The American Journal of Pathology, Vol. 181, No. 5, November 2012

http://dx.doi.org/10.1016/j.ajpath.2012.07.028

Tumorigenesis and Neoplastic Progression

Molecular Characterization of Immunoglobulin Gene Rearrangements in Diffuse Large B-Cell Lymphoma

Antigen-Driven Origin and IGHV4-34 *as a Particular Subgroup of the Non-GCB Subtype*

Elena Sebastián,*† Miguel Alcoceba,*† Ana Balanzategui,*† Luis Marín,*† Santiago Montes-Moreno,‡ Teresa Flores,§ David González,¶ M. Eugenia Sarasquete,*† M. Carmen Chillón,*† Noemí Puig,*† Rocío Corral,*† Emilia Pardal,‖ Alejandro Martín,*†‖ Eva González-Barca,‖ M. Dolores Caballero,*†‖ Jesús F. San Miguel,*†** Ramón García-Sanz,*†** Marcos González*†**

From the Departments of Hematology and Pathology,§ University Hospital of Salamanca, Salamanca, Spain; the Institute of Biomedical Research of Salamanca (IBSAL),† Salamanca, Spain; the Servicio de Anatomía Patológica,‡ University Hospital of Marqués de Valdecilla/IFIMAV, Santander, Spain; the Institute of Cancer Research,¶ Royal Cancer Hospital, London, United Kingdom; the Spanish Lymphoma/Autologous Bone Marrow Transplant Study Group,‖ University Hospital Marqués de Valdecilla, Santander, Spain; and the Cancer Research Center-IBMCC (USAL-CSIC),** Salamanca, Spain*

The pathogenesis of diffuse large B-cell lymphoma (DLBCL) remains partially unknown. The analysis of the B-cell receptor of the malignant cells could contribute to a better understanding of the DLBCL biology. We studied the molecular features of the immunoglobulin heavy chain (IGH) rearrangements in 165 patients diagnosed with DLBCL not otherwise specified. Clonal IGH rearrangements were amplified according to the BIOMED-2 protocol and PCR products were sequenced directly. We also analyzed the criteria for stereotyped patterns in all complete IGHV-IGHD-IGHJ (V-D-J) sequences. Complete V-D-J rearrangements were identified in 130 of 165 patients. Most cases (89%) were highly mutated, but 12 sequences were truly unmutated or minimally mutated. Three genes, *IGHV4-34*, *IGHV3-23*, and *IGHV4-39*, accounted for one third of the whole cohort, including an overrepresentation of *IGHV4-34* (15.5% overall). Interestingly, all *IGHV4-34* rearrangements and all unmutated sequences belonged to the nongerminal center B-cell–like (non-GCB) subtype. Overall, we found three cases following the current criteria for stereotyped heavy chain VH CDR3 sequences, two of them belonging to subsets previously described in CLL. IGHV gene repertoire is remarkably biased, implying an antigen-driven origin in DLBCL. The particular features in the sequence of the immunoglobulins suggest the existence of particular subgroups within the non-GCB subtype. ***(Am J Pathol 2012, 181:1879–1888; http://dx.doi.org/10.1016/j.ajpath.2012.07.028)***

Diffuse large B-cell lymphoma (DLBCL) is a type of aggressive lymphoma that accounts for about 40% of B-cell non-Hodgkin lymphomas in adults,[1] most of them classified as DLBCL not otherwise specified (NOS) according to World Health Organization classification (2008).[1] Genome-wide expression profile technologies have allowed DLBCL NOS to be separated into two biological subtypes: germinal center B-cell–like (GCB) DLBCL and activated B-cell–like (ABC) DLBCL,[2] based on their presumed cell of origin. Immunostaining approaches allow these entities to be classified into GCB or non-GCB subtypes[3,4] and have been proposed as feasible surrogates of genome-wide expression profiles.

Supported in part by grants from the Health Council of Castilla y León (GRS265/A/08) and the Health Research Program (PS0901382 and CM10/00078-Río Hortega, ISCIII).

Accepted for publication July 30, 2012.

Supplemental material for this article can be found at http://ajp.amjpathol.org or at http://dx.doi.org/10.1016/j.ajpath.2012.07.028.

Address reprint requests to Marcos González, M.D., Ph.D., Servicio de Hematología, Hospital Universitario de Salamanca, Paseo de San Vicente, 58-182, 37007, Salamanca, Spain. E-mail: margondi@usal.es.

DLBCL is constituted by large B cells harboring a clonal rearrangement of immunoglobulin (IG) genes. The rearrangement of the immunoglobulin heavy chain (IGH) gene occurs during B-cell ontogeny in an ordered fashion, leading to the assembly of distinct variable (IGHV), diversity (IGHD), and joining (IGHJ) genes.[5] During B-cell differentiation, rearrangements start with an IGHD gene joining to an IGHJ gene, forming a partial IGH-D-IGHJ (D-J) rearrangement, a process that can simultaneously occur in both alleles. This is followed by the joining of an IGHV gene to the partial D-J rearrangement, forming a completely functional IGHV-IGHD-IGHJ (V-D-J) rearrangement in one allele, but can also retain a partial nonfunctional rearrangement in the other allele. Once a rearrangement has become completely functional, the naive B cell can react to antigens in the germinal center (GC) of the secondary lymphoid organs, undergoing affinity maturation by the processes of somatic hypermutation (SHM) and class switch recombination.

The molecular features of IGH rearrangements may provide important information about the ontogeny of B cells. Thus, the intrinsic mutability of IGHV genes has been widely analyzed.[6–8] Since normal B cells undergo SHM as a response to antigen, classically within GC, the pattern of SHM is thought to provide clues about the mutation machinery and the role of antigen selection. Nonrandom, potentially antigen-driven SHM has been described in normal B cells, as well as in chronic lymphocytic leukemia (CLL) and classic hairy cell leukemia (HCL), which has been called canonical SHM.[9–13] However, there are reports suggesting that certain B cells may accumulate IGHV gene mutations in an alternative T-cell–independent fashion.[14–16] The similarities or differences of mutations acquired through an alternative pathway are not yet defined. Several models[17–19] have been used for quantifying antigen selection, but results may not be reliably interpretable.[8] Since precise tools are still evolving,[19] canonical mutation criteria remain as an available method for establishing whether or not a set of sequences are affected by similar processes and selections.

Moreover, IGHV gene usage in IGH rearrangements has been analyzed in normal B cells,[20,21] in autoimmune diseases,[22] and in many B-cell lymphoproliferative disorders (B-LPDs).[13,23–30] Certain IGHV genes have been closely related to some pathogens[31] or linked to autoimmunity.[32] Particular antigens have been identified as possibly being responsible for lymphomagenesis or defined as a secondary event.[30] The involvement of particular antigens in lymphomagenesis may be supported by the identification of stereotyped groups with closely similar complementarity-determining region 3 (VH CDR3).[33] The stereotyped B-cell receptors (BCRs) have been studied extensively in CLL[25,34,35] with a clinical correlation[36] and, more recently, in other lymphomas.[27–30]

Although DLBCL is the most prevalent aggressive lymphoma, few studies have characterized the BCR in this entity, and most are based on small series of patients or in selected populations.[23,23,37–43] In addition, partial D-J rearrangements, the pattern and distribution of mutations, and the existence of stereotyped sequences in DLBCL are still not well defined.[41]

Here, we report a detailed immunogenetic analysis of the complete and partial IGH rearrangements in 165 DLBCL, the largest series to date, to improve our understanding of the biology of this disease.

Materials and Methods

Patient Characteristics and Immunohistochemistry

A total of 165 patients with DLBCL were selected retrospectively on the basis of a diagnosis of *de novo* DLBCL NOS according to the 2008 World Health Organization classification.[1] The median age was 59 years, with a range from 15 to 86 years. Fifty-one percent of the patients were male, and 48% of the total series had an International Prognostic Index score ≥3.[44] To distinguish DLBCL biological subtypes, Hans' algorithm[3] was used as a feasible surrogate of genome-wide expression profiles, since it has demonstrated a reproducibility beyond 80%.[3,4] Accordingly, immunostaining with antibodies to BCL-6, CD10, and melanoma-associated antigen (mutated) 1 (MUM1/IRF4) was performed in 121 patients, which allowed them to be separated into GCB DLBCL (38%) and non-GCB (62%). A total of 159 patients received treatment with curative intent, 125 of them (81%) with rituximab-based regimens (Table 1).

The study was approved by the local ethics review committee in accordance with Spanish legislation, and informed consent was obtained from all participants.

Table 1. Clinical-Biological Characteristics of DLBCL Patients at Diagnosis (n = 165)

Variable	% of Patients
Median follow-up, months (range)	27.8 (0–198)
Age [years, median (range)]	58.7 (15–86)
≤60	54
>60	46
Sex	
Male	51
Female	49
Performance status	
ECOG <2	70
ECOG ≥2	30
β2-microglobulin >3 mg/L	36
Albumin <3.5 g/dL	47
LDH elevated	59
IPI	
IPI 0–1	24
2	28
3	24
IPI 4–5	24
Immunostaining	
GCB	38
Non-GCB	62
Treatment	
Rituximab-based	81
Others	19

ECOG, Eastern Cooperative Oncology Group; GCB, germinal center B-cell–like; IPI, International Prognostic Index; LDH, lactate dehydrogenase.

DNA Extraction

Tumor DNA was extracted from samples collected at the time of diagnosis. These involved lymph nodes (84%), bone marrow (10%), and other tumor tissues (6%). In fresh samples, high-molecular-weight DNA was isolated using the DNAzol reagent (MRC, Cincinnati, Ohio).[45] DNA was extracted from formalin-fixed paraffin-embedded (FFPE) tissue in half of the lymph nodes using RecoverAll Total Nucleic Acid Isolation Kit (Ambion/Applied Biosystems, Foster City, California) or QuickExtract FFPE DNA Extraction Kit (Epicenter Biotechnologies, Madison, Wisconsin), with no differences in terms of recovery capacity.

Identification of Clonal Rearrangements

All samples were tested for the amplification of IGH rearrangements according to the BIOMED-2 Concerted Action protocols, where we actively participated in the standardization.[46] In this strategy, complete V-D-J rearrangement amplification was performed by multiplex PCR with a set of family-specific primers of the framework region 1 (FR1) and one IGHJ consensus primer. In cases using the *IGHV4-34* gene, a second PCR was performed with primers from the IGHV leader region. For the samples with no detectable amplification from FR1, PCR was performed from FR2. Additionally, partial D-J rearrangements were amplified in two different reactions using family-specific primers for the IGHD1-IGHD6 and IGHD7 families, respectively, together with the consensus IGHJ primer. The presence of the monoclonal rearrangement was then confirmed by GeneScan with an ABI 3130 DNA Sequencer (Applied Biosystems).

Sequencing and Identification of IGH Gene Rearrangements

PCR products were sequenced directly, including both forward and reverse reads, using Big-Dye terminators (Applied Biosystems).[46]

Germline IGHV, IGHD, and IGHJ genes from complete V-D-J rearrangements were identified using the IMGT/V-QUEST database (*http://www.imgt.org*, last accessed March 21, 2012). The following information was extracted: IGHV, IGHD, and IGHJ gene usage, the percentage of IGHV identity to the closest germline gene, and the VH CDR3 length and composition. The VH CDR3 length was determined by counting the number of amino acids from the cysteine at position 104 and the tryptophan at position 118, according to the IMGT unique numbering system.[47] Partial D-J rearrangements were analyzed using BLAST database.

Gene usage in DLBCL was compared with that previously described in normal B lymphocytes and in other B-LPDs for complete[13,20,24–28] and partial[26,48–50] rearrangements.

Analysis of Somatic Hypermutation

Percentage Identity with the Closest IGHV Germline

The use of 98% as the percentage identity cutoff value has customarily been used for the clinical consideration of a patient's sequence as mutated or unmutated based on the exclusion of potential genomic polymorphisms. This cutoff has a prognostic value in CLL.[51] However, it has been shown that most sequence differences in the 98% to 99.99% group correspond to low-level SHM,[52] as applied in recent studies.[27,53]

We, therefore, classified the sequences as previously described[27]: "truly unmutated" when no deviation from the germline was observed; "minimal/borderline mutated" when there was 97.0% to 99.9% similarity with the germline; and finally, "highly mutated" when the deviation from the germinal sequence was >97%.

SHM Characteristics

Nucleotide characteristics of IGHV mutations in DLBCL were studied to establish whether the pattern of distribution was consistent with the canonical SHM process.[12] This analysis was also performed separately in both the GCB and non-GCB subtypes, as well as in DLBCL expressing particular IGHV genes.

The defined characteristics of canonical SHM include: i) a higher ratio of replacement (R) versus silent (S) mutations in the CDRs than in the FRs; ii) base change bias for transitions; and iii) more than the expected percentage of mutations in the RGYW hotspots (R = A/G, Y = C/T, and W = A/T) relative to the total number of nucleotides located within these hotspots. In addition, to determine whether these characteristics were due to selection or to mutation machinery, the mutations at redundant wobble bases were analyzed as previously reported for CLL and HCL.[12,13] Wobble bases are the third nucleotide of redundant codons. Finally, mutations at A:T and G:C pairs were studied to determine whether the mutations were targeted with approximately the same frequency as in normal cells.[54] Moreover, recurrent nucleotide and amino acid changes for the most common IGHV genes were analyzed to find evidence of driven SHM in DLBCL.

VH CDR3 Analysis and Identification of Stereotyped Rearrangements

The criteria to define a stereotyped sequence are not yet well established.[35] To perform this analysis, we used the previously described pattern-based method,[55] updated with the recently published, more stringent criteria,[35] to identify possible restrictions in the VH CDR3 amino acid composition of rearrangements. Briefly, the criteria for stereotyped patterns included: i) at least 50% amino acid identity; ii) at least 70% similarity between sequences; iii) IGHV usage of the same clan; iv) identical VH CDR3 length; and v) exact location of the sequence pattern within the VH CDR3 region. Analyses were performed using the ClustalW/ClustalX 2.0 multiple sequence alignment software.

Subsets were named as previously published,[25] those composed of only two sequences being considered "provisional." A letter of the alphabet was assigned to the new subset. These were compared with previously published stereotyped patterns[25,35] and with DLBCL sequences from public databases (EMBL, *http://www.ebi.ac.uk/embl*, last accessed March 28, 2012; NCBI, *http://www.ncbi.nlm.nih.gov*, last accessed March 28, 2012).

Statistical Analysis

The groups of patients were compared by the χ^2 or Fisher exact tests using Graph-Pad 4.0 (GraphPad Software, San Diego, California) and SPSS 15.0 (SPSS, Chicago, Illinois).

Differences were considered to be statistically significant for values of $P < 0.05$.

Results

Detection of Clonal Rearrangements

The overall detection rate of monoclonality by amplifying V-D-J and/or D-J rearrangements was 90% (149 of 165). V-D-J rearrangements were detected in 130 of 165 patients (79%), and D-J rearrangements were seen in 68 of 165 patients (41%).

Complete IGHV-IGHD-IGHJ Rearrangements

A total of 109 of 130 complete V-D-J rearrangements were successfully sequenced. Reasons for not having a sequence in the remaining 21 cases were as follows: high polyclonal background ($n = 2$); double sequences, with no possible individual reading ($n = 1$); unspecific amplification, which means that a sequence could be read but the result did not fit with any possible rearrangement ($n = 2$); low fluorescence (intensity) of the peak rendering an amplification band too weak for accurate sequencing performance ($n = 11$); and partial sequence, enough to be identified as a V-D-J rearrangement but insufficient to be completely identified ($n = 5$).

A total of 103 were productive rearrangements, whereas 6 were nonproductive. Most sequences (82%) were obtained from FR1. Within the productive setting, IGHV gene repertoire analysis revealed that IGHV3 and IGHV4 were the predominant subgroups (Table 2). A total of 27 functional IGHV genes were identified, of which *IGHV4-34* (15.5%), *IGHV3-23* (9.7%), and *IGHV4-39* (8.7%) were the most frequent. Regarding IGHD genes, a total of 25 were identified, with a predominance of *IGHD3* (Table 2), especially *IGHD3-22* and *IGHD3-10*. With respect to IGHJ gene usage, most cases used the *IGHJ4* and *IGHJ6* genes.

Partial IGHD-IGHJ Rearrangements

Sequence analysis was possible in 64 partial D-J rearrangements. The IGHD2 subgroup genes were most frequent (Table 2). In comparison with complete rearrangements, IGHD4 was significantly overrepresented in partial rearrangements (18.8% versus 5.8%, $P < 0.05$), whereas IGHD3 was underrepresented (14.1% versus 35.9%, $P < 0.01$). Overall, 20 IGHD genes were identified, of which *IGHD2-2* was the most frequent. *IGHD3-22* and *IGHD3-3*, two of the most frequent IGHD genes in V-D-J rearrangements, were completely absent in partial rearrangements, suggesting that the use of certain IGHD genes may favor the rearrangement process to obtain a final complete functional rearrangement. Finally, the distribution pattern of the IGHJ genes was similar to the complete rearrangements (Table 2).

Table 2. Distribution of Families in Productive V-D-J and Partial D-J Rearrangements

Variable	V-D-J ($n = 103$) n (%)	D-J ($n = 64$) n (%)
IGHV		
1	17 (16.5)	
2	2 (1.9)	
3	45 (43.7)	
4	34 (33.0)	
5	4 (3.9)	
6	1 (1.0)	
7	0	
IGHD		
1	11 (10.7)	6 (9.4)
2	22 (21.4)	19 (29.7)
3	37 (35.9)	**9 (14.1)**†
4	6 (5.8)	**12 (18.8)***
5	10 (9.7)	5 (7.8)
6	15 (14.6)	12 (18.8)
7	2 (1.9)	1 (1.6)
IGHJ		
1	4 (3.9)	3 (4.9)
2	5 (4.8)	5 (8.2)
3	13 (12.6)	8 (13.1)
4	46 (44.7)	32 (52.5)
5	13 (12.6)	3 (4.9)
6	22 (21.4)	10 (16.4)

Significant differences between productive V-D-J and incomplete D-J rearrangements in DLBCL are depicted in bold.
*$P < 0.05$.
†$P < 0.01$.

Somatic Hypermutation

Percentage Identity to the Closest IGHV Germline

Somatic hypermutation (<98% of identity to the closest germline IGHV gene) was present in 94 of 103 (91%) patients who had productive complete V-D-J rearrangements (median: 10.8%; range: 0% to 26%). As expected, only 9% of the sequences were unmutated, similar to previously reported results.[23,39,42,56] Overall, point mutations predominated, whereas only five cases had deletions ($n = 3$) or insertions ($n = 2$) in their sequences.

According to the criteria for SHM described previously,[27] 4, 8, and 91 cases were "truly unmutated," "borderline mutated," and "highly mutated," respectively. All truly unmutated and borderline mutated sequences were obtained from FR1.

Interestingly, all cases with the *IGHV4-34* gene but one (15 of 16) were "highly mutated." However, *IGHV1-69*,

Table 3. Characteristics of SHM in DLBCL (Global Series) and with Respect to Immunohistochemistry and the Most Common IGHV Families and Genes

	Number of sequences	Number of mutations	R/S ratio			Transition/transversion ratio			% nucleotides in RGYW	% of mutations in RGYW
			Total	CDR	FR	Total	CDR	FR		
DLBCL	96	2485	1.69	2.56	1.42	1.161	1.13	1.18	14.0	24.9
GCB	31	931	1.71	2.77	1.42	1.155	0.97	1.25	13.3	22.9
Non-GCB	46	1037	1.56	2.05	1.39	1.216	1.30	1.18	14.2	25.2
IGHV1	15	631	1.42	**1.45**	**1.42**	1.246	1.29	1.24	17.9	26.7
IGHV3	42	1127	1.81	2.71	1.49	1.159	1.16	1.16	12.0	19.3
IGHV4	32	662	1.70	2.98	1.27	1.115	1.07	1.14	15.9	31.6
IGHV1-69	4	46	0.92	**0.88**	**0.94**	1.300	1.50	1.21	19.8	32.6
IGHV3-23	10	243	1.60	2.24	1.31	1.505	1.47	1.53	13.1	26.3
IGHV4-34	15	304	1.34	1.91	1.22	1.338	1.39	1.32	14.8	30.6

Figures in bold indicate different R/S ratios from that expected.
CDR, complementarity-determining regions; FR, framework region; R, replacement mutation; S, silent mutation; RGYW, hotspots.

IGHV5-51, and *IGHV3-21* were the most frequent genes used in "truly unmutated" and "borderline mutated" cases.

As expected from the physiopathological point of view, mutations were observed in a minority of cases (12%) of partial D-J rearrangements with a low mutation frequency (median: 4.3%, range: 3.5% to 11.1%).

Nucleotide Characteristics of Somatic Hypermutation

S mutations occurred randomly, with a relatively even distribution, whereas R mutations were mainly localized in the CDR1 and CDR2 regions, at the 5′ and 3′ ends of FR2 and the 3′ end of FR3 (see Supplemental Figure S1 at *http://ajp.amjpathol.org*).

Analysis of canonical SHM was possible in 96 DLBCL sequences with <100% identity with the germline. For the 2485 mutations analyzed, the R/S mutation ratio was 1.72, being greater in the CDR (2.56) than in the FR (1.42) regions ($P < 0.01$) (Table 3). The transition-to-transversion ratio was 1.16, significantly higher than the expected ratio of 0.5 if the frequencies of the types of mutation were random ($P < 0.01$), and there was no significant difference between the CDR and FR regions. Finally, 24.9% of the DLBCL mutations were located in RGYW hotspots, a value significantly greater than the expected percentage (14.0%, $P < 0.01$). Thus, SHM in DLBCL displayed all of the established characteristics of canonical SHM. These canonical characteristics persisted in the most common IGHV genes, except for *IGHV1* (mainly *IGHV1-69*), which met these three criteria to a lesser extent (Table 3).

The bias for transitions over transversions was also reflected by the wobble base mutations (1.20, $P < 0.01$) (Figure 1), implying that this bias in DLBCL is due to the same mechanism of SHM.

Finally, G:C pairs were preferentially targeted for SHM relative to A:T (ratio: 1.86) (Figure 1). This observation was especially evident in the *IGHV4-34* gene (ratio: 2.23).

Recurrent Amino Acid Changes

Recurrent amino acid changes were mainly found among *IGHV4-34*, *IGHV4-39*, and *IGHV1-18* rearrangements. For example, a substitution at the FR3 codon 92 occurred in 13 of 16 (81%) cases of *IGHV4-34*; the substitution of S to T (6 of 16, 38%) and S to N (4 of 16, 25%) were the most frequent changes observed. This recurrent amino acid change contrasts with the adjacent codon, 93, which did not undergo any substitution in any of the 16 sequences. Although the *IGHV4-34* genes were highly mutated, all 10 available *IGHV4-34* sequences conserved the specific AVY motif at the end of FR1 codons 24 to 26, and the W at position 7 was intact in the 5

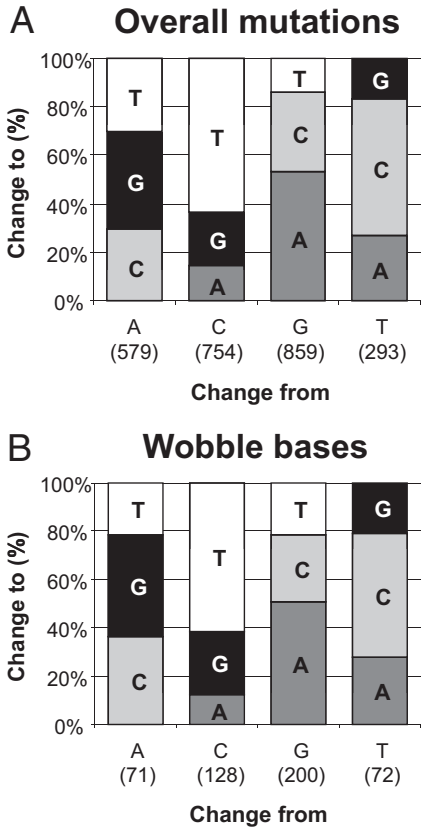

Figure 1. Nucleotides mutated in DLBCL. Relative frequency with which each nucleotide mutates to form each of the other three. The number of mutations occurring in each nucleotide is shown in parentheses. **A:** Total mutations in DLBCL. **B:** Mutations in redundant wobble bases.

Table 4. Stereotyped VH CDR3 Amino Acid Sequences

Subset	IGHV	IGHD	IGHJ	Mutational status	VH CDR3 amino acid sequences			
Subset 5	Published							
11–3436	**IGHV1-69∗13**	**IGHD3-10∗01**	**IGHJ6∗03**	**100**	**AR**	**TMVRGVI**	**NID**	**YYYYYMDV**
SE-01-0444-H1	IGHV1-69∗01	IGHD3-10∗01	IGHJ6∗03	99.6	AR	GMVRGVI	NID	YYYYYMDV
UK-01-0345-H1	IGHV1-69∗01	IGHD3-10∗01	IGHJ6∗03	100	AR	TMVRGVI	RVM	YYYYYMDV
FR-01-0286-H1	IGHV1-69∗01	IGHD3-10∗01	IGHJ6∗03	100	AR	VMVRGVI	SLD	YYYYYMDV
GR-02-0510-H1	IGHV1-69∗01	IGHD3-10∗01	IGHJ6∗03	100	AR	SMVQGVI	NAY	YYYYYMDV
NL-01-0294-H1	IGHV1-69∗01	IGHD3-10∗01	IGHJ6∗03	100	AR	TMVQGVI	RVY	YYYYYMDV
Subset 7D	Published							
11–1949	**IGHV1-69∗01**	**IGHD3-3∗01**	**IGHJ6∗02**	**100**	**AS PVGEG**	**DDFWSGYY**	**PN**	**YYDYGMDV**
CZ-01-0372-H1	IGHV1-69∗01	IGHD3-3∗01	IGHJ6∗02	100	AS SLGEN	YDFWSGYY	PN	YYYYGMDV
FR-01-0143-H1	IGHV1-69∗01	IGHD3-3∗01	IGHJ6∗02	100	AR PVQAK	YDFWSGYY	PN	YYYYGMDV
SE-01-1016-H1	IGHV1-69∗01	IGHD3-3∗01	IGHJ6∗02	99.6	AR PGGLY	YDFWSGYY	PN	YYYYGMDV
CZ-01-0484-H1	IGHV1-69∗01	IGHD3-3∗01	IGHJ6∗02	100	AR QGTLG	YDFWSGYY	PN	YYYYGMDV
UK-01-0427-H1	IGHV1-69∗13	IGHD3-3∗01	IGHJ6∗02	100	AR SGRED	YDFWSGYY	PN	YYYYGMDV
Subset A	Novel							
11–1967	**IGHV4-34∗01**	**IGHD3-22∗01**	**IGHJ2∗01**	**93.8**	**AR AL**	**DYYDNSG**	**RLP**	**GYFDL**
HQ392881.1	IGHV4-34∗01	IGHD3-22∗01	IGHJ2∗01	95.7	AR AL	DYYDSSG	RVLA	YFDL

Stereotyped rearrangements are compared with representative sequences from each subset. A complete list of sequences from subsets 5 and 7D is provided in Supplemental Table S2 at *http://ajp.amjpathol.org*. Sequences derived from the present study are remarked in bold.

IGHV4-34 cases that could be analyzed from the leader region. By contrast, the N-glycosylation site in CDR2 of *IGHV4-34* was lost in 7 of 16 (44%), and only one of the sequences that lost its natural site generated two novel sites in FR3. A summary list of recurrent amino acid substitutions for the predominant IGHV genes is given in Supplemental Table S1 (available at *http://ajp.amjpathol.org*). Some of these changes may indicate specific driven SHM in DLBCL.

VH CDR3 Composition and Stereotyped VH CDR3 Sequences

The median VH CDR3 length was 15 amino acids (range: 7 to 26). VH CDR3 was longer in the IGHV4 family than in the other families (16.8 versus 14.8, $P = 0.01$), and in "truly unmutated" and borderline cases as compared with "highly mutated" cases (18.2 versus 15.1, $P = 0.01$). This was particularly pronounced among the "truly unmutated" cases (21.5 versus 15.1, $P < 0.01$).

Three sequences among the 103 patients in our series met the criteria for stereotyped rearrangements and could be included into different subsets. Two sequences clustered within previously described CLL subsets (subsets 5 and 7D);[25,35] both contained the *IGHV1-69* gene and were "truly unmutated." The other case formed a probably realistic novel provisional subset (Table 4): subset A, sharing the V-D-J IGHV4-34-IGHD3-22-IGHJ2 conformation and with 79% identity of the amino acid sequence with a public DLBCL sequence.[42] A complete list of the sequences published in the literature from subsets 5 and 7D is shown in Supplemental Table S2 (available at *http://ajp.amjpathol.org*).

Molecular and Immunohistochemical Correlations

IGHV usage differed between the immunohistochemical subtypes. *IGHV1* (10 of 14) and *IGHV4* (15 of 21) were more frequent in the non-GCB subtype, especially for *IGHV1-69* (3 of 4) and *IGHV4-34* (12 of 12). A higher percentage of SHM was observed in GCB (13.8%, range: 2.9% to 26.0%) than in non-GCB (10.1%, range: 0.4% to 23.6%, $P < 0.01$) cases. Moreover, 6 of 7 borderline or unmutated cases with available immunostaining were non-GCB, which could explain this difference. By contrast, all *IGHV4-34* cases, despite being highly mutated, belonged to the non-GCB subtype. With respect to other molecular features, VH CDR3 length and the canonical characteristics for SHM showed no differences between the GCB and non-GCB subtypes (Table 3).

Comparative Analysis of IGH Rearrangements in DLBCLs with Other Normal B Cells and B-LPDs

To highlight the singular characteristics of DLBCLs, we compared our results with those reported for normal B cells and other B-LPDs. As shown in Table 5, the *IGHV4-34* gene was significantly overrepresented in DLBCL relative to $CD5^-/IgM^+$ normal B cells (3% to 9%, $P < 0.01$) and to other B-LPDs, such as CLL (8.7%, $P < 0.05$), HCL (6.9%, $P = 0.07$), and especially with multiple myeloma and Waldeström macroglobulinemia (2.7% and 0%, respectively, $P < 0.01$). However, it was found in a similar proportion of patients with splenic marginal zone lymphoma and mantle cell lymphoma.[13,16,20,21,24–26,57] By contrast, some genes frequently used in the rearrangements of normal B cells and other B-LPDs,[13,24–28] such as *IGHV3-30*, were completely absent from the present series ($P < 0.01$). Most complete V-D-J rearrangements were mutated, with a similar percentage of SHM (11.8%) to those previously reported in multiple myeloma (9%), but higher values than in CLL and Waldeström macroglobulinemia (2.4% to 6.6%).[24–26] Regarding SHM, G:C pairs were more often targeted for SHM than A:T pairs (ratio: 1.86) (Figure 1). This differs significantly from the distribution observed in CLL (ratio:

Table 5. Comparison of the IGHV, IGHD, and IGHJ Families and Significantly Different IGHV Genes in Our Series of DLBCL with Those Reported for B-LPDs and Normal B Cells in the Complete V-D-J Setting

Variable	DLBCL (n = 103)	MCL (n = 807)	B-CLL (n = 7596)	SMZL (n = 133)	HCL (n = 102)	MM (n = 270)	WM (n = 58)	$CD5^-/IgM^+$ (n = 206)
IGHV								
1	16.5	15.5	23.8	30.1	14.7	15.6	6.9	13.1
2	1.9	1.5	3.4	0.8	2.9	6.3	0	1.9
3	43.7	51.6	48.2	49.6	53.9	48.9	75.9	53.9
4	33.0	25.8	20.6	17.3	23.5	20.4	13.8	24.8
5	3.9	5.1	2.5	2.3	1.0	7.8	1.7	2.9
6	1.0	0.6	1.2	0	2.9	1.1	1.7	2.4
7	0	0	0.4	0	1.0	0	0	1.0
4–34	15.5	14.6	**8.9***	7.5	7.4	**0.9**†	**0**†	**3.9**†
3–30	0	3.5	**5.5***	**6.0***	**8.5**†	**10.0**†	**8.6**†	**5.8**†
IGHD								
1	10.7	10.6	8.2	5.2	12.7	7.6	27.3	6.4
2	21.4	17.4	19.6	15.5	12.7	24.2	4.5	21.2
3	35.9	34.2	40.3	46.6	41.2	25.8	18.2	37.2
4	5.8	8.5	6.3	7.8	6.9	10.6	13.6	10.9
5	9.7	8.3	8.8	9.5	6.9	13.6	9.1	10.3
6	14.6	20.3	15.9	15.5	18.6	12.1	27.3	10.3
7	1.9	0.6	0.7	0	1.0	6.1	0	3.8
IGHJ								
1	3.9	0.6	1.8	1.5	2.9	1.4	4.4	1.0
2	4.8	3.5	2.3	0	3.9	1.4	8.9	2.0
3	12.6	7.6	9.9	14.3	9.8	17.8	13.3	8.0
4	44.7	43.7	43.3	37.6	46.1	53.4	35.6	55.0
5	12.6	15.0	10.5	19.5	12.7	8.2	13.3	10.0
6	21.4	29.6	32.3	27.1	24.5	17.8	24.4	24.0

Frequencies expressed in percentages. References for comparisons are as follows: MCL,[27] CLL,[55] SMZL,[29] HCL,[13] MM,[24] WM,[26] and normal B cells.[20] Significant differences between DLBCL and other B-LPDs are depicted in bold.

*$P < 0.05$.

†$P < 0.01$.

MM, multiple myeloma; SMZL, splenic marginal zone lymphoma; WM, Waldeström macroglobulinemia.

1.56, $P < 0.01$), HCL (ratio 0.96, $P < 0.01$), and in normal B cells (ratio: 1.00, $P < 0.01$).[12,13,54]

Partial D-J rearrangements were present in 41% of the cases in our series (Table 2), with a similar IGHD usage to that of other B-LPDs (data not shown).

Discussion

Although DLBCL is the most frequent variant of non-Hodgkin lymphoma, its ontogeny is still not well understood. In this study, we analyzed complete V-D-J rearrangements in 165 untreated DLBCL, which is the largest series reported to date. Additionally, we also characterized the partial D-J rearrangements. Our results, based on the characterization of the BCR in DLBCLs, with a biased use of certain IGHV genes, evidence of canonical SHM, and the findings of shared amino acid changes and stereotyped sequences, provide new information about the ontogeny of DLBCL. Our data suggest that certain particular antigens must be involved in the development of DLBCL by stimulating the proliferation of B cells that express surface IG encoded by certain IGHV genes. Moreover, DLBCLs show clear differences from other B-LPDs with respect to IGHV usage (Table 5) and SHM.

A strong bias at the level of IGHV gene usage was observed, with the *IGHV4-34*, *IGHV3-23*, and *IGHV4-39* genes accounting for one third of the cohort. In addition, recurrent amino acid changes were found in certain IGHV genes, which may indicate specific driven SHM in DLBCL. These findings suggest that DLBCL could derive from certain B-cell populations specifically stimulated during the antigenic response. This hypothesis is even more plausible if the presence of some stereotyped sequences is taken into account. We identified three VH CDR3 stereotyped sequences in our series, which is a lower frequency than in CLL.[35] The stereotyped sequences had a rearrangement involving the *IGHV1-69* or *IGHV4-34* genes. Moreover, two of them shared their *IGHV1-69* rearrangement with subgroups previously described in CLL.[25,35] It is well known that there is a relationship between CLL and DLBCL, since some cases of DLBCL occur on a CLL background (Richter transformation). However, the cases reported here are *de novo* DLBCL, indicating "common" stereotypes between DLBCL and CLL, in contrast to other B-cell malignancies.[35] This highly restricted immunoglobulin gene repertoire with stereotyped VH CDR3 and common SHM targeting in DLBCL suggests a role for antigen selection in this entity, at least for some subsets of cases. Some of the antigens could be common to different B-LPDs but could ultimately lead to a different entity, depending on which population they stimulate. Although the results reported here support the notion that stereotyped BCRs in DLBCL do not seem to have as relevant implications as in CLL or other B-LPDs, the frequency of these stereotyped sequences and their possible biological and clinical significance in DLBCL need further analysis in larger and independent series.

The *IGHV4-34* gene was overrepresented in our series (15.5% overall), appearing at a higher frequency than in normal B cells,[16,20] in agreement with previous reports.[23,41,42] However, there are other series in which this percentage was slightly smaller, although this discrepancy could be accounted for in terms of differences in the sample size or the CGB/non-GCB distribution.[37,39,58] Interestingly, the patients showing *IGHV4-34* rearrangements in our series shared several common features: i) all but one were highly mutated (15 of 16 $>$ 2%, median: 9.5%, range: 1.7% to 23.6%); ii) all of them (12 of 12) belonged to the non-GCB subtype; and iii) all (10 of 10) conserved the AVY motif. Moreover, the critical W at position 7 was intact in the five cases that could be analyzed. This specific FR1 motif is thought to be responsible for binding the *N*-acetyllactosamine antigenic determinant, implying that these cells may retain the ability to bind to and be activated by superantigens, despite intense SHM activity.[59] In light of these results, our findings are consistent with those from previous studies,[42,60] and we confirm, in a larger series, the strong association between DLBCL expressing the *IGHV4-34* and non-GCB subtype. However, these results require future confirmation using genome-wide expression profiles as a tool for distinguishing between the various biological subtypes of DLBCL. In addition, we describe a new stereotyped VH CDR3 pattern in DLBCL using *IGHV4-34*. The molecular features of this subgroup differ from other B-LPDs using this gene. For example, in HCL *IGHV4-34* is mutated in only 1 of 17 cases (6%)[13]; in the small proportion of CLLs with the IgG isotype,[61] the *IGHV4-34* is the predominant gene,[25] whereas all of our *IGHV4-34* DLBCL cases belonged to the non-GCB subtype, which is assumed to be IgM^+.[42] Finally, mutations in codons 64 and 80 are found in *IGHV4-34* sequences from DLBCL and from ocular adnexal extranodal marginal zone lymphoma, which is more frequent than in normal B cells or other B-LPDs.[20,25,27,57,62] Taken together, our results suggest that DLBCL expressing IGHV4-34 may constitute a separate subgroup within the non-GCB DLBCL subtype.

Several hypotheses have been proposed to explain the origin and frequency of clonal cells from non-GCB DLBCL. Lossos et al[63] reported that non-GCB DLBCL must derive from B cells that have passed through the GC with no possibility of additional SHM. However, Jardin et al[60] observed ongoing mutations in non-GCB DLBCL, including *IGHV4-34* cases, which would suggest an independent GC mechanism of SHM. Regardless of the GC dependence of these mutations, we observed that SHM of the non-GCB subtype met the canonical criteria (*IGHV4-34* included), which do not support the need to invoke a mutational machinery different from the normal process for IGHV mutations.

In conclusion, on the basis of particular features in the sequence of the immunoglobulin genes, we found that different subgroups exist within the non-GCB subtype. Our data support the idea that DLBCL is characterized by a highly distinctive IG gene repertoire and, to the best of our knowledge, for the first time, we describe stereotyped sequences in DLBCL and "common" stereotypes between CLL and other B-LPD. These results indicate a role for antigen selection in DLBCL development and also open possibilities for future investigations into the biology of DLBCL.

Acknowledgments

We thank Alicia Antón, Felicitación García, Montserrat Hernández-Ruano (Hospital Universitario de Salamanca, Spain), and Mar López-Colomer (Centro Nacional de Investigaciones Oncológicas, Madrid, Spain) for their technical support. In addition, we thank all of the members of GELTAMO and the patients who participated in this study.

References

1. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW (Eds); International Agency for Research on Cancer: WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, 4th ed. Geneva Switzerland, World Health Organization Press, 2008
2. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr., Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM: Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000, 403:503–511
3. Hans CP, Weisenburger DD, Greiner TC, Gascoyne RD, Delabie J, Ott G, Muller-Hermelink HK, Campo E, Braziel RM, Jaffe ES, Pan Z, Farinha P, Smith LM, Falini B, Banham AH, Rosenwald A, Staudt LM, Connors JM, Armitage JO, Chan WC: Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood 2004, 103:275–282
4. Choi WW, Weisenburger DD, Greiner TC, Piris MA, Banham AH, Delabie J, Braziel RM, Geng H, Iqbal J, Lenz G, Vose JM, Hans CP, Fu K, Smith LM, Li M, Liu Z, Gascoyne RD, Rosenwald A, Ott G, Rimsza LM, Campo E, Jaffe ES, Jaye DL, Staudt LM, Chan WC: A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. Clin Cancer Res 2009, 15:5494–5502
5. Tonegawa S: Somatic generation of antibody diversity. Nature 1983, 302:575–581
6. Betz AG, Rada C, Pannell R, Milstein C, Neuberger MS: Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots. Proc Natl Acad Sci U S A 1993, 90:2385–2388
7. Wagner SD, Milstein C, Neuberger MS: Codon bias targets mutation. Nature 1995, 376:732
8. Dunn-Walters DK, Dogan A, Boursier L, MacDonald CM, Spencer J: Base-specific sequences that bias somatic hypermutation deduced by analysis of out-of-frame human IgVH genes. J Immunol 1998, 160:2360–2364
9. Smith DS, Creadon G, Jena PK, Portanova JP, Kotzin BL, Wysocki LJ: Di- and trinucleotide target preferences of somatic mutagenesis in normal and autoreactive B cells. J Immunol 1996, 156:2642–2652
10. Spencer J, Dunn M, Dunn-Walters DK: Characteristics of sequences around individual nucleotide substitutions in IgVH genes suggest different GC and AT mutators. J Immunol 1999, 162:6596–6601
11. Shapiro GS, Aviszus K, Murphy J, Wysocki LJ: Evolution of Ig DNA sequence to target specific base positions within codons for somatic hypermutation. J Immunol 2002, 168:2302–2306
12. Messmer BT, Albesiano E, Messmer D, Chiorazzi N: The pattern and distribution of immunoglobulin VH gene mutations in chronic lymphocytic leukemia B cells are consistent with the canonical somatic hypermutation process. Blood 2004, 103:3490–3495
13. Arons E, Roth L, Sapolsky J, Suntum T, Stetler-Stevenson M, Kreitman RJ: Evidence of canonical somatic hypermutation in hairy cell leukemia. Blood 2011, 117:4844–4851

14. Weller S, Faili A, Garcia C, Braun MC, Le Deist FF, de Saint Basile GG, Hermine O, Fischer A, Reynaud CA, Weill JC: CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. Proc Natl Acad Sci U S A 2001, 98:1166–1170
15. Chiorazzi N, Ferrarini M: B cell chronic lymphocytic leukemia: lessons learned from studies of the B cell antigen receptor. Annu Rev Immunol 2003, 21:841–894
16. Berkowska MA, Driessen GJ, Bikos V, Grosserichter-Wagener C, Stamatopoulos K, Cerutti A, He B, Biermann K, Lange JF, van der Burg M, Van Dongen JJ, van Zelm MC: Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. Blood 2011, 118:2150–2158
17. Chang B, Casali P: The CDR1 sequences of a major proportion of human germline Ig VH genes are inherently susceptible to amino acid replacement. Immunol Today 1994, 15:367–373
18. Lossos IS, Tibshirani R, Narasimhan B, Levy R: The inference of antigen selection on Ig genes. J Immunol 2000, 165:5122–5126
19. Uduman M, Yaari G, Hershberg U, Stern JA, Shlomchik MJ, Kleinstein SH: Detecting selection in immunoglobulin sequences. Nucleic Acids Res 2011, 39:W499–W504
20. Brezinschek HP, Foster SJ, Brezinschek RI, Dorner T, Domiati-Saad R, Lipsky PE: 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 1997, 99: 2488–2501
21. Wu YC, Kipling D, Leong HS, Martin V, Ademokun AA, Dunn-Walters DK: High-throughput immunoglobulin repertoire analysis distinguishes between human IgM memory and switched memory B-cell populations. Blood 2010, 116:1070–1078
22. Collins AM, Sewell WA, Edwards MR: Immunoglobulin gene rearrangement, repertoire diversity, and the allergic response. Pharmacol Ther 2003, 100:157–170
23. Hsu FJ, Levy R: Preferential use of the VH4 Ig gene family by diffuse large-cell lymphoma. Blood 1995, 86:3072–3082
24. Gonzalez D, van der Burg M, Garcia-Sanz R, Fenton JA, Langerak AW, Gonzalez M, Van Dongen JJ, San Miguel JF, Morgan GJ: Immunoglobulin gene rearrangements and the pathogenesis of multiple myeloma. Blood 2007, 110:3112–3121
25. Stamatopoulos K, Belessi C, Moreno C, Boudjograh M, Guida G, Smilevska T, Belhoul L, Stella S, Stavroyianni N, Crespo M, Hadzidimitriou A, Sutton L, Bosch F, Laoutaris N, Anagnostopoulos A, Montserrat E, Fassas A, Dighiero G, Caligaris-Cappio F, Merle-Beral H, Ghia P, Davi F: Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: pathogenetic implications and clinical correlations. Blood 2007, 109:259–270
26. Martin-Jimenez P, Garcia-Sanz R, Balanzategui A, Alcoceba M, Ocio E, Sanchez ML, Gonzalez M, San Miguel J: Molecular characterization of heavy chain immunoglobulin gene rearrangements in Waldenstrom's macroglobulinemia and IgM monoclonal gammopathy of undetermined significance. Haematologica 2007, 92:635–642
27. Hadzidimitriou A, Agathangelidis A, Darzentas N, Murray F, fau-Larue MH, Bredo PL, Navarro LA, Dagklis A, Rombout P, Beldjord K, Kolstad A, Dreyling MH, Anagnostopoulos A, Tsaftaris A, Mavragani-Tsipidou P, Rosenwald A, Ponzoni M, Groenen P, Ghia P, Sander B, Papadaki T, Campo E, Geisler C, Rosenquist R, Davi F, Pott C, Stamatopoulos K: Is there a role for antigen selection in mantle cell lymphoma? Immunogenetic support from a series of 807 cases. Blood 2011, 118:3088–3095
28. Bikos V, Darzentas N, Hadzidimitriou A, Davis Z, Hockley S, Traverse-Glehen A, Algara P, Santoro A, Gonzalez D, Mollejo M, Dagklis A, Gangemi F, Bosler DS, Bourikas G, Anagnostopoulos A, Tsaftaris A, Iannitto E, Ponzoni M, Felman P, Berger F, Belessi C, Ghia P, Papadaki T, Dogan A, Degano M, Matutes E, Piris MA, Oscier D, Stamatopoulos K: Over 30% of patients with splenic marginal zone lymphoma express the same immunoglobulin heavy variable gene: ontogenetic implications. Leukemia 2012, 26:1638–1646
29. Zibellini S, Capello D, Forconi F, Marcatili P, Rossi D, Rattotti S, Franceschetti S, Sozzi E, Cencini E, Marasca R, Baldini L, Tucci A, Bertoni F, Passamonti F, Orlandi E, Varettoni M, Merli M, Rizzi S, Gattei V, Tramontano A, Paulli M, Gaidano G, Arcaini L: Stereotyped patterns of B-cell receptor in splenic marginal zone lymphoma. Haematologica 2010, 95:1792–1796
30. Dagklis A, Ponzoni M, Govi S, Cangi MG, Pasini E, Charlotte F, Vino A, Doglioni C, Davi F, Lossos IS, Ntountas I, Papadaki T, Dolcetti R, Ferreri AJ, Stamatopoulos K, Ghia P: Immunoglobulin gene repertoire in ocular adnexal lymphomas: hints on the nature of the antigenic stimulation. Leukemia 2012, 26:814–821
31. Chan CH, Hadlock KG, Foung SK, Levy S: V(H)1–69 gene is preferentially used by hepatitis C virus-associated B cell lymphomas and by normal B cells responding to the E2 viral antigen. Blood 2001, 97:1023–1026
32. Cappione AJ, Pugh-Bernard AE, Anolik JH, Sanz I: Lupus IgG VH4.34 antibodies bind to a 220-kDa glycoform of CD45/B220 on the surface of human B lymphocytes. J Immunol 2004, 172:4298–4307
33. Messmer BT, Albesiano E, Efremov DG, Ghiotto F, Allen SL, Kolitz J, Foa R, Damle RN, Fais F, Messmer D, Rai KR, Ferrarini M, Chiorazzi N: Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. J Exp Med 2004, 200:519–525
34. Murray F, Darzentas N, Hadzidimitriou A, Tobin G, Boudjogra M, Scielzo C, Laoutaris N, Karlsson K, Baran-Marzsak F, Tsaftaris A, Moreno C, Anagnostopoulos A, Caligaris-Cappio F, Vaur D, Ouzounis C, Belessi C, Ghia P, Davi F, Rosenquist R, Stamatopoulos K: Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. Blood 2008, 111:1524–1533
35. Agathangelidis A, Darzentas N, Hadzidimitriou A, Brochet X, Murray F, Yan XJ, Davis Z, van Gastel-Mol EJ, Tresoldi C, Chu CC, Cahill N, Giudicelli V, Tichy B, Pedersen LB, Foroni L, Bonello L, Janus A, Smedby K, Anagnostopoulos A, Merle-Beral H, Laoutaris N, Juliusson G, di Celle PF, Pospisilova S, Jurlander J, Geisler C, Tsaftaris A, Lefranc MP, Langerak AW, Oscier DG, Chiorazzi N, Belessi C, Davi F, Rosenquist R, Ghia P, Stamatopoulos K: Stereotyped B-cell receptors in one third of chronic lymphocytic leukemia: towards a molecular classification with implications for targeted therapeutic interventions. Blood 2012, 119:4467–4475
36. Tobin G, Thunberg U, Johnson A, Eriksson I, Soderberg O, Karlsson K, Merup M, Juliusson G, Vilpo J, Enblad G, Sundstrom C, Roos G, Rosenquist R: Chronic lymphocytic leukemias utilizing the VH3-21 gene display highly restricted Vlambda2-14 gene use and homologous CDR3s: implicating recognition of a common antigen epitope. Blood 2003, 101:4952–4957
37. Kuppers R, Rajewsky K, Hansmann ML: Diffuse large cell lymphomas are derived from mature B cells carrying V region genes with a high load of somatic mutation and evidence of selection for antibody expression. Eur J Immunol 1997, 27:1398–1405
38. Rosenquist R, Lindstrom A, Holmberg D, Lindh J, Roos G: V(H) gene family utilization in different B-cell lymphoma subgroups. Eur J Haematol 1999, 62:123–128
39. Lossos IS, Okada CY, Tibshirani R, Warnke R, Vose JM, Greiner TC, Levy R: Molecular analysis of immunoglobulin genes in diffuse large B-cell lymphomas. Blood 2000, 95:1797–1803
40. Nakamura N, Kuze T, Hashimoto Y, Hara Y, Hoshi S, Sasaki Y, Shirakawa A, Sato M, Abe M: Analysis of the immunoglobulin heavy chain gene variable region of CD5-positive and -negative diffuse large B cell lymphoma. Leukemia 2001, 15:452–457
41. Capello D, Martini M, Gloghini A, Cerri M, Rasi S, Deambrogi C, Rossi D, Spina M, Tirelli U, Larocca LM, Carbone A, Gaidano G: Molecular analysis of immunoglobulin variable genes in human immunodeficiency virus-related non-Hodgkin's lymphoma reveals implications for disease pathogenesis and histogenesis. Haematologica 2008, 93:1178–1185
42. Ruminy P, Etancelin P, Couronne L, Parmentier F, Rainville V, Mareschal S, Bohers E, Burgot C, Cornic M, Bertrand P, Lenormand B, Picquenot JM, Jardin F, Tilly H, Bastard C: The isotype of the BCR as a surrogate for the GCB and ABC molecular subtypes in diffuse large B-cell lymphoma. Leukemia 2011, 25:681–688
43. Catherwood MA, Gonzalez D, Patton C, Dobbin E, Venkatraman L, Alexander HD: Improved clonality assessment in germinal centre/post-germinal centre non-Hodgkin's lymphomas with high rates of somatic hypermutation. J Clin Pathol 2007, 60:524–528
44. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. Blood 1997, 89:3909–3918
45. Gonzalez M, Gonzalez D, Lopez-Perez R, Garcia-Sanz R, Chillon MC, Balanzategui A, Mateos MV, Alaejos I, Langerak AW, Orfao A, Van Dongen JJ, San Miguel JF: Heteroduplex analysis of VDJ amplified segments from rearranged IgH genes for clonality assessments in

B-cell non-Hodgkin's lymphoma. A comparison between different strategies. Haematologica 1999, 84:779–784
46. Van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL, Delabesse E, Davi F, Schuuring E, Garcia-Sanz R, van Krieken JH, Droese J, Gonzalez D, Bastard C, White HE, Spaargaren M, Gonzalez M, Parreira A, Smith JL, Morgan GJ, Kneba M, Macintyre EA: Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia 2003, 17: 2257–2317
47. Lefranc MP, Pommie C, Ruiz M, Giudicelli V, Foulquier E, Truong L, Thouvenin-Contet V, Lefranc G: IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. Dev Comp Immunol 2003, 27:55–77
48. Tsakou E, Agathagelidis A, Boudjoghra M, Raff T, Dagklis A, Chatzouli M, Smilevska T, Bourikas G, Merle-Beral H, Manioudaki-Kavallieratou E, Anagnostopoulos A, Bruggemann M, Davi F, Stamatopoulos K, Belessi C: Partial versus productive IGH locus rearrangements in chronic lymphocytic leukemia: implications for B-cell receptor stereotypy. Mol Med 2012, 18:138–145
49. Gonzalez D, Gonzalez M, Balanzategui A, Sarasquete ME, Lopez-Perez R, Chillon MC, Garcia-Sanz R, San Miguel JF: Molecular characteristics and gene segment usage in IGH gene rearrangements in multiple myeloma. Haematologica 2005, 90:906–913
50. Martin-Jimenez P, Garcia-Sanz R, Gonzalez D, Balanzategui A, Perez JJ, Caballero MD, Sarasquete ME, Galende J, Orfao A, Lopez-Berges MC, San Miguel JF, Gonzalez M: Molecular characterization of complete and incomplete immunoglobulin heavy chain gene rearrangements in hairy cell leukemia. Clin Lymphoma Myeloma 2007, 7:573–579
51. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK: Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood 1999, 94:1848–1854
52. Davis ZA, Orchard JA, Corcoran MM, Oscier DG: Divergence from the germ-line sequence in unmutated chronic lymphocytic leukemia is due to somatic mutation rather than polymorphisms. Blood 2003, 102:3075
53. Hockley SL, Giannouli S, Morilla A, Wotherspoon A, Morgan GJ, Matutes E, Gonzalez D: Insight into the molecular pathogenesis of hairy cell leukaemia, hairy cell leukaemia variant and splenic marginal zone lymphoma, provided by the analysis of their IGH rearrangements and somatic hypermutation patterns. Br J Haematol 2010, 148:666–669
54. Di Noia JM, Neuberger MS: Molecular mechanisms of antibody somatic hypermutation. Annu Rev Biochem 2007, 76:1–22
55. Darzentas N, Hadzidimitriou A, Murray F, Hatzi K, Josefsson P, Laoutaris N, Moreno C, Anagnostopoulos A, Jurlander J, Tsaftaris A, Chiorazzi N, Belessi C, Ghia P, Rosenquist R, Davi F, Stamatopoulos K: A different ontogenesis for chronic lymphocytic leukemia cases carrying stereotyped antigen receptors: molecular and computational evidence. Leukemia 2010, 24:125–132
56. Smit LA, Bende RJ, Aten J, Guikema JE, Aarts WM, van Noesel CJ: Expression of activation-induced cytidine deaminase is confined to B-cell non-Hodgkin's lymphomas of germinal-center phenotype. Cancer Res 2003, 63:3894–3898
57. Klein U, Rajewsky K, Kuppers R: Human immunoglobulin (Ig)M+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 1998, 188:1679–1689
58. Daley MD, Berinstein NL, Siminovitch KA: Immunoglobulin heavy chain variable gene utilization in human large cell and Burkitt's lymphoma cell lines. Clin Invest Med 1994, 17:522–530
59. Li Y, Spellerberg MB, Stevenson FK, Capra JD, Potter KN: The I binding specificity of human VH 4–34 (VH 4–21) encoded antibodies is determined by both VH framework region 1 and complementarity determining region 3. J Mol Biol 1996, 256:577–589
60. Jardin F, Sahota SS, Ruminy P, Parmentier F, Picquenot JM, Rainville V, Buchonnet G, Lepretre S, Tilly H, Bastard C: Novel Ig V gene features of t(14;18) and t(3;14) de novo diffuse large B-cell lymphoma displaying germinal center-B cell like and non-germinal center-B cell like markers. Leukemia 2006, 20:2070–2074
61. Potter KN, Mockridge CI, Neville L, Wheatley I, Schenk M, Orchard J, Duncombe AS, Packham G, Stevenson FK: Structural and functional features of the B-cell receptor in IgG-positive chronic lymphocytic leukemia. Clin Cancer Res 2006, 12:1672–1679
62. Zhu D, Lossos C, Chapman-Fredricks JR, Matthews JM, Ikpatt OF, Ruiz P, Lossos IS: Biased use of the IGHV4 family and evidence for antigen selection in Chlamydophila psittaci-negative ocular adnexal extranodal marginal zone lymphomas. PLoS One 2011, 6:e29114
63. Lossos IS, Alizadeh AA, Eisen MB, Chan WC, Brown PO, Botstein D, Staudt LM, Levy R: Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. Proc Natl Acad Sci U S A 2000, 97:10209–10213