (Grant No. FDN-143288).



Oncology Research Just Got Easier!



Visit the new ASCO Journals website today and take advantage of:

- Better navigation and readability
- Improved search
- More relevant content—powered by TrendMD
- Altmetric badges

Visit the new ASCO Journals website today at ascopubs.org

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST**MicroRNA Expression-Based Model Indicates Event-Free Survival in Pediatric Acute Myeloid Leukemia**

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/site/ifc.

Emilia L. Lim

No relationship to disclose

Diane L. Trinh

No relationship to disclose

Rhonda E. Ries

No relationship to disclose

Jim Wang

No relationship to disclose

Robert B. Gerbing

Stock or Other Ownership: Pfizer

Yussanne Ma

No relationship to disclose

James Topham

No relationship to disclose

Maya Hughes

Employment: The Everett Clinic

Erin Pleasance

No relationship to disclose

Andrew J. Mungall

No relationship to disclose

Richard Moore

No relationship to disclose

Yongjun Zhao

No relationship to disclose

Richard Aplenc

Honoraria: Sigma-Tau

Travel, Accommodations, Expenses: Sigma-Tau

Lillian Sung

No relationship to disclose

E. Anders Kolb

No relationship to disclose

Alan Gamis

Consulting or Advisory Role: Novartis

Malcolm Smith

No relationship to disclose

Daniela S. Gerhard

No relationship to disclose

Todd A. Alonzo

No relationship to disclose

Soheil Meshinchi

No relationship to disclose

Marco A. Marra

No relationship to disclose

Acknowledgment

We thank the Library Construction, Biospecimen, Sequencing, and Bioinformatics teams at Canada's Michael Smith Genome Sciences Centre for expert technical assistance. We are grateful for expert project management assistance provided by Karen Novik, PhD, and for mentorship and guidance provided by Robert Camfield, PhD. This manuscript is a report from the Children's Oncology Group, and we especially acknowledge the late Robert Arceci, MD, for his leadership in the design of this study. The results published here are in whole or part based on data generated under the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) project managed by the National Cancer Institute. Information about TARGET can be found at <https://ocg.cancer.gov/>. The data generated for this analysis are available via dbGaP (<https://www.ncbi.nlm.nih.gov/gap>) under accession number phs000465 and also via the TARGET Data Matrix at the TARGET Data Coordinating Center (http://target.nci.nih.gov/dataMatrix/TARGET_DataMatrix.html). We thank the TARGET Data Coordination Center team, Leandro Hermida, Patee Gesuwan and Tanja Davidsen, as well as Guidry-Auvil, MD, for TARGET management.