Expression of ribosomal and translation-associated genes is correlated with a favorable clinical course in chronic lymphocytic leukemia

Jan Dürig, Holger Nückel, Andreas Hüttmann, Elisabeth Kruse, Tanja Hölter, Katja Halfmeyer, Anja Führer, Roland Rudolph, Naser Kalhori, Arnd Nusch, Silvia Deaglio, Fabio Malavasi, Tarik Möröy, Ludger Klein-Hitpass, and Ulrich Dührsen

B-cell chronic lymphocytic leukemia (B-CLL) is a heterogenous disease with a highly variable clinical course. Recent studies have shown that CD38 surface expression on the malignant cell clone may serve as a prognostic marker in that CD38+ patients with B-CLL are characterized by advanced disease stage, lesser responsiveness to chemotherapy, and shorter survival than CD38− patients. To further investigate the molecular phenotype of these 2 clinical subgroups, we compared the gene expression profiles of CD38+ (n = 25) with CD38− (n = 45) B-CLL patients using oligonucleotide-based DNA chip microarrays representative of approximately 5600 genes. The results showed that B-CLLs display a common gene expression profile that is largely independent of CD38 expression. Nonetheless, the expression of 14 genes differed significantly between the 2 groups, including genes that are involved in the regulation of cell survival. Furthermore, unsupervised hierarchical cluster analysis of 76 B-CLL samples led to the separation of 2 major subgroups, comprising 20 and 56 patients. Clustering to the smaller group was due in part to the coordinate high expression of a large number of ribosomal and other translation-associated genes, including elongation factors. Importantly, we found that patients with high expression of translation factors were characterized by a more favorable clinical course with significantly longer progression-free survival and reduced chemotherapy requirements than the remaining patients (P < .05). Our data show that gene expression profiling can help identify B-CLL subtypes with different clinical characteristics. Furthermore, our results suggest a role of translation-associated genes in the pathogenesis of B-CLL. (Blood. 2003;101:2748-2755)

© 2003 by The American Society of Hematology

Introduction

B-cell chronic lymphocytic leukemia (B-CLL) is a heterogenous disease with a highly variable clinical course. Staging systems devised by Rai et al[1] and Binet et al[2,3] are useful methods for predicting survival and treatment requirements in patients with CLL. However, these staging systems are of limited prognostic value in early stages of the disease (Binet A or Rai 0-II); this includes most patients at diagnosis. Therefore, studies have focused on identifying novel prognostic markers that may help define patient subgroups with favorable versus poor clinical outcomes in early CLL.[3,4] Recently, 2 independent studies by Damle et al[5] and Hamblin et al[6] have demonstrated that B-CLL may arise from an immature pregerminal center B cell with unmutated immunoglobulin (Ig) variable heavy chain (V_H) genes or from a more mature postgerminal center B cell with somatically mutated immunoglobulin V_H genes. Moreover, Damle et al[5] found a strong correlation between immunoglobulin V_H gene mutation status, CD38 surface expression of the respective B-CLL clone, and clinical outcome in individual patients. B-CLL patients with mutated immunoglobulin V_H genes and low numbers of CD38+ cells exhibit a favorable clinical course, whereas B-CLL patients with unmutated immunoglobulin V_H genes experience poor outcome in terms of reduced survival and reduced responsiveness to chemotherapy.

Two recent gene-expression profiling studies tested the 2-disease model of CLL by correlating gene expression in CLL cells with their immunoglobulin-mutational status.[7,8] Unexpectedly, both studies using unsupervised hierarchical cluster analysis found a common gene-expression profile regardless of the immunoglobulin-mutational status of the CLL patients investigated. Nonetheless, more refined statistical analysis allowed for the detection of a subset of differentially expressed genes that could predict the immunoglobulin-mutational status of CLL cells with high accuracy.[8,9] Collectively, these data suggest that CLL may be viewed as a single disease with 2 common variants that differ with regard to their immunoglobulin-mutational status and clinical course.[9]

Given the dramatic differences in the clinical behavior of CD38+ versus CD38− CLL patients, we compared their gene expression profiles using oligonucleotide-based DNA chip microarrays representative of approximately 5600 genes. In line with the aforementioned results, our data show that B-CLLs display a gene-expression profile largely independent of CD38

Table 1. Clinical characteristics of study patients with B-CLL

Variable	All patients	CD38+ patients	CD38− patients	P
No. patients (%)	76 (100)†	25 (36)	45 (64)	—
Median age, y	67	66	68	.309*
Male sex (%)	55 of 76 (72)	19 of 25 (76)	31 of 45 (69)	.528†
Disease stage				
Binet A (%)	31 of 75 (41)	4 of 25 (16)	24 of 44 (55)	.002†
Binet B (%)	21 of 75 (28)	10 of 25 (40)	10 of 44 (22)	—
Binet C (%)	23 of 75 (31)	11 of 25 (44)	10 of 44 (22)	—
White blood cell count, $\times 10^{-3}/\mu$L				
N = 75	77.5 ± 77.7	41.0 ± 76.0	60.9 ± 82.6	.210*
Hemoglobin level, g/dL; N = 69	12.8 ± 2.4	12.6 ± 2.6	13.6 ± 2.3	.135*
Platelet count, $\times 10^{-3}/\mu$L; N = 68	154.0 ± 74.0	131.0 ± 59.0	161.5 ± 79.6	.124*
Bone marrow histology, %				
infiltration; N = 44	54.9 ± 30.5	80.0 ± 25.9	35.0 ± 24.1	.001*
Thymidine kinase, IU/L; N = 47	19.7 ± 27.2	19.9 ± 39.7	6.0 ± 10.0	.002*
β_2-microglobulin, mg/L; N = 64	3.00 ± 1.9	2.8 ± 2.3	1.9 ± 1.8	.061*
Lactate dehydrogenase level,				
IU/L; N = 63	237.9 ± 67.7	254.0 ± 75.3	215.0 ± 55.0	.020*

A CLL population was considered CD38+ when more than 20% of the gated population (CD19+CD5+ cells) expressed the CD38 antigen.[12,14]

*Calculated using Wilcoxon rank sum test; †calculated using χ^2 test.

†In 6 patients the CD38 expression status was not assessed.

Data are mean ± SD. Because of rounding, percentages do not always add up to 100. — indicates not applicable.

expression. Nonetheless, unsupervised cluster analysis[10] revealed the existence of 2 major subgroups that differed significantly with regard to known clinical prognostic factors including hemoglobin levels, chemotherapy requirements, and progression-free survival.

Patients, materials, and methods

Patients and isolation of CLL cells

Between November 2000 and February 2002, 76 patients with chronic lymphocytic leukemia were enrolled in this retrospective study and were evaluated for several biologic and clinical characteristics: age, sex, Binet stage, white blood cell count, hemoglobin level, platelet count, lactate dehydrogenase level, thymidine kinase level, treatment history, and time from diagnosis to first treatment. The study was approved by the institutional review board of University Hospital Essen; informed consent was obtained from each patient according to the Declaration of Helsinki. In each patient, morphologic diagnosis of B-CLL was confirmed by flow cytometry[11,12] that revealed a typical CD19+CD20+CD5+CD23+ immunoglobulin light chain (κ or λ light chain)-restricted immunophenotype. Unique patient number (UPN) 55 was excluded from the study because the diagnosis of B-CLL could not be confirmed. Patient characteristics are shown in Table 1. Heparinized whole peripheral blood (PB) samples were usually obtained during routine follow-up visits to our institutions with informed consent, according to institutional guidelines. Indications for treatment were based on standard criteria.[13] Thirty-five (46%) of 76 patients had previously received chemotherapy. First-line therapy consisted of chlorambucil in 31 (89%) of 35 or fludarabine in 4 (11%) of 35 patients. Peripheral blood mononuclear cells (PBMNCs) were isolated by Ficoll-Hypaque (Pharmacia, Erlangen, Germany) density centrifugation and washed in Iscove modified Dulbecco medium (Gibco BRL, Karlsruhe, Germany). Proportions of CD19+CD5+ B-CLL cells, CD3+ T cells, and CD14+ monocytes were 89.0% ± 6.0%, 5.3% ± 4.2%, and 1.1% ± 1.0% (mean ± SD), respectively. There were no significant differences in the cellular composition of the PBMNCs between the 2 major subgroups identified by hierarchical cluster analysis (Figure 2; $P > .05$). RNA from 1 to 2×10^8 PBMNCs was extracted, purified using the RNeasy midi kit (Qiagen, Hilden, Germany), and quantified spectrophotometrically. PBMNC aliquots of each patient were viably frozen in fetal calf serum (FCS; Greiner, Limburg, Germany) containing 10% dimethyl sulfoxide (DMSO; Sigma, Deisenhofen, Germany) and were stored in liquid nitrogen.

Cell surface staining and flow cytometry

For the determination of CD38 expression, fresh heparinized PB samples were prepared for flow cytometry by ammonium chloride erythrocyte lysis (Ortho-mune Lysing Reagent; Ortho Diagnostic Systems, Raritan, NJ). Immunophenotype was characterized using a standard 3-color flow cytometry approach as previously described.[12] Antibodies were purchased from DAKO (Glostrup, Denmark; CD19, CD10, IgM, κ and λ light chains), Immunotech (Marseilles, France; CD5) and Becton Dickinson (Heidelberg, Germany; CD38, CD4, CD8, CD3). Negative isotype–matched controls (Becton Dickinson) were used to define the threshold line separating surface marker positive and negative cells such that less than 1% of isotype-positive cells were present to the right of the line. A CLL population was considered CD38+ when more than 20% of the gated population (CD19+CD5+) expressed it.[12] The same method of sample preparation and 3-color staining was used throughout the entire study period. Samples were analyzed on a FACScan flow cytometer (Becton Dickinson) using CellQuest software (Becton Dickinson).

Oligonucleotide microarray analysis

For first-strand cDNA synthesis, 9 μL (13.5 μg) total RNA was mixed with 1 μL mixture of 3 polyadenylated control RNAs, 1 μL 100 μM T7-oligo-d(T)24 primer [5′-GGCCAGTGAATTGTAATACGACTCACT ATAG
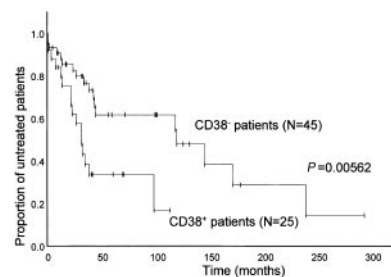

Figure 1. Probability of disease progression, as indicated by treatment-free interval. Kaplan-Meier plot comparing time periods from diagnosis to initiation of chemotherapy in CD38+ CLL patients to CD38− CLL patients. Statistical analysis was performed using the log-rank test. Vertical lines represent follow-up times of individual patients.
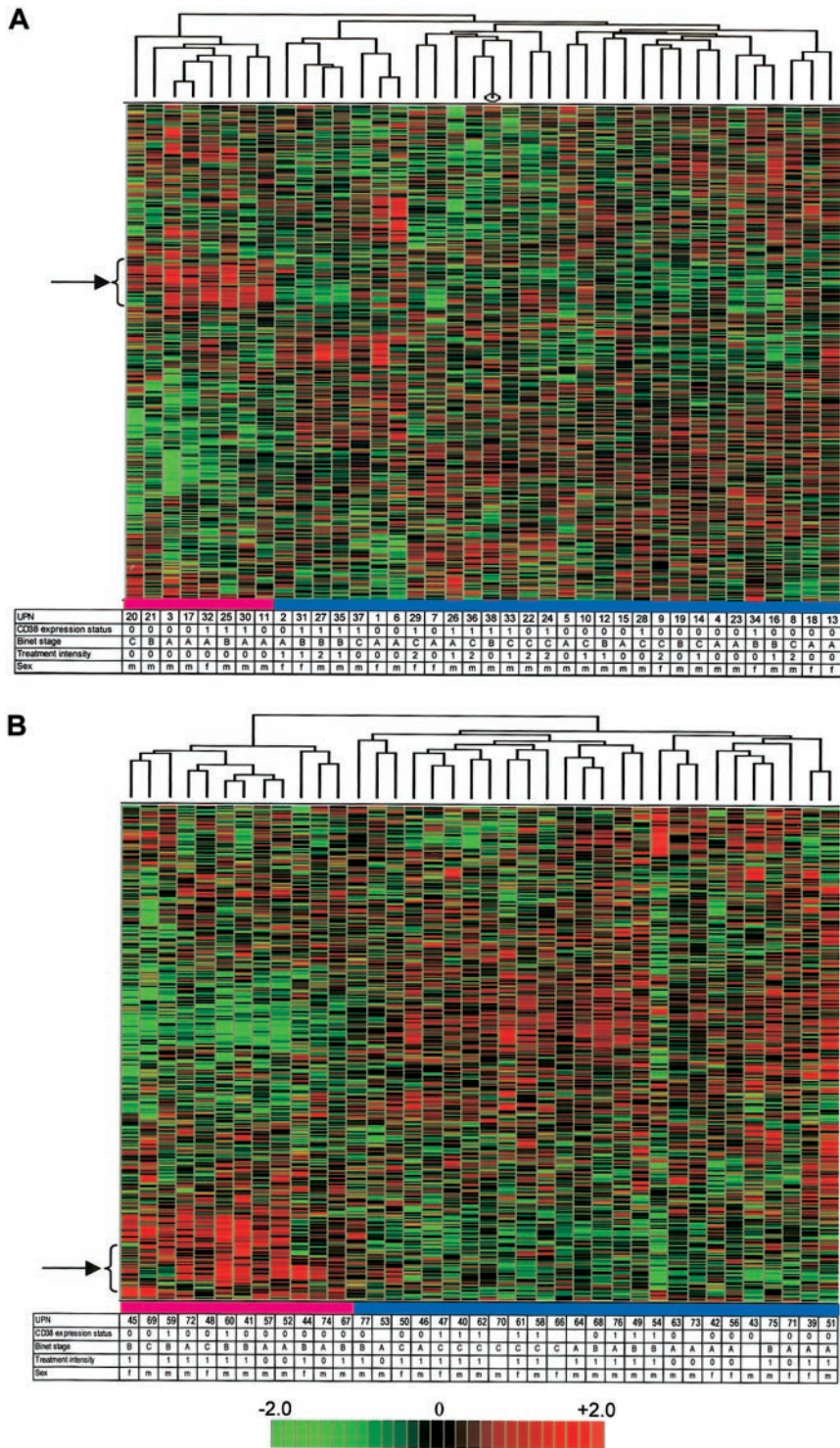
Figure 2. Hierarchical clustering. Color matrices show the hierarchical clustering of 2343 (A, UPN1-38) and 2769 (B, UPN 39-77) selected genes (see "Materials and methods") of gene expression data obtained from 76 B-CLL patients. The hierarchical clustering algorithm used is based on the average-linkage method. Red areas indicate increased expression, and green areas indicate decreased expression. Each column represents a single B-CLL sample and each row a single gene. Tables beneath the heat maps detail the individual patients (UPN) and some of their clinical characteristics. Treatment intensity, 0; no previous treatment, 1; patients requiring minimal treatment, defined as less than 6 months of therapy in the years of follow-up, 2; patients requiring continuous chemotherapy or chemotherapy with 2 or more agents or regimens, 1. Dendrograms at the top show the degree to which each B-CLL patient is related to the others with respect to gene expression. Note the separation into 2 groups comprising 8 and 30 (A) and 12 and 26 patients (B), respectively. Clustering within the leftmost branches of the dendrograms was caused, in part, by the high expression of a large number of ribosomal and translation-associated genes (see arrows at the left side of the heat maps).

GGAGGCGG-(dT24)-3′], incubated at 70°C for 10 minutes and put on ice. Next, 4 μL of 5× first-strand buffer, 2 μL 0.1 M dithiothreitol (DTT), and 1 μL 10 mM dNTPs were added, and the reaction was preincubated at 42°C for 2 minutes. Then, 2 μL (200 units) Superscript II (Life Technologies, Karlsruhe, Germany) was added, and incubation was continued at 42°C for 1 hour.

For second-strand synthesis, 30 μL 5× second-strand buffer, 91 μL RNase-free water, 3 μL 10 mM dNTPs, 4 μL (40 U) *Escherichia coli* DNA polymerase I (Life Technologies), 1 μL (12 U) *E coli* DNA ligase (TaKaRa, Gennevilliers, France), and 1 μL (2 U) RNase H (TaKaRa) were added, and the reaction was incubated at 16°C for 2 hours. Then 2.5 μL (10 U) T4 DNA

polymerase I (TaKaRa) was added at 16°C for 5 minutes. The reaction was stopped by the addition of 10 μL 0.5 M EDTA (ethylenediaminetetraacetic acid), double-stranded (ds) cDNA was extracted with phenol/chloroform, and the aqueous phase was recovered by phase-lock gel separation (Eppendorf, Hamburg, Germany). After precipitation, the cDNA was restored in 12 μL RNase-free water.

Five microliters ds cDNA was used to synthesize biotinylated cRNA using the BioArray High Yield RNA Transcript Labeling Kit (Enzo Diagnostics, NY). Labeled cRNA was purified using the RNeasy mini kit (Qiagen, Hilden, Germany). Fragmentation of cRNA, hybridization to HuGeneFL microarrays (Affymetrix, Santa Clara, CA), and washing,

Table 2. Comparison of clinical and laboratory data between the two major patient subgroups defined by hierarchical cluster analysis

Variable	Ribosomal cluster	Other patients	P
No. patients (%)	20 (26.3)	56 (73.7)	—
Median age, y	64.5	69.5	.133*
Disease stage			
Binet A (%)	9 of 20 (45.0)	22 of 55 (40.0)	.315†
Binet B (%)	8 of 20 (40.0)	13 of 55 (24.0)	—
Binet C (%)	3 of 20 (15.0)	20 of 55 (36.0)	—
White blood cell count, $\times 10^{-3}/\mu L$	61.4 ± 49.9	41.7 ± 85.9	.854*
Hemoglobin level, g/dL	14.6 ± 2.0	12.0 ± 2.2	.0005*
Platelet count, $\times 10^{-3}/\mu L$	171.0 ± 40.6	131.0 ± 83.2	.158*
Bone marrow histology, %			
infiltration	45.0 ± 24.5	60.0 ± 31.8	.473*
CD38+ patients (%)	5 of 20 (25.0)	20 of 50 (40.0)	.280†
Thymidine kinase level, IU/L	7.7 ± 10.1	13.9 ± 30.6	.373*
β_2-microglobulin, mg/L	1.84 ± 1.86	2.76 ± 2.0	.09*
Lactate dehydrogenase			
level, IU/L	205.0 ± 65.9	231.5 ± 68.4	.233*

See also Figure 2.
*Calculated using Wilcoxon rank sum test.
†Calculated using Fisher exact test. Data are mean ± SD.
— indicates not applicable.

staining, and scanning of the arrays in a GeneArray scanner (Agilent, Palo Alto, CA) were performed as recommended in the Affymetrix Gene Expression Analysis Technical Manual. Signal intensities (MAS5 signal) and detection calls for statistical analysis and hierarchical clustering were determined using the Microarray Suite (MAS 5.0) software (Affymetrix, Santa Clara, CA). Scaling across all probe sets of a given array to an average intensity of 1000 U was included to compensate for variations in the amount and quality of the cRNA samples and other experimental variables.

Data processing and hierarchical clustering

For the hierarchical clustering shown in Figure 2, only genes recognized as present by the Affymetrix algorithm in at least one third of the profiles were selected. Gene expression data were ln transformed, normalized to have a mean of 0 and an SD of 1, and subjected to the average linkage clustering method (Array Explorer; Spotfire, Somerville, MA) using correlation (centered) as a similarity measure.

Statistical analysis

Progression-free survival times were measured from the time of diagnosis, plotted by the Kaplan-Meier method, and compared using the log-rank test. Comparison of clinical and laboratory parameters between patient subgroups was performed using the Wilcoxon rank sum test for metric data, Fisher exact test, and the χ^2 test for categorical data. The Cox proportional hazards model was used for multivariate analysis on progression-free survival.

Results

CD38 expression in the B-CLL study group

We evaluated the surface expression of CD38 in our B-CLL study population using a 3-color flow cytometry approach with directly conjugated monoclonal antibodies.[12,14] In accordance with current convention, a given leukemic population was considered positive for CD38 when 20% or more of the B-CLL cells expressed the membrane marker.[12,14] Based on this cutoff value, 25 (36%) patients were defined as CD38+ and 45 (64%) patients as CD38−, respectively. Comparison of clinical and laboratory parameters among the 2 groups is shown in Table 1. Notably, significant differences were found for thymidine kinase, leukemic bone marrow infiltration, Binet stage, lactate dehydrogenase (LDH) serum levels ($P < .05$), and treatment-free survival (Figure 1) confirming our own previous work[12] and that of others.[5,14-16]

Comparison of gene expression profiles among CD38+ and CD38− B-CLL patients

RNA extracted from the PBMNCs of 76 B-CLL patients was converted to labeled cRNA and hybridized to HuGeneFL Affymetrix oligonucleotide chips representative of approximately 5600 genes. The same methods of cell purification, RNA preparation, and chip hybridization were used throughout the entire study period. Two different batches of HuGeneFL Affymetrix oligonucleotide chips were used for 2 independent patient series comprising UPNs 1-38 and UPNs 39-77, respectively. Although no statistically

Table 3. Differentially expressed genes in CD38+ vs CD38− B-CLL patients

Accession no.	Gene description	UPN 1-38				UPN 39-77			
		Median CD38−	Median CD38+	Ratio CD38+/CD38−	P	Median CD38−	Median CD38+	Ratio CD38+/CD38−	P
D84276	CD38	79	212	2.7	.0215	67	185	2.8	.0111
D87119	Osteoblast protein	393	702	1.8	.0215	406	973	2.4	.0047
L29219	CDC-like kinase	1501	939	0.6	.0183	1662	997	0.6	.0296
J04027	ATPase	2961	1766	0.6	.0251	3313	1807	0.5	.0098
L05148	ZAP-70 kinase	1339	3177	2.4	.0001	1594	3032	1.9	.0328
L08895	MEF2C	1803	875	0.5	.0023	1039	589	0.6	.0194
M62302	GDF-1	128	243	1.9	.0423	66	139	2.1	.0296
M62994	Thyroid autoantigen	757	1270	1.7	.0423	698	1351	1.9	.0111
M97963	ISGF-3	995	1544	1.6	.0183	2159	4201	1.9	.0267
S626696	CD21	138	258	1.9	.0366	189	425	2.2	.0296
X16983	Integrin α4, CD49d	186	384	2.1	.0023	215	361	1.7	.0174
X63741	Pilot mRNA	278	77	0.3	.0050	252	49	0.2	.0019
X84373	RIP140	496	176	0.4	.0102	393	111	0.3	.0139
X89986	Apoptotic inducer	99	251	2.5	.0232	125	375	3.0	.0098

For comparative analysis of CD38+ and CD38− B-CLL patients, only genes called present by the Affymetrix algorithm in at least one third of the profiles were selected (approximately 2500 in both series). Differentially regulated genes were identified by comparing the median signal intensities of the two groups using the Wilcoxon rank sum test at a significance level of .05 and defining a cut-off fold change of ± 1.5 (approximately 75 genes in each series [UPNs 1-38 and UPNs 39-77, respectively]). Fourteen of these CD38 distinction candidate genes were found to fulfill the statistical selection criteria in both series and are shown in the table. Data are median signal intensities of the two groups.

Table 4. Ribosomal cluster identified by hierarchical cluster analysis

Accession no.	Gene description
M60784	snRNP-specific protein A
D28589	KIAA00167
M30448	Casein kinase II-β
X14046	CD37
X55733	Initiation factor 4B, EIF-4B
X73358	hAES-1 m
U04241	Homolog of *Drosophila* enhancer of split
D87451	KIAA0262
U07705	p58 GTA protein kinase
D50310	*Cyclin I*
X78136	hnRNP-E2
M84711	Ribosomal protein S3A, component of S40 subunit
X51466	*Elongation factor 2, EF-2*
X57959	*Ribosomal protein L7*
U25789	Ribosomal protein L21
X55954	HL23 ribosomal protein
X53505	*Ribosomal protein S12*
X03689	Elongation factor TU (N-terminus)
M31520	*Ribosomal protein S24*
X16064	Translationally controlled tumor protein
X62691	Ribosomal protein S24
X56932	23-kd highly basic protein
Z21507	*Elongation factor-1 δ, EF-1δ*
X69391	*Ribosomal protein L6*
Z25749	*Ribosomal protein S7*
X69150	*Ribosomal protein S18*
X60489	Elongation factor-1β, EF-1β
Z26876	*Ribosomal protein L38*
J03592	ADP/ATP translocase
D49824	HLA-B
Z12962	*Homolog to yeast ribosomal protein L41*
J04617	Elongation factor-1 EF-1α
L04483	*Ribosomal protein S21 (RPS21)*
M17886	Acidic ribosomal phosphoprotein P1
M36072	*Ribosomal protein L7a (surf3)*
D89667	c-myc binding protein
AB002533	*qip1, importin α3*
14973	*Ribosomal protein S29*
M24194	MHC protein homologous to chicken B-complex protein
D13413	Tumor-associated 120 kd
X06617	*Ribosomal protein S11*
X79234	*Ribosomal protein L11*
X00274	HLA-DRα
J00105	β2-microglobulin
U14972	*Ribosomal protein S10*
U14970	*Ribosomal protein S5*
U14971	*Ribosomal protein S9*
U58682	Ribosomal protein S28
U14969	*Ribosomal protein L28*
X55715	*40S ribosomal protein s3*
X53777	L23 m for putative ribosomal protein
M14199	*Laminin receptor (2H5 epitope)*
U43901	*Laminin receptor/p40 ribosome-associated protein*
Z28407	*Ribosomal protein L8*
M64716	*Ribosomal protein S25*
D87735	*Ribosomal protein L14*
X17206	Ribosomal protein S2
M17885	Acidic ribosomal phosphoprotein P0
D23660	*Ribosomal protein*
S79522	*Ubiquitin carboxyl extension protein*
L06499	*Ribosomal protein L37a (RPL37A)*
L06505	*Ribosomal protein L12*
M77232	*Ribosomal protein S6*
D14530	Homolog of yeast ribosomal protein S28
X67247	*Ribosomal protein S8*

Table 4. Ribosomal cluster identified by hierarchical cluster analysis (continued)

Accession no.	Gene description
M55409	Elongation factor 1-γ, EF-1γ
M81757	*S19 ribosomal protein*
U12465	*Ribosomal protein L35*
U12404	*Ribosomal protein L10a, S60 subunit component*
X80822	*Ribosomal L18A*
L11566	*Ribosomal protein L18 (RPL18)*
X73460	*Ribosomal protein L3*
X63527	*Ribosomal protein L19*
M60854	*Ribosomal protein S16*
X03342	*Ribosomal protein L32*
M13934	Ribosomal protein S14
X15940	*Ribosomal protein L31*
L19527	*Ribosomal protein L27 (RPL27)*
M18000	*Ribosomal protein S17*
U09953	*Ribosomal protein L9*
X64707	*Ribosomal protein L13*
U70439	Silver-stainable protein SSP29
U94855	Translation initiation factor 3, EIF3-P47
Z49148	*Ribosomal protein L29*
X74104	TRAP β-subunit
U06155	Chromosome 1q subtelomeric
L13848	Helicase A, DDX9: DEAD/H
U85611	DNA-PK interaction protein (KIP)

Gene names in italics denote genes that were present in the ribosomal clusters of both series UPN 1-38 (Figure 2A) and UPN 39-77 (Figure 2B).

ADP indicates adenosine diphosphate; ATP, adenosine triphosphate; MHC, major histocompatibility complex.

significant differences were observed between these 2 patient series in terms of sample purity and clinical characteristics (Table 1), preliminary data analysis revealed substantially stronger overall hybridization of the labeled cRNA probes in the second series of experiments than in patient samples UPNs 1-38, possibly because of differences between the 2 chip batches used. For this reason, expression data analysis was performed separately for patients UPNs 1-38 and UPNs 39-77.

Resultant gene expression profiles were analyzed using 2 independent approaches—unsupervised hierarchical clustering,[10] which can identify distinct subgroups that have not been classified beforehand, and the Wilcoxon rank sum test, which is more suitable for comparing subtle differences in gene expression values between 2 predefined groups (eg, CD38+ and CD38− B-CLL patients).

Hierarchical cluster analysis (Figure 2A-B) using an average-linkage algorithm[10] revealed the existence of 2 major subgroups comprising 8 and 30 patients in the first patient series (UPNs 1-38) and 12 and 26 patients in the second set of experiments (UPNs 39-77). However, the incidence of CD38+ patients was comparable in the 2 groups ($P > .05$; Figure 2, Table 2), suggesting a common gene expression profile independent of CD38 expression status for most genes represented on the chip.

For comparative analysis of CD38+ and CD38− B-CLL patients, the following statistical criteria were applied to the array data. As for hierarchical clustering, only those genes called present by the Affymetrix algorithm in at least one third of the profiles were selected (2343 genes in UPNs 1-38 and 2769 genes in UPNs 39-77). Differentially regulated genes were identified by comparing the median signal intensities of the 2 groups and defining a cut-off fold change of ± 1.5. Resultant genes were then further analyzed using the Wilcoxon rank sum test at a significance level of

Table 5. Comparison of treatment histories among the two patient subgroups defined by hierarchical cluster analysis

	Ribosomal cluster	Other patients
No treatment (%)	14 of 20 (70)	23 of 54 (43)
Minimal treatment (%)	5 of 20 (25)	14 of 54 (26)
Continuous chemotherapy or chemotherapy with 2 or more agents or regimens (%)	1 of 20 (5)	17 of 54 (31)
		$P = .041$

Minimal treatment was defined as less than 6 months of therapy in the years of follow-up.[5] P was calculated using χ^2 analysis.

See also Figure 2.

.05, yielding approximately 75 differentially expressed genes in each series. Fourteen of these so-called CD38 distinction genes were found to fulfill the statistical selection criteria in both series and are shown in Table 3.

Characterization of patient subgroups based on gene expression analysis

We next tried to further characterize the previously described gene-expression subgroups identified by hierarchical cluster analysis (Figure 2A-B). Clustering within the leftmost branch of the dendrograms (Figure 2A-B) was caused, in part, by the high expression of a large number of ribosomal and translation-associated genes (ribosomal cluster; see arrows in Figure 2 and Table 4) in both patient series.

For further statistical analysis, the clinical and laboratory data of the 2 series were pooled—that is, 20 patients in the ribosomal cluster were compared with the remaining 56 patients. Treatment histories of the 2 subgroups differed significantly in that the patients with a high expression of ribosomal proteins required less chemotherapy than the remaining patients (Table 5; $P = .041$). Furthermore, we found highly significant differences in disease progression, as indicated by the treatment-free interval (Figure 3). The mean treatment-free interval was longer in patients with a high expression of ribosomal proteins than in the other patients ($P = .0172$; Figure 3). Comparisons of further clinical and laboratory parameters between the 2 groups are shown in Table 2. Finally, we investigated whether the expression of ribosomal and translation-associated genes could further refine the clinical relevance of CD38 expression status (Figure 4). Figure 4A shows that CD38$^+$ patients with high expression of ribosomal genes are characterized by significantly longer progression-free survival than are the remaining CD38$^+$ patients. This observation suggests that a combination of these 2 parameters may increase the prognostic power of either of the 2 factors. However, the subgroup analysis shown in Figure 4A is based on 25 patients and thus must be confirmed in a larger patient cohort.

Univariate analysis of risk factors

Univariate Cox regression analysis was used to assess associations between progression-free survival time and potential risk factors (Table 6). Hemoglobin level, platelet count, CD38 expression, β_2-microglobulin serum concentration, LDH serum activity, and expression of ribosomal genes (ribosomal cluster) were identified as significant factors influencing progression-free survival.

Multivariate analysis

The following patient characteristics, found to impact significantly on treatment-free survival in univariate analysis, were included in

the multivariate Cox regression model: hemoglobin concentration, platelet count, CD38 expression, β_2-microglobulin, LDH serum activity, and expression of ribosomal genes (Table 7). In multivariate analysis, only hemoglobin concentration influenced progression-free survival.

Discussion

It is now well established that CD38 expression is an important prognostic marker in B-CLL.[5,12,14-17] In a recently published retrospective study,[12] we showed that CD38$^+$ CLL patients were characterized by an unfavorable clinical course with advanced disease stage, poor responsiveness to chemotherapy, short time to initiation of first treatment, and shorter survival. In contrast, the CD38$^-$ group required minimal or no treatment, remained treatment free for a longer time period, and had prolonged survival. To identify molecular differences that may be responsible for the differential clinical course of these 2 subgroups, we compared gene expression profiles of CD38$^+$ patients with those of CD38$^-$ patients using oligonucleotide-based microarray technology.

Unsupervised clustering (Figure 2) revealed a common pattern of expression of approximately 5600 genes, independent of the expression of CD38. Detailed statistical analysis identified a set of 14 genes differentially expressed in CD38$^+$ versus CD38$^-$ subtypes. These results are consistent with 2 recent studies comparing gene expression profiles of CLL patients with hypermutated immunoglobulin variable region (immunoglobulin V_H) sequences and a favorable prognosis with that of immunoglobulin V_H–unmutated patients with a more aggressive course of disease.[7,8] Both studies were unable to distinguish the 2 subgroups using cluster analysis and detected only a restricted number of differentially expressed genes using more refined statistical tests. Interestingly, patients with unmutated immunoglobulin V_H genes have been previously found to exhibit a higher percentage of CD38 expression than mutated B-CLL clones.[5,16] In this regard, it is important to note that ZAP-70, found by Rosenwald et al[8,9] to be among the most differentially expressed genes, was also up-regulated in our CD38$^+$ patient cohort (Table 3), indicating an overlap in the molecular characteristics of the 2 subgroups. ZAP-70, a critical kinase involved in T-cell antigen receptor signaling, has only recently been recognized to be expressed in B cells.[8,18] Intriguingly, recent work by Deaglio et al[19] has demonstrated that ZAP-70 is tyrosine phosphorylated in response to antibody-mediated CD38 stimulation in natural killer (NK) cells, raising the possibility
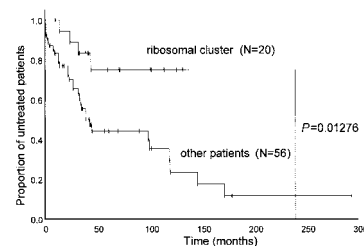

Figure 3. Probability of disease progression, as indicated by the treatment-free interval. Kaplan-Meier plots compared time periods from diagnosis to initiation of chemotherapy in the 2 subgroups identified by unsupervised cluster analysis (Figure 2). Statistical analysis was performed using the log-rank test. Mean follow-up time (observation time from diagnosis) was 72.9 months for both subgroups ($P = .54$).
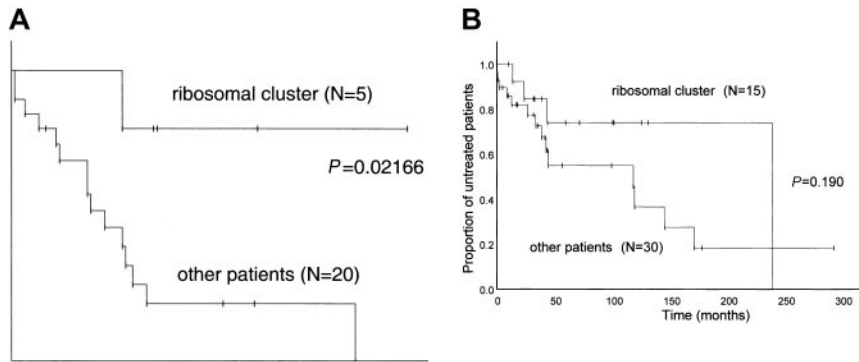
Figure 4. Disease progression. Probability of disease progression was indicated by the treatment-free interval in CD38+ (A) and CD38− (B) B-CLL patients subgrouped according to the expression of ribosomal and translation-associated genes identified by unsupervised cluster analysis (Figure 2). Kaplan-Meier plots compared time periods from diagnosis to initiation of chemotherapy. Statistical analysis was performed using the log-rank test.

that it may play a role in CD38 signaling in B-CLL. Further support for the coordinate expression of ZAP-70 and CD38 in B-CLL comes from a recent study by Chen et al[18] showing that all CD38+ CLL samples analyzed also expressed high levels of ZAP-70 protein; however, ZAP-70 expression was not restricted to CD38+ patients but was also detected in some CD38− patients with unmutated immunoglobulin V_H genes.

Another gene of potential functional importance found to be overexpressed in the CD38+ group is the $\alpha 4$ subunit of the VLA$_4$-integrin (CD49d; Table 3), which plays an important role in cell adhesion to the extracellular matrix molecule fibronectin. Importantly, adhesion to fibronectin through CD49d has recently been shown to protect B-CLL cells from apoptosis induced by serum deprivation[20] and fludarabine treatment in vitro.[21] These observations are consistent with a flow cytometry study showing that CLL cells isolated from early-stage patients (Rai 0-II) exhibit significantly lower CD49d expression levels than CLL patients with advanced disease (Rai III-IV).[22]

Despite the comparatively small differences in gene expression patterns of CD38+ versus CD38− CLL patients, hierarchical clustering of the samples and gene expression levels within the samples led to the separation of 2 major subgroups composed of 20 and 56 patients, respectively (Figure 2). Clustering to the smaller group was attributed, at least in part, to the coordinate high expression of a large number of ribosomal and translation-associated genes (Table 4).

More important, we found that the clinical outcome for patients in the 2 subgroups was strikingly different. Patients with a high expression of translation-associated genes were characterized by a more favorable clinical course with significantly longer progression-free survival and fewer chemotherapy requirements than the remaining patients (Figure 3). Furthermore, the 2 patient subgroups

differed with regard to a panel of known clinical prognostic factors, including peripheral blood hemoglobin levels (Table 2).

To our knowledge, this is the first report showing that unsupervised cluster analysis can identify molecular B-CLL subtypes that differ with regard to the clinical course of the disease. In particular, the ribosomal cluster described here has not been noted in prior gene array studies comparing the gene expression profiles of immunoglobulin V_H mutated with immunoglobulin V_H unmutated CLL patients.[7,8] This discrepancy may be explained at least in part by differences in the study designs, the most important of which is probably the comparatively small number of patients included in these studies (34 and 37 patients in the series reported by Klein et al[7] and Rosenwald et al,[8] respectively). In addition, both studies used immunomagnetically enriched B-CLL cells with a purity exceeding 95%, whereas in our series CLL cells were isolated using density centrifugation yielding only a mean purity of 89% CD19+CD5+ cells. Thus, we cannot exclude the possibility that accessory cells might have contributed to the observed gene expression differences, although the cellular composition of the tumor samples in the 2 major subgroups, defined by differential expression of ribosomal and translation-associated genes, was similar as determined by flow cytometry analysis (see "Patients, materials, and methods").

Overexpression of translation-associated genes in tumors has been previously noted and may reflect a higher metabolic and proliferation rate in the malignant cell population.[23] However, it is now well established that translation factors also play an important role in the regulation of cell death.[24,25] In this context it is important to note that high expression of ribosomal proteins S3A, S29, and elongation factor-1α (Table 4) observed here has been shown to accelerate the apoptotic rate in certain cell types.[24,25] By contrast, high expression levels of elongation factor-2α (Table 4) have previously been found to positively correlate with a high proliferation rate and biologic aggressiveness in non-Hodgkin lymphomas.[26] These results, together with the findings of the recent literature,[27,28] suggest that diverse cellular responses to alterations in translation-associated proteins strongly depend on the expression patterns of the individual factors involved.

Interestingly, high coordinate expression of a large group of ribosomal genes was reported for ovarian tumors that were histologically well differentiated compared with the more poorly differentiated tumors[28] in the same series. Using hierarchical clustering, these authors found that rapidly growing tumor cell lines and the most poorly differentiated ovarian tumors grouped together and exhibited a relative underexpression of ribosomal genes, despite their presumed high metabolic rates. These observations, in combination with our data, raise the possibility that a high

Table 6. Univariate analysis of prognostic variables for progression-free survival (time to first treatment)

Variable	N	P	Relative risk
Sex	76	.9124	0.958
Age	76	.7809	0.995
White blood cell count	75	.3165	1.002
Hemoglobin concentration	69	.0001	0.719
Platelet count	68	.0202	0.992
Bone marrow infiltration	44	.92	1.014
CD38 expression	70	.0295	2.298
Thymidine kinase level	47	.0917	1.012
β_2-microglobulin concentration	64	.0009	1.278
Lactate dehydrogenase activity	63	.013	1.006
Ribosomal cluster	76	.0207	3.069

P was calculated using Cox regression analysis.

Table 7. Multivariate Cox proportional hazards model of the study patients

Variable	Parameter estimate	Standard error	χ^2	P	Hazard ratio
Hemoglobin concentration	0.28789	0.08793	10.7197	.0011	0.75
CD38 expression	0.63252	0.45459	1.9361	.1641	1.882
Lactate dehydrogenase activity	0.00482	0.00315	2.3456	.1256	1.005

expression of ribosomal proteins may be correlated with a less aggressive clinical course in some tumor entities.

In conclusion, our results suggest that oligonucleotide microarray analysis can detect molecular B-CLL subtypes that differ with regard to the clinical course of the disease. A larger study is warranted to confirm our results and to further investigate the potential role of the translational apparatus in the pathogenesis of B-CLL.

Acknowledgments

References

1. Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. Clinical staging of chronic lymphocytic leukemia. Blood. 1975;46: 219-234.

2. Binet JL, Auquier A, Dighiero G, et al. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. Cancer.1981;48:198-206.

3. Binet JL. Prognostic factors in chronic lymphocytic leukaemia. Haematologica. 1999;84:96-107.

4. Zwiebel JA, Cheson BD. Chronic lymphocytic leukemia: staging and prognostic factors. Semin Oncol. 1998;25:42-59.

5. Damle R, Wasil T, Fais F, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood. 1999;94:1840-1847.

6. Hamblin T, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood. 1999;94:1848-1854.

7. Klein U, Tu Y, Stolovitzky G, et al. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. J Exp Med. 2001;194:1625-1638.

8. Rosenwald A, Alizadeh A, Widhopf G, et al. Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. J Exp Med. 2001;194:1639-1647.

9. Rosenwald A, Staudt LM. Clinical translation of gene expression profiling in lymphomas and leukemias. Semin Oncol. 2002;29:258-263.

10. Eisen MB, Spellman P, Brown P, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A. 1998;95: 14863-14868.

11. Dürig J, Schmücker U, Dührsen U. Differential expression of chemokine receptors in B cell malignancies. Leukemia. 2001;15:752-756.

12. Dürig J, Naschar M, Schmücker U, et al. CD38 expression is an important prognostic marker in chronic lymphocytic leukemia. Leukemia. 2002; 16:30-35.

13. Cheson BD, Bennett JM, Grever M, et al. National Cancer Institute-Sponsored Working Group guidelines for chronic lymphocytic leukemia: revised guidelines for diagnosis and treatment. Blood. 1996;87: 4990-4997.

14. Ibrahim S, Keating M, Do K, et al. CD38 expression as an important prognostic factor in B-cell chronic lymphocytic leukemia. Blood. 2001;98: 181-186.

15. Hamblin T, Orchard J, Ibbotson R, et al. CD38 expression and immunoglobulin variable region mutations are independent prognostic variables in chronic lymphocytic leukemia, but CD38 expression may vary during the course of the disease. Blood. 2002;99:1023-1029.

16. Jelinek D, Tschumper R, Geyer S, et al. Analysis of clonal B-cell CD38 and immunoglobulin variable region sequence status in relation to clinical outcome for B-chronic lymphocytic leukaemia. Br J Haematol. 2001;115:854-861.

17. Del Poeta G, Maurillo L, Venditti A, et al. Clinical significance of CD38 expression in chronic lymphocytic leukemia. Blood. 2001;98:2633-2639.

18. Chen L, Widhopf G, Huynh L, et al. Expression of ZAP-70 is associated with increased B cell receptor signaling in chronic lymphocytic leukemia. Blood. 2002;100:4609-4614.

19. Deaglio S, Zubiaur M, Gregorini A, et al. Human CD38 and CD16 are functionally dependent and physically associated in natural killer cells. Blood. 2002;99:2490-2498.

20. Garcia-Gila M, Lopez-Martin EM, Garcia-Pardo A. Adhesion to fibronectin via α 4 integrin (CD49d) protects B cells from apoptosis induced by serum deprivation but not via IgM or Fas/Apo-1 receptors. Clin Exp Immunol. 2002;127: 455-466.

21. De la Fuente MT, Casanova B, Moyano JV, et al. Engagement of $\alpha4\beta1$ integrin by fibronectin induces in vitro resistance of B chronic lymphocytic leukemia cells to fludarabine. J Leukoc Biol. 2002;71:495-502.

22. Eksioglu-Demiralp E, Alpdogan O, Aktan M, et al. Variable expression of CD49d antigen in B cell chronic lymphocytic leukemia is related to disease stages. Leukemia. 1996;10:1331-1339.

23. Dua K, Williams TM, Beretta L. Translational control of the proteome: relevance to cancer. Proteomics. 2001;1:1191-1199.

24. Chen E, Proestou G, Bourbeau D, Wang E. Rapid up-regulation of peptide elongation factor EF-1α protein levels is an immediate early event during oxidative stress-induced apoptosis. Exp Cell Res. 2000;259:140-148.

25. Duttaroy A, Bourbeau D, Wang XL, Wang E. Apoptosis rate can be accelerated or decelerated by overexpression or reduction of the level of elongation factor-1 alpha. Exp Cell Res. 1998;238: 168-176.

26. Wang S, Rosenwald IB, Hutzler MJ, et al. Expression of the eukaryotic translation initiation factors 4E and 2α in non-Hodgkin's lymphomas. Am J Pathol. 1999;155:247-255.

27. Alon U, Barkai N, Notterman DA, et al. Broad pattern of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc Natl Acad Sci U S A. 1999;96:6745-6750.

28. Welsh JB, Zarrinkar PP, Sapinoso LM, et al. Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. Proc Natl Acad Sci U S A. 2001;98:1176-1181.