Identification of novel sequence variations in microRNAs in chronic lymphocytic leukemia

Jana Kminkova[1], Marek Mraz[1], Kristina Zaprazna[1],
Veronika Navrkalova[1], Boris Tichy[1], Karla Plevova[1,2],
Jitka Malcikova[1,2], Katerina Cerna[1], Tobias Rausch[3],
Vladimir Benes[3], Yvona Brychtova[2], Michael Doubek[1,2],
Jiri Mayer[1,2] and Sarka Pospisilova[1,2,*]

[1]CEITEC - Central European Institute of Technology, Center of Molecular
Medicine, Masaryk University, Brno 625 00, Czech Republic, [2]Department
of Internal Medicine – Hematology and Oncology, University Hospital Brno
and Faculty of Medicine, Masaryk University, Brno 625 00, Czech Republic
and [3]European Molecular Biology Laboratory, Genomics Core Facility,
69117 Heidelberg, Germany

*To whom correspondence should be addressed. Tel: +420 532 234 622;
Fax: +420 532 234 623;
Email: sarka.pospisilova@fnbrno.cz

**MicroRNA (miRNA) expression is deregulated in many tumors
including chronic lymphocytic leukemia (CLL). Although the par-
ticular mechanism(s) responsible for their aberrant expression is not
well characterized, the presence of mutations and single-nucleotide
polymorphisms (SNPs) in miRNA genes, possibly affecting their
secondary structure and expression, has been described. In CLL;
however, the impact and frequency of such variations have yet to be
elucidated. Using a custom resequencing microarray, we screened
sequence variations in 109 cancer-related pre-miRNAs in 98 CLL
patients. Additionally, the primary regions of miR-29b-2/29c and
miR-16-1 were analyzed by Sanger sequencing in another cohort of
213 and 193 CLL patients, respectively. Altogether, we describe six
novel miR-sequence variations and the presence of SNPs (*n* = 27),
most of which changed the miR-secondary structure. Moreover,
some of the identified SNPs have a significantly different frequency
in CLL when compared with a control population. Additionally, we
identified a novel variation in miR-16-1 that had not been described
previously in CLL patients. We show that this variation affects the
expression of mature miR-16-1. We also show that the expression
of another miRNA with pathogenetic relevance for CLL, namely
miR-29b-2, is influenced by the presence of a polymorphic inser-
tion, which is more frequent in CLL than in a control population.
Altogether, these data suggest that sequence variations may occur
during CLL development and/or progression.**

Introduction

MicroRNAs (miRNAs) represent an important class of small, non-coding
RNAs regulating expression of at least one-third of human protein-cod-
ing genes and thus play a critical role in a variety of biological processes
(1–4). To date, the miRNA registry contains 1872 human miRNA pre-
cursors, giving rise to 2578 mature miRNAs (MiRBase, Release 20, June
2013). In the human genome, miRNAs constitute ~3–5% of predicted
genes, which can be present in the intergenic, intronic or exonic regions
of either protein-coding or non-protein-coding genes (5). miRNAs are
more frequently located in cancer-associated regions (6), and represent
ideal candidates for cancer predisposition loci since a small genetic
change can lead to widespread defects in normal cell physiology.

The importance of miRNAs in chronic lymphocytic leukemia (CLL)
pathogenesis began with the discovery of miR-15a and miR-16-1 in
the 13q14 region (6,7), which is frequently deleted in CLL (8). This
suggested that miRNAs can act as tumor suppressor genes in CLL and
for the first time demonstrated their direct role in cancer pathogenesis.

Abbreviations: CLL, chronic lymphocytic leukemia; miRNA, microRNA;
SNP, single-nucleotide polymorphism; wt, wild-type.

Subsequent publications not only described germline mutation in miR-
16-1 in two CLL patients (9), but also implicated the potential use of
miRNAs as prognostic markers in CLL due to their aberrant expres-
sion signatures with respect to *IGHV* mutation status and *TP53* abnor-
malities (9,10). A recent publication describing coupled expression of
immunoglobulin genes and miR-650, which is known to be associated
with CLL prognosis and B-cell proliferation (11), further supported
the importance of miRNAs in CLL [reviewed in references (12,13)].

Deregulation of miRNA expression has been observed in hemato-
logical malignancies (9–11,14–16), and many types of solid tumors
(17–20). Although the causes of their aberrant expression in tumor
cells are only partially known, at least three different mechanisms
have been described, including genomic aberrations involving miRNA
genes in cancer-associated regions, epigenetic regulatory mechanisms
and the presence of sequence variations [mutations and single-nucleo-
tide polymorphisms (SNPs)] (9,19,21).

Sequence variations in a miRNA gene can influence the process-
ing of primary transcripts for miRNAs (pri-miRNA, ~100–1000 nt
long) (21) that are processed by the enzyme Drosha in the nucleus
(22,23) into hairpin-shaped precursor miRNAs (pre-miRNA). Pre-
miRNAs (~70 nt long) are cleaved in the cytoplasm by Dicer enzyme
into 18–25 nt long mature miRNAs (24,25). Sequence variations
present in the seed region (7–8 nt of the 5′ end of mature miRNA),
which is responsible for miRNA binding to 3′-untranslated region of
target messenger RNAs, influence miRNA functions by changing the
pattern of targeted genes, and can also affect susceptibility to cancer
(1,3,5,26–28). Moreover, increasing evidence suggests that miRNA–
messenger RNA interactions can be affected by the presence of SNPs
in target gene's 3′-untranslated region, which results in either the
abolishment of existing binding sites or the creation of new, illegiti-
mate ones. Significantly, these SNPs have also been linked to cancer
susceptibility or pathogenesis (29–31).

Although several miRNA mutations and SNPs have been described
in CLL (9,32), their overall frequency and impact is still unresolved.
In 98 CLL patients, we screened sequence variations in 109 pre-
miRNAs (Supplementary Table I, available at *Carcinogenesis* Online)
involved in CLL pathogenesis, other hematological malignancies and
in hematopoesis. Furthermore, the presence of sequence variations
was studied in more detail in the primary regions of miR-29b-2/29c in
another cohort of 213 CLL patients since these two miRNAs belong
to the most important group of miRNAs involved in CLL biology.
Both miR-29b-2 and miR-29c were shown to be downregulated in
aggressive CLL (9,10). Moreover, miR-29 was suggested to target the
expression of *MCL1* (33) and *TCL1*, a critical oncogene in aggressive
CLL (16). Recently, the generation of transgenic mice overexpressing
miR-29 in B cells demonstrated its direct role in CLL pathogenesis
(34). Additionally, mutations in miR-29c and miR-29b-2 have also
been detected in CLL patients (9).

In total, we identified 6 novel variations and 27 SNPs in our CLL
patient cohorts. miRNA–SNP frequency was compared between
CLL patients and a control population, and we show that some of the
identified SNPs may have significantly different frequencies in CLL
patients. Most of the detected variations also affected miRNA second-
ary structure. We show that the expression of miR-16-1 is affected
by the presence of the novel pre-miR-16-1 variation detected in our
study. The effect of miR-29b-2 polymorphic insertion on miR-29b
expression was also observed.

Materials and methods

CLL samples

For the resequencing analysis, 105 DNA samples of 98 high-risk CLL
patients (seven patients were analyzed repeatedly; Table I) were investigated.

Table I. Characteristics of the patients

	Resequencing analysis (n = 98; 105 samples)	miR-16-1 analysis; Sanger sequencing (n = 193)	miR-29b-2/29c analysis; Sanger sequencing (n = 213)
Rai stage (at the time of sample collection)			
0–2 (low/intermediate stage)	72/98 (74%)	135/193 (70%)	153/213 (72%)
3–4 (advanced stage)	17/98 (17%)	46/193 (24%)	48/213 (22%)
NA	9/98 (9%)	12/193 (6%)	12/213 (6%)
Sex ratio (M:F)	60:38	118:75	131:82
Age at diagnosis - years (median)	62.3	60	60
IGHV mutation status			
Unmut	81/98 (83%)	122/193 (63%)	135/213 (63%)
Mut	16/98 (16%)	67/193 (35%)	74/213 (35%)
NA	1/98 (1%)	4/193 (2%)	4/213 (2%)
Cytogenetic aberrations (I-FISH) according to hierarchical cytogenetics (8) (at the time of sample collection)			
Del17p	23/105 (22%)	38/193 (20%)	40/213 (19%)
Del11q	47/105 (45%)	21/193 (11%)	31/213 (15%)
Trisomy 12	3/105 (3%)	29/193 (15%)	30/213 (14%)
Del13q	20/105 (19%)	65/193 (34%)	71/213 (33%)
Normal karyotype	12/105 (11%)	40/193 (21%)	41/213 (19%)
TP53 mutation status (60) (at the time of sample collection)			
del17p + mut*TP53*	19/105 (18%)	32/193 (17%)	34/213 (16%)
Sole mut*TP53*	29/105 (28%)	8/193 (4%)	9/213 (4%)
wt*TP53*	57/105 (54%)	153/193 (79%)	170/213 (80%)

NA, not available.

Additionally, 15 DNA samples of 15 young healthy controls (in their 30s at the time of analysis) were analyzed in order to enable the self-learning algorithm, which produces intensity files in GeneChip Sequence Analysis Software v. 4.1 (GSeq; Affymetrix), to learn as many 'SNP-spots' as possible. Thus, SNP, as the most common type of variation (the presence of which is not age related) can be called more correctly than a novel variation.

Sanger sequencing of miR-16-1 and miR-29b-2/29c was performed in other cohorts of 193 and 213 CLL samples, respectively (Table I). Blood samples were taken in the Department of Internal Medicine, Hematology and Oncology, University Hospital Brno, Czech Republic, with written informed consent in accordance with the Declaration of Helsinki under protocols approved by the Ethical Committee of the University Hospital Brno.

DNA was isolated either from B lymphocytes (purity of CD19+/CD5+ cells >95%) or mononuclear cells (median purity of CD19+/CD5+ cells, 88%) using a DNeasy Blood & Tissue Kit (Qiagen) according to the manufacturer's recommendations. B lymphocytes were separated from peripheral blood using Ficoll-Paque PLUS (GE Healthcare) gradient centrifugation with a depletion of non-B cells (RosetteSep Human B Cell Enrichment Cocktail, RosetteSep Human CD3 Depletion Cocktail; Stemcell Technologies). Mononuclear cells were separated from the peripheral blood using Histopaque (Sigma–Aldrich) gradient centrifugation. The proportion of leukemic cells (CD5+/CD19+) was determined by flow cytometry. DNA extracted from buccal swabs (Quick-gDNAMiniPrep; Zymo Research) was used to detect the germline/somatic origin of sequence variations.

Total RNA was isolated either from B lymphocytes (purity of CD19+/CD5+ cells >95%) or mononuclear cells (median purity of CD19+/CD5+ cells, 85%) from 107 CLL patients' peripheral blood samples (Supplementary Table II, available at *Carcinogenesis* Online) by TriReagent (Molecular Research Center, Inc.) as described previously (11,35). RNA quality was controlled by chip electrophoresis (Bioanalyzer RNA 6000 Nano Assay; Agilent).

Custom microarray resequencing chip design

We designed the resequencing microarray based on a commercially available CustomSeq microarray (format 169, 50 kb; Affymetrix) containing 25mer oligonucleotide probes (36) to detect 1 nt substitutions in miRNAs. For each position of the interrogated sequence, eight 25mer probes are represented on the array: four probes for each strand, each with a different nucleotide in the middle position (A, G, C, T); unrecognized variations were assigned as 'N'.

Reference sequences for miRNAs were downloaded from the UCSC Genome Browser version hg18, which was available at the time of the custom resequencing chip design. The total number of base pairs, representing 109 pre-miRNAs resequenced by microarray, was 13 874. The resequenced parts consisting of the whole pre-miRNAs plus ~20 bp from 5′ and 3′ end of pri-miRNAs were initially analyzed for repeat regions using a RepeatMasker, and any repeats or low complexity regions of >25 bp were excluded. An 814 bp internal control, representing the plasmid control (IQ-EX) provided by Affymetrix, was also tiled on the microarray.

The microarray detection limit was assessed based on the mutation analysis of two plasmids carrying exon 1 of the gene coding for low-density lipoprotein receptor with two types of mutations (i.e. G>T, G>A). The amplicons were hybridized in various proportions of mutated DNA (0, 10, 25, 50, 75, 90 and 100%) on all microarrays.

miRNA resequencing and data analysis

One hundred and nine miRNAs were amplified with 56 primers designed by PerlPrimer v.1.1.17 (37) (Supplementary Table III, available at *Carcinogenesis* Online) using long range PCR (TaKaRa LA Taq™; TaKaRa Bio); the average PCR product size was 4763 bp. The amplicons were quantified with Quant-iT PicoGreen dsDNA Reagent (Invitrogen). Equimolar amplicon amounts from one DNA sample were pooled, the amplicon pool was purified (QIAquick PCR purification Kit; Qiagen), and the DNA concentration was measured (NanoDrop Technologies). The pooled PCR products were fragmented using 0.05 U of fragmentation reagent (Affymetrix) at 37°C/10 min, followed by inactivation at 95°C/15 min. The fragment size was analyzed using chip electrophoresis (Bioanalyzer DNA 1000 Nano Assay; Agilent), and the average size was 50 bp (range 20–200 bp). The pooled and fragmented PCR products were end labeled using a biotin-labeling reagent (Affymetrix) and terminal deoxynucleotidyltransferase (Affymetrix) at 37°C/2 h, followed by inactivation at 95°C/15 min. The labeled amplicons were hybridized to the array (49°C/16 h, 60 r.p.m.). Hybridization was followed by a two-step wash protocol using a FS450 fluidics station (Affymetrix). Finally, the arrays were stained and scanned with the GeneChip 3000 Scanner (Affymetrix).

Intensity files were produced by GeneChip Command Console Software (AGCC; Affymetrix) and processed in GSeq v. 4.1 using version 2 of the resequencing algorithm (38,39). The quality score threshold was set to 3, base reliability threshold was set to 0 (40) and the Modeltype was set to 0 to assess the diploid model, enabling heterozygous calls to be made. Altogether, 120 arrays were analyzed using the mentioned settings. The data were analyzed using Geneious Pro 4.8.2.

Capillary sequencing

The presence of sequence variations detected in miRNAs by microarray was confirmed by Sanger sequencing, which was also used in order to perform miR-16-1 and miR-29b-2/29c mutation analysis. Primers designed by PerlPrimer v.1.1.17 (37) are listed in Supplementary Table IV, available at *Carcinogenesis* Online (cycling conditions are available upon request). Amplicons were sequenced at Macrogen (Seoul, Korea) using an ABI 3730XL DNA Analyzer (Applied Biosystems).

miR-29b-2/29c expression analysis

The effect of sequence variations detected in miR-29c and miR-29b-2 on their expression was evaluated using real-time PCR (TaqMan miRNA Assays; Applied Biosystems) in 107 CLL patients (Supplementary Table II, available at *Carcinogenesis* Online). The obtained miRNA expression levels were normalized to RNU38B, which is uniformly expressed in CLL cells (10). Statistical differences between miRNA levels were evaluated using the non-parametric Mann–Whitney *U*-test (Statistica 6.0; StatSoft).

miRNAs' SNP frequency in CLL patients versus control population

To find out whether allelic SNP frequencies detected in miRNAs in our cohorts of CLL patients differ from the control population, the data from the 1000 genomes project (http://1000genomes.org) were applied since it provides the most comprehensive resource of naturally occurring human variation. At the time of analysis, SNPs were filtered against the April 2012 integrated phase 1 variant release (version 3) of this project (41) containing 38.2M SNPs in total, with phased genotype calls on 1092 samples (525 males, 567 females), of which 379 were Europeans (178 males, 201 females). Since the 1000 genome samples are completely anonymized, sex ratio was the only information available on these samples.

Statistical differences regarding the presence of SNPs between CLL and the control population were evaluated using Fisher's exact test (Statistica 6.0; StatSoft).

miRNAs secondary structures prediction

The RNAfold web tool (http://rna.tbi.univie.ac.at) (42) was used to predict the most stable secondary structures of wild-type (wt) miRNAs and variant sequences. The analyzed sequences included pre-miRNAs and 50 bp upstream and 50 bp downstream flanking sequences at each end of the precursors in case variations were detected by the resequencing microarray. Regarding miR-29b-2/29c, in which the presence of variations was analyzed in a larger area of their primary regions, pre-miRNA regions and 480 bp upstream and 242 bp downstream flanking sequences (miR-29b-2) or 180 bp upstream and 181 bp downstream (miR-29c) at each end of the precursor were studied.

The effects of the novel pre-miR-16-1 variation on its expression

A 760 bp genomic fragment encoding both miR-15a and miR-16-1 was ligated into a pCMV-MIR expression vector (Origene). One construct contained the wt sequence and one contained the novel pre-miR-16-1 variation (83G>C). Both constructs were sequenced to confirm the presence of novel variation. An empty expression vector was used as a negative control. Green fluorescent protein included in the pCMV-MIR expression vector was used as a control to standardize the transfection efficiency.

The above-described constructs were transfected in HEK-293 cells, which have relatively low endogenous expression of miR-15a/16-1 cluster (9), using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. Transfections were performed in quadruplicates. Bright green fluorescent protein-positive cells were sorted 48 h after transfections with the purity of 90–95%.

The expression of wt and mut constructs was evaluated using real-time PCR. miR-16-1 expression levels were normalized to RNU38B. Statistical differences between the constructs were evaluated using the paired *t*-test (Statistica 6.0; StatSoft).

The possible molecular effect of the novel variation on miR-16-1 expression was further analyzed *via* the expression of *BCL2* by real-time PCR. *BCL2* expression levels were normalized to glyceraldehyde-3-phosphate dehydrogenase. Statistical analyses between the constructs were performed with the paired *t*-test (Statistica 6.0; StatSoft).

Results

The resequencing microarray is a convenient tool for sensitive and reliable detection of common SNPs in miRNAs

The array was designed to detect single-nucleotide substitutions in miRNAs. In total, 120 arrays were used to perform mutation analysis of 109 miRNA genes. The average nucleotide call rate (43) was 96.1% (range 81.3–99.4%), demonstrating a good quality of hybridization and overall chip design. Based on a dilution experiment concerning a mutated low-density lipoprotein gene, the custom resequencing microarray detection limit range was 10–25% of mutated amplicons depending on the mutation type.

In 96 CLL patients, 18 SNPs were detected in 15 miRNAs in total (http://1000genomes.org; dbSNP Build 134). Confirmatory Sanger sequencing was done on 93 CLL patients, covering 403 polymorphic sites (i.e. 227 major alleles, 176 minor alleles). Out of 277 homozygous positions, 251 were correctly called by resequencing microarray (90.6%), and 26 were assigned as 'N' (i.e. unrecognized). In the case of 126 heterozygous polymorphic sites, 110 (87.3%) were correctly called by resequencing microarray, 15 were recognized as 'N', whereas 1 was incorrectly called as homozygous by the microarray.

Altogether, the custom-designed array demonstrated a high call rate (>96%), detection limit range 10–25% in the case of mutated low-density lipoprotein genes and small false negativity ~11% for miR-SNP analysis.

Novel sequence variations detected in miRNAs by resequencing microarray and Sanger sequencing

In five miRNAs, the resequencing microarray detected five novel heterozygous 1 nt variations, the presence of which was confirmed by Sanger sequencing, in five CLL patients (Table II, Figure 1). These variations are considered to be novel variations since none was found either in the control population from the 1000 genomes project or the NCBI dbSNP database (build 134). One variation was found in pri-miR-29a, three variations were detected in pre-miRNA region (pre-miR-16-1, pre-miR-372, pre-miR-106b) and one variation was found in the mature miR-142-3p.

Importantly, the resequencing microarray reported 237 additional 1 nt variations in 77 miRNAs. Sanger sequencing neither confirmed the presence of any of these variations, which were thus recognized as false positives, nor found any additional variation. This observation contrasts the relatively high agreement level of correctly called nucleotides in SNP sites (see above). Such a discrepancy can be explained by the fact that the intensity files are produced in GSeq v. 4.1 by a self-learning algorithm and SNP (as the most common type of variation) can be called more correctly than a novel variation.

Among the confirmed sequence variations, miR-16-1 was further selected for consecutive Sanger sequencing since it was the first miRNA described to harbor germline mutation (i.e. +7C>T in pri-miR-16-1) in two CLL patients with 13q deletion (9). In our cohort of 98 CLL patients, the resequencing microarray detected one novel germline variation (83G>C) (see below) in pre-miR-16-1 in one CLL patient (Table II); however, the known mutation (+7C>T) was not found in any CLL patients. We further investigated whether the miR-16-1 sequence variations are as frequent as have been published (9). Therefore, pri-miR-16-1 was analyzed in another 193 CLL patients selected with respect to del13q status (105 patients with del13q). Surprisingly, neither known germline mutation in pri-miR-16-1 nor novel germline variation in pre-miR-16-1 was found in any of these 193 CLL patients. This observation demonstrates that both miR-16-1 sequence variations are extremely rare (<0.5%) in CLL, which contrasts previously published data (9).

The practical absence of sequence variations in mature miRNAs (except for one case with a variation in miR-142-3p) prompted us to check for variations in regions flanking the mature miRNAs. This was performed in 213 CLL patients (20 samples overlapped with those from the resequencing analysis) by sequencing larger genomic regions containing pri-miR-29b-2 and pri-miR-29c. In total, 562 and 129 nts were analyzed in 5′ end of pri-miR-29b-2 and pri-miR-29c, respectively; 298 and 239 nts were sequenced in 3′ end of pri-miR-29b-2 and pri-miR-29c, respectively. Two SNPs were found in pri-miR-29c, seven SNPs (Table III) and one novel variation (Table II) were detected in pri-miR-29b-2; but surprisingly, no variation was found in either pre- or mature miRNA.

Sequence variations are not frequent in the mature miRNA region

In total, 33 variations, i.e. 27 SNPs (18 SNPs in resequencing analysis and 9 SNPs in miR-29b-2/29c) and 6 novel variations, were detected in 22 miRNAs (20 miRNAs in resequencing analysis, miR-29b-2, miR-29c) in our cohorts of CLL patients; 17 variations were found in pri-miRNAs, 12 in pre-miRNAs and 4 in mature miRNAs (Supplementary Table V, available at *Carcinogenesis* Online). Analyzing the frequency of variations per miRNA region (pri-miRNA versus pre-miRNA versus mature miRNA) revealed that the variations were least frequent (P < 0.05) in a mature miRNA region (i.e. 0.9 variation per 1000 nt), and most frequent in a pri-miRNA region (i.e. 3.1 variations per 1000 nt) (Supplementary Table V, available at *Carcinogenesis* Online).

Table II. Novel 1 nt sequence variations detected in miRNAs

miRNA	Location[a]	Chromosomal variant (hg19)	Reference allele	Alternative allele	CLL patients (n = 98/213)	Origin	Control subjects (n = 1092)	Comment
mat-miR-142-3p	62C>T	NC_000017.10:g.56408618G>A	G	A	1/98	NA	0/1092	No cancer history
pre-miR-16-1	83G>C	NC_000013.10:g.50623115C>G	C	G	1/98	Germline	0/1092	Father - bladder cancer
pre-miR-372	13T>C	NC_000019.9:g.54291156T>C	T	C	1/98	Germline	0/1092	No cancer history
pre-miR-106b	45C>T	NC_000007.13:g.99651653G>A	G	A	1/98	NA	0/1092	NA
pri-miR-29a	+22T>C	NC_000007.13:g.130561484A>G	A	G	1/98	NA	0/1092	Father - colon cancer
pri-miR-29b-2	−256A>G	NC_000001.10:g.207976124T>C	T	C	2/213	1 Germline, 1 NA	0/1092	The patient with germline variation - uncle and his son had CLL

[a]Novel variations localized upstream (negative base count, '−') or downstream (positive base count, '+') of the precursor hairpin. NA, not available.

Germline versus somatic origin of variations detected in miRNAs

Using Sanger sequencing of DNA isolated from buccal swabs, the novel variations in pre-miR-16-1, pre-miR-372, and the variation in pri-miR-29b-2 in one patient were proven to be germline. The germline status of the remaining variations (i.e. pri-miR-29b-2 in the second patient, pri-miR-29a, pre-miR-106b and miR-142-3p) is unknown since DNA from buccal swabs was not available for these patients. Notably, the pre-miR-16-1 variation the patient was harboring was also found in two of his offspring (a son and a daughter, 35 and 36 years old, respectively, neither of whom has been diagnosed with CLL/cancer).

Hovever, in the pri-miR-655, we have identified one somatic variation which was present only in CLL cells. We initially considered this a novel variation, although according to dbSNP 134, this is a common SNP (rs139043404) with an unknown allelic frequency (Table III). The patient harboring this SNP had two detectable CLL populations originating from two B-lymphocyte clones (biclonal case) with ~30% of the CLL cells expressing Igκ light chain and ~70% of cells expressing Igλ light chain. Using fluorescence-activated cell sorting, the cells were sorted according to Igλ and Igκ expression, and the pri-miR-655 was sequenced separately in each population. Surprisingly, the miR-655 variation was present only in the subclone of CLL cells expressing Igλ chain. Additionally, by using multiplex ligation-dependent probe amplification, it was found that this variation was accompanied by 11q and 13q deletions. The other subclone, expressing wt pri-miR-655 and Igκ chain, harbored *TP53* mutation and no common chromosomal abnormality (Supplementary Figure 1, available at *Carcinogenesis* Online).

Several SNPs present in miRNAs occur with higher allelic frequency in CLL patients

The resequencing microarray detected 18 SNPs in 15 miRNAs (Table III). Within the miR-29 family, seven SNPs were detected in pri-miR-29b-2, and two SNPs were found in pri-miR-29c (Table III) using Sanger sequencing.

To find out whether there is a difference in allelic frequencies in SNPs detected in miRNAs between CLL patients in our study and a control population, we used data from the 1000 genomes project. The allelic frequency for all samples (n = 1092 samples) from the 1000 genomes project was available for 15 out of 18 SNPs detected by the resequencing microarray, and for all SNPs detected in the miR-29 family (Table III). The allelic frequency was also compared separately between the European population from the 1000 genomes project (n = 379 samples) and CLL patients since they were of European origin. In this case, the allelic frequency for the European population from the 1000 genomes project was known for 13/18 SNPs detected by microarray, and for all SNPs detected in the miR-29 family (Table III).

Altogether, the allelic frequency differed significantly in eight SNPs (P < 0.05) when analyzing SNPs frequency irrespective of population origin (Figure 2). These SNPs are present in two mature miRNAs (miR-412, miR-146a*), pre-miR-656 and three pri-miRNAs (pri-miR-100, pri-miR-154, pri-miR-29b-2). Three SNPs detected in pri-miR-29b-2 (rs141961287: +107+A, rs12401619: −408C>G, rs12410786: −337A>T) were statistically significantly more frequent in analyzed CLL patients than in a control population. However, the comparison between CLL patients and the separate European population indicated that only one SNP in pre-miR-323b was statistically significantly more frequent in our cohort of CLL patients (Figure 2).

Insertion (+107+A) reduces miR-29b expression in CLL patients with unmut IGHV

Since miR-29b/29c are known to be downregulated in aggressive CLL subtypes (9,10) and the particular mechanism responsible for their aberrant expression is not well characterized, we studied the effect of variations detected in miR-29b-2/29c on their expression in 107 CLL patients (Supplementary Table II, available at *Carcinogenesis* Online).
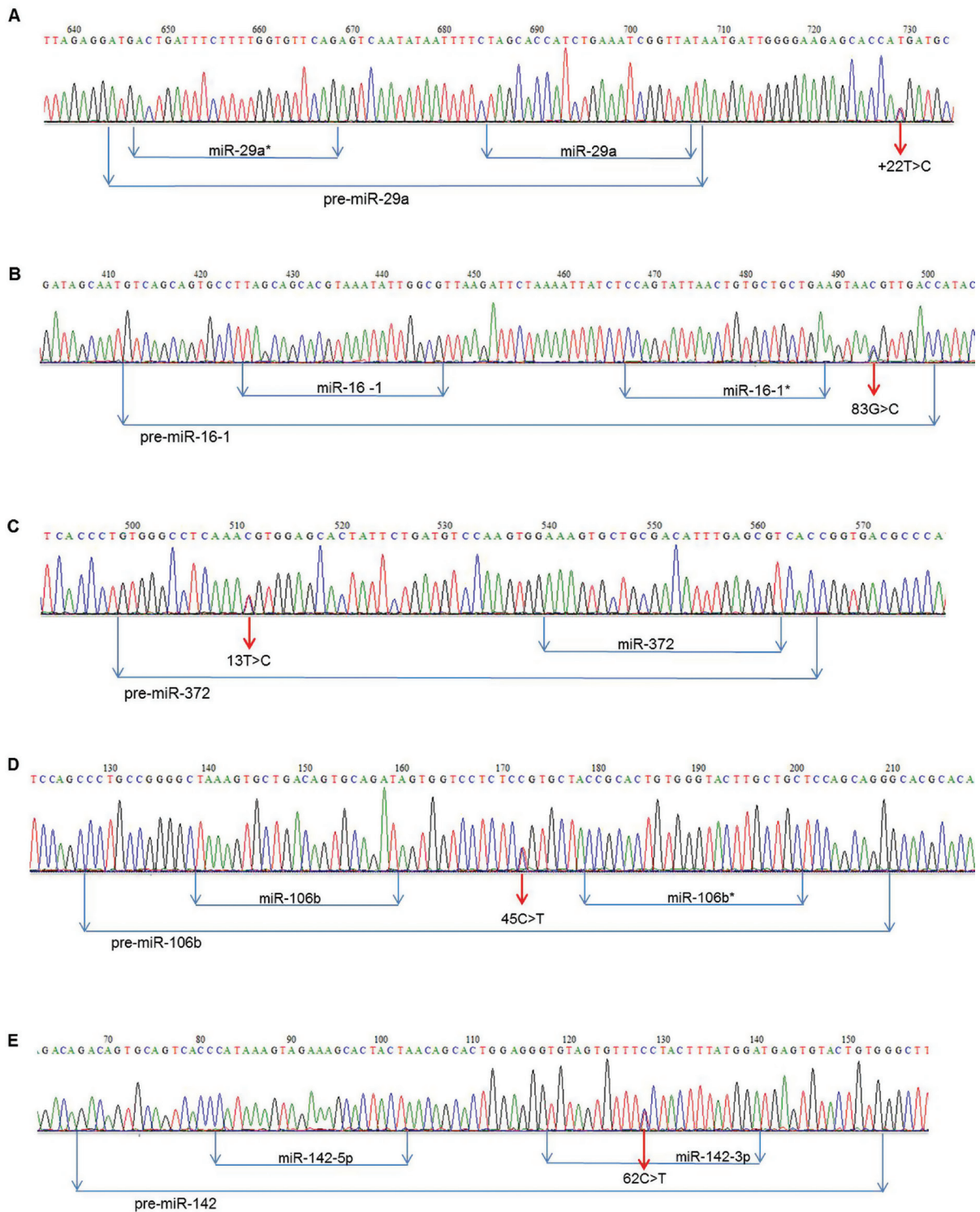
Fig. 1. Novel 1 nt sequence variations detected in miRNAs. Five novel heterozygous variations, localized in various miRNA regions, were detected by resequencing microarray and confirmed by Sanger sequencing in five different CLL patients: (**A**) pri-miR-29a, (**B**) pre-miR-16-1, (**C**) pre-miR-372, (**D**) pre-miR-106b, (**E**) miR-142-3p. Variations localized downstream of the precursor hairpin, which contains the mature miRNA, are assigned '+'.

The expression of both miR-29b (Figure 3A) and miR-29c (Supplementary Figure 2A, available at *Carcinogenesis* Online) was lower in the patients with unmut *IGHV* ($P = 0.01$ and $P < 0.0001$, respectively) and in the patients harboring *TP53* aberration ($P = 0.04$ and $P = 0.052$, respectively), which is in agreement with previously published data (9,10). The expression of miR-29b also tended to be lower in patients with shorter overall survival and time to first

treatment. However, due to the size of the patient cohort, statistical significance was not reached (data not shown).

Because of the apparent correlation between *IGHV* status and miR-29 expression, the cohort was separated in two groups based on *IGHV* status (48 mut *IGHV*, 52 unmut *IGHV*), and patients with the *TP53* mutation/deletion were excluded ($n = 6$) from further analysis. This revealed the effect of variations in miR-29b-2 (Figure 3B

Table III. SNPs detected in miRNAs in CLL patients

miRNA	Chromosome (hg19)	Genomic coordinates (hg19)	Reference allele	Alternative allele	dbSNP ID	SNP–MAF CLL samples (n = 196; 426 alleles)	SNP–MAF 1000 genomes samples (n = 2184 alleles)	SNP–MAF European samples (n = 758 alleles)	P value[a]	P value[b]
pri-miR-29c	chr1	207975060	A	T	rs147139948	0.007	0.003	0.005	ns	ns
pri-miR-29c	chr1	207975315	C	T	rs150749580	0.0047	0.005	0.011	ns	ns
pri-miR-29b-2	chr1	207975968	C	T	rs114790693	0.0141	0.005	0.013	ns	ns
pri-miR-29b-2	chr1	207976276	G	C	rs12401619	0.3685	0.285	0.404	**0.0007**	ns
pri-miR-29b-2	chr1	207976205	T	A	rs12410786	0.3099	0.258	0.338	**0.0320**	ns
pri-miR-29b-2	chr1	207975681	—	T	rs141961287	0.0587	0.027	0.066	**0.0009**	ns
pri-miR-29b-2	chr1	207976037	G	C	rs145834945	0.0141	0.007	0.016	ns	ns
pri-miR-29b-2	chr1	207975949	T	C	rs56075814	0.1972	0.163	0.219	ns	ns
pri-miR-29b-2	chr1	207975905	C	T	rs78876157	0.0047	0.005	0.011	ns	ns
pri-miR-181b-1	chr1	198827998	C	G	rs113162007	0.0051	NA	NA	NA	NA
pri-miR-100	chr11	122022927	C	T	rs543412	0.2602	0.337	0.305	**0.0328**	ns
pri-miR-26a-2	chr12	58218382	G	A	rs41292017	0.0051	0.011	0.009	ns	ns
pri-miR-655	chr14	101515880	G	A	rs139043404	0.0051	0.001	NA	ns	NA
pri-miR-409_miR-412	chr14	101531733	C	T	rs61992670	0.0255	0.023	0.045	ns	ns
mat-miR-412	chr14	101531854	A	G	rs61992671	0.4388	0.234	0.491	**<0.0001**	ns
pri-miR-154	chr14	101526181	G	A	rs41286572	0.0918	0.050	0.088	**0.0195**	ns
pre-miR-656	chr14	101533093	C	T	rs58834075	0.0204	0.074	0.028	**0.0076**	ns
pre-miR-431	chr14	101347355	C	T	rs76090066	0.0051	0.002	0.004	ns	ns
pre-miR-323b	chr14	101522556	T	C	rs56103835	0.2653	0.304	0.183	ns	**0.0123**
mat-miR-323b	chr14	101522589	C	T	rs75330474	0.0051	0.030	0.001	ns	ns
pre-miR-27a	chr19	13947296	G	A	rs11671784	0.0051	0.011	0.024	ns	ns
pre-miR-27a	chr19	13947292	T	C	rs895819	0.3265	0.358	0.328	ns	ns
pre-miR-15b	chr3	160122421	A	G	rs146020563	0.0051	0.0005	NA	ns	NA
mat-miR-146a*	chr5	159912418	C	G	rs2910164	0.199	0.381	0.218	**<0.0001**	ns
pre-miR-96	chr7	129414568	G	A	rs73159662	0.0051	NA	NA	NA	NA
pre-miR-182	chr7	129410227	C	T	rs76481776	0.051	0.049	0.084	ns	ns
pre-miR-223	chrX	65238733	G	A	rs186354597	0.0051	NA	NA	NA	NA

Minor allele frequency (MAF) of detected SNPs. NA, not available; ns, not significant.
[a]P value for all samples from the 1000 genomes project versus CLL patients, [b]P value for European samples from the 1000 genomes project versus CLL patients.
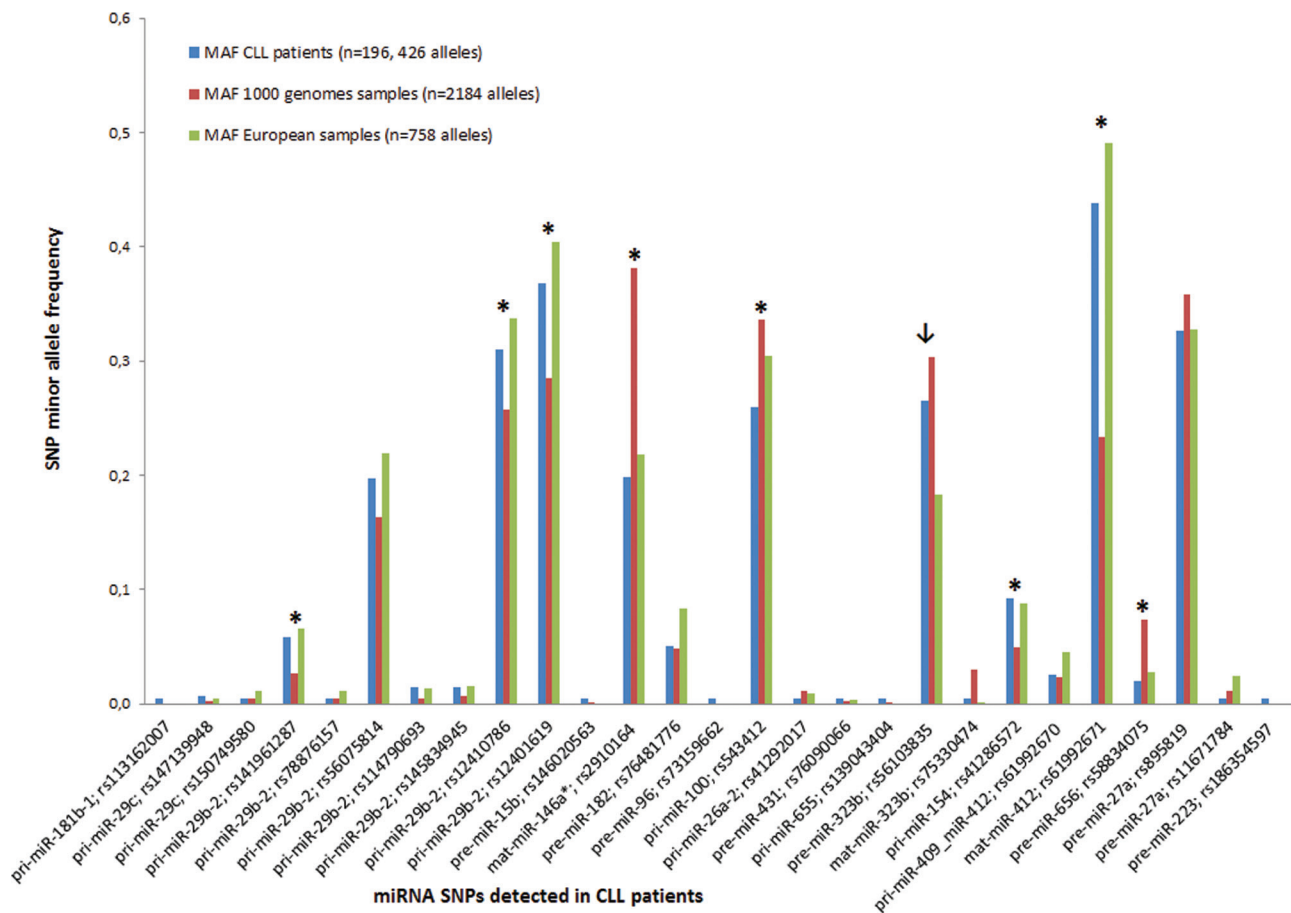
Fig. 2. Frequency of minor alleles in CLL patients versus control population from the 1000 genomes project. The allelic frequencies of SNPs detected in miRNAs in CLL patients differed significantly ($P < 0.05$) between CLL patients and all 1000 genomes project samples in eight SNPs marked with asterisk, and in one SNP (marked with arrow) when a separate European population was analyzed. MAF represents minor allele frequency of detected SNPs. Blue columns represent SNP–MAF of CLL patients, red columns represent SNP–MAF of all samples from the 1000 genomes project and green columns represent SNP–MAF of European samples from the 1000 genomes project.

and C) and miR-29c (Supplementary Figure 2B and C, available at *Carcinogenesis* Online) on their expression with respect to the *IGHV* mutation status. Importantly, the expression of miR-29b-2 harboring insertion (rs141961287: +107+A) was lowered (fold change 0.7) in the patients with unmut *IGHV* ($P = 0.036$; Figure 3B) but not with mut *IGHV* (Figure 3C).

Sequence variations detected in miRNAs alter their secondary structures

It is expected that processing and maturation of a miRNA precursor require appropriate secondary structures and specific sequence elements within pre- or pri-miRNA (21,44–46). Hence, we compared the minimum free energy (dG) for optimal secondary structures of both wt miRNAs and for their variations detected in our study (i.e. novel variations, somatic SNP and SNPs reaching significantly different frequency between CLL and the control population; Supplementary Table VI, available at *Carcinogenesis* Online).

Most of the analyzed variations ($n = 22$) affected miR-secondary structure. In particular, the novel variation in pre-miR-16-1 had the most dramatic effect on its secondary structure, whereas the known germline mutation in pri-miR-16-1 (9) had none (Figure 4A). Therefore, the novel variation is likely to have a larger impact on miR-16-1 maturation/expression.

miR-secondary structures were also changed by the novel variations detected in miR-142-3p, pre-miR-372 and pre-miR-106b, by SNPs, the allelic frequency of which differed between the 1000 genomes project samples and CLL patients (miR-412, miR-146a,

pre-miR-656), and by the pre-miR-323b SNP, which was more frequent in CLL patients than in the European population. A comparison of wt miRNAs and their altered secondary structures is shown in the Supplementary Figure 3A–G, available at *Carcinogenesis* Online.

miR-29b-2 secondary structure was affected by the polymorphic insertion (rs141961287: +107+A) (Figure 4B), and by two remaining SNPs (rs12410786: −337A>T; rs12401619: −408C>G; Supplementary Figure 3H, available at *Carcinogenesis* Online), all of which were more frequent in the analyzed cohort of CLL patients when compared with 1092 samples from the 1000 genomes project. miR-29c secondary structure was affected by one SNP (rs147139948: +137T>A; Supplementary Figure 3I, available at *Carcinogenesis* Online).

Novel pre-miR-16-1 variation affects the expression of mature miR-16-1

The expression vectors containing either the wt allele or the mutated allele of the miR-15a/16-1 cluster were constructed to identify a possible molecular effect of the novel pre-miR-16-1 variation (83G>C) in HEK-293 cells. An empty vector was used as a negative control.

The transfectants harboring the novel variation expressed miR-16-1 at levels that were significantly lower ($P < 0.01$) than the transfectants harboring the wt allele (Figure 5A). This result was further confirmed by the expression analysis of *BCL2*, which is negatively regulated by miR-16-1 (47). *BCL2* expression was higher ($P = 0.037$) in the transfectants harboring the novel variation than in the transfectants
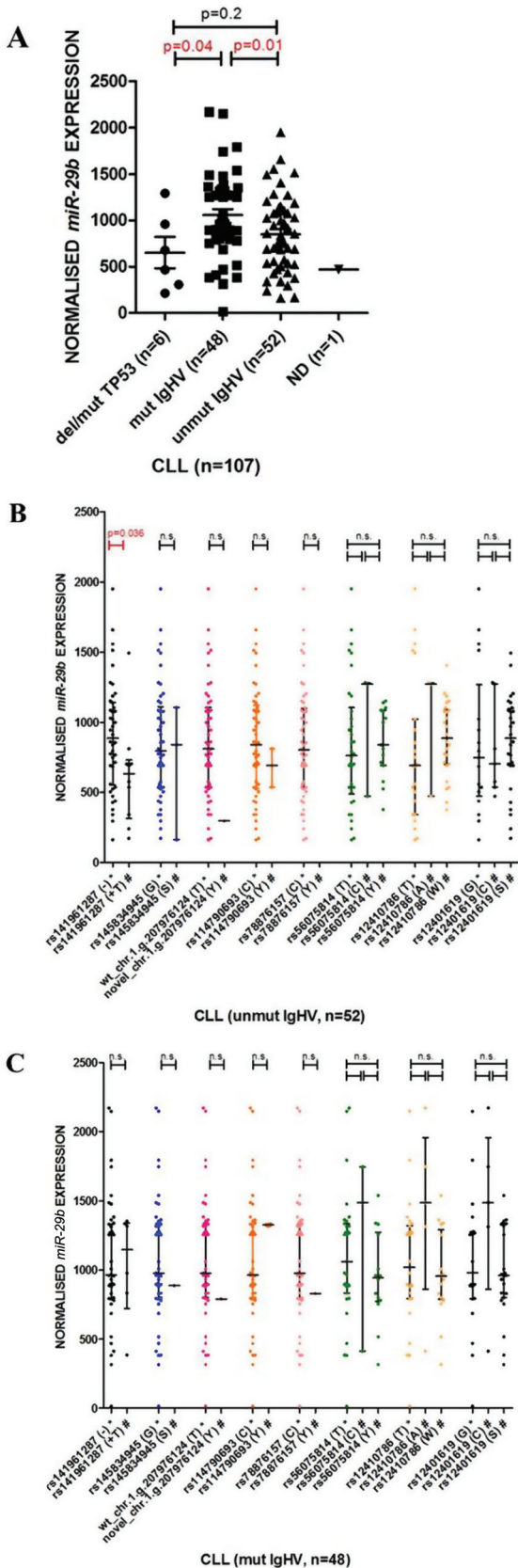
Fig. 3. Expression analysis of miR-29b in 107 CLL patients. (**A**) The expression of miR-29b in CLL patients with respect to *IGHV* mutational status and *TP53* aberrations. (**B** and **C**) The effect of variations detected in pri-miR-29b in separated groups of patients based on *IGHV* status (48 mut *IGHV*, 52 unmut *IGHV*); patients with the *TP53* mutation/deletion

harboring the wt allele (Figure 5B). These results thus indicate that the novel variation (83G>C) affects the mature miR-16-1 expression.

Discussion

We and others have described that miRNAs are abnormally expressed in CLL and have specific expression patterns in CLL subtypes (9–13). However, the mechanism through which miRNA expression is deregulated in cancer and CLL is poorly characterized, and it is not clear whether it is a primary or secondary event caused by deregulation of transcription networks (5), chromatin structure (48,49) or other mechanisms.

It has been repeatedly described that miRNA expression can be influenced by the presence of mutations and SNPs in miRNA genes (9,50–53) and in miRNA seed sequences (3,28,54). In CLL, the overall frequency and impact of miRNA gene sequence variations still need to be understood, although several mutations and SNPs have been described by Sanger sequencing (9,32). Calin *et al.* (9) analyzed 42 miRNAs in 75 CLL patients, and found mutations in 5 miRNAs in 11 patients. Wojcik *et al.* (32) screened sequence variations in 72 miRNAs in a cohort of 39 CLL patients and found both SNPs and novel mutations. Nevertheless, according to the 1000 genomes project data, 11/25 variations that were originally described as 'novel variations' by Calin *et al.* (9) and Wojcik *et al.* (32), are now known to be common SNPs. Thus, we created an updated catalog of sequence variations (novel variations and SNPs) that have been described so far in CLL patients. A detailed description (SNP identifiers, genomic coordinates) of all the variations is provided in Supplementary Table VII, available at *Carcinogenesis* Online.

In this present study, we screened sequence variations not only in more miRNA genes (*n* = 109; selected according to their relevance to CLL pathogenesis or other hematological malignancies), but also in a larger cohort of high-risk CLL patients (*n* = 98). We have demonstrated the feasibility of a custom design microarray platform for screening miRNA variations with low false negativity. However, due to the high false positivity observed in our study, a confirmatory method is necessary.

Altogether, 27 SNPs were detected in our study in 17 miRNAs (Table III); one SNP per one miRNA; except for miR-29c (two SNPs), miR-29b-2 (seven SNPs), miR-323b (two SNPs), miR-412 (two SNPs), miR-27a (two SNPs). Our observation that most of these SNPs were present outside the mature miRNA regions is in agreement with data published by Saunders *et al.* (28). They recorded SNP density in a pri-miRNA region at 3 SNPs per kb and only 1.3 SNPs per kb in a pre-miRNA region, which indicates a strong selective constraint on human pre-miRNAs (28). In total, we detected three SNPs (Table III) and one novel variation (Table II) in four mature miRNAs. Our observation that three of these variations (except for the SNP in miR-146a*) were located outside the seed regions reflects its requirement for target recognition.

The presence of SNPs in miRNAs may be an important source of phenotypic variation and contribute to the susceptibility for complex disorders (55). The 1000 genomes project was used to compare the allelic frequency differences in detected SNPs since it provides the most comprehensive resource of naturally occurring human variation. In our study, only one SNP in pre-miR-323b was notably more frequent in CLL patients (*P* = 0.0123; Figure 2) when its allelic frequency was compared with CLL patients and a separate European population (*n* = 379). Interestingly, when a control population was enlarged by all samples (*n* = 1092) from the 1000 genomes project, the allelic frequency differed significantly for eight other SNPs (*P* < 0.05; Figure 2). This included the SNP located in miR-146a (rs2910164), which is known to alter processing and lower expression

---

(*n* = 6) were excluded (4 patients unmut *IGHV*, 2 patients mut *IGHV*). Not determined (ND): *IGHV* status is not known; *Indicates major allele of the particular SNP or wt pri-miR-29b-2; #Indicates minor allele of the particular SNP or novel variation detected in pri-miR-29b-2.
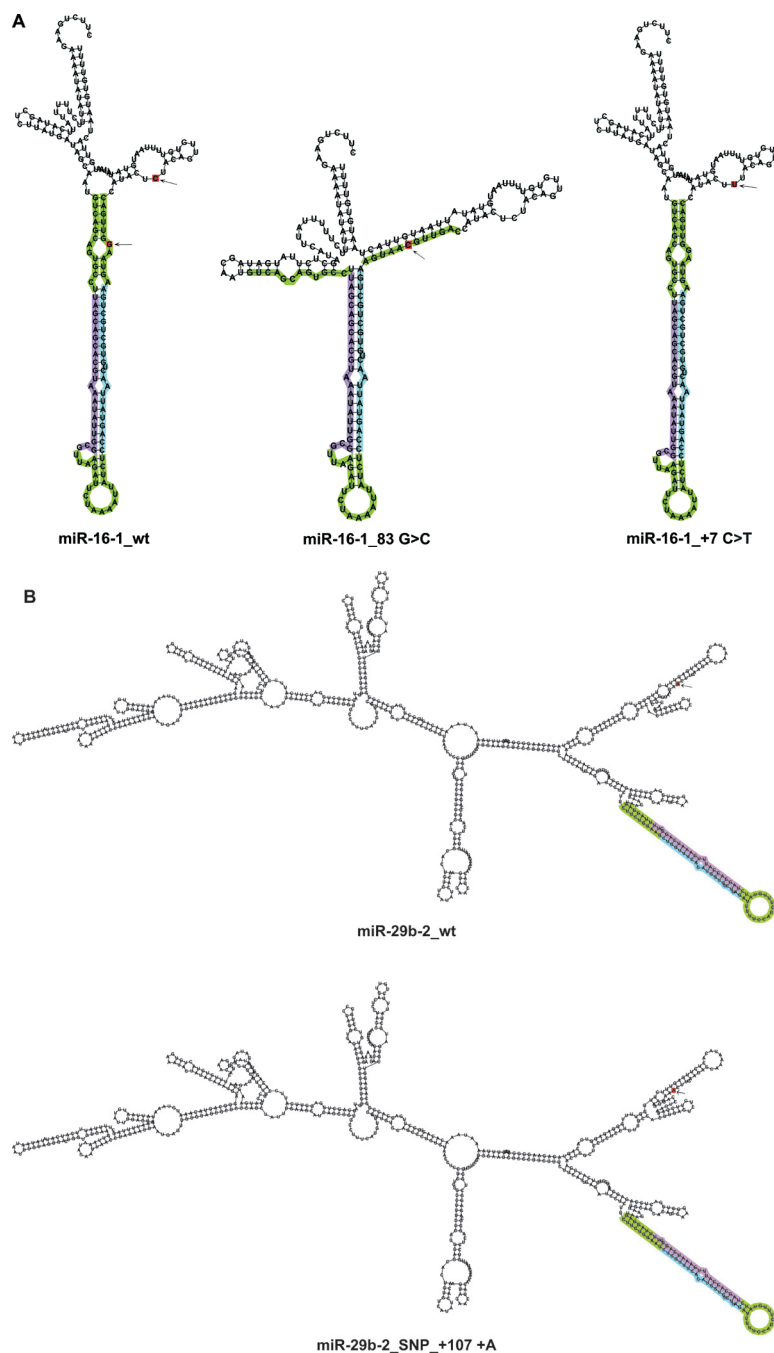
Fig. 4. Secondary structures of wt miRNAs and miRNAs harboring sequence variations (novel variation/SNP) predicted by the RNAfold tool. (**A**) The novel variation found in pre-miR-16-1 (83G>C) had a dramatic effect on its secondary structure, whereas the known mutation present in its primary region (+7C>T) (9) had none. (**B**) pri-miR-29b-2 secondary structure was affected by polymorphic insertion (rs141961287: +107+A). Depicted are the most stable secondary structures with the lowest free energy as predicted by the RNAfold tool. Variations are designated in red and indicated by arrows. Mature miRNAs and miR-5p are designated in violet; mature miRNAs* and miR-3p are designated in blue. Pre-miRNA regions are designated in green; pri-miRNA regions are colorless.

of mature miRNA, and predispose a person to various types of cancer (51,52,56). Although our data show that the frequency of some SNPs present in miRNAs differ significantly between CLL patients and the population represented by individuals from the 1000 genomes project, larger population studies of different ethnic cohorts are necessary to verify their association and possible impact on CLL biology/pathogenesis.

Except for SNPs, in six miRNAs we have also detected six novel 1 nt variations, all of which were present in miRNAs associated either with CLL pathogenesis/prognosis, i.e. miR-16-1 (6,9), miR-29a (9,16), miR-29b-2 (9,16) or implicated in the biology of other tumors,

i.e. miR-142 (57), miR-106b (58) and miR-372 (59). Importantly, the variation in miR-142 is the first novel variation detected in the mature miRNA in CLL, since all the variations described by Calin *et al.* (9) and Wojcik *et al.* (32) were found outside mature miRNA regions.

In particular, we have also found a novel variation in the miR-16-1, which is therefore the second variation of miR-16 described in literature (9). The novel miR-16-1 variation highly affected its secondary structure (Figure 4A) unlike the known germline mutation (+7C>T) (9). It was suggested that the (+7C>T) mutation has a relatively high frequency in CLL (2.7%). However, we presume that miR-16-1 variations observed in Calin's and our study are very rare in CLL (<0.5%)
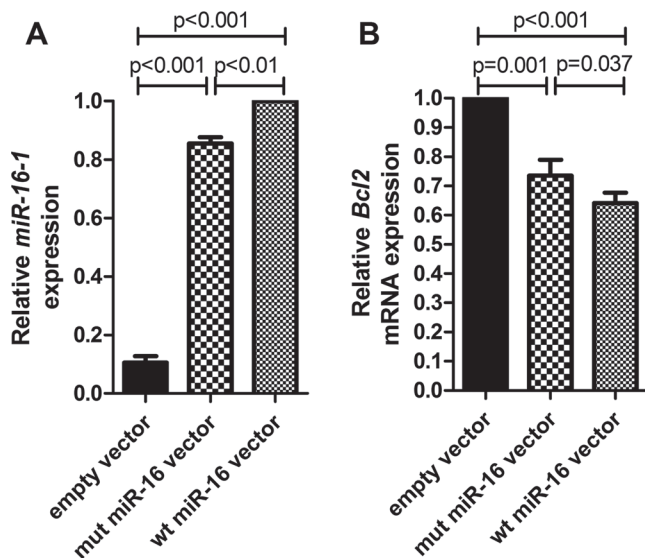
Fig. 5. The effects of the novel pre-miR-16-1 variation on its expression. (**A**) The effect of the novel pre-miR-16-1 variation on the mature miR-16-1 expression. The expression of miR-16 in the cells transfected with the wt vector was set as 1. (**B**) The effect of the novel pre-miR-16-1 variation on the expression of *BCL2*. The expression of *BCL2* messenger RNA (mRNA) in the cells transfected with an empty vector (control) was set as 1. Bars represent mean values; error bars represent standard error of the mean values.

since neither was found in our consecutive analysis of 193 CLL patients.

Surprisingly, our study indicates that miRNA variations are generally rare in CLL samples. The observation that all of the analyzed variations were of germline origin, except for the SNP in pri-miR-655, suggests that although most miRNA variations are not somatic, rare CLL cases exist where the variation is selected only in leukemic cells.

miRNA processing and maturation are highly regulated steps requiring proper secondary structures within pri- or pre-miRNA to be recognized by miR-regulatory machinery (44–46). The sequence variations altering miRNA's secondary structure may thus affect its mature form expression (21). The observation that most of the variations detected in our study altered particular miRNA's secondary structure and free energy values (Supplementary Table VI, available at *Carcinogenesis* Online) implies that variations present in either part of the miRNA region may alter its maturation/expression. This was especially prominent in the novel pre-miR-16-1 variation (Figure 4A). Due to the low frequency of variations in our cohort, we were not able to directly study their effect on miRNA expression. Nevertheless, the possible molecular effect on miR expression was studied in the case of the novel pre-miR-16-1 variation. Our results show that miR-16-1 expression levels were lower ($P < 0.01$; Figure 5A) and *BCL2* expression levels were higher ($P = 0.037$; Figure 5B) in the transfectants harboring mutated allele than in the transfectants harboring wt allele. The variation (83G>C) thus represents a novel miR-16-1 variation, which affects both its secondary structure (Figure 4A) and expression (Figure 5A and B).

The effect of variations on miRNA expression was also analyzed in more detail in miR-29b-2/29c, known to be downregulated in aggressive CLL subtypes (9,10). Among the SNPs detected in miR-29b-2 in our study, three (i.e. rs141961287: +107+A, rs12401619: −408C>G, rs12410786: −337A>T) were not only significantly more frequent in analyzed CLL patients when compared with all samples from the 1000 genomes project (Table III), they also changed the miR-29b-2 secondary structure (Supplementary Table VI, available at *Carcinogenesis* Online). Interestingly, the later two SNPs were found to be present in genetic linkage in most of the CLL patients analyzed (data not shown). Polymorphic insertion +107+A in pri-miR-29b-2 (rs141961287) was described as lowering miR-29b expression when

compared with normal B cells (9). In our cohort of CLL patients, miR-29b expression was lower in the patients with unmut *IGHV* and harboring polymorphic insertion (Figure 3B), which suggests different regulatory mechanisms for miR-29b compared with the patients without this variation.

To conclude, we herein confirm that both SNPs and novel variations, most of which were present outside mature miRNAs and also changed the secondary structure of a particular miRNA, are present in miRNA genes in CLL patients. Among them, we identified a novel variation in miR-16-1 that affected its expression. In addition, our data show that miR-16-1 variations are extremely rare (<0.5%) in CLL, which contrasts previously published data (9). Significantly, certain miR-SNPs, including the polymorphic insertion in pri-miR-29b-2, are more frequent in CLL patients than in a control population. The insertion lowers miR-29b expression in CLL patients with unmut *IGHV* and may therefore affect the biology of CLL B cells. Altogether, this suggests that sequence variations may be related to CLL biology and/or pathogenesis.

Supplementary material

Supplementary Tables I–VII and Figures 1–3 can be found at http://carcin.oxfordjournals.org/

References

1. Lewis,B.P. *et al.* (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **120**, 15–20.
2. Lim,L.P. *et al.* (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, **433**, 769–773.
3. Brennecke,J. *et al.* (2005) Principles of microRNA-target recognition. *PLoS Biol.*, **3**, e85.
4. Krek,A. *et al.* (2005) Combinatorial microRNA target predictions. *Nat. Genet.*, **37**, 495–500.
5. Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.
6. Calin,G.A. *et al.* (2002) Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl Acad. Sci. USA*, **99**, 15524–15529.
7. Klein,U. *et al.* (2010) The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. *Cancer Cell*, **17**, 28–40.
8. Döhner,H. *et al.* (2000) Genomic aberrations and survival in chronic lymphocytic leukemia. *N. Engl. J. Med.*, **343**, 1910–1916.
9. Calin,G.A. *et al.* (2005) A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N. Engl. J. Med.*, **353**, 1793–1801.

10. Mraz,M. *et al.* (2009) miR-34a, miR-29c and miR-17-5p are down-regulated in CLL patients with TP53 abnormalities. *Leukemia*, **23**, 1159–1163.

11. Mraz,M. *et al.* (2012) MicroRNA-650 expression is influenced by immunoglobulin gene rearrangement and affects the biology of chronic lymphocytic leukemia. *Blood*, **119**, 2110–2113.

12. Mraz,M. *et al.* (2012) MicroRNAs in chronic lymphocytic leukemia: from causality to associations and back. *Expert Rev. Hematol.*, **5**, 579–581.

13. Mraz,M. *et al.* (2013) MicroRNAs and B cell receptor signaling in chronic lymphocytic leukemia. *Leuk. Lymphoma*, **54**, 1836–1839.

14. Zenz,T. *et al.* (2009) miR-34a as part of the resistance network in chronic lymphocytic leukemia. *Blood*, **113**, 3801–3808.

15. Asslaber,D. *et al.* (2010) MicroRNA-34a expression correlates with MDM2 SNP309 polymorphism and treatment-free survival in chronic lymphocytic leukemia. *Blood*, **115**, 4191–4197.

16. Pekarsky,Y. *et al.* (2006) Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. *Cancer Res.*, **66**, 11590–11593.

17. Lu,J. *et al.* (2005) MicroRNA expression profiles classify human cancers. *Nature*, **435**, 834–838.

18. Takamizawa,J. *et al.* (2004) Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res.*, **64**, 3753–3756.

19. Calin,G.A. *et al.* (2006) MicroRNA signatures in human cancers. *Nat. Rev. Cancer*, **6**, 857–866.

20. Volinia,S. *et al.* (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl Acad. Sci. USA*, **103**, 2257–2261.

21. Duan,R. *et al.* (2007) Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. *Hum. Mol. Genet.*, **16**, 1124–1131.

22. Denli,A.M. *et al.* (2004) Processing of primary microRNAs by the microprocessor complex. *Nature*, **432**, 231–235.

23. Han,J. *et al.* (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.*, **18**, 3016–3027.

24. Grishok,A. *et al.* (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell*, **106**, 23–34.

25. Hutvágner,G. *et al.* (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science*, **293**, 834–838.

26. Doench,J.G. *et al.* (2004) Specificity of microRNA target selection in translational repression. *Genes Dev.*, **18**, 504–511.

27. Iwai,N. *et al.* (2005) Polymorphisms in human pre-miRNAs. *Biochem. Biophys. Res. Commun.*, **331**, 1439–1444.

28. Saunders,M.A. *et al.* (2007) Human polymorphism at microRNAs and microRNA target sites. *Proc. Natl Acad. Sci. USA*, **104**, 3300–3305.

29. Landi,D. *et al.* (2008) Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer. *Carcinogenesis*, **29**, 579–584.

30. He,H. *et al.* (2005) The role of microRNA genes in papillary thyroid carcinoma. *Proc. Natl Acad. Sci. USA*, **102**, 19075–19080.

31. Chin,L.J. *et al.* (2008) A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. *Cancer Res.*, **68**, 8535–8540.

32. Wojcik,S.E. *et al.* (2010) Non-codingRNA sequence variations in human chronic lymphocytic leukemia and colorectal cancer. *Carcinogenesis*, **31**, 208–215.

33. Mott,J.L. *et al.* (2007) mir-29 regulates Mcl-1 protein expression and apoptosis. *Oncogene*, **26**, 6133–6140.

34. Santanam,U. *et al.* (2010) Chronic lymphocytic leukemia modeled in mouse by targeted miR-29 expression. *Proc. Natl Acad. Sci. USA*, **107**, 12210–12215.

35. Mraz,M. *et al.* (2009) MicroRNA isolation and stability in stored RNA samples. *Biochem. Biophys. Res. Commun.*, **390**, 1–4.

36. Lipshutz,R.J. *et al.* (1999) High density synthetic oligonucleotide arrays. *Nat. Genet.*, **21**(suppl. 1), 20–24.

37. Marshall,O.J. (2004) PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics*, **20**, 2471–2472.

38. Cutler,D.J. *et al.* (2001) High-throughput variation detection and genotyping using microarrays. *Genome Res.*, **11**, 1913–1925.

39. Di,X. *et al.* (2005) Alternative base calling method for resequencing microarrays. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, **3**, 2809–2812.

40. Bruce,C.K. *et al.* (2010) Design and validation of a metabolic disorder resequencing microarray (BRUM1). *Hum. Mutat.*, **31**, 858–865.

41. Consortium,G.P. (2010) A map of human genome variation from population-scale sequencing. *Nature*, **467**, 1061–1073.

42. Gruber,A.R. *et al.* (2008) The Vienna RNA websuite. *Nucleic Acids Res.*, **36**(Web Server issue), W70–W74.

43. Moorhead,M. *et al.* (2006) Optimal genotype determination in highly multiplexed SNP data. *Eur. J. Hum. Genet.*, **14**, 207–215.

44. Zeng,Y. *et al.* (2005) Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J.*, **24**, 138–148.

45. Han,J. *et al.* (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell*, **125**, 887–901.

46. Zeng,Y. *et al.* (2005) Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. *J. Biol. Chem.*, **280**, 27595–27603.

47. Cimmino,A. *et al.* (2005) miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl Acad. Sci. USA*, **102**, 13944–13949.

48. Lujambio,A. *et al.* (2007) CpG island hypermethylation of tumor suppressor microRNAs in human cancer. *Cell Cycle*, **6**, 1455–1459.

49. Bandres,E. *et al.* (2009) Epigenetic regulation of microRNA expression in colorectal cancer. *Int. J. Cancer*, **125**, 2737–2743.

50. Hu,Z. *et al.* (2009) Common genetic variants in pre-microRNAs were associated with increased risk of breast cancer in Chinese women. *Hum. Mutat.*, **30**, 79–84.

51. Xu,B. *et al.* (2010) A functional polymorphism in Pre-miR-146a gene is associated with prostate cancer risk and mature miR-146a expression *in vivo*. *Prostate*, **70**, 467–472.

52. Jazdzewski,K. *et al.* (2008) Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. *Proc. Natl Acad. Sci. USA*, **105**, 7269–7274.

53. Zhou,B. *et al.* (2011) Common genetic polymorphisms in pre-microRNAs and risk of cervical squamous cell carcinoma. *Mol. Carcinog.*, **50**, 499–505.

54. Mencía,A. *et al.* (2009) Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. *Nat. Genet.*, **41**, 609–613.

55. Borel,C. *et al.* (2008) Functional genetic variation of human miRNAs and phenotypic consequences. *Mamm. Genome*, **19**, 503–509.

56. Xu,T. *et al.* (2008) A functional polymorphism in the miR-146a gene is associated with the risk for hepatocellular carcinoma. *Carcinogenesis*, **29**, 2126–2131.

57. Wang,F. *et al.* (2012) miR-29a and miR-142-3p downregulation and diagnostic implication in human acute myeloid leukemia. *Mol. Biol. Rep.*, **39**, 2713–2722.

58. Li,Y. *et al.* (2009) Role of the miR-106b-25 microRNA cluster in hepatocellular carcinoma. *Cancer Sci.*, **100**, 1234–1242.

59. Voorhoeve,P.M. *et al.* (2006) A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. *Cell*, **124**, 1169–1181.

60. Malcikova,J. *et al.* (2009) Monoallelic and biallelic inactivation of TP53 gene in chronic lymphocytic leukemia: selection, impact on survival, and response to DNA damage. *Blood*, **114**, 5307–5314.