Genome-wide identification of human microRNAs located in leukemia-associated genomic alterations

Daniel T. Starczynowski,[1,2] Ryan Morin,[1] Andrew McPherson,[1] Jeff Lam,[1,2] Raj Chari,[1,2] Joanna Wegrzyn,[1,2] Florian Kuchenbauer,[1] Martin Hirst,[1] Kaoru Tohyama,[3] R. Keith Humphries,[1] Wan L. Lam,[1,2] Marco Marra,[1] and Aly Karsan[1,2]

[1]British Columbia Cancer Agency Research Centre, Vancouver, BC; [2]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC; and [3]Clinical Pathology and Laboratory Medicine, Kawasaki Medical School, Kurashiki, Okayama, Japan

Cytogenetic alterations, such as amplifications, deletions, or translocations, contribute to myeloid malignancies. MicroRNAs (miRNAs) have emerged as critical regulators of hematopoiesis, and their aberrant expression has been associated with leukemia. Genomic regions containing sequence alterations and fragile sites in cancers are enriched with miRNAs; however, the relevant miRNAs within these regions have not been evaluated on a global basis. Here, we investigated miRNAs relevant to acute myeloid leukemia (AML) by (1) mapping miRNAs within leukemia-associated genomic alterations in human AML cell lines by high-resolution genome arrays and (2) evaluating absolute expression of these miRNAs by massively parallel small RNA sequencing. Seventy-seven percent (542 of 706) of miRNAs mapped to leukemia-associated copy-number alterations in the cell lines; however, only 18% (99 of 542) of these miRNAs are expressed above background levels. As evidence that this subset of miRNAs is relevant to leukemia, we show that loss of 2 miRNAs identified in our analysis, miR-145 and miR-146a, results in leukemia in a mouse model. Small RNA sequencing identified 28 putative novel miRNAs, 18 of which map to leukemia-associated copy-number alterations. This detailed genomic and small RNA analysis points to a subset of miRNAs that may play a role in myeloid malignancies. (*Blood*. 2011;117(2):595-607)

Introduction

Small noncoding RNAs are conserved and encoded in the genomes of invertebrates, vertebrates, and plants.[1] The largest subset of naturally occurring small RNAs are microRNAs (miRNAs). Mature miRNAs are 19-24 nucleotide transcripts processed from precursor hairpin intermediate RNAs by endonuclease-mediated reactions.[1] miRNAs have been implicated in critical hematopoietic processes, and their deregulation is associated with leukemogenesis. Functional validation of deregulated miRNAs in hematopoiesis has been shown for several miRNAs[2]; however, direct evidence that leukemogenesis is specifically mediated by miRNAs is lacking. In one example, overexpression of miR-155 results in a fatal and aggressive myeloproliferative disorder in mice.[3] In other examples, knockout or overexpression of an miRNA has been shown to deregulate hematopoietic processes related to ≥ 1 steps of leukemogenesis.[2]

Hematologic malignancies often exhibit genomic alterations. Balanced and unbalanced chromosome alterations are found in patients with acute myeloid leukemia (AML). Many miRNA expression studies have been performed to identify differentially expressed miRNAs between normal and leukemic samples.[4-6] These studies are extremely valuable, but they have not examined the direct relationship between genome-wide alterations and the miRNAs within these alterations. Furthermore, these studies, which rely on the annotated sequence information about the miRNAs, are not capable of identifying novel or variant miRNAs.

Analysis of human and mouse genomes reveals that miRNA genes are frequently located at fragile sites and regions of copy number alteration (CNA) associated with cancer.[7] The current approximation states that 50% of miRNAs are within cancer-associated genomic regions or in fragile sites.[7] These findings suggest that a mechanism of miRNA deregulation in oncogenesis is due to genomic instability. Although miRNAs have been mapped to common leukemia-associated alterations with the use of database searches, the potentially relevant miRNAs within these alterations have not been thoroughly investigated. In this study we addressed 2 main objectives: (1) identification of miRNAs that map to common leukemia-associated genomic alterations and (2) identification of relevant miRNAs within the leukemia-associated genomic alterations. We investigated miRNAs relevant to leukemogenesis by first mapping miRNAs within common leukemia-associated genomic alterations in 6 human AML cell lines by high-resolution array comparative genomic hybridization (CGH). We then determined absolute expression of these miRNAs by massively parallel small RNA sequencing. Although 77% (542 of 706) of miRNAs mapped to leukemia-associated CNAs in the 6 cell lines, only 18% (99 of 542) of these miRNAs are expressed at levels above background. In support of our conclusion, knockdown of miR-145 and miR-146a, 2 miRNAs mapping to a commonly deleted region in AML, results in a long-latency myeloid leukemia in mice. This study provides a detailed genomic and miRNA expression analysis of human leukemic cell lines to identify an enriched subset of leukemia-associated miRNAs and facilitates selection of miRNAs to investigate in human leukemogenesis or normal hematopoiesis.

Submitted March 29, 2010; accepted October 6, 2010. Prepublished online as *Blood* First Edition paper, October 20, 2010; DOI 10.1182/blood-2010-03-277012.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

Methods

Cell lines and CD34$^+$ cells

Acute myeloid leukemic cell lines, KG-1, KG-1a, UT-7, HL-60, and THP-1, were purchased from ATCC. The myelodysplastic cell line, MDS-L, was generated in the laboratory of one of the coauthors (K.T.). CD34$^+$ cells were positively selected from cryopreserved marrow or peripheral blood cells by immunomagnetic separation (Stem Cell Technologies).

Whole genome tiling-path aCGH analysis

Details of whole genome array construction and probe prime labeling and hybridization have been described previously.[8,9] The submegabase-resolution tiling set array contains 32 433 overlapping BAC-derived DNA segments that provide tiling coverage over the entire human genome with a theoretical resolution of 50-100 kb.[8] All clones were spotted in duplicate. Sample (cell line) and reference (normal diploid) genomic DNA (50-200 ng of each) were separately labeled with Cyanine 3 and Cyanine 5 deoxycytidine triphosphate fluorescence markers, respectively. The images were captured with a charge-coupled device camera and were analyzed with an ArrayWorx scanner and SoftWorx Tracker Spot Analysis software (Applied Precision). SeeGH custom software was used to visualize all data as log ratio plots.[10] Clones with standard deviation between duplicate spots of > 0.1 were filtered from the raw data. To avoid false-positives due to hybridization noise, a minimum of 2 overlapping consecutive clones showing change was required for a region to be considered altered. Breakpoints of genomic alterations were identified with a hidden Markov model algorithm and were verified by visual inspection.

Small RNA library preparation

Leukemia cell lines were individually harvested, and RNA was extracted with TRIzol (Invitrogen). The extracted RNA was subjected to miRNA library construction with the use of the Illumina sequencing platform according to a published protocol.[11] The sequencing library of small RNAs generated from human CD34$^+$ bone marrow cells was obtained from a recently published analysis.[12]

miRNA isolation and expression analysis

The small RNA fraction was isolated with the mirVana Paris Isolation kit (Ambion). miRNA expression was quantified with miRNA-specific stem-loop primers. The small RNA fraction was used for reverse transcription with the use of reverse transcription primers specific for each independent miRNA. The cDNA synthesis reaction was subsequently used for quantitative polymerase chain reaction (PCR) with the use of miRNA-specific primers

Annotation and prediction of novel miRNAs

The annotation procedure was performed as described but used annotations form miRBase version 14 and the human genome (National Center for Biotechnology Information [NCBI] build 36.3). Novel miRNAs were predicted as previously described.[11] Predicted target genes were identified with the use of TargetScan 5.1. For novel or edited miRNAs, the target genes were predicted by TargetScan Custom 5.1. Binding energies for the miRNA to its mRNA targets were calculated with IntaRNA.[13]

miRNA decoy retroviral vectors, packaging cell lines, and bone marrow transplantation

miRNA decoy sequences (tandem repeats each complementary to miR-145 and miR-146a, respectively) were fused to the 3′ untranslated region (UTR) of the yellow fluorescent protein (YFP) cDNA and then cloned into the dual promoter phosphoglycerate retroviral vector as previously reported.[14] Virus packaging and infection of ecotropic packaging cell lines (GP$^+$E86) were performed as previously described.[14] Marrow transplantation studies were carried out with protocols approved by the University of British Columbia Animal Care Committee.

Peripheral blood and bone marrow analysis

Donor-derived engraftment and reconstitution were monitored by flow cytometric analysis of YFP expression in the peripheral blood. For immunophenotypic analysis, bone marrow cells or peripheral blood were washed and resuspended in phosphate-buffered saline (PBS) containing 4% goat serum, followed by primary monoclonal antibody (labeled with phycoerythrin or allophycocyanin) staining overnight. Details on antibodies are described elsewhere.[14] Samples were run on a FACSCalibur flow cytometer (Beckman Coulter), and data were analyzed with the use of FlowJo software (Version 8.7; TreeStar Inc). Complete blood counts were performed on the peripheral blood with the use of Scil Vet ABC Hematology Analyzer (Scil Animal Care Company). Blood counts were obtained at time of death. Organs and tissues were fixed in PBS with 4% paraformaldehyde, embedded in paraffin, sectioned, and stained with hematoxylin and eosin.

Microarray analysis of miRNAs

The expression data analyzed in this study were accessed at ArrayExpress (www.ebi.ac.uk/microarray-as/ae). The accession numbers for controls ($n = 11$) and AML patient samples ($n = 54$, normal karyotype; $n = 4$, del(5q)) are E-TABM-970 and E-TABM-405, respectively. Analyses were performed with BRB-ArrayTools developed by Dr Richard Simon and BRB-ArrayTools Development Team.[15]

Retroviral transduction of miR-145 and miR-146a into HL-60 myeloid leukemia cells

Amphophoenix cells were transfected with murine stem cell virus–internal ribosome entry site–green fluorescent protein (GFP) (MIG), MIG–miR-145, or MIG–miR-146a constructs. Constructs have been previously described.[14] Virus was harvested and used to infect HL-60 cells. Transduced cell lines were sorted for GFP expression and cultured in Dulbecco minimal essential medium/10% fetal bovine serum.

Apoptosis analysis

Approximately 1×10^5 HL-60 cells were washed in PBS, and resuspended in annexin V–binding buffer (10mM HEPES [N-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid], 140mM NaCl, 2.5mM $CaCl_2$; pH 7.4) and annexin V–conjugated antibody (1:20). After a 15-minute incubation, an additional 500 μL of annexin V–binding buffer was added and analyzed by flow cytometry.

Luciferase assay

hsa-miR-1, hsa-miR-1 (4A:G) mutant, hMEIS1-3′-UTR, and hTP53-3′-UTR were synthesized and cloned into pcDNA3.1 or pMirReport, respectively. A 95-base pair (bp) fragment (134-233) of the MEIS1 3′-UTR encompassing the miR-1 target site or 95-bp fragment (634-735) of the TP53 3′-UTR encompassing the miR-1 (4A:G) predicted target site were inserted downstream of the open reading frame of the luciferase reporter gene. For the 3′-UTR luciferase assays, 50 ng of pMirReport-3′-UTR, 250 ng of pcDNA or pcDNA-miR, and 7.5 ng of thymidine kinase-driven *Renilla* luciferase were cotransfected into human embryonic kidney 294 (HEK293) cells (24-well format) with the use of TransiT transfection reagent (Mirus).

5-Azacytidine treatment of myeloid leukemia cell lines

HL-60, KG-1a, and THP-1 cells were cultured in Dulbecco minimal essential medium/10% fetal bovine serum in 35-mm plates for 24 hours. After plating, cells were treated with 10μM 5-azacytidine (5-aza; Sigma) or dimethyl sulfoxide for 24 hours. Small RNAs were collected with miRVANA Paris (Ambion) isolation kit, and miR-145 and miR-146a were measured by quantitative PCR.

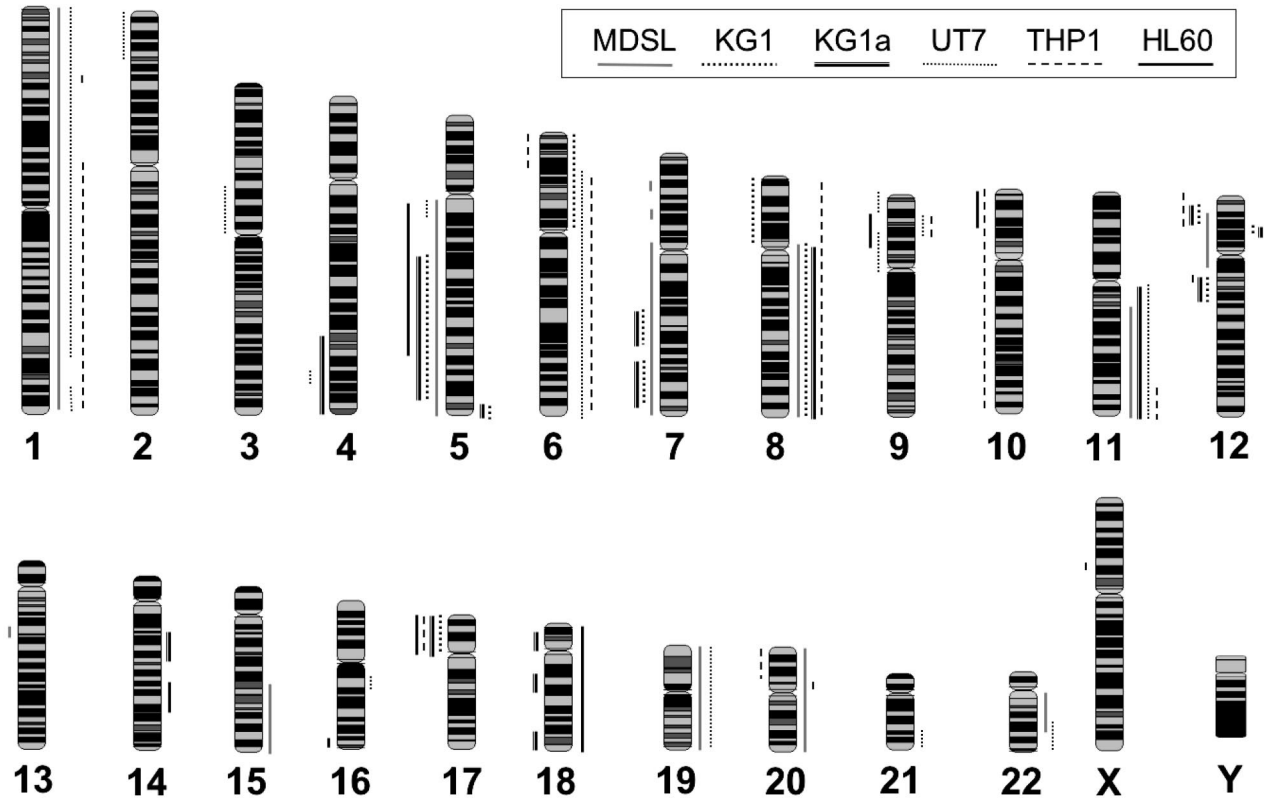

Figure 1. Summary of cytogenetic alterations in 6 human leukemia cell lines by aCGH. Whole-genome frequency distribution of chromosomal alterations in 6 human leukemia cell lines as detected by array CGH and visualized by SeeGH. Lines to the right of the chromosome indicate gain of chromosomal material. Lines to the left of the chromosome indicate loss of chromosomal material. Line patterns correspond to the indicated cell line.

Statistics

Results are depicted as the mean ± standard error of the mean. Statistical analyses were performed with the Student *t* test. Comparison of survival between different groups was done by the Kaplan-Meier test, and *P* value was calculated by the log-rank test. Box-and-whisker plots were used in Figure 5C to depict the range and percentiles of values obtained. The box indicates the 75th (top), 50th (middle line; median), and 25th (bottom) percentiles of the values obtained. GraphPad Prism 4 (GraphPad) was used for statistical analysis.

Results

Mapping of CNAs in human leukemic cell lines by whole genome high-resolution array CGH

We obtained 6 human cell lines representing de novo and secondary acute myeloid malignancies: THP-1 (acute monocytic leukemia), UT-7 (acute megakaryoblastic leukemia), KG-1 (acute myeloid leukemia with myeloblast-like features), KG-1a (a subline of KG-1 that has acquired new karyotypic markers), MDSL (refractory anemia with excess blasts), and HL-60 (promyelocytic leukemia). These cell lines were chosen as models of common AML subtypes and for their large number of cytogenetic abnormalities.[16] To precisely define genomic breakpoints in these leukemic model cell lines, we used high-resolution aCGH.[8] Comprehensive genomic profiles of copy number gains and losses obtained from 6 leukemic cell lines are summarized in supplemental Table 1 (available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article) and Figure 1. The median number of genomic alterations per cell line is 22 (range, 13-29). As expected, all cell lines exhibit genomic instability, including CNAs of all or part of chromosomal arms (Figure 1; supplemental Table 1). UT-7 exhibited the most genomic alterations (29 CNAs, and 1156 Mb of total alterations), whereas HL-60 had the least disrupted genome (429 Mb of total alterations) and KG-1 had the fewest distinct alterations (13 high level CNAs). The most recurrent alterations include gains of chromosomes 1, 6q, 8, 9p, 11q, and 19 and deletions of chromosomes 5q, 7q, 12p, and 17p. Recurring submegabase genomic alterations were also detected and include a deletion at 12q13.11-q13.12 and a gain at 12p12.1-p11.23 (Figure 1; supplemental Table 1). Most of these CNAs have been reported in patients with AML and are thought to contribute to leukemic initiation or progression.[17] Application of high-resolution aCGH provides a detailed and comprehensive analysis of genomic CNAs in 6 model leukemic cell lines.

Discovery of expressed microRNAs mapping to leukemia-associated genomic alterations by massively parallel sequencing of small RNAs

After determining the breakpoints and regions of CNAs in the leukemic cell lines, we next sought to identify miRNAs within the identified CNA. To map miRNAs to the DNA CNA in our leukemic cell lines, we used the most recent human genome build (University of California Santa Cruz [UCSC] Human Mar. 2006) and miRNA database (miRBase 14.0) (supplemental Table 2). According to the current miRNA database, there are 706 human miRNAs.[18] Of these, 542 (77%) are found within leukemia-

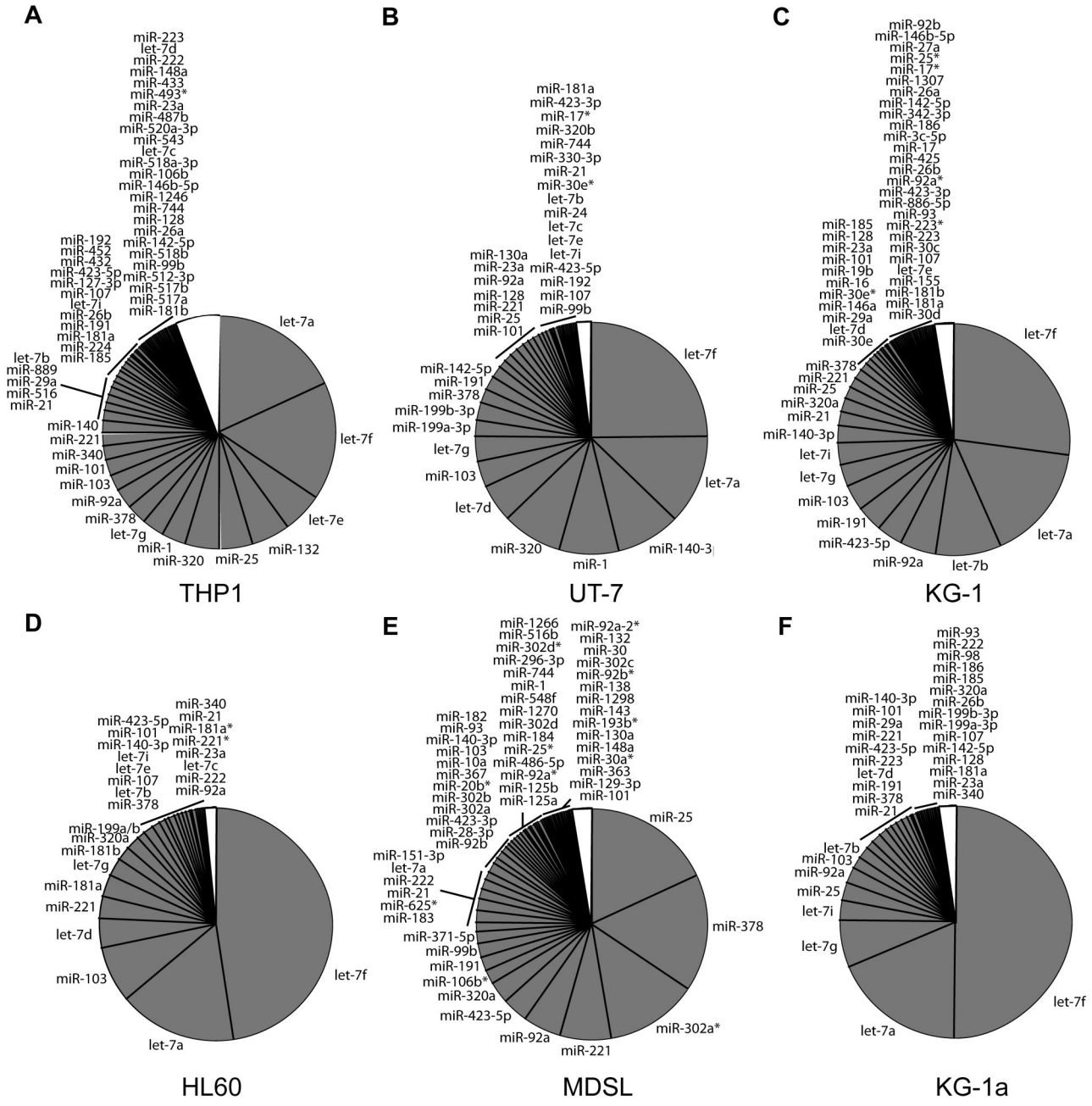

Figure 2. MicroRNA expression in 6 human leukemia cell lines. (A-F) Proportion (in percent) of the most abundantly expressed miRNAs in 6 human leukemia cell lines: (A) THP-1, (B) UT-7, (C) KG-1, (D) HL-60, (E) MDSL, and (F) KG-1a. White portion of the pie chart represents miRNAs reads that are expressed below background.

associated CNAs in the 6 cell lines we studied (see further discussion below).

To determine which of the 542 miRNAs are expressed and potentially relevant to AML, we applied massively parallel sequencing of small RNAs to each of the 6 cell lines. We selected this approach because all variants of mature miRNA transcripts can be identified.[19] In contrast, hybridization-based miRNA arrays rely on annotated miRNA sequences and therefore are not adequate to detect all miRNA variants. Small RNA fractions were isolated from each cell line and prepared for sequencing on the Illumina platform. Sequencing of small RNA libraries resulted in an average of 7.2 million reads (range, 4.5-9.5 million reads) from each of the cell lines. Sequences were mapped according to their overlap with available genomic annotations of miRNAs, resulting in an average of 0.9 million filtered reads from each cell line. The miRNA reads within the small RNA fraction of each cell line represent 40%-90% of all filtered reads (supplemental Tables 3-4). Because the cell lines exhibited numerous miRNAs with single reads, we decided to primarily focus on the miRNAs above an arbitrary background expression level, with a miRNA frequency above 0.001 (1 read per 1000 total reads; Figure 2). By this cutoff, $\sim$ 98% of all miRNA reads and 20% of unique miRNAs are included in our analysis from each cell line (Figure 2). miRNAs with fewer reads (less than a frequency of 0.001) were excluded from further analysis.

To include miRNAs relevant only to myeloid leukemic cells, we used the cell lines as controls to cross-compare absolute miRNA

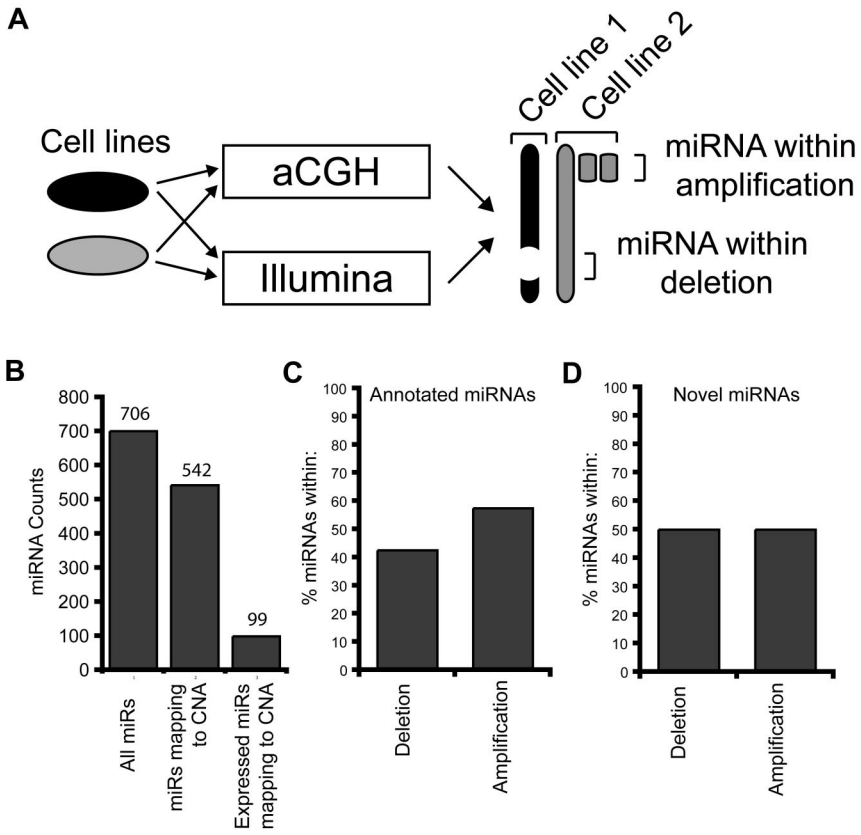

Figure 3. miRNA mapping to CNAs. (A) Schematic representation of the approach to identify miRNAs within CNAs. Each of the 6 cell lines were simultaneously analyzed for CNAs by aCGH and small RNA expression by massively parallel sequencing on the Illumina platform. The CNAs for the 6 cell lines were annotated (Table 1) and subsequently merged. Similarly, miRNA expression data from all 6 cell lines were used to identify expressed miRNAs that map to the CNAs (Table 2). By this approach, we were able to identify miRNAs that map to a homozygous deletion in one cell line by using miRNA expression values from the pooled data. (B) Distribution summary of miRNAs mapping to leukemia-associated genomic alterations found in 6 human leukemia cell lines. According to the current miRNA database (miRBase v13.0), there are 702 human miRNAs. Of these, 542 miRNAs map to CNAs identified in the 6 leukemia cell lines. After small RNA sequencing, 99 miRNAs (of 542) are expressed above background and map to CNAs. (C-D) Proportion (in percent) of the 99 expressed miRNAs within each cell line mapping to genomic regions of amplifications or deletions.

levels within genomic alterations, rather than nonleukemic controls (Figure 3A). By this approach, abundantly expressed miRNAs within genomic alterations that are common and relevant only to myeloid leukemia cells are identified (Figure 3A). miRNAs mapping to the leukemia-associated genomic alterations (n = 542) (supplemental Table 2) were next investigated for their absolute expression levels. Of these 542 miRNAs, 18% (n = 99) were expressed at levels above background (> 1 per 1000 expressed tags) in ≥ 1 of the cell lines (Figure 3B; Table 1). Nearly an equal distribution of miRNAs was mapped within deleted (51%) and amplified (49%) CNAs (Figure 3C-D). These findings suggest that, although many miRNAs map within CNAs, only a subset are expressed at notable levels in myeloid leukemic cell lines. We next wanted to confirm the relevance of this subset of miRNAs in primary AML samples. We used published gene expression data on miRNAs isolated from $CD34^+$ cells from 58 patients with AML and 11 controls.[4] Of 99 expressed miRNAs, 43 were evaluated by microarray analysis, of which 72% (n = 31) were significantly differentially expressed in patients with AML compared with controls (Table 1).

It is possible that tumor suppressor miRNAs may be transcriptionally silenced by mechanisms other than genomic deletions and unintentionally omitted in our analysis above. Therefore, we evaluated miRNAs within deleted regions in the leukemic cell lines and compared their expression levels with nonleukemic human $CD34^+$ cells. We used a recently published sequencing library of small RNAs generated from human $CD34^+$ bone marrow cells[12] (supplemental Table 4). Of the miRNAs mapping to deleted regions, 29 are expressed in human $CD34^+$ cells (Figure 4) and were further analyzed. We compared the expression of these 29 miRNAs between normal $CD34^+$ cells and each leukemic cell line and found that 65%-86% of miRNAs that map to leukemia-associated deleted regions were expressed lower in the cell lines compared with $CD34^+$ cells. A similar pattern of miRNA repression was observed in primary AML samples (n = 58) compared with control $CD34^+$ cells (n = 11) (Figure 4). Collectively, these findings suggest that a subset of miRNAs within deleted regions in leukemic cells correlates with reduced expression (Figure 4). Interestingly, a few miRNAs (eg, miR-378, miR-103-1, miR-25, and let-7f-2) are consistently overexpressed in the leukemic cell lines despite residing within a hemizygous deletion (Figure 4), potentially representing genes essential to cell maintenance.

miRNAs on chromosome 5q may be important in AML

To further explore the role of miRNAs within deleted regions in myeloid malignancies we focused on miRNAs on chromosome 5 (supplemental Figure 1). Deletion of chromosome 5q is the most common cytogenetic alteration in AML and myelodysplastic syndrome (MDS; Figure 1).[20] To evaluate the potential effect of del (5q) on miRNA expression, we compared the relative tag counts for 2 leukemic cell lines, one with a deletion of chromosome 5q (KG-1a) and one diploid at chromosome 5q (THP-1). Effect of the chromosome 5q deletion on miRNA expression is evident by the lower expression levels of miRNAs in KG-1a compared with THP-1 (Figure 5A). As an example, miR-145 and miR-146a, 2 miRNAs within a commonly deleted region in del (5q) myeloid malignancies (supplemental Figure 1), are expressed lower in cell lines with the chromosome 5q deletion or diploid at this locus compared with $CD34^+$ cells (Figures 4 and 5A). To independently confirm this observation, we performed quantitative PCR to access miR-145 and miR-146a expression levels in $CD34^+$ (n = 3), THP-1, and KG-1a cells. As shown in Figure 5A (inset), miR-145

Table 1. Expressed miRNAs mapping to copy number alterations

Chromosome	miR	Bp start	Bp end	*P**
Chr 1 (amplification)	miR-30e	40 s683 565	40 683 656	.08
	miR-30c-1	40 686 494	40 686 582	NA
	miR-101-1	65 296 705	65 296 779	.0008
	miR-186	71 305 902	71 305 987	.32
	miR-9-1	154 656 757	154 656 845	.26
	miR-181b-1	197 094 625	197 094 734	.0001
	miR-181a-1	197 094 796	197 094 905	.0084
	miR-29c	206 041 820	206 041 907	NA
Chr 3 (deletion)	miR-128-2	35 760 972	35 761 055	.0001
	miR-26a-1	37 985 899	37 985 975	.0001
	miR-138-1	44 130 708	44 130 806	.11
	miR-425	49 032 585	49 032 671	NA
	miR-191	49 033 055	49 033 146	.95
	let-7g	52 277 334	52 277 417	.0001
	miR-15b	161 605 070	161 605 167	NA
	miR-16-2	161 605 227	161 605 307	.042
	miR-28	189 889 263	189 889 348	NA
Chr 5 (deletion)	miR-9-2	87 998 427	87 998 513	NA
	miR-886	135 444 076	135 444 196	NA
	miR-143	148 788 674	148 788 779	NA
	miR-378	149 092 581	149 092 646	NA
	miR-146a	159 844 937	159 845 035	.0008
	miR-103-1	167 920 487	167 920 548	NA
	miR-340	179 374 909	179 375 003	.0001
Chr 6 (amplification)	miR-30a	72 169 975	72 170 045	NA
Chr 7 (deletion)	miR-148a	25 956 064	25 956 131	.0001
	miR-25	99 529 119	99 529 202	NA
	miR-93	99 529 327	99 529 406	NA
	miR-106b	99 529 552	99 529 633	NA
	miR-129-1	127 635 161	127 635 232	NA
	miR-182	129 197 459	129 197 568	.0002
	miR-183	129 201 981	129 202 090	.0035
	miR-29a	130 212 046	130 212 109	NA
	miR-26b-1	130 212 758	130 212 838	NA
Chr 8 (amplification)	miR-320a	22 158 420	22 158 501	NA
	miR-486	41 637 116	41 637 183	NA
	miR-30d	135 886 301	135 886 370	NA
	miR-151	141 811 845	141 811 934	.057
Chr 9	miR-101-2	4 840 297	4 840 375	.54
	miR-199b	130 046 821	130 046 930	NA
Chr 10 (deletion)	miR-146b	104 186 259	104 186 331	.0001
	miR-1307	105 144 000	105 144 148	NA
Chr 11 (amplification)	miR-130a	57 165 247	57 165 335	.0001
	miR-192	64 415 185	64 415 294	0.68
	miR-34c	110 889 374	110 889 450	.0001
	let-7a-2	121 522 440	121 522 511	.0001
Chr 12 (deletion)	miR-200c	6 943 123	6 943 190	.027
	miR-26a-2	56 504 659	56 504 742	.0025
	let-7i	61 283 733	61 283 816	.045
Chr 14 (amplification)	miR-342	99 645 745	99 645 843	.0024
	miR-493	100 405 150	100 405 238	NA
	miR-433	100 417 976	100 418 068	NA
	miR-127	100 419 069	100 419 165	.0001
	miR-543	100 568 077	100 568 154	NA
	miR-495	100 569 845	100 569 926	NA
	miR-487b	100 582 545	100 582 628	NA
	miR-889	100 583 991	100 584 069	NA
Chr 15 (amplification)	miR-184	77 289 185	77 289 268	.76
	miR-9–3	87 712 252	87 712 341	NA
Chr 17 (deletion)	miR-132	1 899 952	1 900 052	.67
	miR-744	11 925 941	11 926 038	NA
Chr 18 (amplification)	miR-122	54 269 286	54 269 370	.21

Bp indicates base pair; as, antisense; and NA, not available.

*Value was generated from microarray data when comparing miRNA expression changes between AML patient (n = 58) and control $CD34^+$ samples (n =11).

Table 1. Expressed miRNAs mapping to copy number alterations (continued)

Chromosome	miR	Bp start	Bp end	*P**
Chr 19 (amplification)	miR-199a-1	10 789 102	10 789 172	NA
	miR-24–2	13 808 101	13 808 173	NA
	miR-27a	13 808 254	13 808 331	NA
	miR-23a	13 808 401	13 808 473	NA
	miR-1270	20 371 080	20 371 162	NA
	miR-330	50 834 092	50 834 185	.013
	miR-99b	56 887 677	56 887 746	NA
	let-7e	56 887 851	56 887 929	.52
	miR-125a	56 888 319	56 888 404	.0001
	miR-512-1	58 861 745	58 861 828	NA
	miR-512-2	58 864 223	58 864 320	NA
	miR-1323	58 867 034	58 867 106	NA
	miR-520a	58 885 947	58 886 031	NA
	miR-518b	58 897 803	58 897 885	NA
	miR-517a	58 907 334	58 907 420	NA
	miR-517b	58 916 142	58 916 208	NA
	miR-516b-2	58 920 508	58 920 592	NA
	miR-518a-1	58 926 072	58 926 156	NA
	miR-518a-2	58 934 399	58 934 485	NA
	miR-371	58 982 741	58 982 807	.0001
Chr 20 (amplification)	miR-103-2	3 846 141	3 846 218	.086
	miR-103-2-as	3 846 149	3 846 210	NA
Chr 21 (amplification)	let-7c	16 834 019	16 834 102	.0015
	miR-125b-2	16 884 428	16 884 516	.0001
	miR-155	25 868 163	25 868 227	.0001
Chr 22 (amplification)	miR-185	18 400 662	18 400 743	.58
	miR-130b	20 337 593	20 337 674	.0001
chr X (deletion)	miR-221	45 490 529	45 490 638	.02
	miR-222	45 491 365	45 491 474	.0001
	miR-98	53 599 909	53 600 027	NA
	let-7f-2	53 600 878	53 600 960	NA
	miR-1298	113 855 906	113 856 017	NA
	miR-363	133 131 074	133 131 148	NA
	miR-91a-2	133 131 234	133 131 308	NA
	miR-19b-2	133 131 367	133 131 462	NA
	miR-20b	133 131 505	133 131 573	NA

Bp indicates base pair; as, antisense; and NA, not available.

*Value was generated from microarray data when comparing miRNA expression changes between AML patient (n = 58) and control $CD34^+$ samples (n = 11).

and miR-146a levels are highest in $CD34^+$ cells and lower in leukemic cell lines with del (5q) or diploid (5q). Further, miR-145 and miR-146a are expressed significantly lower in diploid chromosome 5q leukemic cell lines compared with $CD34^+$ cells ($P = .009$ and $P = .02$, respectively; Figure 5A inset). This suggests that miR-145 or miR-146a or both may be silenced in leukemia with a diploid chromosome 5q and further depressed in leukemias with del (5q). To investigate a potential epigenetic suppression (such as by DNA methylation) of miR-145 and miR-146a in AML, we treated 3 AML cell lines with a demethylating agent (5-aza) and examined miRNA expression. As shown in Figure 5B, expression of miR-145 and miR-146a increase ≥ 3-fold after 5-aza treatment of THP-1 and HL-60 but not in KG-1a. To further confirm the importance of these miRNAs in primary AML patient samples, we evaluated the expression of miR-145 and miR-146a in $CD34^+$ cells isolated from patients with AML with a normal karyotype (NK; n = 54) and with del(5q) (n = 4) and controls (n = 11) from previously published expression array data[4] (Figure 5C). Expression of miR-145 is significantly reduced in patients with AML with a normal karyotype compared with controls ($P = .019$) but did not reach significance in the group with del(5q) ($P = .14$). Expression of miR-146a is significantly reduced in patients with AML with a NK ($P = .0006$) and further reduced in patients with del(5q) ($P = .0055$) compared with control samples (Figure 5C). That we observe reduced expression of miRNAs that are associated with deleted regions in AML, even in cell lines or patient samples that may not harbor the specific deletion, suggests that this subset of miRNAs may play a leukemia suppressor role.

Reduced expression of miR-145 and miR-146 results in AML in vivo, whereas re-expression of miR-145 or miR-146a suppresses AML cells in vitro

To determine whether miR-145 and miR-146a function as leukemic suppressors, restoration of their expression in AML cells harboring a del(5q) should therefore negatively effect survival or proliferation or both. We overexpressed miR-145 or miR-146a in HL-60 cells with the use of a retroviral vector coexpressing GFP as a marker. HL-60 cells were chosen because miR-145 and miR-146a are both expressed at low levels (Figure 4) and are also potentially suppressed by epigenetic mechanisms (Figure 5B). After infection of HL-60 cells and sorting for the transduced population, a noticeable impairment in growth and loss of transduced cells (GFP^+) was observed in HL-60 cells expressing miR-145 or miR-146a (supplemental Figure 2; data not shown). To determine whether miR-145 and miR-146a effect the survival of HL-60 cells,

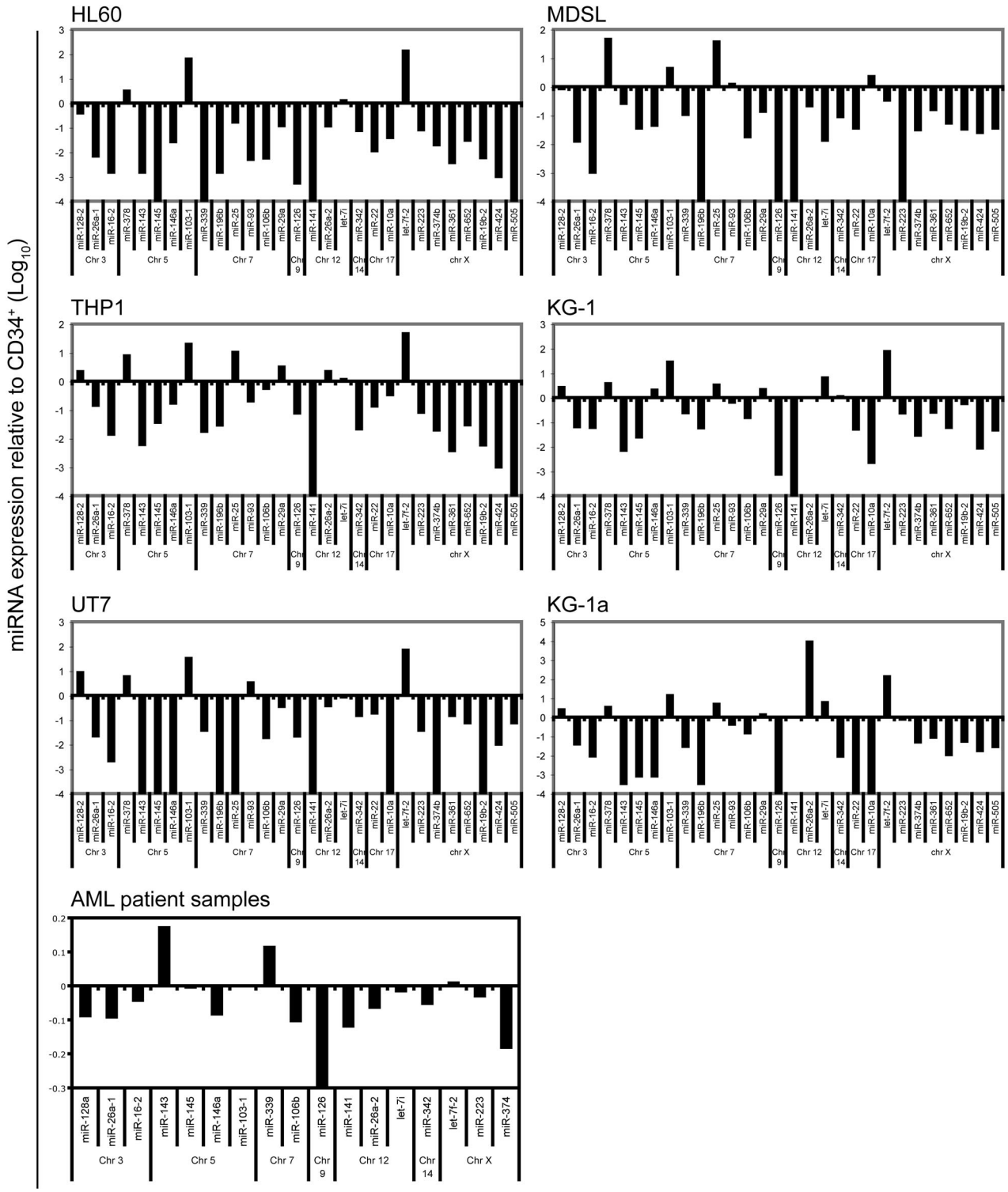

Figure 4. miRNA expression comparison to normal human CD34⁺ cells. Tag counts of miRNAs that encoded from deleted regions for each cell line are compared with normal CD34⁺ cells. Shown are the log expression levels for individual cell lines relative to CD34⁺ cells. The sequencing library of small RNAs generated from human CD34⁺ bone marrow cells was obtained from a recently published analysis.[12] miRNA expression levels from AML patient samples are shown from CD34⁺ marrow cells (n = 58) and shown relative to CD34⁺ cells evaluated from nondiseased persons (n = 11). Data are adapted from published microarray data (E-TABM-970 and E-TABM-405).

we measured annexin V staining. As shown in Figure 5C, ~ 40% and 50% of HL-60 cells transduced with miR-145 or miR-146a, respectively, stained positively for annexin V. In contrast, vector control cells or nontransduced parental cells grew robustly after infection and exhibited ~ 10% annexin V staining (Figure 5C). Re-expression of miR-145 or miR-146a inhibits survival and growth of leukemic cells, therefore implicating them in the pathogenesis of AML.

In a recent study, we showed the knockdown of miR-145 and miR-146a results in nonfatal myelodysplastic syndrome in vivo.[14]

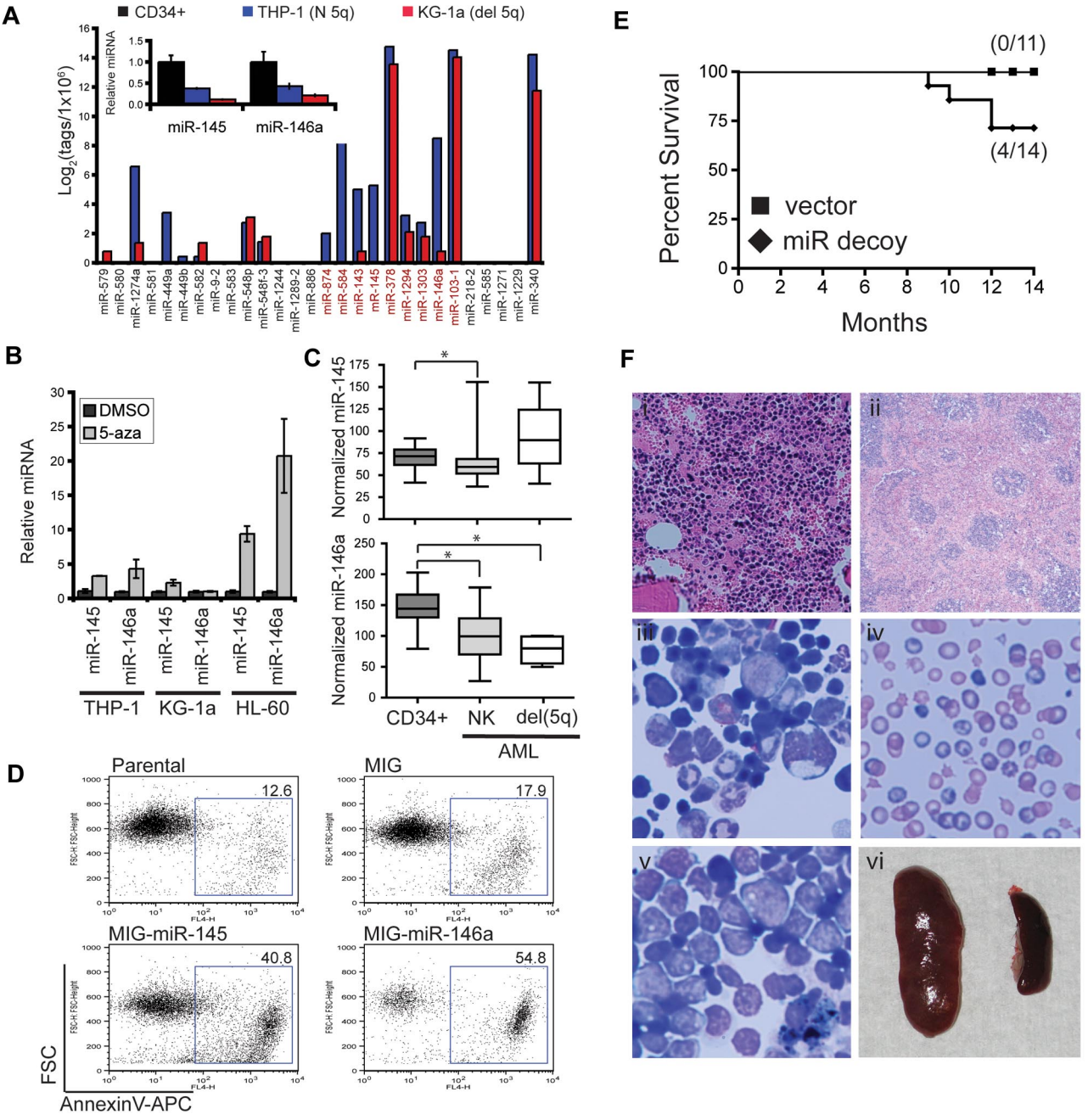

Figure 5. Reduced expression of miR-145 and miR-146a results in myelodysplastic/myeloproliferative features. (A) miRNA tag counts (log_2) are shown for 2 cell lines, one with a chromosome 5q deletion (KG-1a) and one diploid at chromosome 5q (THP-1). (Inset) Relative levels of miR-145 and miR-146a were determined by quantitative PCR in human $CD34^+$ cells (n = 3), KG-1a, and THP-1 cells. (B) KG-1a, THP-1, and HL-60 cells were treated with 10μM 5-azacytadine (5-aza) or dimethyl sulfoxide (DMSO) for 48 hours. Expression of miR-145 and miR-146a was evaluated by quantitative reverse transcription PCR and shown relative to DMSO after normalization to 5S. (C) miR-145 (top) and miR-146a (bottom) expression levels are shown for $CD34^+$ marrow cells isolated from patients with AML with a NK (n = 54) or deletion of chr 5q (del 5q; n = 4). As controls, $CD34^+$ cells were evaluated from nondiseased persons (n = 11). Data are adapted from published microarray data (E-TABM-970 and E-TABM-405). Normalized expression levels are visualized as box-and-whisker plots. (D) HL-60 cells were retrovirally transduced with empty vector (MIG), miR-145, or miR-146a. Transduced cells were isolated by fluorescence-activated cell sorting and analyzed for Annexin V binding after 1 week in culture. Shown is a representative analysis from 2 independent transductions. Gating strategy is provided in supplemental Figure 2. (E) Kaplan-Meier survival curves for mice reconstituted with marrow transduced with vector (n = 11) or miR-145/miR-146a decoy (n = 15) from 3 independent transplants. (F) Hematoxylin and eosin–stained femur (i) and spleen (ii) sections from a myeloproliferative-like diseased mouse. Wright-Giemsa–stained bone marrow cytospins (iii-v) and blood smears (iv) from a bone marrow failure mouse (iii-iv) and a leukemic mouse (v). Spleen image was obtained at time of death from a leukemic mouse (587 mg) and compared with a spleen from a control mouse (100 mg; vi).

Because MDS often precedes evolution to leukemia, we wanted to determine whether miR-145 and miR-146a contribute to an eventual AML by extending the analysis of mice for ≤ 14 months. Stably knocked down miR-145 and miR-146a in mouse hematopoietic stem/progenitor cells was achieved with the use of retroviral-mediated overexpression of miR-145 and miR-146 target sequences (miR decoy) engineered into the 3′-UTR of YFP.[14] Lethally irradiated C57Bl/6 mice received a transplant with 7.5×10^5 marrow cells transduced with the miR decoy construct or vector alone, along with 2×10^5 wild-type competitor marrow cells. We found that starting at 8 months, ~ 30% (4 of 14) of the miR decoy mice began to succumb to hematologic myeloid

Table 2. Novel miRNAs mapping to copy number alterations

miRNA name	Chromosome	Bp start	Bp end	Deletion/amplification	Mature miRNA sequence	Target genes*
miR-472	10	115 041 356	115 041 465	Del	AAATGAATCATGTTGGGCCTGT	ADIPOQ
miR-463	10	103 351 146	103 351 255	Del	AAGGGCTTCCTCTCTGCAGGAC	SYNGR1
miR-431	13	49 468 528	49 468 637	N	ACAAAAAAAAAAGCCCAACCCT	FOXK1, RELB, NF1
miR-443	16	14 902 849	14 902 958	N	AGAAGGGGTGAAATTTAAACGT	SYNGAP1
miR-474	1	159 683 046	159 683 155	Amp	AGCGGAACTTGAGGAGCCGAG	MEX3A
miR-460	4	153 629 916	153 630 025	Del	AGCTTTTGGGAATTCAGGTAG	PTPRT
miR-468	11	61 032 625	61 032 734	Amp	AGGGGGCGGGCTCCGGCG	ZNF385A, AAK1, NFIX
miR-478	2	96 827 730	96 827 839	N	ATCAGGGCTTGTGGAATGGGAAG	ADAM11
miR-432	19	6 367 420	6 367 529	Amp	CAGCCCGGATCCCAGCCCACTTA	FBXL19
miR-457	16	2 260 666	2 260 775	N	CTCGTGGGCTCTGGCCACGGC	MN1
miR-442	3	69 180 778	69 180 887	Del	CTGACTGAATAGGTAGGGTCAT	CADM2
miR-422	15	63 798 621	63 798 730	Amp	GAAGAACTGTTGCATTTGCCC	MSL2L1, ERG, MPL, TRAF7
miR-425	22	18 331 270	18 331 379	Amp	GAGGGCATGCGCACTTTGTCC	RC3H1
miR-427	3	50 687 498	50 687 607	Del	GATGCGCCGCCCACTGCCCCGC	IMPAD1
miR-437	2	207 356 179	207 356 288	N	GCTGCACCGGAGACTGGGTAA	SP8, MAF, TET3
miR-486	22	29 886 043	29 886 152	Amp	GGAGGAACCTTGGAGCTTCGGC	SPTBN1, NFIX, NF2, ADAR
miR-418	19	12 675 394	12 675 503	Amp	GTATTCGTACTGTCTGATGGG	
miR-416	2	12 256 683	12 256 792	Del	TAGTGGATGATGCACTCTGTGC	ABCA5, JAK1
miR-428	8	96 154 304	96 154 413	Amp	TGAGGAGATCGTCGAGGTTGGC	LUZP1
miR-419	16	1 724 971	1 725 080	N	TGCACGGCACTGGGGACACGT	RELT
miR-404	16	14 912 572	14 912 681	N	TGGGGCGGAGCTTCCGGAGGCC	TGRB1, NFIX
miR-408	16	84 505 724	84 505 833	Del	TGGTGTGGAAGTCTAGGCCTG	FBXO41
miR-481	7	29 686 866	29 686 975	Del	TGTCTTACTCCCTCAGGCACAT	CUGBP2, MN1
miR-450	9	136 881 134	136 881 243	N	TGTGATATCATGGTTCCTGGGA	NLGN4X, ROCK1, ITCH
miR-402	9	96 888 098	96 888 207	N	TGTTCCTGCTGAACTGAGCCAG	KLHL9, src
miR-421	19	12 912 277	12 912 386	Amp	TTGGAGGGTGTGGAAGACATC	ACT1, SPARC
miR-414	19	764 560	764 669	N	TTGGCCATGGGGCTGCGCGGGGC	KRAS

Bp indicates base pair; Del, deletion; Amp, amplification; and N, copy number neutral.
*Target Scan v5.1 was used to identify predicted target genes.

malignancies ($P = .049$; Figure 5E). Diseased mice displayed features of bone marrow failure, myeloproliferation, or acute leukemia, as supported by histologic findings and blood counts (Figure 5D; data not shown). Depending on the diseased mouse, evidence of hypercellular marrows, dyshematopoiesis, anemia, immature blastlike cells, and splenomegaly was observed (Figure 5F). To investigate previously defined targets of miR-145 and miR-146a in the mice, we measured protein expression of interleukin-1 receptor–associated kinase 1 (IRAK1; miR-146a target), tumor necrosis factor receptor–associated factor 6 (TRAF6; miR-146a target), and Mal/toll–interleukin-1 receptor domain–containing adaptor protein (TIRAP; miR-145 target). We confirmed by Western blot analysis that IRAK1, TRAF6, and TIRAP were, generally, increased in miR decoy marrow cells before transplantation, after transplantation, and in a leukemic miR-decoy mouse (supplemental Figure 3). Reduced expression of miR-145 and miR-146a, 2 leukemia-associated miRNAs identified in our analysis, results in myeloid diseases preceded by an MDS-like phenotype in vivo.

Novel miRNAs mapping to leukemia-associated genomic alterations

Unclassified sequence reads represented 10%-60% of the filtered libraries (supplemental Table 3). To identify novel candidate miRNAs among the unclassified reads, we used available algorithms.[11] Novel miRNAs from the 6 cell line libraries include 28 unique miRNA sequences (Table 2; supplemental Table 5). Five novel miRNAs were expressed at levels near the background cutoff (> 0.001), including miR-478 and miR-486 (supplemental Table 5). Because we are most interested in miRNAs that map to CNAs, we determined which of the novel miRNA sequences are encoded from leukemia-associated alterations in the 6 cell lines. Eighteen (of 28) of the novel miRNAs are encoded from genomic regions of CNAs, distributed evenly between deletions and amplifications (Figure 2D; Table 2; supplemental Table 5). For example, miR-468 and miR-484 both map to a region of chromosome 11q, a recurring region of amplification in our cell lines (Figure 1; Table 2).

Identification of miRNA variants and mature mRNA variants

Use of small RNA sequencing approaches allowed identification of miRNAs that exhibit variations in nucleotides compared with their miRBase reference sequence. In the 6 libraries, mature mRNA variants (isomiRs) were found more abundant compared with the miRBase reference sequence in some instances. For example, the most abundant sequence of miR-140-3p in the HL-60 library did not match the miRBase reference sequence. The miRBase sequence (**T**ACCACAGGGTAGAACCACGG) was expressed at low levels in this library, but instead an isomiR of miR-140-3p (ACCACAGGGTAGAACCACGG**AC**) with a 5′ nucleotide deletion and 3′ nucleotide additions (underlined) was abundantly detected (913 vs 3 reads). Examination of the libraries revealed that 18 highly expressed miRNAs are overrepresented by an isomiR rather than the reference miRNA sequence (Table 3). Modifications among the isomiRs included nucleotide 5′ and 3′ additions or deletions. The functional effect of the various miRNA modifications has not been studied in depth, but theoretically it could affect the function of miRNAs as well as detection by standard real-time PCR approaches.

Another less-frequent subset of miRNAs are ones that show nucleotide discrepancies of the mature miRNA, other than single

Table 3. Abundantly expressed and overrepresented isomiRNAs

Cell line	miRNA name	miRBase counts	miRBase sequence*	isomiR sequence*	isomiR counts	Ratio (isomR:miRbase)
HL60	miR-107	1136	AGCAGCATTGTACAGGGCTAT**CA**	AGCAGCATTGTACAGGGCTAT	3915	3.4
	miR-181a	840	AACATTCAACGCTGTCGGTGAGT	AACATTCAACGCTGTCGGTGAGT**TT**	3481	4.1
	miR-199a-3p	0	ACAGTAGTCTGCACATTGGTT**A**	ACAGTAGTCTGCACATTGGTT	2896	2896
	miR-181b	296	AACATTCATTGCTGTCGGTGGGT	AACATTCATTGCTGTCGGTGGGT**T**	2455	8.3
	miR-140-3p	3	**T**ACCACAGGGTAGAACCACGG	ACCACAGGGTAGAACCACGG**AC**	913	304.3
	miR-101	74	TACAGTACTGTGATAACTGAA	**G**TACAGTACTGTGATAACTGAA	624	8.4
	miR-222	24	AGCTACATCTGGCTACTGGGT	AGCTACATCTGGCTACTGGGT**CTC**	391	16.3
	miR-23a	79	ATCACATTGCCAGGGATTTCC	ATCACATTGCCAGGGATTTCC**A**	295	3.7
	miR-30d	7	TGTAAACATCCCCGACTGGAAG	TGTAAACATCCCCGACTGGAAG**CT**	193	27.6
THP1	miR-101	1859	TACAGTACTGTGATAACTGAA	GTACAGTACTGTGATAACTGAA	7710	4.1
	miR-140-3p	249	**T**ACCACAGGGTAGAACCACGG	ACCACAGGGTAGAACCACGGAC	6943	27.9
UT-7	miR-140-3p	15	**T**ACCACAGGGTAGAACCACGG	ACCACAGGGTAGAACCACGGAC	642	42.8
	miR-101	10	TACAGTACTGTGATAACTGAA	**G**TACAGTACTGTGATAACTGAA	72	7.2
	miR-142-5p	23	**C**ATAAAGTAGAAAGCACTACT	ATAAAGTAGAAAGCACTACT**AA**	68	3.0
MDSL	miR-183	133	**T**ATGGCACTGGTAGAATTCACT	ATGGCACTGGTAGAATTCACT**G**	367	2.8
	miR-140-3p	0	**T**ACCACAGGGTAGAACCACGG	ACCACAGGGTAGAACCACGG**AC**	148	148
KG-1	miR-101	877	TACAGTACTGTGATAACTGAA	**G**TACAGTACTGTGATAACTGAA	2001	2.3
KG-1a	miR-101	671	TACAGTACTGTGATAACTGAA	**G**TACAGTACTGTGATAACTGAA	1573	2.3

*Bold letters in sequences indicate nucleotides that differ between the miRBase and isomiR sequence.

nucleotide extensions or eliminations. These nucleotide discrepancies are inconsistent with sequencing errors, but instead probably represent posttranscriptional RNA editing of miRNA mature sequences. We detected 7 miRNAs undergoing adenine-to-guanine (A:G) substitutions with a frequency $> 1\%$ of the respective isomiR reads. A:G substitutions are the most commonly identified nucleotide substitutions in our libraries, potentially explained by spontaneous or enzymatic deamination of adenine to guanine. Of the A:G edits, miR-423-5p represented 6% of all miR-423 isomiRs (Table 4). Six other miRNAs (miR-1, miR-30d, let-7d, miR-103, miR-25*, and miR-24) contained edits in $> 1\%$ of their respective isomiR pools (Table 4). Most edits occurred between positions 17 and 20. However, editing events observed in miR-1 were in the fourth position, which includes the seed region (2-8 nucleotides) (Figure 6A). As shown in Figure 6A, a substitution of A:G (position 4) changed the repertoire of predicted miRNA targets.

To demonstrate that the editing event in the seed region of miR-1 (4A:G) alters the binding properties of the miR-1, we measured miRNA binding with the use of 3′-UTR sequence fragments containing the predicted targets of miR-1 and miR-1 (4A:G) inserted downstream of a luciferase reporter. MEIS1, an AML-associated cofactor,[21] is predicted to be a target of miR-1 but not of miR-1 (4A:G) (Figure 6B). In contrast, a tumor suppressor associated with AML,[22,23] TP53, is predicted to be a target of miR-1 (4A:G) but not of miR-1 (Figure 6B). As expected, transient cotransfection of miR-1 or miR-1 (4A:G) and MEIS1 3′-UTR into HEK293 cells yielded a 20% decrease in reporter activity by miR-1 but not miR-1 (4A:G) (Figure 6C).

Table 4. miRNA editing of adenine to guanine (A:G)

miRNA	Cell line	Position	Count	Total	Ratio*
miR-1	UT7	4	23	2306	0.0100
miR-30d	KG1a	20	11	1077	0.0102
let-7d	UT7	1	12	999	0.0120
miR-103	UT7	20	15	1160	0.0129
miR-25*	KG1a	17	13	953	0.0136
miR-24	KG1a	18	32	1774	0.0180
miR-423-5p	KG1a	19	770	11628	0.0662

A:G edits occurring in $> 1\%$ of reads.
*Ratio was calculated from the edited miRNA per all of the miRNA reads.

Transient cotransfection of miR-1 (4A:G) and TP53 3′-UTR into HEK293 cells yielded a 40% decrease ($P = .027$) in reporter activity, but, surprisingly, miR-1 expression also suppressed luciferase activity of TP53 3′-UTR reporter but to a lesser extent (20%; $P = .89$). Further evaluation of the miRNA binding sites revealed that miR-1 is predicted to bind to 2 sites in the TP53 UTR, albeit weakly (-3.7 and -2.9 kcal/mol); however, editing of A to G at position 4 increases the binding affinity of miR-1 to one these sites (-3.7 vs -10.4 kcal/mol) (Figure 6B).

Discussion

Cytogenetic abnormalities represent a large proportion of patients with de novo and therapy-related AML. Identification of AML-relevant genes within cytogenetic alterations is the foremost objective and a daunting challenge to better understand AML. Previous work has shown that miRNAs are frequently located at cancer-associated fragile sites and firmly implicated them as drivers of leukemogenesis. We performed an integrative analysis to identify relevant miRNAs located in leukemia-associated cytogenetic changes. We further corroborated our observations by examining relative expression of the leukemia-associated miRNAs in AML patient samples from published gene expression data (Table 1; Figure 4). Collectively, our finding suggests that, although many miRNAs are located in regions of leukemia-associated cytogenetic changes ($\sim 70\%$), only a subset ($\sim 20\%$) of these miRNAs are expressed and ae probably relevant myeloid malignancies.

To determine whether our subset of miRNAs have been previously implicated in leukemogenesis or hematopoiesis, we searched PubMed for relevant findings. A search showed that miR-143, miR-145, miR-146a, miR-155, miR-181, miR-221, and miR-222 are implicated in cellular processes relevant to AML and thus are consistent with our conclusion that the refined subset of leukemia-associated miRNAs are potentially important[2] (Table 1). miRNAs within the commonly deleted region on chromosome 5q in MDS have recently been evaluated by our group.[14] Previously, we have shown that deletion of miR-145 and miR-146a disrupts hematopoiesis, resulting in a phenotype similar to del (5q) MDS (eg, 5q− syndrome) in mice.[14] Depletion of miR-145 and

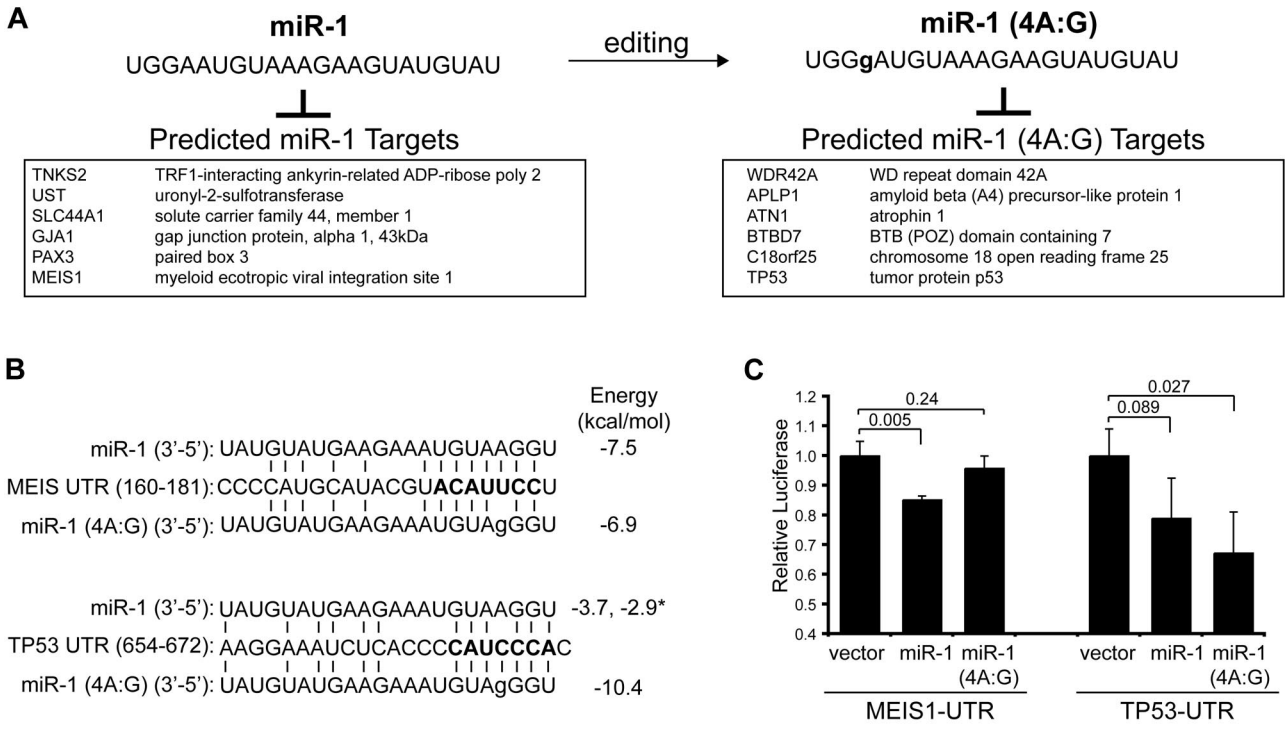

Figure 6. Editing of miR-1 occurs within the seed region and affects binding to mRNA targets. (A) Nucleotide modification at position 4 is shown for miR-1. An adenine to guanine (A:G) change was observed in UT-7, representing 23 of 2306 total miR-1 reads (Table 4). The effect of A:G edit at position 4 (4A:G) on mRNA target specificity is illustrated by comparing select miR-1 and miR-1 (4A:G) targets. The predicted targets for miR-1 (4A:G) are entirely different compared with the unedited miR-1. (B) Shown is a sequence alignment of miR-1 (wild-type and edited) and its binding sites in the 3′-UTR of MEIS1 (position 160-181) and TP53 (position 654-672). The miRNA-mRNA binding energy score (in kcal/mol) was determined for each pairing and shown on the right. The * indicates the binding energy score for a second miR-1 binding site in the TP53 UTR (position 689-698). (C) Luciferase activity was measured for MEIS1-UTR and TP53-UTR luciferase reporters in the presence of miR-1, miR-1 (4A:G), or vector control. Shown are relative values of 2 independent experiments performed in triplicate.

miR-146a also results in a clonal advantage and enhanced survival of hematopoietic stem/progenitor cells.[14] Further analysis revealed that a subset of these mice go onto develop myeloid diseases consistent with leukemia (Figure 5); therefore, deletion of miR-145 and miR-146a, 2 miRNAs identified in our analysis as potentially relevant to human leukemia, results in a long-latency myeloid disease in mice. In addition, reintroduction of miR-145 or iR-146a into AML cells significantly induced cell death (Figure 5D) and prevented growth in vitro (supplemental Figure 2). We also confirmed that miR-145 and miR-146a are down-regulated in AML patient samples and are potentially suppressed by epigenetic mechanisms (Figure 5B-C). Although miR-146a was down-regulated in patients with a NK and further suppressed in patients with del(5q), surprisingly miR-145 was significantly down-regulated only in patients with AML with a NK, but not in patients with AML with del(5q) (Figure 5C). This discrepancy may be explained by the relatively small cohort of patients with del(5q) or that low levels of miR-145 may be less critical for maintaining the AML phenotype and more important for initiation of a premalignant state conducive to development of AML. In support of the latter hypothesis, we previously reported that miR-145 is significantly down-regulated in patients with MDS with del(5q).[14]

Use of a massively parallel sequencing platform revealed a proportion of alternative miRNA species. Identification of miRNA sequence variants, miRNA*, and posttranscriptional editing adds to the complexity of the miRNA transcriptome. The relevance of alternative miRNA species is not precisely known but has also been recently described in a leukemia progression mouse model.[19] Our best understanding of the effect of nucleotide modifications comes from studies on single nucleotide polymorphisms within miRNAs. These single nucleotide polymorphisms have been identified and are associated with altered miRNA biogenesis, stem-loop formation, and strand preference.[24] The functional consequence of nucleotide deletions and additions remains to be known. Although rare occurrences in our libraries, we did identify miRNAs that exhibited editing within sequences of the miRNA:mRNA duplex. miRNA editing has also been reported in mouse embryo, *Oryza sativa*, and *Arabidopsis thaliana* libraries.[25,26] We have shown that an edit within the seed region of miR-1 (4A:G) may alter binding to mRNA targets. In our example, we specifically found that wild-type miR-1, but not the edited version, can bind and repress the 3′-UTR of MEIS1 (Figure 6B-C). This observation is particularly relevant to AML because MEIS1 is a key cofactor of the Hox cluster and critical to leukemogenesis in mice and humans.[21] Not only did we show that the edited version of miR-1 does not bind MEIS1, but it now binds and suppresses the 3′-UTR of TP53, a gene that is mutated in human AML and its loss induces AML in mice.[22,23] Therefore, it is compelling to speculate that editing of miR-1 (4A:G) makes miR-1 proleukemic by de-repressing MEIS1 and simultaneously suppressing TP53.

Massively parallel sequencing also facilitated identification of novel miRNAs. Because this is the first reported attempt to perform massively parallel sequencing on human leukemia cells, we identified 28 novel miRNAs. Approximately 65% of the novel miRNAs are located in leukemia-associated genomic alterations. Notably, miR-481, which is within a deleted region on chromosome 7q, is predicted to target meningioma 1 (Mn1). Elevated expression of Mn1 is a predictor of poor outcome in patients with AML, and expression of a Mn1 transgene in mouse

hematopoietic cells results in aggressive and rapid leukemia.[27,28] As such, deletion of miR-481 provides a compelling mechanism to increase expression of Mn1 in AML. Collectively, we have delineated the genomic alterations and identified leukemia-associated miRNAs.

Acknowledgments

We thank Megan Fuller for animal husbandry.

This work was supported by grants from the Canadian Institutes of Health Research (CIHR; MOP 89 976) and the Leukemia & Lymphoma Society of Canada (A.K.), grants from Genome Canada and CIHR (W.L.L.), and fellowships from CIHR and the Michael Smith Foundation for Health Research (MSFHR; D.T.S.). A.K. is a Senior Scholar of the MSFHR.

Authorship

Contribution: D.T.S. and A.K. participated in designing the research and drafting the manuscript; D.T.S., J.L., and J.W. performed experiments; F.K. and R.K.H. provided technical assistance and assisted in the analysis of data; W.L.L. and R.C. provided valuable reagents and expertise in array CGH; R.M., A.M., M.H., and M.M. performed analysis and provided expertise in small RNA sequencing on the Illumina platform; and K.T. provided the MDS-L cell line.

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

Correspondence: Aly Karsan, British Columbia Cancer Agency Research Centre, 675 West 10th Ave, Vancouver, BC V5Z 1L3, Canada; e-mail: akarsan@bccrc.ca.

References

1. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell.* 2004;116(2):281-297.
2. Baltimore D, Boldin MP, O'Connell RM, Rao DS, Taganov KD. MicroRNAs: new regulators of immune cell development and function. *Nat Immunol.* 2008;9(8):839-845.
3. O'Connell RM, Rao DS, Chaudhuri AA, et al. Sustained expression of microRNA-155 in hematopoietic stem cells causes a myeloproliferative disorder. *J Exp Med.* 2008;205(3):585-594.
4. Garzon R, Volinia S, Liu CG, et al. MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. *Blood.* 2008;111(6):3183-3189.
5. Jongen-Lavrencic M, Sun SM, Dijkstra MK, Valk PJ, Lowenberg B. MicroRNA expression profiling in relation to the genetic heterogeneity of acute myeloid leukemia. *Blood.* 2008;111(10):5078-5085.
6. Marcucci G, Radmacher MD, Maharry K, et al. MicroRNA expression in cytogenetically normal acute myeloid leukemia. *N Engl J Med.* 2008;358(18):1919-1928.
7. Calin GA, Sevignani C, Dumitru CD, et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A.* 2004;101(9):2999-3004.
8. Ishkanian AS, Malloff CA, Watson SK, et al. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet.* 2004;36(3):299-303.
9. Starczynowski DT, Vercauteren S, Telenius A, et al. High-resolution whole genome tiling path array CGH analysis of CD34+ cells from patients with low-risk myelodysplastic syndromes reveals cryptic copy number alterations and predicts overall and leukemia-free survival. *Blood.* 2008;112(8):3412-3424.
10. Chi B, DeLeeuw RJ, Coe BP, MacAulay C, Lam WL. SeeGH–a software tool for visualization of whole genome array comparative genomic hybridization data. *BMC Bioinformatics.* 2004;5:13.
11. Morin RD, O'Connor MD, Griffith M, et al. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* 2008;18(4):610-621.
12. Landgraf P, Rusu M, Sheridan R, et al. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell.* 2007;129(7):1401-1414.
13. Busch A, Richter AS, Backofen R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics.* 2008;24(24):2849-2856.
14. Starczynowski DT, Kuchenbauer F, Argiropoulos B, et al. Identification of miR-145 and miR-146a as mediators of the 5q- syndrome phenotype. *Nat Med.* 2010;16(1):49-58.
15. Simon R, Lam A, Li MC, Ngan M, Menenzes S, Zhao Y. Analysis of gene expression data using BRB-Array Tools. *Cancer Inform.* 2007;3:11-17.
16. Drexler H. *Guide to leukemia-lymphoma cell lines:* Braunschweig; 2005.
17. Alvarez S, Cigudosa JC. Gains, losses and complex karyotypes in myeloid disorders: a light at the end of the tunnel. *Hematol Oncol.* 2005;23(1):18-25.
18. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. *Nucleic Acids Res.* 2008;36(Database issue):D154-158.
19. Kuchenbauer F, Morin RD, Argiropoulos B, et al. In-depth characterization of the microRNA transcriptome in a leukemia progression model. *Genome Res.* 2008;18(11):1787-1797.
20. Giagounidis AA, Germing U, Aul C. Biological and prognostic significance of chromosome 5q deletions in myeloid malignancies. *Clin Cancer Res.* 2006;12(1):5-10.
21. Pineault N, Buske C, Feuring-Buske M, et al. Induction of acute myeloid leukemia in mice by the human leukemia-specific fusion gene NUP98-HOXD13 in concert with Meis1. *Blood.* 2003;101(11):4529-4538.
22. Zhao Z, Zuber J, Diaz-Flores E, et al. p53 loss promotes acute myeloid leukemia by enabling aberrant self-renewal. *Genes Dev.* 2010;24(13):1389-1402.
23. Wattel E, Preudhomme C, Hecquet B, et al. p53 mutations are associated with resistance to chemotherapy and short survival in hematologic malignancies. *Blood.* 1994;84(9):3148-3157.
24. Jazdzewski K, Murray EL, Franssila K, Jarzab B, Schoenberg DR, de la Chapelle A. Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. *Proc Natl Acad Sci U S A.* 2008;105(20):7269-7274.
25. Reid JG, Nagaraja AK, Lynn FC, et al. Mouse let-7 miRNA populations exhibit RNA editing that is constrained in the 5′-seed/ cleavage/anchor regions and stabilize predicted mmu-let-7a:mRNA duplexes. *Genome Res.* 2008;18(10):1571-1581.
26. Ebhardt HA, Tsang HH, Dai DC, Liu Y, Bostan B, Fahlman RP. Meta-analysis of small RNA-sequencing errors reveals ubiquitous post-transcriptional RNA modifications. *Nucleic Acids Res.* 2009;37(8):2461-2470.
27. Heuser M, Argiropoulos B, Kuchenbauer F, et al. MN1 overexpression induces acute myeloid leukemia in mice and predicts ATRA resistance in patients with AML. *Blood.* 2007;110(5):1639-1647.
28. Heuser M, Beutel G, Krauter J, et al. High meningioma 1 (MN1) expression as a predictor for poor outcome in acute myeloid leukemia with normal cytogenetics. *Blood.* 2006;108(12):3898-3905.