REVIEW ARTICLE

Immunoglobulin V Regions and the B Cell

By A. Keith Stewart and Robert S. Schwartz

INVESTIGATIONS of normal and neoplastic B cells are being channeled into new directions by recent work on Ig variable region (V) genes. Fresh impetus to these studies comes from an informative directory of human germline V genes and an improved understanding of how B cells deploy those genes. But substantial questions about the molecular biology of V genes and B-cell maturation remain because of the unexpected finding that a limited group of Ig heavy chain (VH) genes predominates in the preimmune B-cell population. This peculiarity of the VH gene portfolio may hinge on unique properties of certain VH genes that predispose them to frequent rearrangement. It could equally reflect clonal selection in the bone marrow (BM), a process in which clones of immature B cells are selectively deleted (negative selection; clonal deletion) or singled out for preferential growth (positive selection). Emerging evidence suggests that the heavy chain of the B-cell antigen receptor engenders signals that mediate positive and negative selection. In this review we will discuss how the properties of surface Iggs could result in clonal deletion or allow individual B cells to continue their development in the BM.

Limitations on the V gene repertoire may seem to narrow the potential for antibody diversity, but the process of assembling heavy and light chain genes introduces enormous heterogeneity into the third complementarity determining regions (CDR3s) of lgs. For reasons that will become apparent below, variation in CDR3s is especially broad in heavy chains. Moreover, antigen-driven somatic mutations introduce another layer of permutations in antibody structures during the immune response. These diversifying mechanisms have interesting clinical applications. Our understanding of the clonal origins of malignant B cells has been improved by analysis of V gene structures in B-cell lymphomas and leukemias. Detection of very small populations of malignant clones within a larger population of normal B cells is now possible with the cell-specific markers furnished by the unique V gene sequences of the CDR3. In this report we will discuss the clinical utility of the genetic fingerprints of malignant B cells and comment on the limitations of the methods by which they are detected.

Ig GENE REARRANGEMENT

The mechanisms of Ig gene rearrangement have been reviewed in detail elsewhere.4,4 The coding instructions for the antibody V region arise from co-ordinated mechanisms of V gene rearrangement, joining and somatic modification.4,4 Three distinct gene segments—the variable (Vh), diversity (D), and joining region (Jh) genes—encode the heavy chain of the variable region,6,10 whereas two segments—variable (Vk or Vλ) and joining (Jk or JA) region genes—encode the light chains.11-14 The Ig heavy chain (IgH) locus, on chromosome 14, contains an estimated 100 to 150 VH genes, 30 D gene segments, and 6 Jh gene segments.15,16 In a recent census of VH genes compiled from a single individual and from the reported literature, 122 germline VH genes were identified, of which 83 are potentially functional.21,22 As some of these genes are nearly identical, and may represent allelic variants, it is likely that 60 to 70 VH genes are available for rearrangement. These 60 to 70 genes belong to seven families (VH1-VH7) whose members have greater than 80% sequence homology.23 The presence of conserved nucleotide sequences in each family suggests evolutionary pressure to retain certain V gene structures.24,25 Less is known about the germline composition of light chain genes, although analysis of Southern blots suggests that Vκ and Vλ contain approximately 80 and 70 gene segments on chromosomes 2 and 22, respectively.11,12 Recent compilations of sequence data identified only 75 known Vκ sequences, of which only 36 are potentially functional19 and 36 known Vλ sequences of which 24 were potentially functional.4 The database of germline light chain sequences remains incomplete.

Rearrangement of V gene segments6,10 is dependent on the protein products of the recombinase activating genes RAG-1 and RAG-2.1,2 Although these genes are indispensable participants in V gene rearrangement,1,4 the biochemical function of their gene products remains unknown.26 Nevertheless, as all Ig V gene, and indeed T-cell receptor rearrangements, appear to use a common recombinase,27 it follows that targeting of an individual gene for rearrangement may be modulated by the accessibility or exposure of the unrearranged gene to this recombinase.1,2 Transcription of the Ig heavy chain locus is activated in pre-B cells and all rearranging loci in these cells are transcriptionally active at the time of their rearrangement.28,29 Consequently, it has been proposed that transcriptional activation by lymphoid specific enhancers and Ig promoter regions increases accessibility of V genes to recombinase.30,31 Hypomethylation of V genes appears characteristic of this transcriptionally active and thus accessible state.32,33 The importance of transcriptional activation has been confirmed in vitro; activation of germline gene transcription increases the frequency of V gene rearrangement1 whereas deletion of an Ig-specific enhancer (Eμ) dramatically inhibits germline V gene demethylation, transcription, and rearrangement.34,35

Recombination of V genes starts in lymphoid progenitor (both T and B) cells,1 within the IgH locus of either the

From the Division of Hematology-Oncology, The Toronto Hospital, General Division, Toronto, Ontario, Canada; and the Division of Hematology-Oncology, New England Medical Center, Boston, MA.

Submitted September 24, 1993; accepted December 27, 1993.

Supported by grants from The Canadian Red Cross/Cutter, The Medical Research Council of Canada, and National Institutes of Health Grant No. AI28899.

Address reprint requests to A. Keith Stewart, MD, Mulock Larkin Wing, Room 2-024, Toronto Hospital-General Division, 200 Elizabeth St, Toronto, Ontario, MSG 2C4 Canada.

© 1994 by The American Society of Hematology.

0006-4971/94/307-0037$3.00/0

Recombinase proteins must recognize and bind to conserved heptamer-spacer-nonamer recombination recognition sequences flanking all \( V \) gene segments. In B-lineage cells, rearrangement of one D gene to one \( J_H \) gene is followed by the addition of one of the numerous \( V_H \) genes to the fused D-J\( _H \) segment. If successful, the resulting \( V_H\)-D-J\( _H \) rearrangement is transcribed as a single unit into RNA that is spliced to the constant region RNA (C\( _\mu \)) before translation into an \( \mu \) heavy chain in the pre-B cell. If the initial rearrangement yields a sequence that cannot be translated (nonproductive or abortive rearrangement) then rearrangement of the IgH locus proceeds on the other allele.

The presence on the B-cell surface of a fully assembled \( \mu \) heavy chain inhibits further rearrangements that might occur on the opposite allele (allelic exclusion). Thus, each pre-B cell expresses only one kind of heavy chain. However, if malignant transformation arrests B-cell development before expression of the heavy chain on the cell surface, while RAG-1 and RAG-2 are transcriptionally active, a new heavy chain rearrangement may occur. This phenomenon has been inferred from the ongoing VH gene replacement constant region RNA (C\( _\mu \)) before translation into an Ig heavy (\( \mu \)) chain in the pre-B cell. If the initial rearrangement yields a sequence that cannot be translated (nonproductive or abortive rearrangement) then rearrangement of the IgH locus proceeds on the other allele.

After successful heavy chain gene rearrangement in pre-B cells, cytoplasmic \( \mu \) chain attaches to the endoplasmic reticulum predominantly in association with the Ig binding protein BiP. Binding of newly formed \( \mu \) chain to surrogate light chain proteins (\( \psi \) light chain), results in the displacement of small amounts of \( \mu \) from the endoplasmic reticulum and, in some late pre-B cells, allows translocation of the resulting complex to the cell surface. In the progeny of single transformed precursor B cells, RAG-1 and RAG-2 expression is most prominent in early B cells and is limited or absent in later stages of development. Thus, the ability to undergo Ig gene rearrangement decreases as B-cell maturation proceeds.

Following successful heavy chain gene rearrangement in pre-B cells, a fully assembled Ig \( \mu \) chain inhibits further rearrangements that might occur on the opposite allele. Thus, each pre-B cell expresses only one kind of heavy chain. However, if malignant transformation arrests B-cell development before expression of the heavy chain on the cell surface, while RAG-1 and RAG-2 are transcriptionally active, a new heavy chain rearrangement may occur. This phenomenon has been inferred from the ongoing VH gene replacement constant region RNA (C\( _\mu \)) before translation into an Ig heavy chain in the pre-B cell. If the initial rearrangement yields a sequence that cannot be translated (nonproductive or abortive rearrangement) then rearrangement of the IgH locus proceeds on the other allele.

The presence on the B-cell surface of a fully assembled \( \mu \) heavy chain inhibits further rearrangements that might occur on the opposite allele (allelic exclusion). Thus, each pre-B cell expresses only one kind of heavy chain. However, if malignant transformation arrests B-cell development before expression of the heavy chain on the cell surface, while RAG-1 and RAG-2 are transcriptionally active, a new heavy chain rearrangement may occur. This phenomenon has been inferred from the ongoing VH gene replacement constant region RNA (C\( _\mu \)) before translation into an Ig heavy (\( \mu \)) chain in the pre-B cell. If the initial rearrangement yields a sequence that cannot be translated (nonproductive or abortive rearrangement) then rearrangement of the IgH locus proceeds on the other allele.

After successful heavy chain gene rearrangement in pre-B cells, cytoplasmic \( \mu \) chain attaches to the endoplasmic reticulum predominantly in association with the Ig binding protein BiP. Binding of newly formed \( \mu \) chain to surrogate light chain proteins (\( \psi \) light chain), results in the displacement of small amounts of \( \mu \) from the endoplasmic reticulum and, in some late pre-B cells, allows translocation of the resulting complex to the cell surface. In the progeny of single transformed precursor B cells, RAG-1 and RAG-2 expression is most prominent in early B cells and is limited or absent in later stages of development. Thus, the ability to undergo Ig gene rearrangement decreases as B-cell maturation proceeds.

After successful heavy chain gene rearrangement in pre-B cells, cytoplasmic \( \mu \) chain attaches to the endoplasmic reticulum predominantly in association with the Ig binding protein BiP. Binding of newly formed \( \mu \) chain to surrogate light chain proteins (\( \psi \) light chain), results in the displacement of small amounts of \( \mu \) from the endoplasmic reticulum and, in some late pre-B cells, allows translocation of the resulting complex to the cell surface. In the progeny of single transformed precursor B cells, RAG-1 and RAG-2 expression is most prominent in early B cells and is limited or absent in later stages of development. Thus, the ability to undergo Ig gene rearrangement decreases as B-cell maturation proceeds.

After successful heavy chain gene rearrangement in pre-B cells, cytoplasmic \( \mu \) chain attaches to the endoplasmic reticulum predominantly in association with the Ig binding protein BiP. Binding of newly formed \( \mu \) chain to surrogate light chain proteins (\( \psi \) light chain), results in the displacement of small amounts of \( \mu \) from the endoplasmic reticulum and, in some late pre-B cells, allows translocation of the resulting complex to the cell surface. In the progeny of single transformed precursor B cells, RAG-1 and RAG-2 expression is most prominent in early B cells and is limited or absent in later stages of development. Thus, the ability to undergo Ig gene rearrangement decreases as B-cell maturation proceeds.

The presence on the B-cell surface of a fully assembled \( \mu \) heavy chain inhibits further rearrangements that might occur on the opposite allele (allelic exclusion). Thus, each pre-B cell expresses only one kind of heavy chain. However, if malignant transformation arrests B-cell development before expression of the heavy chain on the cell surface, while RAG-1 and RAG-2 are transcriptionally active, a new heavy chain rearrangement may occur. This phenomenon has been inferred from the ongoing VH gene replacement constant region RNA (C\( _\mu \)) before translation into an Ig heavy (\( \mu \)) chain in the pre-B cell. If the initial rearrangement yields a sequence that cannot be translated (nonproductive or abortive rearrangement) then rearrangement of the IgH locus proceeds on the other allele.

The presence on the B-cell surface of a fully assembled \( \mu \) heavy chain inhibits further rearrangements that might occur on the opposite allele (allelic exclusion). Thus, each pre-B cell expresses only one kind of heavy chain. However, if malignant transformation arrests B-cell development before expression of the heavy chain on the cell surface, while RAG-1 and RAG-2 are transcriptionally active, a new heavy chain rearrangement may occur. This phenomenon has been inferred from the ongoing VH gene replacement constant region RNA (C\( _\mu \)) before translation into an Ig heavy (\( \mu \)) chain in the pre-B cell. If the initial rearrangement yields a sequence that cannot be translated (nonproductive or abortive rearrangement) then rearrangement of the IgH locus proceeds on the other allele.
D), and transcribed in any of three potential reading frames depending on the V_H-D junctional sequence. At junctional sites in which no nucleotide deletions occur, there may be templated “p” nucleotides. These sequences, usually only one or two bases long, are complementary copies of the last nucleotides of the coding region they abut. They may be used to repair asymmetrical breaks in the hairpin ends of V genes undergoing recombination. Exceptionally long stretches of “p” nucleotides have been found in thymocytes from mice with SCID, a finding that implicates abnormal DNA repair in the pathogenesis of this disease.

All of these variables contribute to the CDR3 antibody “fingerprint” even before the somatic mutations that further distinguish individual V genes. They result in a CDR3 sequence that is unique to each rearrangement, and that therefore identifies individual B cells or clonal B-cell expansions. Indeed, it has been calculated that only 5 in 100,000 circulating B cells share the same CDR3.

**ANTIGEN SELECTION OF B CELLS**

Although rearrangement stops after the loss of recombinase activity in the BM, further changes in V genes occur during the immune response. The adaptive response to an immunogen occurs in the germinal centers of lymphoid follicles. B cells whose antigen receptors bind to the immunogen react by proliferating. By thus acquiring a growth advantage, they become the targets of clonal selection. Under the influence of T-helper cells, the clonal progeny of these antigen-stimulated B cells mobilize a mechanism that causes somatic mutations (nucleotide substitutions) of rearranged V genes. The mutations occur only in the coding regions of the Ig genes; they may be silent or they may result in amino acid replacements that change the affinity of the antibody’s binding site for the antigen. If the mutation causes a loss of the receptor’s affinity for the antigen, the B cell loses its growth advantage. If the mutation causes a gain in the receptor’s affinity, the B cell gains a proliferative advantage. Because mutations in CDRs are the most likely to affect the antibody’s binding site, a hallmark of an antigen-stimulated B cell is the clustering of replacement mutations in the CDRs. Thus, the B-cell response to antigen is a microcosm of Darwinian evolution in which the environment (antigen) and V gene mutation collaborate to select B-cell clones whose surface Igs best fit the antigenic challenge. The resulting increase in antibody affinity for antigen is the hallmark of a maturing antibody response.

Because V gene mutations are unique to each B cell and its progeny, and because they occur only in B cells that have engaged in an immune response, the V gene nucleotide sequence can show the biography of a B-cell clone. As will
be seen below, the presence or absence of V gene mutations in B-cell malignancies can pinpoint the developmental stage of a neoplastic cell (Fig 2).

Besides somatic mutation, other molecular signs of antigen selection include the repetitive use of particular V genes to encode antibodies of a given specificity and the IgM to IgG isotype switch. Up to 25% of circulating B cells with surface IgM (C\(\mu\)) may show molecular evidence of a previous antigen contact,\,87,\,88 but somatic mutation predominates in the C\(\gamma\),\,89-\,92 C\(\alpha\),\,93 and C\(\epsilon\)\,94 populations of B cells. A relevant example of these principles is a study of 14 V genes used by anti-Rh blood group antibodies.\,95 Both low-affinity IgM and high-affinity IgG anti-Rh antibodies were shown to originate from a limited pool of V genes. The four IgM antibodies were encoded by two minimally mutated V\(\mu\)\(h\) genes, V\(\mu\)A-21 and hv3019b9 (a V\(\mu\)3 family member). Of the IgG antibodies, 6/10 were somatically mutated variants of those same two genes. High affinity in these anti-Rh antibodies correlated with frequent replacement mutations. The J\(\mu\)6 gene (rearranged in 18% of normal human circulating B cells\,97) was used by 79% (11/14) of the antibodies. This very high frequency may reflect structural requirements for binding to the Rh antigen, as 9/11 CDR3 segments encoded by J\(\mu\)6 resulted in the protein sequence tyr-x-met-glu-val. In another study, marked variations in the representation of individual V genes in populations of naive C\(\mu\) and experienced CyB cells were found over time. These population shifts presumably resulted from accumulated antigenic experience; they suggest that antigen selection has a major role in shaping the V gene repertoire in humans.\,96

Ig V GENE REPertoire

The B-cell repertoire may be divided into a naive (preimmune) repertoire, consisting of virginal B cells that have not been exposed to external antigen, and the immune repertoire of mature B cells in peripheral blood and lymphatic tissues. If each functional germline gene segment has an equal chance of undergoing rearrangement to form a V\(\mu\)-D-J\(\mu\) transcription unit, then the V gene repertoire (ie, the distribution of V segments in the population of B-cell clones) of naive B cells should contain all germline V genes in equal frequency. For instance, if there are 83 potentially functional germline V\(\mu\) gene segments,\,2 the frequency of each V\(\mu\) gene in a representative population of preimmune B cells should be 1.2%, and all functional V\(\mu\) genes should be present in rearranged form. However, the actual finding is a pronounced over-representation of certain V genes.\,96,\,97

The naive repertoire has been studied in B cells from fetal lymphoid tissue\,98-\,101 or cord blood\,93,\,102 and in pre-B cells of ALL,\,103,\,104 Normal BM pre-B cells have been studied in mice\,105 but not yet in humans. C\(\mu\) B cells from each of these sources probably rearrange their V genes independently of
the influence of external antigen, an assumption supported by the absence of somatic mutation in V gene coding sequences derived from the Cµ cDNA of these cells.

A consistent pattern of dominance of certain individual \( V_H \), D, and Jµ genes has emerged from these studies. In the naïve V gene repertoire, the use of Jµ genes is biased. The Jµ4 gene segment, which is one of the six potentially available Jµ genes, occurs not in the predicted frequency of 17% (1/6) but in 32% of rearranged Ig cDNA clones from fetal liver,\(^{98,99}\) in 42% of pre-B cell ALL clones,\(^{100,101}\) and in 50% of cord blood cDNA clones.\(^{93,102}\) There is also bias in the use of D gene segments; \( H_Q25 \), is found in 45% of fetal liver B cells and \( D_m, \) with an expected prevalence of 3.3%, was actually found in 15% of pre-B ALL cells. A small group of individual germline \( V_H \) genes, to be mentioned presently, also predominate in fetal liver, cord blood, and in pre-B cell ALL. Studies in the mouse suggest a similar V gene restriction in normal pre-B cells.\(^{105}\)

The small set of recurrent V gene segments that fetal liver B cells rearrange are also ubiquitous in unselected populations of adult peripheral blood B cells,\(^{90,106,107}\) B cells that produce specific antibodies against bacterial or viral pathogens,\(^{99,101,108-111}\) and B cells that produce autoantibodies.\(^{112}\) The \( V_H26 \) gene (also called \( 30P1, 18/2 \), and 3-23),\(^{113}\) one of at least 29 potentially functional germline \( V_H3 \) family members,\(^{21}\) illustrates the point. This \( V_H \) gene was found to be expressed in 24% of fetal B-cell cDNA clones,\(^{98-101}\) 15% of cord blood cDNA clones,\(^{93,102}\) and 6% to 10% of unselected adult peripheral blood B cells.\(^{107}\) It also shows up in 8% of published sequences of autoantibodies with a wide range of specificities,\(^{96}\) including antibodies against platelets, cardiolipin, myelin, and thyroglobulin.\(^{114-118}\) At last count, 23% of published anti-DNA antibody sequences were encoded by this one \( V_H \) gene.\(^{96,114-115}\) Furthermore, \( V_H26 \) also encodes antipathogen antibodies, including antibodies to rai- bies,\(^{91}\) herpes simplex,\(^{108}\) hemophilus B,\(^{109,110}\) and human immunodeficiency virus (HIV),\(^{111}\) often in a somatically mutated form. Thus, the predominance of \( V_H26 \) extends from the fetal liver B-cell to adult germinal center B cells engaged in the secondary immune response. \( V_H \) restriction is not confined to \( V_H26 \). On the contrary, a detailed analysis of the expressed \( V_H \) gene repertoire of two normal subjects showed that one third of over 100 \( C_H \) cDNA clones used only 6 of the 83 available \( V_H \) genes. Moreover, only 23 \( V_H \) genes encoded 90% of all the \( C_H \) clones in unselected adult peripheral blood B cells.\(^{90}\) \( V_H \) gene restriction is not unique to heavy chain genes; for example, a \( k \) light chain gene, \( \text{human}k525 \), which can be identified by the idiotype it encodes, has been shown in 4% of tonsillar B cells,\(^{119}\) in 25% of light-chain bearing B cells in chronic lymphocytic leukemia (CLL),\(^{120,121}\) and frequently in IgM paraproteins with rheumatoid factor activity.\(^{122,123}\)

**FORMATION OF THE V GENE REPERTOIRE**

How can bias in the rearranged V gene repertoire be explained? There are two main possibilities: inherent properties of the overrepresented genes themselves or B-cell selection.

*Germline determinants of V gene rearrangement.* Restricted use of V genes has been found in nonproductive rearrangements in the pre-B cells of ALL\(^{103}\) and in the aborted \( k \) light chain rearrangements of \( \lambda \) light chain-expressing B cells in CLL.\(^{124}\) Therefore, bias in V gene distribution can be introduced into preimmune B-cell populations before expression of the surface Ig, and thus before any possibility of influence by antigen. A clear-cut example has been found for the case of the \( V_K \) gene \( \text{human}k325 \). The expected frequency of rearrangement of \( \text{human}k325 \) is 1.4% (1/70), yet 14% of nonproductive \( V_K \) rearrangements in \( \lambda \) light chain-expressing CLL contained this gene.\(^{124}\) Thus, at least in CLL, \( \text{human}k325 \) has an inherent tendency to rearrange, i.e., its high frequency in the repertoire is independent of antigen selection. Of interest, biased utilization of T-cell receptor variable region genes has also been noted to occur in the absence of any possibility of antigen selection.\(^{125}\)

Repertoire bias that appears before the expression of a surface Ig receptor could be a consequence of the number of copies of individual V genes in the germline, advantageous accessibility of the V gene to recombinase, the presence of V gene specific promoter sequences, or enhanced V gene transcriptional activation. Indeed, one of the frequently expressed \( V_H \) genes (\( V_{w26} \)) is present at two unique chromosomal sites\(^{126}\) and appears to be flanked by sequences that may function as promoters or enhancers of transcription.\(^{127}\) Transcriptional activation greatly increases the probability of V gene rearrangement\(^{28-32}\); therefore, it has been postulated that transcription may make V genes more accessible to recombinase.\(^{12}\) However, no evidence has directly linked gene duplication or preferential transcription to V gene selection bias. The position of \( V_H \) genes within the IgH locus (5' or 3') influences the fetal mouse repertoire,\(^{128}\) perhaps through accessibility to recombinases, but \( V_H \) gene position has no role in the formation of the human B-cell repertoire. Indeed, \( V_H \) genes that rearrange frequently in human fetal B cells are not only dispersed over 890 kb of germline DNA but also interspersed with other \( V_H \) genes that are not developmentally selected.\(^{129,130}\)

There is good experimental evidence that V gene selection occurs in fetal mice during the processes of recombination and ligation that join D to Jµ and \( V_H \) to D-Jµ.\(^{132}\) Short identical nucleotide sequences at the ends of \( V_H \), D, and Jµ appear to promote rearrangement; this results in limited junctional diversity, with a predominance of particular coding sequences at the sites of homology.\(^{132,133}\) The translated products of these rearrangements predominate in the expressed fetal repertoire.\(^{134}\) This form of directed rearrangement may also result in the pairing of homologous segments of \( V_H \) and D genes in a nonproductive reading frame.\(^{135}\) Such a junctional "mistake," if it is the rule rather than the exception, can explain the relative under-representation of certain V genes among the productively rearranged V genes of the fetal mouse repertoire.\(^{136}\) In mice deficient in the enzyme TdT, homology-directed joining is a prominent feature of early V gene rearrangement, prompting speculation that the presence of TdT obscures or inhibits this important mechanism of nonrandom V gene association.\(^{137,138}\)

*Evidence for clonal selection of V genes.* In contrast to the mouse, there is no evidence of directed V gene recombinations in embryonic muscovy ducks.\(^{139}\) On the contrary,
initial rearrangements occur randomly, whereas later in embryonic development there is a progressive positive selection for individual μ chains. This selection favors particular V₁-\(\mu\)-D and D-J₉ \(\mu\) junctional sequences and may be driven by the protein product of the CDR3. Presumably, the antigen-binding characteristics of the favored CDR3 are in some way advantageous to the animal’s developing immune system. In humans, selection for the protein product of CDR3 may explain the predominance of recurrent amino acids in the heavy chain. A dramatic reduction in splenic and lymph node B cells was observed in normal and transgenic mice,126,143,144 suggesting a selective pressure to retain structural features of the heavy chain polypeptide.145 Some frequently expressed \(\text{V}_\text{H}\) genes, such as \(\text{V}_\text{H}26\) and \(\text{V}_\text{H}6\), are conserved between species, exhibit little polymorphism, and occur in up to 98% of the population.126,143,144 Thus, advantageous features of antigen-binding sites encoded by conserved \(\text{V}_\text{H}\) genes may contribute to their over-representation in the preimmune repertoire.

Sequence conservation within human \(\text{V}_\text{H}\) families also suggests a selective pressure to retain structural features of the heavy chain polypeptide.14 Some frequently expressed \(\text{V}_\text{H}\) genes, such as \(\text{V}_\text{H}26\) and \(\text{V}_\text{H}6\), are conserved between species, exhibit little polymorphism, and occur in up to 98% of the population.126,143,144 Thus, advantageous features of antigen-binding sites encoded by conserved \(\text{V}_\text{H}\) genes may contribute to their over-representation in the preimmune repertoire.

Negative selection (clonal deletion) can influence the \(\text{V}_\text{H}\) gene repertoire of early B cells in normal145,146 and transgenic mice.147 Clonal deletion of B cells has been shown unequivocally in cases where the B cells express transgenic \(\text{V}_\text{H}\) genes of potential harmful autoantibodies.148,150 For example, a dramatic reduction in splenic and lymph node B cells was found in mice that were transgenic for the \(\text{V}_\text{H}\) and \(\text{V}_\text{K}\) genes of V₃₃, a hemolytic anti-red blood cell (RBC) autoantibody.149 Furthermore, peritoneal B cells that evaded contact with RBC autoantigens were rapidly destroyed after intraperitoneal injection of syngeneic RBCs.150,151 The \(\text{V}_\text{H}\) gene repertoire of transgenic mice bearing \(\text{V}_\text{H}\) genes for self-reactive antibodies is further limited by light chain "receptor editing,"152,154 during which autoreactive B cells try to escape deletion by replacing a pathogenic light chain with a harmless light chain.

**Autoreactive antibodies and clonal selection.** The majority of polyclonal IgM autoantibodies (natural autoantibodies) that are present in all humans and mice155-157 are encoded by unmutated (germline) \(\text{V}_\text{H}\) genes.158,159 The physiologic role, if any, of these autoantibodies is unknown. It has been proposed that the B-cell surface heavy chain counterparts of natural autoantibodies participate in clonal selection of early B cells.160 In this model, the newly formed surface heavy chain, by reacting with a self antigen, initiates signaling pathways that promote B-cell differentiation. Affinity of the surface heavy chain for self antigens would determine whether the pre-B cell is deleted (high affinity) or allowed to differentiate (low affinity). According to this scheme, the entire preimmune repertoire should be weakly autoreactive. The role of self antigen in clonal selection is supported by recent experimental evidence which shows that strongly autoreactive B cells may modify their surface Ig to escape clonal deletion.152-154 In transgenic mice, developing B cells expressing a surface light chain with high-affinity anti-DNA activity are destined for deletion unless the pathogenic light chain is removed and replaced with a light chain derived from the endogenous mouse Ig locus. This "receptor editing,"152 if successful, offers the developing B cell an opportunity to avoid deletion.152,153

Continuous positive and negative selection of pre-B cells may explain the death of the vast majority of B cells before they leave the BM161; the numerous autoactive surface \(\lg \text{M}^+\) B cells in normal blood; and the restricted heavy and light chain \(\text{V}\) gene repertoire. A parallel model of T-cell differentiation in the thymus also requires T-cell receptor recognition of self antigens.162,165 The molecular events that result in diverse antibody structures and impose limits on the immune repertoire are summarized in Table 1.

**CLINICAL RELEVANCE OF \(\text{V}\) GENE SELECTION**

**Cold agglutinins.** A dramatic example of \(\text{V}_\text{H}\) gene restriction has emerged from molecular studies of cold agglutinins. The first anti-i cold agglutinin to be sequenced was produced by a B-cell lymphoma that used a \(\text{V}_\text{H}4\) family gene segment, \(\text{V}_\text{H}4-21.166\) This \(\text{V}_\text{H}\) gene was subsequently observed in high frequency in normal B cells and in other autoantibodies.112 Analysis of 10 additional monoclonal cold agglutinins showed an absolute \(\text{V}_\text{H}4-21\) gene restriction, irrespective of the anti-i or anti-I specificity of the antibody.140,167,168 That heavy chain structures encoded by \(\text{V}_\text{H}4-21\) itself are essential for cold agglutinin binding was brought out by the unique CDR3 sequences of each cold agglutinin heavy chain and the lack of association with any particular \(\kappa\) light chain. The \(\text{V}_\text{H}4-21\) genes of the anti-i agglutinins were unmutated; in the anti-I antibodies they had few or no mutations. In contrast, the cold agglutinin light chain genes contained many nonrandom somatic mutations of the CDRs. As \(\text{V}_\text{H}4-21\) also predominates in other anti-RBC antibodies,15 it has been hypothesized that it encodes a polyreactive heavy chain with the capacity to bind weakly to a variety of RBC carbohydrates. In the case of cold agglutinins, fine specificity and high avidity binding appear to be conferred on the autoantibody by the \(\text{V}_\text{H}4\) CDR3 and by somatic muta-
tion of the light chain V region gene segments. These latter features signify an antigen-driven immune response. The restriction in the V-D junction of these antibodies (a glycine residue at this site occurred in 6/10 cold agglutinin sequences) probably represents a structural requirement for cold agglutinin binding.

It is important to keep in mind that many of these studies on monoclonal cold agglutinins dealt with the Ig products of malignant B cells. The results we have just summarized imply that the neoplasms which produced the monoclonal cold agglutinins originated from antigen-stimulated progenitor B cells. We will take up below other examples of malignancies that appear to arise from antigen-stimulated B cells.

Although pathogenic monoclonal cold agglutinins show an absolute restriction to V_{\gamma}A-21, the natural cold agglutinins in normal serum do not. B cells from two normal subjects were transformed by Epstein-Barr virus (EBV) and screened for production of anti-i and anti-I cold agglutinins; two anti-I clones and one anti-i clone were found. The anti-I clones used V_{\gamma}A family genes, whereas the anti-i clone used the V_{\gamma}D21 gene. Additional studies confirmed that most serum cold agglutinins from normal subjects were not related to the V_{\gamma}A-21 gene segment, as in monoclonal cold agglutinins. Thus, it appears that the tight correlation between cold agglutinin activity and V_{\gamma}A-21 holds only for B-cell neoplasms.

CDR3 as a clonal marker. The remarkable diversity of V_{\gamma} CD3 has been exploited to identify clonal populations of B cells.\textsuperscript{170-173} The polymerase chain reaction (PCR) is advantageous for this purpose because it amplifies the CDR3 sequence in question thousands of times. To amplify around the CDR3 segment and flanking sequences, primers corresponding to the conserved sequences of the framework 1 or framework 3 regions of V_{\gamma} and the 5' end of J_{\gamma} have been used.\textsuperscript{170-173} With such PCR primers it is possible to detect clonal populations that make up only 1% to 5% of the total population.\textsuperscript{170} However, because the primers are not clone specific, the PCR, if allowed enough cycles, will eventually amplify all CDR3 regions in the sample. Thus, detection of B-cell clonality requires a limited number of PCR cycles to differentiate clonal from polyclonal populations of B cells.\textsuperscript{172} Therefore, blind reliance on the PCR as an indicator of B-cell malignancy is to be discouraged. In up to 15% of reactive B-cell proliferations in lymph nodes or spleen, B-cell clonality has been detected by the PCR reaction.\textsuperscript{174} Furthermore, limitations of the consensus primers result in the failure to amplify clonal sequences in 20% of B-cell malignancies even when the abnormal clone is dominant. This phenomenon reflects loss of primer annealing because of 3' V_{\gamma} or 5' J_{\gamma} nucleotide deletions\textsuperscript{173} or aberrant V gene rearrangements such as those noted in Burkitt's lymphoma.\textsuperscript{175}

Thus, the sensitivity of PCR for identifying B-cell clones is little better than that of Southern analysis. To search for smaller populations of B cells, it is necessary to design tumor-specific primers. This method requires sequencing the clonal CDR3 PCR products and designing PCR primers complimentary to the unique V_{\gamma}D-J_{\gamma} junction (allele-specific primers).\textsuperscript{176-178} Allele-specific primers amplify only rearranged Ig DNA from the particular malignant B cells under study and can detect one malignant B cell in a population of 10^6 normal B cells.\textsuperscript{179} Evaluations of allele-specific PCR to detect minimal residual disease are ongoing; in one early report this technology was unable to predict relapse in ALL.\textsuperscript{180} This result is not surprising because oligoclonal B-cell populations that differ in the CDR3\textsuperscript{181,182} and ongoing V_{\gamma} gene rearrangements that disrupt the CDR3 of the malignant clone\textsuperscript{60-65} are frequent in ALL. Similar problems are likely to be encountered in follicular lymphomas, which have a high CDR3 somatic mutation rate.\textsuperscript{183} Conversely, the more stable V genes in the B-cell clones of intermediate-grade non-Hodgkin's lymphoma (NHL),\textsuperscript{184} multiple myeloma,\textsuperscript{185} and CLL\textsuperscript{186-188} are likely to lend themselves more readily to study with PCR.

ALL. Nucleotide N regions are less frequent in fetal B cells than in neonatal and adult B cells.\textsuperscript{72,73} Therefore, the absence of N nucleotides can serve as a relatively specific temporal marker of B-cell development. Wasserman et al\textsuperscript{175} showed that 39% of sequenced V genes in childhood ALL lacked N insertions, a frequency similar to that of fetal B cells. The lack of N regions was more frequent in leukemias in younger children; 87.5% of the ALL V genes from children under the age of 3 lacked N regions in CDR3, whereas only 11% of leukemic clones from children older than age 3 lacked N insertions. This finding may imply that the transforming event in the younger patients occurred in utero, and that the latency period for development of leukemia was less than 3 years. Other possibilities include low-level TdT expression in a subset of neonatal B cells or more frequent loss of TdT activity in leukemic clones arising before the age of 3.

In 98% of ALLs, clonal IgH rearrangements can be identified, whereas only 48% and 23% show k and \lambda rearrangements, respectively.\textsuperscript{181} There is no correlation between immunophenotype and Ig rearrangement.\textsuperscript{189} As with normal B-cell populations, over-representation of individual V_{\gamma}D, and J_{\gamma} genes occurs in B-cell ALL.\textsuperscript{170,174} For example, when ALL clones with rearranged V_{\gamma}1 family genes were examined, five of nine rearranged V_{\gamma} genes used the V_{\gamma}20P3 gene that was originally identified in fetal liver.\textsuperscript{184} Although exceptions have been published\textsuperscript{42} the majority of V genes in ALL are unmutated.

Multiple clonal Ig rearrangements have been found in 45% of patients with ALL.\textsuperscript{181} After the exclusion of extra copies of chromosome 14 as a contributing factor, 27% of ALL cases were bicalon and 13% oligoclonal. Oligoclonal leukemic populations share the same karyotypic abnormality; thus, the Ig rearrangements must have occurred after neoplastic transformation.\textsuperscript{182} Secondary V gene rearrangements that result in oligoclonal populations reflect ongoing recombination activity in the immature leukemic clone\textsuperscript{65,190} and are most often characterized by identical D-J_{\gamma} rearrangements joined to unique V_{\gamma} genes.\textsuperscript{190} Thus, the oligoclonality of B-cell ALL is caused by ongoing V_{\gamma} gene rearrangement or V_{\gamma} gene replacement. The presence of clonally related yet diverse minor populations of leukemic blasts is clinically significant as these clones may with time predominate and may compromise the ability to detect minimal residual disease at the molecular level.\textsuperscript{130}

Burkitt's lymphoma. The V genes encoding surface Ig
in Burkitt’s lymphoma are germline (unmutated), but in contrast to all, there is no evidence of ongoing junctional changes. This difference presumably reflects the low levels of RAG-1 and RAG-2 mRNA in surface IgM-positive Burkitt’s lymphoma cells. Structural abnormalities of Ig V genes have been reported in cell lines derived from patients with Burkitt’s lymphoma. In these lines the VH heavy chain sequence was truncated and the residual VH (framework 1) was directly rearranged to the \( \mu \) constant region with no intervening V\( _{\mu} \)CDR1-D-J\( _{\mu} \). The site at which many of these breaks occurred may be a potential alternative splice site. These kinds of truncated \( \mu \) chain products are typical of heavy chain disease; they have also been shown in EBV-transformed B-cell lines. The absence of light chain production in malignant B cells with truncated \( \mu \) chains provides further support for the necessity of a functional heavy chain during B-cell development.

**CLL.** V gene use in CLL is highly restricted and frequent pairing of the same germline VH and V\( _{\kappa} \) genes has been noted. The V\( _{\kappa} \) genes in this disease are usually unmutated and expression of surface IgM is weak. The absence of somatic mutation places the CD5\( ^{+} \) (B-1) B cell of CLL in the preimmune repertoire, or at an early stage in the immune response. However, isotype switching in CLL cells has been noted, albeit infrequently, and unusual variants of CLL may have features of the more differentiated B-cell malignancies. Consistent with this observation scattered examples of CLL harbor V genes with somatic mutations; thus, they resemble the malignant cells of non-Hodgkin’s lymphoma. Consistent with clinical patterns, V gene sequences in Richter’s transformation indicate a common clonal origin of the developing high-grade lymphoma and the preceding CLL.

**NHLs.** V gene restriction has been found in follicular NHL. In eight cases that had rearranged a VH\( _{4} \) family gene, three used the VH\( _{4} \)-21 gene and three VH\( _{7} \)-4. As with other B-cell malignancies, the restricted use of V genes in follicular lymphoma is unlikely to be disease specific.

Unlike malignancies arising early B cells, follicular lymphomas characteristically have heavily mutated V genes. The influence of an antigenic stimulus is suggested by prominent replacement mutations in the CSR regions. By contrast, the majority of mutations in the framework regions are silent. Thus, in follicular NHL the forces of clonal selection may preserve the structural framework of the Ig while continuing to modify its CD\( _{R} \)s.

V genes in follicular lymphoma B cells can also exhibit infraclonal diversity. Analysis of the V genes sequences from a single lymphoma strongly suggested that they derived from a common precursor V gene that was modified by somatic mutation, a process that resulted in an oligoclonal population of neoplastic B cells. The ongoing somatic mutation in follicular lymphoma suggests that antigen stimulation not only perpetuates the malignant clone but also causes the waxing and waning course of the disease. Somatic mutation in low-grade lymphoma has further clinical relevance because V gene mutation may result in escape from anti-idiotype sensitivity and thus limit the efficacy of monoclonal antibody therapy.

Serial study of the V genes of follicular lymphomas that progress to diffuse lymphoma has shown numerous shared mutations in both tumors, indicating the origin of the diffuse lymphoma from a single follicular lymphoma cell. Friedman et al showed that a minor clone in a transforming low-grade lymphoma emerged over time to predominate and eventually dominate the diffuse NHL. In contrast, de novo intermediate-grade NHLs show no evidence for ongoing somatic mutation. A similar lack of mutation was noted in a case of gastrointestinal mucosa associated lymphoid tissue (MALT) lymphoma associated with Sjogren’s syndrome that was studied over a period of 13 years. Thus, it would appear that the machinery for somatic mutation is active in follicular B-cell lymphomas, but not in diffuse or MALT-associated NHL.

**Multiple myeloma.** As might be expected, the majority of V genes in multiple myeloma appear heavily mutated, in keeping with the mature B-cell derivation of this malignancy. However, these results are not definitive because the V gene sequences in the myeloma cells were not sought in the patient’s germline. In humans, V gene mutations can be distinguished from polymorphic variants or unknown V genes only by comparisons with unmutated germline sequences. Both idiotypic and sequence analyses suggest that the V gene repertoire of this mature B-cell malignancy is restricted, whereas analysis of tumor-derived V genes indicates that there is little intraclonal variation or ongoing somatic mutation, although isotype switching may occur. The use of CD3 as a clonal marker has been rewarding in myeloma. Circulating malignant cells can be identified in the majority of patients at a frequency of 0.001% to 1.0%. Moreover, circulating \( \kappa \) B cells have been found to be clonally related to, and perhaps precursors of, the plasma cells that secrete \( \kappa \) or \( \kappa \alpha \) paraproteins. These results imply that multiple myeloma may arise from a germinal center \( \kappa \) B cell which, under the influence of an immunogenic stimulus and T-cell help, makes an isotype switch, mutates its V genes, and returns to the BM for clonal expansion.

Analysis of somatic mutations may provide clues to the pathogenic nature of some Igs in this disease. This is illustrated by the structural analysis of a V gene encoding a light chain antibody associated with light chain nephropathy. This light chain V gene showed mutations that created new antigenic sites and thus limit the efficacy of anti-idiotype antibodies. Analysis of somatic mutations may provide clues to the pathogenic nature of some Igs in this disease. This is illustrated by the structural analysis of a V gene encoding a light chain antibody associated with light chain nephropathy. Indeed, the \( \lambda \) light chain protein eluted from diseased kidney was abnormally heavily glycosylated, a finding that suggests a role for aberrant glycosylation in light chain nephropathy. In nonsecretory myeloma, alterations in the coding sequence of the variable region result in a failure of VH transcription and rapid intracellular degradation of the truncated \( \gamma \) chain.

**SUMMARY**

There is now substantial evidence that a small group of V genes predominates in the Ig repertoire of preimmune B cells. This phenomenon of V gene restriction may reflect preferential accessibility of these genes to recombination, homology-directed V gene rearrangement, promoters and enhancers of V gene transcription, or positive and negative
selection mediated by the anti–self binding properties of the B cells surface Ig. These mechanisms may operate alone or in combination to influence V gene rearrangement and populations of immature B cells. Although constraints on the pool of rearranged V genes may seem disadvantageous to the immune system, the mechanisms that generate the CDR3s of heavy and light chains ensure extensive diversity in the pre–B-cell population. In mature B cells, somatic mutation of V genes adds further diversity. CDR3 sequences and somatic mutations not only provide potentially useful clonal markers but also help to identify the normal counterparts of malignant B cells.

REFERENCES

34. Serwe M, Sablotzky F: FDI recombination in B cells is impaired but not blocked by targeted deletion of the immunoglobulin heavy chain intron enhancer. EMBO J 12:2321, 1993
Ordered rearrangement of immunoglobulin heavy chain variable region. EMBO J 3:1209, 1984


119. Kipps TJ, Duffy SF: Relationship to the CD5 B cell to human tonsillar lymphocytes that express autoantibody associated cross reactive idiotypes. J Clin Invest 87:2087, 1991


126. Rubinstein DB, Symann M, Stewart AK, Guillaume T: Restriction fragment length polymorphisms and single germline coding region sequence in VH18/2 a duplicated gene encoding autoantibody. Mol Immunol 30:403, 1993


155. Pisetsky DS, Jelinek DF, McAnally LM, Reich CF, Lipsky...
V REGIONS AND B CELLS

1729


158. Siminovich KA, Misener V, Kwong PC, Song QL, Chen PP: A natural autoantibody is encoded by germline heavy and lambda light chain variable region genes. J Clin Invest 84:1675, 1989


161. Gallagher RB, Osmond DC: To B or not to B: That is the question. Immunol Today 12:1, 1991


169. Jeffries LC, Carchidi CM, Silberstein LE: Naturally occurring anti-i/I cold agglutinins may be encoded by different Vn3 genes as well as the Vn4.21 gene segment. J Immunol 92:2821, 1993


191. Pollok BA, Anker R, Eldridge P, Hendershot L, Levitt D: Molecular basis of the cell-surface expression of immunoglobulin...
m chain without light chain in human B lymphocytes. Proc Natl Acad Sci USA 84:9199, 1987


204. Ebeling SB, Schute MEM, Logtenberg T: Molecular analysis of VH and VL regions expressed in IgG bearing chronic lymphocytic leukemia (CLL): Further evidence that CLL is a heterogeneous group of tumors. Blood 82:1626, 1993


Immunoglobulin V regions and the B cell

AK Stewart and RS Schwartz