Clinical Monoclonal B lymphocytosis versus Rai 0 Chronic Lymphocytic Leukemia: a Comparison of Cellular, Cytogenetic, Molecular, and Clinical Features

Running title: Clinical and molecular profiles of cMBL versus Rai0-CLL

Fortunato Morabito,1* Laura Mosca,2 Giovanna Cutrona,3 Luca Agnelli,2 Giacomo Tuana,2 Manuela Ferracin,4 Barbara Zagatti,4 Marta Lionetti,2 Sonia Fabris,2 Francesco Maura,2 Serena Matis,5 Massimo Gentile,1 Ernesto Vigna,1 Monica Colombo,5 Carlotta Massucco,5 Anna Grazia Recchia,1 Sabrina Bossio,1 Laura De Stefano,1 Fiorella Ilariucci,6 Caterina Musolino,7 Stefano Molica,8 Francesco Di Raimondo,9 Agostino Cortelezzi,2 Pierfrancesco Tassone,10 Massimo Negrini,4 Sara Monti,11 Davide Rossi,11 Gianluca Gaidano,11 Manlio Ferrarini,12 and Antonino Neri2*

1 Department of Onco-hematology, A.O. of Cosenza, Cosenza, Italy; 2 Department of Clinical Sciences and Community Health, University of Milano and Hematology 1 CTMO, Foundation IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Milano, Italy; 3 SS Molecular Diagnostics IRCCS S. Martino-IST, Genova, Italy; 4 Department of Experimental and Diagnostic Medicine, University of Ferrara, Ferrara, Italy; 5 Scientific Direction IRCCS S. Martino-IST, Genova, Italy; 6 Division of Hematology, A.O. Arcispedale Santa Maria Nuova-IRCCS, Reggio Emilia, Italy; 7 Division of Hematology, University of Messina, Messina, Italy; 8 Department of Oncology and Hematology, Pugliese-Ciaccio Hospital, Catanzaro, Italy; 9 Department of Biomedical Sciences, Division of Hematology, University of Catania and Ferrarotto Hospital, Catania, Italy; 10 University of Magna Graecia and Cancer Center, Catanzaro, Italy; 11 Division of Hematology,
Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; 12 Direzione Scientifica IRCCS, San Martino IST, Genoa, Italy

*CORRESPONDING AUTHORS

Fortunato Morabito, MD
UOC Ematologia, Azienda Ospedaliera di Cosenza,
viale della Repubblica snc, 87100 Cosenza, Italy
e-mail: fortunato_morabito@tin.it
ph: +39-0984-681329
fax: +39-0984-681329

Antonino Neri, MD
Dept. of Clinical Sciences and Community Health, University of Milan
Hematology 1 CTMO, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico
Via Francesco Sforza, 35 - 20122 Milan, Italy
e-mail: antonino.neri@unimi.it
ph: +390255033758
fax: +390255034571

Text: 4416 words
Abstract: 233 words
Figures: 5
Tables: 1
References: 42
Translational relevance

The improvement of flow cytometry techniques for the diagnosis of chronic lymphocytic leukemia (CLL) has led to the detection of small circulating B-cell populations (monoclonal B-cell lymphocytosis, MBL) in otherwise normal subjects. A clonal B-cell value of $<5 \times 10^9$/L in the absence of palpable lymphadenopathy and/or organomegaly is usually utilized to differentiate clinical MBL (cMBL) from Rai stage 0 CLL (Rai0-CLL) patients. However, the biological differences between Rai0-CLL and cMBL remain to be clarified. Our study indicates that an increased proportion of cases with unmutated $IGHV$ genes is the only differentiating feature significantly associated with cMBL condition while no specific gene/miRNA signatures were found distinguishing the two patient groups. The diagnosis of cMBL per se was predictive of a longer progression free survival. Overall, our data suggest that cMBL and Rai0-CLL patients may mainly differ in terms of the initial clonal size of a slowly expanding B-cell population.
ABSTRACT

Purpose: To investigate the incidence and clinical relevance of classic and new prognostic markers, IGHV gene mutational status, and chromosomal abnormalities in cMBLs compared to Rai stage 0 chronic lymphocytic leukemia (Rai0-CLL).

Experimental design: A group of 136 clinical monoclonal B lymphocytosis (cMBL) patients and a group of 216 Rai0-CLL cases were investigated prospectively.

Results: IGHV mutated cases were significantly more frequent among cMBLs (P=0.005), while the distribution of CD38 and ZAP-70 positive cases, of patients with NOTCH1 and SF3B1 mutations or exhibiting the major CLL cytogenetic abnormalities was similar in the two groups. Moreover, no significant differences were found neither in IGHV/IGHD/IGHJ gene usage nor in the overall prevalence of stereotyped IGHV gene sequences. Cells from cMBL and Rai0-CLL exhibited similar gene and miRNA signatures; in addition, when grouped according to the IGHV mutational status, IGHV unmutated cases showed different transcriptional signatures compared to IGHV mutated patients, irrespectively of the cMBL or Rai0-CLL classification. cMBL diagnosis per se was predictive of longer progression free survival.

Conclusion: Our study based on a prospective series of patients indicates that no major differences exist between the circulating cells from cMBL and Rai0-CLL, at least based on a comparison of the markers utilized in the study. This possibly suggests that the two conditions mainly differ in the initial size of the monoclonal cell population, which may influence the subsequent timing of clonal expansion and clinical manifestations.

KEY WORDS – clinical MBL, miRNA, gene expression profiling, NOTCH1, prognosis
INTRODUCTION

Chronic lymphocytic leukemia (CLL) is a slowly progressive disease characterized by increasing levels of lymphocytes in the blood, bone marrow and lymphoid tissues (1). The majority of individuals with CLL do not require treatment at presentation and therapy is deferred until there is evidence of bone marrow failure, increasing lymphadenopathy/splenomegaly, or progressive lymphocytosis with a rapid doubling time (2).

Over the past 30 years, the diagnostic criteria for CLL have changed significantly. Notably, the identification of CD5 and CD23 on the surface of leukemic cells as the hallmark of CLL and the use of progressively improved flow cytometry techniques have led to the detection of small amounts of circulating monoclonal B-cell populations (monoclonal B-cell lymphocytosis, MBL) in otherwise normal subjects (2). The reported prevalence of MBL in the general population varies substantially, from 0.6% to as much as 20% in some studies; these differences can be attributed to several variables including the type of study design, the sensitivity of the flow cytometry technique, and the age of the population studied (3-7).

Furthermore, MBL may represent a heterogeneous entity since the circulating cells are indistinguishable from those of classic CLL in some cases, and phenotypically different in others (3, 8, 9). Finally, MBL has been characterized by a progression to CLL at an annual rate of 1.1-2% (10-14).

MBL is subdivided into two groups; i.e. clinical MBL (cMBL) and population-screening MBL (7, 13, 15). The former is generally characterized by values of clonal circulating B-cells greater than ~1500/μL and is mostly detected in individuals who undergo a routine blood test and present a mild lymphocytosis, whereas the latter, observed in screening tests of healthy people using flow cytometry (7) is also denominated low-count MBL (lcMBL) (13).
Although there are no defined cut-off values to distinguish the two entities based upon circulating B-lymphocyte numbers, most lcMBL patients have a B-lymphocyte count below 0.5x10^9/L (7, 14). Finally, cMBL, but not lcMBL has a higher risk of developing an overt CLL requiring treatment than healthy age-matched controls (13).

Early stage CLL (i.e. Rai stage 0) is defined by the presence of lymphocytosis only. In the past, patients were required to present with monoclonal B-cells in the peripheral blood with a minimum absolute lymphocyte count (ALC) of 5.0x10^9/L. However, the recent IWCLL classification led to the replacement of this value by a count of 5.0x10^9 clonal B-cells/L (7-9). Accordingly, many CLL patients (up to 40%) who would have previously been diagnosed as having a Rai stage 0 CLL (Rai0-CLL), can now be classified as having cMBL (16).

The relationships between the cells of CLL and those of cMBL are still undefined. In other words, do the cMBL cells differ from overt CLL cells for a number of intrinsic features and a general lower degree of aggressiveness or are the differences between the two clinical entities mostly related to the clonal size?

Approximately five years ago, a multicenter observational study was started in Italy, recruiting Binet stage A patients at diagnosis. Since, at the onset of the study, the diagnostic criteria were based upon the old National Cancer Institute (NCI) recommendations, a significant fraction of cases had to be re-classified as cMBL according to new IWCLL guidelines. Therefore, our series offered the opportunity to investigate the incidence and clinical relevance of classic prognostic markers, IGHV status, chromosomal abnormalities and NOTCH1 and SF3B1 mutations (17, 18) in cMBLs, comparing them with those of Rai0-CLLs. In addition, gene and microRNA (miRNA) expression patterns were analyzed in both groups. To our knowledge, this is the first report addressing these issues in a prospective series of patients.
Materials and Methods

Patients

Newly diagnosed Binet stage A CLL patients from several Institutions were prospectively enrolled within 12 months of diagnosis (O-CLL1 protocol, clinicaltrial.gov identifier NCT00917540). All participants provided written informed consent and the study was approved by the relevant institutional review boards. The inclusion criteria for CLL diagnosis employed at the time of study design and start followed the 1996 NCI-sponsored Working Group guidelines (16), and required an absolute count >5,000 of mature-appearing lymphocytes/μL in the peripheral blood. Exclusion criteria were: i) CLL diagnosis more than 12 months before registration; ii) leukemic phase of CD5- and/or CD23- B-cell lymphoproliferative disorders; iii) Binet stage B or C; iv) need of therapy according to NCI guidelines; and v) age >70 years. Diagnosis was confirmed by the biological review committee according to flow cytometry analysis centralized at the National Cancer Institute (IST) Laboratory in Genoa, together with the determination of CD38 and ZAP-70 expression and the determination of IGHV mutational status.

Overall, 463 Binet stage A patients have been enrolled in this prospective protocol. Of these, 352 cases were staged as Rai 0 CLL and were included in the present study. Based on the criteria of the 2008 NCI-sponsored workshop (2), the 352 Rai 0 cases could be further subdivided into cMBL (136 cases with <5.0x10^9 B lymphocytes/L in the peripheral blood) and Rai0-CLL (216 cases with >5.0x10^9 B lymphocytes/L in the peripheral blood). The median follow up of living patients is 46 months for cMBL and 38 months for Rai0-CLL patients.

Sample preparation, CD38 and ZAP-70 determination, NOTCH1, SF3B1 and IGHV genes mutational status and cytogenetic analysis
Peripheral blood mononuclear cells from CLL patients were isolated by Ficoll-Hypaque (Seromed, Biochrom KG, Berlin, Germany) density-gradient centrifugation. For gene and miRNA expression profiling experiments CLL cells were enriched by negative selection with the EasySep-Human B-cell enrichment kit without CD43 depletion (Stem Cell Technologies, Voden Medical Instruments S.p.A, Milano, Italy) using the fully automated protocol of immunomagnetic cell separation with RoboSep™ (Stem Cell Technologies) (19). The percentage of purified B-cells (CD19+) exceeded 95%. The proportion of CD5/CD19/CD23 triple positive B-cells in the suspension was determined by flow cytometry using CD19 FITC (BD Biosciences Pharmingen, San José California, USA), CD23 PE (BD Biosciences), and CD5 PC5 (Beckman Coulter Immunotech, Marseille, France) monoclonal antibodies (mABs).

CD38-positive leukemic cells were measured by triple staining with CD19 FITC (BD Biosciences), CD38 PE (BD Biosciences), and CD5 PC5 (Beckman Coulter) mABs. The cells were analyzed using a FACSCalibur flow cytometer (BD Biosciences). A cut-off of ≥ 20% was used for CD38 positivity as previously reported (20, 21).

ZAP-70 was determined by flow cytometry with a ZAP-70-FITC (clone 2F3.2, Millipore, Temecula, California, USA) or an isotype control mAb (mouse IgG2a-FITC BD Biosciences) as previously described (19, 22, 23). In brief peripheral mononuclear cells purified from fresh heparinized CLL samples by Ficoll Hypaque gradient were first incubated with CD3 PE-CY7, CD19 PE, and CD5 allophycocyanin (APC) monoclonal antibodies (mAbs) (Becton Dickinson & Co.), fixed, permeabilized with fix and perm reagents (Caltag Laboratories) and exposed to ZAP-70 or the isotype control mAb. A cut-off of >30% was used for ZAP-70 positivity as previously reported (24) and calculated by receiver-operating characteristic (ROC) analysis as the most suitable ZAP-70 cut-off value for distinguishing IGHV-UM from IGHV-M cases. This cut-off is different from that originally described (around 20%) in previous studies (22, 23);
the difference could be due to the higher sensitivity of the antibody we used (2F3.2 clone) compared to the 1E7.2 clone as already suggested by previous studies (25, 26); see Wang et al. for critical review (27).

The mutation hotspots of NOTCH1 (exon 34; RefSeq NM_017617.2) and SF3B1 (exons 14, 15, 16 including splice sites; RefSeq NM_012433.2) genes were analyzed by PCR amplification and Sanger sequencing of high molecular weight genomic DNA extracted from CD19+ purified B-cells as previously described (17, 18). To assess IGHV mutational status, RNA was obtained from peripheral blood specimens. Sequencing was performed as previously reported and the sequences obtained were aligned to IMGT and analyzed using IMGT/VQUEST software. Sequences differing more than 2% from the corresponding germ-line gene were considered mutated (28). Stereotyped HCDR3 analysis was performed as previously described (29, 30).

Cytogenetic abnormalities involving deletions at chromosomes 11q23, 13q14 and 17p13 and trisomy 12 were evaluated by fluorescence in situ hybridization (FISH) in purified CD19+ population as previously described (31).

Gene expression profiling

The total RNA fraction was obtained from CD19+ purified B-cell samples using Tryzol reagent (Life Technologies, Carlsbad, California, USA). RNA quality was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies). Low quality RNA (RNA integrity number below 7) was excluded from microarray analyses. Preparation of DNA single-stranded sense target, hybridization to GeneChip® Gene 1.0 ST Array (Affymetrix Inc., Santa Clara, CA) and scanning of the chips (7G Scanner, Affymetrix) were carried out according to manufacturer’s protocols. Expression Console Software (Affymetrix) was used for probe set summarization.
and robust multi-array average (RMA) normalization procedures. Different probe sets mapping on the same gene were summarized by median value over the same sample.

The supervised gene expression analyses were performed using the Significance Analysis of Microarrays software (SAM version 4.0; Excel front-end publicly available at http://www-stat.stanford.edu/~tibs/SAM/index.html) (32) as previously reported (33). The stringency level of the 90th percentile of false discovery rate (FDR) being equal to 0 was applied.

miRNA expression profiling

Total RNA extraction and the quality assessment of purified samples were performed as for gene expression profiling. Samples were then prepared and analyzed in conformity with the Agilent manual on the Human miRNA Microarray V2 platform (Agilent Technologies, Santa Clara, CA), which represents miRNAs from the Sanger miRBase (v10.1). The expression values of miRNA were estimated by Agilent Feature Extraction Software 10.1, then summarized and background subtracted. miRNA with low expression (all detection calls absent), non-human miRNA and miRNAs expired according to Sanger miRBase Release 15 (April 2010) were excluded. Raw data obtained were quantile normalized, converted into positive values setting a minimum value of 1 and log2 transformed using the R-2.14 statistical environment (http://www.r-project.org/). Different samples groups were compared by supervised analysis using SAM with a significant threshold set at q-value=0. The unsupervised analyses were performed by hierarchical clustering algorithm, using centroid linkage and Pearson’s correlation as distance measure, on the highly variable genes and miRNAs whose average ratio exceed 2-fold on the mean across the dataset; heatmaps were carried out using dChip software (http://biosun1.harvard.edu/complab/dchip/).
The gene and miRNA expression data have been deposited at NCBI Gene Expression Omnibus repository (http://www.ncbi.nlm.nih.gov/geo) and are accessible through GEO Series accession number GSE40571.

Real-Time Quantitative Polymerase Chain Reaction

LOC400986 gene and miR-130a expression levels were analyzed in purified CD19+ cells by means of real-time quantitative polymerase chain reaction (Q-RT-PCR) assays. LOC400986 gene, total RNA was converted to cDNA using M-MLV reverse transcriptase (Invitrogen). Inventoried TaqMan Gene Expression Assay (Hs02340701_mH) and the TaqMan Fast Universal Master Mix were used according to manufacturer’s instructions (Applied Biosystems, Foster City, CA). UBC TaqMan Gene Expression Control (Hs00824723_m1) (Applied Biosystems) was used as the internal control. miR-130a expression was validated Q-RT-PCR using TaqMan® microRNA assay (Applied Biosystems, Foster City, CA) in accordance with the manufacturer’s protocol. All of the RNA samples were normalized on the basis of the RNU48 TaqMan® microRNA assays control. The measurement of transcript expression was performed using the Applied Biosystems StepONE Real-Time PCR System. All the samples were run in duplicate. Data were expressed as 2-ΔCt (Applied User Bulletin No. 2).

Statistical analysis

The statistical package IBM SPSS Statistics 19 for Windows, (SPSS inc, and IBM company, 2010) was used for all analyses. For categorical variables, statistical comparisons were performed using two-way tables for the Fisher’s exact test and multiway tables for the Pearson’s Chi-square test. The Mann-Whitney U-test was employed to assess differences in the distribution of any specific variable between two groups. Logistic binomial regression
analysis was used to estimate the contribution of single covariates for predicting the response variable (i.e. cMBL or Rai0-CLL), and deviance (D) reported as measure of the likelihood-ratio test for each covariate. Progression Free Survival (PFS) analyses, determined as time to therapy need according to guidelines (16), were performed using the Kaplan-Meier method. Statistical significance of associations between individual variables and survival was calculated using the log-rank test. The prognostic impact for the outcome variable was investigated by univariate and multiple Cox regression analysis. Data are expressed as hazard ratio (HR) and 95% confidence interval (CI). Cox proportional hazards model was used in the globaltest function of R software (34) (under 100,000 permutation) to test the positive or negative association between covariates expression levels, assumed as continuous variable, and clinical outcome as response variable (in terms of progression free survival). A value of $P<.05$ was considered significant for all statistical calculations.

Results

Comparison of biological characteristics between cMBL and Rai0-CLL

One hundred thirty-six (38.6%) cMBL and 216 (61.4%) Rai0-CLL patients were included in the study. While there was no difference in the proportion of CD38 and ZAP-70 positive cases in cMBL and Rai0-CLL, cMBLs were comprised of a significantly lower number of IGHV-UM cases (27/134; 20.15% vs 74/215; 34.4%, $P=0.005$) (Table 1). The IGHV genes most frequently utilized in CLL had the same distribution in cMBLs and in Rai0-CLLs (Supplementary Fig. S1). No significant differences were found in the IGHV/IGHD/IGHJ gene usage as well as in the overall prevalence of stereotyped IGHV gene sequences between cMBL and Rai0-CLL patients (24.2% and 31%, respectively) (29). Interestingly, all patients expressing subset #1 (IGHV1-5-7/IGHD6-19/IGHJ4) were identified in the Rai0-CLL group
(7/215, 3.3%); however there was no difference in the distribution of the other most frequent subsets (#2, #4, #7, #8 and #10) in the whole cMBL and Rai0-CLL groups or following stratification of the two groups according to CD38 or ZAP-70 expression or IGHV gene mutation status.

There were no differences between cMBL and Rai0-CLL in the proportion of patients with del(13)(q14), while the percentage of patients with trisomy 12 was slightly higher in cMBL (20/121, 15.7% vs 18/206, 8.7%; P=.05). Del(11)(q22) was detected in 15/206 (7.2%) Rai0-CLL and in 3/121 (2.5%) cMBL patients and 17p deletion was observed in 1/121 (0.8%) cMBL and 5/207 (2.4%) Rai0-CLL cases, respectively. Overall, no significant difference in the incidence of the major genetic lesions was observed between the two groups (P=.1) (Table 1).

Two-hundred forty-five (85 cMBL and 160 Rai0-CLL) cases were studied for NOTCH1 and SF3B1 mutations, with no significantly different distribution of mutated cases between the cMBL and the Rai0-CLL group (NOTCH1: 8.2% versus 13.1%; SF3B1: 4.7% versus 3.8%; Table 1).

Prognostic markers and disease progression in cMBL and Rai0-CLL

Complete clinical information was available for 329 patients (127 cMBL and 202 Rai0-CLL). The median follow-up was 46 months (range 4-76 months) for cMBLs and 38 months (range 1-82 months) for Rai0-CLLs. Eight cMBL and forty-six Rai0-CLL cases progressed and required therapy. The 3-year PFS probability of cMBL patients was 95% compared to 80% of the Rai0-CLL cases [(HR=4.5, 95%CI 2.1-9.5, P<.0001); Figure 1A]. Moreover, we performed further analyses using different peripheral B-cell count cut-offs. No statistically significant differences were found when cMBL cases were stratified based upon different peripheral B-
cell counts, while the probability of progressing significantly correlated with the peripheral
B-cell counts in Rai0-CLL cases (Figure 1B). Notably, the 3-year PFS probabilities were
significantly smaller for the IGHV-UM, CD38 and ZAP-70 positive Rai0-CLL patients than for
the corresponding cMBL cases having unfavorable prognostic markers (Figure 2A-C,
respectively). Moreover CD38 or ZAP-70 positive or IGHV-UM cMBL cases had a PFS similar
to that of Rai0-CLL cases with favorable prognostic factors (i.e. CD38 or ZAP-70 negative or
IGHV-M).

Despite the lack of significant differences in most comparisons, high risk markers were
slightly overrepresented in the Rai0-CLL group. A regression analysis, aimed at evaluating the
predictive power of the single variables, revealed an association of the Rai0-CLL presentation
with IGHV-UM status (D=4.35, P=0.037) and del(11)(q22) (D=8.91, P=0.0028) only. Therefore,
we investigated whether the occurrence of unfavorable prognostic factors may uncover
significance if they were considered in aggregate. These factors included IGHV-UM, ZAP-70
positivity, CD38 positivity, del(11)(q22), del(17)(p13), absence of del(13)(q14), NOTCH1 and
SF3B1 mutations. More than one lesion was found in 41 cMBL and 85 Rai0-CLL patients.
Although cMBL and Rai0-CLL had the same median values of co-existing alterations (i.e. 2,
Supplementary Fig S2A), the proportion of CLL cases exhibiting 3 or more unfavorable
markers was larger than that of cMBL (Figure 3); the cut-off of 3 alterations represented the
highest difference to indicate progression (Supplementary Fig S2B), and when only patients
with 3 or more alterations were considered, a significant difference between Rai0-CLL and
cMBL was demonstrated (P=0.004, Fisher’s exact test).

Comparison of the gene and miRNA transcription profiles of MBL and Rai0-CLL cells
Gene expression profiling (GEP) analysis was performed on purified B-cells from 45 cMBLs and 115 Rai0-CLLs (Supplementary Table S1). Hierarchical clustering analysis of GEP data (Supplementary Fig S3) indicated that cMBL cases were intermingled among Rai0-CLL cases. Likewise, miRNA global profiling performed on the purified B-cells from 39 cMBLs and 111 Rai0-CLLs (Supplementary Table S1) failed to show specific grouping of cMBL cases by unsupervised analysis (Supplementary Fig S3).

Next, we used a supervised approach in a further attempt to identify differences in genes/miRNAs expression distinguishing the two groups. Notably, no specific signatures were identified for the two conditions at q-value=0; in fact, only a single gene (LOC400986) and a miRNA (miR-130a) transcript were found to be differentially expressed in cMBLs vs Rai0-CLLs (down- and up-regulated in cMBL, respectively) (Supplementary Fig. S4). The expression of miR-130a and LOC400986 was further validated by means of Q-RT-PCR in all samples for which RNA was available (114 and 120 cases, respectively) demonstrating a very good concordance for miR-130a and a borderline correlation for LOC400986 (Pearson correlation coefficients were 0.89 and 0.67, respectively) (Supplementary Fig. S5).

A series of computational analyses were performed to provide insights into the putative implication of these data. First, we failed to detect an inverse correlation between miR-130a and LOC400986 expression which seems to exclude the hypothesis of a regulatory relationship between the two transcripts. Moreover, no miRNA/target relationship was predicted by a number of commonly employed algorithms (MiRanda, PicTar, PITA, TargetScan, DIANA).

Second, to address the possible role of these two transcripts in clinical progression, we tested possible correlations between miR-130a and LOC400986 expression levels and disease progression. Although a clear cut association was not detected when the whole case
cohort was considered, a higher miR-130a expression marginally correlated with an earlier progression when only MBL cases were considered ($P=0.013$). Conversely, no association between LOC400986 and PFS was unraveled in cMBL or in Rai0-CLL (data not shown).

Finally, cMBL and Rai0-CLL cases were grouped according to the IGHV mutational status and multiclass supervised analyses of gene and miRNA expression profiles were performed comparing the four patients groups (IGHV-UM/IGHV-M cMBLs and IGHV-UM/IGHV-M Rai0-CLLs). As shown in Figure 4 these analyses identified 302 genes and 18 miRNAs (Supplementary Table S2), which showed near identical expression pattern in IGHV-UM vs IGHV-M cases, independent of the cMBL or Rai0-CLL condition. Specifically, of the 302 genes, 247 showed the very same directionality in the signs of the class scores all over the four groups; of the remaining 55, 47 showed different signs in the scores but concordant modulation between IGHV-UM and IGHV-M cases. Only eight genes (2.6%) showed discordant behavior between IGHV-UM and IGHV-M cases depending on the cMBL or Rai0-CLL presentation; however, it is worth noting that they all were in the lowest quartiles among the differentially expressed genes. Of the 18 miRNAs, 14 showed the very same directionality in the signs of the class scores all over the four groups, and the remaining 4 showed different signs in the scores but concordant modulation between IGHV-UM and IGHV-M cases (Supplementary Table S3).

Discussion

This study stemmed from an Italian multicenter trial (O-CLL1 trial) aimed at recruiting Binet stage A CLL patients at diagnosis. The patients were characterized for the major genetic abnormalities, cellular and molecular markers and followed prospectively over time. The trial primary endpoint was to evaluate prospectively known markers in the prediction of
the clinical outcome of early stage CLL patients; PFS, intended as time to the first treatment, was considered. This multicenter study began before the publication of the IWCLL guidelines for CLL diagnosis (2) and hence a significant number of Rai0-CLL included in the study fulfilled the new diagnostic criteria proposed for cMBL. This gave us the opportunity to carry out a clinical and biological comparison between these two entities in a large cohort of cases.

The cMBL status \textit{per se} was found to confer a better prognosis than that of Rai0-CLL, since the cMBL condition identified cases characterized by a longer PFS (Figure 1A). These data are in line with those in CLL patients stratification demonstrating a correlation between B-lymphocytosis values at diagnosis and disease progression (see Figure 1B). This study also confirmed the value of prognostic factors given the fact that CD38 or ZAP-70 positive or IGHV-UM cMBL cases had a similar need of therapy as Rai0-CLL cases characterized by IGHV-M status and CD38 or ZAP-70 negativity (Figure 2).

Patient aged more than 70 years were not included in the study. Although we are aware that this may have an impact on our data set since elderly individuals represent near half of the CLL patients, inclusion of elderly patients could have affected the assessment of the natural history of the disease because of the higher frequency of CLL-unrelated deaths. Additional complications for a correct follow-up could have been originated by the fact that elderly patients often choose to be followed by family doctors rather than attending the hospital clinic. Indeed, only an accurate follow-up may allow to dissect the prospective value of predictive markers. The trial also excluded all patients whose diagnosis dated more than 12 months from enrollment. This strategy is likely to have excluded from the study a number of patients, i.e those followed by the participating centers who have progressed from Binet stage A to other stages within one year or less, and those who had been already followed for
long periods having a very indolent disease. Unfortunately, these data have not been quantified and therefore it would be difficult to determine which group was selected against by the study design. However, this exclusion was again suggested by the observation that certain markers may change during disease and hence the assessment of their predictive value requires a determination at or near to the time of diagnosis. According to the study design, determination of markers was carried out every third year and at progression, a strategy capable of assessing potential changes in marker profile.

The whole spectrum of the most frequent chromosomal abnormalities currently detected in CLL was observed in both cMBL and Rai0-CLL (see circular layout plot in Figure 5 that graphically summarized the relationship between Rai0-CLL or cMBL and the major prognostic factors in our cohort). However, unfavorable cytogenetic lesions, such as del(17)(p13) and del(11)(q22) were rare in stage Rai0-CLL and even less frequent in cMBL, although the differences were not significant for the two groups (4, 13). This confirms the notion that these cytogenetic lesions accumulate with time and perhaps with progressive clonal expansion (15, 35, 36). Given the scanty proportion of positive cases, these cytogenetic alterations may have a limited predictive power at the early stages of the disease.

With the exception of IGHV mutational status, there was no significant difference in the prognostic markers of cMBL and Rai0-CLL (including chromosomal lesions), although a trend was noticed in the prevalence of unfavorable prognostic factors in Rai0-CLL vs cMBL. The explanation of this phenomenon is difficult, although it could be speculated that the mechanisms inducing clonal expansion (e.g. antigenic stimulation) are already operative in the transition from cMBL to Rai0-CLL (37). This would explain the prevalence of IGHV-UM cases in the Rai0-CLLs and suggests that larger differences can be detected when cMBL are
compared to CLL cases progressed to more advanced stages, as indicated by Lanasa et al. (38). Moreover, the promoting factors inducing clonal expansion also would facilitate the expression of cellular activation markers (e.g. ZAP-70 and CD38) and the slow accumulation of additional unfavorable lesions (e.g. additional chromosomal abnormalities or NOTCH1 and SF3B1 mutations, appearing late in the expanding clone). Together, these considerations could explain the slight predominance of unfavorable markers in Rai0-CLL when considered in aggregate.

Finally, we have investigated cMBL and Rai0-CLL patients for both gene and miRNA expression profiles. The latter approach could have been particularly relevant to provide new insights, given the proposed role of miRNA involvement in CLL origin and progression (15, 39). However, no differences in gene or miRNA expression profiles were observed between cMBL and Rai0-CLL using unsupervised analyses. Similarly, under our experimental conditions, no significant transcriptional signatures were observed by supervised analyses comparing Rai0-CLL and cMBL, with the only exception of miR-130a and LOC400986 gene, also known as ANKRD36C (ankyrin repeat domain 36C), a putative member of the functionally unknown family of ubiquitin-interacting motif (UIM)-bearing proteins. miR-130a was up-regulated in cMBLs compared to Rai0-CLLs. This miRNA is involved in cell autophagy and survival and in the regulation of expression of other miRNAs (40). From predicted target database analysis, miR-130a may regulate genes involved in the TGF-beta, mTOR and Hedgehog signaling pathways, all of which have been reported to be involved in CLL. In addition, miR-130a is predicted to target genes belonging to the IGF1-PDGFRA signaling pathway involved in cell proliferation and focal adhesion as well as in mechanisms associated to disease progression in CLL (41, 42). In our cohort, the expression of miR-130a was slightly correlated with earlier time to progression in cMBLs but not in Rai0-CLLs.
Although these results are of potential relevance, further investigations are needed to clarify the putative pathogenic role of these two transcripts. Meta-analyses in different patient series would be useful to reinforce our findings, but no public data are available to date that share gene/miRNA information concerning expression levels and absolute lymphocytosis count.

When cases from the whole patient cohort were stratified according to their IGHV mutational status, we found that IGHV-UM cases had both gene and miRNA expression profiles differing from those of IGHV-M cases, irrespectively of their classification as cMBL or Rai0-CLL. This strongly suggests that significant differences in gene and miRNA expression are already present at very early disease stages in IGHV-UM versus IGHV-M patients.

Overall, our study strongly supports the notion that cMBL and Rai0-CLL do not significantly differ in terms of biological and molecular features, and conversely suggests that cMBL and Rai0-CLL mainly differ due to the initial size of the monoclonal cell population and that the slower progression rate observed in cMBL may be primarily related to the time required for the leukemic clone to expand. The predominance of cases with IGHV-UM genes among Rai0-CLL might further reinforce the power of this predictive marker and suggest that stimulation via surface Ig represents an important factor driving clonal expansion.

ACKNOWLEDGEMENTS

This work was supported by Associazione Italiana Ricerca sul Cancro (AIRC, AN-IG10136, MF-IG10492, and FM-RG6432), AIRC–Special Program Molecular Clinical Oncology (5 per Mille, grant 9980, 2010-15; A.N., M.F., P.T., M.N. and F.Mo.) Ricerca Finalizzata from the Italian Ministry of Health 2006 (G.C., F.Mo., M.F., and P.T.) and 2007 (G.C.), Fondo Investimento per la Ricerca di Base (grant RBIP06LCA9; M.F.), progetto Compagnia San Paolo (G.C.), the
Ministero della Salute, Ricerca Corrente, and Ricerca Finalizzata FSN, Rome, Italy (A.P.). S.B. and L.D.S. were supported by fellowships from the AIRC; M.L. was supported by a fellowship from the Fondazione Italiana Ricerca sul Cancro (FIRC).

**AUTHORSHIP CONTRIBUTIONS:**

**Conception and design:** F. Morabito, M. Ferrarini and A. Neri

**Acquisition of data (performed microarray, Q-RT-PCR assays and bio-molecular marker analysis):** L. Mosca, M. Ferracin, B. Zagatti, M. Lionetti, D. Rossi, G. Cutrona, S. Fabris, S. Matis, M. Colombo, S. Monti and C. Massucco

**Analysis and interpretation of data:** L. Mosca, L. Agnelli, G. Tuana, S. Fabris, F. Maura, and M. Negrini

**Writing, review and/or revision of the manuscript:** F. Morabito, L. Mosca, L. Agnelli, A.G. Recchia, G. Gaidano, M. Ferrarini and A. Neri

**Administrative, technical, or material support:** M. Gentile, E. Vigna, S. Bossio, L. De Stefano, F. Ilariucci, C. Musolino, D. Rossi, S. Molica, F. Di Raimondo, A. Cortelezzi, P. Tassone, and F. Morabito.

**Study supervision:** F. Morabito, M. Ferrarini and A. Neri

**CONFLICT OF INTEREST DISCLOSURE**

The authors declare no competing financial interests.
References


<table>
<thead>
<tr>
<th></th>
<th>cMBL (% of total)</th>
<th>Rai-0-CLL (% of total)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age, y (range)</td>
<td>61.5 (38-70)</td>
<td>60.4 (30-70)</td>
<td>.49^a</td>
</tr>
<tr>
<td>Sex, male/female</td>
<td>69/67 (50.7/49.3)</td>
<td>128/88 (59.3/40.7)</td>
<td>.12^a</td>
</tr>
<tr>
<td>Absolute Lymphocyte Count /mm^3 ± SEM</td>
<td>7,701 ± 218</td>
<td>19,374 ± 1,062</td>
<td>&lt;.0001^a</td>
</tr>
<tr>
<td>Absolute B-Lymphocyte Count /mm^3 ± SEM</td>
<td>3,029 ± 102</td>
<td>14,178 ± 931</td>
<td>&lt;.0001^a</td>
</tr>
<tr>
<td>Monoclonal B-cell Count /mm^3 ± SEM</td>
<td>2,872 ± 170</td>
<td>13,969 ± 964</td>
<td>&lt;.0001^a</td>
</tr>
<tr>
<td>CD38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative</td>
<td>109 (80.7)</td>
<td>163 (75.5)</td>
<td></td>
</tr>
<tr>
<td>positive</td>
<td>26 (19.3)</td>
<td>53 (24.5)</td>
<td>.3^c</td>
</tr>
<tr>
<td>Total</td>
<td>135 (100)</td>
<td>216 (100)</td>
<td></td>
</tr>
<tr>
<td>ZAP-70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative</td>
<td>81 (59.6)</td>
<td>131 (60.9)</td>
<td></td>
</tr>
<tr>
<td>positive</td>
<td>55 (40.4)</td>
<td>84 (39.1)</td>
<td>.8^c</td>
</tr>
<tr>
<td>Total</td>
<td>136 (100)</td>
<td>215 (100)</td>
<td></td>
</tr>
<tr>
<td>IGHV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mutated</td>
<td>107 (79.9)</td>
<td>141 (65.6)</td>
<td></td>
</tr>
<tr>
<td>unmutated</td>
<td>27 (20.1)</td>
<td>74 (34.4)</td>
<td>.005^c</td>
</tr>
<tr>
<td>Total</td>
<td>134 (100)</td>
<td>215 (100)</td>
<td></td>
</tr>
<tr>
<td>VH CDR3 stereotype</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>97 (75.8)</td>
<td>148 (68.8)</td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>31 (24.2)</td>
<td>67 (31.2)</td>
<td>.2^c</td>
</tr>
<tr>
<td>Total</td>
<td>128 (100)</td>
<td>215 (100)</td>
<td></td>
</tr>
<tr>
<td>FISH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no abnormality</td>
<td>43 (34.7)</td>
<td>74 (36.7)</td>
<td></td>
</tr>
<tr>
<td>del(13)(q14)</td>
<td>54 (46.3)</td>
<td>95 (44.9)</td>
<td></td>
</tr>
<tr>
<td>trisomy 12</td>
<td>20 (15.7)</td>
<td>18 (8.7)</td>
<td>.1^b</td>
</tr>
<tr>
<td>del(11)(q22)</td>
<td>3 (2.5)</td>
<td>15 (7.2)</td>
<td></td>
</tr>
<tr>
<td>del(17)(p13)</td>
<td>1 (0.8)</td>
<td>4 (2.4)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>121 (100)</td>
<td>206 (100)</td>
<td></td>
</tr>
<tr>
<td>NOTCH1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wild type</td>
<td>78 (91.8)</td>
<td>139 (86.9)</td>
<td></td>
</tr>
<tr>
<td>mutated</td>
<td>7 (8.2)</td>
<td>21 (13.1)</td>
<td>.3^c</td>
</tr>
<tr>
<td>Total</td>
<td>85 (100)</td>
<td>160 (100)</td>
<td></td>
</tr>
<tr>
<td>SF3B1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wild type</td>
<td>81 (95.3)</td>
<td>154 (96.3)</td>
<td></td>
</tr>
<tr>
<td>mutated</td>
<td>4 (4.7)</td>
<td>6 (3.8)</td>
<td>.7</td>
</tr>
<tr>
<td>Total</td>
<td>85 (100)</td>
<td>160 (100)</td>
<td></td>
</tr>
</tbody>
</table>

^aU Mann-Whitney test  
^bPearson’s Chi-square test  
^cFisher’s exact test
Figure Legends

Figure 1. Kaplan-Maier curves of progression free survival (PFS). (A) A statistically significant different PFS was observed for 136 clinical monoclonal B lymphocytosis (cMBL) compared to 216 chronic lymphocytic leukemia (CLL) Rai stage 0 patients (Rai0-CLL). (B) Evaluation of progression stratifying cases using different peripheral B-cell count cut-offs showed statistically significant correlations in Rai0-CLL but not in cMBL patients.

Figure 2. Kaplan-Meier curves for PFS of cMBL versus Rai0-CLL stratified according to: (A) IGHV mutational status, (B) CD38 and (C) ZAP-70 expression.

Figure 3. Bar plot showing the relative percentages of concomitant unfavorable events in cMBL and Rai0-CLL cases.

Figure 4. Multiclass analyses comparing cMBL and Rai0-CLL samples stratified according to IGHV mutational status. Heatmaps of the differentially expressed (A) genes and (B) miRNAs. Information about IGHV mutational status (‘+’ = M, ‘−’ = UM), CD38, ZAP-70, chromosome 12 trisomy, chromosome 17, 11, and 13 deletions (‘+’ = positive, ‘−’ = negative) are included alongside the patient ID. In the legend bar: turquoise indicates Rai0-CLL and yellow cMBL samples, respectively. The color scale bar represents the relative gene expression changes normalized by the standard deviation, and the color changes in each row represent gene expression relative to the mean across the samples.

Figure 5. Circos plot showing the relative percentage of recurrent unfavorable markers and genetic alterations in Rai0-CLL (CLL) and cMBL (MBL) patients. The variables are arranged...
around the circle and distinguished by different colors: the percentage of concomitance with
the other features are indicated by proportional bars filled with the corresponding color
shown in the inner circle, in which the absolute number of samples is also indicated. The
concomitant features are shown as paths emerging from each one to the others, as large as
the number of cases that share the two features.
Figure 2
Clinical Cancer Research

Clinical Monoclonal B lymphocytosis versus Rai 0 Chronic Lymphocytic Leukemia: a Comparison of Cellular, Cytogenetic, Molecular, and Clinical Features

Fortunato Morabito, Laura Mosca, Giovanna Cutrona, et al.

_Clin Cancer Res_ Published OnlineFirst September 13, 2013.

<table>
<thead>
<tr>
<th>Updated version</th>
<th>Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-13-0622</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supplementary Material</td>
<td>Access the most recent supplemental material at: <a href="http://clincancerres.aacrjournals.org/content/suppl/2013/09/13/1078-0432.CCR-13-0622.DC1">http://clincancerres.aacrjournals.org/content/suppl/2013/09/13/1078-0432.CCR-13-0622.DC1</a></td>
</tr>
<tr>
<td>Author Manuscript</td>
<td>Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.</td>
</tr>
</tbody>
</table>

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. |