Analysis of the Human V_H Gene Repertoire

Differential Effects of Selection and Somatic Hypermutation on Human Peripheral $CD5^+/IgM^+$ and $CD5^-/IgM^+$ B cells

Hans-Peter Brezinschek, Sandra J. Foster, Ruth I. Brezinschek, Thomas Dörner, Rana Domiati-Saad, and Peter E. Lipsky
Department of Internal Medicine and Harold C. Simmons Arthritis Research Center, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75235

Abstract

To analyze the immunoglobulin repertoire of human IgM^+ B cells and the $CD5^+$ and $CD5^-$ subsets, individual $CD19^+/IgM^+/CD5^+$ or $CD5^-$ B cells were sorted and non-productive as well as productive V_H gene rearrangements were amplified from genomic DNA and sequenced. In both subsets, the V_H3 family was overrepresented largely as a result of preferential usage of a small number of specific individual family members. In the $CD5^+$ B cell subset, all other V_H families were found at a frequency expected from random usage, whereas in the $CD5^-$ population, V_H4 appeared to be overrepresented in the nonproductive repertoire, and also negatively selected since it was found significantly less often in the productive compared to the nonproductive repertoire; the V_H1 family was significantly diminished in the productive rearrangements of $CD5^-$ B cells. 3-23/DP-47 was the most frequently used V_H gene segment and was found significantly more often than expected from random usage in productive rearrangements of both $CD5^+$ and $CD5^-$ B cells. Evidence for selection based on the D segment and the J_H gene usage was noted in $CD5^+$ B cells. No differences were found between the B cell subsets in CDR3 length, the number of N-nucleotides or evidence of exonuclease activity. Somatically hypermutated V_HDJ_H rearrangements were significantly more frequent and extensive in $CD5^-$ compared to $CD5^+$ IgM^+ B cells, indicating that IgM^+ memory B cells were more frequent in the $CD5^-$ B cell population. Of note, the frequency of specific V_H genes in the mutated population differed from that in the nonmutated population, suggesting that antigen stimulation imposed additional biases on the repertoire of IgM^+ B cells. These results indicate that the expressed repertoire of IgM^+ B cell subsets is shaped by recombinational bias, as well as selection before and after antigen exposure. Moreover, the influences on the repertoires of $CD5^+$ and $CD5^-$ B cells are significantly different, suggesting that human peripheral blood $CD5^+$ and $CD5^-$ B cells represent different B cell lineages, with similarities to murine B-1a and B-2 subsets, respectively. (*J. Clin. Invest.* 1997. 99:2488–2501.) Key words: immunoglobulin • variable heavy chain genes • VDJ rearrangement • selection • hypermutation

Introduction

The V region of Ab is generated by random rearrangement of five gene segments, three for the H-chain, including the V_H, D and the J_H gene segments and the V_L and J_L elements for the L-chain (1–3). The rearrangements occur in an ordered fashion beginning with the heavy chain locus on chromosome 14. A μ chain protein with an antigen combining region encoded by a successfully rearranged V_HDJ_H gene is then expressed on the surface of a B cell progenitor together with a surrogate L chain (ψL). During this stage of differentiation, interactions with self antigens could lead to positive or negative selection of B cells expressing specific V_H germline gene products (4, 5). As pairing of H chains with ψL chain is thought to mask availability of the CDR3 (6), selection at this stage of B cell maturation is proposed to relate largely to the nature of the V_H segment expressed. In the mouse, selection of the Ig repertoire based on V_H expression has been suggested to occur (7). As soon as an L chain gene has been successfully rearranged, B cells express dimers of μ and L chain heterodimers on their surface. At this stage of differentiation, positive and negative selection could also occur, but now based upon the variable regions and the CDR3s of both the μ and the L chains. The impact of these putative selection processes on the expressed Ig repertoire has not been clearly delineated.

Peripheral B cells are usually classified as either naive or memory B cells, based on their antigenic experience and surface phenotype. Naive B cells express surface IgM and IgD with V_H genes in an unmutated germline configuration. These B cells account for the majority of mature, peripheral IgM^+ B cells. It has been shown in humans that this population is generated throughout life (8). Memory B cells, on the other hand, are generally defined as B cells that have undergone Ig heavy chain isotype switching and have been selected for higher affinity binding to specific antigen. They express somatically mutated V region genes in association with downstream heavy chain isotype constant regions and persist in the immune system for a long period of time (9). Whether a population of memory cells is also generated during immune responses that has undergone somatic mutation of V regions, but not isotype switch is a matter of controversy.

IgM^+ B cells can be subgrouped into two subsets, termed B-1 and B-2 that can be separated in the mouse by their phenotype, anatomical location, as well as functional characterization (10, 11). In the mouse, B-1 ($CD5^+$) B cells have been implicated as the major source of serum IgM that binds a variety of antigens, including autoantigens (10–12). It has been shown

Address correspondence to Dr. Peter E. Lipsky, Department of Medicine, Harold C. Simmons Arthritis Research Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd., Dallas, TX 75235. Phone: 214-648-9110; FAX: 214-648-7995; E-mail: peter.lipsky@email.swmed.edu

Received for publication 20 November 1996 and accepted in revised form 7 March 1997.

J. Clin. Invest.

0021-9738/97/05/2488/14 $2.00
Volume 99, Number 10, May 1997, 2488–2501

that this murine B cell subset belongs to a distinct lineage that uses only a limited number of V_H germline genes and exhibits restricted N-region diversity and exonuclease activity at the ends of J_H gene segments. Furthermore, B-1 B cells have been found to contain less mutations and a limited capacity to undergo isotype-switching (13, 14). It has been proposed that this is the result of the limited ability of B-1, compared to B-2 B cells, to form germinal centers and reflects the tendency of B-1 B cells primarily to become IgM-secreting cells (15).

The relationship between the human and murine B-1 B cell populations remains controversial. It has been suggested that human B-1 equivalent B cells defined by expression of CD5 use a limited set of V_H germline genes, as was shown for murine B-1 B cells (10). On the other hand, analysis of V_H 5 and V_H6 containing IgM cDNAs from human peripheral $CD5^+$ and $CD5^-$ B cells indicated that somatic mutation was not limited to either subpopulation (16). Furthermore, analysis of the antigen-binding CDR3 from human cord blood B cells revealed no differences between $CD5^+$ and $CD5^-$ subsets (17). Thus, the relationship between murine B-1 B cells and human $CD5^+$ B cells is not clear, although the molecular analysis of human $CD5^+$ B cells has been incomplete.

The goals of this study were, therefore, to examine the V_H repertoire of human IgM^+ B cells to determine whether selective influences might shape the expressed Ig repertoire. In addition, the molecular organization of the V_H repertoire of human $CD5^+$ B cells was assessed and compared with that of the $CD5^-$ population. Finally, the frequency of mutations in $IgM^+/CD5^+$ and $IgM^+/CD5^-$ B cell subsets was analyzed and compared. A unique single cell PCR technique that permitted amplification and sequence analysis of both the productive and the nonproductive V_H rearrangements from genomic DNA of individual IgM^+ B cells (18) was used to address these issues.

Methods

Cell preparation and sorting. Peripheral blood was obtained from two healthy male donors (donor 1 was a 26-yr-old Hispanic, and donor 2 was a 45-yr-old Caucasian). Mononuclear cells were prepared by ficoll hypaque gradient centrifugation of heparinized blood (19) and B cells were enriched using a commercially available kit (CEPRATE LC Kit; Cellpro, Inc., Bothwell, WA) that permits selection of $CD19^+$ cells. Cells were stained thereafter with PE-labeled anti–human CD19 antibody (Sigma Chemical Co., St. Louis, MO), FITC-labeled anti–human IgM antibody (PharMingen, San Diego, CA), a biotinylated anti–human CD5 antibody (Becton-Dickinson, San Jose, CA) and RED613-labeled streptavidin (Gibco BRL, Gaithersburg, MD). An individual $CD19^+/IgM^+/CD5^+$ or $CD5^-$ B cell was sorted into each well of 96 well PCR plates (Robbins Scientific, Sunnyvale, CA) assembled on a microAmp base (Perkin-Elmer, Norwalk, CT) using a FACStar^{Plus} flow cytometer outfitted with an automatic cell deposition unit (Becton-Dickinson). Each well contained 5 μl of an alkaline lysing solution (200 mM KOH/ 50 mM DTT).

Single cell PCR. Genomic DNA of individual $CD19^+/IgM^+/CD5^+$ or $CD19^+/IgM^+/CD5^-$ B cells was analyzed for rearranged immunoglobulin (Ig) genes using a single cell PCR technique that permits the amplification of both productive and nonproductive V_H rearrangements (18). An initial primer extension preamplification step employing random 15mers and 60 rounds of amplification with Taq polymerase (Promega, Madison, WI) was used to produce sufficient DNA for multiple subsequent specific amplifications. To accomplish this, 5-μl aliquots of template were subjected to nested V_H family specific PCR amplification. In the first step, 35 cycles of amplification with Taq polymerase were used with a battery of V_H-leader region primers, each specific for a V_H family and each with the same annealing temperature, and external J_H region specific primers. In the second round, aliquots were specifically amplified for 32 cycles with Taq polymerase for each V_H-family using family-specific primers complementary to the 5′ end of FR1 and internal J_H region primers (18). Aliquots (10 μl) of the final PCR products were then separated by electrophoresis using a 1.5% agarose gel and were analyzed for products of the anticipated size.

Sequence analysis. PCR products were cut from the gel and purified by gel electrophoresis through 2% SeaKem LE agarose (FMC Bioproducts, Rockland, ME) using GeneCapsule (Geno Technology, Inc., St. Louis, MO) or GenElute (Supelco, Bellefonte, PA) and directly sequenced with an automated DNA sequencer (ABI Prism 377; Perkin Elmer Corp.) using the ABI Prism Dye Termination Cycle Sequencing Kit (Perkin Elmer Corp.) and the 5′ primer used for the second round of amplification. The DNA sequences obtained were analyzed using GeneWorks software (release 2.45; IntelliGenetics, Inc., Mountain View, CA) and Sequencher (Gene Codes Corporation, Ann Arbor, MI) and the V BASE Sequence Directory (20). For identification of a D gene segment, the criteria of Klein et al. (21) were used. In brief, a D segment was identified when it shared at least 7 consecutive nucleotides with a known D segment or 8 nucleotides of sequence identity interrupted by no more than 1 nucleotide of substitution. Whenever a sequence was found that contained stop codons within FR1 to FR4 or an out of frame rearrangement, it was considered to be nonproductive, because it could not encode a protein product. All sequences determined in this study can be found in the GeneBank, EMBL and DDBJ Nucleotide Sequence Database under accession numbers Z80363-Z80770.

Efficiency of the single-cell PCR. A total of 368 cells per B cell subset were sorted and an identifiable PCR product was obtained from 161 wells of $CD5^+/IgM^+$ and 212 wells of $CD5^-/IgM^+$ B cells. The average (±SD) efficiency per PCR-plate was 40.3% (±5.1) and 53.0% (±12.8), respectively. All but 5 V_H germline genes (3-21/DP-77, 2-26/DP-26, 4-28/DP-68, 1-45/DP-4, and 1-58/DP-2; see reference 20) were detected in this study. Furthermore, germline genes that are located in regions of insertion/deletion polymorphism, such as 7-4.1/DP-21, 4-30.2/DP-64, 3-30.3/DP-46, 4-30.4/DP-78, 4-b/DP-67, 4-61/DP-66, and 1-e/DP-88, or that have not been mapped (-/DP-58) were detected. In addition, two pseudogenes (3-22/DP-57 and 3-52/DP-43) were detected in V_HDJ_H rearrangements. None of the nonfunctional germline genes that has an open reading frame and none of the orphan genes were found in this analysis.

Estimation of the Taq-polymerase error. The error-rate of Taq-polymerase in this single-cell PCR technique was calculated by sequencing V_H and V_κ genes from a 4-yr-old patient with Hyper-IgM syndrome, that would be expected to have minimal mutations (22). 23 mutations out of 13,696 V_H nucleotides and 11 out of 9,429 V_κ nucleotides were found, giving a maximal PCR error rate of 1.7×10^{-3} mutations per base for V_H gene segments and 1.2×10^{-3} mutations per base for V_κ gene segments. These error rates correspond to $\sim$ 0.3 to 0.5 mutations per V region and are in agreement with the error rate calculated from sequencing Ig genes of hybridomas and splenic Ab-forming cells (23).

Statistical methods. To determine significant differences in distributions in productive or nonproductive rearrangements, the chi-square test was used. *P*-values equal or less than 0.05 were assumed to be significant. Overall statistical significance between observed and expected frequencies was calculated using the chi-square goodness-of-fit statistic. If significant, each of the single degree of freedom chi-squares was examined for significant contribution to the total. The *P*-value for the significance of these single degree of freedom chi-squares was then adjusted for the accumulation of errors related to multiple testing according to the Bonferonni method (24). Additional statistical analyses included Student's *t* test (CDR3-length, number of N-nucleotides) or Mann-Whitney test (exonuclease-activity).

Table I. Distribution of V_H Families in $CD5^+/IgM^+$ and $CD5^-/IgM^+$ B Cells

		$CD5^+/IgM^+$		$CD5^-/IgM^+$	
	Known functional genes per V_H family*	Productive rearrangements	Nonproductive rearrangements‡	Productive rearrangements	Nonproductive rearrangements
V_H1	11 (21.6%)	27 (18.8%)	6 (20.7%)	27 (13.1%)§	1 (4.0%)
V_H2	3 (5.9%)	3 (2.1%)	1 (3.4%)	4 (1.9%)	2 (8.0%)
V_H3	22 (43.1%)	81 (56.3%)§	13 (44.8%)	111 (53.9%)§	9 (36.0%)
V_H4	11 (21.6%)	26 (18.1%)	6 (20.7%)	51 (24.8%)‖	11 (44.0%)§‖
V_H5	2 (3.9%)	2 (1.4%)	2 (6.9%)	6 (2.9%)	1 (4.0%)
V_H6	1 (2.0%)	2 (1.4%)	1 (3.4%)	5 (2.4%)	1 (4.0%)
V_H7	1 (2.0%)	3 (2.1%)	0	2 (1.0%)	0
Total	51	144	29	206	25

*The number of V_H genes per family observed as V_HDJ_H recombinants was obtained from the V BASE Sequence Directory (20). ‡Three V_H pseudogenes, two in the $CD5^+$ and one in the $CD5^-$ B cell population, were omitted from this analysis. §Significant difference between expected and observed frequency ($P \leq 0.008$). ‖Significant difference between productive and nonproductive rearrangements ($P \leq 0.05$).

Results

Distribution of V_H families in $CD5^+$ and $CD5^-$ IgM^+ peripheral blood B cells. V_H gene sequences were obtained from genomic DNA, and as a result, both the productively and the nonproductively rearranged genes were routinely amplified and could be analyzed (Table I). It should be noted that there was no relationship between the V_H family used for productive and nonproductive rearrangements, other than the infrequent use of the same V_H family for both rearrangements by an individual B cell (18). In general, the distribution of the V_H families in the productive and the non-productive repertoires of the $CD5^+$ B cell population resembled the germline complexity, i.e., V_H3, the largest family was found most often, followed by V_H4 and V_H1. However, the V_H3 family was significantly over-represented in the productive repertoire of $CD5^+$ B cells ($P \leq 0.006$). In contrast, in the $CD5^-$ B cell subset, a distribution of V_H genes reflecting germ-line complexity was only found in the productive rearrangements, whereas in the nonproductive repertoire, the V_H4 family was detected most often and significantly more than expected from random chance ($P \leq 0.008$). The distribution of the V_H3 and the V_H1 families within the productive repertoire was also significantly different than expected, with the V_H1 family found significantly less often and the V_H3 family significantly more often than expected ($P \leq 0.008$).

Evidence for negative and positive selection of specific V_H families was noted. Thus, there was an indication of negative selection of the V_H4 family in the $CD5^-$ B cell subset in that the frequency of V_H4 members in the productive repertoire

Table II. Distribution of D Families and J_H Genes in $CD5^-/IgM^+$ B Cells Using the V_H3 or the V_H4 Family

	Productive rearrangements			Nonproductive rearrangements		
	All V_H families	$V_H3(+)$	$V_H4(+)$	All V_H families	$V_H3(+)$	$V_H4(+)$
DA	17 (8.5%)	11 (9.6%)	4 (8.0%)	1 (2.9%)	0	1 (7.1%)
DHQ52	6 (3.0%)	4 (3.5%)	1 (2.0%)	2 (5.9%)	1 (6.7%)	1 (7.1%)
DIR	35 (17.5%)	16 (14.0%)	9 (18.0%)	6 (17.6%)	4 (26.7%)	1 (7.1%)
DK	16 (8.0%)	11 (9.6%)	4 (8.0%)	3 (8.8%)	1 (6.7%)	2 (14.3%)
DLR	33 (16.5%)	22 (19.3%)	4 (8.0%)	8 (23.5%)	6 (40.0%)	1(7.1%)
DM	9 (4.5%)	6 (5.3%)	3 (6.0%)	1 (2.9%)	0	1 (7.1%)
DN	16 (8.0%)	6 (5.3%)	4 (8.0%)	1 (2.9%)	0	1 (7.1%)
DψA1	10 (5.0%)	8 (7.0%)	1 (2.0%)	3 (8.8%)	1 (6.7%)	2 (14.3%)
DXP	58 (29.0%)	30 (26.3%)	20 (40.0%)	9 (26.5%)	2 (13.3%)	4 (28.6%)
Total	200	114	50	34	15	14
J_H1	2 (1.0%)	1 (0.9%)	0	1 (3.8%)	1 (10.0%)	0
J_H2	5 (2.4%)	2 (1.8%)	2 (3.9%)	0	0	0
J_H3	16 (7.8%)	7 (6.3%)	5 (9.8%)	2 (7.7%)	1 (10.0%)	1 (9.1%)
J_H4	114 (55.3%)	64 (57.7%)	30 (58.8%)	13 (50.0%)	5 (50.0%)	7 (63.6%)
J_H5	20 (9.7%)	8 (7.2%)	6 (11.8%)	4 (15.4%)	2 (20.0%)	1 (9.1%)
J_H6	49 (23.8%)	29 (26.1%)	8 (15.7%)	6 (23.1%)	1 (10.0%)	2 (18.2%)
Total	206	111	51	26	10	11

(51/206, 24.8%) was significantly less than in the nonproductive repertoire (11/25, 44.0%; $P \leq 0.04$). Positive selection of the V_H3 family in both B cell subsets was suggested since this family was observed more frequently than expected in the productive ($P \leq 0.008$), but not the nonproductive repertoires. Moreover, direct comparison of the frequency of the V_H3 family in the productive and the nonproductive repertoires of the combined IgM^+ B cell populations confirmed that it was significantly overrepresented ($P \leq 0.05$) in the productive rearrangements of IgM^+ B cells (data not shown). Of note, the distribution of D-family members and J_H genes of productive and nonproductive rearrangements utilizing V_H3 or V_H4 families was comparable and not different than that predicted from the frequency of these elements in the genome (Table II). These results, therefore, suggest that the frequency of V_H3 and V_H4 family members was altered in the expressed repertoires of both $CD5^+$ and $CD5^-$ B cells by a mechanism that was independent of D and J_H use.

A small number of V_H genes are expressed by a majority of human $CD5^+$ and $CD5^-$ B cells. When the distribution of individual V_H genes was examined, marked differences in the expression of genes compared to their representation in the genome was noted. A small number of individual genes accounted for the majority of the productive repertoire of both B cell subpopulations (Table III). Thus, nine V_H family members, or 18% of the functional V_H genes, were expressed by more than 50% of B cells. In the combined IgM^+ B cell population, 3-23/DP-47 was found significantly more often in the productive versus the nonproductive repertoires (45/350, 12.4% vs. 2/57, 3.5%; $P \leq 0.04$), whereas a total of eight V_H genes were found significantly more often in the productive repertoire of IgM^+ B cells than expected (data not shown). Similarly, seven of these genes, including 3-23/DP-47, were significantly overrepresented in the productive repertoire of $CD5^-$ B cells and when the nine most frequently used V_H genes were grouped together, there was a significant increase in their frequency in the productive versus the nonproductive repertoire of $CD5^-$ B cells (114/206 vs. 9/26, $P \leq 0.046$). Only five of these genes, including 3-23/DP-47, were overrepresented in the $CD5^+$ B cell subset and, as an aggregate, a significant difference between productive and nonproductive repertoires of $CD5^+$ B cells was not found. Of note, and as shown in Table IV, the CDR3s of the productive V_HDJ_H rearrangements employing 3-23/DP-47 showed no preferential use of D or J_H segments, no clonal relationship with one exception (BF2N2-g3B07 and BF2N2-g3F07) and no preferential V_K usage (data not shown). Similar findings were noted for the other V_H genes overrepresented in the productive repertoire. These results, therefore, suggest that biases in the representation of individual V_H genes are likely to be dependent on the V_H gene segment itself. Some V_H genes, such as 3-23/DP-47, were overrepresented in the productive but not the nonproductive repertoire, indicating that they were likely to be positively selected. In contrast, two V_H genes, 4-59/DP-71 in the $CD5^-$ and 3-07/DP-54 in the $CD5^+$ population, were overrepresented in both the productive and the nonproductive repertoires, suggesting that recombinational bias rather than selection accounted for their increased frequency in the B cell subsets.

Among the less frequently used V_H genes, differences in distribution between productive and nonproductive repertoires were also noted. Thus, in the $CD5^+$ B cell population, 3-66/DP-86, was significantly diminished in the productive repertoire (0/144 vs. 1/29; $P \leq 0.025$), whereas in the $CD5^-$ subpopulation three V_H gene segments, namely 4-31/DP-65 (3/206 vs. 2/25; $P \leq 0.034$), 3-43/DP-33 (0/206 vs. 1/25; $P \leq 0.004$) and 2-70/DP-27 (1/206 vs. 2/25; $P \leq 0.002$) were found significantly less frequently in the productive repertoire. It should be emphasized, however, that these genes were found infrequently and, therefore, the biologic significance of these differences is uncertain.

Usage of D segments by $CD5^+$ and $CD5^-$ B cells. The distribution of D gene segments was also found to differ from that expected from their representation in the genome (Table V). There were, however, no significant differences in D segment use between the B cell subsets. In the nonproductive repertoire of $CD5^+$ B cells, the DIR family was observed significantly more often than expected. In the productive repertoire, two D families, DIR and DN, were observed significantly more than expected, whereas one D family, DLR, was found less frequently than expected. Evidence of negative selection of the DLR family in $CD5^+$ B cells was apparent, as it appeared significantly less often in the productive compared to the nonproductive repertoires (13/155 vs. 13/48; $P \leq 0.001$). The variance in the frequency of DLR family members in the productive versus nonproductive repertoires of $CD5^+$ B cells related to altered distributions of specific DLR family members. Thus, DLR1, DLR3, and DLR4, but not DLR2 and DLR5, were significantly more frequent in the non-productive compared to the productive repertoires (DLR1, 2/48 vs. 0/155, $P \leq 0.011$; DLR3, 2/48 vs. 0/155, $P \leq 0.011$; and DLR4, 6/48 vs. 6/155, $P \leq 0.027$). The possibility of selective influences on other D segment members in the $CD5^+$ B cell subset was suggested by the significant difference from predicted in the frequency of the DN family in the productive, but not the nonproductive repertoire. As was seen in the DLR family, two members of the DN family showed a differential distribution. Whereas DN1 appeared to be positively selected, as it was found significantly more often than expected in the productive (14/155, $P \leq 0.002$) but not the nonproductive repertoire

Table III. Most Frequently Employed V_H Genes in $CD5^+/IgM^+$ and $CD5^-/IgM^+$ B Cells

	$CD5^+/IgM^+$		$CD5^-/IgM^+$	
	Productive rearrangements	Nonproductive rearrangements	Productive rearrangements	Nonproductive rearrangements
	n = 144	*n = 31*	*n = 206*	*n = 26*
3-23/DP-47	20 (13.9%)*	1 (3.2%)	25 (12.1%)*	1 (3.8%)
4-59/DP-71	9 (6.3%)*	2 (6.5%)	16 (7.8%)*	4 (15.4%)*
3-30.3/DP-46	12 (8.3%)*	0	10 (4.9%)*	1 (3.8%)
3-30/DP-49	9 (6.3%)*	2 (6.5%)	12 (5.8%)*	0
3-07/DP-54	8 (5.6%)*	3 (9.7%)*	12 (5.8%)*	0
4-39/DP-79	4 (2.8%)	1 (3.2%)	12 (5.8%)*	0
- /DP-58	7 (4.9%)	0	9 (4.4%)	1 (3.8%)
1-18/DP-14	5 (3.5%)	2 (6.5%)	10 (4.9%)*	0
4-34/DP-63	5 (3.5%)	2 (6.5%)	8 (3.9%)	2 (7.7%)
Total	79 (52.8%)	13 (41.9%)	114 (58.7%)‡	9 (34.6%)‡

*Significant difference between expected and observed frequency ($P \leq 0.005$). ‡Significant difference between productive and nonproductive rearrangements ($P \leq 0.05$).

Table IV. CDR3 Composition of CD5⁺/IgM⁺ and CD5⁻/IgM⁺ B Cells Using V_H3-23/DP-47

	B cell	V_H*		J_H‡	D§	J_H‖
CD5⁺	¶ BF2P2-g3C01	YYCAR	EPSGGGS	FDIWG	DIR	J_H3
	BF2P1-g3F04	YYCAK	DNWDYYYDSSH	GWG	21/9	J_H3
	BF2P2-g3H08	YYCAK	AEDLTYYYDRSGWGVHGLL	YYFDYWG	21/9, DIR	J_H4
	¶ BF2P2-g3D05	YYCAT	RGWNY	YFDSWG	ψA1	J_H4
	¶ BF1P-g3C05	YYCAK	APQGYIN	YFHYWG	DIR1	J_H4
	¶ BF1P-g3C03	YYCAK	NLRVTVAGI	FDYWG	–	J_H4
	BF1P-g3B12	YYCAK	KGDRGIAVAGIG	FDYWG	DIR1, N1	J_H4
	¶ BF2P3-g3A07	YYCAR	GAYGRGVLS	FDFWG	A5, DIR1	J_H4
	BF2P1-g3A03	YYCAK	ELPPYYDSSGYP	FDYWG	21/9	J_H4
	BF2P2-g3F11	YYCAK	GGYGDYVV	FDYWG	K4	J_H4
	BF2P1-g3A05	YYCAK	AGGATAK	FDYWG	–	J_H4
	BF2P2-g3G02	YYCAK	VRYSSSWYLF	DYWG	N1	J_H4
	¶ BF2P3-g3B10	YYCAK	VRRWSSGWYGES	DYWG	23/7, XP'1	J_H4
	BF2P3-g3B11	YYCAD	LLTGD	DYWG	HQ52	J_H4
	¶ BF2P3-g3C10	YYCAK	DGRYDPI	DYWG	ψA1	J_H4
	BF2P2-g3B08	YYCAK	VSQILLY	WFDPWG	–	J_H5
	¶ BF2P2-g3F02	YYCAR	DRLVDTG	WFDPWG	K4	J_H5
	BF1P-g3E04	YYCAK	DLGNGSGSHSN	YYYYYYMDVWG	XP'1	J_H6
	¶ BF1P-g3B05	YYCAK	KWGSGD	YYLYMDVWG	HQ52	J_H6
	BF2P2-g3H05	YYCAK	DEGDFIAVAGPYG	MDVWG	N1	J_H6
CD5⁻	¶ BF2N1-g3A12	YYCGK	GTGQDTGGD	AFDIWG	LR3	J_H3
	BF2N2-g3B07	YYCAK	GVLRD	AFDIWG	LR3	J_H3
	BF2N2-g3F07	YYCAK	GVLRH	ALDIWG	–	J_H3
	BF1N-g3D06	YYCAK	SLQQLGPR	YFDYWG	N1	J_H4
	¶ BF2N2-g3B05	YYCAR	MGGITGTTDPI	YFDYWG	M1	J_H4
	¶ BF2N1-g3F08	YYCAK	EGAFSSSISLY	YFDSWG	–	J_H4
	¶ BF1N-g3C12	YYCLQ	WLVEGS	FDYWG	ψA1	J_H4
	¶ BF2N2-g3A02	YYCAK	VHYGRGVEG	FDYWG	–	J_H4
	BF2N2-g3C10	YYCAK	GSSGGSLHHPPK	FDYWG	LR2	J_H4
	BF2N1-g3B03	YYCAK	DRDFIVVVTAV	FDYWG	LR3	J_H4
	¶ BF2N1-g3E08	YYCAR	AFRDTPGNHH	FDYWG	M2	J_H4
	¶ BF2N3-g3B02	YYCAS	AQEGSGYYYL	FDYWG	DIR, 21/9	J_H4
	BF2N1-g3D07	YYCAK	ALWKSSGYGAPEVP	FDYWG	21/9, DIR2	J_H4
	¶ BF2N3-g3C07	YYCAR	PIFYDNAGYYRIGD	FDYWG	21/9	J_H4
	BF2N2-g3B08	YYCAK	DIPSYYDFWSGYYEIRG	FDYWG	XP4, DIR1	J_H4
	BF2N3-g3H01	YYCAK	DGQQLVDLAYLS	DXWG	N1	J_H4
	BF2N2-g3A06	YYCAK	TGGV	DYWG	–	J_H4
	¶ BF2N1-g3D12	YYCAK	GGYGDYYRL	DYWG	K1, A1/4	J_H4
	¶ BF2N2-g3G07	YYCAR	DHYQLRKGF	YWG	N1	J_H4
	BF1N-g3A03	YYCAK	ASGWSAPVG	YWG	K1	J_H4
	¶ BF2N3-g3A05	YYCAK	QVWGGNSSDN	FWG	A5	J_H4
	BF1N-g3B10	YYCAK	GEYE	FDPWG	–	J_H5
	¶ BF2N1-g3H11	YYCAK	FPAVQGAELFKALFNW	FDPWG	–	J_H5
	BF1N-g3A06	YYCAK	DPYSNFYY	YMDVWG	A1/4	J_H6
	¶ BF2N1-g3H05	YYCAK	VKTPYCSNGICYTKPSD	GMDVWG	DIR1, LR1	J_H6

*The last 5 amino acids of the V region according to Kabat et al. (60) are shown. ‡Amino acids of the CDR3 completely encoded by the J_H gene segment and the first two amino acids of FR4 are shown. §D-segments were compiled from Ichiara et al. (47, 61) and Buluwela et al. (62); underlined segments are inverted. ‖Reported by Ravetch et al. (51). ¶V_H gene segments that are ≤ 98% homologous to the respective germline gene 41.

(5/48), DN4 was in the expected range for both the productive as well as the nonproductive repertoire (9/155 and 0/48, respectively).

In the CD5⁻ B cell subset, only two D families differed significantly from the expected frequency. As in the CD5⁺ B cells, the DIR family was found more often in the productive repertoire and, in addition, the DM family was observed less frequently in the productive repertoire (Table V). Of note, DLR4 was also significantly overrepresented in the nonproductive repertoire of CD5⁻ B cells (5/34, $P \leq 0.002$). Furthermore, D21/9, a member of the DXP family, was observed significantly more often than expected in the productive repertoire of both CD5⁺ and CD5⁻ B cell subsets (17/155; $P \leq 0.002$ and 20/200; $P \leq 0.002$, respectively).

Table V. Distribution of D Families in CD5⁺/IgM⁺ and CD5⁻/IgM⁺ B Cells

		CD5⁺/IgM⁺		CD5⁻/IgM⁺	
	Known functional genes per D family*	Productive rearrangements	Nonproductive rearrangements	Productive rearrangements	Nonproductive rearrangements
DA	3 (11.1%)	9 (5.8%)	3 (6.3%)	17 (8.5%)	1 (2.9%)
DHQ52	1 (3.7%)	8 (5.2%)	0	6 (3.0%)	2 (5.9%)
DIR	2 (7.4%)	24 (15.5%)‡	11 (22.9%)‡	35 (17.4%)‡	6 (17.6%)
DK	2 (7.4%)	14 (9.0%)	1 (2.1%)	16 (8.0%)	3 (8.8%)
DLR	5 (18.5%)	13 (8.4%)‡§	13 (27.1%)§	33 (16.4%)	8 (23.5%)
DM	3 (11.1%)	10 (6.5%)	1 (2.1%)	10 (5.0%)‡	1 (2.9%)
DN	2 (7.4%)	23 (14.8%)‡	5 (10.4%)	16 (8.0%)	1 (2.9%)
DψA1	1 (3.7%)	7 (4.5%)	5 (10.4%)	10 (5.0%)	3 (8.8%)
DXP	8 (29.6%)	47 (30.3%)	9 (18.8%)	58 (28.9%)	9 (26.5%)
Total	27	155	48	201	34

*D gene families were compiled from Ichihara et al. (47, 61) and Buluwela et al. (62). ‡Significant difference between expected and observed frequency ($P \leq 0.006$). §Significant difference between productive and nonproductive rearrangements ($P \leq 0.05$).

Distribution of J_H gene segments in CD5⁺ and CD5⁻ B cells. Analysis of the J_H segments indicated that the usage of these components in CD5⁻ B cells did not differ between the productive and the nonproductive repertoires (Table VI). J_H4 was employed most frequently in both the productive and the nonproductive repertoires, followed by J_H6 and J_H5. In contrast, there were significant differences between the distribution of J_H4 and J_H2 in the productive and the nonproductive repertoires of CD5⁺ B cells with J_H4 found more frequently (75/144 vs. 9/30, $P \leq 0.05$) and J_H2 less frequently observed (5/144 vs. 4/30, $P \leq 0.05$) in the productive repertoire. Of note, the distribution of V_H genes in the productively rearranged IgH chain genes containing J_H4 was not different than that found in the entire set of productive V_HDJ_H rearrangements (Table VII).

Composition of the CDR3 in CD5⁺ and CD5⁻ B cells. When the length of the CDR3, the number of N-nucleotides and the exonuclease activity at the J_H-segment were analyzed, no significant differences between the B cell subpopulations were found, but, as noted previously for CD19⁺ B cells (18), there was a significant difference in the CDR3-length of productively and nonproductively rearranged V_H genes in both B cell populations (Table VIII). In addition, the number of N-nucleotides was greater in the nonproductive than in the productive repertoire of the CD5⁻ B cell population ($P \leq 0.016$). The nonproductive V_HDJ_H rearrangements of both CD5⁺ and CD5⁻ B cells were significantly more likely to contain multiple D segments than the productively rearranged genes (Table VIII; CD5⁺ B cells 17/31 vs. 26/144; $P \leq 0.001$; CD5⁻ B cells 9/26 vs. 33/206; $P \leq 0.02$).

Table VI. Distribution of J_H Genes in CD5⁺/IgM⁺ and CD5⁻/IgM⁺ B Cells

	CD5⁺/IgM⁺		CD5⁻/IgM⁺	
J_H-Genes*	Productive rearrangements	Nonproductive rearrangements‡	Productive rearrangements	Nonproductive rearrangements
J_H1	1 (1%)§	1 (3%)	2 (1%)§	1 (4%)
J_H2	5 (4%)§‖	4 (13%)‖	5 (2%)§	0
J_H3	9 (6%)§	2 (7%)	16 (8%)§	2 (8%)
J_H4	75 (52%)§‖	9 (30%)‖	114 (55%)§	13 (50%)§
J_H5	15 (10%)	5 (17%)	20 (10%)§	4 (15%)
J_H6	39 (27%)§	9 (30%)	49 (24%)§	6 (23%)
Total	144	30	206	26

*Reported by Ravetch et al. (51). ‡One sample from the CD5⁺/IgM⁺ B cell subset was excluded, since the J_H gene could not be identified. §Significant difference between expected and observed frequency ($P \leq 0.008$). ‖Significant difference between productive and nonproductive rearrangements ($P < 0.05$).

IgM⁺/CD5⁺ B cells contain significantly fewer hypermutated V_H chain genes compared with the IgM⁺/CD5⁻ B cell population. Hypermutation of productively rearranged V_HDJ_H genes in CD5⁺ and CD5⁻/IgM⁺ B cells was analyzed by assessing the frequency of mutated genes and the number of mutated bases within each V_HDJ_H rearrangement. B cells were subgrouped into unmutated V_H gene segments (99–100% homology to the respective germline gene), mutated sequences sharing 95–98% homology and V_H gene segments that were highly mutated (≤ 94% homology) as previously described (25). CD5⁺ B cells exhibited a significantly lower frequency of mutated and highly mutated genes compared to CD5⁻ B cells (Table IX). This was particularly notable in the older donor. In this donor, 79% of the CD5⁺ B cells were unmutated, whereas less than 50% of CD5⁻ B cells were unmutated ($P \leq 0.001$). Nearly 25% of the CD5⁻ B cells contained highly mutated V_HDJ_H rearrangements, whereas only 5% of the CD5⁺ B cells were mutated to a similar degree. A similar, but less marked pattern, was noted for the younger donor. In this donor, 3/35 (8.6%) of the CD5⁻ B cells contained highly mutated V_HDJ_H rearrangements, whereas 24.6% (42/171) of the CD5⁻ B cell rearrangements of the older donor were highly mutated ($P \leq 0.037$).

Use of the V_H, D and J_H genetic elements in unmutated and mutated rearrangements. To determine whether biases in V_H, D and J_H usage in either population of B cells was limited to

Table VII. Distribution of V_H Families in Productively Rearranged J_H4^+ $CD5^+/IgM^+$ and $CD5^-/IgM^+$ B Cells

	Known functional genes/V_H family*	$CD5^+/IgM^+$		$CD5^-/IgM^+$	
		Total	$J_H4(+)$	Total	$J_H4(+)$
V_H1	11 (21.6%)	27 (18.8%)	7 (9.3%)	27 (13.1%)‡	13 (11.4%)
V_H2	3 (5.9%)	3 (2.1%)	3 (4.0%)	4 (1.9%)	2 (1.8%)
V_H3	22 (43.1%)	81 (56.3%)‡	47 (62.7%)‡	111 (53.9%)‡	64 (56.1%)
V_H4	11 (21.6%)	26 (18.1%)	15 (20.0%)	51 (24.8%)	30 (26.3%)
V_H5	2 (3.9%)	2 (1.4%)	2 (2.7%)	6 (2.9%)	2 (1.8%)
V_H6	1 (2.0%)	2 (1.4%)	0	5 (2.4%)	2 (1.8%)
V_H7	1 (2.0%)	3 (2.1%)	1 (1.3%)	2 (1.0%)	1 (0.9%)
Total	51	144	75	206	114

*The number of V_H genes per family observed as V_HDJ_H recombinants was obtained from the V BASE Sequence Directory (20). ‡Significant difference between expected and observed frequency ($P \leq 0.008$).

the population that contained mutations, the distribution of V_H region gene segments in the unmutated (99–100% homology) and mutated rearrangements (≤ 98% homology) of each B cell subset was compared. The V_H3 family was significantly overrepresented in mutated V_HDJ_H rearrangements of both $CD5^+$ and $CD5^-$ B cells and the V_H1 family underrepresented in the mutated rearrangements of the $CD5^-$ subset (Table X). The distributions of V_H families in the unmutated rearrangements of both subsets were comparable to that found in the nonproductive rearrangements and not different than expected from random usage.

When the distribution of V_H family members was analyzed in IgM^+ B cells with unmutated rearrangements, two of the most commonly used V_H genes, 3-07/DP-54 and -/DP-58, were noted at the expected frequency, whereas the other seven V_H genes were found significantly more often than expected. Of these seven V_H genes, six (3-23/DP-47, 4-59/DP-71, 3-30.3/DP-46, 4-39/DP-79,1-18/DP-14, and 4-34/DP-63) were overrepresented in unmutated rearrangements of $CD5^-$ B cells and three (3-23/DP-47, 4-59/DP-71, and 3-30.3/DP-46) were overrepresented in the unmutated rearrangements of $CD5^+$ B cells. When the distribution of individual V_H genes was compared in mutated versus unmutated rearrangements of $CD5^+$ and $CD5^-$ B cells, significant differences were noted for specific V_H segments. In $CD5^+$ B cells, there was a significantly increased frequency of 3-23/DP-47 in mutated rearrangements (9/31 vs. 11/113, $P \leq 0.006$), whereas in $CD5^-$ B cells there was a significantly increased frequency of 3-07/DP-54 (10/107 vs. 2/99, $P \leq 0.025$) and a significantly decreased frequency of 1-18/DP-14 (2/107 vs. 8/99, $P \leq 0.038$) and 4-34/DP-63 (1/107 vs. 7/99, $P \leq 0.023$) in the mutated rearrangements.

A similar analysis of D and J_H segment use in the mutated and unmutated arrangements of $CD5^+$ and $CD5^-$ B cells was carried out (Table XI). As was noted in the V_H analysis, differences were noted between the frequency of D families in mutated versus unmutated rearrangements, although they were

Table VIII. Composition of the CDR3 of $CD5^+/IgM^+$ and $CD5^-/IgM^+$ B Cells

	Productive rearrangements		Nonproductive rearrangements	
	mean±SE		*mean±SE*	
CDR3 length (nucleotides)				
$CD5^+/IgM^+$ B Cells	41.1±1.1*		52.7±2.8*	
$CD5^-/IgM^+$ B Cells	39.5±0.8*		50.3±3.5*	
Number of N-nucleotides				
$CD5^+/IgM^+$ B Cells	12.7±0.6		15.7±1.6	
$CD5^-IgM^+$ B Cells	13.7±0.6*		19.8±2.2*	
J_H-Exonuclease activity				
(Number of germline nucleotides removed)	Median (range)		Median (range)	
$CD5^+/IgM^+$ B Cells	5.0 (0–21)		5.0 (0–19)	
$CD5^-/IgM^+$ B Cells	5.0 (0–21)		6.5 (0–26)	
No. of D segments used:	$CD5^+/IgM^+$	$CD5^-/IgM^+$	$CD5^+/IgM^+$	$CD5^-/IgM^+$
0	19 (13.2%)	40 (19.4%)	1 (3.2%)	3 (11.5%)
1	99 (68.8%)*	133 (64.6%)	13 (41.9%)*	14 (53.8%)
> 1	26 (18.1%)*	33 (16.0%)*	17 (54.8%)*	9 (34.6%)*
Total	144	206	31	26

*Significant difference between productive and non-productive rearrangements ($P < 0.05$).

Table IX. Proportion of Mutated Productively Rearranged V_H Genes in Individual CD5$^+$/IgM$^+$ and CD5$^-$/IgM$^+$ B Cells

Homology to germline gene	Total		Donor 1 (26 yr)		Donor 2 (45 yr)	
	CD5$^+$/IgM$^+$	CD5$^-$/IgM$^+$	CD5$^+$/IgM$^+$	CD5$^-$/IgM$^+$	CD5$^+$/IgM$^+$	CD5$^-$/IgM$^+$
99–100%	113 (78.5%)*	99 (48.1%)*	25 (78.1%)	22 (62.8%)	88 (78.6%)*	77 (45.0%)*
95–98%	26 (18.1%)*	64 (31.1%)*	7 (21.9%)	10 (28.6%)	19 (17.0%)*	54 (31.6%)*
≤ 94%	5 (3.5%)*	43 (20.9%)*	0	3 (8.6%)‡	5 (4.5%)*	40 (23.4%)*‡
Total	144	206	32	35	112	171

*Significant difference between CD5$^+$ and CD5$^-$ B cells. ($P < 0.05$). ‡Significant difference between donor 1 and donor 2 ($P < 0.05$).

only found in CD5$^-$ B cells. Thus, DψA1 was significantly increased (9/101 vs. 1/100, $P \leq 0.01$) and DXP was significantly decreased (22/101 vs. 36/100, $P \leq 0.026$) in mutated versus nonmutated rearrangements.

Most of the J_H segments were represented in unmutated rearrangements of both B cell subsets at frequencies that differed from expected. J_H4 was found at a significantly increased frequency and J_H6 at a significantly decreased frequency in the mutated rearrangements compared to that observed in unmutated rearrangements of both CD5$^+$ and CD5$^-$ B cells. When J_H6 was analyzed in greater detail (Table XII), it was noted that CDR3s employing this J_H segment were significantly longer than CDR3s containing other J_H segments in the unmutated rearrangements. However, in mutated rearrangements the CDR3s were of comparable length. Most of the difference in the J_H6-containing CDR3 lengths between mutated and unmutated rearrangements could be accounted for by portions of the region encoded by non–J_H nucleotides.

Discussion

The results of this study have provided new information about the shape of the expressed heavy chain repertoire of IgM$^+$ human B cell subsets, identified influences that bias the expressed repertoire, documented the presence of somatically mutated IgM expressing memory B cells in the peripheral B cell repertoire, and demonstrated differences in the frequency of IgM expressing memory B cells in CD5$^+$ and CD5$^-$ subsets. To obtain this information, the V_H repertoire of human IgM$^+$/CD5$^+$ and IgM$^+$/CD5$^-$ B cells was assessed using a technique that analyzed genomic DNA of individual unstimulated B cells. This avoided biases encountered when cDNA is exam-

Table X. Distribution of Unmutated and Mutated V_H Families and Family Members in CD5$^+$/IgM$^+$ and CD5$^-$/IgM$^+$ B Cells

	CD5$^+$/IgM$^+$		CD5$^-$/IgM$^+$	
	Unmutated rearrangements*	Mutated rearrangements	Unmutated rearrangements	Mutated rearrangements
V_H1	24 (21.2%)	3 (9.7%)	20 (20.2%)‡	7 (6.5%)‡§
V_H2	2 (1.8%)	1 (3.2%)	2 (2.0%)	2 (1.9%)
V_H3	59 (52.2%)	22 (71.0%)§	45 (45.5%)‡	66 (61.7%)‡§
V_H4	21 (18.6%)	5 (16.1%)	29 (29.3%)	22 (20.6%)
V_H5	2 (1.8%)	0	3 (3.0%)	3 (2.8%)
V_H6	2 (1.8%)	0	0	5 (4.7%)‡
V_H7	3 (2.7%)	0	0	2 (1.9%)
Total	113	31	99	107
3-23/DP-47	11 (9.7%)‡‖	9 (29.0%)‡‖¶	12 (12.1%)‖	13 (12.1%)‖¶
4-59/DP-71	9 (8.0%)‖	0	8 (8.1%)‖	8 (7.5%)‖
3-30.3/DP-46	8 (7.1%)‖	4 (12.9%)‖	6 (6.1%)‖	4 (3.7%)
3-30/DP-49	6 (5.3%)	3 (9.6%)‖	5 (5.1%)	7 (6.5%)‖
3-07/DP-54	6 (5.3%)	2 (6.5%)	2 (2.0%)‡	10 (9.3%)‡‖
4-39/DP-79	2 (1.8%)¶	2 (6.5%)	9 (9.1%)‖¶	3 (2.8%)
-/DP-58	6 (5.3%)	1 (3.2%)	3 (3.0%)	6 (5.6%)
1-18/DP-14	4 (3.5%)	1 (3.2%)	8 (8.1%)‡‖	2 (1.9%)‡
4-34/DP-63	5 (4.4%)	0	7 (7.1%)‡‖	1 (0.9%)‡
Total	57 (50.4%)‡	22 (70.1%)‡	60 (60.6%)	54 (50.5%)

*Productive rearrangements were grouped into unmutated (99–100% homology to the respective germline gene) and mutated (≤ 98% homology) rearrangements. ‡Significant difference between unmutated and mutated rearrangements ($P < 0.05$). §Significant difference between expected and observed frequency ($P \leq 0.008$). ‖Significant difference between expected and observed frequency ($P \leq 0.005$). ¶Significant differences between CD5$^+$ and CD5$^-$ B cells ($P \leq 0.05$).

Table XI. Distribution of D Families and J_H Genes in Unmutated and Mutated $CD5^+/IgM^+$ and $CD5^-/IgM^+$ B Cells

	$CD5^+/IgM^+$		$CD5^-/IgM^+$	
	Unmutated rearrangements*	Mutated rearrangements	Unmutated rearrangements	Mutated rearrangements
DA	6 (4.8%)	3 (10.3%)	9 (9.0%)	8 (7.9%)
DHQ52	7 (5.6%)	1 (3.4%)	3 (3.0%)	3 (3.0%)
DIR	17 (13.5%)	7 (24.1%)	13 (13.0%)	22 (21.8%)
DK	12 (9.5%)	2 (6.9%)	8 (8.0%)	8 (7.9%)
DLR	12 (9.5%)	1 (3.4%)	15 (15.0%)	18 (17.8%)
DM	9 (7.1%)	1 (3.4%)	4 (4.0%)	6 (5.9%)
DN	21 (13.7%)	2 (6.9%)	11 (11.0%)	5 (5.0%)
DψA1	5 (4.0%)	2 (6.6%)	1 (1.0%)‡	9 (8.9%)‡
DXP	37 (29.4%)	10 (34.5%)	36 (36.0%)‡	22 (21.8%)‡
Total	126	29	100	101
J_H1	0‖	1 (3.2%)	0‖	2 (1.9%)‖
J_H2	4 (3.5%)‖	1 (3.2%)	3 (3.0%)‖	2 (1.9%)‖
J_H3	4 (3.5%)‡‖¶	5 (16.1%)‡	10 (10.1%)¶	6 (5.6%)‖
J_H4	54 (47.8%)‡‖	21 (67.7%)‡‖	46 (46.5%)‡‖	68 (63.6%)‡‖
J_H5	13 (11.5%)	2 (6.5%)	9 (9.1%)	11 (10.2%)
J_H6	38 (33.6%)‡‖	1 (3.2%)‡	31 (31.3%)‡‖	18 (16.8%)‡
Total	113	31	99	107

*Productive rearrangements were grouped into unmutated (99%–100% homology to the respective germline gene) and mutated (≤ 98% homology) rearrangements. ‡Significant difference between unmutated and mutated rearrangements ($P < 0.05$). ‖Significant difference between expected and observed frequency ($P \leq 0.008$). ¶Significant difference between $CD5^+$ and $CD5^-$ B cells ($P \leq 0.05$).

Table XII. CDR3 Length of Unmutated and Mutated $CD5^-/IgM^+$ B Cells

CDR3	Unmutated rearrangements (99–100% homology)	Mutated rearrangements (≤ 98% homology)
	mean nucleotides±SEM	
Entire region		
J_H1-5	38.6±1.3*	37.0±1.1
J_H6	48.4±3.3*‡	40.3±2.7‡
J_H-encoded portion		
J_H1-5	10.5±0.4*	9.5±0.4*
J_H6	20.3±2.4*	16.9±1.3*

*Significant difference between J_H1-5 and J_H6 containing CDR3 ($P \leq 0.0001$). ‡Significant difference between unmutated and mutated rearrangements ($P \leq 0.019$).

ined owing to activation-related differences in immunoglobulin mRNA amount. The importance of this concern was recently emphasized by differences in the repertoire of human peripheral B cells noted when genomic DNA and cDNA were analyzed, with the former, but not the latter correlating with that of pre-B cells (26). Analysis of genomic DNA also made it possible to analyze both the productive and the nonproductive V_HDJ_H rearrangements. Since primary transcripts of nonproductively rearranged genes or genes with nonsense mutations are not spliced into mRNA effectively (27, 28), they are usually underrepresented in cDNA libraries. Analysis of nonproductive rearrangements is essential, however, in understanding the influences that shape the repertoire because these rearrangements do not encode for functional proteins and, therefore, their distribution reflects the molecular events of recombination unaffected by subsequent selective influences dependent on immunoglobulin protein expression.

It is important that the efficiency of this method ranged between 40 and 53% depending on the subpopulation sorted. Analysis of β actin DNA from similarly sorted cells indicated that the cell sorter deposited at least one cell per well in only 60–80% of the wells. This implies that V_H sequences were obtained from more than 70% of B cells. This technique, therefore, should provide an unbiased assessment of the B cell repertoire. Finally, the possibility that some of the biases observed in the current study reflected preferential amplification of certain V_H genes is unlikely, since all but five known V_H genes have been amplified in this study. Furthermore, two of these five genes (3-21/DP-77 and 2-26/DP-26) had previously been amplified from genomic DNA of individual $CD19^+$ B cells using this technique (18). Therefore, this method has detected at least 94% of functional V_H genes, making it unlikely that differential amplification of V_H genes contributed to the biases in V_H usage observed.

It should be noted that the results reported here were obtained from an analysis of two normal volunteers. Although they were racially dissimilar and the findings, in general, were comparable to those obtained from an independent analysis of unselected B cells from a Caucasian female (18), it remains possible that genetic features or individual natural histories of the donors contributed to the results. Analysis of additional donors will be necessary to evaluate this possibility fully.

The V_H repertoire of IgM^+ B cell subsets. As has been noted in several studies using this and other techniques (18, 29, 30), the distribution of V_H families in general reflects their germline complexity. Thus, V_H3, the family with the largest number of functional genes, was found most often. However, the current analysis indicates that V_H3 is not only the most frequently used family, but indeed the V_H3 family is found significantly more often than expected in the productively rearranged repertoire. In contrast, in nonproductive rearrangements, this family was found at the expected frequency, indicating, that selection based upon expression of a μ heavy chain protein influenced its distribution. This is in contrast to our previous report (18), which found the V_H3 family was overrepresented in both the productively and the nonproductively rearranged repertoires. One of the reasons for this discrepancy might be the small number of nonproductive rearrangements ($n = 13$) in the previous analysis of $CD19^+$ B cells. The current analysis of much larger numbers of IgM^+ B cells clearly shows that the V_H3 family is found significantly more frequently than expected in productive, but not nonproductive rearrangements, of both $CD5^+$ and $CD5^-$ subsets. This finding indicates that B cells expressing this V_H segment are positively selected. Of importance, overrepresentation of the V_H3 family in the productive rearranged repertoire was noted only in the population of B cells that contained mutated V_HDJ_H rearrangements. By contrast, V_H3 was found at the expected frequency in B cells with unmutated V_H regions. Since mutations are likely to have developed as a result of antigenic stimulation (31), these data imply that

the V_H3 bias detected resulted from preferential expansion or positive selection of V_H3 expressing B cells by antigen. As the expanded population did not preferentially employ specific D or J_H segments or other specific CDR3 features, the possibility that a B cell superantigen (32), such as staphylococcal protein A that preferentially binds to V_H3 components of Ig (33), may have led to preferential expansion of V_H3 expressing B cells must be considered.

Analysis of the V_H-region repertoire at the level of individual genes segments indicated that the overrepresentation of the V_H3 family largely resulted from the preferential usage of specific individual family members. Five out of the nine most frequently employed V_H germline genes belong to this family and all were found more often in the productive repertoire than expected from random usage. As was noted previously (18), 3-23/DP-47, was the most frequently used individual gene segment. Recently, it has been reported that this gene can be found in multiple copies in the genome (34), leading to the possibility that recombinational bias related to molecular mechanisms and not selection could be responsible for its overrepresentation. This seems unlikely to explain the current data, however, since the frequency of 3-23/DP-47 in the nonproductive repertoire, that is governed by molecular mechanisms, was not different from the frequency expected if there were only one copy of this gene. Therefore, it is more likely that the overrepresentation of the V_H3 family and the 3-23/DP-47 gene in the productive repertoire relates to selective influences after recombination. Since no preference in D or J_H expression was associated with the bias in the V_H3 family in general and 3-23/DP-47 in particular, it is likely that the positive selective influences related to the V_H segment and not to the antigen binding CDR3. Overrepresentation of 3-23/DP-47 was noted in the unmutated as well as the mutated productive rearrangements, suggesting that initial expansion of cells expressing this V_H gene may occur before antigen driven clonal expansion in germinal centers. This conclusion is consistent with the previous suggestion that the 3-23/DP-47 gene may dominate the adult pre–B cell repertoire (26).

Although the numbers are small, the data suggest that there may be a second level of 3-23/DP-47 bias in the $CD5^+$ B cell population, as this V_H gene was significantly overrepresented in mutated compared to unmutated V_HDJ_H rearrangements of this B cell subset. Assuming that mutations found in the V_H sequences were the result of antigen stimulation, this would imply that in the $CD5^+$ B cell subset 3-23/DP-47 is not only positively selected in the bone marrow, but also by subsequent antigen-driven immune responses in the periphery. Since no bias in the D or J_H usage was noted in mutated 3-23/DP-47 containing μ chains of $CD5^+$ B cells, these data imply that superantigen-like stimulation may have been responsible for the overrepresentation of 3-23/DP-47 in the mutated rearrangements of $CD5^+$ B cells. The data are consistent with the conclusion that specific V_H expansions do not occur in parallel in $CD5^+$ and $CD5^-$ B cells.

The V_H4 family was found at a higher frequency in the nonproductive than the productive rearrangements of all IgM^+ B cells (data not shown), although both were within the expected range. Of note, however, when the $CD5^-$ population was analyzed independently, the V_H4 family was found significantly more frequently than expected from random usage in the nonproductive rearrangements, indicating that in this B cell subset the recombination machinery positively influenced rearrangements of this V_H family. A second bias in B cells expressing the V_H4 family was observed in that the frequency of the V_H4 family in the productive rearrangements was significantly diminished compared to that in the nonproductive rearrangements, strongly implying that this family was negatively selected in the expressed repertoire of $CD5^-$ B cells. When the unmutated and mutated productive rearrangements involving the V_H4 family were compared, no significant differences in distribution were found, implying that antigen mediated processes had no additional impact. These findings are consistent with the conclusion that negative selection of the V_H4 family in $CD5^-$ B cells occurs after μ chain expression but before antigen exposure. Since no bias in the D and J_H segments was found in B cells expressing the V_H4 family, the data are most consistent with V_H segment-related negative selection. It is interesting to note that many autoantibodies have been reported to be encoded by members of this family (35, 36) and negative selection might, therefore, be necessary to avoid autoimmunity in normal subjects. This limitation of V_H4 specific negative selection to $CD5^-$ B cells implies that selection affects the B cell subsets differently. In addition, the V_H4 family was employed for rearrangements in the nonproductive repertoire of $CD5^-$ B cells twice as often as in $CD5^+$ B cells (44.0 vs. 20.7%, respectively) and more often in $CD5^-$ B cells than expected from random chance. This suggests that there may be subtle differences in the molecular recombination machinery in the human B cell subsets.

As with the V_H3 family, a small number of V_H4 family members were detected more frequently than expected, including 4-59/DP-71, 4-39/DP-79, and 4-34/DP-63. The same genes had also been found to be overrepresented in the $CD19^+$ B cell population (18), in cDNA libraries from adult human peripheral blood B cells (25) and, with the exception of 4-39/DP-79, in the analysis of a segment of the V_H4 repertoire of a 4-yr-old donor (21). The most common V_H4 family member, 4-59/DP-71, was found significantly more frequently not only in the productive but also in the nonproductive repertoire, indicating that its overrepresentation was most likely the result of molecular rather than selection mechanisms. Nothing about the promotor, flanking genes, location, or copy number of 4-59/DP-71 is known that might explain its preferential rearrangement.

In several reports, 4-34/DP-63 had been claimed to be overrepresented in the expressed repertoire (37, 38). This gene segment is of interest, since it encodes an antibody that recognizes autologous determinants on red blood cells, the I/i antigens (39). Furthermore, it has been shown that a monoclonal anti-idiotype antibody, 9G4, recognizes a conformation-dependent idiotype that is highly associated with the 4-34/DP-63 gene segment (40). In the current study, 4-34/DP-63 was the third most commonly used V_H4 family member, but its frequency in the productive or nonproductive repertoires of peripheral blood B cells was not significantly different than that predicted from random usage. This difference with other reports could be either donor-dependent and/or could be related to the technique used to analyze the repertoire. It is interesting that the analysis of human peripheral blood IgM^+ B cells with the anti-Id 9G4 (38) revealed approximately the same frequency of 4-34/DP-63^+ B cells as was found in this study. In the current analysis, 4-34/DP-63 was found significantly more often than expected only when the unmutated population was analyzed and was consistently underrepresented in the mutated population, im-

plying that it may be less frequently involved in immune responses to exogenous antigens or B cells expressing this V_H gene segment might be deleted during the process of antigen-driven differentiation, as has been suggested (41).

The V_H1 family, which has the same number of functional genes as the V_H4 family, was found significantly less often in the productively rearranged repertoire of $CD5^-$ B cells. In contrast, the $CD5^+$ subset used the V_H1 family at the expected frequency. Overall, the frequency of this family in the peripheral blood was quite stable and ranged from 12.7% in $CD19^+$ (18) to 15.4% in all IgM^+ B cells (data not shown). The observation that the V_H1 family was found at the expected frequency in unmutated, but was significantly decreased in the mutated rearrangements suggests that the V_H1 family is either less often used in responses to antigen or if used might be deleted or more effectively switched to down-stream isotypes and, therefore, not detected in the IgM^+ population. Furthermore, only one member of this family, 1-18/DP-14, was overrepresented in this analysis and, similar to the whole family, was significantly less frequently found in the mutated population. As was previously observed in the analysis of total peripheral $CD19^+$ B cells (18), 1-69/DP-10, which was suggested to be overrepresented in the fetal and the adult repertoire (42, 43), was found at the expected frequency. Recently it has been shown that this V_H gene, which corresponds to the rearranged 51p1 sequence of the fetal repertoire (20), is present in the genome at varying numbers of copies (44). Using an anti-Id antibody, G6, which reacts with V regions encoded by 51p1 (43), the frequency of $G6^+$ IgD^+ tonsillar B cells ranged from zero to 11.4% and correlated with the copy number of 51p1-related genes in different donors (45). This suggests that the differences in the frequency of this gene segment found in different studies may be related to the copy number of the 1-69/DP-10 genetic segment in the donors analyzed. That the current findings are not an artifact of the technique used is supported by an analysis of cDNA libraries from adult human PBL (25), demonstrating that only 1 of 18 clones using a V_H1 family member used 1-69/DP-10.

CDR3 complexity. The CDR3 composition of both IgM^+ B cell subsets was comparable in length, number of N-nucleotides, exonuclease activity at the J_H gene segment and the number of D-segments used. This is in contrast with data obtained from the mouse, in which $CD5^+$ B cells had limited N-nucleotides or exonuclease activity (46). This discrepancy most likely reflects differences in the generation of the CDR3 between mouse and man since data obtained from human cord blood $CD5^+$ and $CD5^-$ B cells were very similar to results obtained in this study (17). Furthermore, in both B cell subpopulations significant differences were found between productive and nonproductive rearrangements in CDR3-length and the number of D-segments per CDR3, as had previously been noted for peripheral $CD19^+$ B cells (18).

It was hypothesized that the increased probability of introducing out of frame rearrangements and/or stop codons with more complex rearrangements might be responsible for these findings (18). However, the current analysis involving many more sequences made this hypothesis unlikely. Although a correlation between CDR3 length and the number of stop codons in nonproductively rearranged V_H genes was found ($r = 0.531$), analysis of the CDR3s of V_HDJ_H rearrangements that contained no stop codons continued to show a significant increase in length in nonproductive compared with productive rearrangements. This finding implies that CDR3 length, independent of the likelihood of introducing stop codons, is the determining factor and suggests the alternative hypothesis that V_HDJ_H rearrangements with longer CDR3 regions are deleted from the expressed repertoire. Whether this might relate to insufficient pairing with surrogate or authentic light chain or altered capacity to bind positively or negatively to selecting ligands is currently unknown.

The distribution of D families and family members in IgM^+ B cell subsets was found to be influenced by selection, inasmuch as several families were either found to be over- or underrepresented or they were significantly different between the productive and nonproductive repertoires. The DIR family, which is characterized by irregular recombination signals (47), was found significantly more often than expected in both the productive and the nonproductive rearrangements, but not at different frequencies in productive and nonproductive repertoires, indicating that recombinational bias related to molecular mechanisms might be responsible. Recent information has indicated that there are three additional DIR elements (DIR3-5) in man (48) than the two originally reported by Ichihara et al. (47). This increase in potential donor segments would alter the frequencies of DIR elements found in the different populations to within the expected range of random usage. In addition, the DIR family was found equally in $CD5^+$ and $CD5^-$ populations with a frequency in unmutated productive rearrangements of $\sim 13\%$. This is in the range that had been postulated for the involvement of this D family in the V_HDJ_H recombination process (49). The frequency of the DIR family in mutated rearrangements was consistently greater than that in unmutated V_HDJ_H rearrangements, suggesting that V_H chains containing a DIR-encoded segment in the CDR3 might be preferentially expanded by antigen.

Evidence for negative selection of the DLR family was found in the $CD5^+$ but not the $CD5^-$ B cell subset. Two stages of selection were noted, with the DLR family being found significantly less often in the productive than the nonproductive rearrangements and additionally decreased in frequency in the mutated V_HDJ_H rearrangements of $CD5^+$ B cells. Since the number of reading frames with stop codons in the DLR family (two or three out of six possible RF) is not different than that in the DA family, for example, that is not underrepresented in the expressed repertoire, the decreased frequency of DLR in the productive rearrangements is unlikely to be related to the use of reading frames with stop codons. Rather the underrepresentation of the DLR family is likely to relate to negative selection both before and after antigenic stimulation, but only within the $CD5^+$ B cell population. The limitation of this selective process to the $CD5^+$ B cell subset strongly suggests that the selective influences that shape the expressed repertoire of $CD5^+$ B cells both before and after antigenic stimulation are different than those that play a role in selecting that of $CD5^-$ B cells.

The distribution of J_H segments in the expressed repertoire of IgM^+ B cell subsets was similar to that reported by Yamada et al. (50), with J_H4 being the most commonly used segment followed by J_H6. The current data document that part of the overrepresentation of J_H4 relates to combinatorial bias, as its frequency is greater than expected in the nonproductive rearrangements of the $CD5^-$ B cell subset. This may reflect the molecular organization of J_H4 that is the only J_H element that possesses the conventional 23-bp spacer length (51). In addi-

tion to the recombinational bias noted in the $CD5^-$ B cells, J_H4 was positively selected in the productive rearrangements of the $CD5^+$ B cell population as well as in the mutated rearrangements of both $CD5^+$ and $CD5^-$ B cells. These results indicate that J_H4 is overrepresented in the $CD5^+$ and $CD5^-$ subsets because of selection and recombinational bias, respectively. In addition, this J_H element is further overrepresented in both populations as a result of antigen-mediated expansion. These latter findings suggest that the antigen-mediated events that result in expansion of J_H4 expressing B cells are comparable in the $CD5^+$ and $CD5^-$ subsets. As opposed to J_H4, the frequency of J_H6 was increased in the unmutated productive rearrangements in both B cell populations and significantly diminished in the mutated V_HDJ_H rearrangements. These results suggest that J_H6 is positively selected in the expressed preimmune repertoire and then not expanded or negatively selected as a result of antigenic stimulation. Of importance, the selective influences on J_H6 were comparable in $CD5^+$ and $CD5^-$ B cells and, therefore, consistent with the conclusion that the selective conditions directed toward J_H segments are comparable in both B cell subsets, whereas the selection of D segments differs. The selective influences on J_H6 may relate to its length, encoding the longest CDR3 of all J_H genes (29 nucleotides). The data obtained from analysis of CDR3 length suggest that longer CDR3s including J_H6 may favor positive selection into the preimmune unmutated repertoire, whereas they may be disadvantageous for expansion by antigen. Analysis of structural characteristics of heavy chains with longer and shorter CDR3s should provide additional information about these possibilities.

Somatic mutation differentially affects human $CD5^+$ and $CD5^-$ B cells. The analysis of genomic DNA from $CD5^+$ and $CD5^-$ B cell subsets revealed no significant differences in the distribution of V_H, D and J_H gene segments, nor in the composition of the CDR3. In the mouse, it had been reported that $CD5^+$ or B-1 B cells resemble the fetal repertoire with respect to the complexity of the CDR3 (46). This difference appears to be related to species-dependent variations in the recombination machinery since the CDR3 profiles of $CD5^+$ and $CD5^-$ B cells observed in the current analysis were comparable and similar to results reported for human cord blood $CD5^+$ and $CD5^-$ B cells (17). Despite these similarities, several differences between the $CD5^+$ and $CD5^-$ B cells became apparent upon more detailed analysis as discussed above for the various influences of selection on the repertoire. A major difference between the $CD5^+$ and the $CD5^-/IgM^+$ B cells was related to the frequency with which V_HDJ_H rearrangements accumulated mutations. Thus, $CD5^+/IgM^+$ B cells were mutated significantly less often than $CD5^-/IgM^+$ B cells. This finding is consistent with the observation that $CD5^+$ B cells home less efficiently to the germinal center (15), where it is thought that hypermutation primarily occurs (52). However, these findings differ from other reports that found no differences in the mutation of V genes of $CD5^+$ and $CD5^-$ human B cells (16, 21, 53). The large sample size of the current study, however, clearly documents that the B cell subsets are different with respect to the frequency of mutations. One previous study supported this conclusion in that the average number and the range of mutations were higher in the $CD5^-$ compared to the $CD5^+$ B cell population (53). Since $CD5^+$ B cells are thought to be primarily involved in immune responses to T-independent antigens and not to participate in germinal center reactions (15), a smaller number of mutated V_HDJ_H rearrangements would be anticipated in this population. Whether the mutations noted in the $CD5^+$ subset are the result of less frequent T-dependent immune responses or the outcome of recurrent T cell independent stimulation outside of germinal centers or even were derived from $CD5^-$ B cells that may have become CD5 positive upon stimulation (54) remains to be determined.

The increase in mutated rearrangements in $CD5^-$ B cells, but not $CD5^+$ B cells in the older donor, was accompanied by an increase in the number of highly mutated rearrangements. The issue of age-related changes in the number of mutated IgM^+ B cells in the peripheral blood is still controversial. Several reports have shown that somatically mutated V region genes accumulate in the peripheral blood during life (25, 55) but since no differences were found in the frequencies of mutations of V_H4 family members of a 4-yr-old (21) and the Vκ3 and Vκ4 genes of a 67-yr-old donor (56) it was assumed that the increase resulted from overrepresentation of mutated transcripts in cDNA libraries. The current data, derived from analysis of genomic DNA of individual B cells, that is not influenced by the activation stage of the B cell, are consistent with the conclusion that there is an increase in IgM memory B cells with age and also indicate that memory B cells can acquire further age-dependent mutations before they switch to downstream isotypes. The latter point is consistent with the conclusion that memory B cells undergo recurrent rounds of somatic mutation (57) and extends this to include the IgM^+ memory B cell pool.

Conclusions. The current data have established differences between the V_H repertoires of adult peripheral blood $CD5^+$ and $CD5^-$ B cells, consistent with the conclusion that they represent different lineages. It should be noted that certain characteristic features of $CD5^+$ murine B cells, such as the limited use of V_H genes and restricted N-nucleotide addition and exonuclease activity were not found in human $CD5^+$ B cells. However, subtle differences in the molecular and selective influences on the $CD5^+$ and $CD5^-$ B cell populations were noted. Most noteworthy was the marked difference in the capacity of $CD5^+$ and $CD5^-$ B cells to acquire somatic mutations. Significant increases in the frequency of mutated V_HDJ_H rearrangements and highly mutated rearrangements were noted in the $CD5^-$ subset, indicating the much greater accumulation of IgM^+ memory B cells in this population. As $CD5^+$ B cells are thought to undergo isotype switching less often than the $CD5^-$ subset (13, 14), the current analysis may have underestimated the difference in the likelihood of the subsets to develop IgM^+ memory B cells. The decreased appearance of mutated IgM^+ cells in the $CD5^+$ B cell population is consistent with the diminished capacity of murine $CD5^+$ B-1 B cells to generate memory (58) and is likely related to their inability to participate in germinal center reactions (15). In this regard, human peripheral blood $CD5^+$ B cells appear to be equivalent to the murine B-1a B cell subpopulation. Since the $CD5^-$ B cell population may also contain B-1b equivalent B cells (59), differences between the two subsets may have been underestimated.

Acknowledgments

We thank Drs. Qin-Chang Cheng and Don McIntire for help in determining the appropriate statistical analyses and Jeff Scholes and Kate Greenway for excellent technical assistance.

This work was supported by National Institutes of Health grant A131229. H.P. Brezinschek is a recipient of Erwin-Schrödinger stipends J0715 and J0929 and T. Dörner is a recipient of a Deutsche Forschungsgemeinschaft grant (Do 491/2-1).

References

1. Kurosawa, Y., and S. Tonegawa. 1982. Organization, structure and assembly of immunoglobulin heavy chain diversity DNA segments. *J. Exp. Med.* 155:201–218.

2. Lewis, S., A. Gifford, and D. Baltimore. 1985. DNA elements are asymmetrically joined during the specific recombination of kappa Ig genes. *Science (Wash. DC).* 228:677–685.

3. Alt, F.W., T.K. Blackwell, and G.D. Yancopoulos. 1987. Development of the primary antibody repertoire. *Science (Wash. DC).* 238:1079–1087.

4. Melchers, F., A. Rolink, U. Grawunder, T.H. Winkler, H. Karasuyama, P. Ghia, and J. Anderson.1995. Positive and negative selection events during B lymphopoiesis. *Curr. Opin. Immunol.* 7:214–227.

5. Schwartz, R.S., and B.D. Stollar. 1994. Heavy-chain directed B-cell maturation: continuous clonal selection beginning at the pre-B cell stage. *Immunol. Today.* 15:27–32.

6. Kirkham, P.M., F. Mortari, J.A. Newton, and H.W. Schroeder, Jr. 1992. Immunoglobulin V_H clan and family identity predicts variable domain structure and may influence antigen binding. *EMBO (Eur. Mol. Biol. Organ.) J.* 11:603–609.

7. Gu, H., D. Tarlinton, W. Müller, K. Rajewsky, and I. Förster. 1991. Most peripheral B cells in mice are ligand selected. *J. Exp. Med.* 173:1357–1371.

8. Nuñez, C., N. Nishimoto, G.L. Gartland, L.G. Billips, P.D. Burrows, H. Kubagawa, and M.D. Cooper. 1996. B cells are generated throughout life in humans. *J. Immunol.* 156:866–872.

9. Schittek, B., and K. Rajewsky. 1990. Maintenance of B-cell memory by long-lived cells generated from proliferating precursors. *Nature (Lond.).* 346: 749–751.

10. Hardy, R.R., and K. Hayakawa. 1994. CD5 B cells, a fetal B cell lineage. *Adv. Immunol.* 55:297–339.

11. Kipps, T.J. 1989. The CD5 B cell. *Adv. Immunol.* 47:117–185.

12. Casali, P., S.E. Burastero, M. Nakamura, G. Inghirami, and A.L. Notkins. 1987. Human lymphocytes making rheumatoid factor and antibody to ssDNA belong to Leu-1+ B-cell subset. *Science (Wash. DC).* 236:77–81.

13. Braun, J., and L. King. 1989. Unique V gene usage by B-Ly1 cell lines, and a discordance between isotype switch commitment and variable region hypermutation. *J. Mol. Cell. Immunol.* 4:121–127.

14. Tarlinton, D.M., M. McLean, and G.J.V. Nossal. 1995. B1 and B2 cells differ in their potential to switch immunoglobulin isotype. *Eur. J. Immunol.* 25: 3388–3393.

15. Stall, A.M., S.M. Wells, and K.-P. Lam. 1996. B-1 cells: unique origins and functions. *Semin. Immunol.* 8:45–59.

16. Ebeling, S.B., M.E.M. Schutte, and T. Logtenberg. 1993. Peripheral human $CD5^+$ and CD^- B cells may express somatically mutated V_H5- and V_H6-encoded IgM receptors. *J. Immunol.* 151:6891–6899.

17. Kiyoi, H., K. Naito, R. Ohno, and T. Naoe. 1995. Comparable profiles of the immunoglobulin heavy chain complementarity determining region (CDR)-3 in $CD5^+$ and $CD5^-$ human cord blood B lymphocytes. *Immunology.* 85:236–240.

18. Brezinschek, H.P., R.I. Brezinschek, and P.E. Lipsky. 1995. Analysis of the heavy chain repertoire of human peripheral B cells using single-cell polymerase chain reaction. *J. Immunol.* 155:190–202.

19. Jelinek, D.F., and P.E. Lipsky. 1987. Comparative activation requirements of human peripheral blood, spleen, and lymph node B cell. *J. Immunol.* 139:1005–1013.

20. Tomlinson, I.M., S.C. Williams, S.J. Corbett, J.B.L. Cox, and G. Winter. 1996. V BASE sequence directory. MRC Centre for Protein Engineering, Cambridge, UK.

21. Klein, U., R. Küppers, and K. Rajewsky. 1994. Variable region gene analysis of B cell subsets derived from a 4-year-old child: somatically mutated memory B cells accumulate in the peripheral blood already at young age. *J. Exp. Med.* 180:1383–1393.

22. Chu,Y.-W., E. Marin, R. Fuleihan, N. Ramesh, F.S. Rosen, R.S. Geha, R.A. Insel. 1995. Somatic mutation of human V genes in the X-linked hyper-IgM syndrome. *J. Clin. Invest.* 95:1389–1393.

23. Jacob, J., J. Przylepa, C. Miller, and G. Kelsoe. 1993. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl) acetyl. III. The kinetics of V region mutation and selection in germinal center B cells. *J. Exp. Med.* 178:1293–1307.

24. Woolson, R.F. 1987. Statistical Methods for the Analysis of Biomedical Data. John Wiley & Sons, New York. pp. 337.

25. Huang, C., and B.D. Stollar. 1993. A majority of Ig H chain cDNA of normal human adult blood lymphocytes resembles cDNA for fetal Ig and natural autoantibodies. *J. Immunol.* 151:5290–5300.

26. Milili, M., C. Schiff, M. Fougereau, and C. Tonnelle. 1996. The VDJ repertoire expressed in human preB cells reflect the selection of bona fide heavy chains. *Eur. J. Immunol.* 26:63–69.

27. Qian, L., M.N. Vu, M.S. Carter, J. Doskow, and M.F. Wilkinson. 1993. T cell receptor-β mRNA splicing during thymic maturation in vivo and in an inducible T cell clone in vitro. *J. Immunol.* 151:6801–6814.

28. Lozano, F., B. Maertzdorf, R. Pannell, and C. Milstein. 1994. Low cytoplasmic mRNA levels of immunoglobulin κ light chain genes containing nonsense codons correlate with inefficient splicing. *EMBO (Eur. Mol. Biol. Organ.) J.* 13:4617–4622.

29. Guigou, V., A.-M. Cuisinier, C. Tonnelle, D. Moinier, M. Fougereau, and F. Fumoux. 1990. Human immunoglobulin V_H and $V_κ$ repertoire revealed by in situ hybridization. *Mol. Immunol.* 27:935–940.

30. Logtenberg, T., M.E.M. Schutte, S.B. Ebeling, F.H.J. Gmelig-Meyling, and J.H. van Es. 1992. Molecular approaches to the study of human B-cell and (auto)antibody repertoire generation and selection. *Immunol. Rev.* 128:23–47.

31. Neuberger, M.S., and C. Milstein. 1995. Somatic hypermutation. *Curr. Opin. Immunol.* 7:248–254.

32. Zouali, M. 1995. B cell superantigens: implications for selection of the human antibody repertoire. *Immunol. Today.* 16:399–405.

33. Vasquez-Kristiansen, S., V. Pascual, and P.E. Lipsky. 1994. Staphylococcal protein A induces biased production of Ig by V_H3-expressing B lymphocytes. *J. Immunol.* 153:2974–2982.

34. Sasso, E.H., J.H. Buckner, and L.A. Suzuki. 1995. Ethnic differences in polymorphism of an immunoglobulin V_H3 gene. *J. Clin. Invest.* 96:1591–1600.

35. Pascual, V., and J.D. Capra. 1991. Human immunoglobulin heavy-chain variable region genes: organization, polymorphism, and expression. *Adv. Immunol.* 49:1–74.

36. Stewart, A.K., C. Huang, A.A. Long, B.D. Stollar, and R.S. Schwartz. 1992. V_H-gene representation in autoantibodies reflects the normal human B cell repertoire. *Immunol. Rev.* 128:101–122.

37. Suzuki, I., L. Pfister, A. Glas, C. Nottenburg, and E.C.B. Milner. 1995. Representation of rearranged V_H gene segments in the human adult antibody repertoire. *J. Immunol.* 154:3902–3911.

38. Kraj, P., D.F. Friedman, F. Stevenson, and L.E. Silberstein. 1995. Evidence for the overexpression of the *VH4–34* (*VH4.21*) Ig gene segment in the normal adult human peripheral blood B cell repertoire. *J. Immunol.* 154:6406–6420.

39. Silberstein, L.E., L.C. Jefferies, J. Goldman, D.F. Friedman, J.S. Moore, P.C. Nowell, D. Roelcke, W. Pruzanski, J. Roudier, and G.J. Silverman. 1991. Variable region gene analysis of pathologic human autoantibodies to the related I and I red blood cell antigens. *Blood.* 78:2372–2386.

40. Potter, K.N., Y. Li, V. Pasqual, R.C. Williams, Jr., L.C. Byres, M. Spellerberg, F.K. Stevenson, and J.D. Capra. 1993. Molecular characterization of the cross-reactive idiotope on human immunoglobulins utilizing the V_H4-21 gene segment. *J. Exp. Med.* 178:1419–1428.

41. Rettig, M.B., R.A. Vescio, J. Cao, C.H. Wu, J.C. Lee, E. Han, M. Der-Danielian, R. Newman, C. Hong, A.K. Lichtenstein, et al. 1996. V_H gene usage in multiple myeloma: complete absence of the $V_H4.21$ (V_H4-34) gene. *Blood.* 87:2846–2852.

42. Schroeder, H., Jr., J. Hillson, and R. Perlmutter. 1987. Early restriction of the human antibody repertoire. *Science (Wash. DC).* 238:791–793.

43. Kipps, T.J., and S.F. Duffy. 1991. Relationship of the CD5 B cell to human tonsillar lymphocytes that express autoantibody-associated cross-reactive idiotypes. *J. Clin. Invest.* 87:2087–2096.

44. Sasso, E.H., K.W. van Dijk, A. Bull, and E.C.B. Milner. 1990. A fetally expressed immunoglobulin V_H1 gene belongs to a complex set of alleles. *J. Clin. Invest.* 91:2358–2367.

45. Sasso, E.H., T. Johnson, and T.J. Kipps. 1996. Expression of the immunoglobulin V_H gene 51p1 is proportional to its germline gene copy number. *J. Clin. Invest.* 97:2074–2080.

46. Tornberg, U.-C., and D. Holmberg. 1995. B-1a, B-1b and B-2 B cells display unique V_HDJ_H repertoires formed at different stages of ontogeny and under different selection pressures. *EMBO (Eur. Mol. Biol. Organ.) J.* 14:1680–1689.

47. Ichihara, Y., H. Matsuoka, and Y. Kurosawa. 1988. Organization of human immunoglobulin heavy chain diversity gene loci. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:4141–4150.

48. Sanz, I., S.-S. Wang, G. Meneses, and M. Fischbach. 1994. Molecular characterization of human Ig heavy chain DIR genes. *J. Immunol.* 152:3958–3969.

49. Sanz, I. 1991. Multiple mechanisms participate in the generation of diversity of human heavy chain CDR3 regions. *J. Immunol.* 147:1720–1729.

50. Yamada, M., R. Wasserman, B.A. Reichard, S. Shane, A.J. Caton, and G. Rovera. 1991. Preferential utilization of specific immunoglobulin heavy chain diversity and joining segments in adult human peripheral blood B lymphocytes. *J. Exp. Med.* 173:395–407.

51. Ravetch, J.V., U. Siebenlist, S. Korsmeyer, T. Waldmann, and P. Leder. 1981. Structure of the human immunoglobulin μ locus: characterization of embryonic and rearranged J and D genes. *Cell.* 27:583–591.

52. Kelsoe, G. 1996. Life and death in germinal centers (redux). *Immunity.* 4:107–111.

53. Ebeling, S.B., M.E.M. Schutte, and T. Logtenberg. 1993. The majority

of human tonsillar $CD5^+$ B cells express somatically mutated $V_\kappa 4$ genes. *Eur. J. Immunol.* 23:1405–1408.

54. Vernino, L.A., D.S. Pisetsky, and P.E. Lipsky. 1992. Analysis of the expression of CD5 by human B cells and correlation with functional activity. *Cell. Immunol.* 139:185–197.

55. van Es, J.H., F.H.J. Gmelig Meyling, and T. Logtenberg. 1992. High frequency of somatically mutated IgM molecules in the human adult blood B cell repertoire. *Eur. J. Immunol.* 22:2761–2764.

56. Klein, U., R. Küppers, and K. Rajewsky. 1993. Human IgM^+IgD^+ B cells, the major B cell subset in the peripheral blood, express V_κ genes with no or little somatic mutation throughout life. *Eur. J. Immunol.* 23:3272–3277.

57. Decker, D.J., P.-J. Linton, S. Zaharevitz, M. Biery, T.R. Gingeras, and N.R. Klinman. 1995. Defining subsets of naive and memory B cells based on the ability of their progeny to somatically mutate in vitro. *Immunity.* 2:195–203.

58. Klinman, N.R. 1996. In vitro analysis of the generation and propagation of memory B cells. *Immunol. Rev.* 150:91–111.

59. Kasaian, M.T., H. Ikematsu, and P. Casali. 1992. Identification and analysis of a novel human surface $CD5^-$ B lymphocyte subset producing natural antibodies. *J. Immunol.* 148:2690–2702.

60. Kabat, E.A., T.T. Wu, H.M. Perry, K.S. Gottesman, and C. Foeller. 1991. Sequences of Proteins of Immunological Interest. U.S. Department of Health and Human Services, Washington, DC.

61. Ichihara, Y., M. Abe, H. Yasui, H. Matsuoka, and Y. Kurosawa. 1988. At least five D_H genes of human immunoglobulin heavy chains are encoded in 9-kilobase DNA fragments. *Eur. J. Immunol.* 18:649–652.

62. Buluwela, L., D.G. Albertson, P. Sherrington, P.H. Rabbitts, N. Spurr, and T.H. Rabbitts. 1988. The use of chromosomal translocations to study human immunoglobulin gene organization: mapping D_H segments within 35 kb of the C μ gene and identification of a new D_H locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:2003–2010.