Immunoglobulin gene analysis reveals 2 distinct cells of origin for EBV-positive and EBV-negative Burkitt lymphomas

Cristiana Bellan, Stefano Lazzi, Michael Hummel, Nazzareno Palumbo, Margherita de Santi, Teresa Amato, Joshua Nyagol, Elena Sabattini, Thierry Lazure, Stefano A. Pileri, Martine Raphael, Harald Stein, Piero Tosi, and Lorenzo Leoncini

The normal counterpart of the neoplastic B cells in Burkitt lymphoma (BL) is still unclear. Based on immunoglobulin gene rearrangement studies, some authors suggest an origin from germinatal center cells and others from memory B cells. However, most of these studies rely on cell lines or on a small series of cases. To help clarify the cell of origin of BL, semi-nested polymerase chain reaction (PCR) was performed to amplify the VDJ rearrangements of the immunoglobulin heavy chain (V_H) genes, and the resultant amplificates were sequenced for comparison with known germline V_H segments. The results of this approach revealed that all cases (15 endemic BL [eBL], 10 sporadic BL [sBL], and 6 AIDS-related BL) harbor mutated V_H genes, with different mutation ranges among the 3 types of BL. The eBL and AIDS-related forms showed considerably higher mutation rates than the sBL form (5.1%, 5.4%, and 1.5%, respectively). The mutations in eBL and AIDS-related BL also showed signs of antigen selection, whereas no signs of antigen selection were found in sBL. Finally, after subcloning the amplificates, sequence analysis revealed no signs of ongoing mutations in any of the cases analyzed. Given that one of the main differences between eBL and AIDS-related BL on the one hand and sBL on the other hand is the association with Epstein-Barr virus (EBV), we compared EBV-positive and EBV-negative BLs independently of their geographic origin and HIV status. The differences in the number of somatic mutations and antigen selection were even more evident when this approach was used. According to our molecular results, it appears that EBV-positive and EBV-negative BL may originate from 2 distinct subsets of B cells, pointing to a particular role for the germinal-center reaction in the pathogenesis of these tumors. The different types of C-MYC translocation reported in BL may also be related to the different stages of B-cell maturation. (Blood. 2005;106:1031-1036) © 2005 by The American Society of Hematology

Introduction

Burkitt lymphoma (BL) is an aggressive B-cell malignancy described for the first time by Dennis Burkitt in 1958. There are 2 major subgroups of BL that differ in geographic distribution. The first, endemic Burkitt lymphoma (eBL), affects African children and, through epidemiological studies, has been demonstrated to be endemic in equatorial Africa, New Guinea, and large areas of South America. The second subgroup, sporadic Burkitt lymphoma (sBL), occurs worldwide and mainly affects young adults. In endemic areas the tumor most often involves facial bones, whereas the terminal ileum and lymph nodes are more commonly involved in sBL. Besides geographic distribution and clinical manifestation, the 2 forms of BL differ in Epstein-Barr virus (EBV) infection status: eBL is almost always associated with EBV, whereas sBL has a more irregular association, ranging from 10% to 30% positivity in different areas.

The incidence of BL has recently increased because it is a common neoplasm in HIV-infected patients, and these lymphomas are currently listed as AIDS-related BL according to the World Health Organization (WHO) classification. However, it is unclear why BL develops frequently in HIV-positive patients but not in patients with other forms of immunodeficiency. The association with EBV is variable among AIDS-related BL. In particular, most AIDS-related BL in Western countries are EBV-negative, whereas in Africa they are strongly associated with EBV.

All types of BLs carry a characteristic reciprocal chromosomal translocation involving chromosome 8 at the site of the C-MYC proto-oncogene and 1 of the 3 immunoglobulin gene loci on chromosome 14, 2, or 22. The translocation event appears to contribute to subsequent malignant transformation, although other molecular abnormalities or epigenetic changes may also be required.

Because BLs have a tendency to home to germinatal centers (GCs), to morphologically resemble GC cells, and to express characteristic GC cell markers such as CD10 and BCL6, some authors favor the derivation of BL from GC blasts. The site of tumor growth is, however, often extranodal, such as in the jaw or ovary, where tissues do not contain germinal centers under physiological conditions.

Despite intensive research, the normal counterpart of the neoplastic B cells occurring in BL is still unclear and may not be the same in eBL and sBL. Based on immunoglobulin gene rearrangement studies, some authors suggest an origin from...
germinal center cells, whereas others suggest memory B cells.\textsuperscript{17-22} However, most of these studies rely on cell lines or on a small series of cases, and eBL has rarely been included.

To help clarify whether BL cases are related to GC cells or to post-GC B cells, we amplified the IgH VDJ rearrangement using DNA extracted from a series of BL (15 eBL, 10 sBL, 6 AIDS-related BL) and analyzed the sequences for the presence of somatic mutations, also in relation to EBV status.

Materials and methods

Selection of cases, immunohistochemistry, Epstein-Barr virus, and C-MYC translocation

Formalin-fixed and paraffin-embedded specimens were collected from the Department of Human Pathology and Oncology, University of Siena, Italy; “L.A. Seragnoli” Haematopathology Unit, Policlinico S. Orsola, Bologna, Italy; Hematology Department and Surgical Pathology Department, Bicêtre Hospital, INSERM E109, University Paris XI, France; and the Department of Human Pathology, Nairobi Hospital, Kenya. Sixty-eight cases were collected; 45 came from hot, moist woodland and savanna country in the Lake Victoria region and from coastal lowlands with endemic malaria in Kenya, and 23 came from Europe. HIV status was determined by enzyme-linked immunosorbent assay on patients’ sera. Five cases of mantle cell lymphoma (MCL) and 8 cases of follicular lymphoma (FL) were collected from the Department of Human Pathology and Oncology, University of Siena, Italy, and were used as controls. The cases were reviewed by expert pathologists (L.L., M.R., S.P.), and diagnoses were established according to the criteria of the WHO classification.

Immunohistochemical studies were performed on representative paraffin sections from each case using microwave pretreatment of slides for antigen retrieval.\textsuperscript{23} A large panel of antibodies recognizing formalin-resistant epitopes of the various antigens was applied (Table 1), in conjunction with the alkaline phosphatase antialkaline phosphatase (APAAP) method, to visualize antibody binding.\textsuperscript{24,25}

EBV RNA in situ hybridization (ISH) was performed using a 30-base oligonucleotide probe complementary to a portion of the EBV-encoded small nuclear RNA (\textit{EBER1}) gene, as described elsewhere.\textsuperscript{26} Five-micrometer–thick paraffin sections were deparaffinized, rehydrated, predigested with proteinase K, and hybridized overnight at a concentration of 0.25 ng/\mu l, biotinylated probe. Detection was accomplished using an avidin-alkaline phosphatase conjugate. A control slide, prepared from a L biotinylated probe. Detection was accomplished using an \textit{H}9262

Table 1. Panel of antibodies used for diagnosis

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-L26</td>
<td>DAKO</td>
</tr>
<tr>
<td>Anti-CD79a</td>
<td>Bio-Optica</td>
</tr>
<tr>
<td>Anti-Ki67</td>
<td>Bio-Optica</td>
</tr>
<tr>
<td>Anti-IgM</td>
<td>DAKO</td>
</tr>
<tr>
<td>Anti-IgA</td>
<td>DAKO</td>
</tr>
<tr>
<td>Anti-IgG</td>
<td>DAKO</td>
</tr>
<tr>
<td>Anti-IgD</td>
<td>DAKO</td>
</tr>
<tr>
<td>Anti-\kappa</td>
<td>DAKO</td>
</tr>
<tr>
<td>Anti-\lambda</td>
<td>DAKO</td>
</tr>
<tr>
<td>Anti-C-MYC</td>
<td>Bio-Optica</td>
</tr>
<tr>
<td>Anti-p53</td>
<td>DAKO</td>
</tr>
<tr>
<td>Anti-Bcl-2</td>
<td>Bio-Optica</td>
</tr>
<tr>
<td>Anti-Bcl-6</td>
<td>DAKO</td>
</tr>
<tr>
<td>Anti-CD10</td>
<td>Neomarkers</td>
</tr>
<tr>
<td>Anti-LMP-1</td>
<td>DAKO</td>
</tr>
</tbody>
</table>

Each source is in Milan, Italy, except Neomarkers, which is located in Fremont, California. LMP indicates latent marker protein.

C-MYC translocation for the cases of European origin was established at the time of diagnosis using different methods according to the facilities of the respective laboratories from which the cases were collected (Southern blot technique, classical cytogenetic analysis, and fluorescence in situ hybridization).\textsuperscript{27,28} For the African cases, chromogen in situ hybridization (CISH) was applied in paraffin-embedded tissue and gave a positive result in 36 cases. The 7 cases in which C-MYC translocation could not be detected were excluded from this study.

DNA extraction and DNA quality control

For \textit{V}_{H} analysis, DNA was extracted from 20-\mu m–thick paraffin sections using a DNA extractor (BioRobot EZ1; Qiagen S.p.a., Milan, Italy), in accordance with the protocols and using the reagents from the same supplier. DNA was then dissolved in 50 to 100 \mu l distilled water, and its concentration was quantified photometrically before it was used as a template for PCR. Control gene primer sets for quality assessment of DNA for paraffin-embedded sections were applied according to the BIOMED 2 protocol.\textsuperscript{29} DNA quality control PCR showed that only 31 of the 68 BL cases gave amplificates of 300 bp or more and, thus, sufficient intact DNA for successful amplification of \textit{IgH} rearrangements.

PCR amplification of the rearranged \textit{V}_{H} gene immunoglobulin

To avoid cross-contamination, all procedures performed before PCR amplification (eg, proteinase K digestion) were performed in a room exclusively dedicated to this purpose and separate from the rooms in which the subsequent steps were carried out.

DNA extracted from paraffin sections was amplified using a PCR method in a Master Cycler (Eppendorf, Hamburg, Germany) with different sets of primers: FR2A, FR2FS, FR2BM, and FR3BM. PCR conditions and primers have been described in detail elsewhere.\textsuperscript{17,29,30}

Cloning and sequence analysis

Purified amplicons were cloned to the pCR-TOPO vector (TOPO TA Cloning Kit; Invitrogen, Paisley, United Kingdom) and were used to transform TOP 10 cells (Invitrogen) according to the manufacturer’s instructions. At least 10 clones per case were sequenced in both directions with an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Weiters- tadt, Germany) using Sp6 reverse and T7 forward primers, respectively. Sequences were compared against the GenBank database using the IgBlast database (http://www.ncbi.nlm.nih.gov/igblast/). Using high-fidelity Taq polymerase (Invitrogen), the PCR error rate for \textit{V}_{H} gene immunoglobulin was less than 1 base pair (bp) per 4 \textit{V}_{H} gene segments (\textasciitilde1/1200 bp), which corresponds to 1.1 \times 10^{-5} mutations/bp per cycle.\textsuperscript{31-33}

Antigen selection

Two methods were applied to assess antigen selection in productive and mutated \textit{V}_{H} gene immunoglobulin rearrangements through their patterns of somatic mutations. First, the ratio of replacement to silent mutations (R/S) in the CDR2 and FWR3 regions was calculated, and a sequence was considered to be antigen selected when the R/S ratio in the CDR2 region was higher than 2.9 and the R/S ratio in the FWR3 region was lower than 1.5. Second, the R/S ratio of the somatic mutations in the FWR3 region alone was considered, and a sequence was regarded as antigen selected when the R/S ratio was less than 1.6.\textsuperscript{34} We also tested other types of calculations, as described previously,\textsuperscript{17} but the calculation of reliable \textit{P} value failed in the cases with a very low number of somatic mutations. Therefore, the R/S calculation was considered most appropriate for the cases in this study.

Results

Morphology, EBV infection, C-MYC translocation, and immunophenotype

This study is based on 31 BLs in which a C-MYC translocation was detectable and in which DNA suitable for \textit{IgH} PCR amplification...
was extractable. Table 2 summarizes the distribution of BLs collected from endemic areas of Kenya according to patient ages, clinicopathologic characteristics, and EBV and HIV status. Interestingly, in this series, 3 out of 4 BLs that developed in young adults were associated with HIV. These cases, which do not conform to the tumor syndrome described by Dennis Burkitt, are now better designated as AIDS-related BL and have been grouped with the AIDS-related BLs from Europe. On the other hand, the case occurring in a 19-year-old man (HIV-negative, EBV-positive, and localized in the testis) was grouped with the other eBLs. As a result, our final case distribution was 15, 10, and 6 cases of eBL, sBL, and AIDS-related BL, respectively. The clinical and pathologic characteristics of these cases are reported in Table 3. EBV was present in 12 out of 15 cases of eBL, 1 out of 10 cases of sBL, and 5 out of 6 cases of AIDS-related BL.

Histologic review of these cases showed that the Burkitt tumors were composed of medium-sized cells with round nuclei, granular chromatin, multiple nucleoli, and abundant basophilic cytoplasm. A “starry-sky” pattern was usually present, imparted by numerous benign macrophages that had ingested apoptotic cells/bodies. Six cases showed more pleomorphism or larger cells than classical BL (3 eBL, 2 sBL, and 1 AIDS-related BL). These cases were classified as atypical BL, according to the WHO classification.

All cases were CD20+, CD79a+, CD10+, Bcl-6+, and Bcl-2+, consistent with the diagnosis of BL, and all had an extremely high rate of proliferation (Ki-67 [Mib1] almost 100%) and of cell death (apoptosis).

**Analysis of IgH rearrangement**

The extracted DNA was amplified using 4 different sets of primers—FR2A, FR2FS, FR2BM, and FR3BM—which resulted in amplification efficiencies ranging from 50% to 100%, the latter obtained with the FR2BM and FR3BM. Using these primers, we found amplification efficiencies ranging from 50% to 100%, the latter obtained with the FR2BM and FR3BM (Figure 1). The amplificates of the lymphomas studied to the FR2BM primer, and all had an extremely high rate of proliferation (Ki-67 [Mib1] almost 100%) and of cell death (apoptosis).

**Table 2. Distribution of BLs collected from endemic areas according to patient age, clinicopathologic characteristics, and EBV and HIV status**

<table>
<thead>
<tr>
<th>Age group</th>
<th>Patients aged 16 y or younger, no.</th>
<th>Patients older than 16 y, no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total cases</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>Male/female</td>
<td>6/8</td>
<td>3/1</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jaw</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Nodal</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Ileum</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>HIV</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>EBV</td>
<td>11</td>
<td>4</td>
</tr>
</tbody>
</table>

For patients aged 16 and younger, the range of ages was 3 to 16 years, and the median age was 7 years. For patients older than 16 years, the range of ages was 19 to 59 years, and the median age was 31 years.

**Table 3. Clinicopathologic characteristics of BL**

<table>
<thead>
<tr>
<th>Category</th>
<th>eBL</th>
<th>sBL</th>
<th>AIDS-related BL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age range, y (median)</td>
<td>3-19 (7)</td>
<td>3-40 (30)</td>
<td>27-69 (47)</td>
</tr>
<tr>
<td>Site, no. nodal, no. extranodal</td>
<td>4/11</td>
<td>2/8</td>
<td>4/2</td>
</tr>
<tr>
<td>EBV cases, no. positive, no. total</td>
<td>12/15</td>
<td>1/10</td>
<td>5/6</td>
</tr>
</tbody>
</table>

Cloning of the IgH PCR products revealed identical V\(\text{H}\) sequences in all 15 cases of eBL, 10 cases of sBL, and 6 cases of AIDS-related BL studied, indicating that the mutation process was not active in these cases. In contrast, all the sequences obtained from FLs showed signs of ongoing mutation. The assignment of the V\(\text{H}\) segments of the lymphomas studied to the V\(\text{H}\) families demonstrated that in sBL, eBL, and AIDS-related BL, the V\(\text{H}\) use was highly similar to that observed in peripheral blood from healthy persons.

The mutation patterns used to determine whether the V\(\text{H}\) mutations detected in the eBL, sBL, and AIDS-related BL cases arose by chance or by antigen selection are also reported in Table 4. In most of the cases of eBL and AIDS-related BL, the number of R mutations in the CDR2 region was higher than expected, and in the FW3 region it was lower than expected. The R/S ratio in the CDR2 region was more than 2.9, and in the FW3 regions it was less than 1.5 in 8 cases of eBL and in 3 cases of AIDS-related BL. These findings strongly suggest that eBL and AIDS-related BL derive, at least in part, from antigen-selected B cells, whereas no signs of antigen selection were found in sBL.

Because one of the main differences between eBL and AIDS-related on the one hand and sBL on the other hand is the association with EBV, we compared EBV-positive and EBV-negative BL, irrespective of geographic origin and HIV status. The differences were even clearer, in that EBV-positive BLs were highly mutated and EBV-negative BLs displayed only a small number of somatic mutations. The differences in antigen selection also held true when we compared EBV-positive and EBV-negative cases: signs of antigen selection were found in 11 of 18 EBV-positive cases, whereas no signs of antigen selection were found among EBV-negative cases (Table 5).

**Discussion**

Previous studies have reported that the neoplastic cells in BL have acquired somatic mutations and, therefore, that the cell of origin is likely to have traversed the germinal center. However, a wide variety in the degree of mutations, antigen selection, and ongoing mutation characteristics have been reported in and among BL types, leaving open the question of whether BLs originate from GC cells or post-GC memory B cells. In addition, the data provided in these previous publications are based almost exclusively on cell
line experiments, and only a few primary tumors have been studied. It is commonly accepted that the results of cell line investigations should be interpreted with caution and that they and require confirmation, whenever possible, through the investigation of a sufficient number of primary tumors.

Hence, a possible approach to this question can be found in the sequence analysis of rearranged immunoglobulin \( V_H \) region genes in a large number of primary tumors. Sequence analysis of the VDJ rearrangements from cells at various stages of B-cell differentiation has revealed that the increase in the affinity of the antibody-binding site is caused by hypermutations in the rearranged \( V_H \) region genes.36,37 This process leads to the introduction of point mutations in the variable, but not in the constant, regions of the immunoglobulin gene sequences.38 If cells expressing the mutated sequences are then exposed to limiting antigens, this interaction will result in the stimulation of only those cells that bind antigen most efficiently, thereby leading to the selection of B cells with \( V_H \) gene sequences encoding amino acids of best fit, with unstimulated cells undergoing apoptosis.39 Therefore, antigen-selected sequences tend to have a concentration of mutations leading to the replacement of amino acids in the complementarity-determining regions (CDRs).

The mutation process starts in the dark zone of GC because the blasts in this zone, designated centroblasts, contain \( V_H \) genes with no or only a few mutations; the mutation process in the GC is ongoing and proceeds when the centroblasts migrate to the light zone and transform into centrocytes because this is associated with an increase in \( V_H \) mutations.32 However, there is evidence that the mutation process in the GC is discontinuous rather than continuous, with mutation-active phases followed by mutation-silent phases during which GC B-cell antigen selection takes place. When cells leave the germinal centers and differentiate into memory B cells, the mutation process is completely switched off.40

Our molecular analysis of rearranged \( V_H \) genes showed that all cases of BL constantly harbor mutated \( V_H \) genes, but no evidence of ongoing mutation was observed by cloning the PCR products. In addition, differences were found in the number of somatic mutations and antigen selections between eBL, sBL, and AIDS-related BL; sBL had a low number of somatic mutations and no signs of antigen selection, whereas eBL and AIDS-related BL showed higher numbers of somatic mutations and signs of antigen selection. Given that one of the main differences among the various forms of BL is the association with EBV, we compared EBV-positive and EBV-negative BL, irrespective of geographic origin and HIV status. Although most EBV-positive cases belonged to the endemic and AIDS-related categories, the difference in the number of somatic mutations and antigen selection became even more evident when EBV-positive and EBV-negative BLs were compared. From our molecular results, it seems that EBV-positive and EBV-negative BLs thus differ in their cells of origin and probably in their pathogenetic mechanisms. Differences in cell kinetics between EBV-positive and EBV-negative BLs and in the expression of cellular genes in EBV-converted BL cell lines have already been acknowledged.41,42 However, there is still no satisfactory explanation of how EBV participates in the pathogenesis of BL.43,44

According to the mutation pattern we found, there are 2 B-cell differentiation stages to which BL could correspond: (1) for sBL (EBV-negative BL), early centroblasts; (2) for eBL and AIDS-related BL (EBV-positive BL), a memory B cell or a late GC B cell in which the mutation process is switched off and the differentiation toward a memory B cell is initiated (Figure 2).

The low number of somatic mutations, the lack of ongoing mutations, and the signs of antigen selection seem to suggest that EBV-negative BL originates from early centroblasts that have gone through a first round of mutation and reached the subsequent mutation-silent phase before antigen selection takes place. Their further differentiation is blocked by the malignant process, which is probably induced or maintained by the \( C-MYC \) deregulation present in all cases of BL.45 It appears that several types of oncogene translocation (such as \( C-MYC \) translocation in BL) may occur as a consequence of genomic instability because of the somatic hypermutation process within the GC.46 Interestingly, recently Zhu et al47 have shown that deregulated expression of \( C-MYC \) drives the development of mouse “Burkitt-like” EBV-negative lymphomas with a very low (insignificant) level of somatic mutation. This observation is similar to our results in human EBV-negative sBL. However, we think the normal cellular counterpart of sBL tumor cells is more likely to be early centroblasts rather than native pre-GC B cells, as hypothesized by Zhu et al.47

![Figure 2](https://example.com/figure2.png)
The higher number of somatic mutations and the signs of antigen selection in the absence of ongoing mutation observed in EBV-positive BL indicate that their cell of origin is more closely related to post-GC memory B cells than to early dark zone B-cell blasts of the GC. A new model of persistent EBV infection has recently been proposed, based on the observation that EBV is found restricted to resting memory B cells in the peripheral blood. The essence of this model is that EBV uses different transcription programs to regulate the behavior of latently infected B cells as they move through the differentiation process to ultimately end up as latently infected resting memory B cells. This is in line with our molecular results, suggesting that EBV-positive BL may originate from a late GC B cell in which the mutation process is switched off and the differentiation toward memory B cells is initiated, or from a memory B cell itself. This may seem to be in contrast with the GC phenotype observed in Burkitt tumors.

Other lymphoma entities besides BL (ie, diffuse large B-cell lymphomas) harbor a GC cell immunophenotype but have molecular and morphologic features not compatible with a GC origin. This means that a distinction between lymphomas that derive from (late) GC cells or from post-germinal (memory) B cells is often not possible with any degree of certainty. In addition, maintenance of the differentiation toward memory B cells is often not demonstrated for the pathogenesis of the various BL subentities: EBV-negative BL (mainly sBL) appears to derive from early centroblasts, and EBV-positive BL (mainly eBL and AIDS-related BL) may originate from late GC B cells or memory B cells. EBV-positive and EBV-negative BLs may thus result from different cellular counterparts with significant implications for the pathogenesis of the various BL subentities: EBV-negative BL (mainly sBL) appears to derive from centroblasts, and EBV-positive BL (mainly eBL and AIDS-related BL) may originate from late GC B cells or memory B cells. EBV-positive and EBV-negative BLs may thus result from different biologic entities. Whether this could have therapeutic implications should also be addressed.

References


Immunoglobulin gene analysis reveals 2 distinct cells of origin for EBV-positive and EBV-negative Burkitt lymphomas

Cristiana Bellan, Stefano Lazzi, Michael Hummel, Nazzareno Palummo, Margherita de Santi, Teresa Amato, Joshua Nyagol, Elena Sabattini, Thierry Lazure, Stefano A. Pileri, Martine Raphael, Harald Stein, Piero Tosi and Lorenzo Leoncini