Consistent Deregulation of Gene Expression between Human and Murine $MLL$ Rearrangement Leukemias

Zejuan Li,$^1$ Roger T. Luo,$^1$ Shuangli Mi,$^1$ Miao Sun,$^1$ Ping Chen,$^1$ Jingyue Bao,$^2$ Mary Beth Neilly,$^1$
Nimanthi Jayathilaka,$^1$ Deborah S. Johnson,$^1$ Lili Wang,$^1$ Catherine Lavau,$^1$ Yanning Zhang,$^1$
Charles Tseng,$^1$ Xiuqing Zhang,$^2$ Jian Wang,$^2$ Jun Yu,$^1$ Huanming Yang,$^2$ San Ming Wang,$^2$
Janet D. Rowley,$^1$ Jianjun Chen,$^1$ and Michael J. Thirman$^1$

$^1$Department of Medicine, University of Chicago, Chicago, Illinois; $^2$Beijing Genomics Institute, Chinese Academy of Sciences, Beijing, China; $^3$Department of Biological Sciences, Purdue University Calumet, Hammond, Indiana; $^4$CNRS UPR9051, IUH, Paris, France; and $^5$Center for Functional Genomics, Division of Medical Genetics, Department of Medicine, ENH Research Institute, Northwestern University, Evanston, Illinois

Abstract

Important biological and pathologic properties are often conserved across species. Although several mouse leukemia models have been well established, the genes deregulated in both human and murine leukemia cells have not been studied systematically. We performed a serial analysis of gene expression in both human and murine $MLL$-ELL or $MLL$-ENL leukemia cells and identified 88 genes that seemed to be significantly deregulated in both types of leukemia cells, including 57 genes not reported previously as being deregulated in $MLL$-associated leukemias. These changes were validated by quantitative PCR. The most up-regulated genes include several $HOX$ genes (e.g., $HOX$A5, $HOX$A9, and $HOX$A10) and $MEIS1$, which are the typical hallmark of $MLL$ rearrangement leukemia. The most down-regulated genes include $LTF$, $LCN2$, $MMP9$, $S100A8$, $S100A9$, $PADI4$, $TGFBI$, and $CYBB$. Notably, the up-regulated genes are enriched in gene ontology terms, such as gene expression and transduction, whereas the down-regulated genes are enriched in signal transduction and apoptosis. We showed that the CpG islands of the down-regulated genes are hypermethylated. We also showed that seven individual microRNAs (miRNA) from the mir-17-92 cluster, which are overexpressed in human $MLL$ rearrangement leukemias, are also consistently overexpressed in mouse $MLL$ rearrangement leukemia cells. Nineteen possible targets of these miRNAs were identified, and two of them (i.e., $APP$ and $RASSF2$) were confirmed further by luciferase reporter and mutagenesis assays. The identification and validation of consistent changes of gene expression in human and murine $MLL$ rearrangement leukemias provide important insights into the genetic base for $MLL$-associated leukemogenesis.

[Cancer Res 2009;69(3):1109–16]

Introduction

Chromosomal translocations are among the most common genetic abnormalities in human leukemia. The $MLL$ (mixed lineage leukemia) gene was identified as a common target of chromosomal translocations associated with human acute leukemias (1, 2). $MLL$ is located on human chromosome 11 band q23 and on mouse chromosome 9. More than 50 different loci are rearranged in11q23 leukemias involving $MLL$ in either acute myelogenous leukemia (AML) or acute lymphoblastic leukemia (ALL; ref. 3). $MLL$ rearrangements are associated with a poor prognosis (4). $MLL$-ELL and $MLL$-ENL that result from t(11;19)(q23;p13.1) and t(11;19)(q23;p13.3), respectively (1, 5), are two common examples of these rearrangements. These two fusions are frequently involved in human AML, whereas $MLL$-ENL is also involved in human ALL. The translocations result in an in-frame fusion of the NH$_2$ terminus of the $MLL$ gene and the COOH terminus of each partner gene. Retroviral-mediated gene transfer of $MLL$-ENL and $MLL$-ELL transforms primary myeloid progenitor cells and cause acute myeloid leukemia in mice (6, 7).

Gene expression profiles differ between distinct subtypes of leukemia and provide specific markers for clinical diagnosis. It is commonly observed that important biological/pathologic properties are often conserved across species (8, 9). Model organisms have contributed substantially to our understanding of the etiology of human disease and the development of new treatment methodologies (10). However, although genetically engineered mouse leukemia models have been well established (6, 7, 11, 12), there are few systematic studies to identify and study the genes that exhibit similar abnormal expression patterns in both human and murine leukemia cells. To perform an interspecies gene expression comparative study in leukemia, we used the serial analysis of gene expression (SAGE) technique (13) to compare gene expression between $MLL$-ELL or $MLL$-ENL myeloid leukemia progenitor cells and normal myeloid progenitor cells in both humans and mice. Herein, we report the identification and validation of differentially expressed genes commonly present in both human and murine $MLL$-ELL or $MLL$-ENL leukemias.

Materials and Methods

Patient samples. The patient samples were obtained at the time of diagnosis and with informed consent at University of Chicago and were stored in liquid nitrogen until use.

SAGE assay and data analysis. Cell purification, total RNA isolation, cDNA synthesis, and SAGE were carried out according to our established procedures (13–15). SAGE tag sequences were extracted with SAGE 2000 software. Tag counts were converted to counts per 100,000, and the expression data were cross-linked to UniGene clusters by extracting the 3’-most NlaIII SAGE tag for each transcript in each UniGene clusters. Only tags that matched to a single gene cluster were taken into account. All SAGE tags mapped to the same gene were then combined, and the sum of
their counts in a given sample represented the expression level/signal of that gene in that sample (16, 17). χ² test was used to identify differentially expressed genes, as Man and colleagues (18) showed that χ² test has the best power and robustness in SAGE data analysis.

**Functional classification and annotation of the candidate genes.** We used the Database for Annotation, Visualization, and Integrated Discovery (DAVID; version 2008) for functional classification and annotation of candidate genes. DAVID provides a comprehensive set of functional annotation tools for investigators to understand the biological meaning behind large list of genes (19). A gene cluster means a set of genes belonging to the same functional category, as defined by DAVID.

**Quantitative real-time PCR assays of protein coding or microRNA genes and data analysis.** We performed quantitative real-time PCR (qPCR) to determine expression of protein coding genes with SYBR green dye using kits from Qiagen and microRNAs (miRNA) using TaqMan qPCR kits from Applied Biosystems. PKGI or glyceraldehyde-3-phosphate dehydrogenase was used as endogenous controls for protein coding genes, whereas U6 RNA was used as an endogenous control for miRNAs. PCR reactions and data analyses were performed, as described previously (20, 21).

**Methylation-specific PCR.** The methylation status of the promoter region was determined by the methylation-specific PCR (MSP) method (22). The primers were designed to anneal specifically to methylated and unmethylated CpG dinucleotides in promoter regions of genes using the Primer3 program. Genomic DNA was isolated using QIAamp DNA mini kit (Qiagen). One microgram of DNA was used for bisulfite modification using the CpGenome DNA modification kit (Chemicon) according to the manufacturer's instructions. The bisulfite-converted DNA was amplified in a total volume of 20 μL using GeneAmp Gold buffer containing 4 mmol/L MgCl₂, 0.5 μmol/L of each primer, 0.2 mmol/L deoxynucleotide triphosphate, 5 μg bovine serum albumin, and 1.25 units of AmpliTaq Gold DNA polymerase (Roche). Hot start PCR was performed for 30 cycles, which consists of denaturation at 95°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 45 s, followed by a final 7-min extension for all primer sets. The products were separated on 10% polyacrylamide gels.

**Cell culture and 5-aza-2′-deoxycytidine (Decitabine) treatment.** Human leukemia cell lines were grown at 37°C under 5% CO₂ in RPMI 1640 (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen), expressed genes, as Man and colleagues (18) showed that best power and robustness in SAGE data analysis. Applied Biosystems.

todetermineexpressionofproteincodinggeneswithSYBRgreendyeusing
genes and data analysis.

to the same functional category, as defined by DAVID.

**Functional classification and annotation tools for investigators to understand the biological meaning** behind large list of genes (19). A gene cluster means a set of genes belonging to the same functional category, as defined by DAVID.

**Quantitative real-time PCR assays of protein coding or microRNA genes and data analysis.** We performed quantitative real-time PCR (qPCR) to determine expression of protein coding genes with SYBR green dye using kits from Qiagen and microRNAs (miRNA) using TaqMan qPCR kits from Applied Biosystems. PKGI or glyceraldehyde-3-phosphate dehydrogenase was used as endogenous controls for protein coding genes, whereas U6 RNA was used as an endogenous control for miRNAs. PCR reactions and data analyses were performed, as described previously (20, 21).

**Methylation-specific PCR.** The methylation status of the promoter region was determined by the methylation-specific PCR (MSP) method (22). The primers were designed to anneal specifically to methylated and unmethylated CpG dinucleotides in promoter regions of genes using the Primer3 program. Genomic DNA was isolated using QIAamp DNA mini kit (Qiagen). One microgram of DNA was used for bisulfite modification using the CpGenome DNA modification kit (Chemicon) according to the manufacturer's instructions. The bisulfite-converted DNA was amplified in a total volume of 20 μL using GeneAmp Gold buffer containing 4 mmol/L MgCl₂, 0.5 μmol/L of each primer, 0.2 mmol/L deoxynucleotide triphosphate, 5 μg bovine serum albumin, and 1.25 units of AmpliTaq Gold DNA polymerase (Roche). Hot start PCR was performed for 30 cycles, which consists of denaturation at 95°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 45 s, followed by a final 7-min extension for all primer sets. The products were separated on 10% polyacrylamide gels.

**Cell culture and 5-aza-2′-deoxycytidine (Decitabine) treatment.** Human leukemia cell lines were grown at 37°C under 5% CO₂ in RPMI 1640 (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen), 1% penicillin/streptomycin, and 1% HEPES. The cells were plated at 2 × 10⁶ per flask in 4 mL of medium. After 24 h, cells were exposed to 5-aza-2′-deoxycytidine (5-Aza-CdR; Fluka) at 1 μmol/L for 96 h. In parallel, untreated cells were used as a control. After 48 h of continuous exposure to 5-Aza-CdR, the medium was changed. Cells were harvested after an additional 48 h of incubation. Then, total RNA and genomic DNA were isolated from the treated and untreated cell lines for further qPCR and MSP assay, respectively.

**Luciferase reporter and mutagenesis assay.** MiR-17 expression during 5-aza-2′-deoxycytidine (5-Aza-CdR; Fluka) from 1 μmol/L to 96 h. In parallel, untreated cells were used as a control. After 48 h of continuous exposure to 5-Aza-CdR, the medium was changed. Cells were harvested after an additional 48 h of incubation. Then, total RNA and genomic DNA were isolated from the treated and untreated cell lines for further qPCR and MSP assay, respectively.

**Results**

Eighty-eight genes were identified by SAGE analysis to be significantly abnormally expressed in both human and murine MLL-ELL and/or MLL-ENL leukemia. We used a modified SAGE technique (13) to compare gene expression profiles between MLL-ELL or MLL-ENL myeloid leukemia progenitor cells and normal myeloid progenitor cells (CD15+ or Gr-1+) in both humans and mice. We analyzed four patient samples (two with each fusion) and two retrovirologically induced mouse leukemias containing either MLL-ELL or MLL-ENL fusions and a leukemic mouse cell line with an MLL-ELL fusion (Table 1). SAGE tags (484,303) were identified from the nine samples (total SAGE tag counts of each sample varies from 40,000 to 100,000, with 53,811 tags per sample on average), yielding 103,899 unique tags in human and 60,993 in mouse samples (see Table 1).

We identified 88 candidate genes that seemed to be deregulated in both human and murine MLL-ELL and/or MLL-ENL leukemias. Of them, 6 and 26 genes are up-regulated and down-regulated, respectively, in both types of leukemia; 3 and 13 genes are up-regulated and down-regulated, respectively, in MLL-ELL leukemias alone; and 23 and 17 genes are up-regulated and down-regulated, respectively, in MLL-ENL leukemias alone (see Fig. 1A and Supplementary Table S1). All the significant genes have at least 3-fold difference in expression (tag counts per 100,000) between each leukemia sample and the normal control and with a χ² test P value of <0.05. As this criterion is stringent, it therefore limits the total number of the candidate genes identified.

Notably, the four most up-regulated genes (each has at least a 30-fold change in expression) in both types of human and mouse leukemia samples are three homeobox (HOX) genes (i.e., HOX5, HOX9, and HOXA10) and one HOX cofactor, MEIS1 (see Fig. 1A). The eight most down-regulated genes (each has at least a 30-fold change in expression on average) in both types of human and mouse leukemia samples include LTF, LCN2, MMP9, S100A8, S100A9, PADI4, TGFBI, and CYBB (see Supplementary Table S1). Note that the fold change of some up-regulated or down-regulated genes may be overestimated because of the limited number of SAGE tags examined (40,000–100,000 tags per sample).

Although 56 genes were identified to be significantly deregulated in only one type of leukemias, many of them exhibited a similar deregulation pattern in both types of leukemias (see Supplementary Table S1), albeit they were significant only in one according to our stringent criterion. If more samples were analyzed and/or a less stringent criterion was used, many of these genes would seem to be deregulated in both MLL-ELL and MLL-ENL leukemias.

Thirty-one of the 88 genes identified by SAGE have been previously reported to be deregulated in MLL rearrangement leukemia. Much effort has been expended previously to identify candidate genes that might be involved in MLL rearrangement leukemia using different methods, such as microarray and qPCR. To obtain a comprehensive list of candidate genes involved in MLL rearrangement leukemia, we searched 35 relevant publications (see Supplementary Table S2A) and collected a total of 1,647 genes that were reported to be significantly deregulated in MLL rearrangement leukemia. The top two most frequently reported genes are MEIS1 and HOX9, having been reported as up-regulated genes by 16 and 15 different publications, respectively. Among the 1,647 genes, 370 were reported by at least two different studies. Of them, only 10 genes (all were reported by only two publications) have contradictory information regarding whether they are up-regulated or down-regulated in MLL rearrangement leukemia. The remaining 360 genes were reported consistently by various different studies, with 213 up-regulated and 147 down-regulated in MLL rearrangement leukemia compared with other subtypes of acute leukemias and/or normal controls. We found that 31 of the 88 genes (35%) identified from our SAGE analysis were reported by previous studies (see Supplementary Table S2B).

**Validation of 81 genes by a large-scale qPCR assay.** To validate the deregulation patterns of the genes identified from our SAGE analysis and those from the previous studies, we performed
a large-scale qPCR validation for 81 genes in 20 samples. The 81 genes (see Supplementary Table S3) include 43 identified from our SAGE analysis and 38 reported only by others. Among the 43 genes identified by SAGE, 19 were also reported by others. The 20 samples include 12 human and 8 mouse samples. The human samples include three MLL-ELL and three MLL-ENL leukemia samples, in addition to six human normal control samples. The mouse samples include two retrovirally induced mouse leukemia samples containing either MLL-ELL or MLL-ENL fusion, two leukemia cell lines with MLL-ELL fusions, and one leukemia cell line with an MLL-ENL fusion, in addition to three murine normal control samples.

We found that 60 of the 81 tested genes (74%) seemed to be significantly deregulated in leukemia samples (see Fig. 1B). Of these, 50 candidate genes were deregulated in human and mouse MLL-ELL leukemia (29 up-regulated and 21 down-regulated) and 48 were deregulated in human and mouse MLL-ENL leukemia (32 up-regulated and 16 down-regulated), whereas 55 genes seemed to be significantly deregulated (34 up-regulated and 21 down-regulated) when combining the MLL-ELL and MLL-ENL leukemia samples together. As shown in Supplementary Table S3, 36 of the 43 SAGE candidate genes (84%) were confirmed to be significantly deregulated in leukemia samples in a similar mode (i.e., up-regulated or down-regulated) to SAGE data, whereas only 31 of the 57 genes (54%) reported previously were confirmed here in a similar mode to previous report(s). Notably, MYB and SMARCA4 were detected by SAGE as being up-regulated but were reported by others as being down-regulated in MLL rearrangement leukemia; the opposite is true for PTPRC, SI00A8, ANXA1, and NCF1. Our qPCR result of these six genes is consistent with our SAGE analysis rather than previous reports (see Supplementary Table S3). In addition, four genes reported only by others, including PROM1, ADCY9, LST1, and TCF4, were detected by qPCR as being deregulated in MLL rearrangement leukemia but in a different mode from previous reports.

**Functional classification and annotation of the candidate genes identified by SAGE.** To gain further insight into the biological importance of these 88 candidate genes, we analyzed the functional categories of known genes in terms of gene ontology (GO) using the DAVID tools (19). One gene (i.e., MAGT1; down-regulated in MLL-ENL leukemia) is not in the DAVID gene database; therefore, only 87 genes were analyzed. We first performed Gene Functional Classification for the 87 genes and found that there were four gene clusters (only the clusters having at least three members were considered). Gene cluster 1 contains S100A4, S100A9, and SI00A11 (enrichment score, 1.98); all of them were down-regulated in leukemia. Gene cluster 2 contains CTSS, CTSK, and MAN2B1 (enrichment score, 1.18); the first two were down-regulated, whereas the last one was up-regulated in leukemia. Gene cluster 3 contains ARHGEF1, CENTB2, and SMAPIL (enrichment score, 0.86); all of them were down-regulated in leukemia. Remarkably, gene cluster 4 contains eight genes, including HOXA5, HOXA9, HOXA10, MEIS1, MYB, SOX4, KLF4, and MLL5 (enrichment score, 0.75); the first six were up-regulated whereas the last two were down-regulated in leukemias.

We also conducted a gene enrichment and functional annotation analysis on the 87 genes with the functional annotation tools. As shown in Supplementary Table S4, among the 288 GO terms analyzed, our SAGE-detected candidate genes are specially enriched in 40 GO terms compared with the whole set of genes (30,000) in the human genome, with at least 2-fold higher enrichment (P < 0.05) in each GO term, such as cell differentiation, cell development, defense response, apoptosis, and signal transduction. Remarkably, we found that our candidate genes enriched in some GO terms are either preferentially up-regulated (e.g., those in gene expression) or down-regulated (e.g., those in apoptosis or signal transduction) in our leukemia samples.

Therefore, we further compared the percentage of up-regulated genes with that of down-regulated genes enriched in each GO term. We found that, in 125 GO terms, the percentage of enriched up-regulated gene in the whole 32 up-regulated genes is at least 2-fold higher or lower than that of enriched down-regulated genes in the whole 55 down-regulated genes (see Supplementary Table S5). As exemplified in Fig. 2, up-regulated genes have a much higher...
percentage of enrichment in GO terms, such as gene expression, transcription, and regulation of transcription compared with down-regulated genes; opposite is true in GO terms, such as apoptosis, cell death, extracellular space, immune response, response to stress, signal transduction, and cell communication (see Supplementary Table S6 for the detailed gene lists). For example, *HOXA5*, *HOXA9*, *HOXA10*, *MEIS1*, *MYB*, *SMARCA2*, *SMARCA4*, *RPUSD1*, *CDK6*, *RPL7A*, *RRAGC*, and *SOX4* are

**Figure 1.** Expression profiles of candidate genes. A, expression profiles of the 88 significant genes from nine human and mouse samples, as detected by SAGE. The genes have at least 3-fold difference in expression value (tag number per 100,000 of total SAGE tag in each library after normalization) between each leukemia sample and the normal control sample (human CD15⁺; mouse:Gr-¹⁺) and with a χ² test P value of <0.05. See Supplementary Table S1 for details. B, expression profiles of the 60 significant genes from 20 human and mouse samples, as detected by qPCR. Because the number of mouse samples is too small, we combined human and mouse quantitative real-time PCR data together for significance analysis with SAM. Data are presented as ΔCt. Unsupervised, average linkage hierarchical clustering was performed with Pearson correlation as distance. All the significant genes have a q value of <0.05, with the overall FDR of <5%. Expression data were mean centered. _ELL, MLL-ELL_; _ENL, MLL-ENL_; _cl, cell line_; _N, normal control_; _m or Mm, mouse_; _h or Hs, human_; _hm, human and mouse_. In the columns of qPCR, “−” means no significant change.
Three candidate genes in small sciences in recent years is the discovery of an abundant class of detected by SAGE.

MLL rearrangement leukemia cells is believed that the central role of miRNAs is to regulate the expression of their target genes (23–25); thus, it is important to identify their targets. Because miRNAs usually function as negative regulators, it is expected that they will exhibit an inverse correlation with the expression of their targets. Recent findings indicate that animal miRNAs can not only repress protein synthesis but also induce mRNA degradation of a large portion of targets (30, 31). To identify targets of the miRNAs in the mir-17-92 cluster, we searched for candidate genes that were significantly down-regulated in MLL rearrangement leukemias as being reported by at least two different studies (including all the previous 35 studies and our SAGE analysis) and/or having been confirmed by our qPCR assay. We obtained a collection of 167 such candidate genes and found that 19 of them were predicted as putative targets of the

A miRNA cluster is abnormally overexpressed in both human and murine leukemia samples with MLL-ELL or MLL-ENL fusion. We recently performed a large-scale, genome-wide miRNA expression profiling assay and observed that all the seven individual miRNAs from a unique polycistronic miRNA cluster, namely mir-17-92, residing in the C13orf25 gene locus at 13q31, including mir-17-5p, mir-17-3p, mir-18a, mir-19a, mir-20a, mir-19b, and mir-92, were significantly up-regulated in MLL rearrangement leukemias relative to normal controls or other leukemias (20). The mir-17-92 is a well-known oncogene (29). To validate whether this miRNA cluster is also abnormally overexpressed in murine leukemia cells with MLL-ELL or MLL-ENL fusion, we performed a qPCR to evaluate the expression levels of all seven individual miRNAs in 33 human and 7 mouse samples. The 33 human samples include 18 MLL rearrangement and 9 non-MLL rearrangement leukemia samples and 6 normal controls. The seven mouse samples include two MLL-ELL and two MLL-ENL leukemia samples and three Gr-1+ normal control samples. As shown in Fig. 4A, as in human samples, all seven individual miRNAs are overexpressed in mouse leukemia cells, although the level is not as high as in corresponding human leukemias.

Candidate targets of the miRNAs in the mir-17-92 cluster. It is believed that the central role of miRNAs is to regulate the expression of their target genes (23–25); thus, it is important to identify their targets. Because miRNAs usually function as negative regulators, it is expected that they will exhibit an inverse correlation with the expression of their targets. Recent findings indicate that animal miRNAs can not only repress protein synthesis but also induce mRNA degradation of a large portion of targets (30, 31). To identify targets of the miRNAs in the mir-17-92 cluster, we searched for candidate genes that were significantly down-regulated in MLL rearrangement leukemias as being reported by at least two different studies (including all the previous 35 studies and our SAGE analysis) and/or having been confirmed by our qPCR assay. We obtained a collection of 167 such candidate genes and found that 19 of them were predicted as putative targets of the

Figure 2. Example of 23 GO terms in which enriched up-regulated genes have an over 2-fold difference of enrichment percentage compared with enriched down-regulated genes. The percentages of the up-regulated and down-regulated genes relative to their total number (i.e., 32 and 55, respectively) in each GO term are shown in the plot for a direct comparison.
miRNAs in the mir-17-92 cluster by at least two of the four prediction programs described above (see Supplementary Table S8). Of these, three genes (i.e., APP, RASSF2, and SH3BRS2) have already been confirmed by our qPCR assay to be significantly down-regulated in both MLL-ELL and MLL-ENL samples (see Fig. 1B).

We further performed a luciferase reporter and mutagenesis assay to validate APP and RASSF2. A significantly negative effect ($P < 0.01$; paired $t$ test) on luciferase activity was observed in the presence of miR-17 on 3′ UTR of APP or RASSF2, and such repression was lost when the predicted miRNA binding site in the 3′ UTR of APP or RASSF2 (in the reporter plasmid) was mutated (Fig. 4B), indicating that both APP and RASSF2 are direct targets of miR-17.

**Discussion**

The power of comparative genomic analysis relies on the assumption that important biological properties are often conserved across species (9). Cross-species sequence comparison has been widely used to infer gene function, but it is becoming apparent that sequence similarity is not always proportional to functional similarity (32, 33). To determine the function of a gene precisely, therefore, we need to investigate not only its sequence characteristics but also its expression characteristics (9). The expression pattern of a gene can thus serve as a sensitive indicator of its function. In the present work, we conducted a genome-wide gene expression profiling assay on both human and mouse leukemia cells bearing MLL fusions using the SAGE method. We obtained 484,303 total SAGE tags for the nine samples and a total of 103,899 unique SAGE tags in the human and 60,993 in the mouse samples. We identified 88 genes that seemed to be significantly deregulated (32 up-regulated and 56 down-regulated) in both human and murine MLL-ELL and/or MLL-ENL leukemia (see Fig. 1A).

In a comparison with previous studies of human MLL rearrangement leukemias that used different methods, such as cDNA microarray, oligonucleotide microarray, and qPCR, we found that 31 of the 88 genes identified by SAGE have been reported previously but 57 are previously undescribed candidate genes that seem to be deregulated in MLL rearrangement leukemias. We further performed a large-scale qPCR assay to validate the deregulation of 81 candidate genes, including 43 genes identified from our SAGE analysis and 38 genes reported only by the previous studies in 20 samples, including 12 human and 8 mouse samples. Eighty-four percent (36 of 43) of the candidate genes detected by SAGE were confirmed, as were 54% (31 of 57) of the genes identified by previous studies. A lower confirmation/validation rate of the genes identified by previous studies relative to genes identified by our SAGE analysis is not surprising, because most of the previous studies compared MLL rearrangement leukemias with other acute leukemias, not with normal controls, and our qPCR focused on MLL-ELL and MLL-ENL samples that might not be included in a large number of previous studies. The fact that only 31 of the 88 genes detected by SAGE were reported previously and 84% of the tested SAGE candidate genes were confirmed by qPCR indicates that, although MLL rearrangement leukemias have been intensively studied previously, there are still many potentially important genes that have not been identified; thus, our study, using a different strategy/method, could provide additional important information to extend our understanding of the complex genetic alterations in MLL fusion-induced leukemogenesis.

Remarkably, HOXA5, HOXA9, HOXA10, and the HOX cofactor MEIS1 are the four most up-regulated genes observed in both types (i.e., MLL-ELL and MLL-ENL) of human and murine leukemia samples (Fig. 1A), which were further confirmed by our large-scale qPCR assay (Fig. 1B). HOX genes and MEIS1 are the best studied downstream targets of MLL and MLL fusion proteins, and their aberrant overexpression has been consistently recognized as the hallmark of MLL rearrangement leukemias (34–41). Interestingly, in a gene functional classification analysis using the DAVID tools (19), we found that HOXA5, HOXA9, HOXA10, and MEIS1 were clustered with four other candidate genes, including MYB, SOX4, KLF4, and MLL5; the first six were up-regulated whereas the last two were down-regulated in MLL-associated leukemias. Notably, all of these eight genes are regulators of gene expression and/or cell cycle; the
first six genes (41–44) are positive regulators, whereas KLF4 (45) and MLL5 (46) are negative regulators. Thus, their up-regulation or down-regulation would promote gene expression and cell proliferation, which in turn would contribute to the development of leukemia.

\textit{LTF, LCN2, MMP9, S100A8, S100A9, PADH4, TGFBI, and CYBB} are the eight most down-regulated genes (each has at least 30-fold change in expression) in both types of human and murine leukemia samples (Fig. 1A) and were also confirmed by our qPCR assay (Fig. 1B). Interestingly, five (i.e., \textit{LTF, MMP9, S100A8, S100A9, and TGFBI}) of these eight genes are enriched in a GO term, namely extracellular space, and the other two genes (i.e., \textit{TNEAIP2 and LYZ}) enriched in this term are also down-regulated in \textit{MLL}-associated leukemias. In addition, \textit{LTF, S100A8, S100A9, and CYBB} are involved in both response to stimulus and defense response, along with \textit{TGFBI} involved in the former (see Supplementary Tables S5 and S6).

As shown in Fig. 2, up-regulated genes have a much higher percentage of enrichment in GO terms related to gene expression and transcription, whereas down-regulated genes are more enriched in GO terms related to apoptosis, signal transduction, and response. Thus, the up-regulation of genes responsible for gene expression and transcription but the down-regulation of genes responsible for apoptosis, signal transduction, and response can promote cell proliferation and inhibit apoptosis and, thereby, contribute to the development of leukemia.

It is well known that \textit{MLL} fusion proteins can promote expression of many target genes, such as \textit{HOX} genes and \textit{MEIS1} (3, 4). However, the mechanism underlying the down-regulation of a large number of genes in \textit{MLL} rearrangement leukemias is unclear. In the present study, we showed that down-regulation of many genes in the leukemias might be associated with DNA methylation (see Fig. 3). Nonetheless, the genetic or epigenetic changes that caused the DNA methylation remain to be explored.

Interestingly, we found that \textit{BMI1} was consistently overexpressed in human and murine \textit{MLL}-associated leukemia cells relative to the normal controls (see Fig. 1B). \textit{BMI1} is a member of polycomb group proteins, which act as negative regulators of transcription (47). Thus, it is important to conduct systematic studies in the future to determine whether down-regulation of many genes in \textit{MLL}-associated leukemias is related to the deregulation of \textit{BMI1} and/or other transcription repressor(s).

We found that all of the 88 candidate genes have putative miRNA regulators (see Supplementary Table S7). Interestingly, 19 putative targets of \textit{mir-17-92} seemed to be down-regulated in \textit{MLL} rearrangement leukemias, as reported previously or detected by our SAGE assay; the down-regulation of \textit{APP, RASSF2, and SH3BP5} and the up-regulation of the \textit{mir-17-92} cluster have been confirmed by our qPCR. We have confirmed both \textit{APP} and \textit{RASSF2} as direct targets of \textit{miR-17} through luciferase reporter and mutagenesis assay (see Fig. 4B). Interestingly, \textit{RASSF2} encodes a protein that contains a Ras association domain and is a novel tumor-suppressor gene, which has been observed to regulate Ras signaling and play a pivotal role in the early stages of colorectal tumorigenesis (48). Therefore, miRNAs may play an important role in the development of \textit{MLL} rearrangement leukemia. In addition, recent studies also showed that some miRNAs, e.g., the \textit{miR-290} cluster, may control de novo DNA methylation through regulating Rbl2-dependent Dnmt expression (49, 50). Thus, besides the direct regulation of the expression of target genes, some miRNAs may also contribute to the DNA methylation of many genes and, thereby, regulate expression of these genes indirectly. Therefore, a combined study of miRNA and mRNA expression profiles would provide a more complete understanding of the complexity of molecular networks in diseases, such as \textit{MLL} rearrangement leukemia.

In all, our study showed that the deregulation patterns of many protein-coding (e.g., \textit{HOX} genes and \textit{MEIS1}) and noncoding (e.g., individual miRNAs within the \textit{mir-17-92} cluster) genes observed in human leukemia cells are also conserved in mouse leukemia cells, highlighting the importance of these genes in leukemogenesis. The identification and validation of consistent changes of gene expression in human and murine \textit{MLL} rearrangement leukemias contribute important information leading to the understanding

---

**Figure 4.** Expression profiles and targets of the miRNAs in the \textit{mir-17-92} cluster. \textbf{A}, expression profiles of the seven miRNAs in 40 human and mouse samples, as detected by TaqMan qPCR. All the significant genes have a q value of \textit{<0.05}, with an overall FDR of \textit{<0.05}. \textit{cl}, cell line; \textit{N}, normal control; \textit{m}, mouse. \textbf{B}, regulation of \textit{APP} and \textit{RASSF2} by \textit{miR-17} was confirmed by luciferase reporter and mutagenesis assay. HEK293T cells were cotransfected with an expression construct for \textit{miR-17} or an empty vector (control); i.e., MSCVpuro; the luciferase reporter construct containing the wild-type or mutated 3' UTR of \textit{APP} or \textit{RASSF2}, along with a \textit{β}-galactosidase reporter control vector to monitor transfection efficiency. Luciferase was measured 42 h after transfection. The firefly luciferase activity was normalized to \textit{β}-galactosidase activity. The normalized luciferase activities represent the firefly/\textit{β}-galactosidase ratios normalized to the control sample. Error bars present SD obtained from three independent experiments. *, \textit{P}<0.01 (paired t test).
References


Consistent Deregulation of Gene Expression between Human and Murine MLL Rearrangement Leukemias

Zejuan Li, Roger T. Luo, Shuangli Mi, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-08-3381

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2009/01/19/0008-5472.CAN-08-3381.DC1

Cited articles
This article cites 50 articles, 19 of which you can access for free at:
http://cancerres.aacrjournals.org/content/69/3/1109.full.html#ref-list-1

Citing articles
This article has been cited by 9 HighWire-hosted articles. Access the articles at:
/content/69/3/1109.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.