Gene Expression Profiling of B Cell Chronic Lymphocytic Leukemia Reveals a Homogeneous Phenotype Related to Memory B Cells

Ulf Klein,¹ Yuhai Tu,² Gustavo A. Stolovitzky,² Michela Mattioli,¹ Giorgio Cattoretti,¹ Hervé Husson,³ Arnold Freedman,³ Giorgio Inghirami,⁴ Lilla Cro,⁵ Luca Baldini,⁶ Antonino Neri,⁵ Andrea Califano,⁶ and Riccardo Dalla-Favera¹

¹Institute for Cancer Genetics, the Departments of Pathology and Genetics & Development, Columbia University, New York, NY 10032
²IBM T.J. Watson Research Center, Yorktown Heights, New York, NY 10598
³Department of Adult Oncology Dana-Farber Cancer Institute, and the Department of Medicine, Harvard Medical School, Boston, MA 02115
⁴Department of Pathology, New York University Medical Center, New York, NY 10016
⁵Ematologia I, Department of Hematology, Ospedale Maggiore, I.R.C.C.S., Milan 20122, Italy
⁶First Genetic Trust, Inc., Lyndhurst, NJ 07071

Abstract

B cell–derived chronic lymphocytic leukemia (B-CLL) represents a common malignancy whose cell derivation and pathogenesis are unknown. Recent studies have shown that ≥50% of CLLs display hypermutated immunoglobulin variable region (IgV) sequences and a more favorable prognosis, suggesting that they may represent a distinct subset of CLLs which have transited through germinal centers (GCs), the physiologic site of IgV hypermutation. To further investigate the phenotype of CLLs, their cellular derivation and their relationship to normal B cells, we have analyzed their gene expression profiles using oligonucleotide-based DNA chip microarrays representative of ~12,000 genes. The results show that CLLs display a common and characteristic gene expression profile that is largely independent of their IgV genotype. Nevertheless, a restricted number of genes (<30) have been identified whose differential expression can distinguish IgV mutated versus unmutated cases and identify them in independent panels of cases. Comparison of CLL profiles with those of purified normal B cell subpopulations indicates that the common CLL profile is more related to memory B cells than to those derived from naïve B cells, CD5⁺ B cells, and GC centroblasts and centrocytes. Finally, this analysis has identified a subset of genes specifically expressed by CLL cells of potential pathogenic and clinical relevance.

Key words: somatic hypermutation • germinal center • CD5 • DNA microarray • cluster analysis

Introduction

B cell chronic lymphocytic leukemia (B-CLL) represents the most common malignancy in the Western countries with an estimated incidence of 1 per 100,000 per y (1). This disease is characterized by the monoclonal expansion of B lymphocytes in the peripheral blood, bone marrow, and lymphoid organs, and by an indolent course which ultimately becomes aggressive and invariably lethal (1). Current knowledge of the pathogenesis of CLL is limited because no specific genetic alteration has yet been associated with this disease. In particular, CLL is not associated...
with reciprocal balanced chromosomal translocations, but rather with specific deletions (2) suggesting the loss of presently unidentified tumor suppressor genes. CLL cells have a low proliferative rate and a prolonged life span, suggesting that their primary alteration may be a defect in apoptosis (1).

The cellular origin of CLL is unknown and has been the object of recent controversy. CLL cells are characterized by the expression of the cell surface markers CD5, CD23, CD19, and low levels of sIgM/sIgD, a pattern not shared by any known B cell subpopulation (1, 3). The expression of the CD5 antigen was originally taken to suggest that B–CLL originates from CD5+ B cells (4), which are usually characterized by unmutated Ig variable region (IgV) genes (5, 6). However, recent studies have shown that 50–70% of CLL have undergone IgV hypermutation (7–9), a phenomenon that characterizes normal B cells undergoing a T cell–dependent germinal center (GC) reaction (10) and malignant B cells derived from GC or post-GC B cells (11, 12). This finding has led to the hypothesis that CLL cases displaying mutated Ig may derive from a cell that had transited through the GC, whereas those with germline IgV may derive from GC-independent cells (8, 9, 13). This hypothesis has both biological and clinical relevance since the two subgroups have different prognoses, with IgV–mutated CLL (M-CLL) displaying a more benign clinical course (13, 14).

To further investigate the phenotype of CLL subtypes, their cellular derivation and their relationship with normal B cells, we have analyzed their gene expression profiles using oligonucleotide-based DNA chip microarrays (Affymetrix) representative of ~12,000 genes. The gene expression profiles have been comparatively analyzed: (i) between M-CLL and IgV-unmutated CLL (UM-CLL) cases, to determine whether the two subgroups are different; (ii) with those of normal B cell subpopulations, including GC-independent (CD5+), preGC (naive), GC (centroblasts [CBs] and centrocytes [CCs]), and post-GC (memory) B cells, to investigate their cellular derivation; and (iii) with those of normal B cell subpopulations and of other malignancies derived from mature B cells to identify genes that are specifically expressed in CLL.

Materials and Methods

**CLL Cases.** Peripheral blood from CLL patients who had not received treatment was taken after informed consent. CLL cells were enriched from PBMCs by magnetic cell separation using the MidiMACS® system (Miltenyi Biotec). After Ficoll-Isopaque™ density centrifugation, PBMCs were incubated with CD19-microbeads (MB) (Miltenyi Biotec). Magnetically labeled cells were isolated by passing the cell suspension over an LS column (Miltenyi Biotec). All isolation steps were performed on ice and using ice-cold solutions. Sequences of IgVH genes of the 34 CLL cases were determined as described previously (15).

**Isolation of Normal Human B Cell Subsets and Tumor Cells.** A detailed description of the gene expression profiles of normal B cell subpopulations, including the methods for purification and characterization, will be reported elsewhere (unpublished data). Briefly, tonsils were obtained from routine tonsillectomies performed at the Children’s Hospital of Columbia University. The specimens were kept on ice immediately after surgical removal, and all cell isolation steps were performed on ice. B cell subsets were isolated by magnetic cell separation (see above). The isolation strategies were based on recent characterizations of the human B cell subsets (16–19). For the isolation of naïve B cells, tonsillar mononuclear cells were first incubated with anti-CD27, anti-CD10 (both BD PharMingen), and anti-CD3 (Beckman Coulter/Immunotech), then with anti-IgG1-MB and anti-CD14-MB. Magnetically labeled cells were depleted by passing the cell suspension over a LD-column (Miltenyi Biotec). The flow-through was incubated first with anti–IgD-FITC (BD PharMingen), then with anti–FITC-MB. IgD-positive B cells were isolated using LS columns. CBs were isolated in a single step by staining tonsillar mononuclear cells with anti-CD77 (Beckman Coulter/Immunotech), followed by incubation with mouse anti-rat IgM (BD PharMingen), and finally anti–IgG1-MB. The cells were passed over a LS column. CC: tonsillar MC were first incubated with anti-CD77, then with anti-CD39 (Beckman Coulter/Immunotech), anti-CD3, and mouse anti–rat IgM, followed by anti–IgG1-MB staining. The cells were passed subsequently over a LS column with G21 needle and a LD column. The resulting flow-through was first incubated with anti–CD10, then with anti–IgG1-MB. CD10+ CCs were isolated using LS columns. Memory B cells: tonsillar mononuclear cells were incubated first with anti–CD10, anti–CD3, and anti–CD38 (BD PharMingen), the latter at a titer of 1:1,000 to facilitate a selective depletion of CD38high plasma cells and GC B cells, then with anti–IgG1-MB and anti–CD14-MB. Labeled cells were depleted using a LD column. The flow-through was first stained with anti–CD27–FITC, then with anti–FITC-MB. CD27+ cells were isolated using a LS column. Cord blood was obtained after informed consent. Cord blood B cells almost exclusively consist of CD5+ B cells. To enrich for CD5– B cells from cord blood, mononuclear cells obtained by Ficoll-Isopaque™ density centrifugation were incubated with CD19-MB (Miltenyi Biotec). The stained cells were passed over a LS column.

Follicular lymphoma (FL) cells were isolated from single cell suspensions from digested lymph nodes or spleen which were histologically involved with FL. These cell suspensions were depleted of T cells, monocytes, NK cells using anti-CD4, -CD8, -CD14, and -CD56, followed by immunomagnetic beads as described previously (20). FL cells were isolated from six individuals with relapsed FL who had not received any therapy for at least 6 mo before biopsy. Diffuse large cell lymphoma (DLCL) and Burkitt lymphoma (BL) tumor cells were enriched by magnetically depleting CD3-, CD4+, CD8-, CD14+, CD15-, CD16-, CD56-, and glycoporphin-expressing nontumor cells. All samples were obtained according to appropriate Human Protection Committee validation and informed patient consent.

**Generation of cRNA and Microarray Hybridization.** Total RNA was isolated in two steps using Trizol (Life Technologies), followed by RNeasy (Qiagen) purification. Double strand cDNA was generated from 5 µg of total RNA using a poly dT oligonucleotide that contains a T7 RNA polymerase initiation site and the SuperScript Choice System kit (Life Technologies). cDNA was phenol/chloroform extracted. Biotinylated cRNA was generated by in vitro transcription using the BioArray High Yield cRNA Transcription Labeling kit (ENZO Diagnostics, Inc.). The cRNA was purified using RNeasy. cRNA was fragmented according to the Affymetrix protocol, and 15 µg of biotinylated...
genes) in a set of random cells whose genes are distributed ac-
based on the probability of observing a similar pattern (i.e. a pat-
statistical significance of a Genes@Work pattern is computed

- **Fold Changes:** The expression values of each selected gene is
-average change in expression level was at least threefold and four-

- **Hierarchical Clustering:** The algorithm builds a cluster with all
- **Unsupervised Clustering:** Following computation of the expression probability of each gene, pattern discovery is performed starting with a support value \( n = n_0 - m \), where \( n_0 \) is equal to the number of experiments in the set and \( m \) is the size of the smallest subset which is expected to be characterized by a discernible subphenotype (typically, \( m = n_0/4 \)). For a given support \( n \), the pattern \( z \) score \( z_p(n) \) of each discovered pattern, \( p \), is computed. (The pattern \( z \) score \( z_p \), is inversely related to the pattern \( p \) value; when \( z_p \) is sufficiently large, this relation is \( \sim pf = 1/z_p \).) The pattern \( P(n) \), corresponding to the maximum pattern \( z \) score \( Z_p(n) \), is selected. \( Z_p(n) = 1 \) if no patterns are found for the given support value. This process is repeated iteratively for decreasing values of \( n \), until a sudden increase in the pattern \( z \) score is detected. This corresponds to a maximum of the first derivative of the pattern \( z \) score function for a support value equal to \( n \). Such extrema can be determined, as a first approximation, when the following condition is satisfied: \( Z_p(n - 1)/Z_p(n) \leq Z_p(n)/Z_p(n + 1) \). The pattern \( P(n) \) for the corresponding value of \( n \) is called the interest pattern. The experiments that support the interest pattern become a subset and the genes that support the interest pattern become the subset-signature. For a discussion of the pattern \( z \) score, see reference 24.

**Supervised Clustering:** the gene expression probability densities are computed and pattern discovery is performed iteratively with a decreasing value of the support, starting at \( N \) (the size of the phenotype group), until either a significant gene cluster size is obtained, or discovered patterns appear to be no longer statistically significant.

- **Unsupervised Clustering:** For the unsupervised clustering analysis, an optimal value for \( \delta = 0.1 \), or 10% of the total probability, is chosen. For supervised clustering, due to the significantly higher sensitivity, a smaller value of \( \delta = 0.01 \), or 1% of the total probability, is used.

**Graphic Representations of Gene Expression Patterns (Matrices).** Columns represent individual experiments, and rows represent individual genes present on the expression microarray. To generate a pseudocolor map, first a gene- and experiment-specific change of variables, from the original measurement \( v \) into \( \xi_{i,v} \), is computed using the formula:

\[
\xi_{i,v} = \nu - (\mu_{i,v} + \mu_{v}) / 2
\]

where \( \mu_{i,v} \) and \( \mu_v \) are respectively the mean and SD computed from the gene expression values for that gene in the phenotype group, and \( \mu_v \) and \( \sigma_v \) are their corresponding values computed from the control group (Eq. 1). The value of this function is then
plotted using a pseudocolor map that represents \( z_i = 0 \) as black, \( z_i > 0 \) as progressively brighter hues of red, and \( z_i < 0 \) as progressively brighter levels of green. \( z_i = 4 \) and \( z_i = -4 \) correspond to complete saturation of the red and the green, respectively. The resulting pseudocolor map associates the same colors to measurements that are off by the same number of SDs from their expected value.

For each gene, the statistical significance of the differential expression across the phenotype and control sets (gene \( z \) score, \( z_i \), is computed using the formula (Eq. 2)

\[
    z_i = \frac{\mu_i - \mu_C}{\sigma_i + \sigma_C}. \tag{2}
\]

In contrast to the generally used fold-ratio (\( \mu_i/\mu_C \)) to describe gene expression changes among cell subsets, \( z_i \) represents the differential expression between phenotype and control samples relative to the variability of their expression levels. Rows are divided into two groups. First, the genes that are overexpressed in the first experiment set are reported by decreasing value of the \( z_i \) score (most significant first). Then, the genes that are underexpressed are reported, ordered by increasing value of the \( z_i \) score (most significant first).

**Classifier and Classification Method.** The classifier is a scoring function based on the values of a set of genes (gene cluster) which are differentially expressed in two sets of cell types and can thus be used for cell type classification. The higher the score, the more likely it is that a cell type is related to the phenotype set. Pattern discovery methods are used here to identify gene clusters. The union of the genes of all statistically significant patterns is used to define the classifier (see above for pattern discovery method). Given a set of genes, the scoring function is defined quantitatively as:

\[
    R_i = \frac{1}{n_i} \sum_{i=1}^{n_i} \left( \frac{\nu_i + (-\mu_i + \mu_C)/2}{(\mu_i - \mu_C)/2} \right)^2 \tag{3}
\]

where \( \nu_i = \max (-1, \min (1, x_i)) \), \( n_i \) is the number of genes in the classifier, and \( i = 1, 2, ..., n_i \) is the label for each individual gene in the classifier, \( \mu_i \) is the expression level of gene \( i \) for the new cell, \( \mu_i^{(1)} \) and \( \mu_i^{(2)} \) represent the mean expression levels of gene \( i \) for cells in the phenotype and control set, respectively (Eq. 3).

**Online Supplemental Material.** GenBank accession no., Affymetrix entry numbers, and the normalized primary data of the genes shown in Figs. 2, 4, and 5 are available at the journal’s website.

**Results**

A panel of 34 CLLs, characterized for their typical cell surface phenotype and presence or absence of IgV mutations, was used for this study. From 20 of these samples, tumor cells were purified by magnetic cell separation of CD19+ cells, while 14 cases that showed a representation of \( >80\% \) of malignant cells in the peripheral blood were used as unpurified cell populations (Table I). As in previous reports (13, 14), CLLs with \( \leq 2\% \) basepair difference to the corresponding germline IgV gene were considered as UMCCLls. 18 of 34 cases, comparably distributed between purified and unpurified samples (11 and 7, respectively), were shown to be M-CLLs carrying 6–33 mutations per case.

**Table I. Summary of the Analysis of CLL Cases for the Level of Somatic Hypermutation in Rearranged IgVH Genes**

| Case     | V_H gene Mutations | Case     | V_H gene Mutations |
|----------|--------------------|----------|--------------------|
| CLL0004  | 3-23               | CLL1215  | 4-34               |
| CLL0014  | 4-39               | CLL1263  | 4-34               |
| CLL1208  | 4-61               | CLL1268  | 1-69               |
| CLL1255  | 4-34               | CLL1285  | 1-69               |
| CLL1259  | 3-9                | CLL1291  | 4-59               |
| CLL1269  | 3-23               | CLL1297  | 4-59               |
| CLL1270  | 4-34               | CLL1302  | 3-73               |
| CLL1272  | 3-23               | CLL1304  | 3-15-rel.         |
| CLL1275  | 3-30               | CLLB2    | 3-64/4-34         | 29/16
| CLL1282  | 3-30               | CLLB3    | 1-2                |
| CLL1310  | 3-43               | CLLB4    | 1-2                |
| CLL1318  | 3-7                | CLLB6    | 1-3/4-59          | 10/17
| CLL1332  | 4-34               | CLLB7    | BF2P2             |
| CLL1333  | 4-39               | CLLB8    | 1-18               |
| CLLB9    | 3-48               | CLLB10   | 3-21               |
| CLLB12   | 3-30               | CLLB13   | 5-51               |
| CLLB14   | 3-21               | CLLB15   | 1-2               |

Nonpurified CLLs comprised 80% or more tumor cells among PBMCs; purified CLLs were isolated from PBMCs by magnetic cell separation of CD19+ cells. If possible, the V_H gene nomenclature of Matsuda et al. was used (reference 56).

*aContains a 1-bp insertion.

*bGenBank accession no. M99406 (3–15-related rearranged IgV gene).

**V_H3-family member**

RNA extracted from these cases was converted into labeled cRNA and hybridized to U95A Affymetrix Gene Chips representative of \( \sim 12,000 \) genes, including mostly known genes (\( >80\% \)). Gene expression profiles were analyzed using two main approaches: (i) unsupervised clustering, which can identify distinct cell types (e.g., CLL cases) which have not been classified a-priori; and (ii) supervised clustering, which allows the identification of differentially expressed genes between cell types (CLL cases) defined a-priori according to a given criterium (e.g., presence of IgV mutations). Unsupervised and supervised clustering were obtained using two algorithms, the average-linkage method (21, 22) or the pattern discovery algorithm SPLASH used by the Genes@Work software platform (see Materials and Methods; references 23 and 24). The latter is capable of capturing subtle differences in gene expression by combining an optimal nonlinear transformation of the gene expression values coupled with an analytical method to com-
pute the statistical significance of the identified clusters (see Materials and Methods and below).

**CLLs Display a Common Gene Expression Profile Independent of IgV Mutations.** To determine whether M- and UM-CLLs are phenotypically different, we first analyzed their gene expression profiles by two independent unsupervised methods. When analyzed using Genes@Work (not shown) or clustering by the average-linkage method (Fig. 1), CLLs displayed a common profile that is clearly distinguishable from that of FL (see first branching in the dendrogram; Fig. 1). Purified and nonpurified cases were readily recognized as different (second branching; Fig. 1), most likely reflecting the contribution of normal cells contaminating the nonpurified cases. However, in both subgroups, M- and UM-CLLs were not distinguishable and appear intermingled (Fig. 1). Lowering the “stringency” of the analysis to the lowest limits of statistical significance for pattern discovery, or varying the selection criteria for the genes used in clustering (see Materials and Methods), did not lead to further separation of M- and UM-cases (data not shown). These results indicate that M- and UM-CLLs have a common pattern of expression for most of the 12,000 genes analyzed, suggesting that they have a largely common phenotype.

A Small Subset of Differentially Expressed Genes Allows the Classification of M- and UM-CLLs. Subtle differences in gene expression among closely related cell populations may escape detection using unsupervised clustering analyses. Thus, to determine whether M- and UM-CLLs have subtle differences in gene expression, the two subgroups were compared by supervised clustering using Genes@Work.

We first compared a set of 20 cases including nine UM- and 11 M-CLL. To avoid the influence of contaminating cells and to obtain the maximum specificity for the CLL phenotype, these cases were selected among the purified ones. Fig. 2 A shows that a set of 23 genes is differentially expressed in M- versus UM-CLLs, the majority of them (20 of 23) being upregulated in UM-CLLs versus M-CLLs. Several of these genes are of unknown function, while the remaining ones encode products of heterogeneous nature.

To validate the specificity of this differential gene expression profile, we tested whether it could be used to classify an independent panel of 14 cases into M- and UM-CLLs. To test the sensitivity of the classifier and its potential clinical use, unpurified cases were used for this analysis. Fig. 2 B shows that the profile shown in Fig. 2 A could correctly classify all seven M-CLLs, and 5 of 7 UM-CLLs (p-value of < 0.025). Thus, the 23 genes differentially expressed in M- versus UM-CLL represent a consistent phenotypic difference between the two subgroups (see Discussion).

**Figure 1.** M- and UM-CLLs share a common gene expression profile. Dendrogram and matrix showing the hierarchical clustering of 2337 selected genes (see Methods) of gene expression data generated from 34 CLL (“P” indicates purified cases) and six purified FL samples. The hierarchical clustering algorithm used is based on the average-linkage method (references 21 and 22). FL, M- and UM-CLL samples are coded by red, blue, and green respectively. The matrix below the dendrogram depicts the gene expression values of the individual samples, with columns representing individual tumor samples and rows representing individual genes ordered according to hierarchical clustering. The color scale identifies relative gene expression changes normalized by the SD, with 0 representing the mean expression level of a given gene across the panel.
1630 Gene Expression Profiling of B Cell Chronic Lymphocytic Leukemia

The Gene Expression Profile of CLL Is Related to Memory B Cells. Phenotypic as well as IgV gene analyses could not conclusively identify the cell of origin of CLL. To address this issue, we compared the gene expression profiles of CLL to those of the major human B cell subsets, namely GC CD77+/H11001 (CBs) and CD77− (CCs) B cells, pre-GC (naive) and post-GC (memory) B cells (16–19), as well as GC-independent CD5-positive B cells. GC, naive (IgD+/H11001), and memory (CD27+/H11001) B cells were purified by magnetic cell separation from tonsillar mononuclear cells, while CD5+/CD27− B cells were isolated from umbilical cord blood (see Materials and Methods). Each B cell subset was isolated from five individuals, and cRNA generated from these fractions was hybridized to the U95A array as described previously.

To determine whether CLLs are more related to GC (CBs and CCs) or nonGC (naive and memory) B cells, we first compared the CLL profiles to those differentially expressed by these two B cell subgroups (Fig. 3 A). Genes that distinguish CB and CC cells from naive and memory B cells as identified by supervised clustering are shown in the left panel of Fig. 3 A, while the CLL samples are aligned to the right to visualize the expression of the respective genes in CLL cells. The classifier (23 genes; A) was applied to score 14 unpurified CLL cases (Table I), each identified by an open circle. The number of IgV gene mutations of each case is indicated within the open circle. The relatedness of each of the test samples to the two CLL subgroups is indicated by their proximity to either subgroup on the vertical axis (p-values are shown). The gray area marks the 95% confidence region, i.e., any sample with a score beyond this range can be assigned to one of the CLL subgroups with >95% confidence (see Materials and Methods). Supplemental table available at http://www.jem.org/cgi/content/full/194/11/1625/F2/DC1.
latedness to memory B cells is significant considering that memory and naive B cells differ in the expression of only \(~\sim 140\) of \(12,000\) genes (data not shown). While the difference was not statistically significant for the remaining six cases, none of the CLL cases was more related to naive than to memory B cells. The degree of relatedness to memory B cells was not significantly different for M- and UM-CLLs. Taken together, these results indicate that, independently of their IgV mutational status, CLLs are more related to memory cells than to naive, GC, or CD5\(^+\) B cells.

This observation prompted a direct analysis of the differences in gene expression profiles between CLL and memory B cells by supervised clustering. Fig. 4 shows the results of this analysis with the differentially expressed genes organized according to putative functional categories, including proliferation, apoptosis, cytokines/chemokines and receptors, adhesion, and cytokinesis. A number of proliferation-associated genes were downregulated in CLL cells (e.g., c-Myc, average fold difference in expression levels [CLL versus memory cells], ninefold; cyclin B, 13-fold; and E2-C, 80-fold). Apoptotic functions appear to be suppressed in CLL with the antiapoptotic gene Bcl-2 upregulated (fourfold) as expected (1), and various genes encoding proapoptotic molecules uniformly downregulated (BID, 33-fold; Rad9 [reference 30], ninefold; DRAK1 [reference 31], threefold; and DRAK2 [reference 31], fourfold). The IL-4 pathway appears to be activated in CLL, based on the observed previously (32) upregulated expression of the gene encoding the IL-4-receptor (sixfold), on the observed downregulation of SOCS-1 (Fig. 4), an inhibitor of the IL-4-signaling pathway (33) (19-fold), and consistent with the fact that CLL cells are responsive to IL-4 in vitro (1) (see Discussion). Finally, the expression of various genes encoding adhesion-associated molecules appears differentially regulated in CLL compared with memory B cells. Overall, these results indicate that CLL differs from memory B cells in the expression of numerous genes that suggest a more quiescent, antiapoptotic phenotype, with distinct cytokine and chemokine response and adhesion properties.

**Identification of Genes Specifically Expressed in CLL.** To identify genes specifically up or downregulated in CLL cells, we used supervised clustering to compare the gene expression profiles of CLL cases to those of normal B cell subsets (naive, CBs, CCs, memory) and to those derived from various nonHodgkin lymphoma subtypes, including FL, DLCL, and BL. Fig. 5 indicates that 32 genes are specifically expressed (or overexpressed), while \(>50\) genes appear downregulated in CLL. Several of the genes upregulated in CLL are involved in signal transduction pathways: CDC25 (average difference in expression level [CLL versus all others]: 14-fold over background) is a Ras guanine nucleotide exchange factor (34, 35). EPAC (35-fold over background) plays a role in the cAMP signal transduction pathway via Rap1 (36), a Ras-related guanine nucleotide exchange factor involved in B cell receptor signaling and oncogenesis (see Discussion). Significantly overexpressed genes include also: (i) the cell surface receptors Ror1 (19-fold), an orphan tyrosine kinase receptor, and the thromboxane A2 receptor (fourfold); and (ii) several genes related to TGF-\(\beta\) signaling, such as fibromodulin (>250 fold over background), associated with modulation of TGF-\(\beta\) signaling and cell adhesion (37), the TGF-\(\beta\)-inducible TIEG2, a SP1-like transcription factor (38) (10-fold over background), and BIGH3/TGFB1 (39) (38-fold). Genes specifically downregulated in CLL include genes involved in cell cycle progression and DNA replication and metabolism (E2-C, CIP2/KAP, CDC2, cyclin B, ribonucleotide reductase, thymidine kinase, dihydrofolate reductase, topoisomerase IIa), suggesting a markedly quiescent phenotype.

**Discussion**

Gene expression profiling allows a more comprehensive examination of cell phenotypes than the ones based on the analysis of individual or small numbers of genes, proteins, or signaling pathways. This technology has been used in this study to address three open questions regarding CLL: does CLL include one or two biological phenotypes since it displays a discordant IgV mutational status and clinical behavior? What is the cellular origin of CLL within the B cell lineage? Can a CLL–specific gene expression profile be identified? The results obtained directly address these questions and have implications relevant for the pathogenesis and, possibly, for the clinical management of this disease.

**CLLs Display a Common Phenotype Independent of IgV Mutations.** The gene expression profiles shown in Fig. 1 demonstrate that CLL has a characteristic gene expression profile that is clearly distinguishable from FL, a malignancy derived from mature B cells and displaying an indolent clinical course similar to CLL. This observation confirms the results obtained by Alizadeh et al. using a different DNA microarray technology (oligonucleotide versus cDNA-based arrays) and a only partially overlapping set of genes examined (12,000 nonlymphocyte-biased versus 18,000 lymphocyte-biased; reference 40).

These results also show that CLLs have a homogeneous phenotype independent of the presence of IgV mutations. This phenotype is defined by the common pattern of expression of 12,000 genes, with only 23 being differentially expressed in the M- versus UM-subtypes. Although significant differences in protein expression or modification cannot be excluded by gene expression profiling, this result does not support the hypothesis that CLL may include distinct biological entities (8, 13), as it is the case, for instance, for DLCL (40). Rather, these findings strongly suggest that all CLLs may derive from a common cell precursor through a common pathogenetic mechanism (see below).

**CLLs Are Related to Memory B Cells.** Based on various phenotypic similarities, mainly the presence of the CD5 marker and on initial reports that CLL displayed unmutated IgV, it has long been proposed that CLL may derive from the malignant transformation of CD5\(^+\) B cells (4), the human equivalent of mouse B1 cells. Subsequently, the identification of M-CLLs led to the hypothesis that a fraction of CLL cases may derive from a GC-experienced B cell, pos-
The gene expression profile of CLL is related to that of memory B cells. Gene expression data sets generated from 20 purified CLL cases are compared with the genes differentially expressed between CB/CC and memory/naive B cells (A), memory/naive B cells vs. CD5+ cells (B), and memory versus naive B cells (C). Genes differentially expressed between the various B cell subpopulations were identified by supervised clustering using Genes@Work. Matrices (a–c) and gene ranking are as in Fig. 2. Genes known to be differentially expressed among GC and non-GC B-cells and among naive and memory B cells, or on CLL cells (IL-4R), are indicated (references 19, 25–29). D shows the quantitative relatedness of the CLLs to the normal B cell populations as derived from panels A–C as described for Fig. 2 B. The gray area marks the 95% confidence region; the lower and upper margins of the gray area each correspond to a p-value of 0.025 (p-values decrease with increasing distance from the x-axis). Open and closed circles represent M- and UM-CLL cases, respectively.

Figure 3. The gene expression profile of CLL is related to that of memory B cells. Gene expression data sets generated from 20 purified CLL cases are compared with the genes differentially expressed between CB/CC and memory/naive B cells (A), memory/naive B cells vs. CD5+ cells (B), and memory versus naive B cells (C). Genes differentially expressed between the various B cell subpopulations were identified by supervised clustering using Genes@Work. Matrices (a–c) and gene ranking are as in Fig. 2. Genes known to be differentially expressed among GC and non-GC B-cells and among naive and memory B cells, or on CLL cells (IL-4R), are indicated (references 19, 25–29). D shows the quantitative relatedness of the CLLs to the normal B cell populations as derived from panels A–C as described for Fig. 2 B. The gray area marks the 95% confidence region; the lower and upper margins of the gray area each correspond to a p-value of 0.025 (p-values decrease with increasing distance from the x-axis). Open and closed circles represent M- and UM-CLL cases, respectively.
with the fact that CLLs lack chromosomal translocations which are thought to occur in developing B cells rearranging their antigen receptors or in GC B cells (see below).

The notion that CLL may all derive from memory B cells leaves open the question of why a sizable fraction of them expresses unmutated IgV genes, which are typical for pre-GC, naive B cells, or GC-independent cells involved in T cell-independent humoral responses. One possible explanation is provided by the observation that a minor fraction of CD27+ B cells carry unmutated IgV (17, 18) and may have entered the memory cell pool without acquiring IgV mutations, possibly due to already high affinity for the antigen. Such CD27+ unmutated memory B cells may therefore represent the normal counterpart of UM-CLLs. Alternatively, the precursor of UM-CLLs may be a cell that has acquired a memory phenotype via encounter with a T-independent (and thus GC-independent and IgV mutation-independent) pathway (9). This hypothesis implies that the antigen recognized by M-CLLs and UM-CLLs may be of different nature, consistent also with the observation that CDRIII length and IgV region usage differs among the two subgroups (9, 43). A difference in the nature and/or time of exposure to antigen may also explain the different clinical behavior of the two subgroups.

Overall, the finding that CLLs, including the UM subtype, display a gene expression profile similar to memory B cells warrants further investigations on the possible heterogeneity of the memory B cell pool (44, 45). The few genes that are differentially expressed in M- versus UM-CLLs (Fig. 2 A) may turn useful in dissecting this heterogeneity, although further characterization is needed because they represent a heterogeneous group of genes, some of unknown function.

Implications for CLL Pathogenesis. The observation that all CLLs share a common gene expression profile suggests that they derive from a common pathogenetic pathway. This notion is consistent with the relatively homogeneous cytogenetic profile of CLL, characterized by few common chromosomal abnormalities and, particularly, by the strong association with 13q14 deletions, present in the majority of cases (up to 70%) (46–48). These deletions are often present as a single chromosomal abnormality and are
thought to reflect the loss or inactivation of a still unknown tumor suppressor gene (46–48). The homogeneity of the gene expression profile of CLL suggests that this alteration or some functional equivalent (e.g., inactivating point mutations) may be present also in those cases lacking cytogenetically or molecularly detectable 13q14 deletions.

The observation that CLL is more related to memory B cells than to any other known normal B cell subset suggests that the multistep process leading to CLL may actually initiate in memory B cells. This hypothesis is consistent with the notion that, unique among lymphoid malignancies, CLL lack reciprocal balanced chromosomal translocations (2). These aberrations are thought to occur either during Ig VDJ recombination in maturing B cells, or during Ig hypermutation and isotype switch in mature B cells within the GC (11), and, accordingly, are common in B cell lymphoma, most of which derive from mature GC B cells (12, 49). Thus, if the transformation process leading to CLL initiates in memory B cells, it cannot involve chromosomal translocations since the mechanisms involved in these aberrations have been inactivated in these cells. This notion is consistent with the predominant presence in CLL of genetic alterations, such as deletions and amplifications, common in tumors deriving from tissues not physiologically subjected to antigen receptor gene rearrangements or hypermutation (50).

The comparative analysis of the gene expression profiles of CLL versus memory B cells (Fig. 4) or other normal and...
neoplastic cells (Fig. 5) provides a significant body of new information to dissect the CLL phenotype. First, the downregulation of a number of proapoptotic genes, together with the already known upregulation of the antiapoptotic molecule Bcl-2, is consistent with the documented long-lived apoptosis-resistant phenotype of CLL (1). Second, unique among normal and neoplastic B cells, CLL cells consistently display a significant overexpression of the EPAC and CDC25 genes, both encoding guanine nucleotide exchange factors that activate, respectively, Rap1 and Ras, the small GTPases that control pleiotropic transcriptional responses via the Raf/ERK pathway (51). This observation has possible pathogenetic relevance and warrants further studies since deregulation of the Ras pathway represents one of the most common alterations in human tumors (52), while recent evidence suggests that Rap1 can be deregulated by chromosomal translocations in lymphoid malignancy (53). Finally, the upregulation of mRNA for several cytokine or chemokine receptors compared with memory cells (IL-4R, TGF-β type III receptor, CCR7; Fig. 4) or to other B cells (Ror1 and thromboxane A2 receptor; Fig. 5) suggests that CLL cells may be abnormally responsive to certain stimuli. While each of the genes specifically expressed in CLL requires validation for differential

Figure 5. Identification of genes specifically expressed in CLL. Gene expression profiles of 10 randomly selected purified CLL cases (five UM- and five M-CLL, respectively) were compared with those generated from normal (CBs, CCs, naive, and memory) B cell subpopulations, purified nonHodgkin lymphoma cells (DLCL, BL, and FL), and DLCL and BL cell lines by supervised clustering using Genes@Work. Matrices and gene ranking are as in Fig. 2. Genes and fold change are indicated at the right; for GenBank accession no., see Supplementary Table III. Supplemental table available at http://www.jem.org/cgi/content/full/194/11/1625/F5/DC1.
expression at the protein level, these observations are potentially important for CLL pathogenesis and are amenable to experimental testing.

**Clinical Implications.** The results of these studies have potential clinical application in several areas. First, the ability to distinguish M- versus UM-CLLs by gene expression profiling (Fig. 2) represents a potential prognostic test for CLL since these two groups have a distinct clinical course. It is unlikely that this analysis will find clinical application using expensive microarrays with 12,000 genes such as the ones used in this study since IgV sequencing is more rapid and economical. Rather, these results should lead to the development of simple and inexpensive cytochemical assays recognizing the products of few of the genes differentially expressed in M- versus UM-CLLs. Second, the products of genes specifically expressed in CLL (Fig. 5) represent potential markers for the diagnosis of CLL, its differential diagnosis from related B cell malignancies, or the detection of small numbers of CLL cells in minimal residual disease contexts. Finally, the products of these same genes represent potential therapeutic targets based on their specific expression in CLL cells versus normal cells. The successful use of mAb therapy targeting the CD20 molecule in B cell lymphoma (54) suggests that cell surface receptors that are abnormally expressed in CLL (e.g., Ror1 and thromboxane A2 receptors) may represent good candidates for initial testing of this approach. Further DNA microarray studies may be aimed at determining the expression profiles of cytogenetically distinct CLL cases or cases displaying different clinical behavior (55).

We are grateful to Vladan Miljkovic for technical assistance. We also thank Laura Pasqualucci for discussions and Richard Baer for comments on the manuscript.

U. Klein was a recipient of fellowships granted by the European Molecular Biology Organization and Human Frontiers Science Program. G. Cattoretti is a recipient of an Abboodi Associate Professorship. H. Husson was supported by the Cure for Lymphoma Society, and the Norman Hirschfield Foundation.

Submitted: 14 June 2001
Revised: 7 August 2001
Accepted: 17 August 2001

**References**

1. Caligaris-Cappio, F., and T.J. Hamblin. 1999. B-cell chronic lymphocytic leukemia: a bird of a different feather. J. Clin. Oncol. 17:399–408.

2. Döhner, H., S. Stilgenbauer, K. Döhner, M. Bentz, and P. Lichter. 1999. Chromosome aberrations in B-cell chronic lymphocytic leukemia: reassessment based on molecular cytogenetic analysis. J. Mol. Med. 77:266–281.

3. Kipps, T.J. 1998. Chronic lymphocytic leukemia. Carr. Opin. Hematol. 5:244–253.

4. Dighiero, G., T. Kipps, H.W. Schroeder, N. Chiorazzi, F. Stevenson, L.E. Silberstein, F. Caligaris-Cappio, and M. Ferrarini. 1996. What is the CLL B-lymphocyte? Leuk. Lymphoma. 22:13–39.

5. Brezinschek, H.P., S.J. Foster, R.I. Brezinschek, T. Dörner, R. Domiati-Saad, and P.E. Lipsky. 1997. Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5+/IgM+ and CD5−/IgM+ B cells. J. Clin. Invest. 99:2488–2501.

6. Fischer, M., U. Klein, and R. Küppers. 1997. Molecular single-cell analysis reveals that CD5-positive peripheral B cells in healthy humans are characterized by rearranged Vκ genes lacking somatic mutation. J. Clin. Invest. 100:1667–1676.

7. Schroeder, H.W., Jr., and G. Dighiero. 1994. The pathogenesis of chronic lymphocytic leukemia: analysis of the antibody repertoire. Immunol. Today. 15:288–294.

8. Oscier, D.G., A. Thompsett, D. Zhu, and F.K. Stevenson. 1997. Differential rates of somatic hypermutation in V(H) genes among subsets of chronic lymphocytic leukemia defined by chromosomal abnormalities. Blood. 89:4153–4160.

9. Fais, F., F. Ghiotto, S. Hashimoto, B. Sellar, A. Valetto, S.L. Allen, P. Schuman, V.P. Vinciguerra, K. Rai, L.Z. Rassenti, et al. 1998. Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. J. Clin. Invest. 102:1515–1525.

10. Rajewsky, K. 1996. Clonal selection and learning in the antibody system. Nature. 381:751–758.

11. Küppers, R., U. Klein, M.L. Hansmann, and K. Rajewsky. 1999. Cellular origin of human B-cell lymphomas. N. Engl. J. Med. 341:1520–1529.

12. Stevenson, F., S. Sahota, D. Zhu, C. Ottensmeier, C. Chapman, D. Oiscer, and T. Hamblin. 1998. Insight into the origin and clonal history of B-cell tumors as revealed by analysis of immunoglobulin variable region genes. Immunol. Rev. 162:247–259.

13. Hamblin, T.J., Z. Davis, A. Gardiner, D.G. Oscier, and F.K. Stevenson. 1999. Unmutated Ig VHI genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood. 94:1848–1854.

14. Damle, R.N., T. Wasil, F. Fais, F. Ghiotto, A. Valetto, S.L. Allen, A. Buchbinder, D. Budman, K. Dittmar, J. Kolitz, et al. 1999. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood. 94:1840–1847.

15. Pasqualucci, L., A. Neri, L. Baldini, R. Dalla-Favera, and A. Migliaccia. 2000. BCL-6 mutations are associated with immunoglobulin variable heavy chain mutations in B-cell chronic lymphocytic leukemia. Cancer Res. 60:5644–5648.

16. Pasqualucci, L., A. Neri, L. Baldini, R. Dalla-Favera, and A. Migliaccia. 2000. BCL-6 mutations are associated with immunoglobulin variable heavy chain mutations in B-cell chronic lymphocytic leukemia. Cancer Res. 60:5644–5648.

17. Klein, U., K. Rajewsky, and R. Küppers. 1998. Human immunoglobulin (IgM+/IgD+) peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. J. Exp. Med. 188:1679–1689.

18. Tangye, S.G., Y.J. Liu, G. Aversa, J.H. Phillips, and J.E. de Vries. 1998. Identification of functional human splenic memory B cells by expression of CD148 and CD27. J. Exp. Med. 188:1691–1703.

19. Agematsu, K., S. Hokibara, H. Nagumo, and A. Komiyama.
2000. CD27: a memory B-cell marker. *Immunol. Today.* 21:204–206.

20. Ghia, P., V.A. Boussiotis, J.L. Schulze, A.A. Cardoso, D.M. Dorfman, J.G. Gibben, A.S. Freedman, and L.M. Nadler. 1998. Unbalanced expression of Bcl-2 family proteins in follicular lymphoma: contribution of CD40 signaling in promoting survival. *Blood.* 91:244–251.

21. Hartigan, J.A. 1975. Clustering Algorithms. Wiley, New York. 351 pp.

22. Eisen, M.B., P.T. Spellman, P.O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA.* 95:14863–14868.

23. Califano, A. 2000. SPLASH: structural pattern localization analysis by sequential histograms. *Bioinformatics.* 16:341–357.

24. Califano, A., G. Stolovitzky, and Y. Tu. 2000. Analysis of gene expression microarrays for phenotype classification. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 8:75–85.

25. Liu, Y.J., C. Barthelemy, O. de Bouteiller, C. Arpin, I. Durand, and J. Banchereau. 1995. Memory B cells from human tonsils colonize mucosal epithelium and directly present antigen to T cells by rapid up-regulation of B7-1 and B7-2. *Immunity.* 2:239–248.

26. Golay, J., V. Broccoli, G. Lamorte, C. Bifulco, C. Parravicini, A. Pizzey, N.S. Thomas, D. Delia, P. Ferrarini, D. Vitolto, and M. Introna. 1998. The A-Myb transcription factor is a marker of centroblasts in vivo. *J. Immunol.* 160:2786–2793.

27. Cattoretti, G., C.C. Chang, K. Cechova, J. Zhang, B.H. Ye, B. Falini, D.C. Louie, K. Offit, R.S. Chaganti, and R. Dalla-Favera. 1995. BCL-6 protein is expressed in germinal-center B cells. *Blood.* 86:45–53.

28. Onizuka, T., M. Moriyama, T. Yamochi, T. Kuroda, A. Kanazawa, N. Kanazawa, K. Sato, T. Kato, H. Ota, and S. Mori. 1995. BCL-6 gene product, a 92- to 98-kD nuclear phosphoprotein, is highly expressed in germinal center B cells and their neoplastic counterparts. *Blood.* 86:28–37.

29. van Der Vuurst De Vries, A.R., and T. Logtenberg. 1995. BCL-6 is a marker of centroblasts in vivo. *J. Immunol.* 160:2786–2793.

30. van Der Vuurst De Vries, A.R., and T. Logtenberg. 1999. A phage antibody identifying an 80-kDa membrane glycoprotein exclusively expressed on a subpopulation of activated B cells and hairy cell leukemia B cells. *Eur. J. Immunol.* 29:3989–3907.

31. Komatsu, K., T. Miyashita, H. Hang, K.M. Hopkins, W. Zheng, S. Cuddeback, M. Yamaeda, H.B. Lieberman, and H.G. Wang. 2000. Human homologue of accharomyces cerevisiae CDC25, a S functionally related to the CDC25 gene of Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA.* 97:7100–7104.

32. de Rooij, J., F.J. Zwartkruis, M.H. Verheijen, R.H. Cool, S.M. Nijman, A. Wittinghofer, and J.L. Bos. 1998. Epac is an Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP. *Nature.* 396:474–477.

33. Soo, C., F.Y. Hu, X. Zhang, Y. Wang, S.R. Beanes, H.P. Lorenz, M.H. Hedrick, R.J. Mackool, A. Plaas, S.J. Kim, et al. 2000. Differential expression of fibromodulin, a transforming growth factor-β modulator, in fetal skin development and scarless repair. *Am. J. Pathol.* 157:423–433.

34. Cook, T., B. Gebelein, K. Mesa, A. Mladek, and R. Urrutia. 1998. Molecular cloning and characterization of TIEG2 reveals a new subfamily of transforming growth factor-β-inducible Sp1-like zinc finger–encoding genes involved in the regulation of cell growth. *J. Biol. Chem.* 273:2929–2936.

35. Skonier, J., M. Neuhauser, L. Madisen, K. Bennett, G.D. Plowman, and A.F. Purchio. 1992. cDNA cloning and sequence analysis of β Ig-h3, a novel gene induced in a human adenocarcinoma cell line after treatment with transforming growth factor-β. *DNA Cell Biol.* 11:511–522.

36. Alizadeh, A.A., M.B. Eisen, R.E. Davis, C. Ma, I.S. Lossos, A. Rosenwald, J.C. Boldrick, H. Sabet, T. Tran, X. Yu, et al. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature.* 403:503–511.

37. Runheim, E.A., M.J. Cantwell, and T.J. Kipps. 1995. Expression of CD27 and its ligand, CD70, on chronic lymphocytic leukemia B cells. *Blood.* 85:3567–3565.

38. Punnoponen, J., G. Avensa, B.G. Cocks, and J.E. de Vries. 1994. Role of interleukin-4 and interleukin-13 in synthesis of IgE and expression of CD23 by human B cells. *Allogry.* 49:576–586.

39. Widhopf, G.F. II, and T.J. Kipps. 2001. Normal B cells express 51p1-encoded Ig heavy chains that are distinct from those expressed by chronic lymphocytic leukemia B cells. *J. Immunol.* 166:95–102.

40. Dono, M., S. Zupo, N. Leanza, G. Melioli, M. Fogli, A. Melagrina, N. Chiorazzi, and M. Ferrarini. 2000. Heterogeneity of tonsillar B lymphocytes, the splenic marginal zone equivalents. *J. Immunol.* 164:5596–5604.

41. Weller, S., A. Faili, C. Garcia, M.C. Braun, F.F. Le Deist, G.G. de Saint Basile, O. Hermine, A. Fischer, C. Reynaud, and J. Weill. 2001. CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. *Proc. Natl. Acad. Sci. USA.* 98:1166–1170.

42. Migliazza, A., F. Bosch, H. Komatsu, E. Cayanis, S. Martiniotti, E. Toniato, E. Guccione, X. Qu, M. Chien, V.V. Murty, et al. 2001. Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia. *Blood.* 97:2098–2104.

43. Mabuchi, H., H. Fujii, G. Calin, H. Alder, M. Negri, L. Rassenti, T.J. Kipps, F. Bullrich, and C.M. Croce. 2001. Cloning and characterization of CLLD6, CLLD7, and CLLD8, novel candidate genes for leukemogenesis at chromosome 13q14, a region commonly deleted in B-cell chronic lymphocytic leukemia. *Cancer Res.* 61:2870–2877.

44. Corcoran, M.M., O. Rasool, Y. Liu, A. Iyengar, D. Grander, R.E. Ibbotson, M. Merup, X. Wu, V. Brodysky, A.C. Gardner, et al. 1998. Detailed molecular delineation of 13q14.3 loss in B-cell chronic lymphocytic leukemia. *Blood.* 91:1637 Klein et al.
91:1382–1390.

49. Klein, U., T. Goossens, M. Fischer, H. Kanzler, A. Braeuninger, K. Rajewsky, and R. Küppers. 1998. Somatic hypermutation in normal and transformed human B cells. *Immunol. Rev.* 162:261–280.

50. Lengauer, C., K.W. Kinzler, and B. Vogelstein. 1998. Genetic instabilities in human cancers. *Nature.* 396:643–649.

51. Zwartkruis, F.J., and J.L. Bos. 1999. Ras and Rap1: two highly related small GTPases with distinct function. *Exp. Cell. Res.* 253:157–165.

52. Bos, J.L. 1989. ras oncogenes in human cancer: a review. *Cancer Res.* 49:4682–4689.

53. Hussey, D.J., M. Nicola, S. Moore, G.B. Peters, and A. Drobotvic. 1999. The (4;11)(q21;p15) translocation fuses the NUP98 and RAP1GDS1 genes and is recurrent in T-cell acute lymphocytic leukemia. *Blood.* 94:2072–2079.

54. White, C.A., R.L. Weaver, and A.J. Grillo-Lopez. 2001. Antibody-targeted immunotherapy for treatment of malignancy. *Annu. Rev. Med.* 52:125–145.

55. Stratowa, C., G. Löffler, P. Lichter, S. Stilgenbauer, P. Habel, N. Schweifer, H. Döhner, and K.K. Wilgenbus. 2001. cDNA microarray gene expression analysis of B cell chronic lymphocytic leukemia proposes potential new prognostic markers involved in lymphocyte trafficking. *Int. J. Cancer.* 91:474–480.

56. Matsuda, F., E.K. Shin, H. Nagaoka, R. Matsumura, M. Haino, Y. Fukita, S. Taka-ishi, T. Imai, J.H. Riley, R. Anand, et al. 1993. Structure and physical map of 64 variable segments in the 3′0.8-megabase region of the human immunoglobulin heavy-chain locus. *Nat. Genet.* 3:88–94.