Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas

Izidore S. Lossos*, Ash A. Alizadeh1, Michael B. Eisen1, Wing C. Chan1, Patrick O. Brown15, David Botstein9, Louis M. Staudt1, and Ronald Levy**.***

*Departments of Medicine, †Biochemistry, and ‡Genetics, Division of Oncology and §Howard Hughes Medical Institute, Stanford University Medical Center, Stanford, CA 94305-5306; †Department of Pathology, University of Nebraska Medical Center, Omaha, NE 68198; and ‡Metabolism Branch, Division of Clinical Sciences, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892

Contributed by David Botstein, July 7, 2000

B cell diffuse large cell lymphoma (B-DLCL) is a heterogeneous group of tumors, based on significant variations in morphology, clinical presentation, and response to treatment. Gene expression profiling has revealed two distinct tumor subtypes of B-DLCL: germinal center B cell-like DLCL and activated B cell-like DLCL. In a separate study, we determined that B-DLCL can also be subdivided into two groups based on the presence or absence of ongoing Ig gene hypermutation. Here, we evaluated the correlation between these B-DLCL subtypes established by the two different methods. Fourteen primary B-DLCL cases were studied by gene expression profiling using DNA microarrays and for the presence of ongoing mutations in their Ig heavy chain gene. All seven cases classified as germinal center B cell-like DLCL by gene expression showed the presence of ongoing mutations in their Ig heavy chain gene. All seven cases classified as germinal center B cell-like DLCL by gene expression showed the presence of ongoing mutations in the Ig genes. Five of the seven cases classified by gene expression as activated B cell-like DLCL had no ongoing somatic mutations, whereas, in the remaining two cases, a single point mutation was observed in only 2 of 15 and 21 examined molecular clones of variable heavy (VH) chain gene, respectively. These two cases were distantly related to the rest of the activated B cell-like DLCL tumors by gene expression. Our findings validate the concept that lymphoid malignancies are derived from cells at discrete stages of normal lymphocyte maturation and that the malignant cells retain the genetic program of those normal cells.

B cell diffuse large cell lymphoma (B-DLCL) is the most common type of human lymphoma, accounting for approximately 40% of all non-Hodgkin’s lymphoma (NHL) (1). Lymphomas classified as B-DLCL, according to the Revised European-American Lymphoma (REAL) classification, are likely to constitute more than one disease entity, as suggested by the striking heterogeneity of morphology, clinical presentation, and response to treatment (2). Indeed, in our recent work on gene expression profiling by DNA microarray techniques, we identified two molecularly distinct forms of B-DLCL: germinal center B cell-like DLCL characterized by expression of genes normally expressed in germinal center B cells, and activated B cell-like DLCL characterized by expression of genes normally induced during in vitro activation of B cells (3). Patients with these two forms of B-DLCL were found to have very different prognosis: those with germinal center B cell-like DLCL had a significantly better overall survival than those with activated B cell-like DLCL.

B cell malignancies are considered to arise from normal lymphocytes at different stages of B cell differentiation. However, the extent to which these tumors maintain the molecular and physiological properties of their normal B-cell counterparts is not clear. The examination of the variable heavy (VH) chain of Ig gene, which is usually rearranged and expressed as a unique clonal surface Ig receptor in majority of B cell lymphomas, is a useful approach to trace the stage of neoplastic transformation. Previous studies of B-DLCL demonstrated that the majority of these tumors contain mutated VH genes in a pattern that suggested antigen selection pressure (4–11). These mutations are considered to be the result of the somatic mutation process that normally occurs only in germinal centers of secondary lymphoid organs. However, whether the mutation process is still ongoing in B-DLCL was a controversial issue (4–5, 10–15).

In our recent study, we demonstrated the presence of two subgroups of B-DLCL: one with and another without ongoing somatic mutation (9). However, it is unclear whether the presence of ongoing somatic mutation distinguishes between B-DLCL derived from different stages of ontogenesis (germinal center or post germinal center lymphocytes) or whether it is a property of malignant cells irrespective of their developmental origin. To answer these questions and to examine the relation between ongoing somatic mutation and the newly recognized distinction between germinal center B cell-like DLCL and activated B cell-like DLCL, we analyzed the relationship between gene expression patterns and the presence of ongoing mutation in the Ig VH region in a series of primary B-DLCL tumors.

Methods

**Tissue Samples, mRNA Preparation, and Microarray Procedures.** Germinal center B cells, centroblasts, and human blood B cells were purified from human tonsils and adult apheresis products, respectively, as previously described (16). In vitro stimulation of peripheral B cells using anti-IgM antibody, IL-4, and/or CD40 ligand-containing membranes was done as described (17).

Tumor tissues frozen in OCT from 42 untreated patients with primary B-DLCL according to the REAL classification and 3 B-DLCL cell lines (SUDHL6, OCI Ly3, and OCI Ly10) were studied. mRNA from the cryopreserved B-DLCL specimens, B-DLCL cell lines, and normal B cells were isolated by using FAST TRACK 2.0 kit (Invitrogen). DNA microarray analysis of gene expression was done as described (3, 18). In each experiment, fluorescent cDNA probes were prepared from an exper-

Abbreviations: B-DLCL, B cell diffuse large cell lymphoma; NHL, non-Hodgkin’s lymphoma; VH, variable heavy; CLL, chronic lymphocytic leukemia; CDR, complementary determining region; FR, framework region; FCL, follicular center lymphoma.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF283779-AF283807).

To whom reprint requests should be addressed. E-mail: levy@leland.stanford.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: Proc. Natl. Acad. Sci. USA, 10.1073/pnas.180316097.

Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.180316097
temporal mRNA sample (Cy5-labeled) and a control mRNA sample (Cy3-labeled) isolated from a pool of nine lymphoma cell lines (Raji, Jurkat, L428, OCI-Ly3, OCI-Ly8, OCI-Ly1, SUDHL5, SUDHL6, and WSU1). The use of a common control cDNA allows the relative expression of each gene to be compared across all of the samples. The results of DNA microarray analysis for all but the specimen DLCL-0047 were previously reported (3). Agglomerative hierarchical clustering was applied to both axes (genes and tumors) by using the CLUSTER program (M. Eisen: http://www.microarrays.org/software) (19). Sufficient mRNA for analysis of somatic mutations was available for 14 specimens (DLCL-0004, 0005, 0006, 0009, 0010, 0011, 0012, 0013, 0015, 0018, 0021, 0026, 0041, and 0047). These 14 B-DLCL samples were used in the present study. The gene expression pattern was reanalyzed for the entire group, now including specimen DLCL-0047. Cytogenetic data were present for 6 of these 14 B-DLCL tumors (specimens 0005, 0009, 0010, 0013, 0015, and 0021). None of these specimens had the chromosomal translocation t(14;18).

**Analysis of Ig V<sub>H</sub> Gene Somatic Mutations.** Analysis of Ig V<sub>H</sub> gene somatic mutations in 14 B-DLCL specimens for which sufficient mRNA was available following DNA microarray experiments was performed as previously described (9). Briefly, the mRNA was reverse transcribed with AMV Reverse Transcriptase (Promega). The cDNA was amplified by Taq DNA polymerase (GIBCO/BRL) with a specific 5′ primer corresponding to one of the six human V<sub>H</sub> family leader genes (V<sub>H</sub>1 through V<sub>H</sub>6) and a 3′ anti-sense joining (J) consensus primer in PCR whose conditions were previously reported. For each PCR, a control with no added template was used to check for contamination. PCR products of appropriate size were excised from 2% agarose gel and purified by adsorption to a silica matrix (QIAquick columns; Qiagen, Chatsworth, CA). Direct DNA sequencing of PCR amplicons was performed on an 373 automatic DNA sequencer (Applied Biosystems) using ABI Prism Big Dye Terminator Kit (Perkin–Elmer) as recommended by the manufacturer. The same primers used for the PCR were used for sequencing. The sequence was defined as clonal if identical or nearly identical CDR3 sequences were obtained from two independent PCR reactions. The clonal V<sub>H</sub> chain amplicons obtained from a third independent PCR reaction were cloned into a TA-PCR cloning vector (Invitrogen). At least 15 molecular clones were sequenced per tumor sample. Sequence analysis was performed by using the MacVector program (Oxford Molecular Group, Campbell, CA). Sequences were aligned with the database of sequences sequenced in our laboratory, to exclude possibility of contamination, and germ line sequences derived from Vβ base database and DNA plot (http://www.mrc-cpe.cam.ac.uk/imt-doc/restricted/ok.html). The V<sub>H</sub> gene sequences were compared with the germ line genes with the highest homology and, accordingly, the number of somatic mutations was determined. Mutations at the last nucleotide position of the sequenced fragment were excluded from the mutational analysis because they might result from nucleotide deletion at the joining sites. Percent of sequence identity was calculated from the aligned sequences from the beginning of FR1 to the end of FR3. For evaluation of ongoing somatic mutations, the following definitions were used (9): unconfirmed mutation—a substitution mutation observed in only one of the V<sub>H</sub> gene molecular clones from the same tumor specimen; confirmed mutation—a mutation observed more than once in the V<sub>H</sub> gene molecular clones from the same tumor specimen. Only the confirmed mutations were considered as an evidence of ongoing somatic mutations, whereas the unconfirmed mutations were disregarded, because they can be caused by reverse transcription-Taq DNA polymerase error rate. The reverse transcription-Taq DNA polymerase error rate in our laboratory is 0.09%, which may amount to 0.36 mutations/V<sub>H</sub> clone. Results of V<sub>H</sub> gene cloning of 5 B-DLCL specimens were previously reported (9).

**Statistical Analysis.** The comparison between B-DLCL groups was performed by two-sided Fisher’s exact test. P value < 0.05 was considered as statistically significant.

The probability that an excess or scarcity of replacement (R) mutations in V<sub>H</sub> complementary determining regions (CDRs) or framework regions (FRs) occurred by chance was calculated by a multinomial distribution model, as was previously described (9).

**Results**

**DNA Microarray Analysis.** The 14 B-DLCL specimens chosen for the current study were analyzed along with a larger set of B-DLCL samples, using the 17,856 known and unknown genes that constitute the lymphochip (3). An agglomerative hierarchical clustering algorithm was used to cluster the tumors, based on their pattern of expression of a subset of 380 genes. These genes were chosen because they had been found to stratify the two B-DLCL subtypes, germinal center-derived and activated B cell-derived (3). As shown in Fig. 1, 7 of the 14 patients occupied a common branch and showed a gene expression profile with persuasive similarities to that of normal germinal center B cells (i.e., GC B cell-like). The remaining seven patients were clustered together on another common branch, with a gene expression pattern that shared many similarities to that of in vitro activated normal peripheral blood B cells (i.e., activated B cell-like).

**V<sub>H</sub> Gene Somatic Mutation Analysis.** Clonal and in frame V<sub>H</sub> gene sequences without stop codons were detected in all of the tested B-DLCL specimens. Clonal V<sub>H</sub> genes from the germinal center B cell-like DLCL derived from the V<sub>H</sub>1 and the V<sub>H</sub>3 families (Table 1), whereas V<sub>H</sub> genes from the activated B cell-like DLCL were derived from the V<sub>H</sub>1, V<sub>H</sub>3, and V<sub>H</sub>4 families. The small numbers of specimens tested preclude any conclusions regarding V<sub>H</sub> gene use in these two subgroups of B-DLCL. All of the V<sub>H</sub> gene sequences contained mutations from their most closely related germ line V<sub>H</sub> gene. No significant difference in the extent of V<sub>H</sub> gene mutation and in the distribution of replacement (R) vs. silent (S) mutations between the CDRs and FRs (Table 2) was observed between the germinal center B cell-like and activated B cell-like DLCL subgroups, strongly suggesting that all of the tumors had arisen from B cells that had resided in a germinal center at one point in time. Evaluation for the presence or absence of ongoing somatic mutations disclosed a significant difference between the two subgroups (P = 0.021). All of the germinal center B cell-like DLCL specimens demonstrated the presence of ongoing somatic mutations in the range of 1 to 37 mutations, that were present in 2 to 8 of the 14–19 tested tumor-derived molecular clones (Table 1). In contrast, five of the seven activated B cell-like DLCL showed no evidence of ongoing somatic mutations. The remaining two activated B cell-like DLCL specimens disclosed only one ongoing mutation per specimen that was present in only two of the 15–21 tested tumor-derived molecular clones (Table 1). These two specimens occupied separate branches on the hierarchical cluster dendrogram from the remaining activated B-like DLCL cases (Fig. 1).

**Discussion**

Ig somatic hypermutation is known to occur in the germinal centers and to cease in postgerminal center cells. We have previously determined that B-DLCL can be divided into two major subclasses based on the pattern of gene expression: a
germinal center B cell-like and an activated peripheral B cell-like type (3). In the present study, we demonstrate that Ig somatic mutation is ongoing in all germinal center B cell-like DLCL but it is switched off in the majority of activated B cell-like DLCL. By using an independent methodology, we support our previous gene expression analysis findings of the presence of two distinct DLCL subtypes.

Analysis of V<sub>H</sub> gene mutations is helpful in assigning the cell of origin of a B cell tumor to a specific developmental stage in the complex process of normal B cell maturation (20). For

**Table 1. Analysis of ongoing somatic mutations and gene profiling in B cell diffuse large cell lymphoma (B-DLCL)**

<table>
<thead>
<tr>
<th>B-DLCL no.</th>
<th>Biopsy site</th>
<th>B-DLCL type by gene profiling</th>
<th>Most similar germ line V&lt;sub&gt;H&lt;/sub&gt; gene</th>
<th>% Homology*</th>
<th>Tumor derived clones/ clones sequenced</th>
<th>No. of confirmed ongoing mutations</th>
<th>No. of molecular clones with ongoing mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>0004</td>
<td>Thyroid</td>
<td>GCBL</td>
<td>1-03</td>
<td>89</td>
<td>16/16</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>0010</td>
<td>LN</td>
<td>GCBL</td>
<td>3-23</td>
<td>97</td>
<td>15/19</td>
<td>37</td>
<td>7</td>
</tr>
<tr>
<td>0012</td>
<td>Spleen</td>
<td>GCBL</td>
<td>3-30</td>
<td>87</td>
<td>19/19</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>0009</td>
<td>LN</td>
<td>GCBL</td>
<td>3-33</td>
<td>82</td>
<td>18/21</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>0015</td>
<td>LN</td>
<td>GCBL</td>
<td>3-48</td>
<td>90</td>
<td>14/27</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0018</td>
<td>AT</td>
<td>GCBL</td>
<td>3-48</td>
<td>87</td>
<td>19/20</td>
<td>1&lt;sup&gt;1&lt;/sup&gt;</td>
<td>2</td>
</tr>
<tr>
<td>0026</td>
<td>LN</td>
<td>GCBL</td>
<td>3-11</td>
<td>90</td>
<td>18/20</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>0005</td>
<td>Lung</td>
<td>ABCL</td>
<td>1-1</td>
<td>82</td>
<td>9/17</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0006</td>
<td>LN</td>
<td>ABCL</td>
<td>1-69</td>
<td>90</td>
<td>19/20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0011</td>
<td>LN</td>
<td>ABCL</td>
<td>3-23</td>
<td>90</td>
<td>21/21</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0013</td>
<td>Lung</td>
<td>ABCL</td>
<td>3-23</td>
<td>96</td>
<td>19/19</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0047</td>
<td>Stomach</td>
<td>ABCL</td>
<td>4-39</td>
<td>88</td>
<td>17/22</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0021</td>
<td>LN</td>
<td>ABCL</td>
<td>4-34</td>
<td>99</td>
<td>15/15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0041</td>
<td>Skin</td>
<td>ABCL</td>
<td>4-34</td>
<td>97</td>
<td>15/15</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

LN, lymph node; AT, abdominal tumor; GCBL, germinal center B cell-like; ABCL, activated B cell-like.
*Percent V<sub>H</sub> gene homology of the most abundant sequence.
<sup>1</sup>Clones sharing identical or near identical CDR3 sequence.
<sup>1</sup>A deletion mutation causing out of frame Ig sequence.

Fig. 1. Clustering of B-DLCL specimens and their relationship to normal B lymphocyte differentiation and activation. (A) Hierarchical clustering of B-DLCL specimens and germinal center B cells using a previously reported set of 380 genes, selectively expressed in GC B-like DLCL and Activated B-like DLCL as was previously reported (3). (Germinal center B-like DLCL in orange and Activated B-like DLCL in blue). The 12 B-DLCL specimens studied for the presence of ongoing somatic mutations of immunoglobulin gene are indicated in red. (B) Gene expression data from germinal center, activated and resting blood B cells, as was previously reported (3). Each row represents a separate cDNA clone on the microarray and each column a separate mRNA sample.
simplicity, the B cell maturation process can be divided into four stages. The first stage is a B cell progenitor stage before V(D)J recombination. The second stage is the generation of naïve B cells following V(D)J recombination, but before encounter with antigen; it is usually considered to be a pre-antigenic stage because VH genes are not mutated. The third stage occurs in the germinal center microenvironment and includes somatic mutation, antigen selection and isotype switch events. The fourth stage is the post germinal center stage, which corresponds to memory B cells and plasma cells, in which the mutational process had been accomplished leading to the production of high affinity antibodies. The stimuli responsible for transition of cells from stage three to memory or plasma cells are unclear. However, it has been suggested that centrocytes treated with antibody against CD40 and IL-4 represent the memory B blasts that proliferate in follicles after the germinal center reaction is over (21). We used these and other stimuli combinations to establish gene expression in activated B cells, which appears to be different from the gene expression pattern of the germinal center B cells (3).

The normal germinal centers are oligoclonal, usually formed by colonization of an average of three B blasts in each germinal center (22, 23). The cells in germinal centers demonstrate an extremely high rate of proliferation and apoptosis. The somatic mutation process is a pathognomonic feature of the germinal center microenvironment and normally requires the presence of both T cells and antigen, presented on follicular dendritic cells (21). It affects Ig genes leading to antibody affinity maturation (24) and it also affects the S' noncoding region of the BCL-6 gene (25–27), a process which may play an important role in the pathogenesis of B-DLCL. The combination of a high proliferative rate, clonal expansion of B cells, and the presence of active mutational machinery may be a perfect environment for malignant transformation. Indeed, follicular center lymphomas are considered to originate from germinal center lymphocytes (28, 29) and continue to express genes characteristic of this stage of differentiation (3). All follicular center lymphomas demonstrate intraclonal heterogeneity of their expressed Ig VH genes—a hallmark of ongoing somatic mutation. There has been some controversy about the cell of origin of B-DLCL as to whether it arises from a germinal center B cell or a post germinal center B cell (4, 5, 8, 10, 30). For example, CD10, a germinal center marker, is expressed in some but not all B-DLCL cases (31, 32). In contrast, follicular center lymphomas (FCL), which represent a germinal center tumor prototype, are characterized by almost uniform CD10 expression. The relationship between immunophenotype and the presence of ongoing mutations has not been tested in B-DLCL.

In the present study, we examined the correlation between ongoing somatic mutations and gene expression profiles of B-DLCL. Gene profiling by DNA microarray methods reveals a molecular phenotype by the simultaneous evaluation of thousands of genes. In our previous work, we found a high correlation between the germinal center gene expression pattern and prolonged survival of patients with B-DLCL. Another germinal center B cell tumor, FCL, is also characterized by a prolonged survival and an indolent clinical course. It is possible that B cell tumors, which are still responding to antigenic selection, are less malignant in their clinical behavior. In addition, these results support the idea that B cell differentiation stage at which malignant transformation occurs carry significant prognostic relevance. Similar findings were recently reported for B cell Chronic Lymphocytic Leukemia (CLL) (33, 34), a malignancy that can apparently originate either from naïve B cells, harboring germ-line Ig genes, or from post germinal center B lymphocytes, harboring mutated Ig genes. CLL cases with mutations of Ig genes were found to have a significantly longer survival.

In conclusion, the present study demonstrates that ongoing somatic mutation of Ig genes is a feature of germinal center B cell-like but not of activated B cell-like DLCL. Gene expression analysis and the ongoing mutation analysis both support the idea that the initial transforming event leading to germinal center B cell-like DLCL occurs in a B cell within the germinal center microenvironment and does not alter the ability of these cells to mutate their Ig genes. In this respect, these germinal center B cell-like DLCL resemble FCL. Conversely, the activated B-like DLCL must either (i) derive from the postgerminal center B-cells in which the somatic hypermutation machinery is no longer active or (ii) derive from germinal center B cell by a transformation event that changes germinal center gene expression pattern and stops the mutation process. Furthermore, our present results raise the possibility of an additional B-DLCL subgroup with gene expression features of activated B cell-like DLCL tumors, but also a low level of ongoing somatic mutation.

Further gene profiling studies may provide genetic explanation for the marked clinical heterogeneity of each lymphoma type and may enable better patient stratification for evaluation of the new therapeutic modalities.

This work was supported by Grants U01 CA84967, CA33399, and CA34233 from the U.S. Public Health Service-National Institutes of Health.
Health and by Howard Hughes Medical Institute. I.S.L. is a Harrold Dobbs Oncology Fellow. P.O.B. is an Associate Investigator of the Howard Hughes Medical Institute. R.L. is an American Cancer Society Clinical Research Professor.