Distinct gene expression profiles in subsets of chronic lymphocytic leukemia expressing stereotyped *IGHV4-34* B-cell receptors

Millaray Marincevic,[1]* Mahmoud Mansouri,[1]* Meena Kanduri,[1] Anders Isaksson,[2] Hanna Göransson,[2] Karin Ekström Smedby,[3] Jesper Jurlander,[4] Gunnar Juliusson,[5] Fred Davi,[6] Kostas Stamatopoulos,[7,8] and Richard Rosenquist[1]

[1]Department of Genetics and Pathology, Uppsala University, Uppsala, Sweden; [2]Department of Medical Sciences, Cancer Pharmacology and Informatics, Uppsala University, Uppsala, Sweden; [3]Department of Medicine, Clinical Epidemiology Unit, Karolinska Institutet, Stockholm, Sweden; [4]Department of Hematology, Leukemia Laboratory, Rigshospitalet, Copenhagen, Denmark; [5]Lund Strategic Research Center for Stem Cell Biology and Cell Therapy, Hematology and Transplantation, Lund University, Lund, Sweden; [6]Laboratoire d'Hématologie, Hôpital Pitié-Salpêtrière et Université Pierre et Marie Curie, Paris, France; [7]Hematology Department and HCT Unit, G. Papanicolaou Hospital, Thessaloniki, Greece; [8]Institute of Agrobiotechnology, Center for Research and Technology, Thessaloniki, Greece

ABSTRACT

Background
Numerous subsets of patients with chronic lymphocytic leukemia display similar immunoglobulin gene usage with almost identical complementarity determining region 3 sequences. Among *IGHV4-34* cases, two such subsets with "stereotyped" B-cell receptors were recently identified, i.e. subset #4 (*IGHV4-34/IGKV2-30*) and subset #16 (*IGHV4-34/IGKV3-20*). Subset #4 patients appear to share biological and clinical features, e.g. young age at diagnosis and indolent disease, whereas little is known about subset #16 at a clinical level.

Design and Methods
We investigated the global gene expression pattern in sorted chronic lymphocytic leukemia cells from 25 subset/non-subset *IGHV4-34* patients using Affymetrix gene expression arrays.

Results
Although generally few differences were found when comparing subset to non-subset 4/16 *IGHV4-34* cases, distinct gene expression profiles were revealed for subset #4 *versus* subset #16. The differentially expressed genes, predominantly with lower expression in subset #4 patients, are involved in important cell regulatory pathways including cell-cycle control, proliferation and immune response, which may partly explain the low-proliferative disease observed in subset #4 patients.

Conclusions
Our novel data demonstrate distinct gene expression profiles among patients with stereotyped *IGHV4-34* B-cell receptors, providing further evidence for biological differences in the pathogenesis of these subsets and underscoring the functional relevance of subset assignment based on B-cell receptor sequence features.

Key words: stereotyped BCR, *IGHV4-34*, chronic lymphocytic leukemia, gene expression.

Citation: Marincevic M, Mansouri M, Kanduri M, Isaksson A, Göransson H, Smedby KE, Jurlander J, Juliusson G, Davi F, Stamatopoulos K, and Rosenquist R. Distinct gene expression profiles in subsets of chronic lymphocytic leukemia expressing stereotyped IGHV4-34 *B-cell receptors. Haematologica 2010;95(12):2072-2079. doi:10.3324/haematol.2010.028639*

**These authors contributed equally.*

Funding: this work was supported by the Swedish Cancer Society, the Swedish Research Council, the Medical Faculty of Uppsala University, Uppsala University Hospital, and Lion's Cancer Research Foundation in Uppsala, Sweden.

Manuscript received on June 4, 2010. Revised version arrived on August 17, 2010. Manuscript accepted on August 25, 2010.

Correspondence:
Richard Rosenquist, MD, PhD, Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, SE-751 85 Uppsala, Sweden.
Phone: international +46.18.6115947.
Fax: international +46.18.554025.
E-mail: richard.rosenquist@genpat.uu.se

The online version of this article has a Supplementary Appendix.

Introduction

The somatic hypermutation status of the immunoglobulin heavy variable (*IGHV*) genes is an independent prognostic marker in chronic lymphocytic leukemia (CLL) and subdivides patients into two subgroups with patients carrying unmutated *IGHV* genes having a poor prognosis.[1,2] It is now well established that CLL displays a remarkably biased *IGHV* gene repertoire with over-representation of a limited number of genes such as *IGHV1-69*, *IGHV4-34*, *IGHV3-23* and *IGHV3-21*.[2-4] Furthermore, several groups have reported the existence of multiple CLL subsets with "stereotyped" B-cell receptors (BCR), with similar *IG* heavy and light-chain gene usage and almost identical complementarity determining region 3 (CDR3) sequences, in up to 30% of patients.[5-9] These findings provide strong evidence for the involvement of antigens in the development of CLL. In addition, a number of studies have also indicated that stereotyped BCR may influence the clinical course in CLL.[8,10]

The *IGHV4-34* gene is detected in approximately 10% of CLL cases and patients carrying this rearrangement generally have highly mutated *IGHV* sequences and a favorable prognosis.[8] Recently, two relatively large subsets carrying stereotyped *IGHV4-34* BCR were reported in CLL. The more common one, occurring at an overall frequency of approximately 1%, is designated as subset #4 and is characterized by *IGHV4-34* rearrangements with highly homologous, 20 amino-acid long heavy variable CDR3 (VH CDR3), exclusive *IGKV2-30* gene usage and uniform IgG-switching.[8] Patients belonging to this subset also have an early age of onset of disease (median age at diagnosis, 43 years) and an indolent disease course, even when compared to cases expressing mutated *IGHV4-34* rearrangements with heterogeneous VH CDR3.[8]

The second subset, occurring at an overall frequency of 0.3%, is known as subset #16 and is characterized by *IGHV4-34* gene rearrangements with a distinct VH CDR3 of 24 amino acids in length, combined with *IGKV3-20* gene usage. Little is known regarding the clinical outcome of this subset, probably because of the limited number of patients identified to date.[8] Intriguingly, however, we recently found that this subset had a different spectrum of genomic aberrations compared to those of subset #4, even though both subsets consist of cases expressing mutated *IGHV4-34* BCR.[11]

In contrast to the more favorable outcome of *IGHV4-34*-expressing cases, we previously showed that patients carrying *IGHV3-21* BCR have a poor prognosis irrespective of the *IGHV* gene mutational status.[4] In a study on global expression profiling, patients with *IGHV3-21* had a different gene expression pattern from that of non-*IGHV3-21* patients.[12] Genes involved in DNA replication/cell cycle control, transcription and protein kinase activity were found to be differentially expressed, which may lead to a higher rate of proliferation and hence underlie the poor outcome of the patients with *IGHV3-21*.[12]

In the present study, we examined the gene expression profile of 25 *IGHV4-34* patients belonging to subset #4, subset #16, or the heterogeneous non-subset 4/16 group with the aim of exploring whether the clustering of cases based on structural features of the antigen-binding site is also reflected in distinct "subset-specific" gene expression profiles.

Design and Methods

Patients' material

Tumor samples, derived from peripheral blood, were collected from 25 CLL patients known to express the *IGHV4-34* gene for microarray expression analysis. The patients were untreated at the time of sample collection and came from collaborating institutes in France (n=3), Greece (n=10) and Sweden (n=12). Samples were classified according to recently revised criteria and displayed the typical CLL immunophenotype.[13] Based on *IG* gene sequence features and following previously established criteria,[8,9] cases were classified into three groups: (i) subset #4: 11 cases with *IGHV4-34/IGKV2-30* usage, 20 amino acid long VH CDR3; (ii) subset #16: 5 cases with *IGHV4-34/IGKV3-20* usage, 24 amino acid long VH CDR3; and, (iii) non-subset 4/16: 9 cases with *IGHV4-34* BCR of varying VH CDR3 length and composition as well as heterogeneous light-chain usage. In keeping with previous reports, all subset #4 cases as well as three subset #16 cases with available information expressed IgG; this information was not available for non-subset 4/16 cases. Although the non-subset 4/16 group included two cases with unmutated BCR, the average *IGHV4-34* gene mutational load in the three groups did not differ. The median age at diagnosis was 55 years (range, 37-73 years) for subset #4, 70 years (range, 55-81 years) for subset #16 and 60 years (range, 45-69 years) for non-subset 4/16 patients. A gender imbalance was observed for subset #16, in that it was primarily composed of women (male:female ratio: 1:4), in contrast to what is seen in CLL in general. Furthermore, two additional subset #16 cases (both IgG-switched) were collected for real-time quantitative polymerase chain reaction (RQ-PCR) validation. The clinical data and molecular characteristics of all patients are summarized in Table 1 and *Online Supplementary Table S1*. Informed consent was obtained according to the Helsinki declaration and the study was approved by the local Ethics Review Committees.

Isolation of chronic lymphocytic leukemia cells and RNA extraction

CLL cells were isolated through negative depletion of non-tumor cells using the Dynal B-Cell Negative Isolation Kit (Invitrogen, Carlsbad, CA, USA). The proportion of tumor cells following isolation was studied by FACS analysis of the cell surface markers CD5 and CD19 and was verified to be 93% or greater (median 98%, Table 1). RNA was extracted from isolated CLL cells using the RNeasy Mini kit (Qiagen, Hilden, Germany). The integrity of the RNA was evaluated using the Agilent 2100 Bioanalyzer system (Agilent Technologies Inc, Palo Alto, CA. USA) and its concentration measured with an ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

Microarray expression analysis

A total of 100 ng RNA from each sample was used to prepare biotinylated fragmented complementary RNA (cRNA) with a two-cycle amplification step, according to the GeneChip® Expression Analysis Technical Manual (Rev. 5, Affymetrix Inc., Santa Clara, CA, USA). The cRNA was hybridized to Affymetrix Human Genome U133 Plus 2.0 GeneChip® expression arrays for 16 h in a 45°C incubator with rotation at 60 rpm. The arrays were then washed and stained using the Fluidics Station 450 and finally scanned using the GeneChip® Scanner 3000 7G.

Microarray data analysis

The gene expression data were subsequently analyzed in R (*http://www.r-project.org*) using packages available from the

Bioconductor project (*www.bioconductor.org*). The raw data were normalized using the robust multi-array average method.[14,15] In order to search for differentially expressed genes between subset #4, subset #16 and non-subset 4/16 samples, an empirical Bayes moderated t-test was applied using the 'limma' package.[16,17] Only probe sets with an average intensity at reliable levels (log_2 value of >5) in at least one subgroup of patients (subset #4, subset #16 and non-subset 4/16) in each pair wise comparison were used for subsequent analyses. To address potential problems with multiple testing, the *P* values were adjusted using the method of Benjamini and Hochberg.[18] Genes with an adjusted *P* value less than 0.05 and an average fold change of at least 1.5 were regarded as differentially expressed.

Results were visualized by heat-maps and dendrograms using Genesis TreeView version 1.2.7 (*www.genome.tugraz.at*). Gene Ontology Tree Machine (GOTM) software was used to identify significant gene ontology categories in the data set (*http://bioinfo.vanderbilt.edu/gotm*). The expression data were further analyzed using ingenuity pathway analysis in order to determine significantly deregulated genes and pathways (Ingenuity® Systems, Mountain View, CA, USA; *www.ingenuity.com*).

Real-time quantitative reverse transcriptase polymerase chain reaction analysis

In order to validate gene expression data obtained from the microarray analysis, we selected three genes (*IL15*, *HOXA1* and *ZHX1*) for RQ-PCR analysis. Primers were designed using the Primer3 software (Broad Institute, Boston, USA) (sequences are available upon request). RQ-PCR analysis was performed using Maxima™ SYBR Green Master Mix according to the manufacturer's protocol (Fermentas, Burlington, Canada). *Beta-actin* expression was used as an internal reference. Differences in expression between subsets were evaluated using student's T-test and presented in box plots using the Statistica 8.0 software (Stat Soft, Tulsa, OK, USA).

Results

Expression profiling of IGHV4-34-expressing chronic lymphocytic leukemia

We performed high-resolution Affymetrix gene expres-

Table 1. Clinical and molecular data for patients included in the study.

Patient	% tumor cells	Subset	*IGHV* identity[a] (%)	Gender	Binet stage	Recurrent aberrations[b]	Additional aberrations[c]
1	99	16	95.0	F	A	del(13q)	del(1)(q32.1q32.1), dup(6)(q26q26)
2	96	16	92.9	F	A	+12	NA
3	94	16	95.8	M	A	none	none
4	99	16	93.0	F	A	none	dup(8)(p11.1p11.2)
5	97	16	93.2	F	A	del(11q)	NA
6	96	4	91.2	M	A	del(13q)x2	none
7	99	4	95.2	F	A	none	none
8	98	4	94.8	M	B	del(13q)	none
9	96	4	94.5	M	B	none	dup(18)(q23q23)
10	95	4	92.9	F	NA	none	del(5)(q15q15), dup(10)(q11.2q11.2)
11	99	4	94.8	M	A	del(13q)	dup(4)(q35.2q35.2)
12	99	4	91.2	M	A	none	none
13	99	4	91.9	M	A	del(13q)	none
14	99	4	93.2	F	A	del(13q)	dup(4)(q35.2q35.2)
15	97	4	93.1	F	A	del(13q)	del(2)(q37.3q37.3), del(8)(p23.1p23.1)
16	99	4	95.9	F	A	none	dup(14)(q23.3q23.3)
17	99	non-subset 4/16	94.4	F	A	del(13q)	del(4)(q34.3q34.3), del(9)(q22.3q22.3)
18	94	non-subset 4/16	92.8	F	A	none	dup(10)(q11.2q11.2), del(13)(q31.1q31.1)
19	99	non-subset 4/16	93.5	M	A	none	none
20	98	non-subset 4/16	93.2	M	B	del(13q)	none
21	98	non-subset 4/16	92.0	M	NA	none	none
22	97	non-subset 4/16	94.6	M	A	none	del(1)(q31.1q31.1), dup(2)(q12.2q12.3), dup(2)(q31.1q31.1), dup(19)(p13.3p13.2)
23	93	non-subset 4/16	92.1	M	A	none	none
24	99	non-subset 4/16	98.3	M	A	none	+3, dup(7)(q11.2q11.2)
25	98	non-subset 4/16	100	M	NA	NA	NA
26*	98	16	96.3	F	A	none	none
27*	99	16	95.1	M	A	NA	NA

*[a]Cases with ≥98% identity to germline were classified as unmutated, whereas cases with <98% identity were considered mutated. [b]Known recurrent aberrations [del(13q), del(11q), del(17p) and trisomy12] were determined by either FISH analysis using commercially available probes or high resolution Affymetrix 250K single nucleotide polymorphism-array analysis. [c]Additional aberrations detected with Affymetrix 250K SNP-arrays using a 200 kbp detection cutoff. NA, not available. *Additional subset #16 samples for RQ-PCR validation only.*

sion arrays on 25 sorted *IGHV4-34*-expressing samples with the aim of identifying biological features capable of distinguishing the different *IGHV4-34* subsets. After robust multi-array average normalization and log transformation, we detected few significantly differentially expressed genes when studying subset #4 and non-subset 4/16 cases (3 genes) and similarly when comparing subset #16 patients and non-subset 4/16 patients (10 genes) (*data not shown*). In contrast, we identified 111 significantly differentially expressed genes, when comparing subsets #4 and #16 (fold change differences are indicated in *Online Supplementary Table S2*). We then clustered patients according to their gene expression profiles for the 111 genes that displayed differential expression and could successfully distinguish subset #4 from subset #16 patients (Figure 1), although two subset #4 patients had an intermediate gene expression profile. The results revealed an overall lower gene expression profile for subset #4 compared to subset #16 (Figure 1 and *Online Supplementary Table S2*).

Next, we evaluated differences in gene expression between subset #4 cases and all other cases, i.e. subset #16 and non-subset 4/16 cases, identifying 14 differentially expressed genes of which 9 were also present in the initial list of 111 genes. These included genes involved in cell cycle regulation, such as *TLK1*, as well as genes implicated in apoptosis, e.g. *RPS27L*.[19,20] Figure 2 illustrates the gene expression profiles for these two groups, where notably, subset #4 cases again appear to have a lower expression of these genes.

Biological annotation of identified genes

In order to determine the biological relevance of the differentially expressed genes, we performed gene ontology enrichment analysis using Gene Ontology Tree Machine software. Here, we focused on the 111 genes that differed in expression between subset #4 and subset #16. Categories that showed a significant enrichment for these are shown in Table 2. Briefly, we found significant overrepresentation of genes affecting proliferation, cell cycle control and regulation of transcription activity including *STOML2*, *PPP2CA* and *HOXA1*.[21-23]

To further analyze our data, a second search using ingenuity pathway analysis was performed, again focusing on genes that differed in expression between subset #4 and subset #16, and identified a number of gene networks which were significantly enriched. These were classified as follows and are listed in Table 3: (i) cellular growth and proliferation (16 genes); (ii) humoral immune response (12 genes); (iii) cell-mediated immune response (13 genes); and (iv) hematopoiesis (7 genes).

Several of the differentially expressed genes studied, displaying lower expression in subset #4, have previously been implicated in cancer and cancer-related pathways where tumorigenesis is most often a consequence of overexpression. However, none of these has previously been

Table 2. Gene ontology categories enriched for differentially expressed genes in subset #4 *versus* subset #16.

Gene ontology categories	Observed #	Expected #	Ratio	Adjusted *P* value
Cellular component				
Intracellular	66	52	1.3	0.0032
Organelle	59	44.2	1.3	0.0035
Intracellular membrane-bounded	55	39.4	1.4	0.0032
Nucleus	43	25.2	1.7	0.0017
Regulation of chromatin assembly or disassembly	8	2.3	3.4	0.0233
Spliceosomal complex	4	0.7	6.0	0.0373
Nulear chromatin	3	0.3	10.1	0.0288
Biological process				
Intracellular transport	68	54.3	1.3	0.0032
RNA splicing	8	1.4	5.7	0.0198

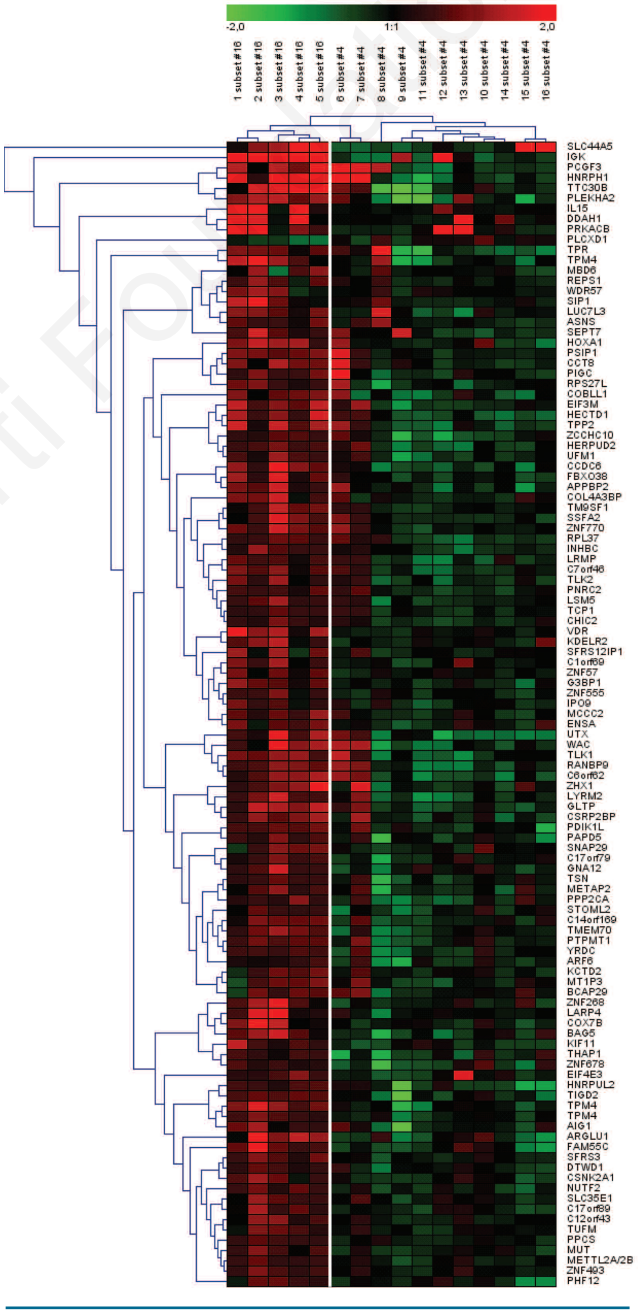

Figure 1. Visualization of differentially expressed genes using hierarchical clustering of genes in patients in subset #4 *versus* subset #16.

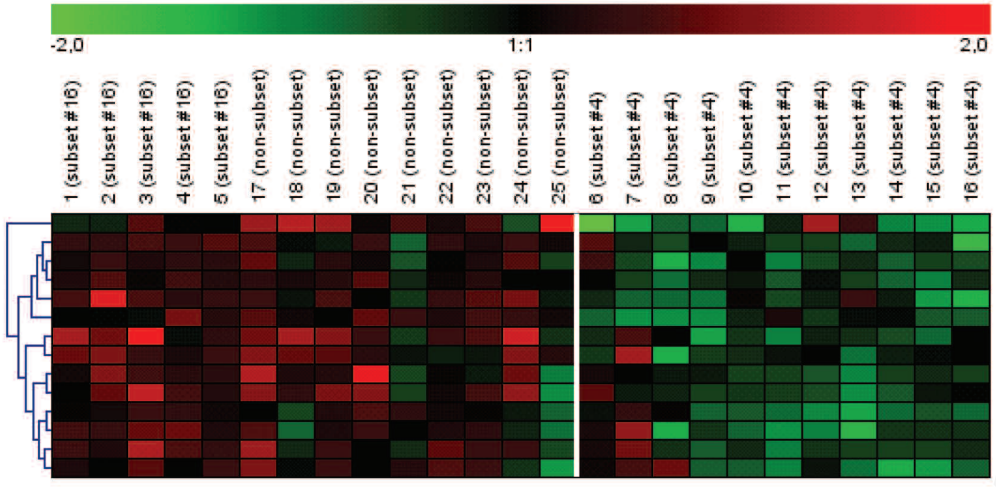

Figure 2. Visualization of differentially expressed genes using hierarchical clustering of genes in subset #4 *versus* all other cases.

associated with CLL. These genes are summarized in *Online Supplementary Table S3* and categorized as follows: (i) tumor proliferation; (ii) PI3K/AKT/NF-κB pathways; (iii) the p53 pathway; (iv) viral adhesion and replication in host cell; (v) apoptosis; and (vi) cancer therapy.

Confirmation of array data using real-time quantitative polymerase chain reaction analysis

To confirm the array data, we selected three differentially expressed genes in functionally relevant pathways, *IL15*, a cytokine involved in immune response,[24-26] *HOXA1*, a DNA binding transcription factor,[23] and *ZHX1*, a regulator of transcription,[27] for RQ-PCR analysis in seven subset #16 cases, of which two had not been analyzed using gene expression microarrays, and seven subset #4 cases. All genes analyzed validated the results from microarray analysis, having significantly higher expression in subset #16 cases (Figure 3).

Discussion

In the present study, we examined the gene expression profile of 25 *IGHV4-34* patients including subset #4, #16 and non-subset 4/16 cases. Initially, we compared the gene expression profiles between subset #4 and non-subset 4/16 patients and between subset #16 patients and non-subset 4/16 patients, and detected only few significant differences. This is probably because, overall, non-subset 4/16 *IGHV4-34* cases exhibited a more heterogeneous gene expression profile, likely reflecting the structural heterogeneity of their BCR, which would be expected to be responsive to a far wider range of antigens than that recognized by stereotyped subsets. Interestingly, however, we detected distinct differences in gene expression patterns when comparing subset #4 and #16 cases, both of which can be reliably defined at the molecular level based on subset-specific VH CDR3 and subset-biased features of somatic hypermutation.[8] This finding is supported by the recent observation that stereotyped subset cases have consistent antigen reactivity profiles.[28] In addition, it may be considered as further evidence that the clustering of cases based on IG primary sequences is biologically and, very likely, clinically relevant.

Intriguingly, subset #4 patients consistently exhibited a lower expression of the genes identified. Detailed characterization of the differently expressed genes in biological processes, using alternative approaches such as gene ontology tree machine and ingenuity pathway analysis, revealed that these genes are involved in cell cycle control and proliferation, such as *STOML2* and *PPP2CA*,[21,22] suggesting that these tumors have a low proliferative capacity compared to other CLL subsets, including subset #16.[29] This is supported by the identification of a subset of genes involved in hematopoiesis as well as humoral and cell-mediated immune responses (Table 3). *IL15* is one such example and acts as a pleiotropic cytokine shown to induce the proliferation of natural killer (NK) cells, B cells, and interferon-producing killer dendritic cells.[24-26]

Several of the differentially expressed genes detected to have consistently lower expression in subset #4 have previously been implicated in cancer and cancer-related pathways (*Online Supplementary Table S3*). In almost all cases, over-expression of these genes was associated with tumorigenesis. For instance, over-expression of *SSFA2* has been found to promote tumorigenicity in multiple myeloma,[30] *KIF11* is over-expressed in several cancers and its inhibition has been shown to induce cell cycle block and cell death via the mitochondrial pathway in acute myeloid leukemia cells,[31] *ARF6* is over-expressed in breast cancer and has been shown to be involved in metastatic cancer[32] and *PPP2CA* is over-expressed in breast tumors and implicated in the negative control of cell growth and division.[22] The finding that subset #4 cases display a distinctly lower gene expression profile may partly explain the favorable prognosis of these cases although the impact of individual genes or pathways in CLL is difficult to discern and needs to be studied in more detail.

Furthermore, when comparing expression profiles for subset #4 with those of other *IGHV4-34*-expressing cases (including subset #16), we detected differences in expression levels for only a few genes, again showing low expression in subset #4 (Figure 2). Several of these genes are important in tumorigenesis, including *RPS27L* and *INHBC*, which play important roles in p53 regulation and oncogenic transformation.[20,33] Nevertheless, when subset #16 cases were pooled with non-subset 4/16 cases, the observed gene expression differences between subset #4

Table 3. Enriched ingenuity pathway analysis (IPA) categories including differentially expressed genes between subset #4 and subset #16.

IPA network	Gene	Adjusted *P* value
(i) Cellular growth and proliferation, cell cycle (16 genes, IPA score 34)	*DTWD1*	2.28E-02
	G3BP1	1.72E-02
	GTPBP8	2.71E-02
	INHBC	2.46E-02
	LARP4	1.99E-02
	LSM5	2.16E-02
	MUT	4.74E-02
	NUTF2	2.17E-02
	PPCS	2.40E-02
	PSIP1	1.57E-03
	SSFA2	1.75E-02
	THAP1	1.72E-02
	TPP2	3.43E-02
	TSN	8.22E-03
	UFM1	4.50E-02
	ZCCHC10	3.84E-02
(ii) Humoral immune response (12 genes, IPA score 22)	*CHIC2*	3.52E-02
	CSNK2A1	2.82E-02
	GNA12	3.76E-02
	HOXA1	2.10E-03
	IL15	8.99E-08
	PIGC	8.44E-03
	PLCXD1	4.56E-02
	PNRC2	1.72E-02
	SIP1	1.57E-03
	VDR	4.76E-06
	ZHX1	2.83E-02
	ZNF268	1.25E-02
(iii) Cell-mediated immune response (13 genes, IPA score 22)	*C17ORF79*	3.76E-02
	CSRP2BP	3.35E-02
	DDAH1	7.04E-03
	HNRNPF	2.06E-02
	KDELR2	3.40E-02
	LRMP	9.00E-03
	PLEKHA2	2.67E-02
	PTPMT1	2.49E-03
	RPL37	9.24E-04
	RPS27L	4.05E-02
	SLC35E1	8.44E-03
	STOML2	3.84E-02
	TPM4	1.64E-02
(iv) Hematopoiesis (7 genes, IPA score 11)	*BAG5*	3.24E-02
	C6ORF62	4.85E-02
	CCT8	8.44E-03
	KDM6A	3.44E-02
	PAPD5	3.66E-02
	PHF12	4.96E-02
	WAC	4.05E-02

All genes were expressed at a lower level in subset #4 than in subset #16.

and #16 were "masked", probably because of the diverse BCR expressed by the latter cases, as discussed above. This result may justifiably be taken to imply that the comparison between distinct CLL subsets defined by BCR structural features is a pre-requisite for ensuring meaningful biological conclusions on the mechanisms underlying the selection and clonal expansion of CLL progenitors with distinctive BCR. Admittedly, this task will not be easy, given that, individually, each subset accounts for only a small fraction of a CLL cohort. Eventually, however, it may elucidate the pathogenesis of a sizable proportion of CLL cases, since, collectively, the stereotyped subsets account for up to 30% of a cohort.

Although the experimental platforms differed significantly, the findings of the present study are consistent with those of our recent work, which showed that subset #4 cases carry remarkably few genomic aberrations, again indicating that the CLL cells are less proliferative. In contrast, subset #16 cases displayed several known recurrent aberrations including the poor-prognostic deletion of 11q and trisomy 12.[11] We consider these findings as different pieces of the biological puzzle underlying the indolent clinical behavior of subset #4 cases. This is also reflected by the fact that only 18% (2/11) subset #4 patients required subsequent treatment while 60% (3/5) subset #16 patients were later treated.

Contrary to the findings for subset #4, we earlier reported a distinct gene expression profile in the group of patients with aggressive CLL carrying *IGHV3-21* BCR. In particular, we observed up-regulation of several genes involved in DNA replication/cell cycle control, transcription and protein kinase activity in *IGHV3-21* CLL cases compared to in non-*IGHV3-21* CLL cases, indicating a proliferative phenotype, which might account for the more aggressive clinical course.[12] This is in turn supported by the high number of copy number alterations detected in this subset.[11] That notwithstanding, both the gene expression and genomic profiles in poor-prognostic *IGHV3-21* patients and good-prognostic subset #4 *IGHV4-34* patients allude to the involvement of different immune-mediated pathways in the development of subgroups of CLL carrying distinct *IG* molecular features and, perhaps, distinct antigen reactivity and BCR signaling activity.

The *IGHV4-34* gene encodes antibodies which are intrinsically autoreactive due to recognition of the N-acetyl-lactosamine epitopes that are present on both various self-antigens (I/i blood group antigen, B-cell isoform of CD45) and microbial pathogens (Epstein-Barr virus, cytomegalovirus, *Mycoplasma pneumoniae*).[34,35] The *IGHV4-34* gene is used at a high frequency in healthy individuals; however, IGHV4-34 antibodies are virtually undetectable in healthy sera.[36] In contrast, IGHV4-34 antibodies are secreted at a high level in patients with systemic lupus erythematosus and in response to acute infections with herpes viruses and *M. pneumoniae*.[37-41] It has, therefore, been proposed that *IGHV4-34* B cells may normally be maintained in a state of diminished responsiveness, sparing the host from potential (auto)immune-mediated pathologies. In line with this, most CLL clonotypic *IGHV4-34* rearrangements are mutated, indicating that these CLL cells do not behave in an inherently different way from non-CLL *IGHV4-34* B cells. However, we previously noted that subset #4 may be distinguished from other *IGHV4-34* CLL by a distinctive somatic hypermutation profile, indicative of selective

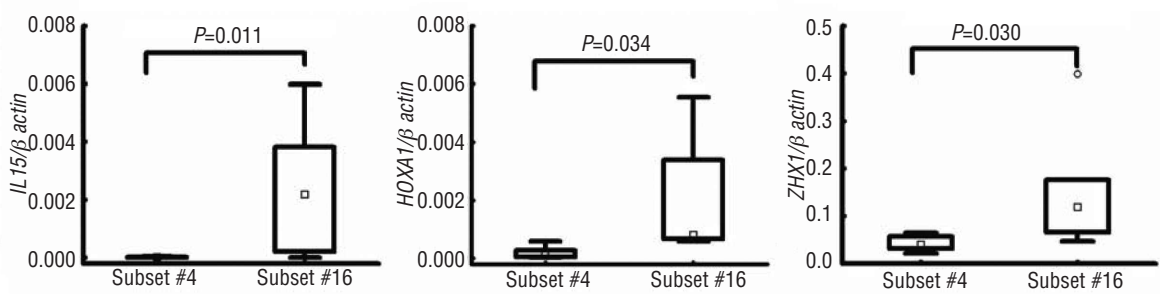

Figure 3. RQ-PCR validation of array data. Relative expression data presented as box plots for the following genes: ***IL15***, ***HOXA1*** and ***ZHX1***. Seven subset #16 and seven subset #4 cases were investigated.

(super)antigenic processes.[9,42] Perhaps relevant to this distinctive somatic hypermutation profile, we obtained evidence of a possible link between persistent infections with Epstein-Barr virus and cytomegalovirus and subset #4 CLL.[43] Additionally, we recently reported that subset #4 cases display extensive intraclonal diversification through precisely targeted somatic hypermutation, possibly as a result of ongoing antigenic stimulation, providing support for these findings.[44,45] Altogether, these observations could be taken to imply that though still actively recognizing antigen(s), subset #4 CLL cells are maintained in a state of diminished responsiveness, perhaps due to as yet unidentified mechanisms affecting signal transduction and eventual biological effects. The elucidation of these mechanisms will not be easy; however, it could be argued that the gene expression results reported here have offered useful suggestive evidence that might assist in designing further research in subset #4.

In conclusion, we show here that the gene expression profiles for subset #4 and subset #16 *IGHV4-34*-expressing CLL cases are distinctly different. Although the number of cases studied was relatively small because of the low overall frequency of patients in the two subsets, our findings provide evidence for differences in the underlying biological mechanisms in these subsets, particularly those involved in tumor maintenance and possibly even malignant transformation. RQ-PCR analysis also confirmed the microarray data for selected genes, validating the use of microarray analysis for gene expression profiling. The differentially expressed genes were found to be involved in important cell regulatory pathways including cell cycle control and proliferation and were commonly found to have lower expression in subset #4 cases which may reflect the lack of genomic complexity and indolent phenotype in this group of patients. Finally, we showed that non-subset 4/16 cases have a heterogeneous gene expression profile, which highlights the need for further sub-characterization of CLL in order to identify differences between biologically distinct subgroups of CLL.

Authorship and Disclosures

The information provided by the authors about contributions from persons listed as authors and in acknowledgments is available with the full text of this paper at www.haematologica.org.

Financial and other disclosures provided by the authors using the ICMJE (www.icmje.org) Uniform Format for Disclosure of Competing Interests are also available at www.haematologica.org.

References

1. Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood. 1999;94(6):1840-7.
2. Hamblin TJ, 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(6):1848-54.
3. Fais F, Ghiotto F, Hashimoto S, Sellars B, Valetto A, Allen SL, et al. Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. J Clin Invest. 1998;102(8):1515-25.
4. Tobin G, Thunberg U, Johnson A, Thorn I, Soderberg O, Hultdin M, et al. Somatically mutated Ig V(H)3-21 genes characterize a new subset of chronic lymphocytic leukemia. Blood. 2002;99(6):2262-4.
5. Tobin G, Thunberg U, Johnson A, Eriksson I, Soderberg O, Karlsson K, et al. Chronic lymphocytic leukemias utilizing the VH3-21 gene display highly restricted Vlambda2-14 gene use and homologous CDR3s: implicating recognition of a common antigen epitope. Blood. 2003;101(12):4952-7.
6. Tobin G, Thunberg U, Karlsson K, Murray F, Laurell A, Willander K, et al. Subsets with restricted immunoglobulin gene rearrangement features indicate a role for antigen selection in the development of chronic lymphocytic leukemia. Blood. 2004;104(9):2879-85.
7. Messmer BT, Albesiano E, Efremov DG, Ghiotto F, Allen SL, Kolitz J, et al. Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. J Exp Med. 2004;200(4):519-25.
8. Stamatopoulos K, Belessi C, Moreno C, Boudjograh M, Guida G, Smilevska T, et al. Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: Pathogenetic implications and clinical correlations. Blood. 2007;109(1):259-70.
9. Murray F, Darzentas N, Hadzidimitriou A, Tobin G, Boudjogra M, Scielzo C, et al. Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. Blood. 2008;111(3):1524-33.
10. Bomben R, Dal Bo M, Capello D, Forconi F, Maffei R, Laurenti L, et al. Molecular and clinical features of chronic lymphocytic leukaemia with stereotyped B cell receptors: results from an Italian multicentre study. Br J Haematol. 2009;144(4):492-506.
11. Marincevic M, Cahill N, Gunnarsson R, Isaksson A, Mansouri M, Göransson H, et al. High-density screening reveals a different spectrum of genomic aberrations in

chronic lymphocytic leukemia patients with 'stereotyped' IGHV3-21 and IGHV4-34 B cell receptors. Haematologica. 2010;95(9):1519-25.
12. Falt S, Merup M, Tobin G, Thunberg U, Gahrton G, Rosenquist R, et al. Distinctive gene expression pattern in VH3-21 utilizing B-cell chronic lymphocytic leukemia. Blood. 2005;106(2):681-9.
13. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. Blood. 2008;111 (12):5446-56.
14. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci USA. 2001;98(1):31-6.
15. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. 2003;4(2):249-64.
16. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol. 2004;3:Article3.
17. Smyth GK. Limma: linear models for microarray data. In: Bioinformatics and Computational Biology Solutions using R and Bioconductor. New York: Springer, 2005.
18. Benjamini Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Statist Soc Ser B. 1995;57(1)289-300.
19. Sillje HHW, Takahashi K, Tanaka K, Van Houwe G, Nigg EA. Mammalian homologues of the plant Tousled gene code for cell-cycle-regulated kinases with maximal activities linked to ongoing DNA replication. EMBO J. 1999;18(20):5691-702.
20. He H, Sun Y. Ribosomal protein S27L is a direct p53 target that regulates apoptosis. Oncogene. 2007;26(19):2707-16.
21. Cao W, Zhang B, Liu Y, Li H, Zhang S, Fu L, et al. High-level SLP-2 expression and HER-2/neu protein expression are associated with decreased breast cancer patient survival. Am J Clin Pathol. 2007;128(3):430-6.
22. Wong LL, Chang CF, Koay ES, Zhang D. Tyrosine phosphorylation of PP2A is regulated by HER-2 signalling and correlates with breast cancer progression. Int J Oncol. 2009;34(5):1291-301.
23. Langston AW, Gudas LJ. Retinoic acid and homeobox gene regulation. Curr Opin Genet Dev. 1994;4(4):550-5.
24. Carson WE, Giri JG, Lindemann MJ, Linett ML, Ahdieh M, Paxton R, et al. Interleukin (IL) 15 is a novel cytokine that activates human natural killer cells via components of the IL-2 receptor. J Exp Med. 1994; 180(4):1395-403.
25. Armitage RJ, Macduff BM, Eisenman J, Paxton R, Grabstein KH. IL-15 has stimulatory activity for the induction of B cell proliferation and differentiation. J Immunol. 1995;154(2):483-90.
26. Arina A, Murillo O, Dubrot J, Azpilikueta A, Gabari I, Perez-Gracia JL, et al. Interleukin-15 liver gene transfer increases the number and function of IKDCs and NK cells. Gene Ther. 2008;15(7):473-83.
27. Kim S-H, Park J, Choi M-C, Kim H-P, Park J-H, Jung Y, et al. Zinc-fingers and homeoboxes 1 (ZHX1) binds DNA methyltransferase (DNMT) 3B to enhance DNMT3B-mediated transcriptional repression. Biochem Biophys Res Commun. 2007;355 (2):318-23.
28. Chu CC, Catera R, Zhang L, Didier S, Agagnina BM, Damle RN, et al. Many chronic lymphocytic leukemia antibodies recognize apoptotic cells with exposed non-muscle myosin heavy chain IIA: implications for patient outcome and cell of origin. Blood. 2010;115(19):3907-15.
29. Herve M, Xu K, Ng YS, Wardemann H, Albesiano E, Messmer BT, et al. Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive B cell precursors despite expressing different antibody reactivity. J Clin Invest. 2005;115(6): 1636-43.
30. Tai YT, Soydan E, Song W, Fulciniti M, Kim K, Hong F, et al. CS1 promotes multiple myeloma cell adhesion, clonogenic growth, and tumorigenicity via c-maf-mediated interactions with bone marrow stromal cells. Blood. 2009;113(18):4309-18.
31. Carter BZ, Mak DH, Woessner R, Gross S, Schober WD, Estrov Z, et al. Inhibition of KSP by ARRY-520 induces cell cycle block and cell death via the mitochondrial pathway in AML cells. Leukemia. 2009;23 (10):1755-62.
32. Sabe H, Hashimoto S, Morishige M, Ogawa E, Hashimoto A, Nam JM, et al. The EGFR-GEP100-Arf6-AMAP1 signaling pathway specific to breast cancer invasion and metastasis. Traffic. 2009;10(8):982-93.
33. Risbridger GP, Mellor SL, McPherson SJ, Schmitt JF. The contribution of inhibins and activins to malignant prostate disease. Mol Cell Endocrinol. 2001;180(1-2):149-53.
34. Bhat NM, Bieber MM, Chapman CJ, Stevenson FK, Teng NN. Human antilipid A monoclonal antibodies bind to human B cells and the i antigen on cord red blood cells. J Immunol. 1993;151(9):5011-21.
35. Silberstein LE, George A, Durdik JM, Kipps TJ. The V4-34 encoded anti-i autoantibodies recognize a large subset of human and mouse B-cells. Blood Cells Mol Dis. 1996;22(2):126-38.
36. van Vollenhoven RF, Bieber MM, Powell MJ, Gupta PK, Bhat NM, Richards KL, et al. VH4-34 encoded antibodies in systemic lupus erythematosus: a specific diagnostic marker that correlates with clinical disease characteristics. J Rheumatol. 1999;26(8): 1727-33.
37. Pugh-Bernard AE, Silverman GJ, Cappione AJ, Villano ME, Ryan DH, Insel RA, et al. Regulation of inherently autoreactive VH4-34 B cells in the maintenance of human B cell tolerance. J Clin Invest. 2001;108(7): 1061-70.
38. Cappione AJ, Pugh-Bernard AE, Anolik JH, Sanz I. Lupus IgG VH4.34 antibodies bind to a 220-kDa glycoform of CD45/B220 on the surface of human B lymphocytes. J Immunol. 2004;172(7):4298-307.
39. Mockridge CI, Rahman A, Buchan S, Hamblin T, Isenberg DA, Stevenson FK, et al. Common patterns of B cell perturbation and expanded V4-34 immunoglobulin gene usage in autoimmunity and infection. Autoimmunity. 2004;37(1):9-15.
40. Bhat NM, Bieber MM, Yang YC, Leu YS, van Vollenhoven RF, Teng NN. B cell lymphoproliferative disorders and VH4-34 gene encoded antibodies. Hum Antibodies. 2004;13(3):63-8.
41. Loomes LM, Uemura K, Childs RA, Paulson JC, Rogers GN, Scudder PR, et al. Erythrocyte receptors for Mycoplasma pneumoniae are sialylated oligosaccharides of Ii antigen type. Nature. 1984;307(5951): 560-3.
42. Hadzidimitriou A, Darzentas N, Murray F, Smilevska T, Arvaniti E, Tresoldi C, et al. Evidence for the significant role of immunoglobulin light chains in antigen recognition and selection in chronic lymphocytic leukemia. Blood. 2009;113(2):403-11.
43. Kostareli E, Hadzidimitriou A, Stavroyianni N, Darzentas N, Athanasiadou A, Gounari M, et al. Molecular evidence for EBV and CMV persistence in a subset of patients with chronic lymphocytic leukemia expressing stereotyped IGHV4-34 B-cell receptors. Leukemia. 2009;23(5):919-24.
44. Sutton LA, Kostareli E, Hadzidimitriou A, Darzentas N, Tsaftaris A, Anagnostopoulos A, et al. Extensive intraclonal diversification in a subgroup of chronic lymphocytic leukemia patients with stereotyped IGHV4-34 receptors: implications for ongoing interactions with antigen. Blood. 2009;114(20):4460-8.
45. Kostareli E, Sutton LA, Hadzidimitriou A, Darzentas N, Kouvatsi A, Tsaftaris A, et al. Intraclonal diversification of immunoglobulin light chains in a subset of chronic lymphocytic leukemia alludes to antigen-driven clonal evolution. Leukemia. 2010;24(7): 1317-24.