Extensive promoter DNA hypermethylation and hypomethylation is associated with aberrant microRNA expression in chronic lymphocytic leukemia

Constance Baer1,*, Rainer Claus1,*, Lukas P. Frenzel2,3,4,*, Manuela Zucknick5, Yoon Jung Park1,6, Lei Gu1, Dieter Weichenhan1, Martina Fischer5, Christian Philipp Pallasch2,3,4,7, Esther Herpel8, Michael Rehli9, John C. Byrd10, Clemens-Martin Wendtner2,3,4, Christoph Plass1

1Department of Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; 2Department I of Internal Medicine, University Hospital of Cologne, 50937 Cologne, Germany; 3Center of Integrated Oncology (CIO), University Hospital of Cologne, 50937 Cologne, Germany; 4Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University Hospital of Cologne, 50937 Cologne, Germany; 5Division of Biostatistics, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; 6Department of Nutritional Science and Food Management, Ewha Womans University, 120-750 Seoul, Republic of Korea; 7Koch Institute for Integrative Cancer Research at MIT, Cambridge, MA 02142, USA; 8Institute of Pathology, University of Heidelberg, 69120 Heidelberg, Germany; 9Department of Hematology and Institute of Pathology, University Hospital Regensburg, 93042 Regensburg, Germany; 10Division of Hematology, Department of Internal Medicine, The Ohio State University, Columbus, OH 43210, USA

*These authors contributed equally

Running Title:
Epigenetic Regulation of miRNA in CLL

Key words:
miRNA, transcriptional regulation, DNA methylation, epigenetics, CLL

Financial support:
This study was funded by the Deutsche José Carreras Leukämie-Stiftung (DJCLS R 10/27) (R.C., C.P., Y.J.P., C.-M.W., C.P.P.) and supported in part by funds and fellowships from the German Funding agency (DFG) to R.C., C.P., L.P.F., C.-M.W.. C.B. holds a stipend of the Helmholtz International Graduate School for Cancer Research. Y.J.P holds a stipend of the Roman Herzog research fellowship from the Herti foundation.

To whom correspondence should be addressed:
Christoph Plass, Ph.D.
German Cancer Research Center (DKFZ)
Im Neuenheimer Feld 280
69120 Heidelberg, Germany
Phone: +49 6221 42 3300
Fax: +49 6221 42 3359
E-mail: c.plass@dkfz-heidelberg.de

Rainer Claus, MD
German Cancer Research Center (DKFZ)
Im Neuenheimer Feld 280
69120 Heidelberg, Germany
Phone: +49 6221 42 3304
Fax: +49 6221 42 3359
E-mail: r.claus@dkfz.de

Disclosure of Potential Conflicts of Interest:
No potential conflicts of interest were disclosed.

Word count (excluding references): 4217
Number of Figures: 4
Number of Tables: 1
Abstract

Dysregulated microRNA (miRNA) expression contributes to the pathogenesis of hematopoietic malignancies, including chronic lymphocytic leukemia (CLL). However, an understanding of the mechanisms that cause aberrant miRNA transcriptional control is lacking. In this study, we comprehensively investigated the role and extent of miRNA epigenetic regulation in CLL. Genome-wide profiling performed on 24 CLL and 10 healthy B cell samples revealed global DNA methylation patterns upstream of miRNA sequences that distinguished malignant from healthy cells and identified putative miRNA promoters. Integration of DNA methylation and miRNA promoter data led to the identification of 128 recurrent miRNA targets for aberrant promoter DNA methylation. DNA hypomethylation accounted for over 60% of all aberrant promoter-associated DNA methylation in CLL, and promoter DNA hypomethylation was restricted to well-defined regions. Individual hyper- and hypomethylated promoters allowed discrimination of CLL samples from healthy controls. Promoter DNA methylation patterns were confirmed in an independent patient cohort, with eleven miRNAs consistently demonstrating an inverse correlation between DNA methylation status and expression level. Together, our findings characterize the role of epigenetic changes in the regulation of miRNA transcription and create a repository of disease-specific promoter regions that may provide additional insights into the pathogenesis of CLL.
Introduction

Chronic lymphocytic leukemia (CLL) is the most frequent leukemia of adults in the Western world and is characterized by clonal accumulation of malignant B cells with a low proliferation rate and disrupted apoptotic mechanisms (1). The frequent deletion of the long arm of chromosome 13 (del(13q14)) harboring the mir-15a/16-1 locus has drawn attention to microRNA (miRNA) involvement in CLL pathogenesis (2).

MiRNAs are small non-coding RNAs that contribute to post-transcriptional gene expression control. They are transcribed by RNA-polymerase II as primary miRNA transcript, processed by the RNase Drosha to pre-miRNAs (referenced as mir in this report) and further cleaved by Dicer to short mature miRNAs (miR). Mature miRNAs are loaded into the RNA-induced silencing complex, where base-pairing of the miRNA to the 3´ untranslated regions of target mRNAs leads to mRNA degradation or inhibition of translation (3).

In recent years, several miRNA expression studies found extensive dysregulation of many different miRNAs in CLL (4-6) and other tumor entities (7). Dysregulation of miRNAs was shown to affect expression of tumor suppressor- or oncogenes and consequently participate in the initiation and progression of malignant phenotypes (7). For example, miR-15a/16-1 target the apoptosis regulator BCL2 and thereby act as tumor suppressors. Deletion of both miRNAs was identified to promote CLL in humans (8) and to recapitulate the CLL phenotype in mice (9). An oncogenic function of miRNAs in CLL has been demonstrated by B cell-specific overexpression of miR-29 giving rise to an indolent lymphocytic leukemia after a latency of approximately two years (10).

The mechanisms leading to aberrant expression of miRNAs are not yet completely understood. Indirect evidences for epigenetic regulation of miRNAs stem from DNA methyltransferases (DNMT) deficient cell lines or studies using demethylating drugs to induce transcriptional reactivation of miRNAs (11-13). Selected candidate miRNAs have been associated with DNA hypermethylation in CLL (6) or solid tumors (14, 15). However, a major drawback of previous reports is the missing identification of miRNA promoters. So far,
studies focused almost exclusively on upstream CpG islands or regions in direct vicinity of pre-miRNAs (16, 17). MiRNA promoters can be identified by the presence of RNA polymerase II (18) or trimethylated histone 3 lysine 4 (H3K4me3) (19, 20), a well known chromatin mark of active transcription. However, considering the constant increase of miRNA annotations (21) as well as the tissue specificity of those surrogate markers, these earlier studies have been able to identify promoters only for a limited set of miRNAs. By now, integration of miRNA promoter regions and epigenetic profiling has been successful in healthy mammary tissue (22).

In the current study, we present a systematic genome-wide profiling for epigenetic regulation of miRNAs in CLL compared to healthy B cells by simultaneous detection of aberrant DNA methylation and miRNA promoters. We characterize the extent and role of epigenetic alterations of miRNA transcriptional regulation in CLL specimens and report discovery of novel aberrantly regulated miRNAs. In addition, we generate a repository of identified putative promoter regions that are of high interest to further miRNA-related research questions.
Materials and Methods

Patient specimens
Peripheral blood was obtained from healthy donors and CLL patients seen at the Department I of Internal Medicine, University of Cologne according to Institutional Review Board-approved protocols after receipt of written informed consent according to the Declaration of Helsinki (Supplementary Table S1). Blood specimens from CLL patients and healthy donors were either enriched for B cells with a purity of more than 95% by applying BRosetteSep (StemCell Technologies) or Ficoll-Hypaque (Seromed) density gradient purification and positive magnetic cell sorting for CD19. Granulocytes and T cells were selected by positive magnetic cell sorting for CD15 and CD3 (Miltenyi Biotec GmbH), respectively. Solid tissue samples were obtained from the tissue bank of the National Center of Tumor Diseases Heidelberg. The following tissues were used: transmural tumor free colon tissue, tumor free pancreatic parenchyma, normal lymph node, univacular adipose tissue, benign prostate hyperplasia, tumor free liver tissue, tumor free renal cortex and tumor free lung parenchyma. DNA and RNA were isolated by DNeasy Blood and Tissue Kit, TRizol (Life Technologies) or AllPrep DNA/RNA Mini Kit (Qiagen) following the manufacturer’s instructions.

Cell lines and 5-aza-2’-deoxycytidine treatment
For identification of H3K4me3 positive regions upstream of miRNA loci, the CLL related cell lines MEC-1, EHEB, GRANTA-519 and the T cell line JURKAT were obtained from the German Collection of Cell Lines and Microorganisms (DSMZ). Authentication was performed by the DSMZ using short tandem repeat profiles less than 6 month before experiments were performed or cells were frozen. The myeloid cell lines KASUMI-1 and HL-60 were obtained from DSMZ and authenticated by the Genomics and Proteomics Core Facility at the German Cancer Research Center using multiplex PCR-based amplification of 24 SNP regions. All cell lines were cultured as recommended by the DSMZ (s. Supplementary Methods) and treated with 1.5 μM (MEC-1, GRANTA-519 and EHEB) or 200 nM (KASUMI-1) 5-aza-2’-
deoxycytidine (Sigma) over 7 days (GRANTA-519 over 8 days) by replacing drug and medium every 24h. DNA demethylation efficiency was evaluated by quantitative DNA methylation analysis of repetitive elements (Supplementary Table S2). HCT116 and HCT116 **DNMT1-/-, DNMT3B-/-** were obtained from Johns Hopkins University, Laboratory of Dr. B. Vogelstein and cultured as previously reported (23).

**Methyl-CpG immunoprecipitation (MCIp)**

MCIp was performed as described previously (24) with modifications (s. Supplementary Information).

**Chromatin Immunoprecipitation (ChIP)**

Cells (2*10^8) were formaldehyde crosslinked as previously described (25) directly after purification (primary cells) or after 5-aza-2’deoxyctydine treatment (cell lines). The cell pellets were resuspended in 1 ml lysis buffer (50 mM Tris, 10 mM EDTA, 0.5% SDS, Protease Inhibitor Cocktails (Roche)) and sonicated with Bioruptor Next Gen (Diagenode). Immunoprecipitation was performed by the SX-8G IP-Star Automated System using 2 µl of polyclonal antibody against trimethylated histone 3 lysine 4 (H3K4me3, pAb-003-050), the Auto ChIP Kit and the Auto IPure Kit (Diagenode).

**Microarray design**

Genomic locations of 939 annotated miRNAs from miRBase 15 (21) and control probes were tiled on an Agilent custom-design 244k array. The probe tiling included 35 kb upstream and 5 kb downstream of the pre-miRNAs and -/+ 2 kb regions around transcriptional start sites of miRNA hosting genes. The Array design is available at Agilent’s earray platform (AMADID27633 for human genome assembly hg18, AMADID029434 for hg19).

**Microarray hybridization and readout**

Fluorescent labelling of enriched DNA fragments was performed using the BioPrime Total Genomic Labeling System (Invitrogen). Hybridization, washing and scanning of microarrays were performed following the manufacturer's protocol for human CpG island microarrays and
mammalian ChIP-on-chip for MCIp and ChIP, respectively. For array scanning, the Agilent G2565BA Microarray Scanner was used. Data from image files were extracted using Agilent’s Feature Extraction Software 10.5.11 (Protocol: ChIP_105_Dec08, Grid: 027633_D_F_20100318, 029434_D_F_20100720). Microarray data are available online from Gene Expression Omnibus (GEO) under accession number GSE33347.

**Analysis of MCIp-chip and ChIP-chip data**

Analysis of H3K4me3 ChIP data was performed with the CoCAS ChIP-on-chip analysis suite (26). MCIp-Agilent array data were analyzed using the statistical environment R v2.11, package RJaCGH v2 and the Bioconductor suite. MCIp-array data were background-corrected by the normal-exponential convolution method (27) with offset 50 and within-array normalised by rank-invariant weighted loess regression (28). For each 40 kb tile, Bayesian non-homogeneous hidden Markov models were fitted to assign posterior probabilities of DNA hypomethylation, hypermethylation and no change, incorporating model uncertainty by Bayesian model averaging (29). Using these probabilities, the most recurrent regions of DNA methylation differences over all samples were identified by weighted averages across the posterior probabilities of all CLL samples. Regions separated by spaces less than 650 bp were merged. Regions defined by single probe hits (<100 bp) were excluded from further analysis. Further details of statistical processing are described in the Supplementary Information.

**Quantitative high resolution DNA methylation analysis**

Quantitative DNA methylation was measured using the MassARRAY system as previously described (30). Primers are given in Supplementary Table S2.

**miRNA expression microarray analysis**

MiRNA expression arrays using BeadArray technology and Human v2 miRNA arrays were processed according to the manufacturer’s instructions (Illumina). Arrays were scanned using a BeadArray reader and data were analyzed with Illumina’s GenomeStudio Gene
Expression Module. Raw data were quantile-normalized. Detailed of statistical analysis are given in Supplementary Information. Microarray data are available online (31).

**Quantitative PCR**

MiRNA expression by quantitative PCR (qPCR) was carried out using the miScript Reverse Transcription Kit, miScript SYBR Green PCR Kit and miScript Primer Assays (Qiagen) according to the manufacturer's instructions. Relative miRNA expression was calculated by normalizing to the average of the following housekeeping RNAs: SCARNA17, SNORD25 and SNORA73A.

**Target gene prediction and pathway analysis**

Targets genes of miRNAs were predicted using miRanda (32) and were considered when reaching a mirSVR (33) score below -0.1. Pathway analysis was conducted using miTALOS (34).
Results

CLL cells exhibit specific distinct DNA methylation profiles upstream of microRNA loci

In order to identify disease-specific differentially methylated miRNA promoters in CLL, we applied a strategy of combined DNA methylation profiling and miRNA promoter detection (Fig. 1A). Signal detection was based on a custom two color tiling array covering -35 kb/+5 kb of the pre-mir locus and the host gene promoter (Fig. 1B). We enriched methylated DNA from 24 CLL patients (Supplementary Table S1) and from B cells of ten age-matched healthy donors combined as reference pool. Hybridization of the methylated DNA fractions of individual CLL patients against the methylated DNA of the healthy B cell reference allowed identification of differentially methylated regions. We detected pronounced DNA methylation changes in individual CLL patients compared to the healthy control references. The extent of differential DNA methylation considerably exceeded the interindividual differences observed between healthy donors compared to the reference pool (Fig. 1C). By fitting a Bayesian non-homogenous hidden Markov model for each individual array, we assigned sample specific posterior probabilities of hyper- and hypomethylation to regions larger than 100 bp. In order to identify regions that harbor a high potential of pathological relevance for CLL, the most recurrent DNA methylation differences were detected by weighted averages across the posterior probabilities of all CLL samples. Excluding the X chromosome, 531 differentially methylated regions (DMRs) with a median size of 385 bp were defined upstream of miRNAs and in host gene promoters. DNA methylation differences (as compared to the reference pool) were subjected to unsupervised hierarchical clustering revealing a clear separation of individual healthy B cell samples from CLL samples and a high abundance of DNA hypomethylation in CLL patients (Fig. 1D). Of all identified DMRs, 408 regions showed loss of DNA methylation in CLL whereas 123 regions displayed increased DNA methylation.

Next, we aimed at comprehensively identifying miRNA promoters to assign functional relevance to the discovered differentially methylated regions. We enriched chromatin for
trimethylated histone 3 lysine 4 (H3K4me3), a hallmark of active promoters, and analyzed the enriched DNA using the custom tiling array platform. To avoid limitations due to tissue-specific expression patterns of miRNAs, we included healthy B cells, CLL cells and a variety of tumor cell lines including CLL related cell lines. In addition, CLL related cell lines were pharmacologically demethylated by 5-aza-2’deoxyctydine and HCT-116 cells carrying a double knockout for DNMT1 and DNMT3B (23) were used for promoter identification. All H3K4me3 peak data sets were integrated in a comprehensive list and non-miRNA associated hits were removed. In total, we discovered putative promoter regions for 781 miRNAs, which are subsequently designated as promoters (Supplementary Table S3). The validity of this approach was confirmed by a high overlap with complementary promoter-associated features: 77% of the identified regions coincided with start sites of expressed sequence tags, 71% exhibited highly accessible chromatin as measured by DNase hypersensitivity and 34% harbored CpG islands.

Next, we focused on those DMRs coinciding with promoters (Fig. 1D lower panel). Individual CLL patients carried DNA hypermethylation at a median of 40 (range 16-55) miRNA promoters and DNA hypomethylation at a median of 60 (range 12-89) miRNA promoters per patient (Fig. 1E). MiRNA promoter associated DMRs differed substantially in their sequence context as 84% of the hypermethylated but only 6% of the hypomethylated promoter sequences co-localized with CpG islands (Fig. 1F).

**Inverse correlation of promoter DNA methylation and expression of mature miRNA**

In total, we identified 128 miRNAs that carry aberrant DNA methylation at a putative promoter (Supplementary Table S4). Some miRNAs harbor multiple differentially methylated H3K4me3 enriched regions and, thus, may possess two or more promoters, e.g. miR-9. In order to address transcriptional consequences of promoter DNA methylation for the 128 candidates, array-based miRNA expression data was generated from the same samples used for the DNA methylation screening (Table 1 and Supplementary Table S5).

Inversely correlating DNA methylation and expression, we identified twelve miRNAs that were candidates for DNA methylation dependant regulation: mir-129-2, mir-551b, mir-708
and mir-21, mir-34a, mir-135a, mir-155, mir-574, mir-664, mir-1204, and the cluster of mir-29a/29b-1 (Table 1). Furthermore mir-9 and mir-124-2 were included as they are known to be frequently epigenetically silenced in various tumor entities (35, 36).

In CLL, aberrant promoter DNA hyper- or hypomethylation in those 14 promoters was confirmed and quantified by high resolution MassARRAY analysis (Fig. 2). Unsupervised hierarchical clustering of the absolute DNA methylation values revealed a clear separation of CLL patients from healthy B cells. Of note, DNA methylation data of single miRNA promoters (e.g. mir-1204) was already enough to clearly separate CLL and healthy B cells by unsupervised clustering.

Furthermore, we could validate 10 out of 11 additional candidates selected from the list of 128 recurrently differentially methylated miRNA promoters (Supplementary Table S4). These miRNAs did not show significant expression changes in our analysis (Supplementary Fig. S1) but were in part previously identified to be differentially expressed in CLL (miR-132, miR-190 (4), miR-451 and miR-598 (6)).

Epigenetically silenced miRNAs in CLL

Promoters of the intergenic mir-124-2 and mir-129-2 as well as the intragenic mir-9-2, mir-551b and mir-708 showed consistent hypermethylation in an independent patient cohort (Supplementary Table S1, Fig. 3A, 3B). DNA hypermethylation correlated with significantly reduced expression as assessed by qPCR (Fig. 3C). None of those have been previously described as being aberrantly regulated in CLL. Mir-129-2 exhibited gain of DNA methylation at two CpG islands: one covering the sequence of the pre-miRNA and one 2 kb upstream (Fig. 3A, 3B). In the initial screening, stronger increase in DNA methylation was detected at the upstream CpG islands. By luciferase promoter assays, we could demonstrate promoter activity of the upstream CpG island compared to neighboring regions (Supplementary Fig. S2). Mir-9 is hosted in the transcript LINC00461, a non-coding RNA that has three annotated transcript variants. All three promoters showed enrichment of H3K4me3 in healthy B cells and increased DNA methylation in CLL patients (Fig. 3A). We tested DNA methylation in two
of these promoters by MassARRAY and found significantly increased DNA methylation. This was in concordance with downregulation of mature miR-9 in the majority of CLL patients. Next, we used the miRanda algorithm (32) to predict target genes of these epigenetically silenced miRNAs. Among all predicted targets, we identified genes that were recently found to have relevance for CLL pathogenesis as they carry somatic mutations in functional domains: NOTCH1, XPO1, POT1 and ZMYM3 (37-39). miR-129 is predicted to target XPO1 and ZMYM3, miR-9 may possibly regulate POT1 and miR-708 NOTCH1. All four genes show increased expression in CLL cells compared to healthy B controls (Supplementary Fig. S3), inversely correlating with the epigenetically reduced expression of the miRNAs.

**Epigenetically reactivated miRNAs in CLL**

In the validation set, the promoters of mir-21, mir-29a mir-34a, mir-155, mir-574, and mir-1204 exhibited significant local DNA methylation decrease accompanied by consistent transcriptional upregulation (Fig. 4). Although, mir-29a and mir-29b-1 are potentially regulated as a cluster by one promoter, no significant expression increase of miR-29b was observed. Expression changes for miR-135a, miR-596 and miR-664 did not reach statistical significance in the validation cohort.

The hypomethylated profiles were specific for CLL and clearly distinguished CLL cells from a large panel of different healthy tissues (Supplementary Fig. S4 and Supplementary Table S6 for expression). We also found that the DNA methylation pattern obtained in those six miRNA promoters exhibited pronounced tissue-specific DNA methylation differences clearly discriminating normal hematopoietic cells (B and T cells and granulocytes) and healthy solid tissues by unsupervised clustering. This strong separation was not observed for the hypermethylated miRNA promoters (Supplementary Fig. S4).

For mir-21, high promoter activity was identified in CLL cells at a locus in the last intron of the VMP1 gene covering the previously described promoter (20) (Fig. 4A). In the independent patient cohort, we validated both significant DNA hypomethylation in a promoter-associated sequence stretch 2 kb upstream of the mir-21 sequence and a two-fold upregulation of miR-21 expression (Fig. 4B, 4C). Mir-34a is located within transcript EF609116 carrying a
promoter CpG island (Fig. 4A). We discovered an alternative site 12 kb upstream of the mir-34a sequence characterized by H3K4me3 enrichment and other promoter-associated features (e.g. DNase hypersensitivity). Regionally limited, significantly reduced DNA methylation was found only at this alternative site in both the screening and the validation cohort (Fig. 4A, 4B). This DNA hypomethylation correlated with upregulation of miR-34a in CLL patients (Fig. 4C). For mir-155, we found DNA hypomethylation in CLL cells clearly limited to a sequence stretch adjacent to the promoter CpG island of the MIR155 host gene (MIR155HG) (Fig. 4A). This was validated in the independent cohort of CLL patients (Fig. 4B). Although not covering the CpG island, the hypomethylated region is restricted to clearly defined CpG dinucleotids and coinciding with numerous transcription factor binding sites including the miR-155 regulator MYB (40) and the miR-155 target PU.1 (41) (Supplementary Fig. S5A). Loss of promoter DNA methylation in this region correlated with increased expression of mature miR-155 (Fig. 4C). The relevance of this region adjacent to the CpG island is further supported by a high correlation of tissue specific DNA methylation with expression of miR-155 in the respective tissue (R²=0.7). The CpG island itself displayed a homogeneously low level of DNA methylation not correlating with expression of miR-155 in the analyzed tissues (Supplementary Fig. S5B).

To evaluate the functional importance of the epigenetically reactivated miRNAs we combined miRNA target prediction with tissue specific pathway analysis by the miTALOS platform (34). We found significant enrichment for targets involved in apoptosis (p=0.005), a pathway known to be defective in CLL.


**Discussion**

Increasing evidence supports the hypothesis that epigenetic mechanisms are involved in regulating miRNA expression (6, 12, 17, 42, 43). In the present study, we used a combined strategy to assign aberrant DNA methylation to putative miRNA promoters. Thereby, we identified extensive disease specific DNA methylation patterns. Moreover we discovered that not only hyper- but also hypomethylation at putative miRNA promoters correlates with expression of the mature miRNAs.

Enrichment of H3K4me3 was previously demonstrated to serve as a valid and reliable surrogate for active promoters of protein coding genes and miRNAs (19, 44). To obtain a large diversity of H3K4me3 signals, we used a variety of cell lines derived from different tissues in addition to primary healthy B cell and CLL samples. In total, we were able to identify putative promoter regions for 781 miRNAs. Most of the promoter regions are novel and co-appearance with additional promoter characteristics demonstrated the validity of our approach. Of the limited number of previously published promoters (18, 20), an overlap of more than 70% with our data set could be noted.

Profiling DNA methylation at the identified promoter regions revealed 38 and 90 miRNAs as consistent targets of promoter DNA hyper- or hypomethylation, respectively. Thus, our results clearly indicate that differential DNA methylation is frequent at miRNA promoters and thereby might constitute a major mechanism leading to transcriptional miRNA deregulation in CLL. We demonstrate that not only aberrantly increased but also decreased DNA methylation levels at miRNA promoters are of functional relevance for the transcriptional control of miRNAs. This finding has so far been underestimated. Particularly, decreased DNA methylation at distinct loci (e.g. mir-1204, see Fig. 4B) generated high contrasts of DNA methylation levels and clearly separated CLL cases from controls. The hypomethylated promoter regions in CLL showed a high degree of tissue specific methylation in a panel of healthy tissues including hematopoietic lineages as well as solid tissues. This further supports the regulatory importance of these regions (see Supplemental Fig. S4). It is
noteworthy that the observed DNA hypomethylation events in CLL were regionally limited with clear boundaries to surrounding sequences as seen for mir-29a/29b-1 (see Fig. 4A, 4B) and mir-155 (see Supplementary Fig. S5). These patterns suggest activity of directed demethylating mechanisms rather than unspecific genome-wide loss of DNA methylation. The frequent appearance of hypomethylated miRNA promoters correlates with the observation that in CLL, more miRNAs are upregulated than downregulated (4, 5), which is in contrast to many other malignancies. While DNA hypermethylation occurred preferentially in CpG islands, decreased DNA methylation was nearly exclusively found in CpG-poor regions (see Fig. 1F). This is in agreement with recent observations reported for protein coding genes (45). The focus on CpG islands and the detection of differentially methylated miRNA promoters by pharmacological DNA demethylation could be one of the reasons why previous studies underestimated the extent of DNA hypomethylation (6, 12, 16).

We focused on 11 miRNAs that were consistently epigenetically deregulated and showed correlation of expression and DNA methylation in an independent validation cohort of CLL patients thus suggesting general relevance for CLL pathogenesis. However, methylation of distinct miRNA promoters could also recapitulate clinically relevant subgroups as it has already been shown for miRNA expression. For example, expression levels of the miR-29 family or miR-34a (46) possess prognostic relevance and form subgroups of differential clinical outcome. Interestingly, their promoter methylation pattern is, although generally lower in CLL, not completely homogenous among patients (see Fig. 4B). Follow-up analyses in large CLL study cohorts may allow dissecting prognostically or therapeutically relevant subgroups based on miRNA promoter methylation.

The epigenetically dysregulated miRNAs show enrichment of target genes involved in apoptosis, a defective key pathway in CLL cells (1) and have predicted targets genes, recently identified to carry mutations in functional domains in CLL.

Among the transcriptionally repressed miRNAs, we identified mir-129-2, which was previously detected to be epigenetically silenced in solid tumors where it functions as a tumor suppressor by targeting the mRNA of the oncogene SOX4 (14). This work, however, focused
on increased DNA methylation at the CpG island that directly covers the pre-mir sequence stretch. By luciferase reporter constructs, we verified promoter activity at a different significantly hypermethylated CpG island located approximately 2 kb upstream of the pre-mir-129-2. The regulatory relevance of this site is further supported by the start of two expressed sequence tags (BI964058, BD120451) that could be identical with the primary miRNA transcript. An alternative epigenetically altered site was also detected for mir-34a, a downstream target of the p53 pathway. DNA hypermethylation at the promoter of the hosting transcript EF609116 has been extensively studied in various tumor entities (47). We detected a potential alternative transcriptional start site approximately 12 kb upstream of the mature miRNA located within the hosting transcript. In CLL, significant DNA hypomethylation accompanied by increased expression could be noted, whereas the hosting transcript promoter did not exhibit differential DNA methylation (see Fig. 4A). Whether this novel regulatory site is also aberrantly methylated in other tumor tissues remains to be determined.

For the epigenetically reactivated miR-155, we identified a regionally restricted significantly hypomethylated region adjacent to the promoter CpG island coinciding with transcription factor binding sites of PU.1, NFκB and a MYB consensus sequence (see Supplementary Fig. S5). MYB was recently shown to bind to the mir-155 promoter and thereby to contribute to its regulation in CLL (40). The high correlation of miR-155 expression and DNA methylation in this region as determined in different healthy tissues emphasizes the regulatory function of the region adjacent to the CpG island. Evidence for the importance of miR-155 in CLL is provided by overexpression in a mouse model, which leads to a high-grade B cell malignancy (48).

In addition to miRNAs previously described in the context of CLL pathogenesis, we also identified epigenetic regulation of novel miRNAs, e.g. miR-551b, demonstrating that combined epigenetic profiling and expression screening is an effective strategy of identifying novel aberrantly transcribed miRNAs. Recent miRNA expression studies generated partially inconsistent candidate lists (5, 6, 49). In future studies, combined analyses of expression and epigenetic profiles could increase sensitivity and improve the detection of significantly and
constantly deregulated miRNAs. Therefore, the generated repository of 781 putative miRNA promoters is a valuable resource for epigenetic and functional analyses also in other entities. Remarkably, the number of deregulated miRNAs in the previously published expression profiling studies (5, 6, 49) was much smaller than the number of epigenetically altered miRNA promoters identified in our work. Several reasons might account for this discrepancy. Firstly, many miRNAs are transcribed from different loci in the genome, e.g. mir-9-1, mir-9-2 and mir-9-3, and share identical sequences of their mature forms. To date, array-based expression analysis does not offer the possibility to discriminate these transcripts and to assign aberrant transcription to altered DNA methylation at distinct promoters. Secondly, it has been shown in T cells that an altered epigenetic status does not necessarily affect the transcription of a miRNA directly but poises miRNA promoters and creates a permissive state for transcription initiation upon activating signals (50). Thus, in CLL, the distinct identified epigenetic profile could be representative for transcriptional activity upon different pathogenesis relevant stimuli e.g. microenvironment contact or cell stress.
Acknowledgments

We are grateful to Oliver Mücke (Heidelberg, Germany) for excellent technical support with MassARRAY DNA methylation analyses as well as Julia Claasen and Reinhild Brinker (Cologne, Germany) for excellent assistance with sample preparation and microarray-based miRNA expression assays.

Grant Support

This study was funded by the Deutsche José Carreras Leukämie-Stiftung (DJCLS R 10/27) (R.C., C.P., Y.J.P., C.-M.W., C.P.P.) and supported in part by funds and fellowships from the Deutsche Forschungsgemeinschaft (DFG) to R.C., C.P., L.P.F., C.-M.W.. C.B. holds a stipend of the Helmholtz International Graduate School for Cancer Research. Y.J.P holds a stipend of the Roman Herzog research fellowship from the Hertie foundation.
References


## Tables

### Table 1. Dysregulated miRNAs in CLL patients versus B cells from healthy donors.

<table>
<thead>
<tr>
<th>microRNA</th>
<th>location</th>
<th>DMR expression difference, log₂ FC (miRNA)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>hypermethylated</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| mir-9-2  | LOC645323| chr5:87968473-87968894, chr5:87970058-87971540,
|          |          | chr5:87973768-87974628, chr5:87975816-87976597,
|          |          | chr5:87980080-87981723                       |
|          |          | literature<sup>a</sup>                       |
| mir-124-2| AK124256 | chr8:65285833-65286806, chr8:65289299-65291897 |
|          |          | literature<sup>a</sup>                       |
| mir-129-2| i        | chr11:43601120-43601393                      |
|          |          | -0.48 (miR-129-3p); - 0.48 (miR-129*)        |
| mir-551b | EGFEM1P  | chr3:167967620-167968099                     |
|          |          | -0.22 (miR-551b)                            |
| mir-708  | ODZ4     | chr11:79147754-79148140                     |
|          |          | -1.23 (miR-708)                             |
| **hypomethylated** |
| mir-21   | i        | chr17:57916274-57916703                     |
|          |          | 0.53 (miR-21*)                              |
| mir-29a, mir-29b-1| anti | chr7:130585935-130586682                      |
|          |          | 0.22 (miR-29a*); 0.40 (miR-29b-1*)          |
| mir-34a  | EF570048 | chr1:9222419-9223806                         |
|          |          | 0.32 (miR-34a); 0.36 (miR-34a*)             |
| mir-135a-1| anti  | chr3:52351422-52351797                      |
|          |          | 0.46 (miR-135a)                             |
| mir-155  | MIR155HG | chr21:26933508-26934239                      |
|          |          | 0.30 (miR-155); 1.017 ( miR-155*)           |
| mir-574  | FAM114A1 | chr4:38869558-38869937                       |
|          |          | 0.88 (miR-574-3p); 0.30 (miR574-5p)          |
| mir-596  | i        | chr8:1736148-1736268                         |
|          |          | 0.27 (miR-596)                              |
| mir-664  | RAB3GAP2 | chr1:220393518-220393902                     |
|          |          | 0.21 (miR-664); 0.36 (miR-664*)             |
| mir-1204 | PVT      | chr8:128808221-128808385                     |
|          |          | 0.15 (miR-1204)<sup>b</sup>                 |

<sup>a</sup>miR-9 and miR-124 are frequently shown to be epigenetically regulated in solid tumors and hematopoietic malignancies. <sup>b</sup>miR-1204 did not show a log₂ fold change larger than 0.2 in this analysis but was included based upon recent expression array data (L.P.F. and C.-M.W., unpublished data, March, 2012). Abbreviations: FC, fold change; mir: pre-microRNA; i, intergenic; anti, antisense to overlapping transcript.
Figure Legends

Figure 1. Aberrant DNA methylation patterns upstream of miRNAs and at miRNA promoters in CLL. (A) Experimental strategy: Differentially methylated regions (DMRs) and putative miRNA promoters defined by enrichment of trimethylated histone 3 lysine 4 (H3K4me3) were overlapped to define regions of interest. Candidate miRNAs were correlated with expression and DNA methylation differences were further validated in an independent sample set. (B) The tilling array used for identification of DMRs and H3K4me3 enriched regions covers the genomic region from -35 kb to +5 kb around 939 pre-miRNAs (red arrow) annotated in miRBase 15 and from -2 kb to +2 kb around the promoters of hosting transcripts for intragenic miRNAs. (C) Average DNA methylation differences of 24 CLL patients individually hybridized vs. a pool of B cell form 10 healthy donors (hB) are displayed in the tiled regions. B cells of one individual from the pool of healthy donors are also compared against the pool of healthy B cells. miRNA locations are indicated in red. (D) Normalized values form the array are shown for 531 regions that are determined by a Hidden Markov model to be different between the CLL cases and the pool of 10 healthy B cell donors. Healthy B cells from two individuals hybridized against the reference pool cluster independently from all CLL patients using unsupervised hierarchical clustering and Pearson correlation as distance metric. Array probes coinciding with putative miRNA promoters regions are highlighted by an orange bar. (E) Bar graphs representing the number of hyper- (blue) and hypomethylated (yellow) putative promoters in individual patients. The order of patient samples corresponds to that of the patient clustering shown in D. (F) Differential overlap of aberrantly methylated promoter regions (as defined above) and CpG islands (CGI).

Figure 2. Validation of DNA hypermethylation or hypomethylation in 14 miRNA promoters was assessed by quantitative, high resolution MassARRAY technique. Heatmaps display quantitative DNA methylation data. Samples are organized in columns by unsupervised hierarchical clustering (Pearson correlation) separating ten healthy B cells (hB, bold black bar).
and 24 CLL patients (bold grey bar). Rows represent single CpG units. In the heatmap, low DNA methylation levels are depicted in yellow, high DNA methylation levels in dark blue, grey stands for missing values. For mir-9-2 two sequence stretches were measured.

**Figure 3.** Association of promoter DNA hypermethylation and miRNA expression. (A) Schematic representation of array based data obtained for the genomic loci of epigenetically silenced miRNAs. Significant enrichment of H3K4me3 in healthy B cells (hB) as marker for promoter activity is displayed by vertical orange bars, bar heights are proportional to the relative enrichment. Average DNA methylation differences in CLL patients versus healthy B cells are shown as black line. Red arrows display the pre-miRNA (mir) indicating the direction of transcription. CpG islands (CGI) are indicated by green bars and regions for MassARRAY based validation by black bars. Host genes (for miR-551b, miR-708 and miR-9-2) are displayed as line with black boxes representing exons. (B) Quantitative DNA methylation measurement by MassARRAY in an independent validation set of B cells from healthy donors (n=15) and CLL samples (n=48) is displayed as heatmap. Heatmap organization and color correspond to Fig. 2. Significance of DNA methylation difference was calculated by non-parametric Mann-Whitney-U test. (C) Mature miRNA expression is measured by qPCR in an independent cohort of B cells from healthy donors (n=15) and CLL samples (n=32). Significance of expression differences was assessed by unpaired non-parametric Mann-Whitney-U test. Black lines represent the median.

**Figure 4:** DNA hypomethylation corresponds with increased miRNA expression. Fig. 4 systematics correspond to Fig. 3. (A) Array data obtained for genomic loci upstream of miRNA. Enriched H3K4me3 profiles were obtained from CLL cells or the related cell line EHEB (mir-34a, mir-574). (B) Quantitative DNA methylation was measured in an independent validation set by MassARRAY and displayed as heatmap. (C) Expression of the mature miRNA was determined by qPCR, median is shown as black line. Significances of DNA methylation and expression differences were tested by Mann-Whitney-U test.
Extensive promoter hypermethylation and hypomethylation is associated with aberrant microRNA expression in chronic lymphocytic leukemia

Constance Baer, Rainer Claus, Lukas P. Frenzel, et al.

Cancer Res Published OnlineFirst June 18, 2012.

Updated version
Access the most recent version of this article at: doi:10.1158/0008-5472.CAN-12-0803

Supplementary Material
Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2012/06/22/0008-5472.CAN-12-0803.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

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.